

LUCIANO ROGÉRIO BRAATZ DE ANDRADE

**VULNERABILIDADE GENÉTICA E RELAÇÃO DOS GERMOPLASMAS
COMERCIAIS DE MILHO NO BRASIL ENTRE SI E COM OS GENITORES
DA POPULAÇÃO NESTED ASSOCIATION MAPPING**

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

VIÇOSA
MINAS GERAIS – BRASIL
2015

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

T

A555v
2015

Andrade, Luciano Rogério Braatz de, 1989-
Vulnerabilidade genética e relação dos germoplasmas
comerciais de milho no Brasil entre si e com os genitores da
população Nested Association Mapping / Luciano Rogério
Braatz de Andrade. – Viçosa, MG, 2015.
vi, 28f. : il. (algumas color.) ; 29 cm.

Orientador: Aluízio Borém de Oliveira.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Referências bibliográficas: f.25-28.

1. Milho. 2. Germoplasma vegetal. 3. Melhoramento
genético. 4. Segurança alimentar. I. Universidade Federal de
Viçosa. Departamento de Fitotecnia. Programa de Pós-graduação
em Genética e Melhoramento. II. Título.

CDD 22. ed. 633.15

LUCIANO ROGÉRIO BRAATZ DE ANDRADE

**VULNERABILIDADE GENÉTICA E RELAÇÃO DOS GERMOPLASMAS
COMERCIAIS DE MILHO NO BRASIL ENTRE SI E COM OS GENITORES
DA POPULAÇÃO NESTED ASSOCIATION MAPPING**

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

APROVADA: 20 de fevereiro de 2015.

Eveline Teixeira Caixeta

Carlos Nick Gomes

Aluízio Borém de Oliveira
(Orientador)

Roberto Fritsche Neto
(Coorientador)

Aos meus pais Ulisses Rogério e Eliane

À Massaine Bandeira

À minha família e amigos

Dedico.

AGRADECIMENTOS

Agradeço a Deus por conceder a honra de trabalhar e conviver com os amigos a quem devo gratidão. Por sua escrita certa, contudo em linhas tortas.

Agradeço aos meus pais Eliane Braatz Arruda de Andrade e Ulisses Rogério Arruda de Andrade, pelo empenho, dedicação e auxílio investidos em mim durante minha jornada.

À minha namorada Massáine Bandeira pela motivação, carinho e apoio.

À minha família por momentos felizes e tristes, de conquistas convividos e jamais esquecidos.

Aos Professores Aluizio Borém e Roberto Fritsche Neto, pelo apoio, dedicação, incentivo e inspiração a mim dedicados.

Ao Professor Cosme Damião Cruz, pela amizade, apoio, conhecimentos e por me acolher em seu laboratório.

À Universidade Federal de Viçosa, e ao Programa de Pós-Graduação em Genética e Melhoramento, pelo excelente ensino e oportunidade doados.

À Universidade de São Paulo, em especial ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas da Escola Superior de Agricultura “Luiz de Queiroz”, por acolher e investir na pesquisa divulgada nesta dissertação.

Ao Miguel Ribeiro de Camargo e demais funcionários da USP/ESALQ e UFV, por ajudar na condução dos experimentos.

Aos amigos e colegas de trabalho do Laboratório de Melhoramento de Alógamas da USP/ESALQ, pelos momentos felizes e grandes conquistas.

Aos amigos do Laboratório de Bioinformática da UFV pelo apoio e amizade.

Ao Laboratório BioCafé e ao Programa Milho da UFV pelo auxílio no presente trabalho.

Aos meus primeiros pais científicos Altamir Frederico Guidolin, Jefferson Luís Meirelles Coimbra e Alfredo do Nascimento Junior, pela motivação e incentivo no melhoramento de plantas.

À Capes pela concessão da bolsa de estudos.

RESUMO

ANDRADE, Luciano Rogério Braatz de, M.Sc., Universidade Federal de Viçosa, fevereiro de 2015. **Vulnerabilidade genética e relação dos germoplasmas comerciais de milho no Brasil entre si e com os genitores da população Nested Association Mapping.** Orientador: Aluizio Borém de Oliveira. Coorientador: Roberto Fritsche Neto.

O germoplasma comercial de milho é oriundo de apenas sete linhagens endogâmicas, o que pode levar a vulnerabilidade genética deste a estresses abióticos e bióticos, e limitar os ganhos com a seleção no Brasil. A vulnerabilidade genética ocorre em culturas melhoradas de baixa variabilidade genética, presentes em grandes áreas. Por exemplo, a infestação de helmintosporiose em 1970 causou preocupação entre os melhoristas Norte-Americanos com a diversidade genética de seus híbridos. No Brasil, o mercado de sementes é dominado por poucas empresas: Monsanto® (35%), DuPont Pioneer® (30%), Dow Agrosiences® (15%), Syngenta® (10%) e Helix Sementes (4%). Assim é importante o monitoramento da diversidade genética que está sendo utilizada, pois as práticas de melhoramento, registro e marketing de novos cultivares podem conduzir a vulnerabilidade genética. Diante do exposto, o objetivo foi avaliar o risco de vulnerabilidade genética e as relações entre os germoplasmas comerciais brasileiros de milho e destes com os genitores da população Nested Association Mapping (NAM). Visando atender a esses objetivos foram utilizados os genótipos dos híbridos elite das empresas de maiores market-share no Brasil e os genitores da população NAM. Os híbridos foram genotipados para 700 marcadores do tipo SNP, por meio da plataforma Illumina GoldenGate®. Já os dados genômicos dos genitores da NAM foram obtidos no banco de dados do Panzea, os quais foram genotipados para 1536 SNP's. Os dados dos marcadores para os genótipos foram submetidos às análises de correlação entre as frequências alélicas, estruturação populacional e diversidade genotípica. Pelo tamanho efetivo populacional evidenciou-se baixo risco de vulnerabilidade genética para o germoplasma comercial brasileiro. A distância genética entre o germoplasma comercial brasileiro é menor que a encontrada na NAM, sendo menor ainda entre os híbridos de mesma empresa. O germoplasma comercial brasileiro não apresentou relações estreitas com as linhagens genitoras da população NAM, com exceção da linhagem Norte-Americana B73, esta que pertenceu a mesma população pela estruturação populacional. Isto evidencia que a linhagem B73 ou seu grupo heterótico (Iowa Stiff Stalk Synthetic) tiveram grande influência na formação do germoplasma comercial brasileiro.

ABSTRACT

ANDRADE, Luciano Rogério Braatz de, M.Sc., Universidade Federal de Viçosa, February, 2015. **Genetic vulnerability and relationship between commercial germplasms of maize in Brazil and progenitors of nested association mapping population.** Adviser: Aluizio Borém de Oliveira. Co-Adviser: Roberto Fritsche Neto.

The limited genetic variability of the Brazilian commercial maize germplasm, originated from only seven inbred lines, could result in genetic vulnerability to abiotic and biotic stress, and also to limited selection gain. The genetic vulnerability happens to situations with low genetic variability, in large areas. The infection of Southern corn leaf blight in 70's, concerned the North-American breeders about their hybrids genetic diversity. In Brazil, Monsanto[®], DuPont Pioneer[®], Dow Agrosiences[®] and Syngenta[®] have the largest seed market share. Therefore, it's important to monitor the genetic diversity available in the country, because breeding practices, registration and marketing of new varieties could result in genetic vulnerability. Thus, the objectives of this study were to evaluate the risk of genetic vulnerability and the relationships between the Brazilian commercial maize germplasm and the Nested Association Mapping (NAM) genitors. Aiming to achieve these objectives, it was include in this work the elite hybrids of the largest market share companies in Brazil and the NAM genitors. The hybrids were genotyped by Illumina GoldenGate[®] platform, while the genomic data of NAM genitors were obtained in Panzea. The genotypes were subjected to linear correlation analysis, population structure and genetic diversity analysis. It was shown lower risk of genetic vulnerability with the Brazilian commercial germplasm. There is a narrow relationship within the Brazilian commercial germplasm, being closer between the hybrids of same company. The Brazilian commercial germplasm did not show close relationship with the NAM genitors, with exception of North-American line B73, which belonged to the same population of the population structure analysis. This showed that the line B73 or its heterotic group (Iowa Stiff Stalk Synthetic) had great influence at the Brazilian commercial germplasm formation.

ÍNDICE

1	INTRODUÇÃO	1
2	OBJETIVOS.....	3
3	MATERIAL E MÉTODOS.....	4
3.1	MATERIAL GENÉTICO	4
3.2	GENOTIPAGEM E CONTROLE DE QUALIDADE DOS DADOS GENÉTICOS.....	5
3.3	ANÁLISES DOS DADOS GENÔMICOS	6
4	RESULTADOS E DISCUSSÕES.....	10
4.1	ANÁLISE DE DIVERSIDADE GENÉTICA VIA MARCADORES SNPS	10
4.2	VULNERABILIDADE GENÉTICA DOS HÍBRIDOS BRASILEIROS DE MILHO.....	15
4.3	RELAÇÃO ENTRE O GERMOPLASMA DOS HÍBRIDOS COMERCIAIS BRASILEIROS E OS GENITORES DA POPULAÇÃO NESTED ASSOCIATION MAPPING	21
5	CONCLUSÕES	24
6	REFERÊNCIAS BIBLIOGRÁFICAS	25

1 INTRODUÇÃO

A cultura do milho mudou drasticamente com o início do uso de híbridos de milho, em 1933 nos Estados Unidos, 1950 na Europa (Le Clerc et al., 2005) e 1944 no Brasil (Galvão et al., 2014). Segundo Duvick (2001), o uso de cultivares híbridos é vantajoso devido a sua elevada produtividade, estabilidade e homogeneidade. A produtividade beneficiou economicamente os agricultores. A estabilidade elevou a tolerância a estresses abióticos e bióticos. Já a homogeneidade favoreceu o processo de mecanização da lavoura, o que gerou um aumento no tamanho das áreas cultivadas e a redução no número de propriedades. Entretanto, o germoplasma comercial de milho possui base genética estreita, sendo que, os principais cultivares híbridos do mundo são oriundos a partir de cruzamentos envolvendo apenas sete linhagens endogâmicas: B73, Mo17, LH82, LH123, PH207, PH59 e PHG39 (Mikel e Dudley, 2006).

