



**AMOSTRAGEM SEQUENCIAL E  
MARCADORES DE MICROSSATÉLITES  
NA AVALIAÇÃO DA QUALIDADE  
GENÉTICA EM LOTES DE SEMENTES DE  
MILHO**

**MAGNÓLIA DE MENDONÇA LOPES**

**2003**

55909

ME/047874

**MAGNÓLIA DE MENDONÇA LOPES**

**AMOSTRAGEM SEQUÊNCIAL E MARCADORES DE  
MICROSSATÉLITES NA AVALIAÇÃO DA QUALIDADE  
GENÉTICA EM LOTES DE SEMENTES DE MILHO**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Agronomia, área de concentração Fitotecnia, para obtenção do título de "Mestre".

Orientadora

Profa. Dra. Maria das Graças Guimarães Carvalho Vieira

LAVRAS  
MINAS GERAIS - BRASIL

2005

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da  
Biblioteca Central da UFLA**

**Lopes, Magnólia de Mendonça**

**Amostragem seqüencial e marcadores de microssatélites na avaliação da  
qualidade genética em lotes de sementes de milho / Magnólia de Mendonça  
Lopes. – Lavras : UFLA, 2003.**

**47 p. : il.**

**Orientadora: Maria das Graças Guimarães Carvalho Vieira.**

**Dissertação (Mestrado) – UFLA.**

**Bibliografia.**

**1. Amostragem. 2. Microssatélite. 3. Semente. 4. Milho. I. Universidade  
Federal de Lavras. II. Título.**

**CDD-633.1521**

**MAGNÓLIA DE MENDONÇA LOPES**

**AMOSTRAGEM SEQUÊNCIAL E MARCADORES DE  
MICROSSATÉLITES NA AVALIAÇÃO DA QUALIDADE  
GENÉTICA EM LOTES DE SEMENTES DE MILHO**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Agronomia, área de concentração Fitotecnia, para obtenção do título de "Mestre".

**APROVADA em 18 de junho de 2003**

**Profa. Dra. Édila Vilela Resende Von Pinho UFLA**

**Prof. Dr. Marcelo Silva de Oliveira UFLA**

  
**Profa. Dra. Maria das Graças Guimarães Carvalho Vieira**

**UFLA**

**(Orientadora)**

**LAVRAS**

**MINAS GERAIS - BRASIL**

## **AGRADEÇO**

A Deus,  
pelo dom da vida  
e bênçãos a mim concedidas.

## **OFEREÇO**

Aos meus pais, Nonato e Leondina,  
exemplos de amor e dedicação, me apoiando incondicionalmente em todos  
os momentos de minha vida.

## **DEDICO**

Aos meus irmãos Rômulo, Lutz, Cristhian e, em especial, à minha irmã e  
amiga Anne, pelo carinho e apoio incondicional.  
Ao Emídio e à minha sobrinha Larissa.

## **AGRADECIMENTOS**

À Universidade Federal de Lavras (UFLA) e a Empresa Brasileira de Pesquisa Agropecuária, Embrapa Milho e Sorgo, pela oportunidade de realização do curso e desenvolvimento científico.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo suporte financeiro.

À Prof. Maria das Graças Guimarães Carvalho Vieira, pela amizade, incentivo constante, orientação e conhecimentos transmitidos.

À Dra. Lilian Padilha, pelo incentivo, colaboração e amizade.

Ao Dr. Marcelo Silva de Oliveira, pela colaboração e sugestões na parte estatística.

À Dra. Cláudia Teixeira Guimarães, pelas sugestões e apoio no desenvolvimento desse trabalho.

À Dra. Édila Vilela de Resende Von Pinho, pelas sugestões e ensinamentos.

Aos professores do Setor de Sementes, pela amizade, convivência e conhecimentos transmitidos.

Aos funcionários setor de sementes, Andrea, Elza e Dalva, pela convivência e amizade.

Aos amigos do LAS e NBA da Embrapa Milho e Sorgo, Tônico, Mônica, Bira e Miguel que, além da amizade, contribuíram para a realização deste trabalho em laboratório.

Aos amigos Maria de Lourdes, Paula, Flávio Henrique, Ildon, Maurício Godoy, Rubens, Val, Enia e Ivânia, pela amizade e apoio.

A todos aqueles que, de alguma forma, contribuíram para a realização deste trabalho.

## SUMÁRIO

	<b>Página</b>
LISTA DE TABELAS .....	i
LISTA DE FIGURAS .....	.ii
RESUMO .....	iii
ABSTRACT .....	iv
1 INTRODUÇÃO.....	1
2 REFERENCIAL TEÓRICO.....	4
2.1 Amostragem seqüencial.....	4
2.1.1 Aspectos teóricos.....	4
2.1.2 Amostragem sequencial na avaliação da qualidade de sementes.....	12
2.2 Pureza genética.....	13
2.3 Marcadores microssatélites.....	17
3 MATERIAL E MÉTODOS.. ..	23
3.1 Amostragem seqüencial.....	23
3.2 Marcadores microssatélites .....	29
4 RESULTADOS E DISCUSSÕES .....	32
4.1 Amostragem seqüencial.....	32
4.2 Marcadores microssatélites.....	36
5 CONCLUSÕES .....	43
6 REFERÊNCIAS BIBLIOGRÁFICAS.....	44

## LISTA DE FIGURAS

Figura	Página
1 Perfil de bandas geradas pelos <i>primers</i> bnlg2291; bnlg1006; umc1037; bnlg669 e bnlg238 que permitem discriminar os híbridos de seus respectivos parentais e a detecção das diferentes proporções de misturas entre os parentais P1 e P2. Embrapa, Sete Lagoas, MG, 2003.....	38
2 Perfil de bandas geradas pelos marcadores umc1019; mmc162 e umc1016 da mistura de DNA (A) ou de discos foliares (B) do híbrido HS3 e da Linhagem L1 em diferentes proporções. Embrapa, Sete Lagoas, MG, 2003.....	40

## LISTA DE TABELAS

Tabela	Página
1	Valores de $p_0$ e $p_1$ para diferentes níveis de contaminação varietal em lotes de sementes híbridas de milho. Embrapa, Sete Lagoas, MG, 2003..... 24
2	Valores calculados dos limites de decisão do plano de amostragem seqüencial para $p_0: 0,9975$ ; $p_1: 0,999$ $\alpha$ e $\beta=0,05$ para pureza genética de 0,25% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003..... 26
3	Valores calculados dos limites de decisão do plano de amostragem seqüencial para $p_0: 0,995$ ; $p_1: 0,999$ $\alpha$ e $\beta=0,05$ para pureza genética de 0,5% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003..... 26
4	Valores calculados dos limites de decisão do plano de amostragem seqüencial para $p_0: 0,99$ ; $p_1: 0,999$ $\alpha$ e $\beta=0,05$ para pureza genética de 1% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003..... 27
5	Valores calculados dos limites de decisão do plano de amostragem seqüencial para $p_0: 0,98$ ; $p_1: 0,999$ $\alpha$ e $\beta=0,05$ para pureza genética de 2% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003..... 27
6	Valores calculados dos limites de decisão do plano de amostragem seqüencial para $p_0: 0,96$ ; $p_1: 0,999$ $\alpha$ e $\beta=0,05$ para pureza genética de 4% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003..... 28
7	Valores calculados dos limites de decisão do plano de amostragem seqüencial para $p_0: 0,94$ ; $p_1: 0,999$ $\alpha$ e $\beta=0,05$ para pureza genética de 6% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003..... 28

8	Híbridos de milho e seus respectivos progenitores utilizados na avaliação da qualidade genética de sementes por meio da técnica de SSR. Embrapa, Sete Lagoas, MG, 2003.....	30
9	Número de sementes analisadas no plano de amostragem seqüencial para a decisão de aceitar ou rejeitar o lote, em função da percentagem de mistura varietal. Embrapa, Sete Lagoas, MG, 2003.....	32
10	Tamanho Médio de Amostra (TMA(p)), para lotes de sementes de milho com diferentes porcentagens de mistura varietal, em função dos valores atribuídos a W. Embrapa, Sete Lagoas, MG, 2003.....	34
11	Conjuntos de <i>primers</i> que permitem a distinção do híbrido de seus parentais. Embrapa, Sete Lagoas, MG, 2003.....	37

## RESUMO

LOPES, Magnólia de Mendonça. **Amostragem sequencial e marcadores de microssatélites na avaliação da qualidade genética em lotes de sementes de milho.** Lavras: UFLA, 2003. 47p. (Dissertação - Mestrado em Fitotecnia)\*

No sistema de produção de sementes, a pureza genética é um dos requisitos fundamentais para a sua comercialização. Nesse sentido são estabelecidos padrões de tolerância e há necessidade de metodologias que detectem contaminação, bem como definir o tamanho de amostra a ser utilizado de forma a garantir ao produtor e ao consumidor, dentro de um determinado nível de segurança, a pureza genética do lote de sementes. O presente trabalho objetivou determinar o tamanho da amostra para a avaliação da pureza genética, de forma a proteger o consumidor e o produtor de sementes, em níveis de segurança conhecidos e predeterminados e avaliar a sensibilidade da técnica de microssatélites para discriminar híbridos de seus respectivos parentais e a sensibilidade em detectar misturas, quando presentes em pequenas proporções na amostra. Para a amostragem seqüencial sementes híbridas foram marcadas e misturadas ao lote de sementes, simulando contaminações de 0,25%; 0,5%; 1%; 2%; 4% e 6%. Posteriormente, foram tomados grupos de 40 sementes em seqüência, até no máximo 400 sementes, visando determinar a quantidade de sementes necessária para detectar as percentagens de mistura acima mencionadas. A sensibilidade da técnica de microssatélite foi avaliada misturando-se diferentes proporções de DNA dos híbridos com o de suas respectivas linhagens. A técnica de microssatélite foi eficiente em detectar pequenas proporções de DNA em mistura na amostra, bem como discriminar o híbridos dos seus respectivos parentais. No entanto, a um nível de mistura superior a 1:8 (1P1:8P2; 8P1:1P2), a sensibilidade do marcador para detectar diferentes proporções das misturas variou em função do *primer* utilizado. Quanto à amostragem seqüencial, verificou-se que para detectar níveis de mistura acima de 1% no lote de sementes, com nível de risco tanto para o produtor quanto para o consumidor de 0,05, o tamanho de amostra necessário foi inferior ao requerido pela amostra de tamanho fixo. Isto possibilitou também a redução do custo, viabilizando o uso da técnica para atestar a pureza genética de lotes de sementes de milho.

---

\* Comitê Orientador: Dra. Maria das Graças Guimarães Carvalho Vieira - UFLA (Orientadora), Dra. Lilian Padilha-Embrapa Milho e Sorgo, Dr. Marcelo Silva de Oliveira – UFLA Dra. Claudia Teixeira Guimarães - Embrapa Milho e Sorgo.

## ABSTRACT

**LOPES, Magnolia de Mendonça. Sequential sampling and markers of microsatellites in the evaluation of the genetic quality in lots of corn seeds. Lavras: UFLA, 2003. 47p (Dissertation - Master in Crop Science)\***

In the seed production system, genetic purity is one of the fundamental requirements for its commercialization. For this reason, standards of tolerance have been established and, then methodologies have to made available to detect contamination and to define the sample size to be used in order to guarantee a certain level of genetic purity within the lot of seeds for both the producer and the consumer. The present work objectified to determine the sample size for the evaluation of genetic purity, in order to protect the consumer and the producer of seeds and to evaluate the sensitivity of the microsatellite technique for discriminating hybrids from their respective relatives and for detecting mixtures when they are present in small amounts in the samples. For the sequential sampling, hybrid seeds were marked and mixed in with the seed lots, simulating the following levels of contamination: 0.25, 0.5, 1.0, 2.0, 4.0, and 6.0%. After this, groups of 40 seeds were taken in sequence, up to a maximum of 400 seeds, with the objective of determining the quantity of seeds necessary to detect the percentage of mixture mentioned above. The sensitivity of the microsatellite technique was evaluated by mixing different proportions of DNA from the hybrids with their respective seed lines. This technique not only efficiently detected the small proportions of DNA mixed in with the sample, it also discriminated the hybrids from their respective relatives. However, when the level of mixture was higher than 1:8 (1P1:8P2; 8P1:1P2), the sensitivity of the marker in detecting different proportions of the mixture varied according to the primer used. In terms of the sequential sampling, it was verified that in order to detect mixture levels higher than 1% within the seed lot- with a risk level for both the producer and the consumer of 0.05- the size of the necessary sample was smaller than the size needed for the fixed-size sample. This also made it possible to reduce the cost, which made possible the use of the microsatellite technique in order to certify to the genetic purity of the lots of corn seeds.

