



VINICIUS SAMUEL MARTINS

**ANALYSIS STRATEGIES FOR OPTIMIZING THE
SELECTION OF POTATO CLONES**

**LAVRAS - MG
2025**

VINICIUS SAMUEL MARTINS

**ANALYSIS STRATEGIES FOR OPTIMIZING THE SELECTION OF POTATO
CLONES**

Tese apresentada à Universidade Federal de
Lavras, como parte das exigências do programa
de Pós-Graduação em Genética e
Melhoramento de Plantas, para obtenção do
título de Doutor.

Prof. Dr. Tiago de Souza Marçal
Orientador

**LAVRAS - MG
2025**

**Ficha Catalográfica elaborada pelo Sistema de Geração
de Ficha Catalográfica da Biblioteca Universitária da UFLA, com
dados informados pelo(a) próprio(a) autor(a).**

Martins, Vinícius Samuel .

Analysis strategies for optimizing the selection of potato clones / Vinícius Samuel
Martins. - 2025.

64 p. : il.

Orientador: Tiago de Souza Marçal

Tese (Doutorado) - Universidade Federal de Lavras, 2025.
Bibliografia.

1. Solanum tuberosum L. 2. acurácia seletiva. 3. visão computacional. 4.
análise multivariada. I. Marçal, Tiago de Souza. II. Universidade Federal de Lavras.
III. Título.

VINICIUS SAMUEL MARTINS

**ANALYSIS STRATEGIES FOR OPTIMIZING THE SELECTION OF POTATO
CLONES**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do programa de Pós-Graduação em Genética e Melhoramento de Plantas, para obtenção do título de Doutor.

APROVADA em 28 de maio de 2025.

Dr. Vinícius Quintão Carneiro UFLA

Dra. Flávia Maria Avelar Gonçalves UFLA

Dr. Pedro Crescêncio Souza Carneiro UFV

Dr. Marcio Fernando Ribeiro de Resende Jr. UF

Prof. Dr. Tiago de Souza Marçal
Orientador

**LAVRAS - MG
2025**

AGRADECIMENTOS

A Deus, em primeiro lugar, por me sustentar em cada etapa desta jornada. Foram muitos os desafios enfrentados ao longo do doutorado, mas em todos eles, pude experimentar a graça, a providência e o cuidado do Senhor. A Ele, a honra, minha gratidão por me conceder força, sabedoria e perseverança.

À minha esposa Letícia, meu amor e companheira incansável nesta caminhada. Mais do que apoio, você foi presença, incentivo e parceria em todas as fases, inclusive como colega de trabalho, onde podemos discutir, trabalhar e planejar vários projetos juntos. Você me inspira a ser um profissional e marido melhor.

Aos meus pais, Marlene e Vicente, por todo amor, exemplo e apoio incondicional ao longo da vida sendo verdadeiros professores.

Aos meus irmãos, Vanessa e Leandro, por estarem sempre ao meu lado com palavras de incentivo e gestos de carinho.

À família da minha esposa, pelo acolhimento e suporte constante ao longo dessa trajetória.

Ao meu orientador, professor Tiago, por sua orientação firme, dedicada e generosa. Sua confiança e direcionamento foram fundamentais para a realização deste trabalho.

Ao professor Márcio, minha gratidão pela oportunidade enriquecedora de vivenciar o intercâmbio na Universidade da Flórida, experiência que expandiu minha visão acadêmica e pessoal.

Ao Ministério Universidade Renovadas e a MECE, por ser expressão concreta do cuidado de Deus em minha vida, por meio dos grupos de oração e amizades que fortaleceram minha fé e esperança durante este período.

À Universidade Federal de Lavras (UFLA), meu sincero agradecimento pelo apoio institucional, pela estrutura física de excelência e, principalmente, pelo ambiente intelectual que tanto contribuiu para minha formação científica e pessoal.

Ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas da UFLA, agradeço a todos os professores, técnicos e funcionários pelo profissionalismo, acolhimento e contribuição em cada etapa desta formação.

Ao PROBATATA, pela parceria técnica essencial à execução destas pesquisas. Em especial, ao técnico Ramon, por todo apoio prático, dedicação e ajuda na condução das etapas experimentais do projeto.

Aos órgãos de fomento CNPq, CAPES e FAPEMIG, pela concessão de bolsas e financiamento, que possibilitaram a dedicação integral à pesquisa e à formação científica.

A todos que, de alguma forma, caminharam comigo nesta jornada, deixo aqui minha gratidão sincera.

“Tudo posso naquele que me fortalece” (Filipenses 4:13)

RESUMO GERAL

A batata (*Solanum tuberosum* L.) destaca-se como a quarta cultura fundamental para a segurança alimentar global, atrás de arroz, trigo e milho. Essa importância é devido ao seu alto potencial produtivo por área, ciclo curto e valor nutricional. No entanto, o melhoramento genético desta cultura enfrenta desafios significativos devido à herança tetrassômica, alta heterozigosidade dos parentais, baixa taxa de multiplicação dos tubérculos, baixa herdabilidade para caracteres de rendimento de tubérculos, pequenos ganhos genéticos e interação genótipos por ambientes (G×E). Com o objetivo de aprimorar a eficiência do melhoramento genético da batata, foram realizados dois estudos no intuito de promover estratégias para incrementar a acurácia seletiva de clones de batata. O primeiro investigou a contribuição do efeito de família na seleção de clones na fase inicial de um programa de melhoramento de batata envolvendo a avaliação clonal em *Single Trial* (ST) ou *Multi-Environment Trials* (MET). Foram avaliados 1280 clones, oriundos de 67 famílias de irmãos completos (cruzamentos biparentais) em diferentes condições de estresse por calor para os caracteres total tuber yield (TTY) e specific gravity (SG). O segundo analisou a partição dos componentes de produção e seu impacto na acurácia seletiva, analisando 620 clones por meio de análises de modelos mistos unicaracterístico (MMUC) e modelo misto multicaracterístico (MMMC) na safra de inverno (sem estresse por calor) para os caracteres massa média de tubérculos (MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT), com o auxílio de imagens digitais. Os resultados demonstraram que a inclusão do efeito de família aumentou a eficiência seletiva tanto em análises ST quanto MET. O modelo MMMC apresentou melhor ajuste para o MMUC, porém não apresentou incremento na acurácia seletiva em análise univariada para safra de inverno. Entretanto, a análise multivariada para componentes de produção de batata, pode ser útil em cenários com baixo controle ambiental, proporcionando incremento na acurácia seletiva de clones em estágios iniciais do programa de melhoramento de batata. A inclusão do efeito de família é vantajosa, pois aumenta acurácia seletiva de clones de batata sem trazer custos adicionais.

Palavras-chave: *Solanum tuberosum* L.; acurácia seletiva; análise multivariada; visão computacional.

ABSTRACT

Potato (*Solanum tuberosum* L.) stands out as the fourth fundamental crop for global food security, behind rice, wheat, and corn, due to its high productive potential per area, short cycle, and nutritional value. However, genetic improvement of this crop faces significant challenges due to tetrasomic inheritance, high parental heterozygosity, low tuber multiplication rate, low heritability for tuber yield traits, small genetic gains, and genotype-by-environment interaction (G×E). With the objective of improving the efficiency of potato genetic improvement, two studies were conducted to promote strategies to increase the selective accuracy of potato clones. The first investigated the contribution of the family effect in the selection of clones in the initial phase of a potato breeding program involving clonal evaluation in Single Trial (ST) or Multi-Environment Trials (MET). 1280 clones from 67 full-sib families (biparental crosses) under different heat stress conditions were evaluated for the traits total tuber yield (TTY) and specific gravity (SG). The second analyzed the partition of yield components and their impact on selective accuracy, evaluating 620 clones using Single-trait Mixed Model (STMM) and Multi-trait Mixed Model (MTMM) in the winter crop (without heat stress) for the traits average tuber mass (ATM), total tuber number (TTN), and total tuber yield (TTY), using a digital imaging approach. The results demonstrated that the inclusion of the family effect increased selective efficiency in both ST and MET analyses. The MTMM model showed better fit than STMM, but showed no increase in selective accuracy in univariate analysis for the winter crop. However, multivariate analysis for potato yield components can be useful in scenarios with low environmental control, providing increased selective accuracy of clones in early stages of the potato breeding program. The inclusion of the family effect is advantageous, as it increases the selective accuracy of potato clones without incurring additional costs.

Keywords: *Solanum tuberosum* L.; selective accuracy; multivariate analysis; computer vision.

INDICADORES DE IMPACTO

Este trabalho apresenta impactos tecnológicos, econômicos e sociais com ferramentas para o melhoramento genético de batata visando o desenvolvimento de novos clones de batata frente as mudanças climáticas e desafios na produção agrícola para segurança alimentar global. A pesquisa desenvolveu e validou metodologias estatísticas avançadas que aumentaram a eficiência seletiva de clones de batata em até 12%, através da incorporação do efeito de família em modelos mistos para análise de ensaios de um único local e multiambientes. Adicionalmente, o auxílio da fenotipagem por imagens digitais possibilitou mensurar em escala os componentes de produtividade com baixo custo, viabilizando a avaliação simultânea de múltiplos caracteres em etapas iniciais do programa de melhoramento de batata. O trabalho possui caráter extensionista ao envolver diretamente no Programa de Melhoramento Genético de Batata da Universidade Federal de Lavras (PROBATATA-UFLA) e poderá beneficiar produtores rurais e técnicos da região Sul de Minas Gerais, bem como todo o estado através do desenvolvimento de clones tolerantes ao estresse térmico e com maior potencial produtivo. Foram avaliados milhares de clones em ambientes contrastantes quanto ao estresse por calor, impactando diretamente a seleção de materiais adaptados às condições tropicais e subtropicais brasileiras. A pesquisa envolveu colaboração internacional com instituições dos Estados Unidos (University of Florida e University of Maine) e nacional (UFV, UFG, CTC), demonstrando impacto científico e tecnológico. Os resultados classificam-se nas áreas temáticas de Tecnologia e Produção (7) e Meio Ambiente (5) da Política Nacional de Extensão, contribuindo para os Objetivos de Desenvolvimento Sustentável da ONU, especialmente ODS 2 (Fome Zero e Agricultura Sustentável), ODS 12 (Consumo e Produção Responsáveis) e ODS 13 (Ação Contra a Mudança Global do Clima), ao desenvolver ferramentas para aumentar a produtividade agrícola de alimentos de forma sustentável e adaptar cultivos às mudanças climáticas na cultura da batata.

IMPACT INDICATORS

This work presents the technological, economic, and social impacts of tools for potato genetic improvement, aiming at the development of new potato clones in the face of climate change and challenges in agricultural production for global food security. The research developed and validated advanced statistical methodologies that increased the selective efficiency of potato clones by up to 12%, through the incorporation of the family effect in mixed models for the analysis of single-site and multi-environment trials. Additionally, the use of digital image phenotyping made it possible to measure productivity components at a low cost, enabling the simultaneous evaluation of multiple traits in the initial stages of the potato breeding program. The work has an extension character as it directly involves the Potato Genetic Improvement Program of the Universidade Federal de Lavras (PROBATATA-UFLA) and may benefit rural producers and technicians in the southern region of Minas Gerais, as well as the entire state, through the development of clones tolerant to thermal stress and with greater productive potential. Thousands of clones were evaluated in contrasting environments regarding heat stress, directly impacting the selection of materials adapted to Brazilian tropical and subtropical conditions. The research involved international collaboration with institutions in the United States (University of Florida and University of Maine) and national institutions (UFV, UFG, CTC), demonstrating scientific and technological impact. The results are classified in the thematic areas of Technology and Production (7) and Environment (5) of the National Extension Policy, contributing to the UN Sustainable Development Goals, especially SDG 2 (Zero Hunger and Sustainable Agriculture), SDG 12 (Responsible Consumption and Production) and SDG 13 (Climate Action), by developing tools to increase agricultural food productivity sustainably and adapt crops to climate change in potato cultivation.

LISTA DE FIGURAS

Figura 1	Modelo da estrutura do estúdio fotográfico usado para capturar as imagens digitais e (B) imagem de uma parcela de batata, contendo uma etiqueta de identificação e um objeto de referência.....	42
Figura 2	Comparação do ajuste de qualidade do Modelo Misto Multicaracterístico (MMMC) usando o AIC em relação ao Modelo Misto Unicaracterístico (MMUC) com efeitos de clone independentes para três populações clonais diferentes de batata (POP).....	46
Figura 3	Partição da variância genética para os caracteres (A) massa média do tubérculo (MMT, g/ tubérculo), (B) número total de tubérculos por parcela (NTT) e (C) rendimento total de tubérculos (PTT, t ha ⁻¹) do Modelo Misto Multicaracterístico (MMMC) e Modelo Misto Unicaracterístico (MMUC) para os componentes de variância: σ_g^2 variância do clone; σ_c^2 variância da coluna; σ_e^2 variância do erro; σ_l^2 variância da linha. Foram avaliadas três populações clonais diferentes de batata (POP).....	46
Figura 4	Correlações parciais genótípicas (ρ_g) (A) e residuais (ρ_e) (B) entre características avaliadas para três populações clonais diferentes de batata (POP). As barras representam as estimativas das correlações e os intervalos do erro padrão ($\pm$ SE). (*) indicam correlações que são significativamente diferentes de zero com 95% probabilidade.....	48
Figura 5	Acurácia seletiva r_{gg} estimada para os modelos: Modelo Misto Unicaracterístico (MMUC) e Modelo Misto Multicaracterístico (MMMC) para os caracteres massa média do tubérculo (MMT, g/ tubérculo), número total de tubérculos por parcela (NTT) e rendimento total de tubérculos (PTT, t ha ⁻¹), sob condições sem estresse térmico em três populações clonais diferentes de batata (POP).....	50

LISTA DE TABELAS

Tabela 1	Caracterização de populações, número de linhas, coluna, clone, testemunha, parcela e porcentagem de parcelas replicadas.....	40
Tabela 2	Resumo dos modelos ajustados: Modelo Misto Unicaracterístico (MMUC) e Modelo Misto Multicaracterístico (MMMC).....	42
Tabela 3	Coeficiente de Czekanowski (CC) entre os vetores de pontuação do índice FAI-BLUP para seleção dos 20% melhores clones das estratégias ($\mathbf{u}_{gMMUC}$ e $\mathbf{u}_{gMMMC}$) para seleção simultânea dos componentes de produção de batata: massa média do tubérculo (MMT, g/ tubérculo), número total de tubérculos por parcela (NTT) e rendimento total de tubérculos (PTT, t ha ⁻¹).....	50
Tabela S1	Resumo do Modelo Misto Unicaracterístico (MMUC), Modelo Misto Multicaracterístico (MMMC): Ponto máximo residual log-likelihood (ℓ) e Akaike Critério de Informação de Akaike (AIC).....	56
Tabela S2	Resumo dos componentes de variância do Modelo Misto Unicaracterístico (MMUC), Modelo misto Multicaracterístico (MMMC), para os caracteres massa média de tubérculos (MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT, t ha ⁻¹) na safra de inverno de 2023).....	56
Tabela S3	Contribuição (%) dos componentes de variância do Modelo Misto Unicaracterístico (MMUC), Modelo Misto Multicaracterístico (MMMC) para os caracteres: massa média de tubérculos (MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT, t ha ⁻¹) na safra de inverno de 2023.....	57
Tabela S4	Resumo das estimativas dos parâmetros de variância genética e não genética para o Modelo Misto Multicaracterístico (MMMC) para populações clonais. Limite inferior de confiança (LIC) e limite superior de confiança (LSC) do intervalo de confiança de 95% para os parâmetros de variância associados às características massa média de tubérculos	

	(MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT, t ha ⁻¹) na safra de inverno de 2023.....	58
Tabela S5	Acurácia seletiva para o Modelo Misto Unicaracterístico (MMUC), Modelo Misto Multicaracterístico (MMMC) para os caracteres massa média de tubérculos (MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT, t ha ⁻¹) na safra de inverno de 2023.....	59

LISTA DE SIGLAS

ABA	Ácido abscísico
AIC	Critério de informação de Akaike
BLUP	Melhores previsões lineares não viesadas
CC	Coeficiente de Czekanowski
DIAG	Estrutura de variância e covariância diagonal
GP	Predição genômica
MET	Ensaio em múltiplos ambientes
METMpF	Modelo de ensaio multiambiental mais efeito familiar
METMwF	Modelo de ensaio multiambiental sem efeito familiar
MMMC	Modelo misto multicaracterístico
MMUC	Modelo misto unicaracterístico
MMT	Massa média de tubérculos
NTT	Número de tubérculos totais
POP	Populações clonais diferentes de batata
PTT	Produtividade total de tubérculos
REML	Verossimilhança máxima residual
SG	Gravidade específica
ST	Ensaio único
STMpf	Modelo de ensaio único mais efeito familiar
TC	Teste clonal
TMAX	Temperatura máxima
TMEAN	Temperatura média compensada
TMIN	Temperatura mínima
TNT	Tecido não tecido
US	Forma não estruturada

SUMÁRIO

PRIMEIRA PARTE.....	16
1 INTRODUÇÃO	17
REFERÊNCIAS	19
SEGUNDA PARTE – ARTIGOS.....	20
ARTIGO 1 - EVALUATING THE IMPACT OF MODELING THE FAMILY EFFECT FOR CLONAL SELECTION IN POTATO BREEDING PROGRAMS	21
ABSTRACT	21
1 INTRODUCTION	22
2 MATERIALS AND METHODS	23
3 RESULTS	27
4 DISCUSSION.....	31
REFERENCES	34
SUPPLEMENTARY MATERIAL	34
SUPPLEMENTARY REFERENCES.....	34
ARTIGO 2 - FENOTIPAGEM POR IMAGENS NA SELEÇÃO DE CLONES DE BATATA COM BASE NA PRODUTIVIDADE E SEUS COMPONENTES.....	37
RESUMO	37
1 INTRODUÇÃO.....	37
2 MATERIAIS E MÉTODOS	40
3 RESULTADOS.....	45
4 DISCUSSÃO	50
5 CONCLUSÃO.....	52
REFERÊNCIAS	53
MATERIAL SUPLEMENTAR	56
2 CONSIDERAÇÕES FINAIS.....	60
REFERÊNCIAS	62

PRIMEIRA PARTE

INTRODUÇÃO

A batata (*Solanum tuberosum* L.) é uma das culturas mais importantes para a segurança alimentar global, ocupando a quarta posição em relevância na alimentação humana, ao lado de arroz, trigo e milho. Sua ampla adaptação, elevada produtividade por unidade de área e versatilidade culinária a tornam essencial para atender à crescente demanda por alimentos, especialmente em um cenário de crescimento populacional (FAO, 2023; Devaux et al., 2021). No entanto, apesar de seu potencial, a produtividade média global (19,1 Mg há⁻¹) apresenta uma pequena parcela da produtividade potencial, indicando oportunidades significativas para incrementos por meio de inovações tecnológicas e ao melhoramento genético (Kunkel e Campbell, 1987).

O melhoramento genético da batata enfrenta desafios complexos, como a segregação tetrassômica, que gera maior variabilidade genética em comparação com a segregação dissômica, e a alta heterozigose dos parentais (Haynes *et al.*, 2012; Bradshaw, 2022). Além disso, as fases iniciais dos programas de melhoramento de batata são limitadas pela escassez de material propagativo, o que restringe o uso de repetições e exige delineamentos experimentais eficientes em parcelas replicadas, como blocos aumentados (DBA) ou delineamentos parcialmente replicados (P-REP) (Cullis *et al.*, 2006; Paget *et al.*, 2017).

A interação genótipo $\times$ ambiente (G $\times$ E) também desempenha um papel crítico, especialmente em regiões tropicais e subtropicais, onde a tolerância ao calor é fundamental para seleção de clones tolerantes ao calor (Fernandes Filho *et al.*, 2021; Patiño-Torres *et al.*, 2021). Modelos mistos têm sido empregados para melhorar a acurácia seletiva nessas condições, contabilizando efeitos aditivos e não aditivos (Amadeu *et al.*, 2020; Voss-Fels *et al.*, 2021).

A seleção de clones por meio de caracteres quantitativos é outro obstáculo que exige abordagens multivariadas para otimizar a eficiência seletiva de clones de batata (Bradshaw, 2017; Thelen *et al.*, 2025). Outra questão é a utilização de modelos mistos multicaracterístico (MMMC) para seleção de clones e na predição genômica. Além disso, modelos MMMC têm demonstrado vantagens sobre modelos mistos unicaracterístico (MMCU), especialmente quando há correlações entre os caracteres (Cuevas *et al.*, 2023).

Avanços em fenotipagem, como a visão computacional aplicada à análise de imagens digitais têm revolucionado a mensuração de caracteres de componentes de produtividade e qualidade no melhoramento de batata. O uso de imagem digitais permite avaliações de forma não destrutiva, identificação de caracteres de maneira quantitativa, o que reduz o viés de

avaliadores, viabiliza análises em escala com baixo custo e de forma semiautomatizada (Thelen *et al.*, 2025).

Neste contexto, a tese contém dois artigos abordando estratégias para seleção de clones de batata. O primeiro artigo teve como objetivo investigar a contribuição do efeito de família na seleção de clones na fase inicial de um programa de melhoramento de batata envolvendo a avaliação clonal em *Single Trial* (ST) ou *Multi-Environment Trials* (MET). O segundo artigo teve como intuito, com auxílio de imagens digitais, investigar a relação entre os componentes da produtividade da batata e seu impacto na acurácia seletiva de clones em um programa de melhoramento genético de batata. Ambas as seções têm como propósito cooperar para a solução dos gargalos atuais do melhoramento da batata e no incremento da acurácia seletiva no avanço da seleção de clones superiores e no entendimento de caracteres complexos, como componentes de produção no melhoramento de batata.

REFERÊNCIAS

- AMADEU, R. R. *et al.* Impact of dominance effects on autotetraploid genomic prediction. **Crop Science**, v. 60, p. 656–665, 2020.
- BRADSHAW, J. E. A Brief History of the Impact of Potato Genetics on the Breeding of Tetraploid Potato Cultivars for Tuber Propagation. **Potato Research**, v. 65, p. 461–501, 2022.
- BRADSHAW, J. E. Revisão e análise das limitações nas formas de melhorar o melhoramento convencional da batata. **Potato Research**. 60, 171–193, 2017.
- CUEVAS, J. *et al.* Modeling genotype× environment interaction for single and multitrait genomic prediction in potato (*Solanum tuberosum* L.). **G3**, v. 13, n. 2, p. 1-13, 2023.
- CULLIS B. R.; SMITH, A. B.; COOMBES, N. E. On the design of early generation variety trials with correlated data. **Journal Agric Biol Environment Stat** v. 11, p. 381–393, 2006.
- DEVAUX, A. *et al.* The potato of the future: opportunities and challenges in Sustainable agri-food systems. **Potato Research**, v. 64, n. 4, p. 681-720, 2021.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). **FAOSTAT**. 2023. Disponível em: www.fao.org/faostat/en/#data/QCL Acesso em: 15 maio 2025.
- FERNANDES FILHO, C. C. *et al.* (2021). Selection of potato clones for heat tolerance and resistance to potato viruses X and Y for processing purposes. **Crop Science**, 61, 552–565, 2021.
- HAYNES, K. G.; GERGELA, D. M.; HUTCHINSON, C. M. Early generation selection at multiple locations may identify potato parents that produce more widely adapted progeny. **Euphytica**, v. 186, n. 2, p. 573-583, 2012.
- KUNKEL, R.; CAMPBELL, G. S. Maximum potential potato yield in the Columbia Basin, USA: Model and measured values. **American Potato Journal**, 64, 355–366, 1987.
- PAGET, M. F. *et al.* Replicate allocation to improve selection efficiency in the early stages of a potato breeding scheme. *Euphytica* 213, 221, 2017.
- PATÍÑO-TORRES, A. J. *et al.* (2021). Performance of superior potato clones under high and mild temperatures in tropical climate. **Agron Journal**, 113, 2349–2360, 2021.
- THELEN, K. *et al.* Variance and covariance components of agronomic and quality traits assessed in tetraploid potato and their implications on practical breeding. **Front Plant Sci**, v. 15, p. 1505193, 2025.
- VOSS-FELS, K. P. *et al.* Strategies and considerations for implementing genomic selection to improve traits with additive and non-additive genetic architectures in sugarcane breeding. **Theor Appl Genet**, v. 134, p. 1493–1511, 2021.