Segundo Romay et al. (2011), a limitada variabilidade genética do atual germoplasma comercial de milho pode levar a vulnerabilidade das lavouras a estresses abióticos e bióticos, e assim gerar insegurança alimentar e limitar os ganhos com a seleção. A vulnerabilidade ocorre principalmente quando a cultura possui baixa variabilidade genética, e está presente em grandes áreas e em sistemas de monocultura, em geral de um único ou poucos cultivares (Brown, 2008). Como exemplo da vulnerabilidade genética em milho, Williams e Hallauer (2000) relatam que os melhoristas americanos somente preocuparam-se com a diversidade genética entre seus híbridos após a infestação de helmintosporiose (*Exserohilum turcicum*) em 1970. A suscetibilidade esta ligada ao citoplasma T, responsável pela macho-esterilidade utilizada na produção de sementes híbridas (Levings, 1990).

Atualmente, o Brasil utiliza sementes certificadas em 90% das áreas cultivadas com milho (ABRASEM, 2013). Destas, a maioria dos cultivares disponíveis são híbridos (88,32%) (Cruz et al., 2013), sendo este mercado dominado pelas empresas Monsanto[®], 35% do Market-share, seguida pela Dupont Pioneer[®] (30%), Dow Agrosiences[®] (15%), Syngenta[®] (10%), Helix Sementes[®] (4%), EMBRAPA (3%) e demais empresas (3%). Neste contexto, Segundo Reif et al. (2005), o monitoramento da diversidade genética disponível é importante, pois as práticas de melhoramento de plantas, o registro e o marketing dos novos cultivares podem conduzir a vulnerabilidade genética. Isto devido as práticas de melhoramento visar a homogeneidade, a estabilidade dos cultivares. O registro de poucos cultivares devido a fatores econômicos ligados ao custo na produção

de sementes. Já o marketing influencia na escolha dos agricultores sobre o cultivar a ser utilizado em suas lavouras.

A diversidade genética pode ser estimada por meios de informações fenotípicas e genotípicas, dentre as análises utilizadas destacam-se: Componentes principais (Technow et al., 2014), Agrupamento (Liu et al., 2010) e estruturação populacional (Wu et al., 2014). O uso de marcadores SNPs (Single Nucleotide Polymorphism) vêm crescendo devido à grande quantidade de marcadores e plataformas de genotipagem disponíveis, os quais possibilitam estudos de análise de diversidade genética e estruturação populacional (Wu et al., 2014), relação e origem de germoplasma (Lu et al., 2009; Semagn et al., 2012), construção de mapas físicos (Gore et al., 2009), detecção de genes (Peiffer et al., 2013), associação genômica (GWAS – Genome Wide Association), e seleção genômica (GWS – Genome Wide Selection) (Frascaroli et al., 2013). Estes marcadores são bialélicos, de baixa informação polimórfica, o que é compensado por terem boa cobertura no genoma do milho (Van Inghelandt et al., 2010) e pelos avanços na automação do processo de genotipagem para diversas espécies (Frascaroli et al., 2013).

Para fins de detectar a origem ou parentesco dos híbridos comerciais brasileiros por meio de análises de diversidade genética e estruturação populacional é interessante usar uma população de referência. Visando prover recursos genéticos para estudos de caracteres quantitativos (QTLs) em milho, McMullen et al. (2009) desenvolveram a população Nested Association Mapping (NAM). A população NAM, consiste de 25 progênies com 200 linhagens endogâmicas recombinantes (RILs) cada. Os genitores utilizados para a formação das progênies foram a linhagem B73, como testadora, e 25 linhagens selecionadas entre 260 linhagens endogâmicas representantes da América do Norte (Estados Unidos e Canadá), Europa, África do Sul, Tailândia e dos bancos de germoplasma do CIMMYT e do Instituto internacional de Agricultura Tropical (IITA – International Institute of Tropical Agriculture), visando avaliar a diversidade genética de cada região e as relações entre os germoplasmas destas regiões (Liu et al., 2003). Estas representam a diversidade mundial de milho e são divergentes com a linhagem B73 (McMullen et al., 2009). A linhagem B73 foi escolhida como genitor comum devido ao seu uso na construção do mapa físico do milho (Wei et al., 2007), no projeto de sequenciamento do milho (www.maizesequence.org) e ser uma das linhagens mais utilizadas em blocos de cruzamentos em todo o mundo, em vista da sua alta capacidade geral de combinação.

2 OBJETIVOS

i) Estimar a variabilidade genotípica entre os híbridos das principais empresas de melhoramento de milho no Brasil.

ii) Identificar o risco de vulnerabilidade genética, por meio do tamanho efetivo populacional da população dos híbridos comerciais brasileiros e do germoplasma que a originou, e as relações de parentesco genético entre os germoplasmas das empresas de melhoramento de milho no Brasil.

iii) Estimar a relação do germoplasma comercial de milho com os genitores do painel de diversidade Nested Association Mapping.

3 MATERIAL E MÉTODOS

3.1 Material genético

Para representar o germoplasma elite de milho brasileiro foram utilizados híbridos comerciais das empresas de maior destaque no mercado de sementes. Estas empresas juntas possuem 94% do market-share, em que, a Monsanto[®] lidera com 35%, seguida pela DuPont Pioneer[®] (30%), Dow Agrosiences[®] (15%), Syngenta[®] (10%) e Helix Sementes[®] (4%). A Monsanto foi representada pela marca comercial Dekalb[®] e a Dow Agrosiences foi representada por si e pela marca Agromen Tecnologia[®]. Já a Helix Sementes foi representada pelas marcas Biomatrix[®] e Santa Helena Sementes[®]. Estes híbridos (população 1) possuem alta adaptabilidade, e são, na sua maioria, transgênicos (Tabela 1).

Tabela 1 – Descrição dos 20 híbridos comerciais de milho deste estudo

Empresa	Híbrido	Tipo	Ciclo	Tipo Grão	Cor do Grão
Agromen Tecnologia	30A37PW ^a	HS	Precoce	Semi-Dentado	Amarelo-Laranja
Biomatrix	BM820	HS	Precoce	Duro	Avermelhado
	BM207	HD	Precoce	Semi-Duro	Avermelhado
	BM915PRO ^b	HS	Super-Precoce	Semi-Dentado	-
Dekalb	DKB 340 PRO ^b	HS	Semi-Precoce	Semi-Duro	Amarelo
	DKB 310 PRO ^b	HS	Normal	Semi-Duro	Amarelo-Laranja
	DKB 390	HS	Precoce	Duro	Amarelo-Laranja
	DKB 177 PRO ^b	HS	Precoce	Semi-Duro	Amarelo-Laranja
Dow Agrosiences	2B688PW ^a	HT	Precoce	Semi-Duro	Alaranja
	2B587PW ^a	HS	Precoce	Semi-Dentado	Amarelo-Laranja
	2B710PW ^a	HS	Precoce	Semi-Duro	Amarelo-Laranja
	2B810PW ^a	HS	Normal	Semi-Duro	Alaranja
	2A550PW ^a	HS	Precoce	Semi-Duro	Alaranja
DuPont Pioneer	P4285H ^c	HS	Precoce	Duro	Amarelo-Laranja
	30F53H ^c	HS	Precoce	Semi-Duro	Alaranja
Santa Helena Sementes	SHS 5560	HT	Precoce	Duro	Laranja
Syngenta	Truck TL ^d	HS'	Precoce	Duro	Alaranja
	Fórmula TL ^d	HS	Super-Precoce	Duro	Alaranja
	Status Viptera ^e	HS	Precoce	Duro	Alaranja
	Impacto TL ^d	HS	Precoce	Duro	Alaranja

HS: híbrido simples, HS': híbrido simples modificado, HD: híbrido duplo, HT: híbrido triplo. Letras em sobrescrito representam a tecnologia transgênica presente no híbrido: ^a PowerCore[®], ^b YieldGard[®], ^c Herculex[®], ^d TL[®], e Viptera[®].

Com a finalidade de obter conhecimento sobre as principais fontes do germoplasma brasileiro atual, utilizou-se, juntamente com os híbridos anteriormente descritos, os genitores da população Nested Association Mapping (NAM) (www.panzea.org) e a linhagem Mo17 (Mcmullen et al., 2009) (população 2; Tabela 2). A linhagem Mo17 foi utilizada devido a sua importância para o melhoramento do milho e pela alta capacidade específica de combinação com a linhagem B73. Desta população foram utilizados, as linhagens Mo17, B73 e outras 25.

Tabela 2 - Painel dos genitores da população NAM utilizados para informar sobre a possível origem do germoplasma brasileiro de milho

Linhagem	Grupo/Origem	Linhagem	Grupo/Origem
B73	stiff stalk	Ky21	non-stiff stalk
B97	non-stiff stalk	M162W	non-stiff stalk
CML103	tropical-subtropical	M37W	-
CML228	tropical-subtropical	Mo17	lancaster
CML247	tropical-subtropical	Mo18W	-
CML277	tropical-subtropical	MS71	non-stiff stalk
CML322	tropical-subtropical	NC350	tropical-subtropical
CML333	tropical-subtropical	NC358	tropical-subtropical
CML52	tropical-subtropical	Oh43	lancaster
CML69	tropical-subtropical	Oh7b	-
Hp301	milho pipoca	P39	milho doce
Il14H	milho doce	Tx303	-
Ki3	tropical-subtropical	Tzi8	tropical-subtropical
Ki11	tropical-subtropical		

Fonte: Liu et al. (2003).