---

\*Supervising Committee: Dra. Maria das Graças Guimarães Carvalho Vieira - UFLA (Adviser), Dra. Lilian Padilha - Embrapa Milho e Sorgo, Dr. Marcelo Silva de Oliveira - UFLA and Dra. Claudia Teixeira Guimarães - Embrapa Milho e Sorgo.

## 1 INTRODUÇÃO

A cultura do milho participa de maneira expressiva nos mercados nacional e mundial. Sua importância resulta num setor de sementes competitivo, em que as empresas de produção de sementes beneficiam-se do fato de a maioria dos materiais comercializados ser híbrida, caracterizando uma barreira natural à utilização das sementes em plantios subsequentes. Tal característica justifica os investimentos nos programas de melhoramento do milho por permitir às empresas o retorno dos investimentos no desenvolvimento de genótipos com ampla adaptação e estabilidade, mesmo quando ainda não existia uma política de proteção de cultivares.

O milho é um exemplo marcante do sucesso do conhecimento de genética aplicada. Nesse sentido, Paterniani & Campos (1999) mencionam que as cultivares mais recentes de milho produzem quatro vezes ou mais quando comparadas às cultivares mais antigas.

O desenvolvimento de variedades com desempenho superior e adaptadas aos ambientes de cultivo é um dos pontos essenciais para a obtenção de alta produtividade e qualidade dos grãos de milho. Na obtenção do material a ser multiplicado, o produtor deve ter garantida a qualidade das sementes, que além de envolver atributos fisiológicos, físicos e sanitários, deve envolver também a pureza genética das sementes.

Dessa forma, a certificação da pureza genética é indispensável para assegurar aos agricultores produto com as características genéticas desenvolvidas pelos melhoristas. No entanto, em função do estreitamento genético que vem ocorrendo para a maioria das espécies, torna-se cada vez mais difícil a diferenciação dos mesmos, utilizando simplesmente características morfológicas. Diante disso, a utilização de técnicas de biologia molecular cresce

em importância, de forma a contribuir de maneira expressiva para o aumento da precisão das informações, viabilizando a certificação das cultivares de interesse.

A escolha do método molecular depende da estrutura genética de cada espécie, do seu modo de reprodução, do tamanho do genoma, entre outros fatores e, sem dúvida alguma, da relação custo-benefício. Na área agrônômica, os marcadores moleculares têm apresentado aplicação direta no monitoramento da pureza genética de sementes. Dentre as técnicas atualmente disponíveis, destaca-se a de microssatélite ou SSR, que amplifica segmentos repetidos ao longo da molécula do DNA sendo está, caracterizada pela simplicidade, rapidez e precisão na geração dos perfis genéticos.

Essa técnica poderá ser viabilizada para o monitoramento da pureza genética devido às várias vantagens que apresenta, como o elevado conteúdo de informação genética por loco ser multialélica, codominante, ser passível de automação, aumentando a resolução e precisão das informações.

No entanto, em rotina em que se tem necessidade de analisar de forma rápida grande quantidade de lotes, a quantidade de sementes a ser utilizada na certificação desse tipo de técnica tem sido um dos principais questionamentos. Neste aspecto, deve-se considerar essencial um sistema que permita monitorar e rastrear contaminações genéticas que possam ocorrer no campo, permitindo o fornecimento de sementes com elevados padrões de pureza genética aos produtores de sementes certificadas.

Um ponto crucial para o sucesso na certificação da pureza genética e que deve ser focado é o tamanho da amostra para garantir resultados rápidos, que proporcionem segurança tanto ao produtor quanto ao consumidor de sementes em níveis conhecidos e preestabelecidos, sem perder de vista a viabilidade operacional. Nesse sentido, pesquisas recentes efetuadas pelo Comitê de Estatística da ISTA (1999) têm apontado, como uma possível alternativa para

análises que requerem rapidez, precisão e redução de custos, a amostragem seqüencial.

A amostragem seqüencial caracteriza-se por utilizar amostras de tamanho variável. Uma vez formulada a hipótese de que o lote está dentro do padrão exigido pelo sistema de certificação e fiscalização, pode-se aceitá-la, rejeitá-la ou continuar amostrando, baseando-se nos resultados acumulados de cada amostragem. As principais vantagens deste método são a redução de tempo e de custos, e assegurar que os riscos sejam os mesmos tanto para o produtor quanto para o consumidor, sem necessitar de um número fixo de amostras (Pieters, 1978). Em certos casos, a amostragem seqüencial requer, em média, amostras com 1/3 do tamanho que seria usado com a amostragem de tamanho fixo (Wald, 1947).

O presente trabalho teve como objetivos determinar o tamanho da amostra para a avaliação da pureza genética, de forma a proteger o consumidor e o produtor de sementes, em níveis de segurança conhecidos e predeterminados, e avaliar a sensibilidade da técnica de microssatélites para discriminar híbridos de seus respectivos parentais. Também foi objetivo avaliar a sensibilidade da técnica de microssatélite em detectar misturas, quando presentes em pequenas proporções na amostra.

## 2 REFERENCIAL TEÓRICO

### 2.1 Amostragem seqüencial

#### 2.1.1 Aspectos teóricos

Na avaliação da qualidade de sementes, o método convencional prevê uma amostra de tamanho fixo. Porém, mais recentemente, em face das técnicas disponíveis e do tempo gasto para amostrar e da quantidade de amostras a serem analisadas em um laboratório de rotina, vem-se notando um maior interesse em estudos relacionados à amostragem seqüencial. A vantagem em relação à amostragem de tamanho fixo é que a amostragem seqüencial caracteriza-se por envolver amostras de tamanho variável, o que pode levar a uma redução nos custos, porém, mantendo-se o controle dos erros que poderiam ser cometidos. Outra diferença é que, enquanto na amostragem convencional procura-se estimar os parâmetros populacionais para depois, eventualmente, testar uma hipótese a respeito dos mesmos, na amostragem seqüencial é testada uma hipótese a respeito desses parâmetros sem a preocupação em estimá-los. A vantagem do tamanho da amostra utilizada na amostragem seqüencial se faz sentir principalmente quando as sementes são de alto valor econômico ou quando o tempo de análise é limitado (Wald, 1947).

Os métodos clássicos para a amostragem de tamanho fixo e os métodos inferenciais estão descritos em diversos trabalhos. A teoria da amostragem seqüencial foi proposta por Wald (1945 e 1947) e, a partir daí, aplicações dessa teoria têm sido apresentadas por muitos autores. Dentre eles Estefanel (1977), Bányai (1978), Ellis & Whitehead, (1987), Jackish et al. (1984) e Santana (1994), que colocam seus elementos de uma maneira que torna possível uma imediata aplicação em situações práticas envolvendo ciências agrárias.

Para elaborar um plano de amostragem seqüencial é necessário conhecer a distribuição da variável que está sendo levantada pela amostragem, formular hipóteses e estabelecer os riscos de tomar decisões erradas. Na amostragem seqüencial, as unidades amostrais são examinadas em seqüência e os resultados são obtidos por meio dos limites previamente determinados. Assim, com base na primeira observação, pode-se chegar a uma das três seguintes decisões: (1) aceitar a hipótese  $H$  formulada, (2) rejeitar a hipótese  $H$  e (3) continuar a amostragem. Se for tomada a decisão (3), outra unidade amostral é examinada e, com base nos resultados acumulados, as três possíveis decisões são examinadas novamente. O processo é continuado até que se possa tomar a decisão (1) ou a (2).

A grande vantagem da amostragem seqüencial consiste na considerável redução de tempo que ela proporciona por evitar a excessiva amostragem quando o parâmetro populacional  $\theta$  é bem maior ou bem menor que o valor  $\theta$  determinado para comparação. Wald (1947) afirma que, em certos casos, a amostragem seqüencial requer, em média, amostras com  $1/3$  do tamanho que seria usado com a amostragem de tamanho fixo.

Para dar início às operações seqüenciais é necessário levar em conta bases estatísticas que irão auxiliar na elaboração do plano. Desse modo, deve-se considerar a inferência estatística, teste de significância e o teste de hipótese.

A avaliação da qualidade de um lote de sementes a partir de uma amostra desse lote é, de fato, um procedimento inferencial, podendo ser classificado em dois tipos: estimação de parâmetros que descrevem a população (o lote) sob estudo ou teste de hipóteses sobre esta população. Na produção de sementes estimar ou testar são considerados procedimentos úteis, porém testar ainda é mais prático e eficiente em muitos casos. Existem dois tipos de teste de hipóteses: os testes de significância, que avaliam as evidências que uma amostra apresenta, contra ou a favor de uma determinada hipótese chamada de hipótese

de nulidade,  $H_0$ , e os testes de hipóteses propriamente ditos, que consideram também uma segunda hipótese  $H_1$ , chamada de hipótese alternativa.

No teste de significância, busca-se controlar a probabilidade de que o erro tipo I não está sendo cometido (nível de significância  $\alpha$ ), isto é, ter a certeza de que não está sendo rejeitada uma hipótese verdadeira. Esse teste só se preocupa com o produtor, não se importando com os riscos do comprador. O custo desse teste é diretamente proporcional ao tamanho da amostra, de modo que o teste tenha um nível de significância  $\alpha$  preestabelecido.

No controle da qualidade de sementes, um dos pontos cruciais é garantir a pureza genética do lote de sementes. Desse modo, um dos parâmetros a serem avaliados é a porcentagem de contaminação desse lote em relação a esses fatores. Pode-se, então, usar um teste de significância para testar  $H_0: p=p_0$ , ou seja, a hipótese de que a porcentagem de contaminação de um dado lote é igual a  $p_0$ . Um teste de significância verificará se uma amostra de  $n$  sementes reúne evidências suficientes para rejeição de  $H_0$ .

O procedimento seqüencial é dado, a princípio, usando uma amostra com tamanho  $n$  fixo, considerando o custo da inferência. No entanto, nada impede que teste-se seqüencialmente, não fixando *à priori* o tamanho da amostra.

A cada semente ensaiada para contaminação, pode-se calcular o tamanho  $n$ , pela fórmula: 
$$n = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Se  $n > z\alpha$  (0,05); rejeita-se. Se  $n < z\alpha$ , continua-se amostrando. Este procedimento seqüencial apresenta como fator limitante a possibilidade de se amostrar indefinidamente, sem tomar uma decisão sobre o lote de sementes. Uma solução para este problema seria limitar o número de sementes ensaiadas em um valor  $n$ -máximo (por exemplo,  $n = 400$  sementes). Desta forma, constrói-se um teste seqüencial truncado, de significância, com risco de cometer o erro de rejeitar uma hipótese verdadeira, prejudicando assim o produtor de sementes.

Ao proceder com o teste de significância por meio do teste de hipóteses corre-se o risco de cometer tanto o erro tipo I, quanto o erro tipo II, este último, com uma probabilidade  $\beta$ . Só que este risco é totalmente descontrolado, somente se controlando a probabilidade  $\alpha$  do erro tipo I. Portanto, nota-se que na versão do tamanho fixo para amostras, como na versão seqüencial, os testes de significância atentam somente para os riscos do produtor das sementes, não levando em consideração que o comprador das sementes também corre riscos, e que estes devem também ser controlados. Enquanto um produtor perde se uma boa semente é classificada como ruim num teste, um comprador será prejudicado se uma semente de má qualidade foi classificada como boa pelo mesmo teste, ou seja, ambos são erros e ambos devem ser controlados.