SEGUNDA PARTE**ARTIGO 1**

Artigo elaborado e publicado na Revista Frontiers https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2023.1253706/full	
Título do artigo:	Evaluating the impact of modeling the family effect for clonal selection in potato-breeding programs
Autores:	Vinícius Samuel Martins e colaboradores



OPEN ACCESS

EDITED BY

Ulrich Schurr,
Helmholtz Association of German
Research Centres (HZ), Germany

REVIEWED BY

Juliano Lino Ferreira,
Embrapa Pecuária Sul, Brazil
Asrat Asfaw,
International Institute of Tropical
Agriculture (IITA), Nigeria

*CORRESPONDENCE

Tiago de Souza Marçal
✉ tiago.marcal@ufpa.br
Vinicius Samuel Martins
✉ viniciusmartins93@outlook.com

RECEIVED 06 July 2023

ACCEPTED 25 September 2023

PUBLISHED 30 October 2023

CITATION

Martins VS, Andrade MHML, Padua LN,
Miguel LA, Fernandes Filho CC,
Guedes ML, Nunes JAR, Hoffmann L Jr,
Zotarelli L, Resende MFRd Jr,
Carneiro PCS and Marçal TdS (2023)
Evaluating the impact of modeling
the family effect for clonal selection
in potato-breeding programs.
Front. Plant Sci. 14:1253706.
doi: 10.3389/fpls.2023.1253706

COPYRIGHT

© 2023 Martins, Andrade, Padua, Miguel,
Fernandes Filho, Guedes, Nunes, Hoffmann,
Zotarelli, Resende, Carneiro and Marçal. This
is an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Evaluating the impact of modeling the family effect for clonal selection in potato-breeding programs

Vinicius Samuel Martins^{1*},
Mario Henrique Murad Leite Andrade²,
Leticia Novais Padua¹, Luciana Aparecida Miguel¹,
Claudio Carlos Fernandes Filho³, Marcio Lisboa Guedes⁴,
Jose Airton Rodrigues Nunes¹, Leo Hoffmann Jr⁵,
Lincoln Zotarelli⁵, Márcio Fernando Ribeiro de Resende Jr⁵,
Pedro Crescêncio Souza Carneiro⁶ and
Tiago de Souza Marçal^{1*}

¹Departamento de Biologia, Universidade Federal de Lavras, Lavras, Brazil, ²School of Food and Agriculture, University of Maine, Orono, ME, United States, ³Centro de Tecnologia Canavieira (CTC), Piracicaba, Brazil, ⁴Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (RIDESA), Universidade Federal de Goiás, Goiânia, Brazil, ⁵Department of Horticultural Sciences, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, United States, ⁶Departamento de Biologia, Universidade Federal de Viçosa, Viçosa, Brazil

Because of its wide distribution, high yield potential, and short cycle, the potato has become essential for global food security. However, the complexity of tetrasomic inheritance, the high level of heterozygosity of the parents, the low multiplication rate of tubers, and the genotype-by-environment interactions impose severe challenges on tetraploid potato-breeding programs. The initial stages of selection take place in experiments with low selection accuracy for many of the quantitative traits of interest, for example, tuber yield. The goal of this study was to investigate the contribution of incorporating a family effect in the estimation of the total genotypic effect and selection of clones in the initial stage of a potato-breeding program. The evaluation included single trials (STs) and multi-environment trials (METs). A total of 1,280 clones from 67 full-sib families from the potato-breeding program at Universidade Federal de Lavras were evaluated for the traits total tuber yield and specific gravity. These clones were distributed in six evaluated trials that varied according to the heat stress level: without heat stress, moderate heat stress, and high heat stress. To verify the importance of the family effect, models with and without the family effect were compared for the analysis of ST and MET data for both traits. The models that included the family effect were better adjusted in the ST and MET data analyses for both traits, except when the family effect was not significant. Furthermore, the inclusion of the family effect increased the selective efficiency of clones in both ST and MET analyses via an increase in the accuracy of the total genotypic value. These same models also allowed the prediction of clone effects more realistically, as the variance components associated with family and clone effects within a family were not confounded. Thus, clonal selection based on the total genotypic value, combining the effects of family and clones within a family,

proved to be a good alternative for potato-breeding programs that can accommodate the logistic and data tracking required in the breeding program.

KEYWORDS

Solanum tuberosum L., nested structure, accuracy, G×E interactions, autotetraploid genetics, tuber yield

1 Introduction

Potato is the third most important crop for human consumption worldwide, playing a central role in global food security. Potato has a wide adaptation and higher yield compared with cereal crops (FAO, 2020). It will certainly continue to have an essential role in food security in the coming years, particularly regarding population growth (Devaux et al., 2014; FAO, 2020; Devaux et al., 2021). The world's average potato production has grown at a rate of 2% per year for the past 20 years, with an average yield of 21.0 Mg ha⁻¹, which represents only 13% of the potential yield (Kunkel and Campbell, 1987; FAO, 2021). The gap in production between average and potential yield presents the potential for increasing global potato production. This potential can be exploited through technological innovations in the potato production system, using new and improved cultivars, and optimizing agricultural practices. Within the framework of genetic improvement, tuber yield can be increased through an accurate selection of clones more tolerant to biotic and abiotic stresses, and more efficient in the use of resources, such as water and nitrogen, meeting the demand for an increasingly sustainable global production system (Foley et al., 2011; Birch et al., 2012; Obidiegwu et al., 2015; Dahal et al., 2019; Devaux et al., 2021).

Tetraploid potato-breeding programs generate thousands of seedlings annually (Stich and Van Inghelandt, 2018). Larger populations are required to increase the probability of selecting superior clones because potato breeders must deal with the complexity of tetrassomic segregation, heterosis, and high level of heterozygosity of parents (Meyer et al., 1998; Gopal, 2015). A small number of seed potatoes are available in the early stages of a potato-breeding program (Haynes et al., 2012; Stich and Van Inghelandt, 2018), which restricts the use of repetitions and the number of plants per plot (Haynes et al., 2012; Paget et al., 2017). In this context, the use of unreplicated designs, such as the augmented block design (ABD) (Federer, 1956), has been frequent (Andrade et al., 2020; Fernandes Filho et al., 2021). Partially replicated design (P-REP) (Cullis et al., 2006) is an efficient alternative in the initial stages of potato breeding (Paget et al., 2017). Furthermore, P-REP can be increased (Williams et al., 2011; Williams et al., 2014), i.e., a proportion of candidates can be replicated in each location, which allows the study of genotype-by-environment interaction (G×E) even with limited seed.

The G×E heavily influences quantitative traits of economic importance in potatoes. Currently, up to 40 traits can be selected in potato-breeding programs (Bradshaw, 2017), where the low

correlation of the main traits between the environments results in a considerable loss of genetic gain (Andrade et al., 2021). This effect is significant for potato-breeding programs in tropical and subtropical regions as the crop is grown in different seasons throughout the year (winter, fall, and summer). Therefore, one goal of this program is the development of heat-tolerant clones by assessing promisor clones in contrasting seasons to determine their ability to withstand heat stress (Fernandes Filho et al., 2021; Patiño-Torres et al., 2021). Mixed model methodologies have an important role in connecting these different experiments and estimating parameters that are useful for the selection process (Henderson et al., 1959; Smith et al., 2001; Smith et al., 2015).

The limited number of repetitions, presence of G×E, and low heritability result in low selection accuracy in the early stages of a potato-breeding program. This implies low genetic progress over the selection cycles once accuracy is directly proportional to expected gains with selection (Cobb et al., 2019). Using a genetic relationship matrix has been demonstrated to increase the accuracy of estimated breeding values for traits with low heritability. This approach leverage information from all relatives (half and full-sibs, parents, etc.) to accurately estimate the breeding values of candidates (Slater et al., 2014a). However, in estimating non-additive effects, such as dominance, you need a balanced mating design to capture general and specific combining ability (Amadeu et al., 2020; Voss-Fels et al., 2021; Yadav et al., 2021).

The prediction of the total genotypic value (additive + non-additive effects) for complex quantitative traits, such as tuber yield, in the initial stage of the potato-breeding programs, relies on the use of genomic resources (Stich and Van Inghelandt, 2018; Amadeu et al., 2020; Voss-Fels et al., 2021; Wilson et al., 2021; Yadav et al., 2021). Nevertheless, early-stage genotyping is more expensive than phenotyping, making unfeasible use of genomic selection in some research, especially in stages where many candidates were evaluated (Stich and Van Inghelandt, 2018; Wilson et al., 2021; Bradshaw, 2022). Alternatively, Piepho et al. (2008) argued that the use of models with nested structure (Family/Clone = Family + Family + Clone) could be advantageous. These models account implicitly for the kinship relationship. Furthermore, using this structure allows us to predict the total genotypic value more easily, without the need for a kinship matrix, which can be valuable in cases where the mating design does not allow an accurate estimation of the specific combining ability.

The dominance effect can be estimated with a kinship matrix using complete mating design or using genomics. A third alternative is the modeling of the family effect. Although the

nested structure appears naturally in the initial stage of potato-breeding programs because seedlings are derived from different crosses (usually bi-parental), the family effect has been neglected (Fernandes Filho et al., 2021), mainly due to the easiness of mass selection in the early stages of selection. Therefore, it is hypothesized that using nested structure models can increase the selective efficiency of clones, both in single-trial (ST) and multi-environment-trial (MET) selection schemes.

Thus, this work aims to investigate the impact of the family effect and the selection accuracy of clones in the initial stage of a tetraploid potato-breeding program in ST and MET clonal selection schemes.

2 Materials and methods

2.1 Field trials

2.1.1 Experimental designs and crop management

A total of six trials from the potato-breeding program at the Universidade Federal de Lavras (PROBATATA-UFLA) were installed at the Center for Scientific and Technological Development, City Lavras, Minas Gerais State, Brazil (21°12'19.8" S, 44°58'48.8" W), located at 919 m American Sign Language (ASL), and soil was classified as red-yellow latosol.

The details of each trial are shown in Table 1. Four trials were designed in an ABD (Federer, 1956), and two trials a P-REP with p_N around 20% were employed (Cullis et al., 2006).

Each plot consisted of five plants spaced 0.30 m between plants and 0.80 m between rows. Crop management practices for all the trials were done according to the recommendations for the state of Minas Gerais, in which 1.5 Mg ha⁻¹ of 08-28-16 fertilizer blend (N-P₂O₅-K₂O) was applied during the planting. Side dress fertilizer application was performed with 0.30 Mg ha⁻¹ 20-00-20 (N-P₂O₅-K₂O). All the trials were irrigated using a sprinkler irrigation system, according to the need of the crop and the incidence of rainfall through the seasons.

2.1.2 Levels of heat stress

The trials were evaluated in three seasons with different levels of heat stress: without heat stress (WHS), moderate heat stress (MHS), and high heat stress (HHS) (Figure 1). The trials WHS (season from May to September), MHS (February to May), and HHS (November to February) were carried out during winter, fall, and summer seasons, respectively (Figure 1; Table 1).

The compensated mean temperatures (TMEAN, °C) were obtained using the expression $TMEAN = (T9am + 2T9pm + TMAX + TMIN)/5$, where T9am, T9pm, TMAX, and TMIN are the air temperature at 9 a.m., 9 p.m., maximum, and minimum, respectively (INMET, 2022).

2.1.3 Phenotyping

Two traits were evaluated in each trial: total tuber yield (TTY – Mg ha⁻¹) and specific gravity (SG). Total tuber yield was estimated by the weight of all tubers harvested in 1.2 m² for each plot. The SG was estimated by the expression $SG = \text{tuber mass in air} / (\text{tuber mass in the air} - \text{tuber mass in water})$, where tuber mass in air and tuber mass in water were measured from fresh samples of tubers, ranging from 2.0 kg and 2.5 kg, using a hydrostatic scale (Schippers, 1976).

2.2 Statistical analysis

A nested genetic treatment structure was evaluated. The clones were obtained from different clonal families. Thus, it is possible to assess family and clone within-family effects from the data.

For the analysis, in which c clones were sampled from s clonal families and evaluated together with p checks, the general form of the linear mixed model is presented in Equation (1). This model is suitable for both ST or MET data. For MET data analysis, appropriate (co)variance structures should be used to model the vectors of the family ($\mathbf{u}_s$) and clone within-family ($\mathbf{u}_c$) effects, aiming to account for the G×E:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_o\tau_o + \mathbf{Z}_s\mathbf{u}_s + \mathbf{Z}_c\mathbf{u}_c + \mathbf{Z}_b\mathbf{u}_b + \mathbf{e} \quad (1)$$

TABLE 1 Trials characterization, size, and number of blocks, family, clone, check, and percentage of plots replicate.

Trial [†]	Year	Design [†]	Number of levels					$p_N^{\S}$
			Block	Family	Clone	Check	Plot	
POP1(WHS)	2013	ABD	48	24	477	3	621	22.71
POP2(WHS)	2017	ABD	20	31	491	2	531	7.16
POP2(MHS)	2017	ABD	20	31	491	2	531	7.16
POP2(HHS)	2018	ABD	20	31	491	2	531	7.16
POP3(WHS)	2021	P-REP	20	12	304	4	400	23.00
POP3(HHS)	2021	P-REP	20	12	312	3	400	21.25

[†]The trial identification: POP1(WHS), POP2(WHS), POP2(MHS), POP2(HHS), POP3(WHS), and POP3(HHS), where the codes POP1, POP2, and POP3 identify different clonal populations and codes WHS, MHS, and HHS identify three different seasons, varying in the function of stress level: without heat stress (WHS) moderate heat stress (MHS), and high heat stress (HHS).

[‡]Experimental designs: augmented block design (ABD) and partially replicated design (P-REP).

[§] p_N : percentage of plots experimental units occupied by replicated clones, given by the expression. $p_N = (N - N_{\text{treat}})/N$, where N is the number of plots and N_{treat} is the number of treatments.

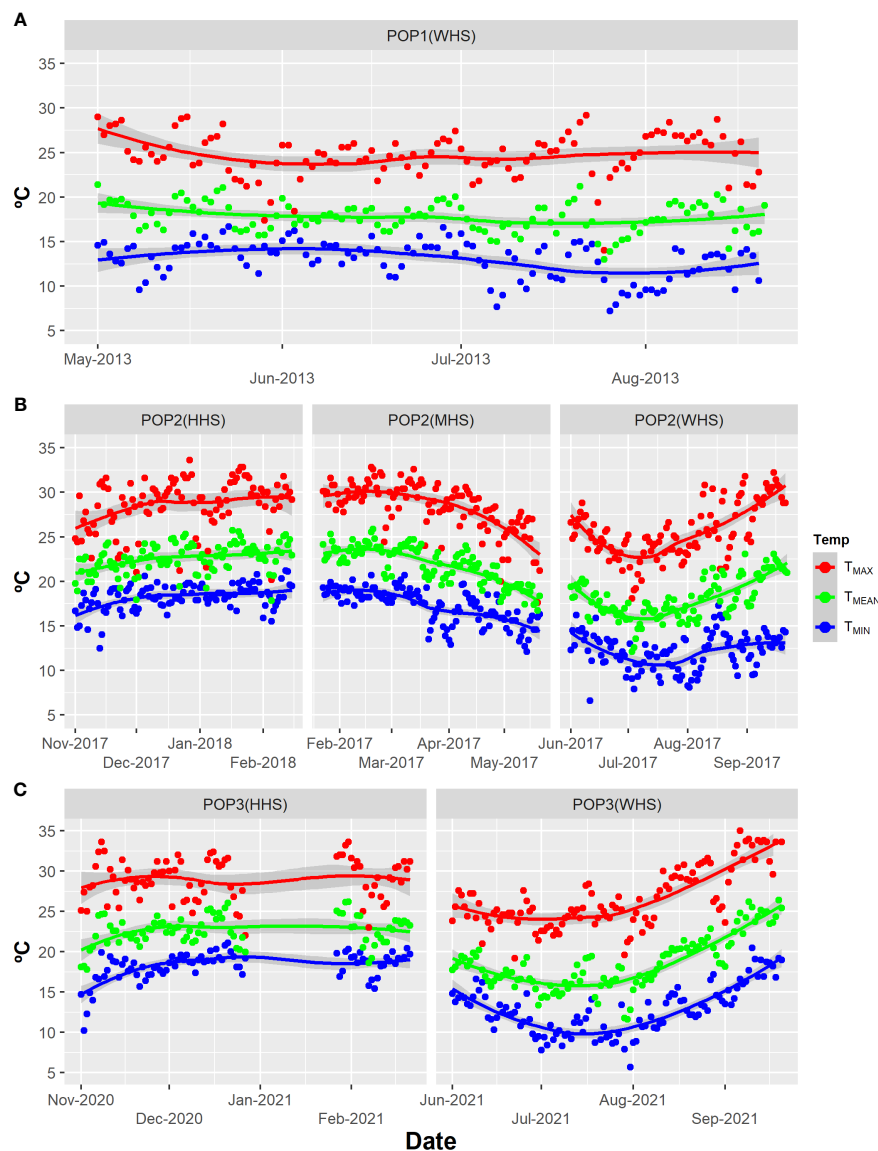


FIGURE 1

Temp: Maximum temperature (T_{MAX}), minimum temperature (T_{MIN}), and compensated mean temperature (T_{MEAN}) in degrees Celsius for all trials: (A) POP1(WHS), ranging from May to August of 2013; (B) POP2(WHS), ranging from June to September 2017; POP2(MHS), ranging from February to May 2017; POP2(HHS), ranging from November 2017 to February 2018; POP3(WHS), ranging from June to September 2021; and (C) POP3(HHS), ranging from November 2020 to February 2021 for Lavras, Minas Gerais State, Brazil (INMET, 2022). Seasons varying in terms of stress level: without heat stress (WHS) moderate heat stress (MHS), and high heat stress (HHS) from 2013 to 2021.

where $\mathbf{y}^{(N \times 1)}$ is the vector of phenotypic observations, where N is the number of plots for ST data or plots by seasons for MET data; $\mathbf{1}^{(N \times 1)}$ is a vector in which all elements are unity; $\mu^{(1 \times 1)}$ is the intercept; $\tau_o^{(o \times 1)}$ is the vector of fixed effects, composed of check, environment, and check by environment interaction effects (the environment and check by environment interaction effects were only used in MET data analysis), associated with matrix the design $\mathbf{X}_o^{(N \times o)}$ (assuming full rank), where o is the number of fixed effects; $\mathbf{u}_s^{(s \times 1)}$ is the vector of random effects of family associated with the design matrix $\mathbf{Z}_s^{(N \times s)}$, where s is the number of families for ST data or families by seasons for MET data; $\mathbf{u}_c^{(cs \times 1)}$ is the vector of random genotypic effects of clone within-family associated with the design matrix $\mathbf{Z}_c^{(N \times cs)}$, where c is the number of clone within-

family for ST data or clone within-family by seasons for MET data; $\mathbf{u}_b^{(b \times 1)}$ is the vector of random effects of block associated with the design matrix $\mathbf{Z}_b^{(N \times b)}$, where b is the number of blocks for ST data or blocks by seasons for MET data; and $\mathbf{e}^{(N \times 1)}$ is the vector of random errors.

We assume that the $\mathbf{u}_c$, $\mathbf{u}_s$, $\mathbf{u}_b$, and $\mathbf{e}$ vectors of random effects are mutually independent and distributed as multivariate Gaussian, with zero means and (co)variance matrices $\text{var}(\mathbf{u}_c) = \mathbf{G}_c$, $\text{var}(\mathbf{u}_s) = \mathbf{G}_s$, $\text{var}(\mathbf{u}_b) = \mathbf{G}_b$, and $\text{var}(\mathbf{e}) = \mathbf{R}$. The structures of these (co)variance matrices are shown in Table 2 for ST (STMpF and STMwF) and MET (METMpF and METMwF) data analysis models, including or not the family effect, respectively. For STMwF and METMwF models, the vector of clone effects was

TABLE 2 Summary of models fitted: single-trial model without family effect (STMwF), single-trial model plus family effect (STMpF), multi-environment-trial model without family effect (METMwF), and multi-environment-trial model plus family effect (METMpF).

(Co)variance matrix	Single-trial analysis		Multi-environment-trial analysis	
	STMwF	STMpF	METMwF	METMpF
$\mathbf{G}_b$	$\sigma_b^2 \mathbf{I}_b$	$\sigma_b^2 \mathbf{I}_b$	$\bigoplus_{j=1}^t \sigma_{bj}^2 \mathbf{I}_{bj}$	$\bigoplus_{j=1}^t \sigma_{bj}^2 \mathbf{I}_{bj}$
$\mathbf{G}_s$		$\sigma_s^2 \mathbf{I}_s$		$\mathbf{G}_{ts} \otimes \mathbf{I}_s$
$\mathbf{G}_c$	$\sigma_c^2 \mathbf{I}_c \otimes \mathbf{I}_s$	$\sigma_c^2 \mathbf{I}_c \otimes \mathbf{I}_s$	$\mathbf{G}_{tc} \otimes \mathbf{I}_c \otimes \mathbf{I}_s$	$\mathbf{G}_{tc} \otimes \mathbf{I}_c \otimes \mathbf{I}_s$
$\mathbf{R}$	$\sigma^2 \mathbf{I}_N$	$\sigma^2 \mathbf{I}_N$	$\bigoplus_{j=1}^t \sigma_j^2 \mathbf{I}_{Nj}$	$\bigoplus_{j=1}^t \sigma_j^2 \mathbf{I}_{Nj}$

$\mathbf{G}_b$, $\mathbf{G}_s$, $\mathbf{G}_c$, and $\mathbf{R}$: (co)variance matrices associated with the block, family, clone within-family, and error effects, respectively; $\mathbf{G}_{ts}$, $\mathbf{G}_{tc}$, and $\mathbf{G}_{te}$: unstructured (co)variance matrices used to accommodate the G×E interaction for the family, clone, and clone within-family effects, respectively; σ_b^2 , σ_{bj}^2 , σ_s^2 , σ_c^2 , σ_{cj}^2 , σ^2 , and σ_j^2 : variance components associated with the block, block in each trial, family, clone, clone within-family, error, and error in each trial effects, respectively; $\mathbf{I}_b$, $\mathbf{I}_{bj}$, $\mathbf{I}_s$, $\mathbf{I}_c$, $\mathbf{I}_{cj}$, $\mathbf{I}_N$, and $\mathbf{I}_{Nj}$: identity matrices associated with the block, block in each trial, family, clone within-family, error, and error in each trial effects, respectively; $\otimes$: Kronecker product operator; $\bigoplus$: direct sum operator.

called $\mathbf{u}_c$. For the MET data analysis, the heterogeneity of variances of block and error effects was accommodated by the direct sum operation ($\bigoplus$), whereas the heterogeneity of the variances and covariances of the G×E interaction for family and clone within-family effects was accommodated by the direct product operation ($\otimes$). The (co)variance matrices of environments for family and clone within-family effects were modeled by an unstructured matrix with $t(t + 1)/2$ covariance parameters, where t is the number of trials.