Os genitores da NAM e a linhagem Mo17 foram genotipadas com um array de 1536 marcadores SNPs pela plataforma Illumina GoldenGate® (Illumina, San Diego, CA). No banco de dados Panzea (www.panzea.org) estão disponibilizadas as informações genóticas das 27 linhagens, os 26 genitores da NAM e a linhagem Mo17, para os 1106 SNPs que foram informativos.

3.2 Genotipagem e controle de qualidade dos dados genéticos

Para coleta do material genético, foram semeadas três sementes por híbrido em vasos de cinco litros, preenchidos com uma mistura de solo e substrato em mesma proporção, no campo experimental Diogo Alves de Melo na Universidade Federal de

Viçosa, no mês de março de 2014. No estádio de três folhas completamente expandidas (V3), foram coletadas folhas destes híbridos para realizar sua liofilização no Bioagro, na Universidade Federal de Viçosa.

As amostras foram então enviadas à empresa DuPont Pioneer[®], para a realização da genotipagem. Essa foi feita por meio da plataforma Illumina GoldenGate[®] com um array de 768 SNPs (Fan et al., 2003). Este painel foi o mesmo utilizado por Nelson et al. (2008) com o objetivo de inferir sobre a estruturação populacional, a origem e as relações entre os genitores de híbridos comerciais Norte-americanos.

De posse dos dados genômicos realizou-se a conversão dos SNPs para algarismos numéricos (1, 2 e 3) por meio do pacote *scrim* (SCHWENDER e FRITSCH, 2013) do programa R (R Core Team, 2014). O controle de qualidade dos dados genômicos foi realizado considerando Call Rate de 90% e Minor Frequency Allele (MAF) de 5%. Para este fim utilizou-se o pacote *HapEstXXR* (KNUEPPEL e ROHDE, 2014) também do programa R

Ao utilizar os dados genotípicos do NAM, foram analisados apenas os SNPs comuns entre a população 1 e os genitores da população 2, originando uma população conjunta (híbridos brasileiros e genitores do NAM). Após a confecção deste painel procedeu-se com o controle de qualidade dos dados semelhante ao descrito anteriormente.

3.3 Análises dos dados genômicos

Foram realizadas para cada SNP, para a população dos híbridos comerciais brasileiros e a população conjunta, as seguintes estimativas:

i) Diversidade gênica ou heterozigidade esperada - D_g ($D_g = 1 - \sum_{i=1}^k p_i^2$), em que

p_i é a frequência do alelo i para o SNP avaliado.

ii) Minor allele Frequency - MAF ($MAF = \frac{H_e + R}{2 \times (D + H_e + R)}$), em que H_e é o total

de locos heterozigotos para o SNP avaliado, D é o total de locos homozigotos para o alelo i , R é o total de locos homozigotos para o alelo i' , este último alelo de menor frequência.

iii) Heterozigosidade observada - H_o ($H_o = \frac{H_e}{D + H_e + R}$), em que H_e é o total de

locos heterozigotos para o SNP avaliado, D é o total de locos homozigotos para o alelo i , R é o total de locos homozigotos para o alelo i' .

iv) Conteúdo de informação polimórfica - PIC (Polimorphism Information Content)

($PIC = 1 - \sum_{i=1}^k p_i^2 - \sum_{i=1}^{k-1} \sum_{j=i+1}^k 2p_i^2 p_j^2$), em que p_i e p_j é a frequência do alelo i e j para o

SNP avaliado.

Estes parâmetros foram estimados por meio do software PowerMarker 3.25 (Liu e Muse, 2005).

A partir das frequências alélicas foram estimadas as distâncias genóticas entre os

genótipos conforme proposto por Nei et al. (1983) ($D_A = \frac{1}{r} \sum_{j=1}^r \left(1 - \sum_{i=1}^{m_j} \sqrt{x_{ij} y_{ij}} \right)$), em que

x_{ij} e y_{ij} são as frequências observadas do alelo i para o loco j nos genótipos x e y , respectivamente.

Em seguida, foi construído um dendrograma de consistência de nós, com os parâmetros de bootstrap (duas mil repetições), pelo método hierárquico da ligação média (UPGMA – Unweighted pair group method average). Estas análises foram realizadas por meio do software PowerMarker 3.25. O dendrograma de consistência de nós foi gerado por meio do software MEGA 5.1 (Tamura et al., 2011).

Para inferir sobre a vulnerabilidade genética do germoplasma comercial brasileiro

foi estimado o tamanho efetivo populacional (N_e) dos híbridos $N_e = \frac{1}{1 + F_{st}}$, e da

população que gerou os híbridos analisados “ $N_e = \frac{1}{2F_{st}}$ ”, em que F_{st} é o coeficiente de

endogamia populacional. Este foi estimado conforme proposto por Resende (2014), via

matriz de parentesco genômico (G), em que G é dado por “ $G = \frac{WW'}{\sum_{i=1}^n 2p_i q_i}$ ”, em que W

é a matriz de incidência para efeito fixo de alelos de marcadores bialélicos, p_i e q_i são as frequências alélicas de marcadores bialélicos para o loco i . O F_{st} é obtido pela média da diagonal da matriz G menos um.

A identificação de alelos privados de cada empresa foi feita pelo software Convert 1.31 (Glaubitz, 2004). São considerados alelos privados de determinado loco, aqueles que estão presentes em uma população e ausente nas demais. Em seguida, foram anotadas as posições dos marcadores que apresentaram estes alelos privados e as informações sobre as funções metabólicas e biológicas dos genes possivelmente associados a estes. Para isso utilizou-se a plataforma gramene (www.gramene.org) (Monaco et al., 2014).

Para quantificar a relação de parentesco entre os germoplasmas das empresas, realizou-se análise de correlação de Pearson entre as frequências alélicas das empresas para os marcadores SNPs selecionados.

Visando identificar a estrutura e as relações entre os germoplasmas elite das empresas de melhoramento de milho no Brasil (população 1) e suas relações com as linhagens genitoras do painel de diversidade NAM (população 2), foram feitas duas estruturações populacionais. Na primeira apenas foram considerados os genótipos da população 1 e na segunda os das populações 1 e 2, em conjunto. A estruturação populacional foi realizada pelo software Structure 2.3.4 (Pritchard et al., 2000). Foram consideradas as seguintes pressuposições: modelo com mistura entre as populações (with admixture), com número de populações (K) variando de 1 a 10, e com dez repetições para cada K. Para cada corrida, o número de cadeias de Markov (MCMC – Markov Chain Monte Carlo) foi de um milhão, com Burn-in das 500 mil primeiras cadeias de MCMC.

Visando estimar o número de populações que melhor estrutura os genótipos, foi utilizado o critério do valor de ΔK (Evanno et al., 2005)

$$\Delta K = \frac{m \left[L(K+1) - 2L(K) + L(K-1) \right]}{\sigma \left[L(K) \right]}$$

”, em que $L(K)$ é a média dos logaritmos

neperianos da probabilidade dos dados de cada passo do MCMC e subtraído pela metade da variância para o número de populações K. Este valor foi estimado pelo software Structure Harvester (Earl e Vonholdt, 2012).

De posse dos resultados da estruturação populacional, foi estimado as atribuições de cada linhagem a cada população de acordo com seus valores médios de probabilidade de associação (Q_{ik}), em que Q_{ik} é a proporção estimada do genoma do genótipo i derivado da população k (Pritchard et al., 2000). A partir dos valores de Q_{ik} de cada genótipo foi construído um gráfico de barras, o qual representa a relação do genótipo i com cada

população K . Considera-se genótipos pertencentes ao mesmo grupo ou população, aqueles cujo Q_k for superior 60%. Foram considerados como pertencentes à população mista os indivíduos que não se encaixarem em nenhuma população conforme descrito acima (Yang et al., 2011).

4 RESULTADOS E DISCUSSÕES

4.1 Análise de diversidade genética via marcadores SNPs

No presente estudo, do total de 768 marcadores SNPs genotipados nos híbridos comerciais, 50 falharam na genotipagem, outros 195 tinham mais de 10% de dados perdidos (call rate de 90%) e 108 tinham MAF inferior a 5%. Com isto, foram utilizados 413 SNPs, 53,8% dos marcadores. Já na população conjunta, formada pelos híbridos comerciais e as linhagens genitoras da população NAM (Mcmullen et al., 2009), foram encontrados 340 marcadores SNPs em comum e polimórficos. Destes, 266 foram selecionados considerando Call Rate mínimo de 90% e MAF à 5%. É comum em trabalhos dessa natureza a retirada de marcadores SNPs com baixo polimorfismo ou elevado número de dados perdidos. Isso é realizado para eliminar os marcadores que não possuem poder de distinguir os genótipos nas populações. Neste contexto, Mammadov et al. (2012) utilizaram call rate de 94% e MAF de 3% em três bancos de dados de milho, sendo dois do Panzea (Zhao et al., 2006; Gore et al., 2009) e um do Maize Assembled Genomics Islands (MAGI), em que de 10.599 SNPs, foram analisados apenas os 1054 informativos, ou seja, 9,94% do total.

Os marcadores SNPs geralmente apresentem valores de PIC relativamente baixos quando comparados, por exemplo, com os marcadores microssatélites, que chegam a valores próximos de um (Sant'ana et al., 2012). Mesmo assim, estes são abundantes e têm relativas taxas de polimorfismo, e podem ser encontrados em 80% dos fragmentos sequenciados (EST) do milho (Borém e Caixeta, 2009). Gore et al. (2009) identificaram cerca de 3,3 milhões de SNPs e deleções/inserções em milho. Segundo os mesmos autores, a cada 44 pares de bases nitrogenadas uma era polimórfica. Os marcadores SNPs selecionados para discriminar a população dos híbridos comerciais brasileiros exibiram estimativas de diversidade gênica, com média de 0,34, e amplitude de 0,09 à 0,50; o PIC médio foi de 0,28, e amplitude de 0,09 à 0,37; a estimativa da heterozigosidade média observada foi de 0,33, variando de 0 à 0,79; já a MAF apresentou média de 0,26, com amplitude de 0,05 à 0,50.

