**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

Predição genômica de híbridos simples de milho

Marcela Pedroso Mendes

Tese apresentada para obtenção do título de
Doutora em Ciências. Área de concentração:
Genética e Melhoramento de Plantas

**Piracicaba
2015**

Marcela Pedroso Mendes
Engenheira Agrônoma

Predição genômica de híbridos simples de milho

Orientador:
Prof. Dr. **CLÁUDIO LOPES DE SOUZA JÚNIOR**

Tese apresentada para obtenção do título de Doutora
em Ciências. Área de concentração: Genética e
Melhoramento de Plantas

**Piracicaba
2015**

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Mendes, Marcela Pedroso

Predição genômica de híbridos simples de milho / Marcela Pedroso Mendes. - -
Piracicaba, 2015.
105 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2015.

1. Predição genômica 2. Milho 3. Híbrido 4. Modelos mistos 5. BLUP 6. Parentesco
7. Seleção genômica ampla 8. Acurácia I. Título

CDD 633.15
M538p

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

Ao meu pai, Nazareno, com amor
Dedico

AGRADECIMENTOS

A realização desta tese de doutorado é fruto de um contexto acadêmico e pessoal altamente favorável. Agradeço primeiramente ao meu orientador, Prof. Cláudio Lopes de Souza Júnior, a quem devo o amadurecimento profissional que obtive ao longo desse doutorado. Particularmente, sou-lhe grata pela solicitude e confiança do trabalho que originou esta tese, significativos para a minha formação.

Ao Prof. Isaias Olívio Geraldi, em nome do qual agradeço aos demais professores do Departamento de Genética e de Estatística da ESALQ-USP pelos ensinamentos e contribuições inestimáveis à minha formação. Sou igualmente grata, de antemão, aos membros da banca examinadora de defesa, que prontamente aceitaram o nosso convite e abriram espaço nas suas agendas para fazerem parte desta importante realização da minha trajetória acadêmica.

No Programa de Pós Graduação em Genética em Melhoramento de Plantas da ESALQ-USP encontrei o melhor ambiente possível para o meu doutoramento. O convívio harmonioso e descontraído dos colegas do Laboratório de Milho transformaram a rotina diária em períodos agradáveis e produtivos. Agradeço carinhosamente à Luciana, Melina, Sueme, Thiaguinho, Paolo, Otávio, Pedro, Ruben, Gustavo, Diego e Alexandre, e aos funcionários Jucelino, Guilherme, Sr. Ary e Têr, Sr. Natálio e D. Maria, Berdan, Valdir, Fernandinho, Léia e demais funcionários do Departamento pelo apoio à realização desta pesquisa e convívio amistoso.

Tenho a felicidade de dizer que em Piracicaba fiz incontáveis amigos e com certeza pecaria se tentasse nomear todos aqui. Agradeço em particular à Izabela, Matheus, Lucas, Fernando, Gui e Augusto, essenciais nos anos iniciais do doutorado. Não posso deixar de agradecer à Marcia Andrade, Renée Danckwardt e Sueli Mantelatto, meus anjos em Piracicaba, sempre muito carinhosas e solícitas em todos os momentos. Sem essas amigadas, a distância familiar e o desafio dessa trajetória particularmente demandante teriam sido impossíveis de suportar.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) pela concessão das bolsas de estudo no Brasil e nos Estados Unidos. Agradeço, também, à Universidade de Wisconsin, *UW-Madison*, pelo valoroso período de estágio sanduíche, e aos amigos Camila, Mariana, Ludmila, Renato, Julia, Luana, Tassy, Vivi e Renato Lima, Luciana e Gerson Mourão,

Natalia, Joe, Scott, Tim, Megan, Stephanie, Mary e Jenna, pelo carinho, amizade e acolhida em Madison.

Em minha vida pessoal sempre fui presenteada com pessoas queridas e especiais. Agradeço à minha mãe, Adriana, pelo amor incondicional e por permitir que eu desenvolvesse todo meu potencial. Ao meu pai, Nazareno, à quem este trabalho é dedicado, agradeço a proteção, o exemplo de caráter e por ter dado os primeiros passos para que hoje eu pudesse conquistar mais essa etapa. Ao meu irmão, Flávio, sou grata pela amizade, o exemplo de competência e as conversas nerds e descontraídas. Também agradeço à Samantha e à Suzem pela torcida e amizade, e aos meus avós Hilda, Messias, Dirce e Fausto, pelo exemplo de simplicidade e trabalho. Agradeço especialmente ao Douglas, que faz parte da minha vida desde o início dessa etapa e com quem eu sou uma pessoa muito melhor. Obrigada pelo apoio, incentivo e confiança, pela simplicidade do nosso amor e pela amizade honesta. Também sou grata aos meus sogros, Bel e Zé, por tanto carinho e apoio sempre.

Finalmente, após percorrer mais esta etapa, posso compreender a importância de expressar a minha gratidão à Deus, por ter colocado cada um de vocês no meu caminho e por ter proporcionado todas as condições para que eu chegasse até aqui. Muito obrigada!

Pão ou pães é uma questão de 'opiniões'

João Guimarães Rosa, Grande Sertão: Veredas

SUMÁRIO

RESUMO.....	11
ABSTRACT	13
1 INTRODUÇÃO.....	15
Referências	19
2 PREDIÇÃO GENÔMICA DE HÍBRIDOS DE MILHO UTILIZANDO INFORMAÇÕES SIMILARIDADE NO ESTADO E PARENTESCO	23
Resumo	23
Abstract.....	24
2.1 Introdução.....	25
2.2 Material e Métodos	26
2.2.1 Dados fenotípicos e moleculares dos híbridos	26
2.2.2 Análises de variância.....	27
2.2.3 Modelo misto de seleção genômica dos híbridos	30
2.2.4 Obtenção das matrizes de parentesco e similaridade no estado	32
2.3 Resultados e discussão.....	37
2.3.1 Análises de variância.....	37
2.3.2 Predição genômica dos híbridos para os diferentes coeficientes de similaridade no estado e parentesco	37
2.3.3 Seleção fenotípica e genômica dos híbridos superiores	42
2.3.4 Considerações finais	44
Referências	45
3 PREDIÇÃO GENÔMICA DE HÍBRIDOS DE MILHO CONSIDERANDO A INTERAÇÃO GENÓTIPOS X AMBIENTES	49
Resumo	49
Abstract.....	50
3.1 Introdução.....	51
3.2 Material e Métodos	53
3.2.1 Dados fenotípicos e moleculares dos híbridos	53
3.2.2 Análises de variância.....	54
3.2.3 Modelo RR-BLUP de seleção genômica ampla para predição dos híbridos de milho.....	57
3.2.4 Validação cruzada do modelo de seleção genômica: efeito do número de marcas e do tamanho da população	58

3.2.5 Validação cruzada entre e dentro de ambientes: efeito da interação H x A	59
3.2.6 Validação cruzada entre e dentro de grupos de híbridos relacionados: efeito da estrutura da população.....	60
3.2.7 Seleção fenotípica e genômica de híbridos superiores	60
3.3 Resultados e Discussão.....	62
3.3.1 Análises de variâncias	62
3.3.2 Validação cruzada do modelo de seleção genômica: efeito do número de marcas e do tamanho da população	63
3.3.3 Validação cruzada entre e dentro de ambientes: efeito da interação H x A	65
3.3.4 Validação cruzada entre e dentro de grupos de híbridos relacionados: efeito da estrutura da população.....	72
3.3.5 Seleção fenotípica e genômica de híbridos superiores	74
3.3.6 Aplicações da seleção genômica nos programas de melhoramento	76
Referências	78
ANEXOS	81

RESUMO

Predição genômica de híbridos simples de milho

Métodos de predição podem aumentar consideravelmente a eficiência dos programas de melhoramento de milho. O objetivo deste estudo foi prever a performance de 250 híbridos simples de milho avaliados em múltiplos ambientes utilizando a informação de marcadores moleculares. Para isso, 50 linhagens endogâmicas provenientes de diferentes populações foram cruzadas com cinco linhagens elite, também endogâmicas, para obtenção dos 250 híbridos simples. As matrizes moleculares das linhagens e dos híbridos foram obtidas a partir da genotipagem das 55 linhagens com 614 marcadores AFLP. Os híbridos simples foram avaliados para produção de grãos em 13 ambientes. A predição dos híbridos foi realizada utilizando o modelo misto BLUP considerando diferentes coeficientes de parentesco e similaridade no estado na predição dos efeitos das capacidades geral e específica de combinação dos genitores. As médias preditas dos híbridos a partir de cada coeficiente foram correlacionadas com as médias fenotípicas para obtenção da acurácia de predição. A predição também foi realizada utilizando o modelo de seleção genômica ampla RR-BLUP. Nesse caso, a matriz molecular dos híbridos foi utilizada diretamente no modelo misto de estimação dos efeitos dos marcadores e da contribuição de cada um deles para o valor genético dos híbridos. Foram realizadas validações cruzadas entre e dentro de ambientes e entre e dentro de grupos de híbridos relacionados a fim de verificar os efeitos do tamanho da população de treinamento (N), número de marcas (N_M), interação híbridos $\times$ ambientes ($H \times A$) e da estrutura da população na estimativa da acurácia de predição. A predição genômica foi comparada com a seleção fenotípica quanto à eficiência em identificar híbridos superiores em um esquema de melhoramento de milho. Todos os coeficientes de parentesco e similaridade no estado apresentaram elevadas estimativas de acurácia, contudo foi possível observar considerável superioridade dos coeficientes Wang e Rogers Modificado tanto na predição quanto na seleção dos híbridos superiores, demonstrando o potencial dessas metodologias como ferramentas a serem utilizadas nos programas de melhoramento de milho. Os resultados da predição utilizando o modelo de seleção genômica ampla indicaram que o aumento de N e N_M não alteraram significativamente as estimativas de acurácia. As estimativas da acurácia na validação cruzada dentro de ambientes foram superiores às obtidas entre ambientes, inferindo que o efeito da interação $H \times A$ foi expressivo. Também foram observadas estimativas de acurácia expressivamente maiores para populações de treinamento e validação compostas por híbridos relacionados. Em todos os casos, as estimativas de acurácia apresentaram amplos intervalos em função da amostra de híbridos utilizada nas populações de treinamento e validação, indicando que a seleção genômica pode não ser eficiente dependendo da população amostrada. Os resultados deste estudo sugerem que a predição genômica é uma ferramenta para aumentar a eficiência da seleção nos programas de melhoramento se utilizada de forma adequada pelo melhorista, considerando os efeitos de estrutura de população e interação $H \times A$ de forma a maximizar a acurácia e, consequentemente, o sucesso da predição.

Palavras-chave: Predição genômica; Milho; Híbrido; Modelos mistos; BLUP; Parentesco; Seleção genômica ampla; Acurácia

ABSTRACT

Genomic prediction of maize single-crosses

Prediction using molecular markers information can greatly increase the efficiency of maize breeding programs. This study aimed to predict the performance of maize single-crosses evaluated in multiple environments and using molecular markers information. Five inbred lines used as testers were crossed to 50 inbred lines from multiple populations to obtain 250 maize single-crosses. 614 AFLP markers were used to assess molecular matrices of the inbred lines and single-crosses. The 250 single-crosses were evaluated for grain yield in 13 environments. Genomic prediction was performed using the mixed model BLUP considering different genomic relationship and similarity in state coefficients to predict the effect of general and specific combining abilities of the parents. Predicted means from each coefficient were correlated with phenotypic means for obtaining prediction accuracy. Genomewide prediction was also performed using the linear regression model RR-BLUP in the estimation of markers genotypic values and its contribution to hybrids genetic values. Cross-validations between and within environments and between and within groups of related single-crosses were performed to verify the effects of training population size (N), number of markers (N_M), genotype-by-environment interaction ($G \times E$) and population structure in estimating accuracy. Genomic prediction was compared with phenotypic selection in efficiency of selecting better hybrids in a maize breeding program. All relationship coefficients and similarity in state coefficients showed high values of accuracy, however we observed superiority of Wang relationship coefficient and Modified Rogers similarity coefficient both in predicting and in identifying the best single-crosses, showing the potential of these approaches as tools to be used in maize breeding programs. Genomewide prediction results showed that increasing N and N_M did not lead to higher accuracy estimates. Predicted accuracies of cross validation analysis within environments were higher than between environments, indicating that the effect of $G \times E$ interaction was significant. Greater accuracies were achieved when training and validation set were from related single-crosses. In all scenarios, wide intervals of accuracy were found, meaning that genomic prediction may not be effective depending on the sample used. Therefore, the results of this study suggest that genomic prediction is a tool to increase the efficiency of selection in breeding programs if used properly by breeders, considering the population structure and $G \times E$ interaction effect so as to reduce sample problems and maximize accuracy and hence the success of prediction.

Keywords: Genomic prediction; Maize; Hybrid; Mixed models; BLUP; Relationship; Genomewide selection; Accuracy

1 INTRODUÇÃO

A obtenção e avaliação de híbridos são as atividades que mais demandam tempo e recursos nos programas de melhoramento de milho. Devido ao pronunciado efeito de dominância presente no controle genético da produção de grãos, a performance dos híbridos não pode ser predita apenas com base no desempenho *per se* das linhagens, sendo necessária avaliação dessas em cruzamentos (HALLAUER; CARENA; MIRANDA FILHO, 2010). Contudo, a partir de poucas linhagens é possível obter um número muito grande de híbridos, limitando o número de híbridos a serem avaliados e de ambientes de avaliação. Dessa forma, métodos de predição com base na informação de marcadores moleculares podem aumentar consideravelmente a eficiência dos programas de melhoramento.

Inicialmente a predição de híbridos utilizando a informação de marcadores moleculares foi realizada através da estimação da distância genética entre as linhagens genitoras. Segundo Falconer e Mackey (1996), existe relação linear entre a distância genética entre as linhagens genitoras, na presença de locos com dominância, e a heterose do híbrido entre estas. Contudo, os estudos realizados nesse sentido apresentaram pouca eficiência e repetibilidade (MELCHINGER, 1999; DIAS et al., 2004). Além disso, a escolha do coeficiente de distância genética não é simples e pode influenciar na magnitude da correlação entre as estimativas da distância genética das linhagens e a heterose do híbrido proveniente do cruzamento entre elas (REIF et al., 2005; BALESTRE et al., 2008).

A metodologia dos modelos mistos proposta por Henderson (1975) consiste na predição linear não viesada (*Best Linear Unbiased Prediction* - BLUP) dos valores genéticos dos híbridos, corrigida para os efeitos fixos do modelo. Uma das principais propriedades dessa metodologia é a possibilidade de prever o valor genético de híbridos não avaliados, sendo que a correlação entre os valores observados e preditos é máxima entre as classes de preditores lineares. White e Hodge (1989) apresentam outras propriedades do BLUP úteis ao melhoramento, como a predição dos efeitos aleatórios e estimação dos efeitos fixos do modelo utilizando um procedimento único, a possibilidade de considerar os efeitos de seleção e endogamia ao longo das gerações, desde que o grau de parentesco entre os híbridos seja conhecido, além de fornecerem estimativas de variâncias residuais inferiores às de outros métodos.

Uma das características do BLUP que o faz um método de predição superior aos demais é a possibilidade de considerar a informação do parentesco genético entre os híbridos, retornando predições mais precisas. Segundo Bernardo (1995), a habilidade do BLUP em prever a performance de híbridos de milho não avaliados, bem como o número de preditores (híbridos avaliados) necessários para a obtenção de predições confiáveis, depende do grau de parentesco entre os híbridos e entre as linhagens genitoras. Como a informação de genealogia nem sempre é conhecida no melhoramento vegetal, o advento dos marcadores moleculares possibilitou a estimação do parentesco entre os híbridos e, conseqüentemente, ampliou a aplicação da metodologia dos modelos mistos nos estudos de predição.

O conceito de parentesco foi estabelecido pelo trabalho clássico de Wright (1922) que considera a probabilidade de dois híbridos compartilharem cópias do mesmo alelo para um dado loco oriundo de um ancestral comum. O cálculo do parentesco deve ter como referência uma população base, ou seja, a probabilidade de dois alelos serem descendentes do mesmo ancestral dessa população. Tradicionalmente, a probabilidade de dois alelos serem idênticos por descendência é calculada com base em um pedigree conhecido, sendo os genótipos do topo desse pedigree considerados os fundadores da população base. A obtenção dos coeficientes de parentesco com base na informação molecular obviamente não considera nenhuma população base, o que pode levar a vários problemas práticos e a obtenção de estimativas viesadas de parentesco (POWELL; VISSCHER; GODDARD, 2010).

Existem inúmeros coeficientes de parentesco e de similaridade no estado que podem ser utilizados nas equações de modelos mistos. Os coeficientes de similaridade são medidas que quantificam o número de marcas que dois híbridos compartilham, ou seja, consideram a probabilidade dos alelos desses híbridos serem idênticos por estado. Nesse caso, os alelos possuem o mesmo segmento de DNA e fornecem o mesmo fenótipo, mas não são originados do mesmo ancestral, sendo, portanto, idênticos por estado, mas não por descendência. No melhoramento de plantas, esses coeficientes têm sido utilizados para estimar a distância genética entre os genótipos (REIF et al., 2005). Já os coeficientes de parentesco utilizam a informação molecular para estimar a probabilidade dos alelos serem idênticos por descendência e, por isso, geralmente são preferidos nos estudos de predição.

Comparados aos coeficientes de parentesco, os coeficientes de similaridade possuem a vantagem de utilizar toda semelhança alélica existente entre os híbridos (NEJATI-JAVEREMI;

SMITH; GIBSON, 1997). Contudo, podem considerar parentesco entre os híbridos quando, na verdade, estes não apresentam nenhuma relação genealógica. Com o intuito de contornar essa limitação, Bernardo (1993) propôs o uso de um coeficiente de coancestria que utiliza a informação de similaridade genética para corrigir possíveis vieses da similaridade no estado. Entretanto, essa metodologia somente é eficiente quando se tem informação prévia da genealogia das linhagens e do grupo heterótico ao qual pertencem.

Meuwissen, Hayes e Goddard (2001) propuseram a seleção genômica ampla que, ao invés de utilizar a informação molecular para obter o parentesco entre os híbridos, utilizam-na diretamente no modelo misto de predição, estimando o valor genotípico dos marcadores e a contribuição de cada um deles para o valor genético dos híbridos. Não há análise de mapeamento ou de ligação entre os marcadores e a característica, assumindo-se que todos os marcadores contribuem para a expressão do caráter, sendo alguns com maiores efeitos e outros com efeitos próximos de zero. Entretanto, mesmo os marcadores com pequeno efeito são considerados na seleção, o que faz com que a seleção genômica ampla seja o método mais indicado para a predição de caracteres quantitativos. Apesar de permanecer discreto por alguns anos devido ao alto custo e à restrição dos marcadores moleculares disponíveis na época, com o desenvolvimento de novos marcadores de baixo custo, esse método tornou-se bastante atrativo (MEUWISSEN, 2007; BERNARDO; YU, 2007).

A seleção genômica pode ser realizada utilizando diferentes abordagens estatísticas que vão desde métodos de regressão linear, como o RR-BLUP e LASSO, regressão não linear, bayesianos e regressão implícita, como os modelos de regressão kernel e os modelos de redes neurais. Segundo Resende et al. (2012), o método ideal para a seleção genômica deve acomodar a arquitetura genética do caráter e considerar problemas de multicolinearidade e número de marcadores. Os modelos BLUP possuem pressuposições mais simples, porém biologicamente incorretas, pois assumem que os marcadores explicam a mesma proporção da variância genética e não consideram a presença de epistasia. Já os modelos bayesianos assumem uma distribuição a priori para os efeitos dos marcadores ou para as variâncias. Assim, é esperado que, apesar de mais complexos, eles sejam mais eficientes (BERNARDO, 2010). Contudo, resultados empíricos mostram que os métodos bayesianos não são mais vantajosos que o BLUP, pelo menos no contexto do melhoramento vegetal. Em um estudo com as culturas do milho, cevada e *Arabidopsis*, Lorenzana e Bernardo (2009) mostraram que mesmo apresentando valores de

predição sutilmente melhores, as correlações entre os valores preditos e os observados foram inconsistentes nos modelos bayesianos. Além disso, a estimativa da acurácia obtida pelo BLUP não apresentou diferença significativa dos outros métodos, sendo então mais vantajoso pela sua simplicidade e rapidez.

Embora essas abordagens sejam mais aplicadas ao melhoramento animal, crescente literatura baseada em dados simulados e empíricos mostraram sua eficiência no melhoramento de plantas (BERNARDO; YU, 2007; LORENZANA; BERNARDO, 2009; CROSSA et al., 2010; HICKEY et al., 2014). Em milho, foram observados valores significativos de acurácia na predição de progênie de um cruzamento (LORENZANA; BERNARDO, 2009; MASSMAN et al., 2013; JACOBSON et al., 2014; LIAN et al., 2014). Nesse caso, os modelos de seleção genômica obtidos a partir dos dados fenotípicos e genotípicos dos genótipos de um determinado cruzamento são utilizados na predição de genótipos não avaliados deste mesmo cruzamento em um ou mais ciclos de seleção. Na maioria dos estudos, a validação do modelo é feita dividindo-se a amostra em uma população de treinamento, da qual são utilizados os dados fenotípicos e moleculares para estimar os efeitos genéticos das marcas, e uma população de validação, da qual são utilizados somente os dados moleculares para predição das médias (GODDARD; HAYES, 2007). As médias preditas da população de validação são correlacionadas com as médias fenotípicas observadas em campo, obtendo-se a acurácia de predição.

Apesar dos estudos até o momento apresentarem valores expressivos de acurácia, uma situação real dos programas de melhoramento de milho avalia híbridos de vários cruzamentos em múltiplos ambientes. Assim, são necessários experimentos para verificar se os valores de acurácia obtidos até o momento podem ser aplicados com o mesmo sucesso em populações e ambientes diferentes daqueles em que os efeitos dos marcadores são estimados. Considerando que para o melhoramento vegetal o que interessa é o ordenamento dos genótipos e o progresso com a seleção, o método utilizado na seleção genômica deve ser simples e ao mesmo tempo capaz de identificar os genótipos superiores com consideráveis valores de acurácia preditiva. Diante disso, o presente estudo teve como objetivos: (i) comparar as informações de similaridade no estado e parentesco no modelo misto de predição de híbridos de milho utilizando as capacidades geral e específica de combinação das linhagens; (ii) determinar a acurácia da seleção genômica ampla em prever a performance de híbridos de milho (a) utilizando diferentes tamanhos de população de treinamento e número de marcadores; (b) considerando a interação genótipos x ambientes; (c)

utilizando populações de treinamento e validação derivadas de cruzamentos diferentes; e (iii) comparar a eficiência da seleção fenotípica e da seleção genômica na identificação de híbridos superiores nos programas de melhoramento de milho.

Referências

BALESTRE, M.; VON PINHO, R.G.; SOUZA, J.C.; LIMA, J.L. Comparison of similarity and dissimilarity genetic coefficients based on microsatellites markers. **Genetics and Molecular Research**, Ribeirão Preto, v.7, n.4, p.695-705, 2008.

BERNARDO, R. **Breeding for quantitative traits in plants**. Woodbury: Stemma Press, 2010. 260 p.

_____. Estimation of coefficient of coancestry using molecular markers in maize. **Theoretical and Applied Genetics**, Berlin, v.85, n.8, p.1055- 1062, 1993.

_____. Genetic models for predicting maize performance in unbalanced yield trial data. **Crop Science**, Madison, v.35, n.1, p.141-147, 1995.

BERNARDO, R; YU, J. Prospects for genome wide selection for quantitative traits in maize. **Crop Science**, Madison, v. 47, p.1082-1090, 2007.

CROSSA, J.; de los CAMPOS, G.; PEREZ, P.; GIANOLA, D.; ATLIN, G.; BURGUEÑO, J.; ARAUS, J. L.; MAKUMBI, D.; YAN, J.; ARIEF, V.; BANZIGER, M.; BRAUN, H. J. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. **Genetics**, Austin, v.186, p.713-724, 2010.

DIAS, L.A.S.; PICOLI, E.A.T.; ROCHA, R.B.; ALFENAS, A.C. A priori choice of hybrid parents in plants. **Genetics and Molecular Research**, Ribeirão Preto, v.3, n.3, p.356-368, 2004.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics**. 4th ed. London: Longman Scientific and Technical, 1996. 464p.

GODDARD, M.E.; HAYES, B.J. Genomic selection. **Journal of Animal Breeding and Genetics**, Oxford, v.124, p. 323-330, 2007.

HALLAUER, A.R.; CARENA, M.J.; MIRANDA FILHO, J.B. **Quantitative genetics in maize breeding**. 3rd ed. Ames: Iowa States University Press, 2010. 664 p.

HENDERSON, C.R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, Arlington, v.31, p.423-447, 1975.

HICKEY, M.J.; DREISIGACKER, S.; CROSSA, J.; HEARNE, S.; BABU, R.; PRASANNA, B.M.; GRONDONA, M.; ZAMBELLI, A.; WINDHAUSEN, V. S.; MATHEWS, K.; GORJANC, G. Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. **Crop Science**, Madison, v. 54, p. 1476-1488, 2014.

JACOBSON, A.; LIAN, L.; ZHONG, S.; BERNARDO, R. General combining ability model for genomewide selection in a biparental cross. **Crop Science**, Madison, v.54, p.895-905, 2014.

LIAN, L.; JACOBSON, A.; ZHONG, S.; BERNARDO, R. Genomewide prediction accuracy within 969 maize biparental populations. **Crop Science**, Madison, v. 54, p.1514-1522, 2014.

LORENZANA, R.E.; BERNARDO, R. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. **Theoretical and Applied Genetics**, Berlin, v.120, p.151-161, 2009.

MASSMAN, J.M.; GORDILLO, A.; LORENZANA, R.E.; BERNARDO, R. Genomewide predictions from maize single-cross data. **Theoretical and Applied Genetics**, Berlin, v.126, p.13-22, 2013.

MELCHINGER A.E. Genetic diversity and heterosis. In: COORS, J.G.; PANDEY, S. (Ed.) **The genetics and exploitation of heterosis in crops**. Madison: ASA, CSSA, and SSSA, 1999 p.99-118.

MEUWISSEN, T.H.E. Genomic selection: marker assisted selection on genome-wide scale. **Journal of Animal Breeding and Genetics**, Oxford, v.124, p.321-322, 2007.

MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, Austin, v.157, p.1819-1829, 2001.

NEJATI-JAVEREMI, A.; SMITH, C.; GIBSON, J.P. Effect of total allelic relationship on accuracy of evaluation and response to selection. **Journal of Animal Science**, Champaign, v.75, n.7, p.1735-1738, 1997.

POWELL, J.E.; VISSCHER, P.M.; GODDARD, M.E. Reconciling the analysis of IBD and IBS in complex trait studies. **Nature Review Genetics**, London, v.11, p.800-805, 2010.

REIF, J.C.; MELCHINGER, A.E.; FRISCH, M. Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. **Crop Science**, Madison, v.45, n.1, p.1-7, 2005.

RESENDE, M.D.V.; SILVA, F.F.; LOPES, P.S.; AZEVEDO, C.F. **Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência Bayesiana (MCMC), Regressão Aleatória Multivariada (RRM) e Estatística Espacial**. Viçosa: Universidade Federal de Viçosa, 2012. 291p.

WHITE, T.L.; HODGE, G.R. **Predicting breeding values with applications in forest tree improvement.** Dordrecht: Kluwer Academic, 1989. 376p.

WRIGHT, S. Coefficient of inbreeding and relationship. **The American Naturalist**, Chicago, v.56, n.645, p. 330-338, 1922.

2 PREDIÇÃO GENÔMICA DE HÍBRIDOS DE MILHO UTILIZANDO INFORMAÇÕES SIMILARIDADE NO ESTADO E PARENTESCO

Resumo

A metodologia de modelos mistos consiste na melhor predição linear não viesada dos valores genéticos dos híbridos e pode considerar a informação do parentesco genético entre eles. Como a informação de genealogia nem sempre é conhecida no melhoramento vegetal, é cada vez mais frequente o uso de marcadores moleculares para obter o parentesco entre os híbridos. Contudo, a escolha do coeficiente a ser utilizado não é simples, e pode levar a estimativas viesadas de parentesco. O objetivo deste estudo foi realizar a predição da performance de híbridos simples de milho considerando diferentes coeficientes de parentesco e de similaridade no estado no modelo misto de predição contendo os efeitos das capacidades geral e específica de combinação. Para isso, 250 híbridos simples de milho, provenientes do cruzamento entre 50 linhagens endogâmicas e cinco linhagens testadoras foram avaliados para produção de grãos em 13 ambientes. Foram utilizados 614 marcadores AFLP para obtenção das matrizes moleculares utilizadas na estimação dos coeficientes de parentesco de Queller e Goodnight, Ritland, Lynch e Ritland, Wang, e de VanRaden, e de similaridade no estado de Jaccard, Sorensen-Dice, de e Rogers Modificado. As médias preditas utilizando cada coeficiente foram correlacionadas com as médias fenotípicas para obtenção da acurácia de predição. A seleção genômica foi posteriormente comparada à seleção fenotípica na identificação dos híbridos superiores. Os resultados mostraram que o coeficiente de parentesco de Wang e o coeficiente de similaridade no estado de Rogers Modificado foram superiores aos demais tanto na predição quanto na seleção de híbridos simples, demonstrando o potencial dessas metodologias como ferramentas a serem utilizadas nos programas de melhoramento de milho.

Palavras-chave: Predição genômica; Milho; Híbrido; Modelos mistos; BLUP; Parentesco; Similaridade no estado

Abstract

The mixed model approach is the best linear unbiased prediction of hybrid breeding values and may use the genetic relatedness between them. Whereas the genealogy information is not always known in plant breeding studies, it is increasingly common the use of molecular markers to estimate the relationship, which can lead to biased estimates of kinship. The aim of this study was to predict the performances of maize single-crosses using a mixed model containing general and specific combining abilities effects and considering different coefficients of relatedness and similarity in state. Five elite inbred lines used as testers were crossed to 50 inbred lines obtaining 250 single-crosses, subsequently evaluated for grain yield in 13 environments. 614 AFLP markers were used to obtain the molecular matrices used to estimate Queller and Goodnight, Ritland, Lynch and Ritland, Wang, and VanRaden relationship coefficients, and Jaccard, Sorensen-Dice, and Modified Rogers similarity coefficients. The predicted means using each coefficient were correlated to phenotypic means to obtain prediction accuracy. Then, genomic prediction models were compared with phenotypic selection in identifying the best single-crosses. The results showed that Wang relationship coefficient and Modified Rogers similarity coefficient were superior both in predicting and in identifying the best single-crosses, showing the potential of these approaches as tools to be used in maize breeding programs.

