

BC/28355
IB/80666

T/UNICAMP

L297_d

**DISTÂNCIAS GENÉTICAS ENTRE LINHAGENS ENDOGÂMICAS
DE MILHO (*Zea mays* L.) E PREDIÇÃO DE HÍBRIDOS SIMPLES
ATRAVÉS DE MARCADORES DO TIPO RAPD**

LUCIANA LASRY BENCHIMOL LANZA

ORIENTADORA: PROFA. DRA. ANETE PEREIRA DE SOUZA

**DISSERTAÇÃO APRESENTADA AO INSTITUTO DE BIOLOGIA
DA UNIVERSIDADE ESTADUAL DE CAMPINAS PARA
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS.
ÁREA DE CONCENTRAÇÃO: GENÉTICA VEGETAL.**

Este exemplar entregue à direção final
da tese, após a defesa, e enviado a
Luciana Lasry Benchimol
Lanza
e aprovada para a obtenção do grau de Mestre.
23/7/96 *Apfouzo*

CAMPINAS - SÃO PAULO - BRASIL

1996

UNIDADE	IB
N.º CHAMADA	7001
N.º	L297d
V.	E. 02
TERMO DO BC/	28355
PROC.	667/96
C	☐
D	☒
PRECISO	R\$ 11,00
DATA	29/08/96
N.º CPD	CM 000 11602-1

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA CENTRAL DA UNICAMP

L297d

Lanza, Luciana Lasry Benchimol

Distâncias genéticas entre linhagens endogâmicas de milho (*Zea mays* L.) e predição de híbridos simples através de marcadores do tipo RAPD / Luciana Lasry Benchimol Lanza. -- Campinas, SP : [s.n.], 1996.

Orientador: Anete Pereira de Souza.

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Biologia.

1. Polimorfismo (Genética). 2. Reação em cadeia de polimerase. 3. Genética vegetal molecular. 4. Milho. 5. Correlação (Estatística). I. Souza, Anete Pereira. de. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

LOCAL E DATA: Campinas, 23 de julho de 1996

BANCA EXAMINADORA:

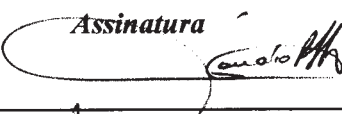
TITULARES:

Profa. Dra. Anete Pereira de Souza (Orientadora)



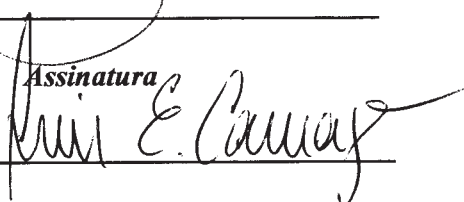
Assinatura

Prof. Dr. Cláudio Lopes de Souza Júnior



Assinatura

Prof. Dr. Luiz Eduardo Aranha Camargo



Assinatura

SUPLENTE:

Prof. Dr. Marcos Antônio Machado

Assinatura

“O homem vale aquilo que seu coração pode dar”.

Dedico,

Ao meu querido Marcos, minha principal razão de viver,

pelo carinho e por estar sempre ao meu lado,

mesmo nas horas mais difíceis;

E

A esta 'sementinha' que cresce dentro de mim, por acrescentar uma nova dimensão à

minha vida.

Ofereço,

Aos meus pais, por serem meu 'porto-seguro',

E

Aos meus irmãos, que sempre

me deram muita força.

Meus sinceros agradecimentos...

À **Profa. Dra. Anete Pereira de Souza**, pela orientação, estímulo científico e confiança depositada na minha capacidade profissional ;

Ao **Prof. Cláudio Lopes de Souza Júnior**, pela prontidão sempre sincera, pela coorientação indispensável em todas as etapas deste trabalho, principalmente na análise dos dados, e pela participação ativa da pré-banca e banca desta tese;

Ao **Prof. Dr. Luiz Eduardo Aranha Camargo** pela sincera contribuição na pré-banca e banca desta tese e pelas discussões e críticas engrandecedoras;

Ao **Prof. Dr. Marcos Antônio Machado** pela disponibilidade em participar da pré-banca e pelas considerações significativas para este trabalho;

À **Profa. Laura Maria Mariscal Ottoboni**, pelos ensinamentos sempre oportunos, e pela participação no exame de qualificação contribuindo para minha formação profissional como um todo;

Ao **Prof Dr. Paulo Arruda** pelo auxílio técnico-administrativo muito importante para o deslanchar deste projeto;

À **Profa. Dra. Maria Lúcia Carneiro Vieira**, pelo auxílio prestado no decorrer do trabalho;

Às **Profas. Dra. Maricilda Palandi de Mello e Dra. Denise Cavalcante Pontes** pela boa disposição em participar como banca do exame de qualificação, oferecendo considerações valiosas que contribuíram para o meu aprimoramento didático-pedagógico;

Ao **Prof Dr. George Sheperd** por ter me introduzido no aprendizado de técnicas de análise multivariada, o que foi muito importante para interpretação dos resultados;

À **Profa. Dra. Cristine Hackel**, pelo apoio e opiniões verdadeiras;

À **Berenice Rosa** e à **Daniela Cabral**, pela amizade sincera;

Ao amigo e colega **Sérgio Tadeu Sibov**, pelo companheirismo, prestatividade e pelos sucessos e fracassos que passamos juntos no dia a dia do “Laboratório de Análises Genéticas Moleculares”;

Aos colegas **Éberson, Patrícia, Susy, Sandra** e, pela convivência agradável e pelo apoio;

Às amigas **Maria Thereza Marques Novo e Maria Elena Infante** pelo carinho constante e pela amizade;

Aos colegas e amigos **Lyriam, Marília, Juliana, Ana Cláudia, “Ed”, Sandrinha, Henrique, Fabiana, Lee, Marcela, Lucianinha, “Ju” e Ivy** pela convivência agradável e descontraída que atenuaram as dificuldades encontradas;

Aos colegas e amigos do Laboratório de Plantas (**André, Sílvia, Ricardo, Felipe, Edson, Germano, Andreas, Celso e Mano**), pela prestatividade e colaboração;

Ao grande **Márcio**, pelos ensinamentos sobre fotografia;

Aos colegas distantes, porém não esquecidos, **Gonçalo, Cláudia (“trem”), Maria do Carmo, Teresa (“carioca”), Bia, Suzi Lanza, Giu, Valéria, Marymar e Fernando** pelo agradável convívio;

À **Zildinha**, pelo seu jeitinho doce de ser e pela sua prontidão em ajudar;

À **Madá**, pela alegria que contagia à todos e à Isa, pelo carinho;

À **Rosângela e Nair**, por serem sempre solícitas;

À **Sandra, Tânia e Cláudia** da secretaria do CBMEG, pelo empenho em atender a meus pedidos;

À **Teresa** da secretaria da Genética, pela grande disposição em solucionar meus problemas
burocráticos;

Aos **funcionários do CBMEG**, pela atenção dispensada;

Ao **CNPq** pela concessão da bolsa de estudos durante à realização do Curso de Pós-Graduação
no nível de Mestrado;

À **FAPESP** e **FAEP** pelo financiamento dado ao projeto;

À **Cooperação CAPES-COFECUB** pela possibilidade de interagir com pesquisadores da
Universidade de Paris XI, Paris-França;

À todos que direta ou indiretamente contribuíram para a realização deste trabalho.

SUMÁRIO

	Pag.
LISTA DE TABELAS.....	xiii
LISTA DE FIGURAS.....	xvi
ABREVIACES.....	xx
RESUMO.....	xxii
SUMMARY.....	xxiv
1. INTRODUO.....	1
2. REVISO DE LITERATURA.....	5
2.1. Milho (<i>Zea mays</i> L.): a espcie cultivada mais.....	6
cosmopolita do mundo	
2.2. Melhoramento do milho: seleo dos parentais.....	8
2.3. Heterose e diversidade gentica.....	10
2.4. Marcadores bioqumicos e diversidade gentica.....	14
2.5. Marcadores genotpicos e diversidade gentica.....	17
2.5.1. RFLP: princpios e aplicaes da tcnica.....	17
2.5.2. Associao das distncias genticas estimadas.....	20
a partir de RFLPs com a performance de hbridos simples	
2.5.3. RAPD: princpios e aplicaes da tcnica.....	22
2.5.4. Associao das distncias genticas estimadas.....	29

a partir de RAPDs com a performance de híbridos simples	
3. OBJETIVO.....	31
3.1. Gerais.....	32
3.2. Específicos.....	32
MATERIAL E MÉTODOS.....	33
4.1. Material Vegetal.....	34
4.2. Avaliações de campo.....	35
4.3. Métodos.....	36
4.3.1. Coleta e preparo do material vegetal.....	36
4.3.2. Extração do DNA genômico.....	38
4.3.3. Quantificação do DNA.....	39
4.3.4. Condições de Amplificação.....	40
4.3.5. Análise dos dados moleculares.....	41
4.3.6. Coeficiente de correlação cofenética ($r_{\text{cof.}}$).....	43
4.3.7. Análise multivariada.....	44
4.3.8. Correlação dos dados moleculares (genéticos).....	46
com os dados de performance híbrida	
5. RESULTADOS.....	47
5.1. Análise molecular via RAPD.....	48
5.1.1. Otimização das condições de amplificação.....	48
5.1.2. Seleção de “primers”.....	51
5.1.3. Produtos do RAPD das 18 linhagens.....	54

5.1.4. Correlação entre distância genética e.....	76
performance do híbrido	
6. DISCUSSÃO.....	82
7. CONCLUSÕES.....	95
8. BIBLIOGRAFIA.....	99
9. ANEXOS.....	128

LISTA DE TABELAS

	Pag.
1. Linhagens endogâmicas de milho (<i>Zea mays</i> L.), seus códigos, nomenclatura e população de origem.....	37
2. “Primers” utilizados no RAPD das 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.).	55
3. Coeficientes de correlações cofenéticas para um número crescente de bandas polimórficas. Dois coeficientes de similaridade (Jaccard e Dice) foram utilizados.....	59
4. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.) obtidas para 262 bandas polimórficas.....	68
5. Médias, mínimos, máximos e desvios-padrões das distâncias genéticas de cada par de linhagens calculados a partir de RAPD com 262 bandas polimórficas para as 18 linhagens endogâmicas de milho.....	70
6. Produção de grãos para os híbridos F1 e heterose para cruzamentos inter- e intrapopulacionais. GDs indicam as distâncias genéticas dos parentais.....	77

7. Correlações lineares (r) encontradas entre distâncias obtidas por RAPD e dados de campo para três caracteres de importância econômica e para suas respectivas heteroses.....	78
Anexo A. Alguns coeficientes de semelhança para variáveis dicotômicas.....	129
Anexo B. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 03 linhagens endogâmicas de milho (<i>Zea mays</i> L.) escolhidas aleatoriamente.....	130
Anexo C. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.) obtidas para 50 bandas polimórficas....	131
Anexo D. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.) obtidas para 98 bandas polimórficas....	132
Anexo E. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.) obtidas para 150 bandas polimórficas..	133
Anexo F. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.) obtidas para 190 bandas polimórficas....	134
Anexo G. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.) obtidas para 215 bandas polimórficas....	135

Anexo H. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (*Zea mays* L.) obtidas para 300 bandas polimórficas.... **136**

Anexo I. Matriz original de dados para as 18 linhagens endogâmicas gerada com 262 bandas polimórficas. A leitura se deu por presença (1) ou ausência (0)..... **137**

LISTA DE FIGURAS

Pag.

- 1A-B-C:** (A) Efeito da quantidade de DNA (10, 20, 30 40 e 50 ng) sobre o perfil de RAPD da linhagem 01 com o “primer” OPB-01 e com 2,0 unidades de enzima; (B) Efeito da quantidade de enzima (2,0 e 3,0 U) sobre o perfil de RAPD da linhagem 01 com o “primer” OPB-01 (GTTTCGCTCC) e 30 ng de DNA; (C)Efeito da quantidade de magnésio (2,5; 3,0; 4,0 e 4,5 mM) sobre o perfil de RAPD da linhagens 01 com o “primer” OPB-01 e 25 ng de DNA. (M) marcador de peso molecular λ HindIII/ ϕ x174Hae III; C é o controle sem DNA..... **50**
- 2A-B:** Perfil de RAPD para 03 linhagens S3 de milho (01, 11 e 17, respectivamente) com dois “primers” para fins de seleção: “primer” OPC-07 (GTCCCGACGA) e OPC-08 (TGGACCGCTG); (B) Perfil de RAPD para as mesmas 03 linhagens descritas acima com dois outros “primers” para fins de seleção: “primer” OPE-01 (CCCAAGGTCC) e OPE-02 (GGTGCGGGAA); (a) e (b) representam duas repetições com 25 ng de DNA cada uma e (M) é marcador de peso molecular λ HindIII/ ϕ x174Hae III..... **52**
- 3:** Dendrograma de 03 linhagens endogâmicas de milho (*Zea mays* L.) usadas para fins de seleção de “primers”. As distâncias genéticas foram obtidas por análise de 253 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao

posicionamento das linhagens nos grupos GI (verde) e GIII (azul) no dendrograma final (Figura 12).....	53
4A-B: Perfil de RAPD para as 18 linhagens endogâmicas de milho com os seguintes “primers”: OPB-09 (TGGGGGACTC) e (B) OPB-11 (GTAGACCCGT). (M) é marcador de peso molecular λ HindIII/ ϕ x174Hae III.....	56
5A-B: Perfil de RAPD para as 18 linhagens endogâmicas de milho com os seguintes “primers”: OPB-12 (CCTTGACGCA) e OPB-17 (AGGGAACGAG). (M) é marcador de peso molecular λ HindIII/ ϕ x174Hae III.....	57
6A-B: Perfil de RAPD para as 18 linhagens endogâmicas de milho com os seguintes “primers”: OPB-18 (CCACAGCAGT) e (B) OPC-07 (GTCCCGACGA). (M) é marcador de peso molecular λ HindIII/ ϕ x174Hae III.....	58
7: Dendrograma das 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.). As distâncias genéticas foram obtidas pela análise de 50 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12).....	63
8: Dendrograma das 18 linhagens endogâmicas de milho (<i>Zea mays</i> L.).As distâncias genéticas foram obtidas pela análise de 98 bandas polimórficas. Os coeficientes de similaridade foram	

- gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12)..... **64**
- 9:** Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 150 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12)..... **65**
- 10:** Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 190 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) do dendrograma final (Figura 12)..... **66**
- 11:** Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 215 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12)..... **67**

- 12:** Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 262 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12)..... **69**
- 13:** Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 300 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12)..... **72**
- 14:** Associações entre as 18 linhagens de milho (*Zea mays* L.) reveladas por análise de coordenadas principais (PCO); baseadas nos coeficientes de similaridade obtidos para 262 bandas polimórficas por RAPD. As cores seguiram o padrão estabelecido na Figura 12: GI (verde), GII (vermelho) e GIII (azul)..... **75**
- 15A-B:** Gráficos de produção de grãos (t/ha) *versus* distâncias genéticas (baseadas nos dados de RAPD com 262 bandas polimórficas) para: (A) 80 híbridos simples interpopulacionais e (B) GI x GII..... **89**
- 15C- D:** Gráficos de produção de grãos (t/ha) *versus* distâncias genéticas (baseadas nos dados de RAPD com 262 bandas polimórficas) para: (C) GI x GIII e (D) GII x GIII..... **80**

ABREVIACES

cm/planta	Centmetro por planta
°C	Graus Celsius
5'→3'	Sentido da fita paralela de DNA
dNTPs	Desoxiborribonucleotdeos trifosfatados
EDTA	cido etileno diaminotetractico
F1	Prognie da primeira gerao de cruzamentos
g	Grama
GD	Distncia gentica
GS	Similaridade gentica
Hg	Mercrio
l	Litro
M	Molaridade
ml	Mililitro
mM	Milimolar
µM	Micromolar
µl	Micro litro
ng	Nanograma
pb	Pares de bases
PCR	Reao de polimerase em cadeia

“Primers”	Oligonucleotídeos
TBE	Ácido bórico/Tris/EDTA
TE	Tris/EDTA
t/ha	Toneladas por hectare
Tris	Tris-(hidroximetil)-aminometano
rpm	Rotações por minuto
U	Unidades de enzima
UV	Ultra-violeta
v/cm	Volts por centímetro

RESUMO

A diversidade genética entre as 18 linhagens endogâmicas de milho foi estimada através da técnica de RAPD (“Random Amplified Polymorphic DNA”) com o objetivo de correlacionar estas medidas de distância genética com as performances dos respectivos híbridos simples.

Oito linhagens obtidas da população BR-105 e dez da população BR-106 foram utilizados como material básico. Estas linhagens foram cruzadas em um esquema dialélico a nível intra- e interpopulacional, e os híbridos simples obtidos foram avaliados em experimentos com repetições. DNA foi extraído de folhas de cada linhagem e usado como molde para a reação de amplificação. As reações foram realizadas em um volume de 25µl e otimizadas para 25 ng de DNA, 3,0 mM de cloreto de magnésio e 2,0 unidades de enzima *Taq*-DNA polimerase (Gibco-BRL). Trinta e dois “primers” diferentes foram usados resultando em 325 produtos de amplificação reprodutíveis, dos quais 262 foram polimórficos.

A distância genética foi determinada usando o coeficiente de Jaccard e cerca de 80,6% de polimorfismo foi detectado. O dendrograma foi construído usando o método de agrupamento de UPGMA (“Unweighted Pair-Group Method with Arithmetical Averages”). A análise do dendrograma final dividiu as amostras em três grupos distintos (GI, GII e GIII) com distâncias genéticas médias de $0,73 \pm 0,06$, $0,52 \pm 0,06$ e $0,46 \pm 0,13$, respectivamente. A análise de coordenadas principais (multivariada) confirmou a divisão das linhagens em três grupos. As estimativas de distâncias genéticas foram correlacionadas com caracteres de interesse agrônômico (produção de grãos, altura de planta e de espiga) de híbridos simples e

com a heterose. As estimativas das correlações, quando a divisão de grupos foi desconsiderada, foram próximas de zero, mas correlações significativas foram detectadas entre as distâncias genéticas envolvendo os grupos GI x GII ($r = 0,70$) e GI x GIII ($r = 0,87$) com os dados correspondentes de produção de grãos. Uma inadequação aos padrões heteróticos previamente estabelecidos foi encontrada já que a população BR-106, de ampla base genética, pôde ser subdividida em dois grupos diferentes (pertencentes à GI e GIII) e a população BR-105, de base genética mais restrita, teve a maioria das linhagens avaliadas alocadas num outro grupo heterótico (GII), enquanto apenas uma de suas linhagens integrou-se ao GI.

Os resultados obtidos indicaram que a técnica de RAPD pode ser utilizada para alocar de forma correta os genótipos a diferentes grupos heteróticos e, também, para direcionar os cruzamentos envolvendo linhagens de grupos heteróticos mais distintos.

SUMMARY

Genetic diversity among the 18 maize S3 inbred lines was determined by using RAPD (Random Amplified Polymorphic DNA) in order to correlate them with the performance of the respective single-crosses.

Eight inbred lines from BR-105 population and ten derived from BR-106 population were used as base material. The inbreds were crossed following a diallel mating design for intra- and interpopulation levels, and the single-crosses were evaluated in replicated experiments. Genomic DNA was isolated from leaves of each inbred line and used in amplification reactions. Reactions were performed in a total volume of 25 μ l with 25 ng of DNA, 3.0 mM of magnesium chloride and 2.0 units of *Taq*-DNA polymerase enzyme (Gibco-BRL). Thirty-two different primers were used giving a total of 325 reproducible amplification products, 262 of them being polymorphic.

Genetic divergence was determined using Jaccard's similarity coefficient and around 80.6% of polymorphism was detected. A final dendrogram was constructed using an unweighted pair-group method with arithmetical averages (UPGMA). Cluster analysis divided the samples into three distinct groups (GI, GII, GIII) with average genetic distances of 0.73 ± 0.06 , 0.52 ± 0.06 e 0.46 ± 0.13 , respectively. The group division was confirmed by principal coordinate analysis. The genetic distances were correlated with important agronomic traits (grain yield, plant and ear height) and with heterosis. No correlation was found when group division was unregarded, but significant correlations were detected between GI x GII and GI x GIII genetic distances with their respective single-crosses' yielding. A non-correspondence to the established heterotic groups were found as BR-106, a broad genetic base population, was

devided in two different groups (belonging to GI and GII) and BR-105, a narrow genetic base population, remained mostly grouped in GIII, except from one inbred line allocated to GI.

The results indicated that RAPD is a powerful tool for allocating genotypes into different heterotic groups, and to guide crosses with distinct lines.

INTRODUÇÃO

1. INTRODUÇÃO

Nas espécies alógamas como o milho (*Zea mays* L.), a obtenção de cultivares híbridos baseia-se no processo de endogamia-hibridação, o qual permite identificar, selecionar e reproduzir indefinidamente o(s) genótipo(s) superior(es) de uma população. Desta forma, plantas de uma população são autofecundadas (em geral, de 6-7 gerações) até um nível de homozigose elevado (99%), obtendo-se assim linhagens homozigóticas ou puras. A seguir, estas linhagens são cruzadas com linhagens de outra população, ou com linhagens elites já existentes, gerando híbridos. Estes híbridos são avaliados em experimentos com repetições em diversos locais e anos, de forma a permitir a identificação do(s) híbrido(s) superior(es).

A principal dificuldade de um programa de melhoramento desta natureza reside na avaliação de linhagens, devido a grande quantidade de esforços e recursos que são alocados nesta fase. Tal fato deve-se à necessidade de instalação de experimentos em extensas áreas, em diferentes locais e épocas de plantio, além da subsequente tomada de dados e processamento estatístico dos mesmos.

Dados experimentais acumulados desde os estudos iniciais de endogamia e produção de híbridos de milho, têm demonstrado que a heterose para a produção de grãos e outras características é função da heterozigosidade de um grande número de locos. Assim, existe uma correlação positiva entre a diversidade genética entre parentais com a produção de grãos do F1 e o nível de heterose em milho (Moll *et al.*, 1962, 1965; Paterniani & Lonnquist, 1963).

Desta forma, medidas indiretas da diversidade genética entre linhagens e populações são importantes para a identificação, reconhecimento e exploração de padrões

heteróticos existentes no germoplasma disponível. Neste sentido, vários critérios ou conjunto de caracteres têm sido utilizados. Medidas da distância genética utilizando-se marcadores morfológicos (Smith & Smith, 1989) e citológicos (Moll *et al.*, 1972; Chughtai & Steffensen, 1987) têm se mostrado pouco eficazes quanto a capacidade de discriminar as diferenças existentes. Marcadores bioquímicos moleculares, tais como isoenzimas e proteínas de reserva (zeínas), também foram utilizados; entretanto, a diversidade genética avaliada por estes marcadores apresentou-se limitada pelo pequeno número de locos polimórficos disponíveis (Price *et al.*, 1986; Lamkey *et al.*, 1987; Smith & Smith, 1989). Com o advento de marcadores moleculares ao nível de DNA, como os RFLPs (“Restriction Fragment Length Polymorphism”) e RAPDs (“Random Amplified Polymorphic DNA”), pôde-se acessar um amplo grau de polimorfismo (Burr *et al.*, 1983; Helentjaris *et al.*, 1986; Helentjaris, 1987).

Marcadores moleculares do tipo RFLPs têm sido amplamente utilizados para avaliar a variabilidade disponível em germoplasmas de diferentes espécies (Helentjaris *et al.*, 1988; Evola *et al.*, 1986; Landry *et al.*, 1987). Em programas de melhoramento de milho, tem-se avaliado a correlação das distâncias genéticas entre linhagens de milho, obtidas a partir de RFLP, com a performance dos respectivos híbridos simples. A existência desta correlação possibilitaria a predição do comportamento das linhagens em híbridos simples, através de análises no laboratório, antes das avaliações no campo. Desta forma, seria possível diminuir o tempo e a quantidade de recursos e esforços alocados na fase de avaliação dos híbridos (Lee *et al.*, 1989; Smith *et al.*, 1990; Melchinger *et al.*, 1991; Messmer *et al.*, 1991). Os resultados da correlação entre as distâncias genéticas das linhagens baseadas em RFLPs e a performance dos respectivos híbridos simples no campo são contraditórios, demonstrando a necessidade de realização de estudos adicionais (Godshalk *et al.*, 1990; Melchinger *et al.*, 1990a,b).

A técnica de RAPD (Williams *et al.*, 1990; Welsh & McClelland, 1990) é hoje uma ferramenta poderosa para acessar a diversidade genética. Devido às suas características, os marcadores moleculares do tipo RAPD expõem elevada parcela do polimorfismo entre os indivíduos, sendo que a técnica necessária para sua utilização é relativamente simples e rápida. Desta forma, estes marcadores têm sido amplamente utilizados para os mais diferentes organismos como procariotos (Welsh & McClelland, 1992) e eucariotos (Torres. *et al.*, 1993; Heun & Helentjaris, 1993; Yang & Quiros, 1993).

Marcadores moleculares do tipo RAPD exploram o polimorfismo entre os indivíduos a partir de um número grande de regiões no genoma, sendo estas estritamente aleatórias. Portanto, tanto regiões codificadoras como não codificadoras podem ser avaliadas. Assim, espera-se que as distâncias genéticas entre linhagens obtidas a partir deste tipo de marcador apresentem maior correlação com a performance dos respectivos híbridos simples.

Neste âmbito, o presente trabalho se propôs a utilizar marcadores moleculares do tipo RAPD para avaliar a diversidade genética entre linhagens endogâmicas de milho (*Zea mays* L) com o objetivo de correlacionar tais medidas com dados de performance de híbridos simples. Havendo correlações significativas entre os dados moleculares e os dados de campo, estas medidas poderiam servir para direcionar ou até mesmo predizer novos cruzamentos entre os mesmos grupos de linhagens.

REVISÃO DE LITERATURA

2. REVISÃO DE LITERATURA

2.1. Milho (*Zea mays* L.): a espécie cultivada mais cosmopolita do mundo

As origens do milho datam provavelmente do surgimento da agricultura moderna no Novo Mundo, cerca de 8.000 anos atrás. Acredita-se que ele tenha como ancestral uma gramínea selvagem, originária do México e da Guatemala, denominada Teosinte [*Z. mexicana* (Schrad.) Kuntze]; portanto, o centro de origem do milho situa-se também nesta região (Galinat, 1988). Quando os europeus chegaram ao Novo Mundo, a divergência entre milho e teosinte já devia estar tão avançada que as duas espécies foram classificadas inicialmente pelos taxonomistas em gêneros separados. Posteriormente, mediante comparações citogenéticas, o erro foi retificado, sendo que o milho e o teosinte pertencem atualmente ao mesmo gênero, denominado *Zea* (Reevis & Mangelsdorf, 1942).

O milho é uma monocotiledônea pertencente à família Gramineae e tribo Andropogoneae, juntamente com o sorgo, o *Coix* (“lágrima de Cristo”) e o *Tripsacum* (Clayton, 1973; Jain & Benerjee, 1974; Clayton, 1983). É uma espécie alógama, monóica, sendo que sua morfologia resulta da supressão, condensação e multiplicação de várias partes da anatomia básica das gramíneas (Anderson, 1945; Alava, 1952; Bonnet, 1954). O número básico de cromossomos do milho, $x = 10$, é igual ao do teosinte (Kuwada, 1911).

Geneticamente, o milho é a espécie mais acessível e melhor caracterizada dentre as plantas superiores, tanto em estudos básicos de genética e citogenética, como em estudos visando o entendimento e manipulação de seu germoplasma.

Desde a diferenciação entre milho e teosinte, a diversidade do milho evoluiu para cerca de 300 raças identificadas até o presente (Paterniani, 1990) e milhares de variedades, sendo portanto, uma espécie altamente politépica. O milho é encontrado nos mais diversos ambientes, havendo genótipos cultivados em diferentes latitudes. Existem também raças e variedades adaptadas às condições ecológicas específicas como, por exemplo, baixas (0 a 100 m), médias (1000 a 2000 m) e altas altitudes (acima de 2000 m). Há igualmente tipos variantes de endospermas, que foram selecionados durante sua domesticação com o objetivo de atender necessidades específicas, tais como endosperma dentado, duro ou cristalino, amiláceo, pipoca e doce. Dentro de cada tipo, existem ainda variações quanto à coloração e certos componentes químicos. Assim, uma enorme variabilidade está presente dentro e entre linhagens, variedades e raças de milho. Apesar de apenas parte desta variabilidade ter sido examinada sistematicamente, vários fatores genéticos já foram identificados, determinando a cor e a morfologia das estruturas constituintes de cada parte ou tecido da planta. Enquanto uma observação casual poderia sugerir que esta variação seria controlada através de efeitos sobrepostos de alguns fatores, que estariam sendo expressos em pequenos incrementos, uma análise genética mais direta tem estabelecido mais de 100 locos condicionando variações morfológicas geneticamente simples na estatura da planta, bem como na forma e estrutura das partes vegetativas e reprodutivas (Coe *et al.*, 1988).

No cenário da produção mundial de grãos, o milho situa-se em segundo lugar, sendo superado apenas pelo trigo. Os Estados Unidos são atualmente os maiores produtores de milho e cultivam essencialmente um só tipo de endosperma, o dentado amarelo, e o doce em

segundo lugar. No Brasil, com o advento do milho híbrido, popularizou-se o endosperma semi-dentado amarelo, o qual representa a maior parcela do milho cultivado, embora o milho dentado amarelo também seja cultivado. Outros tipos de endosperma são cultivados em menor escala. O milho tem grande importância econômica no Brasil e ocupa o primeiro lugar em área cultivada, desconsiderando-se as pastagens, sendo um cereal de múltiplas aplicações, de consumo elástico, e que proporciona o desenvolvimento de indústrias, gerando grande número de empregos diretos e indiretos (Paterniani, 1995).

2.2. Melhoramento do milho: Seleção dos parentais

O melhoramento do milho já era realizado pelos índios americanos há milhares de anos, embora de forma bastante primitiva, baseado apenas na seleção de características de interesse. A partir do final do século XVIII teve início o melhoramento empregando-se algum tipo de metodologia. As primeiras variedades de milho foram obtidas através de seleção massal e, posteriormente, através de seleção com teste de progênie. Após os trabalhos pioneiros de Shull (1908), East (1908) e Jones (1918), teve início o melhoramento de milho visando o desenvolvimento de híbridos, primeiramente, através de programas experimentais de melhoramento em várias estações agrícolas norte-americanas. O melhoramento do milho como prática comercial privada só ocorreu anos depois. O desenvolvimento de linhagens endogâmicas em outras partes do mundo deu-se rapidamente após a Segunda Guerra Mundial (Hallauer *et al.*, 1988).

Com o advento dos híbridos simples comerciais, tornou-se emergente a necessidade de melhorar as linhagens endogâmicas parentais. Vários são os caracteres de importância econômica a serem melhorados nas linhagens, tais como resistência às doenças e insetos, resistência à seca, aumento na produção de grãos, melhoria da qualidade nutricional da semente, entre outros.

