

ANA MARIA CRUZ E OLIVEIRA

**ASSOCIAÇÃO DAS ABORDAGENS DE GENÉTICA QUANTITATIVA E
DE GENÉTICA DE POPULAÇÕES NO MELHORAMENTO DE
PLANTAS**

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

VIÇOSA
MINAS GERAIS – BRASIL
2015

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

O48a
2015
Oliveira, Ana Maria Cruz e, 1982-
Associação das abordagens de genética quantitativa e de
genética de populações no melhoramento de plantas / Ana Maria
Cruz e Oliveira. – Viçosa, MG, 2015.
xi, 67f. : il. ; 29 cm.

Orientador: Pedro Crescêncio de Souza Carneiro.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Plantas - Melhoramento genético. 2. Genética de
populações. 3. Genética quantitativa. 4. Diversidade genética.
5. Simulação (Computadores). I. Universidade Federal de
Viçosa. Departamento de Biologia Geral. Programa de
Pós-graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 631.52

ANA MARIA CRUZ E OLIVEIRA

**ASSOCIAÇÃO DAS ABORDAGENS DE GENÉTICA QUANTITATIVA E
DE GENÉTICA DE POPULAÇÕES NO MELHORAMENTO DE
PLANTAS**

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

APROVADA: 20 de novembro de 2015.

Felipe Lopes da Silva

Marcos Ribeiro Furtado

Renato Domiciano Silva Rosado

Cosme Damião Cruz
(Coorientador)

Pedro Crescêncio Souza Carneiro
(Orientador)

À Deus,
que sempre ouve e atende as minhas preces,
acalma o meu coração, enche a minha vida de
bênçãos e faz de mim uma pessoa melhor.

OFEREÇO

“Se tiveres fé, cumpre saber que tudo é possível àquele que a tem”.

Ao meu filho Mateus,
que é o amor da minha vida
e o motivo que me trouxe até aqui.

E à minha mãe Imaculada,
que me ajudou a cuidar do Mateus
e tornou possível eu chegar até aqui.

DEDICO

AGRADECIMENTO ESPECIAL

Ao Professor Pedro Crescêncio Souza Carneiro,

...por toda sabedoria que dividiu comigo nesses 11 anos de convivência. Por confiar em mim e acreditar no meu potencial. Pela disponibilidade em me atender, tirar minhas dúvidas e me ensinar, com muita paciência, toda vez que precisei de sua ajuda, toda vez que o acompanhei e auxiliei em seu trabalho. Agradeço pela sua amizade, que levarei sempre comigo, lembrando dos seus conselhos e ensinamentos de vida, que me fizeram crescer, não só como profissional, mas também como pessoa!

E ao Professor Cosme Damião Cruz,

...por despertar em mim o gosto pela ciência. Por sempre acompanhar a minha jornada acadêmica, confiar em mim e me possibilitar grandes oportunidades de crescimento. Por ser essa pessoa excepcional, exemplo de dedicação, que nos mostra que nossas conquistas são, e sempre serão, fruto de nosso trabalho. Agradeço por sua simplicidade e disposição ao me ensinar e ajudar, me dando sempre a chance de ver o quanto sou capaz de vencer!

AGRADECIMENTOS

À Deus, por estar sempre comigo, abençoando e iluminando o caminho a minha frente.

Ao meu querido filho Mateus, por esse amor que me fortalece, alegra e ilumina a minha vida.

Aos meus pais Sizernando e Imaculada, pelo amor incondicional e confiança, por dedicarem a mim tantos esforços, sem nunca pedir nada em troca e por fazerem tudo sempre pensando no melhor para a minha vida.

Aos meus irmãos André e Augusto, por todo apoio nessa caminhada, pelas risadas nos momentos em que estamos juntos festejando todo amor, confiança e companheirismo existentes entre nós.

Aos meus sobrinhos Felipe e Helena, que sempre trazem paz, amor e alegria ao meu dia.

À Gal, ajudante da minha mãe, hoje membro da nossa família, por cuidar de mim desde os meus 1 ano e 2 meses de idade, sempre acreditando, rezando e torcendo por mim.

À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento, pela formação acadêmica e oportunidade de crescimento profissional.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos.

Ao Professor Pedro Crescêncio Souza Carneiro, por sua orientação e dedicação despendida a mim no desenvolvimento desse trabalho.

Ao Professor Cosme Damião Cruz, pela sua coorientação, ensinamentos transmitidos e ajuda fundamental na análise dos dados e desenvolvimento da tese.

Ao professor Leonardo Lopes Bhering, pela sua coorientação, pelos ensinamentos, pela amizade e confiança. Por me permitir auxiliá-lo em suas atividades e abrir importantes caminhos em minha jornada profissional.

Aos membros da banca, Professor Felipe Lopes da Silva, Professor Marcos Ribeiro Furtado e Dr. Renato Domiciano Silva Rosado, pela disponibilidade e valiosas sugestões para o enriquecimento do meu trabalho.

A todos os professores do Programa de Pós-Graduação em Genética e Melhoramento pelos ensinamentos e contribuição em minha formação.

Aos secretários do Programa de Pós-Graduação em Genética e Melhoramento, Marco Túlio e Odilon, por estarem sempre dispostos a nos atender e ajudar.

Aos amigos dos Laboratórios de Biometria e Bioinformática, pelo apoio nesse longo caminho que compartilhamos.

Aos amigos da academia que me ajudaram a espantar o stress, especialmente Ana Paula, Marcinho, PC, Rodrigo, Lucas e Letícia, que sempre torceram por mim.

À querida coach e amiga Ethel, que me ajudou a transformar a minha vida e foi peça fundamental nessa conquista.

Ao meu irmão de coração, Galvão, e às minhas irmãs branqueas, Fefe e Mari, por estarem sempre ao meu lado e me mostrarem o valor de uma verdadeira amizade.

Aos grandes amigos, Cássia, Bruno Gomes, Renatinha, Gis, Marciane, Caio, Camila, Ísis, Gustavo, Lúcia, Flávia, Tchola, Thaís, Bruno Brasileiro, Rusthon, Fabiano, Bruno Soares, Rajane, Dê e Cari, por compartilharem comigo momentos que encheram a minha vida de alegria, por me ajudarem a levantar a cabeça, sorrir e seguir em frente... sempre!

A todos que, direta ou indiretamente, contribuíram para a concretização deste trabalho...

MUITO OBRIGADA!

BIOGRAFIA

ANA MARIA CRUZ E OLIVEIRA, mãe de Mateus Oliveira Bomfim e filha de Sizernando Luiz de Oliveira e Maria Imaculada da Conceição Cruz Barbosa, nasceu em Santo Antônio de Jesus, Bahia, em 28 de agosto de 1982.

Em março de 2005 ingressou na Universidade Federal de Viçosa, graduando-se em Agronomia em julho de 2009.

Em março de 2010, iniciou o curso de Mestrado em Genética e Melhoramento da Universidade Federal de Viçosa, recebendo o título de Mestre em fevereiro de 2012.

Em março de 2012, iniciou o curso de Doutorado em Genética e Melhoramento da Universidade Federal de Viçosa, submetendo-se à defesa de tese em novembro de 2015.

SUMÁRIO

	Páginas
RESUMO	x
ABSTRACT	xi
1. INTRODUÇÃO GERAL	1
2. REFERENCIAL TEÓRICO	4
2.1. Divergência genética	4
2.2. Medidas descritivas da estrutura genética da população	5
2.3. Medidas descritivas do grau de diferenciação entre populações	9
2.4. Cruzamentos dialélicos	11
2.5. Simulação de dados	12
3. REFERÊNCIAS BIBLIOGRÁFICAS	14
CAPÍTULO 1: Predição da diversidade genética de combinações híbridas a partir da abordagem de genética de populações em suas populações genitoras	19
1. Resumo	20
2. Abstract	21
3. Introdução	21
4. Material e Métodos	23
4.1. Simulação dos dados	23
4.2. Avaliação do potencial das populações	24
4.3. Grau de diferenciação entre as populações	24
4.4. Correlação entre matrizes	27
5. Resultados e Discussão	28
6. Referências Bibliográficas	35
CAPÍTULO 2: Abordagem de genética de populações na escolha de progenitores potenciais ao melhoramento genético	39
1. Resumo	40
2. Abstract	41
3. Introdução	42
4. Material e Métodos	44
4.1. Simulação dos dados genotípicos.....	44
4.2. Simulação dos dados fenotípicos	44

4.3. Análise dialéctica dos dados fenotípicos	46
4.4. Análise de dissimilaridade entre genitores	46
5. Resultados e Discussão	49
5.1. Análise com base nos dados fenotípicos	49
5.2. Abordagem de genética de populações na escolha de progenitores	58
6. Referências Bibliográficas	64
4. CONCLUSÕES GERAIS	67

RESUMO

OLIVEIRA, Ana Maria Cruz e, D.Sc., Universidade Federal de Viçosa, novembro de 2015. **Associação das abordagens de genética quantitativa e de genética de populações no melhoramento de plantas.** Orientador: Pedro Crescêncio Souza Carneiro. Coorientadores: Cosme Damião Cruz e Leonardo Lopes Bhering.

O presente trabalho teve como objetivo estudar a associação entre os parâmetros que descrevem o potencial de populações e de suas combinações híbridas por meio da abordagem de genética de populações, baseada em dados de marcadores moleculares, com os obtidos por meio da abordagem de genética quantitativa, baseados em dados fenotípicos. Para isso, foram simuladas dez populações genitoras, com 200 indivíduos cada, em equilíbrio de Hardy-Weinberg, para 50 locos independentes e dois alelos codominantes por loco, e obtidas 45 populações híbridas a partir da combinação, aos pares, das dez populações genitoras. Os dados fenotípicos foram obtidos a partir da simulação de 12 variáveis quantitativas pela ação de alelos de 20 locos, selecionados ao acaso entre os 50 previamente simulados, com diferentes graus de herdabilidade, combinados com um efeito aditivo diferencial, de acordo com a importância do loco. As populações foram avaliadas quanto às medidas descritivas da estrutura genética da população e do grau de diferenciação entre pares de populações genitoras, a partir da abordagem de genética de populações, e quanto à complementação gênica, por abordagem de genética quantitativa, considerando aspectos da capacidade geral e específica de combinação. Houve, de modo geral, concordância de indicação de combinações híbridas mais e menos divergentes pelas medidas de dissimilaridade abordadas pela teoria de genética de populações e pode-se afirmar que é possível prever a diversidade genética de combinações híbridas a partir da avaliação da diversidade genética de suas populações genitoras. Observou-se que, em características de efeito aditivo, a escolha de genitores com maior frequência de alelos favoráveis pode ser feita com base em suas médias fenotípicas apenas, sem a necessidade de realização de cruzamentos dialélicos. Além disso, pela abordagem de genética de populações, houve coincidência de 75% na recomendação das populações genitoras potenciais ao melhoramento genético. Nesse sentido, pode-se concluir que a associação da abordagem de genética de populações às informações quantitativas pode incrementar o ganho no melhoramento genético.

ABSTRACT

OLIVEIRA, Ana Maria Cruz e, D.Sc., Universidade Federal de Viçosa, november, 2015. **Association of genetic quantitative and population genetics approaches in plant breeding.** Adviser: Pedro Crescêncio Souza Carneiro. Co-advisers: Cosme Damião Cruz and Leonardo Lopes Bhering.

This work aimed to study the association between the parameters describing the potential of people and their hybrid combinations by means of population genetics approach, based on molecular marker data, with those obtained through quantitative genetic approach, based on phenotypic data. For this, ten progenitors populations were simulated, with 200 individuals each, in Hardy-Weinberg to 50 independent loci and two codominant alleles per locus, and obtained 45 hybrid populations from the combination, in pairs, of ten progenitors populations. Phenotypic data were obtained from the simulation of 12 quantitative variables by the action of alleles from 20 loci, selected randomly among the 50 previously simulated, with different degrees of heritability, combined with a differential additive effect, according to the importance of the locus. The populations were evaluated for descriptive measures of the genetic structure of the population and the degree of differentiation between pairs of progenitors populations from the population genetics approach, and on the genetic complementation, by approach of quantitative genetics, considering aspects of general and specific combining ability. There was, in general, agree indication of more and less divergent hybrids by dissimilarity measures addressed by the theory of population genetics and it can be affirmed that it is possible to predict the genetic diversity of hybrids from the evaluation of their progenitor populations genetic diversity. It was observed that for the additive effect characteristics, the choice of progenitors with highest frequency of favorable alleles can be made based on their phenotypic medium only, without the need of performing diallel. In addition, by the population genetic approach, there was agreement on the recommendation of 75% of the population potential progenitors to genetic improvement. In this sense, it can be concluded that the association of population genetics approach to quantitative information can increase the gain in plant breeding.

1. INTRODUÇÃO GERAL

Em um programa de melhoramento, uma das etapas de maior importância é a escolha de genitores para a formação de potenciais populações segregantes, pois dela depende o sucesso das etapas subsequentes e a eficiência do programa (Bertan et al., 2007; Pereira et al., 2007). Neste contexto, a escolha das populações segregantes de maior potencial otimiza a utilização de recursos humanos e financeiros despendidos ao programa de melhoramento (Pimentel et al., 2013). Para uma escolha eficiente das populações mais promissoras têm sido utilizadas as estimativas de capacidade geral e específica de combinação, obtidas a partir de cruzamentos dialélicos. Neste tipo de análise se avalia o potencial e a diversidade genética de linhagens ou populações, a partir de informações de suas combinações híbridas, o que permite identificar as combinações parentais que gerem populações segregantes com máxima variabilidade genética para a seleção, prevendo, segundo os modelos biométricos abordados pela genética quantitativa, a heterose e as constituições híbridas mais promissoras.

Ênfase também tem sido dada ao estudo de diversidade genética em diversas culturas a fim de, por exemplo, identificar genótipos promissores para o melhoramento genético (Coimbra et al., 2010; Pimentel et al., 2013; Santos et al., 2014), quantificar a variabilidade genética (Rigon et al., 2012) e direcionar o melhorista na escolha mais apropriada para a formação de híbridos superiores (Barili et al., 2011; Ferreira et al., 2012). Para tanto, diferentes conjuntos de dados podem ser utilizados, como dados de

pedigree, dados fenotípicos, bioquímicos, fitopatológicos, moleculares, entre outros. A diversidade genética pode ser analisada em diferentes níveis, incluindo genótipos individuais, acessos de germoplasma e populações, variando especialmente de acordo com a natureza do material genético avaliado em cada caso. A análise da diversidade em nível de população é considerada a mais complexa, por sofrer influência do número de indivíduos amostrados, número de locos, constituição genotípica e tamanho efetivo (Cruz et al., 2011).

Atualmente, os dados de marcadores moleculares tornaram-se ferramentas úteis que podem aumentar a acurácia da seleção, reduzindo cargas de trabalho em campo e aumentando o sucesso em programas de melhoramento (Zhang et al., 2013). Muitos são os trabalhos que utilizam informações de marcadores moleculares, fundamentados nas teorias de genética de populações, para estimar a diversidade genética entre e dentro de populações. De forma geral, estes trabalhos obtêm estimativas de parâmetros como o grau de endogamia, sistema reprodutivo predominante e heterozigosidade a fim de estabelecer a estratégia mais adequada na quantificação da variabilidade e conservação das espécies (Viegas et al., 2011; Pimentel et al., 2013). Porém, poucos são os estudos que associam a diversidade predita a partir das populações com o desempenho observado em seus híbridos por estas duas correntes de fundamentos e informações, ou seja, genética de populações e genética quantitativa. Pode-se afirmar que uma das grandes dificuldades em realizar tal associação é a falta de informação conjunta de dados moleculares e fenotípicos de qualidade, além da falta de conhecimento a respeito das associações gênicas, interações e efeitos ambientais. Nesse contexto, destaca-se a simulação de dados que fornece resultados para análises mais elaboradas a respeito do sistema estudado, com a possibilidade de interpretação mais profunda e abrangente, sem perda de confiabilidade (Cruz, 2006).

Assim, o objetivo com o presente trabalho foi estudar a associação entre os parâmetros que descrevem o potencial de populações e de suas combinações híbridas por meio da abordagem de genética de populações, baseada em dados de marcadores moleculares, com os obtidos por meio da abordagem de genética quantitativa, baseados em dados fenotípicos.

2. REFERENCIAL TEÓRICO

2.1. Divergência Genética

Vários autores já relataram a importância da avaliação da diversidade genética para a escolha de materiais promissores para a formação de populações segregantes de maior potencial (Barili et al., 2011; Ferreira et al., 2012; Silva et al., 2013). Nesse contexto, em se tratando de um programa de melhoramento vegetal, essa decisão é extremamente importante, uma vez que está diretamente relacionada à otimização da utilização de recursos humanos e financeiros despendidos no programa (Pimentel et al., 2013).