Keywords: Genomic prediction; Maize; Hybrids; Mixed models; BLUP; Relationship coefficient; Similarity in state

2.1 Introdução

A obtenção e avaliação de híbridos estão entre as atividades que mais demandam tempo e recursos nos programas de melhoramento de milho. Devido ao pronunciado efeito de dominância presente no controle genético da produção de grãos, a performance dos híbridos não pode ser predita apenas com base no desempenho *per se* das linhagens, sendo necessária avaliação dessas em cruzamentos (HALLAUER; CARENA; MIRANDA FILHO, 2010). Contudo, a partir de poucas linhagens é possível obter um número muito grande de híbridos, o que torna o processo impraticável e limita o número de ambientes de avaliação. Assim, métodos de predição de híbridos podem aumentar consideravelmente a eficiência dos programas de melhoramento de milho.

A metodologia dos modelos mistos foi proposta por Henderson (1975) e consiste na predição linear não viesada (*Best Linear Unbiased Prediction* - BLUP) dos valores genéticos dos híbridos corrigida para os efeitos fixos do modelo. Uma das principais propriedades dessa metodologia é a possibilidade de considerar a informação do parentesco genético entre os híbridos, sendo que a correlação entre os valores observados e preditos é máxima entre as classes de preditores lineares. Segundo Bernardo (1995), a habilidade do BLUP em prever a performance de híbridos de milho depende do grau de parentesco entre os híbridos e entre as linhagens genitoras.

O conceito de parentesco foi estabelecido pelo trabalho clássico de Wright (1922) que considera a probabilidade de dois alelos de um loco serem oriundos de um ancestral comum a partir das informações de genealogia. Esses alelos possuem o mesmo segmento de DNA e são considerados idênticos por descendência. Porém, dois alelos podem ter o mesmo segmento de DNA e produzir o mesmo fenótipo sem ter a mesma origem. Esses alelos são idênticos por estado, originados a partir de mutações em alelos ancestrais independentes. Assim, dois híbridos podem ter alelos idênticos por estado e não possuir parentesco.

Como a informação de genealogia nem sempre é conhecida no melhoramento vegetal, é cada vez mais frequente o uso de marcadores moleculares para estimar o parentesco entre os híbridos avaliados. Contudo, a definição da probabilidade de dois alelos serem idênticos por descendência a partir de dados moleculares pode ser inconsistente e levar a vários problemas práticos (POWELL; VISSCHER; GODDARD, 2010). Nesse caso, podem ser geradas

estimativas de parentesco iguais a zero quando, na verdade, os híbridos podem ter herdado alelos do mesmo ancestral mas em tempos diferentes. Assim, esses estimadores perdem muita informação contida nos dados e levam a conclusões incorretas sobre o parentesco entre os híbridos.

Os coeficientes de similaridade no estado são medidas que quantificam o número de marcas que dois híbridos compartilham, ou seja, consideram a probabilidade das marcas serem idênticas por estado, possuindo a vantagem de utilizar toda semelhança alélica existente entre os híbridos (NEJATI-JAVEREMI; SMITH; GIBSON, 1997). Contudo, esses coeficientes podem detectar similaridade entre híbridos que não apresentam nenhuma relação genealógica, podendo, portanto, superestimar o parentesco. Assim, a escolha do coeficiente a ser utilizado no modelo misto de predição nem sempre é simples, podendo influenciar expressivamente o resultado da predição. Diante do exposto, o objetivo deste estudo foi comparar a utilização de diferentes coeficientes de parentesco e de similaridade no estado no modelo misto de predição de híbridos simples de milho.

2.2 Material e Métodos

2.2.1 Dados fenotípicos e moleculares dos híbridos

Foram realizados cruzamentos entre 50 linhagens endogâmicas e cinco linhagens testadoras para obtenção dos 250 híbridos simples de milho. As 50 linhagens endogâmicas são provenientes de diferentes populações (IG-1, IG-2 e CMS-05), híbridos comerciais (HS-1, XL-560, BR-201, C-606 e C-601) e de progênies de meios-irmãos interpopulacionais (BR-105 e BR-106). As cinco linhagens previamente selecionadas como testadoras também são endogâmicas, provenientes de diferentes populações e possuem germoplasma elite (AGUIAR et al., 2003): L-08-05F (Testador 1), L-38-05D (Testador 2), L-36-07F (Testador 3), L-49-02D (Testador 4) e L-46-10D (Testador 5). A escolha das 55 linhagens objetivou representar a variabilidade existente em um programa de melhoramento de milho, em que os híbridos avaliados são provenientes de múltiplos cruzamentos de origens diferentes. Todas as linhagens endogâmicas e os 250 híbridos foram obtidos no Departamento de Genética pelo Programa de Melhoramento de Milho da

ESALQ/USP, em Piracicaba - SP. As origens das 50 linhagens endogâmicas bem como as origens e características das cinco linhagens testadoras estão apresentadas no Anexo A.

Os híbridos foram avaliados para produção de grãos em 13 ambientes, sendo cada ambiente uma combinação local/ano: Estação Experimental Fazenda Areão (E.E. Areão), Estação Experimental Fazenda Caterpillar (E.E. Caterpillar) e Estação Experimental do Departamento de Genética da ESALQ/USP (E.E. Departamento de Genética.), nos anos agrícolas de 2002/2003, 2003/2004 e 2004/2005; Estação Experimental Fazenda Anhembi (E.E. Anhembi), nos anos agrícolas 2003/2004 e 2004/2005; e Estação Experimental da Empresa Biomatrix em Patos de Minas-MG (E.E. Patos de Minas), no ano agrícola de 2004/2005 (dois experimentos). Utilizou-se o delineamento experimental látice simples 16 x 16, composto pelos 250 híbridos e seis testemunhas: os híbridos comerciais Z-8420, Z-8550, Z-8460 e Z-8480, da empresa Zeneca Sementes, e os híbridos BRS-2020 e BRS-1010 da EMBRAPA. Foram utilizadas parcelas de uma linha de 4m espaçadas 0,80m entre linhas e 0,20m entre plantas, e estande de 20 plantas por parcela, correspondendo a 62.500 plantas por hectare. A produção de grãos de cada parcela foi corrigida para teor de umidade de 15% e para estande e, posteriormente, os dados obtidos em kg parcela⁻¹ foram convertidos para t ha⁻¹.

A genotipagem das linhagens endogâmicas foi realizada pelo Centro de Biologia Molecular e Engenharia Genética da Universidade Estadual de Campinas (CBMEG/UNICAMP). Para isso, foram feitas extrações de DNA de folhas adultas das 50 linhagens e dos cinco testadores separadamente, de acordo com procedimentos descritos por Hoisington, Khairllah e Gonzales-de-Leon (1994). A genotipagem foi realizada utilizando-se 614 marcadores AFLP segundo protocolo de Vos et al. (1995). Apesar deste ser um marcador dominante, foi possível distinguir os genótipos heterozigotos dos híbridos, pois todas as 50 linhagens e os cinco testadores utilizados como genitores são endogâmicos. Os dados moleculares apresentaram desbalanceamento inferior a 30%.

2.2.2 Análises de variância

Foram realizadas as análises de variância para cada ambiente, considerando todos os efeitos aleatórios, exceto a média, segundo o modelo, eq. (1):

$$y_{ijk} = m + t_i + b_{j(k)} + r_k + e_{ijk} \quad (1)$$

em que:

y_{ijk} é o valor da produção em $t \text{ ha}^{-1}$ do híbrido i no bloco j na repetição k ;

m é a média geral do experimento;

t_i é o efeito do híbrido i ($i = 1, 2, 3, \dots, 256$);

$b_{j(k)}$ é o efeito do bloco j dentro da repetição k ($j = 1, 2, 3, \dots, 16$);

r_k é o efeito da repetição k ($k = 1, 2$);

e_{ijk} é o erro experimental associado à observação y_{ijk} com $e_{ijk} \sim N(0, \sigma^2)$.

Posteriormente foi realizada análise de variância conjunta também considerando todos os efeitos aleatórios, exceto a média, segundo o modelo, eq. (2):

$$y_{ijkl} = m + t_i + a_l + b_{j(kl)} + r_{k(l)} + (ta)_{il} + e_{ijkl} \quad (2)$$

em que:

y_{ijkl} é a produção em $t \text{ ha}^{-1}$ do híbrido i no bloco j na repetição k no ambiente l ;

m é a média geral do experimento;

t_i é o efeito do híbrido i ($i = 1, 2, 3, \dots, 256$);

a_l é o efeito do ambiente l ($l = 1, 2, 3, \dots, 13$);

$b_{j(kl)}$ é o efeito do bloco j dentro da repetição k no ambiente l ($j = 1, 2, 3, \dots, 16$);

$r_{k(l)}$ é o efeito da repetição k dentro do ambiente l ($k = 1, 2$);

$(ta)_{il}$ é o efeito da interação híbridos x ambientes;

e_{ijkl} é o erro experimental associado à observação y_{ijkl} com $e_{ijkl} \sim N(0, \sigma^2)$.

A análise conjunta foi realizada utilizando-se as médias ajustadas dos híbridos em cada ambiente. Como estas foram obtidas com recuperação da informação do resíduo intrabloco, foi calculado um erro efetivo para cada análise de variância individual, conforme apresentado por Cochran e Cox (1966), eq. (3):

$$QM_{EF} = QM_{EI} \left[1 + \frac{kj\mu}{j+1} \right] \quad (3)$$

em que:

QM_{EF} é o quadrado médio do erro efetivo;

QM_{EI} é o quadrado médio do erro intrabloco;

$QM_{B(R)}$ é o quadrado médio de bloco dentro de repetição; e

$$\mu = \frac{QM_{B/R} - QM_{EI}}{j(k-1)QM_{B(R)}}$$

O erro médio da análise conjunta foi obtido da média dos quadrados médios dos erros efetivos das análises individuais. Todas as análises de variância foram realizadas utilizando-se o software SAS (SAS INSTITUTE, 2008). Os esquemas das análises individuais e conjunta estão apresentados nos Anexos B e C. Os quadrados médios das análises de variância foram igualados às suas respectivas esperanças para obtenção das estimativas dos coeficientes de herdabilidade no sentido amplo em nível de média dos híbridos para cada ambiente ($\hat{h}_A^2$) e para os ambientes em conjunto ($\hat{h}_{Conj}^2$), eq. (4 e 5):

$$\hat{h}_A^2 = \frac{\sigma_H^2}{\sigma_F^2} = \frac{\sigma_H^2}{\left(\sigma_H^2 + \frac{\sigma^2}{K}\right)} = \frac{(QM_H - QM_{EI})/K}{QM_H/K} \quad (4)$$

$$\hat{h}_{Conj}^2 = \frac{\sigma_H^2}{\sigma_F^2} = \frac{\sigma_H^2}{\left(\sigma_H^2 + \frac{\sigma_{HA}^2}{L} + \frac{\sigma^2}{LK}\right)} = \frac{(QM_H - QM_{HA})/LK}{QM_H/LK} \quad (5)$$

em que σ_H^2 é a variância genética, σ_F^2 é a variância fenotípica em nível de média, σ_{HA}^2 é a variância da interação híbridos x ambientes, σ^2 é a variância residual, QM_H é o quadrado médio dos híbridos, QM_{HA} é o quadrado médio da interação híbridos x ambientes. Os intervalos de confiança para as estimativas dos componentes de herdabilidade foram obtidos segundo Knapp, Stroup e Ross (1985), eq. (6 e 7).

$$IC_{(\hat{h}_A^2)_{0,95}} = \left\{ 1 - \frac{1}{\left(\frac{QM_H}{QM_{EI}}\right) F_{0,025;GL_H;GL_{EI}}} \leq \hat{h}_A^2 \leq 1 - \frac{1}{\left(\frac{QM_H}{QM_{EI}}\right) F_{0,975;GL_H;GL_{EI}}} \right\} \quad (6)$$

$$IC_{(\hat{h}_{Conj}^2)_{0,95}} = \left\{ 1 - \frac{1}{\left(\frac{QM_H}{QM_{HA}}\right) F_{0,025;GL_H;GL_{HA}}} \leq \hat{h}_{Conj}^2 \leq 1 - \frac{1}{\left(\frac{QM_H}{QM_{HA}}\right) F_{0,975;GL_H;GL_{HA}}} \right\} \quad (7)$$

em que $IC_{(\hat{h}_A^2)_{0,95}}$ e $IC_{(\hat{h}_{Conj}^2)_{0,95}}$ são os intervalos de confiança para os coeficientes de herdabilidade no sentido amplo em nível de média dos híbridos para cada ambiente ($\hat{h}_A^2$) e para os ambientes em conjunto ($\hat{h}_{Conj}^2$), respectivamente, ao nível de 95% de probabilidade. $F_{0,975;GL_H;GL_{EI}}$ e $F_{0,975;GL_H;GL_{HA}}$ referem-se aos valores tabelados de F a 97,5% de probabilidade, com os graus de liberdade de híbridos (GL_H), do erro intrablocos (GL_{EI}) e da interação híbridos x ambientes (GL_{HA}). $F_{0,025;GL_H;GL_{EI}}$ e $F_{0,025;GL_H;GL_{HA}}$ referem-se aos valores tabelados de F a 2,5% de probabilidade, com os mesmos graus de liberdade. Os intervalos de confiança foram utilizados na comparação das estimativas dos coeficientes de herdabilidade, sendo que duas estimativas foram consideradas diferentes entre si ($P \leq 0,05$) quando uma estimativa não estiver contida no intervalo de confiança da outra.

2.2.3 Modelo misto de seleção genômica dos híbridos

O modelo linear misto de predição das médias dos híbridos foi o modelo apresentado por Bernardo (2010), eq. (8):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}\mathbf{g}_T + \mathbf{W}\mathbf{g}_L + \mathbf{Z}\mathbf{s} + \mathbf{e} \quad (8)$$

em que:

$\mathbf{y}$ é o vetor das médias dos híbridos no 13 ambientes, de dimensão 250 x 1;

$\mathbf{X}$ é a matriz de incidência que relaciona $\mathbf{y}$ ao efeito $\boldsymbol{\beta}$, de dimensão 250 x 1;

$\boldsymbol{\beta}$ é a média geral dos híbridos e constitui o único efeito fixo do modelo;

$\mathbf{U}$ é a matriz de incidência que relaciona $\mathbf{y}$ ao efeito $\mathbf{g}_T$, de dimensão 250×5 ;

$\mathbf{g}_T$ é o vetor dos efeitos aleatórios da capacidade geral de combinação (CGC) das linhagens testadoras, de dimensão 5×1 ;

$\mathbf{W}$ é a matriz de incidência que relaciona $\mathbf{y}$ ao efeito $\mathbf{g}_L$, de dimensão 250×50 ;

$\mathbf{g}_L$ é o vetor dos efeitos aleatórios da CGC das linhagens endogâmicas, de dimensão 50×1 ;

$\mathbf{Z}$ é a matriz de incidência que relaciona $\mathbf{y}$ ao efeito $\mathbf{s}$, de dimensão 250×250 ;

$\mathbf{s}$ é o vetor dos efeitos aleatórios da capacidade específica de combinação (CEC) dos híbridos, de dimensão 250×1 ;

$\mathbf{e}$ é o vetor dos efeitos residuais, sendo $\mathbf{e} \sim N(0, \mathbf{I}\sigma_R^2)$,

As equações de modelos mistos que resolvem esse sistema matricial são, eq. (9):

$$\begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{g}}_T \\ \hat{\mathbf{g}}_L \\ \hat{\mathbf{s}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{U} & \mathbf{X}'\mathbf{W} & \mathbf{X}'\mathbf{Z} \\ \mathbf{U}'\mathbf{X} & \mathbf{U}'\mathbf{U} + \boldsymbol{\theta}_T & \mathbf{U}'\mathbf{W} & \mathbf{U}'\mathbf{Z} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{U} & \mathbf{W}'\mathbf{W} + \boldsymbol{\theta}_L & \mathbf{W}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{U} & \mathbf{Z}'\mathbf{W} & \mathbf{Z}'\mathbf{Z} + \boldsymbol{\theta}_S \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{U}'\mathbf{y} \\ \mathbf{W}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix} \quad (9)$$

As soluções referem-se às estimativas da média geral ($\hat{\beta}$), dos efeitos das CGC das linhagens testadoras ($\hat{\mathbf{g}}_T$), dos efeitos das CGC das linhagens endogâmicas ($\hat{\mathbf{g}}_L$) e dos efeitos das CEC dos híbridos ($\hat{\mathbf{s}}$). As informações de parentesco entre os híbridos, linhagens endogâmicas e linhagens testadoras estão inseridas matricialmente nos componentes $\boldsymbol{\theta}_T$, $\boldsymbol{\theta}_L$ e $\boldsymbol{\theta}_S$ conforme, eq.(10, 11 e 12):

$$\boldsymbol{\theta}_T = \mathbf{G}_T^{-1}[\sigma_R^2/\sigma_{CGCT}^2] \quad (10)$$

$$\boldsymbol{\theta}_L = \mathbf{G}_L^{-1}[\sigma_R^2/\sigma_{CGCL}^2] \quad (11)$$

$$\boldsymbol{\theta}_S = \mathbf{S}^{-1}[\sigma_R^2/\sigma_{CEC}^2] \quad (12)$$

em que $\mathbf{G}_T^{-1}$ é o inverso da matriz de parentesco entre as linhagens testadoras, de dimensão 5×5 ; $\mathbf{G}_L^{-1}$ é o inverso da matriz de parentesco entre as linhagens endogâmicas, de dimensão 50×50 ; $\mathbf{S}^{-1}$ é o inverso da matriz de parentesco entre os híbridos, de dimensão 250×250 ; e σ_{CGCT}^2 , σ_{CGCL}^2 , σ_{CEC}^2 e σ_R^2 são os componentes de variância associados à CGC das linhagens testadoras, à

CGC das linhagens endogâmicas, à CEC dos híbridos e à variância residual, respectivamente. As matrizes de $\mathbf{G}_T$, $\mathbf{G}_L$ e $\mathbf{S}$ foram obtidas utilizando diferentes coeficientes de parentesco e similaridade no estado, conforme descrito no item seguinte.

Os componentes de variância não são conhecidos na prática. Assim, eles foram estimados utilizando o algoritmo de máxima verossimilhança restrita REML-EM (*Restricted Maximum Likelihood – Expectation-Maximization*) de Dempster, Laird e Rubin (1977). Esse é um processo iterativo no qual foram utilizados valores arbitrários iniciais de σ_{CGCT}^2 , σ_{CGCL}^2 , σ_{CEC}^2 e σ_R^2 para calcular o vetor de soluções $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{g}}_T$, $\hat{\mathbf{g}}_L$ e $\hat{\mathbf{s}}$. As soluções foram, posteriormente, utilizadas para obter novas estimativas dos componentes de variância. Assim, o processo iterativo acontece até a convergência das estimativas de σ_{CGCT}^2 , σ_{CGCL}^2 , σ_{CEC}^2 e σ_R^2 , e obtenção do vetor final de soluções.

Os valores iniciais dos componentes de variância foram os mesmos utilizados por Bernardo (2010): $\sigma_{CGCT}^2 = 0,325$; $\sigma_{CGCL}^2 = 0,30$; $\sigma_{CEC}^2 = 0,20$; e $\sigma_R^2 = 1,8$. As médias preditas dos híbridos ($\hat{\mathbf{y}}$) foram, então, obtidas a partir do somatório dos efeitos $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{g}}_T$, $\hat{\mathbf{g}}_L$ e $\hat{\mathbf{s}}$ do vetor final de soluções, eq.(13):

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{U}\hat{\mathbf{g}}_T + \mathbf{W}\hat{\mathbf{g}}_L + \mathbf{Z}\hat{\mathbf{s}} \quad (13)$$

em que as matrizes $\mathbf{X}$, $\mathbf{U}$, $\mathbf{W}$ e $\mathbf{Z}$ são as mesmas descritas para eq. (8). As operações matriciais de predição das médias, bem como as análises iterativas para estimação dos componentes de variância foram realizadas utilizando-se o software R (R DEVELOPMENT CORE TEAM, 2012).

2.2.4 Obtenção das matrizes de parentesco e similaridade no estado

As matrizes de parentesco entre as linhagens testadoras ($\mathbf{G}_T$), entre as linhagens endogâmicas ($\mathbf{G}_L$) e entre os híbridos ($\mathbf{S}$) foram obtidas utilizando cinco coeficientes de parentesco, eq. (14, 15, 16, 17, 18), e três coeficientes de similaridade no estado, eq. (19, 20, 21):

i. QG – Queller e Goodnight (1989)

$$\hat{r}_{QG_{XY}} = \frac{0,5(S_{ac} + S_{ad} + S_{bc} + S_{bd}) - p_a - p_b}{1 + S_{ab} - p_a - p_b} \quad (14)$$

em que:

- $\hat{r}_{QG_{XY}}$ é o estimador do parentesco genético entre o genótipo X (tomado como referência), que possui os alelos “a” e “b”, e o genótipo Y, que possui os alelos “c” e “d”;
- S_{ac} é igual à 1 se “a” e “c” forem iguais, e igual à 0 caso contrário;
- S_{ad} é igual à 1 se “a” e “d” forem iguais, e igual à 0 caso contrário;
- S_{bc} é igual à 1 se “b” e “c” forem iguais, e igual à 0 caso contrário;
- S_{bd} é igual à 1 se “b” e “d” forem iguais, e igual à 0 caso contrário;
- p_a e p_b são as frequências dos alelos “a” e “b” dos genótipos.

ii. R – Ritland (1996)

$$\hat{r}_{R_{XY}} = \frac{2(S'_{XY} - J_o)}{1 - J_o} \quad (15)$$

em que:

- $\hat{r}_{R_{XY}}$ é o estimador do parentesco genético entre os genótipos X e Y, considerando que X e Y podem ter até quatro alelos diferentes: ‘i’, ‘j’, ‘k’ e ‘l’;
- S'_{XY} é igual à 1,0 se $(x = ii; y = ii)$ ou $(x = ij; y = ij)$;
- S'_{XY} é igual à 0,75 se $(x = ii; y = ij)$;
- S'_{XY} é igual à 0,5 se $(x = ij; y = ik)$;
- S'_{XY} é igual à 0 se $(x = ii; y = kl)$;
- J_o é a homozigiosidade esperada, sendo $J_o = \sum_{i=1}^n p_i^2$, em que n é o número de locos.

iii. LR – Lynch e Ritland (1999)

$$\hat{r}_{LRXY(k)} = \frac{p_a(S_{bc} + S_{bd}) + p_b(S_{ac} + S_{ad}) - 4p_ap_b}{(1 + S_{ab})(p_a + p_b) + 2p_ap_b} \quad (16)$$

em que:

- $\hat{r}_{LRXY(k)}$ é o estimador do parentesco genético entre o genótipo X (tomado como referência), que possui os alelos “a” e “b”, e o genótipo Y, que possui os alelos “c” e “d” para o loco k;
- S_{bc} é igual à 1 se “b” e “c” forem iguais, e igual à 0 caso contrário;
- S_{bd} é igual à 1 se “b” e “d” forem iguais, e igual à 0 caso contrário;
- S_{ac} é igual à 1 se “a” e “c” forem iguais, e igual à 0 caso contrário;
- S_{ad} é igual à 1 se “a” e “d” forem iguais, e igual à 0 caso contrário;
- p_a e p_b são as frequências dos alelos “a” e “b” dos genótipos.

iv. W – Wang (2002)

$$\hat{r}_{WXY} = \frac{(1 + S_{ab}) + (p_a + p_b) - 4p_ap_b}{2p_ap_b} \quad (17)$$

em que:

- $\hat{r}_{WXY}$ é o estimador do parentesco genético entre o genótipo X (tomado como referência), que possui os alelos “a” e “b”, e o genótipo Y, que possui os alelos “c” e “d”;
- S_{ab} é igual à 1 se “a” e “b” forem iguais, e igual à 0 caso contrário;
- p_a e p_b são as frequências dos alelos “a” e “b” dos genótipos.

v. VR – Van Raden (2007)

$$\mathbf{G} = \frac{(\mathbf{MM}')}{\sum_{k=1}^m 2\hat{p}_k(1 - \hat{p}_k)} \quad (18)$$

em que:

- **G** é a matriz quadrada contendo o parentesco entre os genótipos;
- **M** é a matriz dos dados dos marcadores moleculares, com valores 1, 0 e -1;
- **m** é o número de marcas;
- $\hat{p}_k$ é a frequência alélica para o loco **k**.

vi. J – Jaccard (1908)

$$\hat{r}_{J_{XY}} = \frac{a}{a + b + c} \quad (19)$$

em que:

- $\hat{r}_{J_{XY}}$ é o estimador da similaridade no estado entre os genótipos X e Y;
- **a** é o número de bandas em comum entre X e Y;
- **b** é o número de bandas presentes em X e ausentes em Y;
- **c** é o número de bandas presentes em Y e ausentes em X.

vii. SD - Sorensen-Dice (DICE, 1945)

$$\hat{r}_{SD_{XY}} = \frac{2a}{2a + b + c} \quad (20)$$

em que:

- $\hat{r}_{SD_{XY}}$ é o estimador da similaridade no estado entre os genótipos X e Y;
- **a** é o número de bandas em comum entre X e Y;
- **b** é o número de bandas presentes em X e ausentes em Y;
- **c** é o número de bandas presentes em Y e ausentes em X.

viii. MR – Rogers Modificado (WRIGHT, 1978)

$$\hat{d}_{MR_{XY}} = \sqrt{\frac{1}{2m} \sum_k (x_{kx} - x_{ky})^2} \quad (21)$$

em que:

- $\hat{d}_{MR_{XY}}$ é o estimador da similaridade no estado entre os genótipos X e Y;
- m é o número de marcas;
- x_{kx} e x_{ky} são as frequências do alelo k em X e Y.

Os coeficientes QG, R, LR e W foram estimados utilizando o software Mark (RITLAND, 2006). Os demais coeficientes foram estimados utilizando-se o software R (R DEVELOPMENT CORE TEAM, 2012), sendo que para obtenção de J e SD foi utilizado o pacote ‘ecodist’ (GOSLEE; URBAN, 2007). Além das matrizes de parentesco e similaridade no estado originadas por esses coeficientes, também foram utilizadas matrizes identidades no lugar de $\mathbf{G}_T$, $\mathbf{G}_L$ e $\mathbf{S}$, assumindo-se, portanto, nenhum parentesco ou covariância genética entre os testadores, linhagens e híbridos.

As predições dos híbridos de milho foram realizadas utilizando cada uma das matrizes de parentesco e similaridade no estado no modelo misto apresentado anteriormente. A eficiência dos modelos de seleção genômica em prever a performance dos híbridos foi verificada a partir da estimativa da acurácia de predição (r_{MG}), que consiste na correlação entre as médias preditas e as médias fenotípicas dos híbridos, dividida pela raiz quadrada do coeficiente de herdabilidade da seleção genômica (h_{SG}^2), (DEKKERS, 2007). O coeficiente h_{SG}^2 foi obtido no sentido amplo em nível de média dos híbridos, (eq. 22):

$$h_{SG}^2 = \frac{\sigma_{CGCT}^2 + \sigma_{CGCL}^2 + \sigma_{CEC}^2}{\sigma_{CGCT}^2 + \sigma_{CGCL}^2 + \sigma_{CEC}^2 + \frac{\sigma_R^2}{L}} \quad (22)$$

O componente de variância residual (σ_R^2) é dividido pelo número de ambientes (L) e considera tanto a variância do erro dentro de ambientes (σ_e^2), dividida pelo número de repetições (K), quanto a variância da interação híbridos x ambientes (σ_{ta}^2), (eq.23):

$$\sigma_R^2 = \frac{\sigma_e^2}{K} + \sigma_{ta}^2 \quad (23)$$

2.3 Resultados e discussão

2.3.1 Análises de variância

Foram observadas diferenças significativas entre os híbridos ($p \leq 0,01$) nas análises de variância individuais em todos os ambientes. As médias da produção de grãos variaram de 6,75 t ha⁻¹ na E.E. Caterpillar - 2002/2003, a 10,32 t ha⁻¹ na E.E. Departamento de Genética - 2004/2005. Os coeficientes de variação (CV%) variaram de 9,65% na E.E. Areão - 2003/2004, a 22,36% na E.E. Areão - 2004/2005. Somente em dois ambientes as estimativas do coeficiente de variação foram superiores à 20% em função da ocorrência de estresse hídrico após o florescimento dos híbridos nesses ambientes (20,31% na E.E. Caterpillar - 2004/2005 e 22,36% na E.E. Areão - 2004/2005). Nos demais ambientes, as estimativas foram sempre inferiores a 15%. Os valores de CV(%) obtidos estão dentro dos limites reportados na literatura (AGUIAR et al., 2003; RAPOSO, 2002). Os coeficientes de herdabilidade ampla em nível de média dos híbridos para cada ambiente (h_A^2) variaram de 0,63, com intervalo de confiança de 0,53 - 0,71, na E.E. Caterpillar - 2004/2005, a 0,81 com intervalo de confiança variando de 0,76 - 0,85 na E.E. Departamento de Genética - 2002/2003, portanto, diferentes estatisticamente (Anexo D).

Também foram observadas diferenças significativas ($p \leq 0,01$) na análise de variância conjunta dos 13 ambientes para as fontes de variação híbridos, ambientes e para a interação híbridos x ambientes (H x A). Esses resultados mostram que há variabilidade genética entre os híbridos e que os híbridos apresentaram desempenho diferencial entre os ambientes em que foram avaliados. A média de produção de grãos dos híbridos na análise conjunta foi de 8,55 t ha⁻¹, e o coeficiente de variação (CV%) foi de 13,21%. O coeficiente de herdabilidade ampla em nível de média dos híbridos na análise conjunta (h_{Conj}^2) foi de 0,95, com intervalo de confiança de 0,94 - 0,96 (Anexo E).