Já os testes de hipótese são baseados no Teorema de Neyman e Pearson; tanto o  $\alpha$  quanto  $\beta$  são controlados, mas não limita o risco. O usuário determina *a priori* o valor que ele admite para o  $\alpha$  e sabe que o  $\beta$  está acontecendo no menor valor possível para aquele  $\alpha$  escolhido. Usualmente, mesmo sabendo que  $\beta$  é mínimo, não se sabe o valor de  $\beta$ , sendo este uma função da verdadeira porcentagem de contaminação  $p$  do lote e da amostra. Dessa forma, permanece ainda sem solução uma parte do problema, que é controlar  $\beta$  a um nível conhecido. Observa-se que a curva característica de operação do teste daria os valores de  $\beta$  para cada valor possível da porcentagem  $p$  de contaminação do lote se este valor fosse conhecido. Como  $p$  não é conhecido, uma solução seria calcular o tamanho  $n$  da amostra necessário para que tanto o  $\alpha$  quanto o  $\beta$  alcançassem níveis pré-fixados: 
$$n = \hat{p}(1 - \hat{p}) \frac{(z_\beta - z_\alpha)^2}{(p_1 - p_0)^2}$$

Este procedimento é forçosamente seqüencial, pois depende de  $\hat{p}$ , o valor amostral de  $p$ . Assim que o valor de  $n$  alcançar, ou ultrapassar, o valor dado pela expressão acima, pode-se parar de amostrar e tomar uma decisão

(aceitando ou rejeitando  $H_0$ ): se  $H_0$  é rejeitada, corre-se o risco  $\alpha$ , ou seja, rejeitar  $\alpha$  erroneamente, e se  $H_0$  é aceita, corre-se um risco  $\beta$  de  $H_0$  ter sido aceita quando na verdade deveria ser rejeitada.

Quando se trabalha com amostragem seqüencial, seja em testes de significância ou em testes de hipóteses, deve-se considerar o tamanho de amostra necessária para se tomar uma decisão. O interesse maior é conhecer, na amostragem seqüencial, o tamanho médio de amostra (*TMA*) que informará quantas sementes em média terão que ser analisadas para finalmente se tomar uma decisão quanto à pureza do lote. O interessante é tornar mínimo esse tamanho de amostra. Portanto, em testes de hipóteses envolvendo sementes, três elementos devem ser considerados: as probabilidades  $\alpha$  e  $\beta$  de se tomar decisões erradas, que prejudicam tanto ao produtor quanto ao consumidor e o tamanho da amostra utilizada.

A teoria de Wald (1947) é apresentada nesta revisão para o teste seqüencial da razão de probabilidade e função do tamanho médio da amostra.

A primeira idéia da amostragem seqüencial foi lançada por Dooge et al. (1929) ao sugerirem a amostragem em duas etapas sucessivas. Posteriormente, outros autores aperfeiçoaram a idéia, porém, o procedimento ainda necessitaria de bases teóricas.

Wald (1945,1947), trabalhando para a Marinha dos Estados Unidos, iniciou a pesquisa em 1940, visando ao controle de qualidade do material bélico. A partir de então, muitas indústrias passaram a usar o procedimento desenvolvido por Wald, baseado no teste seqüencial da razão de probabilidade (TSRP), sendo a técnica mais utilizada e difundida da amostragem seqüencial, posteriormente adaptada por biólogos e entomologistas (Oakland, 1950; Morris, 1954). Desde então, diversos trabalhos têm sido realizados sobre amostragem seqüencial. Essa teoria assume que os dados que estão sendo analisados são coletados como uma seqüência de variáveis aleatórias  $X_1, X_2, X_3, \dots$ , e que

finalmente deseja-se decidir entre duas hipóteses,  $H_0$  e  $H_1$ . Estabelecido isto, o TSRP de Wald determina que, em qualquer estágio da amostragem seqüencial, onde já se coletou uma amostra de tamanho  $n$ :

( i )  $p(x_1, x_2, \dots, x_n; \theta_1) / p(x_1, x_2, \dots, x_n; \theta_0) \geq A$ , então, a amostragem deve ser cessada e deve-se rejeitar  $H_0$  (e aceitar  $H_1$ ).

( ii )  $p(x_1, x_2, \dots, x_n; \theta_1) / p(x_1, x_2, \dots, x_n; \theta_0) \leq B$ , então, a amostragem deve ser cessada e deve-se rejeitar  $H_1$  (e aceitar  $H_0$ ).

( iii )  $A < p(x_1, x_2, \dots, x_n; \theta_1) / p(x_1, x_2, \dots, x_n; \theta_0) < B$  então deve-se continuar amostrando seqüencialmente, isto é, observar o valor  $x_{n+1}$  da variável sob estudo.

As constantes  $A$  e  $B$  são determinadas de tal maneira que as probabilidades de se cometer os erros tipo I e II sejam, respectivamente,  $\alpha$  e  $\beta$ , valores pré-definidos pelo pesquisador /usuário:

- ✓ Prob  $\alpha$  (erro tipo I) = risco do produtor, ou seja, rejeitar um lote de sementes que atenda as exigências mínimas de pureza.
- ✓ Prob  $\beta$  (erro tipo II) = risco do consumidor, ou seja, aceitar um lote de sementes que não atenda as exigências mínimas de pureza.

Pela definição acima, pode-se observar que a TSRP prescreve uma amostragem seqüencial com tamanho de amostra indefinida (uma amostragem seqüencial “aberta”). Em alguma situação específica, o tamanho da amostra pode ser demasiadamente grande. Apesar de ser possível demonstrar que toda amostragem seqüencial determinada pelo TSRP termina ou seja, para algum valor de  $n$  ela cessará, tem-se ainda considerada a possibilidade de se utilizar procedimentos truncados (Wald, 1947; Ellis & Whitehead, 1987), em que um número máximo de elementos amostrados é imposto.

Para esclarecer a aplicação desta teoria na área de sementes, considere-se o parâmetro porcentagem  $p$  de pureza genética de um lote de sementes. Deseja-se decidir entre dois valores para  $p$ . Ou o lote está com uma porcentagem de pureza  $p_0$  abaixo do desejado, ou está com uma porcentagem  $p_1$  dentro do aceitável. Para se evitar que a avaliação da pureza seja feita semente a semente, amostram-se seqüencialmente grupos de sementes de tamanho  $n$ . O desenvolvimento de TSRP para esta situação determina que em cada grupo amostrado conte-se o número de sementes puras ( $X_1$  para o primeiro grupo,  $X_2$  para o segundo, etc.) e comparem-se os dois limites de decisão:  $li$ , limite inferior e  $ls$ , limite superior. Se para um dado tamanho  $n$  de amostra, oriundo de  $k$  grupos de  $m$  sementes, isto é,  $n = mk$ , o total acumulado de sementes puras é igual ou menor do que  $li$ , então toma-se a decisão de rejeitar  $H_1$  e aceitar  $H_0$ , considerando o lote com porcentagem de pureza abaixo do desejado. Se  $\sum_{i=1}^k X_i$  é igual ou maior do que  $ls$ , rejeita-se  $H_0$ , considerando o lote como aprovado quanto à porcentagem de pureza.

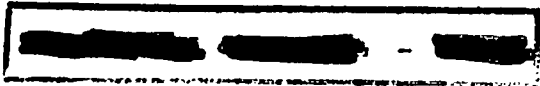
Se  $li < \sum_{i=1}^k X_i < ls$ , continua-se amostrando. Para o estudo da porcentagem de pureza de sementes pode-se demonstrar que os limites,  $li$  e  $ls$ , são funções lineares de  $n$ . Com o intuito de facilitar os cálculos é vantajoso usar o logaritmo da razão de probabilidades no lugar da razão propriamente dita. Assim pode-se escrever onde  $L$  indica o logaritmo neperiano. Pode ser usado o logaritmo em qualquer base, entretanto, para a distribuição de Poisson há vantagens em usar o logaritmo neperiano, segundo a fórmula:

$$l_i = b + an \quad \text{e} \quad l_s = c + an$$

$$a = \frac{Ln \frac{1 - P_0}{1 - P_1}}{Ln \frac{P_1 - (1 - P_0)}{P_0(1 - P_1)}}$$

$$b = \frac{Ln \frac{\beta}{1 - \alpha}}{Ln \frac{P_1(1 - P_0)}{P_0(1 - P_1)}}$$

$$c = \frac{Ln \frac{1 - \beta}{\alpha}}{Ln \frac{P_1(1 - P_0)}{P_0(1 - P_1)}}$$



O estudo do tamanho médio de amostra (*TMA*) para um TSRP para porcentagem de contaminação genética em sementes em função de *p* (*TMA(p)*) pode ser feito a partir da expressão:

$$TMA(p) = \frac{P(p) \cdot (b - c) + c}{p - a}$$

Em que  $P(p) = \frac{\left(\frac{1-\beta}{\alpha}\right)^W - 1}{\left(\frac{1-\beta}{\alpha}\right)^W - \left(\frac{\beta}{1-\alpha}\right)^W}$  é a probabilidade de aceitação de  $H_0: p=p_0$

Os valores de *p* e *W* estão associados pela expressão:

$$p = \frac{1 - \left(\frac{1-p_1}{1-p_0}\right)^W}{\left(\frac{p_1}{p_0}\right)^W - \left(\frac{1-p_1}{1-p_0}\right)^W}$$

Uma tabela com os valores de *TMA* (*p*) em função de *p* pode ser obtida atribuindo-se valores a *W* num intervalo adequado, o que é um expediente matemático para facilitar os cálculos. Em determinadas situações, às vezes é necessário garantir que o tamanho da amostra de sementes não ultrapassará um determinado valor *n*<sub>max</sub>. Neste caso, procedimentos seqüenciais truncados podem ser usados, como, por exemplo, aqueles tratados em Ellis & Whitehead (1987).

Finalmente, os conceitos de  $\alpha$  e  $\beta$  podem ser generalizados também aqui em amostragem seqüencial usando o TSRP de Wald como feito anteriormente, em função da probabilidade (aceitação  $H_0$ , dada a porcentagem *p* de pureza do lote). As fórmulas para construir esta função são um tanto complexas e não serão

apresentadas nesta revisão. Esta função pode ser um critério útil para comparar procedimentos que testam qualidade de sementes.

### **2.1.2 Amostragem seqüencial na avaliação da qualidade de sementes**

No sistema de produção de sementes, reduzir custos e evitar gastos adicionais é, hoje, um dos pontos fundamentais no sucesso do empreendimento. É imprescindível a obtenção de amostras representativas do lote, de modo que os resultados dos testes retratem a realidade do lote. Portanto, a decisão de aceitar ou rejeitar um lote de sementes é dependente da amostragem efetuada.

Alguns testes determinantes da qualidade de sementes requerem a utilização de materiais que elevam o custo da análise. Desse modo, o Comitê de Estatística da ISTA (1999) está atentando para a redução do tamanho da amostra, sem levar a uma baixa precisão do teste. A amostragem seqüencial tem sido uma das alternativas para esse problema. Entretanto, pela necessidade do conhecimento do método e de técnicas estatísticas, ainda não está sendo recomendada para a análise de rotina.

Estudos envolvendo amostragem seqüencial em sementes permitiram a redução do tamanho de amostra fixa de 400 sementes para 100 no plano seqüencial, utilizando a metodologia do ácido fucsina para a capacidade germinativa. Lotes com baixa e alta capacidade germinativa apresentaram valores médios de amostras menores em comparação aos requeridos por lotes de capacidade intermediária (Jackish et al., 1984).