Ênfase tem sido dada ao estudo de diversidade genética em diversas culturas a fim de identificar genótipos promissores para o melhoramento genético (Coimbra et al., 2010; Santos et al., 2014), quantificar a variabilidade genética (Almeida et al., 2011; Rigon et al., 2012), direcionar o melhorista na escolha mais apropriada de genitores para a formação de híbridos superiores (Barili et al., 2011; Ferreira et al., 2012), entre outros. Para tanto, diferentes técnicas e conjuntos de dados podem ser utilizados. Coimbra et al. (2010), por exemplo, utilizaram descritores quantitativos e qualitativos para avaliar a divergência genética entre populações de milho e identificar genótipos promissores ao melhoramento genético, enquanto Signorini et al. (2013) utilizaram dados bioquímicos para análise da variabilidade genética entre espécies do gênero *Capsicum* e Moulin et al.

(2013) utilizaram dados oriundos de marcadores moleculares com o intuito de caracterizar acessos de pimenta quanto a variabilidade genética.

As medidas de dissimilaridade comumente utilizadas em variáveis quantitativas são a distância Euclidiana média e a distância generalizada de Mahalanobis, que é a mais apropriada quando se dispõe da matriz de covariâncias residuais estimada a partir de ensaios experimentais com repetição (Cruz et al. 2012).

Com o advento dos marcadores de DNA, os dados de marcadores moleculares tornaram-se informações úteis que podem aumentar a acurácia da seleção, reduzindo cargas de trabalho em campo e aumentando o sucesso em programas de melhoramento (Zhang et al., 2013). Muitos são os trabalhos que utilizam informações de marcadores moleculares, fundamentados nas teorias de genética de populações, para estimar a diversidade genética entre e dentro de populações (Viegas et al., 2011; Pimentel et al., 2013). Vale ressaltar que a análise da diversidade em nível de população é considerada a mais complexa, por sofrer influência do número de indivíduos amostrados, número de locos, constituição genotípica e tamanho efetivo (Cruz et al., 2011).

2.2. Medidas descritivas da estrutura genética da população

A estrutura de uma população é definida pela frequência dos alelos que compõem os diferentes genótipos dos indivíduos que a constituem e seu entendimento pode direcionar ações em programas de melhoramento (Malhado et al., 2008; Cruz et al., 2011). No estabelecimento da estrutura populacional, identificar a ocorrência de equilíbrio de Hardy-Weinberg (EHW) e estimar parâmetros como conteúdo de informação polimórfica (PIC), coeficiente de endogamia (F) e heterozigosidade (H) são primordiais (Santos et al., 2012).

i) Equilíbrio de Hardy-Weinberg

No início do século 20, Hardy (1908) e Weinberg (1908) afirmaram que, na ausência de migração, mutação, seleção ou deriva genética e sob condições de acasalamento ao acaso, a frequência genotípica de uma população é determinada pelo quadrado de suas frequências alélicas. Tal fato é hoje conhecido por Equilíbrio de Hardy-Weinberg (EHW) e a possibilidade de predição das frequências alélicas e genotípicas de gerações futuras permite aos pesquisadores estabelecer estratégias de melhoramento e manipulação da população (Cruz et al., 2011).

Desvios no equilíbrio de Hardy-Weinberg podem sugerir problemas com a estrutura populacional (Wigginton et al., 2005; Ziegler et al., 2011) e atenção deve ser dada à ocorrência de fluxo gênico, mutações, deriva genética, acasalamentos preferenciais, dentre outros, que são fatores perturbadores do equilíbrio, porém, podem não ser estatisticamente detectáveis. A ocorrência de deriva genética, por exemplo, é especificamente perceptível apenas em pequenas populações, desta forma, é importante considerar informações de vários locos, a fim de se obter maior credibilidade sobre as reais condições da população em estudo (Cruz et al., 2011). Neste sentido destacam-se os dados de marcadores moleculares.

Há grande importância em avaliar se uma determinada população, ou melhor, se cada loco independentemente, encontra-se, ou não, em equilíbrio de Hardy-Weinberg. Havendo informações sobre a frequência dos genótipos, pode-se verificar as condições de equilíbrio por meio do teste de qui-quadrado, teste da razão de verossimilhança e, ainda, teste exato de Fisher (Cruz et al., 2011). A hipótese de nulidade a ser testada é a de que os genótipos observados são produtos da união aleatória de gametas masculinos e femininos.

O teste de qui-quadrado compara a quantidade de indivíduos observados na população com a quantidade de indivíduos esperada sob condição de EHW. É uma estimativa altamente influenciada pelo tamanho amostral, portanto, amostras pequenas podem não ser suficientes para quantificar a verdadeira ocorrência de certas classes genotípicas. O teste da razão de verossimilhança é indicado para situações em que se tem mais de dois alelos por loco e se deseja trabalhar com as frequências das classes genotípicas. Este é baseado na razão entre duas funções de máxima verossimilhança, uma definida pelas frequências genotípicas esperadas e outra pelas frequências observadas. O teste exato de Fisher, por sua vez, é indicado quando se tem apenas dois alelos por loco. Entretanto, quando da utilização na ocorrência de alelos múltiplos, recomenda-se transformar os dados. Este está baseado na probabilidade assumida por um possível arranjo genotípico, condicionado às frequências alélicas amostradas na população (Weir, 1996).

ii) Endogamia

Endogamia significa o acasalamento entre indivíduos aparentados e o grau de parentesco dos indivíduos numa população será medido por meio do coeficiente de endogamia, que é a probabilidade de dois genes em qualquer loco, num indivíduo, serem idênticos por ascendência. Os genes são considerados idênticos por ascendência quando derivam ou são cópias de um alelo comum encontrado nos ancestrais daquele indivíduo (Cruz et al., 2011).

Numa população de indivíduos bissexuados, todo indivíduo tem 2^t ancestrais (t = gerações passadas) e, segundo Falconer (1987), o número de indivíduos necessários para proporcionar ancestrais separados, para todos os indivíduos atuais, torna-se maior do que o que qualquer população real poderia conter. Portanto, qualquer par de indivíduo deve

estar relacionado com um ou mais ancestrais comuns e, quanto menor for o tamanho da população da geração anterior, maior será o seu número.

A principal consequência de dois indivíduos terem um ancestral comum é que os dois podem carregar réplicas de um dos genes presentes no ancestral e, no caso de se acasalarem, podem transmitir essas réplicas à sua progênie.

O grau de parentesco, expresso pelo coeficiente de endogamia (F), compara a frequência de heterozigotos da população em estudo com a frequência de heterozigotos esperada em uma população supostamente em equilíbrio de Hardy-Weinberg (EHW). De acordo com Cruz et al (2011), os valores obtidos geralmente variam de zero a um, entretanto, valores negativos também podem ocorrer e os mesmos devem ser interpretados como nulos, de forma a caracterizar a ausência de acasalamento entre aparentados na população.

Em programas de melhoramento de variedades, por exemplo, a depressão por endogamia é a principal consequência resultante da autofecundação artificial visando obter linhagens para formação de híbridos (Scapim et al., 2006). Além disso, características agronômicas importantes, como rendimento de grãos, podem ter seu desempenho altamente prejudicado pela diminuição do vigor das linhagens obtidas. Segundo Vencovsky & Barriga (1992), para aumentar a probabilidade de obtenção de linhagens vigorosas, a população deve apresentar alta frequência de alelos favoráveis em relação à característica de interesse.

iii) Heterozigosidade e Conteúdo de Informação Polimórfica

A heterozigosidade retrata a proporção de indivíduos heterozigotos na população e é considerada um indicativo de variabilidade genética (Ferreira et al., 2012), essencial para a manifestação da heterose na população (Cruz et al., 2008).

O conteúdo de informação polimórfica (PIC) apresenta papel relevante na seleção de genitores e marcadores informativos, permitindo avaliar o grau de heterozigosidade de um loco, calculado com base na frequência do loco em estudo. Seu valor fornece uma estimativa do poder discriminatório do marcador, por considerar não somente o número de alelos por loco, mas também a frequência relativa desses alelos. De maneira geral, os valores de PIC são inferiores aos obtidos para a heterozigosidade.

Segundo Ott (1992), um loco será considerado polimórfico quando sua heterozigosidade observada for maior do que 10% (situação em que o alelo mais frequente tem frequência menor que 95%) e altamente polimórfico quando sua heterozigosidade observada for maior do que 70%.

iv) Média e Variância

O cruzamento entre duas populações em equilíbrio de Hardy-Weinberg produzirá uma população híbrida de maior desempenho, pelo incremento da heterose manifestada e maior variabilidade, quanto mais divergentes forem as populações genitoras. Vários estudos de avaliação da diversidade genética entre populações têm sido realizados com o objetivo de identificar aquelas divergentes e de bom desempenho, que possibilite a formação de populações híbridas com variabilidade, melhor desempenho e presença de indivíduos transgressivos, com valores fenotípicos que superam os limites estabelecidos pelas populações genitoras (Cruz et al., 2011).

2.3. Medidas descritivas do grau de diferenciação entre populações

Vários são os métodos para estimação da distância entre e dentro de populações a partir de informações moleculares e a escolha do mais apropriado é um importante

componente da análise da diversidade genética. Os estimadores podem ser baseados em medidas de dissimilaridade ou em índices de fixação.

As medidas de dissimilaridade podem considerar tanto a frequência alélica das populações, como no caso da Distância Euclidiana e da Distância Angular, como a frequência genotípica, através da Distância Genotípica de Hedrick, ferramenta particularmente importante, uma vez que populações podem ser erroneamente classificadas como idênticas, quando a análise está baseada apenas em suas frequências alélicas. Segundo Cruz et al. (2011), a utilização de informações de frequências alélicas, por sua vez, é de fundamental importância em análises de diversidade genética de materiais elites, como linhagens endogâmicas, a fim de estabelecer grupos heteróticos para o desenvolvimento de híbridos promissores.

Os índices de fixação são usados para medir qualquer desvio da condição de equilíbrio de Hardy-Weinberg quando há dois alelos para um determinado loco (Wright, 1951). O índice de fixação segundo o modelo de Nei (1973) estima parâmetros de identidade, heterozigosidade e diversidade genética para estabelecimento do Índice de Diversidade, avaliando sempre o que ocorre dentro da população e entre pares de populações.

O índice de fixação de Wright (1951) compara as frequências dos genes entre as populações e descreve três coeficientes aplicáveis a uma população – F_{IT} , F_{ST} e F_{IS} , medindo, respectivamente, o desvio das frequências genotípicas da população em relação ao equilíbrio de Hardy-Weinberg (EHW), a probabilidade de dois indivíduos de populações diferentes possuírem um alelo idêntico por ascendência e a probabilidade de dois alelos em um loco do mesmo indivíduo serem idênticos por ascendência.

Vale ressaltar que na decomposição do índice de fixação de Wright (1951) são consideradas algumas pressuposições básicas:

- a) A população e as subpopulações sob investigação são derivadas de um ancestral comum ao mesmo tempo;
- b) Todas as subpopulações são igualmente relacionadas umas com as outras ou não há migração entre elas e
- c) As populações são consideradas amostras ao acaso de um conjunto infinito de subpopulações igualmente relacionadas.

2.4. Cruzamentos Dialélicos

Quando se deseja obter populações segregantes com alto potencial genético, o estabelecimento de quais genitores devem ser cruzados é de fundamental importância para o sucesso do programa de melhoramento. Escolher os genitores apenas com base no desempenho per se pode não ser suficiente para formação de populações segregantes promissoras. Aliado ao comportamento per se das variedades também deve ser considerada a divergência entre elas.

O termo dialélico tem sido utilizado para expressar um conjunto de híbridos resultante do acasalamento entre genitores, que podem incluir, além dos respectivos pais, os híbridos recíprocos e outras gerações relacionadas (Cruz et al., 2012). Por meio da análise dialélica, o melhorista acessa informações em relação ao comportamento dos genitores entre si e suas combinações híbridas, especificados pelas capacidades geral e específica de combinação, respectivamente.

Entre os métodos propostos de análise dialélica, o mais conhecido é o de Griffing (1956), que permite estimar as capacidades geral e específica de combinação, além do efeito recíproco, detectando variações genéticas a partir do comportamento do híbrido F1 e do seu recíproco.

A capacidade geral de combinação (CGC) representa o comportamento médio de um genitor em suas combinações híbridas, e seu efeito é função da ação de genes de efeito

predominantemente aditivos (Sprague e Tatum, 1942). Os efeitos aditivos são responsáveis pela proporção da variação fenotípica que pode ser transmitida para as sucessivas gerações, sendo, portanto, responsável pelo desempenho das linhas puras obtidas em cada população quando a homozigose é alcançada (Ramalho et al., 1988).

A capacidade específica de combinação (CEC) está relacionada com efeitos gênicos não aditivos (Sprague e Tatum, 1942) e indica o grau de complementaridade entre os genitores envolvidos nos cruzamentos (Cruz et al., 2012). O efeito da CEC é útil, pois sua magnitude indica a dimensão da variabilidade que pode ser expressa em cada população, esperando-se maior probabilidade de obtenção de segregantes transgressivos quando os efeitos da CEC forem elevados (Barelli et al., 2000).

2.5. Simulação de dados

O desenvolvimento da tecnologia permitiu o avanço da ciência baseada na área da bioinformática, responsável por estudar e relacionar dados biológicos, com o auxílio de métodos computacionais e simulação de dados (Prosdocimi, 2007). Segundo Banks (2000), a simulação de dados permite representar um sistema real a partir de um modelo computacional e realizar observações e inferências nas características avaliadas. Além disso, no processo de simulação de um fenômeno biológico são estabelecidos parâmetros e restrições que tornam o cenário emulado o mais fiel possível à realidade, de forma que possibilite a descrição do fenômeno conforme o sistema natural (Naylor, 1971; Dachs, 1988; Cruz, 2006).

O processo de simulação tem sido utilizado desde 1950, com diversas áreas de aplicação, incluindo os sistemas computacionais e de telecomunicações, logística, militar, treinamento e científica (Banks et al., 1998; Harrell et al., 2000; Law e Kelton, 2000). Atualmente, a simulação de dados vem sendo utilizada em vários ramos da genética como em análise discriminante em estruturas de populações (Jombart et al., 2010), estudos de

genoma (Price et al., 2006, Cruz et al., 2013) e análise da variância genética (Bhattacharya et al., 2011). Os estudos são baseados em modelos mais simples que os sistemas reais, entretanto, tal simplicidade não afeta seu desempenho em comparação ao modelo real.

A simulação de dados merece grande destaque, uma vez que nem sempre é possível realizar experimentos com todas as observações e repetições necessárias à obtenção de uma estimativa precisa. Desse modo, as técnicas de ampliação dos dados, por meio da simulação, têm poupado recursos demandados na instalação de um experimento em campo e permitido estudos mais diversificados e abrangentes.

Diversos softwares podem ser utilizados para realizar a simulação dos dados, existindo dois deles que são diretamente relacionados com estudos genéticos – os programas GENES e GQMOL, disponíveis para download gratuito, sem restrição de uso ou divulgação, no endereço <<http://www.ufv.br/dbg/biodata.htm>>. Os dois programas têm sido de grande utilidade na área de melhoramento genético, pois contam com recursos para análise e processamento de dados baseados em diferentes metodologias biométricas. Além disso, diferentes genomas, tipos de marcadores, tipos de populações segregantes, características quantitativas, entre outros, podem ser estabelecidos através do módulo de simulação, proporcionando valiosas contribuições na experimentação científica (Cruz et al., 2013).

3. REFERÊNCIAS BIBLIOGRÁFICAS

Almeida, R.D. de; Peluzio, J.M.; Afféri, F.S.. Divergência genética entre cultivares de soja, sob condições de várzea irrigada, no sul do Estado Tocantins. **Revista Ciência Agronômica**, v.42, n.1, p.108-115, 2011.

Banks, J. **Handbook of simulation: principles, methodology, advances, applications, and Practice**. New York: John Wiley & Sons, 1998.

Banks, J. **Introduction to simulation**. Proceedings of the Winter Simulation Conference. Atlanta, 2000.

Barelli, M.A.A.; Vidigal, M.C.G.; Amaral Júnior, A.T.; Vidigal Filho, P.S.; Scapim, C.A.; Sagrilo, E. Diallel analysis for grain yield and yield components in *Phaseolus vulgaris* L. **Acta Scientiarum Agronomy**, v.22, p.883-887, 2000.

Barili, L.D.; Vale, N.M. do; Arruda, B.; Toaldo, D.; Rocha, F. da; Coimbra, J.L.M.; Bertoldo, J.G.; Guidolin, A.F. Escolha de genitores contrastantes para compor o bloco de cruzamentos de genótipos de feijão. **Revista Brasileira de Agrociência**, v.17, n.3-4, p.303-310, 2011.

Bertan, I.; Carvalho, F.I.F.; Oliveira, A.C. Parental selection strategies in plant breeding programs. **Journal of Crop Science and Biotechnology**, v.10, p.211–222, 2007.

Bhattacharya, S.; Heitmann, K.; White, M.; Lukić, Z.; Wagner, C.; Habib, S. Mass Function Predictions Beyond LCDM. In: **Bulletin of the American Astronomical Society**, p.12704, 2011.