2.3.2 Predição genômica dos híbridos para os diferentes coeficientes de similaridade no estado e parentesco

Os componentes de variância associados à capacidade geral de combinação das linhagens testadoras (σ_{CGCT}^2), capacidade geral de combinação das linhagens endogâmicas (σ_{CGCL}^2),

capacidade específica de combinação dos híbridos (σ_{CEC}^2) e variância residual (σ_{R}^2) variaram expressivamente entre os coeficientes. As estimativas de σ_{CGCT}^2 variaram de 0,002, para o coeficiente de parentesco de VanRaden (VR), a 2,18, para o coeficiente de similaridade de Sorense-Dice (SD); as estimativas de σ_{CGCL}^2 variaram de 0,02, para o coeficiente de parentesco VR, a 2,28, para o coeficiente de similaridade SD; e as estimativas de σ_{CEC}^2 variaram de 0,16, para o coeficiente de parentesco de Ritland (R), a 3,22, para o coeficiente de similaridade de Rogers Modificado (MR). Apesar dessas estimativas variarem entre os coeficientes, o componente de variância residual (σ_{R}^2) variou proporcionalmente em cada caso, originando estimativas de herdabilidade da seleção genômica (h_{SG}^2) similares entre os coeficientes. As estimativas de h_{SG}^2 variaram de 0,95 para o coeficiente de parentesco VR, a 1,00 para os coeficientes de parentesco de Wang (W) e de similaridade de Rogers Modificado (MR), (Tabela 1).

Foram observadas elevadas estimativas de acurácia de predição (r_{MG}) para todos os coeficientes. Entre os coeficientes de parentesco, as estimativas de r_{MG} variaram de 0,62 para o coeficiente R, a 1,00 para o coeficiente W. Entre os coeficientes de similaridade no estado, as estimativas de r_{MG} variaram de 0,92 para o coeficiente SD, a 1,00 para MR. Quando se utilizou a matriz identidade (I) no modelo de predição, ou seja, quando se considerou que não há parentesco entre as linhagens e híbridos, a estimativa de r_{MG} foi de 0,85, superior às obtidas para os coeficientes R e VR (Tabela 1).

Tabela 1 – Estimativas dos componentes de variância σ_{CGCT}^2 , σ_{CGCL}^2 , σ_{CEC}^2 , σ_{R}^2 , do coeficiente de herdabilidade ampla da seleção genômica em nível de média dos híbridos (h_{SG}^2) e da acurácia de predição (r_{MG}) dos coeficientes de parentesco e similaridade no estado utilizados no modelo misto de predição

	Coeficientes de parentesco*					Coeficientes de similaridade no estado*			I*
	QG	R	LR	W	VR	J	SD	MR	
σ_{CGCT}^2	0.34	0.42	0.34	0.19	0.002	1.40	2.18	0.23	0.68
σ_{CGCL}^2	0.32	2.08	0.31	0.19	0.02	1.37	2.28	0.10	0.36
σ_{CEC}^2	0.66	0.16	0.67	1.45	1.39	1.33	1.38	3.22	0.10
σ_{R}^2	0.54	1.28	0.54	0.01	0.90	0.37	0.56	0.01	0.92
h_{SG}^2	0.97	0.96	0.97	1.00	0.95	0.99	0.99	1.00	0.94
r_{MG}	0.96	0.62	0.96	1.00	0.82	0.97	0.92	1.00	0.85

*QG: Queller e Goodnight; R: Ritland; LR: Lynch e Ritland; W: Wang; VR: VanRaden; J: Jaccard; SD: Sorensen-Dice; MR: Rogers Modificado; I: Matriz Identidade

O vetor final de soluções com as estimativas da média geral ($\hat{\beta}$), dos efeitos da capacidade geral de combinação (CGC) das linhagens testadoras ($\hat{g}_T$), dos efeitos da CGC das linhagens endogâmicas ($\hat{g}_L$) e dos efeitos da capacidade específica de combinação (CEC) dos híbridos ($\hat{s}$) estão apresentados nos Anexos G e H. Assim como para σ^2_{CGCT} , σ^2_{CGCL} e σ^2_{CEC} , as soluções $\hat{\beta}$, $\hat{g}_T$, $\hat{g}_L$ e $\hat{s}$ variaram expressivamente entre os coeficientes.

A acurácia de predição (r_{MG}) é uma medida da eficiência do estimador, pois ela considera a correlação entre as médias fenotípicas e as médias preditas a partir das informações moleculares, ponderada pela estimativa de herdabilidade (DEKKERS, 2007). Um estimador acurado apresenta menor diferença quadrática entre os valores observados e preditos, ou seja, apresenta mínimo erro quadrático médio e maior acurácia de predição. Neste estudo, foram observadas elevadas estimativas de r_{MG} para todos os coeficientes de parentesco e similaridade no estado. Esses resultados indicam que o modelo de seleção genômica contendo os efeitos das capacidades geral e específica de combinação apresentado por Bernardo (2010) foi eficiente em prever a performance dos híbridos para todos os coeficientes.

É importante salientar, contudo, que as médias fenotípicas utilizadas no modelo de predição foram obtidas por experimentos com repetições e avaliados em múltiplos ambientes, ou seja, minimizando os efeitos não genéticos do modelo. Além disso, todos os 250 híbridos foram considerados no modelo de predição. Quando são utilizados dados desbalanceados ou quando apenas uma amostra dos híbridos é fenotipada para predição de híbridos não avaliados, a estimativa de acurácia de predição tende a ser bem menor, e o modelo pode não ser eficiente em prever a média dos híbridos não avaliados (LORENZANA; BERNARDO, 2009; LIAN et al., 2014). Nesse caso, a diferença das estimativas de acurácia apresentada entre os coeficientes, mesmo que pequena, pode ser bastante relevante.

A superioridade do coeficiente W em relação aos coeficientes de parentesco R, RL e QG confirma o trabalho de Wang (2002), que compara esses coeficientes em precisão e acurácia, considerando diferentes distribuições de frequência alélica, número de alelos por loco, tamanho amostral e inclusão de parentesco real entre os híbridos. Quando a frequência alélica da população é conhecida, todos os coeficientes originam estimativas não viesadas do parentesco. Contudo, o autor afirma que, quando a frequência alélica da população é obtida a partir de uma amostra, o erro amostral deve ser considerado na estimação. O desequilíbrio de ligação entre os locos interfere na estimação das frequências alélicas e, dependendo dos locos utilizados na

amostragem, o parentesco pode ser consideravelmente alterado. O coeficiente W é o único que considera o efeito da ligação diferencial entre os locos, enquanto que o coeficiente QG ignora essa informação e os coeficientes R e LR consideram-na apenas parcialmente. O autor conclui que o coeficiente W é superior aos demais especialmente quando são utilizadas pequenas amostras.

Já era esperado que o coeficiente LR fosse superior aos coeficientes QG e R, pois o LR foi proposto para superar as limitações de aplicação e computacionais desses coeficientes (LYNCH; RITLAND, 1999). Contudo, o que chamou atenção foi a menor estimativa de r_{MG} para o coeficiente VR, semelhante a obtida quando se utilizaram matrizes identidade no lugar de G_T , G_L e S , ou seja, quando não se considerou parentesco. Segundo Van Raden (2008), esse coeficiente propõe a estimação do efeito genético total a partir de operações matriciais de modelos lineares mistos, sendo equivalente à predição via seleção genômica ampla apresentada por Meuwissen, Hayes e Goddard (2001). Essa abordagem já havia sido proposta anteriormente por Nejati-Javanemi, Smith e Gibson (1997) e Fernando (1998). Além de ser computacionalmente simples, contempla toda informação genotípica, sendo frequentemente utilizada nos estudos de seleção genômica em melhoramento animal. Contudo, as matrizes de parentesco originadas por esse coeficiente podem ser singulares, ou seja, podem não ter solução única (VAN RADEN, 2008). Isso acontece quando as matrizes possuem linhas ou colunas repetidas, por exemplo quando dois híbridos ou linhagens apresentam o mesmo genótipo para um dado marcador molecular.

É importante salientar que todos os coeficientes de parentesco obtidos a partir de marcadores moleculares são, de alguma forma, viesados (LYNCH; RITLAND, 1999). O cálculo do parentesco entre dois híbridos deve ter como referência uma população base, ou seja, a probabilidade de dois alelos serem descendentes do mesmo ancestral dessa população de referência. Tradicionalmente, a probabilidade de dois alelos serem idênticos por descendência é calculada com base em um *pedigree* conhecido, sendo os genótipos do topo desse *pedigree* considerados os fundadores da população base. A obtenção desses coeficientes com base na informação molecular obviamente não considera nenhuma população base, o que pode levar a estimativas viesadas de parentesco (POWELL; VISSCHER; GODDARD, 2010). Isso porque, durante a meiose, os segmentos de DNA são quebrados para que haja recombinação. A existência de segmentos de DNA herdados, então, depende do número de gerações desde o ancestral

comum mais recente. Assim, métodos que consideram o parentesco entre dois híbridos igual a zero, por exemplo, podem não estar considerando todas as informações contidas nos dados e, com isso, estar subestimando o parentesco. Segundo Lynch e Ritland (1999), a magnitude desse viés depende principalmente do real parentesco entre os híbridos.

Os coeficientes de similaridade são medidas que quantificam o número de marcas que dois híbridos compartilham. De modo geral, esses coeficientes baseiam-se em comparações entre a ocorrência de bandas comuns e diferentes para o mesmo loco, e geralmente são utilizados para obter a divergência genética entre linhagens e distinção de grupos heteróticos (MELCHINGER, 1999; REIF et al., 2003; BALESTRE et al., 2008). Existem alguns relatos na literatura do uso de coeficientes de similaridade na predição de híbridos de milho (BALESTRE; VON PINHO; SOUZA, 2010; FERREIRA et al., 2010). Comparados aos coeficientes de parentesco, os coeficientes de similaridade possuem a vantagem de utilizar toda semelhança alélica existente entre os híbridos e linhagens (NEJATI-JAVEREMI; SMITH; GIBSON, 1997). Contudo, também são estimativas viesadas do parentesco e podem detectar similaridade entre os híbridos quando, na verdade, estes não apresentam nenhuma relação genealógica. Com o intuito de contornar essa limitação, Bernardo (1993) propõe o uso de um coeficiente de coancestria que utiliza a informação de similaridade genética para corrigir possíveis vieses da similaridade no estado. Entretanto, essa metodologia somente é válida para dialelos parciais nos quais não ocorrem cruzamentos endogâmicos, podendo ser utilizada somente quando se tem a informação prévia do grupo heterótico ao qual pertencem os híbridos.

Dentre os coeficientes de similaridade utilizados, o coeficiente MR é o único que considera a frequência alélica dos marcadores, o que o torna mais informativo que os demais. Segundo Reif, Melchinger e Frisch (2005), esse coeficiente possui as propriedades genéticas desejáveis na predição da heterose e determinação de grupos heteróticos. Esses autores afirmam que a distância de Rogers, proposta inicialmente por Rogers (1972), é linearmente relacionada ao coeficiente de coancestria de Malecot (1948), que é o parentesco genealógico utilizado nas equações de modelos mistos de Henderson (1975). Visando aprimorá-la, Wright (1978) propôs o coeficiente de Rogers Modificado (MR), apenas atribuindo uma dimensão diferente para cada alelo. Neste estudo, o coeficiente MR também foi superior quando utilizado na predição dos híbridos.

Apesar de menos informativos, os coeficientes J e SD também apresentaram boas estimativas de acurácia. Esses coeficientes não consideram a ausência comum de bandas que, segundo Meyer et al. (2004), não significa necessariamente que essas regiões do DNA são idênticas, o que poderia levar a estimativas ainda mais viesadas da similaridade entre os híbridos. Foi possível observar que, entre os coeficientes de similaridade, o SD foi o que apresentou menor r_{MG} . Isso pode ter acontecido porque ele considera o dobro da ocorrência comum de bandas de J, superestimando a similaridade entre os híbridos.

2.3.3 Seleção fenotípica e genômica dos híbridos superiores

Os híbridos preditos via seleção genômica (SG) utilizando as diferentes matrizes de parentesco e similaridade no estado foram ranqueados e identificados os 25 híbridos superiores em cada caso. Também foi realizada seleção fenotípica (SF) a partir das médias dos híbridos nos 13 ambientes. A proporção de híbridos selecionados em comum pela SF e SG utilizando os diferentes coeficientes está apresentada na Tabela 2. O coeficiente R foi o que apresentou menor coincidência de híbridos selecionados com os demais coeficientes, em média 62%, e com a SF, 52%. Os coeficientes LR e QG apresentaram 100% de coincidência de híbridos selecionados entre si, e 76% com a SF. A utilização de matrizes identidade (I) no lugar das matrizes de parentesco, assumindo-se nenhum parentesco entre as linhagens e híbridos, apresentou coincidência de 76% de híbridos selecionados com a SF, proporção igual ou maior às apresentadas entre a SF e os coeficientes QG, R, RL, VR e SD. Os coeficientes W e MR apresentaram 100% de coincidência entre si e com a seleção fenotípica, ou seja, W, MR e a SF selecionaram os mesmos híbridos (Tabela 2).

Tabela 2 – Coincidência de híbridos selecionados (%) pela seleção genômica utilizando os coeficientes de parentesco e similaridade no estado, ausência de parentesco (matriz identidade) e seleção fenotípica

Coeficientes*	QG	R	RL	W	VR	J	SD	MR	I
R	64								
RL	100	64							
W	76	52	76						
VR	88	72	88	68					
J	88	60	88	84	80				
SD	84	68	84	72	84	84			
MR	76	48	76	100	68	84	72		
I	88	60	88	76	88	84	80	76	
SF	76	52	76	100	68	84	76	100	76

*QG: Queller e Goodnight; R: Ritland; LR: Lynch e Ritland; W: Wang; VR: VanRaden; J: Jaccard; SD: Sorensen e Dice; MR: Rogers Modificado; I: matriz identidade; SF: seleção fenotípica

A comparação da proporção de híbridos superiores identificados pela SG utilizando os diferentes coeficientes e a SF é um excelente indicador da eficiência desses coeficientes, pois a SF foi realizada baseando-se nas médias ajustadas dos híbridos avaliados em experimentos com repetições e avaliados em múltiplos ambientes. A estimativa da herdabilidade no sentido amplo em nível de média dos híbridos da análise de variância conjunta dos híbridos nos 13 ambientes ($h^2_{\text{Conj}} = 0,95$), indica que 95% da variação presente entre os híbridos é devida a causas genéticas. Portanto, espera-se que os híbridos selecionados pela SF sejam realmente superiores geneticamente. Assim, os coeficiente W e MR, que obtiveram maior estimativa de r_{MG} na predição, confirmaram sua superioridade também na identificação de híbridos superiores.

É interessante observar que, além dos coeficientes W e MR, apenas o coeficiente J teve coincidência de híbridos selecionados com a SF (84%) superior à apresentada quando utilizou-se a matriz I (76%). Os outros coeficientes QG, R, RL, VR e SD apresentaram coincidência igual ou inferior a 76%, indicando que o parentesco obtido por esses coeficientes pode ter prejudicado a identificação de híbridos superiores.

Comparando a eficiência dos coeficientes de similaridade e parentesco na predição de híbridos simples de milho não avaliados, Ferreira et al. (2010) observaram que a correlação entre as médias preditas e observadas foi ligeiramente superior para os coeficientes de similaridade, principalmente para grandes níveis de desbalanceamento. Os autores explicam que esses resultados podem estar associados aos baixos valores de parentesco estimados para as linhagens

genitoras, o que dificulta a predição quando os níveis de desbalanceamento são maiores. De forma geral, a similaridade no estado mostrou melhor ranqueamento dos híbridos, confirmado pelo cálculo da eficiência de seleção, indicando que esses coeficientes podem ser melhor aplicados no contexto do melhoramento. Contudo, é importante observar que os autores utilizaram uma amostra muito pequena de genótipos (29 híbridos) e marcas (48 marcadores microssatélites), e os coeficientes de parentesco foram propostos para o estudo de populações naturais, em que o tamanho da amostra pode interferir muito na estimativa do parentesco (RITLAND, 1996).

2.3.4 Considerações finais

Os programas de melhoramento de milho visam identificar e selecionar híbridos superiores para serem liberados posteriormente como cultivares comerciais. A eficiência da seleção, admitindo-se adequada experimentação, depende diretamente de métodos estatísticos que aproveitem ao máximo as informações disponíveis. Nesse sentido, o enfoque da análise de modelos mistos proposta por Henderson (1975) tem a vantagem de possibilitar a inclusão da informação de parentesco genético entre os híbridos, sendo que a correlação entre os valores observados e preditos é máxima entre as classes de preditores lineares.

A obtenção do parentesco a partir da informação de marcadores moleculares pode ser inconsistente e levar a vários problemas práticos (POWELL; VISSCHER; GODDARD, 2010). Os coeficientes de parentesco podem gerar estimativas de parentesco iguais a zero quando, na verdade, os híbridos podem ter herdado os alelos do mesmo ancestral mas em tempos diferentes. Os coeficientes de similaridade no estado, por outro lado, possuem a vantagem de utilizar toda semelhança alélica existente entre os híbridos (NEJATI-JAVEREMI; SMITH; GIBSON, 1997). Contudo, podem superestimar o parentesco, detectando similaridade entre híbridos que não possuem relação genealógica.

Todos os coeficientes de parentesco e similaridade no estado apresentaram elevadas estimativas de acurácia de predição neste estudo. Contudo, considerando que o milho é uma cultura que vem sendo intensivamente melhorada há muitos anos, é importante que sejam estudados métodos capazes de identificar diferenças significativas entre os híbridos. Nesse

sentido, a diferença das estimativas de acurácia apresentada entre os coeficientes, mesmo que pequena, pode ser bastante relevante.

Os coeficientes Wang e Rogers Modificado foram os que apresentaram maior acurácia de predição e coincidência na identificação de híbridos superiores em relação à seleção fenotípica. Essa comparação é um excelente indicador da eficiência desses coeficientes, pois a seleção fenotípica foi realizada baseando-se nas médias ajustadas dos híbridos avaliados em experimentos com repetições e avaliados em múltiplos ambientes. Portanto, espera-se que os híbridos selecionados pela seleção fenotípica sejam realmente superiores geneticamente.

A superioridade do coeficiente Wang sobre os demais coeficientes de parentesco é relatada por Wang (2002), principalmente nos casos em que são avaliadas pequenas amostras, como é o caso das populações de melhoramento. O coeficiente de similaridade de Rogers Modificado é linearmente relacionado ao coeficiente de coancestria de Malecot (1948), que é o parentesco genealógico utilizado nas equações de modelos mistos de Henderson (1975), sendo, portanto, a medida de similaridade que mais se aproxima do parentesco genético entre os híbridos (REIF; MELCHINGER; FRISCH, 2005). Assim, os resultados deste estudo indicam que a diferença conceitual entre coeficientes de parentesco e coeficientes de similaridade no estado não tem tanta importância nos estudos de predição genômica em plantas, pois ambos consideram somente a informação molecular dos genótipos, e não sua ascendência ou descendência. Apesar disso, a facilidade da utilização da informação molecular na obtenção dos coeficientes e as elevadas estimativas de acurácia obtidas neste estudo demonstram o potencial dessas metodologias como ferramentas a serem utilizadas na predição de híbridos, podendo aumentar consideravelmente a eficiência dos programas de melhoramento de milho.

Referências

AGUIAR, A.M.; CARLINI-GARCIA, L.A.; SILVA, A.R.; SANTOS, M.F.; GARCIA, A.A.F.; SOUZA JUNIOR, C.L. Combining ability of inbred lines of maize and stability of their respective single-crosses. **Scientia Agricola**, Piracicaba, v.60, n.1, p.83-89, 2003.

BALESTRE, M.; VON PINHO, R.G.; SOUZA, J.C. Prediction of maize single-cross performance by mixed linear models with microsatellite marker information. **Genetics and Molecular Research**, Ribeirão Preto, v.9, n.2, p.1054-1068, 2010.

BALESTRE, M.; VON PINHO, R.G.; SOUZA, J.C.; LIMA, J.L. Comparison of similarity and dissimilarity genetic coefficients based on microsatellites markers. **Genetics and Molecular Research**, Ribeirão Preto, v.7, n.4, p.695-705, 2008.

BERNARDO, R. **Breeding for quantitative traits in plants**. Woodbury: Stemma Press, 2010. 260p.

_____. Estimation of coefficient of coancestry using molecular markers in maize. **Theoretical and Applied Genetics**, Berlin, v.85, n.8, p.1055-1062, 1993.

_____. Genetic models for predicting maize performance in unbalanced yield trial data. **Crop Science**, Madison, v.35, n.1, p.141-147, 1995.

DEKKERS, J.C.M. Prediction of response to marker-assisted and genomic selection using selection index theory. **Journal of Animal Breeding and Genetics**, Oxford, v.124, p.331-341, 2007.

DEMPSTER, A.P.; LAIRD, N.M.; RUBIN, D.B. Maximum likelihood from incomplete data via EM algorithm. **Journal of the Royal Statistical Society**, London, v.39, p.1-38, 1977.

DICE, L.R. Measures of the amount of ecologic association between species. **Ecology**, Washington, v.26, p.297-302, 1945.

FERNANDO, R. L. Genetic evaluation and selection using genotypic, phenotypic and pedigree information. In: 6th WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 1998, Armidale. **Proceedings of the 6th World Congress on Genetics Applied to Livestock Production**. Armidale: NSW, 1998. p. 329-336.

FERREIRA, D.V.; VON PINHO, R.G.; BALESTRE, M.; OLIVEIRA, R.L. Prediction of maize hybrid performance using similarity in state and similarity by decent information. **Genetics and Molecular Research**, Ribeirão Preto, v.9, n.4, p.2381-2394, 2010.

GOSLEE, S.C.; URBAN, D.L. The ecodist package for dissimilarity-based analysis of ecological data. **Journal of Statistical Software**, v.22, p.1-19, 2007.

HALLAUER, A.R.; CARENA, M.J.; MIRANDA FILHO, J.B. **Quantitative genetics in maize breeding**. 3rd ed. Ames: Iowa States University Press, 2010. 664 p.

HENDERSON, C.R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, Washington, v.31, n.2, p.423-447, 1975.

HOISINGTON, D.; KHAIRLLAH, M.; GONZALES-DE-LEON, D. **Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory**. 2nd ed. Mexico: CIMMYT, 1994, 102 p.

JACCARD, P. Nouvelles recherches sur la distribution florale. **Bulletin de la Société vaudoise des sciences naturelles**, Lausanne, v.44, p. 223-270, 1908.

KNAPP, S.J.; STROUP, W.W.; ROSS, W.M. Exact confidence intervals for heritability on a progeny mean basis. **Crop Science**, Madison, v. 25, n. 1, p. 192-194, 1985.

LIAN, L.; JACOBSON, A.; ZHONG, S.; BERNARDO, R. Genomewide prediction accuracy within 969 maize biparental populations. **Crop Science**, Madison, v. 54, p. 1514-1522, 2014.

LORENZANA, R.E.; BERNARDO, R. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. **Theoretical and Applied Genetics**, Berlin, v. 120, p. 151-161, 2009.

LYNCH, M.; RITLAND, K. Estimation of pairwise relatedness with molecular markers. **Genetics**, Baltimore, v.152, n.4, p.1753-1766, 1999.

MALECOT, G. **Les Mathématiques de l'Hérédité**, Paris: Masson et Cie, 1948. 64 p.

MELCHINGER A.E. Genetic diversity and heterosis. In: COORS, J.G.; PANDEY, S. (Ed.) **The genetics and exploitation of heterosis in crops**. Madison: ASA, CSSA, and SSSA, 1999. p. 99–118.

MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, Austin, v. 157, p. 1819-1829, 2001.

MEYER, A.S.; GARCIA, A.A.F.; SOUZA, A.P.; SOUZA JUNIOR, C.L. Comparison of similarity coefficients used for cluster analysis with dominant markers in maize (*Zea mays L.*). **Genetic and Molecular Biology**, Ribeirão Preto, v.27, n.1, p.83-91, 2004.

NEJATI-JAVEREMI, A.; SMITH, C.; GIBSON, J.P. Effect of total allelic relationship on accuracy of evaluation and response to selection. **Journal of Animal Science**, Champaign, v.75, n.7, p.1735-1738, 1997.

POWELL, J.E.; VISSCHER, P.M.; GODDARD, M.E. Reconciling the analysis of IBD and IBS in complex trait studies. **Nature Review Genetics**, London, v.11, p.800-805, 2010.

QUELLER, D.C.; GOODNIGHT, K.F. Estimating relatedness using genetic markers. **Evolution**, Houston, v.43, n.2, p.258-275, 1989.

R DEVELOPMENT CORE TEAM. **R**: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2012. 3604 p.

REIF, J.C.; MELCHINGER, A.E.; FRISCH, M. Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. **Crop Science**, Madison, v.45, n.1, p.1-7, 2005.

REIF, J.C.; MELCHINGER, A.E.; XIA, X.C.; WARBURTON, M.L. Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. **Crop Science**, Madison, v.43, n.4, p.1275-1282, 2003.

RITLAND, K. Estimators for pairwise relatedness and individual inbreeding coefficients. **Genetical Research**, London, v.67, n.2, p.175-185, 1996.

RITLAND, K. **Mark genetic marker analysis program**. Version 3.1. Columbia: University of British Columbia, Software, 2006.

ROGERS, J.S. **Measures of genetic similarity and genetic distance**. Austin: Univ. of Texas, 1972. p. 145–153. (Studies in genetics VII)

SAS INSTITUTE. **Statistical analysis software: SAS 9.2**. Cary, 2008.

VAN RADEN, P.M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, Champaign, v. 91, n. 11, p. 4414-4423, 2008.

_____. Genomic Measures of Relationship and Inbreeding, **Interbull Bull**, Uppsala, v.37, p. 33–36, 2007.

VOS, P.; ROGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRITJERS, A.; POT, J.; PELEMAN, J.; KUIPE, M.; ZABEAU, M. AFLP: a new technic for DNA fingerprint. **Nucleic Acids Research**, London, v.23, n.21, p.4407-4414, 1995.

WANG, J. An estimator for pairwise relatedness using molecular markers. **Genetics**, Berlin, v.160, n.1, p.1203-1215, 2002.

WRIGHT, S. Coefficient of inbreeding and relationship. **The American Naturalist**, Chicago, v.56, n.645, p. 330-338, 1922.

_____. **Evolution and genetics of populations**. Chicago: The Univ. of Chicago Press, 1978. 590p.

3 PREDIÇÃO GENÔMICA DE HÍBRIDOS DE MILHO CONSIDERANDO A INTERAÇÃO GENÓTIPOS X AMBIENTES

Resumo

A predição genômica de híbridos de milho tem sido proposta a fim de maximizar a eficiência dos programas de melhoramento, sendo que apenas os híbridos mais promissores, de acordo com a predição, seriam efetivamente avaliados em campo. A maioria das pesquisas avaliam os genótipos de um determinado cruzamento em alguns ambientes. Contudo, os programas de melhoramento de milho geralmente avaliam inúmeros cruzamentos em vários ambientes a fim de identificar os melhores híbridos para uma determinada região. Neste estudo, 250 híbridos simples de milho foram avaliados para produção de grãos em 13 ambientes a fim de verificar os efeitos do tamanho da população de treinamento (N), número de marcas (N_M), interação híbridos x ambientes ($H \times A$) e da estrutura da população na acurácia de predição (r_{MG}) no modelo de seleção genômica ampla RR-BLUP, utilizando-se 614 marcadores AFLP. Foram realizadas validações cruzadas entre e dentro de ambientes e entre e dentro de grupos de híbridos relacionados. A seleção genômica foi posteriormente comparada com a seleção fenotípica quanto à eficiência em identificar híbridos superiores em um esquema de melhoramento de milho. De forma geral, o aumento de N e N_M não alterou significativamente as estimativas da acurácia de predição. As estimativas da acurácia na validação cruzada dentro de ambientes foram superiores às obtidas entre ambientes, indicando que o efeito da interação $H \times A$ foi significativo. Observou-se aumento expressivo nas estimativas de acurácia para os híbridos provenientes de cruzamentos relacionados. Em todos os casos, as estimativas de acurácia apresentaram amplos intervalos, indicando que a seleção genômica pode não ser eficiente dependendo dos híbridos amostrados. A utilização da seleção genômica nos programas de melhoramento de milho é discutida no sentido de considerar os efeitos da estrutura da população e da interação genótipos x ambientes de forma a maximizar a estimativa da acurácia e, conseqüentemente, o sucesso da predição.

Palavras-chave: Seleção genômica; Milho; Híbrido; Interação genótipos x ambientes; Estrutura populacional

Abstract

Genomewide selection offers great potential for increasing the genetic improvement in plant breeding, since only the hybrids with better performance, according to prediction, would be evaluated in field trials. Most genomic selection application uses a single-environment model to predict progenies from a biparental cross. However, in a hybrid maize breeding program, numerous crosses between inbred lines need to be evaluated in extensive field trials to identify hybrids with greater yield potential in a target environment. In this study, 614 AFLP marker effects estimated in 250 maize hybrids evaluated in 13 environments were used to assess the influence of training population size (N), number of markers (N_M), genotype-by-environment interaction and population structure on the prediction accuracy (r_{MG}) for grain yield using RR-BLUP model. Cross validation analysis were performed within and across environments and within and across groups of related hybrids. Genomic selection and phenotypic selection were compared in effectiveness of identifying superior hybrids in maize breeding program scheme. In general, increasing N and N_M did not lead to higher accuracy estimates. Predicted accuracies of cross validation analysis within environments were higher than between environments, indicating that the effect of $G \times E$ interaction was significant. Greater accuracy estimates were achieved when training and validation set were derived from related hybrids. In all scenarios, wide intervals of accuracy were found, meaning that genomic prediction may not be effective depending on the sample used. Therefore, the use of genomic prediction in maize hybrid breeding is discussed emphasizing a need of a training set evaluated in multiple environments, and design a genomic selection experiment according to the population structure so as to reduce sample problems and maximize accuracy and hence the success of prediction.

Keywords: Genomewide selection, Maize, Hybrid, Genotype-by-environment interaction, Population structure

3.1 Introdução

Métodos de predição podem aumentar consideravelmente a eficiência dos programas de melhoramento de milho. Meuwissen, Hayes e Goddard (2001) propuseram a seleção genômica ampla, que consiste na predição dos valores genéticos dos híbridos através da análise de um grande número de marcadores amplamente distribuídos no genoma. Por essa metodologia, estima-se o valor genotípico dos marcadores e a contribuição de cada um deles para o valor genético dos híbridos. Nesse caso, não há análise de mapeamento ou ligação entre os marcadores e a característica. Assume-se que todos os marcadores contribuem para a expressão do caráter, alguns com maiores efeitos, outros com efeitos próximos de zero. Entretanto, mesmo os marcadores com pequeno efeito são considerados na predição, o que faz com que a seleção genômica ampla seja o método mais indicado para a predição de caracteres quantitativos.