Santana et al. (1994) constataram a viabilidade de se utilizar a amostragem seqüencial, com redução do tamanho de amostra, para se determinar pelo teste de pH do exsudado a qualidade fisiológica de sementes de milho. Trabalhando com amostragem seqüencial na detecção de contaminação varietal em sementes, Bányai (1987 e 1978) observou que é inviável utilizar o plano de

amostra fixo, principalmente em análises que envolvem métodos eletroforéticos para avaliação da pureza varietal, devido ao custo dessa análise. Em 1992, o mesmo autor ressaltava a decisão do Comitê de Estatística da ISTA em relação ao uso da amostragem seqüencial, na tentativa de diminuir os riscos do plano, estudando ainda o desenvolvimento do plano de amostragem seqüencial truncado e triangular na expectativa de que essas variações do plano pudessem reduzir o tamanho da amostra.

Segundo Ellis & Whitehead (1987), tanto o plano seqüencial quanto o plano de amostra fixa, em função da escolha entre o truncamento e o teste triangular, dependem dos valores de  $p$ . Os limites de aceitação e rejeição do teste seqüencial triangular são inicialmente muito afastados, quando comparados ao do truncamento. Para o triangular é mais difícil uma tomada de decisão quanto a qualidade do lote nas primeiras amostras, enquanto no método truncado, o término do teste ocorre quando o número de sementes testadas é igual a 400, valor este correspondente ao plano de amostra fixo.

O sucesso da amostragem seqüencial depende do parâmetro a ser adotado, pois a seleção dos parâmetros mais indicados resultará na economia do processo de amostragem, no controle dos riscos do produtor e do consumidor com base na curva característica de operação e na curva de tamanho médio de amostras nas tomadas de decisão (Dodge, 1972; Jackisch et al., 1984).

## **2.2 Pureza genética**

A produtividade tem sido um fator fundamental na agricultura moderna, devido ao aprimoramento tecnológico dos agricultores e o uso de sementes de alta qualidade que contribuem diretamente para um processo produtivo eficiente.

conseqüentemente, empresas de sementes têm investido em programas de melhoramento genético que requerem elevados investimentos financeiros e intelectuais, resultando no lançamento de uma grande quantidade de materiais com características desejadas. Nesse sentido torna-se essencial o monitoramento de todas as fases da produção, objetivando garantir aos agricultores cultivares com as mesmas características desenvolvidas pelos melhoristas.

Um dos fatores de maior importância para a obtenção do aumento de produtividade é a utilização de materiais genéticos estáveis e adaptados, atendendo às exigências de seus consumidores. Os investimentos no desenvolvimento das cultivares devem estar associados a eficientes sistemas de multiplicação de sementes com alta qualidade genética, física, fisiológica e sanitária (Andreoli, 1992).

A obtenção e comercialização de sementes com alto grau de pureza genética envolvem um rigoroso controle dos vários fatores contaminantes da cultivar a ser produzida. Estes fatores determinam uma contaminação do tipo varietal, que é a presença de sementes da mesma espécie, porém, de outra cultivar, ou do tipo genética, decorrente de polinização não controlada. Na cultura do milho, a grande maioria das cultivares comercializada é híbrida, obtidas a partir de cruzamentos entre linhagens ou híbridos simples. Neste sistema de produção são necessários a manutenção de estoques de linhagens com altos níveis de homozigose e também um controle de qualidade rígido nas etapas de multiplicação das sementes para garantir apenas a ocorrência dos cruzamentos desejados.

O controle da polinização garante os efeitos da heterose, que são obtidos por meio de cruzamentos entre linhagens com boa capacidade combinatória, gerando híbridos com uma performance agrônômica superior (Sprague & Heberhart, 1977). Sendo assim, a pureza genética é um dos requisitos para a comercialização de sementes e a não observação dos níveis de pureza

recomendados leva à redução na produtividade e à degeneração de materiais promissores.

A principal causa da contaminação genética é a ocorrência de autofecundação da linhagem fêmea. Para que isso não ocorra, é necessária a eliminação do pólen produzido pelo parental fêmea utilizando a macho esterilidade, tratamento químico ou despendoamento manual ou mecânico (Smith & Wych, 1986, Hutchcroft, 1959).

A presença de sementes de outras cultivares em um lote resulta em efeitos negativos na produção de grãos e na uniformidade da lavoura. Portanto, é imprescindível, para o controle de qualidade na produção de sementes, saber a identidade genética e porcentagem de mistura varietal para cada classe de semente. Assim, as empresas de sementes e órgãos oficiais podem administrar estrategicamente as operações no sistema de produção, certificando-se da pureza genética das sementes para posterior comercialização.

Em um sistema de produção de sementes são adotados padrões de campo e laboratório. Estas normas estabelecem padrões que estipulam valores máximos de mistura varietal para assegurar a pureza, a identidade e a uniformidade genética da semente nas gerações subseqüentes (Andreoli, 1992).

Porém, para a produção de sementes de milho no Brasil, não existem padrões de tolerância estabelecidos para a porcentagem de contaminação proveniente do parental fêmea autofecundado

As conseqüências da autofecundação indesejável na produção de sementes híbridas de milho foram estudadas por Von Pinho (1995). Para esse estudo, observou-se que a utilização de sementes com qualidade genética comprometida em função da falta de metodologias e de padrões para a identificação destas contaminações, pode reduzir a qualidade fisiológica de sementes e a produção de grãos. A referida autora avaliou a qualidade fisiológica de lotes de sementes de híbridos simples, triplo e duplos,

contaminados com sementes híbridas provenientes do parental fêmea autofecundado nas taxas de 0%, 5%, 10% e 100%, de modo a estudar as conseqüências da autofecundação indesejável na produção de sementes de milho. Os resultados demonstraram redução na qualidade fisiológica das sementes e para cada 1% de contaminação de sementes provenientes da autofecundação misturadas nas sementes híbridas, houve também uma redução média de 0,5% na produção de grãos dos quatro híbridos estudados.

A redução no rendimento da produção poderia ser maior, caso fosse adotado um padrão de tolerância de 6% de sementes provenientes do parental fêmea autofecundado, como adotado por algumas empresas estrangeiras.

O estreitamento genético que vem ocorrendo para a maioria das cultivares em diferentes espécies tem levado à obtenção de cultivares muito próximas, dificultando a diferenciação desses materiais. Para certificação da pureza genética, têm sido utilizados marcadores morfológicos, baseados em características das sementes, plântulas, plantas em desenvolvimento e espigas. Além destes, sistemas isoenzimáticos também têm sido utilizados para o monitoramento da pureza genética das sementes (AOSA, 1991). No entanto, as técnicas moleculares fundamentadas na amplificação de fragmentos de DNA vêm tornando-se mais eficientes em discriminar cultivares, uma vez que não são influenciáveis pelo ambiente e grandes avanços têm sido alcançados nas diferentes áreas da biologia. Nesse sentido, diversos autores, a exemplo de Salgado (2001), Gethi et al. (2002), Padilha (2002), Prasad et al. (2000), Senior et al. (1998), Smith et al. (1997), avaliaram diferentes técnicas com o objetivo de certificação da pureza genética.

Controlar a pureza genética das sementes na cultura do milho é, hoje, um dos pontos cruciais para garantir uma alta produtividade, monitorando todas as fases de produção comercial de sementes híbridas e certificando-se da pureza

genética, de modo a contribuir substancialmente para a expansão da cultura no país (Koranyi, 1989).

### 2.3 Marcadores microssatélites

As técnicas de marcadores moleculares são dinâmicas e estão evoluindo rapidamente na área agronômica, apresentando aplicação direta em processos de caracterização de cultivares para obtenção de *fingerprintings* e monitoramento da pureza genética e varietal de sementes. Os marcadores moleculares são segmentos de DNA que diferenciam dois ou mais indivíduos e são herdados geneticamente. Os vários tipos de marcadores moleculares hoje disponíveis diferenciam-se pela tecnologia utilizada em indicar variabilidade em DNA, ou seja, hibridação ou amplificação de DNA.

Esses marcadores são abundantes, apresentam herança mendeliana simples, não sofrem ação do ambiente, podem ser detectados em qualquer tecido e fase de desenvolvimento da planta, não apresentam efeito pleitrópico e propiciam um número ilimitado de polimorfismo distribuído aleatoriamente ao longo de todo o genoma. (Barros, 1998; Vasconcelos, 1995; Lanza et al., 2000).

A associação desses marcadores com os descritores morfo-fisiológicos pode ser indicada no processo de proteção intelectual de genótipos elite, o que representa um forte incentivo a novos investimentos e novas oportunidades tecnológicas.

A escolha do método molecular depende da estrutura genética de cada espécie, do seu modo de reprodução, do tamanho do seu genoma, entre outros fatores, como a baixa relação custo-benefício. Todavia, não existe, ainda, uma técnica que é mais eficiente e recomendada para todo e qualquer tipo de estudo. Pelo contrário, as técnicas apresentam vantagens e desvantagens quando comparadas (Powell et al., 1996).

Dentre as técnicas moleculares que têm apresentado confiabilidade em prover uma identificação precisa por meio de um perfil genético detalhado, em menor período de tempo e custo, e passível de automatização, cita-se a técnica de microssatélite, também conhecida como seqüências simples repetidas (Simple Sequence Repeats - SSR). Esses marcadores atualmente estão sendo os mais indicados para o monitoramento da pureza genética devido às suas características de multialelismo e co-dominância (Hamada et al., 1982 Weissenbach et al., 1992).

Regiões contendo os microssatélites podem ser amplificadas individualmente, por meio de um par de iniciadores específicos que são complementares às seqüências altamente conservadas que flanqueiam os microssatélites (Litt & Lutz, 1989). Cada microssatélite constitui um loco genético altamente variável, multialélico e com um elevado conteúdo informativo de polimorfismo (Ferreira & Grattapaglia, 1998), apresentando-se estáveis ao longo de várias gerações (Russel et al., 1997). Os genomas eucariotos são densamente povoados por diferentes classes de seqüências repetidas. Os microssatélites são seqüências curtas com um a seis nucleotídeos repetidos de 10 a 60 vezes em tandem ao longo da molécula de DNA, os quais são flanqueados por seqüências conservadas (Rafalski et al., 1996).

A identificação dos SSR é um processo trabalhoso e de elevado custo, envolvendo a construção de bibliotecas genômicas, seleção de clones contendo os microssatélites, o seqüenciamento em larga escala desses clones, desenho dos *primers* específicos e sua confirmação por PCR (Rafalski et al., 1996).

Chin et al. (1996) concordam que a grande limitação da técnica é a necessidade de informações da seqüência do DNA para construção dos iniciadores. No entanto, esta limitação está sendo vencida rapidamente com disponibilidade de *primers* SSR para diversas espécies de interesse econômico,

como, por exemplo, o milho. Atualmente, já se encontram disponíveis 1.855 seqüências de pares de *primers* SSR no “*Maize Genome DataBank*”.

A implementação desta técnica torna-se ainda mais atrativa para os programas de melhoramento de milho, uma vez que o desenvolvimento dos iniciadores não é mais limitante (Lanza et al., 2000). Com o aumento da disponibilidade de *primers* de microssatélites para várias espécies, será possível obter um padrão único para cada cultivar, possibilitando com isso sua utilização em processos de identificação de cultivares (Chin et al., 1996; Senior et al., 1996, Smith et al., 1997).

Os microssatélites são importantes como suporte ao processo de proteção intelectual de materiais genéticos de milho (Smith et al., 1997). Além disso, a característica de co-dominância viabiliza a utilização da técnica de SSR como ferramenta para o monitoramento rápido e eficiente em análises de pureza genética de sementes e tem se apresentado como uma das mais promissoras, dentre os marcadores de DNA, na identificação de contaminação genética, gerando resultados de fácil análise e interpretação. Outra vantagem é o fato de possuírem vários locos com potencial para gerar um perfil alélico único, ou seja, um *fingerprinting* genético, que pode ser utilizado para o controle da pureza genética de sementes híbridas ou de linhagens, tornando-se uma técnica importante na certificação ou fiscalização de sementes. Além disso, a disponibilidade de informações das seqüências dos pares de *primers* supera as dificuldades na transferência de sondas entre laboratórios (Chin et al., 1996).