Os métodos de seleção de parentais podem ser divididos em duas categorias: (a) *a priori* ou avaliação direta dos parentais e (b) *a posteriori* ou por algum tipo de teste de progênie (Baenziger & Paterson, 1992). Os programas de melhoramento de espécies anuais têm se baseado predominantemente na última categoria, especialmente para o desenvolvimento de híbridos F1. Os métodos *a priori* têm sido utilizados mais freqüentemente para avaliação de caracteres herdáveis simples, por melhoristas experientes, empregando-se populações sob melhoramento ou germoplasma exótico.

Apesar de investigações teóricas e práticas, o desenvolvimento de métodos para identificar, escolher e usar parentais capazes de produzir híbridos superiores tem sido uma “luta constante”. No caso do milho, não é possível prever as performances dos híbridos a partir das linhagens parentais, isto é, a correlação entre linhagens e híbridos é muito baixa para se ter algum valor preditivo, principalmente para a produção de grãos (Hallauer *et al.*, 1988).

2.3. Heterose e a diversidade genética

O fenômeno do vigor híbrido em milho, foi primeiramente descrito por Darwin no século XIX. Resultados dos estudos de autofecundação de milho realizados por Shull (1908), East (1908) e Jones (1918) fizeram com que os melhoristas despertassem para o fato de poder haver uma relação entre produção de grãos ou heterose para esta característica e diversidade genética.

A heterose em cruzamentos entre cultivares e linhagens endogâmicas tem sido medida e predita através do emprego de metodologias genético-estatísticas, mas sua base genética específica não foi completamente elucidada. Visões divergentes à respeito do fenômeno levaram a extensivas pesquisas relativas à métodos de melhoramento, de desenvolvimento de linhagens e de escolha de testadores, além de outros métodos de seleção. Duas premissas confrontantes, tentando explicar a base do fenômeno, atribuíram suas causas como sendo devido à efeitos primários de sobredominância ou à efeitos de dominância parcial/completa.

Em milho, há fortes evidências de que a heterose nesta espécie é resultante do acúmulo de locos com alelos expressando dominância parcial/completa (Hallauer *et al.*, 1988). Smith & Smith (1989) relataram, para milho, que a heterose seria uma função da heterozigosidade em um grande número de locos.

Beal (1877) relatou o primeiro dado de produção em cruzamentos controlados de cultivares de milho. Depois destes experimentos, a performance dos cruzamentos entre cultivares passou a ser comparada à performance dos cultivares parentais. Richey (1922)

analisou dados de 244 cruzamentos de milho e relatou que 86,5% estavam acima da média dos parentais.

Estudos realizados por Moll *et al.* (1962, 1965) e Paterniani & Lonnquist (1963) mostraram uma correlação positiva entre divergência genética (inferida em função das relações ancestrais, separação geográfica e adaptação à diferentes ambientes), produção de grãos do F1 e heterose para esta característica em milho. Estes estudos indicaram que a heterose, medida entre cruzamentos de populações e variedades de diferentes regiões geográficas, era maior a medida em que a divergência genética entre as populações aumentava. Moll *et al.* (1962) relataram que a heterose expressa numa variedade de cruzamentos concordava com a classificação original dos parentais, baseada na divergência genética presumida a partir das distâncias geográficas. Mas este conceito de divergência genética também parecia ter limites relativos a um máximo de expressão da heterose. Moll *et al.* (1965) descobriram que a heterose expressa nos cruzamentos entre as variedades mais geneticamente diversas era menor do que a heterose entre as consideradas geneticamente menos diversas. Esta redução na heterose foi atribuída a incompatibilidade de combinações gênicas que poderiam resultar do cruzamento de duas populações parentais altamente divergentes (Moll *et al.*, 1965)

Melhoristas que desenvolvem híbridos têm usado predominantemente o conceito de grupos heteróticos na seleção de parentais visando estabelecer combinações híbridas superiores. Um grupo heterótico é uma coleção de germoplasma que, quando cruzado com um germoplasma externo ao seu grupo, tende a exibir um maior grau de heterose (em média) do que quando cruzado com um membro de seu próprio grupo (Lee, 1995). É claro que exceções importantes à esta tendência têm sido observadas, mas o conceito tem fornecido uma ferramenta simples e conveniente para o manejo de germoplasma, e tem sido aplicada na manutenção de

germoplasma e criação de populações usadas na identificação das melhores combinações híbridas, e, também, na classificação de germoplasma exótico.

O estabelecimento de grupos heteróticos em milho tem sido um processo laborioso, requerendo décadas para evoluir nos programas de melhoramento norte-americanos. Sem nenhum benefício da sistemática dos métodos de melhoramento, várias combinações de parentais foram testadas até que “pedigrees” dominantes emergiram, estabelecendo uma identidade para um grupo heterótico (Hallauer *et al.*, 1988). Divergências de opiniões surgiram em estudos recomendando cruzamentos de cultivares norte-americanos com diferentes tipos de endospermas (Richey, 1922), enquanto outros estudos mostravam que a heterose expressa era maior quando cultivares com tipos similares de endospermas eram cruzados (Paterniani & Lonquist, 1963). A escolha de qual padrão heterótico deve ser usado depende da frequência de alta performance dos híbridos, derivados de linhagens oriundas das respectivas populações de base. Neste sentido, espera-se que as linhagens obtidas de populações com padrões heteróticos distintos, tendam a se complementar até um máximo de performance híbrida. No Brasil, cruzamentos entre linhagens de origem dentada e dura são amplamente utilizados com esta finalidade (Hallauer *et al.*, 1988).

A obtenção de híbridos tem sido o principal objetivo de programas de melhoramento, convencidos dos méritos da heterose, uniformidade e estabilidade dos híbridos. Para algumas espécies, tais como as produtoras de cacau, borracha e óleo de palmáceas, um dos principais problemas tem sido o tempo e espaço para testar o grande número de combinações híbridas e a falta de registros de “pedigree” e de performance que ajudariam a diminuir o possível número de combinações a serem testadas (Lee, 1995).

Para milho, o único meio de produção de híbridos para uso comercial é o cruzamento de linhagens endogâmicas. Diferentes esquemas de cruzamentos estão disponíveis para aumentar a homozigosidade das linhagens, mas a mais rápida e intensa forma de endocruzamento é através da autofecundação (Hallauer, 1990).

Os melhoristas têm utilizado vários métodos para acessar a diversidade genética e estabelecer relações entre germoplasma elite, sobretudo com a finalidade de aumentar o ganho genético. Eles têm tido acesso, geralmente, à informações de genealogia, dados de performance (capacidade de combinação, avaliação da progênie, seleção, etc.) e também por meio de inferências feitas através de vários delineamentos experimentais (Dudley, 1987; Troyer *et al.*, 1988).

O coeficiente de Malècot (1948) ou de coancestralidade (f), o qual fornece uma estimativa do grau de similaridade genética entre dois indivíduos, é a probabilidade de que um alelo de um indivíduo X seja idêntico, por descendência, a um alelo de um indivíduo Y, sendo uma probabilidade média baseada em relações de “pedigree” (Dudley, 1994). Esta medida tem sido vastamente utilizada para estimar níveis de diversidade genética, bem como de relacionamento genético entre plantas autógamas como aveia (Rodgers *et al.*, 1983), trigo (Cox *et al.*, 1985b) e soja (Cox *et al.*, 1985a), entre outras.

Para plantas alógamas, incluindo o milho, coeficientes de coancestralidade não têm sido facilmente determinados, já que os dados de “pedigrees” das linhagens são muitas vezes escassos ou ainda imprecisos, especialmente quando a seleção é feita a partir de populações com ampla base genética. Além disto, estimativas das relações genéticas baseadas neste coeficiente podem ser inadequadas devido à simplificações do modelo, tais como: (1) ausência de seleção, mutação, migração e deriva; (2) meiose regular diplóide e (3) nenhum

parentesco entre os indivíduos ($f = 0$) a não ser os ancestrais comuns observados (Melchinger, 1993). No entanto, vários aspectos comuns aos programas de melhoramento fogem a estas premissas: (1) intensa seleção; (2) deriva devido ao pequeno tamanho de amostra; (3) meiose irregular em algumas culturas e (4) registros desconhecidos ou incorretos de “pedigree” (Bernardo, 1993; Messmer *et al.*, 1993).

2.4. Marcadores bioquímicos e diversidade genética

Dados morfológicos têm sido tradicionalmente utilizados como descritores de variedades de plantas e para a distinção de cultivares. Entretanto, os caracteres morfológicos, muitas vezes, não retratam fielmente as relações genéticas existentes devido às suas interações ambientais e ao grande desconhecimento do controle genético dos caracteres avaliados (Smith & Smith, 1989). Além disto, a classificação de materiais em grupos heteróticos com base neste grupo de descritores tem sido reconhecida como inadequada já que alguns tipos de endosperma diferem em apenas um único gene (Coe *et al.*, 1988).

Com o advento das técnicas de eletroforese e de HPLC-fase reversa (“reversed-phase High Performance Liquid Chromatography”), no início da década de 60, surgiu uma nova classe de marcadores bioquímicos revelados através da análise de proteínas: as variantes isoenzimáticas. Elas logo foram utilizadas em estudos para caracterizar a variação genética de várias espécies, inclusive milho (Stuber & Moll, 1972; Stuber *et al.*, 1980). Esta nova classe de marcadores representou um avanço frente aos descritores morfológicos e citológicos, já que ela

não é tão profundamente afetada pela interação com o ambiente, como observado nos casos anteriores.

Em milho, isoenzimas têm sido extensivamente utilizadas para caracterizar a variação genética existente entre linhagens endogâmicas (Stuber & Goodman, 1983; Smith *et al.*, 1985a,b), híbridos comerciais (Smith, 1984, 1986), populações exóticas e de polinização aberta (Khaler *et al.*, 1986; Smith, 1986) e coleções de germoplasma (Goodman & Stuber, 1983; Doble *et al.*, 1983).

Vários estudos tentaram relacionar dados de divergência genética entre linhagens endogâmicas de milho, medidos através de isoenzimas, com a performance dos respectivos híbridos simples. Em experimentos dialélicos envolvendo linhagens selecionadas e 8 a 15 locos isoenzimáticos, a correlação entre o número de locos isoenzimáticos heterozigotos e a produção de híbridos ou a capacidade específica de combinação (CEC) dos mesmos foram positivas, mas muito pequenas para serem preditivas (Hunter & Kannenberg, 1971; Heidrich-Sobrinho & Cordeiro, 1975; Gonella & Peterson, 1978; Hadjinov *et al.*, 1982). Price *et al.* (1986) relacionaram a produção de grãos de híbridos simples com a heterozigosidade em 13 locos isoenzimáticos polimórficos. Os autores não encontraram uma associação entre performance de cruzamento e número de locos isoenzimáticos heterozigóticos. Lamkey *et al.* (1987) também encontraram que as diferenças alélicas dos locos isoenzimáticos não eram preditivas da performance dos cruzamentos. Frei *et al.* (1986) relacionaram dados de isoenzimas para linhagens com “pedigrees” similares e diversos e encontraram uma associação entre a dissimilaridade de dados isoenzimáticos com alta produção. No entanto, o valor preditivo das linhagens avaliadas ficou restrito a linhagens com “pedigrees” similares e linhagens específicas mostraram desvios do padrão geral.

Tersac *et al.*(1994) avaliaram nove sistemas isoenzimáticos de 39 populações de girassóis (*Helianthus annuus* L.) originárias de dez países diferentes. Nestas populações, analisou-se também a capacidade de combinação com quatro linhagens testadoras, para produção, umidade e conteúdo de óleo das sementes. Quatro sistemas isoenzimáticos dentre os demais testados, geraram a melhor discriminação da capacidade específica de combinação com as quatro linhagens testadoras. Dois destes sistemas (malato desidrogenase e glutamato oxaloacetato transaminase) agruparam as populações de acordo com a origem geográfica, sendo consistentes com dados de heterose e capazes de predizer a qual grupo heterótico cada população pertence.

Existem ressalvas quanto ao uso de descritores bioquímicos na avaliação das relações filogenéticas entre linhagens de milho. Dados isoenzimáticos coletados para 21 locos de 72 híbridos de milho norte-americanos revelaram associações que foram incongruentes com os seus grupos heteróticos (Smith *et al.*, 1985a). A partir de dados isoenzimáticos e cromatográficos, Smith & Smith (1987, 1988) compararam análises multivariadas e agrupamentos e foram capazes de distinguir subgrupos de linhagens de milho relacionadas por “pedigree”. Entretanto, os dados isoenzimáticos foram piores estimadores das relações genéticas entre linhagens elites de milho do que medidas de distâncias baseadas em “pedigree” e heterose (Smith & Smith, 1989).

Uma das maiores limitações dos estudos que se utilizam de isoenzimas reside no número limitado de locos examinados e no baixo nível de polimorfismo existente (Melchinger *et al.*, 1990a). Desta forma, não há uma boa amostragem da variação genética. Este fato é característica intrínseca do marcador, uma vez que a eletroforese de proteínas apenas permite o reconhecimento de polimorfismos de regiões expressas, as quais representam uma pequena

parte do genoma, na maioria das espécies (Marshall & Brown, 1975; Edwards *et al.*, 1987). Além disto, um outro problema envolvendo esta classe de marcadores é a possibilidade de que um único gene codifique diferentes proteínas (policistrônico) ou de que uma proteína seja formada a partir de porções codificadas por diferentes genes, as quais são unidas num único produto final. De forma geral, os marcadores bioquímicos, juntamente com os moleculares, são presumivelmente neutros; entretanto, casos de não-neutralidade foram relatados, como por exemplo o da *Adh* em *Drosophila* (Anderson & MacDonald, 1983). Assim, ao que parece, o uso deste tipo de marcadores para fins preditivos parece um tanto quanto controverso.

2.5. Marcadores genotípicos e diversidade genética

2.5.1. RFLP: princípios e aplicações da técnica

A variação na sequência de DNA é a base da diversidade dentro de uma espécie. Em muitos casos, a variação da sequência é fenotipicamente neutra, pois grande parte das mutações que alteram a sequência de nucleotídeos somente é mantida durante a evolução porque são mutações silenciosas, ou seja, não causam nenhum efeito no indivíduo. O número destas variações é elevado, podendo chegar a 1 em cada 100 nucleotídeos. Este tipo de variação não está sujeito à influência do ambiente, representando assim uma excelente fonte de polimorfismo a ser explorada.

O desenvolvimento de técnicas para o estudo do DNA provocou uma verdadeira revolução no campo da biossistemática, da evolução e da genética e melhoramento de plantas e

animais. A grande sensibilidade destas técnicas associada à descoberta de regiões com alto grau de variabilidade tem permitido avanços significativos para estudos de polimorfismos em populações humanas com aplicações na clínica forense e na determinação de paternidade.

Marcadores moleculares do tipo RFLP (“Restriction Fragment Length Polymorphism”) apresentam herança mendeliana, sendo em sua maioria codominantes. A detecção do polimorfismo entre indivíduos é baseada na diferença do tamanho do fragmento de restrição, observada em autoradiogramas, após hibridização de uma membrana (Southern, 1975) contendo o DNA digerido por uma dada enzima, com uma sonda específica. Assim, todo marcador molecular do tipo RFLP é caracterizado por uma associação específica sonda/enzima (Souza, 1995).

A medida que enzimas de restrição cortam o DNA em sítios específicos, mutações pontuais dentro de um sítio resultam na perda ou ganho de um sítio de reconhecimento, gerando, nesta região, fragmentos de restrição de tamanhos diferentes. Mutações causadas por inserções, deleções ou inversões de segmentos de DNA também podem gerar variação de comprimento nos fragmentos de restrição (Gebhardt & Salamini, 1992).

A variação na sequência de DNA detectada por este método foi determinada de RFLP por Botstein *et al.* (1980). O fenômeno foi primeiramente descrito para linhagens mutantes de adenovírus (Grodzicker *et al.*, 1974) e sua herança mendeliana foi demonstrada para DNA de levedura por Peter & Botstein (1977). Entretanto, foi com estudos de genes de globinas humanas que o potencial dos RFLPs como marcadores de diagnóstico se tornou evidente, através da correlação entre uma mutação de um alelo específico de β -globina e a presença de certos fragmentos de RFLPs (Jeffreys, 1979).

Botstein *et al.* (1980) sugeriram o uso desta classe de marcadores genéticos para a construção de mapas de ligação altamente saturados do genoma humano. O potencial desta estratégia para o melhoramento de plantas foi reconhecido poucos anos mais tarde (Beckmann & Soller, 1983; Tanksley, 1983). Sugeriu-se sua aplicação na construção de mapas de ligação e a utilização destes mapas para a localização de características de interesse agrônômico. Além disso, os marcadores moleculares do tipo RFLP associados à características agrônômicas de interesse, podem ser utilizados em programas de melhoramento, com a finalidade de monitorar a seleção a ser realizada (“MAS - Marker Assisted Selection”; Gimelfarb & Lande, 1994).

Burr *et al.* (1983) foram os primeiros a sugerirem a utilização dos RFLPs para explorar as relações evolutivas entre populações e para o estudo da similaridade entre linhagens, permitindo, assim, a estimativa das distâncias genéticas entre elas. Hoje há vários trabalhos na literatura utilizando a técnica de RFLP para avaliar o polimorfismo de diversas espécies vegetais tais como milho (Messmer *et al.*, 1991, 1992), aveia (Besse *et al.*, 1994), cebola (Bark & Harvey, 1995) e feijão (Velásquez & Gepts, 1993), entre outras. Mediante o uso deste tipo de marcador, verificou-se o alto grau de polimorfismo intraespecífico em espécies de polinização cruzada, tais como milho (*Zea mays* L.), batata doce (*Ipomoea batata* L.), e crucíferas (*Brassica campestris*). Estas espécies aparentam ter maior grau de polimorfismo do que as que fazem autofecundação. Gebhart & Salamini (1992) propuseram que, em milho, a presença de elementos de transposição poderia estar ampliando esta variabilidade.

Os RFLPs apresentam uma série de vantagens em relação a todos os outros tipos de marcadores (morfológicos, citológicos ou bioquímicos) usados inicialmente: (a) os marcadores do tipo RFLP são codominantes e livres de efeitos pleiotrópicos; (b) não são afetados pelo ambiente; (c) são praticamente ilimitados em número; (d) têm distribuição

aleatória no genoma e sua estabilidade independe do estágio de desenvolvimento da planta (Evola *et al.*, 1986; Helentjaris *et al.*, 1985). Desta forma, eles permitem uma estimativa precisa da distância genética entre genótipos, devido a possibilidade de amostrarem o genoma uniforme e completamente.

2.5.2. Associação das distâncias genéticas estimadas a partir de RFLPs com a performance de híbridos simples

Visto que a diversidade genética entre linhagens está relacionada à produção de híbridos F1 e ao nível de heterose (Moll *et al.*, 1962,1965; Paterniani & Lonnquist, 1963), vários estudos foram desenvolvidos com a finalidade de verificar se as distâncias genéticas estimadas através dos dados de RFLPs estariam positivamente correlacionadas com a produção de grãos de híbridos F1 e/ou com o nível de heterose para esta característica. Também tem se avaliado a utilidade destes dados para prever a performance de híbridos simples em milho. A eficiência de um programa de melhoramento de híbridos poderia ser aumentada se cruzamentos superiores pudessem ser preditos antes da avaliação no campo, através da seleção de linhagens parentais.

Os resultados obtidos até o momento são conflitantes. Lee *et al.* (1989) avaliaram a utilidade dos dados de RFLP para elucidar padrões heteróticos de linhagens e híbridos simples de milho e correlacionaram estes dados com medidas de produção de grãos e com a capacidade específica de combinação, obtendo correlações significativas para 06 dos 10 cromossomos de milho. Melchinger *et al.* (1990a,b) também relataram correlações positivas entre medidas de distâncias genéticas estimadas por RFLPs com a performance dos híbridos F1 para a produção de grãos. Embora significativos, os valores não foram preditivos. Smith *et al.* (1990) usaram 287

combinações clone-enzima para determinar distâncias genéticas entre linhagens endogâmicas de milho relacionadas e não-relacionadas à linhagens elites do cinturão norte-americano. Os autores relataram correlações altamente significativas entre as distâncias genéticas medidas por RFLPs com heterose ($r = 0,87$) e com performance dos híbridos F1 para produção de grãos ($r = 0,93$).

Outros trabalhos não encontraram nenhuma correlação direta entre heterose ou qualquer caráter agrônômico examinado com medidas de distância genética obtidas por RFLPs (Godshalk *et al.*, 1990; Dudley *et al.*, 1991; Boppenmaier *et al.*, 1992). Estes resultados sugerem prudência na utilização destes dados para estimativas de produção de híbridos simples e de heterose.

Vários argumentos têm sido apresentados para explicar a ausência de correlação encontrada. Entre eles está o fato de que as sondas utilizadas não representem necessariamente seqüências codificadoras de importância para a produção. Esta correlação poderia ser positiva se as sondas utilizadas fossem reconhecidamente ligadas à locos de interesse agrônômico, tais como QTLs (“Quantitative Trait Loci”) (Dudley *et al.*, 1991; Melchinger *et al.*, 1990a,b; Godshalk *et al.*, 1990). Embora até o momento exista bastante conhecimento sobre sondas ligadas à QTLs em milho, a questão de como estas regiões são reguladas e quais os genes nelas presentes ainda permanece não esclarecida.

Charcosset *et al.* (1991) propuseram, teoricamente, a necessidade de haver um desequilíbrio de ligação entre o grau médio de heterozigosidade num loco marcador e num QTL para a associação entre eles ocorrer. Bernardo (1992) demonstrou a necessidade de certas condições para que haja a predição efetiva da performance híbrida utilizando-se a heterozigosidade num loco marcador: (1) fortes efeitos de dominância; (2) as frequências

alélicas em locos individuais nas linhagens parentais deveriam ser negativamente correlacionadas (comportamentos contrários); (3) alta herdabilidade do caráter agrônômico avaliado; (4) pequena faixa de variação da média de frequências alélicas parentais; (5) 30 - 50% dos QTLs (ao menos) deveriam estar ligados aos marcadores moleculares e (6) não mais de 20 - 30% dos marcadores moleculares deveriam estar dispersos no genoma ou ligados à QTLs.

2.5.3 RAPD: princípios e aplicações da técnica

Desde o seu desenvolvimento, a técnica de PCR (“Polymerase Chain Reaction” - Saiki *et al.*, 1988) tem revolucionado vários ramos da Biologia Molecular. Modificações desta técnica produziram uma classe específica de marcadores moleculares dita RAPD (“Random Amplified Polymorphic DNA”), (Williams *et al.*, 1990; Welsh & McClelland, 1990).

A técnica de RAPD é baseada na amplificação por PCR de fragmentos de DNA, utilizando oligonucleotídeos de bases aleatórias como “primers” para amplificar o DNA genômico. Os “primers” usados, em geral, têm de 9 a 10 pares de base (pb) de tamanho, têm alto conteúdo de GC (guanina/citosina - de 50 a 80%) e não contêm seqüências palindrômicas. Os produtos de PCR são produzidos em regiões do genoma flanqueadas por dois sítios que devem estar a uma distância máxima de 4.000 pb, já que este é o limite para que ocorra a amplificação por PCR. Os produtos de amplificação são analisados por eletroforese em géis de agarose, corados com brometo de etídeo. Pelas condições da reação de PCR, o limite do tamanho das bandas varia de 100 a 3000 pb (Kang *et al.*, 1993).

Num ensaio de RAPD, um único “primer” de DNA de seqüência aleatória é misturado com o DNA genômico, na presença de uma polimerase termoestável (*Taq*-DNA

polimerase) e de um tampão de reação. Na reação de PCR, a mistura é submetida a vários ciclos de desnaturação, anelamento e extensão, amplificando o fragmento flanqueado pelo par de “primers”.

Apesar da aparente facilidade de sua metodologia, a técnica requer uma otimização empírica entre uma dada combinação DNA molde - “primer”, o que pode ser problemático devido à uma série de variáveis a serem consideradas: (a) temperatura de anelamento; (b) duração da desnaturação; (c) intervalos de anelamento e extensão; (d) qualidade e concentração de DNA; (e) concentração de magnésio (cofator da enzima); (f) concentração da polimerase; (g) sequência do “primer” e (i) termociclador. Recentemente, Schweder *et al.* (1995) indicaram também o efeito do intervalo de transição entre a temperatura de desnaturação e a temperatura de anelamento.

Marcadores do tipo RAPD são herdáveis e de caráter dominante, em sua maioria. Os polimorfismos podem ser devido a alterações de uma única base, deleções nos sítios dos “primers” ou inserções que aumentem a distância entre os sítios de anelamento dos “primers” acima de 3000 pb. A frequência de polimorfismos varia de 2,5 bandas/“primer” em *Neurospora* para 0,5 banda/“primer” em milho (Kang *et al.*, 1993). Os polimorfismos geralmente são reconhecidos pela presença de um fragmento amplificado em um dos genótipos em relação à ausência deste mesmo fragmento em outro genótipo. A reação de amplificação é repetida com vários “primers” diferentes em condições que resultem na produção de várias bandas para cada “primer”. Um único “primer” pode identificar várias marcas, cada uma representando um loco (Williams *et al.*, 1993).

Esta técnica distingue-se da técnica de RFLP pela sua rapidez e facilidade de trabalho, já que não envolve hibridização ou radioatividade. Além disto, utiliza uma quantidade bem menor de DNA (10 - 50 ng) do que a técnica de RFLP (2 - 10 µg) (Waugh & Wayne, 1992).

Ragot & Hoisington (1993) compararam três protocolos de marcadores moleculares: c-RFLPs (c = quimioluminescente), r-RFLPs (r = radioativos) e RAPDs quanto ao custo e tempo requerido. Os experimentos de avaliação foram feitos com milho. O aumento do custo total com aumento do número de indivíduos e dos marcadores analisados foi maior para RAPDs do que RFLPs. Entretanto, RAPDs foram os de maior eficiência e custo para estudos envolvendo pequenos tamanhos de amostra, enquanto os RFLPs mostraram vantagem para grandes tamanhos de amostra. c-RFLPs requiseram menos tempo do que r-RFLPs (pelo menos de exposição).

Recentemente, Micheltore (1991) mostrou que o perfil de bandas de RAPD ligado a um caráter de interesse poderia ser facilmente identificado usando-se duas amostras de DNA: uma, de indivíduos que expressem o caráter e outra, de indivíduos que não o expressem. Qualquer polimorfismo entre os dois “pools” deveria estar ligado ao caráter. Os marcadores identificados como polimórficos nos “pools”, poderiam ser confirmados quanto a sua co-segregação com o caráter, através de sua análise na população segregante. Esta técnica foi denominada de “Bulk Segregant Analysis” e hoje é amplamente usada para mapear caracteres simples com marcadores do tipo RAPD (Rafalski & Tingey, 1993).

Uma vez que um marcador ligado a um caráter é encontrado, é relativamente simples transformar a técnica de RAPD em outro ensaio de PCR mais reprodutível, pelo uso de: AS - PCR (“Allele - Specific PCR”- Wu *et al.*, 1989), Ligação Alelo Específica (Nickerson *et al.*, 1990) ou SCAR (“Sequence - Characterized Amplified Region”- Paran & Micheltore,

1993). Outras técnicas que também são utilizadas atualmente para avaliar a divergência na sequência de DNA são a de SSCP (“Single - Strand Conformational Polymorphism Assay”- Masato *et al.*, 1989) e DGGE (“Denaturing Gradient Gel Eletrophoresis”- Myers *et al.*, 1987), as quais podem ser aplicadas à produtos individuais de amplificação derivados de bandas de RAPD. Os microssatélites, que revelam o polimorfismo no número de pequenas repetições em “tandem”, têm sido utilizados mais recentemente para acessar a diversidade de genomas de eucariotos (Rafalski & Tingey, 1993).

Entre as principais aplicações do RAPD estão o “fingerprint” (Welsh & McClelland, 1991; Lanham *et al.*, 1995), o mapeamento genético (Ruter *et al.*, 1992; Binelli & Bucci, 1994; Plomion *et al.*, 1995) e a tipagem molecular de genótipos e de espécies para estudo de genética de populações, melhoramento e biossistemática (Halward *et al.*, 1992; Huff *et al.*, 1993; Hilu, 1995).

Assim, na clínica médica, a técnica de RAPD tem sido utilizada para distinguir isolados clínicos de microrganismos patogênicos, como é o caso de *Helicobacter pylori*, bactéria associada à etiologia de carcinomas gástricos: 64 isolados puderam ser diferenciados com um único “primer”, sugerindo a existência de um elevado nível de diversidade genética dentro desta espécie (Akopyanz *et al.*, 1992).

Na sistemática de fungos patogênicos, a identificação da variabilidade genética é feita segundo padrões de morfologia e de comportamento fisiológico do fungo. Entretanto, estas abordagens não são suficientes para a correta identificação de espécies patogênicas, o que é essencial para o estabelecimento de espécies, variedades e isolados de fungos entomo- ou fitopatogênicos como por exemplo *Leptosphaeria* (Goodwin & Annis, 1991) e *Fusarium* (Crowhurst *et al.*, 1991).

Na sistemática de insetos, a diagnose de espécies e a filogenia de um grupo de coleópteros (*Ips grandicollis*) foi feita com base em dados de RAPD. As espécies têm morfologia muito semelhante e são simpátricas; logo, o diagnóstico por DNA resolveu o problema. A filogenia conseguida foi congruente com a reconstrução por dados de isoenzimas (Cognato *et al.*, 1995). Também, utilizou-se a técnica de RAPD para diferenciar espécies de traças (Lepdoptera: Plutellidae) susceptíveis ou resistentes à *Bacillus turingensis* (Heckel *et al.*, 1995).

Em plantas, os RAPDs têm sido utilizados para detectar variação genética intra-específica (Williams *et al.*, 1990; Welsh *et al.*, 1991; Caetano-Anollis *et al.*, 1991; Hu & Quiros, 1991; Kazan *et al.*, 1992; Chalmers *et al.*, 1994; Abdelnoor *et al.*, 1995), para a avaliação comparativa da variabilidade inter- e intra-específica (Huff *et al.*, 1993; Haley *et al.*, 1994), para a determinação da variabilidade genética de plantas micropropagadas *in vitro* (Martín & Pérez, 1994; Consoli, 1995), para descarte de híbridos somáticos, como é o caso de híbridos de batata gerados por fusão de protoplastos (Newbury & Ford-Lloyd, 1993), para a construção de mapas genéticos (Carlson *et al.*, 1991; Klein-Lnakhorst *et al.*, 1991; Michelmore *et al.*, 1991; Giovannoni *et al.*, 1991; Reiter *et al.*, 1992; Al-Janabi *et al.*, 1993) e para avaliar o fluxo gênico entre espécies (Arnold *et al.*, 1991). A técnica também tem sido utilizada na determinação de parentais (Welsh *et al.*, 1990), com propósitos de classificação e sistemática (Honeycutt *et al.*, 1992; Khider, 1994) e para descarte de contaminantes em progênies de cruzamentos controlados em espécies florestais (Roy *et al.*, 1992).