A seleção genômica pode ser realizada utilizando diferentes abordagens estatísticas que vão desde métodos de regressão linear, modelos bayesianos ou regressão implícita. De acordo com Bernardo (2010), o modelo misto de regressão linear RR-BLUP possui pressuposições mais simples, porém biologicamente incorretas, pois assume que os marcadores explicam a mesma proporção da variância genética e ignoram a presença de epistasia. Já os modelos bayesianos assumem uma distribuição a priori para os efeitos dos marcadores ou para suas variâncias. Assim, é esperado que eles sejam mais eficientes. Entretanto, resultados empíricos mostram que, para o contexto do melhoramento vegetal, os métodos bayesianos não apresentam acurácia significativamente superior à predita pelo RR-BLUP, sendo este mais vantajoso pela sua simplicidade e rapidez (LORENZANA; BERNARDO, 2009).

Embora estas abordagens sejam mais aplicadas ao melhoramento animal, crescente literatura baseada em dados simulados e empíricos comprovam a sua importância no melhoramento de plantas (BERNARDO; YU, 2007; CROSSA et al., 2010; HICKEY et al., 2014). Em milho, foram observados valores significativos de acurácia na predição de genótipos provenientes do mesmo cruzamento (LORENZANA; BERNARDO, 2009; MASSMAN et al., 2013; JACOBSON et al., 2014). Nesse caso, dados fenotípicos e moleculares de genótipos de um determinado cruzamento são utilizados para obter os efeitos genotípicos dos marcadores que, posteriormente, são utilizados na predição de genótipos não avaliados desse mesmo cruzamento em um ou mais ciclos de seleção. Estudos envolvendo híbridos não relacionados, entretanto,

apresentaram expressiva queda nos valores de acurácia (ALBRECH et al., 2011; ALBRECH et al., 2014; WINDHAUSEN et al., 2012).

Na maioria dos estudos, os híbridos são avaliados dividindo-se a amostra em uma população de treinamento e uma população de validação. Os dados fenotípicos e moleculares da população de treinamento são utilizados na estimação dos efeitos genéticos das marcas que, posteriormente, são utilizados para prever as médias dos híbridos da população de validação. As médias preditas são então correlacionadas com as médias fenotípicas obtenção da acurácia de predição (GODDARD; HAYES, 2007). Contudo, a aplicação prática da seleção genômica visa prever os valores genotípicos em ambientes não avaliados. Como as populações de validação e treinamento são avaliadas no mesmo ambiente, o efeito da interação híbridos x ambientes (H x A) não é considerado e a acurácia de predição pode estar superestimada. A interação H x A é particularmente importante no melhoramento de plantas. Os genótipos são rotineiramente avaliados em diferentes locais e anos e a interação é quase sempre presente e substancial (HALLAUER; CARENA; MIRANDA FILHO, 2010).

Apesar dos estudos até o momento apresentarem valores expressivos de acurácia, uma situação real dos programas de melhoramento de milho avalia híbridos procedentes de diferentes populações e em diferentes ambientes. Assim, são necessários experimentos para verificar se os valores de acurácia obtidos até o momento podem ser aplicados com o mesmo sucesso em populações e ambientes diferentes daqueles em que os efeitos dos marcadores são estimados. Os objetivos desse trabalho foram (1) determinar a acurácia da seleção genômica em prever a performance de híbridos de milho (a) utilizando diferentes tamanhos de população de treinamento e considerando desbalanceamento de marcas; (b) considerando a interação genótipos x ambientes; (c) utilizando populações de treinamento e validação compostas por híbridos não relacionados; e (2) comparar a eficiência da seleção fenotípica e da seleção genômica na seleção de híbridos superiores em programas de melhoramento de milho.

3.2 Material e Métodos

3.2.1 Dados fenotípicos e moleculares dos híbridos

Foram realizados cruzamentos entre cinco linhagens testadoras e 50 linhagens endogâmicas para obtenção dos 250 híbridos simples de milho. As 50 linhagens endogâmicas são provenientes de diferentes populações (IG-1, IG-2 e CMS-05), híbridos comerciais (HS-1, XL-560, BR-201, C-606 e C-601) e progênies de meios-irmãos interpulacionais (BR-105 e BR-106). As cinco linhagens previamente selecionadas como testadoras também são endogâmicas, provenientes de diferentes populações e possuem germoplasma elite (AGUIAR et al., 2003): L-08-05F (Testador 1), L-38-05D (Testador 2), L-36-07F (Testador 3), L-49-02D (Testador 4) e L-46-10D (Testador 5). A escolha das 55 linhagens objetivou representar a variabilidade existente em um programa de melhoramento de milho, em que os híbridos avaliados são provenientes de múltiplos cruzamentos de origens diferentes. Todas as linhagens endogâmicas e os 250 híbridos foram obtidos no Departamento de Genética pelo Programa de Melhoramento de Milho da ESALQ/USP, em Piracicaba - SP. As origens das 50 linhagens endogâmicas bem como as origens e características das cinco linhagens testadoras estão apresentadas no Anexo A.

Os híbridos foram avaliados para produção de grãos em 13 ambientes, sendo cada ambiente uma combinação local/ano: Estação Experimental Fazenda Areão (E.E. Areão), Estação Experimental Fazenda Caterpillar (E.E. Caterpillar) e Estação Experimental do Departamento de Genética da ESALQ/USP (E.E. Departamento de Genética.), nos anos agrícolas de 2002/2003, 2003/2004 e 2004/2005; Estação Experimental Fazenda Anhembi (E.E. Anhembi), nos anos agrícolas 2003/2004 e 2004/2005; e Estação Experimental da Empresa Biomatrix em Patos de Minas-MG (E.E. Patos de Minas), no ano agrícola de 2004/2005 (dois experimentos). Utilizou-se o delineamento experimental látice simples 16 x 16, composto pelos 250 híbridos e seis testemunhas: os híbridos comerciais Z-8420, Z-8550, Z-8460 e Z-8480, da empresa Zeneca Sementes, e os híbridos BRS-2020 e BRS-1010 da EMBRAPA. Foram utilizadas parcelas de uma linha de 4m espaçadas 0,80m entre linhas e 0,20m entre plantas, e estande de 20 plantas por parcela, correspondendo a 62.500 plantas por hectare. A produção de grãos de cada parcela foi corrigida para teor de umidade de 15% e para o estande e, posteriormente, os dados obtidos em kg parcela⁻¹ foram convertidos para t ha⁻¹.

A genotipagem das linhagens endogâmicas foi realizada pelo Centro de Biologia Molecular e Engenharia Genética da Universidade Estadual de Campinas (CBMEG/UNICAMP). Para isso, foram feitas extrações de DNA de folhas adultas das 50 linhagens e dos cinco testadores separadamente, de acordo com procedimentos descritos por Hoisington, Khairllah e Gonzales-de-Leon (1994). A genotipagem foi realizada utilizando-se 614 marcadores AFLP segundo protocolo de Vos et al. (1995). Apesar desse ser um marcador dominante, foi possível distinguir os genótipos heterozigotos dos híbridos, pois todas as 50 linhagens e os cinco testadores utilizados como genitores são endogâmicos. Os dados moleculares apresentaram desbalanceamento inferior a 30%.

3.2.2 Análises de variância

Foram realizadas as análises de variância para cada ambiente, considerando todos os efeitos aleatórios, exceto a média, segundo o modelo, eq. (1):

$$y_{ijk} = m + t_i + b_{j(k)} + r_k + e_{ijk} \quad (1)$$

em que:

y_{ijk} é o valor da produção em $t \text{ ha}^{-1}$ do híbrido i no bloco j na repetição k ;

m é a média geral do experimento;

t_i é o efeito do híbrido i ($i = 1, 2, 3, \dots, 256$);

$b_{j(k)}$ é o efeito do bloco j dentro da repetição k ($j = 1, 2, 3, \dots, 16$);

r_k é o efeito da repetição k ($k = 1, 2$);

e_{ijk} é o erro experimental associado à observação y_{ijk} com $e_{ijk} \sim N(0, \sigma^2)$.

Posteriormente foi realizada análise de variância conjunta também considerando todos os efeitos aleatórios, exceto a média, segundo o modelo, eq. (2):

$$y_{ijkl} = m + t_i + a_l + b_{j(kl)} + r_{k(l)} + (ta)_{il} + e_{ijkl} \quad (2)$$

em que:

y_{ijkl} é a produção em $t \text{ ha}^{-1}$ do híbrido i no bloco j na repetição k no ambiente l ;

m é a média geral do experimento;

t_i é o efeito do híbrido i ($i = 1, 2, 3, \dots, 256$);

a_l é o efeito do ambiente l ;

$b_{j(kl)}$ é o efeito do bloco j dentro da repetição k no ambiente l ($j = 1, 2, 3, \dots, 16$);

$r_{k(l)}$ é o efeito da repetição k dentro do ambiente l ($k = 1, 2$);

$(ta)_{il}$ é o efeito da interação híbridos x ambientes;

e_{ijkl} é o erro experimental associado à observação y_{ijkl} com $e_{ijkl} \sim N(0, \sigma^2)$.

Foram realizadas análises conjuntas de todos os ambientes ($l = 13$) e dos ambientes dois a dois ($l = 2$) utilizando-se as médias das análises individuais. Como estas foram obtidas com recuperação da informação do resíduo intrabloco, foi calculado um erro efetivo para cada uma delas, conforme apresentado por Cochran e Cox (1966), eq. (3):

$$QM_{EF} = QM_{EI} \left[1 + \frac{kj\mu}{j+1} \right] \quad (3)$$

em que:

QM_{EF} é o quadrado médio do erro efetivo;

QM_{EI} é o quadrado médio do erro intrabloco;

$QM_{B(R)}$ é o quadrado médio de bloco dentro de repetição; e

$$\mu = \frac{QM_{B/R} - QM_{EI}}{j(k-1)QM_{B(R)}}$$

O erro médio da análise conjunta foi obtido da média dos quadrados médios dos erros efetivos das análises individuais. Todas as análises de variância foram realizadas utilizando-se o software SAS (SAS INSTITUTE, 2008). Os esquemas das análises individuais e conjunta estão apresentados nos Anexos B e C. Os quadrados médios das análises de variância foram iguais

às suas respectivas esperanças para obtenção dos coeficientes de herdabilidade que, posteriormente, foram utilizados no cálculo das acurácias de predição (r_{MG}) da seleção genômica. Os coeficientes de herdabilidades foram calculados no sentido amplo para cada ambiente ($\hat{h}_A^2$) e para os ambientes em conjunto ($\hat{h}_{Conj}^2$) em nível de média dos híbridos de acordo com as equações, eq. (4 e 5):

$$\hat{h}_A^2 = \frac{\sigma_H^2}{\sigma_F^2} = \frac{\sigma_H^2}{\left(\sigma_H^2 + \frac{\sigma^2}{k}\right)} = \frac{(QM_H - QM_{EI})/k}{QM_H/k} \quad (4)$$

$$\hat{h}_{Conj}^2 = \frac{\sigma_H^2}{\sigma_F^2} = \frac{\sigma_H^2}{\left(\sigma_H^2 + \frac{\sigma_{HA}^2}{l} + \frac{\sigma^2}{lk}\right)} = \frac{(QM_H - QM_{HA})/lk}{QM_H/lk} \quad (5)$$

em que σ_H^2 é a variância genética dos híbridos, σ_F^2 é a variância fenotípica média dos híbridos, σ_{HA}^2 é a variância da interação híbridos x ambientes, σ^2 é a variância residual, QM_H é o quadrado médio dos híbridos, QM_{HA} é o quadrado médio da interação híbridos x ambientes. Os intervalos de confiança para as estimativas dos componentes de herdabilidade foram obtidos segundo Knapp, Stroup e Ross (1985), eq. (6 e 7).

$$IC_{(\hat{h}_A^2)_{0,95}} = \left\{ 1 - \frac{1}{\left(\frac{QM_H}{QM_{EI}}\right) F_{0,025;GL_H;GL_{EI}}} \leq \hat{h}_A^2 \leq 1 - \frac{1}{\left(\frac{QM_H}{QM_{EI}}\right) F_{0,975;GL_H;GL_{EI}}} \right\} \quad (6)$$

$$IC_{(\hat{h}_{Conj}^2)_{0,95}} = \left\{ 1 - \frac{1}{\left(\frac{QM_H}{QM_{HA}}\right) F_{0,025;GL_H;GL_{HA}}} \leq \hat{h}_{Conj}^2 \leq 1 - \frac{1}{\left(\frac{QM_H}{QM_{HA}}\right) F_{0,975;GL_H;GL_{HA}}} \right\} \quad (7)$$

em que $IC_{(\hat{h}_A^2)_{0,95}}$ e $IC_{(\hat{h}_{Conj}^2)_{0,95}}$ são os intervalos de confiança para os coeficientes de herdabilidade no sentido amplo em nível de média dos híbridos para cada ambiente ($\hat{h}_A^2$) e para os ambientes em conjunto ($\hat{h}_{Conj}^2$), respectivamente, ao nível de 95% de probabilidade.

$F_{0,975;GL_H;GL_{EI}}$ e $F_{0,975;GL_H;GL_{HA}}$ referem-se aos valores tabelados de F a 97,5% de probabilidade, com os graus de liberdade de híbridos (GL_H), do erro intrabloco (GL_{EI}) e da interação híbridos x ambientes (GL_{HA}). $F_{0,025;GL_H;GL_{EI}}$ e $F_{0,025;GL_H;GL_{HA}}$ referem-se aos valores tabelados de F a 2,5% de probabilidade, com os mesmos graus de liberdade. Os intervalos de confiança foram utilizados na comparação das estimativas dos coeficientes de herdabilidade, sendo que duas estimativas foram consideradas diferentes entre si ($P \leq 0,05$) quando uma estimativa não estiver dentro do intervalo de confiança da outra.

3.2.3 Modelo RR-BLUP de seleção genômica ampla para predição dos híbridos de milho

O modelo de seleção genômica ampla para predição de médias utilizado foi o modelo misto RR-BLUP descrito por Meuwissen, Hayes e Goddard (2001). Esse modelo calcula os efeitos das marcas por regressão (*ridge-regression*) e assume que os marcadores explicam a mesma proporção da variância genética.

Sendo N o número de híbridos e N_M o número de marcas, o modelo linear misto RR-BLUP é, (eq. 8):

$$\mathbf{y} = \mu \mathbf{1} + \mathbf{X}\mathbf{g} + \mathbf{e} \quad (8)$$

em que:

$\mathbf{y}$ é o vetor das médias fenotípicas dos híbridos, de dimensão $N \times 1$;

μ é a média fenotípica geral dos híbridos;

$\mathbf{1}$ é o vetor de uns que relaciona a média ao vetor $\mathbf{y}$, de dimensão $N \times 1$;

$\mathbf{X}$ é a matriz de incidência relacionada aos efeitos aleatórios $\mathbf{g}$, de dimensão $N \times N_M$;

$\mathbf{g}$ é o vetor dos efeitos aleatórios das marcas, de dimensão $N_M \times 1$;

$\mathbf{e}$ é o vetor dos efeitos aleatórios dos resíduos, de dimensão $N \times 1$.

A matriz de incidência $\mathbf{X}$ é composta pelos valores das marcas 1, 0 e -1 que se referem aos genótipos homozigoto dominante, heterozigoto e homozigoto recessivo dos locos,

respectivamente. As equações de modelos mistos que resolvem esse sistema matricial são, (eq. 9):

$$\begin{bmatrix} \hat{\mu} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} \mathbf{1}' \mathbf{1} & \mathbf{1}' \mathbf{X} \\ \mathbf{X}' \mathbf{1} & \mathbf{X}' \mathbf{X} + \mathbf{I} [(\sigma_F^2 / \sigma_{M_i}^2)] \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}' \mathbf{y} \\ \mathbf{X}' \mathbf{y} \end{bmatrix} \quad (9)$$

em que:

$\hat{\mu}$ é a estimativa da média fenotípica populacional;

$\hat{g}$ é a estimativa dos efeitos genotípicos dos marcadores;

$\mathbf{I}$ é uma matriz identidade, de dimensão $N_M \times N_M$;

σ_F^2 é a variância fenotípica média dos híbridos;

$\sigma_{M_i}^2$ é a variância de cada marcador.

Dado que a variância de cada marcador é desconhecida e sendo σ_H^2 a variância genética entre os híbridos, o RR-BLUP assume que $\sigma_{M_i}^2 = \sigma_H^2 / N_M$, ou seja, cada marcador explica $1/N_M$ da variância genética. Isso não significa, entretanto, que os efeitos preditos dos marcadores serão iguais, apenas que eles possuem a mesma distribuição de variância (BERNARDO, 2010). As médias preditas dos híbridos ($\hat{y}$) foram então obtidas pelo somatório dos efeitos genotípicos estimados dos marcadores ($\hat{g}$), eq. (10):

$$\hat{y} = \hat{\mu} \mathbf{1} + \mathbf{X} \hat{g} \quad (10)$$

sendo as matrizes $\mathbf{1}$ e $\mathbf{X}$ as mesmas descritas para eq. (8). As operações matriciais das análises de seleção genômica foram realizadas utilizando-se a função “mixed.solve” do pacote rrBLUP (ENDELMAN, 2011) do software R (R DEVELOPMENT CORE TEAM, 2012).

3.2.4 Validação cruzada do modelo de seleção genômica: efeito do número de marcas e do tamanho da população

Inicialmente as análises de seleção genômica foram feitas utilizando-se as médias dos híbridos nos 13 ambientes considerando diferentes números de marcas (N_M) e diferentes

tamanhos da população de treinamento (N). Alguns genótipos dos 614 marcadores não foram obtidos para algumas linhagens na genotipagem. Por isso, foi realizada a estimação das marcas não observadas utilizando-se a função “A.mat” do pacote rrBLUP (ENDELMAN, 2011). O método de estimação utilizado foi “impute.method=mean”, em que a estimação do genótipo não observado de uma marca é a média dos genótipos observados para essa marca nos demais híbridos. Assim, foram considerados $N_M = 124$, em que somente foram utilizadas as marcas que estavam completas para todos os híbridos; $N_M = 418$, em que 10% das marcas foram estimadas; $N_M = 537$, em que 20% das marcas foram estimadas; e $N_M = 614$, em que todas as marcas não observadas foram estimadas, que correspondem a cerca de 30% do total de marcas.

A validação cruzada do modelo foi feita dividindo-se os 250 híbridos em duas populações: uma população de treinamento (N), da qual foram utilizados os dados fenotípicos e moleculares para estimar os efeitos genotípicos das marcas; e uma população de validação de tamanho $250 - N$, da qual foram preditas as médias. Foram utilizados quatro tamanhos de população de treinamento: $N = 50$, $N = 100$, $N = 150$ e $N = 200$. Também utilizou-se $N = 250$, ou seja, não houve separação das populações de treinamento e validação, e todos os híbridos foram utilizados para estimar os efeitos genotípicos das marcas e prever as médias. A acurácia de predição (r_{MG}) foi estimada pela correlação entre as médias preditas e as médias fenotípicas da população de validação, dividida pela raiz quadrada da herdabilidade no sentido amplo em nível de média para os 13 ambientes (h_{Conj}^2), (DEKKERS, 2007). Para composição das populações de treinamento e validação foram realizadas 500 amostragens aleatórias dos híbridos e, posteriormente, foram obtidos a média e o intervalo das estimativas de r_{MG} com 95% de probabilidade, sendo que duas estimativas foram consideradas estatisticamente diferentes entre si ($P \leq 0,05$) somente quando uma estimativa não estiver dentro do intervalo de confiança da outra. As amostragens dos híbridos foram realizadas utilizando-se o software R (R DEVELOPMENT CORE TEAM, 2012).

3.2.5 Validação cruzada entre e dentro de ambientes: efeito da interação H x A

Para verificar o efeito da interação híbridos por ambientes (H x A) na estimativa de r_{MG} , utilizaram-se as médias fenotípicas dos híbridos em cada ambiente, ou seja, não ajustadas para o efeito de ambiente e da interação H x A. A validação cruzada foi feita dentro de ambientes e entre

os ambientes dois a dois. As r_{MG} entre ambientes foram estimadas correlacionando-se as médias preditas da população de validação em um ambiente com as médias fenotípicas da mesma população de validação em outro ambiente, dividindo-se pela raiz quadrada do coeficiente de herdabilidade (h_A^2) do ambiente de treinamento. Assumiu-se $N_M = 614$ e $N = 150$ e, da mesma forma, foram obtidos a média e o intervalo das estimativas de r_{MG} nos 500 ciclos de amostragem.

3.2.6 Validação cruzada entre e dentro de grupos de híbridos relacionados: efeito da estrutura da população

Os 250 híbridos simples foram separados em cinco grupos de 50 híbridos de acordo com os cinco testadores utilizados no cruzamento. Os grupos foram denominados G1, G2, G3, G4 e G5, em que os híbridos pertencentes ao mesmo grupo possuem o testador como genitor comum, sendo, portanto, geneticamente relacionados. As validações cruzadas foram realizadas entre e dentro de grupos, assumindo-se $N_M = 614$. Dentro dos grupos, todos os 50 híbridos foram utilizados para estimar os efeitos genotípicos das marcas e prever as médias, não havendo separação das populações de treinamento e validação. Entre os grupos, as validações foram realizadas utilizando-se os 50 híbridos de um grupo como população de treinamento, e os 50 híbridos de outro grupo como população de validação, todas as combinações dos grupos dois a dois. Para o cálculo das estimativas de r_{MG} , utilizou-se a herdabilidade no sentido amplo em nível de média dos híbridos para os 13 ambientes (h_{Conj}^2).

3.2.7 Seleção fenotípica e genômica de híbridos superiores

Nos programas de melhoramento, os híbridos geralmente são avaliados em experimentos com repetições em múltiplos ambientes, e aqueles que apresentam performance superior são selecionados para serem liberados nos anos subsequentes. A interação entre híbridos e ambientes ($H \times A$) que ocorre nesse caso é consequência apenas do efeito de ano agrícola, e os híbridos são, então, recomendados para regiões específicas. Para verificar a eficiência da predição genômica em identificar híbridos superiores, foi proposto um esquema de seleção semelhante ao que é feito pelos programas de melhoramento de milho. Foram realizadas as seleções fenotípica (SF) e genômica (SG) dos 25 híbridos superiores avaliados nos anos agrícolas 2002/2003 e 2003/2004

nos locais E.E. Areão, E.E. Caterpillar, E.E. Departamento de Genética e E.E. Anhembi. A seleção genômica foi realizada utilizando-se as médias fenotípicas dos 250 híbridos conjuntamente às informações moleculares no modelo de predição genômica RR-BLUP apresentado anteriormente, assumindo-se $N_M = 614$. Posteriormente realizou-se a SF dos 25 híbridos superiores no ano agrícola 2004/2005 avaliados nos mesmos locais. Foi estimado o coeficiente de correlação e a porcentagem de coincidência dos híbridos selecionados via SF e SG nos anos agrícolas 2002/2003-2003/2004, e destes com os híbridos selecionados fenotipicamente no ano agrícola 2004/2005. Assim, foi possível observar se os híbridos selecionados pelos dois métodos mantiveram a performance no ano agrícola subsequente. O esquema das seleções genômica e fenotípica está apresentado na Figura 1.

Figura 1 - Esquema das seleções fenotípica e genômica nos locais avaliados nos anos agrícolas 2002/2003 e 2003/2004 e posterior correlação (r) com os híbridos selecionados fenotipicamente no ano agrícola 2004/2005. As siglas referem-se: ARE: E.E. Areão; CAT: E.E. Caterpillar; DEP: E.E. Departamento de Genética; ANH: E.E. Anhembi; 0203: ano agrícola 2002/2003; 0304: ano agrícola 2003/2004; 0405: ano agrícola 2004/2005

3.3 Resultados e Discussão

3.3.1 Análises de variâncias

Foram observadas diferenças significativas entre os híbridos ($p \leq 0,01$) nas análises de variâncias individuais em todos os ambientes. As médias da produção de grãos variaram de 6,75 t ha⁻¹ na E.E. Caterpillar - 2002/2003, a 10,32 t ha⁻¹ na E.E. Departamento de Genética - 2004/2005. Os coeficientes de variação (CV%) variaram de 9,65% na E.E. Areão - 2003/2004, a 22,36% na E.E. Areão - 2004/2005. Somente em dois ambientes as estimativas do coeficiente de variação foram superiores à 20% em função da ocorrência de estresse hídrico após o florescimento dos híbridos nesses ambientes (20,31% na E.E. Caterpillar - 2004/2005 e 22,36% na E.E. Areão - 2004/2005). Nos demais ambientes, as estimativas foram sempre inferiores a 15%. Os valores de CV(%) obtidos estão dentro dos limites reportados na literatura (AGUIAR et al., 2003; RAPOSO, 2002). Os coeficientes de herdabilidade ampla em nível de média dos híbridos para cada ambiente (h_A^2) variaram de 0,63, com intervalo de confiança de 0,53 - 0,71, na E.E. Caterpillar - 2004/2005, a 0,81 com intervalo de confiança variando de 0,76 - 0,85 na E.E. Departamento de Genética - 2002/2003, portanto, diferentes estatisticamente (Anexo D).

Também foram observadas diferenças significativas ($p \leq 0,01$) na análise de variância conjunta dos 13 ambientes para as fontes de variação híbridos, ambientes e para a interação híbridos x ambientes (H x A). Esses resultados mostram que há variabilidade genética entre os híbridos e que os híbridos apresentaram desempenho diferencial entre os ambientes em que foram avaliados. A média de produção de grãos dos híbridos na análise conjunta foi de 8,55 t ha⁻¹, e o coeficiente de variação (CV%) foi de 13,21%. O coeficiente de herdabilidade ampla em nível de média dos híbridos na análise conjunta (h_{Conj}^2) foi de 0,95, com intervalo de confiança de 0,94 - 0,96 (Anexo E).

Todas as análises de variância conjunta dos ambientes dois a dois apresentaram diferenças significativas entre os híbridos. Apenas duas das 78 combinações dos ambientes dois a dois não apresentaram diferenças significativas entre os ambientes. Esse resultado indica que os ambientes apresentam diferenças edafo-climáticas que podem influenciar na produção dos híbridos. O efeito da interação H x A foi significativo em 73 das 78 combinações dos ambientes dois a dois, reforçando que há performance diferencial dos híbridos nos diferentes ambientes (Anexo F).

3.3.2 Validação cruzada do modelo de seleção genômica: efeito do número de marcas e do tamanho da população

As médias das acurácias de predição (r_{MG}) obtidas nas análises de seleção genômica para os híbridos nos 13 ambientes considerando diferentes números de marcas (N_M) e diferentes tamanhos da população de treinamento (N) variaram de 0,51, para $N = 50$ e $N_M = 124$, a 0,82, para $N = 250$ e $N_M = 614$. Assim, a menor estimativa média de r_{MG} foi obtida para os menores valores de N e N_M e, da mesma forma, a maior estimativa média de r_{MG} foi obtida para N e N_M maiores (Tabela 1).

De forma geral, para os valores médios das acurácias, foi possível observar que o aumento de N teve maior influência na r_{MG} média que o aumento de N_M . Também observou-se que o incremento na estimativa de r_{MG} com o aumento de N_M foi mais expressivo nas análises em que N também foi maior. Para $N = 50$, o aumento de $N_M = 124$ para $N_M = 614$ levou a um incremento de 0,01 unidades na estimativa média de r_{MG} , enquanto que para $N = 250$, o aumento de $N_M = 124$ para $N_M = 614$ levou a um incremento de 0,08 unidades na estimativa média de r_{MG} , indicando que há uma relação positiva entre o número de híbridos na população de treinamento e o número de marcas utilizados na obtenção do modelo de seleção genômica (Tabela 1).

Tabela 1 – Médias e intervalos (entre parênteses) das estimativas de acurácia de predição (r_{MG}) dos 500 ciclos de amostragem dos híbridos nas análises de seleção genômica considerando diferentes números de marcas (N_M) e tamanhos de populações de treinamento (N)

N	N_M			
	614	537	418	124
50	0,54 (0,41 – 0,62)	0,52 (0,38 – 0,60)	0,51 (0,37 – 0,60)	0,53 (0,38 – 0,62)
100	0,61 (0,52 – 0,68)	0,59 (0,49 – 0,66)	0,57 (0,49 – 0,65)	0,59 (0,52 – 0,67)
150	0,64 (0,52 – 0,73)	0,61 (0,51 – 0,69)	0,60 (0,49 – 0,69)	0,61 (0,50 – 0,70)
200	0,66 (0,51 – 0,79)	0,63 (0,48 – 0,75)	0,61 (0,45 – 0,76)	0,64 (0,48 – 0,77)
250	0,82	0,79	0,76	0,74

Esses resultados podem indicar que quanto maior N e N_M , maior a estimativa de r_{MG} . Contudo, observando os intervalos das estimativas de r_{MG} obtidos pelos ciclos de amostragens dos híbridos, foi possível observar que o aumento de N e N_M não influenciou significativamente as estimativas de r_{MG} . Somente para $N = 250$ e $N_M = 614$ a estimativa média de r_{MG} ($r_{MG} = 0,82$) foi maior que os limites superiores dos demais intervalos. Nas demais combinações de N e N_M ,

quando se analisam somente os intervalos, os mesmos se sobrepõem e, portanto, não se diferem significativamente. É importante salientar que no caso de $N = 250$ não houve separação das populações de treinamento e validação, e todos os híbridos foram utilizados na predição dos efeitos genotípicos das marcas. Na prática, a seleção genômica é realizada utilizando-se apenas uma parte dos híbridos na população de treinamento e, por isso, são esperadas estimativas de r_{MG} semelhantes às obtidas para os demais tamanhos de N .

Mesmo os híbridos utilizados neste estudo sendo provenientes de cruzamentos entre linhagens de diferentes origens e a produção de grãos em milho um caráter quantitativo complexo, foi possível observar valores de r_{MG} de alta magnitude comparados com a literatura (LORENZANA; BERNARDO, 2009; MASSMAN et al., 2013; CROSSA et al., 2014; JACOBSON et al., 2014; LIAN et al., 2014). É interessante observar, contudo, a amplitude dos intervalos das estimativas de r_{MG} . Esses intervalos foram obtidos à partir dos ciclos de amostragem dos híbridos, ou seja, as populações de treinamento e validação foram amostradas 500 vezes, sendo obtida uma estimativa de r_{MG} para cada amostra. Observou-se que os valores de r_{MG} podem variar muito em função da amostra. Para $N = 200$ e $N_M = 418$, por exemplo, o intervalo das estimativas de acurácia variou de 0,45 a 0,76, originando uma amplitude de 0,31 unidades. Como a amostragem dos híbridos foi aleatória, esse resultado indica que dependendo da população amostrada a seleção genômica pode ou não ser eficiente em prever as performances dos híbridos.