Estes resultados evidenciam que o conjunto de SNPs usados apresentaram diversidade gênica superior a outros trabalhos que utilizaram a mesma plataforma de genotipagem, como por exemplo, o de Nelson et al. (2008). Neste trabalho observou-se médias de diversidade gênica e PIC de 0,33 e 0,26, respectivamente. A diversidade gênica

nos 20 híbridos comerciais brasileiros é similar a observada entre as 103 linhagens do germoplasma comercial Norte-Americano da década de 80 encontrada por Nelson et al. (2008). Além disso, os SNPs identificados no germoplasma comercial brasileiro de milho apresentaram maior PIC, sugerindo maior diversidade genética do germoplasma brasileiro em relação ao americano.

A estimativa de heterozigosidade observada foi inferior a esperada tanto para os híbridos simples, como duplos. Logo, parte dos locos avaliados pode apresentar interação alélica aditiva e estar ligados a genes de importância agrônômica, pois os alelos estão presentes em ambos os genitores de cada híbrido. Além disso, este fenômeno pode ser consequência da seleção, na fixação de alelos favoráveis nas populações melhoradas (Carneiro et al., 2007).

As frequências observadas das estimativas de PIC, e de diversidade gênica (Figuras 1a e 1c; barras pretas) seguiram uma distribuição exponencial, ou seja, a maior parte dos SNPs selecionados possui altas estimativas de PIC e de diversidade gênica, com mais de 40% dos SNPs com estimativas de 0,41 a 0,50 para diversidade gênica. Já para PIC, mais de 50% dos SNPs apresentaram estimativas entre 0,31 e 0,40. Essas informações reiteram as previamente descritas, sobre o poder destes SNPs em discriminar genótipos.

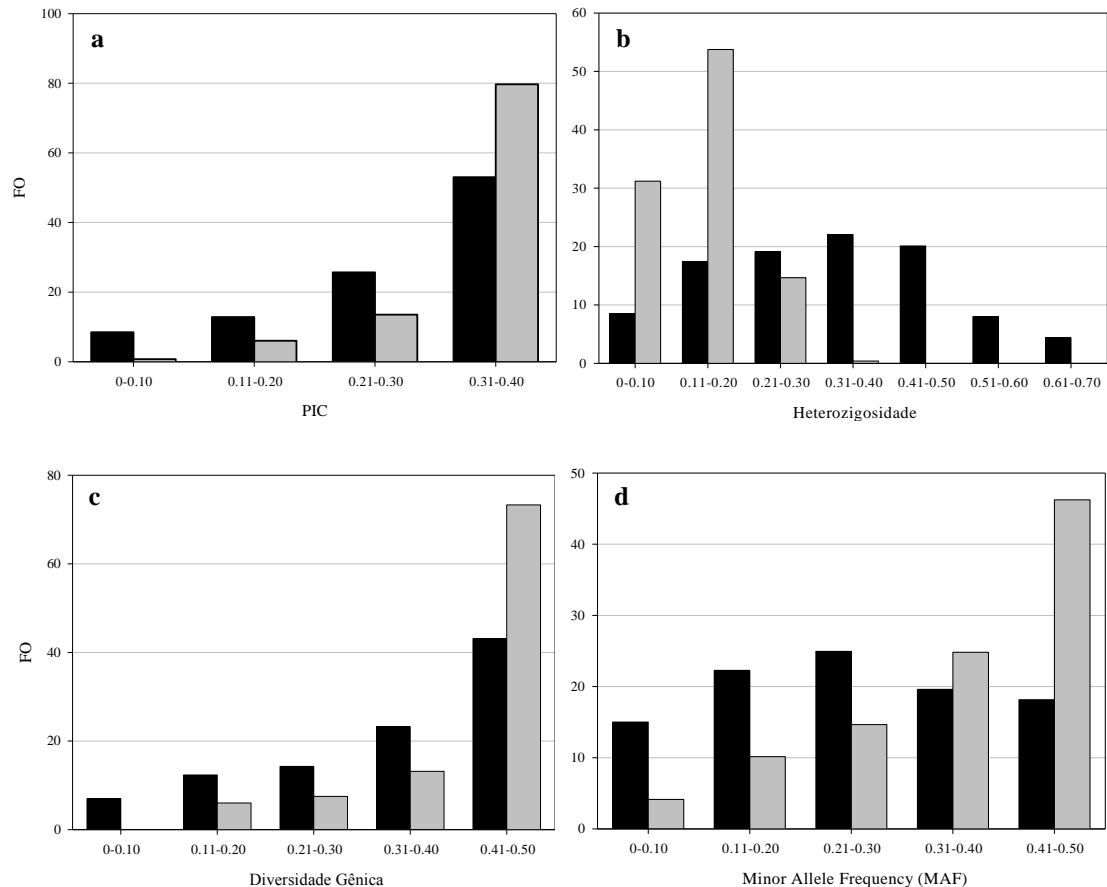


Figura 1 - Frequências observadas (FO) de marcadores SNPs para as estimativas de PIC (conteúdo de informação polimórfica) (a), heterozigidade observada (b), diversidade gênica (c) e MAF (Minor Allele Frequency) (d). Barras pretas referem-se as estimativas dos SNP SNPs na população dos 20 híbridos comerciais e as cinzas a população composta pelos 20 híbridos e os genitores da NAM.

Em relação às frequências de heterozigidade observadas e MAF (Figuras 1b e 1d, barras pretas), estas seguem uma distribuição normal platicúrtica, de baixo efeito de curtose. Estes marcadores SNPs apresentam alto desvio padrão quanto às estimativas de heterozigidade observada e MAF. A distribuição de heterozigidade observada é um pouco assimétrica à direita, pois possui moda e mediana menores que a média. Esses valores indicam que há predominância de homozigose para os locos avaliados, como já discutido anteriormente.

Outra relação interessante é o efeito que a heterozigidade tem sobre a MAF, especificamente quando os marcadores são bialélicos. Quando o loco é heterozigoto e sob as condições previamente descritas, os alelos possuem a mesma frequência alélica da população. Logo quanto maior a heterozigidade maior será a MAF. No entanto, o

aumento da MAF não implica necessariamente em altos valores de heterozigosidade observada.

Os marcadores utilizados para avaliar a relação entre os híbridos comerciais tiveram estimativas altas de diversidade gênica, com média de 0,43 e amplitude de 0,10 à 0,50; a estimativa média de PIC foi de 0,33, com amplitude de 0,09 à 0,37; a estimativa da heterozigosidade média observada foi de 0,13, variando de zero à 0,30; a MAF teve estimativas altas com média de 0,35 com amplitude de 0,05 à 0,50.

As médias da heterozigosidade observada foram diferentes entre as populações, nos híbridos de 0,33 e na população conjunta de 0,13. Isto ocorreu devido aos genitores da população NAM serem linhagens endogâmicas. No entanto, os índices de diversidade gênica e PIC, tiveram um incremento sensível em suas estimativas, ou seja, os SNPs comuns selecionados para as duas populações foram suficientes para discriminar os genótipos. O aumento na MAF de 0,25 para 0,35, indica que os marcadores na população NAM apresentaram, em média, elevados valores de MAF, muito próximo de 0,50.

As distribuições das estimativas de PIC e da diversidade gênica, para a população conjunta (20 híbridos e NAM), também seguiram distribuição exponencial (Figura 1a e 1c; barras cinza). Porém, houve aumento das frequências das estimativas de PIC de 0,31 a 0,4 de 53% para 79,7% e das de diversidade gênica de 0,41 a 0,5 de 43% para 73,3%. Isto ocorreu devido ao aumento do número de indivíduos analisados e também pela alta diversidade genética entre os genitores da população NAM.

Quando analisada a população conjunta, houve uma queda nas estimativas de heterozigosidade observada (Figura 1b; barras cinzas), o que era esperado devido aos genitores serem linhagens endogâmicas e estarem em maior número que os híbridos na população conjunta. Com isso, a distribuição das estimativas de heterozigosidade observada ficou com maior efeito de assimetria a direita. Embora estas estimativas tenham reduzido suas magnitudes, as de MAF aumentaram a ponto de alterar a distribuição normal para exponencial (Figura 1d; barras cinzas). As frequências das estimativas de MAF que variavam de 0,41 a 0,5 eram de 18,1% de todos os locos avaliados na população 1 e alteraram para 46,2% na população conjunta.

Os índices de diversidade gênica, PIC, heterozigosidade observada e MAF foram estimados também para os germoplasmas de cada empresa de melhoramento de milho no Brasil (Tabela 3). Para isso, os híbridos foram agrupados pelas suas empresas detentoras. As empresas apresentaram índices similares para todos os parâmetros. No entanto, a

Syngenta apresentou estimativas um pouco distintas das demais, com índices de diversidade gênica (0,31), PIC (0,24), heterozigidade observada (0,38) e MAF (0,24), maiores que as demais empresas. A diversidade gênica e a PIC foram baixas para as empresas Agromen Tecnologia e Santa Helena Sementes principalmente devido ao baixo número de híbridos amostrados de cada uma destas empresas. No entanto, os índices MAF e heterozigidade observada foram informativos, indicando que não há predomínio de locos heterozigotos nos híbridos destas empresas.