Abdelnoor *et al.* (1995) utilizaram a técnica de RAPD para avaliar a diversidade genética entre 38 cultivares de soja brasileiros (*Glycine max* L. Merrill). A análise possibilitou a separação dos cultivares em cinco grupos distintos, baseados nas distâncias genéticas entre eles.

Comparações entre os dados de RAPD e os dados de “pedigree” para estes mesmos cultivares foram coerentes na maioria dos casos. Os autores recomendaram o uso da técnica de RAPD como uma poderosa ferramenta para o estudo da diversidade genética entre indivíduos da mesma espécie.

Huen & Helentjaris (1993) analisaram a herança de RAPDs em híbridos F1 de milho. Usando 21 “primers”, eles obtiveram um total de 140 fragmentos diferentes para a presença/ausência e para a intensidade de variação. Considerando todos os genótipos analisados, 20,7% apresentaram fragmentos não-polimórficos, enquanto 31,1% foram polimórficos, indubitavelmente, distintos pela simples presença e ausência de um fragmento específico dos genótipos das linhagens. Por outro lado, 42,7% foram polimorfismos quantitativos os quais tiveram variação de intensidade do fragmento. Os autores atribuem parte desta diferenciação de intensidade à um padrão intermediário entre os padrões dos parentais (dominância parcial). Davis *et al.* (1995) identificaram dez marcadores codominantes (de 300 a 1350 pb) no mapeamento de populações de *Cicer arietum* L. e *Fragaria vesca* L. Atribuíram este fenômeno à presença de banda heteroduplex (não-parental), migrando mais vagarosamente que as outras, presente em indivíduos heterozigotos. Sugeriram que uma mistura de DNAs dos parentais antes da reação de PCR (“Template mixing”) poderia propiciar a geração de uma banda não-parental. Existem outros relatos a cerca deste fenômeno (Hunt & Page, 1992; Aylife *et al.*, 1994). Assim, ao que parece, o “dogma” de que RAPDs são exclusivamente marcadores dominantes parece estar sendo abalado pela descoberta de bandas não-parentais, ainda que ocasionais. É claro que a visualização da codominância é ainda um fenômeno raro em RAPDs quando comparado à RFLPs, e requer estratégias especiais de otimização como o “Template mixing”.

Vários trabalhos propuseram-se a fazer estudos comparativos entre isoenzimas, RFLPs e RAPDs, com a finalidade de avaliar a utilidade de cada tipo de marcador na determinação de relações genéticas entre diferentes espécies (Liu & Furnier, 1993; Thormann *et al.*, 1994; Halldén *et al.*, 1994; Heun *et al.*, 1994; Baruffi *et al.*, 1995; Vicario *et al.*, 1995).

A variação genética em locos isoenzimáticos foi examinada nas espécies *Populus tremuloides* (A) e *Populus grandidentata* (B) (Liu & Furnier, 1993). Os resultados de RFLPS foram consistentes com o de isoenzimas, de forma que a espécie (A) teve maior variabilidade genética do que a espécie (B). Os padrões de RAPD, em ambas as espécies, mostraram maior variabilidade do que as estimativas de outros marcadores. Entretanto, os RAPDs se adequaram plenamente ao “fingerprinting” dos indivíduos. Heun *et al.* (1994) comparando dados de isoenzimas e RAPDs a fim de determinar as relações genéticas entre acessos de aveia (*Avena sterilis* L.) mostraram que, embora a correspondência entre os elementos das matrizes de distâncias para os dois marcadores tenha sido moderada, as relações entre os acessos, representadas pelos agrupamentos e pelo método de ordenação, foram de considerável acordo. Os autores afirmaram que a análise dos dados de RAPD resultou numa separação mais definitiva dos acessos. Também dados de RAPD para isolados de *Giardia duodenalis* se correlacionaram significativamente com dados isoenzimáticos para os mesmos isolados (Morgan *et al.*, 1993). A diferenciação de 16 eco-espécies de arroz (*Oryza sativa* L.) utilizando a técnica de RAPD se mostrou de acordo com dados anteriores de isoenzimas e de RFLPs (Fukuoka *et al.*, 1992).

Thormann *et al.* (1994) compararam a eficiência de marcadores do tipo RFLP e RAPD para estimar relações filogenéticas entre e dentro de espécies crucíferas (*Brassica* sp., *Raphanus sativus*). Baseados em comparações entre as matrizes de similaridade genética e

valores cofenéticos, os marcadores do tipo RAPD foram muito similares aos do tipo RFLP para estimar relações genéticas intraespecíficas. Entretanto, a nível interespecífico, os autores encontraram diferenças entre estes dois tipos de marcadores, de forma que, a este nível, os RAPDs pareceram ser menos confiáveis do que os RFLPs. Este fato deve-se à amplificação, em indivíduos diferentes, de fragmentos que possuem o mesmo tamanho mas não correspondem ao mesmo loco, ou seja, não são homólogos. Os autores descartaram a possibilidade da presença de DNA mitocondrial e de cloroplasto poderem estar contribuindo para as diferenças entre os dendrogramas obtidos por RAPDs e por RFLPs.

2.5.4. Associação das distâncias genéticas estimadas a partir de RAPDs com a performance de híbridos simples

Há poucos trabalhos na literatura que relatem o uso da técnica de RAPD para avaliar as distâncias genéticas de espécies vegetais, com a finalidade de correlacionar estas estimativas com dados de performance de híbridos simples ou de heterose. Jain *et al.* (1994) exploraram o potencial dos ensaios de RAPD para avaliar as relações de genótipos de *Brassica juncea* de diversas origens geográficas e para predizer a performance dos híbridos F1. Embora nenhuma correlação direta tenha sido encontrada entre as medidas de distância genética e o grau de heterose, os autores encontraram que 90% dos cruzamentos entre genótipos indianos e exóticos tiveram performance de grãos acima da média. Assim, defenderam o uso da técnica de RAPD para alocar genótipos de origem desconhecida para grupos heteróticos já estabelecidos, mas sugeriram cautela quanto a utilização das distâncias genéticas baseadas unicamente em dados de RAPD para predizer a performance híbrida.

Desta forma, até o presente momento, o trabalho, que foi objeto desta tese, é pioneiro na investigação das distâncias genéticas através de dados de RAPD com linhagens endogâmicas de milho (*Zea mays* L.) visando correlacionar estas medidas com dados de performance dos respectivos híbridos simples.

OBJETIVOS

3. OBJETIVOS

3.1. Gerais

3.1.1. Conhecimento mais profundo sobre a diversidade genética existente em duas populações tropicais de milho (BR-105 e BR-106), de importância em programas de melhoramento no Brasil;

3.1.2. Predição da heterose antes da análise de híbridos simples no campo, reduzindo assim, recursos e esforços alocados na fase de avaliação dos híbridos;

3.2. Específicos

3.2.1. Avaliar o polimorfismo entre linhagens obtidas de duas populações divergentes através de marcadores moleculares do tipo RAPD;

3.2.2. Estimar as distâncias genéticas entre estas linhagens.

3.2.3. Examinar a correlação entre as distâncias genéticas (estimadas a partir da diversidade detectada por RAPDs) entre as linhagens, com a performance para produção de grãos, altura de planta e de espiga medida para seus respectivos híbridos simples e para a heterose medida para cada uma destas características.

MATERIAL E MÉTODOS

4. MATERIAL E MÉTODOS

4.1. Material Vegetal:

Foram utilizadas 18 linhagens endogâmicas (S3) de milho (*Zea mays* L.), obtidas no programa de melhoramento desenvolvido pelo professor Dr. Cláudio Lopes de Souza Jr., do Departamento de Genética da ESALQ/USP. Oito destas linhagens são provenientes de uma população tailandesa, que foi trazida para o CIMMYT (México) e depois, para o Brasil (BR-105), enquanto que as outras dez restantes são originadas de uma população brasileira (BR-106) (Tabela 01). BR-105 e BR-106 são populações liberadas pelo Centro Nacional de Pesquisa de Milho e Sorgo - EMBRAPA - Sete Lagoas - M.G.

A população BR-106 foi obtida pelo inter cruzamento das populações Centralmex, Maya e Composto Dentado e o material resultante foi cruzado com BR-108, uma população de porte baixo e ciclo precoce. A população BR-106, portanto, apresenta porte baixo e grãos dentados amarelos.

A população BR-105, originalmente denominada “Suwan-DMR”, foi obtida na Tailândia, através de seleção recorrente com progênies S1, utilizando-se alta intensidade de seleção. Apresenta porte baixo, ciclo precoce e grãos duros alaranjados.

Por ser resultante de cruzamentos entre populações e compostos, BR-106 é considerada uma população de alta variabilidade genética, enquanto que BR-105, por ter sido submetida a seleção recorrente com progênies S1 e alta intensidade de seleção, pode ser considerada um sintético que apresenta menor variabilidade genética em relação a BR-106

(Souza Jr. *et al.*, 1993). Estas populações apresentam, em cruzamento, elevados valores de heterose, o que as aloca em grupos heteróticos distintos (Souza Jr. *et al.*, 1993; Naspolini F.^o *et al.*, 1981).

4.2. Avaliações de campo

Das populações BR-106 e BR-105, foram selecionadas 10 e 08 linhagens endogâmicas, respectivamente, as quais foram cruzadas entre si num esquema dialélico, gerando 80 híbridos simples. Estes híbridos foram avaliados em látices 10 x 10, em três locais, com duas épocas de plantio por local e duas repetições por experimento, em 1993/1994. Foram incluídas 20 testemunhas para completar os cem tratamentos. As parcelas foram constituídas de 3,6 m de comprimento, espaçadas de 0,90 m entre linhas e 0,20 m entre plantas dentro de linhas, com um total de 20 plantas por parcela após o desbaste. Tomaram-se dados de produção de grãos por parcela, sendo estes transformados em toneladas por hectare (t/ha), altura da planta (cm/planta) e altura da espiga (cm/espiga) de cinco plantas tomadas aleatoriamente por parcela. Simultaneamente, as linhagens foram avaliadas nas mesmas condições que os experimentos dos híbridos em blocos casualizados. De posse dos valores médios das linhagens e seus respectivos híbridos, estimaram-se os valores de heterose para cada cruzamento como: $h = F1 - (1/2) (L1 + L2)$, onde $L1$ e $L2$ são os valores das linhagens parentais, e $F1$ corresponde ao valor do híbrido entre as duas linhagens. Foram também avaliados em blocos casualizados (nas mesmas condições acima descritas) os híbridos simples intrapopulacionais, isto é, resultantes dos cruzamentos de linhagens oriundas da mesma população.

A instalação dos experimentos de campo, bem como a tomada de dados, foram realizadas pelo Prof. Dr. Cláudio Lopes de Souza Jr., na Estação Experimental do Departamento de Genética da ESALQ-USP, Piracicaba, SP.

4.3. Métodos

4.3.1. Coleta e preparo do material vegetal

Vinte sementes de cada linhagem foram plantadas em casa-de-vegetação (Departamento de Genética da ESALQ/USP, Piracicaba, S.P.), no primeiro semestre de 94, tendo recebido adubação adequada. Plantas, com cerca de 05 semanas de crescimento, tiveram suas folhas coletadas e colocadas rapidamente em gelo, trazidas para o CBMEG/UNICAMP onde foram congeladas em nitrogênio líquido. Em seguida as folhas congeladas foram liofilizadas (72h, -60°C, 05 a 10 microns de Hg), moídas num moinho mecânico (Ciclotec-1093 Sample Mill, Tecator), acondicionadas em frascos plásticos hermeticamente fechados e armazenadas em freezer - 20°C.

Para cada linhagem foram misturadas quantidades iguais de folhas de dez plantas, em média, (“bulk”) no momento da coleta do material. Desta forma, cada uma das dezoito amostras era constituída da mistura de, pelo menos, 10 plantas da linhagem. Este procedimento foi adotado para evitar que polimorfismos dentro de uma linhagem, pudessem interferir nos resultados de diversidade entre as linhagens.

Tabela 1. Linhagens endogâmicas de milho (*Zea mays* L.),
seus códigos de uso, nomenclatura e população
de origem.

CÓDIGO DE USO	NOMENCLATURA	ORIGEM
01	<i>05-01.4B</i>	BR-105
02	<i>05-05.2A</i>	BR-105
03	<i>05-17.1A</i>	BR-105
04	<i>05-18.6A</i>	BR-105
05	<i>05-19.1B</i>	BR-105
06	<i>05-23.2B</i>	BR-105
07	<i>05-33.5B</i>	BR-105
08	<i>05-34.2B</i>	BR-105
09	<i>06-03.5B</i>	<u>BR-106</u>
10	<i>06-06.3A</i>	<u>BR-106</u>
11	<i>06-08.1A</i>	<u>BR-106</u>
12	<i>06-08.2A</i>	<u>BR-106</u>
13	<i>06-14.4B</i>	<u>BR-106</u>
14	<i>06-24.7B</i>	<u>BR-106</u>
15	<i>06-28.1A</i>	<u>BR-106</u>
16	<i>06-29.7B</i>	<u>BR-106</u>
17	<i>06-37.5B</i>	<u>BR-106</u>
18	<i>06-44.1B</i>	<u>BR-106</u>

4.3.2. Extração do DNA genômico

DNA genômico foi isolado de acordo com Hoisington *et al.* (1994). Adicionou-se 9,0 ml de tampão de extração CTAB (100 mM Tris-7,5, 700 mM NaCl, 50 mM EDTA-8,0, 140 mM BME, 1% CTAB) quente (65°C) à 300 mg de tecido liofilizado e moído em tubos de propileno de 15,0 ml para centrifuga. O tecido foi homogeneizado com tampão de extração, com auxílio de bastão de vidro e também através de inversões suaves e, em seguida, incubado por 60 minutos a 65°C com agitação a cada 15 minutos. Após serem removidos do banho, esperou-se 4 a 5 minutos e adicionou-se 4,5 ml de clorofórmio/álcool isoamílico (24:1) a cada tubo. A mistura foi conseguida através inversão suave por 5 a 10 minutos.

Em seguida, as amostras foram centrifugadas durante 10 minutos à 2.500 rpm. A fase aquosa foi transferida para novos tubos de 15 ml. Repetiu-se a extração com clorofórmio/álcool isoamílico (24:1). O sobrenadante foi transferido para novos tubos de 15 ml contendo 25,0 µl de RNase A à 10,0 mg/ml. A solução foi misturada através de lentas inversões e os tubos foram incubados durante 30 minutos a temperatura ambiente.

Um total de 6,0 ml de isopropanol foi adicionado a cada tubo e misturado por inversões suaves. O DNA precipitado de cada linhagem foi removido com o auxílio de “anzóis” construídos utilizando-se pipetas Pasteur.

O DNA foi recuperado, colocado em tubo plástico de 5,0 ml, contendo 3,0 ml de solução “Wash 1” (76% EtOH; 0,2M NaOAc) onde permaneceu em imersão por 20 minutos.

Em seguida, o DNA foi transferido para outro tubo de 5,0 ml, brevemente lavado em 2,0 ml de solução “Wash 2” (76% EtOH, 10mM NH₄OAc) e transferido para novo tubo

contendo 1,0 ml de TE (10 mM Tris-HCL, 1 mM EDTA, pH 8,0). O DNA de cada linhagem permaneceu neste tubo “overnight”, sob agitação lenta para sua ressuspensão completa. Em seguida, foi realizada uma extração com fenol para se obter uma maior pureza de DNA, fator crucial para um bom desempenho das reações de amplificação. Adicionou-se ao DNA de cada linhagem 1,0 ml de fenol equilibrado (0,1 mM Tris-HCl, pH 8,0 e 0,2% β -mercaptoetanol). Os tubos foram centrifugados (10 min., 2.500 rpm) e o sobrenadante foi transferido para novo tubo. Foi feita uma nova extração de cada DNA com 1,0 ml de clorofórmio/álcool isoamílico (24:1). Após centrifugação, o sobrenadante foi transferido para um tubo novo, sendo precipitado por adição de 50,0 μ l de NaCl 5M e 2,5 ml de etanol absoluto. A solução foi misturada através de inversões e o DNA precipitado foi removido com o auxílio de um bastão de vidro. Novamente, foram feitas lavagens em solução de “Wash 1” e “Wash 2”, tal como previamente descrito.

Finalmente, tendo atingido maior grau de pureza, o DNA foi ressuspendido “overnight” em TE, sob agitação lenta, à temperatura ambiente. O DNA obtido foi quantificado e diluído para utilização nas reações de RAPD.

4.3.3. Quantificação do DNA

Inicialmente, o DNA de cada uma das linhagens foi quantificado em espectrofotômetro (Beckman - Du-65 Spectrophotometer) obtendo-se uma estimativa grosseira de sua concentração. Mediante esta primeira estimativa, alíquotas de 10,0 μ l foram retiradas de cada amostra e diluídas à uma concentração estimada de “40,0 ng/ μ l”. Alíquotas de 5,0 μ l (200 ng) retiradas das soluções de “40,0 ng/ μ l” migraram em gel de agarose 0,8% ao lado de um padrão de concentração de DNA, construído a partir de DNA de fago λ , numa gama crescente

de 50 a 300 ng. Mediante comparação visual da intensidade das bandas do DNA das amostras e dos padrões, novas concentrações foram estimadas. O processo de quantificação foi repetido com o DNA de cada linhagem até se ter certeza de suas respectivas concentrações. Uma diluição de 2,5 ng/ μ l (aproximadamente 300 μ l) foi feita e novas alíquotas foram migradas em gel junto com o padrão de concentração de DNA. Uma vez verificado que as diluições estavam corretas, foram feitas novas diluições para um volume final de 4,0 ml a fim de serem usadas ao longo de todos os experimentos de RAPD. As diluições finais foram feitas em “tampão de diluição de DNA”, contendo apenas 10 mM Tris-HCl (pH 8,0) e 0,2 mM de EDTA (pH 8,0) ao invés de TE, já que este último contém uma concentração muito alta de EDTA. O EDTA é um agente quelante que, em excesso, interfere na reação de amplificação.

O DNA, tanto no estoque concentrado como diluído, foi armazenado a 4°C sendo periodicamente requantificado para avaliar possível degradação.

4.3.4. Condições de Amplificação

As reações de amplificação foram realizadas como descrito por Williams *et al.* (1990) com pequenas modificações. As amplificações foram efetuadas num Termociclador 480 (Perkin Elmer). Uma desnaturação inicial do DNA foi conduzida a 94°C por 5 minutos; esta etapa foi seguida por 45 ciclos de amplificação (94°C, 1 min - desnaturação; 35°C, 1 minuto - anelamento; 72°C, 2 minutos - extensão). Ao final dos 45 ciclos, adicionou-se um passo final de extensão (7 minutos a 72°C), seguido de um passo de incubação (15°C, ∞).

As reações foram realizadas em 25,0 μ l contendo 0,2 μ M de “primer” (Kits Operon, Technologies Inc.- Alameda, CA), 2,0 unidades de *Taq*-DNA polimerase (Gibco BRL-

5U/ μ l), 2,5 μ l de tampão 10X fornecido pelo fabricante (200 mM Tris-HCl, pH 8,4 e 500 mM KCL), 0,1 mM de cada dNTP e 3,0 mM de $MgCl_2$. A quantidade de DNA utilizada foi de 25 ng para cada linhagem. Todas as reações foram feitas com repetições.

Os produtos de amplificação foram separados por eletroforese em géis de agarose 1,5% usando tampão TBE 1X (89 mM Tris, 89 mM Ácido Bórico e 2 mM EDTA dissódico) e, como marcador de peso molecular, foi usado o DNA do fago ϕ x174 digerido com *HaeIII* e misturado ao DNA do fago λ digerido com *HindIII*. Os géis foram corridos a uma taxa inferior a 3 V/cm, corados com brometo de etídeo e fotografados com máquina polaróide sob luz UV.

4.3.5. Análise dos dados moleculares

A análise dos dados foi realizada, a partir da leitura dos géis, comparando o perfil de bandas amplificadas com um mesmo “primer” para as diferentes linhagens endogâmicas de milho. Os registros foram tomados de acordo com a presença ou ausência (1 ou 0, respectivamente) de cada banda de DNA reprodutível com o mesmo peso molecular. Os dados foram organizados numa matriz binária, sendo que as linhagens endogâmicas constituíram as colunas da matriz, enquanto os dados referentes à presença ou ausência de bandas, obtidos a partir de todos os produtos de amplificação gerados pelos “primers”, foram organizados como linhas da respectiva matriz.

Coeficientes de similaridade entre todas as linhagens foram gerados a partir da matriz binária (Anexo I) de acordo com Jaccard (1908): $\frac{a}{n-d}$ onde a = número de coincidências positivas, b = número de bandas que estão simultaneamente presentes na

linhagem i e ausentes na linhagem i' , c = número de bandas que estão simultaneamente presentes na linhagem i' e ausentes na linhagem i , n = tamanho total da amostra ($a + b + c + d$), e d = número de coincidências negativas, sendo:

$i \backslash i'$	1	0
1	a	b
0	c	d

onde: 1 e 0 referem-se a presença ou ausência de bandas, respectivamente nas linhagens i e i' .

Como pode ser visto através da fórmula, o coeficiente de Jaccard exclui a dupla ausência de bandas de mesmo peso molecular como um fator que assemelharia duas linhagens (Jaccard, 1908). Outros coeficientes (Anexo A) usados para estimar semelhança para variáveis dicotômicas como o de Coincidência Simples (“Simple Matching”) consideram “ d ” (coincidência de zeros). Além disso, o coeficiente de Jaccard em relação ao de Dice (Dice, 1945) não dá peso duplo para as coincidências positivas o que, no caso de linhagens endogâmicas, poderia superestimar a similaridade entre duas linhagens. Tanto o coeficiente de similaridade de Jaccard quanto o de Dice foram testados com os mesmos conjuntos de dados. As matrizes de parença ou similaridade foram utilizadas para a construção de dendrogramas através de métodos de agrupamento.

Um número aleatório e crescente de bandas polimórficas (50, 98, 150, 190, 215 e 262 bandas) foi utilizado para construir dendrogramas sucessivos, com o objetivo de avaliar as mudanças nos agrupamentos com o acréscimo de bandas, e de acompanhar a “estabilidade” dos agrupamentos gerados. O método de médias das distâncias ou UPGMA (“Unweighted Pair-Group Method with Arithmetical Averages”) foi utilizado para a construção dos agrupamentos.

Este é um processo hierárquico, onde a cada passo diminui-se uma dimensão da matriz de similaridade através da união dos pares mais semelhantes, até se reunir todos os pontos em um único grupo (Bussab *et al.*, 1990).

4.3.6. Coeficiente de correlação cofenética (r_{cof})

Como um indicativo quantitativo do desempenho da análise de agrupamento, utilizou-se o coeficiente de correlação cofenética, computando uma correlação entre os correspondentes elementos da matriz de similaridade original e a matriz de similaridade gerada por um método de agrupamento hierárquico, dita matriz cofenética. Este coeficiente inicialmente proposto para avaliar a concordância entre um agrupamento obtido e sua estrutura original, passou a ser empregado, posteriormente, como um método de avaliação para diferentes agrupamentos obtidos a partir do mesmo conjunto de dados (Bussab *et al.*, 1990). Quanto mais próximo da unidade, melhor é a representação do coeficiente; quanto mais próximo de zero, pior. Entretanto, a interpretação do quanto significa um coeficiente de correlação cofenética entre duas variáveis (alto/baixo), depende da área de estudo e da natureza dos dados. Pode-se dizer que um valor em torno de 0,8 pode ser considerado um bom ajuste (Bussab *et al.*, 1990). Rholf (1992) apresentou uma sugestão para interpretação do grau de ajuste destas medidas:

<i>Nível</i>	<i>Interpretação</i>
$0,9 \leq r_{\text{cof}}$	Muito Bom
$0,8 \leq r_{\text{cof}} < 0,9$	Bom
$0,7 \leq r_{\text{cof}} < 0,8$	Regular
$r_{\text{cof}} < 0,7$	Péssimo

Rholf & Fisher (1968) mostraram que os procedimentos estatísticos normais não podem ser utilizados para testar a significância deste tipo de correlação já que uma matriz é construída a partir de outra. Estes autores também demonstraram que a maioria das correlações cofenéticas encontradas são em geral significativas.

Correlações cofenéticas foram calculadas para os sucessivos dendrogramas obtidos tanto com o coeficiente de similaridade de Jaccard (Jaccard, 1908) quanto com o de Dice (Dice, 1945). Os resultados foram utilizados para sustentar as escolhas do melhor coeficiente e do agrupamento mais “estável”.

4.3.7. Análise multivariada

Complementando a análise de agrupamento dos dados, utilizou-se o método de ordenação de coordenadas principais calculadas a partir da matriz de similaridade original.

Entre os métodos de ordenação estão o de componentes principais (PCA - Pearson, 1901; Hotelling, 1933,1936) e o de coordenadas principais (PCO - Gower, 1966). A análise dos dados através da ordenação consiste em transformar um conjunto original de variáveis em outro conjunto de dimensão equivalente, mas com propriedades importantes, que

são de grande interesse em certos estudos de melhoramento. Cada coordenada principal é uma combinação linear das variáveis originais. Além disso, são independentes entre si e estimadas com o propósito de reter, em ordem de estimação, o máximo da informação, em termos de variação total, contida nos dados iniciais (Manly, 1994). Quando as dissimilaridades são medidas em termos de distâncias euclidianas, a análise de componentes principais equivale à de coordenadas principais (Dunn & Everitt, 1982).

De forma geral, a viabilidade de utilizar coordenadas principais em estudos de divergência genética depende da possibilidade de resumir o conjunto de variáveis originais em poucas coordenadas, o que significa ter uma boa aproximação do comportamento dos indivíduos (progenitores), oriundo de espaço n -dimensional (n = número de caracteres estudados) para um espaço bi- ou tridimensional, simplificando a interpretação dos resultados.

Cada coordenada explica um percentual da variação total e as duas primeiras coordenadas são utilizadas quando elas envolvem pelo menos 80% da variação total (Cruz & Regazzi, 1994). No caso deste limite não ter sido atingido nas duas primeiras coordenadas, a análise é complementada com a dispersão gráfica em relação ao terceiro e quarto eixos.

Utilizou-se, para fim de análise das similaridades das 18 linhagens endogâmicas, um sistema de três coordenadas principais. Todas as análises dos dados moleculares foram obtidas utilizando-se o pacote de análise estatística NTSYS (“Numerical Taxonomy and Multivariate Analysis System”), versão 1,70 (Rohlf, 1992).

4.3.8. Correlação dos dados moleculares (genéticos) com os dados de performance dos híbridos simples

Distâncias genéticas (GD) foram calculadas a partir das medidas de similaridade (GS) como: $GD = 1 - GS$. As distâncias genéticas para cada par de linhagens endogâmicas foram correlacionadas com os seguintes caracteres de interesse agrônomo, bem como com a heterose avaliada para cada uma destas características:

- Produção de grãos (t/ha); • Altura de planta (cm/planta); • Altura de espiga (cm/planta).

O coeficiente de correlação utilizado foi o linear (ρ) e sua estimativa foi indicada por r , calculado pela fórmula:

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{[\sum X^2 - \frac{(\sum X)^2}{N}] [\sum Y^2 - \frac{(\sum Y)^2}{N}]}} , \text{ onde X e Y são as variáveis e r varia de } [+1, -1].$$

Inicialmente, as distâncias genéticas obtidas para todos os pares de linhagens endogâmicas foram correlacionadas com os dados de campo obtidos para os respectivos híbridos. Depois, outra metodologia foi adotada mediante a separação das distâncias genéticas de acordo com o padrão de agrupamento observado no dendrograma final e com o método de ordenação. As linhagens 01 e 16 foram excluídas do cálculo das correlações entre grupos, devido à falta de registros de campo apropriados para o cruzamento destas linhagens com outras num mesmo conjunto de experimentos. Ainda, para o cálculo da correlação entre os grupos GI x GIII, foi também desconsiderado o par de linhagens irmãs 12 x 11.

RESULTADOS

5. RESULTADOS

5.1. Análise molecular via RAPD

5.1.1. Otimização das condições de amplificação

O DNA genômico da linhagem 01 de milho foi escolhido para a otimização das condições de amplificação. Foram avaliadas concentrações de 10 ng, 20 ng, 30 ng, 40 ng e 50 ng (com três repetições para cada) utilizando o “primer” OPB-01 (GTTTCGTCC) e 2,0 unidades (U) de enzima (Figura 1A). O perfil de bandas gerado foi o mesmo de 10 ng a 30 ng, apresentando alteração com o aumento da quantidade de DNA (perda de bandas de alto peso molecular). Desta forma, estabeleceu-se a concentração ótima de trabalho como sendo 25 ng de DNA/25 µL de reação de amplificação. Foram feitas sempre duas amplificações com a mesma quantidade de DNA (25ng) para cada linhagem e apenas as bandas reprodutíveis nas duas repetições foram selecionadas e interpretadas como marcadores polimórficos.

Reações com esta quantidade de DNA foram desenvolvidas avaliando-se 2,0 e 3,0 unidades de enzima (Figura 1B). Foi encontrado um perfil reprodutível de bandas com 2,0 U tendo-se adotado esta concentração para as demais amplificações. Em seguida, foram avaliadas diferentes concentrações de magnésio (2,5 mM, 3,0 mM, 3,5mM, 4,0 mM e 4,5mM). O melhor padrão de bandas foi obtido com 3,0 mM (Figura 1C). Um aumento na concentração de magnésio produziu bandas menos reprodutíveis. Isto é esperado já que a concentração de íons

magnésio afeta, entre outras coisas, o anelamento dos “primers”, a dissociação das fitas e a especificidade do produto final (Innis & Gelfand, 1990).