Resultados semelhantes foram obtidos por Lian et al. (2014). Esses autores avaliaram os efeitos de N e N_M na estimativa da acurácia de predição de populações de milho utilizando o modelo de predição genômica RR-BLUP. Eles utilizaram $N = 60$, $N = 90$, $N = 120$, $N = 150$, $N = 180$, e $N_M = 36$, $N_M = 72$ e $N_M = 107$. Observou-se que o aumento de N levou a diferenças significativas das estimativas de r_{MG} , exceto para $N = 150$ e $N = 180$. O aumento de N_M , contudo, não alterou significativamente as estimativas de r_{MG} e, da mesma forma, foram observados amplos intervalos para as estimativas de r_{MG} . Para produção de grãos, a estimativa de r_{MG} média foi 0,45, variando de -0,59 a 1,03. Também foram observados amplos intervalos das estimativas de acurácia para os caracteres umidade (-0,34 a 0,96), acamamento (-0,30 a 0,93), quebramento (-0,22 a 1,04), altura de plantas (-0,45 a 0,97) e altura de espiga (-0,43 a 0,94). Os autores discutem que, mesmo para os caracteres que apresentam maior coeficiente de herdabilidade, a amplitude

dos intervalos das estimativas de r_{MG} é expressiva e, portanto, deve-se ter cuidado com a conclusão dos resultados de r_{MG} .

De acordo com Lorenzana e Bernardo (2009), menos de 100 marcadores devem ser suficientes para realizar a seleção genômica em genótipos derivados de um cruzamento em milho, pois a acurácia permanece relativamente estável com a alteração de N_M . Em estudo envolvendo a simulação de genótipos F_2 para um determinado cruzamento, Hickey et al. (2014) mostraram que para obtenção de valores razoáveis de r_{MG} são necessários cerca de $N = 200-1.000$ e $N_M = 300-500$ quando as populações de treinamento e validação são provenientes do mesmo cruzamento ou de cruzamentos relacionados, e $N = 5.000-20.000$ e $N_M = 10.000$ quando as populações são provenientes de cruzamentos não relacionados. Os autores mostraram ainda que aumento de N_M acima de 10.000 marcadores não beneficiou e até reduziu a estimativa de r_{MG} em alguns cenários. Em função dos valores médios de r_{MG} , da amplitude dos intervalos observados e da estrutura da população deste estudo, considerou-se razoável utilizar $N = 150$ e $N_M = 614$ nas análises de validação cruzada entre ambientes.

3.3.3 Validação cruzada entre e dentro de ambientes: efeito da interação H x A

As estimativas médias de acurácia (r_{MG}) das análises de validação cruzada dentro dos ambientes variaram de 0,49 para o ambiente E.E. Departamento de Genética - 2004/2005, a 0,72 para os ambientes E.E. Anhembi - 2003/2004 e E.E. Patos de Minas - 2004/2005. As análises de validação cruzada entre os ambientes dois a dois apresentaram estimativas médias de r_{MG} variando de 0,22 entre os ambientes E.E. Departamento de Genética - 2003/2004 e E.E. Areão - 2004/2005, a 0,75 entre os ambientes E.E. Anhembi - 2003/2004 e E.E. Anhembi - 2004/2005. O ambiente que apresentou menor estimativa de r_{MG} em média com os demais foi E.E. Departamento de Genética - 2003/2004 ($r_{MG} = 0,42$), e o ambiente E.E. Patos de Minas - 2004/2005 foi o que apresentou a maior estimativa de r_{MG} em média com os demais ($r_{MG} = 0,56$). A média geral das estimativas de r_{MG} foi de 0,63 dentro de ambientes e de 0,48 entre ambientes (Tabela 2).

Assim como na validação cruzada utilizando as médias dos 13 ambientes, as validações entre e dentro de ambientes também apresentaram amplos intervalos das estimativas de r_{MG} em função da amostragem dos híbridos. Foi possível observar que os intervalos das estimativas de

r_{MG} entre ambientes foram bem mais amplos que os intervalos das estimativas de r_{MG} dentro de ambientes. Além disso, nesse caso foram observados intervalos que não se sobrepuseram, indicando que há diferenças significativas entre as r_{MG} dos ambientes dois a dois, ou seja, existem combinações de ambientes de treinamento e validação que foram mais eficientes em prever as médias dos híbridos que outras. Quando se utilizou o ambiente E.E. Areão - 2004/2005 como ambiente de treinamento, por exemplo, foi obtido um intervalo para as estimativas de r_{MG} variando de 0,08 a 0,49 para o ambiente de validação E.E. Departamento de Genética - 2002/2003, e um intervalo de r_{MG} variando de 0,54 a 0,78 para o ambiente de validação E.E. Caterpillar - 2004/2005. Assim, quando os efeitos dos marcadores foram estimados no ambiente E.E. Areão - 2004/2005, a acurácia de predição foi significativamente maior quando as médias foram preditas no ambiente de validação E.E. Caterpillar - 2004/2005 comprado ao ambiente E.E. Departamento de Genética - 2002/2003. Estimativas significativas de r_{MG} também foram obtidas mantendo-se fixos os ambientes de treinamento E.E. Anhembi - 2004/2005, E.E. Caterpillar - 2004/2005, E.E. Departamento de Genética - 2003/2004 e E.E. Patos de Minas - 2004/2005. Para todos os outros ambientes, mantendo-se o ambiente de treinamento constante, os intervalos das estimativas de r_{MG} se sobrepuseram, ou seja, o ambiente no qual foi feita a validação não alterou significativamente as estimativas de r_{MG} (Tabela 2).

Tabela 2 - Médias e intervalos (entre parênteses) das estimativas de acurácia de predição (r_{MG}) dos 500 ciclos de amostragem das análises de validação cruzada entre os ambientes dois a dois

Amb Valid*	Ambiente de Treinamento					
	1	2	3	4	5	6
1	0,72 (0,59 – 0,82)	0,75 (0,62 – 0,88)	0,47 (0,31 – 0,60)	0,43 (0,27 – 0,58)	0,59 (0,44 – 0,73)	0,46 (0,31 – 0,60)
2	0,64 (0,49 – 0,76)	0,68 (0,53 – 0,81)	0,50 (0,37 – 0,64)	0,48 (0,31 – 0,61)	0,50 (0,30 – 0,67)	0,46 (0,32 – 0,62)
3	0,42 (0,22 – 0,59)	0,53 (0,31 – 0,69)	0,52 (0,38 – 0,64)	0,53 (0,38 – 0,64)	0,43 (0,22 – 0,60)	0,51 (0,31 – 0,67)
4	0,40 (0,19 – 0,57)	0,52 (0,32 – 0,70)	0,56 (0,41 – 0,69)	0,62 (0,49 – 0,73)	0,40 (0,20 – 0,59)	0,54 (0,38 – 0,70)
5	0,55 (0,36 – 0,69)	0,57 (0,38 – 0,73)	0,45 (0,28 – 0,60)	0,40 (0,23 – 0,54)	0,71 (0,57 – 0,82)	0,52 (0,35 – 0,68)
6	0,45 (0,26 – 0,61)	0,53 (0,33 – 0,69)	0,56 (0,41 – 0,67)	0,57 (0,44 – 0,70)	0,55 (0,36 – 0,71)	0,63 (0,49 – 0,74)
7	0,50 (0,34 – 0,65)	0,58 (0,37 – 0,75)	0,50 (0,34 – 0,63)	0,55 (0,41 – 0,68)	0,54 (0,36 – 0,70)	0,58 (0,43 – 0,70)
8	0,51 (0,35 – 0,65)	0,54 (0,38 – 0,70)	0,43 (0,29 – 0,58)	0,40 (0,21 – 0,54)	0,66 (0,54 – 0,78)	0,52 (0,36 – 0,67)
9	0,36 (0,16 – 0,55)	0,43 (0,23 – 0,60)	0,50 (0,32 – 0,64)	0,54 (0,41 – 0,67)	0,29 (0,08 – 0,49)	0,45 (0,27 – 0,61)
10	0,30 (0,10 – 0,49)	0,35 (0,13 – 0,57)	0,42 (0,23 – 0,59)	0,50 (0,34 – 0,63)	0,22 (0,008 – 0,41)	0,38 (0,19 – 0,54)
11	0,34 (0,13 – 0,51)	0,47 (0,26 – 0,64)	0,43 (0,25 – 0,58)	0,45 (0,30 – 0,60)	0,23 (0,01 – 0,41)	0,39 (0,20 – 0,54)
12	0,36 (0,14 – 0,53)	0,48 (0,28 – 0,66)	0,54 (0,39 – 0,67)	0,61 (0,49 – 0,73)	0,36 (0,16 – 0,56)	0,52 (0,36 – 0,67)
13	0,41 (0,20 – 0,59)	0,54 (0,33 – 0,73)	0,55 (0,38 – 0,68)	0,62 (0,50 – 0,73)	0,41 (0,23 – 0,61)	0,57 (0,39 – 0,70)

Tabela 2 - Médias e intervalos (entre parênteses) das estimativas de acurácia de predição (r_{MG}) dos 500 ciclos de amostragem das análises de validação cruzada entre os ambientes dois a dois

Amb Valid*	Ambiente de Treinamento							(conclusão)
	7	8	9	10	11	12	13	
1	0,55 (0,40 – 0,69)	0,59 (0,39 – 0,76)	0,40 (0,22 – 0,55)	0,36 (0,12 – 0,53)	0,41 (0,21 – 0,58)	0,38 (0,23 – 0,53)	0,46 (0,29 – 0,60)	
2	0,54 (0,41 – 0,68)	0,52 (0,30 – 0,69)	0,41 (0,24 – 0,55)	0,35 (0,16 – 0,53)	0,48 (0,31 – 0,64)	0,44 (0,27 – 0,60)	0,53 (0,40 – 0,67)	
3	0,49 (0,33 – 0,62)	0,45 (0,22 – 0,64)	0,48 (0,35 – 0,61)	0,44 (0,28 – 0,56)	0,48 (0,33 – 0,62)	0,49 (0,37 – 0,62)	0,54 (0,38 – 0,69)	
4	0,57 (0,38 – 0,71)	0,41 (0,17 – 0,61)	0,55 (0,40 – 0,68)	0,53 (0,37 – 0,66)	0,51 (0,34 – 0,66)	0,61 (0,47 – 0,73)	0,65 (0,52 – 0,78)	
5	0,54 (0,36 – 0,69)	0,70 (0,52 – 0,84)	0,31 (0,13 – 0,48)	0,24 (0,05 – 0,41)	0,26 (0,06 – 0,45)	0,37 (0,21 – 0,52)	0,46 (0,29 – 0,60)	
6	0,61 (0,46 – 0,72)	0,57 (0,36 – 0,75)	0,46 (0,31 – 0,60)	0,41 (0,22 – 0,57)	0,46 (0,26 – 0,62)	0,55 (0,41 – 0,66)	0,61 (0,48 – 0,76)	
7	0,62 (0,48 – 0,74)	0,59 (0,40 – 0,75)	0,44 (0,28 – 0,57)	0,41 (0,24 – 0,57)	0,45 (0,30 – 0,60)	0,56 (0,43 – 0,69)	0,64 (0,49 – 0,76)	
8	0,57 (0,42 – 0,70)	0,69 (0,52 – 0,81)	0,36 (0,18 – 0,53)	0,29 (0,10 – 0,45)	0,31 (0,12 – 0,47)	0,41 (0,24 – 0,55)	0,48 (0,33 – 0,64)	
9	0,44 (0,26 – 0,59)	0,36 (0,13 – 0,56)	0,60 (0,48 – 0,71)	0,57 (0,45 – 0,70)	0,53 (0,36 – 0,66)	0,53 (0,39 – 0,65)	0,54 (0,38 – 0,70)	
10	0,40 (0,22 – 0,55)	0,27 (0,05 – 0,50)	0,55 (0,40 – 0,67)	0,57 (0,45 – 0,70)	0,49 (0,31 – 0,65)	0,47 (0,32 – 0,62)	0,49 (0,32 – 0,64)	
11	0,41 (0,22 – 0,57)	0,28 (0,05 – 0,49)	0,48 (0,34 – 0,61)	0,47 (0,31 – 0,60)	0,49 (0,31 – 0,63)	0,47 (0,33 – 0,60)	0,48 (0,33 – 0,63)	
12	0,58 (0,42 – 0,71)	0,44 (0,22 – 0,62)	0,55 (0,41 – 0,67)	0,52 (0,33 – 0,66)	0,55 (0,40 – 0,68)	0,67 (0,54 – 0,77)	0,71 (0,56 – 0,83)	
13	0,62 (0,46 – 0,74)	0,49 (0,26 – 0,69)	0,53 (0,38 – 0,66)	0,50 (0,35 – 0,64)	0,53 (0,34 – 0,66)	0,67 (0,54 – 0,78)	0,72 (0,60 – 0,83)	

*Ambiente de Validação. Amb 1: E.E. Anhembi - 2003/2004; Amb 2: E.E. Anhembi - 2004/2005; Amb 3: E.E. Areão - 2002/2003; Amb 4: E.E. Areão - 2003/2004; Amb 5: E.E. Areão - 2004/2005; Amb 6: E.E. Caterpillar - 2002/2003; Amb 7: E.E. Caterpillar - 2003/2004; Amb 8: E.E. Caterpillar - 2004/2005; Amb 9: E.E. Departamento de Genética - 2002/2003; Amb 10: E.E. Departamento de Genética - 2003/2004; Amb 11: E.E. Departamento de Genética - 2004/2005; Amb 12: exp 1 da E.E. Patos de Minas - 2004/2005; Amb 13: exp 2 da E.E. Patos de Minas - 2004/2005.

Estes resultados indicam que, da mesma forma que a seleção fenotípica, a seleção genômica não está livre do efeito da interação híbridos por ambientes (H x A). A interação H x A ocorre sempre que a resposta fenotípica difere em função das variações ambientais, ou seja, quando os híbridos apresentam comportamento diferencial entre os ambientes. Isto acontece porque a expressão dos genes que controlam a característica de interesse varia entre os ambientes. Assim, dois ambientes apresentam correlação alta se os genes que expressam a característica forem os mesmos nesses ambientes (FALCONER; MACKAY, 1996). Se a correlação é baixa, os genes expressos em um ambiente são diferentes daqueles expressos no outro ambiente. Da mesma forma ocorre para os locos dos marcadores. Os efeitos dos marcadores obtidos pelos híbridos avaliados no ambiente de treinamento podem não ser consistentes no ambiente de validação, o que faz com que a predição das médias esteja sujeita aos efeitos da interação H x A.

Uma alternativa seria obter o efeito da interação marcadores x ambientes (M x A) e incluí-lo no modelo linear misto de predição dos efeitos genotípicos das marcas, de forma que somente aqueles marcadores com efeito consistente entre os ambientes seriam efetivamente utilizados na seleção genômica. Contudo, Schulz-Streeck et al. (2013) observaram que a contribuição da interação M x A na estimativa de r_{MG} é insignificante, pois ela influencia apenas na estimação do componente de variância associado ao efeito das marcas, mas não influencia expressivamente no valor da média predita.

A interação H x A pode ser descomposta em dois componentes: um devido à diferença na variabilidade genética dos híbridos em cada ambiente, denominado interação simples, e outro devido à falta de correlação genética entre os híbridos de um ambiente para outro, denominada interação complexa (VENCOVSKY; BARRIGA, 1992). Quando há predomínio da parte simples, a classificação dos híbridos nos ambientes de treinamento e validação não se altera, e a predição das médias no ambiente de validação pode ser realizada com base nas médias fenotípicas e nos efeitos dos marcadores obtidos no ambiente de treinamento. Contudo, quando a interação complexa é mais expressiva, os híbridos avaliados no ambiente de treinamento tem performance diferencial no ambiente de validação, e os efeitos genotípicos dos marcadores podem não ser boas estimativas a serem utilizadas na obtenção das médias preditas.

A relação entre as estimativas da acurácia de predição (r_{MG}) e o efeito da interação H x A pode ser facilmente visualizada pela magnitude dos componentes de variância da interação H x A (σ_{HA}^2) obtidas nas análises de variância conjunta entre os ambientes dois a dois. Com exceção da

análise conjunta dos ambientes E.E. Areão - 2004/2005 (Ambiente 5) e E.E. Departamento de Genética - 2004/2005 (Ambiente 11), os outros quatro pares de ambientes que apresentaram valores não significativos de σ_{HA}^2 , ou seja, onde não houve interação H x A, também apresentaram altas estimativas de r_{MG} na validação cruzada. Além disso, esses pares de ambiente também apresentaram altas estimativas de correlação fenotípica (r_F) entre as médias dos híbridos. Os ambientes E.E. Anhembi - 2003/2004 (Ambiente 1) e E.E. Anhembi - 2004/2005 (Ambiente 2), por exemplo, apresentaram acurácia de predição média de $r_{MG} = 0,70$. A análise conjunta entre esse dois ambientes apresentou $\sigma_{HA}^2 = 0,10^{ns}$ e $r_F = 0,70$. Esses resultados também foram observados para os pares de ambiente E.E. Areão - 2004/2005 (Ambiente 5) e E.E. Caterpillar - 2004/2005 (Ambiente 8), ($r_{MG} = 0,68$, $\sigma_{HA}^2 = 0,10^{ns}$ e $r_F = 0,70$); E.E. Departamento de Genética - 2002/2003 (Ambiente 9) e E.E. Departamento de Genética - 2003/2004 (Ambiente 10), ($r_{MG} = 0,56$, $\sigma_{HA}^2 = 0,08^{ns}$ e $r_F = 0,79$); e entre os dois experimentos avaliados em E.E. Patos de Minas - 2004/2005, (Ambientes 12 e 13) ($r_{MG} = 0,69$, $\sigma_{HA}^2 = -0,04^{ns}$ e $r_F = 0,78$), (Tabela 3).

Para os pares de ambientes que apresentaram estimativas significativas de σ_{HA}^2 , ou seja, onde houve interação H x A, observou-se relação das magnitudes de σ_{HA}^2 e r_F com as estimativas de r_{MG} . Os ambientes que apresentaram maior estimativa significativa de σ_{HA}^2 também apresentaram menores estimativas para r_F e r_{MG} . A maior estimativa de σ_{HA}^2 foi obtida para os ambientes E.E. Departamento de Genética - 2002/2003 (Ambiente 9) e E.E. Caterpillar - 2002/2003 (Ambiente 6), ($\sigma_{HA}^2 = 1,67^{**}$), que apresentaram baixos valores de $r_F = 0,62$ e $r_{MG} = 0,46$. Já os ambientes E.E. Areão - 2003/2004 (Ambiente 4) e E.E. Caterpillar - 2003/2004 (Ambiente 7) apresentaram menor estimativa significativa de σ_{HA}^2 ($0,12^{**}$) e maiores estimativas de $r_F = 0,76$ e $r_{MG} = 0,56$ (Tabela 3).

Tabela 3 – Correlação fenotípica (r_F) entre as médias de produção de grãos em $t\ ha^{-1}$ dos ambientes (abaixo da diagonal) e estimativas dos componentes de variância da interação híbridos x ambientes (σ_{HA}^2) das análises de variância conjunta entre os ambientes dois a dois, e suas respectivas significâncias (acima da diagonal)

Amb	1	2	3	4	5	6	7	8	9	10	11	12	13
1	-	0,10 ^{ns}	0,68**	0,52**	0,46**	0,43**	0,31**	0,68**	0,91**	0,89**	0,57**	0,74**	0,44**
2	0,70	-	0,54**	0,43**	0,29*	0,37**	0,20**	0,66**	0,82**	0,87**	0,58**	0,71**	0,51**
3	0,59	0,61	-	0,31**	0,67**	0,32**	0,40**	0,85**	0,49**	0,41**	0,30**	0,65**	0,53**
4	0,59	0,57	0,73	-	0,69**	0,17**	0,12**	0,91**	0,36**	0,29**	0,30**	0,32**	0,15*
5	0,60	0,62	0,58	0,55	-	0,56**	0,43**	0,08 ^{ns}	1,17**	1,20**	-0,25 ^{ns}	0,43**	0,24**
6	0,64	0,61	0,73	0,76	0,59	-	1,12**	0,96**	1,67**	0,66**	0,59**	0,55**	0,40**
7	0,65	0,65	0,71	0,76	0,63	0,77	-	0,52**	0,55**	0,53**	0,38**	0,39**	0,17**
8	0,59	0,56	0,56	0,53	0,68	0,63	0,65	-	1,14**	0,83**	0,85**	0,87**	0,72**
9	0,48	0,45	0,69	0,69	0,42	0,62	0,62	0,49	-	0,08 ^{ns}	0,44**	0,62**	0,55**
10	0,48	0,43	0,69	0,71	0,41	0,59	0,61	0,48	0,79	-	0,19*	0,68**	0,48**
11	0,58	0,53	0,71	0,68	0,45	0,59	0,65	0,53	0,66	0,72	-	0,34**	0,21*
12	0,55	0,53	0,63	0,72	0,44	0,64	0,68	0,55	0,64	0,61	0,68	-	-0,04 ^{ns}
13	0,58	0,52	0,62	0,72	0,52	0,63	0,70	0,55	0,59	0,61	0,67	0,78	-

**/ significativo à 1% de probabilidade pelo teste F; */ significativo à 5% de probabilidade pelo teste F; ^{ns}/ não significativo. Amb 1: E.E. Anhembi - 2003/2004; Amb 2: E.E. Anhembi - 2004/2005; Amb 3: E.E. Areão - 2002/2003; Amb 4: E.E. Areão - 2003/2004; Amb 5: E.E. Areão - 2004/2005; Amb 6: E.E. Caterpillar - 2002/2003; Amb 7: E.E. Caterpillar - 2003/2004; Amb 8: E.E. Caterpillar - 2004/2005; Amb 9: E.E. Departamento de Genética - 2002/2003; Amb 10: E.E. Departamento de Genética - 2003/2004; Amb 11: E.E. Departamento de Genética - 2004/2005; Amb 12: exp 1 da E.E. Patos de Minas - 2004/2005; Amb 13: exp 2 da E.E. Patos de Minas - 2004/2005.

Esses resultados indicam que o efeito da interação H x A na seleção genômica deve ser considerado da mesma forma que vem sendo feito na seleção fenotípica. O ideal é que os híbridos sejam avaliados em vários ambientes e anos, e a média fenotípica dessas avaliações sejam utilizadas como população de treinamento para estimar os efeitos genotípicos das marcas. Além disso, os ambientes onde são conduzidos os experimentos podem ser estratificados em grupos homogêneos, sendo que os ambientes dentro dos grupos devem possuir as mesmas características edafo-climáticas. As validações cruzadas, então, devem ser realizadas entre ambientes de treinamento e validação alocados dentro do mesmo grupo de forma a minimizar o efeito da interação H x A e, conseqüentemente, maximizar a estimativa da acurácia de predição (r_{MG}). Segundo Bernardo (2010), as análises cluster e componentes principais constituem métodos eficientes para estratificação de ambientes.

3.3.4 Validação cruzada entre e dentro de grupos de híbridos relacionados: efeito da estrutura da população

As análises de validação cruzada dentro de grupos de híbridos relacionados apresentaram elevadas estimativas de acurácia de predição (r_{MG}), variando de 0,90, para o grupo G4, a 1,02, para o grupo G3. As estimativas de r_{MG} para as análises de validação cruzada variaram expressivamente entre os grupos dois a dois. Foram observadas elevadas estimativas de r_{MG} entre os grupos G2 e G3 ($r_{MG} = 0,90$ e $r_{MG} = 0,86$), entre G2 e G5 ($r_{MG} = 0,72$ e $r_{MG} = 0,73$) e entre G3 e G5 ($r_{MG} = 0,77$ e $r_{MG} = 0,74$). Os grupos G3 e G4 foram os que apresentaram as menores estimativas de r_{MG} ($r_{MG} = 0,02$ e $r_{MG} = 0,03$). O grupo G5 apresentou maior r_{MG} em média com os demais ($r_{MG} = 0,49$), e os grupos G1 e G4 apresentaram as menores r_{MG} em média com os demais, $r_{MG} = 0,21$ e $r_{MG} = 0,19$, respectivamente. Apesar disso, os grupos G1 e G4 apresentaram estimativas de r_{MG} entre si, $r_{MG} = 0,44$ e $r_{MG} = 0,48$, relativamente superiores às apresentadas entre G1 e G4 e os demais grupos (Tabela 4).

Tabela 4 – Estimativas da acurácia de predição (r_{MG}) obtidas nas análises de validação cruzada dentro de grupos de híbridos relacionados (diagonal da tabela) e entre os grupos dois a dois

Grupo utilizado como pop. val. *	Grupo utilizado como população de treinamento				
	G1 (L x T1)	G2 (L x T2)	G3 (L x T3)	G4 (L x T4)	G5 (L x T5)
G1 (L x T1)	0.91	0.13	0.11	0.48	0.26
G2 (L x T2)	0.07	1.01	0.86	0.09	0.72
G3 (L x T3)	0.03	0.90	1.02	0.03	0.77
G4 (L x T4)	0.44	0.09	0.02	0.90	0.19
G5 (L x T5)	0.27	0.73	0.74	0.27	0.97
r_{MG} média	0,21	0,45	0,43	0,19	0,49

* Grupo utilizado como população de validação; L: 50 linhagens endogâmicas; T1: Testador L-08-05F; T2: Testador L-38-05D; T3: Testador L-36-07f; T4: Testador L-49-02D; T5: Testador L-46-10D.

Das dez combinações dos grupos dois a dois, as combinações G2-G3, G2-G5, G3-G5, e G1-G4 apresentaram estimativas de r_{MG} que sugerem alguma relação genética entre esses grupos. Como as 50 linhagens endogâmicas utilizadas na obtenção dos híbridos em cada grupo são as mesmas, a performance diferencial dos grupos se deve à linhagem testadora utilizada e à interação entre as linhagens endogâmicas x linhagens testadoras. Utilizando o mesmo conjunto de dados deste estudo, Alves (2006) avaliou a produção de grãos das 50 linhagens endogâmicas e

das cinco linhagens testadoras em quatro experimentos com duas repetições. Foi realizada análise de variância conjunta dos experimentos desdobrando a fonte de variação ‘tratamento’ nos componentes testadores, linhagens e interação linhagens x testadores. Foi possível observar que a interação linhagens x testadores foi significativa, indicando que a performance das linhagens endogâmicas realmente foi diferencial quando elas foram cruzadas com as diferentes linhagens testadoras.

Alves (2006) também avaliou a relação existente entre as cinco linhagens testadoras em função das suas divergências genéticas. Utilizando o método UPGMA, os testadores foram alocados em dois grupos heteróticos distintos com base nas estimativas das capacidades específicas de combinação (CEC) da análise dialélica envolvendo as 50 linhagens endogâmicas, as cinco linhagens testadoras e os 250 híbridos. Os dois grupos heteróticos foram entre as linhagens testadoras T1-T4, e entre T2-T3-T5, sendo que os testadores T2 e T3 apresentaram menor estimativa de CEC que às obtidas entre T2 e T5 e entre T3 e T5. Menores estimativas de CEC indicam menor divergência, ou seja, os testadores T2 e T3, são mais correlacionados entre si que com T5. Os grupos heteróticos obtidos pelo autor confirmam a relação genética sugerida pelas estimativas de r_{MG} entre os grupos G2-G3, G2-G5, G3-G5, e G1-G4. Esses resultados podem ainda ser verificados pelas correlações das médias fenotípicas dos híbridos (r_F) entre os grupos dois a dois. Foram obtidas maiores estimativas de r_F para os grupos G2-G3 ($r_F = 0,86$), G2-G5 ($r_F = 0,69$), G3-G5 ($r_F = 0,71$) e moderada estimativa de r_F para os grupos G1-G4 ($r_F = 0,42$). As demais estimativas de r_F foram inferiores à 0,33 (Tabela 5).

Tabela 5 – Correlações fenotípicas (r_F) entre as médias dos híbridos dos grupos dois a dois

Grupos	G2: L x T2	G3: L x T3	G4: L x T4	G5: L x T5
G1: L x T1	0,16	0,14	0,42	0,33
G2: L x T2		0,86	0,10	0,69
G3: L x T3			0,02	0,71
G4: L x T4				0,22

L: refere-se às 50 linhagens endogâmicas. T1: Testador L-08-05F; T2: Testador L-38-05D; T3: Testador L-36-07f; T4: Testador L-49-02D; T5: Testador L-46-10D.

Fica evidente que a seleção genômica em grupos de híbridos relacionados é expressivamente mais eficiente que a realizada entre híbridos não relacionados geneticamente. Assim como ocorre na validação cruzada entre ambientes, os efeitos dos marcadores podem não

ser consistentes entre os grupos de híbridos utilizados nas populações de treinamento e validação (SCHULZ-STREECK et al., 2012). Isso significa que os efeitos genotípicos dos marcadores preditos pela população de treinamento podem não ter capacidade preditiva na população de validação se essas populações não apresentaram correlação genética. Nessa análise, todos os 50 híbridos de cada grupo foram utilizados na predição dos efeitos genotípicos dos marcadores e validação do modelo, portanto toda variabilidade genética da população de treinamento esta representada na população de validação. Como discutido anteriormente, se forem amostrados apenas alguns híbridos para comporem a população de treinamento, as estimativas de r_{MG} da validação cruzada entre grupos podem ser ainda menores ou até negativas em função da amostragem, pois pode ser que os híbridos amostrados nas populações de treinamento e validação não apresentem nenhum parentesco. A influência da relação genética entre os genótipos das populações de treinamento e validação também foi reportada na predição de linhagens duplo haploides (ALBRECHT et al, 2014) e genótipos F_2 em milho (WINDHAUSEN et al., 2012).

3.3.5 Seleção fenotípica e genômica de híbridos superiores

A fim de verificar a eficiência da seleção genômica nos programas de melhoramento de milho, foram realizadas as seleções fenotípica (SF) e genômica (SG) dos híbridos avaliados nos anos agrícolas 2002/2003-2003/2004, e posterior SF dos híbridos avaliados no ano agrícola 2004/2005. A correlação entre as médias fenotípicas e as médias preditas para produção de grãos ($t\ ha^{-1}$) dos 250 híbridos avaliados nos anos agrícolas 2002/2003-2003/2004 foi de 0,79. O coeficiente de correlação entre as médias fenotípicas do ano agrícola 2004/2005 foi de 0,69 com as médias preditas e de 0,83 com as médias fenotípicas dos anos agrícolas anteriores.