Salgado (2001), avaliando sementes milho, observou que dez locos de microssatélites são eficientes para distinguir o híbrido de seus progenitores masculino e feminino. Para a diferenciação entre o híbrido e o parental fêmea autofecundado outros 14 locos apresentaram-se eficientes. Smith et al. (1997), ao avaliarem 131 locos SSR em 13 conjuntos de parentais e seus respectivos híbridos, observaram a incidência de apenas 2,2% desses locos com herança não

mendeliana, ou seja, ausência de uma banda parental ou presença de uma banda não parental no híbrido. Os microssatélites são úteis também na detecção da heterozigosidade de um loco específico, o que é importante para o acompanhamento da pureza genética e do nível de endogamia dos materiais elites utilizados no melhoramento genético (Prasad et al., 2000; Senior et al., 1998; Smith et al., 1997). Os referidos autores citam ainda como vantagem o fato de normalmente poucos locos garantirem a completa diferenciação dos genótipos de interesse. Essa característica é um fator importante, considerando a necessidade de discriminação de cultivares muito próximas ou essencialmente derivadas.

Segundo Padilha (2002), a discriminação de 35 linhagens de milho foi possível com apenas dois locos SSR. Além disto, as diferenças na amplitude alélica entre os locos, associadas à precisão conferida pelo sistema semi-automatizado, permitem a combinação de locos em única linha do gel, garantindo a obtenção de um *fingerprinting* com alto poder de discriminação. Warburton et al. (2002) trabalharam com marcadores SSR para a caracterização genética de linhagens e populações de milho do CIMMYT e selecionaram 53 locos com alto poder discriminatório, passíveis de serem utilizados no processo de genotipagem dos materiais envolvidos no estudo. Apenas cinco locos de SSR foram necessários para a definição de *fingerprints* únicos para cada uma das 94 linhagens de milho avaliadas por Senior et al. (1998).

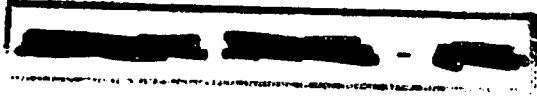
Avaliando nove linhagens de milho de diversos pedigrees foi possível identificar 200 locos de microssatélites com alto poder discriminatório dentre as 576 seqüências examinadas (Chin et al., 1996). O referido autor ainda estudou a presença e a distribuição dos microssatélites em milho com base nas informações das seqüências obtidas no GenBank e no EMBL. Dentre os locos microssatélites, 69 apresentaram polimorfismo moderado gerando 2 e 4 alelos, sendo os mais abundantes os motivos  $(AG/TC)_n$ ,  $(CCT/GGA)_n$  e  $(CCG/GGC)_n$ .

Meesang et al. (2001), utilizando 11 cultivares e 133 acessos de sementes de soja, estudaram a diversidade e a pureza genética desse material em diferentes fontes. As análises foram feitas com 19 locos de microssatélites, sendo que 15 dos 19 locos discriminaram uma alta diversidade entre 10 cultivares e com um só alelo de cada cultivar foi possível, por meio dos 19 locos de microssatélites, determinar a pureza genética do material.

Diwan & Cregan (1997), trabalhando com soja, selecionaram 20 locos SSR com motivos di ou trinucleotídeos. *Primers forward* foram marcados fluorescentemente com as cores azul, verde e amarela. As cultivares dentro dos grupos de maturidade I, II, IV e VI foram facilmente distintas utilizando-se os 20 locos SSR. Os autores acreditam que alta diversidade genética baseada nesses marcadores SSR utilizando locos tri-nucleotídeos juntamente com outros 10-15 marcadores SSR que estão sendo desenvolvidos podem permitir a identificação rápida de genótipos muito próximos de soja, podendo ser utilizados no sistema de proteção intelectual do material.

Guilford et al. (1997) verificaram que os motivos  $(GA)_n$  e  $(GT)_n$  repetidos eram abundantes em macieira, ocorrendo aproximadamente a cada 120 e 190Kb, respectivamente. *Primers* para estes locos microssatélites foram passíveis de promoverem ampliações em 21 cultivares diferentes. A maioria das marcas foi altamente polimórfica, diplóide e mostrou herança mendeliana simples, embora cerca de 25 das marcas geraram complexos padrões de bandas consistentes com a amplificação de mais do que um loco. Três microssatélites foram eficientes para diferenciar todas as 21 cultivares.

Segundo Rongwen et al. (1995), a seleção de outros locos de microssatélites com alto polimorfismo possibilitará a distinção de cultivares, mesmo com pedigree similares. Para tal afirmação, os autores avaliaram 96 cultivares de soja com sete locos de microssatélites e, destas, somente duas



cultivares apresentaram perfis de bandas idênticas, devido ao fato de possuírem pedigree similares.

Onze locos de microssatélites foram usados para examinar 24 genótipos de cevada, representados por 23 cultivares e uma linhagem. Foram encontradas três combinações separadas de quatro microssatélites que puderam distinguir entre todos os genótipos de cevada (Russel et al., 1997). Foi mostrada a estabilidade dos microssatélites ao longo de diferentes geração por meio de análise de pedigree para a cultivar Golden Promise, resolvendo conflitos de publicações anteriores com relação ao parentesco desta cultivar.

A literatura indica a grande possibilidade de utilização da técnica de microssatélite na certificação da pureza genética em sementes. Ainda assim, a adoção desse marcador molecular ainda não é uma técnica viável quando são consideradas análises rotineiras de laboratório, devido ao tamanho da amostra fixa de 400 sementes utilizado, o que gera a necessidade de períodos prolongados nas análises e elevados custos. No entanto, a exemplo do que tem acontecido em outras áreas, o crescente investimento nesta tecnologia certamente implicará adaptações que possibilitarão a automatização, de forma a permitir seu uso nos programas de produção de sementes, de forma rápida e segura. Nesse sentido, é necessário adotar uma técnica alternativa de amostragem como meio de agilização, redução da mão-de-obra utilizada e, conseqüentemente, minimizando os custos.

## 3 MATERIAL E MÉTODOS

### 3.1 Amostragem seqüencial

O trabalho foi conduzido no laboratório de Análise de Sementes da Embrapa Milho e Sorgo em Sete Lagoas, MG.

A amostragem seqüencial foi simulada em lotes de sementes de 1kg com conhecidos percentuais de contaminantes. Para a obtenção destes lotes, sementes de milho híbrido foram marcadas e misturadas às sementes sem marcação e, em seguida procedeu-se à homogeneização em um divisor centrífugo Gamet. Os níveis de contaminação de 0,25%, 0,5%, 1%, 2%, 4% e 6% foram considerados para as análises estatísticas do processo de amostragem.

Na elaboração do plano de amostragem seqüencial, a pureza genética de sementes de milho foi considerada como variável. Foram formuladas as hipóteses  $H_0$ : o lote de sementes híbridas está dentro dos padrões aceitáveis para a característica de pureza genética e  $H_1$ : o lote de sementes híbridas não está dentro dos padrões aceitáveis para a característica de pureza genética. Os riscos envolvidos nas tomadas de decisão para rejeição ou aceitação do lote de sementes foram  $\alpha$  e  $\beta$ . O primeiro se refere ao risco do produtor em rejeitar um lote de sementes que atende às exigências mínimas de pureza genética. Por outro lado,  $\beta$  se refere ao risco do consumidor em comprar um lote que não atende às exigências mínimas de pureza genética. Os valores de  $\alpha$  e  $\beta$  foram fixados em 0,05.

Foram considerados os limites para os valores de pureza genética, representados por  $p_0$  e  $p_1$ , em que  $p_0$  é o limite inferior e se refere à percentagem de pureza genética mínima tolerada para as sementes de milho híbrido simples, duplo ou triplo e  $p_1$ , limite superior, se refere à percentagem de pureza genética

ideal para as sementes híbridas de milho. Os valores calculados para  $p_0$  e  $p_1$  variam com os níveis de contaminação nos lotes de sementes (Tabela 1).

**TABELA 1** Valores de  $p_0$  e  $p_1$  para diferentes níveis de contaminação varietal em lotes de sementes híbridas de milho. Embrapa, Sete Lagoas, MG, 2003.

Níveis de contaminação (%)	$p_0$	$p_1$
0,25	0,9975	0,999
0,5	0,995	0,999
1	0,99	0,999
2	0,98	0,999
4	0,96	0,999
6	0,94	0,999

$p_0$ : percentagem de pureza genética mínima tolerada  
 $p_1$ : percentagem de pureza genética ideal

Baseou-se no tamanho máximo de amostra de 400 sementes, segundo recomendação da RAS (Brasil, 1992). A amostragem foi realizada em seqüência crescente de grupos de 40 sementes, os quais foram fixados e analisados até totalizar as 400 sementes.

Nas tabelas a seguir, estão os valores calculados para os limites inferiores ( $l_i$ ) e superiores ( $l_s$ ), adotando-se os números anteriormente mencionados. Esses limites e o tamanho médio de amostra ( $TMA$ ) foram calculados segundo as fórmulas:

$$l_i = b + an \quad e \quad l_s = c + an$$

$$a = \frac{\text{Ln} \frac{1 - P_0}{1 - P_1}}{\text{Ln} \frac{P_1 - (1 - P_0)}{P_0(1 - P_1)}} \quad b = \frac{\text{Ln} \frac{\beta}{1 - \alpha}}{\text{Ln} \frac{P_1(1 - P_0)}{P_0(1 - P_1)}} \quad c = \frac{\text{Ln} \frac{1 - \beta}{\alpha}}{\text{Ln} \frac{P_1(1 - P_0)}{P_0(1 - P_1)}}$$

Em que:

$n$ : tamanho de amostra

$l_i$ : quando  $n \leq$  ao limite inferior, decisão de rejeitar o lote

$l_s$ : quando  $n \geq$  ao limite superior, decisão de aceitar o lote

$a, b$  e  $c$ : coeficientes das duas retas, aceitação, continuação ou rejeição

$p_0$ : limite aceitável

$p_1$ : limite ideal

$\alpha = \beta = 0,05$ , ou seja, 5% de probabilidade de erro.

O tamanho médio de amostra ( $TMA$ ) é calculado em função de  $p$ , podendo ser feita a partir da expressão:

$$TMA(p) = \frac{P(p) \cdot (b - c) + c}{p - a}$$

Em que  $P(p) = \frac{\left(\frac{1 - \beta}{\alpha}\right)^W - 1}{\left(\frac{1 - \beta}{\alpha}\right)^W - \left(\frac{\beta}{1 - \alpha}\right)^W}$  é a probabilidade de aceitação  
de  $H_0$ :  $p = p_0$

Os valores de  $p$  e  $W$  estão associados pela expressão:

$$p = \frac{1 - \left(\frac{1 - p_1}{1 - p_0}\right)^W}{\left(\frac{p_1}{p_0}\right)^W - \left(\frac{1 - p_1}{1 - p_0}\right)^W}$$

TABELA 2 Valores calculados dos limites de decisão do plano de amostragem seqüencial para  $p_0$ : 0,9975;  $p_1$ : 0,999 e  $\alpha = \beta = 0,05$  para pureza genética de 0,25% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003.

N	Limite inferior ( li )	Limite superior ( ls )
40	37	40
80	77	80
120	117	120
160	157	160
200	197	200
240	236	240
280	276	280
320	316	320
360	356	360
400	396	400

TABELA 3 Valores calculados dos limites de decisão do plano de amostragem seqüencial para  $p_0$ : 0,995;  $p_1$ : 0,999 e  $\alpha = \beta = 0,05$  para pureza genética de 0,5% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003.

N	Limite inferior ( li )	Limite superior ( ls )
40	38	40
80	78	80
120	118	120
160	158	160
200	198	200
240	238	240
280	277	280
320	317	320
360	357	360
400	397	400

TABELA 4 Valores calculados dos limites de decisão do plano de amostragem seqüencial para  $p_0: 0,99$ ;  $p_1: 0,999$  e  $\alpha = \beta = 0,05$  para pureza genética de 1% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003.

