

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E MELHORAMENTO**

JOSÉ HENRIQUE SOLER GUILHEN

**SELEÇÃO DE MODELOS E A INFLUÊNCIA DOS TESTADORES NA
SELEÇÃO DE HÍBRIDOS DE MILHO DE ENSAIOS PRELIMINARES**

**ALEGRE-ES
2016**

JOSÉ HENRIQUE SOLER GUILHEN

**SELEÇÃO DE MODELOS E A INFLUÊNCIA DOS TESTADORES NA
SELEÇÃO DE HÍBRIDOS DE MILHO DE ENSAIOS PRELIMINARES**

Dissertação apresentada ao Programa de Pós-Graduação em Genética e Melhoramento da Universidade Federal do Espírito Santo, como requisito para obtenção do título de mestre em Genética e Melhoramento.

Orientador: Prof. Dr. Adésio Ferreira.

ALEGRE-ES
2016

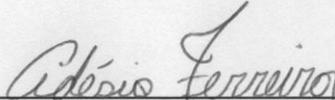
JOSÉ HENRIQUE SOLER GUILHEN

**COMPARAÇÃO DE MODELOS GENÉTICO-ESTATÍSTICOS PARA A
SELEÇÃO DE HÍBRIDOS DE MILHO EM ENSAIOS PRELIMINARES**

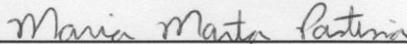
Dissertação apresentada ao Programa de Pós-Graduação em Genética e Melhoramento da Universidade Federal do Espírito Santo, como requisito para obtenção do título de mestre em Genética e Melhoramento.

Aprovada: 30/06/2016

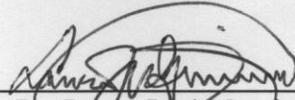
COMISSÃO EXAMINADORA



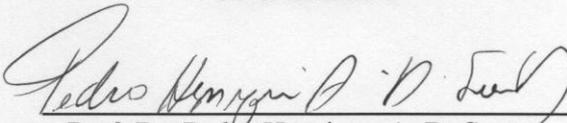
Prof. Dr. Adésio Ferreira
Universidade Federal do Espírito Santo
Orientador



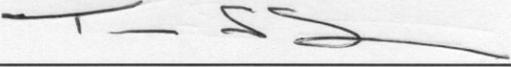
Pesq (a). Dr (a). Maria Marta Pastina
Embrapa Milho e Sorgo
Co-orientadora



Pesq. Dr. Lauro José Moreira Guimarães
Embrapa Milho e Sorgo
Co-orientador



Prof. Dr. Pedro Henrique A. D. Santos
Pós-Doutorado UENF
Membro externo



Prof. Dr. Tércio da Silva de Souza
Instituto Federal do Espírito Santo
Membro externo

Dedico este trabalho á minha família e, em especial, aos meus pais Luis Henrique Guilhen e Marisa Ferraro Soler Guilhen, a minha irmã Ana Carolina Soler Guilhen, a minha namorada Ingrid Schimidt Kaiser, pelo amor e confiança de todos.

AGRADECIMENTOS

À Deus!

À Universidade Federal do Espírito Santo e ao Programa de Pós-Graduação em Genética e Melhoramento, pela oportunidade de realizar o mestrado.

Aos órgãos Fundação de Amparo a Pesquisa do Espírito Santo (FAPES), Conselho Nacional de Pesquisa e Desenvolvimento Científico e Tecnológico (CNPq) e Universidade Federal do Espírito Santo (UFES) pela concessão da bolsa e apoio financeiro e a Embrapa pela parceria na execução do trabalho.

Ao meu orientador Dr. Adésio Ferreira, pela orientação, ao apoio incondicional e principalmente pela amizade durante o curso.

Aos Coorientadores Dr. Lauro José Moreira Guimarães e Dr (a). Maria Marta Pastina, pelo apoio, auxílio no desenvolvimento dessa pesquisa e disponibilidade sempre prestadas.

A professora Dr (a). Marcia Flores Ferreira, pelo apoio incondicional e principalmente pela amizade durante o curso.

Aos outros membros da banca o Dr. Tércio da Silva de Souza e Dr. Pedro Henrique Araújo Diniz Santos.

Aos meus pais, Luis Henrique Guilhen e Marisa Ferraro Soler Guilhen e minha irmã Ana Carolina Soler Guilhen, pelo apoio, dedicação, compreensão nos momentos difíceis, ao amor, aos esforços e sacrifícios que tiveram até eu atingir mais esta etapa de minha vida.

A minha namorada Ingrid Schimidt Kaiser, pela compreensão nos momentos de falta de tempo e atenção e por sempre me motivar.

Ao meu amigo Tiago de Souza Marçal, pelo apoio nos momentos de dificuldade, pelos trabalhos realizados em parceria durante a graduação e o mestrado.

Aos meus amigos Wagner Bastos dos Santos Oliveira, Rafael Fonseca Zanotti e Lorena Contarini Machado, pelas conversas e trabalhos realizados em parcerias.

Aos amigos do Laboratório de Genética e Melhoramento Vegetal e Biometria: Guilherme, Ramon, Cintia, Lidiane, Ana Beatriz, Liana, Clemilton, Sephora, Sabrina, Paula Henrique, Luina, Paula Mauri, Drielli, Carolina, Matheus e Iana, pelo companheirismo e os momentos de descontração.

A todos os colegas de mestrado e de universidade, pela amizade e o alegre convívio.

A todos aqueles não citados aqui, que de alguma forma estiveram torcendo por mim durante o desenvolvimento deste trabalho.

BIOGRAFIA

JOSÉ HENRIQUE SOLER GUILHEN, filho de Marisa Ferraro Soler Guilhen e Luis Henrique Guilhen, nasceu em 21 de fevereiro de 1989, na Cidade de Tupã, estado de São Paulo. Em 2004, ingressou no Colégio Técnico Agrícola de Pompéia no estado de São Paulo, graduou no ano de 2008 e durante 18 meses (2007-2008) estagiou nos Estados Unidos em uma fazenda agrícola. No 2009, ingressou no curso de Bacharel em Agronomia do Centro de Ciências Agrárias da Universidade Federal do Espírito Santo, vindo a se graduar em julho de 2014, recebendo o título de Bacharel em Agronomia. Em agosto de 2014, ingressou no curso de mestrado do Programa de Pós-Graduação em Genética e Melhoramento do Centro de Ciências Agrárias da Universidade Federal do Espírito Santo (CCA– UFES/Alegre), submetendo-se à defesa de dissertação em junho de 2016.

“O sol nasce para todos, mas os que lutam podem chegar á ele.”

“A maior derrota é não saber levantar depois de um tombo.”

SELEÇÃO DE MODELOS E INFLUÊNCIA DOS TESTADORES NA SELEÇÃO DE HÍBRIDOS DE MILHO DE ENSAIOS PRELIMINARES

RESUMO

As dificuldade que a população poderá passar com a falta de alimento e os desafios para as produção nos anos futuros, vem preocupando e desenvolvendo diversas pesquisas no meio academico sobre o assunto e os possíveis resultados. Dentre as preocupações e esperaça ao mesmo tempo para uma amortização das perspectivas contestualizadas dos pesquisadores se encontra a cultura do milho (*Zea mays*). O milho é uma das principais fontes de alimento de forma direta e indireta para a população mundial. Trabalhos que busquem melhor eficiencia estatísticas, como a tecnica de modelos mistos (REML/BLUP), vem sendo adotados como uma possível amortização da falta alimentar no futuro. Além disto, modelos que explicuer melhor os valores estimados em relação aos observados e ensaios que resultem em híbridos com maiores maiores produtividades são de esta menessidade no panorana atual. Mediante está contestualização os objetivos do presente estudo foi selecionar o modelo que explique melhor os resultados observados e selecionar híbridos altamente produtivos. O experimento foi montado em um delineamento de blocos aumentados, dois ensaios separados, o ensaio 1 com 1801 linhagens hibridizada com o testador G24 e o ensaio 2 formado por 1551 linhagens hibridizado com o testador G8. A análise estatística foi realizada através da metodologia de modelos mistos (REML/BLUP). Foram testados 8 diferentes modelos, divididos em três classes. Os modelos selecionados foram os que apresentaram uma melhor precisão, baseado no log(L), AIC, BIC, acurácia e medidas de tempo e com e sem informação de parentesco. Com os modelos selecionados foram realizados testes de seleção de híbridos dentro de cada ensaio e no conjunto e estudou-se a coincidência entre os ensaios em avaliação. Os melhores ajustes foram para os modelos M3, M6 e M8. Os menores tempos foram respostas dos modelos que não consideram informação de parentesco. A coincidência entre os 5% selecionados fenotipicamente e genotipicamente para todas as análises foi de 0,41 a 0,60. A coincidência entre os diferentes modelos e ensaios foi de 43,45 a 96,67%. Com este trabalho podemos concluir que os melhores modelos de modo geral entre ajustes e tempo foram os modelos M3 e M6 e o ensaio 1 que resultou nos melhores híbridos provenientes no experimento.

Palavras-chave: Modelos Mistos, Algoritimos; *Zea mays*; Heterose; Topcross.

LISTA DE FIGURAS

Capítulo 1

Figura 1. Fixando-se uma intensidade de seleção em 5% os modelos sem informação de parentesco selecionam praticamente os mesmos indivíduos, modelos M1.I, M4.II e M7.III com a coincidência entre eles de 99,40% e suas coincidência com os modelos com informação de parentesco variaram entre 36,90 e 38,10 (Tabela 2). A coincidência entre os modelos que consideraram matriz de parentesco variou entre 88,10 (M2.I-M6.II e M2.I-M8.III) a 97,62 (M6.II-M8.III).....35

Capítulo 2

Figura 1 Dispersão gráfica das médias fenotípicas do rendimento de grãos (Figuras – A1, B1, C1) e acurácias (Figuras – A2, B2, C2) em relação aos valores genotípicos preditos para o ensaio 1, 2 e conjunto respectivamente considerando o modelo M1.....58

Figura 2. Dispersão gráfica das médias fenotípicas do rendimento de grãos (Figuras – A1, B1, C1) e acurácias (Figuras – A2, B2, C2) em relação aos valores genotípicos preditos para o ensaio 1, 2 e conjunto respectivamente considerando o modelo M2.....59

Figura 3. Boxplot do comportamento dos híbridos ao longo dos blocos nos ensaios 1 e 2.....60

Figura 4. Boxplot do comportamento das testemunhas nos ensaios 1 e 2.....61

Figura 5. BLUP dos testadores para os modelos M1 ($y = Xu + Za + Zd + Tc + Wb + \epsilon$) e M2 ($y = Xb + Za + Zd + Tc + \epsilon$).....62

Figura 6. Gráfico de radar entre médias fenotípicas e genotípicas de produção (Kg ha⁻¹) dos híbridos selecionados em cada um dos Ensaios (E1 e E2) e da análise conjunta para cada um dos modelos (M1 e M2) considerando uma intensidade de seleção de 5%. ρ – correlação de Spearman.....63

LISTA DE TABELAS

Capítulo 1

Tabela 1. Logaritmo de máxima verossimilhança residual ($\log(L)$), critério de informação de Akaike (AIC), critério de informação bayesiano (BIC), acurácia média dos híbridos candidatos a seleção (\bar{r}), número de iterações (I), tempo médio por iteração (T/I) em segundos e tempo total por análise (T) em minutos, para os modelos ajustados considerando as diferentes classes de efeitos fixos (EF).....	31
Tabela 2 Porcentagem de coincidência na classificação do valor genético dos diferentes modelos considerando uma intensidade de seleção de 5%.....	36

Capítulo 2

Tabela 1 Análise de deviance para os ensaios E1 (G24) e E2 (G8) e conjunta(E1 + E2) considerando os modelos M1 ($y = Xu + Za + Zd + Tc + Wb + \varepsilon$) e M2 ($y = Xb + Za + Zd + Tc + \varepsilon$).....	57
Tabela 2. Os elementos da diagonal representam a quantidade de híbridos selecionados em cada modelo/ensaio para a intensidade de seleção especificada (5% e 100%). Os demais elementos fora da diagonal representam a porcentagem de coincidência entre a classificação dos modelos M1 e M2 nos ensaios (E1 e E2).....	64
Tabela 3. Médias dos valores genéticos e fenotípicos dos 5% selecionados, dos valores genéticos e fenotípicos das testemunhas e a média geral dentro de cada um dos modelos (M1 e M2) e ensaios (E1-ensaio1, E2-ensaio2 e Conjunto).....	65

SUMÁRIO

CONTEÚDO

1.	INTRODUÇÃO GERAL	12
2.	OBJETIVOS	14
2.1	OBJETIVO GERAL	14
2.2	OBJETIVOS ESPECÍFICOS	14
3.	Referencia.....	15
CAPITULO 1.....		20
Resumo		20
1.	Introdução.....	21
2.	Material e Métodos.....	24
2.1	Dados Biológicos	24
2.2	Descrição dos Modelos	24
I.	Efeito fixo: Média geral + Bloco	25
II.	Efeito fixo: Média geral	25
III.	Efeito fixo: Média geral + Testemunha.....	26
2.3	Modelos Mistos.....	27
2.3.1	Algoritmo <i>Expectation Maximization</i> (EM).....	28
2.4	Análises aplicadas	29
2.4.1	Modelos das análises aplicadas	29
3.	Resultado	31
4.	Discussão	37
5.	Conclusão	39
6.	Referencias	40
CAPITULO 2.....		46
Resumo		46
1.	Introdução.....	47
2.	Material e Métodos.....	50
2.1	Dados Biológicos	50
2.2	Modelos empregados	50
2.3	Modelos Mistos.....	51
2.3.1	Algoritmo <i>Expectation Maximization</i> (EM).....	53

2.4	Análises estatísticas	53
3.	Resultado	56
4.	DISCUSSÃO	66
5.	Conclusão	69
6.	Referencia.....	70

1. INTRODUÇÃO GERAL

O mundo vem passando por um período que diversos trabalhos estão discorrendo sobre uma preocupação, a respeito de uma possível falta de alimento para a população e diversos fatores que vem contribuindo para este fato, como o crescimento da população ate o ano de 2050 e mudanças climativas que vem causando perdas em diversas regiões produtoras do mundo. Está preocupação é de forma geral para toda a cadeia produtiva agricola e principalmente para alguns produtos básicos da alimentação que são essenciais como o arroz, trigo e milho (Godfray et al., 2010, Strzepek and Boehlert, 2010, Foley et al., 2011, Oury et al. 2012, Tilmana et al., 2011, Ray et al., 2013, Bajželj et al., 2014, Laurance et al., 2014, Lobell, 2014, Bassi et al., 2016 and Davis et al., 2016.)

Dentre as culturas base para a alimentação mundial, o milho (*Zea mays*) é a única que vem apresentando incremento na produtividade (Long and Ort, 2010). Para isto continuar acontecendo, as técnicas adotados no melhoramento genético serão cada vez mais cruciais para a superação das adversidades (Tester and Langridge, 2010 and Araus and Cairns, 2014).

A utilização dos milhos híbridos, que resultam em ganhos produtivos decorrente do vigor híbrido, foi adotado e comprovado por inúmeros países, que esta técnica gerou um crescimento em produtividade em Kg ha⁻¹ em uma mesma área (East, 1908, Shull, 1908, Shull, 1911, Duvick et al., 2004, Nerling et al., 2013, Hallauer et al., 2010 and Shapira and David, 2016).

Atualmente acredita-se que para se continuar com os avanços o melhoramento deverá tomar maiores atenções em alguns aspectos para continuar com a mesma efetividade encontrada, como a implementação de projetos mais precisos de experimentais a campo, métodos estatísticos que resultem em resultados mais próximos da verdadeira resposta, a aplicação em maior escala de conceitos de genética quantitativa e populacional, e uma maior integração de várias disciplinas de ciências (Tester and Langridge, 2010, Araus and Cairns, 2014 and Bassi et al., 2016). Esta junção poderão contribuir para contornar ou amenizar os problemas das possíveis perspectivas futuras sobre a alimentação da população.

No melhoramento, mais precisamente na área de métodos estatísticos, a metodologia de modelos mistos vem trazendo resultados promissores para a área florestal e animal, e com estes resultados observados, esta metodologia nos últimos anos vem sendo implementada em diversas culturas de ciclo curto (Grattapaglia and Resende, 2011, Burgueño et al., 2012, Rincent et al., 2012, Isik, 2014 and Bassia et al., 2016), como é o caso do milho.

1 A metodologia de modelos mistos tem sua estimação de componentes de variâncias,
2 pela máxima verossimilhança restrita (REML- *Restricted Maximum Likelihood*) (Patterson
3 and Thompson, 1971) e a estimação dos valores genotípicos preditos, pelo melhor preditor
4 linear não viesado (BLUP- *Best Linear Unbiased Prediction*) (Henderson, 1975 and
5 Rodrigues et al, 2013). A técnica de modelos mistos nos permite uma melhor exploração dos
6 dados e descrição que quais os efeitos fixos e aleatórios de interesse no modelo.

7 Os melhoristas ou pesquisadores em questão, devem sempre pensar quanto a seleção
8 de um modelo, que o melhor modelo é o que melhor aproxime os valores estimados pelo
9 modelo dos reais observados a campo (Szydlowski et al., 2015), para isto existem algumas
10 critérios de precisão experimental para a comparação de modelos existentes, quando o
11 interesse é a seleção do que melhor se adequou a seus dados, como é o caso do AIC, BIC
12 (Szydlowski et al., 2015), Log(L) e acurácia (Resende, 2008).

13 A seleção de um modelo não é o único fator em questão a ser tomado como cautela, o
14 conhecimento da resposta do testador no banco de dados é um desafio para os melhoristas,
15 pois este deve fornecer a melhor discriminação entre os genótipos de acordo com os
16 propósitos da seleção (Hallauer, 1975, Li et al., 2007, Edwards, 2009, Sreckov et al., 2011,
17 Viana et al., 2011; Barreto et al., 2012, Grieder et al., 2012, Guimarães et al., 2012,
18 Rodovalho et al., 2012, and Clovis, et al., 2015).

19 O aumento da eficiência de seleção de híbridos de milho provenientes dos testadores, é
20 obtida com a utilização do método de topcross em conjunto com a utilização de tecnologias
21 que permitam a melhor caracterização genética dos traços fenotípicos e o aumento da
22 eficiência de seleção (Cobb et al., 2013, Araus and Cairns, 2014, Casale et al., 2015 and Loh et
23 al., 2015). Além disto, para uma melhor seleção o conhecimento a resposta do banco de dados
24 e o modo de como será feito o processo de seleção, deve sempre ser adotado como um critério
25 durante o processo de análise dos dados.

26 Ensaios diferentes, ou seja, coleta de banco de dados separados ou gerados por
27 testadores diferentes, deve ser estudados a parte, para depois saber se pode ou não fazer a
28 junção para a seleção dos melhores híbridos e conhecer se as diferentes resposta observados
29 em cada ensaio é decorrente de fatores ambientais ou da heterose.

2. OBJETIVOS

2.1 OBJETIVO GERAL

Avaliar modelos com diferentes arranjos de efeitos fixos e aleatórios que se expliquem melhor os híbridos de milho de primeiro ciclo de seleção e com base nos melhores modelos selecionar os híbridos mais promissores em ensaios individuais e conjuntos para a características produtividade em Kg ha^{-1} .

2.2 OBJETIVOS ESPECÍFICOS

- I. Testar modelos com e sem matriz de parentesco que resultem em uma melhor precisão dos parâmetros AIC, BIC, $\log(L)$ e Acurácia;
- II. Testar modelos com diferentes efeitos fixos e aleatórios que resultem em uma melhor precisão dos AIC, BIC, $\log(L)$ e Acurácia e com otimização de tempo computacional;
- III. Selecionar os 5% melhores híbridos de primeiro ciclo de seleção no ensaio 1 genotipicamente e fenotipicamente;
- IV. Selecionar os 5% melhores híbridos de primeiro ciclo de seleção no ensaio 2 genotipicamente e fenotipicamente;
- V. Selecionar os 5% melhores híbridos de primeiro ciclo de seleção no ensaio conjunto genotipicamente e fenotipicamente;
- VI. Verificar qual ensaio contribuiu para um maior número de híbridos selecionados no ensaio conjunto genotipicamente;
- VII. Testar se os 5% selecionados no ensaio 1, 2 e conjunto foram superiores as médias fenotípicas e genotípicas que os testadores.

3. REFERENCIA

Araus, J.L., Cairns J.E. Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*, 19(1) (2014), pp. 52-61. doi:10.1016/j.tplants.2013.09.008

Bajželj B., Richards K.S., Allwood J.M., Smith P., Dennis J.S., Curmi E., Gilligan C.A. Importance of food-demand management for climate mitigation. *Nature Climate Change*, 4(10) (2014), pp. 924-929. doi:10.1038/nclimate2353

Barreto R. R., Scapim, C. A., Amaral Júnior, A. T. do, de Araújo Rodovalho, M., Vieira, R. A., & Schuelter, A. R. Avaliação da capacidade de combinação de famílias S2 de milho-pipoca por meio de diferentes testadores. *Semina*, 33(3) (2012), 873-890. doi:http://dx.doi.org/10.5433/1679-0359.2012v33n3p873

Bassi F.M., Bentley A.R., Charvet G., Ortiz R., Crossa J. Breeding schemes for the implementation of genomic selection in wheat (*Triticum* spp.). *Plant Science*, 242 (2016), PP. 23-36. doi:10.1016/j.plantsci.2015.08.021

Burgueño J., Campos G. de los, Weigel K., Crossa J. Genomic prediction of breeding values when modeling genotype \times environment interaction using pedigree and dense molecular markers. *Crop Science*, 52(2) (2012), pp. 707-719. doi:10.2135/cropsci2011.06.0299

Burgueño J., Campos G. de los, Weigel K., Crossa J. Genomic prediction of breeding values when modeling genotype \times environment interaction using pedigree and dense molecular markers. *Crop Science*, 52(2) (2012), pp. 707-719. 2012.doi:10.2135/cropsci2011.06.0299

Casale F.P., Rakitsch B., Lippert C., Stegle O. Efficient set tests for the genetic analysis of correlated traits. *Nature methods*, 12(8) (2015), pp. 755-758. doi:10.1038/nmeth.3439

Clovis L.R., Scapim C.A., Pinto R.J.B., Bolson E., Senhorinho, H.J.C. Avaliação De Linhagens S3 De Milho Por Meio De Testadores Adaptados À Safrinha. *Revista Caatinga*, 28(1) (2015), pp. 109-120.