Coimbra, R.R.; Miranda, G.V.; Cruz, C.D.; Melo, A.V.; Eckert, F.R. Caracterização e divergência genética de populações de milho resgatadas do Sudeste de Minas Gerais. **Revista Ciência Agronômica**, v.41, n.1, p.159–166, 2010.

Cruz, C.D. **Princípios de genética quantitativa**. Viçosa, MG: Ed. UFV, 2005.

Cruz, C.D. **Programa Genes: Análise Multivariada e Simulação**. Viçosa, MG: Ed. UFV, 2006.

Cruz, C.D. **Programa Genes: Diversidade Genética**. Viçosa, MG: Ed. UFV, 2008.

Cruz, C.D.; Ferreira, F.M.; Pessoni, L.A. **Biometria aplicada ao estudo da diversidade genética**. 1ed, Ed. Suprema, 2011.

Cruz, C.D.; Regazzi, A.J.; Carneiro, P.C.S. **Modelos Biométricos Aplicados ao Melhoramento Genético**. V.1, 4ed, Viçosa, MG: Ed. UFV, 2012.

Cruz, C.D.; Salgado, C.C.; Bhering, L.L. **Genômica Aplicada**. Ed. Suprema, 2013.

Dachs, J.N.W. **Estatística computacional: uma introdução em turbo pascal**. Rio de Janeiro. LTC, 1988.

Falconer, D.S. **Introdução à genética quantitativa**. 279p. UFV, 1987.

Ferreira C.B.B.; Lopes, M.T.G.; Lopes, R.; Cunha, R.N.V. da; Moreira, D.A.M.; Barros, W.S.; Matiello, R.R. Diversidade genética molecular de progênies de dendezeiro. **Pesquisa Agropecuária Brasileira**, v.47, n.3, p.378-384, 2012.

Griffing, B. Concept of general and specific ability in relation to diallel crossing systems. **Australian Journal of Biological Sciences**, v.9, n.4, p.462-93, 1956.

Hardy, H.G. Mendelian proportions in a mixed population. **Science**, v.28, p.49–50, 1908.

Harrel, C.R.; Ghosh, B.K.; Bowden, R. **Simulation Using ProModel**. McGraw-Hill, 2000.

Jombart, T.; Devillard, S.; Balloux, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. **BMC genetics**, v.11, n.1, p.94-108, 2010.

Law, A.M., Kelton, W.D. **Simulation Modeling and Analysis**, 3rd ed. McGraw-Hill, New York, 2000.

Malhado, C.H.M.; Carneiro, P.L.S.; Pereira, D.G.; Martins Filho, R.; Souza, J.C. Progresso genético e estrutura populacional do rebanho Nelore no Estado da Bahia. **Pesquisa Agropecuária Brasileira**, v.43, n.9, p.1163-1169, 2008.

Moulin, M.M.; Oliveira, F.L.; Bianchi, P.A.; Pimenta, S.; Rodrigues, R. Uso de marcadores ISSR para estimar a diversidade genética entre acessos de pimenta coletados no sul do estado do Espírito Santo. **Perspectivas online: Ciências Biológicas e da Saúde**, v.10, n.3, p.35-43, 2013.

Nei, M. Analysis of gene diversity in subdivided populations. **Proceedings of the National Academy of Sciences of USA**, v.70, n.12, p.3321-3323, 1973.

Naylor, T.H., Balintfy, J.L., Burdick, D.S., Chu, K. **Técnicas de simulação em computadores**. São Paulo, Vozes, 1971.

Ott, J. Strategies for characterizing highly polymorphic markers in human gene mapping. **American journal of human genetics**, v.51, n.2, p.283, 1992.

Pimentel, A.J.B.; Ribeiro, G.; Souza, M.A.; Moura, L.M.; Assis, J.C.; Machado, J.C. Comparação de métodos de seleção de genitores e populações segregantes aplicados ao melhoramento de trigo. **Bragantia**, v.72, n.2, p.113-121, 2013.

Price, A.L.; Patterson, N.J.; Plenge, R.M.; Weinblatt, M.E.; Shadick, N.A.; Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. **Nat Genet**, v.38, n.8, p.904-909, 2006.

Prosdocimi, F. Curso de Bioinformática. Biotecnologia Ciência & Desenvolvimento, 77f., 2007.

Ramalho, M.A.P.; Santos, J.B.; Pereira Filho, I.A. Choice of parents for dry bean (*Phaseolus vulgaris* L.) breeding. I. Interactions of mean components by generation and by location. **Revista Brasileira de Genética**, v.11, p.391-400, 1988.

Rigon, J.P.G.; Capuani, S.; Brito Neto, J.F. de; Rosa, G.M. da; Wastowski, A.D.; Rigon, C.A.G. Dissimilaridade genética e análise de trilha de cultivares de soja avaliada por meio de descritores quantitativos. **Revista Ceres**, v.59, n.2, p.233-240, 2012.

Santos, L.H.; Oliveira, S.M.P. de; Malhado, C.H.M.; Carneiro, P.L. de; Martins Filho, R.; Lôbo, R.N.B.; Rodrigues, D.S. Estrutura populacional e tendências genéticas e fenotípicas da raça Guzerá no Nordeste do Brasil. **Revista Brasileira Saúde Produção Animal**, v.13, n.4, p.1032-1043, 2012.

Santos, J.A.S.; Teodoro, P.E.; Correa, A.M.; Soares, C.M.G.; Ribeiro, L.P.; Abreu, H.K.A. de. Desempenho agrônomico e divergência genética entre genótipos de feijão-caupi cultivados no ecótono Cerrado/Pantanal. **Bragantia**, v.73, n.4, p.377-382, 2014.

Scapim, C.A.; Braccini, A.L.; Pinto, R.J.B.; Amaral Júnior, A.T.; Rodovalho, M.A.; Silva, R.M.; Moterle, L.M. Componentes genéticos de médias e depressão por endogamia em populações de milho-pipoca. **Ciência Rural**, v.36, n.1, p.36-41, 2006.

Signorini, T.; Renesto, E.; Machado, M.F.P.S.; Bespalhok, D.N.; Monteiro, E.R. Diversidade genética de espécies de *Capsicum* com base em dados de isozimas. **Horticultura Brasileira**, v.31, p.534-539, 2013.

Silva, V.M.P.; Carneiro, P.C.S.; Menezes Júnior, J.A.N.; Carneiro, V.Q.; Carneiro, J.E.S.; Cruz, C.D.; Borém, A. Genetic potential of common bean parents for plant architecture improvement. **Scientia Agricola**, v.70, n.3, p.167-175, 2013.

Sprague, G.F. e Tatum, L.A. General vs. specific combining ability in single crosses of corn. **J. am. soc. agron**, v.34, n.10, p.923-932, 1942.

Vencovsky, R. e Barriga, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 486p., 1992.

Viegas, M.P.; Silva, C.L.S.P.; Moreira, J.P.; Cardin, L.T.; Azevedo, V.C.R.; Ciampi, A.Y.; Freitas, M.L.M.; Moraes, M.L.T.; Sebbenn, A.M. Diversidade genética e tamanho efetivo de duas populações de *Myracrodruon urundeuva* fr. all., sob conservação ex situ. **Revista Árvore**, v.35, p.769–779, 2011.

Weinberg, W. Über vererbungsgesetze beim menschen. **Molecular and General Genetics**, v.1, n.1, p.440-460, 1908.

Weir, B.S. **Genetic data analysis II: methods for discrete population genetic data**. Sunderland, Massachusetts: Sinauer Associates, 455p., 1996.

Wigginton, J.E.; Cutler, D.J.; Abecasis, G.R. A note on exact tests of Hardy-Weinberg equilibrium. **The American Journal Human Genetics**, v.76, p.887–893, 2005.

Wright, S. The general structure of populations. **Annals of Eugenics**, v.15, p.323-354, 1951.

Zhang, G.; Xu, S.; Mao, W.; Hu, Q.; Hu, Y. Determination of the genetic diversity of vegetable soybean [*Glycine max* (L.) Merr.] using EST-SSR markers. **Journal of Zhejiang University Science B**, v.14, n.4, p.279-288, 2013.

Ziegler, A.; Steen, K.V.; Wellek, S. Investigation Hardy-Weinberg equilibrium in case-control or cohort studies or meta-analysis. **Breast Cancer Research and Treatment**, v.128, n.1, p.197-201, 2011.

CAPÍTULO 1:

PREDIÇÃO DA DIVERSIDADE GENÉTICA DE COMBINAÇÕES HÍBRIDAS A PARTIR DA ABORDAGEM DE GENÉTICA DE POPULAÇÕES EM SUAS POPULAÇÕES GENITORAS

Predição da diversidade genética de combinações híbridas a partir da abordagem de genética de populações em suas populações genitoras

1. Resumo

Com o objetivo de avaliar a estrutura populacional e o grau de diferenciação entre pares de populações para prever o potencial de suas combinações híbridas por abordagem de Genética de Populações, foram simuladas dez populações genitoras e obtidas 45 combinações híbridas a partir do cruzamento, aos pares, das populações genitoras. As populações eram constituídas de 200 indivíduos com 50 locos independentes e dois alelos codominantes por loco e foram avaliadas quanto à condição de equilíbrio de Hardy-Weinberg (EHW), coeficiente de endogamia (F), heterozigosidade (H) e conteúdo médio de informação polimórfica (PIC). Além disso, foram utilizadas cinco medidas de dissimilaridade – Distância Euclidiana, Angular e Genotípica de Hedrick e Índice G_{ST} de Nei e F_{ST} de Wright – para avaliação da diversidade genética entre pares de populações genitoras e obtida as estimativas de correlação de Mantel entre estas medidas e os valores de PIC, H e F observados nas combinações híbridas. As populações genitoras estavam todas em EHW e a combinação híbrida que mais saiu desta condição foi a 3x5, com apenas 26% dos locos manifestando EHW. Esta combinação está também entre as de menores estimativas de F e maiores valores de H, indicando maior divergência entre as suas populações genitoras. Houve, de modo geral, concordância de indicação de combinações híbridas mais e menos divergentes pelas medidas de dissimilaridade e, pela correlação de Mantel, pode-se afirmar que é possível prever a diversidade genética de combinações híbridas a partir da avaliação da diversidade genética de suas populações genitoras.

2. Abstract

In order to evaluate population structure and differentiation degree between pairs of populations to predict the potential of their hybrid combinations by population genetics approach, ten progenitor populations were simulated, and 45 hybrids were obtained from crossing, in pairs, of the progenitor populations. Population consisted of 200 individuals with 50 independent loci and two codominant alleles per locus, and were evaluated for the condition of Hardy-Weinberg equilibrium (HWE), inbreeding coefficient (F), heterozygosity (H) and mean polymorphic information content (PIC). In addition, five dissimilarity measures were used – Euclidean Distance, Angle Distance and Genotypic of Hedrick Distance, G_{ST} Index of Nei, and F_{ST} Index of Wright – in order to evaluate the genetic diversity between pairs of progenitor populations and obtain the Mantel correlation estimates between these measures and the PIC, H and F values observed in hybrid combinations. All progenitor populations were in HWE. Hybrid combination 3x5 was the one that least fit in this condition, and had only 26% of the loci expressing HWE. This combination also had the lowest F estimates, and the highest H values, indicating greater divergence between the progenitor populations. In general, there was agreement of indication of more and less divergent hybrid combinations by the dissimilarity measures. Moreover, by the Mantel correlation, it can be affirmed that it is possible to predict the genetic diversity of hybrid combinations from the evaluation of the genetic diversity of their progenitor populations.

3. Introdução

Em um programa de melhoramento vegetal, uma das etapas de maior importância é a escolha de genitores para a formação de potenciais populações segregantes, pois dela depende o sucesso das etapas subsequentes e a eficiência do programa (Bertan et al., 2007; Pereira et al., 2007). Neste contexto, a escolha das populações segregantes de maior

potencial otimiza a utilização de recursos humanos e financeiros despendidos no programa de melhoramento (Pimentel et al., 2013). Para uma escolha eficiente das populações mais promissoras têm sido utilizadas as estimativas de capacidade geral e específica de combinação, obtidas a partir de cruzamentos dialélicos. Neste tipo de análise avalia-se o potencial e a diversidade genética de linhagens ou populações, a partir de informações de suas combinações híbridas, o que permite identificar as combinações parentais que gerem populações segregantes com máxima variabilidade genética para a seleção, prevendo, segundo os modelos biométricos abordados pela Genética Quantitativa, a heterose e as constituições híbridas mais promissoras (Rocha et al., 2014; Baldissera et al., 2012).

Ênfase tem sido dada ao estudo de diversidade genética em diversas culturas a fim de, por exemplo, identificar genótipos promissores para o melhoramento genético (Coimbra et al., 2010; Pimentel et al., 2013; Santos et al., 2014), quantificar a variabilidade genética (Rigon et al., 2012) e direcionar o melhorista na escolha mais apropriada para a formação de híbridos superiores (Barili et al., 2011; Ferreira et al., 2012). Para tanto, diferentes conjuntos de dados podem ser utilizados, como dados de pedigree, dados fenotípicos, bioquímicos, fitopatológicos, moleculares, entre outros. A diversidade genética pode ser analisada em diferentes níveis, incluindo genótipos individuais, acessos de germoplasma e populações; variando especialmente de acordo com a natureza do material genético avaliado em cada caso. A análise da diversidade em nível de população é considerada a mais complexa, por sofrer influência do número de indivíduos amostrados, número de locos, constituição genotípica e tamanho efetivo (Cruz et al., 2011).

Atualmente, os dados de marcadores moleculares tornaram-se informações úteis que podem aumentar a acurácia da seleção, reduzindo cargas de trabalho em campo e aumentando o sucesso em programas de melhoramento (Zhang et al., 2013). Muitos são

os trabalhos que utilizam informações de marcadores moleculares, fundamentados nas teorias de Genética de Populações, para estimar a diversidade genética entre e dentro de populações (Viegas et al., 2011; Pimentel et al., 2013). De forma geral, estes trabalhos obtêm estimativas de parâmetros como o grau de endogamia, sistema reprodutivo predominante e heterozigiosidade a fim de estabelecer a estratégia mais adequada na quantificação da variabilidade e conservação das espécies.

Porém, poucos são os estudos que associam a diversidade predita a partir das populações com o desempenho observado em seus híbridos a partir dos fundamentos da Genética de Populações, como endogamia, heterozigiosidade, equilíbrio de Hardy-Weinberg e diferenciação. Pode-se afirmar que uma das grandes dificuldades em realizar tal associação é a falta de informação conjunta de dados moleculares e fenotípicos de qualidade. Nesse contexto, destaca-se a simulação de dados que fornece resultados para análises mais elaboradas a respeito do sistema estudado, com a possibilidade de interpretação mais profunda e abrangente, sem perda de confiabilidade (Cruz, 2006).

Assim, o objetivo com este trabalho é estudar, via simulação de dados, a estrutura populacional e o grau de diferenciação entre pares de populações, a fim de prever a diversidade de suas combinações híbridas por abordagem de genética de populações, considerando aspectos de endogamia, heterozigiosidade, equilíbrio de Hardy-Weinberg e diferenciação.

4. Material e Métodos

4.1. Simulação dos dados

Foram simuladas dez populações genitoras, com 200 indivíduos (Bhering et al., 2008) cada, em equilíbrio de Hardy-Weinberg, para 50 locos independentes e dois alelos codominantes por loco, e obtidas 45 populações híbridas a partir da combinação, aos pares, das dez populações genitoras.

4.2. Avaliação do potencial das populações

A fim de validar o processo de simulação, as dez populações genitoras foram avaliadas em relação à sua condição de equilíbrio de Hardy-Weinberg (EHW) pelo teste de qui-quadrado (Cruz et al., 2011). Após confirmação da condição de EHW, os dados das combinações híbridas foram utilizados para estimação dos parâmetros descritores da estrutura populacional, incluindo o coeficiente médio de endogamia (Cruz et al., 2011) e o conteúdo médio de informação polimórfica (Botstein et al., 1980).

Considerando H_e como a frequência de heterozigotos esperada em uma população supostamente em equilíbrio de Hardy-Weinberg ($H_e = 2pq$) e H_o como a frequência de heterozigotos observada na população em avaliação, para locos com dois alelos, temos o coeficiente médio de endogamia (F) estimado por meio de:

$$F = 1 - \frac{H_o}{H_e}$$

E o conteúdo médio de informação polimórfica (PIC) foi estimado por:

$$PIC = 1 - \sum_{i=1}^a p_i^2 - \sum_{i,j=1}^a \sum_{(i \neq j)} p_i^2 p_j^2$$

em que p_i é a frequência do i -ésimo alelo (a) do loco estudado.

