

# DESENVOLVIMENTO E ANÁLISE GENÉTICA DE UM BANCO DE DADOS DE PERFIS GENÉTICOS DE MARCADORES MICROSSATÉLITES DAS VARIEDADES DE SOJA PROTEGIDAS NO BRASIL

## DEVELOPMENT AND GENETIC ANALYSIS OF A DATABASE OF MULTILOCUS MICROSATELLITE DNA FINGERPRINTS OF THE BRAZILIAN PROTECTED SOYBEAN VARIETIES

Grattapaglia, D.; Soares, C.; Ferreira, M.E.

GENOMAX™ Tecnologia Genômica, SHIN Centro de Atividades 02 Lote 19, Lago Norte, 70505-000 -Brasília, DF.

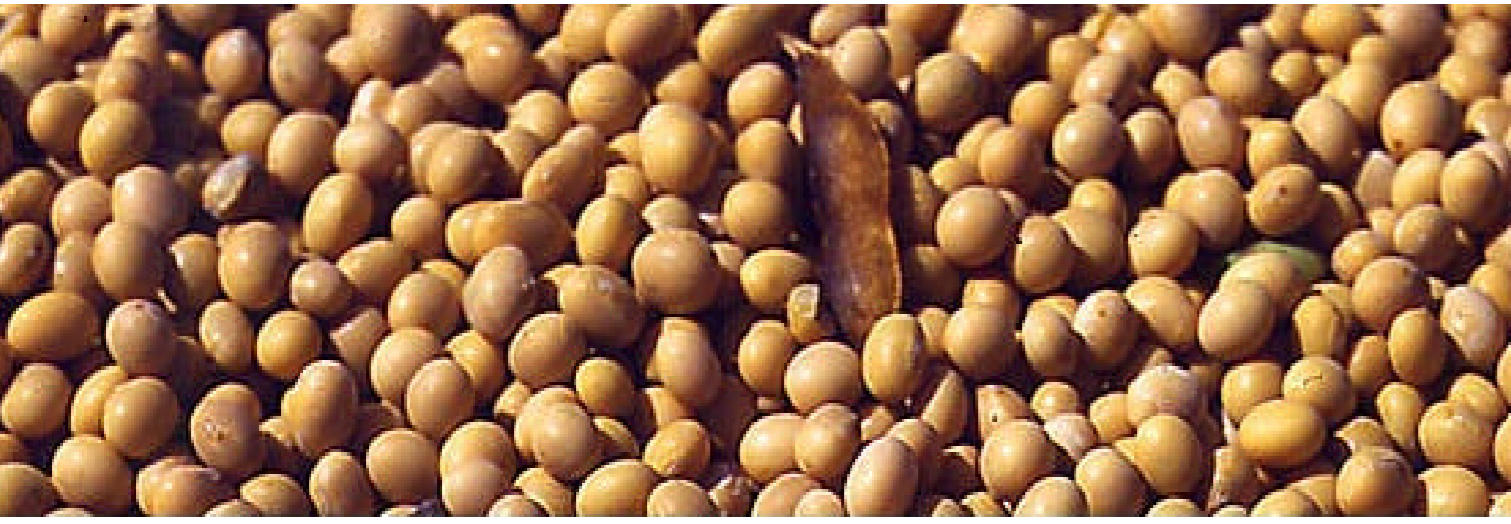


### OBJETIVOS

- Otimização e implementação em rotina de sistemas multiplex com detecção fluorescente para a análise de polimorfismos de 15 locos microsatélites autossômicos em soja;
- Desenvolvimento de um banco de dados dos perfis genéticos multiloco de 15 marcadores microsatélites das 125 cultivares registradas e protegidas junto ao Serviço Nacional de Proteção de Cultivares do MAPA.
- Avaliação do poder de resolução da tecnologia de marcadores microsatélites, para testes de DHE (distingibilidade, homogeneidade e estabilidade) em um conjunto extenso de variedades protegidas de soja de estreita base genética.