Cobb J.N., DeClerck G., Greenberg A., Clark R., McCouch, S. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*, 126(4) (2013), pp. 867-887. DOI: 10.1007/s00122-013-2066-0

1 Davis K.F., Gephart J.A., Emery K.A., Leach A.M., Galloway J.N., D’Odorico, P.
2 Meeting future food demand with current agricultural resources. *Global Environmental*
3 *Change*, 39 (2016), pp. 125-132. doi:10.1016/j.gloenvcha.2016.05.004

4 Duvick D.N., Smith J.S.C., Cooper M. Long term selection in a commercial hybrid
5 maize breeding program. *Janick. I. Plant Breeding Reviews. Part, 24(2) (2010)*, pp. 109-152.
6 doi:10.1002/9780470650288.ch4

7 EAST, E. M. Inbreeding in corn. *Rep. Conn. Agric. Exp. Stn. (1908)*, pp. 419–428.

8 Edwards, J. Testcross Response to Four Cycles of Half-sib and S Recurrent Selection
9 in the BS13 Maize (L.) Population. *Crop science*, 50(5) (2010), pp. 1840-1847.
10 doi:10.2135/cropsci2009.09.0557

11 Foley J.A., Ramankutty N., Brauman K.A., Cassidy E.S., Gerber J.S., Johnston M.,
12 Mueller N.D., O’connell C., Ray D.K., West P.C., Balzer C., Bennett E.M., Carpenter S.R.,
13 Hill J., Monfreda C., Polasky S., Rockstrom J., Sheehan J., Siebert S., Tilman D., Zaks
14 D.P.M. Solutions for a cultivated planet. *Solutions for a cultivated planet. Nature*, 478(7369)
15 (2011), pp. 337-342. doi:10.1038/nature10452

16 Godfray H.C.J., Beddington J.R., Crute I.R., Haddad L., Lawrence D., Muir J. F.,
17 PRETTY J., ROBINSON S., THOMAS S.M., Toulmin, C. Food security: the challenge of
18 feeding 9 billion people. *Science*, 327(5967) (2010), pp. 812-818.
19 doi:10.1126/science.1185383

20 Grattapaglia D., Resende, M.D. Genomic selection in forest tree breeding. *Tree*
21 *Genetics & Genomes*, 7(2) (2011), pp. 241-255. doi:10.1007/s11295-010-0328-4

22 Grieder C., Dhillon B.S., Schipprack W., Melchinger A.E. Breeding maize as biogas
23 substrate in Central Europe: I. Quantitative-genetic parameters for testcross performance.
24 *Theoretical and Applied Genetics*, 124(6) (2012), pp. 971-980. doi:10.1007/s00122-011-
25 1761-y

26 Guimarães L.J.M., Miranda G.V., DeLima R.O., Maia C., Oliveira L.R.D., Souza
27 L.V.D. Performance of testers with different genetic structure for evaluation of maize inbred
28 lines. *Ciência Rural*, 42(5) (2012), pp. 770-776. [http://dx.doi.org/10.1590/S0103-](http://dx.doi.org/10.1590/S0103-84782012000500002)
29 [84782012000500002](http://dx.doi.org/10.1590/S0103-84782012000500002)

30 Hallauer A.R. Relation of gene action and type of tester in maize breeding procedures.
31 *In Proc Annu Corn Sor Res Conf. Chicago: ASTA (1975)*. pp. 150-165.

32 Hallauer, A.R.; Carena, M.J.; Miranda Filho, J. D. *Quantitative genetics in maize*
33 *breeding. Iowa: Springer (2010)*.

1 Henderson C.R. Best linear unbiased estimation and prediction under a selection
2 model. *Biometrics* (1975), pp. 423-447.

3 Isik F. Genomic selection in forest tree breeding: the concept and an outlook to the
4 future. *New forests*, 45(3) (2014), pp. 379-401. doi:10.1007/s11056-014-9422-z

5 Laurance W.F., Sayer J., Cassman K.G. Agricultural expansion and its impacts on
6 tropical nature. *Trends in ecology & evolution*, 29(2) (2014), pp. 107-116.
7 doi:10.1016/j.tree.2013.12.001

8 Li M.S., Li X.H., Deng L.W., Zhang D.G., Bai L., Zhang S.H. Comparisons of four
9 testers in evaluating 27 CIMMYT and Chinese maize populations. *Maydica*, 52(2) (2007), pp.
10 173-179.

11 Lobell D.B. Climate change adaptation in crop production: Beware of illusions. *Global*
12 *Food Security*, 3(2) (2014), pp. 72-76. doi:10.1016/j.gfs.2014.05.002

13 Loh, P. R., Bhatia, G., Gusev, A., Finucane, H. K., Bulik-Sullivan, B. K., Pollack, S.
14 J., Consortium S.W.G.P.G., Candia T.R., Lee S.H., Wray N.R., Kendler K.S., O'Donovan
15 M.C, Neale B.M., Patterson N., Price A.L. Contrasting genetic architectures of schizophrenia
16 and other complex diseases using fast variance-components analysis. *Nature genetics*. 47(12)
17 (2015), pp. 1385-1392. doi:10.1038/ng.3431

18 Long S.P., Ort D.R. More than taking the heat: crops and global change. *Current*
19 *opinion in plant biology*, 13(3) (2010), pp. 240-247. doi:10.1016/J.PBI.2010.04.008

20 Nerling D., Coelho C.M.M., Nodari R.O. Genetic diversity for physiological quality of
21 seeds from corn (*Zea mays* L.) intervarietal crossbreeds. *Journal of Seed Science*, 35(4)
22 (2013), pp. 449-456. doi.org/10.1590/S2317-15372013000400006

23 Oury F.X., Godin C., Mailliard A., Chassin A., Gardet O., Giraud A., Heumez E.,
24 Morlais J.Y., B., Rousset M., Trottet M., Charmet G. A study of genetic progress due to
25 selection reveals a negative effect of climate change on bread wheat yield in France. *European*
26 *Journal of Agronomy*, 40 (2012), pp. 28-38. doi:10.1016/j.eja.2012.02.007

27 Patterson H.D., Thompson R. Recovery of inter-block information when block sizes
28 are unequal. *Biometrika*, 58(3) (1971), pp. 545-554. doi: 10.1093/biomet/58.3.545

29 Ray D.K., Mueller N.D., West P.C., Foley, J.A. Yield trends are insufficient to double
30 global crop production by 2050. *PloS one*, 8(6) (2013), pp. 1-8.
31 <http://dx.doi.org/10.1371/journal.pone.0066428>

1 Resende R.M.S, Resende M.D.V, Jank L., (2008). Experimentação e análise estatística
2 no melhoramento de forrageiras. In: Resende R.M.S., Valle C.B do, Jank L. (Eds.)
3 Melhoramento de Forrageiras Tropicais. Campo Grande, Embrapa Gado de Corte. p.195-284.

4 Rincent R., Laloë D., Nicolas S., Altmann T., Brunel D., Revilla P., Rodríguez V.M.,
5 Moreno-Gonzalez J., Melchinger A., Bauer E., Schoen C.C., Meyer N., Giauffret C., Bauland
6 C., Jamin P., Laborde J., Monod H., Flament P., Charcosset A., Moreau L. Maximizing the
7 reliability of genomic selection by optimizing the calibration set of reference individuals:
8 comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). *Genetics*,
9 192(2) (2012), pp. 715-728. doi: 10.1534/genetics.112.141473

10 Rodvalho M. de A., Scapim C.A., Pinto R.J.B., Barreto R.R., Ferreira F.R.A., Clóvis
11 L.R. Comparação de testadores em famílias S2 obtidas do híbrido simples de milho-pipoca
12 IAC-112= Comparison of testers in S2 families obtained from the popcorn single hybrid
13 ‘IAC-112’. *Bioscience Journal*, 28(2) (2012), pp. 145-154.

14 Rodrigues W.P., Vieira H.D., Barbosa D.H., Souza Filho G.R., Candido L.S.
15 Adaptability and genotypic stability of *Coffea arabica* genotypes based on REML/BLUP
16 analysis in Rio de Janeiro State, Brazil. *Genetics and Molecular Research*, 12(3) (2013), pp.
17 2391-2399.

18 Shapira R., David L. Genes with a Combination of Over-Dominant and Epistatic
19 Effects Underlie Heterosis in Growth of *Saccharomyces cerevisiae* at High Temperature.
20 *Frontiers in genetics*, 7 (2016), pp. 1-13. doi.org/10.3389/fgene.2016.00072

21 Shull, G. H. The composition of a field of maize. *Journal of Heredity*, (1) (1908), pp.
22 296-301. doi: 10.1093/jhered/os-4.1.296

23 Shull, G. H. The genotypes of maize. *The American Naturalist*, 45(532) (1911), pp.
24 234-252. doi: 10.1086/279207

25 Sreckov Z.O.R.A.N.A., Nastasic A., Bocanski J., Djalovic I., Vukosavljev M.,
26 Jockovic B. Correlation and path analysis of grain yield and morphological traits in test-cross
27 populations of maize. *Pak. J. Bot*, 43(3) (2011), pp. 1729-1731.

28 Strzepek K., Boehlert B. Competition for water for the food system. *Philosophical*
29 *Transactions of the Royal Society of London B: Biological Sciences*, 365(1554) (2010), pp.
30 2927-2940. doi: 10.1098/rstb.2010.0152.

31 Szydłowski M., Krawiec A., Kurek A., Kamionka, M. AIC, BIC, Bayesian evidence
32 against the interacting dark energy model. *The European Physical Journal C*, 75(1) (2015), pp.
33 1-11. doi:10.1140/epjc/s10052-014-3236-1

1 Tester M., Langridge P. Breeding technologies to increase crop production in a
2 changing world. *Science*, 327(5967) (2010), pp. 818-822. doi:10.1126/science.1183700

3 Tilman D., Balzer C., Hill J., Befort B.L. Global food demand and the sustainable
4 intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108(50)
5 (2011), pp. 20260-20264. doi:10.1073/pnas.1116437108

6 Viana J.M.S., Almeida R.V., Faria V.R., Resende M.D.V., Silva F.F. Genetic
7 evaluation of inbred plants based on BLUP of breeding value and general combining ability.
8 *Crop and Pasture Science*, 62(6) (2011), pp. 515-522. doi:10.1071/CP11016 1836-
9 0947/11/060515

10

11

CAPITULO 1

Capacidade preditiva e eficiência computacional entre diferentes modelos na seleção de híbridos de milho

RESUMO

A possível falta de alimento no mundo é um tema que vem sendo bem discutido entre os pesquisadores e as possíveis formas de se contornar está previsibilidade. Dentre as previsões a cultura do milho por estar entre as três principais culturas da base da alimentação do mundial, vem ganhando peso nos estudos e nas repercussões. Uma alternativa para se contorna este problema seria a seleção de híbridos mais produtivos, pois com o mesmo tamanho de área seria possível atingir maiores produtividade. A seleção de modelos que resulte em uma melhor seleção de híbridos é uma alternativa possível e viável de ser adotada, na contribuição dos esforços para serem atingidas as metas necessárias. Mediante aos contextos apresentados o objetivo do presente estudo foi avaliar os diferentes arranjos de modelos, com o intuito de se selecionar os que respondam a uma maior e melhor precisão entre os dados ajustados pelo modelo e os reais. As análises estatísticas foram feitas adotando-se os modelos mistos (REML/BLUP). Foram testados 8 diferentes modelos, estes divididos em três classes, a divisão foi baseado nos diferentes tipos de efeito fixos considerados. Os testes de precisão na seleção de modelos foram baseados no log(L), AIC, BIC, acurácia e medidas de tempo. Os modelos com informação de matriz parentesco, responderam com, pelo menos, o dobro da acurácia dos modelos com ausência da matriz de parentesco. Os melhores ajustes encontrados foram para os modelos M3, M6 e M8. Os menores tempos de respostas nos modelos que não consideram informação de parentesco. Com este trabalho podemos concluir que os melhores modelos, são os modelos de declararam entre os efeitos aleatórios a informação de parentesco e de testemunha e como fixo o modelo M3 o bloco mais a média geral e o M6 a média geral, por resultarem nos melhores ajustes de log(L), AIC, BIC e tempo.

Palavra chave: Melhoramento; Modelos Mistos; Seleção; *Zea mays*

1. INTRODUÇÃO

No mundo a principal fonte de alimento para a população mundial é baseada em três grãos, sendo estes o trigo (*Triticum spp.*), o arroz (*Oryza sativa*) e o milho (*Zea mays*) (Long and Ort, 2010). Nas últimas décadas a cultura do milho é a única dentre as três que vem respondendo com crescimento de produtividade de forma progressiva, o trigo e o arroz vêm até acrescentando na produtividade, mas não como ocorreu a alguns anos (Long and Ort, 2010 and Bassi et al., 2016).

Atualmente diversos autores vêm pesquisando a respeito de uma possível falta de alimento ou mudanças que preocupam sobre as principais fontes de alimento da população mundial (Godfray et al., 2010, Strzepek and Boehlert, 2010, Foley et al., 2011, Oury et al. 2012, Tilmana et al., 2011, Ray et al., 2013, Bajželj et al., 2014, Lurance et al., 2014, Lobell, 2014 and Davis et al., 2016.)

Melhoramento combinado com a implementação de projetos experimentais precisos de campo, métodos estatísticos, aplicação dos conceitos de genética quantitativa e de população, e a integração de várias disciplinas de ciências da planta (Tester and Langridge, 2010, Araus and Cairns, 2014 and Bassi et al., 2016), poderão contribuir para contornar ou amenizar os problemas das possíveis perspectivas futuras sobre a alimentação da população.

O milho dentre as três principais culturas é a de maior flexibilidade de utilização, pela possibilidade de ser fonte alimentar de forma direta e indireta para os homens. De forma indireta como fonte de alimento para aves, suínos, bovinos, caprinos e outros, que servem de alimento para a população.

Mediante a flexibilidade de possíveis formas de utilização do milho, verifica-se a necessidade de se buscar técnicas e métodos que apresentem maior ganhos. No início, com a descoberta que o milho respondia ao vigor híbrido, ou seja, à produção de híbridos propiciou resultados positivos e muitos países adotaram a utilização do milho híbrido, em que no início o incremento produtividade de grãos foi de modo expressivo. Porém, com o tempo não tendo aumentado muito (Duvick et al., 2004 and Hallauer et al., 2010).

O que aconteceu é que no início sem a utilização de técnicas e modelos precisos para a seleção dos melhores híbridos, ainda resultava na seleção de materiais com ganhos em produção, porém com o passar dos anos foi diminuindo os ganhos obtidos e necessitando a utilização de tecnologias que permitam a melhor caracterização genética dos traços fenotípicos e o aumento da eficiência de seleção (Cobb et al., 2013, Araus and Cairns, 2014, Casale et al., 2015 and Loh et al., 2015).

1 A metodologia de modelos mistos a vários anos vem resultando em progressos no
2 melhoramento nas áreas florestal e animal, e nos últimos anos vem sendo implementada em
3 culturas de ciclo curto (Grattapaglia and Resende, 2011, Burgueño et al., 2012, Rincent et al.,
4 2012, Isik, 2014 and Bassia et al., 2016), como é o caso do milho. Na metodologia de
5 modelos mistos a estimação dos componentes de variâncias é feita pela máxima
6 verossimilhança restrita (REML- *Restricted Maximum Likelihood*) (Patterson and Thompson,
7 1971) e a estimação dos valores genotípicos preditos pelo melhor preditor linear não viesado
8 (BLUP- *Best Linear Unbiased Prediction*) (Henderson, 1975 and Rodrigues et al, 2013).

9 Decorrente da adoção de modelos mistos na análise dos dados, é possível descrever
10 diferentes modelos com um mesmo banco de dados. As informações sobre os dados e a
11 necessidade das respostas serão de interesse para determinar qual o efeito fixo e aleatório da
12 análise. Sendo assim, com um banco de dados podemos descrever vários modelos diferentes.

13 Em teoria não existem modelos verdadeiros, o que existe é apenas modelos que podem
14 ser aproximados da realidade, dependendo de um certo número de parâmetros. O melhor do
15 conjunto sob consideração deve ser a melhor aproximação da verdade, por métodos existentes
16 de comparação de modelos (Szydlowski et al., 2015). Algumas das precisões experimentais
17 para a comparação de modelos existentes é o AIC, BIC (Szydlowski et al., 2015), Log(L) e
18 acurácia (Resende, 2014).

19 A informação predita é aproximada da verdadeira pelo modelo em questão, sendo
20 medido pela chamada informação de KullbackLeibler (KL), de modo que quanto melhor o
21 valor observado menor este valor. É impossível calcular as informações KL diretamente,
22 porque é depende do valor verdadeiro que é desconhecido. O método proposto por Akaike
23 (1974) encontrou uma aproximação à quantidade KL, que é chamado de critério de
24 informação de Akaike (AIC), que considera o valor máximo da função de probabilidade e o
25 número de parâmetros do modelo. O melhor modelo é o que melhor se aproximar da verdade,
26 ou seja, o que tem o menor valor de AIC (Szydlowski et al., 2015).

27 No contexto Bayesiano, com as dificuldades de mensurações das medidas de
28 precisões, o pesquisador Schwarz (1978) propôs uma aproximação, a qual é chamada de
29 critério de informação Bayesiano (BIC). O melhor modelo de um conjunto sob consideração é
30 o que minimiza a quantidade BIC (Szydlowski et al., 2015).

31 Existe uma semelhança entre a critério de avaliação de precisão de AIC e BIC, embora
32 eles vêm de diferentes abordagens para o problema de seleção de modelos (Szydlowski et al.,

1 2015). Os critérios de informação de AIC ou BIC, tem como vantagem o caráter de ter a
2 forma funcional de um modelo em consideração (Heck et al., 2014).

3 De modo geral, os valores genéticos preditos não são iguais aos valores genéticos
4 verdadeiros dos indivíduos, sendo que a acurácia é uma medida de precisão que avalia a
5 proximidade destes dois. O cálculo da acurácia completa simultaneamente o coeficiente de
6 variação experimental, coeficiente de variação genotípico e o número de repetições, e esta
7 simultaneidade de parâmetros para o cálculo, nos permite aferir sobre a precisão (Resende et
8 al., 2008).

9 O $\log(L)$ (Log Likelihood), é o logaritmo função para atinge o seu máximo valor para
10 os mesmos pontos da função em si. Sendo que encontrar o máximo de uma função, muitas
11 vezes envolve tomar a derivada de uma função e resolvendo para o parâmetro a ser
12 maximizado, e isso é mais fácil quando a função a ser maximizada é um log-verossimilhança
13 (Resende et al., 2014).

14 As diferentes formas de se descrever um modelo são de quais efeitos se interessa
15 conhecer e qual a importância da fragmentação deste efeito, e o que o pesquisador tem de
16 disponibilidade para montar, ou seja, a flexibilidade em descrever um modelo depende das
17 informações existentes e o interesse que o pesquisador tem sobre estes.

18 O objetivo do nosso estudo com híbridos de milho de primeiro ciclo de seleção, foi
19 avaliar os melhores modelos com diferentes efeitos fixos e aleatórios e a influência matriz de
20 parentesco para uma melhor precisão de estimativas, a fim de se obter modelos com maior
21 adequabilidade que expressem híbridos potenciais. Almejando a possibilidade dos melhoristas
22 selecionarem híbridos de alto potencial, que responderão a altas produtividades.

2. MATERIAL E MÉTODOS

2.1 DADOS BIOLÓGICOS

Os dados experimentais foram obtidos a partir de ensaios montados em blocos aumentados, com um grupo de experimentos que agrupou híbridos topcrosses (TCs) derivados de progênies endogâmicas (S1) do grupo heterótico “Dent” em cruzamento com um testador “Flint”, e outro grupo de experimentos que reuniu híbridos TCs derivados de cruzamentos entre progênies (S1) “Flint” em cruzamento com um testador “Dent”. Os ensaios foram conduzidos no ano agrícola de 2013/14, na Embrapa Milho e Sorgo, localizada no município de Sete Lagoas no estado de Minas Gerais.

Na estrutura de blocos aumentados, utilizou-se 64 híbridos como número básico de tratamentos dentro de cada bloco, incluindo cinco testemunhas comuns a todos os blocos, mas houve experimentos com número diferente. Nessa estrutura, foram avaliados 3352 híbridos TCs (de primeiro ciclo) e mais cinco híbridos comerciais (utilizados como testemunhas comuns), totalizando 3357 tratamentos. As parcelas úteis foram compostas de duas linhas de 4,2 m espaçadas em 0,7 m. A característica analisada foi a produtividade de grãos (PG), em kg ha^{-1} , determinada com base no peso de grãos da parcela, e corrigida para a umidade de 13%.

2.2 DESCRIÇÃO DOS MODELOS

Os modelos adotados foram divididos quanto ao tipo do efeito fixo, sendo estes em três grupos de modelos e um total de 8 diferentes modelos (3, 3 e 2 modelos por classe de grupo I, II e III respectivamente). A diferenciação entre os modelos teve como objetivo a compreensão e a maior exploração diferenciadas dos dados, para a maximização da predição de variáveis que remetem confiabilidade durante a seleção dos híbridos.

Como descrito a cima, o que diferencia os grupos são o tipo de efeitos, sendo que o pesquisador considera como efeito é fixo o que de interesse irá estimar e testar as hipóteses sobre combinações lineares dos níveis do mesmo.

O que diferencia os modelos dentro de cada grupo são os efeitos aleatórios, tendo como função o interesse da estimação dos componentes de variâncias e covariâncias associada a esse fator.

1 I. Efeito fixo: Média geral + Bloco

2 O grupo I tem como característica a utilização da média geral somada aos blocos como
3 o efeito fixo da análise e é subdividido em três tipos de modelos (M1[1], M2[2] e M3[3]) com
4 efeitos aleatórios diferenciados.

5
6
$$y = Xb + Zg + \varepsilon [1]$$

7
$$y = Xb + Za + Zd + \varepsilon [2]$$

8
$$y = Xb + Za + Zd + Tc + \varepsilon [3]$$

9 onde: “**y**” é o vetor da variável a ser analisada; “**b**” é o vetor de efeitos da média geral mais
10 os efeitos dos blocos assumidos como fixo; “**a**” é o vetor de efeitos aditivos assumidos como
11 aleatórios; “**d**” é o vetor de efeitos de dominância assumidos como aleatórios; “**g**” é o vetor
12 de efeitos genotípicos (aditivo + dominância) assumidos como aleatórios; “**c**” é o vetor de
13 efeitos das testemunhas assumidos como aleatórios; “**ε**” é o vetor de erros aleatórios; “**X**” é a
14 matriz de incidência para o efeito fixo; “**Z**” e “**T**” são as matrizes de incidências para os
15 feitos aleatórios descritos anteriormente.

16
17 No M1[1], o efeito aleatório é decorrente dos efeitos genotípicos (aditivo +
18 dominância), no M2[2] o efeito aleatório de variância genotípica como estimado no anterior
19 foi subdividido em aditiva e de dominância (matriz de parentesco - *pedigree*) e no M3[3],
20 além do efeito aleatório da matriz de parentesco como no M2 foi acrescido a matriz de
21 informação de testemunha.

22
23 II. Efeito fixo: Média geral

24 O grupo II tem como característica a utilização da média geral como o efeito fixo da
25 análise e é subdividida em três tipos de modelos (M4[4], M5[5] e M6[6]) com efeitos
26 aleatórios diferenciados.

27
28
$$y = Xu + Zg + Wb + \varepsilon [4]$$

29
$$y = Xu + Za + Zd + Wb + \varepsilon [5]$$

30
$$y = Xu + Za + Zd + Tc + Wb + \varepsilon [6]$$

31 onde: “**y**” é o vetor da variável a ser analisada; “**u**” é o vetor de efeitos da média geral
32 assumido como fixo; “**a**” é o vetor dos efeitos aditivos assumidos como aleatórios; “**d**” é o

1 vetor de efeitos de dominância assumidos como aleatórios; “g” é o vetor de efeitos
2 genotípicos (aditivo + dominância) assumidos como aleatórios; “c” é o vetor de efeitos da
3 testemunhas assumidos como aleatórios; “b” é o vetor de efeitos dos blocos assumidos como
4 aleatórios; “ε” é o vetor de erros aleatórios; “X” é a matriz de incidência para o efeito fixo;
5 “Z”, “T” e “W” são as matrizes de incidências para os feitos aleatórios descritos
6 anteriormente.