4.3. Grau de diferenciação entre as populações

Para estimar o grau de diferenciação entre os pares de populações genitoras foram utilizadas três medidas de dissimilaridade, duas baseadas em informação de frequência alélica (Distância Euclidiana e Distância Angular) e uma em informação de frequência genotípica (Distância Genotípica de Hedrick), e dois índices de fixação (Índice G_{ST} de Nei e F_{ST} de Wright), conforme descrito em Cruz et al. (2011), apresentado a seguir:

Considerando p_{ijk} e $p_{i'jk}$ a frequência do k -ésimo alelo para o j -ésimo loco das respectivas populações i e i' ; a_j , o número de alelos do j -ésimo loco; e L , o número de locos analisados, a distância euclidiana foi estimada por:

$$D_{Eii'} = \sqrt{\sum_{j=1}^L \sum_{k=1}^{a_j} (p_{ijk} - p_{i'jk})^2}$$

A Distância Angular é outra forma de representação da distância geométrica entre duas populações, cuja medida de distância é dada pelo ângulo (ψ) formado entre as linhas de projeção que unem a origem aos respectivos pontos das populações (Cavalli-Sforza e Edwards, 1967; Edwards, 1971). Esse ângulo corresponde ao comprimento da corda que liga as duas populações ($\sqrt{2D_{CE}^2}$). Considerando apenas um loco com dois alelos para cada uma das populações, i e i' , temos que:

$$\cos(\psi) = \sqrt{p_{i1}p_{i'1}} + \sqrt{p_{i2}p_{i'2}}$$

Se as populações i e i' não apresentam alelos em comum, então $\cos(\psi) = 0$. Caso as frequências alélicas sejam as mesmas ($i = i'$), então $\cos(\psi) = 1$. Assim, é adequado caracterizar o $\cos(\psi)$ como uma medida de similaridade e:

$$D_{CE}^2 = 1 - \cos(\psi)$$

A estimativa da Distância Genotípica de Hedrick (1971) foi dada por:

$$I_{Hii'} = \frac{\sum_{g \geq}^{a_j} \sum_{g'}^{a_{j'}} P_{ijgg'} P_{i'jgg'}}{\frac{1}{2} \left(\sum_{g \geq}^{a_j} \sum_{g'}^{a_j} P_{ijgg'}^2 + \sum_{g \geq}^{a_{j'}} \sum_{g'}^{a_{j'}} P_{i'jgg'}^2 \right)}$$

em que $P_{ijgg'}$ e $P_{i'jgg'}$ são as frequências do gg' -ésimo genótipo nas populações i e i' , respectivamente, em relação ao loco j ; e a_j é o número de alelos para o loco em consideração.

O numerador expressa a probabilidade de se tomar ao acaso genótipos idênticos provenientes da população i e i' , e o denominador é a probabilidade média de extrair

genótipos idênticos provenientes de uma mesma população em duas sucessivas retiradas independentes.

O índice G_{ST} de Nei (1973) é estimado baseando-se na identidade, heterozigosidade e diversidade genética. A identidade genética entre os pares de populações foi fornecida por:

$$J_{ii'} = \sum_{k=1}^{a_j} p_{ijk} p_{i'jk}$$

em que p_{ijk} é a frequência do k-ésimo alelo, do j-ésimo loco, na i-ésima população e $p_{i'jk}$, a frequência do k-ésimo alelo, do j-ésimo loco, na i'-ésima população.

Para os 50 locos foi calculada a heterozigosidade média da população, dada por:

$$\bar{H}_i = \frac{1}{L} \sum_{j=1}^L H_{i(j)}$$

em que $H_{i(j)}$ é a estimativa de heterozigosidade no j-ésimo loco e L, o número de locos amostrados. Em uma população em equilíbrio de Hardy-Weinberg, a estimativa h é igual à proporção esperada de heterozigotos.

A diversidade genética foi estimada por:

$$D_{ii'} = H_{ii'} - \frac{H_i + H_{i'}}{2}, \text{ com } D_{ii} = 0 \text{ e } D_{ii'} = D_{i'i}$$

A partir desses valores, foram estabelecidas as seguintes estatísticas:

FV	Identidade	Heterozigosidade	Diversidade
Dentro	\bar{J}_{IS}	$\bar{H}_{IS} = 1 - \bar{J}_{IS}$	$D_{IS} = 0$
Entre	\bar{J}_{ST}	$\bar{H}_{ST} = 1 - \bar{J}_{ST}$	$\bar{D}_{ST} = \bar{J}_{IS} - \bar{J}_{ST}$ $\bar{D}_{ST} = \bar{H}_{ST} - \bar{H}_{IS}$

Então,

$$G_{ST} = 100 \frac{\bar{D}_{ST}}{\bar{H}_{ST}}$$

O índice F_{ST} de Wright (1951) é uma medida da correlação de gametas entre as populações, dada por:

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

em que H_T corresponde à heterozigosidade na população total:

$$H_T = 1 - \sum_{k=1}^{a_j} \bar{p}_{.jk}^2 \quad \text{e} \quad \bar{p}_{.jk} = \sum_{i=1}^g w_i p_{ijk}$$

sendo p_{ijk} a frequência do alelo k , do loco j , na população i e w_i , o tamanho relativo da i -ésima população e H_S corresponde à heterozigosidade esperada entre as populações:

$$H_S = 1 - \sum_{i=1}^g w_i \left(\sum_{k=1}^{a_j} p_{ijk}^2 \right),$$

com $w_i = 1/g$ para todo i (populações com o mesmo número de indivíduos).

4.4. Correlação entre matrizes

O teste de Mantel (Manly, 1997) foi utilizado para correlacionar as matrizes de distância das populações genitoras com as matrizes dos valores de conteúdo médio de informação polimórfica (PIC), heterozigosidade (H) e coeficiente de endogamia (F) das combinações híbridas, a fim de prever o desempenho dessas combinações a partir de suas populações genitoras.

A simulação e análise dos dados foram realizadas no Laboratório de Bioinformática da Universidade Federal de Viçosa, localizado no Instituto de Biotecnologia aplicada a Agropecuária (BIOAGRO) utilizando aplicativo computacional Genes (Cruz, 2013).

5. Resultados e Discussão

Observou-se que todas as populações genitoras encontram-se em equilíbrio de Hardy-Weinberg (EHW), com pelo menos 86% de locos manifestando EHW (Tabela 1). Dentre os fatores que perturbam o EHW, como seleção, migração, mutação e deriva genética, apenas o último, função de problemas com amostragem, seria fator perturbador no processo de simulação. Assim, estes resultados validam o processo de simulação dessas populações. O processo de simulação de dados fornece resultados para análises mais elaboradas a respeito do sistema estudado, com a possibilidade de interpretação mais profunda e abrangente, sem perda de confiabilidade (Harrel, 2000). Almeida et al. (2012) também relatam elevada acurácia no processo de simulação de populações utilizando o software GENES.

Considerando as combinações híbridas, observou-se que nenhuma das 45 encontram-se em EHW, com porcentagem máxima de 62% de locos por população manifestando EHW (Tabela 1), indicando elevada diversidade genética entre as populações genitoras. As combinações híbridas 3x5, 1x9, 2x4 e 8x10 foram as que apresentaram menor porcentagem de locos manifestando EHW (Tabela 1).

Estes resultados indicam maior divergência entre as populações genitoras destas combinações, uma vez que se espera EHW em uma população híbrida apenas quando esta for resultante de populações genitoras que apresentem as mesmas frequências alélicas. Neste contexto, as combinações 4x9, 7x8 e 7x10, foram oriundas de populações menos divergentes, com 62% dos locos por combinação híbrida manifestando EHW (Tabela 1).

Quanto ao coeficiente de endogamia (F), observou-se que todos os valores estimados para as populações genitoras foram próximos a zero e que, entre as combinações híbridas, as estimativas foram todas negativas, variando entre -0,1146 (1x7) e -0,2967 (8x10) (Tabela 1).

Tabela 1 – Porcentagem de locos manifestando equilíbrio de Hard-Weinberg (EHW) (abaixo da diagonal) e coeficiente de endogamia (F) (acima da diagonal) das populações híbridas e valores referentes às populações genitoras (nas marginais)

Populações	1	2	3	4	5	6	7	8	9	10	EHW
1		-0,2219	-0,2404	-0,2037	-0,2416	-0,2593	-0,1146	-0,2263	-0,2651	-0,1967	100%
2	48%		-0,1860	-0,2348	-0,1888	-0,2225	-0,2323	-0,1582	-0,1814	-0,2214	86%
3	44%	50%		-0,1936	-0,2918	-0,1932	-0,2449	-0,1548	-0,1829	-0,2429	98%
4	54%	32%	56%		-0,1771	-0,2190	-0,1452	-0,2174	-0,1317	-0,2059	90%
5	56%	54%	26%	60%		-0,1968	-0,2034	-0,2031	-0,2011	-0,2248	88%
6	48%	54%	52%	52%	54%		-0,2123	-0,2134	-0,2156	-0,2641	96%
7	46%	50%	34%	58%	56%	46%		-0,1603	-0,2154	-0,1887	94%
8	46%	42%	56%	44%	58%	52%	62%		-0,2146	-0,2967	98%
9	32%	60%	50%	62%	48%	56%	36%	56%		-0,1940	96%
10	52%	40%	46%	60%	54%	52%	62%	32%	56%		88%
F	0,0048	0,0479	0,0217	0,0587	0,0179	-0,0066	-0,0425	-0,0291	-0,0352	0,0556	

Vale ressaltar que o cálculo do coeficiente de endogamia, neste caso, foi realizado com base na frequência de heterozigotos observados e a frequência de heterozigotos esperada na população sob hipótese de EHW. Desta forma, os valores obtidos podem variar de zero a um negativo (-1). Assim, valores mais negativos de F nas populações indicam maior variabilidade entre suas populações genitoras. Neste sentido, as combinações 3x5 e 8x10 seriam as mais divergentes, enquanto as 1x7, 4x9 e 4x7, as menos divergentes.

Na Tabela 2 são apresentados os valores de heterozigosidade (H) e o conteúdo médio de informação polimórfica (PIC) observados nas populações genitoras e suas respectivas combinações híbridas. De modo geral, foram observados maiores valores de H para as combinações híbridas, comparados aos observados para as populações genitoras. As combinações 8x10 (0,5645), 3x5 (0,5620) e 3x7 (0,5507) apresentaram as maiores estimativas de H. Vale ressaltar que as populações genitoras 10 e 5 apresentaram os menores valores de H, com estimativas de 0,2725 e 0,2942, respectivamente, e as populações genitoras 8, 3 e 7 estão entre as quatro de maiores valores de H.

Quanto ao PIC, observou-se que os valores variam entre 0,2353 e 0,2890 nas populações genitoras e entre 0,2933 e 0,3477 nas populações híbridas (Tabela 2), todos inferiores aos valores de H. Ott (1992) relata que os valores de PIC estimados devem ser sempre inferiores aos estimados para heterozigosidade (H) e que o valor máximo de PIC esperado para locos com dois alelos é de 0,375.

Vários autores confirmam tal expectativa, incluindo Vieira et al. (2009), Reis et al. (2011) e Crispim et al. (2012).

Tabela 2 – Valores de heterozigosidade (H) (acima da diagonal) e conteúdo médio de informação polimórfica (PIC) (abaixo da diagonal) das populações híbridas e valores referentes às populações genitoras (nas marginais)

Populações	1	2	3	4	5	6	7	8	9	10	PIC
1		0,5074	0,5275	0,5094	0,5004	0,5144	0,4442	0,5200	0,5280	0,4434	0,2467
2	0,3234		0,5043	0,5303	0,4707	0,4956	0,5433	0,4958	0,5009	0,5158	0,2760
3	0,3309	0,3309		0,5325	0,5620	0,4883	0,5507	0,4912	0,4643	0,5239	0,2817
4	0,3272	0,3332	0,3440		0,4850	0,5153	0,4922	0,5409	0,4765	0,4883	0,2890
5	0,3157	0,3117	0,3370	0,3228		0,4553	0,4939	0,4931	0,4874	0,4727	0,2433
6	0,3202	0,3184	0,3186	0,3310	0,3006		0,5097	0,5089	0,5168	0,5076	0,2457
7	0,3124	0,3415	0,3420	0,3345	0,3195	0,3275		0,4864	0,5107	0,4703	0,2756
8	0,3295	0,3324	0,3302	0,3438	0,3204	0,3268	0,3272		0,5484	0,5645	0,2844
9	0,3261	0,3312	0,3340	0,3274	0,3187	0,3297	0,3282	0,3477		0,4847	0,2870
10	0,2933	0,3285	0,3273	0,3160	0,3047	0,3142	0,3109	0,3367	0,3182		0,2353
H	0,3072	0,3279	0,3498	0,3435	0,2942	0,3084	0,3602	0,3723	0,3730	0,2725	

Entretanto, o conteúdo médio de informação polimórfica (PIC) dos marcadores se mostrou pouco informativo, uma vez que, segundo Botstein et al. (1980), o mesmo deve estar entre 0,25 e 0,50 para o marcador ser considerado moderadamente polimórfico e superior a 0,50 para altamente polimórfico. Estes resultados eram esperados em razão de ter sido considerado neste trabalho apenas dois alelos para cada um dos 50 locos em questão.

Na Tabela 3 são apresentadas as estimativas do coeficiente de correlação de Mantel entre a matriz de distâncias estimadas nos pares de populações genitoras, por diferentes medidas de dissimilaridade, com os valores de PIC, H e F das combinações híbridas. Esta é uma medida de predição da diversidade das combinações híbridas a partir da diversidade observada nas populações genitoras. De modo geral, foram observadas estimativas de correlação de baixa magnitude entre as medidas de dissimilaridade dos pares de populações com os valores de PIC das combinações híbridas, com alguns valores positivos e outros negativos. Estes baixos valores podem ser devidos à baixa variação entre os valores de PIC.

Tabela 3 – Correlação da distância genética entre pares de populações genitoras com a heterozigosidade (H) e o coeficiente de endogamia (F) observado nas populações híbridas

Medidas de dissimilaridade	Descritores populacionais		
	PIC	H	F
Distância Euclidiana	0,1566	0,6137**	-0,7910**
Distância Angular	0,2000	0,6320**	-0,7786**
Distância Genotípica de Hedrick	0,2950*	0,6717**	-0,7239**
Estatística G_{ST} de Nei	-0,1138	0,3941**	-0,7641**
Estatística F_{ST} de Wright	-0,1138	0,3941**	-0,7641**

* e ** significativo a 5 e 1% de probabilidade pelo teste de Mantel, respectivamente.

Já com os valores de H, os coeficientes de correlação foram todos positivos e maiores para a distância euclidiana, distância angular e distância genotípica de Hedrick comparados aos índices G_{ST} e F_{ST} , embora todas tenham sido significativas pelo teste de Mantel a 1% de

probabilidade. Já considerando os valores de F , todas as estimativas foram elevadas e negativas independentemente da medida de distância utilizada. Estes resultados demonstram que, para populações em EHW e locos com dois alelos, as distâncias euclidiana, angular e de Hedrick apresentam superioridade na predição da diversidade das populações híbridas a partir da diversidade observada nas populações genitoras, comparadas aos índices G_{ST} e F_{ST} .

Na Tabela 4 são apresentadas os três pares de combinações mais e menos divergentes com base nas cinco medidas de dissimilaridade. Houve, de modo geral, concordância de indicação de combinações híbridas mais e menos divergentes pelas medidas de distância euclidiana, angular e genotípica de Hedrick. As distâncias Euclidiana e Angular são baseadas em propriedades geométricas e consideram as informações das frequências alélicas para discriminação da diversidade genética. Já a distância genotípica de Hedrick propõe uma maneira alternativa de quantificar a dissimilaridade entre populações, considerando estatísticas fundamentadas nas frequências genotípicas e não nas frequências alélicas.

Tal metodologia pode ser vantajosa em relação às demais, uma vez que populações que diferem genotipicamente podem ser erroneamente consideradas idênticas, se a análise for baseada apenas em suas frequências alélicas. Os índices de fixação são usados para medir qualquer desvio da condição de equilíbrio de Hardy-Weinberg quando há dois alelos para um determinado loco (Wright, 1951). O índice de fixação segundo o modelo de Nei (1973) estima parâmetros de identidade, heterozigosidade e diversidade genética para estabelecimento do Índice de Diversidade, avaliando sempre o que ocorre dentro da população e entre pares de populações. A medida de fixação de Wright (1951) compara as frequências dos genes entre as populações e descreve a probabilidade de dois indivíduos de populações diferentes possuírem um alelo idêntico por ascendência.

Tabela 4 – Pares de populações genitoras simuladas mais e menos divergentes de acordo com as estimativas de medidas de distância Euclidiana, Angular e Genotípica de Hedrick e com os Índices de Fixação G_{ST} de Nei e F_{ST} de Wright

Medidas de dissimilaridade	Combinações Híbridas					
	Mais Divergentes			Menos Divergentes		
Distância Euclidiana	8x10	2x10	6x10	4x9	3x8	7x8
	(0,6642)	(0,6550)	(0,6542)	(0,4766)	(0,5117)	(0,5190)
Distância Angular	8x10	2x10	3x5	4x9	3x8	3x9
	(0,4588)	(0,4466)	(0,4447)	(0,3179)	(0,3438)	(0,3471)
Distância Genotípica de Hedrick	8x10	2x10	2x7	4x9	1x10	3x8
	(0,4207)	(0,4012)	(0,4010)	(0,2685)	(0,2891)	(0,3021)
Índice G_{ST} de Nei	6x10	8x10	2x10	4x9	3x8	7x8
	(0,2645)	(0,2533)	(0,2531)	(0,1354)	(0,1540)	(0,1604)
Índice F_{ST} de Wright	6x10	8x10	2x10	4x9	3x8	7x8
	(0,2645)	(0,2533)	(0,2531)	(0,1354)	(0,1540)	(0,1604)

Os valores entre parênteses referem-se às medidas de dissimilaridade estimadas.