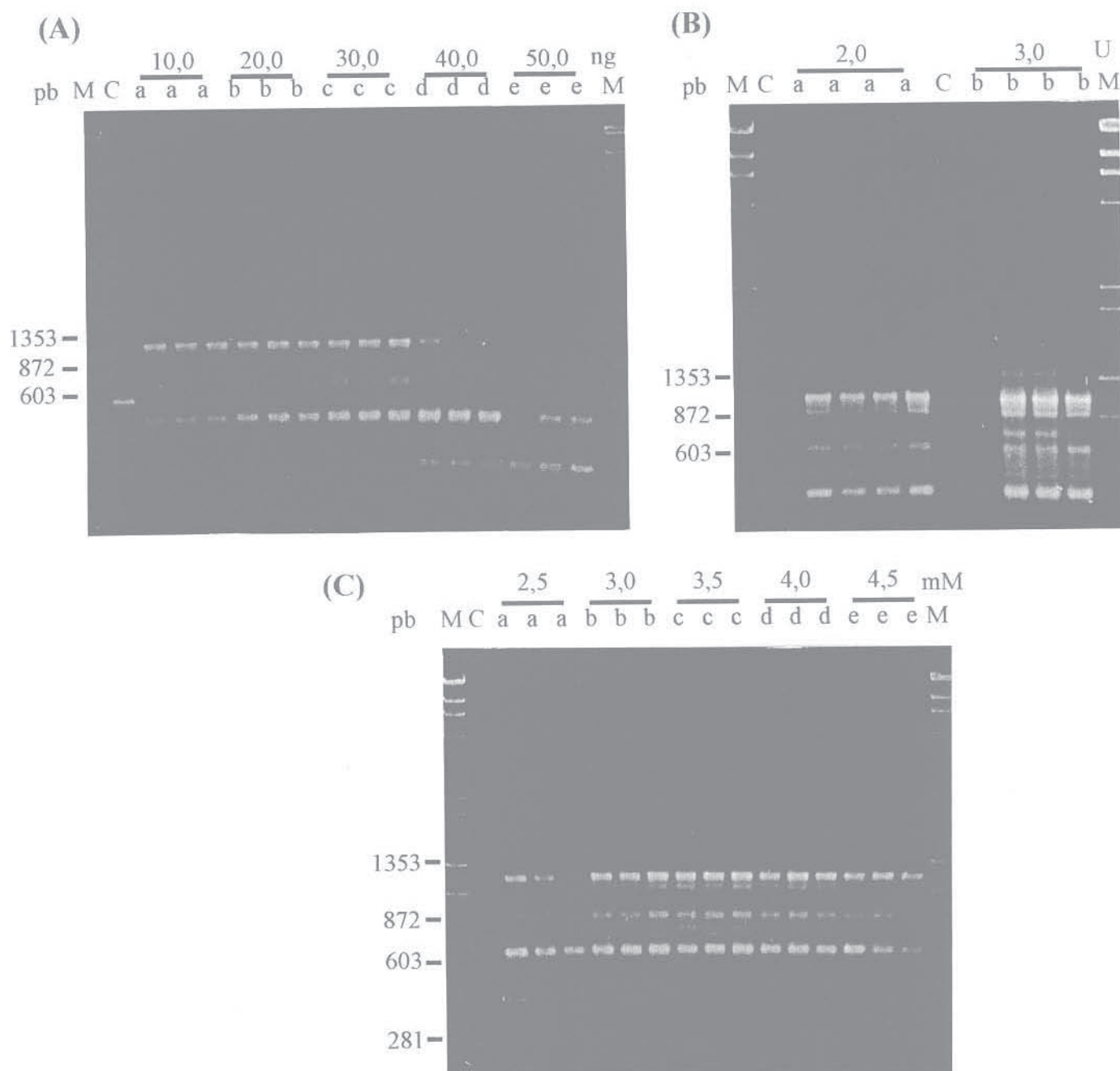


Figura 1A-B-C: (A) Efeito da quantidade de DNA (10, 20, 30, 40 e 50 ng) sobre o perfil de RAPD da linhagem 01 com o "primer" OPB-01 e com 2,0 unidades (U) de enzima; (B) Efeito da quantidade de enzima (2,0 e 3,0 U) sobre o perfil de RAPD da linhagem 01 com o "primer" OPB-01(GTTTCGCTCC) e 30 ng de DNA; (C) Efeito da quantidade de magnésio (2,5; 3,0; 4,0 e 4,5 mM) sobre o perfil de RAPD da linhagem 01 com o "primer" OPB-01 e 25 ng de DNA. (M) marcador de peso molecular λ HindIII/ ϕ x174HaeIII; C é o controle sem DNA.

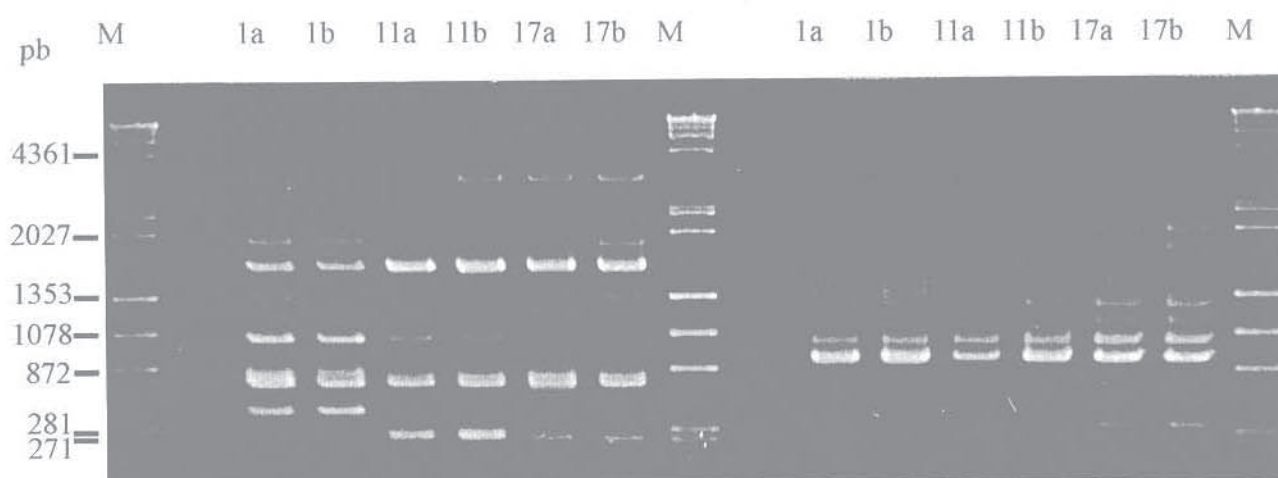
4.1.2. Seleção de “primers”

Após os procedimentos de otimização da reação de amplificação, foi feita uma seleção entre 68 “primers” sendo 19 do Kit B (OPB-02 a OPB-20), 19 do Kit C (OPC-02 a OPC-20), 20 do Kit D (OPD-01 a OPD-20) e 10 do Kit E (OPE-01 a OPE-10). Esta seleção foi feita utilizando-se DNA de apenas 03 linhagens tomadas arbitrariamente: 01, 11 e 17. Com este procedimento foi possível conhecer o perfil dos “primers” relacionados acima, viabilizando a seleção dos mais polimórficos para serem usados em amplificações com todas as linhagens (Figuras 2A e 2B). Os “primers” que revelaram monomorfismos não foram usados para a análise das progênies.

Cerca de 10,2% dos “primers” testados foram completamente monomórficos para as três linhagens. Do total de 604 bandas, 253 foram polimórficas correspondendo à 41,9% de polimorfismo. O número de bandas para cada “primer” variou de 03 a 21 com uma média de 9 bandas por “primer”. Todos os “primers” testados (com exceção de OPD-14) apresentaram um padrão reprodutível de bandas. O peso molecular dos produtos amplificados variou de 230 pb a 3000 pb.

A distância genética destas 03 linhagens foi estimada usando o coeficiente de Jaccard (Anexo B). A linhagem 11 apresentou 52% de similaridade em relação à linhagem 17, enquanto a linhagem 01 proveniente de outra população (BR105) apresentou apenas 6,8% de similaridade em relação às outras duas (Figura 3). Estes dados foram obtidos a partir do experimento de seleção de “primers”. Eles foram analisados com o objetivo de ter-se uma idéia preliminar sobre o comportamento do RAPD em linhagens de milho.

(A)



(B)

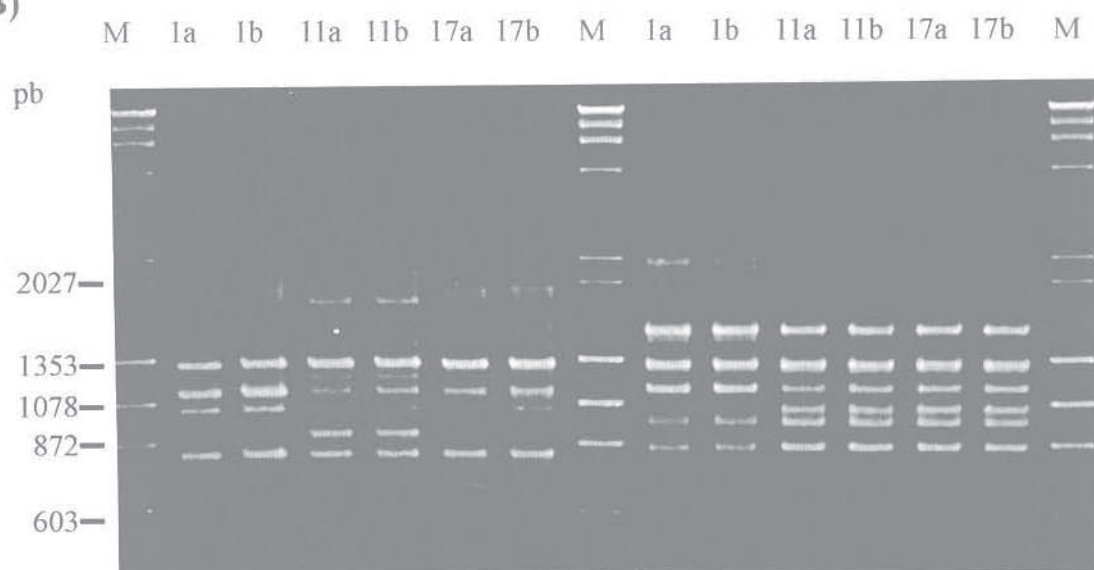


Figura 2A-B: (A) Perfil de RAPD para 03 linhagens endogâmicas de milho (01, 11 e 17, respectivamente) com dois "primers" para fins de seleção: OPC-07 (GTCCCGACGA) e OPC-08 (TGGACCGCTG); (B) Perfil de RAPD para as mesmas 03 linhagens descritas acima com dois outros "primers" para fins de seleção: OPE-01 (CCCAAGGTCC) e OPE-02 (GGTGCGGGAA); (a) e (b) representam duas repetições com 25 ng de DNA cada uma e (M) é marcador de peso molecular λ HindIII/ ϕ x174HaeIII.

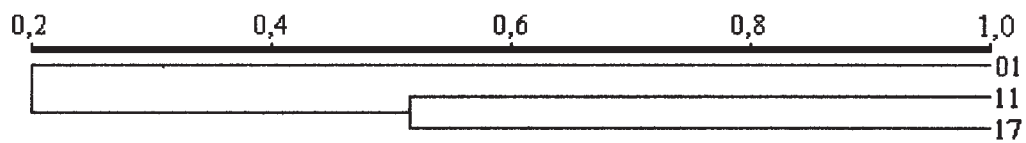


Figura 3. Dendrograma de 03 linhagens endogâmicas de milho (*Zea mays* L.) usadas para fins de seleção de “primers”. As distâncias genéticas foram obtidas pela análise de 253 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens nos grupos GI (verde) e GIII (azul) do dendrograma obtido após a análise das 18 linhagens (Figura 12).

4.1.3- Produtos do RAPD das 18 linhagens

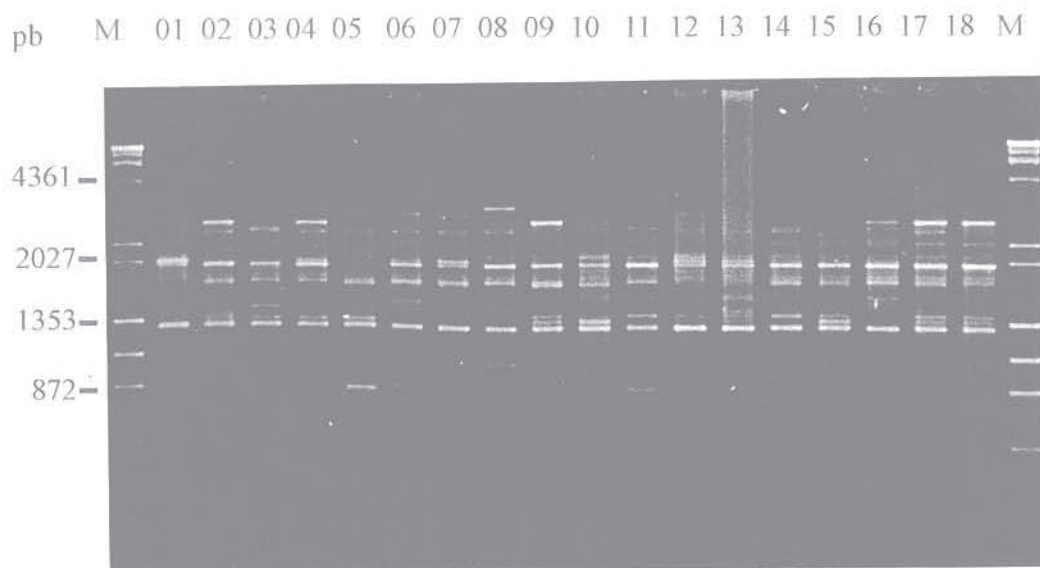
O DNA genômico de cada linhagem foi amplificado em duas repetições, seguindo-se as condições de amplificação estabelecidas, com um total de 32 “primers” incluindo 16 do Kit B (OPB-03 a OPB-18), 8 do Kit C (OPC-03, OPC-04, OPC-05, OPC-07, OPC-09, OPC-13, OPC-14, e OPC-16), 6 do Kit D (OPD-01, OPD-02, OPD-10, OPD-11, OPD-19, OPD-20) e 2 do Kit E (OPE-01 e OPE-08), (Tabela 2). Um total de 325 produtos de amplificação legíveis e reprodutíveis foram gerados, sendo que destes, 262 foram polimórficos. Isto representa 80,62% de polimorfismo gerado para estes dados. O número de bandas para cada “primer” variou de 3 a 19, obtendo-se um número médio de 11 bandas por “primer” (Figuras 4, 5 e 6). Os produtos amplificados apresentaram a mesma amplitude de peso molecular já verificada.

Tabela 2. “Primers” utilizados no RAPD das 18 linhagens endogâmicas de milho.

“Primer” (Operon)	Seqüência (5'→3')	Bandas totais (n ° de bandas polimórficas)*
OPB-03	CATCCCCCTG	13 (11)
OPB-04	GGACTGGAGT	11 (07)
OPB-05	TGCGCCCTTC	11 (05)
OPB-06	TGCTCTGCCC	09 (08)
OPB-07	GGTGACGCAG	07 (05)
OPB-08	GTCCACACGG	16 (14)
OPB-09	TGGGGGACTC	13 (12)
OPB-10	CTGCTGGGAC	15 (13)
OPB-11	GTAGACCCGT	15 (13)
OPB-12	CCTTGACGCA	12 (10)
OPB-13	TTCCCCCGCT	11 (09)
OPB-14	TCCGCTCTGG	07 (05)
OPB-15	GGAGGGTGTT	07 (03)
OPB-16	TTTGCCCGGA	09 (05)
OPB-17	AGGGAACGAG	19 (12)
OPB-18	CCACAGCAGT	14 (11)
OPC-03	GGGGGTCTTT	09 (05)
OPC-04	CCGCATCTAC	19 (15)
OPC-05	GATGACCGCC	04 (02)
OPC-07	GTCCCGACGA	11 (09)
OPC-09	CTCACCGTCC	09 (01)
OPC-13	AAGCCTCGTC	14 (10)
OPC-14	TGCGTGCTTG	14 (12)
OPC-16	CACACTCCAG	14 (11)
OPD-01	ACCGCGAAGG	05 (02)
OPD-02	GGACCCAACC	07 (05)
OPD-10	GGTCTACACC	07 (05)
OPD-11	AGCGCCATTG	10 (06)
OPD-19	CTGGGGACTT	11 (11)
OPD-20	ACCCGGTCAC	13 (08)
OPE-01	CCCAAGGTCC	09 (04)
OPE-08	TCACCACGGT	08 (05)

*Apenas bandas reprodutíveis foram contadas.

(A)



(B)

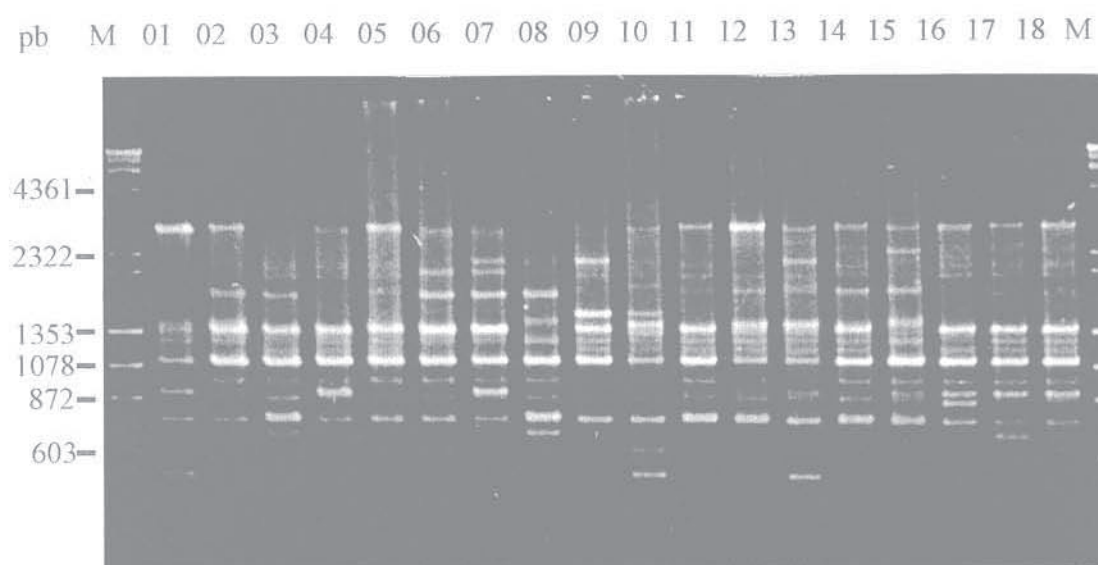


Figura 4A-B: Perfil de RAPD para as 18 linhagens endogâmicas de milho com os seguintes "primers": (A) OPB-09 (TGGGGGACTC) e (B) OPB-11 (GTAGACCCGT), (M) é marcador de peso molecular λ HindIII/ ϕ x174HaeIII.

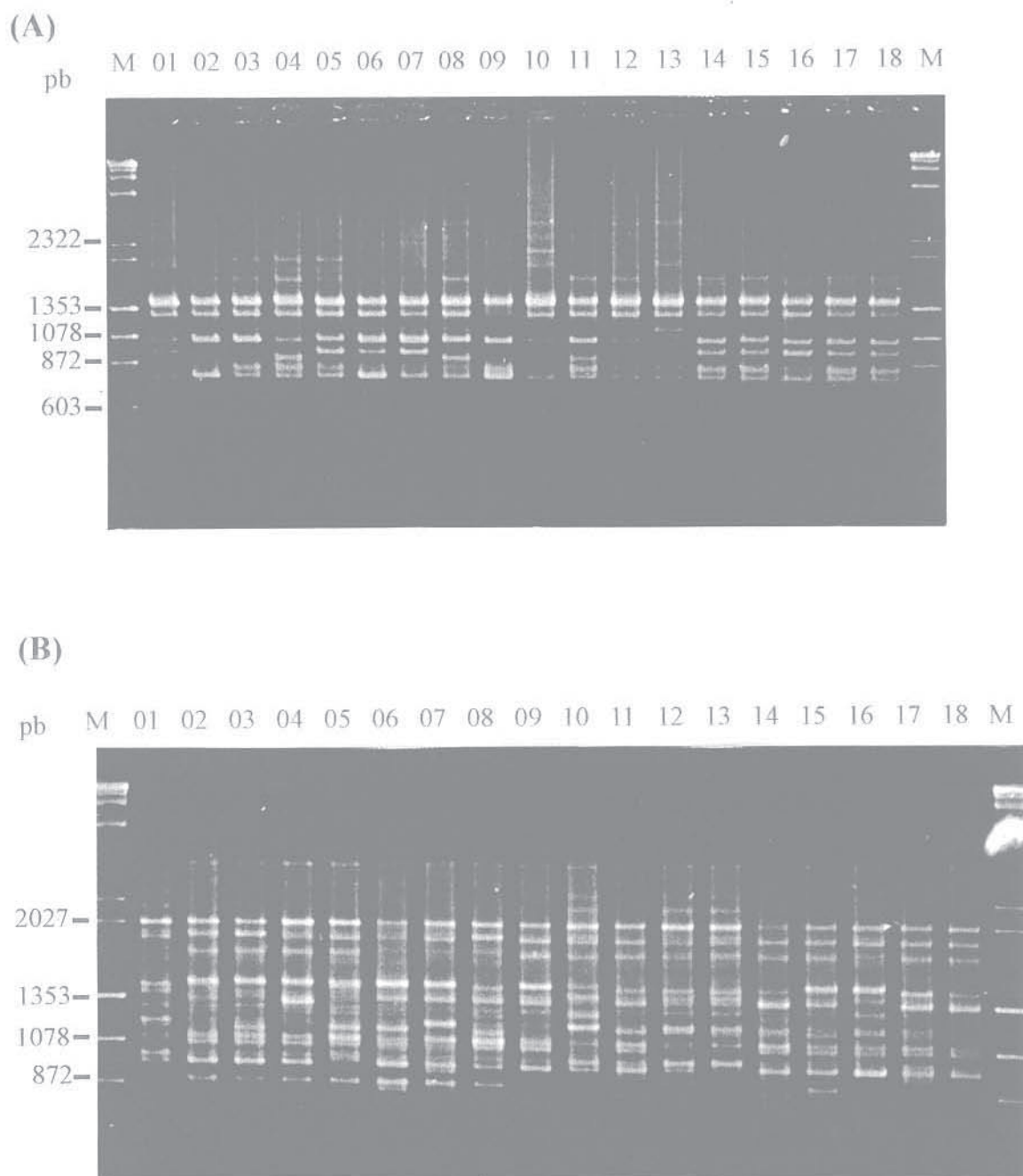
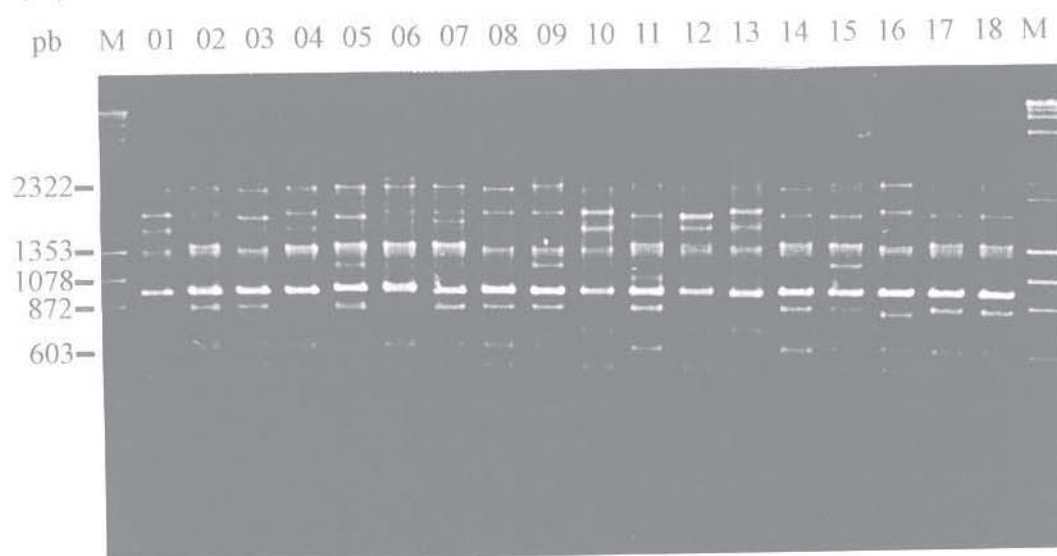


Figura 5A-B: Perfil de RAPD para as 18 linhagens endogâmicas de milho com os seguintes "primers": (A) OPB-12 (CCTTGACGCA) e (B) OPB-17 (AGGGAACGAG). (M) é marcador de peso molecular λ HindIII/ ϕ x174HaeIII.

(A)



(B)

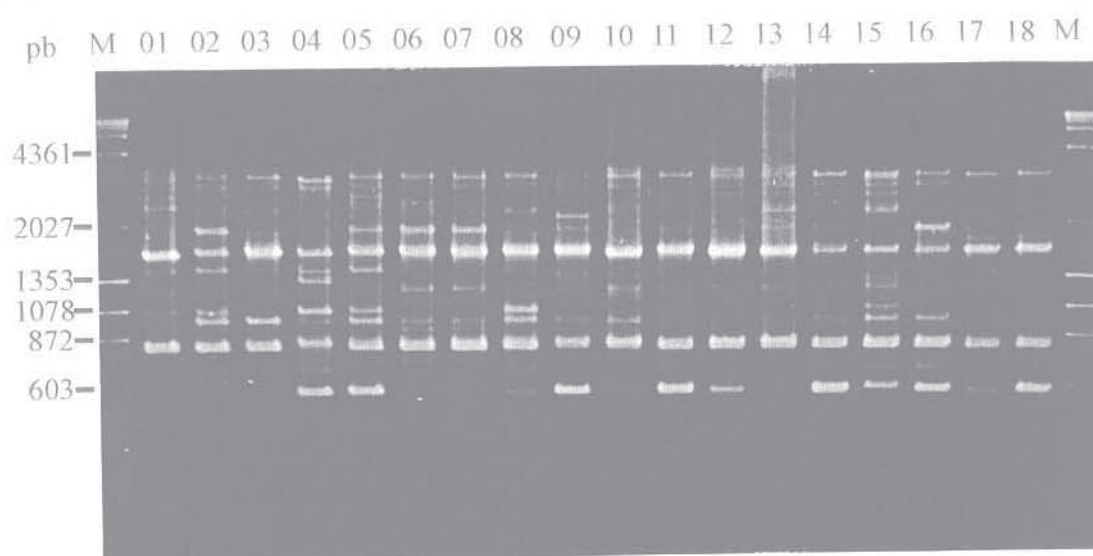


Figura 6A-B: Perfil de RAPD para as 18 linhagens endogâmicas de milho com os seguintes "primers": (A) OPB-18 (CCACAGCAGT) e (B) OPC-07 (GTCCCGACGA). (M) é marcador de peso molecular λ HindIII/ ϕ x174HaeIII.

Os perfis de RAPD das 18 linhagens de milho foram comparados, e a variação nos perfis de bandas foi observada para cada “primer”. As relações genéticas entre as linhagens foram estimadas usando os coeficientes de similaridade de Dice (1945) e Jaccard (1908). Ambos produziram perfis idênticos de agrupamentos com números crescentes de bandas polimórficas. As correlações cofenéticas foram muito próximas, sendo que as geradas por Jaccard foram ligeiramente superiores às de Dice indicando um melhor ajuste dos dendrogramas produzidos por UPGMA em relação às respectivas matrizes de similaridade produzidas por Jaccard, do que em relação àquelas geradas por Dice (Tabela 3). Assim, o coeficiente de Jaccard foi eleito como o melhor para se produzir a matriz de similaridade final dos dados.

Tabela 3. Coeficientes de correlações cofenéticas (r_{cof}) para um número crescente de bandas polimórficas. Dois coeficientes de similaridade (Jaccard e Dice) foram utilizados.

	50 bandas	98 bandas	150 bandas	190 bandas	215 bandas	262 bandas
$r_{\text{cof. Jaccard}}$	0,83	0,88	0,93	0,94	0,95	0,95
$r_{\text{cof. Dice}}$	0,80	0,86	0,91	0,93	0,94	0,94

Dendrogramas com números crescentes de bandas polimórficas (50, 98, 150, 190, 215 e 262) foram construídos a partir de suas respectivas matrizes de similaridades (Anexos C, D, E, F, G e Tabela 4, respectivamente). As Figuras 7, 8, 9, 10, 11 e 12 representam os respectivos dendrogramas. Tal metodologia foi adotada com o objetivo de verificar a reprodutibilidade dos sucessivos agrupamentos gerados, de tal forma a estabelecer o número

necessário de bandas contidas na matriz original de dados para produzir um dendrograma final estável. O dendrograma formado a partir de 262 bandas polimórficas (Figura 12) mostrou a estabilidade desejada para a agrupamento das 18 linhagens endogâmicas e foi considerado como o dendrograma final.

No dendrograma final, o primeiro grupo (Grupo I ou GI) incluiu uma linhagem da população BR-105 (linhagem 01) e três da população BR-106 (linhagens 10, 12 e 13). Este foi o grupo mais consistente, mantendo o mesmo perfil desde o dendrograma de 150 bandas polimórficas (Figura 9). A linhagem 01 apresentou a maior divergência genética, mesmo com relação à linhagens presentes no mesmo grupo, e especialmente para com aquelas do segundo grupo (Grupo II ou GII). O segundo grupo (GII), no dendrograma final, foi composto de sete linhagens da população BR-105 (linhagens 02, 03, 04, 05, 07 e 08) e uma da população BR-106 (linhagem 16) e manteve o mesmo agrupamento desde o dendrograma de 150 bandas polimórficas (Figura 9), exceto pela linhagem 16, que oscilou entre o Grupo III e o Grupo II, permanecendo a partir de 215 bandas polimórficas no Grupo II. O terceiro grupo (GIII) foi constituído apenas de linhagens da população BR-106 (linhagens 09, 11, 14, 15, 17 e 18) no dendrograma final. A composição do Grupo III esteve alterada no dendrograma com 215 bandas (Figura 11) pela saída das linhagens 06 e 07 (sempre agrupadas juntas) para o Grupo II. Estas duas linhagens da população BR-105 mantiveram sua associação com linhagens do GII também nos dendrogramas de 150 e 262 bandas polimórficas.

De forma geral, observou-se a manutenção do agrupamento gerado já com 215 bandas polimórficas (Figura 11) em relação àquele de 190 bandas polimórficas (Figura 10). Houve um aumento progressivo da correlação cofenética entre a matriz de similaridade e a respectiva matriz cofenética produzida para cada um dos vários dendrogramas (Tabela 3).

As matrizes de similaridade dos dendrogramas gerados com números de bandas crescentes foram comparadas e verificou-se também um gradual aumento na correlação das mesmas em relação à matriz gerada com 215 bandas polimórficas. Assim, a matriz de similaridade produzida com 150 bandas (Anexo E) apresentou um coeficiente de correlação de 0,97 em relação à de 215 bandas (Anexo G).

A matriz de similaridade produzida a partir de 190 bandas polimórficas (Anexo F) ajusta-se perfeitamente àquela produzida com 215 bandas polimórficas (Anexo G) ($r = 0,99$), enquanto que a última ajusta-se perfeitamente àquela produzida por 262 bandas polimórficas (Tabela 4) ($r = 0,99$). Isto corrobora a correlação cofenética encontrada para 262 bandas polimórficas e atesta a estabilidade dos coeficientes de similaridade gerados a partir de 150 bandas polimórficas. Houve alta correlação entre o dendrograma gerado para 262 bandas polimórficas (Figura 12) e a matriz de similaridade dos dados ($r_{\text{cof. Jaccard}} = 0,95$) (Tabela 4).