7
8 Nos modelos estudados da classe II, o M4 tem como efeito aleatório o genotípico e o
9 de bloco, no M5 os efeitos aleatórios presentes são decorrentes da informação da matriz de
10 parentesco e do bloco e no M6, além dos efeitos aleatórios iguais ao M5 tem mais um efeito
11 que é decorrente da matriz de testemunha.

12 13 III. Efeito fixo: Média geral + Testemunha

14 O grupo III tem como característica a utilização da média geral mais a testemunha
15 como o efeito fixo da análise e é subdividido em dois tipos de modelos (M7[7] e M8[8]) com
16 efeitos aleatórios diferenciados.

$$17 \quad y = Xc + Zg + Wb + \varepsilon \quad [7]$$

$$18 \quad y = Xc + Za + Zd + Wb + \varepsilon \quad [8]$$

19 onde: “y” é o vetor da variável a ser analisada; “c” é o vetor de efeitos da média geral mais o
20 efeito das testemunhas assumido como fixo; “a” é o vetor dos efeitos aditivos assumidos
21 como aleatórios; “d” é o vetor de efeitos de dominância assumidos como aleatórios; “g” é o
22 vetor de efeitos genotípicos (aditivo + dominância) assumidos como aleatórios; “b” é o vetor
23 de efeitos dos blocos assumidos como aleatórios; “ε” é o vetor de erros aleatórios; “X” é a
24 matriz de incidência para o efeito fixo; “Z” e “W” são as matrizes de incidências para os
25 feitos aleatórios descritos anteriormente.

26
27
28 Na classe de modelo classe III, o M7 tem como efeito aleatório o genotípico e o de
29 bloco e o M8 os efeitos aleatórios presentes são decorrentes da informação da matriz de
30 parentesco e do bloco.

31

2.3 MODELOS MISTOS

Os dados obtidos foram analisados utilizando modelos mistos. A obtenção das estimativas dos parâmetros dos modelos mistos foi via REML pelo processo iterativo, onde a função de verossimilhança residual foi maximizada, e o algoritmo numérico utilizado foi o EM (*Expectation Maximization*) (Dempster et al., 1977) obtido a partir da primeira deriva da função de verossimilhança residual (MRODE, 2014).

Os modelos lineares mistos utilizados neste estudo são uma generalização do originalmente proposto por Henderson et al. (1959) e pode ser escrito na forma matricial usando a notação [9]

$$y = X\beta + Zu + \varepsilon \text{ com } N_u = \sum_{i=1}^I N_{u_i} \text{ [9]}$$

onde y : vetor de $N \times 1$ observações; “ β ” vetor de $N_\beta \times 1$ efeitos fixos e covariáveis fixas; “ X ” matriz de incidência para efeitos fixos e covariáveis fixas de dimensão $N \times N_\beta$; “ u ” vetor que concatena i vetores de efeitos aleatórios $u = (u_1, u_2, u_3 \dots u_i)$ de dimensão $N_u \times 1$; “ Z ” matriz de incidência que agrupa i matrizes de efeitos aleatórios $Z = (Z_1, Z_2, Z_3 \dots Z_i)$ de dimensão $N \times N_u$ e “ ε ” vetor de $N \times 1$ erros aleatórios.

A matriz de estruturas de (co)variância do modelo [9] pode ser expressa por

$$\begin{bmatrix} u \\ \varepsilon \end{bmatrix} \sim N \left(0, \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \right), \quad G = \bigoplus_{i=1}^I G_i \text{ [10]}$$

$$\text{var}(u_i) = H_i \sigma_{u_i}^2 = G_i \text{ ou } I \sigma_{u_i}^2 = G_i \text{ para } \forall H_i = I,$$

$$\text{var}(\varepsilon) = I \sigma_\varepsilon^2 = R, \quad \text{cov}(u_i, \varepsilon) = 0, \quad \text{var}(y) = V = ZGZ' + R,$$

onde “ G ” matriz de (co)variância dos efeitos aleatórios; “ R ” matriz de (co)variância residual; “ V ” matriz de (co)variância de y , “ H_i ” matriz de correlação entre os efeitos aleatórios de u_i , “ I ” matriz identidade, “ $\sigma_{u_i}^2$ ” variância de u_i e “ σ_ε^2 ” variância do erro.

As soluções de máxima verossimilhança de β e $u = (u_1, u_2, u_3 \dots u_i)$ do modelo [9], podem ser obtidas através da primeira derivada da função de densidade de probabilidade conjunta $f(y, u)$, proposta por Henderson et al. (1959), em relação aos parâmetros β e u e então obtemos o sistema de equações de modelos mistos (SEMM), assumindo que a estrutura de R^{-1} é uma matriz identidade (MRODE, 2014) podemos reescrever o SEMM como [11]

$$\begin{bmatrix} X'X & X'Z_1 & X'Z_2 & \cdots & X'Z_j \\ Z_1'X & Z_1'Z_1 + H_1^{-1}\lambda_1 & Z_1'Z_2 & \cdots & Z_1'Z_j \\ Z_2'X & Z_2'Z_1 & Z_2'Z_2 + H_2^{-1}\lambda_2 & \cdots & Z_2'Z_j \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_i'X & Z_i'Z_1 & Z_i'Z_2 & \cdots & Z_i'Z_j + H_i^{-1}\lambda_i \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \tilde{u}_1 \\ \tilde{u}_2 \\ \vdots \\ \tilde{u}_i \end{bmatrix} = \begin{bmatrix} X'y \\ Z_1'y \\ Z_2'y \\ \vdots \\ Z_i'y \end{bmatrix} \quad [11]$$

$$\lambda_1, \lambda_2 \cdots \lambda_i = \frac{\sigma_\varepsilon^2}{\sigma_{u_1}^2}, \frac{\sigma_\varepsilon^2}{\sigma_{u_2}^2} \cdots \frac{\sigma_\varepsilon^2}{\sigma_{u_i}^2} \quad [12]$$

A solução dos SEMM apresentados em [11] assume que $(\sigma_{u_1}^2, \sigma_{u_2}^2 \cdots \sigma_{u_i}^2)$ e σ_ε^2 são conhecidos. Porém, na prática tais parâmetros são desconhecidos e podem ser substituídos por suas estimativas, $(\hat{\sigma}_{u_1}^2, \hat{\sigma}_{u_2}^2 \cdots \hat{\sigma}_{u_i}^2)$ e $\hat{\sigma}_\varepsilon^2$, obtidas através de REML.

2.3.1 Algoritmo *Expectation Maximization* (EM)

A utilização do algoritmo EM, considerou o SEMM apresentado em [11] e uma generalização das equações apresentadas por Mrode (2014), da função de verossimilhança residual, obtidas através do rearranjo das expressões de primeira derivada

$$[N - r(X)][\hat{\sigma}_\varepsilon^2] = \left[(y - X\hat{\beta} - \sum_{i=1}^I Z_i \tilde{u}_i)' y \right] \quad [13]$$

$$\begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_i \end{bmatrix} \begin{bmatrix} \hat{\sigma}_{u_1}^2 \\ \hat{\sigma}_{u_2}^2 \\ \vdots \\ \hat{\sigma}_{u_i}^2 \end{bmatrix} = \begin{bmatrix} \tilde{u}_1' H_1^{-1} \tilde{u}_1 + \text{tr}(H_1^{-1} C^{22}) \hat{\sigma}_\varepsilon^2 \\ \tilde{u}_2' H_2^{-1} \tilde{u}_2 + \text{tr}(H_2^{-1} C^{33}) \hat{\sigma}_\varepsilon^2 \\ \vdots \\ \tilde{u}_i' H_i^{-1} \tilde{u}_i + \text{tr}(H_i^{-1} C^{kk}) \hat{\sigma}_\varepsilon^2 \end{bmatrix} \quad [14]$$

onde “ $r(X)$ ” é o *rank* da matriz X e $(C^{22}, C^{33} \dots C^{kk})$ são obtidas da inversa da matriz de coeficientes do SEMM [11]

$$C^{-1} = \begin{bmatrix} C^{11} & C^{12} & C^{13} & \cdots & C^{1j} \\ C^{21} & C^{22} & C^{23} & \cdots & C^{2j} \\ C^{31} & C^{32} & C^{33} & \cdots & C^{3j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C^{i1} & C^{i2} & C^{i3} & \cdots & C^{ij} \end{bmatrix} \quad [15]$$

O cálculo dos traços matriciais foi feito por meio da soma do produto de Hadamard [16]

$$\text{tr}(H_i^{-1} C^{kk}) = \sum H_i^{-1} \circ C^{kk} = \sum_{i=1, j=1}^{I, J} h_{ij} c_{ij} \quad [16]$$

2.4 ANÁLISES APLICADAS

Mediante os modelos adotados foram observados alguns valores como o Logaritmo de máxima verossimilhança residual ($\log(L)$) [17], o critério de informação de Akaike (AIC) [18], o critério de informação bayesiano (BIC) [19], a acurácia média dos híbridos candidatos a seleção (\hat{r}_{gg}) [20], o número de iterações (I) e tempo médio por iteração (T/I), em segundos.

Com os valores preditivos de cada modelo, plotamos gráficos de dispersão entre o valor genotípico e o valor as médias fenotípicas do rendimento de grãos em Kg ha^{-1} e da acurácia. Também com os dados genotípicos preditos selecionamos 5% dos melhores híbridos em cada um dos modelos e com os selecionados de cada modelo verificou as coincidências entre modelos.

Todas as análises estatísticas foram feitas no programa R (TEAM, 2016)

2.4.1 Modelos das análises aplicadas

A função de verossimilhança residual ($\log(L)$) assume que $y \sim N(X\beta, V)$ pode ser escrita como (Meyer, 1989; Johnson and Thompson, 1995)

$$\begin{aligned}\log(L) &= -\frac{1}{2}(\text{const.} + \ln|V| + \ln|X'V^{-1}X| + y'Py) \\ &= -\frac{1}{2}(\text{const.} + \ln|R| + \ln|G| + \ln|C^*| + y'Py)\end{aligned}\quad [17]$$

onde “**log(L)**” logaritmo da função de verossimilhança residual (Patterson, Thompson, 1971), const.: constante da função de verossimilhança residual; “**P**” projetor ortogonal de y no espaço coluna de X (Gilmour et al., 1995), “**C***” matriz de coeficientes do SEMM [11] pelo inverso da variância residual.

O critério de Akaike (AIC) (Akaike, 1974) [18] pode ser escrito como

$$\text{AIC} = -2\log(L) + 2p \quad [18]$$

onde “**AIC**” critério de informação de Akaike (1974); “**log(L)**” logaritmo da função de verossimilhança residual; “**p**” número de parâmetros do modelos avaliado (fixo + aleatório)

Critério de informação bayesiano (BIC) (Schwarz, 1978) [19] pode ser escrito como

$$\text{BIC} = -2\log(L) + p \cdot \log(v) \quad [19]$$

onde “**BIC**” critério de informação bayesiano de Schwarz (1978); “**log(L)**” logaritmo da função de verossimilhança residual; “**p**” número de parâmetros do modelos avaliado (fixo + aleatório); “**v**” é o número de graus de liberdade do resíduo.

Acurácia média dos híbridos candidatos a seleção (\hat{r}_{gg}) [20] pode ser escrito como

1
$$\hat{r}_{\tilde{g}g} = [1 - (\text{PEV} / \sigma_g^2)]^{1/2} \quad [20]$$

2
$$\text{PEV} = \text{Var}(g - \tilde{g}) = C^{22} \sigma_\varepsilon^2 = (1 - r_{\tilde{g}g}^2) \sigma_g^2 \quad [21]$$

3 onde “**PEV**” é um estimador do erro de predição; “ σ_g^2 ” é a variância genética; “ σ_ε^2 ”
4 variância do erro.

3. RESULTADO

Os critérios de precisão de log(L), AIC e BIC entre os diferentes modelos resultaram em variações de -28297,48 à -29003,59; 56715,87 à 58014,57 e 57133,03 à 58039,37 respectivamente (Tabela 1). Para a classe de efeito fixo I, os menores valores observados de log(L), AIC e BIC foi para o M3, com os respectivos valores -28297,48; 56715,87 e 57099,29. Dentro da classe II, os menores valores observados foi do M6, com os respectivos valores -28721,74; 57455,49 e 57492,69. E, dentro da classe III no M7 foram observados os menores valores estimados, sendo -28873,66; 57765,32 e 57821,10 para as respectivas variáveis.

Tabela 1. Logaritmo de máxima verossimilhança residual (log(L)), critério de informação de Akaike (AIC), critério de informação bayesiano (BIC), acurácia média dos híbridos candidatos a seleção (\hat{r}_{gg}), número de iterações (I), tempo médio por iteração (T/I) em segundos e tempo total por análise (T) em minutos, para os modelos ajustados considerando as diferentes classes de efeitos fixos (EF).

EF	Modelos	log(L)	AIC	BIC	\hat{r}_{gg}	I	T/I	T
I	M1	-28543,04	57205,80	57576,85	0,37	116	1,48	2,86
	M2	-28316,89	56755,79	57133,03	0,64	1132	236,78	4467,25
	M3	-28297,48	56715,87	57099,29	0,59	433	232,54	1678,16
II	M4	-29003,59	58014,57	58039,37	0,37	92	2,63	4,03
	M5	-28740,19	57541,03	57572,03	0,65	981	349,21	5709,58
	M6	-28721,74	57455,49	57492,69	0,61	417	358,52	2491,71
III	M7	-28873,66	57765,32	57821,10	0,11	499	3,30	27,45
	M8	-28679,38	57378,77	57440,77	0,61	969	240,94	3891,18

Classe de efeitos fixos (EF): I) Média geral + bloco; II) Média geral; III) Média geral + Testemunha. Modelos: M1: $y = Xb + Zg + \varepsilon$; M2: $y = Xb + Za + Zd + \varepsilon$; M3: $y = Xb + Za + Zd + Tc + \varepsilon$; M4: $y = Xu + Zg + Wb + \varepsilon$; M5: $y = Xu + Za + Zd + Wb + \varepsilon$; M6: $y = Xu + Za + Zd + Tc + Wb + \varepsilon$; M7: $y = Xc + Zg + Wb + \varepsilon$; M8: $y = Xc + Za + Zd + Wb + \varepsilon$.

A acurácia média dos híbridos candidatos a seleção, variou de 0,11 a 0,65 entre os modelos, sendo o menor e o maior valor referente ao M7 e M5 respectivamente. Dentro das classes I, II e III a acurácia média variara de 0,37 a 0,59 (M1 e M2); 0,37 a 0,65 (M4 e M5) e 0,11 a 0,61 (M7 e M8) respectivamente.

A taxa de convergência pelo critério de herdabilidade adotado é de 10^{-5} . Os modelos para está taxa de convergência variaram de 92 a 1132 iterações, nos respectivos modelos M4 e M2. Dentro de cada uma das três das classes, o menor número de interação observado foi no modelo M1 (116), M4 (92) e M7 (499), nas respectivas classes I, II e III.

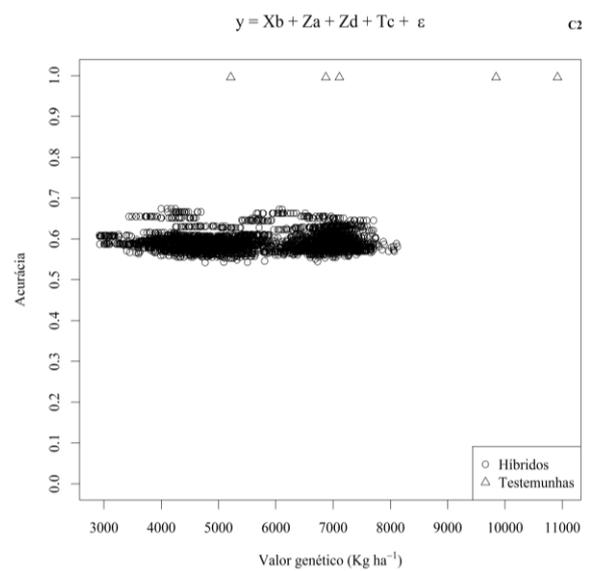
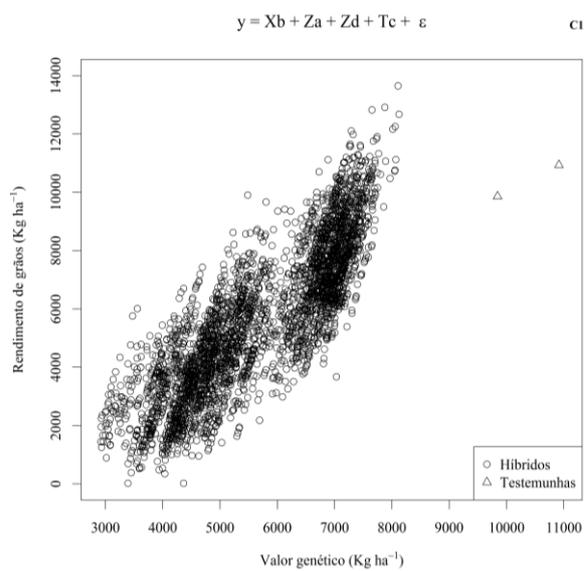
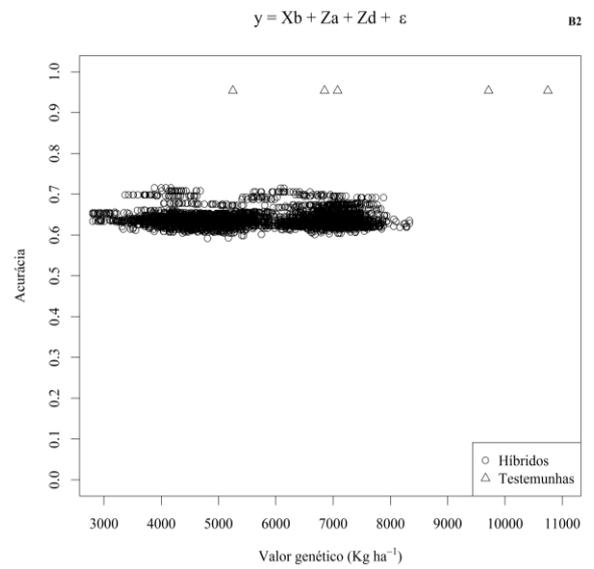
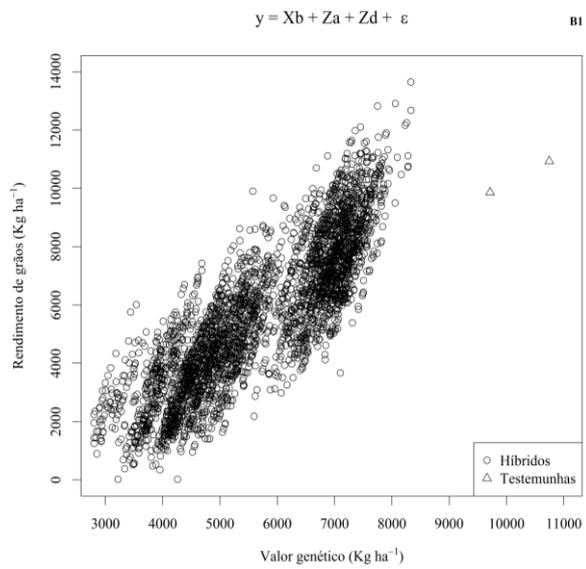
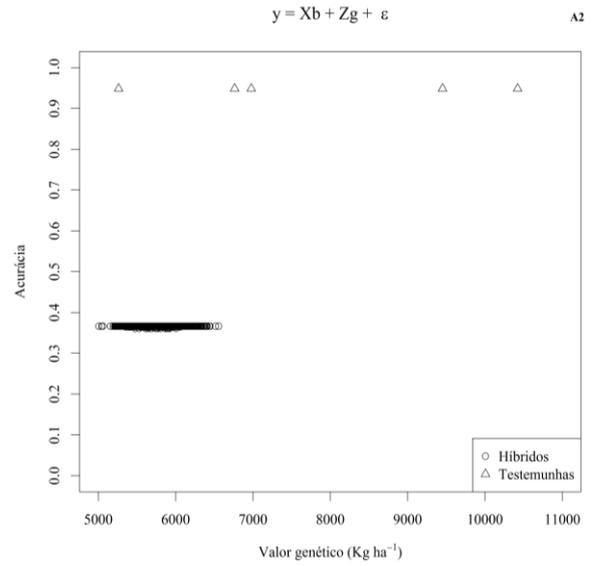
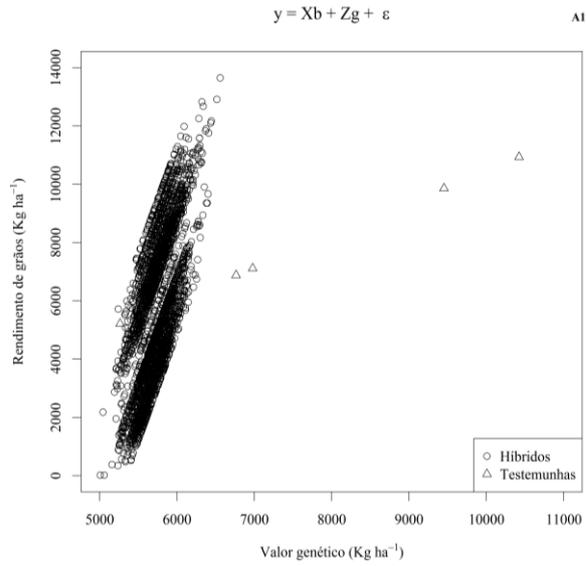
1 A variável tempo médio por interação, referente a quantidade de tempo demandado
2 por ciclo de ajuste, variou entre os modelos de 1,48 a 358,52 segundos, para os modelos M1 e
3 M6, respectivamente. Os menores tempos médios dentro de cada classe foram observados no
4 modelo M1 (1,48 segundos), M4 (2,63 segundos) e M7 (3,3 segundos), para as classes I, II e
5 III respectivamente.