Tabela 3 - Número de híbridos (NH) e estimativas por empresas da diversidade gênica (D_g), conteúdo de informação polimórfica (PIC), heterozigidade observada (H_o), e menor frequência alélica (MAF)

Empresas	NH	D_g	PIC	H_o	MAF
Agromen	1	0,04	0,04	0,37	0,04
Biomatrix	3	0,28 (0-0,5)	0,22 (0-0,37)	0,31 (0-1)	0,21 (0-0,5)
Dekalb	4	0,28 (0-0,5)	0,22 (0-0,37)	0,34 (0-1)	0,21 (0-0,5)
Dow Agrosiences	5	0,26 (0-0,5)	0,21 (0-0,37)	0,30 (0-1)	0,20 (0-0,5)
DuPont Pioneer	2	0,27 (0-0,5)	0,21 (0-0,37)	0,32 (0-1)	0,22 (0-0,5)
Santa Helena	1	0,07	0,07	0,31	0,07
Syngenta	4	0,31 (0-0,5)	0,24 (0-0,37)	0,38 (0-1)	0,24 (0-0,5)

Por meio dos índices de diversidade gênica e PIC pode-se inferir também sobre a diversidade genética do germoplasma e o grau de parentesco entre os híbridos dentro de cada empresa. Quanto maior a estimativa destes índices maior é a diversidade genética do germoplasma da empresa e, portanto, menor o grau de parentesco entre os híbridos desta (Oyekunle et al., 2015). Todas as empresas apresentaram baixas estimativas de diversidade gênica e PIC para os locos avaliados. Isto ocorreu, possivelmente, devido ao baixo número de híbridos avaliados por empresa.

Foram identificados um, três e 12 alelos privados nas empresas Biomatrix, DuPont Pioneer e Syngenta, respectivamente (Tabela 4). O grande número de alelos privados observados no germoplasma da empresa Syngenta, evidencia uma maior divergência genética desta em relação às demais. Os híbridos da Syngenta apresentam alelos privados nos cromossomos um, quatro, cinco, seis, sete e nove (Panzea - www.panzea.org). Estes alelos estão relacionados indiretamente a diversas funções biológicas como a sinalização celular e a organelas celulares como cloroplasto e vacúolo (Gramene - www.gramene.org). No entanto, até o momento nenhum destes alelos possui função direta com caracteres agronômicos ou componentes de produção (Gramene - www.gramene.org).

Tabela 4 - Alelos privados do germoplasma de cada empresa, cromossomo (Chr), posição no cromossomo em kb (Pos), gene associado, distância do gene associado em kb (Dist) e funções do gene associado

Empresas	Marcador	Chr	Pos (kb)	Gene associado	Dist (kb)	Funções
Biomatrix	PZA02970.9	6	98.315,9	GRMZM2G701028	29,7	-*
D. Pioneer	PZA00065.2	1	49.621,4	GRMZM2G056373	1,9	-*
	PZA00427.3	6	80.974,9	GRMZM2G103033	89,8	Ligação proteica.
	PZA02519.7	10	142.184,1	GRMZM2G028902	23,8	-*
Syngenta	PHM4531.46	1	22.892,8	GRMZM2G312226	íntron	Processo catabólico da quitina, Processo catabólico da membrana celular, Atividade da quitina.
	PHM11000.2	1	43.554,6	GRMZM2G030880	íntron	Cloroplasto, Membrana tilacóide do cloroplasto.
	PZA00975.1	4	9.759,8	GRMZM2G097469	0,04	-*
	PHM14055.6	4	53.441,5	GRMZM2G049525	26,6	Ciclo Celular Mitótico, Recombinação do DNA.
	PHM1307.11	4	56.312,8	GRMZM2G568636	22,9	Vacúolo, Processo de Oxi-redução.
	PZA00401.1	5	56.019,4	GRMZM2G048455	23,4	Cascata de MAP quinase.
	PZA03063.18	6	35.896,3	GRMZM2G035741	íntron	-*
	PHM10225.15	7	167.937,9	GRMZM2G150758	8,3	-*
	PZA03416.7	9	17.000,2	GRMZM5G891465	0,4	-*
	PHM13183.1	9	107.796,8	GRMZM2G055320	13,4	Atividade Catalítica.
PHM2278.86	9	112.333,8	GRMZM2G093270	18,3	-*	

Fonte: Gramene (www.gramene.org/ acessado em 11 de janeiro de 2015). *Nenhum relato ontológico encontrado até o momento.

4.2 Vulnerabilidade genética dos híbridos brasileiros de milho

A estimativa de tamanho efetivo populacional é muito usada para representar a diversidade genética encontrada na população, o risco de deriva genética, e até mesmo a vulnerabilidade genética de uma população (Sebbenn e Seoane, 2005). Neste sentido, a coeficiente de endogamia populacional (F_{st}) nos 20 híbridos foi de 0,08, o que corresponde a um tamanho efetivo populacional estimado de 18,54 indivíduos. Já a estimativa do tamanho efetivo populacional dos genitores dos 20 híbridos comerciais avaliados foi de 127,22 indivíduos. Este alto tamanho efetivo populacional permite inferir que o quadro atual dos híbridos comerciais brasileiros oferece pouco risco de vulnerabilidade genética, considerando a variabilidade genética e o coeficiente de parentesco entre eles. Segundo Sebbenn e Seoane (2005) tamanhos efetivos superiores ou similares à população real estudada ocorrem quando há estimativas de correlação genéticas negativas ou próximas de zero entre os genótipos avaliados. Todavia, é

importante ressaltar que existem outros fatores ligados a vulnerabilidade genética que não foram considerados neste estudo como exemplo: a área de plantio, a proporção de mercado de cada híbrido e a vulnerabilidade devido a genes de resistência a pragas e doenças, o que sugere, que a baixa vulnerabilidade aqui encontrada deve ser considerada com cautela.

Neste mesmo contexto, o germoplasma comercial Norte-Americano de milho da década de 80 (Nelson et al., 2008), apresentou tamanho efetivo populacional de 55,9 indivíduos, correspondente a estimativa de coeficiente endogâmico populacional (F_{st}) de 0,84. Esta estimativa evidencia maior vulnerabilidade genética deste germoplasma em relação ao germoplasma comercial brasileiro.

As correlações entre as frequências alélicas dos marcadores SNP's entre as empresas foram significativas ao nível de 1% de significância (Tabela 5). Isso indica que há uma relação de parentesco entre o germoplasma das diferentes empresas consideradas. Destes, a empresa Dow Agrosciences obteve os maiores índices de correlação alélica com as empresas Agromen Tecnologia (0,74), Santa Helena Sementes (0,77) e Biomatrix (0,68). A correlação entre a Dow Agrosciences e a Agromen Tecnologia pode ser explicada pelo fato de a segunda empresa ser afiliada da primeira, porém com administração e estrutura produtiva independentes. No entanto, há a possibilidade de agregar as tecnologias e acessar o banco de germoplasma unificado (www.dowagro.com/br/agromen/nossa/brasil.htm, acessado em 28 de janeiro de 2015). Já as correlações entre a empresa Dow Agrosciences com a Biomatrix e a Santa Helena Sementes são altas apesar de não existir relações conhecidas entre elas. Entretanto, as empresas Biomatrix e Santa Helena Sementes são pertencentes a Helix Sementes. A Santa Helena Sementes foi adquirida pela Helix Sementes em 2011 e transferiu os estoques de semente, o banco de germoplasma e a estrutura comercial para a Biomatrix (www.agroceres.com.br/negocios_sementessh.jsp, acessado em 28 de janeiro de 2015). Embora estas empresas sejam do mesmo grupo, elas não possuem uma correlação tão alta entre si como a correlação delas com a empresa Dow Agrosciences. Isto provavelmente ocorreu pelo pouco tempo de intercâmbio de material, sendo posterior ao desenvolvimento e lançamento dos cultivares avaliados.

Tabela 5 – Correlação entre as frequências alélicas de 413 SNPs presentes no germoplasma das principais empresas de melhoramento de milho no Brasil

Empresas	Biomatrix	Dekalb	Dow	Pioneer	SH	Syngenta	Média
Agromen	0,57**	0,53**	0,74**	0,52**	0,48**	0,47**	0,55
Biomatrix		0,65**	0,68**	0,58**	0,55**	0,53**	0,59
Dekalb			0,60**	0,54**	0,51**	0,54**	0,56
Dow Agrosiences				0,58**	0,77**	0,59**	0,66
DuPont Pioneer					0,44**	0,50**	0,52
Santa Helena (SH)						0,52**	0,54
Syngenta							0,52

** significativo a 1% de probabilidade pelo teste t.

O fato das intensidades das correlações alélicas terem sido de média a alta, pode ser também explicado pelos objetivos comuns das empresas, as quais buscam obter genótipos de alta adaptabilidade, precoces, resistentes a pragas e doenças, e produtivos. Além disso, muitas empresas tiveram a mesma fonte de germoplasma inicial. Outro fator é a extração de linhagens pelas empresas a partir de cultivares elite de empresas concorrentes, o que é permitido no Brasil pelo artigo 10, inciso quatro da lei ordinária Brasileira de Nº 9.456 de 27 de março de 1997. O inciso deste artigo elucida que o uso de cultivar protegido com a finalidade de fonte de variação de melhoramento genético ou pesquisa científica não fere o direito de propriedade sobre a cultivar protegida.

Outro parâmetro analisado no presente trabalho foi a estruturação populacional. A análise de estruturação populacional indica o número provável de populações que possam ter dado origem à população em estudo. Para isso, geralmente é utilizado a estatística bayesiana, a qual segrega a população de acordo com a sua frequência alélica, separando os indivíduos da população em populações em equilíbrio de Hardy-Weinberg. Neste estudo, o número de populações ótimo estimado para estruturar a população dos 20 híbridos comerciais de milho, baseado no teste de Δk (Earl e Vonholdt, 2012), foi cinco (Figura 2a).