Distâncias genéticas entre linhagens das duas populações (BR-105 e BR-106) variaram de 0,47 a 0,75, apresentando $0,63 \pm 0,08$, como distância genética média. As linhagens da população BR-105 apresentaram distância genética média de $0,58 \pm 0,12$, variando de 0,29 a 0,77, enquanto que as linhagens provenientes da população BR-106 apresentaram distância genética média de $0,59 \pm 0,14$, variando de 0,11 a 0,77. Embora as médias de GDs das duas populações sejam similares, houve uma maior amplitude de variação na população BR-106, indicando que esta população tem maior variabilidade genética (Tabela 5).

O dendrograma final produzido pela análise de agrupamento a partir de 262 bandas polimórficas ($r_{\text{cof. Jaccard}} = 0,95$) subdividiu as linhagens em três grupos heteróticos distintos que foram denominados de Grupo I (GI), Grupo II (GII) e Grupo III (GIII) (Figura 12), cujas distâncias genéticas médias foram de $0,73 \pm 0,06$, $0,52 \pm 0,06$ e $0,46 \pm 0,13$,

respectivamente (Tabela 5). O Grupo I (GI) incluiu uma linhagem da população BR-105 e três outras da população BR-106. A linhagem da população BR-105 (linhagem 01) apresentou as maiores distâncias genéticas em relação às outras linhagens do mesmo grupo (linhagens 10, 12 e 13) e para com outras linhagens da mesma população (BR-105) reunidas no GII. O Grupo II (GII) incluiu 7 linhagens derivadas da população BR-105 (linhagens 02, 03, 04, 05, 06, 07 e 08) e uma derivada da população BR-106 (linhagem 16). O último grupo (GIII) incluiu 6 linhagens (09, 11, 14, 15, 17 e 18) da população BR-106.

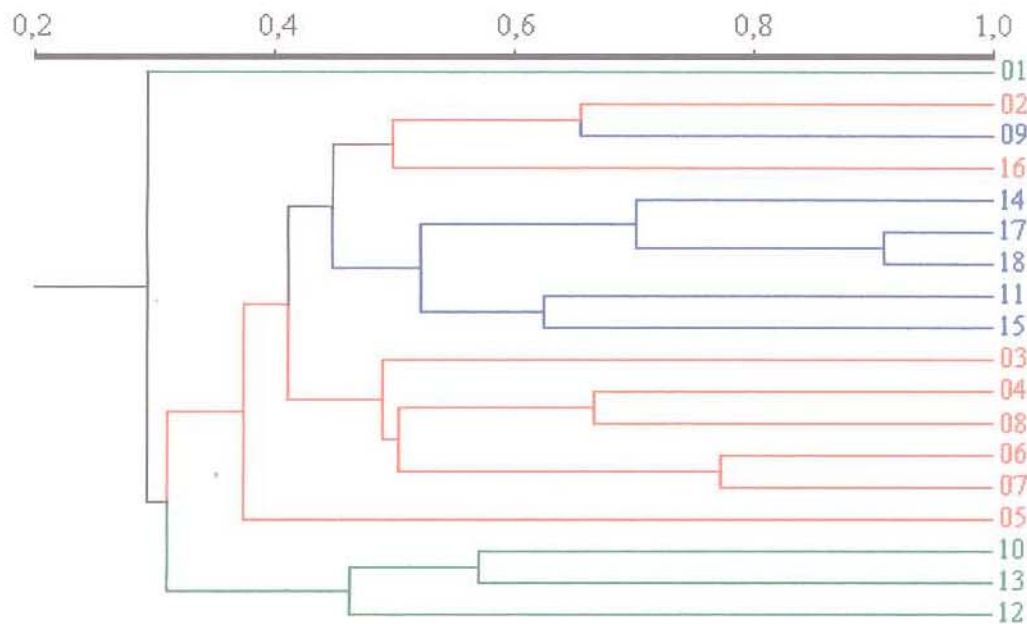


Figura 7: Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 50 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram utilizadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12).

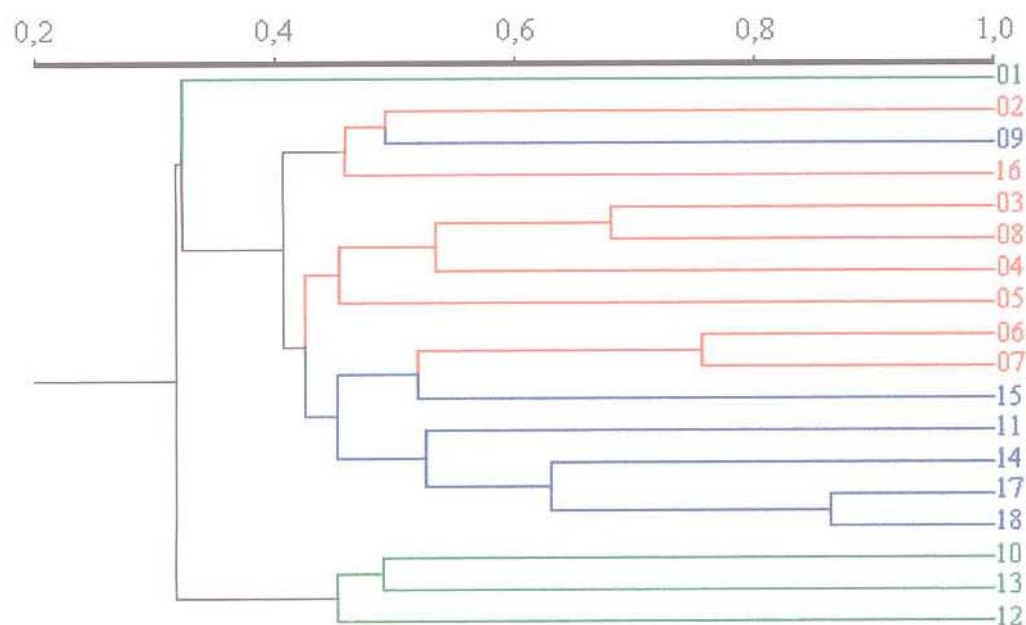


Figura 8: Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 98 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens nos grupos GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12).

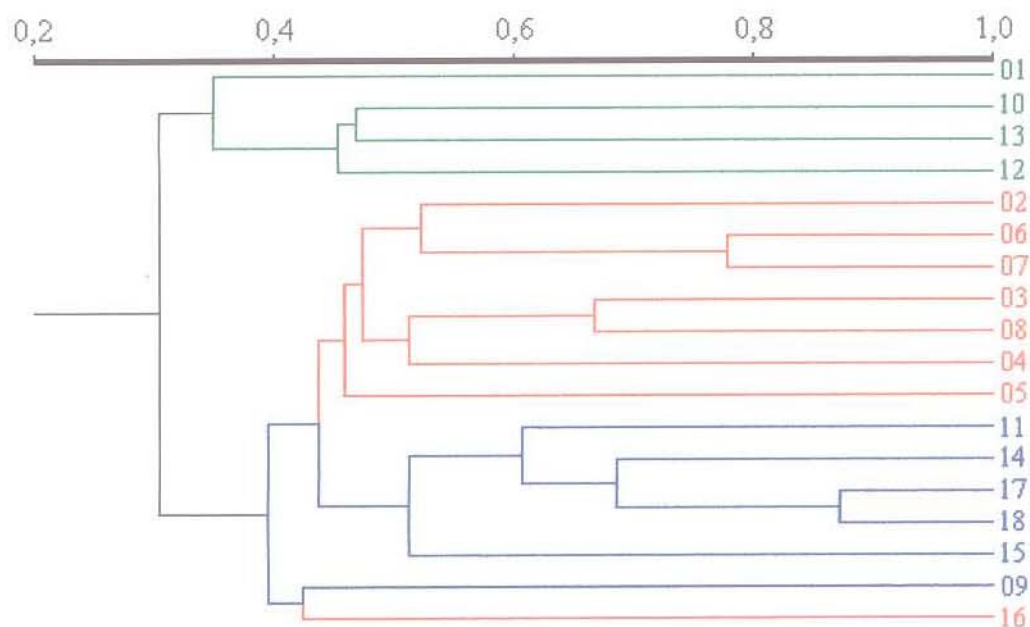


Figura 9: Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 150 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12).

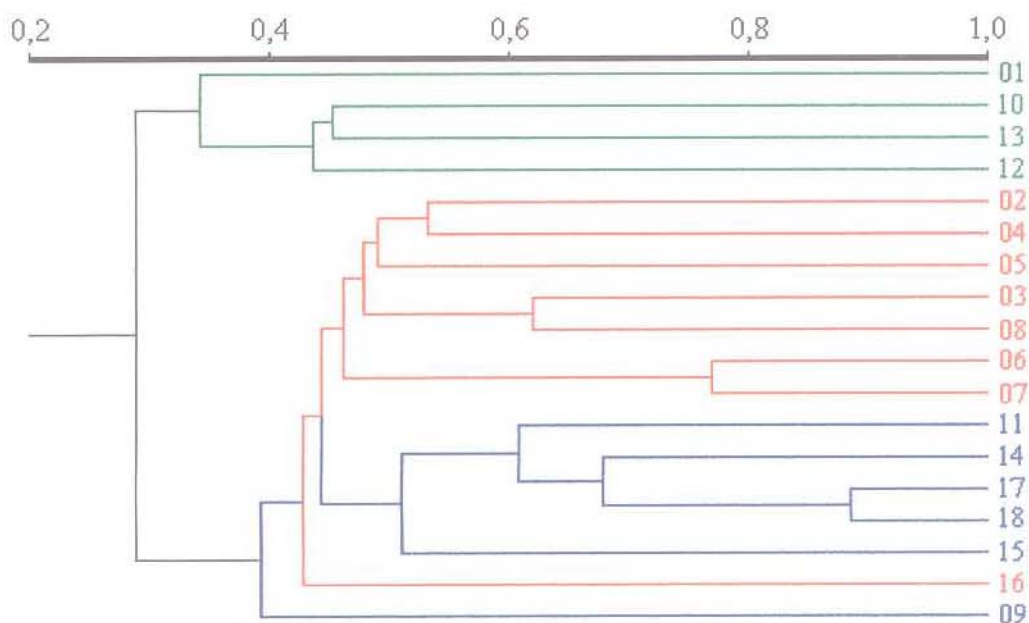


Figura 10: Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 190 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens nos grupos GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12).

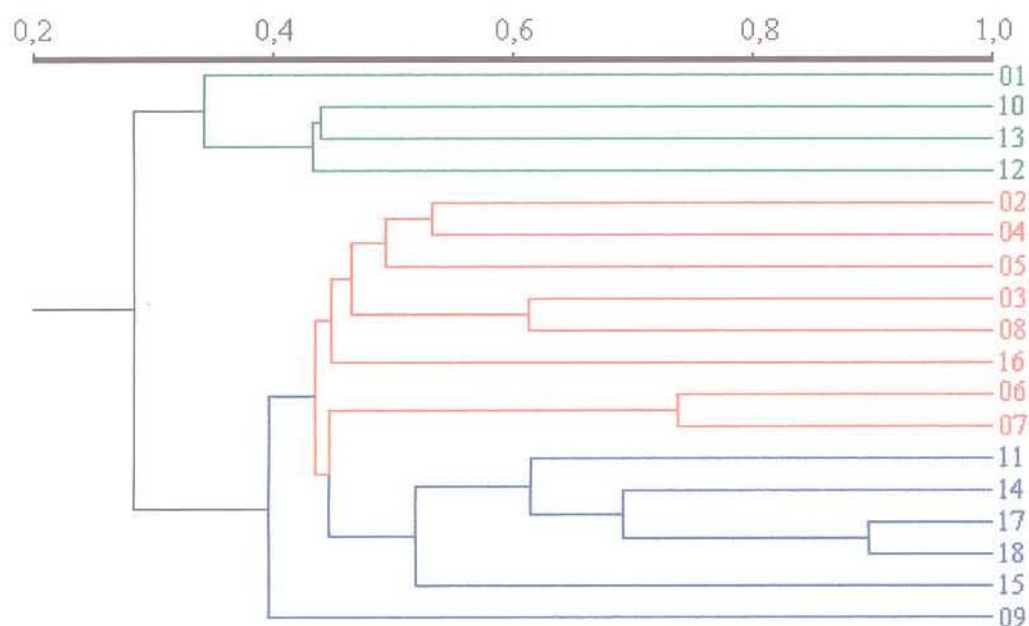


Figura 11: Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 215 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12).

Tabela 4. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (*Zea mays* L.) obtidas para 262 bandas polimórficas.

Linhagens	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
(01) 05-01.4B	0																	
(02) 05-05.2A	0,70	0																
(03) 05-17.1A	0,72	0,50	0															
(04) 05-18.6A	0,69	0,44	0,52	0														
(05) 05-19.1B	0,77	0,46	0,55	0,51	0													
(06) 05-23.2B	0,76	0,50	0,55	0,56	0,54	0												
(07) 05-33.5B	0,75	0,51	0,56	0,58	0,53	0,29	0											
(08) 05-34.2B	0,74	0,52	0,39	0,50	0,48	0,54	0,47	0										
(09) 06-03.5B	0,73	0,54	0,59	0,59	0,57	0,64	0,65	0,61	0									
(10) 06-06.3A	0,62	0,69	0,69	0,66	0,70	0,71	0,72	0,72	0,66	0								
(11) 06-08.1A	0,75	0,55	0,61	0,61	0,55	0,55	0,60	0,62	0,55	0,69	0							
(12) 06-08.2A	0,63	0,70	0,71	0,69	0,72	0,73	0,75	0,73	0,71	0,56	0,55	0						
(13) 06-14.4B	0,71	0,76	0,72	0,73	0,73	0,73	0,75	0,71	0,70	0,55	0,71	0,56	0					
(14) 06-24.7B	0,73	0,54	0,58	0,59	0,60	0,55	0,55	0,59	0,54	0,68	0,35	0,62	0,73	0				
(15) 06-28.1A	0,76	0,55	0,56	0,58	0,51	0,52	0,55	0,55	0,61	0,68	0,49	0,67	0,74	0,46	0			
(16) 06-29.7B	0,68	0,47	0,50	0,54	0,50	0,56	0,55	0,54	0,54	0,72	0,58	0,73	0,77	0,53	0,58	0		
(17) 06-37.5B	0,73	0,50	0,54	0,55	0,53	0,55	0,58	0,56	0,51	0,68	0,38	0,62	0,67	0,29	0,47	0,52	0	
(18) 06-44.1B	0,74	0,55	0,58	0,59	0,56	0,60	0,59	0,58	0,55	0,67	0,39	0,60	0,72	0,31	0,45	0,56	0,11	0

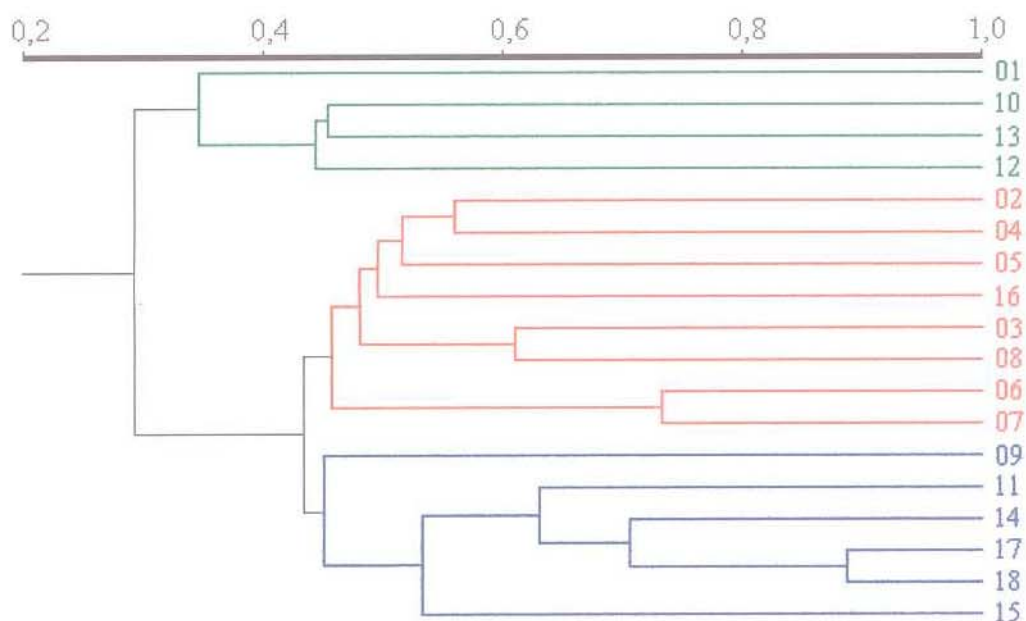


Figura 12. Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 262 bandas polimórficas. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. Os três grupos principais foram ilustrados com cores diferentes: GI (verde), GII (vermelho) e GIII (azul).

Tabela 5. Valores médios, mínimos, máximos e desvios-padrões das distâncias genéticas de cada par de linhagens calculados a partir de RAPD com 262 bandas polimórficas para 18 linhagens endogâmicas de milho.

Grupos de combinações de linhagens	N^a	Médias	Min.	Max.	Desvios-padrões
BR-105 x BR-105	28	0,58	0,29	0,77	0,12
BR-106 x BR-106	45	0,59	0,11	0,77	0,14
BR-105 x BR-106^b	80	0,63	0,47	0,75	0,08
GI x GI	06	0,73	0,55	0,71	0,06
GII x GII	28	0,52	0,29	0,58	0,06
GIII x GIII	15	0,46	0,11	0,61	0,13
GI x GII^c	21	0,75	0,66	0,76	0,02
GI x GIII^d	17	0,72	0,60	0,74	0,04
GII x GIII	42	0,58	0,50	0,65	0,03

^a Número de combinações de linhagens não relacionadas ($f < 0,1$)

^b 80 Híbridos simples F1

^c Excluiu as combinações contendo as linhagens 01 e/ou 16

^d Excluiu as combinações contendo a linhagem 01 e o par 12 x 11 de linhagens irmãs

Com a finalidade de avaliar a precisão e a fidelidade da leitura das bandas dos géis e de garantir que não tenha havido nenhuma distorção no critério de seleção das bandas que foram lidas como polimórficas, uma nova leitura dos géis foi realizada por uma outra pessoa. Em decorrência da diferença de critérios pessoais estabelecidos para leitura das bandas, um total de 300 bandas polimórficas foram lidas desta vez ao invés das 262 anteriores. O coeficiente de similaridade de Jaccard foi novamente utilizado (Anexo H) e o dendrograma correspondente foi gerado através de UPGMA (Figura 13). Também três grupos maiores puderam ser visualizados e mantiveram o mesmo conteúdo das suas linhagens constituintes que os GI, GII e GIII do dendrograma de 262 bandas. Houve apenas uma mudança na posição de agrupamento das linhagens 4 e 5 dentro do segundo grupo em relação ao GII, e da linhagem 9 dentro do terceiro grupo em relação ao GIII. A comparação das matrizes cofenéticas geradas para estes dois dendrogramas (com 300 e 262 bandas polimórficas respectivamente) produziu uma correlação de 0,98, o que indica uma altíssima similaridade entre os agrupamentos gerados.

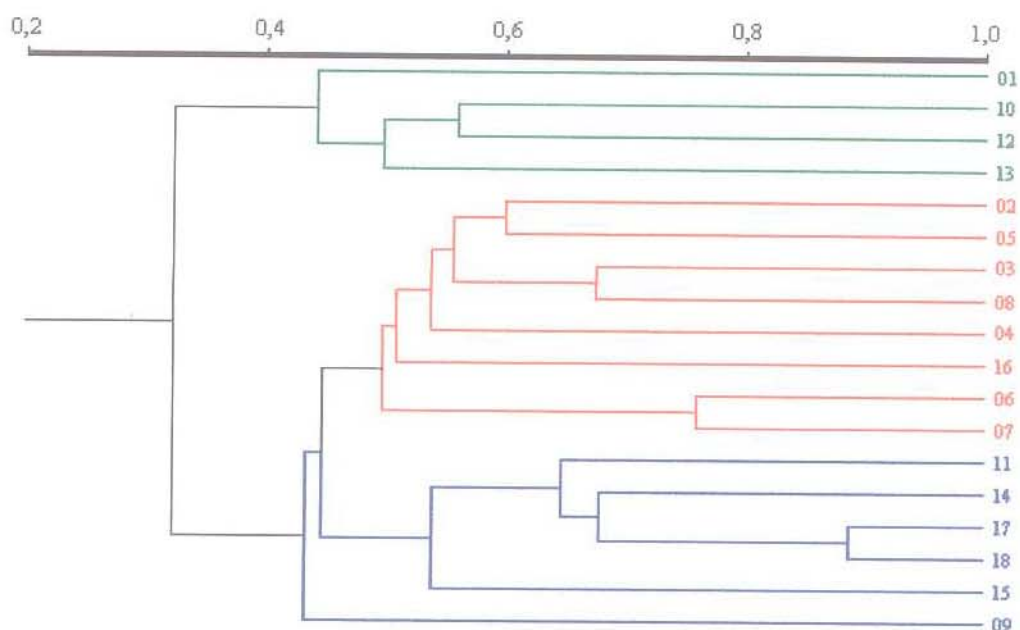


Figura 13: Dendrograma das 18 linhagens endogâmicas de milho (*Zea mays* L.). As distâncias genéticas foram obtidas pela análise de 300 bandas polimórficas através de releitura dos dados. Os coeficientes de similaridade foram gerados de acordo com Jaccard e o agrupamento foi feito segundo o método de UPGMA. As cores foram usadas em relação ao posicionamento das linhagens GI (verde), GII (vermelho) e GIII (azul) no dendrograma final (Figura 12).

A associação entre as linhagens foi também investigada através da análise multivariada. A análise de coordenadas principais foi desenvolvida com base nas similaridades genéticas (Figura 14). As coordenadas principais 1 e 2 (PC1 e PC2) explicaram 27,5% da variância total, enquanto o terceiro eixo (PC3) contribuiu com apenas 8,5% da variância total. Embora nem toda a variância foi explicada, as três coordenadas principais dividiram as linhagens endogâmicas em três grupos heteróticos (Figura 14), refletindo o mesmo padrão de associação em grupos observados no dendrograma final (Figura 12). Os grupos GI e GIII mostraram um perfil mais espalhado de suas linhagens constituintes do que o grupo GII, o que poderia indicar que as linhagens do GII teriam maior similaridade genética do que as dos outros grupos. Entretanto, GII apresentou uma maior distância genética média do que GIII ($52,4\% \pm 6,0$ para GII e $46,1 \pm 13,0$ para o GIII). Como as linhagens 06 e 07 apresentam-se bem distanciadas em relação as demais do Grupo II poder-se-ia assumir que estas estivessem causando este desvio do comportamento esperado. Entretanto, o cálculo da média do Grupo II excluindo as linhagens 06 e 07 produziu um valor bem próximo ao anterior ($53,0\% \pm 4,2$) demonstrando que a premissa anterior não estaria correta.

Note-se que a análise de coordenadas principais com três eixos para estes dados reproduz a separação dos três grandes grupos visualizados no dendrograma final. Entretanto, este sistema de três coordenadas exprimiui apenas uma pequena parte da variância total (36%), e somente com o 11º vetor (dos 18 possíveis) 81,6% da variância total seria explicada. Como um espaço de 11 dimensões torna-se inviável de se analisar graficamente, o uso do espaço tridimensional para ilustrar o esquema de separação maior dos grupos foi aceitável. Cautela, todavia, deve ser mantida frente a análise das variâncias internas à estes grupos, pois poderia

poderia estar havendo sobreposição dos planos vistos a este nível, dando uma falsa idéia de homogeneidade interna em um ou outro grupo.

A análise dos três grandes grupos em relação às três coordenadas principais mostra que a distância genética do GI para o GII e para o GIII é maior do que a observada entre os grupos GII e GIII. A estimativa das médias de distâncias genéticas entre os pares de combinações de linhagens GI x GII, GI x GIII e GII x GIII foi de $75,3\% \pm 2,4$, $72,4\% \pm 3,9$ e $58,2 \pm 3,4$, respectivamente. Desde que haja uma correlação positiva entre a heterose da geração F1 e a divergência genética parental (Moll *et al.*, 1962, 1965; Paterniani & Lonnquist, 1963), poder-se-ia esperar que os híbridos obtidos de cruzamentos entre linhagens endogâmicas de GI e GII ou GI e GIII teriam uma melhor performance do que híbridos obtidos de cruzamentos entre linhagens GII e GIII.

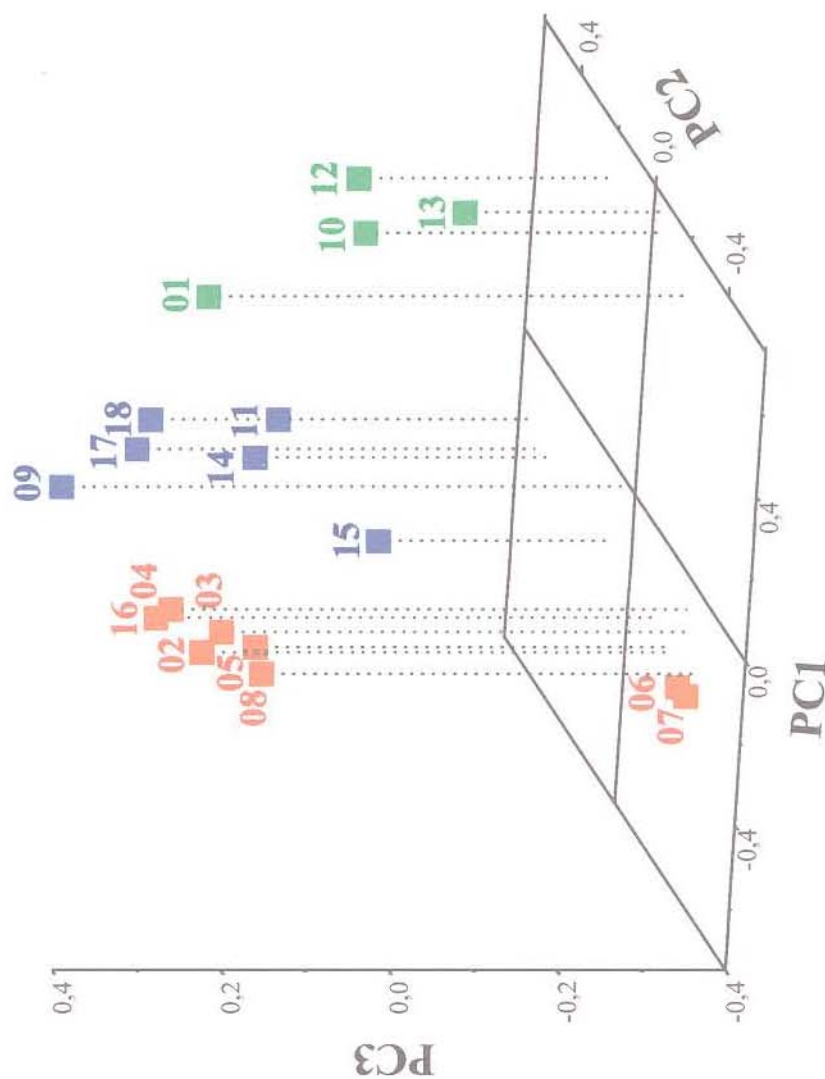


Figura 14: Associações entre as 18 linhagens endogâmicas de milho (*Zea mays* L.) reveladas por análise de coordenadas principais (PCO); baseadas nos coeficientes de similaridade obtidos para 262 bandas polimórficas por RAPD. As cores seguem o padrão da Figura 12: GI (verde), GII(vermelho) e GIII (azul).

5.1.4. Correlação entre distância genética e performance do híbrido

Distâncias genéticas de GI x GII foram correlacionadas com seus cruzamentos obtidos de uma avaliação de 80 híbridos F1 interpopulacionais (Tabela 6). O mesmo procedimento foi usado para as distâncias entre GII x GIII e GI x GIII, sendo que as últimas foram correlacionadas com outro conjunto de experimentos, conduzidos para cruzamentos intrapopulacionais (Tabela 6).

Nenhuma correlação consistente foi encontrada entre as GDs e os caracteres agronômicos avaliados para híbridos simples, produzidos a partir das 18 linhagens endogâmicas, quando a divisão de grupos foi desconsiderada. A Figura 15A ilustra o padrão encontrado entre as GDs totais e a produção de grãos para F1 ($r = 0,16$). A heterose para a produção total de grãos também foi correlacionada com GDs, mas produziu correlação inconsistente e negativa, isto é, $r = -0,15$ (Tabela 8).

Quando os grupos foram considerados, as GDs entre GI x GII apresentaram elevado valor de correlação ($r = 0,70$) com a produção de grãos em F1, mas ainda três cruzamentos mostraram produção abaixo da média e distância genética acima da média de distância (Figura 15B - Quadrante IV). Os pares de linhagens GI x GII apresentaram baixa correlação para outros caracteres agronômicos (altura de planta e altura de espiga) ou com a heterose medida para estas características. As distâncias genéticas entre GI x GIII apresentaram alta correlação com a produção de grãos para F1 ($r = 0,87$) e nenhuma distância genética acima da média associada com produção de grãos abaixo da média (Figura 15C - Quadrante IV).

Tabela 6. Produção de grãos para os híbridos F1 e heterose para cruzamentos inter- e intrapopulacionais. GDs indicam as distâncias genéticas entre os parentais.