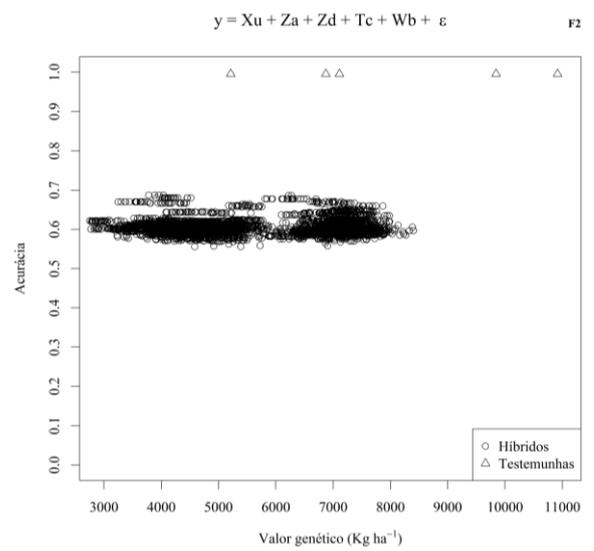
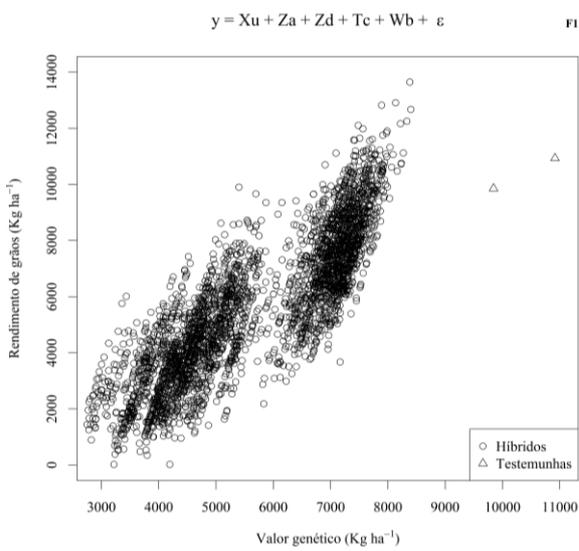
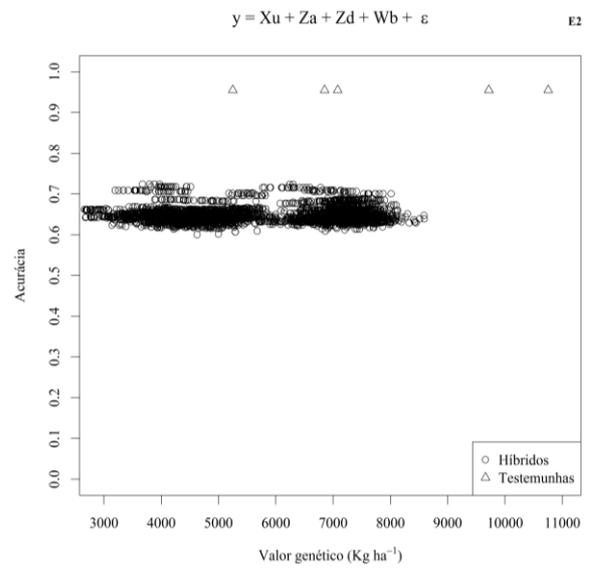
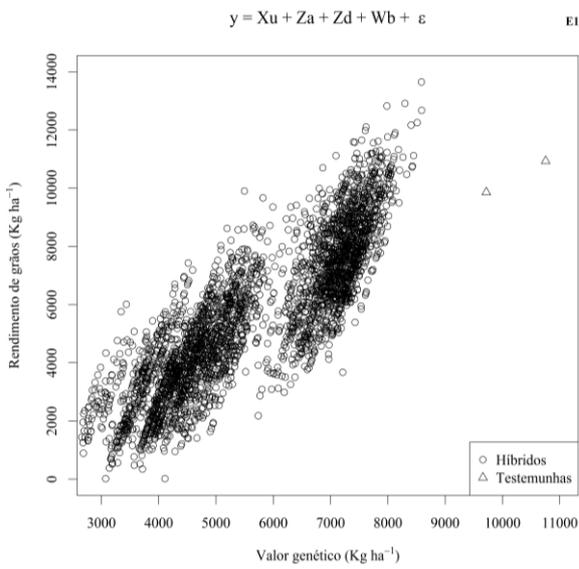
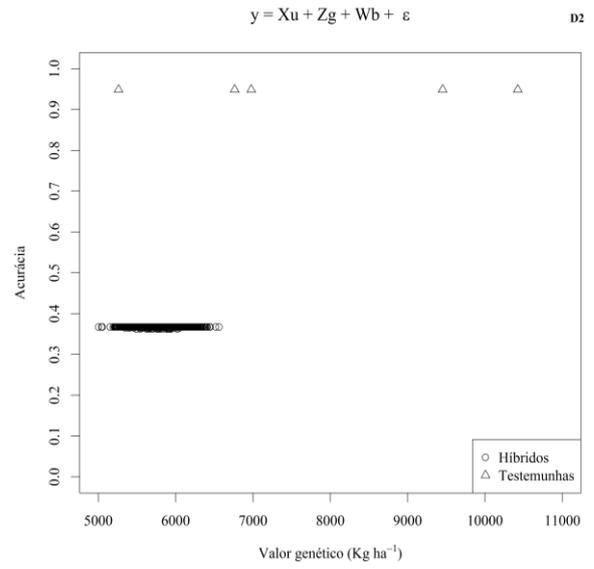
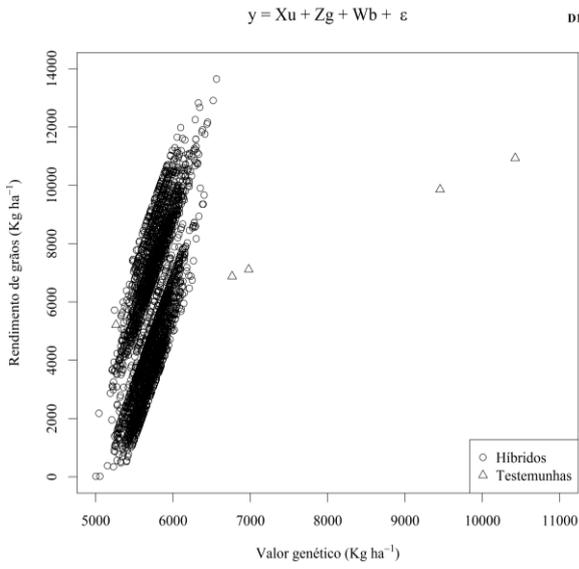
6 Avaliando-se a variável tempo total estimado por modelo em minutos, pode-se dizer
7 que dentre todos os modelos avaliados o menor tempo observado foi no M1 (2,86 minutos),
8 presente na classe I e o maior tempo o M5 (5709,58 minutos) da classe II. Os modelos M4
9 (4,03 minutos) e M7 (27,45 minutos), referem-se aos menores tempos totais para a
10 convergência das classes II e III, respectivamente.

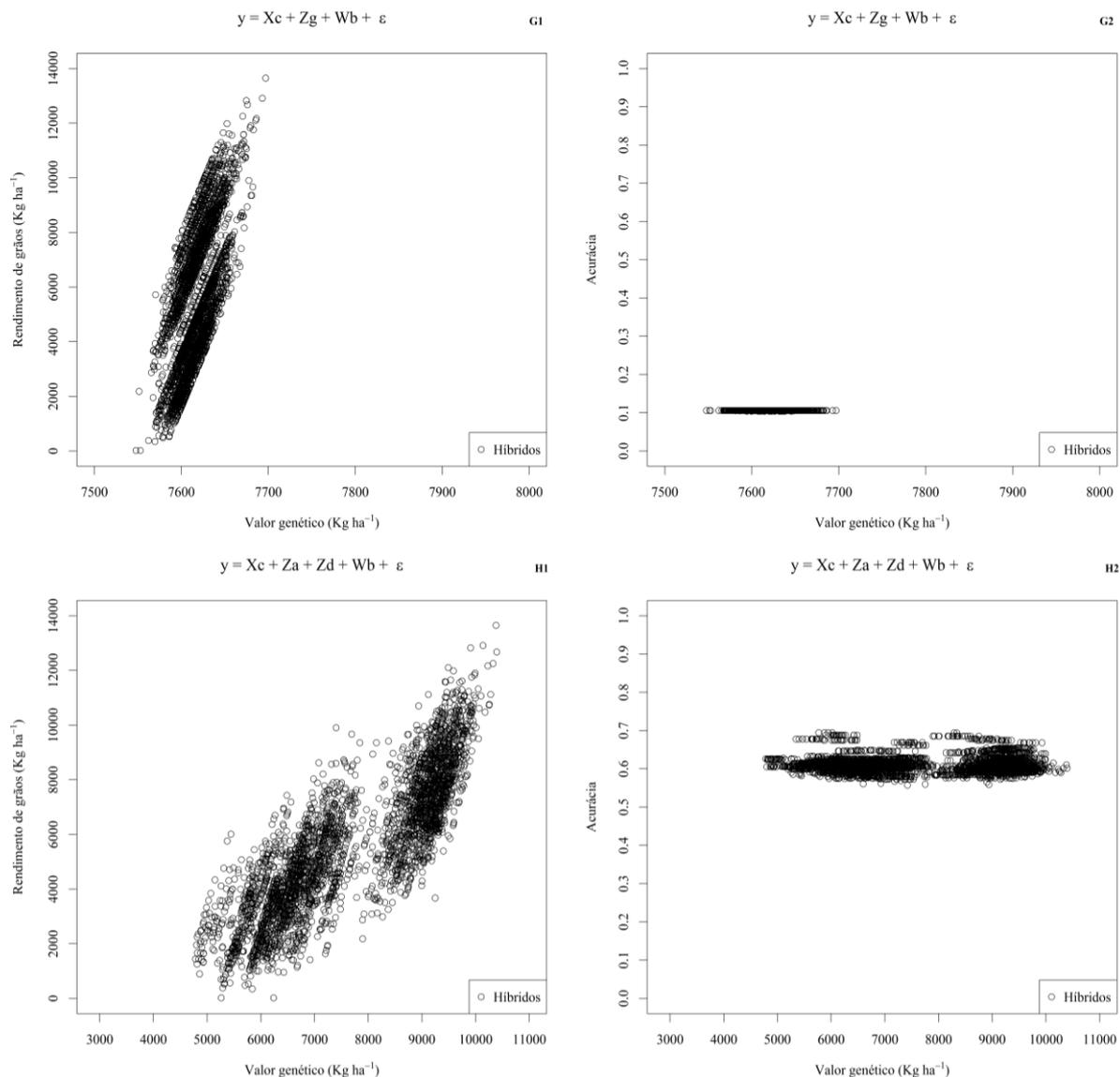
11 Os valores genotípicos estimados foram mais condensados nos modelos que não
12 adotaram a informação de parentesco, M1 (A), M4 (D) e M7 (G) entre aproximadamente
13 5000 a 6500, 5000 a 6500 e 7550 a 7700 respectivamente e estes modelos resultaram nos
14 menores valores de acurácia seletivas dos híbridos (Figura 1). As maiores amplitudes de
15 valores genotípicos observados variaram de 3000 a 8000 (B, C, E e F) e 5000 a 1000 (H) Kg
16 ha⁻¹. Sendo estes resultados sem considerar as testemunhas nos modelos que as apresentavam.

17 Os valores de acurácia para os testadores foram acima de 0,9 para todos os modelos
18 que foram estimadas (A2, B2, C2, D2, E2 e F2). A menor acurácia seletiva foi do M7 (G2) e a
19 maior do M5 (E2). Para a maioria dos modelos a acurácia variou entre 0,55 a 0,75 (62,5%).

20 Fixando-se uma intensidade de seleção em 5% os modelos sem informação de
21 parentesco selecionam praticamente os mesmos indivíduos, modelos M1.I, M4.II e M7.III
22 com a coincidência entre eles de 99,40% e suas coincidências com os modelos com
23 informação de parentesco variaram entre 36,90 e 38,10 (Tabela 2). A coincidência entre os
24 modelos que consideraram matriz de parentesco variou entre 88,10 (M2.I-M6.II e M2.I-
25 M8.III) a 97,62 (M6.II-M8.III).







1 Figura 1. Fixando-se uma intensidade de seleção em 5% os modelos sem informação de
 2 parentesco selecionam praticamente os mesmos indivíduos, modelos M1.I, M4.II e M7.III
 3 com a coincidência entre eles de 99,40% e suas coincidência com os modelos com informação
 4 de parentesco variaram entre 36,90 e 38,10 (Tabela 2). A coincidência entre os modelos que
 5 consideraram matriz de parentesco variou entre 88,10 (M2.I-M6.II e M2.I-M8.III) a 97,62
 6 (M6.II-M8.III).

7
 8
 9
 10
 11
 12
 13

1 Tabela 2 Porcentagem de coincidência na classificação do valor genético dos diferentes
 2 modelos considerando uma intensidade de seleção de 5%.

Modelos	M1.I	M2.I	M3.I	M4.II	M5.II	M6.II	M7.III	M8.III
M1.I	100,00	36,90	36,90	99,40	38,10	37,50	99,40	37,50
M2.I		100,00	96,43	36,90	88,69	88,10	36,90	88,10
M3.I			100,00	36,90	91,67	91,07	36,90	89,88
M4.II				100,00	38,10	37,50	99,40	37,50
M5.II					100,00	95,83	38,10	95,24
M6.II						100,00	37,50	97,62
M7.III							100,00	37,50
M8.III								100,00

3 .I – modelos considerando a média geral + os blocos como efeito fixo; .II – modelos considerando a média geral
 4 como efeito fixo e .III – modelos considerando a média geral + as testemunhas como efeito fixo. Modelos: M1.I:
 5 $y = Xb + Zg + \varepsilon$; M2.I: $y = Xb + Za + Zd + \varepsilon$; M3.I: $y = Xb + Za + Zd + Tc + \varepsilon$; M4.II: $y = Xu + Zg + Wb + \varepsilon$;
 6 M5.II: $y = Xu + Za + Zd + Wb + \varepsilon$; M6.II: $y = Xu + Za + Zd + Tc + Wb + \varepsilon$; M7.III: $y = Xc + Zg + Wb + \varepsilon$;
 7 M8.III: $y = Xc + Za + Zd + Wb + \varepsilon$.

8

1 4. DISCUSSÃO

2 Segundo Resende et al. (2012) quanto menor o valor de $\log(L)$, AIC e BIC, melhores
3 serão as estimativas dos modelos. Mediante este critério de classificação, dentro de cada
4 classe os modelos que melhor se ajustaram foram o M3, M6 e M8 nas respectivas classes I, II
5 e III (Tabela 1). Porém, os modelos que representaram o melhor ajuste aos dados reais, foram
6 o M3 e M6, por responderem aos melhores valores de $\log(L)$, AIC e BIC; e a acurácia sendo
7 próximas das melhores dentro de suas classes.

8 A acurácia quanto maior o valor observado melhor, pois maior a confiança na
9 avaliação e no valor preditivo (Resende et al., 2008). Esta é outra medida de precisão que
10 deve ser observado na escolha do modelo. Nos modelos M3 e M6 a acurácia, pode não ter
11 sido as maiores dentro de suas classes, decorrente do confundimento nas estimativas em
12 relação a outros modelos. Este confundimento é decorrente da não declaração das
13 testemunhas como efeito aleatório separado dos híbridos, que podem ocasionar em uma
14 superestimação da média geral da acurácia, pois os valores da acurácia das testemunhas
15 entram para o calculo da acurácia média.

16 Outro indício desta superestimação é decorrente das maiores acurácias não serem dos
17 modelos dos melhores ajuste de $\log(L)$, AIC e BIC. Neste contexto, existem diversos outros
18 trabalhos na literatura, utilizando os mesmos ou outros critérios na avaliação da precisão para
19 a seleção de modelos, que se ajustem mais próximo dos dados reais (Grün and Leisch, 2008,
20 Feddag and Bacci, 2009 and Feddag, 2013).

21 O número de iterações (I), tempo médio por iteração (T/I) e tempo total por análise
22 (T), vem corroborando com os resultados apresentados anteriormente na seleção dos modelos
23 M3 e M6, pois dentre os modelos que consideraram a matriz de pedigree e resultaram em
24 acurácia seletiva média, estes foram os dois mais eficientes computacionalmente. A eficiência
25 computacional vem sendo por inúmeros trabalhos abordados, pois modelos que resultam em
26 menor tempo e eficientes vem sendo buscados (Meyer, 1991, Harville, 2004 and Misztal,
27 2008).

28 Os modelos M1, M4 e M7, resultarão nos menores tempos totais por análise, porém ao
29 inserir a matriz de parentescos diversos trabalhos discorrem uma perda computacional por se
30 tratar de modelos mais complexos, ou seja, um gasto de memória e tempo de processamento
31 dos computadores (Misztal and Perez-Enciso, 1993, Johnson and Thompson, 1995,
32 Thompson et al., 2003, Meyer, 2008, Diffey et al., 2013 and Li and Pourahmadi, 2013). Por

1 outro lado este gasto é vantajoso por resultar em uma melhor predição de seleção, como foi
2 observado nos valores preditivos.

3 A incorporação das matrizes de parentesco nos modelos M2 e M3 da classe I e M5 e
4 M6 da classe II, praticamente provocou o dobramento da acurácia seletiva individual,
5 melhorando a capacidade preditiva dos modelos (Figura 1). O pedigree causou o dobramento
6 da acurácia, pois ele colabora na estimativa dos valores preditos (Hasstedt, 1982). Diversos
7 trabalhos na área de melhoramento animal, florestal e outras inúmeras culturas vêm
8 adicionando a informação de parentesco nos modelos mistos para melhores predições (Piepho
9 et al., 2008, Piepho, 2009, Li et al., 2012, Boysen et al., 2013 and Cheung et al., 2013).

10 Além disto, a matriz de parentesco elevou os valores genéticos compensando a
11 ausência de repetições dos híbridos diminuindo a discrepância entre o valor genético das
12 testemunhas e dos híbridos.

13 Os modelos com informação de pedigree e sem a informação resultaram em baixas
14 taxas de coincidências entre modelos, e isto é devido à ausência de genótipos selecionados por
15 uma classe de modelos e não pela outra classe (Tabela 3). Esta resposta poderá ser traduzida
16 como uma falta de híbridos promissores para uma das classes de modelos testados.

17 Por estes motivos apresentados anteriormente, deve-se considerar os modelos mais
18 precisos e acurados, sendo este, que resultem em informações com os menores log(L), AIC e
19 BIC, maiores acurácias e com gastos computacionais moderados. Pois estes modelos serão os
20 mais aconselhados para a adoção nos programas de melhoramentos de milho de primeiro ciclo
21 de seleção, como pode ser observados nos trabalhos de Piepho (2009), Borges et al. (2010),
22 Albrecht et al. (2011) and Zhang et al. (2016), buscando os melhores modelos para a seleção
23 das culturas milho, batata-doce, milho e soja respectivamente.

24 Outro fato interessante e que corroborou com os resultados apresentados no trabalho, é
25 que os modelos M3.I e M6.II resultaram em uma coincidência de 91,07%, ou seja, chegaram
26 a uma similaridade entre os selecionados muito grande.

27

1 5. CONCLUSÃO

2 Os melhores valores de precisão dos critérios de AIC, BIC, log(L), acurácia e
3 parâmetros de tempo avaliados foram para os modelos M3 e M6, pertencentes às classes I e
4 II, respectivamente.

5 Os modelos com ausência de matriz de parentesco apresentaram uma acurácia de no
6 mínimo 50% inferior aos modelos com a matriz de parentesco ponderando na estimação e
7 decomposição dos efeitos para nos parâmetros. Além disso, os modelos com ausência de
8 pedigree apresentaram uma baixa coincidência dos selecionados comparando com os outros
9 modelos, menor que 34% e todos os modelos considerando o pedigree apresentaram uma
10 coincidência superior a 88%.

11 Mediante a estes resultados, podemos afirmar que a matriz de pedigree é de suma
12 importância para uma estimação de ajustes dos dados estimados em relação aos verdadeiros,
13 quando presente existente a informação.

14

6. REFERENCIAS

Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6) (1974), pp. 716 - 723. doi: 10.1109/TAC.1974.1100705

Albrecht T., Wimmer V., Auinger H.J., Erbe M., Knaak C., Ouzunova M., Simianer H., Schön C.C. Genome-based prediction of testcross values in maize. Theoretical and Applied Genetics, 123(2) (2011), pp. 339-350. doi:10.1007/s00122-011-1587-7

Araus, J.L., Cairns J.E. Field high-throughput phenotyping: the new crop breeding frontier. Trends in Plant Science, 19(1) (2014), pp. 52-61. doi:10.1016/j.tplants.2013.09.008

Bajželj B., Richards K.S., Allwood J.M., Smith P., Dennis J.S., Curmi E., Gilligan C.A. Importance of food-demand management for climate mitigation. Nature Climate Change, 4(10) (2014), pp. 924-929. doi:10.1038/nclimate2353

Bassi F.M., Bentley A.R., Charmet G., Ortiz R., Crossa J. Breeding schemes for the implementation of genomic selection in wheat (*Triticum* spp.). Plant Science, 242 (2016), PP. 23-36. doi:10.1016/j.plantsci.2015.08.021

Borges V., Ferreira P.V., Soares L., Santos G.M., Santos A.M.M. Seleção de clones de batata-doce pelo procedimento REML/BLUP. Acta Scientiarum. Agronomy, 32(4) (2010), 643-649. doi: 10.4025/actasciagron. v32i4. 4837

Boysen, T. J., Heuer, C., Tetens, J., Reinhardt, F., & Thaller, G. (2013). Novel use of derived genotype probabilities to discover significant dominance effects for milk production traits in dairy cattle. Genetics, 193(2), 431-442. doi:10,1534 / genetics.112.144535

Burgueño J., Campos G. de los, Weigel K., Crossa J. Genomic prediction of breeding values when modeling genotype \times environment interaction using pedigree and dense molecular markers. Crop Science, 52(2) (2012), pp. 707-719. doi:10.2135/cropsci2011.06.0299

Casale F.P., Rakitsch B., Lippert C., Stegle O. Efficient set tests for the genetic analysis of correlated traits. Nature methods, 12(8) (2015), pp. 755-758. doi:10.1038/nmeth.3439

Cheung C.Y., Thompson E.A., Wijsman E.M. GIGI: an approach to effective imputation of dense genotypes on large pedigrees. The American Journal of Human Genetics, 92(4) (2013), pp. 504-516. doi:10.1016/j.ajhg.2013.02.011

Cobb J.N., DeClerck G., Greenberg A., Clark R., McCouch, S. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–

1 phenotype relationships and its relevance to crop improvement. *Theoretical and Applied*
2 *Genetics*, 126(4) (2013), pp. 867-887. DOI: 10.1007/s00122-013-2066-0

3 Davis K.F., Gephart J.A., Emery K.A., Leach A.M., Galloway J.N., D’Odorico, P.
4 Meeting future food demand with current agricultural resources. *Global Environmental*
5 *Change*, 39 (2016), pp. 125-132. doi:10.1016/j.gloenvcha.2016.05.004

6 Dempster A.P., Laird N.M., Rubin D.B. Maximum likelihood from incomplete data
7 via the EM algorithm. *Journal of the royal statistical society*, 39(1) (1977), pp. 1-38.
8 doi:http://www.jstor.org/stable/2984875

9 Diffey S., Welsh A., Cullis B. A faster and computationally more efficient REML
10 (PX) EM algorithm for linear mixed models. *Centre for Statistical and Survey Methodology*,
11 8 (2013), pp. 2-13.