The covariance parameters of models shown in Table 2 were estimated by the residual maximum likelihood (REML) method (Patterson and Thompson, 1971) and best linear unbiased predictions (BLUP) of the random effects by Henderson's mixed model equations (Henderson et al., 1959) through the software Echidna Mixed Models (Gilmour, 2021) version 1.61. The graphic plots and other analyses discussed in the following sections were performed using the R software (R Core Team, 2021) and R packages base and ggplot2 (Wickham, 2016).

2.2.1 Single-trial analysis

From the linear mixed model for STMpF presented in Table 2, the vector of random total genotypic effects of clones can be predicted for ST analysis ($\mathbf{u}_{gST}$) by combining the vectors of family ($\mathbf{u}_s$) and clone within-family ($\mathbf{u}_c$) effects as presented in Equation (2). The (co)variance structure of the $\mathbf{u}_{gST}$ vector is given by the composite symmetry (CS) form as shown in Equation (3):

$$\mathbf{u}_{gST} = (\mathbf{I}_c \otimes \mathbf{I}_s) \mathbf{u}_s + \mathbf{u}_c \quad (2)$$

$$\text{var}(\mathbf{u}_{gST}) = (\sigma_s^2 \mathbf{J}_c + \sigma_c^2 \mathbf{I}_c) \otimes \mathbf{I}_s \quad (3)$$

where $\mathbf{I}_c$ is a vector in which all elements are unity, $\mathbf{J}_c$ is a matrix in which all elements are unity, and σ_s^2 and σ_c^2 are variance components of family and clone within-family effects, respectively.

The simple reparameterization of Equation (3), for correlation scale, allows obtaining ρ_s correlation by Equation (4). This correlation ranges from 0 to 1 (assuming $\sigma_s^2 > 0$ and $\sigma_c^2 > 0$) and measures the proportion of total genetic variance due to variation among families.

$$\rho_s = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_c^2} \quad (4)$$

The vector of clone effects ($\mathbf{u}_c$) from STMwF model was also predicted. However, for STMwF, both the variance component and the BLUP of clones are confounded with family effect. Thus, the comparison between STMpF and STMwF models in terms of accuracy of total genotypic values of clones is inadequate if $\sigma_s^2 > 0$ (Supplementary Information, Note S1). In this context, the STMpF and STMwF models were compared using the Akaike information criterion (AIC) presented in Equation (5) (Akaike, 1974) and through the correspondence of the top 20% best clones by Czekanowski coefficient (Qiao et al., 2000) [CC, Equation (6)], as well as ranking concordance through the Spearman correlation coefficient (r_s) between $\mathbf{u}_c$ and $\mathbf{u}_{gST}$. Furthermore, the variance components of the STMwF and STMpF models were tested using the likelihood ratio test (LRT) shown in Equation (7):

$$\text{AIC} = -2\ell + 2p \quad (5)$$

$$\text{CC} = a/(a + b) \quad (6)$$

$$\text{LRT} = -2\ln(\ell_1/\ell_2) \quad (7)$$

where ℓ is the maximum point of residual log-likelihood function, p is the number of variance parameters, a is the number of coincident clones by both selection strategies, b is the number of divergent clones by both selection strategies, ℓ_1 is the maximum point of residual log-likelihood function from reduced model (without the effect tested), and ℓ_2 is the maximum point of residual log-likelihood function from complete model.

Although the variance component σ_c^2 represents the average within-families genotypic variance, the BLUP of the clone within-family effect is coded to the overall mean and adjusted for the family structure, which allows for comparison of clones from different families (Supplementary Information, Note S1). Thus, we also compared the selection by $\mathbf{u}_c$ and $\mathbf{u}_{gST}$ vectors through the correspondence of the top 20% best clones by Czekanowski coefficient [Equation (6)], as well as ranking concordance through the Spearman correlation coefficient (r_s).

To facilitate the visualization of the results, the variance components of the STMwF and STMpF models have been presented at percentage of total variation (sum of all variance components of ST analysis of the STMwF or STMpF model). To compare the efficiency of the selection strategies based on total genotypic ($\mathbf{u}_{gST}$) and clone within-family ($\mathbf{u}_c$) effects from STMpF, we assessed the efficiency through the accuracy ratio of respective effects.

2.2.2 Multi-environment-trial analysis

The models METMwF and METMpF are extensions of the models STMwF and STMpF for MET data, respectively (Table 2). Similarly, to what was done for the STMpF model, the vector of total genotypic effects may also be predicted for the MET analysis using the METMpF model, combining the vectors of family ($\mathbf{u}_s$) and clone within-family ($\mathbf{u}_c$) effects as presented in Equation (8). However, unlike Equation (2), Equation (8) capitalizes the $G \times E$. The (co)variance structure of vector $\mathbf{u}_{gMET}$ is given by multivariate compound symmetry form as shown in Equation (9):

$$\mathbf{u}_{gMET} = (\mathbf{I}_c \otimes \mathbf{I}_s \otimes \mathbf{I}_t) \mathbf{u}_s + \mathbf{u}_c \quad (8)$$

$$\text{var}(\mathbf{u}_{gMET}) = (\mathbf{G}_{ts} \otimes \mathbf{J}_c + \mathbf{G}_{tc} \otimes \mathbf{I}_c) \otimes \mathbf{I}_s \quad (9)$$

where $\mathbf{I}_t$ is an identity matrix of trials and $\mathbf{G}_{ts}$ and $\mathbf{G}_{tc}$ are (co) variance matrices of trials for family and clone within-family effects.

A similar vector to the vector $\mathbf{u}_{gMET}$ may also be predicted from the METMwF model, which was called $\mathbf{u}_c$. However, because of the reasons highlighted in Section 2.2.1, the models METMpF and METMwF models were compared using the AIC presented in Equation (5) (Akaike, 1974) and through the correspondence of the top 20% best clones by Czekanowski's coefficient (CC) (Qiao et al., 2000) [Equation (6)], as well as ranking concordance through the Spearman correlation coefficient (r_s) based on FAI-BLUP index score (Rocha et al., 2018), between the strategies ($\mathbf{u}_{gMET}$ vs. $\mathbf{u}_c$).

Because of the difficulty of performing the LRT test for parameters of the unstructured matrices ($\mathbf{G}_{tc}$, $\mathbf{G}_{ts}$, and $\mathbf{G}_{tc}$), two 95% confidence intervals were used for parameters of the METMwF and METMpF models. The first one, based on Chi-Square distribution (SAS Institute, 2016), used the variance components for family and clones within-family. The second is based on the Normal distribution (Meyer, 2008) for the correlations between pairs of environments for family and clone within-family effects (Supplementary Information, Note S2).

The genotypic correlations between the environment pairs for clone ($\rho_{G_c'}$), family (ρ_{G_s}), and clone within-family (ρ_{G_c}) effects were estimated from parameters of matrices $\mathbf{G}_{tc}$, $\mathbf{vG}_{ts}$, and $\mathbf{G}_{tc}$ using the expressions (10), (11), and (12):

$$\rho_{G_c'ij} = \frac{\sigma_{c'ij}}{\sqrt{\sigma_{c'i}^2 \times \sigma_{c'j}^2}} \quad (10)$$

$$\rho_{G_{sij}} = \frac{\sigma_{sij}}{\sqrt{\sigma_{s_i}^2 \times \sigma_{s_j}^2}} \quad (11)$$

$$\rho_{G_{cij}} = \frac{\sigma_{cij}}{\sqrt{\sigma_{c_i}^2 \times \sigma_{c_j}^2}} \quad (12)$$

where $\sigma_{c'ij}$, σ_{sij} , and σ_{cij} are covariances between environments pairs i and j for clone, family, and clone within-family effects; $\sigma_{c'i}^2$ and $\sigma_{c'j}^2$ are variance components of environments i and j for clone effect; $\sigma_{s_i}^2$ and $\sigma_{s_j}^2$ are variance components of environments i and j for family effect; and $\sigma_{c_i}^2$ and $\sigma_{c_j}^2$ are variance components of environments i and j for clone within-family effect.

The visualization of the results and the variance components of the models METMwF and METMpF were presented in percentage of total variation for each trial (sum of all variance components for each trial of the MET analysis of the model METMwF or METMpF). To compare the effectiveness of the selection strategies based on total genotypic ($\mathbf{u}_{gMET}$) and clone within-family ($\mathbf{u}_c$) effects from METMpF, we assessed the effectiveness through the accuracy ratio of respective effects.

Finally, we used the FAI-BLUP index (Rocha et al., 2018), to rank the clones based on predicted BLUPs of MET data for both strategies ($\mathbf{u}_{gMET}$ vs. $\mathbf{u}_c$). The two selection strategies have been compared as described before. Conducted the exploratory factor analysis, with Factor Analysis (FA) and Principal Component Analysis (PCA). PCA was used to extract the factorial loads from the genetic correlation matrix, obtained by the predicted values (BLUPs). The analysis used is varimax criterion (Kaiser, 1958) for the analytic rotation and the calculation of the factor scores of the weighted least squares method (Bartlett, 1938). Thus, PCA and FA were performed on the set of BLUP mean for six variables from population 2, three seasons (POP2HHS, POP2MHS, and POP2WHS), and two traits (TTY and SG) of each vector $\mathbf{u}_c$, $\mathbf{u}_c$, and $\mathbf{u}_{gMET}$. They were estimated from multi-environment-trial model without the family effect (METMwF) and multi-environment-trial model plus the family effect (METMpF).

2.2.3 Accuracy of family, clone within-family, total genotypic effects, and relative efficiency

The accuracy of family (r_{ss}), clone within-family (r_{cc}), total genotypic effects (r_{gg}), and relative efficiency (RE) were obtained by expressions (13), (14), (15), and (16) for both ST and MET data analysis:

$$r_{ss} = \sqrt{1 - \frac{v_s}{\sigma_s^2}} \quad (13)$$

$$r_{cc} = \sqrt{1 - \frac{v_c}{\sigma_c^2}} \quad (14)$$

$$r_{gg} = \sqrt{1 - \frac{v_g}{\sigma_s^2 + \sigma_c^2}} \quad (15)$$

$$RE = \frac{r_{gg}}{r_{cc}} \quad (16)$$

where v_s , v_c , and v_g are the average prediction error variance of family, clone within-family, and total genotypic effects, respectively.

The RE was used to measure the difference between the proposed models to ST and MET data.

3 Results

3.1 Comparison of the STMwF and STMpF models

The inclusion of the family effect improved the goodness of fit of the models in all trials, on which the STMpF model showed lower AIC than the STMwF, and the only exception was for TTY from POP3(WHS) (Table 3). In addition, the inclusion of family effect also increased the log-likelihood (ℓ) in all trials for both traits. Only for the TTY trait in the POP3(WHS) trial did this increment does not exceed 1.92 units (critical point for the detection of significant

σ_s^2 effect) (Table 3). These results are reinforced by the results of the LRT test (Table 4) and indicate the presence of genetic variability between families in most trials.

Clone (σ_c^2 , STMwF) and clone within-family ($\sigma_{c'}^2$, STMpF) variances were significant in all trials for both traits (Table 4; Supplementary Material, Table S2), revealing the existence of genetic variability among clones. The contribution of clone within-family variance (C) to the total phenotypic variance was always lower than the clone effect (C') (Table 4). However, the magnitude of the difference between C' and C was directly proportional to the contribution of family variance to the total genetic variation (ρ_s). The amplitude of ρ_s for the SG trait (0.09 and 0.35) exceeded the amplitude for the TTY trait (0.03 and 0.26) (Table 4).

CC and Spearman's correlation coefficient (r_s) were used as comparison criterion for both selection strategies tested (u_c vs. u_{gst} and u_c vs. u_{gst}). It was observed that, regardless of the selection strategy

TABLE 3 Log-likelihood residual (ℓ) and Akaike information criterion (AIC) for single-trial model without family effect (STMwF) and single-trial model plus family effect (STMpF), for all trials and traits.

Trait [†]	Model	Trial [†]	ℓ	AIC
TTY	STMwF	POP1(WHS)	-1,826.81	3,659.63
		POP2(WHS)	-1,459.44	2,924.89
		POP2(MHS)	-1,584.44	3,174.88
		POP2(HHS)	-1,191.74	2,389.47
		POP3(WHS)	-1,031.07	2,068.14
		POP3(HHS)	-1,155.63	2,317.27
		POP1(WHS)	-1,820.36	3,648.72
	STMpF	POP2(WHS)	-1,450.92	2,909.84
		POP2(MHS)	-1,581.43	3,170.85
		POP2(HHS)	-1,189.63	2,387.26
		POP3(WHS)	-1,030.43	2,068.85
		POP3(HHS)	-1,149.94	2,307.87
SG	STMwF	POP1(WHS)	2,587.78	-5,169.57
		POP2(WHS)	2,200.75	-4,395.51
		POP2(MHS)	2,262.34	-4,518.68
		POP2(HHS)	1,822.53	-3,639.06
		POP3(WHS)	1,278.39	-2,550.77
		POP3(HHS)	1,591.35	-3,176.70
		POP1(WHS)	2,613.73	-5,219.45
	STMpF	POP2(WHS)	2,209.84	-4,411.68
		POP2(MHS)	2,276.28	-4,544.56
		POP2(HHS)	1,833.07	-3,658.15
		POP3(WHS)	1,281.73	-2,555.45
		POP3(HHS)	1,594.28	-3,180.55

[†]Total tuber yield (TTY; Mg ha⁻¹) and specific gravity (SG).

[†]The trial identification: POP1(WHS), POP2(WHS), POP2(MHS), POP2(HHS), POP3(WHS), and POP3(HHS), where the codes POP1, POP2, and POP3 identify the different clonal populations and the codes WHS, MHS, and HHS identify three different seasons, varying in function of stress level: without heat stress (WHS) moderate heat stress (MHS), and high heat stress (HHS).

TABLE 4 Contribution (%) of the variances of block, clone (C'), family (F), clone within-family (C), and residuals (Res.) for the phenotypic variance of the traits total tuber yield (TTY; Mg ha⁻¹) and specific gravity (SG) estimated from single-trial model without family effect (STMwF) and single-trial model plus family effect (STMpF) in different seasons.

Trait	Trial [†]	STMwF			STMpF				
		Block	Clone (C')	Res.	Block	Family (F)	Clone (C)	Res.	ρ_S [‡]
TTY	POP1(WHS)	2.67 ^{ns}	40.97**	56.35	2.05 ^{ns}	7.18**	34.42**	56.35	0.17
	POP2(WHS)	7.88**	47.78**	44.34	5.50**	11.56**	35.56*	47.38	0.26
	POP2(MHS)	3.88**	37.74*	58.38	3.77**	4.43**	32.62*	59.18	0.12
	POP2(HHS)	0.22 ^{ns}	69.55**	30.23	0.19 ^{ns}	4.38*	65.25**	30.18	0.06
	POP3(WHS)	7.13**	77.27**	15.61	7.26**	2.32 ^{ns}	74.90**	15.52	0.03
	POP3(HHS)	1.09 ^{ns}	68.44**	30.47	1.61 ^{ns}	8.76**	59.62**	30.01	0.13
SG	POP1(WHS)	9.83**	69.87**	20.30	7.69**	18.56**	53.15**	20.60	0.26
	POP2(WHS)	12.76**	35.00*	52.24	5.03 ^{ns}	15.11**	28.43*	51.43	0.35
	POP2(MHS)	2.26*	43.40*	54.34	2.10 ^{ns}	13.60**	29.94*	54.36	0.31
	POP2(HHS)	3.91**	56.78**	39.31	3.57*	12.97**	44.10**	39.36	0.23
	POP3(WHS)	3.29*	57.63**	39.09	3.42*	6.75**	49.88**	39.96	0.12
	POP3(HHS)	8.01**	58.58**	33.41	7.97**	5.04**	52.42**	34.56	0.09

[†]The trial identification: POP1(WHS), POP2(WHS), POP2(MHS), POP2(HHS), POP3(WHS), and POP3(HHS), where the codes POP1, POP2, and POP3 identify the different clonal populations and the codes WHS, MHS, and HHS identify three different seasons, varying in function of stress level: without heat stress (WHS) moderate heat stress (MHS), and high heat stress (HHS).

[‡] ρ_S correlation: measures the proportion of total genetic variance due to variation among families.

Significance by the likelihood-ratio test (LRT): p-value < 0.01 “**” and 0.05 “*” and p-value > 0.05 not significant “ns”.

and trait, both coefficients showed an inverse relationship with ρ_S , suggesting that an increase in genetic variability among families reduces the similarity of the $\mathbf{u}_C$ and $\mathbf{u}_C$ vectors with the $\mathbf{u}_{gST}$ vector in the ranking of clones. Furthermore, the magnitude of the correlations of the CC and

r_S coefficients with ρ_S was higher for the TTY (−0.82 and −0.78) compared with that for the SG (−0.72 and −0.71) (Figure 2).

In general, both CC and r_S showed lower magnitude for the second selection strategy with family effect ($\mathbf{u}_C$ vs. $\mathbf{u}_{gST}$), which

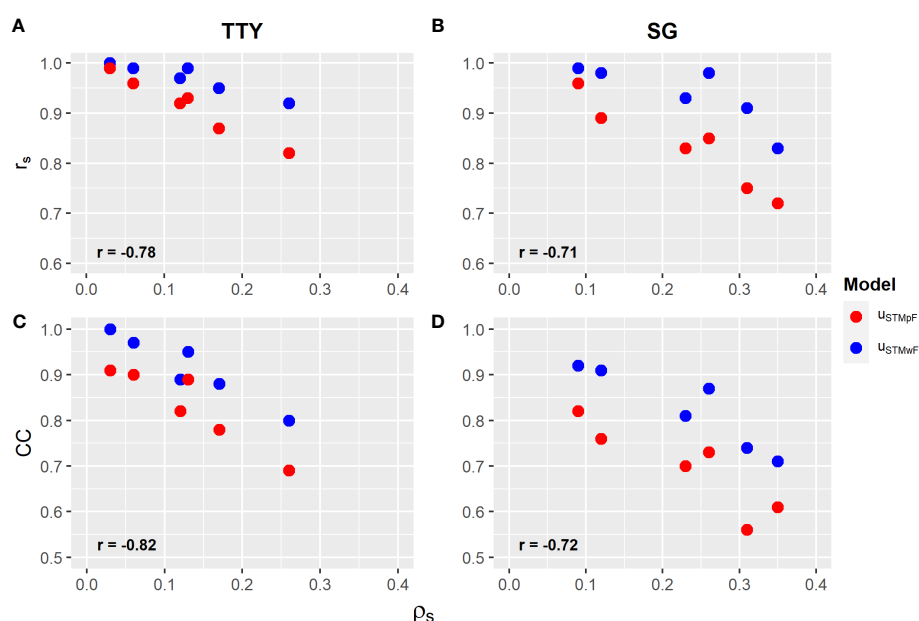


FIGURE 2

(A, B) Spearman correlation coefficient (r_S), and (C, D) Czekanowski coefficient (CC) (Supplementary Material, Tables S3, S4) between the proportion of total genetic variance of families (ρ_S) (Table 4) for the vector of clone effects of models from the single trial without family (U_{STMwF}) and plus family effect (U_{STMPF}) to the total tuber yield (TTY; Mg ha⁻¹) and specific gravity (SG). Labels in bold are correlations between selective efficiency and ρ_S .

suggests a greater agreement of the $\mathbf{u}_c$ vector without family effect with $\mathbf{u}_{gst}$ in the clone ranking (Figure 2; Supplementary Material, Tables S4, S5). Furthermore, independent of the selection strategy, an increment in CC and r_s can be observed with increasing heat stress for both traits in POP2 and POP3, except for the TTY in population POP3 (Supplementary Material, Tables S4, S5).

3.2 Relative efficiency of selection based on total genotypic effect of clone for ST analysis

Regardless of the trait, selection based on the vector of total genotypic effects of clone ($\mathbf{u}_{gst}$) was found to be greater than the selection based on the vector of clone effects within family ($\mathbf{u}_c$) (Table 5). The efficiency was directly proportional to ρ_s , showing that the increment in genetic variability among families increases the selective efficiency of clones. The correlation between efficiency and ρ_s was higher for SG (0.88) when compared with that for TTY (0.72) (Figure 3).

On average, the accuracies r_{ss} , r_{cc} , and r_{gg} were higher for SG (0.80, 0.70, and 0.78, respectively) compared with that for TTY (0.71, 0.69 and 0.74), respectively. The difference between the accuracies r_{gg} and r_{cc} was higher for SG (0.78 and 0.70) than that for TTY (0.74 and 0.69), which resulted in higher average selective efficiency for the SG (11%) compared with that for TTY (7%) (Table 5).

3.3 Comparison of the METMwF and METMpF models

Variance component associated with family effect in the trial POP3(WHS) was not significant by the LRT test for TTY (Table 4) and, thus, was not included in the MET analysis. The family effect (METMpF) did not improve the goodness of fit compared with the METMwF model for POP3 (Table S5); thus, the MET analysis results were presented only for POP2 (Table 6).

The variance estimates associated with the effects of clone (σ_c^2 , METMwF), family (σ_s^2 , METMpF), and clone within-family (σ_c^2 , METMpF), as well as their respective contributions to the phenotypic variance, were similar to those obtained in the ST analyses in all trials and traits. The variance components associated with clone, family, and clone within-family were higher than zero in all scenarios (Table S6), confirming the results found in the ST analyses (Tables 4, 6; Supplementary Material, Tables S6, S7). It is worth noting that the values of ρ_s were also like those obtained in the ST analyses (Tables 4, 6).

Overall, variation was observed in the estimates of variance components σ_c^2 , σ_s^2 , and σ_c^2 across the different seasons for both traits, indicating that populations that have G×E can be attributed to the interaction of a simple nature (Supplementary Material, Table S8). Genetic correlation between seasons for the effects of clone, family, and clone within-family was positive in all scenarios for SG, with values higher than 0.50 in most cases. This suggests a low contribution of complex type G×E for this trait. In contrast, most genetic correlation estimates for TTY were lower than 0.50,

TABLE 5 Accuracy of the family (r_{ss}), clone within-family (r_{cc}), and total genotypic (r_{gg}) effects for traits total tuber yield (TTY; Mg ha⁻¹) and specific gravity (SG) estimated from single-trial model plus family effect (STMPF) in different seasons.