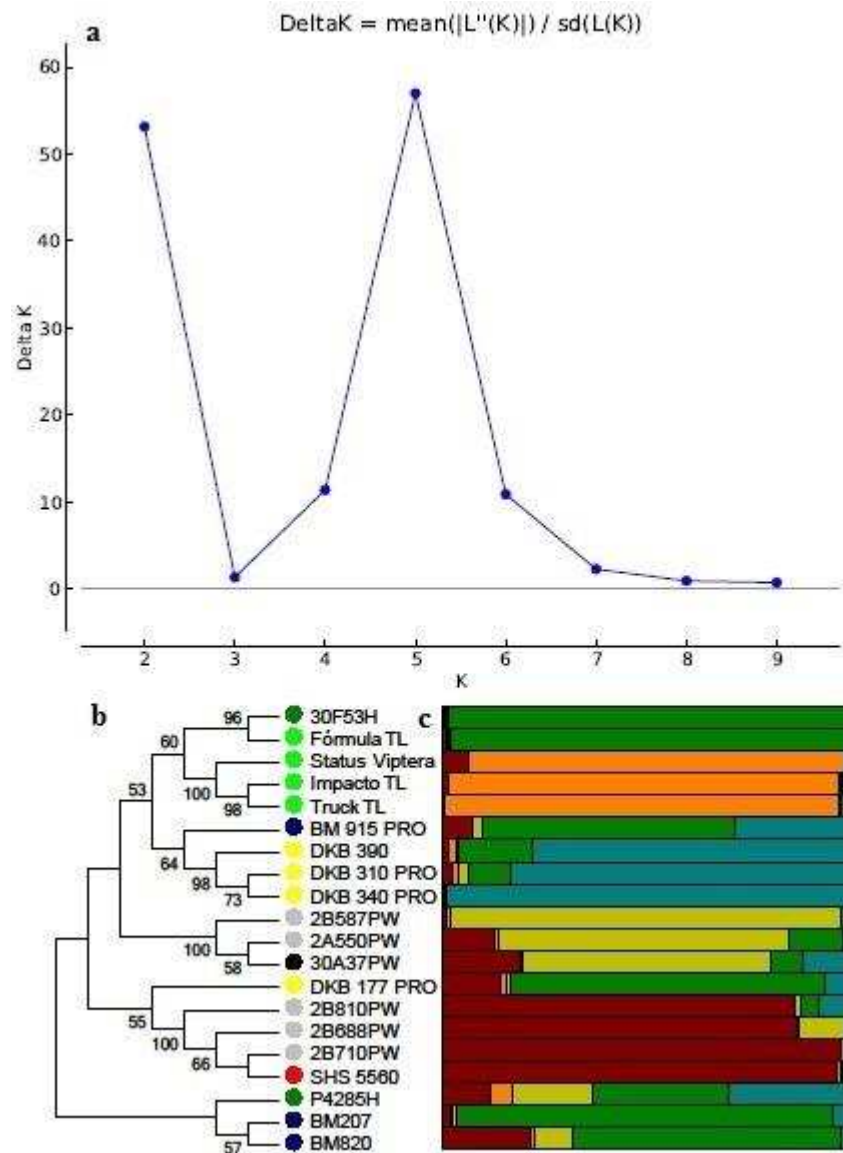


Figura 2 - Delta K (Δk) obtido por meio do Structure Harvester (a), Dendrograma e consistência de nós (b) e estruturação populacional dos 20 híbridos comerciais de milho em cinco populações (c). Círculos coloridos representam a empresa do genótipo. Agromen Tecnologia - preto; Biomatrix - azul-escuro; Dekalb - amarelo; Dow Agrosiences - cinza; DuPont Pioneer - verde-escuro; Santa Helena Sementes - vermelho; e Syngenta - verde-claro.

No entanto, estruturação segregou os híbridos em cinco populações (Figura 2c), no entanto o híbrido P4285H (DuPont Pioneer) não pertenceu a nenhuma destas populações, sendo uma mistura desta. A primeira foi composta por três híbridos da Dow Agrosiences

(2B710PW, 2B688PW, 2B810PW) e o híbrido da Santa Helena Sementes (SHS 5560). A segunda população foi composta por três híbridos da Syngenta (Truck TL, Impacto TL e Status Viptera). A terceira população foi formada pelos dois híbridos restantes da Dow Agrosciences (2B587PW, 2A550PW) e o híbrido da Agromen Tecnologia (30A37PW). A quarta população foi formada, pelos híbridos da Biomatrix (BM 207, BM820, BM915 PRO), o último da Syngenta (Fórmula TL), um da DuPont Pioneer (30F53H) e um da Dekalb (DKB 177 PRO). A quinta população foi composta pelos híbridos restantes da Dekalb (DKB 340 PRO, DKB 310 PRO e DKB 390). Para ser considerado como membro de uma população é necessário que a estrutura do híbrido tenha no mínimo 60% do Q_k de uma das populações (Yang et al., 2011). Os híbridos representantes das populações um, dois, três, quatro e cinco, foram os híbridos 2B710PW, Truck TL, 2B587PW, 30F53H e DKB 340 PRO, respectivamente, devido a apresentar os maiores Q_k de cada população. Assim, as populações que mais contribuíram para o germoplasma comercial brasileiro de milho foram a quarta (29,2%), a primeira (28,0%), a terceira (16,2%), a quinta (15,4%).e a segunda (11,2%).

Destas cinco populações estruturadas pela análise bayesiana, três são exclusivas de empresas. Isso reitera que há diversidade genética satisfatória na população brasileira de híbridos comerciais de milho, sendo esta maior entre as empresas do que dentro destas. As empresas que formaram populações exclusivas foram a Dekalb (DKB 340 PRO), a Dow Agrosciences (2B710PW e 2B587PW) e a Syngenta (Truck TL). No entanto, os híbridos destas empresas não pertencem todos a uma mesma população, tendo, como no caso da Dow Agrosciences, duas populações, uma exclusiva e uma formada com o híbrido da Santa Helena Sementes e os demais híbridos da Dow Agrosciences. A empresa Dupont Pioneer foi uma exceção, pois todos os seus híbridos foram alocados em populações diferentes, o que evidencia maior variabilidade genética do seu germoplasma desta em relação ao das outras empresas.

Já a empresa Biomatrix, quando comparada as outras empresas, apresentou baixa diversidade genética dentro de seu germoplasma, devido ao agrupamento de todos os híbridos em apenas uma população. O híbrido 30A37PW da Agromen Tecnologia, mais uma vez mostrou similaridade com os híbridos da Dow Agrosciences, reiterando os resultados obtidos na correlação das frequências alélicas entre as empresas. A estruturação populacional não só indicou que existe similaridade entre as empresas, Dow Agrosciences e Agromen Tecnologia, mas também qual o híbrido mais similar com 30A37PW, o 2A550PW. Isso corrobora com os fatos do livre acesso que a empresa

Agromen Tecnologia possui ao banco de germoplasma da empresa Dow Agrosiences (www.dowagro.com/br/agromen/nossa/brasil.htm, acessado em 28 de janeiro de 2015). O híbrido DKB 177 PRO da Dekalb foi distinto dos demais híbridos da empresa.

Apesar de a empresa Syngenta ter maior número de alelos privados, o que evidencia uma maior divergência genética desta empresa em relação as outras, a estruturação indicou divergência genética dentro do germoplasma da Syngenta, entre o híbrido Fórmula TL com os demais híbridos da empresa. Todavia existe variabilidade genética dentro do germoplasma da Syngenta.

Em estudos anteriores, utilizando a mesma genotipagem, Nelson et al. (2008) analisaram a estruturação populacional de 92 linhagens de milho com o direito de proteção Norte-Americano expirado. O número de populações ótimo estimado foi sete, representados pelas linhagens com suas contribuições para cada uma das sete populações avaliadas B73 (15,9%), Mo17 (12,9%), PH207 (12,7%), A632 (14,8%), Oh43 (11,2%), B37 (8,0%) e a população mista (24,4%). Os autores citaram ainda que, ao avaliar por grupo heterótico, as contribuições destes grupos para o germoplasma comercial americano foram de 38,7% do Stiff Stalk (B73, B37 e A632), 24,1% do Lancaster (Mo17 e Oh43) e 12,7% do Iodent (PH207).

Construído a partir das distâncias genéticas entre os híbridos comerciais, o dendrograma (Figura 2b) separou estes em quatro grupos. O grupo um foi o mais diverso, sendo formado por nove híbridos, quatro da Syngenta (Fórmula TL, Status Viptera, Impacto TL, Truck TL), três da Dekalb (DKB 390, DKB 310 PRO, DKB 340 PRO), um da DuPont Pioneer (30F53H) e um da Biomatrix (BM 915 PRO). O segundo grupo foi formado por dois híbridos da Dow Agrosiences (2B587PW e 2A550PW) e um da Agromen Tecnologia (30A37PW). No grupo três estavam os demais híbridos da Dow Agrosiences (2B810PW, 2B688PW, 2B710 PW), um da Dekalb (DKB 177 PRO) e o da Santa Helena Sementes (SHS 5560). O último grupo foi formado por apenas dois híbridos da Biomatrix (BM 207 e BM 820). O Híbrido P4285H da DuPont Pioneer não pertenceu a nenhum grupo, pois não obteve um nó com consistência superior a 50%. Assim, ele foi o mais geneticamente distinto dos demais, incluindo o outro híbrido da DuPont Pioneer, (30F53H).

O uso da estruturação populacional juntamente a dendrogramas de dissimilariedade auxilia na diferenciação dos grupos genéticos. Utilizando estes métodos percebe-se a formação de seis grupos e que três híbridos não se ajustaram a nenhum destes grupos. Já

o dendrograma reitera que existe variabilidade genética satisfatória entre os híbridos comerciais brasileiros, pois os separa em, pelo menos, quatro grupos distintos.