Na Tabela 5 são apresentadas as populações genitoras simuladas mais e menos divergentes, entre si, com base nas cinco medidas de dissimilaridade – Distância Euclidiana, Distância Angular, Distância Genotípica de Hedrick, Índice G_{ST} de Nei e Índice F_{ST} de Wright. Nota-se que houve, de modo geral, concordância na indicação de quais são as populações genitoras mais e menos divergentes pelas medidas de dissimilaridade. Tal resultado corrobora com o apresentado na Tabela 3, que indica que as medidas de dissimilaridade são equivalentes na predição da diversidade das combinações híbridas a partir da diversidade observada nas populações genitoras.

Tabela 5 – Diversidade relativa para cada uma das 10 populações genitoras simuladas, com base nas cinco medidas de dissimilaridade genética: Distância Euclidiana (D_E), Distância Angular (D_A), Distância Genotípica de Hedrick (D_H), Índice GST de Nei (G_{ST}) e Índice FST de Wright (F_{ST})

População Genitora	Mais Divergentes					População Genitora	Menos Divergentes				
	D_E	D_A	D_H	G_{ST}	F_{ST}		D_E	D_A	D_H	G_{ST}	F_{ST}
1	6	6	6	6	6	1	7	10	10	7	7
2	10	10	10	10	10	2	8	5	4	4	4
3	5	5	5	5	5	3	8	8	8	8	8
4	6	6	6	6	6	4	9	9	9	9	9
5	3	3	3	3	3	5	2	2	6	2	2
6	10	10	10	10	10	6	3	3	5	3	3
7	2	2	2	2	2	7	8	8	9	8	8
8	10	10	10	10	10	8	3	3	3	3	3
9	6	1	6	6	6	9	4	4	4	4	4
10	8	8	8	6	8	10	1	1	1	4	4

Nota-se ainda que a diversidade relativa das populações genitoras, com base nas cinco medidas de dissimilaridade genética, concorda com os descritores populacionais das respectivas combinações híbridas. Deste modo, as estimativas de heterozigosidade, conteúdo médio de informação polimórfica, porcentagem de locos manifestando EHW e coeficiente de endogamia (Tabelas 1 e 2), medidas nas populações genitoras, podem auxiliar na predição da diversidade genética de combinações híbridas.

6. Referências Bibliográficas

Almeida, I.F.; Cruz, C.D.; Resende, M.D.V.; Almeida, R.V.; Mendes, F.R. Métodos de estimação e validação na seleção genômica. **Revista Agrotecnologia**, v.3, n.2, p.1-11, 2012.

Baldissera, J.N.C.; Valentini, G.; Coan, M.M.D.; Almeida, C.B.; Guidolin, A.F.; Coimbra, J.L.M. Capacidade combinatória e efeito recíproco em características agrônômicas do feijão. **Semina: Ciências Agrárias**, v.33, n.2, p.471-480, 2012.

Barili, L.D.; Vale, N.M.; Arruda, B.; Toaldo, D.; Rocha, F.; Coimbra, J.L.M.; Bertoldo, J.G.; Guidolin, A.F. Escolha de genitores contrastantes para compor o bloco de cruzamentos de genótipos de feijão. **Revista Brasileira de Agrociência**, v.17, n.3-4, p.303-310, 2011.

Bertan, I.; Carvalho, F.I.F.; Oliveira, A.C. Parental selection strategies in plant breeding programs. **Journal of Crop Science and Biotechnology**, v.10, p.211–222, 2007.

Bhering, L.L.; Cruz, C.D. Tamanho de população ideal para mapeamento genético em famílias de irmãos completos. **Pesquisa Agropecuária Brasileira**, v.43, n.3, p.379-385, 2008.

Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, v. 32, n. 3, p. 314, 1980.

Cavalli-Sforza, L.L.; Edwards, A.W.F. Phylogenetic analysis. Models and estimation procedures. **American Journal of Human Genetics**, v.19, n.3, p.233, 1967.

Coimbra, R.R.; Miranda, G.V.; Cruz, C.D.; Melo, A.V.; Eckert, F.R. Caracterização e divergência genética de populações de milho resgatadas do Sudeste de Minas Gerais. **Revista Ciência Agronômica**, v.41, n.1, p.159–166, 2010.

Crispim, B.A.; Silva, D.B.S.; Banari, A.C.B.; Seno, L.O.; Grisolia, A.B. Discriminação alélica em ovinos naturalizados do Pantanal Sul-Matogrossense por meio de marcadores microssatélites. **Journal of the Selva Andina Research Society**, v.3, n.1, p.3-13, 2012.

Cruz, C.D. **Programa Genes: Análise Multivariada e Simulação**. Viçosa, MG: Ed. UFV, 2006.

Cruz, C.D.; Ferreira, F.M.; Pessoni, L.A. **Biometria aplicada ao estudo da diversidade genética**. 1ed, Ed. Suprema, 2011.

Cruz, C.D. GENES: a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum Agronomy**, v.35, p.271-276, 2013.

Edwards, A.W.F. Distances between populations on the basis of gene frequencies. **Biometrics**, p.873-881, 1971.

Ferreira, C.B.B.; Lopes, M.T.G.; Lopes, R.; Cunha, R.N.V.; Moreira, D.A.M.; Barros, W.S.; Matiello, R.R. Diversidade genética molecular de progênies de dendezeiro. **Pesquisa Agropecuária Brasileira**, v.47, n.3, p.378-384, 2012.

Harrel, C.R.; Ghosh, B.K.; Bowden, R. **Simulation Using ProModel**. McGraw-Hill, 2000.

Hedrick, P.W. A new approach to measuring genetic similarity. **Evolution**, v.25, n.2, p.276-280, 1971.

Manly, B.F.J. Randomization, bootstrap and Monte Carlo methods in biology. London: Chapman & Hall, 1997.

Nei, M. Analysis of gene diversity in subdivided populations. **Proceedings of the National Academy of Sciences of USA**, v.70, n.12, p.3321-3323, 1973.

Ott, J. Strategies for characterizing highly polymorphic markers in human gene mapping. **American journal of human genetics**, v.51, n.2, p.283, 1992.

Pereira, H.S.; Santos, J.B.; Abreu, A.F.B.; Couto, K.R. Informações fenotípicas e marcadores microssatélites de QTL na escolha de populações segregantes de feijoeiro. **Pesquisa Agropecuária Brasileira**, v.42, p.707-713, 2007.

Pimentel, A.J.B.; Ribeiro, G.; Souza, M.A.; Moura, L.M.; Assis, J.C.; Machado, J.C. Comparação de métodos de seleção de genitores e populações segregantes aplicados ao melhoramento de trigo. **Bragantia**, v.72, n.2, p.113-121, 2013.

Reis, R.V.; Oliveira, E.J.; Viana, A.P.; Pereira, T.N.S.; Pereira, M.G.; Silva, M.G.M. Diversidade genética em seleção recorrente de maracujazeiro amarelo- detectada por marcadores microssatélites. **Pesquisa Agropecuária Brasileira**, v.46, n.1, p.51-57, 2011.

Rigon, J.P.G.; Capuani, S.; Brito Neto, J.F.; Rosa, G.M.; Wastowski, A.D.; Rigon, C.A.G. Dissimilaridade genética e análise de trilha de cultivares de soja avaliada por meio de descritores quantitativos. **Revista Ceres**, v.59, n.2, p.233-240, 2012.

Rocha, F.; Stinghen, J.C.; Gemeli, M.S.; Coimbra, J.L.M.; Guidolin, A.F. Análise dialéctica como ferramenta na seleção de genitores em feijão. **Revista Ciência Agronômica**, v.45, n.1, p.74-81, 2014.

Santos, J.A.S.; Teodoro, P.E.; Correa, A.M.; Soares, C.M.G.; Ribeiro, L.P.; Abreu, H.K.A. Desempenho agronômico e divergência genética entre genótipos de feijão-caupi cultivados no ecótono Cerrado/Pantanal. **Bragantia**, v.73, n.4, p.377-382, 2014.

Viegas, M.P.; Silva, C.L.S.P.; Moreira, J.P.; Cardin, L.T.; Azevedo, V.C.R.; Ciampi, A.Y.; Freitas, M.L.M.; Moraes, M.L.T.; Sebbenn, A.M. Diversidade genética e tamanho efetivo de duas populações de *Myracrodruon urundeuva* fr. all., sob conservação ex situ. **Revista Árvore**, v.35, p.769-779, 2011.

Vieira, E.S.N.; Schuster, I.; Silva, R.D.; Oliveira, M.D. Variabilidade genética em cultivares de soja determinada com marcadores microssatélites em gel de agarose. **Pesquisa Agropecuária Brasileira**, v.44, n.11, p.1460-1466, 2009.

Wright, S. The general structure of populations. **Annals of Eugenics**, v.15, p.323-354, 1951.

Zhang, G.; Xu, S.; Mao, W.; Hu, Q.; Hu, Y. Determination of the genetic diversity of vegetable soybean [*Glycine max* (L.) Merr.] using EST-SSR markers. **Journal of Zhejiang University Science B**, v.14, n.4, p.279-288, 2013.

CAPÍTULO 2:

ABORDAGEM DE GENÉTICA DE POPULAÇÕES NA ESCOLHA DE PROGENITORES POTENCIAIS AO MELHORAMENTO GENÉTICO

Abordagem de genética de populações na escolha de progenitores potenciais ao melhoramento genético

1. Resumo

O objetivo com o presente trabalho foi avaliar a abordagem de genética de populações na escolha de progenitores potenciais ao melhoramento genético. Para obtenção das informações genotípicas, foram simuladas dez populações genitoras, com 200 indivíduos cada, em equilíbrio de Hardy-Weinberg, para 50 locos independentes e dois alelos codominantes por loco, e obtidas 45 combinações híbridas a partir do cruzamento, aos pares, das dez populações genitoras. Além disso, para as informações fenotípicas, foram simuladas 12 variáveis quantitativas pela ação de alelos de 20 locos, selecionados ao acaso entre os 50 previamente simulados, em delineamento inteiramente ao acaso, com valores de herdabilidade de 20, 40, 60 e 80%, combinados com um efeito aditivo diferencial, de acordo com a importância do loco, estabelecidos a partir de uma distribuição binomial e grau médio de dominância com valores iguais a zero, meio e um. Observou-se que, em características de efeito aditivo ($gmd = 0$), a escolha de genitores com maior frequência de alelos favoráveis pode ser feita com base em suas médias fenotípicas apenas, sem a necessidade de realização de cruzamentos dialélicos. Desta forma, baseando-se nas informações fenotípicas, as populações genitoras 2, 4, 8 e 7 são as que apresentam potencial para formação de combinações híbridas heteróticas, e as populações 9, 8, 7 e 4, baseando-se nas informações moleculares. Nota-se que, pela abordagem de genética de populações, houve coincidência de 75% na recomendação das populações genitoras potenciais ao melhoramento genético, sem a necessidade de realização dos cruzamentos dialélicos. Nesse sentido, conclui-se que a associação da abordagem de genética de populações às informações quantitativas pode incrementar o ganho no melhoramento genético.

2. Abstract

The aim of this study was to evaluate the population genetics approach in selecting potential progenitors for genetic improvement. To obtain the genotypic information, ten progenitor populations were simulated, consisted of 200 individuals each, all in Hardy-Weinberg equilibrium, with 50 independent loci and two codominant alleles per locus, and 45 hybrids were obtained from crossing, in pairs, of the progenitor populations. To obtain the phenotypic information, 12 quantitative variables were simulated by the action of alleles of 20 loci, randomly selected among the 50 loci previously simulated, in a completely randomized design. Heritability values were 20, 40, 60 and 80%, combined with differential additive effect, according to the importance of the locus, established from a binomial distribution, with medium degree of dominance of 0, 0.5 and 1. It was observed that, in additive effect characteristics ($gmd = 0$), the choice of parents with highest frequency of favorable alleles can be made based on their phenotypic means only, without the need of diallel crossing to evaluate the hybrid combinations. In addition, diallel analysis for such characteristics was insufficient to detect the genetic divergence between the progenitor populations. Thus, progenitor populations 2, 4, 8 and 7 were chosen based on phenotypic information; and populations 9, 8, 7 and 4, were chosen based on molecular information. It was observed that, by the population genetics approach, there was agreement of 75% on the recommendation of potential progenitor populations for genetic improvement, without the need of diallel crosses. In this sense, it is concluded that the association of genetic approach of populations with quantitative information may increase gain in plant breeding.

3. Introdução

A divergência genética tem grande importância em programas de melhoramento genético envolvendo hibridações, pois as populações segregantes de maior efeito heterótico associado àquelas de bom desempenho proporcionam maior chance de sucesso na obtenção de linhagens superiores (Cruz et al., 2012).

A divergência genética de um grupo de genótipos ou populações pode ser avaliada por técnicas biométricas e, ou processos preditivos. Os cruzamentos dialélicos, fundamentados no comportamento das combinações híbridas oriundas dos cruzamentos entre as populações genitoras, estão entre as técnicas biométricas desenvolvidas com essa finalidade (Baldissera et al., 2012; Pimentel et al., 2013). Nos processos preditivos a diversidade genética é estimada a partir apenas do desempenho per se dos genótipos ou populações. Teorias relativas às análises dialélicas foram apresentadas a bastante tempo (Sprague e Tatum, 1942). Esta técnica permite estimar a capacidade geral e específica de combinação, que são funções, respectivamente, da frequência de alelos favoráveis envolvidos no controle genético da característica em questão e da diversidade dos genitores cruzados (Cruz et al., 2012; Engelsing et al., 2011).

De maneira geral, as análises dialélicas têm sido utilizadas com sucesso nos programas de melhoramento (Silva et al., 2013; Rocha et al., 2014). Entretanto, as causas genéticas que determinam a máxima expressão da complementariedade gênica entre duas populações, no contexto da estruturação da população, ainda são desconhecidas. Os estudos da capacidade combinatória são realizados a partir de informações fenotípicas dos potenciais genitores e, portanto, os efeitos que são estimados estão restritos a esta natureza de dados, não permitindo qualificar e quantificar quais as propriedades populacionais são mais determinantes do sucesso da combinação híbrida.

Com o advento dos marcadores de DNA, muitos foram os trabalhos desenvolvidos, fundamentados nas teorias de Genética de Populações, os quais possibilitam melhor entendimento da estruturação das populações de trabalho. De forma geral, estes trabalhos obtêm estimativas de parâmetros como o grau de endogamia, sistema reprodutivo predominante, heterozigosidade, equilíbrio de Hardy-Weinberg e diferenciação (Viegas et al., 2011; Pimentel et al., 2013).

É conhecido pela biometria, em que se consideram efeitos médios ou ação monogênica, que o vigor híbrido é determinado pela diferença de frequência gênica das populações que se cruzam e pelo desvio da dominância. Estudos práticos, fundamentados apenas em valores fenotípicos, são incapazes de prover informações sobre a dinâmica populacional de cada população genitora, envolvendo informações sobre endogamia, heterozigose, equilíbrio, desequilíbrio de ligação, dentre outros. Assim, nestes estudos, não há complementariedade de informação molecular e fenotípica.

Frente ao exposto, torna-se fundamental conduzir estudos biométricos sobre capacidade geral e específica de combinação, agregando ou associando conhecimento sobre as populações de trabalho. Espera-se com esta atitude ter maior entendimento de causa e efeito sobre o sucesso e insucesso no investimento em cruzamentos de potenciais genitores. Assim, o objetivo com este trabalho foi avaliar a abordagem de genética de populações na escolha de progenitores.

4. Material e Métodos

4.1. Simulação dos dados genotípicos

Foram simuladas dez populações genitoras, com 200 indivíduos cada (Bhering et al., 2008), em equilíbrio de Hardy-Weinberg, para 50 locos independentes e dois alelos codominantes por loco, e obtidas 45 populações híbridas a partir da combinação, aos pares, das dez populações genitoras.

4.2. Simulação dos dados fenotípicos

Foram simuladas 12 variáveis quantitativas (Quadro 1) pela ação de alelos de 20 locos, tomados ao acaso entre os 50 previamente simulados, com média igual a 100 e valores de herdabilidade de 20, 40, 60 e 80%, combinados com grau médio de dominância com valores iguais a zero, meio e um.

Neste trabalho, o interesse é pelos genótipos com as maiores médias fenotípicas e as maiores capacidades geral e específica de combinação para as características simuladas.