Trait	Trial [†]	r_{ss}	r_{cc}	r_{gg}	Efficiency [‡]
TTY	POP1(WHS)	0.74	0.60	0.67	1.12
	POP2(WHS)	0.76	0.64	0.71	1.11
	POP2(MHS)	0.63	0.58	0.63	1.09
	POP2(HHS)	0.60	0.81	0.83	1.02
	POP3(WHS)		0.90 [§]		
	POP3(HHS)	0.81	0.81	0.84	1.02
	Average	0.71	0.69	0.74	1.07
SG	POP1(WHS)	0.87	0.82	0.87	1.06
	POP2(WHS)	0.80	0.58	0.71	1.22
	POP2(MHS)	0.82	0.58	0.71	1.22
	POP2(HHS)	0.79	0.70	0.78	1.11
	POP3(WHS)	0.77	0.75	0.78	1.04
	POP3(HHS)	0.73	0.78	0.80	1.03
	Average	0.80	0.70	0.78	1.11

[†]The trial identification: POP1(WHS), POP2(WHS), POP2(MHS), POP2(HHS), POP3(WHS), and POP3(HHS), where the codes POP1, POP2, and POP3 identify the different clonal populations and the codes WHS, MHS, and HHS identify three different seasons, varying in function of stress level: without heat stress (WHS) moderate heat stress (MHS), and high heat stress (HHS).

[‡]Relative efficiency: r_{gg}/r_{cc} ratio.

[§]Not included in estimate of average accuracy.

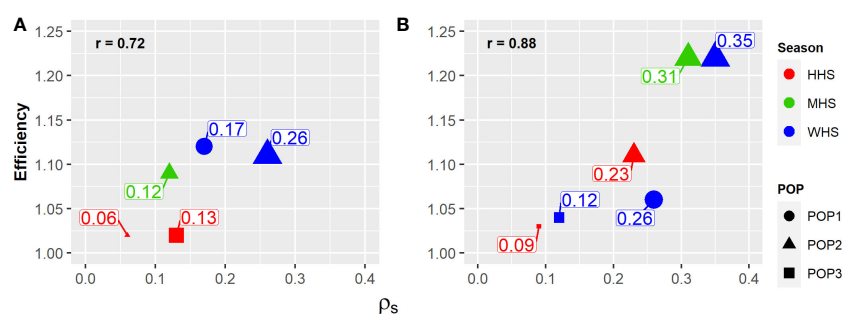


FIGURE 3

Family effect response (ρ_s) (size) on the selective accuracy of potato clones (relative efficiency) in the function of heat stress level (color): without heat stress (WHS), moderate heat stress (MHS), and high heat stress (HHS), in three populations evaluated (shape): POP1, POP2, and POP3. (A) Total tuber yield (TTY; Mg ha⁻¹) and (B) specific gravity (SG).

suggesting a greater contribution of the complex G×E (Supplementary Material, Tables S6, S7).

Regardless of the model and trait, the variance component estimates associated the effects of clone (σ_c^2) and clone within-family ($\sigma_{c'}^2$) were always higher in the higher heat stress season (HHS) when compared with that in the no heat stress season (WHS), indicating an increase in genetic variability under extreme heat stress (Supplementary Material, Table S1). The variance component associated with family effect (σ_f^2) showed a behavior inversely proportional to the increase of heat stress for the TTY trait, reducing about 50% with the increment of heat stress [11.32 (WHS), 7.32 (MHS), and 3.64 (HHS)] (Supplementary Material, Table S3). Furthermore, regardless of the model adopted, we recorded reductions of 42% and 3% in the average of the clones in the POP2 population for the traits TTY and SG under HHS, respectively (Supplementary Material, Table S3). It is worth noting that, considering the period from the beginning of tuberization (about 30 days after planting) until harvest, the average daily temperature exceeded 20°C on 64% of the days in the MHS season and 88% of the days in the HHS season (88%) (Figure 1).

The exploratory factor analysis showed mean communality of 0.91, 0.91, and 0.87 for the respective vectors $\mathbf{u}_c$, $\mathbf{u}_c$, and $\mathbf{u}_{gMET}$,

respectively, indicating that the three factors were sufficient to explain more than 87% of the relationship between seasons. Regardless of the effect, Factor 1 represented the three seasons for SG, Factor 2 represented the WHS and MHS seasons for TTY, and Factor 3 represented only the HHS season (Supplementary Material, Table S8).

After obtaining the FAI-BLUP index scores, which included both traits and all seasons, the Czarnowski's coefficients (CC) and Spearman's correlation (r_s) were utilized to compare the two selection strategies tested ($\mathbf{u}_c$ vs. $\mathbf{u}_{gMET}$ and $\mathbf{u}_c$ vs. $\mathbf{u}_{gMET}$). Similarly, to what was observed for the ST analysis, both CC and r_s showed lower magnitude for the second selection strategy ($\mathbf{u}_c$ vs. $\mathbf{u}_{gMET}$), which suggests closer concordance of the vector $\mathbf{u}_c$ with $\mathbf{u}_{gMET}$ in the ranking of the clones (Table 7).

3.4 Relative efficiency of selection based on the total genotypic effect of clone for MET analysis

Similar to the ST results, independent of trait or season, the selection based on the vector of genotypic effect of clones in the

TABLE 6 Contribution (%) of the variances of blocks, clones (C'), families (F), clone within-family (C), and residuals (Res.) for the phenotypic variance of the traits total tuber yield (TTY; Mg ha⁻¹) and specific gravity (SG) estimated from multi-environment-trial model without family effect (METMwF) and multi-environment-trial model plus family effect (METMpF) in different seasons.

Trait	Trial [†]	METMwF			METMpF				
		Block	Clone (C')	Res.	Block	Family (F)	Clone (C)	Res.	ρ_s [‡]
TTY	POP2(WHS)	8.75 ⁺	48.17 ⁺	43.09 ⁺	5.80 ⁺	11.26 ⁺	36.45 ⁺	46.48 ⁺	0.24
	POP2(MHS)	3.22 ⁺	38.34 ⁺	58.44 ⁺	3.00 ⁺	4.81 ⁺	33.08 ⁺	59.11 ⁺	0.13
	POP2(HHS)	0.00 [#]	70.02 ⁺	29.98 ⁺	0.00 [#]	4.38 ⁺	65.70 ⁺	29.92 ⁺	0.06
SG	POP2(WHS)	12.63 ⁺	33.33 ⁺	54.04 ⁺	6.37 ⁺	13.50 ⁺	28.03 ⁺	52.10 ⁺	0.33
	POP2(MHS)	1.53 ⁺	43.06 ⁺	55.41 ⁺	1.52 ⁺	13.79 ⁺	29.70 ⁺	54.99 ⁺	0.32
	POP2(HHS)	3.81 ⁺	56.31 ⁺	39.88 ⁺	3.47 ⁺	12.71 ⁺	44.13 ⁺	39.69 ⁺	0.22

[†]The trial identification: POP2WHS, POP2MHS, and POP2HHS, where the code POP2 identify the clonal population and codes WHS, MHS, and HHS identify three different seasons, varying in function of stress level: without heat stress WHS, moderate heat stress MHS, and high heat stress HHS.

[‡] ρ_s correlation: measures the proportion of total genetic variance due to variation among families.

⁺Variance component does not intercept the zero by 95% Chi-Squared confidence intervals.

[#]Variance component intercept the zero by 95% Chi-Squared confidence intervals.

TABLE 7 Czekanowski's coefficient (CC) and Spearman's correlation (r_s) between the FAI-BLUP index score vectors of different strategies (u_c vs. u_{gMET} and u_c vs. u_{gMET}), for traits total tuber yield (TTY; Mg ha⁻¹) and specific gravity (SG).

Strategies [†]	CC	r_s
u_c vs. u_{gMET}	0.94	0.98
u_c vs. u_{gMET}	0.82	0.92

[†] u_c is the vector of clone effects from multi-environment-trial model without family effect (METMwF), u_c is the vector of clone's effects from multi-environment-trial model with family effect (METMpF), and u_{gMET} is the vector of total genotypic effects of clones from METMpF.

MET (u_{gMET}) was higher than the selection based on the vector of clone within-family in the MET (u_c) (Table 8). It is worth noting that, because of the relationship between the season's pairs, the accuracy associated with the within-family clone effect was increased in all seasons for both traits, although the accuracy associated the family effect was maintained or was increased (Table 8).

For the TTY trait, efficiency estimates were progressively reduced with increasing heat stress levels [1.09 (WHS), 1.06 (MHS), and 1.02 (HHS)], although, for the SG trait, efficiencies were similar in all seasons. On average, the accuracies r_{ss} , r_{cc} , and r_{gg} were higher for the SG trait (0.82, 0.70, and 0.78) in comparison with the TTY trait (0.70, 0.71, and 0.75), respectively. Furthermore, the magnitude of the difference between the accuracies r_{gg} and r_{cc} was higher for the SG trait (0.78 and 0.70) detriment of TTY (0.75 and 0.71), which resulted in higher average selective efficiency for the SG trait (12%) when compared to the TTY trait (6%) (Table 8).

4 Discussion

4.1 ST analyses

The results presented here indicate that adjusting for the structure created by family effects can improve the estimation of

genetic values. However, it is important to highlight that phenotyping individual data in the early stages of the process may not be compatible with certain breeding pipelines. For example, in certain programs, the first-year evaluation takes place in the field, with mass selection, and using single-hill trials (one seed potato). In this case, the extra phenotyping and modeling of the data would incur additional resources, and the feasibility of such change needs to be evaluated on a case-by-case basis. One technology that potentially can be used to facilitate data collection in the early stages of the program would be the use of aerial images for phenotyping tubers that are dug out of the ground but left in the field (Matias et al., 2020). In contrast, in cases where the early stages are already established with larger plots and randomization of clones, the inclusion of the family effect in the quantitative genetics model would come at no additional effort in data collection.

The genetic progress obtained through conventional potato breeding is slow due to the complexity of the tetrasomic inheritance, the elevated level of heterozygosity of the genitors, and the large number of traits to be assessed (~40) (Haynes et al., 2012; Slater et al., 2014b; Gopal, 2015; Bradshaw, 2017). Potato-breeding programs generate thousands of seedlings annually to overcome these challenges (Stich and Van Inghelandt, 2018). Furthermore, initial stage of potato breeding has low availability of seeds because of the slow tuber multiplication rates, which restricts the use of repetitions and plot size (Paget et al., 2017). These factors contribute to lower experimental precision and selection accuracy.

Typically, the clones being evaluated come from various families, which are often created through biparental crossings. Therefore, the selection process involves making comparisons not only between clones from the same family but also between those from different families. Thus, the effect of clones, and its variance component, are confounded with the family effect. To address this problem and achieve higher selection accuracy, is the use of a nested model, on which the effect of clones is nested within the family effect (Family/Clone = Family + Family + Clone). According to Piepho et al. (2008), models with this structure can be advantageous as they

TABLE 8 Accuracy of family (r_{ss}), clone within-family (r_{cc}), and total genotypic (r_{gg}) effects for traits total tuber yield (TTY; Mg ha⁻¹) and specific gravity (SG) estimated from multi-environment-trial model plus family effect (METMpF) in different seasons.

Trait	Trial [†]	r_{ss}	r_{cc}	r_{gg}	Efficiency [‡]
TTY	POP2(WHS)	0.76	0.67	0.73	1.09
	POP2(MHS)	0.68	0.65	0.69	1.06
	POP2(HHS)	0.65	0.82	0.84	1.02
	Average	0.70	0.71	0.75	1.06
SG	POP2(WHS)	0.80	0.67	0.75	1.12
	POP2(MHS)	0.84	0.69	0.78	1.13
	POP2(HHS)	0.82	0.73	0.80	1.10
	Average	0.82	0.70	0.78	1.12

[†]The trial identification: POP2(WHS), POP2(MHS), POP2(HHS), where the code POP2 identify the clonal population and codes WHS, MHS, and HHS identify three different seasons, varying in function of stress level: without heat stress (WHS) moderate heat stress (MHS), and high heat stress (HHS).

[‡]Relative efficiency: r_{gg}/r_{cc} ratio.

implicitly account for the genetic relatedness of clones. It is important to highlight that including the family effect enhances the model performance due to a better representation of the families, once the clones within a family work as repetitions.

Resende et al. (2016) conducted a study on the effectiveness of the nested structure in a dry bean breeding program using progenies from various populations. They emphasized the significance of the population effect, as the use of the nested model resulted in 20 times more repetitions for each population compared to the progenies. Resulted from an increase in the selective efficiency of progenies. Some papers that considered the effect of populations are reported on bean (Resende et al., 2016; Paula et al., 2020) and soybean (Pereira et al., 2017; Volpato et al., 2018) crops. Thus, the nested structure allows obtaining a more accurate BLUP of the family, which enables a more reliable selection of the best clones. In addition, it allows for the selection of the best families, which enables the evaluation of a greater number of clones per family and increased plot size, thereby increasing the probability of obtaining superior clones.

The utilization of a nested structure presents a significant advantage in providing precise estimations of genetic parameters. This is because the variance components within and between families are not confounded, which can happen in models that do not take family effects into account (Supplementary Material, Note S1). In the former, the heritability of clones is overestimated due to the variance component of family, whereas, in the nested models, the family structure gives a better estimate of the clone effect (Supplementary Material, Note S1). The effects of clone and clone within-family are only equivalent when the variance component of families is zero or close to zero. Pereira et al. (2017) indicates that incorporating the population effect into the estimation of genetic and non-genetic components in soybean breeding data offers a more accurate and realistic methodology.

Higher accuracy was observed using the vector of total genotypic effects ($\mathbf{u}_{gST}$), obtained by combining the vectors of family effects ($\mathbf{u}_s$) and clone within-family ($\mathbf{u}_c$), and it was directly proportional to the contribution of the variance component associated with family in the total genetic variation (ρ_s). Resende et al. (2016) in a simulation study showed that the selection accuracy increases in function of the increased contribution of the between-population variance component to the total genetic variation. Therefore, although, in potato breeding, the genetic variance between families is lower than the genetic variance within families, the inclusion of the effect of family can increase the selection accuracy because the heritability of families has been, in general, higher than the heritability of clone within-family (Diniz et al., 2006; Melo et al., 2011). Greater average RE recorded for the SG trait is associated with the higher mean accuracy for the effect of families (0.80) and the higher ρ_s mean (0.23).

The utilization of the family effect in potato breeding, which models the nested effect of clones within the family, has been shown to be advantageous. This methodology has led to an improvement in the accuracy of predicting the genotypic values of clones and has achieved a greater degree of RE for the two traits that were studied. The increase in accuracy achieved by the nested model does not reflect increased costs in cases where the data are already recorded.

We expect that animal modeling properly accounting for additive and dominance effects would also result in more efficient results. However, the results showed that family effect inclusion is a simpler approach, mainly in cases where the mating design is not complete.

4.2 MET analyses

Understanding and effectively managing the G×E interaction is essential for achieving long-term improvements in plant breeding programs because the success of a new cultivar depends on its improved performance on different traits (e.g., TTY and SG) while also presenting good adaptability and stability. The G×E assumes a particular importance for potato breeding, mainly under tropical conditions, because heat stress limits yield and quality in the hottest periods of the year (Fleisher et al., 2017; Fernandes Filho et al., 2021; Patiño-Torres et al., 2021).

Potato crop generally presents better performance in regions with temperate climate, with average temperatures between 5°C and 21°C (Haverkort and Verhagen, 2008). With temperature exceeding 21°C, heat stress significantly reduces tuber yield and quality, due to a series of physiological changes, including an increase in tuber disorders (e.g., internal heat necrosis, knobs, and tuber chain), reduced plant growth, increased respiration rate, greater allocation of dry biomass to leaves at the expense of the tubers, and the reduction in photosynthetic pigments (Lambert et al., 2006; Haverkort and Verhagen, 2008; Hancock et al., 2014; Rykaczewska, 2015; Patiño-Torres et al., 2021). Thus, the selection of heat-tolerant clones is vital to increase potato tuber yield and quality.

Brazil potato season is carried out in three distinct seasons: dry (January to March), winter (April to July), and water (August to December) seasons. Temperatures above the critical threshold, 21°C, are commonly recorded in the dry and water seasons (Andrade et al., 2021). Fernandes Filho et al. (2021) classified these seasons according to their level of heat stress: MHS, WHS, and HHS. One of the major limiting factors for Brazilian potato yield is the use of cultivars that are poorly adapted for tropical conditions and hot temperatures.

The Universidade Federal de Lavras's potato-breeding program has worked intensively developing heat-tolerant clones (Benites and Pinto, 2011; Figueiredo et al., 2015; Fernandes Filho et al., 2021; Patiño-Torres et al., 2021). To determine which clones, have high heat tolerance, they are tested under varying levels of heat stress. Only clones that perform well under mild and high temperatures are selected (Benites and Pinto, 2011; Rykaczewska, 2015; Fernandes Filho et al., 2021; Patiño-Torres et al., 2021). Furthermore, Andrade et al. (2021) reported that environments with a higher level of heat stress showed a greater capacity to differentiate clones according to their performance.

The strategy mentioned above for selecting heat-tolerant clones requires clones to be evaluated in two or more contrasting environments regarding heat stress. According to Smith et al. (2001), to better evaluate G×E, it is more realistic to use models that consider variance heterogeneity and genetic covariances when analyzing MET data. Cullis et al. (1998) and Smith et al. (2001) reported a significant reduction in the log-likelihood REML when homogenous variance was assumed for the G×E. In multiplicative

models for MET data, the interaction of simple nature is accounted by the heterogeneity of genetic variances along the environments, whereas the complex interaction is accounted by the heterogeneity of genetic correlations between pairs of environments (Crossa et al., 2004; Eeuwijk et al., 2016). According to Eeuwijk et al. (2016), when the genetic correlations between two environments are high, it means that there is a smaller percentage of complex interaction involved.

MET data analysis can be carry out in one or two stages (Cullis et al., 1998; Smith et al., 2001; Smith et al., 2015; Gogel et al., 2018). Although the two-stage analysis is often employed, under unbalance (varying number of genotypes between environments or varying number of repetitions between and within environments), a common condition in the early stage of potato-breeding programs, the MET data analysis in one stage is more efficient (Cullis et al., 1998; Smith et al., 2001; Paget et al., 2017; Gogel et al., 2018).

Although the analysis of MET data is advantageous because of the advantage of the interrelationship between environments, it is increasing selective accuracy and allows a better interpretation of the G×E interaction (Smith et al., 2001; Kelly et al., 2007). The family effect, naturally present in the initial stage of potato-breeding programs, has also been neglected in this type of analysis (Fernandes Filho et al., 2021). However, one can combine the advantages cited above with those described in Section 4.1. Making MET analysis a powerful tool for clone selection in the early stages of potato-breeding programs, especially under tropical conditions, it is opportune to highlight that the recovery of inter-environmental information was more pronounced for the effect of clones within-family because of the higher magnitude of genetic correlations between seasons. This corroborates with the lower magnitude of the selective efficiency in MET analysis comparison to ST analysis. Selective efficiency for the TTY trait progressively reduced with increasing heat stress levels, whereas the selective efficiency for the SG trait was maintained practically constant. These results ratify the importance of environments with higher heat stress levels to discriminate potato clones for the TTY trait, as well as the greater contribution of the family effect to the selection of superior clones for the SG trait.

The unstructured model can be used in cases when only a few trials are included, due to the smaller number of parameters that need to be estimated, avoiding the use of factor analysis (Smith et al., 2001; Melo et al., 2020; Fernandes Filho et al., 2021). However, the use of the unstructured model implies the prediction of genotypic effects for each trial (Kelly et al., 2007; Eeuwijk et al., 2016; Melo et al., 2020; Fernandes Filho et al., 2021). Hence, it is desirable to use a selection index to capture the G×E interaction and to combine MET analyses realized for multiple traits (Kelly et al., 2007; Melo et al., 2020; Fernandes Filho et al., 2021).

There are many options for selection indexes (Mendes et al., 2009; Rocha et al., 2018; Yan and Frégeau-Reid, 2018; Melo et al., 2020). Among the different indexes, the FAI-BLUP index stands out, because it can incorporate data from various environments and traits without the need for weights. It also avoids issues with multicollinearity and makes it easier to understand the G×E

interaction through exploratory factor analysis (Rocha et al., 2018; Oliveira et al., 2019). In the present study, factor analysis showed a superior performance in summarizing the relationships among environments (WHS, MHS, and HHS) and traits (TTY and SG) for each effect (u_c , u_e , and u_{gMET}). In addition, the factor analysis showed that the SG trait had a lower level of interaction because all seasons were grouped under Factor 1. On the other hand, the TTY trait had a higher level of interaction as the WHS and HHS seasons were grouped under distinct factors. Thus, including the effect of families in the MET analysis is a helpful strategy to increase the accuracy of superior clone selection.

Finally, the inclusion of the family effect increased the selective efficiency of clones in ST and MET selection on schemes through an increment in the accuracy of the total genotypic value. On average, the selective efficiency of clones was 11% and 7% for ST and 12% and 6% in MET for the traits SG and TTY, respectively. An expressive reduction of the family effect under heat stress for TTY and of lower magnitude for SG was observed.