Dos 25 híbridos selecionados no ano agrícola 2004/2005, 14 (56%) também foram selecionados pela SF, e 13 (52%) foram selecionados pela SG nos anos agrícolas anteriores. Destes, 10 (40%) foram coincidentes, ou seja, foram selecionados pela SG e SF dos anos agrícolas 2002/2003-2003/2004 e pela SF do ano agrícola 2004/2005. Assim, 4 foram selecionados somente pela SF (híbridos H-69, H-77, H-105 e H-249) e 3 foram selecionados somente pela SG (híbridos H-95, H-207 e H-230) nos dois anos agrícolas (2002/2003 e 2003/2004). Considerando que nos estágios finais dos programas de melhoramento poucos híbridos são selecionados para serem realmente lançados como cultivares comerciais, observando

apenas os cinco melhores híbridos do ano agrícola 2004/2005, apenas dois seriam selecionados pela SF (híbridos H-226 e H-219) e um seria selecionado pela SG (híbrido H-205), (Tabela 6).

Tabela 6 – Identificação dos híbridos selecionados via seleção fenotípica (SF) e seleção genômica (SG) nos anos agrícolas 2002/2003-2003/2004 e híbridos selecionados fenotipicamente (SF) no ano agrícola 2004/2005

Rank	2002/2003-2003/2004		2004/2005	Rank	2002/2003-2003/2004		2004/2005
	SF	SG			SF	SG	
1	H-226	H-208	H-210	13	H-206	H-211	H-76
2	H-55	H-205	H-219	14	H-58	H-60	H-105
3	H-105	H-207	H-205	15	H-35	H-210	H-250
4	H-219	H-204	H-226	16	H-77	H-57	H-73
5	H-88	H-250	H-69	17	H-100	H-219	H-203
6	H-208	H-206	H-207	18	H-245	H-54	H-229
7	H-210	H-227	H-204	19	H-249	H-203	H-95
8	H-205	H-244	H-202	20	H-79	H-243	H-230
9	H-204	H-201	H-55	21	H-67	H-58	H-44
10	H-250	H-217	H-209	22	H-90	H-95	H-249
11	H-211	H-55	H-208	23	H-150	H-245	H-77
12	H-203	H-238	H-245	24	H-93	H-230	H-108
				25	H-69	H-226	H-10

Esse esquema de seleção representou uma situação prática dos programas de melhoramento de milho. Os 250 híbridos foram avaliados em sete ambientes, considerando os locais onde os experimentos foram conduzidos nos anos agrícolas 2002/2003-2003/2004, e as suas performances foram verificadas em quatro ambientes no ano agrícola seguinte. Mesmo a seleção genômica tendo apresentado as elevadas estimativas de acurácia de predição apresentadas anteriormente, ela não foi superior à seleção fenotípica em identificar os híbridos superiores nesse esquema de seleção, indicando que a SG está sujeita aos mesmos efeitos que a SF. Contudo, foi possível afirmar que a SG foi equivalente à SF, pois mesmo não identificando os mesmos híbridos, essas metodologias selecionaram praticamente a mesma proporção de híbridos em relação aos selecionados no ano agrícola seguinte.

3.3.6 Aplicações da seleção genômica nos programas de melhoramento

Segundo Massman et al. (2013), em um programa de melhoramento de milho os híbridos simples possuem duas características que os tornam atrativos para serem utilizados em populações de treinamento nos estudos de seleção genômica: (1) constituem-se de germoplasma elite e são provenientes do cruzamento entre as linhagens que geralmente são utilizadas como genitoras nos demais cruzamentos do programa; (2) são avaliados em vários ambientes, permitindo a estimação do efeito da interação $H \times A$. Com isso, inúmeros estudos têm sido realizados utilizando diferentes modelos de predição genômica, caracteres avaliados e estruturas de população em milho (LORENZANA; BERNARDO, 2009; JACOBSON et al., 2014; LIAN et al., 2014; ALBRECHT et al., 2011; WINDHAUSEN et al., 2012). Embora tenham sido relatados valores expressivos de acurácia de predição, os resultados até o momento não dão suporte à aplicação prática da seleção genômica em híbridos simples de milho avaliados em múltiplos ambientes.

No que se refere ao número de híbridos da população de treinamento (N) e o número de marcadores utilizados no modelo de seleção (N_M), os resultados do presente estudo indicaram que o aumento de N e N_M não alteraram significativamente as estimativas da acurácia de predição. Na prática, a escolha do número de híbridos a ser avaliado em campo e do número de marcas utilizadas é função da estrutura da população avaliada, da herdabilidade do caráter e da quantidade de recursos disponíveis no programa (LORENZANA; BERNARDO, 2009). De forma geral, quanto menor a relação genética entre a população de treinamento e validação, maior número de híbridos devem ser fenotipados e maior número de marcas utilizadas. Da mesma forma, se a característica em questão for um caráter quantitativo complexo com baixa herdabilidade, maior N e N_M são necessários.

Atualmente, com a disponibilidade de marcadores novos e baratos, o número de marcas não deve ser um problema. Contudo, o custo da fenotipagem geralmente é elevado. O que se deseja é avaliar o menor número de híbridos possível, mas o melhorista deve ter o cuidado de avaliar um número de híbridos mínimo que represente a variabilidade da população de validação e minimize os problemas de amostragem. As amplitudes dos intervalos das estimativas de acurácia observadas neste estudo indicam que a seleção genômica pode ou não ser eficiente em prever as performances dos híbridos dependendo da população amostrada.

Assim como na seleção fenotípica, os efeitos da interação híbridos x ambientes (H x A) também interferem na seleção genômica, pois os valores genotípicos dos marcadores preditos no ambiente de treinamento podem não ser consistentes no ambiente de validação. Dessa forma, eles podem não fornecer boas estimativas a serem utilizadas na predição das médias dos híbridos (SCHULZ-STREECK et al., 1013). A melhor estratégia nesse caso é estratificar previamente os ambientes em grupos homogêneos para que sejam utilizados ambientes de treinamento e validação do mesmo grupo e, com isso, minimizar o efeito da interação H x A. Para isso, metodologias como cluster e análise de componentes principais se constituem boas alternativas de estratificação de ambientes (BERNARDO, 2010). Além disso, o ideal é que os híbridos sejam avaliados em vários locais e anos, e as médias fenotípicas de todas essas avaliações sejam utilizadas como população de treinamento. A eficiência dessa estratégia pode ser facilmente visualizada pela estimativa da acurácia de predição (r_{MG}) da validação cruzada utilizando as médias fenotípicas nos 13 ambientes ($r_{MG} = 0,64$), consideravelmente superior à acurácia média das validações cruzadas utilizando as médias fenotípicas em cada ambiente ($r_{MG} = 0,48$).

Com relação à estrutura da população, observou-se que a seleção genômica só foi eficiente quando as populações de treinamento e validação eram constituídas de híbridos relacionados. Idealmente, uma população composta por híbridos geneticamente diversos poderia ser utilizada no ambiente de treinamento dos programas de melhoramento. Assim, seria garantido que a população de validação estaria sempre representada no ambiente de treinamento, independente da amostra. Segundo Windhausen et al. (2012), para ser efetiva, essa estratégia requer a obtenção de uma população de treinamento e número de marcas bem maior que o comumente utilizado. Simulando uma população F_2 , Hickey et al. (2014) mostraram que seria necessário avaliação fenotípica de 5.000 a 20.000 genótipos para obtenção de estimativas razoáveis de acurácia de predição quando essas progênies são provenientes de cruzamentos não relacionados. Considerando os tipos de população dos programas de melhoramento de milho, o custo da fenotipagem de um número tão alto híbridos não deve compensar o ganho na estimativa da acurácia. Assim, a forma mais eficiente de aplicar a seleção genômica nesse caso é realizar um estudo prévio da divergência genética e grupos heteróticos aos quais pertencem as linhagens genitoras para que os híbridos avaliados e preditos sejam derivados de linhagens pertencentes ao mesmo grupo heterótico e, portanto, relacionados geneticamente.

Mesmo a seleção genômica tendo apresentado elevadas estimativas de acurácia de predição neste estudo, ela foi equivalente à seleção fenotípica em identificar híbridos superiores em um esquema de seleção semelhante ao realizado pelos programas de melhoramento de milho. Esses resultados indicam que a seleção genômica está sujeita aos mesmos efeitos que a seleção fenotípica no que se refere à variabilidade e estrutura da população avaliada e ao efeito da interação genótipos x ambientes. Contudo, a facilidade de aplicação e o baixo custo da seleção genômica atualmente fazem que esta seja uma estratégia promissora para aumentar a eficiência dos programas de melhoramento. Assim, os resultados deste trabalho sugerem que a predição genômica de híbridos de milho seja uma ferramenta útil se forem tomados os cuidados de obter as médias fenotípicas com alta precisão em múltiplos ambientes, utilizar ambientes de treinamento e validação relacionados e delinear as populações de treinamento e validação de acordo com a variabilidade e correlação genética dos híbridos, a fim de maximizar a acurácia e, consequentemente, o sucesso da predição.

Referências

- AGUIAR, A.M.; CARLINI-GARCIA, L.A.; SILVA, A.R.; SANTOS, M.F.; GARCIA, A.A.F.; SOUZA JUNIOR, C.L. Combining ability of inbred lines of maize and stability of their respective single-crosses. **Scientia Agricola**, Piracicaba, v.60, n.1, p.83-89, 2003.
- ALBRECHT, T.; AUINGER, H. J.; WIMMER, V.; OGUTU, J.O.; KNAACK, C.; OUZUNOVA, M.; PIEPHO, H.P.; SCHON, C. C. Genome-based prediction of maize hybrid performance across genetic groups, testers, locations and years. **Theoretical and Applied Genetics**, Berlin, v. 127, p. 1375-1386, 2014.
- ALBRECHT, T.; WIMMER, V.; AUINGER, H. J.; ERBE, M.; KNAACK, C.; OUZUNOVA, M.; SIMIANER, H.; SCHON, C. C. Genome-based prediction of testcross values in maize. **Theoretical and Applied Genetics**, Berlin, v. 123, p. 339-350, 2011.
- ALVES, G.F. **Relações entre seleção de testadores de milho e suas divergências genéticas**. 2006. 145p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2006.
- BERNARDO, R. **Breeding for quantitative traits in plants**. Woodbury: Stemma Press, 2010. 260 p.
- BERNARDO, R.; YU, J. Prospects for genome wide selection for quantitative traits in maize. **Crop Science**, Madison, v. 47, p. 1082-1090, 2007.

COCHRAN, W.G.; COX, G.M. **Experimental design**. 2nd ed. New York: John Wiley, 1966. 611 p.

CROSSA, J.; de los CAMPOS, G.; PEREZ, P.; GIANOLA, D.; ATLIN, G.; BURGUEÑO, J.; ARAUS, J. L.; MAKUMBI, D.; YAN, J.; ARIEF, V.; BANZIGER, M.; BRAUN, H. J. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. **Genetics**, Austin, v. 186, p. 713-724, 2010.

CROSSA, J.; PEREZ, P.; HICKEY, J.; BURGUEÑO, J.; ORNELLA, L.; CERÓN-ROJAS, J.; ZHANG, X.; DREISIGACKER, S.; BABU, R.; LI, Y.; BONNETT, D.; MATHEWS, K. Genomic prediction in CIMMYT maize and wheat breeding programs. **Heredity**, New York, v.112, p.48-60, 2014.

DEKKERS, J.C.M. Prediction of response to marker-assisted and genomic selection using selection index theory. **Journal of Animal Breeding and Genetics**, Oxford, v.124, p.331-341, 2007.

ENDELMAN, J. B. Ridge regression and other kernels for genomic selection with R package rrBLUP. **The Plant Genome**, Madison, v.4, p. 250-255, 2011.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics**. 4th ed. London: Longman Scientific e Technical, 1996. 464 p.

GODDARD, M.E.; HAYES, B.J. Genomic selection. **Journal of Animal Breeding and Genetics**, Oxford, v. 124, p. 323-330, 2007.

HALLAUER, A.R.; CARENA, M.J.; MIRANDA FILHO, J.B. **Quantitative genetics in maize breeding**. 3rd ed. Ames: Iowa States University Press, 2010. 664 p.

HICKEY, M.J.; DREISIGACKER, S.; CROSSA, J.; HEARNE, S.; BABU, R.; PRASANNA, B.M.; GRONDONA, M.; ZAMBELLI, A.; WINDHAUSEN, V. S.; MATHEWS, K.; GORJANC, G. Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. **Crop Science**, Madison, v. 54, p. 1476-1488, 2014.

HOISINGTON, D.; KHAIRLLAH, M.; GONZALES-DE-LEON, D. **Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory**. 2nd ed. Mexico: CIMMYT, 1994. 102 p.

JACOBSON, A.; LIAN, L.; ZHONG, S.; BERNARDO, R. General combining ability model for genomewide selection in a biparental cross. **Crop Science**, Madison, v. 54, p. 895-905, 2014.

KNAPP, S.J.; STROUP, W.W.; ROSS, W.M. Exact confidence intervals for heritability on a progeny mean basis. **Crop Science**, Madison, v. 25, n. 1, p. 192-194, 1985.

LIAN, L.; JACOBSON, A.; ZHONG, S.; BERNARDO, R. Genomewide prediction accuracy within 969 maize biparental populations. **Crop Science**, Madison, v. 54, p. 1514-1522, 2014.

LORENZANA, R.E.; BERNARDO, R. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. **Theoretical and Applied Genetics**, Berlin, v. 120, p. 151-161, 2009.

MASSMAN, J.M.; GORDILLO, A.; LORENZANA, R.E.; BERNARDO, R. Genomewide predictions from maize single-cross data. **Theoretical and Applied Genetics**, Berlin, v.126, p.13–22, 2013.

MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, Austin, v. 157, p. 1819-1829, 2001.

R DEVELOPMENT CORE TEAM. **R**: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2012. 3604 p.

RAPOSO, F.V. **Seleção recorrente recíproca em populações derivadas de híbridos simples de milho**. 2002. 106p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, 2002.

SAS INSTITUTE. **Statistical analysis software: SAS 9.2**. Cary, 2008.

SCHULZ-STREECK, T.; OGUTU, J.O.; GORDILLO, A.; KARAMAN, Z.; KNAAK, C.; PIEPHO H.P. Genomic selection allowing for marker-by-environment interaction. **Plant Breeding**, Malden, v.132, p.532-538, 2013.

SCHULZ-STREECK, T.; OGUTU, J.O.; KARAMAN, Z.; KNAAK, C.; PIEPHO H.P. Genomic selection using multiple populations. **Crop Science**, Madison, v.52, p.2453-2461, 2012.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica aplicada ao fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 1992. 496p.

VOS, P.; ROGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRITJERS, A.; POT, J.; PELEMAN, J.; KUIPE, M.; ZABEAU, M. AFLP: a new technic for DNA fingerprint. **Nucleic Acids Research**, London, v.23, n.21, p.4407-4414, 1995.

WINDHAUSEN, V.S.; ATLIN, G.N.; HICKEY, J.M.; CROSSA, J.; JANNINK, J.L.; SORRELLS, M.E.; RAMAN, B.; CAIRNS, J.E.; TAREKEGNE, A.; SEMAGN, K.; BEYENE, Y.; GRUDLOYMA, P.; TECHNOW, F.; RIEDELSHEIMER, C.; MELCHINGER, A.E. Effectiveness of genomic prediction of maize hybrid performance in different breeding populations and environments. **Genes Genomes Genetics**, Bethesda, v. 2, p. 1427-1436, 2012.

ANEXOS

Anexo A - Origens das linhagens endogâmicas e testadoras utilizadas na obtenção dos híbridos simples de milho

Número	Linhagem ¹	Origem	Número	Linhagem ¹	Origem
01	16-02 D	IG-2	29	131-01 F	Top Cross
02	25-04 D	IG-2	30	148-03 F	Top Cross
03	27-01 D	IG-2	31	16-04 R	BR-201
04	39-05 D	IG-2	32	16-07 F	BR-201
05	55-02 D	IG-2	33	22-02 D	BR-201
06	66-08 D	IG-2	34	23-05 D	BR-201
07	94-02 D	IG-2	35	24-03 R	BR-201
08	102-02 D	IG-2	36	29-03 D	BR-201
09	120-04 F	IG-2	37	35-04 F	BR-201
10	149-05 D	IG-2	38	44-01 F	BR-201
11	168 F	IG-2	39	48-03 D	BR-201
12	03-01 D	HS-1	40	18-08A F	CMS-05
13	20-02 R	HS-1	41	37-02B D	CMS-05
14	30-07 F	HS-1	42	37-03B D	CMS-05
15	31-01 F	HS-1	43	37-04B F	CMS-05
16	33-04 D	HS-1	44	37-07B D	CMS-05
17	04-05 F	IG-1	45	31 F	606
18	88-05 F	IG-1	46	53 F	606
19	90-05 F	IG-1	47	06-05 D	601
20	8 F	XL-560	48	61-02 F	Top Cross
21	14 D	XL-560	49	84-03 F	Top Cross
22	16 D	XL-560	50	08-04B D	CMS-05
23	56 D	XL-560	51 (T1)	L-08-05 F	IG-1
24	128 D	XL-560	52 (T2)	L-38-05 D	HS-1
25	14-01 F	Top Cross	53 (T3)	L-36-07 F	HS-1
26	45-03 D	Top Cross	54 (T4)	L-49-02 D	Top Cross
27	53-01 F	Top Cross	55 (T5)	L-46-10 D	BR-201
28	82-01 D	Top Cross			

¹D: dentado; F: flint ou duro e R: roxa

Características das linhagens testadoras:

- T1: L-08-05F, com grãos alaranjados e duros, obtida da população IG-1;
- T2: L-38-05D, com grãos amarelos e dentados, obtida do híbrido simples utilizado como genitor feminino do híbrido duplo comercial BR-201;
- T3: L-36-07F, com grãos alaranjados e duros, também obtida do híbrido simples utilizado como genitor feminino do híbrido duplo comercial BR-201;
- T4: L-49-02D, com grãos amarelos e dentados, obtida do cruzamento das populações BR-105 e BR-106;
- T5: L-46-10D, com grãos amarelos e dentados, obtida do híbrido duplo comercial BR-201.

Anexo B - Esquema das análises de variância individuais e esperanças dos quadrados médios

FV	GL	QM	E(QM)	F
Repetição	$k - 1$	-	-	-
Bloco/Repetição	$k(j - 1)$	-	-	-
Híbrido	$i - 1$	QM_H	$\sigma^2 + k \sigma_H^2$	QM_H/QM_{EI}
Erro Intra	$(j - 1)(kj - j - 1)$	QM_{EI}	σ^2	

k: número de repetições; i: número de tratamentos; j: número de blocos

Anexo C – Esquema da análise de variância conjunta e esperanças dos quadrados médios

FV	GL	QM	E(QM)	F
Ambiente	$l - 1$	QM_A	$\sigma^2 + k \sigma_{HA}^2 + ki \sigma_A^2$	QM_A/QM_{HA}
Híbrido	$i - 1$	QM_H	$\sigma^2 + k \sigma_{HA}^2 + kl \sigma_H^2$	QM_H/QM_{HA}
H x A	$(i - 1)(l - 1)$	QM_{HA}	$\sigma^2 + k \sigma_{HA}^2$	QM_{HA}/QM_{EM}
Erro Médio	$l(j - 1)(kj - j - 1)$	QM_{EM}	σ^2	

l: número de ambientes; k: número de repetições; i: número de tratamentos; j: número de blocos

Anexo D – Estimativas dos quadrados médios (QM) de híbridos e do erro intrabloco, coeficientes de variação (CV%), médias ($t\ ha^{-1}$), herdabilidades no sentido amplo em nível de média dos híbridos (h_A^2) e intervalos de confiança das estimativas de herdabilidade a 95% de probabilidade ($IC_{(h_A^2)0,95}$) das análises de variância individuais para produção de grãos dos híbridos de milho avaliados em 13 ambientes

FV	GL	QM			
		ARE_0203 ^a	CAT_0203 ^a	DEP_0203 ^a	ARE_0304 ^a
Híbridos	255	6,51 ^{**}	4,57 ^{**}	4,71 ^{**}	4,19 ^{**}
Erro Intra	224	1,36	0,89	0,87	0,81
CV %		13,39	14,00	12,04	9,65
Média		8,70	6,75	7,77	9,33
h_A^2		0,79	0,80	0,81	0,80
$IC_{(h_A^2)0,95}$		0,73 - 0,83	0,74 - 0,84	0,76 - 0,85	0,75 - 0,84

FV	GL	QM			
		CAT_0304 ^a	DEP_0304 ^a	ANH_0304 ^a	ARE_0405 ^a
Híbridos	255	3,58 ^{**}	5,68 ^{**}	4,13 ^{**}	6,86 ^{**}
Erro Intra	224	0,71	1,20	0,96	2,29
CV %		10,52	11,55	13,68	22,36
Média		8,04	9,48	7,19	6,76
h_A^2		0,80	0,78	0,76	0,66
$IC_{(h_A^2)0,95}$		0,74 - 0,84	0,72 - 0,83	0,69 - 0,81	0,56 - 0,74

FV	GL	QM				
		CAT_0405 ^a	DEP_0405 ^a	ANH_0405 ^a	PTM1_0405 ^a	PTM2_0405 ^a
Híbridos	255	9,05 ^{**}	5,73 ^{**}	4,19 ^{**}	6,35 ^{**}	4,95 ^{**}
Erro Intra	224	3,28	1,42	1,44	1,36	1,54
CV %		20,31	11,54	14,22	11,53	13,22
Média		8,91	10,32	8,46	10,12	9,38
h_A^2		0,63	0,75	0,65	0,78	0,68
$IC_{(h_A^2)0,95}$		0,53 - 0,71	0,68 - 0,80	0,55 - 0,73	0,72 - 0,83	0,59 - 0,75

^{a/} As siglas referem-se às Estações Experimentais da Fazenda Areão, Caterpillar, Departamento de Genética, Fazenda Anhembí e Patos de Minas (2 experimentos), e aos anos agrícolas 2002/2003, 2003/2004 e 2004/2005

^{**} significativo à 1% de probabilidade

Anexo E – Estimativas dos quadrados médios (QM) de ambientes (A), híbridos (H), interação híbridos x ambientes (H x A) e erro médio, coeficiente de variação (CV%), média ($t \text{ ha}^{-1}$), herdabilidade no sentido amplo em nível de média dos híbridos (h^2_{Conj}) e intervalo de confiança da estimativa de herdabilidade a 95% de probabilidade ($IC_{(h^2_{\text{Conj}})0,95}$) da análise de variância conjunta da produção de grãos dos híbridos de milho nos 13 ambientes

FV	GL	QM
Ambientes (A)	12	730,66**
Híbridos (H)	255	51,52**
H x A	3060	2,56**
Erro Médio	2912	1,41
CV %		13,21
Média		8,55
h^2_{Conj}		0,95
$IC_{(h^2_{\text{Conj}})0,95}$		0,94 - 0,96

** significativo à 1% de probabilidade

Anexo F – Quadrados médios das análises de variância conjunta para produção de grãos (t ha⁻¹) de híbridos de milho dos treze ambientes dois a dois

(continua)							
FV	GL	1 x 2	1 x 3	1 x 4	1 x 5	1 x 6	1 x 7
Híbrido	255	8.40**	10.04**	7.66**	10.56**	8.52**	7.48**
Ambiente	1	410.40**	588.62**	1172.46**	46.22**	51.20**	184.28**
H x A	255	1.50 ^{ns}	2.60**	2.00**	2.70**	1.86**	1.52**
Erro Médio	448	1.30	1.22	0.95	1.76	0.98	0.90
FV	GL	1 x 8	1 x 9	1 x 10	1 x 11	1 x 12	1 x 13
Híbrido	255	12.20**	7.54**	8.26**	8.96**	9.40**	8.42**
Ambiente	1	761.12**	84.02**	1338.36**	2515.26**	2198.02**	1231.96**
H x A	255	3.56**	2.82**	2.96**	2.44**	2.72**	2.16**
Erro Médio	448	2.20	0.99	1.17	1.28	1.23	1.27
FV	GL	2 x 3	2 x 4	2 x 5	2 x 6	2 x 7	2 x 8
Híbrido	255	10.12**	7.62**	10.70**	8.44**	7.50**	12.02**
Ambiente	1	16.02*	195.50**	732.14**	751.58**	44.66**	53.72**
H x A	255	2.58**	2.08**	2.62*	2.00**	1.58**	3.80**
Erro Médio	448	1.49	1.22	2.03	1.25	1.17	2.47
FV	GL	2 e 9	2 x 10	2 e 11	2 x 12	2 e 13	3 x 4
Híbrido	255	7.52**	8.10**	8.76**	9.24**	8.06**	10.64**
Ambiente	1	123.02**	266.50**	893.62**	708.84**	220.24**	99.58**
H x A	255	2.90**	3.18**	2.72**	2.94**	2.58**	1.78**
Erro Médio	448	1.26	1.44	1.55	1.50	1.54	1.14
FV	GL	3 x 5	3 x 6	3 x 7	3 x 8	3 x 9	3 x 10
Híbrido	255	12.74**	11.36**	9.90**	14.46**	10.96**	11.82**
Ambiente	1	964.80**	987.08**	114.18**	11.06 ^{ns}	227.84**	151.82**
H x A	255	3.30**	1.82**	1.90**	4.10**	2.18**	2.18**
Erro Médio	448	1.95	1.17	1.09	2.39	1.18	1.36
FV	GL	3 x 11	3 x 12	3 x 13	4 x 5	4 x 6	4 x 7
Híbrido	255	12.12**	12.16**	10.82**	9.96**	8.94**	7.76**
Ambiente	1	670.32**	511.72**	117.44**	1685.44**	1713.74**	427.06**
H x A	255	2.08**	2.74**	2.54**	3.08**	1.24**	1.06**
Erro Médio	448	1.47	1.43	1.47	1.68	0.90	0.82
FV	GL	4 x 8	4 x 9	4 x 10	4 x 11	4 x 12	4 x 13
Híbrido	255	11.62**	8.52**	9.34**	9.38**	10.14**	8.88**
Ambiente	1	44.24**	628.72**	5.48**	253.16**	159.80**	0.72 ^{ns}
H x A	255	3.94**	1.64**	1.68**	1.82**	1.80**	1.50*
Erro Médio	448	2.12	0.91	1.09	1.20	1.15	1.19

Anexo F – Quadrados médios das análises de variância conjunta para produção de grãos ($t\ ha^{-1}$) de híbridos de milho dos treze ambientes dois a dois

(conclusão)							
FV	GL	5 x 6	5 x 7	5 x 8	5 x 9	5 x 10	5 x 11
Híbrido	255	10.96**	9.92**	16.08**	9.68**	10.32**	10.76**
Ambiente	1	0.12 ^{ns}	415.14**	1182.54**	254.92**	1882.08**	3243.50**
H x A	255	2.84**	2.50**	3.10 ^{ns}	4.08**	4.30**	1.50 ^{ns}
Erro Médio	448	1.71	1.63	2.93	1.72	1.90	2.01
FV	GL	5 x 12	5 x 13	6 x 7	6 x 8	6 x 9	6 x 10
Híbrido	255	11.22**	10.62**	8.42**	13.00**	8.76**	9.32**
Ambiente	1	2881.80**	1755.50**	429.80**	1207.20**	266.44**	1913.18**
H x A	255	2.84**	2.50**	3.10**	4.08**	4.30**	2.44**
Erro Médio	448	1.97	2.01	0.85	2.15	0.94	1.12
FV	GL	6 x 11	6 x 12	6 x 13	7 x 8	7 x 9	7 x 10
Híbrido	255	9.52**	10.38**	9.08**	11.80**	7.56**	8.28**
Ambiente	1	3284.28**	2920.24**	1785.52**	196.36**	19.42**	529.36**
H x A	255	2.42**	2.30**	2.04**	3.12**	1.96**	2.10**
Erro Médio	448	1.23	1.19	1.22	2.07	0.86	1.04
FV	GL	7 x 11	7 x 12	7 x 13	8 x 9	8 x 10	8 x 11
Híbrido	255	8.64**	9.40**	8.24**	11.84**	12.54**	13.16**
Ambiente	1	1337.86**	1109.38**	463.26**	339.36**	80.90**	509.12**
H x A	255	1.92**	1.90**	1.50**	4.44**	4.00**	4.16**
Erro Médio	448	1.15	1.10	1.14	2.16	2.33	2.45
FV	GL	8 e 12	8 x 13	9 x 10	9 x 11	9 x 12	9 x 13
Híbrido	255	13.88**	12.62**	10.42**	9.76**	10.20**	8.74**
Ambiente	1	372.26**	56.40**	751.68**	1679.82**	1422.50**	672.48**
H x A	255	4.16**	3.88**	1.30 ^{ns}	2.14**	2.44**	2.34**
Erro Médio	448	2.40	2.44	1.13	1.24	1.19	1.23
FV	GL	10 x 11	10 x 12	10 x 13	11 x 12	11 x 13	12 x 13
Híbrido	255	10.98**	10.76**	9.56**	11.52**	10.16**	11.48**
Ambiente	1	184.10**	106.06**	2.20 ^{ns}	10.68**	226.58**	138.84**
H x A	255	1.80**	2.74**	2.38**	2.18**	1.96*	1.38 ^{ns}
Erro Médio	448	1.42	1.37	1.41	1.49	1.52	1.48

**/ significativo à 1% de probabilidade; */ significativo à 5% de probabilidade; ^{ns}/ não significativo

Ambiente 1 = ANH_0304; Ambiente 2 = ANH_0405; Ambiente 3 = ARE_0203; Ambiente 4 = ARE_0304; Ambiente 5 = ARE_0405; Ambiente 6 = CAT_0203; Ambiente 7 = CAT_0304; Ambiente 8 = CAT_0405; Ambiente 9 = DEP_0203; Ambiente 10 = DEP_0304; Ambiente 11 = DEP_0405; Ambiente 12 = PTM1_0405; Ambiente 13 = PTM2_0405

ANEXO G – Estimativas da média geral ($\hat{\beta}$) em t ha⁻¹, capacidades geral de combinação das linhagens testadoras ($\hat{g}_T$), das linhagens endogâmicas ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{g}_H$) obtidos com os diversos os coeficientes de parentesco