O modelo estatístico utilizado para simular o valor fenotípico foi o aditivo-dominante:

$$Y_{ij} = \mu + G_i + \varepsilon_{ij}$$

em que:

Y_{ij} : observação simulada de uma dada característica para o i -ésimo indivíduo pertencente à j -ésima população;

μ : média geral da característica, estabelecido a priori;

G_i : efeito associado ao i -ésimo genótipo e

ε_{ij} : erro aleatório, sendo $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Quadro 1: Parâmetros para simulação dos dados fenotípicos

Característica	Número de Locos	Critério	Locos controladores da característica	Efeito a e d	gmd	média	h ²	Modelo
X1	20	Aleatório	40 42 16 4 9 39 31 22 18 49 32 38 8 5 33 12 35 48 6 28	Binomial	0.0	100	0.2	Aditivo-Dominante
X2			5 47 37 7 36 8 32 25 20 41 49 29 15 4 39 26 1 13 2 18				0.4	
X3			24 36 40 32 11 10 50 47 48 46 42 8 20 30 5 34 14 35 4 21				0.6	
X4			4 45 37 31 17 5 47 9 21 1 29 23 11 7 3 33 25 36 39 44				0.8	
X5	20	Aleatório	4 15 48 28 6 30 8 24 25 41 5 12 18 42 32 17 47 31 40 34	Binomial	0.5	100	0.2	Aditivo-Dominante
X6			48 1 49 16 50 5 39 36 25 22 24 27 43 2 15 37 29 26 9 20				0.4	
X7			37 47 50 19 14 5 10 21 16 4 20 6 40 48 31 12 27 30 46 42				0.6	
X8			30 15 23 45 16 20 17 21 46 3 7 38 34 39 35 26 24 8 37 10				0.8	
X9	20	Aleatório	31 20 26 27 48 44 24 32 49 11 34 18 12 8 22 38 3 42 47 40	Binomial	1.0	100	0.2	Aditivo-Dominante
X10			33 9 23 27 11 32 1 46 47 3 41 43 15 44 30 31 34 14 39 7				0.4	
X11			32 14 29 2 30 16 21 20 17 4 39 15 1 26 44 24 3 45 40 11				0.6	
X12			21 15 49 33 43 34 6 14 40 30 2 28 8 50 46 20 36 37 19 22				0.8	

a = aditivo; d = dominante; gmd = grau médio de dominância (d/a); h² = herdabilidade

4.3. Análise dialélica dos dados fenotípicos

Com base nas informações obtidas a partir da simulação fenotípica, foi realizada a análise dialélica proposta por Griffing (1956), para estimação das capacidades geral e específica de combinação. O modelo estatístico proposto por Griffing (1956) decompõe a média Y_{ij} da combinação híbrida ($i \neq j$) ou da população genitora ($i = j$) nos seguintes efeitos:

$$Y_{ij} = m + g_i + g_j + s_{ij} + \bar{e}_{ij}$$

em que:

m : efeito médio de todos os tratamentos;

g_i : efeito da capacidade geral de combinação da população genitora i ;

g_j : efeito da capacidade geral de combinação da população genitora j ;

s_{ij} : efeito da capacidade específica de combinação do cruzamento Y_{ij} ;

\bar{e}_{ij} : erro experimental médio associado às médias da tabela dialélica.

4.4. Análise de dissimilaridade entre genitores

a. Diversidade fenotípica

Foi calculada a distância generalizada de Mahalanobis (D^2) entre as populações i e i' , dada por:

$$D_{ii'}^2 = (\bar{X}_i - \bar{X}_{i'})' S^{-1} (\bar{X}_i - \bar{X}_{i'})$$

em que:

S : é a matriz de variâncias e covariâncias amostral comum a todos os indivíduos e

\bar{X}_i : vetor de médias fenotípicas observadas nas populações i e i' .

b. Diversidade molecular

b.1. Coeficiente de Heterozigosidade (H) e Endogamia (F)

Considerando H_e como a frequência de heterozigotos esperada em uma população supostamente em equilíbrio de Hardy-Weinberg ($H_e = 2pq$) e H_o como a frequência de heterozigotos observada na população em avaliação, para locos com dois alelos, temos o coeficiente médio de endogamia (F) estimado por meio de:

$$F = 1 - \frac{H_o}{H_e}$$

b.2. Distância Euclidiana

Considerando p_{ijk} e $p_{i'jk}$ a frequência do k-ésimo alelo para o j-ésimo loco das respectivas populações i e i'; a_j , o número de alelos do j-ésimo loco; e L, o número de locos analisados, a distância euclidiana foi estimada por:

$$D_{Eii'} = \sqrt{\sum_{j=1}^L \sum_{k=1}^{a_j} (p_{ijk} - p_{i'jk})^2}$$

b.3. Distância Genotípica de Hedrick

A estimativa da Distância Genotípica de Hedrick (1971) foi dada por:

$$I_{Hii'} = \frac{\sum_{g \geq}^{a_j} \sum_{g'}^{a_g} P_{ijgg'} P_{i'jgg'}}{\frac{1}{2} \left(\sum_{g \geq}^{a_j} \sum_{g'}^{a_j} P_{ijgg'}^2 + \sum_{g \geq}^{a_j} \sum_{g'}^{a_j} P_{i'jgg'}^2 \right)}$$

em que $P_{ijgg'}$ e $P_{i'jgg'}$ são as frequências do gg'-ésimo genótipo nas populações i e i', respectivamente, em relação ao loco j; e a_j é o número de alelos para o loco em consideração.

O numerador expressa a probabilidade de se tomar ao acaso genótipos idênticos provenientes da população i e i' , e o denominador é a probabilidade média de extrair genótipos idênticos provenientes de uma mesma população em duas sucessivas retiradas independentes.

b.4. Índice G_{ST} de Nei

O índice G_{ST} de Nei (1973) é estimado baseando-se na identidade, heterozigosidade e diversidade genética. A identidade genética entre os pares de populações foi fornecida por:

$$J_{ii'} = \sum_{k=1}^{a_j} p_{ijk} p_{i'jk}$$

em que p_{ijk} é a frequência do k -ésimo alelo, do j -ésimo loco, na i -ésima população e $p_{i'jk}$, a frequência do k -ésimo alelo, do j -ésimo loco, na i' -ésima população.

Para os 50 locos foi calculada a heterozigosidade média da população, dada por:

$$\bar{H}_i = \frac{1}{L} \sum_{j=1}^L H_{i(j)}$$

em que $H_{i(j)}$ é a estimativa de heterozigosidade no j -ésimo loco e L , o número de locos amostrados. Em uma população em equilíbrio de Hardy-Weinberg, a estimativa h é igual à proporção esperada de heterozigotos.

A diversidade genética foi estimada por:

$$D_{ii'} = H_{ii'} - \frac{H_i + H_{i'}}{2}, \text{ com } D_{ii} = 0 \quad \text{e} \quad D_{ii'} = D_{i'i}$$

A partir desses valores, foram estabelecidas as seguintes estatísticas:

FV	Identidade	Heterozigosidade	Diversidade
Dentro	\bar{J}_{IS}	$\bar{H}_{IS} = 1 - \bar{J}_{IS}$	$D_{IS} = 0$
Entre	\bar{J}_{ST}	$\bar{H}_{ST} = 1 - \bar{J}_{ST}$	$\bar{D}_{ST} = \bar{J}_{IS} - \bar{J}_{ST}$ $\bar{D}_{ST} = \bar{H}_{ST} - \bar{H}_{IS}$

Então,

$$G_{ST} = 100 \frac{\bar{D}_{ST}}{\bar{H}_{ST}}$$

Todo processo de simulação e análise dos dados fenotípicos e genotípicos foram realizados no Laboratório de Bioinformática da Universidade Federal de Viçosa, localizado no Instituto de Biotecnologia aplicada a Agropecuária (BIOAGRO), utilizando aplicativo computacional Genes (Cruz, 2013).

5. Resultados e Discussão

5.1. Análise com base nos dados fenotípicos

I. Análise dialélica

Pela análise de variância dialélica (Tabela 1) observou-se efeito significativo de tratamentos ($P \leq 0.01$) para todas as características simuladas. A capacidade geral de combinação (CGC) também foi significativa ($P \leq 0.01$) em todas as características avaliadas, indicando a existência de pelo menos um genitor superior aos demais, com relação ao comportamento médio das combinações híbridas.

Tabela 1 – Resumo da análise de variância dialélica para as características quantitativas simuladas (X1 a X12) das populações genitoras e suas combinações híbridas

V	GL	QM												
		X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	
Tratamentos	54	52410,27**	22820,35**	56669,55**	51250,10**	8543,27**	18122,33**	68497,24**	44303,53**	60248,79**	79244,54**	103934,12**	47440,78**	
C.G.C.	9	311961,39**	134430,81**	336730,36**	303863,97**	29090,15**	80685,00**	381506,72**	245129,99**	293086,86**	370859,13**	484097,43**	194078,60**	
C.E.C.	45	500,04 ^{ns}	498,25 ^{ns}	657,38 ^{ns}	727,33 ^{ns}	4433,90**	5609,80**	5895,34**	4138,24**	13681,18**	20921,62**	27901,46**	18113,22**	
Resíduo	1094	5	1933,59	955,97	770,48	580,56	2560,62	1320,36	785,23	708,24	4150,96	1775,74	1208,04	1058,37
Média Geral			91,28	98,72	97,18	91,43	130,41	117,03	123,34	103,08	138,64	131,32	143,27	136,64

* e **: significativos a 5 e 1% de probabilidade, pelo teste F

^{ns}: não significativo

Já a capacidade específica de combinação (CEC) foi significativa ($P \leq 0.01$) para as características simuladas com grau médio de dominância (gmd) iguais a 0,5 e 1, e não significativa para gmd igual a zero (Tabela 1). Estes resultados indicam que, de fato, os efeitos gênicos não-aditivos são determinantes no estabelecimento de diferenças entre o vigor manifestado nas combinações híbridas quando genes que exibem dominância estão envolvidos no controle da característica (Sibiya et al., 2010). Além disso, indicam que o processo de simulação, baseado no modelo biométrico adotado, foi suficiente e válido para gerar as informações genotípicas e fenotípicas necessárias para a inferência sobre importantes fenômenos genéticos.

Segundo Cruz et al. (2012), a CGC é função direta dos efeitos aditivos enquanto a CEC depende dos desvios de dominância, indicando, respectivamente, o potencial médio e diversidade dos genitores. Segundo Pinto et al. (2007) e Guimarães et al. (2009), conhecer e explorar os efeitos aditivos e não aditivos de genes no melhoramento de populações é crucial para a escolha dos métodos a serem empregados na condução do programa de melhoramento, pois possibilita alcançar maiores ganhos com a seleção. Assim, percebe-se que os resultados obtidos ratificam tais afirmativas pois contribuem para identificar e qualificar as populações e híbridos de melhor condição de utilização para fins de melhoramento.

Vários autores têm utilizado as estimativas de CGC e CEC na experimentação com plantas e animais. Bastos et al. (2003) obtiveram informações das capacidades geral e específica de combinação, a fim de determinar a ação gênica predominante em caracteres de importância econômica da cana-de-açúcar e identificar cruzamentos com potencial genético favorável para o desenvolvimento de novos cultivares. Silva et al. (2013), avaliando 14 linhagens de feijão, também utilizaram informações de suas capacidades geral e específica de combinação, a fim de obter populações segregantes promissoras que associassem alta

produtividade, boa arquitetura de plantas e grãos do tipo carioca. Rocha et al. (2014), avaliaram o potencial de quatro genótipos do feijoeiro, a fim de obter as estimativas de CGC e CEC e determinar quais seriam os mais promissores para uso em programas de melhoramento.

Capacidade geral de combinação

Na Tabela 2 são apresentados os valores das estimativas de capacidade geral de combinação (CGC) (\hat{g}_i) das dez populações genitoras, de acordo com as 12 variáveis quantitativas simuladas. A avaliação da CGC é importante, pois ela representa o comportamento médio de um genitor em suas combinações híbridas, e seu efeito é função da ação de genes de efeito predominantemente aditivos (Sprague e Tatum, 1942). Os efeitos aditivos são responsáveis pela proporção da variação fenotípica que pode ser transmitida aos descendentes (Ramalho et al., 2008). Considerando que se deseja selecionar as populações genitoras que proporcionem aumento da característica, observou-se que, de modo geral, as populações genitoras 4, 7, 9 e 10 se destacaram com as maiores estimativas de CGC para a maioria das características e que a população genitora 1 apresentou as menores estimativas de CGC em todas as características avaliadas (Tabela 2), não sendo, portanto, recomendada. De acordo com Cruz et al. (2012), quando se tem baixa estimativa de \hat{g}_i , ou seja valores próximos a zero, o valor da CGC do progenitor não difere muito da média geral dos cruzamentos dialélicos, entretanto, quando essas estimativas são altas, positivas ou negativas, o progenitor em questão é superior, ou inferior, aos demais, com relação ao comportamento médio dos cruzamentos.

Tabela 2 – Estimativa dos efeitos da capacidade geral de combinação (CGC) entre as populações genitoras para as características quantitativas simuladas (X1 a X12)

Populações Genitoras	Efeito da CGC											
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
1	-14,3018	3,1292	-20,2763	-13,9505	-0,6324	-2,6820	-4,3225	-9,8324	-14,1407	-14,2085	0,7544	-18,2034
2	10,3208	2,6008	-14,1992	12,0308	2,7664	1,2586	-1,9601	1,9563	6,9577	-13,5981	-20,5742	8,7258
3	-1,6826	-0,4777	17,1914	-3,7744	-5,9041	4,6698	17,7484	3,0286	5,4404	7,4313	-3,1692	3,6840
4	14,4421	1,8156	4,4603	23,9787	4,1381	-5,6288	9,2916	10,7878	5,4212	-11,6681	6,1912	3,9780
5	-15,6596	-6,1910	-3,0159	-10,3280	3,0941	-3,3278	-13,8426	-6,5050	-13,9923	-9,1799	-5,8327	4,0014
6	4,8476	-15,8124	12,1609	-5,9662	-3,4984	-7,4448	-11,7668	-3,5346	-18,1811	16,5097	-23,7730	-11,1336
7	5,0583	9,4343	4,8157	2,0136	-1,5716	-2,7348	-15,3028	1,8537	9,6491	17,9200	16,6030	6,6699
8	14,2595	-6,1477	-8,3314	-5,3818	-2,7571	11,7577	-3,0580	-11,5242	12,0176	-6,4950	3,4402	8,6748
9	-9,0732	5,1078	-1,8380	4,8530	0,2276	5,1919	2,5574	21,1560	1,9076	8,0796	5,7969	-0,8831
10	-8,2112	6,5410	9,0325	-3,4753	4,1373	-1,0598	20,6554	-7,3862	4,9207	5,2098	20,5633	-5,5136
DP (Gi)	0,8515	0,5987	0,5375	0,4666	0,9799	0,7037	0,5426	0,5154	1,2476	0,8160	0,6731	0,6300
DP (Gi - Gj)	1,2694	0,8925	0,8013	0,6956	1,4608	1,0490	0,8089	0,7682	1,8599	1,2165	1,0033	0,9391

Capacidade específica de combinação

Na Tabela 3 são apresentadas as estimativas da capacidade específica de combinação (CEC) para as populações genitoras (\hat{s}_{ii}) e combinações híbridas (\hat{s}_{ij}). A estimativa do efeito da CEC é útil, pois sua magnitude indica a dimensão da variabilidade que pode ser expressa em cada população, esperando-se maior probabilidade de obtenção de segregantes transgressivos quando os efeitos da CEC forem elevados (Barelli et al., 2000).

Segundo Cruz e Vencovsky (1989), os valores de \hat{s}_{ii} têm grande significado genético, tanto no seu sinal, quanto na sua magnitude, e seu somatório é função linear da heterose média. O sinal negativo dos \hat{s}_{ii} 's (Tabela 3) indica a existência de desvios da dominância unidirecionais e predominantemente positivos e sua magnitude é indicativo da divergência genética do progenitor i em relação à média dos outros progenitores considerados no dialelo – quanto maior for o valor de \hat{s}_{ii} , maior será o efeito da heterose varietal, que é manifestada em todos os seus híbridos.

Considerando que o somatório de \hat{s}_{ii} ($\sum \hat{s}_{ii}$) é função linear da heterose média, observa-se na Tabela 3 que as estimativas do $\sum \hat{s}_{ii}$ estão conforme o esperado, uma vez que as características X1 a X4 foram simuladas sob condição de grau médio da dominância (gmd) nulo; X5 a X8, gmd = 0.5 e X9 a X12, gmd igual a um.

Em relação às combinações híbridas, baixos valores de \hat{s}_{ij} indicam que os híbridos F1's, dentre os progenitores em questão, comportam-se como o esperado com base na capacidade geral de combinação (CGC) dos progenitores; enquanto altos valores de \hat{s}_{ij} demonstram que o comportamento de um cruzamento em particular é relativamente melhor do que o esperado com base na CGC de seus progenitores (Cruz et al., 2012).