12 Duvick D.N., Smith J.S.C., Cooper M. Long term selection in a commercial hybrid
13 maize breeding program. *Janick. I. Plant Breeding Reviews. Part*, 24(2) (2010), pp. 109-152.
14 doi:10.1002/9780470650288.ch4

15 Feddag M.L. Composite likelihood estimation for multivariate probit latent traits
16 models. *Communications in Statistics-Theory and Methods*, 42(14) (2013), pp. 2551-2566.
17 doi:10.1080/03610926.2010.538793

18 Feddag M.L., Bacci S. Pairwise likelihood for the longitudinal mixed Rasch model.
19 *Computational Statistics & Data Analysis*, 53(4) (2009), pp. 1027-1037.
20 doi:10.1016/j.csda.2008.08.031

21 Foley J.A., Ramankutty N., Brauman K.A., Cassidy E.S., Gerber J.S., Johnston M.,
22 Mueller N.D., O’connell C., Ray D.K., West P.C., Balzer C., Bennett E.M., Carpenter S.R.,
23 Hill J., Monfreda C., Polasky S., Rockstrom J., Sheehan J., Siebert S., Tilman D., Zaks
24 D.P.M. Solutions for a cultivated planet. *Solutions for a cultivated planet. Nature*, 478(7369)
25 (2011), pp. 337-342. doi:10.1038/nature10452

26 Gilmour A.R., Thompson R., Cullis B.R. Average information REML: an efficient
27 algorithm for variance parameter estimation in linear mixed models. *Biometrics*, v. 51, n. 4, p.
28 1440-1450, 1995. doi: http://www.jstor.org/stable/2533274

29 Godfray H.C.J., Beddington J.R., Crute I.R., Haddad L., Lawrence D., Muir J. F.,
30 PRETTY J., ROBINSON S., THOMAS S.M., Toulmin, C. Food security: the challenge of
31 feeding 9 billion people. *Science*, 327(5967) (2010), pp. 812-818.
32 doi:10.1126/science.1185383

1 Grattapaglia D., Resende, M.D. Genomic selection in forest tree breeding. *Tree*
2 *Genetics & Genomes*, 7(2) (2011), pp. 241-255. doi:10.1007/s11295-010-0328-4

3 Grün B., Leisch F. Identifiability of finite mixtures of multinomial logit models with
4 varying and fixed effects. *Journal of classification*, 25(2) (2008), pp. 225-247. doi:
5 10.1007/s00357-008-9022-8

6 HALLAUER, A. R.; CARENA, M. J.; MIRANDA FILHO, J. D. *Quantitative genetics*
7 *in maize breeding*. Iowa: Springer (2010).

8 Harville D.A. Making REML computationally feasible for large data sets: use of the
9 Gibbs sampler. *Journal of Statistical Computation & Simulation*, 74(2) (2004), pp. 135-153.
10 doi:10.1080/0094965031000110588

11 Hasstedt, S.J. A Mixed-Model Likelihood Approximation on Large Pedigrees. 15
12 (1982), pp. 295-307.

13 Heck D.W., Moshagen M., Erdfelder E. Model selection by minimum description
14 length: Lower-bound sample sizes for the Fisher information approximation. *Journal of*
15 *Mathematical Psychology*, 60 (2014), pp. 29-34. doi:10.1016/j.jmp.2014.06.002

16 Henderson C.R. Best linear unbiased estimation and prediction under a selection
17 model. *Biometrics* (1975), pp. 423-447.

18 Henderson C.R., Kempthorne O., Searle S.R., Von Krosigk C.M. The estimation of
19 environmental and genetic trends from records subject to culling. *Biometrics*, 15(2) (1959),
20 pp. 192-218. doi: <http://www.jstor.org/stable/2527669>

21 Isik F. Genomic selection in forest tree breeding: the concept and an outlook to the
22 future. *New forests*, 45(3) (2014), pp. 379-401. doi:10.1007/s11056-014-9422-z

23 Jeffreys H. *Teoria da Probabilidade*. Oxford University Press, Oxford, 3 ed. (1961).
24 doi: <http://www.emis.de/MATHitem?0116.34904>

25 Johnson D.L., Thompson R. Restricted maximum likelihood estimation of variance
26 components for univariate animal models using sparse matrix techniques and average
27 information. *Journal of dairy science*, 78(2) (1995), pp. 449-456. doi:10.3168/jds.S0022-
28 0302(95)76654-1

29 JOHNSON, D. L.; THOMPSON, R. Restricted maximum likelihood estimation of
30 variance components for univariate animal models using sparse matrix techniques and
31 average information. *Journal of dairy science*, v. 78, n. 2, p. 449-456, 1995. doi:
32 doi:10.3168/jds.S0022-0302(95)76654-1

1 Laurance W.F., Sayer J., Cassman K.G. Agricultural expansion and its impacts on
2 tropical nature. *Trends in ecology & evolution*, 29(2) (2014), pp. 107-116.
3 doi:10.1016/j.tree.2013.12.001

4 Li B., Chen W., Zhan X., Busonero F., Sanna S., Sidore C., Cucca F., Kang H.M.,
5 Abecasis G.R. A likelihood-based framework for variant calling and de novo mutation
6 detection in families. *PLoS Genet*, 8(10) (2012), e1002944.
7 <http://dx.doi.org/10.1371/journal.pgen.1002944>

8 Li, E.; Pourahmadi M. An alternative REML estimation of covariance matrices in
9 linear mixed models. *Statistics & Probability Letters*, 83(4) (2013), pp. 1071-1077.
10 doi:10.1016/j.spl.2012.12.028

11 Lobell D.B. Climate change adaptation in crop production: Beware of illusions. *Global*
12 *Food Security*, 3(2) (2014), pp. 72-76. doi:10.1016/j.gfs.2014.05.002

13 Loh, P. R., Bhatia, G., Gusev, A., Finucane, H. K., Bulik-Sullivan, B. K., Pollack, S.
14 J., Consortium S.W.G.P.G., Candia T.R., Lee S.H., Wray N.R., Kendler K.S., O'Donovan
15 M.C, Neale B.M., Patterson N., Price A.L. Contrasting genetic architectures of schizophrenia
16 and other complex diseases using fast variance-components analysis. *Nature genetics*. 47(12)
17 (2015), pp. 1385-1392. doi:10.1038/ng.3431

18 Long S.P., Ort D.R. More than taking the heat: crops and global change. *Current*
19 *opinion in plant biology*, 13(3) (2010), pp. 240-247. doi:10.1016/J.PBI.2010.04.008

20 Meyer K. Estimating variances and covariances for multivariate animal models by
21 restricted maximum likelihood. *Genetics Selection Evolution*, 23(1) (1991), pp. 67-83.

22 Meyer K. Parameter expansion for estimation of reduced rank covariance
23 matrices. *Genetics Selection Evolution*, 40(3) (2008), pp. 3-24. doi:10.1186/1297-9686-40-1-
24 3

25 MEYER, K. Restricted maximum likelihood to estimate variance components for
26 animal models with several random effects using a derivative-free algorithm. *Genetics*
27 *Selection Evolution*, v. 21, n. 3, p. 1-24, 1989.

28 Misztal I. Reliable computing in estimation of variance components. *Journal of animal*
29 *breeding and genetics*, 6(125) (2008), pp. 363-370. doi:10.1111/j.1439-0388.2008.00774.x

30 Misztal I., Perez-Enciso M. Sparse matrix inversion for restricted maximum likelihood
31 estimation of variance components by expectation-maximization. *Journal of dairy science*,
32 76(5) (1993), pp. 1479-1483. doi:10.3168/jds.S0022-0302(93)77478-0

33 Mrode R. A. *Linear models for the prediction of animal breeding values*. Cabi (2014).

1 Oury F.X., Godin C., Mailliard A., Chassin A., Gardet O., Giraud A., Heumez E.,
2 Morlais J.Y., B., Rousset M., Trottet M., Charmet G. A study of genetic progress due to
3 selection reveals a negative effect of climate change on bread wheat yield in France. *European*
4 *Journal of Agronomy*, 40 (2012), pp. 28-38. doi:10.1016/j.eja.2012.02.007

5 Patterson H.D., Thompson R. Recovery of inter-block information when block sizes
6 are unequal. *Biometrika*, 58(3) (1971), pp. 545-554. doi: 10.1093/biomet/58.3.545

7 Patterson H.D., Thompson R. Recovery of inter-block information when block sizes
8 are unequal. *Biometrika*, v. 58, n. 3, p. 545-554, 1971. doi: doi: 10.1093/biomet/58.3.545

9 Piepho H.P. Ridge regression and extensions for genomewide selection in maize. *Crop*
10 *Science*, 49(4) (2009), pp. 1165-1176. doi:10.2135/cropsci2008.10.0595

11 Piepho, H. P., Möhring, J., Melchinger, A. E., & Büchse, A. (2008). BLUP for
12 phenotypic selection in plant breeding and variety testing. *Euphytica*, 161(1) (2008), pp. 209-
13 228. doi:10.1007/s10681-007-9449-8

14 R CORE TEAM . R: A language and environment for statistical computing. R
15 Foundation for Statistical Computing, Vienna, Austria, 2016. Disponível em:
16 <<https://www.R-project.org/>>. Acesso em: 10 de maio de 2016.

17 Ray D.K., Mueller N.D., West P.C., Foley, J.A. Yield trends are insufficient to double
18 global crop production by 2050. *PloS one*, 8(6) (2013), pp. 1-8.
19 <http://dx.doi.org/10.1371/journal.pone.0066428>

20 Resende M.D.V de, Fonseca-Silva F., Lopes P.S., Azevedo C.F. Fundamentos
21 Estatísticos de Seleção Genética. Resende M.D.V de, Fonseca-Silva F., Lopes P.S., Azevedo
22 C.F. Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência
23 Bayesiana (MCMC), Regressão Aleatória Multivariada e Estatística Espacial. 2012, pp. 46

24 Resende M.D.V. de. REML e componentes de variância. Resende M.D.Vde,
25 Fonsenca-Silva F. Azevedo C.F. Estatística Matemática, Biométrica e Computacional. pp.
26 404-405.

27 Resende R.M.S, Resende M.D.V, Jank L. Experimentação e análise estatística no
28 melhoramento de forrageiras. In: Resende R.M.S., Valle C.B do, Jank L. (Eds.)
29 Melhoramento de Forrageiras Tropicais. Campo Grande, Embrapa Gado de Corte (2008),
30 pp.195-284.

31 Rincen R., Laloë D., Nicolas S., Altmann T., Brunel D., Revilla P., Rodríguez V.M.,
32 Moreno-Gonzalez J., Melchinger A., Bauer E., Schoen C.C., Meyer N., Giauffret C., Bauland
33 C., Jamin P., Laborde J., Monod H., Flament P., Charcosset A., Moreau L. Maximizing the

1 reliability of genomic selection by optimizing the calibration set of reference individuals:
2 comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). *Genetics*,
3 192(2) (2012), pp. 715-728. doi: 10.1534/genetics.112.141473

4 Rodrigues W.P., Vieira H.D., Barbosa D.H., Souza Filho G.R., Candido L.S.
5 Adaptability and genotypic stability of *Coffea arabica* genotypes based on REML/BLUP
6 analysis in Rio de Janeiro State, Brazil. *Genetics and Molecular Research*, 12(3) (2013), pp.
7 2391-2399.

8 Schwarz G. Estimating the dimension of a model. *The annals of statistics*, 6(2) (1978),
9 pp. 461-464.

10 Strzepek K., Boehlert B. Competition for water for the food system. *Philosophical*
11 *Transactions of the Royal Society of London B: Biological Sciences*, 365(1554) (2010), pp.
12 2927-2940. doi: 10.1098/rstb.2010.0152.

13 Szydłowski M., Krawiec A., Kurek A., Kamionka, M. AIC, BIC, Bayesian evidence
14 against the interacting dark energy model. *The European Physical Journal C*, 75(1) (2015), pp.
15 1-11. doi:10.1140/epjc/s10052-014-3236-1

16 Tester M., Langridge P. Breeding technologies to increase crop production in a
17 changing world. *Science*, 327(5967) (2010), pp. 818-822. doi:10.1126/science.1183700

18 Thompson R., Cullis B., Smith A., Gilmour A. A sparse implementation of the
19 average information algorithm for factor analytic and reduced rank variance models.
20 *Australian & New Zealand Journal of Statistics*, 45(4) (2003), pp. 445-459. doi:10.1111/1467-
21 842X.00297

22 Tilman D., Balzer C., Hill J., Befort B.L. Global food demand and the sustainable
23 intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108(50)
24 (2011), pp. 20260-20264. doi:10.1073/pnas.1116437108

25 Zhang, J., Song, Q., Cregan, P. B., & Jiang, G. L. (2016). Genome-wide association
26 study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine*
27 *max*). *Theoretical and Applied Genetics*, 129(1) (2016), pp. 117-130. doi:10.1007/s00122-
28 015-2614-x

29
30
31

1 **CAPITULO 2**

2
3 **IMPACTO ENTRE A COMBINAÇÃO DOS TESTADORES E AS LINHAGENS NA**
4 **SELEÇÃO DE HÍBRIDOS DE MILHO**

5
6 **RESUMO**

7 O milho (*Zea mays*) está entre as três principais culturas bases da alimentação mundial e é a
8 única que vem apresentando incremento de produtividade. Para que este progresso se
9 mantenha, a utilização de modelos que proporcionem uma melhor seleção dos híbridos e o
10 conhecimento da resposta da hibridização entre os testadores e as linhagens é de suma
11 importância. O objetivo do presente trabalho foi avaliar os efeitos das combinações dos dois
12 testadores em cruzamentos com progênies endogâmicas derivadas de populações de diferentes
13 grupos heteróticos, sobre os resultados de análises individuais e conjuntas, visando a seleção
14 de híbridos topcrosses superiores quanto à produtividade de grãos, considerando-se dois
15 diferentes modelos estatísticos. O conjunto de dados foi constituído por dois grupos de
16 experimentos, sendo o grupo 1 (E1) formado por 1801 híbridos TCs derivados de linhagens
17 S1 Flint hibridizadas com o testador G24 (Dent) e o grupo de ensaios 2 (E2) formado por
18 1551 Híbridos TCs derivados de linhagens Dent hibridizadas com o testador G8 (Flint). As
19 análises estatísticas foram realizadas através da metodologia de modelos mistos
20 (REML/BLUP). A coincidência entre os 5% selecionados fenotipicamente e genotipicamente
21 para todas as análises foi de 0,41 a 0,60. A coincidência entre os diferentes modelos e ensaios
22 foi de 43,45 a 96,67%. No grupo de ensaios 1 e na análise conjunta geral os valores dos
23 híbridos TCs selecionados foram superiores a das testemunhas do ensaio. Com este trabalho
24 podemos concluir que os melhores híbridos estavam alocados no grupo de ensaios E1, e que
25 o testador influencia a formação de híbridos.

26
27 Palavra Chave: Melhoramento; Heterose; Modelos Mistos; *Topcross*; Testador; Seleção

1. INTRODUÇÃO

A produção de alimentos para a população mundial deverá, pelo menos, ser duplicada caso não ocorram mudanças drásticas na alimentação da população, para manter a disponibilidade de alimento ao ritmo da demanda decorrente do incremento populacional (Godfray et al., 2010, Foley et al., 2011, Tilmana et al., 2011, Bajželj et al., 2014 and Davis et al., 2016). Além disso, existem diversos outros fatores que vem aumentando a pressão sobre os sistemas produtivos, como as mudanças climáticas, a perda da biodiversidade, degradação de terras e a demanda por bicompostíveis (Godfray et al., 2010, Strzepek and Boehlert, 2010, Ray et al., 2013 and Laurance et al., 2014), de modo que há a necessidade de aumentar na eficiência na produção agrícola, mesmo sob condições adversas.

O milho (*Zea mays*), ao lado do trigo e do arroz, está entre as três principais fontes de alimentos para a população mundial, sendo que o milho é a única que vem apresentando incremento ano após ano em produtividade (Long and Ort, 2010). Além disso, o seu uso é bastante diversificado decorrente das suas múltiplas aplicações (Teixeira et al., 2014), incluindo usos na alimentação humana e na formulação de rações para animais, usos industriais, e na geração de energia animal.

O melhoramento genético com seus avanços tecnológicos vem gerando incremento ano após ano, sendo extremamente crucial para a superação das adversidades (Tester and Langridge, 2010 and Araus and Cairns, 2014). Neste contexto, a exploração do vigor híbrido, ou heterose, é sem dúvida uma das maiores contribuições da ciência para o agronegócio mundial (East, 1908, Shull, 1908, Shull, 1911, Nerling et al., 2013 and Shapira and David, 2016).

A heterose é um fenômeno genético relativo à superioridade fenotípica de uma progênie híbrida em relação à média de seus pais (Li et al., 2012). Foi por meio da utilização da heterose que foram possíveis os incrementos expressivo em vários países em produtividade de grãos de milho, devido ao direcionamento dos programas de melhoramento para o desenvolvimento e lançamento de híbridos (Duvick et al., 2004 and Hallauer et al., 2010).

Os híbridos de milho, além de apresentar elevado potencial produtivo, devem ser portadores de caracteres agrônômicos interessantes para os agricultores, como ciclo e porte adequados, resistência a doenças e pragas, estabilidade de produção. Além disso, os parentais dos híbridos (que, para híbridos simples, são linhagens endogâmicas) precisam ser produtivos, de fácil manejo em termos de coincidência no florescimento, apresentar bom rendimento na classificação de sementes, dentre outros atributos, para que haja viabilidade técnica e econômica no processo de produção de sementes.

1 A equipe de melhoramento deve ainda considerar as influências ambientais durante a
2 seleção, as metodologia estatísticas aplicadas para a discriminação entre os materiais, o
3 potencial de heterose das novas linhagens derivadas de diversas populações fonte, a adequada
4 escolha de testadores (linhagens elites utilizadas em cruzamentos com grande número de
5 novas linhagens) e a viabilidade das novas linhagens como parentais, entre outros fatores.

6 Um dos fatores que exercem grande impacto na eficiência dos programas de
7 melhoramento é a utilização de testadores adequados durante o processo de seleção. A
8 escolha dos testadores é um desafio para os melhoristas, pois este deve fornecer a melhor
9 discriminação entre os genótipos de acordo com os propósitos da seleção (Hallauer, 1975, Li
10 et al., 2007, Edwards, 2009, Sreckov et al., 2011, Viana et al., 2011; Barreto et al., 2012,
11 Grieder et al., 2012, Guimarães et al., 2012, Rodovalho et al., 2012, and Clovis, et al., 2015).

12 Nos programas de melhoramento há a necessidade de identificação híbridos superiores
13 dentre uma gama de opções de cruzamentos entre as linhagens disponíveis, sendo que o
14 grande número de cruzamentos possíveis entre as novas linhagens constitui um grande
15 desafio. No entanto, o método de topcross foi desenvolvido com o intuito de otimizar este
16 desafio com a redução do tempo e recursos demandados (Elias et al., 2000 and Scapim et al.,
17 2008), pois consiste em avaliar um grande número de linhagens parcialmente endogâmicas, a
18 partir de seus cruzamentos com um testador comum, que atua como genitor masculino em
19 campos isolados (Davis, 1927).

20 O aumento da eficiência de seleção em programas de melhoramento de milho pode ser
21 obtida com a utilização do método de topcross em conjunto com a utilização de
22 delineamentos experimentais adequados e de metodologias estatísticas que permitam a melhor
23 discriminação entre híbridos portadores de caracteres de interesse, possibilitando o aumento
24 da eficiência de seleção sem aumento de custos (Cobb et al., 2013, Araus and Cairns, 2014,
25 Casale et al., 2015 and Loh et al., 2015).

26 Para maior eficiência no processo de seleção de cultivares superiores deve priorizar os
27 valores genotípicos, em detrimento de médias fenotípicas, tanto na etapa inicial quanto na
28 etapa final de um programa de melhoramento genético (Borges et al. 2010). Nas espécies
29 perenes e em animais uma alternativa empregada e altamente precisa é a estimação dos
30 componentes de variâncias pela máxima verossimilhança restrita (REML- *Restricted*
31 *Maximum Likelihood*) e a estimação dos valores genotípicos preditos pelo melhor preditor
32 linear não viesado (BLUP- *Best Linear Unbiased Prediction*) (Henderson, 1975; Rodrigues et
33 al, 2013).

1 O método REML/BLUP possibilita trabalhar com dados desbalanceados e incorporar
2 matrizes de parentesco na comparação dos genótipos. Esta forma de utilização é amplamente
3 utilizada no melhoramento animal, em espécies florestais e perenes, contudo, pesquisas com
4 culturas anuais ainda são recentes, mas estão se tornando uma alternativa eficiente na
5 tentativa de otimizar os recursos dos programas de melhoramentos, proporcionando bons
6 resultados (Grattapaglia and Resende, 2011, Burgueño et al., 2012, Rincent et al., 2012, Isik,
7 2014 and Bassia et al., 2016).

8 Os objetivos do nosso estudo com híbridos de milho de primeiro ciclo de seleção
9 foram avaliar os efeitos de dois testadores utilizados como parentais masculinos, em
10 cruzamentos com progênie endogâmicas (S1), extraídas de diferentes populações sob os
11 resultados de análises individuais e conjuntas, e verificar as implicações da utilização de
12 modelos estatísticos que incorporação de informações de parentesco via pedigree, mas com
13 diferentes estruturas para os fatores fixos, sobre a seleção de híbridos superiores para
14 produtividade de grãos.

2. MATERIAL E MÉTODOS

2.1 DADOS EXPERIMENTAIS

Os dados experimentais foram obtidos a partir de ensaios montados em blocos aumentados, com um grupo de experimentos que agrupou híbridos topcrosses (TCs) derivados de progênies endogâmicas (S1) do grupo heterótico “Dent” em cruzamento com um testador “Flint”, e outro grupo de experimentos que reuniu híbridos TCs derivados de cruzamentos entre progênies (S1) “Flint” em cruzamento com um testador “Dent”. Os ensaios foram conduzidos no ano agrícola de 2013/14, na Embrapa Milho e Sorgo, localizada no município de Sete Lagoas no estado de Minas Gerais. Na estrutura de blocos aumentados, utilizou-se 64 híbridos como número básico de tratamentos dentro de cada bloco, incluindo cinco testemunhas comuns a todos os blocos, mas houve experimentos com número diferente. Nessa estrutura, foram avaliados 3352 híbridos TCs (de primeiro ciclo) e mais cinco híbridos comerciais (utilizados como testemunhas comuns), totalizando 3357 tratamentos. As parcelas úteis foram compostas de duas linhas de 4,2 m espaçadas em 0,7 m. A característica analisada foi o rendimento de grãos, em kg ha⁻¹, determinada com base no peso de grãos da parcela, e corrigida para a umidade de 13%.

O grupo de ensaios 1 (E1) corresponde ao grupo de híbridos derivados de populações do grupo heterótico Flint, composto por 1801 híbridos TCs provenientes da hibridação de 1801 progênies S1 Flint e o testador DentG24 (521236) e o grupo de ensaios 2 (E2) é relativo ao grupo de 1551 híbridos TCs derivados de progênies S1 Dent, cruzadas com o testador Flint G8 (521283). Em ambos os grupos de ensaios foram utilizadas as mesmas cinco testemunhas, visando a comparação dos ensaios e a possibilidade da realização da análise conjunta (Totais).