4.3 Relação entre o germoplasma dos híbridos comerciais brasileiros e os genitores da população Nested Association Mapping

Para inferir sobre a relação do germoplasma comercial brasileiro com os demais usados no mundo é interessante confrontar os dados genômicos dos híbridos nacionais com uma população que representa a diversidade genética mundial. Para isto, a população utilizada foi a dos genitores do Nested Association Mapping.

Por meio do teste de Δk (Earl e Vonholdt, 2012), foi estimada a estruturação populacional ótima de duas populações (Figura 3b). Os genitores do NAM foram alocados em apenas um grupo, com exceção da linhagem B73, a qual ficou no mesmo grupo que os híbridos brasileiros. Isso indica que os genitores da NAM são geneticamente distintas da linhagem B73, como era esperado pela própria formação desta população (Mcmullen et al., 2009). Também foi evidente o contraste entre as linhagens B73 e Mo17, as quais são as mais utilizadas em programas de melhoramento a nível mundial visando explorar o vigor híbrido. Nelson et al. (2008) também observou uma significativa distância genética entre estas duas linhagens.

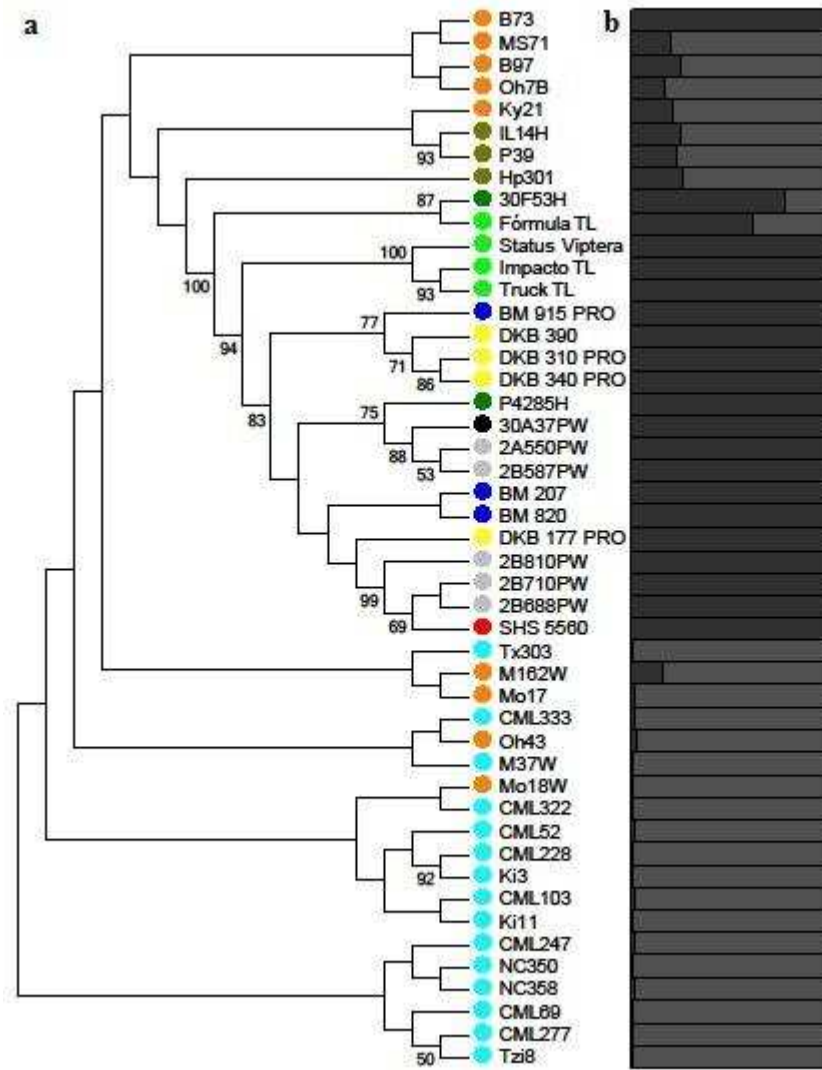


Figura 3 - Dendrograma de consistência de nós (a) e estruturação populacional dos híbridos e linhagens genitoras do Nested Association Mapping (b). Círculos coloridos representam a empresa ou origem do genótipo. Agromen Tecnologia - preto; Biomatrix – azul-escuro; Dekalb – amarelo; Dow Agrosiences – cinza; DuPont Pioneer - verde-escuro; Santa Helena Sementes – vermelho; Syngenta – verde-claro; linhagens tropicais – azul-claro; linhagens temperadas –laranja e linhagens de milho pipoca ou doce.

O dendrograma de distâncias genéticas entre os híbridos comerciais e a população NAM (Figura 3a) reitera estes resultados, que os híbridos foram separados das linhagens genitoras do NAM. No entanto, evidenciou-se que as linhagens do NAM têm tanta variabilidade genética que houve uma baixa ocorrência de nós consistentes entre elas. As

maiores taxas de nós coincidentes foram observadas para os híbridos brasileiros, os quais foram os únicos a formarem grupos e subgrupos. Isto indica que o processo de melhoramento de milho no Brasil levou a uma redução na variabilidade genética do germoplasma. Assim, apesar de indicar haver um elevado tamanho efetivo populacional é necessário ter cuidado com a seleção para não estreitar a base genética entre os híbridos comerciais brasileiros, pois intensidades elevadas podem conduzir a perda de variabilidade genética, não apenas para a característica selecionada (Bernardo, 2010).

Por fim, ao analisar o dendrograma e a estruturação populacional percebe-se que apesar da linhagem B73 ser divergente dos híbridos brasileiros pela análise de dissimilaridade genética, essa foi agrupada na mesma população pela análise de estruturação populacional com os híbridos brasileiros. Mostrando que existe efeito de coalescência. Segundo Hartl e Clark (2010), a coalescência ocorre quando os indivíduos da população possuem um único ancestral comum. A coalescência ocorreu devido ao efeito fundador da linhagem B73, ou de seu grupo heterótico de origem (Iowa Stiff Stalk Synthetic), quando introduzidas por empresas privadas na década de 90 (Parentoni et al., 2013) visando redução do ciclo na cultura do milho. O efeito fundador evidencia a influência do grupo heterótico Iowa Stiff Stalk Synthetic na formação do germoplasma comercial brasileiro de milho.

5 CONCLUSÕES

Considerando os resultados obtidos não há evidências de imediata vulnerabilidade genética para o germoplasma comercial brasileiro de milho.

O germoplasma comercial brasileiro de milho atual evidencia menor vulnerabilidade genética que o germoplasma comercial Norte-Americano da década de 80.

Existem relações de parentesco significativas, positivas e de alta magnitude entre os germoplasmas das empresas de melhoramento de milho no Brasil.

O germoplasma comercial brasileiro de milho possui relação estreita com a linhagem B73 ou de linhagens relacionadas a esta, possivelmente devido ao efeito fundador do grupo heterótico Norte-Americano Stiff Stalk.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ABRASEM. **Anuário Abrasem 2013**. Brasília, 2013.

BORÉM, A.; CAIXETA, E. T. **Marcadores moleculares**. 2ª Ed. Viçosa, 2009. 532p.

BROWN, A. H. D. Indicators of genetic diversity, genetic erosion and genetic vulnerability for plant genetic resources for food and agriculture. **FAO**. 2008. URL <http://www.fao.org/docrep/013/i1500e/i1500e20.pdf>

BERNARDO, R. **Breeding for Quantitative Traits in Plants**. Stemma Press, 2010. ISBN 9780972072410. Disponível em: < <http://books.google.com.br/books?id=Aw9kbwAACAAJ> >.

CARNEIRO, P. L. S. et al. Inbreeding, alleles fixation and selection limit in populations under traditional or marker assisted selection methods. **Revista Brasileira De Zootecnia-Brazilian Journal of Animal Science**, v. 36, n. 2, p. 369-375, Mar-Apr 2007. ISSN 1516-3598. Disponível em: < <Go to ISI>://WOS:000246540500013 >.

Commission on genetic resources for food and agriculture. Plant genetic resources use them or lose them. **FAO**. 2014. URL http://www.fao.org/fileadmin/templates/nr/documents/CGRFA/factsheets_plant_en.pdf

CRUZ, J. C. et al. Quatrocentas e sessenta e sete cultivares de milho estão disponíveis no mercado de sementes do Brasil para a safra 2013/14. **EMBRAPA**. Sete lagoas. 2013. URL <http://www.cnpms.embrapa.br/milho/cultivares/index.php>

DUVICK, D. N. Biotechnology in the 1930s: the development of hybrid maize. **Nature Reviews Genetics**, v. 2, n. 1, p. 69-74, Jan 2001. ISSN 1471-0056. Disponível em: < <Go to ISI>://WOS:000166056000019 >.

EARL, D. A.; VONHOLDT, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation Genetics Resources**, v. 4, n. 2, p. 359-361, Jun 2012. ISSN 1877-7252. Disponível em: < <Go to ISI>://WOS:000303536400036 >.

EVANNO, G.; REGNAUT, S.; GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. **Molecular Ecology**, v. 14, n. 8, p. 2611-2620, Jul 2005. ISSN 0962-1083. Disponível em: < <Go to ISI>://WOS:000229961500029 >.

FAN, J. B. et al. Highly parallel SNP genotyping. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 68, p. 69-78, 2003. ISSN 0091-7451. Disponível em: < <Go to ISI>://WOS:000222969300010 >.

FRASCAROLI, E.; SCHRAG, T. A.; MELCHINGER, A. E. Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. **Theoretical and Applied Genetics**, v. 126, n. 1, p. 133-141, Jan 2013. ISSN 0040-5752. Disponível em: < <Go to ISI>://WOS:000313056900013 >.