(continua)

Estimativas	LR	QG	R	W	MR
$\hat{\beta}$	8.7427	8.7391	8.6230	8.2848	6.0120
$\hat{g}_{L_1}$	0.3732	0.3999	0.8722	-0.0516	-0.0372
$\hat{g}_{L_2}$	0.0532	0.0909	0.6393	-0.3332	-0.2312
$\hat{g}_{L_3}$	0.4908	0.5266	0.3077	0.0682	0.0595
$\hat{g}_{L_4}$	0.5798	0.6187	0.4558	0.1694	0.0968
$\hat{g}_{L_5}$	1.0408	1.0549	0.6862	0.4330	0.3668
$\hat{g}_{L_6}$	0.3928	0.4205	0.6487	-0.0391	-0.0620
$\hat{g}_{L_7}$	0.7038	0.7201	0.6981	0.2231	0.1567
$\hat{g}_{L_8}$	0.9943	1.0062	0.3328	0.3745	0.2879
$\hat{g}_{L_9}$	0.4241	0.4379	0.4437	0.1103	-0.0002
$\hat{g}_{L_{10}}$	0.7992	0.7898	-0.0346	0.3965	0.2345
$\hat{g}_{L_{11}}$	0.5888	0.6119	0.2905	0.0389	0.0646
$\hat{g}_{L_{12}}$	-0.9525	-0.9758	-0.7206	-0.6862	-0.6486
$\hat{g}_{L_{13}}$	-0.6013	-0.6035	-1.2064	-0.3730	-0.3870
$\hat{g}_{L_{14}}$	-0.6451	-0.6695	-1.2275	-0.2003	-0.3787
$\hat{g}_{L_{15}}$	-0.6598	-0.6656	-1.5623	-0.1700	-0.4159
$\hat{g}_{L_{16}}$	-0.4105	-0.4380	-1.0570	-0.2892	-0.3249
$\hat{g}_{L_{17}}$	0.1729	0.1603	0.4587	0.1134	0.0269
$\hat{g}_{L_{18}}$	0.1390	0.1411	0.0703	0.0996	0.0754
$\hat{g}_{L_{19}}$	0.3439	0.3332	0.2163	0.2213	0.1308
$\hat{g}_{L_{20}}$	-0.3732	-0.3838	0.3505	-0.1398	-0.2932
$\hat{g}_{L_{21}}$	-0.2511	-0.2632	0.2302	-0.1781	-0.1907
$\hat{g}_{L_{22}}$	-0.4510	-0.4647	-0.4617	-0.1648	-0.2285
$\hat{g}_{L_{23}}$	-0.0790	-0.0994	0.2837	-0.1288	-0.0848
$\hat{g}_{L_{24}}$	-0.1732	-0.1718	-0.8539	0.0690	-0.0734
$\hat{g}_{L_{25}}$	-0.1758	-0.1658	-0.4243	0.0882	-0.0645
$\hat{g}_{L_{26}}$	0.4202	0.4295	-0.1318	0.2549	0.0906
$\hat{g}_{L_{27}}$	0.5141	0.4955	-0.1876	0.3830	0.1289
$\hat{g}_{L_{28}}$	0.1208	0.1302	-0.3552	0.0015	0.0444
$\hat{g}_{L_{29}}$	0.0501	0.0401	0.0528	0.0420	0.0205
$\hat{g}_{L_{30}}$	0.2694	0.2667	0.1644	0.2332	0.1003
$\hat{g}_{L_{31}}$	-0.1749	-0.1846	-0.3763	-0.1499	-0.1918
$\hat{g}_{L_{32}}$	-0.5735	-0.5610	-0.4613	-0.2450	-0.2862
$\hat{g}_{L_{33}}$	-0.4945	-0.5068	-0.3880	-0.3176	-0.3174
$\hat{g}_{L_{34}}$	-0.4171	-0.4413	-0.6407	-0.1767	-0.2809
$\hat{g}_{L_{35}}$	0.0136	0.0016	-0.6222	0.1595	-0.1423
$\hat{g}_{L_{36}}$	-0.8325	-0.8191	-0.4811	-0.2582	-0.4352
$\hat{g}_{L_{37}}$	-0.2605	-0.2611	-0.9611	-0.0432	-0.2033
$\hat{g}_{L_{38}}$	0.3986	0.4225	-0.5232	0.3731	0.1256

ANEXO G – Estimativas da média geral ($\hat{\beta}$) em t ha⁻¹, capacidades geral de combinação das linhagens testadoras ($\hat{g}_T$), das linhagens endogâmicas ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos os coeficientes de parentesco

(continuação)

Estimativas	LR	QG	R	W	MR
$\hat{g}_{L39}$	-0.5124	-0.5213	-1.0449	-0.1906	-0.2580
$\hat{g}_{L40}$	-0.2658	-0.2842	0.4519	-0.0103	-0.0323
$\hat{g}_{L41}$	0.1733	0.1526	0.4039	0.0432	-0.0115
$\hat{g}_{L42}$	0.1203	0.1021	0.1331	0.1084	-0.0297
$\hat{g}_{L43}$	0.1405	0.1505	0.1359	0.0716	0.0135
$\hat{g}_{L44}$	0.3340	0.3341	0.1611	0.1917	0.1486
$\hat{g}_{L45}$	0.3858	0.3850	0.0513	0.2378	0.0691
$\hat{g}_{L46}$	0.1215	0.1263	0.0363	0.1018	0.0251
$\hat{g}_{L47}$	-0.1539	-0.1567	-0.3413	-0.0600	-0.1443
$\hat{g}_{L48}$	-0.3747	-0.3763	0.1411	-0.1014	-0.2078
$\hat{g}_{L49}$	0.2056	0.1995	0.0478	0.1880	0.0979
$\hat{g}_{L50}$	0.3264	0.3161	0.1644	0.2017	0.1267
$\hat{g}_{T1}$	-0.2484	-0.2467	-0.1051	-0.1638	0.0400
$\hat{g}_{T2}$	0.5157	0.5190	0.2508	0.2784	0.0689
$\hat{g}_{T3}$	-0.0904	-0.0820	-0.4504	-0.0555	-0.0913
$\hat{g}_{T4}$	-0.5819	-0.5814	-0.5177	-0.3762	-0.1411
$\hat{g}_{T5}$	0.4372	0.4410	0.7445	0.3170	0.1214
$\hat{s}_{H1}$	-0.3626	-0.3815	0.1373	0.4499	2.5063
$\hat{s}_{H2}$	-0.5782	-0.6004	0.1341	-0.2088	1.7560
$\hat{s}_{H3}$	-0.0347	-0.0585	0.0782	1.1820	3.2624
$\hat{s}_{H4}$	-0.2293	-0.2559	0.1225	0.4158	2.5537
$\hat{s}_{H5}$	-0.1927	-0.2056	0.1573	0.8536	2.9873
$\hat{s}_{H6}$	-0.3008	-0.3207	0.1731	0.1787	2.2682
$\hat{s}_{H7}$	0.1247	0.1092	0.1892	1.2513	3.3848
$\hat{s}_{H8}$	0.1447	0.1326	0.1728	1.3898	3.5477
$\hat{s}_{H9}$	-0.5400	-0.5503	0.1227	-0.0230	2.1517
$\hat{s}_{H10}$	0.0875	0.0952	0.1860	1.3392	3.5689
$\hat{s}_{H11}$	-0.0455	-0.0607	0.1266	1.0361	3.0817
$\hat{s}_{H12}$	-0.0961	-0.0761	-0.1990	1.0143	3.0508
$\hat{s}_{H13}$	-0.1719	-0.1665	-0.1677	0.6600	2.7505
$\hat{s}_{H14}$	-0.2624	-0.2429	-0.2416	0.6393	2.8911
$\hat{s}_{H15}$	-0.1511	-0.1438	-0.2422	0.4730	2.7889
$\hat{s}_{H16}$	0.0453	0.0669	-0.1303	0.6908	2.7995
$\hat{s}_{H17}$	-0.1961	-0.1861	0.1683	-0.0924	2.0639
$\hat{s}_{H18}$	-0.1578	-0.1575	0.2128	-0.1018	1.9933
$\hat{s}_{H19}$	-0.6867	-0.6743	0.2951	-1.4392	0.7113
$\hat{s}_{H20}$	-0.5216	-0.5119	0.0039	-0.8869	1.3274
$\hat{s}_{H21}$	-0.3361	-0.3263	0.1375	-0.3467	1.7324

ANEXO G – Estimativas da média geral ($\hat{\beta}$) em t ha⁻¹, capacidades geral de combinação das linhagens testadoras ($\hat{g}_T$), das linhagens endogâmicas ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{g}_H$) obtidos com os diversos os coeficientes de parentesco

(continuação)					
Estimativas	LR	QG	R	W	MR
$\hat{S}_{H22}$	-0.5553	-0.5440	-0.0099	-0.9364	1.1900
$\hat{S}_{H23}$	-0.3860	-0.3701	0.2918	-0.5751	1.4414
$\hat{S}_{H24}$	-0.2517	-0.2498	0.0554	-0.3683	1.8419
$\hat{S}_{H25}$	-0.5207	-0.5254	0.0870	-1.0826	1.1297
$\hat{S}_{H26}$	0.2876	0.2826	0.1289	1.0180	3.2550
$\hat{S}_{H27}$	0.2387	0.2536	0.0816	0.4101	2.7329
$\hat{S}_{H28}$	0.0320	0.0286	0.1434	0.5528	2.5836
$\hat{S}_{H29}$	-0.3239	-0.3143	0.1113	-0.1077	1.9803
$\hat{S}_{H30}$	-0.0469	-0.0420	0.1238	0.3392	2.5376
$\hat{S}_{H31}$	0.4345	0.4444	-0.0694	1.1328	3.2528
$\hat{S}_{H32}$	-0.0846	-0.0912	-0.1456	0.0146	2.1219
$\hat{S}_{H33}$	0.0705	0.0813	-0.1009	0.3694	2.4434
$\hat{S}_{H34}$	0.2863	0.3047	-0.1178	0.5590	2.7365
$\hat{S}_{H35}$	0.7331	0.7423	-0.0739	1.5824	3.9723
$\hat{S}_{H36}$	-0.0590	-0.0660	-0.1819	-0.2369	2.0025
$\hat{S}_{H37}$	0.2920	0.2948	-0.0552	0.4987	2.7304
$\hat{S}_{H38}$	0.5741	0.5589	0.1084	1.0568	3.3702
$\hat{S}_{H39}$	-0.0478	-0.0403	-0.0784	-0.3429	1.7874
$\hat{S}_{H40}$	-0.6705	-0.6545	-0.0401	-0.1623	1.9212
$\hat{S}_{H41}$	0.3418	0.3548	0.0631	0.9318	3.0591
$\hat{S}_{H42}$	0.2183	0.2327	-0.0035	0.6557	2.8679
$\hat{S}_{H43}$	0.4095	0.4059	0.0826	0.9345	3.0625
$\hat{S}_{H44}$	0.5824	0.5833	0.1475	1.4481	3.5600
$\hat{S}_{H45}$	0.0481	0.0508	0.1350	-0.0137	2.2189
$\hat{S}_{H46}$	0.0695	0.0688	0.0938	0.3173	2.4617
$\hat{S}_{H47}$	0.3775	0.3819	-0.0518	0.8692	3.0282
$\hat{S}_{H48}$	-0.6970	-0.6918	0.0017	-1.6194	0.5452
$\hat{S}_{H49}$	0.0071	0.0145	0.1163	0.4309	2.5874
$\hat{S}_{H50}$	-0.2771	-0.2664	0.1701	-0.1195	2.0263
$\hat{S}_{H51}$	-0.1704	-0.1921	0.1368	0.4688	2.9305
$\hat{S}_{H52}$	-0.2730	-0.3011	0.0873	0.4331	2.8088
$\hat{S}_{H53}$	-0.2825	-0.3084	0.0938	0.5048	2.9929
$\hat{S}_{H54}$	-0.2239	-0.2531	0.1205	0.6283	3.1853
$\hat{S}_{H55}$	0.4785	0.4599	0.1223	2.2906	4.8403
$\hat{S}_{H56}$	-0.0690	-0.0908	0.1437	1.1794	3.6877
$\hat{S}_{H57}$	0.1549	0.1374	0.1769	1.4283	3.9793
$\hat{S}_{H58}$	0.0934	0.0801	0.1600	1.4600	4.0279
$\hat{S}_{H59}$	0.0865	0.0724	0.0747	1.3671	3.9633
$\hat{S}_{H60}$	-0.0327	-0.0258	0.1909	1.0124	3.6557

ANEXO G – Estimativas da média geral ($\hat{\beta}$) em t ha⁻¹, capacidades geral de combinação das linhagens testadoras ($\hat{g}_T$), das linhagens endogâmicas ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos os coeficientes de parentesco

(continuação)					
Estimativas	LR	QG	R	W	MR
$\hat{s}_{H61}$	-0.7197	-0.7369	0.1537	-0.2137	2.2361
$\hat{s}_{H62}$	-2.4920	-2.4738	0.0110	-3.4141	-0.9793
$\hat{s}_{H63}$	-2.3250	-2.3197	0.0104	-2.9288	-0.4419
$\hat{s}_{H64}$	-2.1732	-2.1544	-0.0282	-2.4311	0.2189
$\hat{s}_{H65}$	-1.7790	-1.7729	-0.0219	-1.4604	1.2697
$\hat{s}_{H66}$	-1.3319	-1.3115	-0.0277	-0.8975	1.6147
$\hat{s}_{H67}$	0.3284	0.3370	0.0469	1.4665	4.0376
$\hat{s}_{H68}$	0.1475	0.1461	0.0606	0.6864	3.1870
$\hat{s}_{H69}$	0.8754	0.8842	0.0482	2.2188	4.7898
$\hat{s}_{H70}$	-0.3129	-0.3055	-0.0480	0.5121	3.1478
$\hat{s}_{H71}$	-0.2854	-0.2762	0.0881	0.2278	2.7192
$\hat{s}_{H72}$	-0.7434	-0.7326	-0.0734	-0.7153	1.8216
$\hat{s}_{H73}$	0.5109	0.5251	0.0380	1.4012	3.8352
$\hat{s}_{H74}$	-0.3127	-0.3123	-0.0270	0.3476	2.9795
$\hat{s}_{H75}$	0.0443	0.0383	-0.0069	0.6661	3.3112
$\hat{s}_{H76}$	0.1078	0.1007	0.1053	1.2664	3.9146
$\hat{s}_{H77}$	-0.1080	-0.0945	0.0748	1.4661	4.2128
$\hat{s}_{H78}$	0.1211	0.1157	0.1333	1.1954	3.6319
$\hat{s}_{H79}$	0.2275	0.2352	0.0234	1.4089	3.9147
$\hat{s}_{H80}$	-0.3061	-0.3013	0.0839	0.1200	2.7334
$\hat{s}_{H81}$	-0.8804	-0.8691	-0.0028	-0.2906	2.2337
$\hat{s}_{H82}$	-1.2009	-1.2076	-0.0868	-1.2151	1.2997
$\hat{s}_{H83}$	-1.1352	-1.1246	-0.0582	-0.7123	1.7684
$\hat{s}_{H84}$	-1.3062	-1.2883	-0.0364	-0.8959	1.6918
$\hat{s}_{H85}$	-1.1979	-1.1885	0.0567	-1.1073	1.6762
$\hat{s}_{H86}$	-1.2193	-1.2266	-0.1419	-0.8424	1.8216
$\hat{s}_{H87}$	-0.8574	-0.8545	-0.0510	0.1567	2.8075
$\hat{s}_{H88}$	0.0556	0.0390	0.0786	1.7598	4.5034
$\hat{s}_{H89}$	-1.1052	-1.0978	-0.0520	-0.4212	2.1359
$\hat{s}_{H90}$	0.0500	0.0656	-0.0588	1.0813	3.5967
$\hat{s}_{H91}$	-0.3464	-0.3335	0.0769	0.5113	3.0521
$\hat{s}_{H92}$	-0.5130	-0.4986	0.0373	0.4415	3.0676
$\hat{s}_{H93}$	0.0584	0.0543	0.0705	1.2663	3.8101
$\hat{s}_{H94}$	-0.0759	-0.0750	0.1159	0.7870	3.3123
$\hat{s}_{H95}$	0.0273	0.0276	0.1005	1.1944	3.8473
$\hat{s}_{H96}$	-0.0927	-0.0946	0.0077	0.8031	3.3634
$\hat{s}_{H97}$	-0.8910	-0.8868	-0.0106	-0.5470	2.0200
$\hat{s}_{H98}$	-0.6083	-0.6056	-0.0613	-0.1825	2.4089
$\hat{s}_{H99}$	-0.1839	-0.1775	0.0748	0.2730	2.8434

ANEXO G – Estimativas da média geral ($\hat{\beta}$) em t ha⁻¹, capacidades geral de combinação das linhagens testadoras ($\hat{g}_T$), das linhagens endogâmicas ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos os coeficientes de parentesco

(continuação)					
Estimativas	LR	QG	R	W	MR
$\hat{s}_{H100}$	0.1404	0.1488	0.0714	1.0279	3.5819
$\hat{s}_{H101}$	-0.5093	-0.5376	-0.0093	0.4195	2.7087
$\hat{s}_{H102}$	-0.9022	-0.9350	-0.0509	-0.4469	1.7551
$\hat{s}_{H103}$	-0.8702	-0.9011	-0.0371	-0.2376	2.0743
$\hat{s}_{H104}$	-0.2554	-0.2915	-0.0324	1.2679	3.6555
$\hat{s}_{H105}$	0.0852	0.0611	-0.0167	2.1393	4.5153
$\hat{s}_{H106}$	-0.8016	-0.8285	0.0143	-0.0352	2.2946
$\hat{s}_{H107}$	-0.5521	-0.5746	0.0427	0.2723	2.6486
$\hat{s}_{H108}$	-0.1056	-0.1252	0.0039	1.7605	4.1611
$\hat{s}_{H109}$	-0.6614	-0.6793	-0.0510	0.0425	2.4621
$\hat{s}_{H110}$	-0.5321	-0.5293	0.0639	0.2704	2.7406
$\hat{s}_{H111}$	-0.2850	-0.3091	-0.0193	1.2771	3.5669
$\hat{s}_{H112}$	-2.9087	-2.8930	-0.1726	-4.1542	-1.8949
$\hat{s}_{H113}$	-2.4991	-2.4979	-0.0854	-3.2985	-0.9888
$\hat{s}_{H114}$	-2.9546	-2.9387	-0.0226	-4.8221	-2.3486
$\hat{s}_{H115}$	-3.0093	-3.0044	-0.0556	-5.0075	-2.4674
$\hat{s}_{H116}$	-1.9919	-1.9746	-0.0905	-2.8553	-0.5244
$\hat{s}_{H117}$	0.0166	0.0218	-0.0853	0.8357	3.2296
$\hat{s}_{H118}$	0.3636	0.3571	-0.0964	1.4778	3.8067
$\hat{s}_{H119}$	0.3360	0.3418	-0.0839	1.0816	3.4777
$\hat{s}_{H120}$	-0.6116	-0.6078	-0.1770	-0.0988	2.3689
$\hat{s}_{H121}$	-0.3377	-0.3328	-0.0890	0.2239	2.5495
$\hat{s}_{H122}$	-0.8356	-0.8287	-0.2197	-0.7511	1.6183
$\hat{s}_{H123}$	0.3432	0.3533	-0.0831	1.2060	3.4696
$\hat{s}_{H124}$	-0.8896	-0.8928	-0.1529	-0.9171	1.5357
$\hat{s}_{H125}$	-0.2683	-0.2784	-0.1471	0.0224	2.4924
$\hat{s}_{H126}$	-0.2773	-0.2893	-0.0094	0.4332	2.9081
$\hat{s}_{H127}$	-0.4519	-0.4424	-0.0569	0.6313	3.2034
$\hat{s}_{H128}$	-0.0147	-0.0242	-0.0021	1.0309	3.2986
$\hat{s}_{H129}$	-0.0193	-0.0160	-0.1143	1.0379	3.3729
$\hat{s}_{H130}$	-0.1705	-0.1703	-0.0935	0.6971	3.1460
$\hat{s}_{H131}$	-1.4057	-1.3983	-0.1050	-1.6160	0.7301
$\hat{s}_{H132}$	-1.7008	-1.7116	-0.2152	-2.4427	-0.1059
$\hat{s}_{H133}$	-1.4980	-1.4908	-0.2013	-1.6726	0.6288
$\hat{s}_{H134}$	-1.8419	-1.8271	-0.1060	-2.6162	-0.2100
$\hat{s}_{H135}$	-1.3923	-1.3870	-0.0894	-1.3297	1.2843
$\hat{s}_{H136}$	-1.6977	-1.7088	-0.2210	-1.8565	0.6330
$\hat{s}_{H137}$	-1.0555	-1.0570	-0.1616	-0.4464	2.0270
$\hat{s}_{H138}$	-1.0077	-1.0267	-0.0160	-0.7796	1.7778

ANEXO G – Estimativas da média geral ($\hat{\beta}$) em t ha⁻¹, capacidades geral de combinação das linhagens testadoras ($\hat{g}_T$), das linhagens endogâmicas ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos os coeficientes de parentesco

(continuação)					
Estimativas	LR	QG	R	W	MR
$\hat{s}_{H139}$	-1.6255	-1.6216	-0.1887	-1.8836	0.4913
$\hat{s}_{H140}$	-0.3173	-0.3043	-0.1949	0.1833	2.5204
$\hat{s}_{H141}$	-0.5389	-0.5308	-0.0373	0.0951	2.4596
$\hat{s}_{H142}$	-0.7243	-0.7145	-0.0810	-0.2155	2.2354
$\hat{s}_{H143}$	-0.5211	-0.5286	-0.0575	0.1076	2.4753
$\hat{s}_{H144}$	-0.2320	-0.2359	-0.0174	0.6131	2.9669
$\hat{s}_{H145}$	-0.2145	-0.2185	-0.0281	0.5866	3.0697
$\hat{s}_{H146}$	-0.0523	-0.0594	-0.1410	0.9972	3.3880
$\hat{s}_{H147}$	-0.4193	-0.4213	-0.1547	0.6173	3.0219
$\hat{s}_{H148}$	-0.4544	-0.4567	-0.2105	0.0598	2.4835
$\hat{s}_{H149}$	-0.1938	-0.1920	-0.1011	0.5364	2.9393
$\hat{s}_{H150}$	0.2602	0.2628	-0.0914	1.4469	3.8338
$\hat{s}_{H151}$	-0.2679	-0.2864	0.0914	0.9062	2.9320
$\hat{s}_{H152}$	-0.3383	-0.3628	0.0676	0.5063	2.4376
$\hat{s}_{H153}$	-0.4291	-0.4514	0.0601	0.5240	2.5685
$\hat{s}_{H154}$	-0.2456	-0.2719	0.0932	0.4573	2.5633
$\hat{s}_{H155}$	-0.3940	-0.4068	0.1097	0.4144	2.5178
$\hat{s}_{H156}$	-0.1888	-0.2078	0.1116	0.8586	2.9181
$\hat{s}_{H157}$	-0.2352	-0.2489	0.1505	0.5521	2.6550
$\hat{s}_{H158}$	-0.2235	-0.2329	0.1429	0.6420	2.7652
$\hat{s}_{H159}$	-0.5590	-0.5692	0.0518	-0.0022	2.1432
$\hat{s}_{H160}$	0.0874	0.0951	0.1821	1.0343	3.2402
$\hat{s}_{H161}$	-0.2144	-0.2296	0.0942	0.8748	2.8831
$\hat{s}_{H162}$	-0.6637	-0.6423	-0.2316	-0.3759	1.6282
$\hat{s}_{H163}$	-0.0933	-0.0860	-0.2081	1.1632	3.2241
$\hat{s}_{H164}$	-0.2812	-0.2598	-0.2541	0.2830	2.5119
$\hat{s}_{H165}$	-0.1847	-0.1756	-0.2522	0.0082	2.3015
$\hat{s}_{H166}$	0.0921	0.1149	-0.1344	1.0713	3.1550
$\hat{s}_{H167}$	-0.2130	-0.2008	0.0752	-0.1685	1.9551
$\hat{s}_{H168}$	-0.1987	-0.1965	0.0893	-0.0709	1.9932
$\hat{s}_{H169}$	-0.2473	-0.2331	0.1405	0.1035	2.2308
$\hat{s}_{H170}$	-0.6406	-0.6300	-0.0493	-1.1194	1.0692
$\hat{s}_{H171}$	-0.1351	-0.1241	0.0519	0.0105	2.0591
$\hat{s}_{H172}$	-0.5424	-0.5295	-0.1158	-0.5969	1.5074
$\hat{s}_{H173}$	-0.4097	-0.3909	0.1257	-0.3701	1.6170
$\hat{s}_{H174}$	-0.1036	-0.1004	-0.0740	0.4451	2.6257
$\hat{s}_{H175}$	-0.1300	-0.1331	-0.0041	-0.1941	1.9923
$\hat{s}_{H176}$	-0.6664	-0.6686	0.1363	-1.0527	1.1419
$\hat{s}_{H177}$	-0.3051	-0.2870	0.0724	0.0929	2.3834

ANEXO G – Estimativas da média geral ($\hat{\beta}$) em t ha⁻¹, capacidades geral de combinação das linhagens testadoras ($\hat{g}_T$), das linhagens endogâmicas ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos os coeficientes de parentesco

(continuação)					
Estimativas	LR	QG	R	W	MR
$\hat{s}_{H178}$	-0.2955	-0.2969	0.1648	-0.9840	1.0036
$\hat{s}_{H179}$	-0.6633	-0.6510	0.1250	-1.8255	0.2214
$\hat{s}_{H180}$	0.2041	0.2101	0.0444	0.7121	2.8845
$\hat{s}_{H181}$	0.0640	0.0771	-0.0449	0.9167	3.0016
$\hat{s}_{H182}$	0.0771	0.0709	-0.2081	0.7115	2.8046
$\hat{s}_{H183}$	-0.1323	-0.1196	-0.1230	0.1415	2.1839
$\hat{s}_{H184}$	-0.1138	-0.0926	-0.1263	0.3196	2.4677
$\hat{s}_{H185}$	0.2199	0.2321	-0.0343	1.0647	3.4063
$\hat{s}_{H186}$	-0.0584	-0.0651	-0.2089	-0.0062	2.2172
$\hat{s}_{H187}$	-0.3326	-0.3263	-0.0866	-0.6202	1.5794
$\hat{s}_{H188}$	0.4040	0.3899	0.0792	0.7559	3.0418
$\hat{s}_{H189}$	0.0403	0.0491	-0.1260	0.4201	2.5280
$\hat{s}_{H190}$	0.4163	0.4295	0.0822	-0.9923	1.0616
$\hat{s}_{H191}$	-0.2077	-0.1922	0.0966	-0.1618	1.9249
$\hat{s}_{H192}$	-0.3820	-0.3652	0.0077	-0.4777	1.6957
$\hat{s}_{H193}$	-0.2205	-0.2216	0.1229	-0.4782	1.6094
$\hat{s}_{H194}$	-0.1186	-0.1152	0.1610	-0.1792	1.8943
$\hat{s}_{H195}$	0.2128	0.2154	0.1484	0.6431	2.8470
$\hat{s}_{H196}$	-0.3700	-0.3695	0.0327	-0.8578	1.2514
$\hat{s}_{H197}$	-0.5921	-0.5850	-0.0245	-1.2474	0.8638
$\hat{s}_{H198}$	-0.1313	-0.1269	-0.0708	-0.5037	1.6377
$\hat{s}_{H199}$	-0.0299	-0.0216	0.0348	-0.0119	2.1142
$\hat{s}_{H200}$	0.0669	0.0785	0.1205	-0.2223	1.8866
$\hat{s}_{H201}$	0.5929	0.5691	0.1561	1.3972	3.8560
$\hat{s}_{H202}$	0.6740	0.6445	0.1136	1.5446	3.9174
$\hat{s}_{H203}$	0.7672	0.7390	0.1086	1.9292	4.4139
$\hat{s}_{H204}$	0.8661	0.8339	0.1422	1.8435	4.3890
$\hat{s}_{H205}$	0.9058	0.8866	0.1495	2.1241	4.6677
$\hat{s}_{H206}$	0.7973	0.7732	0.1674	1.6077	4.1006
$\hat{s}_{H207}$	1.0485	1.0288	0.1960	2.1975	4.7366
$\hat{s}_{H208}$	0.8988	0.8834	0.2005	2.0714	4.6283
$\hat{s}_{H209}$	0.7919	0.7758	0.1046	1.8442	4.4332
$\hat{s}_{H210}$	0.8753	0.8794	0.2024	2.5339	5.1731
$\hat{s}_{H211}$	0.8489	0.8283	0.1610	1.9700	4.4106
$\hat{s}_{H212}$	-0.1400	-0.1241	-0.1228	0.3278	2.7513
$\hat{s}_{H213}$	-0.4581	-0.4552	-0.1079	-0.4026	2.0667
$\hat{s}_{H214}$	-0.6359	-0.6188	-0.1533	-0.3596	2.2834
$\hat{s}_{H215}$	-0.7896	-0.7835	-0.1512	-0.6370	2.0729

ANEXO G – Estimativas da média geral ($\hat{\beta}$) em t ha⁻¹, capacidades geral de combinação das linhagens testadoras ($\hat{g}_T$), das linhagens endogâmicas ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos os coeficientes de parentesco