Cruzamento	GD	Híbrido (t/ha)	Heterose (t/ha)	Cruzamento	GD	Híbrido (t/ha)	Heterose (t/ha)
Interpopulacionais (06 ^a x 05 ^b)				Interpopulacionais (06 ^a x 05 ^b)			
09 x 01	0.73	7.10	4.82	15 x 03	0.56	5.15	3.24
09 x 02	0.54	6.73	4.55	15 x 04	0.58	7.89	5.87
09 x 03	0.59	6.15	3.91	15 x 05	0.51	8.30	6.40
09 x 04	0.59	7.39	5.03	15 x 06	0.52	7.07	5.69
09 x 05	0.57	5.68	3.45	15 x 07	0.54	7.17	5.32
09 x 06	0.64	6.98	5.26	15 x 08	0.55	6.55	4.93
09 x 07	0.65	7.42	5.23	16 x 01	0.68	6.33	4.46
09 x 08	0.61	6.67	4.72	16 x 02	0.47	5.90	4.14
10 x 01	0.62	6.03	3.88	16 x 03	0.56	4.65	2.84
10 x 02	0.69	6.19	4.14	16 x 04	0.54	3.81	1.87
10 x 03	0.69	5.96	3.86	16 x 05	0.50	7.20	5.38
10 x 04	0.66	6.03	3.80	16 x 06	0.56	6.53	5.23
10 x 05	0.70	6.77	4.66	16 x 07	0.55	5.79	4.03
10 x 06	0.71	6.90	5.31	16 x 08	0.54	6.19	4.66
10 x 07	0.72	6.80	4.75	17 x 01	0.73	6.72	4.11
10 x 08	0.72	6.29	4.46	17 x 02	0.50	6.61	4.10
11 x 01	0.74	7.13	4.6	17 x 03	0.54	6.42	3.85
11 x 02	0.55	6.60	4.17	17 x 04	0.55	6.93	4.24
11 x 03	0.61	6.86	4.37	17 x 05	0.53	7.07	4.51
11 x 04	0.61	6.95	4.34	17 x 06	0.55	6.85	4.81
11 x 05	0.55	7.60	5.12	17 x 07	0.58	7.35	4.84
11 x 06	0.55	8.00	6.04	17 x 08	0.56	6.69	4.01
11 x 07	0.60	7.44	5.00	18 x 01	0.74	6.49	4.00
11 x 08	0.62	7.16	4.95	18 x 02	0.55	7.20	4.64
12 x 01	0.63	6.79	4.33	18 x 03	0.58	6.35	6.78
12 x 02	0.70	6.57	4.22	18 x 04	0.59	6.74	4.18
12 x 03	0.71	6.60	4.20	18 x 05	0.56	6.78	4.35
12 x 04	0.69	6.60	4.07	18 x 06	0.60	7.31	5.40
12 x 05	0.72	7.59	5.19	18 x 07	0.59	7.00	4.60
12 x 06	0.73	7.68	5.79	18 x 08	0.58	6.92	4.77
12 x 07	0.75	7.82	5.47	Intrapopulacionais (06 ^a x 06 ^a)			
12 x 08	0.73	6.88	4.76	10 x 09	0.66	5.66	3.66
13 x 01	0.71	7.52	3.91	10 x 11	0.69	6.94	4.69
13 x 02	0.76	6.80	3.30	10 x 14	0.68	6.19	4.10
13 x 03	0.72	6.60	3.04	10 x 15	0.68	6.82	5.14
13 x 04	0.73	6.97	3.29	10 x 17	0.68	6.57	5.56
13 x 05	0.73	7.41	3.86	10 x 18	0.67	7.75	6.82
13 x 06	0.73	7.47	4.43	12 x 09	0.71	6.55	4.24
13 x 07	0.75	7.72	4.21	12 x 11	0.55	1.49	-1.06
13 x 08	0.71	7.16	3.89	12 x 14	0.62	3.71	1.31
14 x 01	0.73	6.69	4.31	12 x 15	0.67	6.09	4.09
14 x 02	0.54	7.16	4.89	12 x 17	0.62	4.25	1.61
14 x 03	0.58	6.84	4.51	12 x 18	0.60	3.80	1.29
14 x 04	0.60	6.82	4.37	13 x 09	0.70	6.87	3.41
14 x 05	0.54	7.53	5.21	13 x 11	0.71	7.08	3.37
14 x 06	0.52	6.35	4.54	13 x 14	0.73	7.68	4.13
14 x 07	0.55	6.92	4.64	13 x 15	0.74	7.22	4.07
14 x 08	0.59	6.75	4.71	13 x 17	0.70	7.78	3.99
15 x 01	0.76	6.94	4.99	13 x 18	0.72	7.00	3.34
15 x 02	0.55	7.50	5.66				

^a 06: BR-106/ ^b 05: BR-105

Tabela 7. Correlações lineares (r) entre distâncias genéticas obtidas por RAPD e dados de campo para três caracteres de importância econômica e para suas respectivas heteroses.

Grupos	Altura de Planta (AP) ^{*1}		Altura de Espiga (AE) ^{*2}		Produção de Grãos (PG) ^{*3}	
	Híbrido AP	Heterose AP	Híbrido AE	Heterose AE	Híbrido PG	Heterose PG
80 dados	0,05	0,11	0,02	0,18	0,16	-0,15
GI x GII	0,23	0,19	0,22	0,19	0,70	0,13
GI x GIII	0,57	0,45	0,33	0,18	0,87	0,63
GII x GIII	0,11	0,09	0,06	0,15	0,00	-0,05

^{*1}cm/planta / ^{*2} cm/planta / ^{*3} t/ha

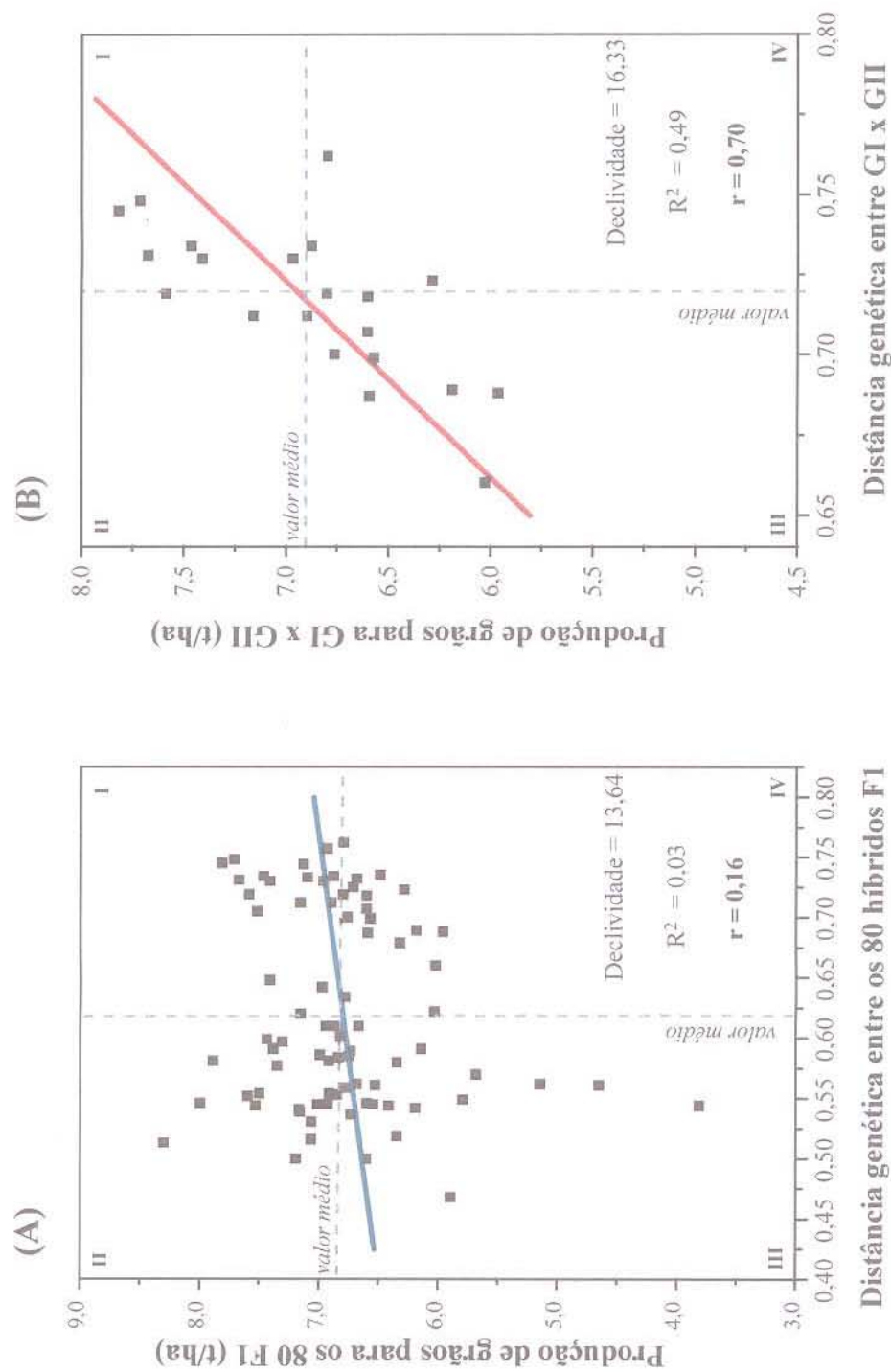


Figura 9 A-B: Gráficos de produção de grãos (t/ha) *versus* distâncias genéticas (baseadas nos dados de RAPD com 262 bandas polimórficas) para: (A) 80 híbridos simples interpopulacionais e (B) GI x GII

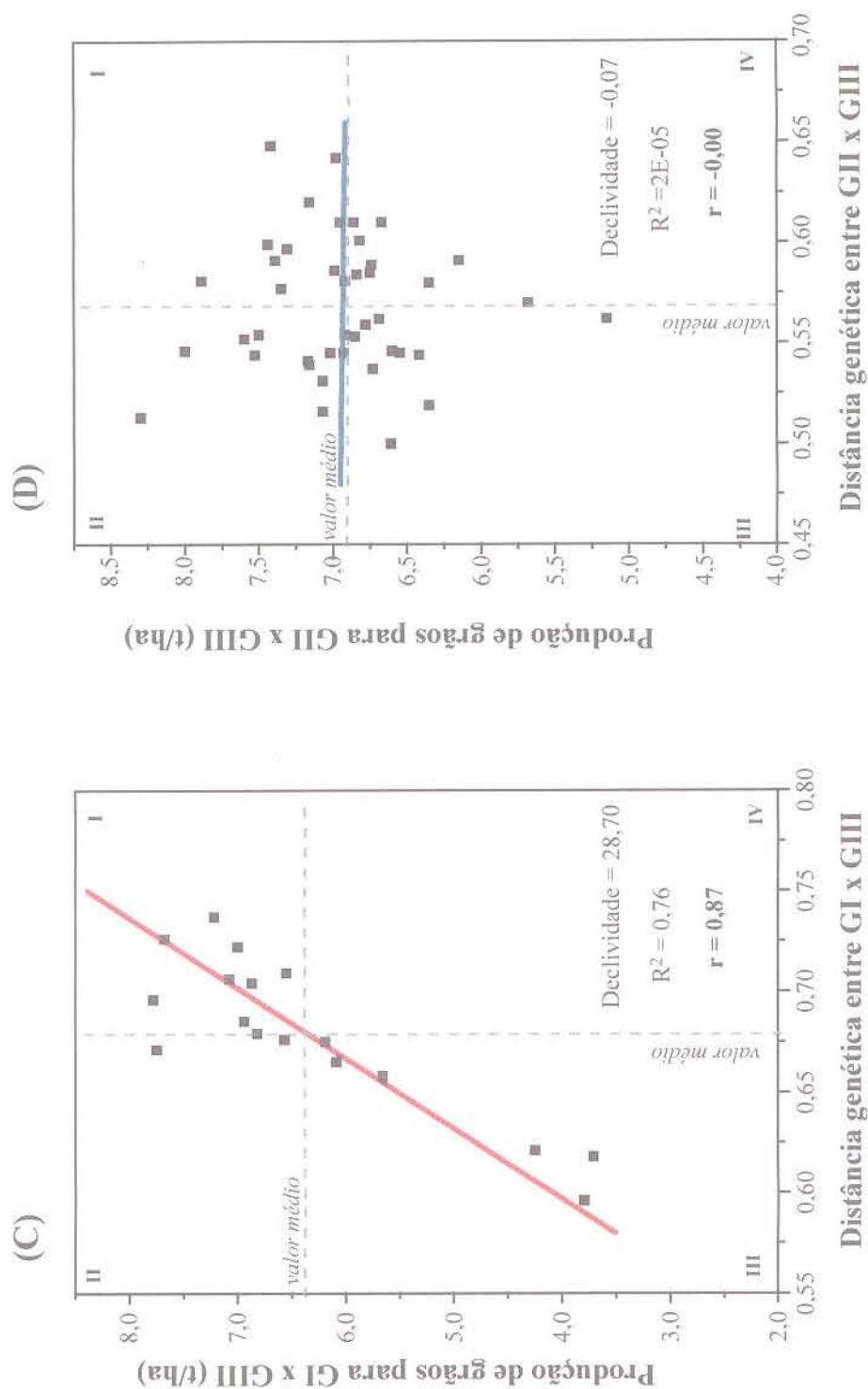


Figura 9 C-D: Gráficos de produção de grãos (t/ha) *versus* distâncias genéticas (baseadas nos dados de RAPD com 262 bandas polimórficas) para: (C) G1 x GIII e (D) GII x GIII

Todas as altas distâncias genéticas se mostraram associadas com altas produções quando envolvem cruzamentos de linhagens do GI x GII e GI x GIII. A heterose para produção de grãos correlacionou-se positivamente com os pares de linhagens GI x GIII ($r = 0,63$, Tabela 7). O caráter altura de planta também exibiu uma melhor associação dos pares GI x GIII com híbridos F1 ($r = 0,57$, Tabela 7) do que com os outros pares de linhagens e com a heterose para esta característica ($r = 0,45$) do que com os pares GI x GII. Para as distâncias GII x GIII, nenhuma correlação consistente foi encontrada para qualquer um dos caracteres avaliados; a Figura 15D ilustra esta ausência de correlação para produção de grãos. O caráter altura de espiga mostrou baixa associação com todas as combinações de linhagens examinadas (Tabela 7).

DISCUSSÃO

6. DISCUSSÃO

Melhoristas de milho utilizam-se de um certo grau de subjetividade para a seleção de alguns caracteres (tipo de planta, conformação de espiga, cor e tipo do endosperma, etc...); outros, entretanto, são mais objetivos por requererem o uso de índices simplificados de seleção (produção de grãos, maturidade, resistência à doenças, etc...). Paralelamente ao grande número de caracteres a considerar, os melhoristas, ainda têm de descontar os efeitos ambientais atuantes nos genótipos avaliados, o que pode reduzir a efetividade da seleção.

Um dos objetivos primordiais em programas de melhoramento de milho tem sido o de identificar linhagens endogâmicas que produzam cruzamentos exibindo ótimos níveis de heterose. Desde que se demonstrou uma correlação positiva entre a heterose da progênie F1 e a divergência genética dos parentais (Moll *et al.*, 1962, 1965; Paterniani & Lonnquist, 1963), a busca por meios capazes de avaliar a divergência genética de linhagens endogâmicas ou variedades de milho tem sido uma preocupação constante, sobretudo com a finalidade de prever o comportamento da performance híbrida. Com este intuito e com o advento da Biologia Molecular, marcadores do tipo isoenzimas, RFLPs e RAPDs, livres dos efeitos pleiotrópicos característicos dos marcadores fenotípicos, passaram a ser utilizados. Os resultados têm gerado polêmica entre os dois segmentos envolvidos: melhoristas de campo, propensos a assumir, muitas vezes, que marcadores moleculares teriam sua única aplicação na tipagem de variedades, raças ou linhagens e receosos em acreditar nas suas potencialidades práticas, podendo abreviar etapas do melhoramento; geneticistas moleculares, cientes do ritmo acelerado dos avanços das técnicas moleculares e esperançosos de que os obstáculos ainda

existentes possam ser logo sanados. Assim se encaixa o desapontamento de Haullauer (1995) frente à tecnologia de marcadores de DNA, no estado em que se encontra no presente momento. Do outro lado, temos Lee (1995) exaltando os marcadores de DNA como o elo fundamental entre o melhoramento e a biologia de plantas, e afirmando que tais marcadores têm trazido um novo modo de percepção da diversidade genética aos melhoristas.

Bernardo (1994) sugeriu que a produção de híbridos simples não avaliados pode ser eficazmente predita baseando-se em dados de RFLP parental e dados de produção para o conjunto de híbridos relacionados. Segundo Lee (1995), desafios significativos para a predição da performance híbrida têm permanecido devido à alguns fatores: (a) linhagens derivadas de populações base devem ainda ser testadas em inúmeras combinações entre grupos; (b) um grupo heterótico pode ter uma subestrutura genética importante e imperceptível (Melchinger *et al.*, 1991; Livini *et al.*, 1992) ou (c) linhagens endogâmicas de origem não-caracterizada ou misturada podem não se enquadrar em grupos estabelecidos.

Tentativas iniciais de predição da performance híbrida usando proteínas e marcadores de DNA, em milho, foram resumidas por Stuber (1994). Associações entre a performance híbrida e a heterozigosidade predita têm sido maiores para cruzamentos entre linhagens de “pedigrees” similares (Frei *et al.*, 1986; Lee *et al.*, 1989; Smith *et al.*, 1990). Estas observações podem ser um indício dos poderes preditivos relativos ao conhecimento dos estados alélicos: “idênticos por descendência” *versus* “idênticos em estado” (Smith & Smith, 1992). Se “alelos” de marcadores do tipo RFLP ou RAPD forem “idênticos por descendência”, estes devem ser informativos da região do loco e das regiões ligadas à ele. Entretanto, para parentais não-relacionados, “alelos” de marcadores podem ser apenas semelhantes em estado, traduzindo muito pouco sobre seus vizinhos. Assim, a força das associações diminui nos limites superiores

da heterozigosidade predita e da dissimilaridade genética, e também para cruzamentos entre linhagens de “pedigrees” desconhecidos (Melchinger *et al.*, 1990a,b; Godshalk *et al.*, 1990; Dudley *et al.*, 1991, 1992; Boppenmaier *et al.*, 1992). Melchinger (1993) comparou vários trabalhos com milho e observou uma relação íntima entre estimativas de distâncias genéticas e heterose, capacidade de combinação específica ou performance dos híbridos simples somente se o grupo de parentais estivesse relacionado por “pedigree”; as correlações observadas eram fracas ou inexistentes para parentais não-relacionados ($f_{ij} < 0,1$, sendo f_{ij} = coeficiente de parentesco).

A falta de associação entre diversidade medida com marcadores moleculares e performance dos cruzamentos entre linhagens não-relacionadas poderia ser associada à diversas causas, tais como baixa associação entre a heterozigosidade estimada pelos marcadores e a heterose nos QTLs (“Quantitative Trait Loci”) nos cruzamentos examinados ou ainda ocorrência de epistasia. Moser & Lee (1994) observaram que distâncias genéticas baseadas na genealogia e em dados de RFLPs para aveia correlacionavam-se positivamente, mas não se correlacionavam com as distâncias estimadas em caracteres quantitativos. Algumas correlações de baixas magnitudes foram obtidas entre as variâncias genéticas e os dois estimadores (RFLPs e genealógicos) para seis características agrônômicas avaliadas, entre as quais estavam a produção de grãos e altura de plantas. As discrepâncias entre medidas de genealogia e aquelas estimadas por RFLPs poderiam ser esperadas já que ambos os métodos utilizados estão sujeitos à erros de amostragem (poucas sondas e genótipos). Os autores sugeriram que a utilidade dos marcadores de DNA para as predições da variação genética poderia ser ampliada quando: (a) os alelos do marcador de DNA estão suficientemente ligados à QTLs bem caracterizados e (b) as inferências são delimitadas à populações genéticas bem definidas.

Um dos problemas que se encontra no uso de marcadores moleculares é o fato que embora estes representem amostragens do genoma das plantas, são utilizados para inferir similaridades do genoma inteiro, entre um conjunto de linhagens ou populações. O número de marcadores ou sondas utilizados numa análise, e a variabilidade da população de referência, pode afetar diretamente a precisão das estimativas feitas (Moser & Lee, 1994). Autores como Messmer *et al.* (1991) e Smith *et al.* (1992) ressaltaram que no mínimo 100 sondas (locos), no caso de RFLPs, deveriam ser utilizados para gerar estimativas capazes de traduzir relacionamentos genéticos verdadeiros entre linhagens de milho.

Dados de literatura indicam que linhagens de milho requerem mais locos amostrados para gerar estimativas acuradas de distância genética, devido à grande variação dos locos de RFLPs, gerando maiores intervalos de confiança e menores correlações entre subgrupos de sondas (Messmer *et al.*, 1991; Smith *et al.*, 1992; Moser & Lee, 1994). Um levantamento feito revela, por exemplo, que a correlação entre distâncias baseadas em “pedigrees” e distâncias de marcadores moleculares (isoenzimas e RFLPs, no caso) melhora, a medida em que mais sondas ou locos marcadores são empregados na análise. Assim, Eloibu *et al.* (1990) obtiveram uma correlação não significativa de 0,41 baseada em seis sondas em *Drosophila*; Cox *et al.* (1985) obtiveram correlações variando de 0,15 a 0,45 baseadas em 13 marcadores isoenzimáticos e morfológicos em soja; Smith & Smith (1989) obtiveram uma correlação 0,51 baseada em 31 locos isoenzimáticos em milho; Melchinger *et al.* (1991) estimaram uma correlação de 0,74 baseada na análise de dados de RFLP com 83 sondas em milho; e Smith *et al.* (1990) obtiveram uma correlação de 0,90 usando 257 combinações sonda-enzima em milho. Ainda em aveia, espécie relativamente menos polimórfica que milho, uma correlação de 0,63

foi conseguida para 26 sondas de RFLP com dados de “pedigree” estimados pelo coeficiente de parentesco (Moser & Lee, 1994).

Para dados de RAPD, Sobral & Honeycutt (1994) mostraram uma correlação de 0,55 obtida entre as distâncias, estimadas por RAPD com 130 bandas polimórficas e àquelas medidas experimentalmente com dados de campo para 22 linhagens endogâmicas de milho.

Lee (1995) relatou alguns pontos polêmicos sobre as aplicações de marcadores moleculares em estudos básicos e aplicados de plantas quanto ao desconhecimento: (1) das relações alélicas dos fragmentos de DNA usados como dados; (2) das frequências alélicas nas populações de referência; (3) das condições em que um polimorfismo de DNA refletiria uma diferença genética significativa; (4) de como diferentes marcadores de DNA acessariam diferentes porções do genoma e se estas diferenças comprometeriam a amostragem do genoma e interpretação dos dados; (5) da capacidade de prever quando o genoma alvo estaria saturado com marcadores e de se saber a melhor maneira quando isto já ocorreu e (6) dos limites de resolução de um método de detecção de polimorfismo de DNA.

Charcosset *et al.* (1991) propuseram a necessidade de haver um desequilíbrio de ligação entre o grau de heterozigosidade média medida num loco marcador (para isoenzimas e RFLPs) e nos QTLs para que haja associação entre eles. Bernardo (1992) demonstrou a necessidade de certas condições para uma predição efetiva da performance híbrida ocorrer usando a heterozigosidade molecular entre as quais que: (1) 30 a 50% dos QTLs deveriam estar ligados ao marcador molecular e (2) não mais do que 20 a 30% dos marcadores moleculares deveriam estar aleatoriamente dispersos no genoma ou não ligados à QTLs.

De acordo com os resultados obtidos neste trabalho, apenas o caráter produção de grãos correlacionou-se consistentemente com as medidas de distância genética obtidas por

Considerando-se plantas individuais, este é um caráter de herança complexa e com baixo coeficiente de herdabilidade, mas quando estimado através da média de indivíduos, passa a apresentar coeficientes de herdabilidade de valores médios a elevados. Produção de grãos é ainda um caráter que apresenta elevado nível de dominância. Tais características sustentam alguns pressupostos de Bernardo (1992) para que haja uma predição efetiva da performance híbrida, tais como fortes efeitos de dominância e alta herdabilidade do caráter agrônomo avaliado.

Polimorfismos gerados pela reação de polimerase em cadeia (PCR - “Polymerase Chain Reaction”) com oligonucleotídeos arbitrários são baseados em pareamentos irregulares (“mismatches”) nos sítios de anelamentos dos oligonucleotídeos ou por eventos de inserção/deleção, resultando, então, na presença ou ausência de um produto de amplificação de um único loco. RAPDs têm natureza dominante e podem eficientemente gerar marcadores dispersos aleatoriamente por todo genoma, bem como também ligados a genes específicos (Welsh & McClelland, 1990; Williams *et al.*, 1990). Desta forma, considerando-se a natureza da técnica de RAPD podemos supor que um outro requisito proposto por Bernardo (1992), pelo qual no mínimo 30 a 50 % dos QTLs devem estar ligados ao marcador, tenha sido atingido neste trabalho.

Smith & Smith (1989) encontraram muitos híbridos com mesmas distâncias medidas por isoenzimas, mas com grandes diferenças na quantidade de heterose e levantaram a questão de que isto poderia ser devido a inabilidade de se amostrar aleatoriamente todo o genoma quando marcadores bioquímicos são utilizados. A técnica de RAPD poderia atingir qualquer loco que estivesse afetando a produção de grãos aleatoriamente pelo genoma, podendo superar neste ponto, algumas restrições de outros marcadores sítio-específicos. Para tanto, há a

necessidade de que uma grande quantidade de oligonucleotídeos sejam utilizados, de modo que todo o genoma seja analisado quanto ao polimorfismo presente.

Uma questão pertinente a ser levantada em estudos que se utilizam da técnica de RAPD refere-se ao número mínimo de produtos de amplificação necessários para se ter uma boa confiabilidade das estimativas de distâncias genéticas. Desta forma é possível atingir a estabilidade dos agrupamentos gerados e conseguir boas correlações deste tipo de marcador com outros, ou mesmo, um poder preditivo da performance híbrida.

Sendo assim, a estabilidade do dendrograma deve ser uma tarefa importante num estudo de diversidade genética utilizando dados de RAPD. De acordo com os resultados obtidos neste trabalho, o padrão de agrupamento foi mantido desde do dendrograma produzido a partir de 150 bandas polimórficas até o de 262 bandas polimórficas (de um total de 325 bandas). Thormann *et al.* (1994) relataram que um total de 327 bandas seria suficiente para produzir um coeficiente médio de variação de 10% para RAPD, em *Brassica*. Jain *et al* (1994) compararam dendrogramas gerados usando 200, 300 e 595 produtos de amplificação, em *Brassica juncea*, e concluíram que os genótipos averiguados podiam ser separados nos mesmos grupos heteróticos usando apenas um número limitado de bandas (200). Assim, os resultados obtidos estão de acordo com dados da literatura, podendo-se supor que se tenha atingido uma adequada saturação do genoma.

As três primeiras coordenadas principais da análise multivariada (PCO) conseguiram reproduzir o padrão de associação em três grupos encontrado no dendrograma final. Entretanto, muito da variância presente nos dados permaneceu não resolvida. Apenas 36% da variância total foi explicada pelos três primeiros eixos, o que não necessariamente implica que a análise de coordenadas principais de produtos gerados por RAPD seja pouco confiável; ao

contrário, indica que uma menor proporção da informação contida na distância genética entre as linhagens poderia ser representada por três coordenadas principais. Apesar de haver 18 possíveis eixos a serem formados (cada um correspondente a uma variável), dificilmente se consegue uma boa visualização de um espaço acima do tridimensional. Como as três primeiras coordenadas foram capazes de separar os dados binários de RAPD em três grupos principais bem definidos, os mesmos grupos visualizados no dendrograma final obtido para os mesmos dados, pôde-se reconhecer um padrão consistente de variação existente entre os genomas. Pode ser que o resíduo de variação não representada nos três eixos seja decorrente de subgrupos menores visualizados no dendrograma final o que enfatiza a grande variabilidade presente nos dados de RAPD entre e dentro dos grupos de linhagens das populações estudadas (BR-105 e BR-106) . Messmer *et al.* (1991) também não conseguiram explicar a variância total presente nos dados usando um procedimento de ordenação para comparar marcadores de RFLP e de isoenzimas. Nienhuis *et al.* (1994) explicaram apenas 53% da variação observada entre dados de RAPD obtidos para feijões vermelhos e pretos num espaço bidimensional de coordenadas principais; entretanto, conseguiram a separação dos feijões pretos e vermelhos em dois grupos distintos e consistentes o que foi surpreendente se se considerar que os feijões pretos foram usados durante muitos anos para melhorar os vermelhos, servindo de fonte de variação genética para produção e resistência à pragas (Beebe *et al.*, 1994).

De acordo com os dados apresentados, o dendrograma e a análise multivariada foram fundamentais para a percepção de uma subestrutura presente na classificação anterior. Pelos registros de campo anteriores, as populações BR-105 e BR-106 foram alocadas em dois grupos heteróticos (Naspolini Fº. *et al.*, 1981; Souza Jr. *et al.*, 1993). Entretanto, através da análise das distâncias genéticas obtidas por RAPD, a população BR-106 pôde ser dividida em

dois grupos heteróticos diferentes pertencentes à GI e GIII, definidos pelo dendrograma final. As medidas de distância genética não foram capazes de prever o melhor cruzamento ($r < 1,0$), mas mostraram que a divisão heterótica tradicional era inadequada, já que os pares de distâncias genéticas obtidos entre GII (composto essencialmente por linhagens BR-105) e GIII (composto exclusivamente de linhagens BR-106) não se correlacionavam com a performance de seus cruzamentos. Por outro lado, combinações entre GI e GII ou ainda as combinações intrapopulacionais entre GI e GIII mostraram correlações significativas entre as distâncias genéticas estimadas por RAPD e a performance híbrida para produção de grãos.

A população BR-106, como previamente descrito, derivou de diferentes populações e compostos, e poderia ser considerada uma população composta de ampla base genética, com maior variabilidade do que a população BR-105. Sendo assim, é compreensível que dados eminentemente fenotípicos sujeitos à interações ambientais, encobrissem a diversidade genética presente na população BR-106, consequentemente remetendo-a a um único grupo heterótico. Entretanto, marcadores do tipo RAPD, pelo seu caráter aleatório, podem acessar uma gama de regiões do genoma (expressas ou não) fornecendo uma maior representatividade da diversidade existente.

Jain *et al.* (1994) realizaram ensaios de RAPD com 12 genótipos indianos e 11 exóticos de *Brassica juncea* a fim de detectar possíveis variações inerentes e estabelecer relações genéticas. Um baixo grau de variabilidade genética foi encontrado entre os genótipos indianos, enquanto os genótipos exóticos exibiram um considerável nível de polimorfismo entre eles. Cerca de 80% dos cruzamentos entre genótipos indianos ou exóticos apresentaram heterose negativa. Os autores não conseguiram detectar qualquer correlação direta entre os dados de distância genética com o grau de heterose; entretanto, através de uma análise gráfica entre a

performance de produção de grãos em F1 *versus* a distância genética estimada por RAPD, puderam verificar que cerca de 90% dos cruzamentos entre genótipos indianos e exóticos apresentaram uma performance de produção de grãos acima da média. Assim, embora tal estudo não tenha reconhecido a possibilidade de se empregar distâncias genéticas baseadas em RAPDs para fins de predição, reconheceu a capacidade da técnica de RAPD poder ser utilizada para alocar genótipos à grupos heteróticos.

De todos os caracteres agronômicos avaliados, apenas produção de grãos correlacionou-se consistentemente com distâncias obtidas por RAPD. Este é um dos caracteres de importância econômica mais valorizado no melhoramento de milho; entretanto, sua herança complexa (quantitativa) e sua baixa herdabilidade dificultam a existência de marcadores fenotípicos diretos que possam ser usados para selecionar linhagens/variedades mais produtivas. Assim, a busca pelo uso de marcadores moleculares poderia ampliar a eficiência de seleção para este caráter.

Ajmore-Marsan *et al.* (1995) descreveram que o QTL de maior expressão para produção de grãos em milho estaria localizado nas proximidades da região organizadora do nucléolo. Entretanto, estudos de mapeamento com QTLs utilizando RFLPs sugerem que arranjos únicos de QTLs poderiam ser detectados dentro de uma mesma população, com certo grau de sobreposição. Beavis *et al.* (1991) relataram comparações de QTLs para altura de planta entre quatro populações não identificaram qualquer QTL comum à todas; entretanto, apenas poucos QTLs foram comuns à subgrupos de populações. Estas dissimilaridades resultariam numa dificuldade de uso consistente de marcadores do tipo RFLPs associados à QTLs para uma dada característica. Além disso, como marcadores deste tipo não são QTLs, mas estão ligados à eles, a ligação entre o alelo marcador e o QTL seria importante. Havendo equilíbrio de ligação ou

permuta entre os alelos, associações entre marcadores e QTLs não poderiam ser identificadas (Dudley, 1993). RAPDs, neste sentido, podem acessar um grande número de seqüências espalhadas aleatoriamente pelo genoma de tal forma a poder atingir diferentes QTLs, mesmo que estes apresentem alteração de posição em diferentes populações.