### RESUMO

A análise da variabilidade genética na sequência do DNA pela tecnologia de marcadores moleculares permite uma elevada precisão na identificação e discriminação varietal. Perfis genéticos de marcadores moleculares poderão, em breve, tornar-se importantes descritores complementares no requerimento de proteção de novo cultivar, particularmente em espécies autógamas de base genética estreita. Amostras oficiais de sementes de cada cultivar de soja foram fornecidas para análise em sacos lacrados e identificados pelo LADIC/SNPC/MAPA. A análise genética foi realizada em sistema de prova e contraprova em experimentos "cegos", onde a relação de vínculo genético e/ou comercial entre as amostras não era conhecida. Análise de sementes individuais da variedade foi realizada em casos onde havia indícios de contaminação ou mistura de sementes. Para a determinação do perfil genético dos cultivares foi realizada inicialmente uma triagem de 21 locos microsatélites internacionalmente recomendados para soja, dos quais foram selecionados 15 por apresentarem maior conteúdo informativo e robustez analítica. Os 15 locos selecionados, sendo 12 de trinucleotídeos e 3 dinucleotídeos, foram genotipados em quatro sistemas "multiplexes" de análise simultânea de vários locos em um sequenciador automático de DNA ABI Prism 377XL, e os genótipos descritos com os alelos em pares de bases. As diferenças genéticas entre os cultivares foram claramente reveladas pelas diferenças no comprimento das sequências de DNA amplificadas, ou seja dos alelos. Das 125 amostras analisadas, 123 foram declaradas geneticamente distintas e apenas duas foram geneticamente indistinguíveis com os 15 marcadores utilizados. Outros três pares de amostras apresentaram apenas uma diferença que requer uma confirmação com análises de locos adicionais para verificar se as amostras que compõem cada par são de fato diferentes, ou se a diferença observada é devida a uma rara mutação genética na sequência de DNA no marcador testado. Neste sentido, recomenda-se adotar uma abordagem conservadora e exigir no mínimo duas diferenças genéticas para declarar a distigibilidade com elevada confiabilidade. Três amostras apresentaram fortes indícios de mistura de linhas e uma amostra apresentou evidências de heterozigosidade residual. Em ambos os casos uma análise complementar com maior número de sementes individuais deverá ser realizada para confirmar definitivamente estas hipóteses. Foi observada uma elevada diversidade alélica entre os 125 cultivares, embora em alguns locos determinados alelos apresentem frequências elevadas. Em média, os locos utilizados apresentaram 5,9 alelos com uma variação de três a nove. Apenas três dos quinze locos apresentaram uma diversidade gênica abaixo de 0,5 enquanto que quatro locos mostraram uma diversidade acima de 0,7. Tanto a faixa de tamanho de alelos bem como o poder de discriminação observado para estes 15 locos no germoplasma brasileiro estão de acordo com o que tem sido descrito para o germoplasma norte-americano. É interessante observar, entretanto, que no conjunto de 125 cultivares, foram observados diversos alelos não reportados anteriormente o que confirma o fato de material genético exclusivo do Brasil estar presente no pool gênico protegido. Em conclusão, este trabalho estabelece de forma pioneira um sistema de identificação genética de cultivares de soja baseado em uma tecnologia rápida, econômica e, mais importante, de alto poder de resolução, que poderá se constituir em uma ferramenta inovadora para testes de DHE no processo de proteção varietal. O banco de dados de perfis genéticos dos cultivares de soja construído poderá ser imediatamente utilizado para o estabelecimento de um sistema eletrônico de comparação na condução de testes de identidade genética entre amostras questionadas e amostras padrão no controle de qualidade e fiscalização na cadeia de comercialização de sementes.