Thus, the results of the present work suggest that the inclusion of the family effect in clone selection models, in the initial stage of potato-breeding programs, is desirable because it contributes to increasing the selective efficiency of clones without generating additional costs, especially for the SG trait.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

VM: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. MA: Supervision, Writing – review & editing. LP: Data curation, Writing – review & editing. LM: Data curation, Writing – review & editing. CF: Data curation, Writing – review & editing. MG: Data curation, Writing – review & editing. JN: Supervision, Writing – review & editing. LJ: Supervision, Writing – review & editing. LZ: Supervision, Writing – review & editing. MR: Supervision, Writing – review & editing. PC: Supervision, Writing – review & editing. TM: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

This work was partially supported by the Brazilian funding agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1253706/full#supplementary-material>

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* 19, 716–723. doi: 10.1109/TAC.1974.1100705
- Amadeu, R. R., Ferrão, L. F. V., Oliveira, I., de, B., Benevenuto, J., Endelman, J. B., et al. (2020). Impact of dominance effects on autotetraploid genomic prediction. *Crop Sci.* 60, 656–665. doi: 10.1002/csc2.20075
- Andrade, M. H. M. L., Fernandes Filho, C. C., Fernandes, M. O., Bastos, A. J. R., Guedes, M. L., Marçal, T., et al. (2020). Accounting for spatial trends to increase the selection efficiency in potato breeding. *Crop Sci.* 60, 2354–2372. doi: 10.1002/csc2.20226
- Andrade, M. H. M. L., Patiño-Torres, A. J., Cavallin, I. C., Guedes, M. L., Carvalho, R. P., Gonçalves, F. M. A., et al. (2021). Stability of potato clones resistant to potato virus Y under suboptimal conditions. *Crop Breed. Appl. Biotechnol.* 21, 1–9. doi: 10.1590/1984-70332021v21n1a8
- Bartlett, M. S. (1938). Methods of estimating mental factors. *Nature* 141, 609–610.
- Benites, F. R. G., and Pinto, C. A. B. P. (2011). Genetic gains for heat tolerance in potato in three cycles of recurrent selection. *Crop Breed. Appl. Biotechnol.* 11, 133–140. doi: 10.1590/S1984-70332011000200005
- Birch, P. R. J., Bryan, G., Fenton, B., Gilroy, E. M., Hein, I., Jones, J. T., et al. (2012). Crops that feed the world 8: Potato: are the trends of increased global production sustainable? *Food Secur.* 4, 477–508. doi: 10.1007/s12571-012-0220-1
- Bradshaw, J. E. (2017). Review and analysis of limitations in ways to improve conventional potato breeding. *Potato Res.* 60, 171–193. doi: 10.1007/s11540-017-9346-z
- Bradshaw, J. E. (2022). A brief history of the impact of potato genetics on the breeding of tetraploid potato cultivars for tuber propagation. *Potato Res.* 65, 461–501. doi: 10.1007/s11540-021-09517-w
- Cobb, J. N., Juma, R. U., Biswas, P. S., Arbelaez, J. D., Rutkoski, J., Atlin, G., et al. (2019). Enhancing the rate of genetic gain in public-sector plant breeding programs: lessons from the breeder's equation. *Theor. Appl. Genet.* 132, 627–645. doi: 10.1007/s00122-019-03317-0
- Crossa, J., Yang, R.-C., and Cornelius, P. L. (2004). Studying crossover genotype × environment interaction using linear-bilinear models and mixed models. *J. Agric. Biol. Environ. Stat.* 9, 362–380. doi: 10.1198/108571104X4423
- Cullis, B., Gogel, B., Verbyla, A., and Thompson, R. (1998). Spatial analysis of multi-environment early generation variety trials. *Biometrics* 54, 1. doi: 10.2307/2533991
- Cullis, B. R., Smith, A. B., and Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11, 381–393. doi: 10.1198/108571106X154443
- Dahal, K., Li, X.-Q., Tai, H., Creelman, A., and Bizimungu, B. (2019). Improving potato stress tolerance and tuber yield under a climate change scenario – A current overview. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00563
- Devaux, A., Goffart, J.-P., Kromann, P., Andrade-Piedra, J., Polar, V., and Hareau, G. (2021). The potato of the future: opportunities and challenges in sustainable agri-food systems. *Potato Res.* 64, 681–720. doi: 10.1007/s11540-021-09501-4
- Devaux, A., Kromann, P., and Ortiz, O. (2014). Potatoes for sustainable global food security. *Potato Res.* 57, 185–199. doi: 10.1007/s11540-014-9265-1
- Diniz, M. C. D. R., Pinto, C. A. B. P., and de Lambert, E. S. (2006). Sample size for family evaluation in potato breeding programs. *Cienc. Agrotecnologia.* 30, 277–282. doi: 10.1590/S1413-70542006000200013
- Eeuwijk, F. A., Bustos-Korts, D. V., and Malosetti, M. (2016). What should students in plant breeding know about the statistical aspects of genotype × Environment interactions? *Crop Sci.* 56, 2119–2140. doi: 10.2135/cropsci2015.06.0375
- FAO. (2020). *Food balance sheet* (Food and Agriculture Organization of the United Nations). Available at: www.fao.org/faostat/en/#data/FBS (Accessed July 01, 2023).
- FAO. (2021). (Food and Agriculture Organization of the United Nations. FAOSTAT). Available at: www.fao.org/faostat/en/#data/QCL (Accessed July 01, 2023).
- Federer, W. T. (1956). “Augmented (or hoonuiku) designs,” in *Biometrics unit technical reports*, vol. 33. Available at: <https://ecommons.cornell.edu/server/api/core/bitstreams/7e626471-05ad-4f2b-858e-be100a513b5e/content>
- Fernandes Filho, C. C., Andrade, M. H. M. L., Souza Marçal, T., Fernandes, M. O., Bastos, A. J. R., Guedes, M. L., et al. (2021). Selection of potato clones for heat tolerance and resistance to potato viruses X and Y for processing purposes. *Crop Sci.* 61, 552–565. doi: 10.1002/csc2.20361
- Figueiredo, I. C. R., Pinto, C. A. B. P., Ribeiro, G. H. M. R., de O., L., Lyra, D. H., and Moreira, C. M. (2015). Efficiency of selection in early generations of potato families with a view toward heat tolerance. *Crop Breed. Appl. Biotechnol.* 15, 210–217. doi: 10.1590/1984-70332015v15n4a37
- Fleisher, D. H., Condori, B., Quiroz, R., Alva, A., Asseng, S., Barreda, C., et al. (2017). A potato model intercomparison across varying climates and productivity levels. *Glob. Chang. Biol.* 23, 1258–1281. doi: 10.1111/gcb.13411
- Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., et al. (2011). Solutions for a cultivated planet. *Nature* 478, 337–342. doi: 10.1038/nature10452
- Gilmour, A. R. (2021). *Echidna mixed model software*. Available at: www.echidnamms.org/ (Accessed March 05, 2021).
- Gogel, B., Smith, A., and Cullis, B. (2018). Comparison of a one- and two-stage mixed model analysis of Australia's National Variety Trial Southern Region wheat data. *Euphytica* 214, 44. doi: 10.1007/s10681-018-2116-4
- Gopal, J. (2015). Challenges and way-forward in selection of superior parents, crosses, and clones in potato breeding. *Potato Res.* 58, 165–188. doi: 10.1007/s11540-015-9292-6
- Hancock, R. D., Morris, W. L., Ducreux, L. J. M., Morris, J. A., Usman, M., Verrall, S. R., et al. (2014). Physiological, biochemical, and molecular responses of the potato (*Solanum tuberosum* L.) plant to moderately elevated temperature. *Plant Cell Environ.* 37, 439–450. doi: 10.1111/pce.12168
- Haverkort, A. J., and Verhagen, A. (2008). Climate change and its repercussions for the potato supply chain. *Potato Res.* 51, 223–237. doi: 10.1007/s11540-008-9107-0
- Haynes, K. G., Gergela, D. M., Hutchinson, C. M., Yencho, G. C., Clough, M. E., Henninger, M. R., et al. (2012). Early generation selection at multiple locations may identify potato parents that produce more widely adapted progeny. *Euphytica* 186, 573–583. doi: 10.1007/s10681-012-0685-1
- Henderson, C. R., Kempthorne, O., Searle, S. R., and von Krosigk, C. M. (1959). The estimation of environmental and genetic trends from records subject to culling. *Biometrics* 15, 192. doi: 10.2307/2527669
- INMET. (2022). *Instituto Nacional de Meteorologia (INMET)*. Available at: www.gov.br/agricultura/pt-br/assuntos/inmet (Accessed July 04, 2022).
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 187–200. doi: 10.1007/BF02289233
- Kelly, A. M., Smith, A. B., Eccleston, J. A., and Cullis, B. R. (2007). The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Sci.* 47, 1063–1070. doi: 10.2135/cropsci2006.08.0540
- Kunkel, R., and Campbell, G. S. (1987). Maximum potential potato yield in the Columbia Basin, USA: Model and measured values. *Am. Potato J.* 64, 355–366. doi: 10.1007/BF02853597
- Lambert, E. S., Pinto, C. A. B. P., and Menezes, C. B. (2006). Potato improvement for tropical conditions: I. Analysis of stability. *Crop Breed. Appl. Biotechnol.* 6, 129–135. doi: 10.12702/1984-7033.v06n02a03
- Matias, F. I., Caraza-Harter, M. V., and Endelman, J. B. (2020). FIELDImageR: An R package to analyze orthomosaic images from agricultural field trials. *TPPJ* 3:e20005. doi: 10.1002/ppj2.20005

- Melo, D. S., Pinto, C. A. B. P., Peixoto, L. S., Neder, D. G., and de Assis, J. C. (2011). Early selection of full-sib potato families. *Cienc. Agrotecnologia* 35, 1101–1109. doi: 10.1590/S1413-70542011000600009
- Melo, V. L., de Marçal, T. S., de Rocha, J. R. A. S. C., dos Anjos, R. S. R., Carneiro, P. C. S., and de Carneiro, J. E. S. (2020). Modeling (co)variance structures for genetic and non-genetic effects in the selection of common bean progenies. *Euphytica* 216, 77. doi: 10.1007/s10681-020-02607-9
- Mendes, F. F., Ramalho, M. A. P., and de Abreu, Â.F.B. (2009). Selection index for choosing segregating populations in common bean. *Pesqui Agropecu Bras.* 44, 1312–1318. doi: 10.1590/S0100-204X2009001000015
- Meyer, K. (2008). Likelihood calculations to evaluate experimental designs to estimate genetic variances. *Heredity (Edinb)* 101, 212–221. doi: 10.1038/hdy.2008.46
- Meyer, R. C., Milbourne, D., Hackett, C. A., Bradshaw, J. E., McNichol, J. W., and Waugh, R. (1998). Linkage analysis in tetraploid potato and association of markers with quantitative resistance to late blight (*Phytophthora infestans*). *Mol. Gen. Genet.* 259, 150–160. doi: 10.1007/s004380050800
- Obidiegwu, J. E., Bryan, G. J., Jones, H. G., and Prashar, A. (2015). Coping with drought: stress and adaptive responses in potato and perspectives for improvement. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00542
- Oliveira, I. C. M., Marçal, T., de, S., da Bernardino, K. C., de Ribeiro, P. C. O., da Parrella, R. A. C., et al. (2019). Combining ability of biomass sorghum lines for agroindustrial characters and multitrait selection of photosensitive hybrids for energy cogeneration. *Crop Sci.* 59, 1554–1566. doi: 10.2135/cropsci2018.11.0693
- Paget, M. F., Alspach, P. A., Anderson, J. A. D., Genet, R. A., Braam, W. F., and Apolaza, L. A. (2017). Replicate allocation to improve selection efficiency in the early stages of a potato breeding scheme. *Euphytica* 213, 221. doi: 10.1007/s10681-017-2004-3
- Patiño-Torres, A. J., Andrade, M. H. M. L., Guedes, M. L., Cavallin, I. C., Pinto, C. A. B. P., Souza, J. C., et al. (2021). Performance of superior potato clones under high and mild temperatures in tropical climate. *Agron. J.* 113, 2349–2360. doi: 10.1002/agj2.20704
- Patterson, H. D., and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545–554. doi: 10.1093/biomet/58.3.545
- Paula, R. G., Pereira, G. S., Paula, I. G., Carneiro, A. L. N., Carneiro, P. C. S., dos Anjos, R. S. R., et al. (2020). Multipopulation recurrent selection: An approach with generation and population effects in selection of self-pollinated progenies. *Agron. J.* 112, 4602–4612. doi: 10.1002/agj2.20422
- Pereira, F. C., Bruzi, A. T., de Matos, J. W., Rezende, B. A., Prado, L. C., and Nunes, J. A. R. (2017). Implications of the population effect in the selection of soybean progeny. *Plant Breed* 136, 679–687. doi: 10.1111/pbr.12512
- Piepho, H. P., Möhring, J., Melchinger, A. E., and Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161, 209–228. doi: 10.1007/s10681-007-9449-8
- Qiao, C. G., Basford, K. E., DeLacy, I. H., and Cooper, M. (2000). Evaluation of experimental designs and spatial analyses in wheat breeding trials. *Theor. Appl. Genet.* 100, 9–16. doi: 10.1007/s001220050002
- R Core Team. (2021). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: www.Rproject.org/.
- Resende, M. D. V., Ramalho, M. A. P., Carneiro, P. C. S., Carneiro, J. E. S., Batista, L. G., and Gois, I. B. (2016). Selection index with parents, populations, progenies, and generations effects in autogamous plant breeding. *Crop Sci.* 56, 530–546. doi: 10.2135/cropsci2015.05.0303
- Rocha, J. R., do, A. S., de, C., MaChado, J. C., and Carneiro, P. C. S. (2018). Multitrait index based on factor analysis and ideotype-design: proposal and application on elephant grass breeding for bioenergy. *GCB Bioenergy* 10, 52–60. doi: 10.1111/gcbb.12443
- Rykaczewska, K. (2015). The effect of high temperature occurring in subsequent stages of plant development on potato yield and tuber physiological defects. *Am. J. Potato Res.* 92, 339–349. doi: 10.1007/s12230-015-9436-x
- SAS Institute Inc. (2016). “SAS/STAT® 14.2 user's guide,” in *Software* (Cary, NC: SAS Institute Inc).
- Schippers, P. A. (1976). The relationship between specific gravity and percentage dry matter in potato tubers. *Am. Potato J.* 53, 111–122. doi: 10.1007/BF02854115
- Slater, A. T., Cogan, N. O., Hayes, B. J., Schultz, L., Dale, M. F. B., Bryan, G. J., et al. (2014b). Improving breeding efficiency in potato using molecular and quantitative genetics. *Theor. Appl. Genet.* 127, 2279–2292. doi: 10.1007/s00122-014-2386-8
- Slater, A. T., Wilson, G. M., Cogan, N. O. I., Forster, J. W., and Hayes, B. J. (2014a). Improving the analysis of low heritability complex traits for enhanced genetic gain in potato. *Theor. Appl. Genet.* 127, 809–820. doi: 10.1007/s00122-013-2258-7
- Smith, A., Cullis, B., and Thompson, R. (2001). Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57, 1138–1147. doi: 10.1111/j.0006-341X.2001.01138.x
- Smith, A. B., Ganesalingam, A., Kuchel, H., and Cullis, B. R. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theor. Appl. Genet.* 128, 55–72. doi: 10.1007/s00122-014-2412-x
- Stich, B., and Van Inghelandt, D. (2018). Prospects and potential uses of genomic prediction of key performance traits in tetraploid potato. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00159
- Volpato, L., Simiqueli, G. F., Alves, R. S., Rocha, J. R., do, A. S., de, C., et al. (2018). Selection of inbred soybean progeny (*Glycine max*): an approach with population effect. *Plant Breed* 137, 865–872. doi: 10.1111/pbr.12648
- Voss-Fels, K. P., Wei, X., Ross, E. M., Frisch, M., Aitken, K. S., Cooper, M., et al. (2021). Strategies and considerations for implementing genomic selection to improve traits with additive and non-additive genetic architectures in sugarcane breeding. *Theor. Appl. Genet.* 134, 1493–1511. doi: 10.1007/s00122-021-03785-3
- Wickham, H. (2016). “ggplot2,” in *Data analysis* (Cham: Springer), 189–201. doi: 10.1007/978-3-319-24277-4_9
- Williams, E. R., John, J. A., and Whitaker, D. (2014). Construction of more flexible and efficient P-rep designs. *Aust. N Z J. Stat.* 56, 89–96. doi: 10.1111/anzs.12068
- Williams, E., Piepho, H.-P., and Whitaker, D. (2011). Augmented p-rep designs. *Biom. J.* 53, 19–27. doi: 10.1002/bimj.201000102
- Wilson, S., Zheng, C., Maliepaard, C., Mulder, H. A., Visser, R. G. F., van der Burgt, A., et al. (2021). Understanding the effectiveness of genomic prediction in tetraploid potato. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.672417
- Yadav, S., Wei, X., Joyce, P., Atkin, F., Deomano, E., Sun, Y., et al. (2021). Improved genomic prediction of clonal performance in sugarcane by exploiting non-additive genetic effects. *Theor. Appl. Genet.* 134, 2235–2252. doi: 10.1007/s00122-021-03822-1
- Yan, W., and Frégeau-Reid, J. (2018). Genotype by yield*Trait (GYT) biplot: a novel approach for genotype selection based on multiple traits. *Sci. Rep.* 8, 8242. doi: 10.1038/s41598-018-26688-8

ARTIGO 2

Artigo - Elaborado conforme a norma NBR 6022	
Título do artigo:	Fenotipagem por imagens na seleção de clones de batata com base na produtividade e seus componentes
Autor:	Vinícius Samuel Martins



FENOTIPAGEM POR IMAGENS NA SELEÇÃO DE CLONES DE BATATA COM BASE NA PRODUTIVIDADE E SEUS COMPONENTES

PHENOTYPING BY IMAGES IN THE SELECTION OF POTATO CLONES BASED ON PRODUCTIVITY AND ITS COMPONENTS¹

Vinícius Samuel Martins

<https://orcid.org/0000-0002-4878-6888>

Resumo: O cultivo da batata (*Solanum tuberosum* L.) estabeleceu-se como um dos pilares da segurança alimentar mundial, equiparando-se em importância ao arroz e trigo, graças à sua elevada produtividade por área e ao expressivo valor nutricional. O processo de melhoramento desta cultura apresenta desafios ao avaliar diversas características de forma simultânea, que podem ser reguladas tanto por mecanismos genéticos simples quanto por interações gênicas complexas na herança dos caracteres. Para otimizar este processo, tecnologias de visão computacional têm sido implementadas no melhoramento de batata, viabilizando a fenotipagem em escala, com baixo custo e precisão dos múltiplos caracteres. O objetivo deste trabalho foi investigar os componentes de produção de batata e seu efeito na acurácia seletiva de clones em um programa de melhoramento de batata envolvendo análise clonal em modelo misto unicaracterístico (MMUC) e modelo misto multicaracterístico (MMMC). Foram avaliados 620 clones oriundos do programa de melhoramento de batata da Universidade Federal de Lavras para os caracteres massa média de tubérculos (MMT), número de tubérculos totais (NTT) e produtividade total de tubérculos (PTT). Estes clones foram distribuídos em três ensaios avaliados na safra de inverno do ano de 2023. No intuito de investigar o efeito de dos componentes da produtividade da batata e seu impacto na acurácia seletiva de clones em análises MMUC e MMMC, os resultados demonstraram a superioridade do modelo MMMC sobre a MMUC no melhoramento de batata proporcionando melhor ajuste e estimativas mais robustas dos componentes de variância genética e melhor controle ambiental. As correlações genéticas se confirmaram em positivas entre o caráter PTT e NTT, e antagônicas entre MMT e NTT. As correlações genotípicas e residuais confirmam associação positiva entre rendimento total de tubérculos e NTT e MMT. Adicionalmente, controlando PTT a correlação parcial entre MMT e NTT foi alta e negativa. As correlações parciais residuais permaneceram com mesmo sinal das genotípicas, porém com menor magnitude. Ambos os modelos demonstram alta acurácia seletiva para os componentes de produção MMT, NTT e PTT. Apesar das acurácias e variâncias de clone estar próximas o modelo MMMC pode ser útil em cenários que não se tem um bom controle ambiental.

¹ Artigo científico apresentado à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração Melhoramento Genético de Plantas, Genética Quantitativa e Biometria.

Palavras-Chave: *Solanum tuberosum* L.; Fenotipagem por imagem; Acurácia; Análise multivariada.

Abstract: Potato (*Solanum tuberosum* L.) cultivation has established itself as one of the pillars of global food security, equaling rice and wheat in importance thanks to its high productivity per area and significant nutritional value. The breeding process of this crop presents challenges in evaluating several characteristics simultaneously, which can be regulated by both simple genetic mechanisms and complex gene interactions in the inheritance of traits. To optimize this process, computer vision technologies have been implemented in potato breeding, enabling phenotyping at scale, with low cost and precision of multiple traits. The objective of this work was to investigate the potato production components and their effect on the selective accuracy of clones in a potato breeding program involving clonal analysis in a single-trait mixed model (STM) and a multi-trait mixed model (MMMC). Six hundred and twenty clones from the potato breeding program at the Federal University of Lavras were evaluated for the traits average tuber mass (MMT), total number of tubers (NTT), and total tuber yield (PTT). These clones were distributed across three trials evaluated during the winter crop season of 2023. In order to investigate the effect of potato yield components and their impact on the selective accuracy of clones in MMUC and MMMC analyses, the results demonstrated the superiority of the MMMC model over the MMUC model in potato breeding, providing better fit and more robust estimates of genetic variance components and better environmental control. Genetic correlations were confirmed as positive between PTT and NTT, and antagonistic between MMT and NTT. Genotypic and residual correlations confirm a positive association between total tuber yield and NTT and MMT. Additionally, controlling for PTT, the partial correlation between MTT and NTT was high and negative. The residual partial correlations remained with the same sign as the genotypic correlations, but with a lower magnitude. Both models demonstrate high selective accuracy for the production components MMT, NTT, and PTT. Despite the accuracies and clone variances being close, the MMMC model may be useful in scenarios where there is poor environmental control.

Keywords: *Solanum tuberosum* L.; Image phenotyping; Accuracy; Multivariate analysis.

1 INTRODUÇÃO

A batata (*Solanum tuberosum* L.) é um alimento básico da cadeia alimentar, sendo comparável ao arroz e ao trigo em termos de importância para a segurança alimentar global. Com alto valor nutricional e alta eficiência produtiva por unidade de área, o tubérculo se destaca como uma estratégia promissora para atender à crescente demanda por alimentos, notabilizando-se por sua versatilidade culinária e adaptabilidade a diferentes condições de cultivo (Andrison 2017; Devaux *et al.*, 2021).

Apesar de sua grande importância na cadeia alimentar, o melhoramento genético da batata tem apresentado ganhos relativamente modestos em relação a caracteres associados à produção de tubérculos (Bonierbale *et al.*, 2020; Bradshaw, 2022; Ortiz *et al.*, 2022). Todavia, Douches *et al.* (1996) relataram um incremento no rendimento de tubérculos até cinco vezes

superior ao longo do século XX nos EUA. As cultivares com maior produtividade foram lançadas na primeira metade do século, evidenciando a perda de ganho genético na cultura em um período de cinco décadas. Um exemplo é a cultivar 'Atlantic' que se destaca como a principal variedade utilizada no processamento de batata *chips* no mundo, tendo sido registrada na década de 1970 (Schumaker *et al.*, 1977).

Cabello *et al.* (2014) investigaram a associação entre caracteres dos componentes de produção em batatas do grupo *Andigenum*, sob condições de irrigação e seca, revelando variações significativas entre caracteres e ambientes. Observaram que a seleção para número de tubérculos e tamanho tem se mostrado mais eficaz, enquanto o rendimento total demanda estratégias que considerem as interações genótipo $\times$ ambiente (G $\times$ E) e possuem baixa herdabilidade analisando o caráter separadamente.

Diante disso, os melhoristas de batata enfrentam o desafio de selecionar não apenas características fenotípicas individuais, mas também um conjunto de até 40 caracteres que apresentem simultaneamente boa performance agrônômica e aceitação comercial (Bradshaw, 2017; Gopal, 2015). A compreensão das correlações genótípicas e fenotípicas de modelos univariado e multivariado, bem como seus componentes entre caracteres, é essencial para o planejamento e estratégia do programa de melhoramento de batata (Thelen *et al.*, 2025).

A seleção simultânea de múltiplos caracteres tem se tornado uma abordagem desejável, sendo a análise multivariada uma ferramenta adequada para identificar genótipos que reúnam as características de interesse, com o intuito de selecionar indivíduos superiores (Aksoy *et al.*, 2021). Cuevas *et al.* (2023) destacam que modelos multicaracterísticos (MMMC) podem aumentar a acurácia da predição genômica (GP), especialmente quando há correlação entre os caracteres, superando modelos univariados (MMUC).