O estudo realizado é de importância fundamental já que, primeiramente, utiliza germoplasma tropical com ampla base genética enquanto que a maioria dos artigos em literatura internacional exploram linhagens do cinturão de milho norte-americano e européias, ou seja, linhagens de regiões temperadas, oriundas de grupos heteróticos bem definidos e de estreita base genética. Neste trabalho, mostrou-se a importância de se considerar o padrão de agrupamento do dendrograma além de se empregar outras análises gráficas, como a multivariada, para se poder interpretar os resultados de maneira mais informativa que a matriz original de dados. A interpretação dos dados obtidos por meio de marcadores moleculares requer métodos que possam revelar a sua estrutura intrínseca.

De acordo com os resultados obtidos, distâncias genéticas entre linhagens endogâmicas estimadas por RAPD podem se correlacionar positivamente com dados de produção de grãos para os híbridos simples. Estes resultados concordam com o fato que linhagens divergentes, quando cruzadas, produzem híbridos que exibem melhor performance (Paterniani & Lonnquist, 1963; Moll *et al.*, 1965). Em programas de melhoramento convencional, vários cruzamentos são geralmente realizados e a produção de grãos dos híbridos simples tem que ser avaliada em delineamentos experimentais.

Os resultados sugerem que as distâncias genéticas obtidas por RAPD poderiam ser utilizadas para predizer grupos de linhagens capazes de, quando cruzados entre si, gerar

híbridos superiores, embora não se tenha conseguido, com este tipo de medida, predizer o melhor cruzamento dentre todos os possíveis.

Este trabalho é, portanto, pioneiro no uso da tecnologia de RAPD para acessar a divergência genética entre linhagens de milho de germoplasma tropical com finalidades preditivas, e por ser bem sucedido em mostrar altas correlações de medidas avaliadas por esta classe de marcadores com performances dos respectivos híbridos simples; na verdade, as mais altas já relatadas em literatura. De tal forma, o RAPD mostrou ser uma ferramenta adequada, podendo ser de grande utilidade para abreviar etapas de seleção de combinações híbridas em programas de melhoramento de milho.

CONCLUSÕES

7. CONCLUSÕES

Os resultados obtidos neste trabalho permitem concluir que:

(A) A técnica de RAPD mostrou ser uma ferramenta adequada para alocar linhagens endogâmicas de milho à diferentes grupos heteróticos, como foi o caso da população BR-106, que foi subdividida em dois grupos heteróticos enquanto a BR-105, permaneceu preferencialmente (com exceção de uma única linhagem) alocada num único grupo heterótico;

(B) As distâncias genéticas baseadas em RAPD mostraram-se capazes de preferencializar os melhores cruzamentos dentro de um programa de melhoramento, podendo contribuir para abreviar sobretudo a etapa de seleção de combinações de linhagens.

Várias conclusões parciais e metodológicas mostraram-se importantes para conduzir às conclusões principais citadas acima:

- O perfil de amplificação de DNA manteve-se constante de 10 a 30 ng havendo alteração do perfil com aumento da quantidade de DNA;

- A técnica de RAPD mostrou-se efetiva para a distinção das linhagens endogâmicas de milho (*Zea mays* L.), revelando amplo grau de polimorfismo (80,6%) entre elas;
- O acompanhamento dos dendrogramas crescentes, gerados pelo acúmulo de bandas polimórficas, foi fator primordial para determinar a estabilidade dos agrupamentos produzidos;
- O dendrograma final produzido com 262 bandas polimórficas dividiu as linhagens em três grupos principais: GI, GII e GIII. Esta divisão foi confirmada pela análise de coordenadas principais (PCO);
- O Grupo I (GI) foi o mais estável dos grupos gerados ao longo dos sucessivos dendrogramas e inclui linhagens de ambas as populações de origem (BR-105 e BR-106);
- O Grupo II (GII) apresentou uma preponderância de linhagens da população BR-105;
- O Grupo III (GIII) foi constituído apenas de linhagens da população BR-106;
- Quando a divisão de grupos foi desconsiderada, nenhuma correlação com a performance dos híbridos simples foi encontrada;

- Quando se respeitou a divisão de grupos encontrada com os métodos de agrupamento e ordenação, obteve-se uma correlação significativa ($r = 0,70$) entre as distâncias genéticas para os pares de linhagens GI x GII com os dados de produção de grãos dos respectivos híbridos simples;
- Alta correlação foi obtida entre as distâncias dos pares de linhagens GI x GIII ($r=0,87$) com os dados de produção de grãos de seus respectivos híbridos simples;
- Apenas o caráter produção de grãos correlacionou-se com os dados de distância genética obtidos por RAPD; as demais características agronômicas avaliadas (altura de planta e espiga), bem como com as heteroses respectivas à estas características, não exibiram associações significativas com os dados moleculares;
- As distâncias entre os pares de linhagens dos grupos GII e GIII não se correlacionaram com nenhum dos caracteres agronômicos avaliados nem com as heteroses respectivas à cada um destes caracteres;
- A divisão heterótica tradicional em dois grupos heteróticos se mostrou inadequada já que os pares de distâncias genéticas entre os grupos GII (preponderantemente linhagens BR-105) e GIII (composto exclusivamente por linhagens BR-106) não se correlacionaram com a performance de seus cruzamentos.

BIBLIOGRAFIA

8.BIBLIOGRAFIA

- Abdelnoor, R.V.; Barros, E.G.; Moreira, M.A (1995).** Determination of genetic diversity within Brazilian soybean germplasm using random amplified polymorphic DNA techniques and comparative analysis with pedigree data. *Rev. Brasil. Genet.*, 18(2): 265-273.
- Akopyanz, N.; Bukanov, N.O; Westblom, T.U.; Kresovich, S. & Berg, D.E. (1992).** DNA diversity among clinical isolates of *Helicobacter pylori* detected by PCR-based RAPD fingerprinting. *Nucleic Acids Res.*, 20: 5137-5142.
- Alava, R.O (1952).** Spikelet variation in *Zea mays* L. *Ann. Mo. Bot. Grd.*, 39: 65-96.
- Al-Janabi, S.M.; Honeycutt, R.J.; McClelland, M. & Sobral, B.W.S. (1993).** A genetic linkage map of *Saccharum spontaneum* (L.) "SES-208". *Genetics*, 134: 1249-1260.
- Anderson, E. (1945).** Maize in the new world. Pp. 27-42. *In*: Wilson, C.M. (Ed.), *New crops for the new world*. McMillan Co., New York.
- Anderson, S.M. & MacDonald J.S. (1983).** Biochemical and Molecular Analyses of Naturally occurrent *Adh* variant in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA.*, 80: 4798-4802.

- Arnold, M.L.; Buckner, C.M. & Robenson, J.J. (1991).** Polen-mediated introgression and hybrid speciation in *Lousiana erises*. *Proc. Natl. Acad. Sci. USA*, 88: 1398-1402.
- Ayliffe, M.A; Lawrence, G.J.; Ellis, J.G. & Pryor, A.J. (1994).** Heteroduplex molecules formed between allelic sequences cause nonparental RAPD bands. *Nucleic Acids Res.*, 22: 1632-1636.
- Baenzinger, P.S. & Peterson, C.J. (1992).** Genetic variation: Its origin and use for breeding self-pollinated species. Pp. 62-69. *In*: Stalker, H.T. & Murphy, J.P. (Eds.). *Plant Breedings in the 1990s*. CAB International, Wallinford, U.K.
- Bark, O H. & Havey, M.J. (1995).** Similarities and relationships among populations of the bulb onion as estimated by nuclear RFLPs. *Theor. Appl. Genet.*, 90: 407-414.
- Baruffi, L.; Damiani, G.; Guglielmino, C.R.; Bandi, C.; Malacrida, AR. & Gasperi, G. (1994).** Polymorphism within and between populations of *Ceratitis capitata*: comparison between RAPD and multilocus enzyme electrophoresis data. *Heredity*, 74: 425-437.
- Beal, W.J. (1877).** Report of the professor of botany and horticulture. *Michigan Board of Agric.*, Lansing.

- Beavis, W.D.; Grant, D.; Albertson, M. & Fincher, R. (1991).** Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theor. Appl. Genet.*, 83: 141-145.
- Beckman, J.S. & Soller, M. (1983).** Restriction fragment length polymorphisms and genetic improvement: methodologies, mapping and costs. *Theor. Appl. Genet.*, 67: 35-45.
- Beebe, S.E.; Ochoa, I.; Nienhuis, P.; Skroch, P. & Tivang, J. (1994).** Genetic diversity among common bean breeding lines developed for Central America. *Crop Sci.* (In review).
- Bernardo, R. (1992).** Relationship between single-cross performance and molecular marker heterozygosity. *Theor. Appl. Genet.*, 83: 628-634.
- Bernardo, R. (1993).** Estimation of coefficient of coancestry using molecular markers in maize. *Theor. Appl. Genet.*, 85: 1055-1062.
- Bernardo, R. (1994).** Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Sci.*, 34: 20-25.
- Besse, P.; Seguin, M.; Lebrun, P.; Chevalier, M.H.; Nicolas, D. & Lanaud, C. (1994).** Genetic diversity among wild and cultivated population of *Hevea brasiliensis* assessed by nuclear RFLP analysis. *Theor. Appl. Genet.*, 88: 199-207.

- Binelli, G. & Bucci, G. (1994).** A genetic linkage map of *Picea abies* Karst., based on RAPD markers, as a tool in population genetics. *Theor. Appl. Genet.*, 88: 283-288.
- Bonnet, O T. (1954).** The inflorescences of maize. *Science*, 120: 77-87.
- Boppenmaier, J.; Melchinger, A E.; Brunklaus-Jung, E.; Geiger, H.H. & Herrmann, R.G. (1991).** Genetic diversity for RFLPs in European maize inbreds: I. Relation to performance of flint X dent crosses for forage traits. *Crop Sci.*, 32: 895-902.
- Botstein, D.; White, R.L.; Skolnick, M. & Davis, R.W. (1980).** Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, 32: 314-331.
- Burr, B.; Evola, S.V.; Burr, F.A. & Beckman, J.S. (1983).** The application of restriction fragment length polymorphism to plant breeding. Pp. 45-59. In: Setlow, J.K. e Hollander, A. (Eds.). *Genetic engineering principles and methods*, Vol. 5., Plenum, New York.
- Bussab, W.O; Miazaki, E.S. & Andrade, D.F. (1990).** Introdução à análise de agrupamentos. Associação Brasileira de Estatística - ABE, 9 ° Simpósio Nacional de Probabilidade e Estatística, São Paulo.
- Caetano-Anollis, G.; Bassam, B.J. & Gresshoff, P.M. (1991).** DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Bio technology*, 9: 553-557.

- Carlson, J.E.; Tulsieram, L.K.; Glaubitz, J.C.; Kauffeldt, C. & Rutledge, R. (1991).** Segregation of random amplified DNA markers in F1 progeny of conifers. *Theor. Appl. Genet.*, 83: 194-200.
- Clayton, W.D. (1973).** The awnless species of Andropogoneae. *Kew Bull.*, 28:49-58.
- Clayton, W.D. (1983).** Notes on tribe Andropogoneae (Gramineae). *Kew Bull.*, 35:813-818.
- Chalmers, K.J.; Newton, A.C.; Waugh, R.; Wilson, J. & Powell, W. (1994).** Evaluation of extent of genetic variation in mahoganies (Meliceae) using RAPD markers. *Theor. Appl. Genet.*, 89: 504-508.
- Charcosset, A; Lefort-Buson, M. & Gallais, A (1991).** Relationship between heterosis and heterozygosity at marker loci: a theoretical computation. *Theor. Appl. Genet.*, 81: 571-575.
- Chughtai, S. R. & Seffensen, D.M. (1987).** Heterochromatic knob composition of commercial inbred lines of maize. *Maydica*, 32:171-187.
- Coe Jr., E.H.; Neuffer, M.G. & Hoisington, S.A (1988).** The genetics of corn. Pp. 81-258. *In*: Sprague, G.F. & Dudley, J.W. (Eds.). *Corn and Corn Improvement*, 3^a.ed., Agron. Monogr. 18 ASA., CSSA e SSSA, Madison, WI.

- Cognato, A.I.; Rodgers, S.O. & Teale, S.A (1995).** Species diagnosis and phylogeny of *Ips grandicollis* group (Coleoptera: Scolytidae) using random amplified polymorphic DNA. *Annals of Entomological Society of America*, 88(4): 397-405.
- Consoli, L. (1995).** Análise genética e molecular da variação somaclonal em *Stylosanthes guinensis* (Aubl.) Sw. (Leguminosae). Dissertação apresentada à Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo para obtenção do título de “Mestre em Agronomia”. Piracicaba, São Paulo. 108 pp.
- Cox, T.S.; Kang, Y.T.; Gorman, M.B. & Rodgers, D.M. (1985a).** Relationship between coefficient of parentage and genetic similarity indices in soybean. *Crop Sci.*, 25: 529-532.
- Cox, T.S.; Lockhart, G.L.; Walker, D.W.; Harrill, L.G.; Albers, L.D. & Rodgers, D.M. (1985b).** Genetic relationships among hard red winter wheat cultivars as evaluated by pedigree analysis and glicidin polyacrylamide gel electrophoresis patterns. *Crop Sci.*, 25: 1058-1063.
- Crowhurst, R.N.; Hawthorne, B.T.; Rikkerink, E.H.A & Templeton, M.D. (1991).** Differentiation of *Fusarium solani* f. sp. *Curcubitae* races 1 and 2 by random amplified polymorphic DNA. *Current Genet.*, 20: 391-396.

- Cruz, C.D. & Regazzi, A.J. (1994).** Estudo da divergência genética por componentes principais. Pp. 306-313. In: Cruz, C.D. & Regazzi, A.J.(Eds.). *Modelos Biométricos Aplicados ao Melhoramento Genético*. Viçosa, UFV, Imp. Univ.
- Davis, T.M.; Yu, H.; Haigis, K.M. & McGowan, P.J. (1995).** Template mixing: a method of enhancing detection and interpretation of codominant RAPD markers. *Theor. Appl. Genet.*, 91: 582-588.
- Dice, L.R. (1945).** Measures of the amount of ecologic association between species. *Ecology*, 26: 297-302.
- Doebley, J.F.; Goodman, M. M. & Stuber, C.W. (1983).** Isozyme variation in maize from the southwestern United States: taxonomic and anthropological implications. *Maydica*, 28:97-120.
- Dudley, J.W. (1987).** Modification of methods for identifying populations to be used for improving parents of elite single crosses. *Crop Sci.*, 27: 940-943.
- Dudley, J.W (1992).** Theory for identification of marker locus - QTL associations in population by line crosses. *Theor. Appl. Genet.*, 85: 101-104.
- Dudley, J.W. (1993).** Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. *Crop Sci.*, 33: 660-668.

- Dudley, J.W. (1994).** Comparison of genetic distance estimators using molecular marker data. Pp. 3-7. *In: Proceedings of the Symposium "Analysis of Molecular Marker Data". Joint Plant Breeding Symposia Series, American Society for Horticultural Science, Crop Science Society of America, Corvallis, Oregon. 46 pp.*
- Dudley, J.W; Saghai-Maroo, M.A; Rufener, G.K. (1991).** Molecular markers and grouping of parents in maize breeding programs. *Crop Sci.*, 31: 718-723.
- Dunn, G. & Everitt, B.S. (1982).** An introduction to mathematical taxonomy. Cambridge, New York.
- East, E.M. (1908).** Inbreeding in corn. 1907. Pp. 419-428. *In: Connecticut Agric. Exp. Stn. Rep.*, 1936. Heterosis. *Genetics*, 21: 375-397.
- Edwards, M.D.; Stuber, C. W. & Wendel, J.F. (1987).** Molecular-marker facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distribution and types of gene action. *Genetics*, 116: 113-125.
- Eloibu, N.G.; Goddard, M. & Taylor, J.F. (1990).** Prediction of heterosis in crosses between inbred lines of *Drosophila melanogaster*. *Theor. Appl. Genet.*, 80: 321-325.
- Evola, S.V. & Burr, B. (1986).** The suitability of restriction fragment length polymorphisms as genetic markers in maize. *Theor. Appl. Genet.*, 71: 765-771.

- Frei, O.M.; Stuber, C.W & Goodman, M.M.(1986).** Uses of allozymes as genetic markers for predicting performance in maize single-cross hybrids. *Crop Sci.*, 26: 37-42.
- Galinat, W.C. (1988).** The origin of corn. Pp. 1-31. *In: Sprague, G.F. & Dudley J.W. (Eds.). Corn and Corn Improvement*, 3^a ed., Agron Monogr. 18 ASA., CSSA. e SSSA., Madison, WI.
- Gebhardt, C. & Salamini, F. (1992).** Restriction fragment length polymorphisms analysis of plant genomes and its application to plant breeding. *Int. Review of Cytology* , 135: 201-251.
- Gimelfarb, A & Lande, R. (1994).** Simulation of marker assisted selection in hybrid populations. *Genet. Res. Camb.*, 63: 39-47.
- Giovannoni, J.J.; Wing, R.A; Canal, M.W. & Tanksley, S.D. (1991).** Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucl. Acids Res.*, 19: 6553-6558.
- Godshalk, E.B.; Lee, M. & Lamkey, K.R. (1990).** Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. *Theor. Appl. Genet.*, 80: 273-280.
- Gonella, J.A & Peterson, P.A (1978).** Isozyme relatedness of inbred lines of maize and performance of their hybrids. *Maydica*, 23: 55-61.

- Goodman, M.M. & Stuber, C.W. (1983).** Races of maize. VI. Isozyme variation among races of maize in Bolivia. *Maydica*, 28:169-187.
- Goodwin, P.H. & Annis, S.L. (1991).** RAPD identification of genetic variation and pathotype of *Leptosphaeria maculans* by random amplified polymorphic DNA assay. *Appl. Environ. Microbiol.*, 57: 2482-2486.
- Grodzicker, T.; Williams, J; Sharp, P. & Sambrook, J. (1974).** Physical mapping of temperature-sensitive mutations of adenovirus. Cold Spring Harbor Symp. *Quant. Biol.*, 39:439-446.
- Hadjinov, M.I.; Scherbak, V.S.; Benko, N.I.; Gusev, V.P.; Sukhorchevshaya, T.B. & Voronova, L.P. (1982).** Interrelationships between isozyme diversity and combining ability in maize lines. *Maydica*, 27: 135-149.
- Haley, S.D.; Miklas, P.N. Afanador, L. & Kelly, J.D. (1994).** Random amplified polymorphic DNA (RAPD). Marker variability between and within gene pools of common bean. *J. Amer. Soc. Hort. Sci.*, 119(1): 122-125.
- Hallauer, A R. (1990).** Methods used in developing maize inbreds. *Maydica*, 35:1-16.

- Hallauer, A R. (1995).** Application of Biotechnology to maize improvement. Pp. 519-533. In: L. Avila, G. & P. Céspedes, L.M. (Eds.). *Memorias de la III Reunion Latino Americana y XVI Reunion de la Zona Andina de Investigadores en Maiz*, Tomo I, Cochabamba, Santa Cruz, Bolivia.
- Hallauer, A R.; Russel, W.A & Lamkey, K.R. (1988).** Corn breeding. Pp. 463-564. In: Sprague, G.F. & Dudley, J.W. (Eds.). *Corn and Corn Improvement*, 3^a.ed., Agron. Monogr. 18 ASA., CSSA e SSSA, Madison, WI.
- Halldén, C.; Nilsson, N-O; Rading, I.M. & Säll, T. (1994).** Evaluation of RFLP and RAPD markers in a comparison of *Brassica napus* breeding lines. *Theor. Appl. Genet.*, 88: 123-128.
- Halward, T.; Stalker, T.; Larue, E. & Kochert, G. (1992).** Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). *Plant Mol. Biol.*, 18: 315-325.
- Heckel, D.G.; Gahan, L.J.; Tabashnik, B.E. & Johnson, M.W. (1995).** Randomly amplified polymorphic DNA differences between strains of diamondback moth (Lepidoptera: Plutellidae) susceptible or resistant to *Bacillus thuringiensis*. *Annals of Entomological Society of America*, 88(4): 531-537.

- Heidrich-Sobrinho, E. & Cordeiro, A R. (1975).** Codominant isozymic alleles as markers of genetic diversity correlated with heterosis in maize (*Zea mays* L.). *Theor. Appl. Genet.*, 46: 197-199.
- Helentjaris, T. (1987).** A genetic linkage map for maize based on RFLPs. *TIG.*, 3:217-221.
- Helentjaris, T.; King, G.; Slocum, M.; Siedestrang, C.& Wegman, S. (1985).** Restriction fragments polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. *Plant Mol. Biol.*, 5: 109-118.
- Helentjaris, T.; Slocum, M.; Wrigth, S.; Schaefer, A. & Nienhuis, J. (1986).** Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theor. Appl. Genet.*, 72: 761-769.
- Helentjaris, T.; Weber, D. & Wrigth, S.(1988).** Identification of the genomic locations of duplicated sequences in maize by analysis of restriction fragment length polymorphisms. *Genetics*, 188: 353-363.
- Heun, M. & Helentjaris, T. (1993).** Inheritance of RAPDs in F1 hybrids of corn. *Theor. Appl. Genet.*, 85: 961-968.

- Heun, M.; Murphy, J.P. & Phillips, T.D. (1994).** A comparison of RAPD and isozyme analyses for determining the genetic relationships among *Avena sterilis* L. accessions. *Theor. Appl. Genet.*, 87: 689-696.
- Hilu, K.W. (1995).** Evolution of finger millet: evidence from random amplified polymorphic DNA. *Genome*, 38: 232-238.
- Hoisington, D.; Khairallah, M. & Gonzalez-de-Lion, D. (1994).** Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory. Second Edition, Mexico, D.F.: CIMMYT.
- Honeycutt, R.J.; Smith, S. & Sobral, B.W.S. (1992).** Reconstructing histories of maize inbreds using molecular characters. *Plant Genome I Abstracts*: 29, San Diego, CA.
- Hotelling, H. (1993).** Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 24: 417-441.
- Hotelling, H. (1936).** Simplified calculation of principal components. *Psychometrika*, 1: 27-35.
- Hu, J. & Quiros, C.F. (1991).** Identification of broccoli and cauliflower cultivars with RAPD markers. *Plant Cell Rep.*, 10: 505-511.

- Huff, D R.; Peakall, R. & Smouse, P.E.(1993).** RAPD variation within and among natural populations of outcrossing buffalograss [*Buchlœe dactyloides* (Nutt.) Englem.]. *Theor. Appl. Genet.*, 86: 927-934.
- Hunter, R.B. & Kannenberg, L.W. (1971).** Isozyme characterization of corn (*Zea mays* L.) inbreds and its relationship to single-cross hybrid performance. *Can. J. Genet. Cytol.*, 13: 649-655.
- Hunt, G.J. & Page, R.E. Jr. (1992).** Patterns of inheritance with RAPD molecular markers reveal novel types of polymorphism in the honeybee. *Theor. Appl. Genet.*, 85: 15-20.
- Innis, M.A & Gelfand, D.H. (1990).** Optimization of PCRs. Pp. 3-20. *In: Sninsky, J.J. & White, T.J. (Eds.). PCR Protocols: A Guide to Methods and Applications.* Academic Press, Inc., San Diego.
- Jaccard, P. (1908).** Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaud. Nat.*, 44: 223-270.
- Jain, A; Bhatai, S.; Banga, S.S.; Prakash, S. & Lakhmikumaran, M. (1994).** Potential use of random amplified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard (*Brassica juncea*) and its relationship to heterosis. *Theor. Appl. Genet.*, 88: 116-122.

- Jain, S.K. & Benerjee, D.K. (1974).** Preliminary observations on the ethnobotany of the genus *Coix*. *Econ Bot.*, 28: 38-42.
- Jeffreys, A J. (1979).** DNA sequence variants in the G_α, A_α, δ and β-globin of man. *Cell*, 18: 1-10.
- Jones, D.F. (1918).** The effects of inbreeding and cross-breeding upon development. Pp. 5-100. *In: Connecticut Agric. Exp. Stn. Bull.* 207.
- Khahler, A L.; Hallauer, A R.& Gardner, C.O (1986).** Allozyme polymorphisms within and among open-pollinated and adapted exotic populations of maize. *Theor. Appl. Genet.*, 72: 592-601.
- Kang, F.Y.; Deynze, A V. & Pauls, K.P. (1993).** Random amplified polymorphic DNA (RAPD) analysis. Pp. 287-301. *In: Gliab, R.B. & Thompson, J. (Eds.). Methods in Plant Molecular Biology and Biotechnology.* CRC Press. Inc.
- Kazan, K.; Manners, J.M. & Cameron, D.F. (1992).** Genetic relationships and variation in *Stylosanthes guinensis* species complex assessed by random amplification polymorphic DNA. *Genome*, 36: 43-49.

- Klein-Lankhorst, R.M.; Vermut, A; Weide, R.; Lharska, T. & Zabel, P. (1991).** Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). *Theor. Appl. Genet.*, 83: 108-114.
- Kuwada, Y. (1911).** Meiosis in pollen mother cells of *Zea Mays* L. *Bot. Mag. Tokio*, 25:163-181.
- Lamkey, K.R.; Hallauer, A.R. & Kahler, A.L.(1987).** Allelic differences at enzyme loci and hybrid performance in maize. *J. Hered.*, 78: 231-234.
- Landry, .B.S.; Kesseli, R.V.; Farrara, B. & Michelmore, R. W. (1987).** A genetic map for lettuce (*Lactuca sativa* L.) with restriction length polymorphism, isozyme, disease resistance and morphological markers. *Genetics*, 116:331-337.
- Lanham, P.G.; Brennan, R.M.; Mackett, C. & McNicol, R.J. (1995).** RAPD fingerprinting of blackcurrant (*Ribes nigrum* L.) cultivars. *Theor. Appl. Genet.*, 90: 166-172.
- Lee, M. (1995).** DNA markers and plant breeding programs. *Advances in Agronomy*, 55: 265-344.
- Lee, M.; Godshalk, E.B.; Lamkey, K. R. & Woodman, W.W. (1989).** Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. *Crop Sci.*, 29: 1067-1071.

- Liu, Z. & Fournier, G.R. (1993).** Comparison of allozyme, RFLP, and RAPD markers for revealing genetic variation within and between trembling aspen and bigtooth aspen. *Theor. Appl. Genet.*, 87: 97-105.
- Livini, C.; Ajmone-Marsan, P.; Melchinger, A.E.; Messmer, M.M. & Motto, M. (1992).** Genetic diversity of maize inbred lines within and among heterotic groups by RFLPs. *Theor. Appl. Genet.*, 84: 17-25.
- Malécot, G. (1948).** Les Mathématiques de l'hérédité. Masson et Cie, Paris.
- Manly, B.F.J. (1994).** Principal component analysis. Pp. 76-92. In: *Multivariate Statistical Methods: a primer*, second edition, Chapman e Hall, London.
- Marshall, O.R. & Brown, A.H.D. (1975).** The charge state model of protein polymorphism in natural populations. *J. Mol. Evol.*, 6: 149-163.
- Martín, C. & Pérez, C. (1994).** The use of RAPD to determine the genetic variability of micropropagated plants from endangered species. Application to the Spanish endemism *Limonium estevi*. *phyton*, 56: 65-72.

Melchinger, A. E. (1993). Use of RFLP markers for analysis of genetic relationships among breeding materials and prediction of hybrid performance. Pp. 621-628. *In*: Buxlon, D.R.; Shibles, R.; Fosberg, R.A.; Blad, B.L.; Asay, K.H.; Paulson, G.M. & Wilson, R.F. (Eds.). *Crop Science of America*, Madison, WI.

Melchinger, A.E.; Lee, M.; Lamkey, K.R. & Woodman, W.L. (1990a). Genetic diversity for restriction fragment length polymorphisms: relation to estimated genetic effects in maize inbreds. *Crop Sci.*, 30: 1033-1040.

Melchinger, A.E.; Lee, M.; Lamkey, K.R.; Hallauer, A.R. & Woodman, W.L. (1990b). Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. *Theor. Appl. Genet.*, 80: 488-496.

Melchinger, A.E.; Messmer, M.M.; Lee, M.; Woodman, W.L. & Lamkey, K.R. (1991). Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. *Crop Sci.*, 31: 669-678.

Messmer, M.M.; Melchinger, A. E; Herrmann, R. & Boppenmaier, J. (1993). Relationships among early European maize inbreds: II. Comparison of pedigree and RFLP data. *Crop Sci.*, 33: 944-950.

- Messmer, M.M.; Melchinger, A. E.; Boppenmaier, J.; Brunklaus-Jung, E. & Herrmann, R.G. (1992).** Relationships among early European maize inbreds. I. Genetic diversity among flint and dent lines revealed by RFLPs. *Crop Sci.*, 32: 1301-1309.
- Messmer, M.M.; Melchinger, A. E.; Lee, M.; Woodman, W. L.; Lee, E.A. & Lamkey, K. R. (1991).** Genetic diversity among progenitors and elite lines from Iowa Stiff Stalk Synthetic (BSSS) maize population: Comparison of allozyme and RFLP data. *Theor. Appl. Genet.*, 83: 97-107.
- Michelmore, R.W.; Paran, I. & Kessel, R.V. (1991).** Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. *Proc. Natl. Acad. Sci.*, 88: 9828-9832.
- Moll, R.H.; Salhuna, W.S. & Robinson, H.F.(1962).** Heterosis and genetic diversity in variety crosses of maize. *Crop Sci.*, 2: 197-198.
- Moll, R.H.; Lonnquist, J.H.; Fortuna, J.V. & Johnson, C. E. (1965).** The relationship of heterosis and genetic divergence in maize. *Genetics*, 52: 139-144.
- Moll, R.H.; Hanson, W.D.; Levings, C.S. III & Ohta, Y. (1972).** Associations between chromosome knobs of *Zea mays* L. and agronomic performance. *Crop Sci.*, 12:585-589.

- Morgan, U.M.; Constantini, C.C.; Grune, W.K. & Thompson, R.C.A. (1993).** RAPD (random amplified polymorphic DNA) analysis of *Giardia* DNA and correlation with isozyme data. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 87: 702-705.
- Moser, H. & Lee, M. (1994).** RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variance in oats. *Theor. Appl. Genet.*, 87: 947-956.
- Myers, R.M.; Maniatis, T. & Lerman, L. S. (1987).** Detection and localization of single base changes by denaturing gradient gel electrophoresis. *Methods in Enzymology*, 155: 501-527.
- Naspolini F.º, V.; Gama, E.E.G.; Vianna, R.T. & MÔro, J.R. (1981).** General and specific combining ability for yield in a diallel cross among 18 maize populations. *Revista Brasileira de Genética*, 4: 571-577.
- Newburry, H.J. & Ford-Lloyd, B.V. (1993).** The use of RAPD for assessing variation in plants. *Plant Growth Regulation*, 12: 43-51.
- Nickerson, D.A; Kaiser, R.; Lappin, S.; Stewart, J. & Hood, L. (1990).** Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. *Proc. Natl. Acad. Sci.*, 87: 8923-8927.