### ABSTRACT

The analysis of genetic variation at the DNA sequence level is a powerful tool to resolve questions of varietal identification and discrimination in crop plants. Genetic profiles of highly polymorphic molecular markers are becoming an important complementary descriptor when requiring the protection of a new variety, particularly in selfing species of narrow genetic base. With this perspective in mind we developed a database of molecular profiles of the 125 soybean varieties currently registered and protected at the National Service of Varietal Protection of the Brazilian Ministry of Agriculture (LADIC/SNPC/MAPA). The objective of this work was to evaluate the performance of microsatellite markers for tests of DHE (distinctiveness, homogeneity and stability) in an extensive but closed set of protected soybean varieties of narrow genetic base. Official seed samples from each variety were supplied in sealed numbered bags. The genetic analysis was carried out in double blind experiments and sample identity was not informed. Fifteen microsatellite markers, selected based on map position, robustness and genetic information content, were typed by multicolor fluorescence detection in four optimized multiplex systems in a DNA sequencer. Of the 125 samples analyzed, 120 could be unequivocally distinguished based on at least two but typically more than four genotype differences. Two samples were indistinguishable, i.e. displayed the same exact multilocus genotype, and three pairs of samples were different by only one marker locus, requiring additional markers to test the hypothesis of mutation or to confirm the distinctiveness with further genetic differences. Three samples showed evidences of mixture of lines and one sample had evidences of residual heterozygosity. On average, the 15 loci had 5.9 alleles and only three out of 15 loci had a genetic diversity (GD) of less than 0.5 while the more informative loci had GD above 0.7. Both the allelic size range and the power of discrimination of these 15 loci are in close agreement with previous estimates from studies on US germplasm. Not surprisingly, however, a number of alleles previously not reported in the US germplasm were detected in the set of 125 samples studied, suggesting the existence of novel germplasm. This work establishes a rapid and economical system for the genetic identification of all currently registered Brazilian soybean varieties. This system could add a significant power of resolution for DHE tests especially when closely related varieties are under scrutiny. The database of genetic profiles and allele frequencies could be immediately used to implement genetic identity tests by electronic comparison of multilocus profiles between questioned and reference samples in QA/QC procedures along the seed commercialization chain.

### INTRODUÇÃO

Marcadores de DNA são extremamente potentes para a resolução de questões de identidade e discriminação e por isso atendem com perfeição os requerimentos do teste de DHE – Distingibilidade, Homogeneidade e Estabilidade. O teste DHE é o procedimento técnico de comprovação de que a nova cultivar ou a cultivar essencialmente derivada são distinguíveis de outra cujos descritores sejam conhecidos, homogêneas quanto às suas características em cada ciclo reprodutivo e estáveis quanto à repetição das mesmas características ao longo de gerações sucessivas. Isto se deve à natureza analítica dos marcadores que se baseia na detecção direta da variabilidade existente na sequência de DNA. Fazendo um paralelo, é possível verificar que os marcadores apresentam características convergentes com os requerimentos legais para a proteção de um cultivar. Os marcadores se caracterizam pela:

- Unicidade**: cada indivíduo tem um perfil de DNA único com exceção de clones de uma planta ou gêmeos univitelinos no caso de animais;
- Mutabilidade**: ou melhor, mutabilidade extremamente baixa (da ordem de 10-4 a 10-6/geração); o DNA de um indivíduo é o mesmo em todas as células e este é estável ao longo de sua vida;
- Estabilidade física**: moléculas de DNA podem ser recuperadas, purificadas e analisadas com sucesso a partir de praticamente todo e qualquer material biológico da planta, independentemente do tecido, idade, condição fisiológica ou mesmo em degradação;
- Variabilidade**: apesar do fato de que a maior parte do genoma é muito conservada entre indivíduos, algumas regiões genômicas específicas são altamente variáveis permitindo uma discriminação altamente precisa, quantificável e reproduzível com alta precisão entre laboratórios distintos.

Diversas tecnologias para a detecção de marcadores moleculares têm sido desenvolvidas nos últimos anos. Para a proteção de cultivares, além dos aspectos de co-dominância, estabilidade genética, baixa mutabilidade e robustez analítica, é absolutamente crucial que o marcador molecular utilizado permita determinar com precisão a identidade dos alelos e com isso fazer comparações acuradas de resultados em diferentes momentos e entre diferentes laboratórios. Marcadores RAPD e AFLP, embora altamente polimórficos, apresentam sérias limitações de robustez violando uma das premissas básicas de estabilidade do descritor. Além disso, estes marcadores apresentam herança dominante o que complica sobremaneira a análise de transmissão de marcadores entre gerações para estudos, por exemplo, de cultivares essencialmente derivados.