2.2 MODELOS EMPREGADOS

As análises estatísticas foram processadas utilizando-se procedimentos de modelos mistos, sendo consideradas diferentes fontes de variação entre os efeitos fixos. Os dois modelos adotados foram baseados na seleção dos melhores modelos do capítulo 1, sendo estes os modelos M1 [1] e M2 [2]:

$$y = Xu + Za + Zd + Tc + Wb + \varepsilon [1]$$

onde: y é o vetor da variável a ser analisada; u é o vetor de efeitos da média geral assumido como fixo; a é o vetor de efeitos aditivos, assumidos como aleatórios; d é o vetor de efeitos de dominância, assumidos como aleatórios; c é o vetor de efeitos da testemunhas, assumidos como aleatórios; b é o vetor de efeitos dos blocos, assumidos como aleatórios; ε é o vetor de

1 erros aleatórios; X é a matriz de incidência para o efeito fixo; Z, T e W são as matrizes de
2 incidências para os feitos aleatórios descritos anteriormente.

$$3 \quad y = Xb + Za + Zd + Tc + \varepsilon [2]$$

4 onde: y é o vetor da variável a ser analisada; b é o vetor de efeitos fixos do modelo, que
5 considera a média geral somada aos efeitos de blocos; a é o vetor de efeitos aditivos,
6 assumidos como aleatórios; d é o vetor de efeitos de dominância, assumidos como aleatórios;
7 c é o vetor de efeitos de testemunhas, assumidos como aleatórios; X é a matriz de incidência
8 para o efeito fixo; Z e T são as matrizes de incidências para os feitos aleatórios descritos
9 anteriormente.

10
11 A diferença existente entre os modelos M1 e M2 é baseada na diferenciação dos
12 vetores de efeitos fixos e aleatórios. No M1, apenas a média geral foi considerada para o vetor
13 de efeito fixo e no M2 o vetor de efeito fixo foi composto pela média geral somada aos efeitos
14 de blocos, além disso, no M1 o bloco entra como efeito aleatório. Os demais efeitos aleatórios
15 em ambos os modelos são iguais.

17 2.3 MODELOS MISTOS

18 Na análise de modelos mistos, são obtidas estimativas dos valores dos efeitos fixos e
19 são obtidos valores preditos para os efeitos aleatórios do modelo estatístico (FRITSCHÉ-
20 NETO, 2008). O método de BLUP (*Best Linear Unbiased Prediction*- Melhor Predição Linear
21 não Viesado) foi proposto por Henderson (1975) com a função da predição dos efeitos
22 aleatórios, ajustando os dados aos efeitos fixos e, quando há desbalanceamento nos dados, os
23 valores preditos consideram número desigual de informações relacionadas aos diferentes
24 níveis do fator aleatórios.

25 O método de modelos mistos foi utilizado para a estimação dos parâmetros. O método
26 REML por um processo iterativo estima os parâmetros genéticos, através da maximização da
27 função de verossimilhança residual, sendo que um dos procedimentos amplamente utilizados
28 para este propósito é dada pela utilização do algoritmo numérico EM
29 (*Expectation Maximization*), que se baseia na primeira derivada (Dempster et al., 1977).

30 A obtenção da estimativa dos parâmetros de um modelo misto via REML ocorre em
31 um processo iterativo, onde a função de verossimilhança residual é maximizada. No presente
32 trabalho, o algoritmo numérico utilizado foi o EM (*Expectation Maximization*), caracterizado
33 como de primeira derivada (Dempster et al., 1977)

Os modelos lineares mistos utilizados neste estudo são uma generalização do originalmente proposto por Henderson et al. (1959) e pode ser escrito na forma matricial usando a notação [9]

$$y = X\beta + Zu + \varepsilon \text{ com } N_u = \sum_{i=1}^I N_{u_i} \quad [3]$$

onde y : vetor de $N \times 1$ observações; “ β ” vetor de $N_\beta \times 1$ efeitos fixos e covariáveis fixas; “ X ” matriz de incidência para efeitos fixos e covariáveis fixas de dimensão $N \times N_\beta$; “ u ” vetor que concatena i vetores de efeitos aleatórios $u = (u_1, u_2, u_3 \dots u_i)$ de dimensão $N_u \times 1$; “ Z ” matriz de incidência que agrupa i matrizes de efeitos aleatórios $Z = (Z_1, Z_2, Z_3 \dots Z_i)$ de dimensão $N \times N_u$ e “ ε ” vetor de $N \times 1$ erros aleatórios.

A matriz de estruturas de (co)variância do modelo [3] pode ser expressa por

$$\begin{bmatrix} u \\ \varepsilon \end{bmatrix} \sim N \left(0, \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \right), \quad G = \bigoplus_{i=1}^I G_i \quad [4]$$

$$\text{var}(u_i) = H_i \sigma_{u_i}^2 = G_i \text{ ou } I \sigma_{u_i}^2 = G_i \text{ para } \forall H_i = I,$$

$$\text{var}(\varepsilon) = I \sigma_\varepsilon^2 = R, \quad \text{cov}(u_i, \varepsilon) = 0, \quad \text{var}(y) = V = ZGZ' + R,$$

onde “ G ” matriz de (co)variância dos efeitos aleatórios; “ R ” matriz de (co)variância residual; “ V ” matriz de (co)variância de y , “ H_i ” matriz de correlação entre os efeitos aleatórios de u_i , “ I ” matriz identidade, “ $\sigma_{u_i}^2$ ” variância de u_i e “ σ_ε^2 ” variância do erro.

As soluções de máxima verossimilhança de β e $u = (u_1, u_2, u_3 \dots u_i)$ do modelo [3], podem ser obtidas através da primeira derivada da função de densidade de probabilidade conjunta $f(y, u)$, proposta por Henderson et al. (1959), em relação aos parâmetros β e u e então obtemos o sistema de equações de modelos mistos (SEMM), assumindo que a estrutura de R^{-1} é uma matriz identidade (MRODE, 2014) podemos reescrever o SEMM como [5]

$$\begin{bmatrix} X'X & X'Z_1 & X'Z_2 & \dots & X'Z_j \\ Z_1'X & Z_1'Z_1 + H_1^{-1}\lambda_1 & Z_1'Z_2 & \dots & Z_1'Z_j \\ Z_2'X & Z_2'Z_1 & Z_2'Z_2 + H_2^{-1}\lambda_2 & \dots & Z_2'Z_j \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_i'X & Z_i'Z_1 & Z_i'Z_2 & \dots & Z_i'Z_j + H_i^{-1}\lambda_i \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \tilde{u}_1 \\ \tilde{u}_2 \\ \vdots \\ \tilde{u}_i \end{bmatrix} = \begin{bmatrix} X'y \\ Z_1'y \\ Z_2'y \\ \vdots \\ Z_i'y \end{bmatrix} \quad [5]$$

$$\lambda_1, \lambda_2 \dots \lambda_i = \frac{\sigma_\varepsilon^2}{\sigma_{u_1}^2}, \frac{\sigma_\varepsilon^2}{\sigma_{u_2}^2} \dots \frac{\sigma_\varepsilon^2}{\sigma_{u_i}^2} \quad [6]$$

1 A solução dos SEMM apresentados em [5] assume que $(\sigma_{u_1}^2, \sigma_{u_2}^2 \dots \sigma_{u_i}^2)$ e σ_ε^2 são
 2 conhecidos. Porém, na prática tais parâmetros são desconhecidos e podem ser substituídos por
 3 suas estimativas, $(\hat{\sigma}_{u_1}^2, \hat{\sigma}_{u_2}^2 \dots \hat{\sigma}_{u_i}^2)$ e $\hat{\sigma}_\varepsilon^2$, obtidas através de REML.

5 2.3.1 Algoritmo *Expectation Maximization* (EM)

6 A utilização do algoritmo EM, considerou o SEMM apresentado em [5] e uma
 7 generalização das equações apresentadas por Mrode (2014), da função de verossimilhança
 8 residual, obtidas através do rearranjo das expressões de primeira derivada

$$9 \quad [N - r(\mathbf{X})][\hat{\sigma}_\varepsilon^2] = \left[(y - \mathbf{X}\hat{\beta} - \sum_{i=1}^I \mathbf{Z}_i \tilde{u}_i)' y \right] [7]$$

$$10 \quad \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_i \end{bmatrix} \begin{bmatrix} \hat{\sigma}_{u_1}^2 \\ \hat{\sigma}_{u_2}^2 \\ \vdots \\ \hat{\sigma}_{u_i}^2 \end{bmatrix} = \begin{bmatrix} \tilde{u}_1' \mathbf{H}_1^{-1} \tilde{u}_1 + \text{tr}(\mathbf{H}_1^{-1} \mathbf{C}^{22}) \hat{\sigma}_\varepsilon^2 \\ \tilde{u}_2' \mathbf{H}_2^{-1} \tilde{u}_2 + \text{tr}(\mathbf{H}_2^{-1} \mathbf{C}^{33}) \hat{\sigma}_\varepsilon^2 \\ \vdots \\ \tilde{u}_i' \mathbf{H}_i^{-1} \tilde{u}_i + \text{tr}(\mathbf{H}_i^{-1} \mathbf{C}^{kk}) \hat{\sigma}_\varepsilon^2 \end{bmatrix} [8]$$

11 onde “ $r(\mathbf{X})$ ” é o *rank* da matriz \mathbf{X} e $(\mathbf{C}^{22}, \mathbf{C}^{33} \dots \mathbf{C}^{kk})$ são obtidas da inversa da matriz de
 12 coeficientes do SEMM [5]

$$13 \quad \mathbf{C}^{-1} = \begin{bmatrix} \mathbf{C}^{11} & \mathbf{C}^{12} & \mathbf{C}^{13} & \dots & \mathbf{C}^{1j} \\ \mathbf{C}^{21} & \mathbf{C}^{22} & \mathbf{C}^{23} & \dots & \mathbf{C}^{2j} \\ \mathbf{C}^{31} & \mathbf{C}^{32} & \mathbf{C}^{33} & \dots & \mathbf{C}^{3j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}^{i1} & \mathbf{C}^{i2} & \mathbf{C}^{i3} & \dots & \mathbf{C}^{ij} \end{bmatrix} [9]$$

14
 15 O cálculo dos traços matriciais foi feitos por meio da soma do produto de Hadamard
 16 [10]

$$17 \quad \text{tr}(\mathbf{H}_i^{-1} \mathbf{C}^{kk}) = \sum \mathbf{H}_i^{-1} \circ \mathbf{C}^{kk} = \sum_{i=1, j=1}^{I, J} h_{ij} c_{ij} [10]$$

19 2.4 ANÁLISES ESTATÍSTICAS

20 Os grupos de ensaios E1 e E2 e o modelo completo (E1+E2), foram submetidos a
 21 análise em dois modelos (M1 e M2) e foram estimados os componentes de variâncias (CV), o
 22 coeficiente de determinação (CD), a *deviance* [11] e o teste de razão de verossimilhança
 23 (LTR) [12] para as fontes de variações dos blocos (Blocos ou Ensaios), das testemunhas
 24 (Testemunhas), a variação genética devido ao efeito aditivo (a) e a variação genética devido
 25 ao efeito de dominância (d).

1 Mediante as estimativas dos componentes de variância genotípicos preditos para os
2 diferentes modelos (M1 e M2) e grupos de ensaios (M1, M2 e conjunto), aplicou-se a análise
3 de dispersão gráfica entre os valores genotípicos preditos e as médias fenotípicas do
4 rendimento de grãos (Kg ha^{-1}) e acurácia.

5 Além disso, para se verificar o comportamento da distribuição dos dados dentro de
6 cada bloco ao longo dos grupos de ensaios E1 e E2 e entre as testemunhas nos diferentes
7 ensaios, foram demonstrado os resultados por gráficos de Boxplot e os BLUPs dos testadores
8 utilizados como parentais dos híbridos TCs (E1 e E2) para cada modelo (M1e M2) por
9 gráficos de barra.

10 Considerando-se os dados genotípicos estimados pelos BLUPs em cada um dos
11 modelos (M1 e M2), aplicou-se uma intensidade de seleção de 5% dentro de cada grupo de
12 híbridos TCs (E1 e E2) e na análiseconjunta total (E1+E2), sendo assim no E1 foram
13 selecionados 90 híbridos, no E2 foram selecionados 78 híbridos e na análiseconjunta total
14 foram selecionados 168 híbridos.

15 Para os híbridos selecionados genotipicamente aplicou-se um gráfico de radar entre as
16 médias genotípicas e fenotípicas de produção (Kg ha^{-1}) e calculou-se a correlação destes pares
17 de dados pelo método de Spearman (ρ) [13] e testou a significância por χ^2 .

18 Baseando-se nos 5% selecionados verificou as possíveis coincidências entre os
19 híbridos dos diferentes ensaios e entre os selecionados pela análise conjunta total e os ensaios
20 individuais. Também foram estimadas as médias dos valores genéticos e dos rendimentos de
21 grãos dos 5% selecionados, para cada um dos modelos (M1 e M2) e ensaios (E1, E2 e na
22 conjunta total).

23 Todas as análises estatísticas foram realizadas pelo programa R, versão 3.2.2 (R Core
24 Team, 2015).

25 A *Deviance* é descrita como [11]

$$\text{Deviança} = -2\log(L) [11]$$

27 onde “L” é o ponto de máxima da função de verossimilhança residual.

28 A LTR (Teste de Razão de Verossimilhança) (Wilks, 1938) é descrita como [12]

$$\text{LTR} = \lambda = 2\log_e L(V) - 2\log_e L(U)$$

30 onde “LTR” igual a “ λ ” é a significância da diferença no ajuste de diferentes modelos aos
31 dados testados pela Razão de Verossimilhança de Wilks (1938); “V” equivale ao modelo
32 saturado a reprodução dos dados; “U” equivale ao modelo saturado a reprodução dos dados

33 A correlação pelo método de Spearman (ρ) é descrita como [13]

1
$$\rho = 1 - \frac{6 \sum d_i^2}{(n^3 - n)} \quad [13]$$

2 onde “ ρ ” é a correlação; “ d_i ” é a diferença entre cada posto de valor correspondentes de x e y;

3 “n” é o número de pares dos valores

4

3. RESULTADO

Os componentes de variância relativos aos efeitos genéticos aditivos mostraram significância estatística, pelo teste LRT, nos grupos de ensaios E1 e E2 e na análise conjunta total, para os dois modelos testados. Por outro lado, os componentes de variância genéticas associados aos efeitos de dominância, nos grupos de ensaios individuais foram não significativos, mas, para nas análises conjuntas totais, as variâncias dos efeitos de dominância foram significativas para ambos os modelos testados (Tabela 1). Os efeitos fixos de blocos foram significativos, indicando heterogeneidade entre as faixas de cultivo; e o efeito da testemunha, de modo geral, para todos os modelos e ensaios absorveu grande parte da variação fenotípica.

Os valores de acurácia seletiva individual para os híbridos TCs variou de 0,55 a 0,75, e para as testemunhas os valores ficaram próximos de 1,00 para os modelos M1 e M2 no grupo de ensaios E1 e na análise conjunta, enquanto que os valores de acurácia foram de 0,9 no E2 para as testemunhas nos modelos M1 e M2 (Figuras 1 e 2, respectivamente).

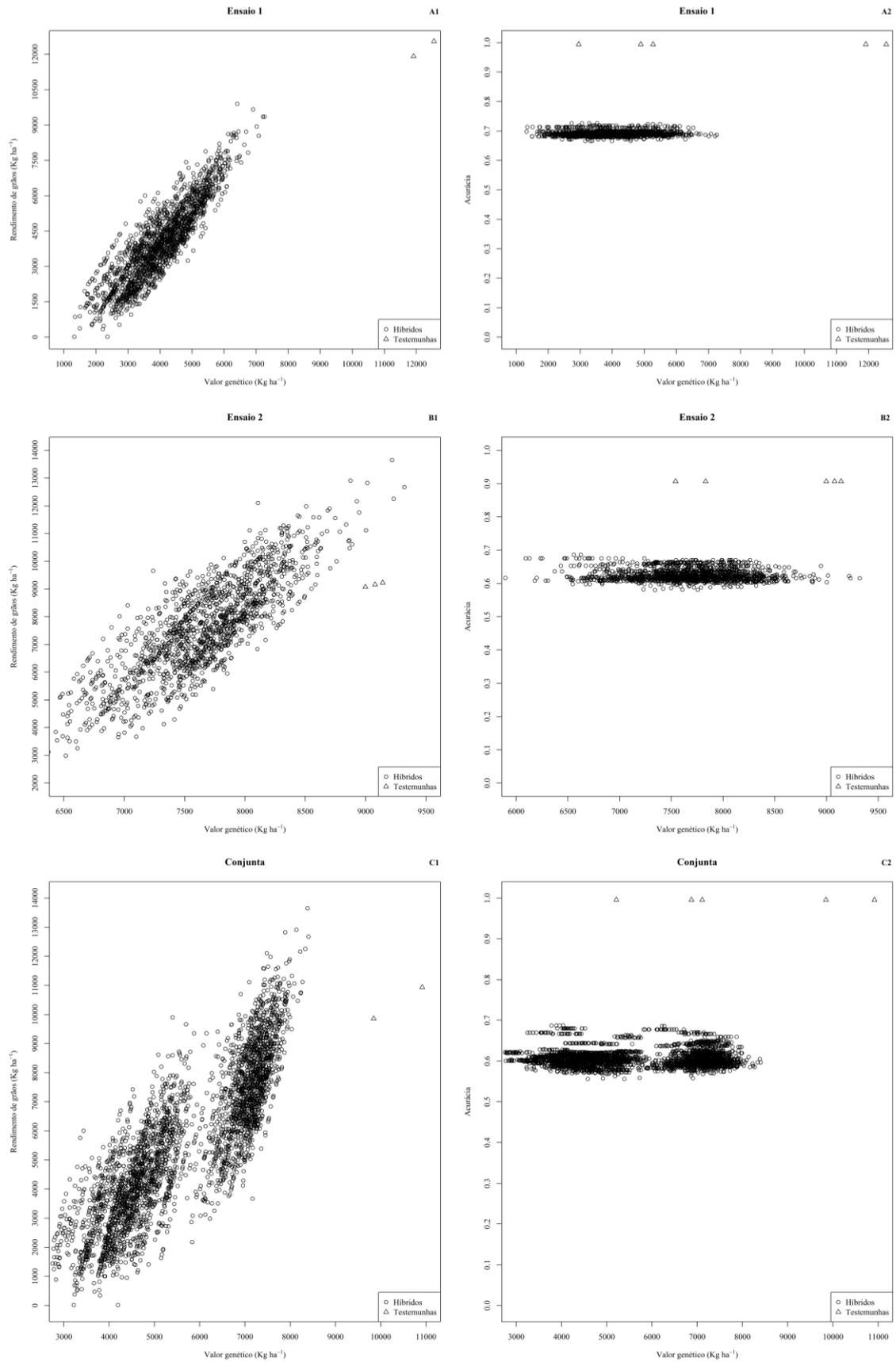
Dentre os ensaios pode-se verificar a existência da diferenciação entre os valores genéticos preditos. O grupo de ensaios E1 apresentou menores valores genéticos para os híbridos TCs comparado ao E2 em ambos os modelos (Figura 1 e 2 - A1, B1 respectivamente), podendo-se evidenciar melhor a diferenciação na análise conjunta, onde há uma separação visual em dois conjuntos de dados, sendo E1 e E2 respectivamente (Figura 1 e 2, C1). O grupo de ensaios E2 contém maior número de híbridos TCs com maiores valores genéticos preditos, ou seja, com alto potencial para rendimento de grãos.

Na Figura 3, pode-se observar que os híbridos TCs alocados no grupo de ensaios E2 apresentaram maior rendimento de grãos. Este conjunto de dados também apresentou maior quantidade “outliers” em todos os blocos (32 a 58) comparados com o grupo de ensaios E1 (blocos 1 a 31), apesar disto os dois ensaios apresentaram amplitudes médias próximas.

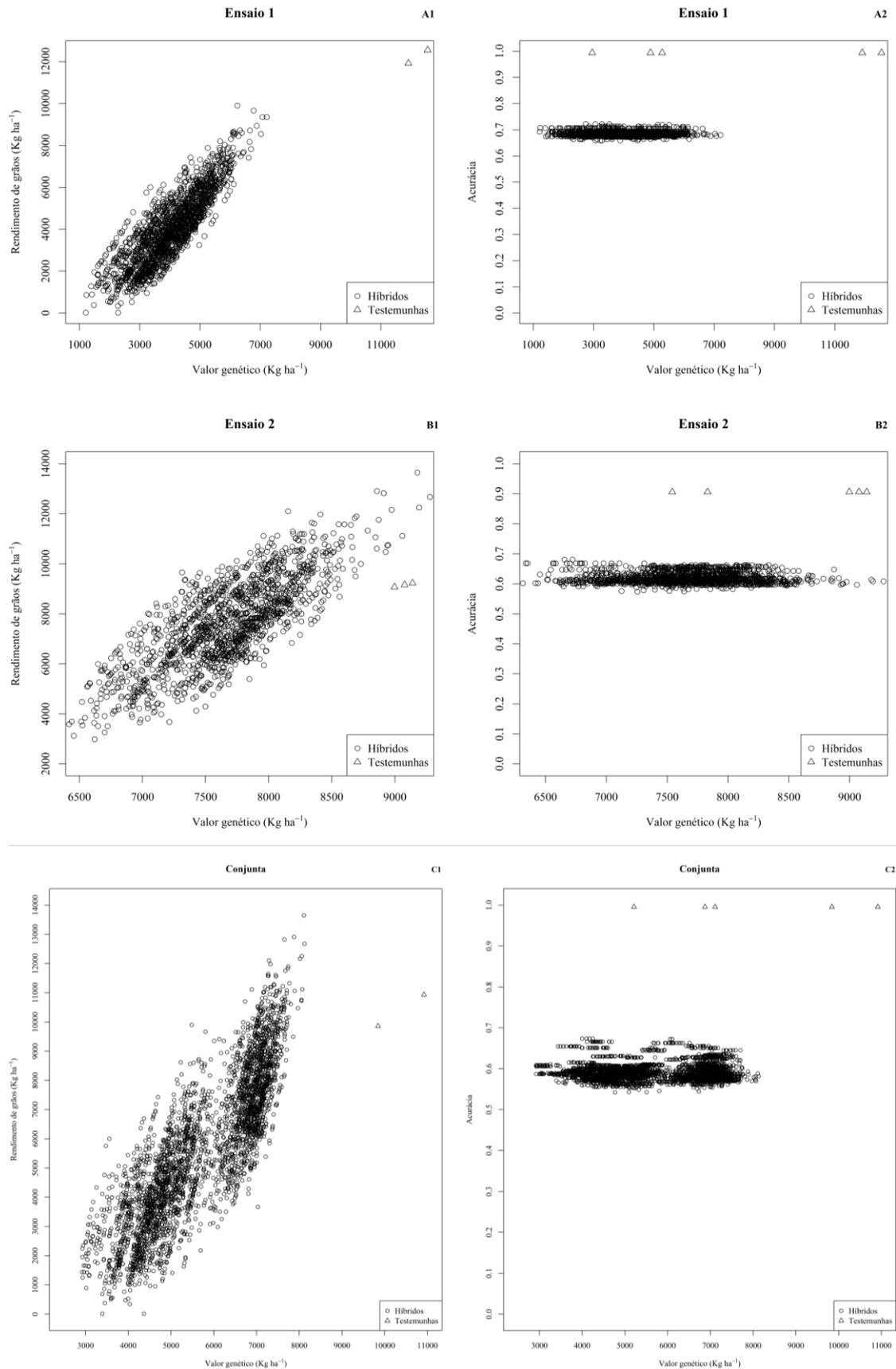
1 Tabela 1 Análise de *deviance* para os ensaios E1 (G24) e E2 (G8) e conjunta(E1 + E2)
 2 considerando os modelos M1 ($y = Xu + Za + Zd + Tc + Wb + \varepsilon$) e M2 ($y = Xb + Za + Zd +$
 3 $Tc + \varepsilon$).