N	Limite inferior ( li )	Limite superior ( ls )
40	35	40
80	75	80
120	115	120
160	155	160
200	194	200
240	234	240
280	274	280
320	313	320
360	353	360
400	393	400

TABELA 5 Valores calculados dos limites de decisão do plano de amostragem seqüencial para  $p_0: 0,98$ ;  $p_1: 0,999$  e  $\alpha = \beta = 0,05$  para pureza genética de 2% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003.

N	Limite inferior ( li )	Limite superior ( ls )
40	37	40
80	77	80
120	117	120
160	156	160
200	196	200
240	235	240
280	275	280
320	314	320
360	354	360
400	394	400

**TABELA 6** Valores calculados dos limites de decisão do plano de amostragem seqüencial para  $p_0: 0,96$ ;  $p_1: 0,999$  e  $\alpha = \beta = 0,05$  para pureza genética de 4% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003.

N	Limite inferior ( li )	Limite superior ( ls )
40	38	40
80	77	80
120	116	120
160	156	160
200	195	200
240	235	240
280	274	280
320	313	320
360	353	360
400	392	400

**TABELA 7** Valores calculados dos limites de decisão do plano de amostragem seqüencial para  $p_0: 0,94$ ;  $p_1: 0,999$  e  $\alpha = \beta = 0,05$  para pureza genética de 6% em lotes de sementes de milho. Embrapa, Sete Lagoas, MG, 2003.

N	Limite inferior ( li )	Limite superior ( ls )
40	38	40
80	77	80
120	116	120
160	155	160
200	194	200
240	233	240
280	273	280
320	312	320
360	351	360
400	390	400

### 3.2 Marcadores microssatélites

As análises moleculares foram realizadas no Núcleo de Biologia Aplicada (NBA) da Embrapa Milho e Sorgo em Sete Lagoas, MG.

A extração do DNA foi efetuada da parte aérea de plântulas oriundas de sementes germinadas em rolo de papel por cinco dias, à temperatura constante de 25°C. O material vegetal foi macerado em presença de nitrogênio líquido até a obtenção de um pó fino, ao qual foram adicionados 800 µL do tampão de extração CTAB (2% de CTAB; 1M Tris-HCl pH 8,0; 0,5mM de EDTA pH 8,0; 5M NaCl, 2% de β-Mercaptoetanol). A mistura foi mantida a 65°C por 1 hora, sendo homogeneizada a cada 15 minutos. Em seguida, foram adicionados 800µL da mistura clorofórmio-álcool octanol (24:1). As amostras foram homogeneizadas por 10 minutos e centrifugadas a 16.000xg por 10 minutos. O sobrenadante foi transferido para outro tubo e repetiu-se a etapa de lavagem com clorofórmio-octanol e de centrifugação. O DNA foi precipitado da fase sobrenadante com a adição de 800µL de isopropanol. Após a lavagem em etanol 70%, o DNA foi ressuscitado em 200µL de TE pH 8,0 (Tris-HCL 10 mM; EDTA 1 mM). A quantificação do DNA foi realizada em gel de agarose 0,8%, por meio da comparação com um DNA de concentração conhecida.

O DNA estoque foi mantido a -20°C e o DNA diluído, na concentração de trabalho de 10 ng/µL, foi armazenado a 4°C.

Em uma primeira etapa, foi realizada a seleção de conjuntos de *primers* para avaliação da qualidade genética. Foram utilizados 17 genótipos de milho, sendo cinco híbridos simples, um híbrido triplo e onze linhagens (Tabela 8), provenientes do programa de melhoramento da Embrapa Milho e Sorgo, Sete Lagoas, MG. O DNA de cada material foi obtido a partir da parte área de 10 plântulas.

TABELA 8 Híbridos de milho e seus respectivos progenitores utilizados na avaliação da qualidade genética de sementes por meio da técnica de SSR. Embrapa Milho e Sorgo, Sete Lagoas, MG, 2003.

Híbridos	Linhagens
HS1	L9 x L8
HS2	L5 x L6
HS3	L3 x L4
HS4	L7 x L2
HS5	L10 x L11
HT	HS2 x L1

HS - híbrido simples HT - híbrido triplo L - linhagem

Foram testados 100 *primers* com o objetivo de selecionar aqueles mais informativos, com melhor padrão de amplificação. Foram avaliados *primers* capazes de gerar perfis que permitissem a distinção do híbrido de seus parentais.

Um vez selecionado o conjunto de *primers*, foram efetuados misturas de DNA provenientes dos parentais, visando avaliar a sensibilidade da técnica SSR em detectar misturas varietais em lotes de sementes, em diferentes proporções. Assim, foi realizada contaminação com misturas do DNA diluído de cada um dos parentais do híbridos HS1, HS2, HS3, HS4 e HT nas proporções 1:1, 1:5, 1:8, 5:1 e 8:1.

A reação de SSR com volume final de 10  $\mu$ L, realizada em termociclador ABI 9600, foi realizada com 30 ng de DNA, 1,0 mM Tris-HCl, 50 mM KCl, 2,0mM MgCl<sub>2</sub>, 125  $\mu$ M de cada um dos dNTPs, 0,6  $\mu$ M de cada um dos *primers* e 1 U da enzima Taq polimerase. Os ciclos da reação de PCR consistiram de uma desnaturação inicial a 95°C por 2 minutos, seguida por nove ciclos de 94°C por 20 segundos, 68°C por 20 segundos, com redução de 1°C a cada ciclo e 72°C por 20 segundos; seguidos de 25 ciclos a 94°C por 20 segundos, 60°C por 20 segundos e 72°C por 20 segundos. A extensão final foi

realizada a 72°C por 10 minutos. Os fragmentos amplificados foram resolvidos em géis de agarose 4%, sob eletroforese a 100 v por 2,5 horas. O gel foi corado com brometo de etídio (1 mg/mL) e a imagem visualizada sob luz ultravioleta.

Em uma segunda etapa foi verificada a sensibilidade da técnica de SSR em detectar diferentes concentrações de DNA, numa mesma reação de PCR. Para tanto, foram utilizados o híbrido HS3 e linhagem L1. A seleção de *primers* nessa etapa foi baseada em informações já disponíveis no laboratório do NBA. O conteúdo de informação por loco dos marcadores SSR e a qualidade de amplificação dos fragmentos em gel de agarose foram considerados nesta primeira seleção. Nesta etapa foram avaliados 24 *primers*.

As misturas do híbrido HS3 e a linhagem L1, e vice-versa, obedeceram às seguintes proporções: 1:0; 1:1; 1:5; 1:10; 1:15; 1:20; 1:30 e 1:40. Para a obtenção das proporções citadas procedeu-se à: *i*) mistura de volumes de DNA cuja concentração foi determinada em gel de agarose 0,8% e *ii*) mistura de discos foliares com diâmetro de 0,8cm, sendo o DNA extraído desta mistura de discos, seguindo o protocolo do CTAB (Saghai-Marroof et al., 1984).

A avaliação da sensibilidade foi feita visualmente, levando-se em conta a resolução das bandas no gel.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Amostragem seqüencial

Observa-se, pelos resultados contidos na Tabela 9, que, num total de 400 sementes, o que corresponde ao tamanho de amostra fixo exigido pela RAS (Brasil,1992), não foi possível tomar nenhuma decisão de aceitação e/ou rejeição de lotes de sementes de milho em relação à pureza genética, quando estes lotes tinham 0,25% e 0,5% de mistura. No entanto, para lotes com mistura varietal acima de 1%, foi necessário um número menor de sementes a serem analisadas que o exigido pelo tamanho de amostra fixo  $n=400$ , para se tomar uma decisão.

TABELA 9. Número de sementes analisadas no plano de amostragem seqüencial para a decisão de aceitar ou rejeitar o lote, em função da percentagem de mistura varietal. Embrapa, Sete Lagoas, MG, 2003.

Lotes	Contaminantes (%)	Número de sementes necessárias para tomada de decisão de rejeite do lote
1	0,25	> 400
2	0,5	> 400
3	1	280
4	2	120
5	4	40
6	6	40

Os lotes que apresentavam mistura varietal igual ou superior a 0,5% precisariam de um tamanho de amostra de 1,6 a 8 vezes superior ao tamanho de amostra fixo  $n=400$  (Tabela 10), para se tomar uma decisão sobre o destino destes lotes. Por ocasião da avaliação das 400 sementes exigidas pelo plano truncado, nenhuma decisão foi tomada ao final do teste, pois sempre caía-se na região de indecisão, entre os dois limites de controle. Como optou-se pelo truncamento do teste, quando o tamanho da amostra atingiu 400 sementes, a amostragem foi finalizada e nenhuma decisão foi tomada. O resultado da média teórica (Tabela 10) indicou que, para uma mistura varietal de 0,25%, é necessário analisar, em média, 3.347 sementes para que se possa tomar uma decisão de aceite ou rejeite do lote de sementes, com um risco de 5% de erro para o produtor e consumidor. Ao analisar apenas 400 sementes somente haverá preocupação com os riscos do produtor, não considerando que o comprador das sementes também corre riscos e que estes devem também ser controlados. Da mesma forma que o produtor perde quando um lote de sementes é classificada como fora do padrão num teste, o comprador também será prejudicado se um lote de má qualidade for classificado dentro do padrão para o mesmo teste.

Já para lotes com 0,5% de mistura varietal, segundo os resultados da média teórica (Tabela 10), é necessário analisar, em média, 654 sementes para atestar a pureza genética desse lote, com o risco de 5% tanto para o produtor de sementes quanto para o consumidor. Vale ressaltar, que apesar da exigência de se analisar um número de sementes superior ao exigido pela amostragem de tamanho fixo ( $n = 400$  sementes), a amostragem sequencial garante, em um mesmo nível de segurança proteção tanto para o consumidor quanto para o produtor. Ainda pela Tabela 10, para aceitar ou rejeitar lotes com mistura de 1% (com nível de probabilidade de erro de 5% para o produtor e consumidor é necessário analisar, em média, 188 sementes. Neste experimento, foi necessário avaliar 280 sementes para se tomar a decisão de rejeite, com o mesmo risco de

probabilidade. Isso ocorreu devido à flutuação do tamanho de amostra usada no plano aleatório e pelo agrupamento das sementes de 40 em 40. Como o grupo de sementes a ser ensaiado é fixado a critério do analista, se tivesse estabelecido grupos com menor quantidade de sementes, provavelmente a decisão de rejeitar o lote seria tomada com um número de sementes total, mais próximo ao da média teórica.

Isto também ocorreu para os lotes de milho com mistura varietal de 2%,4% e 6%, ou seja, na prática houve necessidade de analisar um número superior ao estabelecido pela média teórica, para se rejeitar os lotes com as mistura acima referidas, com probabilidade de risco de 5% de erro para o consumidor e produtor (Tabela 10). Resultados semelhantes ocorreram no trabalho de Santana (1994). Ao adotar agrupamento de sementes de 20 em 20, lotes que pela média teórica precisariam de analisar em média 217 sementes, a referida autora, na prática, conseguiu rejeitar o lote com 95% de segurança, apenas quando analisou 400 sementes.

**TABELA 10** Tamanho médio de amostra (TMA(p)) para lotes de sementes de milho com diferentes porcentagem de mistura varietal, em função dos valores atribuídos a W. Embrapa, Sete Lagoas, MG, 2003.

Contaminação (%)	W	p	P(p)	TMA (p)
0,25	1	0,9975	0,9500	3.347
0,5	1	0,9950	0,9500	654
1	1	0,9900	0,9500	188
2	1	0,9800	0,9500	65
4	1	0,9600	0,9500	24
6	1	0,9400	0,9500	14

W: expediente matemático

p: porcentagem de pureza mínima tolerada

P(p): probabilidade de aceitar  $H_0$

Ressalta-se, no entanto, a eficiência da amostragem seqüencial em reduzir o tamanho da amostra para a determinação da pureza varietal em lotes com mais de 1% de mistura, sem aumentar o nível de risco de erro para o produtor e consumidor. Estes resultados reforçam outros trabalhos envolvendo o uso da amostragem seqüencial em sementes, como os obtidos por Jackish et al., (1984), Santana (1994), Bányai (1987) e Ellis & Whitehead (1987). Estes autores constataram a vantagem do uso da amostragem seqüencial em relação à amostragem de tamanho fixo, quanto à redução do número de sementes utilizadas e redução do tempo gasto para se tomar uma decisão sobre a qualidade de um determinado lote sob avaliação.