- Nienhuis, J.; Tivang, J. & Tkoroch, P. (1994).** Analysis of genetic relationships among genotypes based on molecular marker data. Pp. 8-14. *In: Proceedings of the Symposium "Analysis of Molecular Marker Data", Joint Plant Breeding Symposia Series, American Society for Horticultural Science/ Crop Science of America, Corvallis, Oregon.*
- Orita, M.; Iwahana, H.; Kanazawa, H.; Hayashi, K. & Sekiya, T. (1989).** Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA*, 86: 2766-2770.
- Paran, I & Michelmore R.W. (1993).** Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theor. Appl. Genet.*, 85: 985-993.
- Paterniani, E. (1990).** Maize breeding in the tropics. *Critical Reviews in Plant Sciences*, 9: 125-154.
- Paterniani, E. (1995).** Importância do milho na Agroindústria. Pp.1-12. *In: Osuna, J.A & Moro, J.R. (Eds.). Produção e melhoramento de milho, Jaboticabal, FUNEP.*
- Paterniani, E. & Lonnquist, J.H. (1963).** Heterosis in interracial crosses of corn (*Zea mays* L.). *Crop Sci.*, 3:504-507.
- Pearson, K. (1901).** On lines and planes of closest fit to systems of points in space. *Philos mag.*, 2: 559-572.

Peter, T. & Botstein, D. (1977). Simple mendelian inheritance of the reiterated ribosomal DNA of yeast. *Proc. Nat. Acad. Sci.*, 74: 5091-5095.

Plomion, C.; Bahman, N.; Durel, C-E. & O'Malley, D.M. (1995). Genomic mapping in *Pinus pinaster* (maritime pine) using RAPD and protein markers. *Heredity*, 74: 661-668.

Price, S.C.; Kahler, A.L.; Hallauer, A.R.; Charmey, P. & Giegel, D.A. (1986). Relationships between performance and multilocus heterozygosity at enzyme loci in single-cross hybrids of maize. *J.Hered.*, 78:341-344.

Ragot, M. & Hoisington, D.A. (1993). Molecular markers for plant breeding: comparisons of RFLP and RAPD genotyping costs. *Theor. Appl. Genet*, 86: 975-984.

Rafalski, J.A. & Tingey, S.V. (1993). Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *TIG.*, 9(8): 275-280.

Reiter, R.S.; Williams, J.G.K.; Feldmann, K.A; Rafalski, J.A; Tingey, S.V. & Scolnik, P.A (1992). Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci. USA*, 89: 1477-1481.

Rholf, F.J. (1992). Numerical Taxonomy and Multivariate Analysis System, Version 1.70, Exeter Publ., New York.

Rholf, F.J. & Fisher, D.L. (1968). Test for hierarchical structure in random data sets. *Systematic Zool.*, 17: 407-412.

Richey, F.D. (1922). The experimental basis for the present status of corn breeding. *J. Am. Soc. Agron.*, 14: 1-17.

Rievis, R.G. & Mangelsdorf, P.C. (1942). A proposed taxonomic change in the tribe Maydea. *Am. J. Bot.*, 29: 815-817.

Rodgers, D.M.; Murphy, J. P. & Frey, K. J. (1983). Impact of plant breeding on the grain yield and genetic diversity of spring oats. *Crop Sci.*, 23: 737-740.

Romesburg, H.C. (1984). Cluster analysis for researchers. *Lifetime Learning Publications*, California.

Roy, A; Frascaria, N.; Mackay, J. & Bousquet, J. (1992). Segregating random amplified polymorphic DNAs (RAPDs) in *Betula alleghaniensis*. *Theor. Appl. Genet.*, 85: 173-180.

Schweder, M.E.; Shatters, R.G.; Jr.; West,, S.H. & Smith, R.L. (1995). Effect of transition interval between melting and annealing temperatures on RAPD analysis. *Biotechniques*, 19(1): 153-156.

- Shull, G.H.(1908).** The composition of a field of maize. *Am. Breeders' Assoc. Rep.*, 4: 296-301.
- Sobral, B.W.S. & Honeycutt, R.J. (1994).** Genetics, plants and the polymerase chain reaction. Pp. 304-319. *In: Mullis, K.B.; Ferré, F; & Gibbs, R.A. (Eds.). The Polymerase Chain Reaction*, Birkhauser, Boston.
- Smith, J.S.C. (1984).** Genetic variability within U.S. hybrid maize: multivariate analysis of isozyme data. *Crop Sci.*, 24:1041-1046.
- Smith, J.S.C (1986).** Genetic diversity within the corn belt dent racial complex of maize (*Zea mays* L.). *Maydica*, 31:349-367.
- Smith, J.S.C.; Goodman, M.M & Stuber, C.W.(1985a).** Genetic variability within U.S. hybrid maize germplasm .I. Historically important lines. *Crop Sci.*, 25: 550-555.
- Smith, J.S.C.; Goodman, M.M. & Stuber, C.W.(1985b).** Genetic variability within U.S. hybrid maize germoplasm. II. Widely used inbred lines from 1970 to 1979. *Crop Sci.*, 25: 681-685.
- Smith, J.S.C. & Smith, O.S. (1987).** Associations among inbred lines of maize using electrophoretic, chromatographic, and pedigree data. 1. Multivariate and cluster analysis of data from "Lancaster Sure Crop" derived lines. *Theor. Appl. Genet.*, 73: 654-664.

- Smith, J.S.C. & Smith, O.S. (1988).** Associations among inbred lines using electrophoretic, chromatographic, and pedigree data. 2. Multivariate and cluster analysis of data from Iowa Stiff Stalk Synthetic derived lines. *Theor. Appl. Genet.*, 76: 39-44.
- Smith, J.S.C. & Smith, O.S. (1989).** The description and assessment of distances between inbred lines of maize. II. The utility of morphological, biochemical, and genetic descriptors and a scheme for testing of distinctiveness between inbred lines. *Maydica*, 34: 151-161.
- Smith, J.S.C. & Smith, O.S. (1992).** Fingerprinting crop varieties. *Adv. Agron.*, 47: 85-140.
- Smith, J.S.C.; Smith, O.S.; Browen, S.L.; Tenborg, R.A. & Wall, S.J.(1990).** Similarities among a group of elite maize inbreds as measured by pedigree, F₁ grain yield, grain yield, heterosis and RFLPs. *Theor. Appl. Genet.*, 80: 833-840.
- Smith, J.S.C.; Smith, O.S.; Browen, S.L. & Tenborg, R.A.(1992).** Numbers of RFLP necessary to show associations between lines. *Maize Genet. Coop. Newslett.*, 66:66.
- Southern, E.M. 1975).** Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.*, 98: 503-517.
- Souza, A P. (1995).** Mapeamento genético com marcadores moleculares. Pp. 165-176. In: Osuna, J. A & Moro, J. R. (Eds.). *Produção e Melhoramento do Milho*. Jaboticabal, FUNEP.

- Souza Jr., C.L.; Santos, M.X.; Mangavaca, R. & Gama, E.E.G. (1993).** Estimativas de parâmetros genéticos na interpopulação de milho BR-105 x BR-106 e suas implicações no melhoramento. *Pesquisa Agropecuária Brasileira*, 28: 473-479.
- Stuber, C.W. (1994).** Heterosis in plant breeding. Pp. 227-251. *In: Janick, J. (Ed.). Plant Breeding Reviews*, Wiley, New York.
- Stuber, O.S. & Goodman, M.M.(1983).** Allozyme genotypes for popular and historically important inbred lines of corn (*Zea mays* L). *U.S. Dept. Agric., Agric. Res. Serv., N ° . 16*, pp.28.
- Stuber, O S. & Moll, R.H. (1972).** Frequency changes of isozyme alleles in a selection experiment for grain yield in maize (*Zea mays* L.). *Crop Sci.*, 12: 240-337.
- Stuber, O S.; Moll, R.H.; Goodman, M.M.; Schaffer, H.E.; Weir, B.S. (1980).** Allozyme frequency changes associated with selection for increased grain yield in maize (*Zea mays* L.). *Genetics*, 95: 225-236.
- Tersac, M.; Blanchard, P.; Brunel, D. & Vincourt, P. (1994).** Relations between heterosis and enzymatic polymorphism in populations of cultivated sunflowers (*Helianthus annuus* L.). *Theor. Appl. Genet.*, 88: 49-55.

- Thormann, C.E.; Ferreira, M.E.; Camargo, L.E.A; Tivang, J.G. & Osborn, T.C. (1994).** Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. *Theor. Appl. Genet.*, 88: 973-980.
- Torres, A.M.; Weeden, N.F. & Martin, A (1993).** Linkage among isozyme, RFLP and RAPD markers in *Vicia faba*. *Theor. Appl. Genet.*, 85: 973-945.
- Troyer, P.F.; Openshaw, S. J. & Knittle, K. H. (1988).** Measurement of genetic diversity among popular commercial corn hybrids. *Crop Sci.*, 28: 481-485.
- Velásquez, V.L.B. & Gepts, P. (1993).** RFLP diversity of common bean (*Phaseolus vulgaris*) in its centres of origin. *Genome*: 37: 256-263.
- Vicario, F.; Vendramin, G.G.; Rossi, P.; Lio, P. & Giovannini, R. (1995).** Allozyme, chloroplast DNA and RAPD markers for determining genetic relationships between *Abies alba* and the relic population of *Abies hedrodensis*. *Theor. Appl. Genet.*, 90: 1012-1018.
- Yang, X. & Quiros, C. (1993).** Identification and classification of celery cultivars with RAPD markers. *Theor. Appl. Genet.*, 86: 205-212.
- Waugh, R. & Powell, W. (1992).** Using RAPD markers for crop improvement. *Tibtech*, 10: 186-191.

Welsh, J.; Chada, K.; Dalal, S.S.; Cheng, R.; Rhalp, D. & McClelland, M. (1992).

Arbitrarily primed PCR Fingerprinting of RNA. *Nuc. Acids. Res.*, 20: 4965-4970.

Welsh, J. & McClelland, M.(1990). Fingerprinting genomes using PCR with arbitrary primers.

Nucleic Acids Res., 18: 7213-7218.

Welsh, J. & McClelland, M. (1991). Genomic fingerprinting using arbitrarily primed PCR and

a matrix of pairwise combinations of primers. *Nucleic Acids Res.*, 19(19): 5275-5279.c

Williams, J.G.K.; Hanafey, M.K.; Rafalski, J.A & Tingey, S.V. (1993). Genetic analyses

using random amplified DNA markers. Pp. 704-740. *In: Wu, R. (Ed.). Methods in Enzymology, Recombinant DNA*, Part I, Vol. 218, Academic Press, San Diego.

Williams, J.G.K.; Kubelik, A.R.; Livak, K.J.; Rafalski, J.A & Tingey, S.V.(1990). DNA

polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.*, 18: 6531-6535.

Wu, D. Y.; Ugozzoli, L.; Bijai, K. P. & Wallace, R.B. (1989). Allele-specific enzymatic

amplification of b-globin DNA for diagnosis of sickle cell anemia. *Proc. Natl. Acad. Sci. USA*, 86: 2757-2760.

ANEXOS

9. Anexos

Anexo A. Alguns coeficientes de semelhança para variáveis dicotômicas

Nome	Expressão	Variação
Distância Binária de Sokal	$\sqrt{\frac{b+c}{a+b+c+d}}$	0-1
Coincidência Simples	$\frac{a+d}{a+b+c+d}$	0-1
Rogers e Tanimoto	$\frac{a+d}{a+2(b+c)+d}$	0-1
Sokal e Sneath	$\frac{2(a+d)}{2(a+d)+b+c}$	0-1
Russel e Rao	$\frac{a}{a+b+c+d}$	0-1
Jaccard	$\frac{a}{a+b+c}$	0-1
Sorenson	$\frac{2a}{2a+b+c}$	0-1
Ochiai	$\frac{a}{\sqrt{(a+b)(a+c)}}$	0-1
Baroni-Urbani-Buser	$\frac{a + \sqrt{ad}}{a+b+c + \sqrt{ad}}$	0-1
Haman	$\frac{(a+d)-(b+c)}{a+b+c+d}$	0-1
Yule	$\frac{ad-bc}{ad+bc}$	0-1
ϕ	$\frac{ad-bc}{\sqrt{[(a+b)(a+c)(b+d)(c+d)']}}$	0-1
Ochiai II	$\frac{ad}{\sqrt{[(a+b)(a+c)(b+d)(c+d)']}}$	0-1
	$\sqrt{\left[\left(\frac{a}{a+b}\right) + \left(\frac{a}{a+c}\right)\right]}$	0-1
	$\sqrt[4]{\left[\left(\frac{a}{a+b}\right) + \left(\frac{a}{a+c}\right) + \left(\frac{d}{d+b}\right) + \left(\frac{d}{d+c}\right)\right]}$	0-1

Anexo B. Distâncias genéticas (complemento do coeficiente de Jaccard) para 03 linhagens de milho (*Zea mays* L.), escolhidas aleatoriamente.

Linhagens	01	11	17
(01) 05-01.4B	0		
(11) 06-08.1A	0,93	0	
(17) 06-37.5B	0,93	0,52	0

Anexo C. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho

(*Zea mays* L.) obtidas para 50 bandas polimórficas.

Linhagens	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
(01) 05-01.4B	0																	
(02) 05-05.2A	0,59	0																
(03) 05-17.1A	0,76	0,50	0															
(04) 05-18.6A	0,67	0,48	0,52	0														
(05) 05-19.1B	0,86	0,57	0,62	0,64	0													
(06) 05-23.2B	0,75	0,40	0,56	0,48	0,68	0												
(07) 05-33.5B	0,83	0,44	0,48	0,52	0,56	0,23	0											
(08) 05-34.2B	0,69	0,50	0,48	0,33	0,44	0,50	0,48	0										
(09) 06-03.5B	0,62	0,35	0,59	0,60	0,56	0,59	0,58	0,58	0									
(10) 06-06.3A	0,69	0,69	0,74	0,63	0,75	0,77	0,73	0,68	0,67	0								
(11) 06-08.1A	0,60	0,46	0,61	0,58	0,72	0,46	0,56	0,66	0,50	0,61	0							
(12) 06-08.2A	0,74	0,65	0,75	0,71	0,76	0,74	0,73	0,73	0,70	0,50	0,44	0						
(13) 06-14.4B	0,76	0,74	0,63	0,60	0,59	0,72	0,71	0,57	0,72	0,43	0,68	0,57	0					
(14) 06-24.7B	0,64	0,44	0,54	0,62	0,62	0,56	0,59	0,59	0,53	0,68	0,50	0,69	0,79	0				
(15) 06-28.1A	0,69	0,55	0,70	0,67	0,71	0,56	0,59	0,64	0,58	0,55	0,68	0,54	0,67	0,48	0			
(16) 06-29.7B	0,71	0,53	0,68	0,69	0,59	0,68	0,67	0,62	0,47	0,78	0,68	0,75	0,80	0,52	0,71	0		
(17) 06-37.5B	0,67	0,48	0,52	0,60	0,69	0,54	0,57	0,62	0,56	0,67	0,42	0,67	0,77	0,26	0,52	0,55	0	
(18) 06-44.1B	0,71	0,53	0,58	0,64	0,73	0,59	0,62	0,67	0,67	0,67	0,42	0,67	0,80	0,33	0,52	0,60	0,09	0

Anexo D. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (*Zea mays* L.) obtidas para 98 bandas polimórficas.

Linhagens	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
(01) 05-01.4B	0																	
(02) 05-05.2A	0,65	0																
(03) 05-17.1A	0,72	0,56	0															
(04) 05-18.6A	0,61	0,51	0,51	0														
(05) 05-19.1B	0,41	0,58	0,57	0,56	0													
(06) 05-23.2B	0,65	0,50	0,56	0,53	0,58	0												
(07) 05-33.5B	0,70	0,49	0,55	0,55	0,55	0,25	0											
(08) 05-34.2B	0,71	0,58	0,32	0,42	0,48	0,52	0,54	0										
(09) 06-03.5B	0,72	0,48	0,63	0,65	0,66	0,73	0,70	0,67	0									
(10) 06-06.3A	0,65	0,66	0,70	0,63	0,70	0,73	0,73	0,69	0,67	0								
(11) 06-08.1A	0,69	0,53	0,62	0,59	0,64	0,55	0,60	0,63	0,61	0,69	0							
(12) 06-08.2A	0,73	0,66	0,70	0,70	0,70	0,69	0,73	0,72	0,71	0,54	0,49	0						
(13) 06-14.4B	0,78	0,73	0,62	0,66	0,64	0,72	0,73	0,64	0,66	0,55	0,68	0,56	0					
(14) 06-24.7B	0,70	0,52	0,55	0,65	0,55	0,52	0,56	0,58	0,60	0,72	0,47	0,62	0,69	0				
(15) 06-28.1A	0,64	0,56	0,59	0,58	0,59	0,47	0,49	0,56	0,69	0,60	0,51	0,57	0,66	0,60	0			
(16) 06-29.7B	0,62	0,52	0,64	0,59	0,60	0,59	0,60	0,64	0,57	0,76	0,62	0,72	0,77	0,53	0,64	0		
(17) 06-37.5B	0,68	0,52	0,54	0,54	0,59	0,56	0,57	0,59	0,61	0,71	0,47	0,65	0,67	0,34	0,52	0,55	0	
(18) 06-44.1B	0,67	0,53	0,56	0,56	0,56	0,60	0,61	0,59	0,60	0,60	0,48	0,62	0,71	0,39	0,51	0,57	0,14	0

Anexo E. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho
(*Zea mays* L.) obtidas para 150 bandas polimórficas.

Linhagens	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
(01) 05-01.4B	0																	
(02) 05-05.2A	0,35	0																
(03) 05-17.1A	0,72	0,50	0															
(04) 05-18.6A	0,61	0,51	0,51	0														
(05) 05-19.1B	0,76	0,54	0,53	0,55	0													
(06) 05-23.2B	0,70	0,47	0,55	0,52	0,57	0												
(07) 05-33.5B	0,72	0,48	0,53	0,55	0,54	0,22	0											
(08) 05-34.2B	0,69	0,54	0,33	0,47	0,51	0,52	0,53	0										
(09) 06-03.5B	0,69	0,52	0,63	0,63	0,64	0,68	0,67	0,63	0									
(10) 06-06.3A	0,62	0,69	0,71	0,65	0,75	0,75	0,73	0,72	0,69	0								
(11) 06-08.1A	0,70	0,54	0,60	0,60	0,61	0,44	0,42	0,61	0,59	0,73	0							
(12) 06-08.2A	0,63	0,62	0,71	0,69	0,27	0,71	0,71	0,28	0,26	0,56	0,45	0						
(13) 06-14.4B	0,70	0,77	0,70	0,71	0,74	0,75	0,75	0,67	0,67	0,53	0,70	0,53	0					
(14) 06-24.7B	0,67	0,37	0,58	0,62	0,59	0,56	0,57	0,59	0,57	0,72	0,39	0,59	0,70	0				
(15) 06-28.1A	0,71	0,55	0,56	0,58	0,52	0,48	0,49	0,56	0,65	0,68	0,48	0,62	0,72	0,49	0			
(16) 06-29.7B	0,61	0,50	0,60	0,56	0,61	0,60	0,60	0,61	0,58	0,74	0,65	0,74	0,79	0,59	0,65	0		
(17) 06-37.5B	0,66	0,50	0,54	0,54	0,58	0,55	0,56	0,57	0,57	0,72	0,39	0,61	0,68	0,29	0,50	0,58	0	
(18) 06-44.1B	0,67	0,53	0,55	0,57	0,58	0,60	0,58	0,59	0,58	0,70	0,40	0,58	0,71	0,35	0,48	0,61	0,13	0

Anexo F. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (*Zea mays* L.) obtidas para 190 bandas polimórficas.

Linhagens	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
(01) 05-01.4B	0																	
(02) 05-05.2A	0,67	0																
(03) 05-17.1A	0,70	0,53	0															
(04) 05-18.6A	0,64	0,47	0,53	0														
(05) 05-19.1B	0,76	0,48	0,55	0,54	0													
(06) 05-23.2B	0,72	0,49	0,55	0,54	0,56	0												
(07) 05-33.5B	0,74	0,51	0,56	0,57	0,53	0,23	0											
(08) 05-34.2B	0,71	0,53	0,38	0,50	0,49	0,53	0,54	0										
(09) 06-03.5B	0,70	0,54	0,62	0,63	0,64	0,67	0,67	0,64	0									
(10) 06-06.3A	0,64	0,70	0,70	0,68	0,75	0,73	0,73	0,72	0,70	0								
(11) 06-08.1A	0,71	0,53	0,60	0,60	0,58	0,53	0,57	0,61	0,57	0,71	0							
(12) 06-08.2A	0,63	0,71	0,71	0,70	0,74	0,72	0,74	0,75	0,72	0,57	0,55	0						
(13) 06-14.4B	0,71	0,79	0,72	0,73	0,75	0,75	0,76	0,70	0,70	0,55	0,71	0,56	0					
(14) 06-24.7B	0,70	0,50	0,57	0,60	0,54	0,51	0,53	0,57	0,56	0,70	0,37	0,61	0,71	0				
(15) 06-28.1A	0,74	0,57	0,56	0,59	0,51	0,48	0,51	0,53	0,64	0,69	0,49	0,67	0,74	0,47	0			
(16) 06-29.7B	0,67	0,48	0,60	0,55	0,55	0,57	0,57	0,58	0,59	0,75	0,61	0,76	0,80	0,54	0,61	0		
(17) 06-37.5B	0,69	0,51	0,56	0,56	0,57	0,55	0,56	0,58	0,55	0,71	0,39	0,63	0,70	0,29	0,50	0,58	0	
(18) 06-44.1B	0,71	0,54	0,58	0,59	0,58	0,60	0,56	0,59	0,57	0,70	0,41	0,61	0,73	0,35	0,50	0,60	0,11	0

Anexo G. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (*Zea mays* L.) obtidas para 215 bandas polimórficas.

Linhagens	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
(01) 05-01.4B	0																	
(02) 05-05.2A	0,68	0																
(03) 05-17.1A	0,72	0,52	0															
(04) 05-18.6A	0,66	0,47	0,56	0														
(05) 05-19.1B	0,77	0,51	0,56	0,53	0													
(06) 05-23.2B	0,74	0,51	0,57	0,57	0,55	0												
(07) 05-33.5B	0,73	0,52	0,57	0,59	0,54	0,26	0											
(08) 05-34.2B	0,73	0,53	0,39	0,52	0,51	0,55	0,54	0										
(09) 06-03.5B	0,73	0,56	0,62	0,63	0,63	0,67	0,67	0,65	0									
(10) 06-06.3A	0,63	0,70	0,71	0,67	0,73	0,72	0,71	0,73	0,70	0								
(11) 06-08.1A	0,73	0,54	0,60	0,61	0,57	0,52	0,48	0,62	0,57	0,70	0							
(12) 06-08.2A	0,63	0,71	0,71	0,70	0,73	0,73	0,74	0,73	0,73	0,56	0,56	0						
(13) 06-14.4B	0,70	0,80	0,73	0,75	0,76	0,75	0,75	0,71	0,72	0,56	0,73	0,57	0					
(14) 06-24.7B	0,73	0,52	0,57	0,61	0,54	0,51	0,54	0,59	0,54	0,68	0,36	0,63	0,73	0				
(15) 06-28.1A	0,75	0,55	0,56	0,58	0,52	0,50	0,54	0,54	0,61	0,31	0,50	0,67	0,24	0,47	0			
(16) 06-29.7B	0,68	0,48	0,58	0,57	0,46	0,57	0,57	0,57	0,58	0,74	0,61	0,65	0,81	0,53	0,59	0		
(17) 06-37.5B	0,71	0,51	0,55	0,56	0,57	0,56	0,57	0,58	0,55	0,70	0,39	0,63	0,73	0,28	0,48	0,56	0	
(18) 06-44.1B	0,73	0,54	0,57	0,59	0,58	0,60	0,58	0,59	0,57	0,69	0,41	0,62	0,75	0,33	0,48	0,58	0,11	0

Anexo H. Distâncias genéticas (complemento do coeficiente de Jaccard) para as 18 linhagens endogâmicas de milho (*Zea mays* L.) obtidas para 300 bandas polimórficas.

Linhagens	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
(01) 05-01.4B	0																	
(02) 05-05.2A	0,63	0																
(03) 05-17.1A	0,63	0,41	0															
(04) 05-18.6A	0,66	0,43	0,47	0														
(05) 05-19.1B	0,69	0,40	0,49	0,48	0													
(06) 05-23.2B	0,69	0,43	0,49	0,53	0,51	0												
(07) 05-33.5B	0,66	0,46	0,47	0,55	0,53	0,24	0											
(08) 05-34.2B	0,64	0,41	0,33	0,48	0,47	0,50	0,52	0										
(09) 06-03.5B	0,68	0,55	0,54	0,56	0,55	0,69	0,59	0,58	0									
(10) 06-06.3A	0,53	0,68	0,72	0,66	0,70	0,72	0,71	0,71	0,67	0								
(11) 06-08.1A	0,69	0,50	0,57	0,56	0,53	0,55	0,57	0,60	0,55	0,63	0							
(12) 06-08.2A	0,59	0,67	0,71	0,68	0,70	0,72	0,71	0,70	0,68	0,44	0,49	0						
(13) 06-14.4B	0,59	0,74	0,71	0,74	0,70	0,72	0,69	0,70	0,68	0,50	0,69	0,51	0					
(14) 06-24.7B	0,69	0,52	0,58	0,56	0,52	0,57	0,59	0,58	0,54	0,66	0,36	0,56	0,70	0				
(15) 06-28.1A	0,67	0,49	0,55	0,55	0,48	0,54	0,55	0,53	0,57	0,66	0,45	0,62	0,72	0,46	0			
(16) 06-29.7B	0,62	0,49	0,49	0,48	0,48	0,56	0,54	0,50	0,55	0,72	0,58	0,69	0,75	0,56	0,55	0		
(17) 06-37.5B	0,69	0,50	0,57	0,56	0,53	0,56	0,57	0,59	0,58	0,66	0,34	0,57	0,71	0,32	0,47	0,56	0	
(18) 06-44.1B	0,69	0,50	0,58	0,58	0,54	0,57	0,58	0,59	0,58	0,64	0,36	0,56	0,71	0,33	0,46	0,58	0,11	0

48) 100000001111101000
 49) 010000001000000100
 50) 000000010000000000
 51) 010100001000000111
 52) 110001100000000100
 53) 010010001100000000
 54) 000001001000110110
 55) 010000001010110011
 56) 011110011100101011
 57) 111111110100011101
 58) 111111111010011111
 59) 011111110000111010
 60) 001000010010010010
 61) 111101111111111111
 62) 010000001101110011
 63) 101101010101101111
 64) 001000011101110001
 65) 000100000011000011
 66) 101110011000000111
 67) 000010000001110011
 68) 010000000000001000
 69) 100011100011011011
 70) 010000000010000100
 71) 001000001000100000
 72) 111110001111100100
 73) 010111101111111111
 74) 001000111010111111
 75) 110111100111111111
 76) 000000000000001000
 77) 000000101000100000
 78) 000001100000000000
 79) 011001110001011000
 80) 000000001100000000
 81) 011111110010010111
 82) 110100100000001111
 83) 001011010011111100
 84) 011000010000000011
 85) 101111110100001000
 86) 100000000100100000
 87) 011111110010011111
 88) 111111110111111111
 89) 000000000000010000
 90) 111111111111101111
 91) 100011100000011111
 92) 000100010010000000
 93) 001110011011111011
 94) 011011110111101011
 95) 000010000010010011
 96) 000000000100000000
 97) 010000001011010100
 98) 111111111111101111
 99) 000100000000000000
 100) 001100100000000100
 101) 000011100010001000
 102) 100100000101010000

```

104) 000110001110111011
105) 000001100111111011
106) 11111111011110011
107) 100000000000000000
108) 11111111011100111
109) 011111011000001100
110) 100000000101100000
111) 100000001000000000
112) 000000000011101001
113) 100000000101100000
114) 01101111011011111
115) 11111110011011111
116) 000000000010010000
117) 011100000110010111
118) 000000001000000000
119) 111101110111010111
120) 0000000000000000100
121) 100000000101100000
122) 01111101100001100
123) 100000010011110011
124) 000101111011111011
125) 100000100101110100
126) 111110011010011111
127) 111111111110011111
128) 011111110000001000
129) 011001100010011011
130) 011111111010011111
131) 000000000010000000
132) 000000001010011000
133) 000001101000000000
134) 110111101010011111
135) 001010110101001001
136) 000001100000000000
137) 110101111100100100
138) 001010000011011011
139) 101100000101100000
140) 010111100011011011
141) 101000011100100100
142) 011010111010011011
143) 0000000000000000100
144) 000001100000000000
145) 000000001110010011
146) 100000001011110011
147) 110000001101110011
148) 100100001001011111
149) 111101111111111110
150) 001001111010111011
151) 011010111010011111
152) 101000000101100000
153) 111111110010011111
154) 0000000000000000100
155) 00000000000000001000
156) 001000010000001000
157) 100000000001100000
158) 111111001111111010

```

```

159) 00000000000000001100
160) 0000000001000001011
161) 010111111010011111
162) 010100001000000011
163) 0000000000000001000
164) 010011101010010100
165) 111110000000000100
166) 000100000000001000
167) 110110010000001000
168) 011101010010011100
169) 010000000100000000
170) 010110011011011111
171) 100000001011000000
172) 011111111100011100
173) 101001101111011111
174) 000000000100000000
175) 011011110000011111
176) 000001100000011011
177) 100000000000100000
178) 000010010000101100
179) 010110000011010111
180) 011111111010011111
181) 001111111110111111
182) 001001001111111011
183) 000100100000110011
184) 000101100000000100
185) 000000000111000011
186) 010111100111111100
187) 000001100000000000
188) 101100001001000000
189) 110011110110111111
190) 111111111111011111
191) 000100001000001000
192) 001000010000000000
193) 011001110000100100
194) 001001001010011111
195) 010110001000001000
196) 011011111010011111
197) 000000000010011011
198) 110000100000000100
199) 111100111100011111
200) 010111000110010011
201) 100000110101100000
202) 010100000000000000
203) 011110010001001111
204) 000011101110010000
205) 111000000000000000
206) 000001010000001000
207) 100000000101100000
208) 000100010000000000
209) 000000001100011111
210) 011111111111011111
211) 111110011111011111
212) 100111100100100000
213) 011111111111111111

```