Tabela 3 – Desempenho das populações genitoras e suas combinações híbridas, por meio de informações fenotípicas – Distância generalizada de Mahalanobis (D^2) e capacidade específica de combinação (CEC)

Cruzamentos	Efeito da CEC												D^2
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	
1x1	0.7785	0.3492	0.6873	0.5204	-4.3571	-7.1362	-10.3742	-3.0413	-9.7880	-15.6217	-12.2696	-26.1184	-
2x2	1.2881	0.8876	0.6458	1.9361	-6.8140	-2.4039	-10.6467	-3.9561	-14.7778	-6.1095	-17.3058	-12.3599	-
3x3	-1.4667	-1.2371	-1.0056	-1.5005	-8.5348	-12.7263	-8.6926	-9.2197	-12.0281	-19.1302	-20.9585	-7.4591	-
4x4	-1.3722	-2.1557	-1.7760	-1.0998	-6.3774	-6.8159	-7.8085	-6.2119	-9.8636	-17.8855	-18.1138	-14.8417	-
5x5	0.1402	0.6857	1.2115	0.7064	-7.0802	-5.4410	-6.3196	-4.8918	-10.5059	-9.3936	-16.5553	-17.4495	-
6x6	-0.8026	-0.3623	0.3212	-0.1854	-10.0498	-9.7167	-7.0698	-5.8760	-16.7139	-19.4672	-16.6098	-12.2931	-
7x7	-0.8031	-0.9760	-0.5174	-0.5606	-7.4437	-8.8108	-6.5410	-6.8153	-13.7490	-15.2990	-16.5691	-7.9596	-
8x8	-1.5655	-1.4955	-2.8212	-2.3364	-9.9350	-8.5741	-5.5428	-6.2943	-18.5256	-19.1217	-10.6367	-18.9513	-
9x9	0.2415	0.2671	0.1222	0.9262	-4.6835	-4.4907	-3.1901	-6.3682	-10.3677	-8.6323	-9.3308	-5.3422	-
10x10	0.7375	0.6516	-0.2655	0.1278	-4.7292	-7.0356	-9.9845	-7.2102	-11.8800	-14.2371	-23.7052	-14.3875	-
$\sum S_{ii}$	-2.8243	-3.3854	-3.3977	-1.4658	-70.0047	-73.1512	-76.1698	-59.8848	-128.1996	-144.8978	-162.0546	-137.1623	-
DP (S _{ii})	2.5674	1.8052	1.6206	1.4068	2.9545	2.1216	1.6361	1.5538	3.7617	2.4604	2.0293	1.8994	-
1x2	0.3272	-0.6710	-0.7298	0.3089	-0.8138	-2.2053	7.1815	0.9370	3.2344	-0.8882	2.2287	8.0315	13.5634
1x3	-0.1158	0.6067	1.0455	0.3035	4.4995	4.9430	3.2424	1.7426	7.9981	14.5901	-2.1871	6.2816	13.2397
1x4	-0.1370	-0.1682	-1.3598	-0.9891	1.7901	-1.4128	0.2663	-0.2119	0.3113	-8.9554	-6.3528	3.9923	16.9213
1x5	0.0174	0.5732	1.5779	2.4577	-0.6903	-0.2969	7.4786	4.9711	2.3341	-2.6074	12.4212	18.5698	4.9745
1x6	-1.3846	-1.0124	-0.2727	-2.2314	1.7115	4.4337	4.4726	-2.0273	-0.7411	14.7950	1.1666	6.5225	13.1955
1x7	-0.3621	1.0146	-0.4667	-0.3835	-1.6796	4.5361	-2.4879	-0.7870	-4.2067	3.3106	9.2890	5.3950	12.6694
1x8	0.5704	-0.6965	-1.2342	-1.6880	5.9346	4.7785	-0.6505	-1.3388	3.2597	12.3616	6.2565	5.3024	7.6149
1x9	0.2516	0.7351	0.5158	1.0076	0.9855	4.0341	0.6471	3.9515	9.6573	-0.5011	3.1619	5.7068	10.1269
1x10	-0.7242	-1.0799	-0.4504	0.1735	-3.0234	-4.5379	0.5983	-1.1544	-2.2712	-0.8618	-1.4448	-7.5653	9.0746
2x3	-1.3057	-1.1162	-2.1737	-2.0062	2.2776	-0.8454	2.1821	-0.0692	2.8673	-2.3280	-5.5305	-2.9214	11.3749
2x4	-0.2627	1.5175	2.2228	0.0088	5.4543	1.8906	4.7981	1.4540	10.9711	2.3644	15.2709	3.6928	5.4353
2x5	-0.6795	-1.6430	-1.9735	-0.8280	-3.0032	-1.0087	-2.4823	-1.6993	-5.1572	-2.1330	1.5680	4.8944	8.8052
2x6	-4.2603	-5.3510	-3.6207	-5.3217	-4.8646	-0.4116	-5.3746	-7.0048	6.2900	-4.070	-16.5282	-3.3467	13.0455
2x7	0.4319	1.0488	1.5296	0.6114	7.0470	5.8240	1.8882	5.0079	4.1619	7.1373	12.9345	2.9346	11.7371
2x8	0.5835	1.4606	1.3085	0.5957	-0.0225	1.7839	0.2855	2.1032	-1.3908	-1.0428	-5.8124	10.4901	8.1860
2x9	0.5862	0.4340	0.6153	1.0036	4.3883	-1.9108	4.6880	0.0228	4.3884	1.4192	6.2192	6.0772	7.4092

Continuação da Tabela 3

2x10	2.0030	2.5451	1.5297	1.7551	3.1651	1.6913	8.1270	7.1605	4.1903	8.0973	24.2614	-5.1327	15.5768
3x4	1.2456	0.1711	-0.3619	0.4558	-0.5334	5.3827	2.2042	1.9929	2.2116	8.9143	7.3377	-0.5858	9.5892
3x5	0.9507	-0.5289	1.0320	-0.6768	6.1485	9.9672	5.9802	6.8472	7.0002	10.1930	16.4215	8.1316	7.2707
3x6	-0.0339	1.6663	1.7499	1.8706	6.1861	2.6280	2.7301	2.5959	2.4555	-4.3793	-1.3515	-5.4375	7.2114
3x7	3.1072	1.6936	1.5615	1.3195	3.5741	6.9119	7.2234	5.2815	11.3253	9.2739	14.5865	8.7165	10.8902
3x8	-0.6743	0.4887	-0.6416	1.3339	-4.4383	-6.1771	-0.0910	-0.0802	0.2417	-11.2943	3.6173	6.2323	7.8964
3x9	0.3790	0.4506	-1.0089	-0.0423	-1.7260	-0.9223	-0.4423	-0.0867	-9.4198	0.4595	-1.3010	-5.7550	7.1594
3x10	-0.6193	-0.9577	0.8083	0.4431	1.0815	3.5648	-5.6439	0.2153	-0.6237	12.8312	10.3242	0.2560	3.5292
4x5	-1.4911	-0.7678	-1.8784	-3.9815	0.3335	-3.5294	-2.0373	-1.1502	1.5294	1.9148	4.7673	5.5852	11.9367
4x6	2.5160	1.7968	3.6045	1.7388	2.2378	2.9170	6.3553	5.5121	-0.2197	17.6688	8.9531	9.0078	14.6957
4x7	-1.1790	-0.5666	-1.8079	0.1348	0.0681	-2.5155	-0.9538	-3.4591	-3.9454	0.4119	-0.5659	-2.7816	10.8166
4x8	4.4150	4.3174	6.0697	5.9374	3.5718	9.6721	8.5436	9.2340	4.1513	9.4049	16.1939	3.2212	10.4122
4x9	-1.6222	-1.8208	-2.1559	-1.5673	-1.3060	-1.8223	-2.9783	-3.9019	0.6291	-0.1410	-1.8058	1.1907	6.5177
4x10	-0.7399	-0.1679	-0.7807	0.4618	1.1384	3.0494	-0.5812	2.9540	4.0883	4.1883	-7.5706	6.3606	11.1610
5x6	0.8111	2.1841	-0.2434	2.2840	1.4907	3.3823	-2.2520	-1.4422	-2.6998	2.7969	4.3633	5.4565	6.1669
5x7	-1.2071	-0.6812	-1.3762	0.1357	2.7584	-1.6365	-0.1907	1.5513	5.9484	5.6505	-5.9941	-5.8651	5.6920
5x8	-0.0921	-1.0355	-0.1625	0.1043	3.9401	4.8467	0.8559	-0.6345	9.2679	8.5918	-5.8229	-6.0698	4.6298
5x9	-0.0150	-0.3987	-1.2270	-2.0614	0.0194	-0.7379	-0.5256	-0.0659	-3.3462	-2.8470	2.2091	-4.1083	6.7466
5x10	1.4254	0.9264	1.8282	1.1531	3.1632	-0.1047	5.8124	1.4062	6.1351	-2.7724	3.1771	8.3048	8.1029
6x7	3.2782	1.6310	-0.5185	2.0722	3.3689	-0.0399	2.9034	6.6734	6.0826	6.6453	11.1184	2.5211	11.7025
6x8	1.2547	0.0434	-0.5178	1.7498	2.9688	1.7528	-1.2986	0.6622	6.4193	-5.6627	-0.4487	7.4071	11.2660
6x9	0.9518	1.3748	0.3171	0.5133	4.9474	3.3766	2.4619	3.6095	7.6417	9.4915	9.7608	0.4840	12.7322
6x10	-1.5279	-1.6083	-1.1408	-2.3046	2.0529	1.3943	4.1412	3.1731	8.1993	-2.0140	16.1858	1.9711	11.9493
7x8	-0.2778	-0.4025	1.7172	-1.1333	2.3775	-0.6387	-2.3856	-0.5125	4.4764	4.7988	-3.6680	-4.8955	6.7646
7x9	-0.6832	-0.9114	0.2462	-0.8967	-0.2535	3.6739	1.0263	0.4940	8.2600	-2.4492	-1.7670	0.8081	6.7780
7x10	-1.5018	-0.8743	0.1497	-0.7389	-2.3736	1.5061	6.0588	-0.6189	-4.6045	-4.1812	-2.7951	9.0860	9.6688
8x9	-1.5945	-0.7481	1.4845	-0.4187	1.7978	-1.5439	2.9364	5.2920	2.4521	9.8659	3.9348	3.5005	8.8360
8x10	-1.0537	-0.4365	-2.3812	-1.8083	3.7400	2.6736	2.8899	-2.1366	8.1735	11.2202	7.0230	12.7142	7.1836
9x10	1.2633	0.3499	0.9683	0.6094	0.5141	4.8343	-1.4333	3.4210	0.4727	1.9668	-1.7503	2.7803	8.6597
DP (Sij)	2.8641	2.0138	1.8079	1.5693	3.2959	2.3668	1.8252	1.7333	4.1964	2.7447	2.2638	2.1189	-
DP (Sii - Sij)	3.5903	2.5245	2.2663	1.9673	4.1316	2.9669	2.2880	2.1729	5.2605	3.4406	2.8378	2.6562	-
Singh (1981) (%)	5.3203	4.3734	14.2019	17.3987	0.6240	2.0288	15.2158	10.2796	2.0719	6.0572	13.5832	8.8453	-

As combinações híbridas de maiores estimativas de CEC foram as 2x7, 2x4 e 6x9, para a característica X5; 4x8, 3x7 e 2x7, para X6; 4x8, 2x10 e 3x7, para X7; 4x8, 2x10 e 6x7, para X8; 3x7, 2x4 e 1x9, para X9; 4x6, 3x10 e 8x10, para X10; 2x10, 4x8 e 6x10, para X11 e 8x10, 7x10 e 4x6, para X12 (Tabela 3). Vale ressaltar que essas combinações híbridas foram indicadas com base nas populações genitoras com as maiores estimativas de CGC (4, 7, 9 e 10) e que para os caracteres X1 a X4 ($gmd = 0$) não se indicou nenhuma combinação, pois a CEC foi não significativa.

II. Diversidade fenotípica das populações

Na Tabela 3 também são apresentadas as medidas de dissimilaridade genética entre os genitores, a partir da distância generalizada de Mahalanobis (D^2), e a importância relativa de cada característica para diferenciação das populações genitoras, de acordo com Singh (1981). Pela medida de distância generalizada de Mahalanobis (D^2), que é uma medida preditiva do desempenho da combinação híbrida a partir da diversidade dos genitores, destacam-se as combinações híbridas 1x4 ($\widehat{D}^2 = 16,92$), 2x10 ($\widehat{D}^2 = 15,58$), 4x6 ($\widehat{D}^2 = 14,70$), 1x2 ($\widehat{D}^2 = 13,56$) e 1x3 ($\widehat{D}^2 = 13,24$) entre as mais divergentes. Entretanto, apenas as combinações 2x10 e 4x6 foram divergentes e complementares, com elevadas estimativas de \hat{s}_{ij} para a maioria das características (Tabela 3). Tal resultado confirma a importância de se levar em consideração não apenas a diversidade entre os genitores, mas a capacidade geral de combinação (CGC) dos genitores que participam do cruzamento, ao se avaliar qual cruzamento é considerado favorável para o aumento da característica. Assim, se considerarmos as populações genitoras 4, 7, 9 e 10, de maiores estimativas de CGC, as combinações híbridas 2x10 ($D^2 = 15,58$), 4x6 ($D^2 = 14,70$), 6x9 ($D^2 = 12,73$), 1x7 ($D^2 = 12,67$) e 6x10 ($D^2 = 11,95$) seriam as de maior desempenho pela medida de distância

generalizada de Mahalanobis (D^2). Observa-se que quatro, das cinco combinações híbridas indicadas pelas estimativas de D^2 , estão entre as indicadas pela capacidade específica de combinação (CEC).

Avaliou-se ainda a importância relativa das características na análise da diversidade genética entre as populações genitoras, de acordo com a metodologia proposta por Singh (1981) (Tabela 3). Vários autores recomendam essa metodologia, que é uma técnica multivariada baseada na distância generalizada de Mahalanobis, para tal finalidade (Silva et al., 2009; Coimbra et al., 2010; Cardoso et al., 2015; Domiciniano et al., 2015). Entre as 12 características simuladas, as características que mais contribuíram para a dissimilaridade entre as populações genitoras foram aquelas de maior herdabilidade ($60\% \leq h^2 \leq 80\%$), independente do grau médio de dominância (gmd) dos locos (Tabela 3). A característica que menos contribuiu para a diferenciação das populações genitoras foi a X5, com apenas 0,6% de contribuição. No Quadro 1 estão apresentados os locos controladores das características em estudo e pode-se observar a correlação existente entre a característica X5 e os outros caracteres avaliados, tornando-a redundante e, dessa forma, dispensável ao estudo da diversidade genética (Cruz et al., 2012).

5.2. Abordagem de genética de populações na escolha de progenitores

Na Tabela 4 são apresentadas as estimativas das médias fenotípicas, capacidades geral e específica de combinação, baseadas em informações biométricas, além das estimativas de heterozigosidade e coeficiente de endogamia, baseadas em informações moleculares. Observa-se que a média fenotípica das características X1 e X4, de efeito aditivo (gmd = 0), é diretamente proporcional à capacidade geral de combinação (CGC) das populações genitoras, independente da herdabilidade da característica.

Tabela 4 - Estimativas relacionadas à abordagem quantitativa – médias, capacidade geral de combinação (CGC) e capacidade específica de combinação (CEC) e à abordagem molecular – heterozigiosidade (H) e coeficiente de endogamia (F) das populações genitoras para as características quantitativas simuladas X1, X4, X9 e X12

Populações Genitoras	Médias				CGC				CEC				H	F
	X1	X4	X9	X12	X1	X4	X9	X12	X1	X4	X9	X12		
1	63.4499	64.0446	100.5697	74.1168	-14.3018	-13.9505	-14.1407	-18.2034	0.7785	0.5204	-9.7880	-26.1184	0.3072	0.0048
2	113.2049	117.4228	137.7769	141.7340	10.3208	12.0308	6.9577	8.7258	1.2881	1.9361	-14.7778	-12.3599	0.3279	0.0479
3	86.4430	82.3758	137.4918	136.5513	-1.6826	-3.7744	5.4404	3.6840	-1.4667	-1.5005	-12.0281	-7.4591	0.3498	0.0217
4	118.7869	138.2826	139.6179	129.7566	14.4421	23.9787	5.4212	3.9780	-1.3722	-1.0998	-9.8636	-14.8417	0.3435	0.0587
5	60.0960	71.4756	100.1486	127.1957	-15.6596	-10.3280	-13.9923	4.0014	0.1402	0.7064	-10.5059	-17.4495	0.2942	0.0179
6	100.1676	79.3073	85.5630	102.0820	4.8476	-5.9662	-18.1811	-11.1336	-0.8026	-0.1854	-16.7139	-12.2931	0.3084	-0.0066
7	100.5885	94.8917	144.1883	142.0226	5.0583	2.0136	9.6491	6.6699	-0.8031	-0.5606	-13.7490	-7.9596	0.3602	-0.0425
8	118.2285	78.3250	144.1487	135.0407	14.2595	-5.3818	12.0176	8.6748	-1.5655	-2.3364	-18.5256	-18.9513	0.3723	-0.0291
9	73.3701	102.0575	132.0867	129.5340	-9.0732	4.8530	1.9076	-0.8831	0.2415	0.9262	-10.3677	-5.3422	0.3730	-0.0352
10	75.5902	84.6023	136.6005	111.2276	-8.2112	-3.4753	4.9207	-5.5136	0.7375	0.1278	-11.8800	-14.3875	0.2725	0.0556

As populações genitoras 4, 8 e 2 são as que se destacam para a característica X1 e 4, 2 e 9, para a X4, com base tanto em suas médias fenotípicas, quanto em suas capacidades gerais de combinação. Assim, para características de efeito aditivo, nota-se que a escolha de populações genitoras com maior frequência de alelos favoráveis pode ser feita sem a necessidade de realização dos cruzamentos dialélicos, podendo-se escolher apenas com base em suas médias fenotípicas.