Essa diminuição significativa do número de sementes a serem analisadas é de suma importância quando se utilizam técnicas rápidas e precisas, como as moleculares, tipo microssatélites. Quando métodos eletroforéticos são considerados para avaliar pureza varietal, é inviável a utilização do plano de amostragem fixa, pois estes métodos requerem a utilização de produtos químicos que elevam o custo da análise (Bányai, 1978).

Segundo Von Pinho (1995), a utilização de sementes com qualidade genética comprometida em função da falta de metodologias e de padrões para a identificação de contaminações proveniente da autofecundação do parental fêmea, pode reduzir a qualidade fisiológica de sementes e a produção de grãos.

Para Salgado (2001), uma análise de pureza genética em sementes de milho pode ser feita utilizando tanto sementes inteiras como folhas para a extração de DNA, o que torna o teste mais rápido e de fácil execução.

No entanto, para viabilizar o uso de análise molecular em laboratorial, é fundamental adotar uma amostragem que propicie a mesma porcentagem de risco, tanto para o consumidor quanto para o produtor, possibilitando a utilização de menor número de sementes visando rapidez e menor custo. Nesse sentido, a amostragem seqüencial é uma alternativa, tendo ainda como vantagem

o aspecto que a tomada de decisão a respeito do destino do lote não é unilateral, ou seja, os riscos são os mesmos para o produtor e o consumidor.

Considerando também que quando se trabalha com as probabilidades  $\alpha$  e  $\beta$  predeterminadas de se cometer o erro tipo I e tipo II, respectivamente, nas decisões tomadas, não há nenhum outro plano de amostragem que reduza o tamanho da amostra quanto à amostragem seqüencial (Santana, 1994).

#### 4.2 Marcadores microssatélites

Em relação à seleção de conjunto de *primers*, visando discriminar os híbridos de seus respectivos parentais, entre os 100 *primers* testados, apenas cinco para cada híbrido geraram perfis de amplificação que permitiram essa discriminação, conforme consta da Tabela 11. Para os demais *primers* não foi observada herança mendeliana, ou seja, ausência da banda na linhagem parental e presença no híbrido ou vice-versa, não havendo diferenciação do híbrido com relação aos seus respectivos progenitores. Segundo Smith et al., (1997), dentre as prováveis causas estão a heterozigidade residual nas linhagens parentais, contaminação genética da progênie, mutações que podem eliminar o sítio de anelamento do *primer* ou erros durante a replicação pela taq polimerase. Na Figura 1, encontram-se exemplos de padrão de amplificação gerados por alguns dos *primers* selecionados. Salgado (2001) também conseguiu distinguir o híbrido de milho de seus respectivos progenitores por meio de marcadores SSR.

TABELA 11. Conjuntos de *primers* que permitem a distinção do híbrido de seus parentais. Embrapa, Sete Lagoas, MG, 2003.

Material	<i>Primers</i>
HS1	bnlg125, bnlgl66, bnlgl325, mmc1312, bnlg2291
HS2	bnlg2291, bnlgl006, umc1037, bnlg589, mmc041
HS3	mmc312, umc1008, phi 072, umc1055, umc1037
HS4	bnlg125, bnlgl66, bnlgl520, umc1016, bnlg669
HT	bnlg2291, bnlgl006, umc1008, umc1016, bnlg238

Pode-se observar ainda que a técnica de microssatélite apresentou-se eficiente em detectar pequenas proporções de DNA em mistura na amostra. Observa-se, pelos padrões eletroforéticos, que não existe diferença no tamanho das bandas referentes as misturas 1:1, 1:5 e 1:8 das linhagens. Dessa forma, fica evidenciada a sensibilidade da técnica em detectar mistura em qualquer uma dessas proporções (Figura 1), o que viabiliza a utilização da técnica para este fim. Isto vem ao encontro do sistema de certificação e fiscalização de sementes, pois os SSR são marcadores precisos e muito informativos em função de suas características de multialéllismo e co-dominância. Estas características permitem que os SSR sejam úteis na detecção da heterozigiosidade de um loco específico, sendo importante para o monitoramento da pureza genética (Padilha 2002, Salgado 2001, Prasad et al., 2000; Senior et al., 1998; Smith et al., 1997). Trata-se de uma técnica que é caracterizada pela simplicidade e que, após a etapa de extração do DNA, é necessária apenas a realização da PCR e eletroforese dos fragmentos.

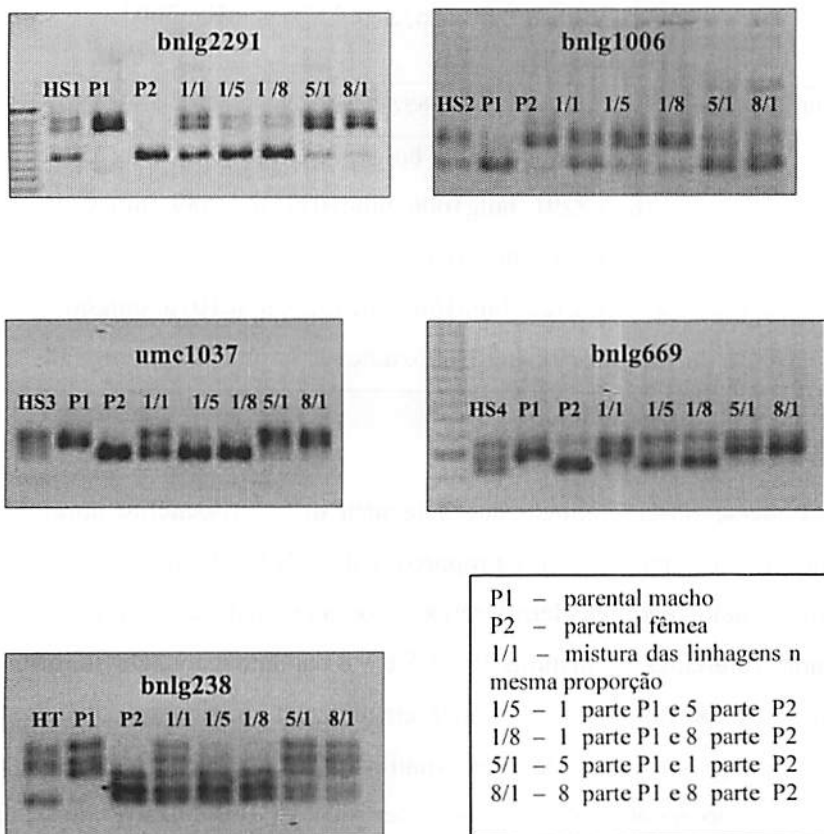
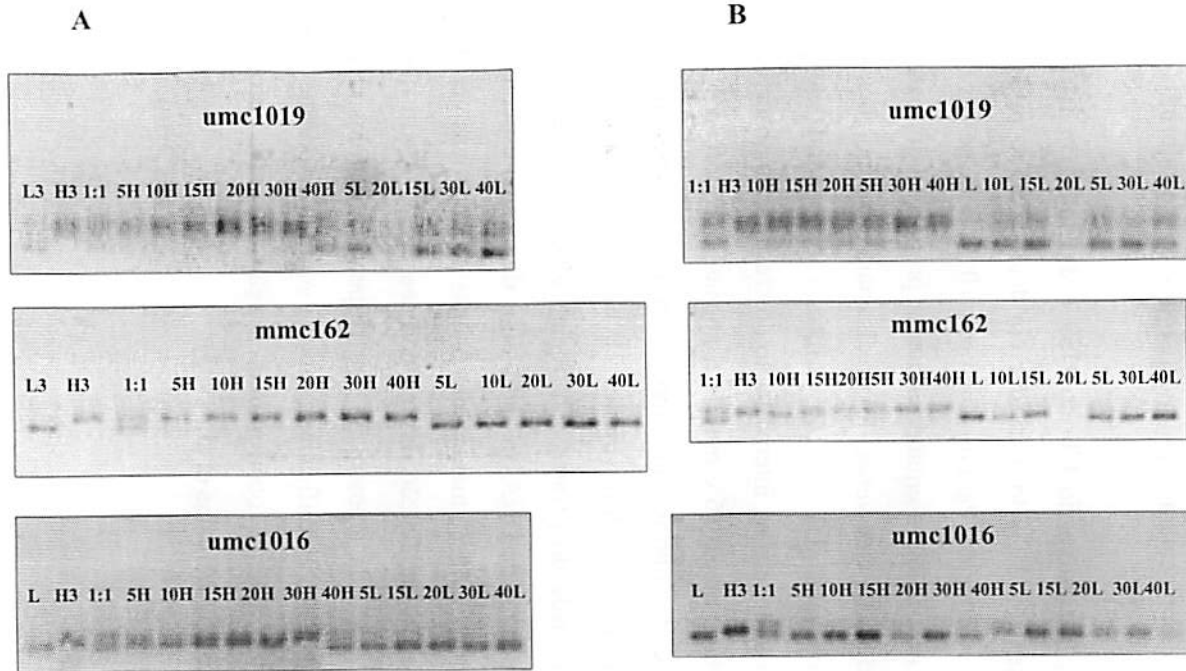


FIGURA 1 Perfil de bandas geradas pelos *primers* bnlg2291; bnlg1006; umc1037; bnlg669 e bnlg238 que permitem discriminar os híbridos de seus respectivos parentais e a detecção das diferentes proporções de misturas entre os parentais P1 e P2. Embrapa, Sete Lagoas-MG, 2003.

Na etapa de avaliação da sensibilidade dos *primers*, entre os 24 testados, três foram selecionados em função da sua capacidade de diferenciar o híbrido de seus respectivos parentais. Os demais não amplificaram alelos para um ou mais indivíduos e 6 deles não apresentaram polimorfismo entre o híbrido e a linhagem.

Quando utilizou-se a mistura do híbrido HS3 e da linhagem L1, em diferentes proporções, a sensibilidade do marcador em detectar essas misturas variou em função do *primer* utilizado. Para o *primer* umc1019, pôde-se verificar que a resolução da banda no gel, representando a menor proporção do material na mistura, da linhagem no HS3 e vice-versa, foi precisa para as proporções 1:1; 1:5; 1:10 e 1:20. Nas proporções 1:30 e 1:40, a resolução da banda no gel diminuiu e houve comprometimento na interpretação dos resultados observados (Figura 2). Já para os *primers* umc1016 e mmc162, apenas a proporção 1:1 de mistura da linhagem L1 e do híbrido HS3 apresentou banda nítida no gel, possibilitando detectar a referida mistura (Figura 2). Este fato provavelmente ocorreu em função da afinidade do *primer* pela região alvo. Ainda há de se considerar que quando amostras de DNA de indivíduos da mesma espécie são misturadas em diferentes proporções, a maioria das bandas polimórficas geradas são resultado da amplificação em função das proporções de DNA em que o fragmento se encontrava em cada amostra no seu respectivo genoma. Estas bandas, teoricamente, devem ser amplificadas proporcionalmente, porque elas estão presentes igualmente, podendo ocorrer uma ligação do *primer* com a região alvo do DNA genômico (Williams et al., 1993)



L – linhagem; HS3 - híbrido simples; 1:1 - 1 parte da L e 1 parte HS3; 5H - 5 parte HS3 e 1 parte L; 10H - 10 parte HS3 e 1 parte L; 15H - 15 parte HS3 e 1 parte L; 20H - 20 parte HS3 e 1 parte L; 30H - 30 parte HS3 e 1 parte L; 40H - 40 parte HS3 e 1 parte L; 5L - 5 parte L e 1 parte HS3; 10L - 10 parte L e 1 parte HS3; 15L - 15 parte L e 1 parte HS3; 20L - 20 parte L e 1 parte HS3; 30L - 30 parte L e 1 parte HS3; 40L - 40 parte L e 1 parte HS3.