Mesmo com a relevância desses caracteres para análises multivariadas, a compreensão total de sua herança é frequentemente limitada pela dificuldade em mensurá-los de forma precisa e quantitativa, especialmente nas populações clonais dentro dos estágios do programa de melhoramento. Uma alternativa promissora é a utilização da visão computacional aplicada à análise de imagens digitais, que auxilia na mensuração de caracteres em estágios de difícil mensuração, com trabalho semiautomático, em escala e de baixo custo, além de possibilitar o armazenamento e análise de dados de forma não destrutiva (Thelen *et al.*, 2025).

A utilização de imagens em estágios iniciais do programa de melhoramento de plantas, se torna essencial em cenários com avaliação de milhares de indivíduos viabilizando análises multivariadas para componentes de produção. Nesse sentido, o objetivo deste trabalho foi investigar a associação de caracteres de produtividade da batata em análise univariada e

multivariada e seu impacto na acurácia seletiva de clones em um programa de melhoramento genético da batata.

2 MATERIAIS E MÉTODOS

2.1 Ensaio de campo

Três populações clonais de batata (*Solanum tuberosum* L.) foram avaliadas no Programa de Melhoramento Genético da Batata da Universidade Federal de Lavras (PROBATATA-UFLA), conduzidos no Centro de Desenvolvimento Científico e Tecnológico (21°12'19,8"S, 44°58'48,8"W, 919 m de altitude, e com solo classificado como latossolo vermelho-amarelo), localizado no município de Lavras, Minas Gerais, Brasil. As diferentes populações clonais foram identificadas em POP1, POP2 e POP3. As populações são oriundas do PROBATATA-UFLA, caracterizadas nos segmentos de produção *in natura*, *chips* e pré-fritos congelados. Os ensaios foram avaliados na safra de inverno de 2023, ambiente caracterizado sem estresse por calor, iniciando o plantio em abril e colheita em outubro.

O número de linhas, colunas, clones, testemunhas, parcelas e porcentagem de parcelas repetidas (p_N) são mostrados na Tabela 1. Os ensaios foram conduzidos no delineamento parcialmente replicado (p-rep) (Cullis *et al.*, 2006) e percentual de parcelas replicadas (p_N) próximo de 40%. As parcelas foram compostas por cinco plantas espaçadas 0,30 m entre plantas e 0,80 m entre linhas.

Tabela 1 - Caracterização de populações, número de linhas, coluna, clone, testemunha, parcela e porcentagem de parcelas replicadas.

População [†]	Número de níveis					p_N [§]
	Linha	Coluna	Clone	Testemunha	Parcela	
POP1	9	29	211	2	350	39.14
POP2	15	27	303	2	543	43.83
POP3	15	12	100	2	173	41.04

[†]Identificação da população clonal: POP1, POP2 e POP3, sendo os códigos POP1, POP2 e POP3 identificam as diferentes populações clonais.

[§] p_N - Porcentagem de unidades experimentais de parcelas ocupadas por clones replicados, dada pela expressão. $p_N = (N - N_{trat})/N$, onde N é o número de parcelas e N_{trat} é o número de tratamentos.

Fonte: Elaboração própria (2025).

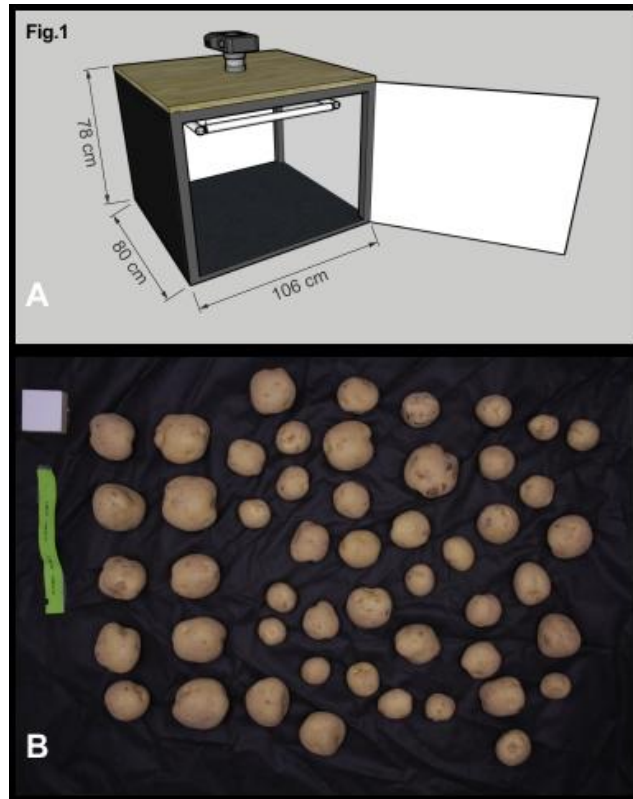
O manejo da cultura da batata seguiu as recomendações técnicas preconizadas para o estado de Minas Gerais em todos os experimentos. A adubação de plantio consistiu na aplicação de $1,5 \text{ t ha}^{-1}$ do fertilizante formulado 28-08-16 (N-P₂O₅-K₂O) no plantio. A adubação em cobertura foi realizada com $0,30 \text{ t ha}^{-1}$ 20-00-20 (N-P₂O₅-K₂O). O preparo do solo envolveu três operações sequenciais: aração, gradagem e uso de enxada rotativa. A irrigação complementar foi fornecida por sistema de aspersão, sendo seu manejo ajustado de acordo com as necessidades hídricas da cultura e o regime pluviométrico observado ao longo do ciclo de cultivo.

As características avaliadas neste estudo foram: massa média de tubérculos (MMT, g/tubérculo), número total de tubérculos por parcela (NTT) e produção total de tubérculos (PTT, t ha^{-1}). Em uma área de $1,2 \text{ m}^2$ foi obtida a massa de tubérculos totais MTT (g/parcela), convertendo em seguida para toneladas por hectare (PTT, t ha^{-1}). A MMT, foi calculado pela expressão $MMT = \text{MTT (g/parcela)} / NTT$. Devido ao grande número tubérculos e parcelas para avaliação, o NTT foi estimado com auxílio de imagens digitais pela contagem indireta dos tubérculos.

2.2. Aquisição de imagem digitais

A avaliação do carácter NTT, foi realizada por meio de imagens digitais capturadas de todas as unidades experimentais, em formato “.jpg”. As imagens foram capturadas usando um sistema de visão computacional no estúdio fotográfico no Laboratório de Fenômica do Departamento de Biologia da UFLA. O estúdio fotográfico contém iluminação controlada e uniforme, fabricado com estrutura metálica e parte superior de MDF tipo “cru” de 15 mm de espessura, contém duas lâmpadas de LED de 55 cm de comprimento, 9W, 900 lm, e luz cor branco frio centralizado no topo do estúdio. Nas partes laterais são revestidas de lona dupla face de polietileno de baixa densidade, sendo a parte branca voltada para o interior do estúdio. O fundo é revestido de material tecido não tecido (TNT) de polipropileno, cor preta fosca. A câmera profissional modelo Canon EOS 60D, lente Canon Zoom Lens EF-S, com distância focal de 19mm, acoplada na área central do topo na parte externa do estúdio fotográfico, ajustada com as seguintes configurações: ISO 200, abertura do diafragma 1/13 e com escala de número F10. (Figura 1).

Figura 1 - Modelo da estrutura do estúdio fotográfico usado para capturar as imagens digitais e (B) imagem de uma parcela de batata, contendo uma etiqueta de identificação e um objeto de referência.



Fonte: Elaboração própria (2025).

As imagens digitais de cada parcela experimental foram capturadas individualmente. A quantificação do NTT originou de técnicas de visão computacional, utilizando o pacote OpenCV do Pynton. O ordenamento padrão das cores é RGB (*Red, Green, Blue*), a segmentação das imagens dos tubérculos de batata iniciou com a conversão do espaço de cor RGB para YCrCb, com foco específico no canal Cb, seguida pela técnica de limiarização automática proposta por Otsu (1975). A discriminação entre tubérculos e outros objetos foi realizada pelo processamento das imagens e análise dos atributos morfológicos dos tubérculos, bem como aplicação de máscara binária, permitindo a exclusão de elementos não pertencentes à classe de interesse, seguida por detecção de contornos.

2.3. Análises Estatísticas

Considerando um teste clonal (TC), sendo parcelas arranjadas em um *grid* com r linhas e c colunas, com g clones oriundos de diferentes famílias e avaliados juntamente com p testemunhas. Para análise de dados dos modelos mistos unicaracterístico (MMUC) e modelo misto multicaracterístico (MMMC), a forma matricial do modelo misto é apresentada na equação (1).

$$\mathbf{y} = \mathbf{1}\boldsymbol{\mu} + \mathbf{X}_o\boldsymbol{\tau}_o + \mathbf{Z}_g\mathbf{u}_g + \mathbf{Z}_r\mathbf{u}_r + \mathbf{Z}_c\mathbf{u}_c + \mathbf{e} \quad (1)$$

em que: $\mathbf{y}^{(N \times 1)}$ é o vetor de observações fenotípicas, sendo N o número de parcelas avaliadas; $\mathbf{1}^{(N \times 1)}$ é um vetor de unidades para todos os elementos; $\boldsymbol{\mu}^{(1 \times 1)}$ é o intercepto; $\boldsymbol{\tau}_o^{(o \times 1)}$ é o vetor de efeitos fixos, associado a matriz $\mathbf{X}_o^{(N \times o)}$ (assumindo posto completo), onde o é o número de efeitos fixos; $\mathbf{u}_g^{(g \times 1)}$ é o vetor de efeitos aleatórios de clone associado a matriz $\mathbf{Z}_g^{(N \times g)}$; $\mathbf{u}_r^{(r \times 1)}$ é o vetor de efeitos aleatórios de linha associados a matriz $\mathbf{Z}_r^{(N \times r)}$; $\mathbf{u}_c^{(c \times 1)}$ é o vetor de efeitos genotípicos aleatórios de coluna associados a matriz $\mathbf{Z}_c^{(N \times c)}$; $\mathbf{e}^{(N \times 1)}$ é o vetor de erros aleatórios.

Foi assumido que os vetores $\mathbf{u}_r$, $\mathbf{u}_c$, $\mathbf{u}_g$, e $\mathbf{e}$ são efeitos aleatórios e mutuamente independentes com distribuição multivariada gaussiana, com média zero e matrizes de (co)variâncias $\text{var}(\mathbf{u}_g) = \mathbf{G}_g$, $\text{var}(\mathbf{u}_r) = \mathbf{G}_r$, $\text{var}(\mathbf{u}_c) = \mathbf{G}_c$, e $\text{var}(\mathbf{e}) = \mathbf{R}$. As estruturas dessas matrizes de (co)variância são mostradas na Tabela 2. O modelo MMUC apresenta heterogeneidade das variâncias com estrutura de variância e covariância diagonal (DIAG) para os efeitos de linha, coluna, clone e erro, utilizando a operação de soma direta ($\oplus$). Já o modelo MMMC apresenta heterogeneidade de variâncias e (co)variâncias, ou seja, as matrizes $\mathbf{G}_{tg}$ e $\mathbf{R}_t$ assumiram a forma não estruturada (US), acomodados pela operação direta do produto ($\otimes$). Desta maneira foi possível modelar as variâncias e correlações para efeitos de clone e erro entre os caracteres.

Tabela 2 - Resumo dos modelos ajustados: Modelo Misto Unicaracterístico (MMUC) e Modelo Misto Multicaracterístico (MMMC).

Matriz de (Co)variância	MMUC [‡]	MMMC [‡]
$\mathbf{G}_r$	$\oplus_{y=1}^t \sigma_{ry}^2 \mathbf{I}_{ry}$	$\oplus_{y=1}^t \sigma_{ry}^2 \mathbf{I}_{ry}$
$\mathbf{G}_c$	$\oplus_{y=1}^t \sigma_{cy}^2 \mathbf{I}_{cy}$	$\oplus_{y=1}^t \sigma_{cy}^2 \mathbf{I}_{cy}$
$\mathbf{G}_g$	$\oplus_{y=1}^t \sigma_{gy}^2 \mathbf{I}_{gy}$	$\mathbf{G}_{tg} \otimes \mathbf{I}_g$
$\mathbf{R}$	$\oplus_{y=1}^t \sigma_{ey}^2 \mathbf{I}_{Ny}$	$\mathbf{R}_t \otimes \mathbf{I}_N$

$\mathbf{G}_b$, $\mathbf{G}_s$, $\mathbf{G}_c$ e $\mathbf{R}$ - matrizes de (co)variância associadas aos efeitos de linha, coluna, clone e erro, respectivamente; $\mathbf{G}_{tg}$ e $\mathbf{R}_t$ - matrizes de (co)variância não estruturada usadas para acomodar a interação de clone: genótipo $\times$ característica ($G \times T$) e erro: resíduo $\times$ característica ($R \times T$), respectivamente; σ_{ry}^2 , σ_{cy}^2 , σ_{gy}^2 e σ_{ey}^2 - componentes de variância associados a linha, coluna, clone e erro, sendo cada efeito em cada característica, respectivamente; $\mathbf{I}_{ry}$, $\mathbf{I}_{cy}$, $\mathbf{I}_g$, $\mathbf{I}_{gy}$, $\mathbf{I}_N$, $\mathbf{I}_{Ny}$ - Matrizes de identidade associadas à linha e coluna em cada característica, clone, clone em cada característica, erro e erro em cada característica, respectivamente $\otimes$ - operador de produto Kronecker; $\oplus$ - operador de soma direta.

Fonte: Elaboração própria (2025).

Os parâmetros de covariância dos modelos mostrados na Tabela 2 foram estimados pelo método de *residual maximum likelihood* (REML) (Patterson; Thompson, 1971) e *best linear unbiased predictions* (BLUP) de efeitos aleatórios pelas equações de modelos mistos de

Henderson (Henderson *et al.*, 1959) usando o *software* modelos mistos Echidna (Gilmour, 2021) versão 1.61. As plotagens gráficas e outras análises discutidas nas seções a seguir foram realizadas usando o *software* R (R Core Team, 2025), por meio de pacotes básicos do R e ggplot2 (Wickham, 2016).

2.3.1. Análise dos modelos MMUC e MTM

Os modelos MMUC e MMMC foram comparados usando o critério de informação de Akaike (AIC) apresentado na equação (2) (Akaike, 1974) e por meio da correspondência dos 20% melhores clones pelo coeficiente Czekanowski (Qiao *et al.*, 2000) [CC, equação (3)], baseado na pontuação do índice FAI-BLUP (Rocha *et al.*, 2018), entre as estratégias ($\mathbf{u}_{gMMUC}$ vs $\mathbf{u}_{gMMMC}$). Adicionalmente, os componentes de variância do modelo MMUC foi testado usando o teste da razão de verossimilhança (LRT) mostrado na equação (4):

$$AIC = -2\ell + 2p \quad (2)$$

$$CC = a/(a + b) \quad (3)$$

$$LRT = -2\ln(\ell_1/\ell_2) \quad (4)$$

Em que: ℓ é o logaritmo do ponto de máximo da função de log-verossimilhança residual, p é o número de parâmetros de variância, ℓ_1 é o ponto de máximo da função de log-verossimilhança residual do modelo reduzido (sem o efeito testado), ℓ_2 é o ponto de máximo da função de log-verossimilhança residual do modelo completo.

Logo em seguida, para os parâmetros das matrizes US ($\mathbf{G}_{tg}$ e $\mathbf{R}_t$), foram utilizados dois intervalos de confiança de 95% para os parâmetros do modelo MMMC. O primeiro, baseado na distribuição Qui-Quadrado (SAS Institute, 2016), foram utilizados componentes de variância de linha, coluna, clone e resíduo. O segundo baseado na distribuição Normal Padrão (Meyer, 2008) para as correlações entre os pares de caracteres.

As correlações genotípicas e residuais entre pares de caracteres para efeitos de clone (ρ_{gij}) e resíduo (ρ_{eij}) foram estimadas a partir dos parâmetros das matrizes $\mathbf{G}_{tg}$ e $\mathbf{R}_t$, utilizando as expressões (5) e (6).

$$\rho_{gij} = \frac{\sigma_{gij}}{\sqrt{\sigma_{g_i}^2 \times \sigma_{g_j}^2}} \quad (5)$$

$$\rho_{eij} = \frac{\sigma_{eij}}{\sqrt{\sigma_{e_i}^2 \times \sigma_{e_j}^2}} \quad (6)$$

Onde:

$\sigma_{g_{ij}}$ e $\sigma_{e_{ij}}$ são (co)variâncias entre pares do caráter i e para efeitos de clone e resíduo; $\sigma_{g_i}^2$ e $\sigma_{g_j}^2$ são componentes de variância de caráter i e j para efeitos de clone; $\sigma_{e_i}^2$ e $\sigma_{e_j}^2$ são componentes de variância de caráter i e j para efeito de resíduo.

Os componentes de variância dos modelos MMUC e MMMC foram apresentados em percentagem da proporção da variação total para cada caráter (soma de todos os componentes para cada característica nas análises MMUC e MMMC). Para comparar a eficácia das estratégias de seleção baseadas nos efeitos de clone das estruturas DIAG ($\mathbf{u}_{gMMUC}$) e US ($\mathbf{u}_{gMMMC}$), avaliou-se a eficiência por meio da razão da acurácia dos respectivos efeitos.

2.3.2. Acurácia de clone

A acurácia dos efeitos genotípicos de clone do modelo MMUC ($r_{\hat{g}_{gSTMM}}$) e do modelo MMMC ($r_{\hat{g}_{gMTMM}}$) foram estimadas por meio das expressões (7) e (8), respectivamente.

$$r_{\hat{g}_{gSTMM}} = \sqrt{1 - \frac{v_g}{\sigma_g^2}} \quad (7)$$

$$r_{\hat{g}_{gMTMM}} = \sqrt{1 - \frac{v_{g'}}{\sigma_g^2}} \quad (8)$$

Em que: v_g e $v_{g'}$ são os efeitos da variância médio do erro de predição de clone e σ_g^2 é o componente de variância genética de clone dos modelos MMUC e MMMC, respectivamente.

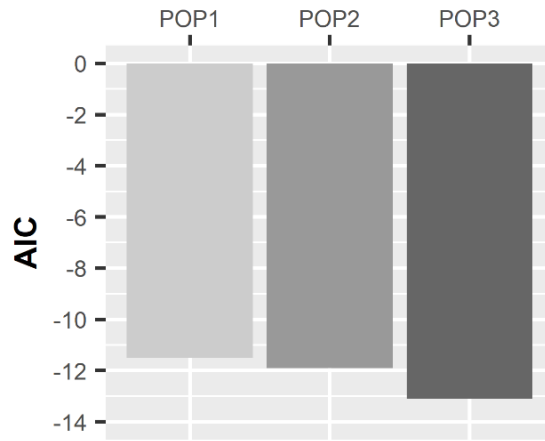
3 RESULTADOS

3.1. Comparação dos modelos MMUC e MMMC

O MMMC proporcionou o melhor ajuste para todas as populações, na comparação com modelo MMUC, o modelo MMMC apresentou AIC inferior a 10 unidades e proporcionou incremento médio de 14% ao log-verossimilhança (ℓ) para ambas as populações (Tabela S1, Figura 2).

Figura 2 - Comparação do ajuste de qualidade do Modelo Misto Multicaracterístico (MMMC) usando o AIC em relação ao Modelo Misto Unicaracterístico (MMUC) com efeitos de clone independentes para três populações clonais diferentes de batata (POP).

Fig.2

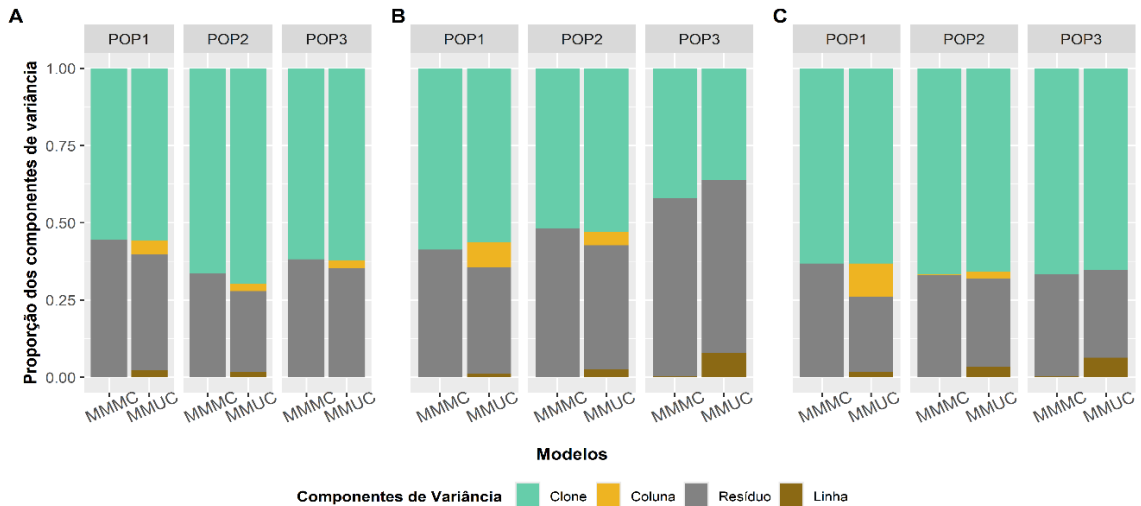


Fonte: Elaboração própria (2025).

Os componentes de variância entre clone, apresentaram diferenças significativas para ambas populações e caracteres, indicando a existência de variabilidade genética entre os clones nas populações, o qual pode ser evidenciado pelas diferenças significativas detectadas pelo teste LRT (Material Suplementar S3), como mostra a Figura 3.

Figura 3 - Partição da variância genética para os caracteres (A) massa média do tubérculo (MMT, g/ tubérculo), (B) número total de tubérculos por parcela (NTT) e (C) rendimento total de tubérculos (PTT, t ha⁻¹) do Modelo Misto Multicaracterístico (MMMC) e Modelo Misto Unicaracterístico (MMUC) para os componentes de variância: σ_g^2 variância do clone; σ_c^2 variância da coluna; σ_e^2 variância do erro; σ_l^2 variância da linha. Foram avaliadas três populações clonais diferentes de batata (POP).

Fig.3



Fonte: Elaboração própria (2025).

Comparando-se os modelos MMUC e MMMC, observou-se que para o modelo MMMC, a média das populações apresentou, em PTT, uma contribuição relativa 0,8% maior da variância genética entre clones. Para o carácter NTT, comprovou-se um incremento de 2,4 % no componente genético entre clones, enquanto para MMT uma redução de 1,4%. Vale destacar que, apesar das diferenças nos valores de AIC entre os modelos e do incremento ao log-verossimilhança (ℓ), as proporções dos componentes de variância foram similares (Figura 3, Material Suplementar S2 e S3).