Modelo	Ensaio	FV	CV	CD	<i>Deviance</i>	LRT	PV
M1	E1	Blocos	598625,18	0,15	26590,54	143,45	0,00
		Testemunhas	776110,99	0,19	26475,52	28,44	0,00
		Híbridos (a)	768471,95	0,19	26452,39	5,31	0,02
		Híbridos (d)	112492,76	0,03	26447,11	0,03	0,87
		Resíduos	1762439,9	0,44	---	---	---
		Modelo Completo	---	---	26447,08	---	---
	E2	Blocos	384703,81	0,01	30264,21	111,99	0,00
		Testemunhas	27574298,74	0,91	31419,58	1267,36	0,00
		Híbridos (a)	1301456,48	0,04	62431,39	32279,17	0,00
		Híbridos (d)	178022,1	0,01	30152,56	0,33	0,56
		Resíduos	902354,17	0,03	---	---	---
		Modelo Completo	---	---	30152,22	---	---
	Conjunta	Blocos/Grupos	849736,11	0,06	57771,33	327,84	0,00
		Testemunhas	9937675,98	0,73	58062,90	619,41	0,00
		Híbridos (a)	602583,89	0,04	114915,05	57471,56	0,00
Híbridos (d)		126150,67	0,01	57447,20	3,71	0,05	
Resíduos		2023497,26	0,15	---	---	---	
Modelo Completo		---	---	57443,49	---	---	
M2	E1	Testemunhas	27337325,48	0,92	30991,00	1263,25	0,00
		Híbridos (a)	1285474,01	0,04	29755,85	28,10	0,00
		Híbridos (d)	173816,26	0,01	29728,09	0,34	0,56
		Resíduos	912088,78	0,03	---	---	---
		Modelo Completo	---	---	29727,75	---	---
	E2	Blocos*			306,46		0,00
		Testemunhas	766717,17	0,23	26096,81	28,23	0,00
		Híbridos (a)	758729,29	0,22	26073,82	5,25	0,02
		Híbridos (d)	102911,99	0,03	26068,60	0,03	0,87
		Resíduos	1774023	0,52	---	---	---
Modelo Completo	---	---	26068,58	---	---		
Conjunta	Blocos*			326,64		0,00	
	Testemunhas	9986490,52	0,79	57213,75	618,79	0,00	
	Híbridos (a)	558741,86	0,04	56745,86	150,90	0,00	
	Híbridos (d)	131449,44	0,01	56598,77	3,81	0,05	
	Resíduos	2039353,74	0,16	---	---	---	
	Modelo Completo	---	---	56594,96	---	---	
	Blocos/Grupos *			3867,70		0,00	

4 FV – Fontes de variação, CV – componentes de variância, CD – coeficientes de determinação, LRT – Razão de
 5 verossimilhança, pv – valor p, * efeitos fixos, (a) – Variação genética devida aos efeitos aditivos, (d) – variação
 6 genética devida aos efeitos de dominância.



1 Figura 1 Dispersão gráfica das médias fenotípicas do rendimento de grãos (Figuras – A1, B1, 2 C1) e acurácias (Figuras – A2, B2, C2) em relação aos valores genotípicos preditos para o 3 ensaio 1, 2 e conjunto respectivamente considerando o modelo M1.



1 Figura 2. Dispersão gráfica das médias fenotípicas do rendimento de grãos (Figuras – A1, B1,
 2 C1) e acurácias (Figuras – A2, B2, C2) em relação aos valores genotípicos preditos para os
 3 grupos de ensaios E1, E2 e na análise conjunta, respectivamente, considerando o modelo M2.

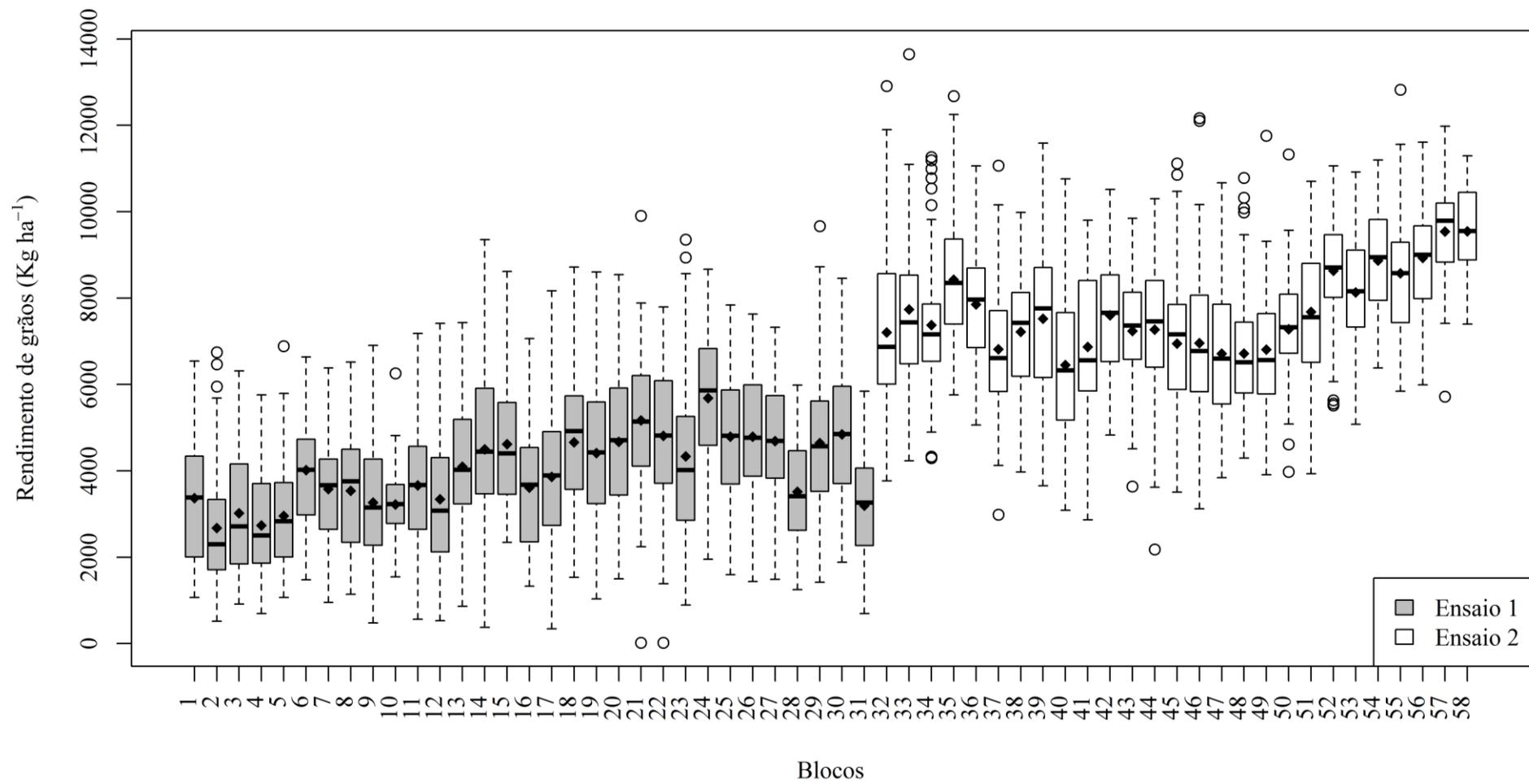
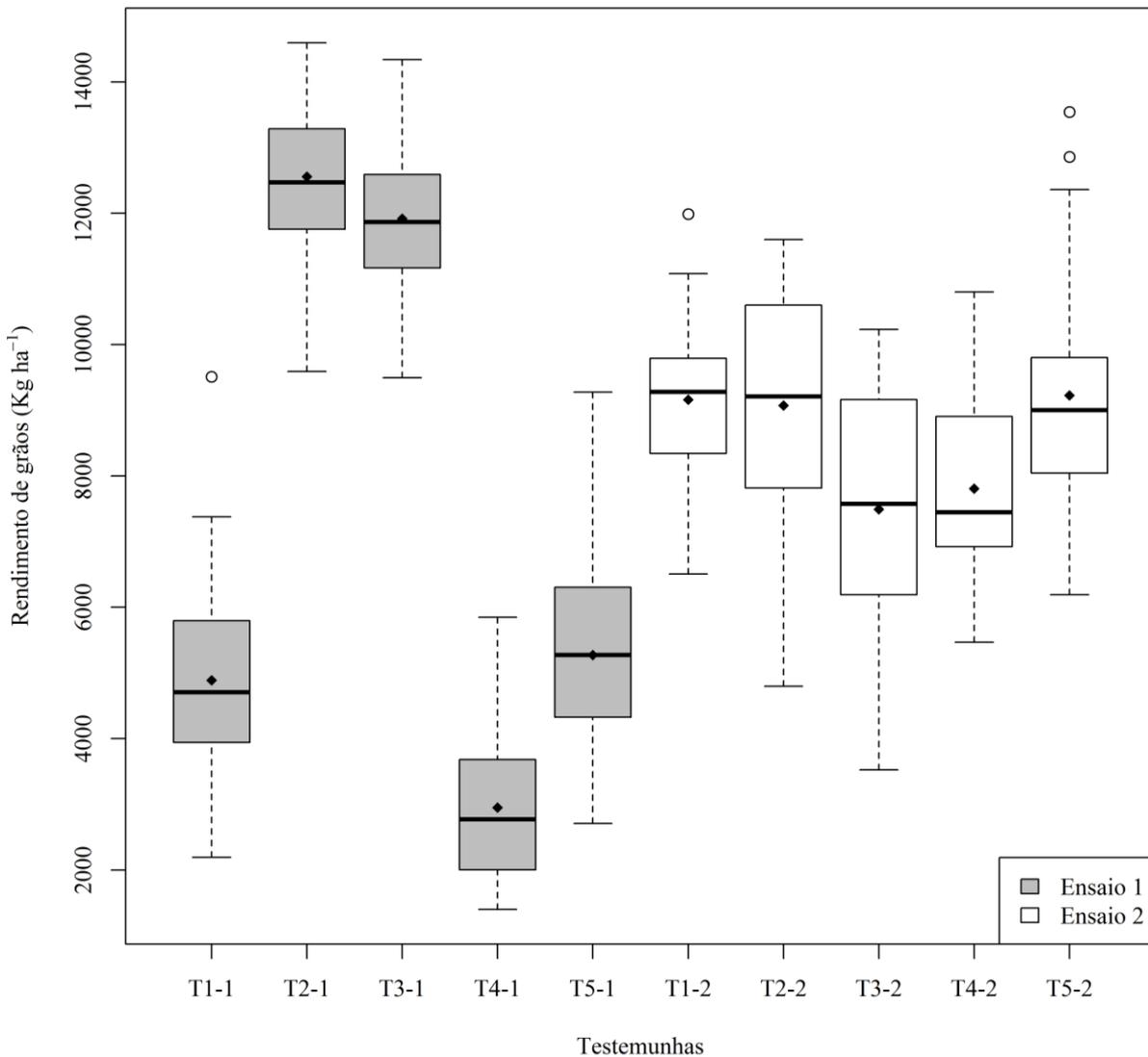


Figura 3. Boxplot do comportamento dos híbridos TC ao longo dos blocos nos grupos de ensaios E1 e E2

1 As testemunhas (cinco híbridos comerciais) foram avaliados separadamente entre si e
2 entre os diferentes grupos de ensaios (Figura 4). No grupo de ensaios E1 as testemunhas
3 demonstraram maior variação média, com maior valor de rendimento de grãos de,
4 aproximadamente, 15.000 Kg ha⁻¹ no quarto quartil da testemunha 2 (T2-1) e o menor valor
5 menor que 2.000 Kg ha⁻¹, paraa testemunha (T4-1). No grupo E1, as testemunhas T2-1 e T3-1
6 apresentaram rendimento de grãos maior que no grupo E2 e as testemunhas T1-2, T4-2 e T5-2
7 no E2 apresentaram maior rendimento de grãos que no grupo de ensaios E1.

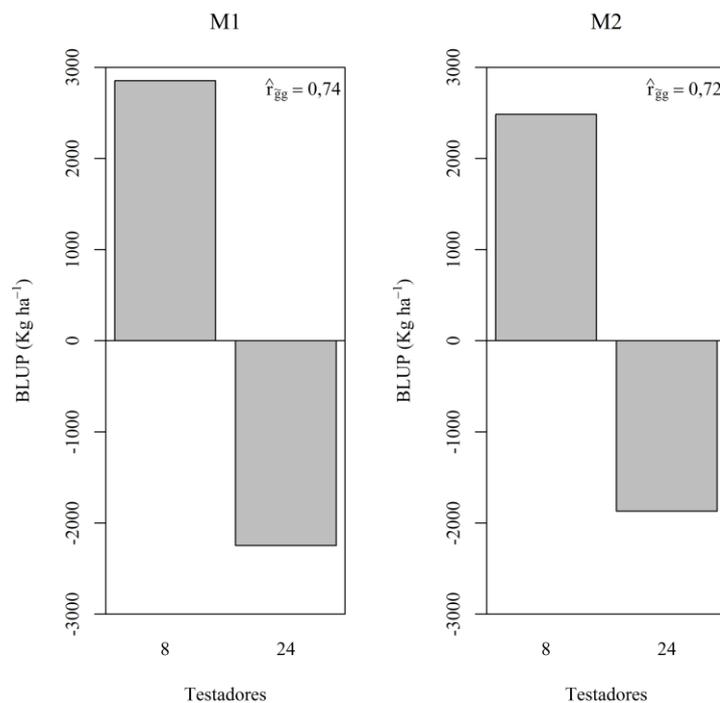
8
9
10



11
12
13

Figura 4. Boxplot do comportamento das testemunhas nos grupos de ensaios E1 e E2.

1 Analisando os dados conjuntamente para ambos os modelos, verifica-se que o testador
 2 Flint (521283) contribui para aumentar a média genotípica de rendimento de grãos dos
 3 híbridos TCs, enquanto o testador Dent (521236) provocou a diminuição da média genotípica
 4 geral (Figura 5). Os valores de acurácia, nas análises conjuntas totais, foram de 0,74 e 0,72
 5 para os modelos M1 e M2, respectivamente. O modelo M1 possibilitou a predição de valores
 6 genotípicos de rendimento de grãos com maior amplitude, com valores de desvios da média
 7 geral para os BLUPs entre -2000 Kg ha⁻¹ e 3000 Kg ha⁻¹.
 8

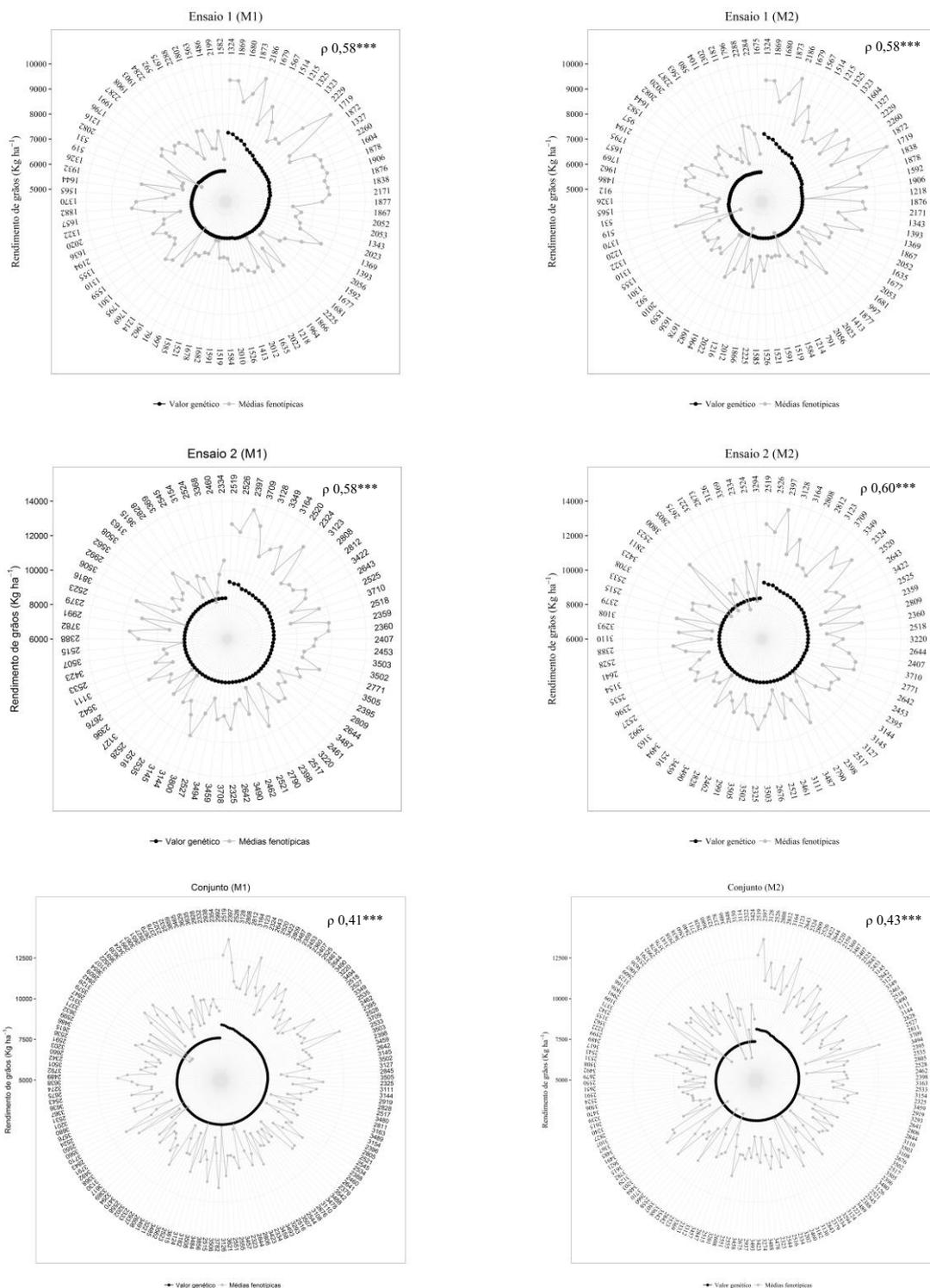


9
 10 Figura 5. BLUP dos testadores para os modelos M1 ($y = Xu + Za + Zd + Tc + Wb + \epsilon$) e M2
 11 ($y = Xb + Za + Zd + Tc + \epsilon$).
 12

13 Considerando a classificação pelos valores genotípicos preditos, via BLUP, para os 5%
 14 melhores híbridos, quanto ao rendimento de grãos, pode-se observar a oscilação dos dados
 15 fenotípica nos diferentes grupos de ensaios (E1, E2 e Conjunta Total) e para diferentes
 16 modelos estudados (M1 e M2) (Figura 6). Pode-se observar a existência de híbridos TCs com
 17 maiores ou com menores médias fenotípicas para rendimentos de grãos que os valores
 18 genotípicos preditos via BLUP.

19 Comparando-se os ranks dos selecionados fenotipicamente e genotipicamente, foram
 20 obtidas estimativas de coeficiente de correlação de Spearman de 0,58; 0,58; 0,58; 0,60; 0,41 e

- 1 0,43, nos respectivos grupos de ensaios e modelos: E1 (M1); E1 (M2); E2 (M1); E2 (M2);
- 2 Conjunta Total (M1); e Conjunta Total (M2).



3 Figura 6. Gráfico de radar entre médias fenotípicas e genotípicas de rendimento de grãos (Kg
4 ha⁻¹) dos híbridos TCs selecionados em cada um dos grupos de ensaios (E1 e E2) e na análise
5 conjunta total para os modelos (M1 e M2), considerando uma intensidade de seleção de 5%. ρ
6 – correlação de Spearman e testado a significância por χ^2 .

Na Tabela 2, podem-se verificar diferentes porcentagens de coincidência na seleção dos 5% híbridos TCs superiores quanto ao rendimento de grãos. A coincidência na seleção de híbridos TCs entre os modelos M1 e M2, no grupo de ensaios E1, foi de 96,67%, e no grupo de ensaios E2 a coincidência foi de 98,72%. Considerando-se o modelo M1, obteve-se índices de coincidência de 0% para a seleção praticada na com base na análise conjunta total (M1 + M2) e o grupo E1, e de 43,45%, entre a seleção na análise conjunta e no grupo de híbridos selecionados em E2.

Tabela 2. Índices de coincidência na seleção de 5% dos híbridos TCs superiores, quanto ao rendimento de grãos, para os modelos M1, M2 e análise conjunta total diferentes grupos de ensaios (E1, E2 e conjunta)

Ensaio/Modelo	E1 (5%)		E2 (5%)		Conjunto (5%)		
	M1	M2	M1	M2	M1	M2	
E1 (5%)	M1	90	96,67	---	---	0	0
	M2		90	---	---	0	0
E2 (5%)	M1			78	98,72	43,45 [#]	43,45 [#]
	M2				78	43,45 [#]	43,45 [#]
Conjunto (5%)	M1					168	96,43
	M2						168

Os valores fenotípicos para o rendimento de grãos (em kg.ha⁻¹) dos híbridos TCs selecionados entre os 5% mais produtivos no modelo M1 nos grupos de ensaios E2 (10.466) e conjunta (9.861) e no modelo M2 nos ensaios E2 (10.310) e conjunta (9.690) foram superiores ao rendimento de grãos das testemunhas (Tabela 3). Observa-se, ainda, que os valores genéticos obtidos via BLUP para os 5% selecionados no grupo de híbridos alocados no grupo de ensaios E2 (TCs derivados de progênies S1 Dent cruzadas com o testador Flint), para os modelos M1 e M2 (8597 e 8587, respectivamente) foram superiores aos valores genéticos obtidos para as testemunhas (8517 e 8517, respectivamente), mesmo com menores valores de acurácia para os grupos de híbridos TCs em relação às testemunhas, demonstrando o alto potencial de rendimento de grãos dos híbridos de primeiro ciclo selecionados neste trabalho.

1 Tabela 3. Médias dos valores genéticos e fenotípicos dos híbridos TCs selecionados entre os
 2 5% mais produtivos, das testemunhas e a média geral dentro de cada modelo estatístico (M1
 3 e M2), para os grupos de ensaios (E1, E2 e naconjunta total)

Modelo	Médias	E1	E2	Conjunta
M1	Valor genético (5%)	6080	8597	7774
	Valor genético testemunhas	7514	8517	7991
	Valor Fenotípico (5%)	7445	10466	9861
	Valor Fenotípico das testemunhas	7518	8551	7999
	Média geral	4083	8109	5644
M2	Valor genético (5%)	6029	8587	7544
	Valor genético testemunhas	7514	8517	7990
	Valor Fenotípico (5%)	7334	10310	9690
	Valor Fenotípico das testemunhas	7518	8551	7999
	Média geral	4117	8118	5634

4

5

1 4. DISCUSSÃO

2 A influência significativa da variação genética aditiva pode ser um indício de
3 endogamia na população de híbridos estudada (Charlesworth and Willis, 2009 and Reid et al.,
4 2011) e o efeito significativo dos blocos demonstra a existência de efeitos controlados e não
5 controlados atuando de forma significativa para o resultado observado nos fenótipos dos híbridos
6 testados, por este motivo a utilização do delineamento adequado é de extrema importância
7 para uma melhor estimação dos componentes de variâncias genotípicos (Yang, 2007,
8 Mohammadi et al., 2015) (Tabela 1).