Os descritores moleculares internacionalmente recomendados para a identificação individual de plantas, animais e seres humanos são baseados em segmentos curtos de DNA de 1 a 7 pares de bases repetidos em tandem conhecidos como microsatélites. Para a resolução de diferenças de 2 pares de bases entre estes produtos, que correspondem aos alelos, a análise é realizada em géis de eletroforese de alta resolução com detecção automatizada de fluorescências em sequenciador automático de DNA. Os alelos são identificados pelo seu tamanho em pares de bases de DNA. Os marcadores microsatélites são altamente polimórficos, ricos em formas alélicas, possibilitando uma análise genética acurada e um elevado poder de resolução para testes de DHE bem como em testes de parentesco e identidade genética mesmo em situações de forte parentesco entre clones. Adicionalmente, marcadores microsatélites são transferíveis de genoma para genoma dentro de uma espécie e entre espécies geneticamente relacionadas.

A análise de marcadores microsatélites é realizada identificando-se as formas alélicas que um determinada planta possui em locos definidos dos cromossomos, comparando-as com formas alélicas conhecidas para a espécie, permitindo, portanto, estimativas precisas e estatisticamente formais de vínculo genético e/ou identidade, bem como declarações categóricas de não-identidade ou parentesco. A técnica de análise de microsatélites é utilizada rotineiramente com sucesso na geração de estimativas robustas de individualidade genética para inúmeras espécies vegetais, animais e em humanos, constituindo hoje o padrão internacional em investigação genética forense.

### MATERIAL E MÉTODOS

**Processamento das amostras.** Amostras de sementes de cada cultivar de soja foram fornecidas para análise em sacos lacrados e devidamente identificados pelo LADIC/SNPC/MA. As amostras foram identificadas originalmente pelo LADIC/SNPC/MA e esta numeração foi seguida durante toda a análise. A análise genética foi realizada em experimentos "cegos", onde a relação de vínculo genético e/ou comercial entre as amostras não era conhecida, o que evitou qualquer tendenciosidade na declaração de identidade e/ou diferença genética entre as amostras. Para cada cultivar foram utilizadas inicialmente duas sub-amostras de trabalho de 50 sementes para a extração do DNA genômico total, sendo uma para a prova e outra para a contra-prova. Análise de sementes individuais da variedade foi realizada em casos onde havia indícios de contaminação ou mistura de sementes. A amostra de sementes foi macerada integralmente (pulverizada) para a obtenção de uma amostra homogênea de farelo. Para as amostras de 50 sementes maceradas, uma amostra de 250 miligramas de farelo foi utilizada no procedimento de extração de DNA. Para as sementes individuais a semente inteira foi utilizada no processo de extração.

**Análise de amostra de prova e contra-prova.** Foi utilizado um sistema de prova e contra-prova no processamento e análise das amostras. A primeira sub-amostra analisada foi utilizada para a prova. A segunda sub-amostra foi utilizada para a contra-prova e as duas armazenadas em locais distintos uma da outra. A prova foi processada independentemente da contra-prova em dias diferentes e por equipes técnicas distintas para a confirmação do resultado.

**Extração e quantificação de DNA.** Foi utilizado o procedimento de extração e quantificação de DNA genômico de tecidos vegetais descrito por Ferreira & Grattapaglia (1998) modificado com a adição de uma etapa de purificação do DNA em coluna de troca iônica.

**Locos microsatélites.** Foi realizada uma triagem de um total de 21 locos microsatélites, dos quais foram selecionados 15 por apresentarem maior conteúdo informativo para discriminação genética individual, independência, melhor robustez e repetibilidade (Tabela 1).

**PCR de microsatélites.** Os iniciadores dos locos microsatélites foram marcados com fluorescência azul (FAM), verde (HEX), ou amarela (NED) para detecção no filtro virtual D. As reações de PCR para cada loco individualmente foram realizadas com 10 a 50 ng de DNA genômico de soja, em um volume total de 20 µl.

**Determinação de genótipos.** Os produtos amplificados foram carregados em gel de poliacrilamida e separados via eletroforese em um sequenciador ABI Prism 377XL e analisados com os softwares Genescan Analysis v. 2.1 e Genotyper 2.0.