Estimativas	(conclusão)				
	LR	QG	R	W	MR
$\hat{s}_{H216}$	-0.4711	-0.4522	-0.0311	-0.2699	2.2252
$\hat{s}_{H217}$	0.3055	0.3147	0.1381	1.0611	3.6201
$\hat{s}_{H218}$	-0.0375	-0.0377	0.1247	0.8608	3.3614
$\hat{s}_{H219}$	0.9874	0.9961	0.1684	2.9373	5.5190
$\hat{s}_{H220}$	-0.1345	-0.1267	0.0617	-0.0338	2.5930
$\hat{s}_{H221}$	0.0256	0.0346	0.1174	-0.0309	2.4487
$\hat{s}_{H222}$	0.0392	0.0483	-0.0436	0.6397	3.1815
$\hat{s}_{H223}$	0.0899	0.1051	0.1525	0.5816	3.0203
$\hat{s}_{H224}$	-0.2172	-0.2166	0.0192	-0.3004	2.3043
$\hat{s}_{H225}$	0.1432	0.1383	0.0292	0.6880	3.3089
$\hat{s}_{H226}$	0.9577	0.9496	0.1352	2.5915	5.2333
$\hat{s}_{H227}$	0.3435	0.3573	0.1533	1.1279	3.8456
$\hat{s}_{H228}$	0.3844	0.3801	0.1716	1.2545	3.6815
$\hat{s}_{H229}$	0.4646	0.4725	0.0919	1.6100	4.1091
$\hat{s}_{H230}$	0.6320	0.6358	0.1056	1.5157	4.1220
$\hat{s}_{H231}$	-0.5651	-0.5541	0.0780	-0.7251	1.7678
$\hat{s}_{H232}$	-0.6377	-0.6451	-0.0843	-0.8030	1.7088
$\hat{s}_{H233}$	-0.6240	-0.6133	-0.0202	-0.6732	1.7813
$\hat{s}_{H234}$	-0.1973	-0.1801	-0.0236	0.2435	2.8040
$\hat{s}_{H235}$	-0.2975	-0.2885	0.0982	-0.1811	2.5728
$\hat{s}_{H236}$	-1.6538	-1.6598	-0.0218	-3.4558	-0.8277
$\hat{s}_{H237}$	-0.6725	-0.6692	0.0421	-0.9182	1.6946
$\hat{s}_{H238}$	0.0765	0.0603	0.1685	0.6059	3.3231
$\hat{s}_{H239}$	-0.7648	-0.7575	0.0122	-1.1114	1.4106
$\hat{s}_{H240}$	-0.0086	0.0054	0.0364	-0.1518	2.3334
$\hat{s}_{H241}$	0.3664	0.3790	0.1452	1.2624	3.7841
$\hat{s}_{H242}$	0.0927	0.1066	0.1023	0.8202	3.4219
$\hat{s}_{H243}$	0.0579	0.0549	0.1502	0.5949	3.1266
$\hat{s}_{H244}$	0.3884	0.3895	0.1902	1.2119	3.7293
$\hat{s}_{H245}$	0.5912	0.5920	0.1968	1.7754	4.4215
$\hat{s}_{H246}$	0.2021	0.1996	0.0733	1.0710	3.6201
$\hat{s}_{H247}$	-0.1840	-0.1802	0.0330	0.3476	2.8908
$\hat{s}_{H248}$	-0.0334	-0.0310	0.0026	0.6648	3.2405
$\hat{s}_{H249}$	0.5143	0.5202	0.0963	1.7151	4.2800
$\hat{s}_{H250}$	0.5161	0.5251	0.1105	2.0229	4.5739

* QG: Queller e Goodnight; R: Ritland; LR: Lynch e Ritland; W: Wang; VR: VanRaden

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

(continua)

Estimativas	J	SD	MR	I
$\hat{\beta}$	9.1068	9.2795	6.0120	8.5240
$\hat{g}_{L_1}$	0.3833	0.3053	-0.0372	0.3362
$\hat{g}_{L_2}$	-0.0739	-0.0521	-0.2312	-0.1650
$\hat{g}_{L_3}$	0.3335	0.1994	0.0595	0.4703
$\hat{g}_{L_4}$	0.8090	0.7002	0.0968	0.6585
$\hat{g}_{L_5}$	0.9511	0.7440	0.3668	1.3616
$\hat{g}_{L_6}$	0.3783	0.3169	-0.0620	0.3686
$\hat{g}_{L_7}$	0.8162	0.7117	0.1567	0.8682
$\hat{g}_{L_8}$	1.0494	0.8767	0.2879	1.2367
$\hat{g}_{L_9}$	0.3557	0.2554	-0.0002	0.3994
$\hat{g}_{L_{10}}$	1.0149	0.9145	0.2345	1.0794
$\hat{g}_{L_{11}}$	0.4547	0.3656	0.0646	0.6076
$\hat{g}_{L_{12}}$	-1.3427	-1.4972	-0.6486	-1.7379
$\hat{g}_{L_{13}}$	-1.3561	-1.5472	-0.3870	-1.2167
$\hat{g}_{L_{14}}$	-1.6774	-1.8959	-0.3787	-1.3724
$\hat{g}_{L_{15}}$	-1.6149	-1.7891	-0.4159	-1.3385
$\hat{g}_{L_{16}}$	-0.6506	-0.7938	-0.3249	-0.7587
$\hat{g}_{L_{17}}$	0.5072	0.4444	0.0269	0.3824
$\hat{g}_{L_{18}}$	0.4133	0.3249	0.0754	0.3331
$\hat{g}_{L_{19}}$	0.6266	0.4661	0.1308	0.7448
$\hat{g}_{L_{20}}$	-0.4160	-0.4545	-0.2932	-0.5443
$\hat{g}_{L_{21}}$	-0.0873	-0.0589	-0.1907	-0.3115
$\hat{g}_{L_{22}}$	-0.5746	-0.6883	-0.2285	-0.6770
$\hat{g}_{L_{23}}$	0.0711	0.0430	-0.0848	0.0601
$\hat{g}_{L_{24}}$	-0.2102	-0.4044	-0.0734	-0.2529
$\hat{g}_{L_{25}}$	-0.0105	-0.0904	-0.0645	-0.1002
$\hat{g}_{L_{26}}$	0.8138	0.7007	0.0906	0.6708
$\hat{g}_{L_{27}}$	0.5895	0.5371	0.1289	0.6893
$\hat{g}_{L_{28}}$	0.2373	0.1896	0.0444	0.2868
$\hat{g}_{L_{29}}$	0.3730	0.2853	0.0205	0.1758
$\hat{g}_{L_{30}}$	0.5831	0.4780	0.1003	0.5197
$\hat{g}_{L_{31}}$	-0.3767	-0.4218	-0.1918	-0.3917
$\hat{g}_{L_{32}}$	-0.9458	-0.9852	-0.2862	-0.9515
$\hat{g}_{L_{33}}$	-0.9508	-1.0564	-0.3174	-0.8250
$\hat{g}_{L_{34}}$	-0.6927	-0.8143	-0.2809	-0.6908
$\hat{g}_{L_{35}}$	0.1473	-0.0305	-0.1423	-0.0555
$\hat{g}_{L_{36}}$	-1.5466	-1.6688	-0.4352	-1.3724
$\hat{g}_{L_{37}}$	-0.2422	-0.3533	-0.2033	-0.4225
$\hat{g}_{L_{38}}$	0.5016	0.3718	0.1256	0.6307

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

(continuação)				
Estimativas	J	SD	MR	I
$\hat{g}_{L_{39}}$	-0.9592	-1.0213	-0.2580	-0.8481
$\hat{g}_{L_{40}}$	-0.2358	-0.3077	-0.0323	-0.1989
$\hat{g}_{L_{41}}$	0.2129	0.0875	-0.0115	0.2560
$\hat{g}_{L_{42}}$	-0.1043	-0.0911	-0.0297	0.0894
$\hat{g}_{L_{43}}$	0.1160	-0.0015	0.0135	0.2452
$\hat{g}_{L_{44}}$	0.4307	0.3204	0.1486	0.5629
$\hat{g}_{L_{45}}$	0.3528	0.2740	0.0691	0.6461
$\hat{g}_{L_{46}}$	0.5203	0.4644	0.0251	0.2544
$\hat{g}_{L_{47}}$	-0.2548	-0.3558	-0.1443	-0.2251
$\hat{g}_{L_{48}}$	-0.5021	-0.5676	-0.2078	-0.5073
$\hat{g}_{L_{49}}$	0.2802	0.1821	0.0979	0.4164
$\hat{g}_{L_{50}}$	0.8172	0.7342	0.1267	0.6137
$\hat{g}_{T_1}$	0.1571	0.1038	0.0400	-0.0651
$\hat{g}_{T_2}$	0.1418	0.1569	0.0689	0.3305
$\hat{g}_{T_3}$	-0.8745	-0.8839	-0.0913	-0.5712
$\hat{g}_{T_4}$	-0.4175	-0.3772	-0.1411	-0.5111
$\hat{g}_{T_5}$	1.0009	1.0523	0.1214	0.8170
$\hat{s}_{H_1}$	-1.0926	-1.0717	2.5063	-0.0267
$\hat{s}_{H_2}$	-1.1001	-0.9603	1.7560	-0.0702
$\hat{s}_{H_3}$	-0.5667	-0.7059	3.2624	0.0437
$\hat{s}_{H_4}$	-1.2495	-1.1960	2.5537	-0.0405
$\hat{s}_{H_5}$	-1.0379	-0.9450	2.9873	-0.0398
$\hat{s}_{H_6}$	-0.8822	-0.7928	2.2682	-0.0560
$\hat{s}_{H_7}$	-0.7839	-0.7952	3.3848	0.0265
$\hat{s}_{H_8}$	-0.6401	-0.6578	3.5477	0.0189
$\hat{s}_{H_9}$	-1.3234	-1.2611	2.1517	-0.0639
$\hat{s}_{H_{10}}$	-0.7129	-0.9013	3.5689	0.0312
$\hat{s}_{H_{11}}$	-0.4402	-0.4228	3.0817	0.0130
$\hat{s}_{H_{12}}$	-0.3282	-0.6000	3.0508	0.1687
$\hat{s}_{H_{13}}$	0.4221	0.3109	2.7505	0.1143
$\hat{s}_{H_{14}}$	0.6613	0.5125	2.8911	0.1439
$\hat{s}_{H_{15}}$	0.5090	0.3321	2.7889	0.1271
$\hat{s}_{H_{16}}$	-0.2731	-0.3956	2.7995	0.0805
$\hat{s}_{H_{17}}$	-1.4441	-1.3740	2.0639	-0.0680
$\hat{s}_{H_{18}}$	-1.3576	-1.2476	1.9933	-0.0652
$\hat{s}_{H_{19}}$	-2.1197	-1.6915	0.7113	-0.2254
$\hat{s}_{H_{20}}$	-1.2720	-1.1554	1.3274	-0.0810
$\hat{s}_{H_{21}}$	-1.3968	-1.4396	1.7324	-0.0541

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

(continuação)				
Estimativas	J	SD	MR	I
$\hat{s}_{H22}$	-1.4771	-1.2981	1.1900	-0.0749
$\hat{s}_{H23}$	-1.8671	-1.6159	1.4414	-0.1076
$\hat{s}_{H24}$	-1.2322	-1.1319	1.8419	-0.0374
$\hat{s}_{H25}$	-1.6025	-1.3611	1.1297	-0.1211
$\hat{s}_{H26}$	-1.0094	-1.2113	3.2550	0.0262
$\hat{s}_{H27}$	-0.5768	-0.5516	2.7329	-0.0231
$\hat{s}_{H28}$	-0.9429	-1.1233	2.5836	-0.0064
$\hat{s}_{H29}$	-1.3071	-1.2439	1.9803	-0.0567
$\hat{s}_{H30}$	-1.1800	-1.1850	2.5376	-0.0280
$\hat{s}_{H31}$	-0.2189	-0.4888	3.2528	0.1021
$\hat{s}_{H32}$	-0.7359	-0.9185	2.1219	0.0371
$\hat{s}_{H33}$	-0.2839	-0.3801	2.4434	0.0530
$\hat{s}_{H34}$	-0.1830	-0.2687	2.7365	0.0720
$\hat{s}_{H35}$	-0.0861	-0.2960	3.9723	0.1442
$\hat{s}_{H36}$	0.1788	0.2828	2.0025	0.0518
$\hat{s}_{H37}$	-0.7861	-0.9914	2.7304	0.0527
$\hat{s}_{H38}$	-0.3904	-0.4919	3.3702	0.0447
$\hat{s}_{H39}$	-0.4592	-0.4666	1.7874	-0.0030
$\hat{s}_{H40}$	-1.1063	-1.1744	1.9212	-0.0310
$\hat{s}_{H41}$	-0.6583	-0.7791	3.0591	0.0374
$\hat{s}_{H42}$	0.0218	-0.0473	2.8679	0.0331
$\hat{s}_{H43}$	-0.3831	-0.4952	3.0625	0.0413
$\hat{s}_{H44}$	-0.5069	-0.6410	3.5600	0.0726
$\hat{s}_{H45}$	-0.6500	-0.5696	2.2189	-0.0752
$\hat{s}_{H46}$	-1.2897	-1.3392	2.4617	-0.0168
$\hat{s}_{H47}$	-0.2864	-0.4238	3.0282	0.0685
$\hat{s}_{H48}$	-1.6836	-1.3611	0.5452	-0.1525
$\hat{s}_{H49}$	-0.9836	-1.1282	2.5874	-0.0131
$\hat{s}_{H50}$	-1.6557	-1.5556	2.0263	-0.0847
$\hat{s}_{H51}$	-0.2973	-0.1891	2.9305	-0.0214
$\hat{s}_{H52}$	0.1167	0.2524	2.8088	-0.0038
$\hat{s}_{H53}$	-0.0588	0.1329	2.9929	-0.0189
$\hat{s}_{H54}$	-0.4805	-0.3763	3.1853	-0.0148
$\hat{s}_{H55}$	0.3751	0.2080	4.8403	0.1051
$\hat{s}_{H56}$	-0.1270	-0.2285	3.6877	0.0472
$\hat{s}_{H57}$	-0.0279	-0.1523	3.9793	0.0482
$\hat{s}_{H58}$	0.0329	-0.0626	4.0279	0.0300
$\hat{s}_{H59}$	0.0022	-0.2027	3.9633	0.0772

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{s}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

(continuação)				
Estimativas	J	SD	MR	I
$\hat{s}_{H60}$	-0.2620	-0.4152	3.6557	0.0035
$\hat{s}_{H61}$	-0.7101	-0.5348	2.2361	-0.1059
$\hat{s}_{H62}$	-2.4424	-1.9108	-0.9793	-0.2597
$\hat{s}_{H63}$	-2.1985	-1.9662	-0.4419	-0.2326
$\hat{s}_{H64}$	-1.4486	-1.2298	0.2189	-0.1515
$\hat{s}_{H65}$	-1.0259	-1.1331	1.2697	-0.0559
$\hat{s}_{H66}$	-1.2594	-1.3475	1.6147	-0.0704
$\hat{s}_{H67}$	0.1467	-0.1014	4.0376	0.0886
$\hat{s}_{H68}$	-0.1775	-0.2171	3.1870	0.0148
$\hat{s}_{H69}$	0.4606	-0.0141	4.7898	0.1369
$\hat{s}_{H70}$	-0.1887	-0.5295	3.1478	0.0611
$\hat{s}_{H71}$	-0.5460	-0.5433	2.7192	0.0065
$\hat{s}_{H72}$	-0.6404	-0.4956	1.8216	-0.0492
$\hat{s}_{H73}$	0.4565	0.0968	3.8352	0.0888
$\hat{s}_{H74}$	-0.4285	-0.6897	2.9795	0.0377
$\hat{s}_{H75}$	-0.0624	-0.2604	3.3112	0.0559
$\hat{s}_{H76}$	-0.3387	-0.6045	3.9146	0.0548
$\hat{s}_{H77}$	-0.1384	-0.6308	4.2128	0.0860
$\hat{s}_{H78}$	0.1642	-0.0178	3.6319	0.0600
$\hat{s}_{H79}$	0.2743	0.1383	3.9147	0.0960
$\hat{s}_{H80}$	-0.6153	-0.5023	2.7334	-0.0450
$\hat{s}_{H81}$	-0.4539	-0.5664	2.2337	-0.0333
$\hat{s}_{H82}$	-0.5909	-0.3999	1.2997	-0.0789
$\hat{s}_{H83}$	-0.6287	-0.5925	1.7684	-0.0485
$\hat{s}_{H84}$	-1.0583	-1.1378	1.6918	-0.0653
$\hat{s}_{H85}$	-1.6062	-1.5769	1.6762	-0.1153
$\hat{s}_{H86}$	-0.4306	-0.7096	1.8216	-0.0012
$\hat{s}_{H87}$	-0.7136	-1.0487	2.8075	0.0250
$\hat{s}_{H88}$	0.3093	-0.1439	4.5034	0.1198
$\hat{s}_{H89}$	-0.3577	-0.5993	2.1359	-0.0045
$\hat{s}_{H90}$	0.5582	0.3905	3.5967	0.0965
$\hat{s}_{H91}$	-0.2409	-0.3460	3.0521	0.0009
$\hat{s}_{H92}$	-0.1685	-0.4522	3.0676	0.0171
$\hat{s}_{H93}$	0.2084	-0.0517	3.8101	0.0786
$\hat{s}_{H94}$	-0.0339	-0.1195	3.3123	0.0119
$\hat{s}_{H95}$	0.3212	0.1545	3.8473	0.0484
$\hat{s}_{H96}$	-0.2799	-0.4124	3.3634	0.0350
$\hat{s}_{H97}$	-0.8549	-0.8475	2.0200	-0.0659

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{g}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

(continuação)				
Estimativas	J	SD	MR	I
$\hat{g}_{H98}$	-0.4575	-0.5793	2.4089	-0.0065
$\hat{g}_{H99}$	-0.3017	-0.2179	2.8434	-0.0243
$\hat{g}_{H100}$	-0.1670	-0.2541	3.5819	0.0312
$\hat{g}_{H101}$	0.2696	0.3706	2.7087	0.0292
$\hat{g}_{H102}$	0.0753	0.0667	1.7551	-0.0347
$\hat{g}_{H103}$	-0.1836	-0.0719	2.0743	-0.0362
$\hat{g}_{H104}$	-0.0298	-0.3526	3.6555	0.1037
$\hat{g}_{H105}$	0.6937	0.4572	4.5153	0.1460
$\hat{g}_{H106}$	-0.2328	-0.1723	2.2946	-0.0166
$\hat{g}_{H107}$	0.0825	0.0769	2.6486	-0.0098
$\hat{g}_{H108}$	0.3390	-0.0932	4.1611	0.1155
$\hat{g}_{H109}$	0.0096	0.0628	2.4621	0.0027
$\hat{g}_{H110}$	-0.1471	-0.1749	2.7406	-0.0138
$\hat{g}_{H111}$	0.4207	0.1283	3.5669	0.0970
$\hat{g}_{H112}$	-2.5049	-1.9716	-1.8949	-0.2770
$\hat{g}_{H113}$	-2.0189	-1.7394	-0.9888	-0.2130
$\hat{g}_{H114}$	-2.6442	-2.1067	-2.3486	-0.3299
$\hat{g}_{H115}$	-2.7003	-2.2600	-2.4674	-0.3487
$\hat{g}_{H116}$	-2.0720	-1.7970	-0.5244	-0.2070
$\hat{g}_{H117}$	0.1282	-0.1273	3.2296	0.0819
$\hat{g}_{H118}$	0.7188	0.4424	3.8067	0.1478
$\hat{g}_{H119}$	0.6459	0.4624	3.4777	0.0807
$\hat{g}_{H120}$	-0.2591	-0.5919	2.3689	0.0574
$\hat{g}_{H121}$	0.1958	0.0469	2.5495	0.0620
$\hat{g}_{H122}$	-0.0769	-0.0245	1.6183	0.0033
$\hat{g}_{H123}$	0.7512	0.5646	3.4696	0.1258
$\hat{g}_{H124}$	-0.3106	-0.1194	1.5357	-0.0310
$\hat{g}_{H125}$	0.1836	0.0991	2.4924	0.0483
$\hat{g}_{H126}$	-0.1708	-0.2889	2.9081	0.0287
$\hat{g}_{H127}$	-0.0344	-0.4382	3.2034	0.0599
$\hat{g}_{H128}$	0.4963	0.3236	3.2986	0.1000
$\hat{g}_{H129}$	0.4209	0.2368	3.3729	0.1156
$\hat{g}_{H130}$	0.1697	0.0254	3.1460	0.0677
$\hat{g}_{H131}$	-0.8841	-0.8294	0.7301	-0.1078
$\hat{g}_{H132}$	-1.1325	-0.8880	-0.1059	-0.1437
$\hat{g}_{H133}$	-0.7495	-0.5882	0.6288	-0.0871
$\hat{g}_{H134}$	-1.5591	-1.3101	-0.2100	-0.1787
$\hat{g}_{H135}$	-1.4526	-1.5107	1.2843	-0.0812

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{g}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

(continuação)				
Estimativas	J	SD	MR	I
$\hat{H}_{136}$	-0.8018	-0.9701	0.6330	-0.0447
$\hat{H}_{137}$	-0.7510	-0.9345	2.0270	0.0213
$\hat{H}_{138}$	-0.6542	-0.5931	1.7778	-0.0741
$\hat{H}_{139}$	-0.5689	-0.4653	0.4913	-0.0926
$\hat{H}_{140}$	0.5814	0.5078	2.5204	0.0636
$\hat{H}_{141}$	-0.1587	-0.2405	2.4596	0.0156
$\hat{H}_{142}$	-0.2296	-0.4536	2.2354	0.0085
$\hat{H}_{143}$	0.0799	-0.0211	2.4753	0.0206
$\hat{H}_{144}$	0.2497	0.0750	2.9669	0.0509
$\hat{H}_{145}$	0.2572	0.1206	3.0697	0.0447
$\hat{H}_{146}$	0.2312	-0.1448	3.3880	0.1099
$\hat{H}_{147}$	0.1529	-0.2300	3.0219	0.1050
$\hat{H}_{148}$	0.1138	-0.0918	2.4835	0.0732
$\hat{H}_{149}$	0.3394	0.2030	2.9393	0.0573
$\hat{H}_{150}$	0.3882	0.0892	3.8338	0.1284
$\hat{H}_{151}$	-0.7266	-0.9788	2.9320	0.0408
$\hat{H}_{152}$	-0.5484	-0.5798	2.4376	0.0225
$\hat{H}_{153}$	-0.7344	-0.9019	2.5685	0.0016
$\hat{H}_{154}$	-0.7736	-0.8473	2.5633	-0.0137
$\hat{H}_{155}$	-0.5044	-0.4705	2.5178	-0.0606
$\hat{H}_{156}$	-0.6602	-0.8395	2.9181	0.0338
$\hat{H}_{157}$	-0.7832	-0.8984	2.6550	-0.0195
$\hat{H}_{158}$	-0.6706	-0.7316	2.7652	-0.0320
$\hat{H}_{159}$	-0.6776	-0.6484	2.1432	-0.0390
$\hat{H}_{160}$	-0.6441	-0.9095	3.2402	0.0250
$\hat{H}_{161}$	-0.5059	-0.6799	2.8831	0.0194
$\hat{H}_{162}$	-0.7006	-0.8228	1.6282	0.0558
$\hat{H}_{163}$	0.7264	0.3058	3.2241	0.1866
$\hat{H}_{164}$	0.7466	0.4997	2.5119	0.1329
$\hat{H}_{165}$	0.6375	0.4189	2.3015	0.1053
$\hat{H}_{166}$	0.1362	-0.2027	3.1550	0.1412
$\hat{H}_{167}$	-0.9801	-0.9677	1.9551	-0.0529
$\hat{H}_{168}$	-1.2417	-1.4101	1.9932	-0.0394
$\hat{H}_{169}$	-1.0044	-1.0525	2.2308	-0.0512
$\hat{H}_{170}$	-1.0395	-0.8616	1.0692	-0.0804
$\hat{H}_{171}$	-0.9520	-1.2170	2.0591	0.0037
$\hat{H}_{172}$	-1.1082	-1.3318	1.5074	-0.0180
$\hat{H}_{173}$	-1.0648	-0.9833	1.6170	-0.0653

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{g}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

(continuação)				
Estimativas	J	SD	MR	I
$\hat{H}_{174}$	-0.5323	-0.8421	2.6257	0.0650
$\hat{H}_{175}$	-0.9147	-1.0601	1.9923	-0.0109
$\hat{H}_{176}$	-1.2516	-0.8291	1.1419	-0.1546
$\hat{H}_{177}$	-0.3117	-0.1566	2.3834	-0.0313
$\hat{H}_{178}$	-1.3777	-1.2680	1.0036	-0.1348
$\hat{H}_{179}$	-2.3476	-2.1953	0.2214	-0.2026
$\hat{H}_{180}$	-0.8111	-1.2205	2.8845	0.0318
$\hat{H}_{181}$	-0.3793	-0.7438	3.0016	0.1037
$\hat{H}_{182}$	-0.1033	-0.6131	2.8046	0.1299
$\hat{H}_{183}$	-0.1078	-0.2912	2.1839	0.0536
$\hat{H}_{184}$	-0.1658	-0.4116	2.4677	0.0716
$\hat{H}_{185}$	-0.1770	-0.4638	3.4063	0.1147
$\hat{H}_{186}$	-0.1568	-0.5127	2.2172	0.0989
$\hat{H}_{187}$	-1.0039	-1.1458	1.5794	-0.0340
$\hat{H}_{188}$	-0.2672	-0.4046	3.0418	0.0385
$\hat{H}_{189}$	-0.1733	-0.5942	2.5280	0.0956
$\hat{H}_{190}$	-1.7896	-1.8329	1.0616	-0.0887
$\hat{H}_{191}$	-0.7899	-0.6071	1.9249	-0.0474
$\hat{H}_{192}$	-0.4035	-0.3746	1.6957	-0.0555
$\hat{H}_{193}$	-0.9409	-0.7219	1.6094	-0.0745
$\hat{H}_{194}$	-0.7921	-0.6395	1.8943	-0.0646
$\hat{H}_{195}$	-0.4881	-0.6768	2.8470	0.0127
$\hat{H}_{196}$	-1.5862	-1.4413	1.2514	-0.1094
$\hat{H}_{197}$	-1.0258	-0.7680	0.8638	-0.1172
$\hat{H}_{198}$	-0.8228	-0.9684	1.6377	-0.0199
$\hat{H}_{199}$	-0.7632	-1.0142	2.1142	-0.0339
$\hat{H}_{200}$	-1.6298	-1.7712	1.8866	-0.0724
$\hat{H}_{201}$	-0.4674	-0.5841	3.8560	0.0265
$\hat{H}_{202}$	-0.4309	-0.8913	3.9174	0.0625
$\hat{H}_{203}$	-0.0844	-0.4097	4.4139	0.0775
$\hat{H}_{204}$	-0.3679	-0.5865	4.3890	0.0602
$\hat{H}_{205}$	-0.1885	-0.5882	4.6677	0.0453
$\hat{H}_{206}$	-0.1434	-0.2928	4.1006	0.0447
$\hat{H}_{207}$	-0.0747	-0.2505	4.7366	0.0796
$\hat{H}_{208}$	0.0143	-0.0931	4.6283	0.0458
$\hat{H}_{209}$	-0.1135	-0.4465	4.4332	0.0805
$\hat{H}_{210}$	-0.0877	-0.4790	5.1731	0.1096
$\hat{H}_{211}$	0.0988	-0.0297	4.4106	0.0642

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{g}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

(continuação)				
Estimativas	J	SD	MR	I
$\hat{H}_{212}$	-0.6948	-0.8135	2.7513	0.0618
$\hat{H}_{213}$	-0.7330	-0.6430	2.0667	-0.0305
$\hat{H}_{214}$	-0.4652	-0.5506	2.2834	0.0069
$\hat{H}_{215}$	-0.5792	-0.4580	2.0729	-0.0206
$\hat{H}_{216}$	-1.2002	-1.2222	2.2252	-0.0536
$\hat{H}_{217}$	-1.0123	-1.2531	3.6201	0.0055
$\hat{H}_{218}$	-0.9170	-1.1012	3.3614	-0.0101
$\hat{H}_{219}$	-0.1247	-0.5552	5.5190	0.1663
$\hat{H}_{220}$	-1.2558	-1.3193	2.5930	-0.0356
$\hat{H}_{221}$	-0.8224	-0.8149	2.4487	-0.0630
$\hat{H}_{222}$	-0.8105	-1.1072	3.1815	0.0413
$\hat{H}_{223}$	-0.3892	-0.4334	3.0203	-0.0331
$\hat{H}_{224}$	-1.1895	-1.0693	2.3043	-0.0706
$\hat{H}_{225}$	-1.1325	-1.4822	3.3089	0.0135
$\hat{H}_{226}$	-0.3903	-0.8778	5.2333	0.1415
$\hat{H}_{227}$	-0.9150	-1.2271	3.8456	0.0077
$\hat{H}_{228}$	-0.4547	-0.6111	3.6815	0.0225
$\hat{H}_{229}$	-0.7741	-1.1401	4.1091	0.0731
$\hat{H}_{230}$	-0.6219	-0.8715	4.1220	0.0484
$\hat{H}_{231}$	-1.4664	-1.4631	1.7678	-0.1212
$\hat{H}_{232}$	-1.0703	-1.0009	1.7088	-0.0814
$\hat{H}_{233}$	-1.1820	-1.0658	1.7813	-0.0898
$\hat{H}_{234}$	-0.8240	-0.8750	2.8040	0.0010
$\hat{H}_{235}$	-1.2695	-1.1986	2.5728	-0.0704
$\hat{H}_{236}$	-2.1792	-1.3640	-0.8277	-0.3025
$\hat{H}_{237}$	-1.8029	-1.6827	1.6946	-0.1260
$\hat{H}_{238}$	-0.8533	-0.7969	3.3231	-0.0380
$\hat{H}_{239}$	-1.5106	-1.3522	1.4106	-0.1177
$\hat{H}_{240}$	-1.0158	-1.0206	2.3334	-0.0691
$\hat{H}_{241}$	-0.4857	-0.4743	3.7841	0.0304
$\hat{H}_{242}$	-0.9268	-1.1962	3.4219	0.0097
$\hat{H}_{243}$	-0.8705	-0.9703	3.1266	-0.0307
$\hat{H}_{244}$	-0.5141	-0.6957	3.7293	0.0103
$\hat{H}_{245}$	-0.2380	-0.6676	4.4215	0.0624
$\hat{H}_{246}$	-1.0895	-1.3398	3.6201	0.0179
$\hat{H}_{247}$	-1.0563	-1.1209	2.8908	-0.0229
$\hat{H}_{248}$	-0.5837	-0.8489	3.2405	0.0326
$\hat{H}_{249}$	-0.3422	-0.5438	4.2800	0.0740

ANEXO H – Estimativas da média geral ($\hat{\beta}$), capacidades geral de combinação dos testadores ($\hat{g}_T$), das linhagens ($\hat{g}_L$) e capacidade específica de combinação dos híbridos ($\hat{g}_H$) obtidos com os diversos coeficientes de similaridade no estado e com a matriz identidade

				(conclusão)
Estimativas	J	SD	MR	I
$\hat{g}_{H_{250}}$	-0.3692	-0.6540	4.5739	0.0859

* J: Jaccard; SD: Sorensen e Dice; MR: Rogers Modificado; I: matriz identidade