9 A acurácia seletiva individual do experimento pode ser considerada alta, sendo que
10 quanto maior melhor, pois nos remete informações de precisão experimental (Resende et al.,
11 2008), assim esta informação dá uma ideia de confiabilidade da informação apresentada ao
12 pesquisador (Figura 1 e 2).

13 A diferenciação entre o rendimento de grãos e a quantidade de outliner pode ser
14 decorrente da interação entre o testador e a população de linhagens formadoras dos híbridos,
15 efeito conhecido como vigor híbrido ou heterose (Charlesworth and Willis, 2009, Li et al.,
16 2012 and Khaleida et al., 2013), ou causado pela influência ambiental durante o cultivo dos
17 diferentes ensaios (Torres et al., 2015) (Figura 3).

18 Porém como amplitude de modo geral entre os híbridos dentro de cada bloco em
19 ambos os ensaios são próximas, podemos destacar que a influência entre as médias dos
20 ensaios é devido a interação genótipos ambientes, pois quanto maior a intensidade dos efeitos
21 não controlados maiores as variações dentro das parcelas avaliadas.

22 Mediante a informação anterior chegamos à conclusão que a diferenciação foi devido
23 a interação proveniente do testador e das linhagens para a formação dos híbridos testados,
24 além disso, resultados similares foram apresentados em diversos trabalhos (Li et al., 2012 and
25 Bertoiaa and Aulicinob, 2014).

26 Estudando-se o comportamento entre as testemunhas nos diferentes ensaios, observou-
27 se que a amplitude de um mesmo testador nos diferentes ensaios não exibiu uma oscilação de
28 amplitude. Porém os testadores responderam diferenciadamente nos ensaios. Esta observação
29 corrobora com as discussões anteriores que a diferenciação entre as médias dos ensaios foi
30 decorrente dos híbridos de cada população e não dos efeitos não controlados sofridos, pois as
31 influências causadas por fatores não controlados resultam em respostas com amplitudes de
32 variações elevadas. O trabalho de Oliveira et al. (2009), verificou na literatura que diversos

1 trabalhos apresentam elevados valores de coeficiente de variação, o que pode-se remeter que
2 o efeito não controlados foram bem responsivos na respostas observadas.

3 Os maiores rendimentos de algumas testemunhas em um ensaio e de outras
4 testemunhas em outro ensaio, nos remete a existência de possível interação genótipo ambiente
5 e que ambos os ensaios apresentaram as condições adequadas para o desenvolvimento dos
6 materiais genéticos. Pois, quando o ambiente apresenta condições propícias para a expressão
7 do potencial genético existem genótipos que iram se sobressair em algumas condições e em
8 outras condições outros genótipos sobressairão, considerando-se os mesmos materiais (Cruz
9 et al., 2014).

10 Na análise conjunta, o testador 8 foi o responsivo contribuindo para o aumento da
11 média geral pelo maior vigor híbrido (Charlesworth and Willis, 2009 and Khaleda et al.,
12 2013) (Figura 5). O vigor híbrido é a informação resultante da hibridação entre dois genótipos
13 e quanto maior o valor do vigor híbrido quanto maior a resposta do ganho na característica em
14 estudo (Khaleda et al., 2013).

15 A acurácia do BLUPs dos testadores nos dois modelos podem ser classificadas como
16 alta segundo (Resende et al., 2009). Os diferentes valores estimados de BLUPs observados
17 entre os modelos é devido a diferenciação entre os efeitos considerados como fixos e
18 aleatórios no modelo, além disso, não é preocupante a diferença existente, pois é baixa e para
19 ambos os modelos a acurácia foi considerada alta.

20 Deve ser ratificado que o vigor híbrido elevado, resulta em híbridos potenciais, porém
21 caso a seleção seja ineficiente, resultará na seleção de híbridos falsos positivos, estes falsos
22 positivos são os híbridos que apresentam uma boa produtividade fenotípica mas
23 genotipicamente não são promissores, ou seja, a seleção é decorrente do efeito da interação
24 genótipos *versus* ambiente (Ewert et al., 2011, Hatfield and Walthall, 2015, Torres et al.,
25 2015, Messina et al., 2015 and Teixeira et al., 2016).

26 Resultados referentes aos falsos positivos podem ser observados por coeficientes de
27 correlação baixos entre a seleção via genotípica e fenotípica na Figura 6. Estes resultados
28 reforçam a necessidade da adoção dos modelos mistos para a análise estatísticas, junto com a
29 adesão das informações de pedigree para minimizar os efeitos não controlados decorrentes de
30 inúmeros fatores. Como foram adotados em inúmeros outros trabalhos de seleção
31 (Grattapaglia and Resende, 2011, Burgueño et al., 2012, Rincen et al., 2012, Bassia et al.,
32 2016, Isik, 2014, Pereira et al., 2016 and Bassia et al., 2016).

1 Com uma intensidade de seleção de 5%, os resultados observados entre os mesmos
2 ensaios e os diferentes modelos chegaram a uma coincidência é acima de 96% (Tabela 2).
3 Decorrente da elevada similaridade entre os modelos testados reforça que a seleção está
4 ocorrendo em sentido correto.

5 A resposta entre as coincidências dos diferentes ensaios dentro de um mesmo modelo
6 resultou em que 100% ou próximo disto foi decorrente a coincidência do E2. Mediante está
7 observação e as demais citadas á cima é possível verificar que o E2 responde pelos melhores
8 híbridos potenciais.

9 Os melhores híbridos podem ser decorrentes de diversos fatores, sendo que entre eles a
10 presença de um bom testador, acarretando em um maior vigor híbrido ou heterose entre as
11 linhagens e o testador (Bertoiaa and Aulicinob, 2014). Além disso, outro ponto importante
12 observado é que o conhecimento do banco de dados é de suma importância para que a seleção
13 não seja comprometida sem o conhecimento do pesquisado.

14 Os valores dos rendimentos de grãos dos selecionados no E2 superiores aos do E1 e
15 conjunto, corroboram com os resultados apresentados durante todo o trabalho, referindo-se
16 que o E2 resultou nos melhores híbridos e que se fossem feitos análise de seleção conjunta,
17 maior parte dos selecionados ou todos seriam do respectivo ensaio. Além disso, estes
18 resultados podem ser observados nos valores fenotípicos (rendimento de grãos) como nos
19 genotípicos (valor genético) (Tabela 3).

20 O E2 apresenta diversos híbridos potenciais, pois os valores fenotípicos e genotípicos
21 de rendimentos de grãos dos 5% selecionados foram superiores que as testemunhas para
22 ambos os modelos em estudo, sendo que são materiais genéticos elites do mercado atual.
23 Porém no E1, os selecionados foram inferiores os resultados observados que os testadores.

24 Na avaliação de dados conjunto foi observado a superioridade dos 5% selecionados
25 em cima das testemunhas, este resultado corrobora com a Tabela 2, que demonstra a
26 coincidência entre os selecionados se assemelha próximo de 100% ao do E2 e 0% ao E1.

1 5. CONCLUSÃO

2 As estratégias de seleção adotada pelos melhoristas sempre devem ser mediadas de
3 diversas cautelas, pois mesmo como no presente trabalho que ambos os testadores foram
4 responsivos em produzir híbridos de primeiro ciclo com alto vigor híbrido dentro de cada uma
5 das populações dos diferentes ensaios (testadores). Na análise conjunta, pode ser observado
6 que o testador do ensaio 2 respondeu a uma população com maior vigor híbrido (maiores
7 produtividades), o que resultou em um número de híbridos suficiente para que os 5% dos
8 híbridos selecionados da análise conjunta todos seriam do ensaio 2 e nenhum do ensaio 1.

9 No ensaio 2 os valores genotípicos e fenotípicos de produtividade dos 5%
10 selecionados foram superiores a das testemunhas adotadas para comparação e no ensaio 1 as
11 testemunhas que responderam a uma maior produtividade que os 5% selecionados. Os
12 resultados de ambos os ensaios foram iguais para os dois modelos adotados.

6. REFERENCIA

Araus, J.L., Cairns J.E. Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*, 19(1) (2014), pp. 52-61. doi:10.1016/j.tplants.2013.09.008

Bajželj B., Richards K.S., Allwood J.M., Smith P., Dennis J.S., Curmi E., Gilligan C.A. Importance of food-demand management for climate mitigation. *Nature Climate Change*, 4(10) (2014), pp. 924-929. doi:10.1038/nclimate2353

Barreto R. R., Scapim, C. A., Amaral Júnior, A. T. do, de Araújo Rodovalho, M., Vieira, R. A., & Schuelter, A. R. Avaliação da capacidade de combinação de famílias S2 de milho-pipoca por meio de diferentes testadores. *Semina*, 33(3) (2012), 873-890. doi:http://dx.doi.org/10.5433/1679-0359.2012v33n3p873

Bassi F.M., Bentley A.R., Charmet G., Ortiz R., Crossa J. Breeding schemes for the implementation of genomic selection in wheat (*Triticum* spp.). *Plant Science*, 242 (2016), pp. 23-36. doi:10.1016/j.plantsci.2015.08.021

Bertoia L.M., Aulicino, M.B. Maize forage aptitude: Combining ability of inbred lines and stability of hybrids. *The Crop Journal*, 2(6) (2014), pp. 407-418. doi:10.1016/j.cj.2014.07.002

Borges V., Ferreira P.V., Soares L., Santos G.M., Santos A.M.M. Sweet potato clone selection by REML/BLUP procedure. *Acta Scientiarum. Agronomy*, 32(4) (2010), pp. 643-649. doi.org/10.4025/actasciagron.v32i4.4837

Burgueño J., Campos G. de los, Weigel K., Crossa J. Genomic prediction of breeding values when modeling genotype× environment interaction using pedigree and dense molecular markers. *Crop Science*, 52(2) (2012), pp. 707-719. 2012.doi:10.2135/cropsci2011.06.0299

Casale F.P., Rakitsch B., Lippert C., Stegle O. Efficient set tests for the genetic analysis of correlated traits. *Nature methods*, 12(8) (2015), pp. 755-758. doi:10.1038/nmeth.3439

Charlesworth D., Willis J.H. The genetics of inbreeding depression. *Nature Reviews Genetics*, 10(11) (2009), pp. 783-796. doi:10.1038/nrg2664

Clovis L.R., Scapim C.A., Pinto R.J.B., Bolson E., Senhorinho, H.J.C. AVALIAÇÃO DE LINHAGENS S3 DE MILHO POR MEIO DE TESTADORES ADAPTADOS À SAFRINHA. *Revista Caatinga*, 28(1) (2015), pp. 109-120.

Cobb J.N., DeClerck G., Greenberg A., Clark R., McCouch, S. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–

phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*, 126(4) (2013), pp. 867-887. DOI: 10.1007/s00122-013-2066-0

Cruz C.D., Carneiro, P.C.S., Regazzi A.J. Modelos biométricos aplicados ao melhoramento genético (3. ed.). Viçosa: Editora UFV. (2014), pp. 668.

Davis K.F., Gephart J.A., Emery K.A., Leach A.M., Galloway J.N., D’Odorico, P. Meeting future food demand with current agricultural resources. *Global Environmental Change*, 39 (2016), pp. 125-132. doi:10.1016/j.gloenvcha.2016.05.004

DAVIS, R. L. Report of the plant breeder. Puerto Rico: Agricultural Experiments Annual Report, (1927), pp. 13-22.

Dempster A.P., Laird N.M. Rubin D.B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society*, 39(1) (1977), pp. 1-38.

Dempster A.P., Laird, N.M., Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society*, 39(1) (1977), pp. 1-38.

Duvick D.N., Smith J.S.C., Cooper M. Long term selection in a commercial hybrid maize breeding program. Janick. I. *Plant Breeding Reviews*. Part, 24(2) (2010), pp. 109-152. DOI: 10.1002/9780470650288.ch4

EAST, E. M. Inbreeding in corn. Rep. Conn. Agric. Exp. Stn. (1908), pp. 419–428.

Edwards, J. Testcross Response to Four Cycles of Half-sib and S Recurrent Selection in the BS13 Maize (L.) Population. *Crop science*, 50(5) (2010), pp. 1840-1847. doi:10.2135/cropsci2009.09.0557

Elias H.T., Carvalho S.P., André C.G.M. Comparação de testadores na avaliação de famílias S2 de milho. *Pesquisa Agropecuária Brasileira*, 35(6) (2000), pp. 1135-1142.

Ewert F., Van Ittersum M.K., Heckeleei T., Therond O., Bezlepkina I., Andersen E. Scale changes and model linking methods for integrated assessment of agri-environmental systems. *Agriculture, Ecosystems & Environment* 142(1) (2011), pp. 6–17 <http://dx.doi.org/10.1016/j.agee.2011.05.016>

Foley J.A., Ramankutty N., Brauman K.A., Cassidy E.S., Gerber J.S., Johnston M., Mueller N.D., O’connell C., Ray D.K., West P.C., Balzer C., Bennett E.M., Carpenter S.R., Hill J., Monfreda C., Polasky S., Rockstrom J., Sheehan J., Siebert S., Tilman D., Zaks D.P.M. Solutions for a cultivated planet. *Solutions for a cultivated planet. Nature*, 478(7369) (2011), pp. 337-342. doi:10.1038/nature10452

Godfray H.C.J., Beddington J.R., Crute I.R., Haddad L., Lawrence D., Muir J. F., PRETTY J., ROBINSON S., THOMAS S.M., Toulmin, C. Food security: the challenge of

feeding 9 billion people. *Science*, 327(5967) (2010), pp. 812-818. doi:10.1126/science.1185383

Grattapaglia D., Resende, M.D. Genomic selection in forest tree breeding. *Tree Genetics & Genomes*, 7(2) (2011), pp. 241-255. doi:10.1007/s11295-010-0328-4

Grieder C., Dhillon B.S., Schipprack W., Melchinger A.E. Breeding maize as biogas substrate in Central Europe: I. Quantitative-genetic parameters for testcross performance. *Theoretical and Applied Genetics*, 124(6) (2012), pp. 971-980. doi:10.1007/s00122-011-1761-y

Guimarães L.J.M., Miranda G.V., DeLima R.O., Maia C., Oliveira L.R.D., Souza L.V.D. Performance of testers with different genetic structure for evaluation of maize inbred lines. *Ciência Rural*, 42(5) (2012), pp. 770-776. <http://dx.doi.org/10.1590/S0103-84782012000500002>

Hallauer A.R. Relation of gene action and type of tester in maize breeding procedures. In *Proc Annu Corn Sor Res Conf*. Chicago: ASTA (1975). pp. 150-165.

Hallauer, A.R.; Carena, M.J.; Miranda Filho, J. D. *Quantitative genetics in maize breeding*. Iowa: Springer (2010).

Hatfield J.L., Walthall C.L. Meeting global food needs: realizing the potential via genetics \times environment \times management interactions. *Agronomy Journal*, 107(4) (2015), pp. 1215-1226. <http://dx.doi.org/10.2134/agronj15.0076>

Henderson C.R. Best linear unbiased estimation and prediction under a selection model. *Biometrics* (1975), pp. 423-447.

Isik F. Genomic selection in forest tree breeding: the concept and an outlook to the future. *New forests*, 45(3) (2014), pp. 379-401. doi:10.1007/s11056-014-9422-z

Khaled A.G.A., Hamam K.A., Motawea M.H., El-Sherbeny G.A.R. Genetic studies on tissue culture response and some agronomical traits in Egyptian bread wheat. *Journal of Genetic Engineering and Biotechnology*, 11(2) (2013), 79-86. doi:10.1016/j.jgeb.2013.11.002

Laurance W.F., Sayer J., Cassman K.G. Agricultural expansion and its impacts on tropical nature. *Trends in ecology & evolution*, 29(2) (2014), pp. 107-116. doi:10.1016/j.tree.2013.12.001

Li M.S., Li X.H., Deng L.W., Zhang D.G., Bai L., Zhang S.H. Comparisons of four testers in evaluating 27 CIMMYT and Chinese maize populations. *Maydica*, 52(2) (2007), pp. 173-179.

Li Z.Y., Zhang T.F., Wang S.C. Transcriptomic analysis of the highly heterotic maize hybrid zhengdan 958 and its parents during spikelet and floscule differentiation. *Journal of Integrative Agriculture*, 11(11) (2012), pp. 1783-1793. doi:10.1016/S2095-3119(12)60183-X

Loh, P. R., Bhatia, G., Gusev, A., Finucane, H. K., Bulik-Sullivan, B. K., Pollack, S. J., Consortium S.W.G.P.G., Candia T.R., Lee S.H., Wray N.R., Kendler K.S., O'Donovan M.C, Neale B.M., Patterson N., Price A.L. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nature genetics*. 47(12) (2015), pp. 1385-1392. doi:10.1038/ng.3431

Long S.P., Ort D.R. More than taking the heat: crops and global change. *Current opinion in plant biology*, 13(3) (2010), pp. 240-247. doi:10.1016/J.PBI.2010.04.008

Messina C.D., Sinclair T.R., Hammer G.L., Curan D., Thompson J., Oler Z., Ghod C. and Cooper M. Limited-Transpiration Trait May Increase Maize Drought Tolerance in the US Corn Belt. *Agronomy Journal*, 107(6) (2015), pp. 1978-1986. doi:10.2134/agronj15.0016

Mohammadi R., Farshadfar, E., Amri, A. Interpreting genotype× environment interactions for grain yield of rainfed durum wheat in Iran. *The Crop Journal*, 3(6) (2015), pp. 526-535. doi:10.1016/j.cj.2015.08.003

Nerling D., Coelho C.M.M., Nodari R.O. Genetic diversity for physiological quality of seeds from corn (*Zea mays* L.) intervarietal crossbreeds. *Journal of Seed Science*, 35(4) (2013), pp. 449-456. doi.org/10.1590/S2317-15372013000400006

Oliveira, R. L., Muniz, J. A., de Andrade, M. J. B., & dos Reis, R. L. (2009). Precisão experimental em ensaios com a cultura do feijão. *Ciência e Agrotecnologia*, 33(1) (2009), pp. 113-119.

Ray D.K., Mueller N.D., West P.C., Foley, J.A. Yield trends are insufficient to double global crop production by 2050. *PloS one*, 8(6) (2013), pp. 1-8. <http://dx.doi.org/10.1371/journal.pone.0066428>

Reid J.M., Arcese P., Sardell R.J., Keller, L.F. Additive genetic variance, heritability, and inbreeding depression in male extra-pair reproductive success. *The American Naturalist*, 177(2) (2011), pp. 177-187. doi:10.1086/657977

Resende R.M.S, Resende M.D.V, Jank L., (2008). Experimentação e análise estatística no melhoramento de forrageiras. In: Resende R.M.S., Valle C.B do, Jank L. (Eds.) *Melhoramento de Forrageiras Tropicais*. Campo Grande, Embrapa Gado de Corte. p.195-284.

Rincent R., Laloë D., Nicolas S., Altmann T., Brunel D., Revilla P., Rodríguez V.M., Moreno-Gonzalez J., Melchinger A., Bauer E., Schoen C.C., Meyer N., Giauffret C., Bauland

C., Jamin P., Laborde J., Monod H., Flament P., Charcosset A., Moreau L. Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). *Genetics*, 192(2) (2012), pp. 715-728. doi: 10.1534/genetics.112.141473

Rodvalho M. de A., Scapim C.A., Pinto R.J.B., Barreto R.R., Ferreira F.R.A., Clóvis L.R. Comparação de testadores em famílias S2 obtidas do híbrido simples de milho-pipoca IAC-112= Comparison of testers in S2 families obtained from the popcorn single hybrid 'IAC-112'. *Bioscience Journal*, 28(2) (2012), pp. 145-154.

Rodrigues W.P., Vieira H.D., Barbosa D.H., Souza Filho G.R., Candido L.S. Adaptability and genotypic stability of *Coffea arabica* genotypes based on REML/BLUP analysis in Rio de Janeiro State, Brazil. *Genetics and Molecular Research*, 12(3) (2013), pp. 2391-2399.

SCAPIM C.A., ROYER M.R., PINTO R.J.B., JÚNIOR A.T., PACHECO C.A.P., MOTERLE L. Comparação de testadores na avaliação da capacidade de combinação de famílias S2 de milho-pipoca. *Revista Brasileira de Milho e Sorgo*, 7(1) (2010). pp. 83-91 doi:<http://dx.doi.org/10.18512/1980-6477/rbms.v7n1p83-91>

Shapira R., David L. Genes with a Combination of Over-Dominant and Epistatic Effects Underlie Heterosis in Growth of *Saccharomyces cerevisiae* at High Temperature. *Frontiers in genetics*, 7 (2016), pp. 1-13. doi.org/10.3389/fgene.2016.00072

Shull, G. H. The composition of a field of maize. *Journal of Heredity*, (1) (1908), pp. 296-301. doi: 10.1093/jhered/os-4.1.296

Shull, G. H. The genotypes of maize. *The American Naturalist*, 45(532) (1911), pp. 234-252. doi: 10.1086/279207

Sreckov Z.O.R.A.N.A., Nastasic A., Bocanski J., Djalovic I., Vukosavljev M., Jockovic B. Correlation and path analysis of grain yield and morphological traits in test-cross populations of maize. *Pak. J. Bot*, 43(3) (2011), pp. 1729-1731.

Strzepek K., Boehlert B. Competition for water for the food system. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1554) (2010), pp. 2927-2940. doi: 10.1098/rstb.2010.0152.

Teixeira F.F., Paes M.C.D., Pereira Filho I.A., Miranda R.A., Parentoni S.N., Cotta L.V., Meirelles W.F., Pacheco C.A.P., Guimaraes L.J.M., Silva A.R., Machado J.R.A. BRS Vivi: single-cross super sweet corn hybrid. *Crop Breeding and Applied Biotechnology*, 14(2) (2014), pp. 124-127. <http://dx.doi.org/10.1590/1984-70332014v14n2c21>

Teixeira, E. I., Zhao, G., de Ruiter, J., Brown, H., Ausseil, A. G., Meenken, E., & Ewert, F. (2016). The interactions between genotype, management and environment in regional crop modelling. *European Journal of Agronomy*, (2016).

Tester M., Langridge P. Breeding technologies to increase crop production in a changing world. *Science*, 327(5967) (2010), pp. 818-822. doi:10.1126/science.1183700

Tilman D., Balzer C., Hill J., Befort B.L. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108(50) (2011), pp. 20260-20264. doi:10.1073/pnas.1116437108

Torres F.E., Teodoro P.E., Sagrilo E., Ceccon G., Correa, A.M. Interação genótipo x ambiente em genótipos de feijão-caupi semiprostrado via modelos mistos. *Bragantia*, 74(3) (2015), pp. 255-260. <http://dx.doi.org/10.1590/1678-4499.0099>

Viana J.M.S., Almeida R.V., Faria V.R., Resende M.D.V., Silva F.F. Genetic evaluation of inbred plants based on BLUP of breeding value and general combining ability. *Crop and Pasture Science*, 62(6) (2011), pp. 515-522. doi:10.1071/CP11016 1836-0947/11/060515

Yang R. C. Mixed-model analysis of crossover genotype–environment interactions. *Crop science*, 47(3) (2007), pp. 1051-1062. doi:10.2135/cropsci2006.09.0611

Wilks S.S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9(1) (1938), pp. 60-62. URL: <http://www.jstor.org/stable/29576>