FIGURA 2 Perfil de bandas geradas pelos marcadores umc1019; mmc162 e umc1016 da mistura de DNA (A) ou de discos foliares (B) do híbrido HS3 e da linhagem L1 em diferentes proporções. Embrapa, Sete Lagoas, MG, 2003.

Uma forma possível de aumentar a sensibilidade da técnica para detectar misturas quando um dos materiais encontra-se em menor proporção seria a NESTED-PCR, sugerida por Fungaro (2000). Segundo a autora, esta técnica, que consiste numa reação de PCR em duas etapas, pode aumentar a sensibilidade em até 1000 vezes. Ainda foi observado (Figura 2) que os padrões de bandas foram mais informativos e nítidos ao utilizar-se mistura de discos de folhas. Isso sugere que a técnica pode ser utilizada para detectar mistura de *bulk* de sementes para extração de DNA, o que facilita e diminui o custo das análises para determinação da pureza genética em lotes de sementes de milho.

Outro aspecto importante a ser considerado é o custo da análise, que gira em torno de R\$ 2,80 por reação. Isso levando-se em conta que foi necessária a utilização de apenas um *primer* de microssatélite para diferenciar os materiais considerados de seus respectivos parentais.

Por outro lado, se for necessário determinar o percentual de mistura para atender a uma possível exigência do sistema de certificação e fiscalização de sementes em relação à pureza genética de um determinado lote, o custo aumentaria em função do percentual de contaminação permitido e do tamanho da amostra (número de sementes) necessário para detectar um referido percentual. Dessa forma, se for considerada uma percentagem de 0,25% ou 0,5% de contaminação, segundo os resultados desta pesquisa, para atestar-se a pureza genética do lote via marcador molecular, seria necessário analisar 3.347 ou 654 sementes, respectivamente, o que resultaria num custo de R\$ 9.370,00 ou 1.830,00 respectivamente.

De acordo com Von Pinho (1995), lotes com 1% de contaminação têm a qualidade fisiológica reduzida, bem como a produção de grãos. Dessa forma, simulando-se um percentual de mistura permitido de 1%, para se analisar um lote de sementes por meio da técnica de microssatélite, adotando o sistema de amostra de tamanho fixo (400 sementes) que é exigido pela RAS (Brasil, 1992),

o custo por amostra (lote) é de R\$ 1.120,00. Segundo resultados da presente pesquisa, o tamanho de amostra necessário quando se utiliza amostragem seqüencial para se detectar o mesmo 1% de mistura, é de 280 sementes, o que propiciaria uma redução de 30% no custo por amostra, ou seja, ficaria em R\$ 340,00, utilizando-se a mesma técnica com a mesma percentagem de segurança (5% de risco para o consumidor e o produtor). Para se determinar níveis mais elevados de contaminação, como 2%, 4% e 6% , o uso da amostragem seqüencial reduziria ainda mais o custo da análise em relação ao uso da amostragem de tamanho fixo. Isto porque seria necessário analisar 120, 40 e 40 sementes por lote, respectivamente, o que ocasionaria uma redução de 70% e 90%, respectivamente, no custo por amostra.

## 5 CONCLUSÕES

A amostragem seqüencial possibilita atestar a pureza genética de lotes de sementes, com riscos iguais para o produtor e o consumidor

A amostragem seqüencial possibilita a redução do tamanho de amostra em relação ao tamanho de amostra fixa, para lotes de sementes acima de 1% de mistura varietal.

Amostragem seqüencial viabiliza o uso da técnica de microsatélite para atestar a qualidade genética de lotes de sementes de milho.

A adoção da amostragem seqüencial reduz os custos na determinação da qualidade genética por meio da técnica de microsatélite.

A técnica de microsatélite é eficiente para discriminar híbridos dos seus respectivos parentais, bem como em detectar misturas de proporções de DNA até um nível de 1:8. A detecção de mistura de até 1:40 variou em função do *primer*.

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

ANDREOLI, C. Mistura varietal: aspectos genéticos e físicos na produção de sementes. **Informativo ABRATES**, Brasília, v.1, n.3, p.32-36,1992.

ASSOCIATION OF OFICIAL SEED ANALYSTS. **Cultivar purity testing**. Lansing, 1991. 371p.

BÁNYAI, J. Sequenzanalyse in der saatgutprufung. **Seed Science & Technology**, New Delhi, v.6, n.2, p.505-515, 1978.

BÁNYAI, J. Report of the statistics committee 1983-1986. **Seed Science & Technology**, New Delhi, v.15, n.2, p.499-505, 1987.

BARROS, E.G. **Métodos bioquímicos e moleculares aplicados à identificação e avaliação da pureza genética de plantas**. Sete Lagoas, MG: EMBRAPA/CNPMS. 1998. (Palestra do Curso de Qualidade e Análise de Sementes).

BRASIL. Ministério da Agricultura. **Regras para análise de sementes**. Brasília: LANARV/SNAD/MA, 1992.360p.

CHIN, E.C.L. et al. Maise simple repetitive DNA sequences abundance and allelic variation. **Genome**, Ottawa, v.39, p.866-873,1996.

DIWAN, N.; GREGAN, P.B. Automated sizing of fluorescent-labelled simple sequence repeats (SSR) markers to assay genetic variation in soybean. **Theoretical and Applied Genetics**, Berlim, v.95, p.723-733, 1997.

DODGE, Y. Application of sequential analysis to tests for noxious weed seeds. **Proceedings of the Association, of Official Seed Analysts of North America**, New Brunswick, v.62, p.332-329, June 1972.

ELLIS, R.H.; WHITEHEAD, J. Open, truncated and triangular sequential seed testing procedures. **Seed Science & Technology**, New Delhi, v.15, n.1, p.1-17, 1987.

ESTEFANEL, V. **A amostragem sequencial baseada no teste sequencial da razão de probabilidades e seu uso no controle das lagartas da soja no estado**

**do Rio Grande do Sul.** 1977. 117p. Dissertação (Mestrado em Experimentação e Estatística)-Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

**FERREIRA, M. E.; GRATAPAGLIA, D. Introdução ao uso de marcadores moleculares em análise genética.** Brasília: EMBRAPA/CENARGEM, 1998. 220p.

**FUNGARO, M.H.P. PCR na micologia diagnóstico e análise de variabilidade. Biotecnologia Ciência e Desenvolvimento,** 2002.

**GUILFORD, P. E. et al. Microsatellites in *Malus x domestica* (apple): abundance, polymorphism and cultivar identification. Theoretical and Applied Genetics,** Berlin, v. 94, n. 2, p. 249-254, Feb. 1997.

**HAMADA, H., PETRINO, M. G. & KAKUNAGA, T. 1982 A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA 79: 6465-6469.**

**HUTCHCROFT, C.D. Contamination in seed fields of corn resulting from incomplete detasseling. Agronomy Journal,** Madison, v.51, p.267-271,1959.

**INTERNACIONAL SEED TESTING ASSOCIATION. Rules for seed testing.** Switzerland, 1999. 44p.

**JACKISH, W.; LINDNER, H.; SIEG, M. Sequentielle verfahren bei der keimfähigkeits-bestimmung von getreide unter anwendung der saurefuchsin-methode. Seed Science & Technology,** New Delhi, v.12, n.2, p.697-706, 1984.

**KORANYI, P. Chacaracterisation of maize (*Zea mays L.*) seed samples by the eletrophoretic patterns of their protein monomers. Seed Science & Tecnology,** Zurichi, v.10, p.153-159, 1989.

**LANZA, M.B.; GUIMARÃES, C.T.; SCHUSTER, I. Aplicação de marcadores moleculares no melhoramento genético. Informe Agropecuário,** Belo Horizonte, v.21,n.204, p.97-108, 2000.

**LITT, M.; LUTZ, J. A. A hypevariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetic,** v.44, p.397-401, 1989.

**MORRIS, R.F. A sequencial sampling tecinique for spruce budworm egg surveys. Canadian Journal Zool.,** Ottawa, v.32, p.302-313, 1954.

OAKLAND, G.B. Na application of sequential analysis to whitefish sampling. **Biometrics**, Alexandria, v.6, n.1, p.56-67, 1950.

PADILHA, L. **Marcadores moleculares semi-automatizados na caracterização e determinação da diversidade genética entre linhagens de milho tropical**. 2002. 85p. Tese(Doutorado em Fitotecnia)-Universidade Federal de Lavras, Lavras.

PATERNIANI, E.; CAMPOS, M. S.; Melhoria do milho. In: BORÉM, A. **Melhoria de espécies cultivadas**. Viçosa: UFV, 1999. p. 429-485.

PIETERS, E.P. Bibliography of sequential sampling plans for insects. **Bull. Entomol. Soc. Am.**, Lanham, v.30, p.35-7, 1978.

POWELL, W. et al. Genepool variation in genus *Glycine* Subgenus soja revealed by polymorphic nuclear and chloroplast microsatellites. **Genetics**, Princeton, v.144, p.793-803, 1996.

PRASAD, M. et al. The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. **Theoretical and Applied Genetics**, Berlin, v. 100, n. 3/4, p. 584-592, Feb. 2000.

RAFALSKI, D. J. A. et al Generating and using DNA markers in plants. In: BIRREN, B.; LAI, E. **Nonmammalian genomic analysis: a practical guide**. New York, 1996. p. 75-134.

RONGWEN, J. et al. The use of microsatellite DNA markers for soybean genotype identification. **Theoretical and Applied Genetics**, Berlin, v.90, p.43-48, 1995.

RUSSEL, J. et al Discriminating between barley genotypes using microsatellite markers. **Genome**, Ottawa, v.40, p.442-450, 1997.

SAGHAI-MAROOF, M. A. et al. Ribosomal DNA spacer length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. **Proceeding of National Academic Science of the United States of America Biological Science**, New York, v.81, n.24, p.8014-8018, 1984.

SALGADO, K.C.C. **Certificação da pureza genética em sementes híbridas de milho por meio de marcadores morfológicos e moleculares**. 2001. 67p. Dissertação (Mestrado em Fitotecnia)-Universidade Federal de Lavras, Lavras.

**SANTANA, D.G. de Adaptação do teste do pH do exsudado e viabilidade do uso da amostragem sequencial na rápida definição sobre o destino de lotes de sementes de milho.** 1996. 79p. Dissertação (Mestrado em Fitotecnia)-Universidade Federal de Lavras, Lavras.

**SENIOR, M.L. et al. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system.** *Crop Science*, Madison, v. 38, n. 4 p. 1088-1098, July/Aug. 1998.

**SMITH, J. S. C. ; WICH, R. D. The identification of female selfs in hybrid maize: a comparison using electrophoresis and morphology.** *Seed Science & Technology*, Zurich, v.14, p.1-8, 1986.

**SMITH, J.S.C. et al. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays L.*): comparisons with data from RFLPS and pedigree.** *Theoretical and Applied Genetics*, Berlim, v.95, p.163-173, 1997.

**SPRAGUE, G.F.; EBERHART, S.A. Corn Breeding In: SPRAGUE, G.F Corn and corn improvement.** American Society of Agronomy, Madison, p.305-362, 1977.

**VASCONCELOS, M. J. V. Avaliação da variabilidade genética de cultivares de feijão pelo uso de marcadores moleculares.** Viçosa: UFV, 1995. 54p. (Tese-Mestrado).

**VON PINHO, E.V.R. Conseqüências da autofecundação indesejável na produção de sementes híbridas de milho.** 1995. 130p. Tese(Doutorado em Fitotecnia)-Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

**WALD, A. Sequential analysis.** New York: John Wiley, 1947. 21p.

**WALD, A. Sequential analysis.** New York: John Wiley, 1945. 23p.

**WIEL, C. et al. Microsatellite fingerprinting in lettuce (*Lactuca sativa L.*) and wild relatives.** *Plant Cell Reports*, Berlin, v. 17, n. 11, p. 837-842, Nov. 1998.

**WILLIAMS, J. G.K.; RAFALKI, J.A. & TINGEY, S.V. 1993 Genetic analysis using RAPD markers.** *Methods Enzymol.* 218: 704-740.