As populações POP1 e POP2, do modelo MMUC, apresentaram diferenças significativas para ambos os caracteres nos componentes de linha e coluna, exceto para o componente de linha da POP1 no carácter NTT. De maneira semelhante, a POP3 não foi detectada diferenças no componente de coluna para ambos os caracteres, nem o componente de linha para o carácter MMT. Nesse sentido, no modelo MMMC foram observadas contribuições relativas de menor magnitude para os efeitos de linhas e coluna, e não significativos, sendo que para vários caracteres, não foi possível estimar o componente, resultando em valor igual a zero (Material Suplementar S2 e S3).

3.2 Correlações genótípicas parciais e residuais do modelo MMMC

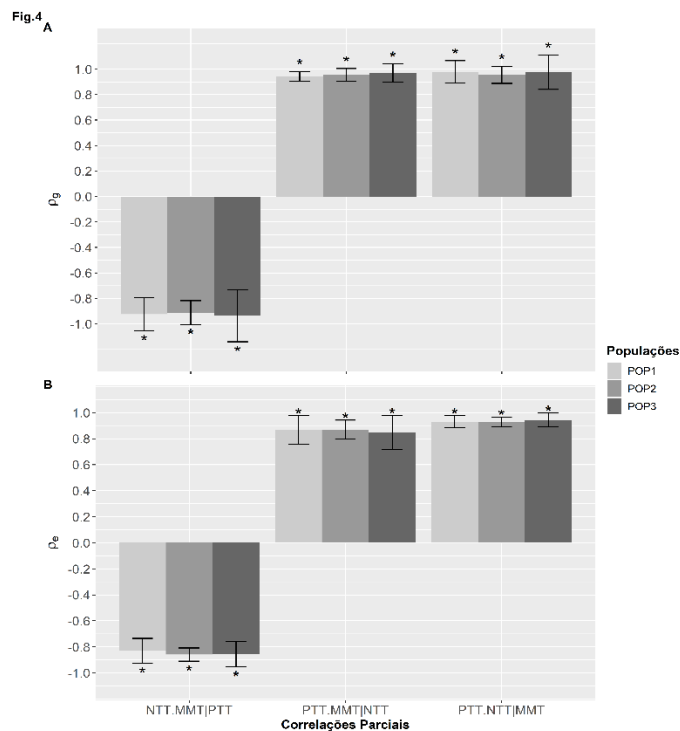
O modelo MMMC revelou correlações genótípicas parciais de alta magnitude dos componentes de produção PTT para NTT e MMT. As correlações genótípicas entre PTT e NTT variaram de 0,95 a 0,98, enquanto entre PTT e MMT oscilaram de 0,94 a 0,97 nas três populações avaliadas. Os intervalos de confiança associados a essas estimativas não incluíram zero, confirmando a significância entre as associações genéticas, indicando forte controle genético destes componentes. Esses resultados demonstram que ambos os componentes contribuem substancialmente para a expressão da produtividade de tubérculos, sugerindo alta eficiência da seleção indireta via NTT ou MMT para incremento de PTT.

Em contraste, a correlação genotípica entre NTT e MMT foi fortemente negativa, variando de -0,91 a -0,94 quando controlada pela PTT. Esse padrão evidencia contribuição diferencial para os componentes da produtividade, no qual genótipos que alocam recursos para produzir maior NTT tendem a apresentar tubérculos de menor MMT, e vice-versa, mantendo-se constante a produtividade total. Essa resposta diferencial de pesos genéticos foi consistente nas três populações avaliadas, demonstrando que a seleção simultânea para aumento de ambos os componentes é geneticamente desafiadora e requer estratégias específicas de melhoramento, como o uso de índices de seleção ou a busca por genótipos recombinantes que superem essa

compensação. A estabilidade das associações genéticas entre populações foi notável, com amplitudes reduzidas entre as estimativas de correlação genética: 0,03 para o par PTT-NTT (0,95 a 0,98) e 0,03 para PTT-MMT (0,94 a 0,97). Essa consistência indica que as associações genéticas observadas são robustas e entre os diferentes *backgrounds* genéticos, reforçando a confiabilidade dessas relações para aplicação em programas de melhoramento genético de batata.

As correlações residuais seguiram padrão similar às genéticas, porém com magnitudes ligeiramente inferiores, sugerindo que fatores ambientais também contribuem para as associações fenotípicas. Entre PTT e NTT, as correlações residuais foram positivas e de alta magnitude (0,93 a 0,94), enquanto entre PTT e MMT variaram de moderadas a altas (0,85 a 0,87). Para o par NTT-MMT, observaram-se correlações residuais negativas de magnitude moderada (-0,83 a -0,86), reforçando o efeito ambiental diferencial. Os intervalos de confiança das correlações residuais não interceptam no zero, indicando que, embora os efeitos ambientais foram de menor magnitude influenciam diretamente em ganhos genéticos no melhoramento de batata (Figura 4).

Figura 4 - Correlações parciais genotípicas (ρ_g) (A) e residuais (ρ_e) (B) entre características avaliadas para três populações clonais diferentes de batata (POP). As barras representam as estimativas das correlações e os intervalos do erro padrão ($\pm$ SE). (*) indicam correlações que são significativamente diferentes de zero com 95% probabilidade.



Fonte: Elaboração própria (2025).

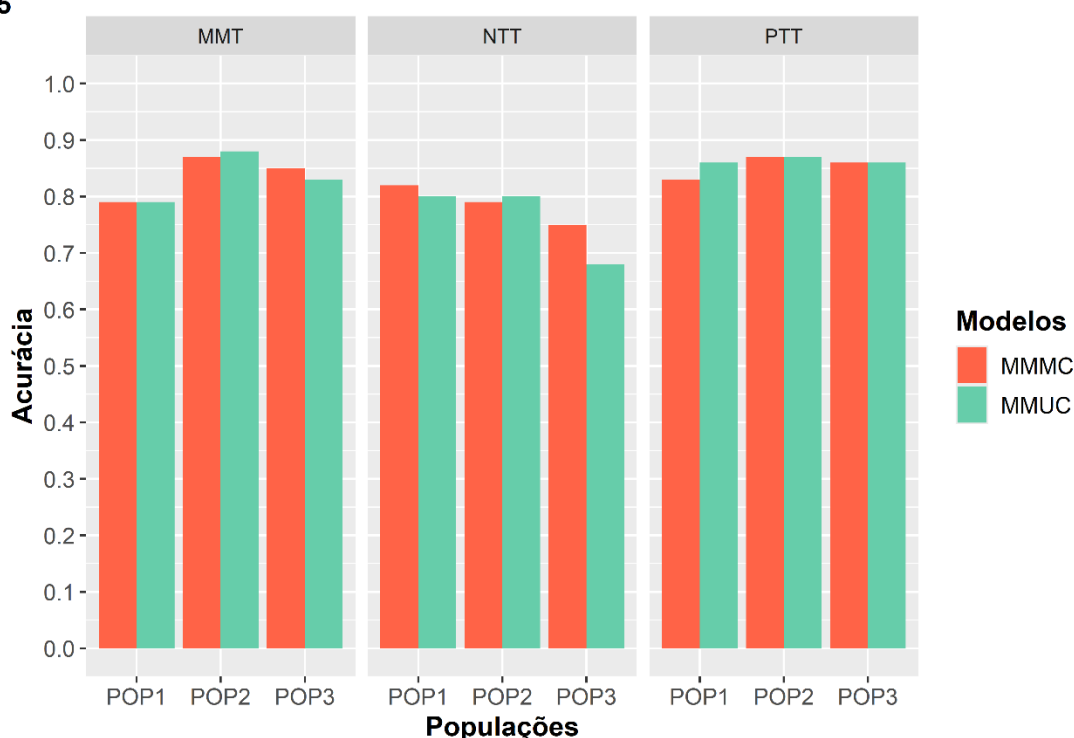
3.3 Acurácia seletiva modelo MMUC e MMMC

Os resultados de r_{gg} do modelo MMUC foram, em média, semelhantes aos do modelo MMMC. Os valores de acurácia para o caráter MMT apresentaram amplitude de 0,09 (0,79-0,88) no MMUC e 0,08 (0,79-0,87) no MMMC, com médias de 0,83 e 0,84, respectivamente. Para NTT, a acurácia variou de 0,12 (0,68-0,80) no MMUC e 0,07 (0,75-0,82) no MMMC, resultando em médias de 0,76 e 0,79, respectivamente (Figura 5, Material suplementar S4).

Todavia, menores amplitudes de acurácia de clone foram detectadas para o caráter PTT, de 0,01 (0,86-0,87) no MMUC e 0,04 (0,83-0,87) no MMMC, com médias de 0,86 e 0,85, respectivamente. Em geral, ambos os modelos apresentaram alta acurácia seletiva ($>0,70$) para a maioria das características avaliadas, com exceção do NTT na POP3 no modelo MMUC (0,68). Vale destacar que, a magnitude da diferença de r_{gg} foi maior para NTT e menor para PTT, sendo que o modelo MMMC apresentou menores amplitudes de r_{gg} em comparação ao modelo MMUC (Figura 5, Material suplementar S4).

Figura 5 - Acurácia seletiva r_{gg} estimada para os modelos: Modelo Misto Unicaracterístico (MMUC) e Modelo Misto Multicaracterístico (MMMC) para os caracteres massa média do tubérculo (MMT, g/ tubérculo), número total de tubérculos por parcela (NTT) e produtividade total de tubérculos (PTT, t ha⁻¹), sob condições sem estresse térmico em três populações clonais diferentes de batata (POP).

Fig.5



Fonte: Elaboração própria (2025).

Após a obtenção dos escores do índice FAI-BLUP, que incluiu os caracteres de produtividade de tubérculos (MMT, NTT e PTT), os modelos MMUC e MMMC foram comparados por meio dos coeficientes de Czekanowski (CC) visando avaliar as estratégias de seleção ($\mathbf{u}_{gMMUC}$ vs $\mathbf{u}_{gMMMC}$). De maneira semelhante às acurácias observadas nos modelos multivariado e univariado (Figura 5), os valores de CC foram similares na concordância na seleção dos 20% melhores clones entre as estratégias ($\mathbf{u}_{gMMUC}$ vs $\mathbf{u}_{gMMMC}$), o que sugere uma alta concordância entre vetores ($\mathbf{u}_{gMMUC}$ com $\mathbf{u}_{gMMMC}$) de ordenamento dos clones (Tabela 3).

Tabela 3 - Coeficiente de Czekanowski (CC) entre os vetores de pontuação do índice FAI-BLUP para seleção dos 20% melhores clones das estratégias ($\mathbf{u}_{gMMUC}$ e $\mathbf{u}_{gMMMC}$) para seleção simultânea dos componentes de produção de batata: massa média do tubérculo (MMT, g/ tubérculo), número total de tubérculos por parcela (NTT) e produtividade total de tubérculos (PTT, t ha⁻¹).

Estratégias [†]	Populações [§]	CC
$\mathbf{u}_{gMMUC}$ vs $\mathbf{u}_{gMMMC}$	POP1	0,88
	POP2	0,90
	POP3	0,90

[†] $\mathbf{u}_{gMMUC}$ é o vetor de efeitos de clones do Modelo Misto unicaracterístico (MMUC), $\mathbf{u}_{gMMMC}$ é o vetor de efeitos de clones do Modelo Misto Multicaracterístico (MMMC).

[§]Identificação das populações: POP1, POP2 e POP3, em que os códigos POP1, POP2 e POP3 identificam as diferentes populações clonais.

Fonte: Elaboração própria (2025).

4 DISCUSSÃO

O início do melhoramento de batata envolve o cruzamento biparental de genótipos heterozigotos para o desenvolvimento da primeira geração clonal (F₁), na qual os clones serão selecionados com base em seu desempenho *per se*, e propagados vegetativamente, seguindo um esquema de melhoramento clonal (Stich; Van Inghelandt, 2018). Contudo, o número de indivíduos selecionados para avaliação nessas etapas é elevado o que inviabiliza a avaliação fenotípica de múltiplas características quantitativas (Feldman *et al.*, 2024).

Os principais atributos de um clone de batata são de natureza quantitativa, ou seja, influenciadas por centenas de genes. Os programas de melhoramento de batata realizam a fenotipagem simultânea de dezenas de características com o intuito de selecionar genótipos superiores. Apesar da importância desses caracteres, a compreensão completa da sua herança é limitada pela dificuldade operacional de mensurá-la quantitativamente, com precisão e em escala, em cada fase do programa de melhoramento. O uso de imagens digitais pode auxiliar nas estimativas e estratégias adotadas no melhoramento de batata.

Neste sentido, a utilização de imagens digitais permite o desenvolvimento de um trabalho semiautomatizado para capturar dados, armazenar e mensurar cada uma das características alvo por meio de visão computacional (Feldman *et al.*, 2024). A abordagem utilizada possibilitou a mensuração, em escala e com baixo custo, dos caracteres NTT e MMT. Desta maneira, foi viável adotar a análise multivariada do MMMC para componentes de produção de tubérculos, bem como analisá-los no modelo MMUC.

O modelo MMMC apresentou consistência das correlações genótípicas entre as três populações de PTT para NTT e MMT demonstraram robustez dessas associações genéticas entre diferentes backgrounds genéticos (Figura 4). A forte correlação positiva entre PTT e ambos os componentes (NTT e MMT) permite que melhoristas utilizem seleção indireta em gerações iniciais, uma vez que, nesta fase é caracterizada pelo grande número de genótipos avaliados e de forma onerosa (Stich; Van Inghelandt, 2018).

Todavia, a associação negativa observada entre NTT e MMT sugere que a definição de ideótipos deve ser considerado no objetivo do programa de melhoramento, obtendo populações melhoradas para NTT e MMT separadamente e depois inter cruzadas explorando a heterose e quebra de ligação gênica (Haynes *et al.*, 2012; Slater *et al.*, 2016). Além da alocação de recursos divergentes essa independência genética também pode ser explicada pela regulação hormonal a qual influencia diretamente o número e a massa média de tubérculos. Altos níveis de giberelinas tendem a reduzir a quantidade de tubérculos formados, pois mantêm os estolões no crescimento vegetativo. Por outro lado, a redução das giberelinas e aumento de ácido abscísico (ABA) e citocininas favorecem a indução e o desenvolvimento dos tubérculos, impactando positivamente o número e a massa média de tubérculos (Xu *et al.*, 1998; Aksenova *et al.*, 2012; Saidi; Hajibarat, 2021).

A correlação residual negativa entre NTT e MMT reflete uma associação inversa induzida pelo ambiente, possivelmente devido à competição por recursos. Em condições de alta densidade de tubérculos, a disponibilidade limitada de fotoassimilados pode reduzir o crescimento individual de tubérculos e consequentemente sua massa média (MMT) (Alam *et al.*, 2016). Estresses abióticos como deficiência hídrica ou nutricional pode exacerbar essa competição, como observado em estudos com irrigação controlada (Tsegaw; Hammes; Robbertse, 2005). Em consonância, Bradshaw (2022) reportou ρ_e associação negativa entre NTT e MMT em ensaios com restrição de nitrogênio e Cuevas *et al.* (2023) notaram que ρ_e negativas são mais pronunciadas em ambientes subótimos, reforçando a influência ambiental para estes caracteres de produtividade de tubérculos.

As correlações genéticas e residuais entre PTT e NTT foram igualmente altas e positivas, indicando uma base genética compartilhada, no qual loci pleiotrópicos ou ligados controlam ambos os caracteres, como genes reguladores do ciclo celular (Endelman *et al.*, 2018). Além disso observar-se uma resposta ambiental sincronizada, fatores como irrigação e adubação aumentam simultaneamente NTT e PTT (Fantaw *et al.*, 2019). A alta correlação residual sugere que a seleção indireta para NTT é eficaz para melhorar PTT, mesmo em diferentes ambientes.

Os modelos MMMC e MMUC apresentaram acurácias seletivas altas e similares para os caracteres avaliados (MMT, NTT, PTT), mesmo com o MMMC exibindo ajuste significativamente melhor. Estudos como o de Cuevas *et al.* (2023) reforçam que o MMMC é particularmente vantajoso quando os caracteres têm correlações genóticas moderadas a altas ($\rho_g > 0.5$), mas seu impacto na acurácia é limitado para caracteres com alta herdabilidade (Endelman *et al.*, 2018), e delineamentos experimentais com maior controle da variação ambiental como por exemplo p-rep (Cullis *et al.*, 2006).

Neste sentido, o emprego de análises multivariadas em comparação a univariadas no melhoramento de batata tetraploide, auxilia na estimativa de valores genéticos (Endelman *et al.*, 2018; Cuevas *et al.*, 2023). Todavia, para análises de ensaios únicos (*Single-trial*) e sem estresse por calor, não foram observados incremento nos componentes de variância, bem como em acurácia seletiva (Figuras 2 e 5, respectivamente).

Assim, a abordagem de análises multivariada aproveita as correlações entre os caracteres, resultando no maior poder preditivo para caracteres de baixa herdabilidade. Contudo, quando a herdabilidade é alta e a variância genética é predominante o modelo MMUC consegue estimar valores genéticos com maior precisão (Endelman *et al.*, 2018).

5 CONCLUSÃO

Este estudo demonstrou que, mesmo considerando a penalização pelo maior número de parâmetros, o modelo MMMC apresentou melhor ajuste ao modelo MMUC, tendo AIC inferior a 10 unidades para todas as populações. O modelo MMMC mostrou-se mais eficiente que o MMUC por proporcionar estimativas mais precisas dos componentes de variância genética, além de um controle ambiental mais eficaz. As correlações parciais genóticas e residuais em batata confirmaram a relação positiva entre PTT para NTT e MMT, bem como relação inversa entre MMT e NTT. Os componentes de produção MMT, NTT e PTT apresentaram alta acurácia seletiva em ambos os modelos MMUC e MMMC.

REFERÊNCIAS

- AKAIKE, H. Uma nova perspectiva sobre a identificação de modelos estatísticos. **IEEE Transactions on Automatic Control**, **AC-19**, p. 716-723, 1974.
- AKSENOVA, N. P. *et al.* Hormonal regulation of tuber formation in potato plants. **Russian Journal of Plant Physiology**, v. 59, p. 451-466, 2012.
- AKSOY, E. *et al.* (2021). Recent advances in potato (*Solanum tuberosum* L.) breeding. In: AL-KHAYRI, J. M.; JAIN, S. M.; JOHNSON, D. V. (ed). **Advances in Plant Breeding Strategies: Vegetable Crops**. v. 8. Suíça: Springer, 2021. p. 409-487.
- ALAM, S. *et al.* Optimizing crop geometry for processing-grade tuber yield, quality and economics of potato in Grey Terrace Soil. **Archives of Agronomy and Soil Science**, v. 62, n. 11, p. 1496-1507, 2016.
- ANDRIVON, D. Potato facing global challenges: how, how much, how well?. **Potato Research**, v. 60, n. 3, p. 389-400, 2017.
- BONIERBALE, M. W. Potato breeding. In: CAMPOS, H.; ORTIZ, O. **The potato crop: Its agricultural, nutritional and social contribution to humankind**. Suíça: Springer, 2020. p.163-217.
- BRADSHAW, J. E. A Brief History of the Impact of Potato Genetics on the Breeding of Tetraploid Potato Cultivars for Tuber Propagation. **Potato Research**, v. 65, p. 461–501, 2022.
- BRADSHAW, J. E. Revisão e análise das limitações nas formas de melhorar o melhoramento convencional da batata. **Potato Research**. 60, 171–193, 2017.
- CABELLO, R. *et al.* Heritability of yield components under irrigated and drought conditions in andigenum potatoes. **American Journal of Potato Research**, v. 91, n. 5, p. 492-499, 2014.
- CUEVAS, J. *et al.* Modeling genotype× environment interaction for single and multitrait genomic prediction in potato (*Solanum tuberosum* L.). **G3**, v. 13, n. 2, p. 1-13, 2023.
- CULLIS B. R.; SMITH, A. B.; COOMBES, N. E. On the design of early generation variety trials with correlated data. **Journal Agric Biol Environment Stat** v. 11, p. 381–393, 2006.
- DEVAUX, A. *et al.* The potato of the future: opportunities and challenges in Sustainable agri-food systems. **Potato Research**, v. 64, n. 4, p. 681-720, 2021.
- DOUCHES, D.S. *et al.* Assessment of potato breeding progress in the USA over the last century. **Crop Science**, v. 36, n. 6, p. 1544–1552, 1996.
- ENDELMAN, J. B. *et al.* Genetic variance partitioning and genome-wide prediction with allele dosage information in autotetraploid potato. **Genetics**, v. 209, n. 1, p. 77-87, 2018.
- FANTAW, S. *et al.* Evaluation of potato (*Solanum tuberosum* L.) varieties for yield and yield components. **Journal of Horticulture and Forestry**, v. 11, n. 3, p. 48-53, 2019.

- FELDMAN, M. J. *et al.* A scalable, low-cost phenotyping strategy to assess tuber size, shape, and the colorimetric features of tuber skin and flesh in potato breeding populations. **The Plant Phenome Journal**, v. 7, n. 1, p. 1-20, 2024.
- HAYNES, K. G.; GERGELA, D. M.; HUTCHINSON, C. M. Early generation selection at multiple locations may identify potato parents that produce more widely adapted progeny. **Euphytica**, v. 186, n. 2, p. 573-583, 2012.
- HENDERSON, C. R. *et al.* The Estimation of Environmental and Genetic Trends from Records Subject to Culling. **Biometrics**, v. 15, v. 2, p. 192-218, 1959.
- GILMOUR, A.R. **Echidna Mixed Model Software**, 2021. Disponível em: www.echidnamms.org/ Acesso em: 10 mar 2021.
- GOPAL, J. Challenges and Way-forward in Selection of Superior Parents, Crosses, and Clones in Potato Breeding. **Potato Research**, v. 58, p. 165–188, 2015.
- MEYER, K. Likelihood calculations to evaluate experimental designs to estimate genetic variances. **Heredity (Edinb)** v. 101, p. 212–221, 2008.
- ORTIZ, R. *et al.* Genetic gains in potato breeding as measured by field testing of cultivars released during the last 200 years in the Nordic Region of Europe. **The Journal of Agricultural Science**, v. 160, n. 5, p. 310-316, 2022.
- OTSU, N. A threshold selection method from gray-level histograms. **Automatica**, v. 11, n. 285-296, p. 23-27, 1975.
- PATTERSON, H. D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v. 58, p. 545–554, 1971.
- QIAO, C. G. *et al.* Evaluation of experimental designs and spatial analyses in wheat breeding trials. **Theoretical and Applied Genetics**, v. 100, p. 9–16. 2000.
- R CORE TEAM. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria, 2025. Disponível em: <https://www.r-project.org/> Acesso em: 10 fev. 2025.
- SAIDI, A.; HAJIBARAT, Z. Phytohormones: plant switchers in developmental and growth stages in potato. **Journal of Genetic Engineering and Biotechnology**, v. 19, n. 1, p. 89, 2021.
- SCHUMAKER, J. R. *et al.* 'Atlantic': A New Potato Cultivar for North Florida. **Proceedings of the Florida State Horticultural Society**, v. 90, p. 370-372, 1977.
- SAS INSTITUTE INC. SAS/STAT® 14.2 User's Guide. **Software** (Cary, NC: SAS Institute Inc.), 2016.
- SLATER, A. T *et al.* Improving genetic with genomic selection in autotetraploid potato. **The Plant Genome**, v. 9, n. 3, p. 1-15, 2016.