GALVAO, J. C. C. et al. Seven decades of development of the maize production system. **Rev. Ceres**, Viçosa, v. 61, supl. dez. 2014. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-

737X2014000700007&lng=pt&nrm=iso>. acessos em 16 jan. 2015.
<http://dx.doi.org/10.1590/0034-737x201461000007>.

GLAUBITZ, J. C. CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. **Molecular Ecology Notes**, v. 4, n. 2, p. 309-310, Jun 2004. ISSN 1471-8278. Disponível em: <<Go to ISI>://WOS:000221636200054 >.

GORE, M. A. et al. A First-Generation Haplotype Map of Maize. **Science**, v. 326, n. 5956, p. 1115-1117, Nov 20 2009. ISSN 0036-8075. Disponível em: <<Go to ISI>://WOS:000271951000045 >.

HARTL, D. L. & CLARK, A. G. **Princípios de Genética de Populações**. Artmed. Porto Alegre. 4ªed. 2010. 660p.

KNUEPPEL, S. & ROHDE, K. HapEstXXR: Multi-locus stepwise regression (MSR). R package version 0.1-7. 2014. URL <http://CRAN.R-project.org/package=HapEstXXR>

LE CLERC, V. et al. Assessing temporal changes in genetic diversity of maize varieties using microsatellite markers. **Theoretical and Applied Genetics**, v. 110, n. 2, p. 294-302, Jan 2005. ISSN 0040-5752. Disponível em: <<Go to ISI>://WOS:000226553900012 >.

LEVINGS, C. S. The Texas Cytoplasm of Maize - Cytoplasmic Male-Sterility and Disease Susceptibility. **Science**, v. 250, n. 4983, p. 942-947, Nov 16 1990. ISSN 0036-8075. Disponível em: <<Go to ISI>://WOS:A1990EH79300031 >.

LIU, K. J. et al. Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. **Genetics**, v. 165, n. 4, p. 2117-2128, Dec 2003. ISSN 0016-6731. Disponível em: <<Go to ISI>://WOS:000187805300038 >.

LIU, K. J.; MUSE, S. V. PowerMarker: an integrated analysis environment for genetic marker analysis. **Bioinformatics**, v. 21, n. 9, p. 2128-2129, May 1 2005. ISSN 1367-4803. Disponível em: <<Go to ISI>://WOS:000228783000061 >.

LIU, Z. Z. et al. Analysis of Genetic Diversity and Population Structure of Maize Landraces from the South Maize Region of China. **Agricultural Sciences in China**, v. 9, n. 9, p. 1251-1262, Sep 20 2010. ISSN 1671-2927. Disponível em: <<Go to ISI>://WOS:000281919800002 >.

LU, Y. L. et al. Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. **Theoretical and Applied Genetics**, v. 120, n. 1, p. 93-115, Dec 2009. ISSN 0040-5752. Disponível em: <<Go to ISI>://WOS:000271939900009 >.

MAMMADOV, J. et al. Development of versatile gene-based SNP assays in maize (*Zea mays* L.). **Molecular Breeding**, v. 29, n. 3, p. 779-790, Mar 2012. ISSN 1380-3743. Disponível em: <<Go to ISI>://WOS:000300778000017 >.

MCMULLEN, M. D. et al. Genetic Properties of the Maize Nested Association Mapping Population. **Science**, v. 325, n. 5941, p. 737-740, Aug 7 2009. ISSN 0036-8075. Disponível em: <<Go to ISI>://WOS:000268723700046 >.

MIKEL, M. A.; DUDLEY, J. W. Evolution of north American dent corn from public to proprietary germplasm. **Crop Science**, v. 46, n. 3, p. 1193-1205, May-Jun 2006. ISSN 0011-183X. Disponível em: <<Go to ISI>://WOS:000237375900023 >.

MONACO, M. K. et al. Gramene 2013: comparative plant genomics resources. **Nucleic Acids Research**, v. 42, n. D1, p. D1193-D1199, Jan 2014. ISSN 0305-1048. Disponível em: <<Go to ISI>://WOS:000331139800175 >.

NEI, M.; TAJIMA, F.; TATENO, Y. Accuracy of Estimated Phylogenetic Trees from Molecular-Data .2. Gene-Frequency Data. **Journal of Molecular Evolution**, v. 19, n. 2, p. 153-170, 1983. ISSN 0022-2844. Disponível em: <<Go to ISI>://WOS:A1983QL21300006 >.

NELSON, P. T. et al. Molecular characterization of maize inbreds with expired US plant variety protection. **Crop Science**, v. 48, n. 5, p. 1673-1685, Sep-Oct 2008. ISSN 0011-183X. Disponível em: <<Go to ISI>://WOS:000259792100003 >.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014. URL <http://www.R-project.org/>.

DE RESENDE, M. D. V. **Seleção genômica – software para seleção genômica ampla (GWS)**. 1 ed. Viçosa: Universidade Federal de Viçosa. 2014. 34p

SCHWENDER, H. & FRITSCH, A. scrime: Analysis of High-Dimensional Categorical Data such as SNP Data. R package version 1.3.3. 2013. URL <http://CRAN.R-project.org/package=scime>

OYEKUNLE, M. et al. Genetic diversity of tropical early-maturing maize inbreds and their performance in hybrid combinations under drought and optimum growing conditions. **Field Crops Research**, v. 170, p. 55-65, Jan 2015. ISSN 0378-4290. Disponível em: <<Go to ISI>://WOS:000347132000007 >.

PARENTONI, S. N.; DE MIRANDA, R. A.; GARCIA, J. C. Implications on the introduction of transgenics in Brazilian maize breeding programs. **Crop Breeding and Applied Biotechnology**, v. 13, n. 1, p. 9-22, Mar 2013. ISSN 1984-7033. Disponível em: <<Go to ISI>://WOS:000318216700002 >.

PEIFFER, J. A. et al. The Genetic Architecture of Maize Stalk Strength. **Plos One**, v. 8, n. 6, Jun 20 2013. ISSN 1932-6203. Disponível em: <<Go to ISI>://WOS:000322342800117 >.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945-959, Jun 2000. ISSN 0016-6731. Disponível em: <<Go to ISI>://WOS:000087475100039 >.

REIF, J. C. et al. Trends in genetic diversity among European maize cultivars and their parental components during the past 50 years. **Theoretical and Applied Genetics**, v. 111, n. 5, p. 838-845, Sep 2005. ISSN 0040-5752. Disponível em: <<Go to ISI>://WOS:000232598600004 >.

ROMAY, M. C. et al. Three Cycles of Full-Sib Reciprocal Recurrent Selection in Two Spanish Maize Populations. **Crop Science**, v. 51, n. 3, p. 1016-1022, May 2011. ISSN 0011-183X. Disponível em: <<Go to ISI>://WOS:000289713900011 >.

SANT'ANA, G. C. et al. Comparison of a retrotransposon-based marker with microsatellite markers for discriminating accessions of *Vitis vinifera*. **Genetics and Molecular Research**, v. 11, n. 2, p. 1507-1525, 2012. ISSN 1676-5680. Disponível em: <<Go to ISI>://WOS:000305519600079 >.

SEBBENN, A. M.; SEOANE, C. E. S. Estimativa de tamanho efetivo de endogamia por marcadores genéticos. **Revista Árvore**, v. 29, p. 1-8, 2005. ISSN 0100-6762. Disponível

em: < http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622005000100001&nrm=iso >.

SEMAGN, K. et al. Molecular characterization of diverse CIMMYT maize inbred lines from eastern and southern Africa using single nucleotide polymorphic markers. **Bmc Genomics**, v. 13, Mar 25 2012. ISSN 1471-2164. Disponível em: < <Go to ISI>://WOS:000304567500001 >.

TAMURA, K. et al. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. **Molecular Biology and Evolution**, v. 28, n. 10, p. 2731-2739, Oct 2011. ISSN 0737-4038. Disponível em: < <Go to ISI>://WOS:000295184200003 >.

TECHNOW, F. et al. Identification of key ancestors of modern germplasm in a breeding program of maize. **Theoretical and Applied Genetics**, v. 127, n. 12, p. 2545-2553, Dec 2014. ISSN 0040-5752. Disponível em: < <Go to ISI>://WOS:000345336300002 >.

VAN INGHELANDT, D. et al. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. **Theoretical and Applied Genetics**, v. 120, n. 7, p. 1289-1299, May 2010. ISSN 0040-5752. Disponível em: < <Go to ISI>://WOS:000276674500002 >.

WEI, F. et al. Physical and genetic structure of the maize genome reflects its complex evolutionary history. **Plos Genetics**, v. 3, n. 7, p. 1254-1263, Jul 2007. ISSN 1553-7390. Disponível em: < <Go to ISI>://WOS:000248350000012 >.

WILLIAMS, T. R.; HALLAUER, A. R. Genetic diversity among maize hybrids. **Maydica**, v. 45, n. 3, p. 163-171, 2000. ISSN 0025-6153. Disponível em: < <Go to ISI>://WOS:000165651600003 >.

WU, X. et al. Fine genetic characterization of elite maize germplasm using high-throughput SNP genotyping. **Theoretical and Applied Genetics**, v. 127, n. 3, p. 621-631, Mar 2014. ISSN 0040-5752. Disponível em: < <Go to ISI>://WOS:000331962400009 >.

YANG, X. H. et al. Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. **Molecular Breeding**, v. 28, n. 4, p. 511-526, Dec 2011. ISSN 1380-3743. Disponível em: < <Go to ISI>://WOS:000297173800008 >.

ZHAO, W. et al. Panzea: a database and resource for molecular and functional diversity in the maize genome. **Nucleic Acids Research**, v. 34, p. D752-D757, Jan 1 2006. ISSN 0305-1048. Disponível em: < <Go to ISI>://WOS:000239307700158 >.