Para as características X9 e X12, simuladas sob condição de dominância completa ($gmd = 1$), podemos observar que essa relação não ocorre (Tabela 4). Para a característica X9, as populações genitoras recomendadas, com base nas maiores médias fenotípicas, seriam as populações 7, 8 e 4, com médias iguais a 144,19; 144,15 e 139,62, respectivamente. Tal recomendação não seria coincidente, se for baseada nas estimativas de CGC, devendo-se recomendar as populações genitoras 8, 7 e 2, com estimativas de CGC iguais a 12,02; 9,65 e 6,96, respectivamente. Nota-se que a população genitora 7, que apresenta a maior média fenotípica (144,19), não apresenta a maior estimativa de CGC (9,65). O mesmo acontece para a característica X12 e, nesse caso, as populações genitoras recomendadas, com base nas maiores médias fenotípicas, seriam as populações 7, 2 e 3, entretanto, se basearmos nas maiores estimativas de CGC, as populações recomendadas seriam as 2, 8 e 7. De maneira geral, baseando-se apenas em informações fenotípicas, as populações genitoras com maior destaque foram a 2, 4, 8 e 7.

Entretanto, apenas o desempenho das populações genitoras não é indicativo suficiente do potencial das combinações híbridas quanto à segregação transgressiva. Considerando as estimativas de heterozigosidade (Tabela 4), as populações genitoras potenciais são as 9, 8, 7 e 4. Assim, nota-se que há coincidência de três das quatro populações recomendadas com base nas informações fenotípicas.

Pela Tabela 5 podemos observar o desempenho e a divergência das combinações híbridas, baseadas em informações fenotípicas e genotípicas, respectivamente. As populações genitoras 1, 5 e 6 foram as que apresentaram as menores médias fenotípicas, com a maioria das combinações híbridas com média inferior à média geral. Considerando a diversidade entre as populações pela abordagem de genética de populações, por meio das medidas de distância Euclidiana, genotípica de Hedrick e índice G_{ST} de Nei (1973), baseadas em dados moleculares, as combinações híbridas que mais se destacaram foram: 2x10, 2x7, 3x5, 3x10, 6x10, 8x9 e 8x10. Entretanto, muitas dessas combinações híbridas não apresentaram bom desempenho, destacando as populações 2, 4, 7 e 8 como as de melhor desempenho. Assim, considerando o desempenho médio das populações associado à diversidade de suas combinações, as populações 8x10, 2x10 e 2x7 foram as que mais se destacaram.

Comparando estes resultados com os obtidos na análise dialélica, percebe-se que o desempenho médio das populações genitoras associado com a diversidade molecular refletem suas capacidades geral e específica de combinação, possibilitando escolher as populações com maior potencial para a obtenção de populações segregantes promissoras para a extração de linhagens superiores em programas de melhoramento.

Em geral, o melhorista dispõe de um elevado número de populações ou linhagens elites em um programa de melhoramento e muitas vezes se depara com a limitação do uso de cruzamentos dialélicos em virtude da quantidade de cruzamentos necessária para avaliar determinado grupo de genitores, como também dificuldade na obtenção de sementes para a avaliação das combinações híbridas ou impossibilidade de avaliar todas as combinações híbridas possíveis (Freitas Junior et al., 2006; Silva et al, 2013; Rocha et al., 2014). Para contornar tal situação, o melhorista pode optar por utilizar um número reduzido de genitores em esquema de dialelo parcial, por exemplo (Carvalho et al., 2014).

Tabela 5 – Estimativas relacionadas à abordagem quantitativa – médias e distância generalizada de Mahalanobis (D^2) e à abordagem molecular – distância euclidiana ($D_{\text{Euclidiana}}$), distância genotípica de Hedrick (D_{Hedrick}) e Índice G_{ST} de Nei (G_{ST}) das populações genitoras para as características quantitativas simuladas X1, X4, X9 e X12

Combinções Híbridas	Médias				D^2	$D_{\text{Euclidiana}}$	D_{Hedrick}	G_{ST}
	X1	X4	X9	X12				
1x2	87.62117	89.81431	134.6907	135.1961	13.5634	0.5853	0.3397	0.2078
1x3	75.17473	74.00381	137.937	128.4045	13.2397	0.6022	0.3699	0.2139
1x4	91.27827	100.4642	130.231	126.4092	16.9213	0.5841	0.3388	0.2021
1x5	61.33097	69.60436	112.8403	141.0101	4.9745	0.6248	0.3643	0.2430
1x6	80.43618	69.27695	105.5762	113.8278	13.1955	0.6451	0.3752	0.2528
1x7	81.6694	79.10462	129.9408	130.5038	12.6694	0.5282	0.3237	0.1757
1x8	91.80314	70.40477	139.7758	132.4161	7.6149	0.5982	0.3573	0.2106
1x9	68.15157	83.33537	136.0634	123.2626	10.1269	0.5754	0.3591	0.1984
1x10	68.03781	74.17282	127.1478	105.3598	9.0746	0.5398	0.2891	0.1961
2x3	98.60747	97.67526	153.9047	146.1307	11.3749	0.5500	0.3307	0.1773
2x4	115.7752	127.4434	161.9893	153.039	5.4353	0.5432	0.3130	0.1722
2x5	85.25667	92.29986	126.4473	154.264	8.8052	0.5479	0.3133	0.1890
2x6	102.183	92.16793	133.7058	130.8878	13.0455	0.5841	0.3517	0.2077
2x7	107.0861	106.081	159.408	154.9728	11.7371	0.6232	0.4010	0.2196
2x8	116.4389	98.66988	156.2237	164.5331	8.1860	0.5426	0.3281	0.1725
2x9	93.10884	109.3126	151.893	150.5623	7.4092	0.5433	0.3253	0.1731
2x10	95.38775	101.7357	154.708	134.7218	15.5768	0.6550	0.4012	0.2531
3x4	105.28	112.0853	151.7124	143.7186	9.5892	0.5860	0.3396	0.1920
3x5	74.88356	76.64589	137.0874	152.4595	7.2707	0.6503	0.4002	0.2434
3x6	94.40601	83.55516	128.354	123.7552	7.2114	0.5552	0.3240	0.1884
3x7	97.75794	90.98388	165.054	155.7129	10.8902	0.5939	0.3527	0.2005
3x8	103.1776	83.60286	156.3389	155.2336	7.8964	0.5117	0.3021	0.1540
3x9	80.89816	92.4614	136.5673	133.6884	7.1594	0.5283	0.3039	0.1628
3x10	80.76187	84.61855	148.3765	135.0689	3.5292	0.6328	0.3695	0.2366

Continuação da Tabela 5

4x5	88.56631	101.0943	131.5975	150.2071	11.9367	0.5702	0.3427	0.1966
4x6	113.0807	111.1765	125.6594	138.4946	14.6957	0.5862	0.3493	0.2038
4x7	109.5964	117.5522	149.764	144.5087	10.8166	0.5520	0.3443	0.1766
4x8	124.3917	115.9594	160.2293	152.5164	10.4122	0.5785	0.3396	0.1872
4x9	95.0216	118.6895	146.5971	140.9281	6.5177	0.4766	0.2685	0.1354
4x10	96.76598	112.3903	153.0694	141.4674	11.1610	0.5616	0.3171	0.1944
5x6	81.27419	77.41497	103.7659	134.9667	6.1669	0.5556	0.3123	0.2030
5x7	79.46656	83.24642	140.2444	141.4486	5.6920	0.5913	0.3452	0.2132
5x8	89.78283	75.81968	145.9324	143.2488	4.6298	0.5614	0.3223	0.1924
5x9	66.52712	83.88866	123.2082	135.6524	6.7466	0.5549	0.3379	0.1892
5x10	68.82969	78.7749	135.7026	143.435	8.1029	0.6109	0.3573	0.2408
6x7	104.4591	89.54475	136.1898	134.6998	11.7025	0.6064	0.3675	0.2200
6x8	111.6369	81.82694	138.895	141.5907	11.2660	0.5843	0.3568	0.2035
6x9	88.00126	90.8253	130.0074	125.1097	12.7322	0.6098	0.3709	0.2181
6x10	86.38347	79.67892	133.5781	121.9663	11.9493	0.6542	0.3857	0.2645
7x8	110.315	86.92355	164.7823	147.0916	6.7646	0.5199	0.3065	0.1604
7x9	86.5769	97.39504	158.456	143.2374	6.7780	0.5255	0.3057	0.1636
7x10	86.62029	89.22444	148.6044	146.8847	9.6688	0.5550	0.3273	0.1954
8x9	94.86678	90.47759	155.0165	147.9347	8.8360	0.6049	0.3580	0.2021
8x10	96.26957	80.75965	163.751	152.5178	7.1836	0.6642	0.4207	0.2533
9x10	75.25402	93.41224	145.9402	133.0261	8.6597	0.5647	0.3422	0.1973
Média Geral	91.28	91.43	138.64	136.64	-	-	-	-

Desta forma, a ferramenta de predição proposta pela utilização conjunta das informações biométricas e moleculares torna-se fundamental na escolha dos genitores potenciais à formação de híbridos superiores.

A análise dialélica também apresenta limitações na escolha de genitores na formação de populações segregantes promissoras quando no controle das características predomina ausência de dominância. Também neste caso, esta limitação pode ser superada pela

associação do desempenho das populações segregantes com a diversidade das populações pela abordagem de genética de populações.

6. Referências Bibliográficas

Baldissera, J.N.C.; Valentini, G.; Coan, M.M.D.; Almeida, C.B.; Guidolin, A.F.; Coimbra, J.L.M. Capacidade combinatória e efeito recíproco em características agronômicas do feijão. **Semina: Ciências Agrárias**, v.33, n.2, p.471-480, 2012.

Barelli, M.A.A.; Vidigal, M.C.G.; Amaral Júnior, A.T.; Vidigal Filho, P.S.; Scapim, C.A.; Sagrilo, E. Diallel analysis for grain yield and yield components in *Phaseolus vulgaris* L. **Acta Scientiarum Agronomy**, v.22, p.883-887, 2000.

Bastos, I.T.; Barbosa, M.H.P.; Cruz, C.D.; Burnquist, W.L.; Bressiani, J.A.; Silva, F.D. Análise dialélica em clones de cana-de-açúcar. **Bragantia**, v.62, n.2, p.199-206, 2003.

Cardoso, D.L.; Silva, R.F.; Pereira, M.G.; Viana, A.P.; Araújo, E.F. Diversidade genética e parâmetros genéticos relacionados à qualidade fisiológica de sementes em germoplasma de mamoeiro. **Ceres**, v.56, n.5, 2015.

Carvalho, A.D.F.; Silva, G. O.; Pereira, R.B.; Pinheiro, J. B.; Vieira, J.V. Análise da capacidade de combinação em cruzamentos dialélicos em linhagens de cenoura. **Horticultura Brasileira**, v.32, n.2, 2014.

Coimbra, R.R.; Miranda, G.V.; Cruz, C.D.; Melo, A.V.; Eckert, F.R. Caracterização e divergência genética de populações de milho resgatadas do Sudeste de Minas Gerais. **Revista Ciência Agronômica**, v.41, n.1, p.159–166, 2010.

Cruz, C.D.; Vencovsky, R. Comparação de alguns métodos de análise dialélica. **Revista Brasileira de Genética**, v.12, n.2, p.425-438, 1989.

Cruz, C.D. GENES: a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum Agronomy**, v.35, p.271-276, 2013.

Cruz, C.D.; Regazzi, A.J.; Carneiro, P.C.S. **Modelos Biométricos Aplicados ao Melhoramento Genético**. V.1, 4ed, Viçosa, MG: Ed. UFV, 2012.

Domiciano, G.P.; Alves, A.A.; Laviola, B.G.; Conceição, L.D.H.C.S. Parâmetros genéticos e diversidade em progênies de Macaúba com base em características morfológicas e fisiológicas. **Ciência Rural**, v.45, n.9, 2015.

Engelsing, M.J., Rozzetto, D.S., Coimbra, J.L.M., Zanin, C.G., Guidolin, A.F. Capacidade de combinação em milho para resistência a *Cercospora zea-maydis*. **Revista Ciência Agronômica**, v.42, p.232-241, 2011.

Freitas Júnior, S.P.; Amaral Júnior, A.T.; Pereira, M.G.; Cruz, C.D.; Scapim, C.A. Capacidade combinatória de milho pipoca por meio de dialelo circulante. **Pesquisa Agropecuária Brasileira**, v.41, n.11, p.1599-1607, 2006.

Griffing, B. Concept of general and specific ability in relation to diallel crossing systems. **Australian Journal of Biological Sciences**, v.9, n.4, p.462-93, 1956.

Guimarães, P.S.; Paterniani, M.E.A.G.Z.; Dudienas, C.; Lüders, R.R.; Gallo, P.B. Capacidade combinatória para resistência à mancha branca em linhagens endogâmicas de milho. **Summa phytopathol**, v.35, p.282-287, 2009.

Pimentel, A.J.B.; Ribeiro, G.; Souza, M.A.; Moura, L.M.; Assis, J.C.; Machado, J.C. Comparação de métodos de seleção de genitores e populações segregantes aplicados ao melhoramento de trigo. **Bragantia**, v.72, n.2, p.113-121, 2013.

Pinto, R.J.B.; Scapim, C.A.; Barreto, R.R.; Rodovalho, M.A.; Esteves, N.; Lopes, A.D. Análise dialélica de linhagens de milho-pipoca. **Ceres**, v.54, n.315, p.471-477, 2007.

Ramalho, M.A.P.; Santos, J.B.; Pinto, C.A.B.P. **Genética na Agropecuária**. 4.ed. Lavras: UFLA, p.460, 2008.

Rocha, F.; Stinghen, J.C.; Gemeli, M.S.; Coimbra, J.L.M.; Guidolin, A.F. Análise dialéctica como ferramenta na seleção de genitores em feijão. **Revista Ciência Agronômica**, v.45, n.1, p.74-81, 2014.

Sibiya, J.; Tongoona, P.; Derera, J; van Rij, N; Makanda, I. Combining ability analysis for Phaeosphaeria leaf spot resistance and grain yield in tropical advanced maize inbred lines. **Field Crops Research**, v.120, n.1, p.86-93, 2011.

Silva, G.O.; Vieira, J.V.; Vilela, M.S. Seleção de caracteres de cenoura cultivada em dois sistemas de produção agroecológicos no Distrito Federal. **Ceres**, v.56, n.5, 2009.

Silva, V.M.P.; Carneiro, P.C.S.; Menezes Júnior, J.A.N.; Carneiro, V.Q.; Carneiro, J.E.S., Cruz, C.D.; Borém, A. Genetic potential of common bean parents for plant architecture improvement. **Scientia Agricola**, v.70, n.3, 167-175, 2013.

Singh, D. The relative importance of characters affecting genetic divergence. **The Indian Journal of Genetics and Plant Breeding**, v.41, p.237-245, 1981.

Sprague, G.F. e Tatum, L.A. General vs. specific combining ability in single crosses of corn. **J. am. soc. agron**, v.34, n.10, p.923-932, 1942.

Viegas, M.P.; Silva, C.L.S.P.; Moreira, J.P.; Cardin, L.T.; Azevedo, V.C.R.; Ciampi, A.Y.; Freitas, M.L.M.; Moraes, M.L.T.; Sebbenn, A.M. Diversidade genética e tamanho efetivo de duas populações de *Myracrodruon urundeuva* fr. all., sob conservação ex situ. **Revista Árvore**, v.35, p.769–779, 2011.

4. CONCLUSÕES GERAIS

Os parâmetros que descrevem a estrutura populacional, como equilíbrio de Hardy-Weinberg, coeficiente de endogamia e heterozigosidade são ferramentas úteis na avaliação da diversidade genética da população.

É possível prever a diversidade genética de combinações híbridas a partir da diversidade genética de suas populações genitoras.

A diversidade molecular das populações genitoras, por abordagem de genética de populações, associada aos seus desempenhos médios, refletem suas capacidades geral e específica de combinação, possibilitando economia de recursos em um programa de melhoramento genético.

A abordagem de genética de populações associada às informações quantitativas aprimoram a escolha de populações genitoras pela análise dialética.