STICH, B.; VAN INGHELANDT, D. Prospects and potential uses of genomic prediction of key performance traits in tetraploid potato. **Frontiers in Plant Science**, v. 9, n. 159, p. 1-12, 2018.

THELEN, K. *et al.* Variance and covariance components of agronomic and quality traits assessed in tetraploid potato and their implications on practical breeding. **Frontiers in Plant Science**, 15, p. 1-15, 2025.

TSEGAW, T.; HAMMES, S.; ROBBERTSE, J. Paclobutrazol-induced leaf, stem, and root anatomical modifications in potato. *HortScience*, v. 40, n. 5, p. 1343-1346, 2005.

WICKHAM, H. “**Data Analysis,**” in **ggplot2**. Use R! Springer Verlag, New York, 2016.

XU, X. *et al.* The role of gibberellin, abscisic acid, and sucrose in the regulation of potato tuber formation in vitro. **Plant physiology**, v. 117, n. 2, p. 575-584, 1998.

MATERIAL SUPLEMENTAR

Tabelas suplementares

Tabela S1 - Resumo do Modelo Misto Unicaracterístico (MMUC), Modelo Misto Multicaracterístico (MMMC): Ponto máximo residual log-likelihood (ℓ) e Akaike Critério de Informação de Akaike (AIC).

População [‡]	Modelo	p	ℓ	AIC
POP1	MMUC	12	-2706,54	5437,07
	MMMC	18	-2387,55	4811,10
POP2	MMUC	12	-4333,61	8691,23
	MMMC	18	-3809,20	7654,39
POP3	MMUC	12	-1376,65	2777,30
	MMMC	18	-1188,28	2412,56

[‡]Identificação de população: POP1, POP2 e POP3, em que os códigos POP1, POP2 e POP3 identificam as diferentes populações clonais.

Fonte: Elaboração própria (2025).

Tabela S2 - Resumo dos componentes de variância do Modelo Misto Unicaracterístico (MMUC), Modelo misto Multicaracterístico (MMMC), para os caracteres massa média de tubérculos (MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT, t ha⁻¹) na safra de inverno de 2023.

Caráter	População [†]	MMUC [‡]				MMMC [‡]			
		σ_r^2	σ_c^2	σ_g^2	σ^2	σ_r^2	σ_c^2	σ_g^2	σ^2
MMT	POP1	24,31	45,03	504,14	337,05	0,00	0,00	518,38	416,34
	POP2	14,55	19,77	670,93	255,42	0,00	1,71	639,80	323,03
	POP3	0,00	37,94	903,66	513,94	0,00	0,00	898,24	555,29
NTT	POP1	4,21	27,69	177,99	109,87	0,00	0,00	190,36	134,07
	POP2	3,98	7,46	93,54	70,85	0,12	0,11	91,46	84,81
	POP3	9,57	0,00	43,51	67,17	0,46	0,00	50,88	69,43
PTT	POP1	4,64	27,17	157,21	60,40	0,00	0,00	156,63	91,04
	POP2	4,95	3,26	94,05	41,08	0,00	0,32	96,29	47,66
	POP3	10,08	0,00	103,79	45,16	0,46	0,00	105,91	52,66

[†]Identificação de população: POP1, POP2 e POP3, em que os códigos POP1, POP2 e POP3 identificam as diferentes populações clonais.

[‡] σ_r^2 , σ_c^2 , σ_g^2 e σ^2 : componentes de variância associados aos efeitos de linha, coluna, clone e resíduo.

Fonte: Elaboração própria (2025).

Tabela S3 - Contribuição (%) dos componentes de variância do Modelo Misto Unicaracterístico (MMUC), Modelo Misto Multicaracterístico (MMMC) para os caracteres: massa média de tubérculos (MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT, t ha⁻¹) na safra de inverno de 2023.

Carácter	População [†]	MMUC				MMMC			
		Linha	Coluna	Clone	Res.	Linha	Coluna	Clone	Res.
MMT	POP1	2,29 [*]	4,61 [*]	55,74 ^{**}	37,37	0,00	0,00	55,46	44,54
	POP2	1,51 [*]	2,06 [*]	69,84 ^{**}	26,59	0,00	0,18	66,33	33,49
	POP3	0,00 ^{ns}	2,61 ^{ns}	62,08 ^{**}	35,31	0,00	0,00	61,80	38,20
NTT	POP1	1,08 ^{ns}	8,25 [*]	56,18 ^{**}	34,49	0,00	0,00	58,67	41,33
	POP2	2,26 [*]	4,24 ^{**}	53,20 ^{**}	40,30	0,07	0,06	51,82	48,05
	POP3	7,96 [*]	0,00 ^{ns}	36,18 ^{**}	55,86	0,38	0,00	42,13	57,49
PTT	POP1	1,76 [*]	10,76 ^{**}	63,19 ^{**}	24,29	0,00	0,00	63,24	36,76
	POP2	3,45 ^{**}	2,28 [*]	65,61 ^{**}	28,66	0,00	0,22	66,74	33,04
	POP3	6,34 [*]	0,00 ^{ns}	65,26 ^{**}	28,40	0,29	0,00	66,60	33,11

[†]Identificação de população: POP1, POP2 e POP3, em que os códigos POP1, POP2 e POP3 identificam as diferentes populações clonais.

Significância pelo teste da razão de verossimilhança (LRT), p-valor < 0.01^{***}, 0.05^{**} e p-valor > 0.05 não significativo^{ns}.

Fonte: Elaboração própria (2025).

Tabela S4 - Resumo das estimativas dos parâmetros de variância genética e não genética para o Modelo Misto Multicaracterístico (MMMC) para populações clonais. Limite inferior de confiança (LIC) e limite superior de confiança (LSC) do intervalo de confiança de 95% para os parâmetros de variância associados às características massa média de tubérculos (MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT, t ha⁻¹) na safra de inverno de 2023.

Parâmetros [†]	POP1 [‡]			POP2 [‡]			POP3 [‡]		
	LIC	Estimativas	LSC	LIC	Estimativas	LSC	LIC	Estimativas	LSC
$\sigma_{r_{ATM}}^2$									
$\sigma_{r_{TTN}}^2$				0,01	0,12	$8,26 \times 10^{+7}$	0,06	0,46	$1,28 \times 10^{+8}$
$\sigma_{r_{TTY}}^2$							0,06	0,46	$1,05 \times 10^{+6}$
$\sigma_{c_{ATM}}^2$				0,19	1,71	$3,20 \times 10^{+11}$			
$\sigma_{c_{TTN}}^2$				0,01	0,11	$1,05 \times 10^{+73}$			
$\sigma_{c_{TTY}}^2$				0,06	0,32	707,20			
$\sigma_{g_{ATM}}^2$	367,41	518,38	786,54	511,62	639,8	823,28	596,26	898,24	1506,51
$\sigma_{g_{TTN}}^2$	138,60	190,36	277,84	69,55	91,46	125,72	29,11	50,88	110,83
$\sigma_{g_{TTY}}^2$	117,60	156,63	219,01	77,75	96,29	122,41	72,54	105,91	169,17
$\sigma_{e_{ATM}}^2$	311,00	416,34	586,27	262,33	323,03	407,67	383,14	555,29	877,02
$\sigma_{e_{TTN}}^2$	101,23	134,07	186,06	69,35	84,81	106,12	49,04	69,43	105,91
$\sigma_{e_{TTY}}^2$	69,82	91,04	123,70	39,10	47,66	59,41	36,53	52,66	82,50
$\rho_{g_{ATM.TTY}}$	0,32	0,51	0,69	0,60	0,70	0,79	0,55	0,74	0,93
$\rho_{g_{TTN.TTY}}$	0,77	0,85	0,92	0,59	0,69	0,79	0,67	0,80	0,94
$\rho_{g_{TTN.ATM}}$	-0,25	0,01	0,26	-0,18	0,01	0,19	-0,18	0,22	0,62
$\rho_{e_{ATM.TTY}}$	0,30	0,47	0,64	0,25	0,38	0,51	-0,01	0,25	0,52
$\rho_{e_{TTN.TTY}}$	0,66	0,76	0,85	0,65	0,72	0,80	0,69	0,79	0,90
$\rho_{e_{TTN.ATM}}$	-0,33	-0,12	0,09	-0,41	-0,27	-0,13	-0,56	-0,31	-0,05

[†] $\sigma_{r_{MMT}}^2$, $\sigma_{r_{NTT}}^2$, $\sigma_{r_{PTT}}^2$, $\sigma_{c_{MMT}}^2$, $\sigma_{c_{NTT}}^2$, $\sigma_{c_{PTT}}^2$, $\sigma_{g_{MMT}}^2$, $\sigma_{g_{NTT}}^2$, $\sigma_{g_{PTT}}^2$, $\sigma_{e_{MMT}}^2$, $\sigma_{e_{NTT}}^2$ e $\sigma_{e_{PTT}}^2$ - Componentes de variância associados aos efeitos de linha, coluna, clone e resíduo, para os caracteres: MMT, NTT e PTT, respectivamente;

$\rho_{g_{ATM.TTY}}$, $\rho_{g_{TTN.TTY}}$, $\rho_{g_{TTN.ATM}}$, $\rho_{e_{ATM.TTY}}$, $\rho_{e_{TTN.TTY}}$ e $\rho_{e_{TTN.ATM}}$ - Correlações genotípicas e residuais entre os pares de caracteres para o efeito de clone e resíduo (ρ_g , ρ_e), para os caracteres MMT, NTT e PTT.

[‡] Identificação de população: POP1, POP2 e POP3, em que os códigos POP1, POP2 e POP3 identificam as diferentes populações clonais.

Fonte: Elaboração própria (2025).

Tabela S5 - Acurácia seletiva para o Modelo Misto Unicaracterístico (MMUC), Modelo Misto Multicaracterístico (MMMC) para os caracteres massa média de tubérculos (MMT), número total de tubérculos (NTT) e produtividade total de tubérculos (PTT, t ha⁻¹) na safra de inverno de 2023.

População [†]	Caráter	Modelos	
		$r_{gg}^{(MMUC)}$	$r_{gg}^{(MMMC)}$
POP1	MMT	0,79	0,79
POP2	MMT	0,88	0,87
POP3	MMT	0,83	0,85
Média		0,83	0,84
POP1	NTT	0,8	0,82
POP2	NTT	0,8	0,79
POP3	NTT	0,68	0,75
Média		0,76	0,79
POP1	PTT	0,86	0,83
POP2	PTT	0,87	0,87
POP3	PTT	0,86	0,86
Média		0,86	0,85

[†]Identificação de população: POP1, POP2 e POP3, em que os códigos POP1, POP2 e POP3 identificam as diferentes populações clonais.

Fonte: Elaboração própria (2025).

2 CONSIDERAÇÕES FINAIS

A presente tese teve como propósito contribuir significativamente para o avanço metodológico, científico e prático do melhoramento genético de batata, abordando os desafios enfrentados pelos melhoristas, especialmente em programas conduzidos em condições tropicais e subtropicais. Os resultados obtidos demonstram que a integração de abordagens estatísticas robustas em ensaios de local único e multiambientes, bem como análises unicaracterístico e multicaracterístico, aumenta a eficiência seletiva de clones para componentes de produção de batata.

O primeiro estudo demonstrou de maneira consistente e robusta a relevância da incorporação do efeito de família na estrutura de modelos de seleção clonal. A estrutura aninhada (Família/Clone = Família + Família·Clone) permitiu uma partição mais realista dos componentes de variância genética, evitando a confusão entre os efeitos de família e de clones dentro das famílias que ocorre em modelos tradicionais. Esta abordagem resultou em incrementos consistentes na acurácia seletiva, com ganhos médios de 11% e 7% para gravidade específica e produtividade total de tubérculos, respectivamente, nas análises de ensaios únicos, e de 12% e 6% para análises multiambientes.

A magnitude do efeito de família mostrou-se variável entre caracteres e ambientes, com a proporção da variância genética total atribuída ao efeito de família (ρ_S) variando de 0,03 a 0,35. Observou-se redução expressiva do efeito de família sob condições de estresse térmico elevado para produtividade total de tubérculos e de menor proporção para gravidade específica. Este padrão sugere que diferentes mecanismos genéticos e fisiológicos regulam a resposta ao estresse térmico.

A análise multiambientes revelou a complexidade da interação genótipo $\times$ ambiente (G $\times$ E) na cultura da batata sob diferentes níveis de estresse térmico. Os resultados evidenciaram uma natureza de interação complexa para produtividade total de tubérculos e uma interação simples para gravidade específica. O uso do índice FAI-BLUP para integração de informações de múltiplos ambientes e caracteres mostrou-se eficaz na seleção de clones superiores. Assim, a incorporação do efeito de família é vantajosa e não implica em custos adicionais, aproveitando os dados já coletados proporcionando o aumento da eficiência seletiva e a acurácia na identificação de clones superiores.

O segundo estudo, com auxílio da fenotipagem por imagens digitais, proporcionou a estimativa em larga escala de componentes de produção de batata e de baixo custo. O uso de imagens viabilizou a quantificação precisa e eficiente do número total de tubérculos e massa

média de tubérculos, especialmente em populações clonais numerosas encontradas em estágios iniciais de melhoramento. Esta abordagem permite a avaliação de múltiplos caracteres, em escala e sem aumento proporcional de custos.

A comparação entre modelos mistos unicaracterístico (MMUC) e multicaracterístico (MMMC) revelou que, embora ambos apresentem acurácias seletivas elevadas e similares para os componentes de produção avaliados, o modelo MMMC proporcionou melhor ajuste estatístico, evidenciado por valores de AIC inferiores.

A consistência das correlações genéticas entre as três populações avaliadas demonstra a robustez destas associações genéticas em diferentes *backgrounds* genéticos. As acurácias seletivas obtidas foram consistentemente elevadas ($> 0,70$), refletindo a eficiência do delineamento experimental parcialmente replicado (p-rep) adotado e a adequação dos modelos estatísticos empregados.

Integrando os resultados dos dois estudos emergiu uma visão abrangente de estratégias com maior eficiência para programas de melhoramento de batata em condições tropicais e subtropicais. A combinação da inclusão do efeito de família e das análises multivariadas de componentes de produção de batata com auxílio de imagens digitais aumenta, simultaneamente, a acurácia seletiva de clones em populações numerosas dos estágios iniciais dos programas de melhoramento de batata.

A incorporação do efeito de família nos modelos de seleção clonais deve ser considerada rotineiramente, especialmente para caracteres com maior contribuição da variância entre famílias (gravidade específica). A utilização de fenotipagem por imagens deve ser considerada para análises multivariadas de componentes de produção de batata, permitindo avaliações em larga escala e a baixo custo.

Em conclusão, esta tese demonstrou que a integração de metodologias estatísticas robustas e o uso de análises multivariadas com auxílio da fenotipagem por imagem podem promover o aumento da eficiência seletiva de clones em programas iniciais de melhoramento de batata. Os resultados obtidos são significativos para atender à crescente demanda por novos clones de batata adaptados às condições tropicais e subtropicais, contribuindo para a segurança alimentar mundial e na sustentabilidade da produção do tubérculo diante de cenários de mudanças climáticas.

REFERÊNCIAS

- AKAIKE, H. Uma nova perspectiva sobre a identificação de modelos estatísticos. **IEEE Transactions on Automatic Control**, AC-19, p. 716-723, 1974.
- AKSENOVA, N. P. *et al.* Hormonal regulation of tuber formation in potato plants. **Russian Journal of Plant Physiology**, v. 59, p. 451-466, 2012.
- AKSOY, E. *et al.* (2021). Recent advances in potato (*Solanum tuberosum* L.) breeding. In: AL-KHAYRI, J. M.; JAIN, S. M.; JOHNSON, D. V. (ed). **Advances in Plant Breeding Strategies: Vegetable Crops**. v. 8. Suíça: Springer, 2021. p. 409-487.
- ALAM, S. *et al.* Optimizing crop geometry for processing-grade tuber yield, quality and economics of potato in Grey Terrace Soil. **Archives of Agronomy and Soil Science**, v. 62, n. 11, p. 1496-1507, 2016.
- AMADEU, R. R. *et al.* Impact of dominance effects on autotetraploid genomic prediction. **Crop Science**, v. 60, p. 656–665, 2020.
- ANDRIVON, D. Potato facing global challenges: how, how much, how well?. *Potato Research*, v. 60, n. 3, p. 389-400, 2017.
- BONIERBALE, M. W. Potato breeding. In: CAMPOS, H.; ORTIZ, O. **The potato crop: Its agricultural, nutritional and social contribution to humankind**. Suíça: Springer, 2020. p.163-217.
- BRADSHAW, J. E. A Brief History of the Impact of Potato Genetics on the Breeding of Tetraploid Potato Cultivars for Tuber Propagation. **Potato Research**, v. 65, p. 461–501, 2022.
- BRADSHAW, J. E. Revisão e análise das limitações nas formas de melhorar o melhoramento convencional da batata. **Potato Research**. 60, 171–193, 2017.
- CABELLO, R. *et al.* Heritability of yield components under irrigated and drought conditions in andigenum potatoes. **American Journal of Potato Research**, v. 91, n. 5, p. 492-499, 2014.
- CUEVAS, J. *et al.* Modeling genotype× environment interaction for single and multitrait genomic prediction in potato (*Solanum tuberosum* L.). **G3**, v. 13, n. 2, p. 1-13, 2023.
- CULLIS B. R.; SMITH, A. B.; COOMBES, N. E. On the design of early generation variety trials with correlated data. **Journal Agric Biol Environment Stat** v. 11, p. 381–393, 2006.
- DEVAUX, A. *et al.* The potato of the future: opportunities and challenges in Sustainable agri-food systems. **Potato Research**, v. 64, n. 4, p. 681-720, 2021.
- DOUCHES, D.S. *et al.* Assessment of potato breeding progress in the USA over the last century. **Crop Science**, v. 36, n. 6, p. 1544–1552, 1996.
- ENDELMAN, J. B. *et al.* Genetic variance partitioning and genome-wide prediction with allele dosage information in autotetraploid potato. **Genetics**, v. 209, n. 1, p. 77-87, 2018.

FANTAW, S. *et al.* Evaluation of potato (*Solanum tuberosum* L.) varieties for yield and yield components. **Journal of Horticulture and Forestry**, v. 11, n. 3, p. 48-53, 2019.

FELDMAN, M. J. *et al.* A scalable, low-cost phenotyping strategy to assess tuber size, shape, and the colorimetric features of tuber skin and flesh in potato breeding populations. **The Plant Phenome Journal**, v. 7, n. 1, p. 1-20, 2024.

FERNANDES FILHO, C. C. *et al.* (2021). Selection of potato clones for heat tolerance and resistance to potato viruses X and Y for processing purposes. **Crop Science**, 61, 552–565, 2021.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). **FAOSTAT**. 2023. Disponível em: www.fao.org/faostat/en/#data/QCL Acesso em: 15 maio 2025.

HAYNES, K. G.; GERGELA, D. M.; HUTCHINSON, C. M. Early generation selection at multiple locations may identify potato parents that produce more widely adapted progeny. **Euphytica**, v. 186, n. 2, p. 573-583, 2012.

HENDERSON, C. R. *et al.* The Estimation of Environmental and Genetic Trends from Records Subject to **Culling**. **Biometrics**, v. 15, v. 2, p. 192-218, 1959.

GILMOUR, A.R. **Echidna Mixed Model Software**, 2021. Disponível em: www.echidnamms.org/ Acesso em: 10 mar 2021.

GOPAL, J. Challenges and Way-forward in Selection of Superior Parents, Crosses, and Clones in Potato Breeding. **Potato Research**, v. 58, p. 165–188, 2015.

KUNKEL, R.; CAMPBELL, G. S. Maximum potential potato yield in the Columbia Basin, USA: Model and measured values. **American Potato Journal**, 64, 355–366, 1987.

MEYER, K. Likelihood calculations to evaluate experimental designs to estimate genetic variances. **Heredity (Edinb)** v. 101, p. 212–221, 2008.

ORTIZ, R. *et al.* Genetic gains in potato breeding as measured by field testing of cultivars released during the last 200 years in the Nordic Region of Europe. **The Journal of Agricultural Science**, v. 160, n. 5, p. 310-316, 2022.

OTSU, N. A threshold selection method from gray-level histograms. **Automatica**, v. 11, n. 285-296, p. 23-27, 1975.

PAGET, M. F. *et al.* Replicate allocation to improve selection efficiency in the early stages of a potato breeding scheme. *Euphytica* 213, 221, 2017.

PATÍÑO-TORRES, A. J. *et al.* (2021). Performance of superior potato clones under high and mild temperatures in tropical climate. **Agron Journal**, 113, 2349–2360, 2021.

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

QIAO, C. G. *et al.* Evaluation of experimental designs and spatial analyses in wheat breeding trials. **Theoretical and Applied Genetics**, v. 100, p. 9–16. 2000.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria, 2025. Disponível em: <https://www.r-project.org/> Acesso em: 10 fev. 2025.

SAIDI, A.; HAJIBARAT, Z. Phytohormones: plant switchers in developmental and growth stages in potato. **Journal of Genetic Engineering and Biotechnology**, v. 19, n. 1, p. 89, 2021.

SCHUMAKER, J. R. et al. 'Atlantic': A New Potato Cultivar for North Florida. **Proceedings of the Florida State Horticultural Society**, v. 90, p. 370-372, 1977.

SAS INSTITUTE INC. SAS/STAT® 14.2 User's Guide. **Software** (Cary, NC: SAS Institute Inc.), 2016.

SLATER, A. T *et al.* Improving genetic with genomic selection in autotetraploid potato. **The Plant Genome**, v. 9, n. 3, p. 1-15, 2016.

STICH, B.; VAN INGHELANDT, D. Prospects and potential uses of genomic prediction of key performance traits in tetraploid potato. **Frontiers in Plant Science**, v. 9, n. 159, p. 1-12, 2018.

THELEN, K. *et al.* Variance and covariance components of agronomic and quality traits assessed in tetraploid potato and their implications on practical breeding. **Frontiers in Plant Science**, 15, p. 1-15, 2025.

TSEGAW, T.; HAMMES, S.; ROBBERTSE, J. Paclobutrazol-induced leaf, stem, and root anatomical modifications in potato. *HortScience*, v. 40, n. 5, p. 1343-1346, 2005.

VOSS-FELS, K. P. *et al.* Strategies and considerations for implementing genomic selection to improve traits with additive and non-additive genetic architectures in sugarcane breeding. **Theor Appl Genet**, v. 134, p. 1493–1511, 2021.

WICKHAM, H. **“Data Analysis,” in ggplot2**. Use R! Springer Verlag, New York, 2016.

XU, X. et al. The role of gibberellin, abscisic acid, and sucrose in the regulation of potato tuber formation in vitro. **Plant physiology**, v. 117, n. 2, p. 575-584, 1998.