**Interpretação Genética.** Os genótipos foram descritos com os alelos identificados em pares de bases, respeitando a natureza do microsatélite seja ele tri ou dinucleotídeo. Por ser a soja uma espécie autógama, o genótipo observado em cada loco é tipicamente homozigoto. Ocasionalmente foram observados dois alelos em um loco, o que seria evidência de heterozigosidade residual em caso de semente individual ou mistura de linhas em caso de amostras de bulks de sementes. Para controle experimental, sempre foi utilizada a amostra 119 em todos os géis de eletroforese visando a comparação de estimativas de tamanhos de alelos em pares de bases entre experimentos distintos.

**Análises estatísticas.** Os dados experimentais gerados foram utilizados para definir o alelo presente em cada loco em cada uma das 125 amostras, levando em consideração a natureza molecular do loco microsatélite. Os dados experimentais foram submetidos a uma análise de distância genética utilizando-se a distância Euclidiana definida pela soma de quadrados das diferenças entre os alelos em pares de bases a todos os locos. Esta estimativa de distância é a mais adequada para este tipo de dado pois ela leva em conta não apenas o compartilhamento ou não do mesmo alelo ao loco, mas também a magnitude da diferença quantitativa entre os alelos em pares de bases o que é um reflexo da divergência genética como resultado do processo recorrente de mutação ("step-wise mutation model"). A matriz de distâncias Euclidianas foi em seguida submetida a uma análise de agrupamento utilizando o algoritmo UPGMA e os dados plotados em um dendrograma para melhor visualização das distâncias genéticas. Além da análise de distância, os dados genotípicos foram utilizados para a estimativa do número de alelos e das frequências dos alelos observados neste conjunto de 125 amostras representativas do material depositado no SNPC. Sendo a soja uma espécie autógama, parte-se da premissa de que o coeficiente de endocruzamento é F=1 e, portanto, a frequência de ocorrência de um genótipo homozigoto (p<sup>2</sup>) é igual à própria frequência alélica p. Os dados de frequências alélicas foram utilizados para estimar a Diversidade Gênica de cada loco o que representa um indicador do poder de discriminação genética do loco em testes de identidade e é o parâmetro adequado para a avaliação de diversidade em plantas autógamas.

### RESULTADOS E DISCUSSÃO

São apresentados a seguir os resultados das análises genéticas das 125 amostras de cultivares de soja. A grande maioria dos locos selecionados (12 em 15) é constituída de trinucleotídeos o que, de fato, permite uma melhor resolução analítica dos alelos. As diferenças genéticas entre as amostras são claramente reveladas pelas diferenças no comprimento das sequências amplificadas, ou seja dos alelos, identificados como picos nos eletroferogramas (Figura 1). Nos exemplos apresentados, podem ser facilmente verificadas as diferenças de três pares de bases dos tamanhos dos alelos entre as amostras. Por exemplo, no painel A, para o marcador Satt070 a amostra 24 possui genótipo 172/172 enquanto que a amostra 27 é homozigota para o alelo 193, sete unidades repetitivas de três pares de bases totalizando 21 pares de bases maior.

Na Tabela 2 são apresentados os dados que compõem os perfis genéticos de 30 das 125 amostras analisadas. Observa-se uma significativa diversidade alélica entre amostras, embora em alguns locos determinados alelos apresentem frequências elevadas. Este banco de dados constitui a referência para uma possível comparação de novas cultivares submetidas à registro bem como para trabalhos de controle e certificação genética de origem de lotes de sementes.

Os 15 locos utilizados na análise apresentaram em média um elevado poder de discriminação. Em média, os locos utilizados apresentaram 5,9 alelos com uma variação de três a nove. Apenas 3 dos 15 locos apresentaram uma diversidade gênica abaixo de 0,5 enquanto que 4 locos mostraram uma diversidade acima de 0,7 para o conjunto de 125 amostras analisadas (Tabela 3). Tanto a faixa de tamanho de alelos em pares de bases, como o poder de discriminação observado para estes 15 locos no germoplasma brasileiro estão de acordo com o que foi publicado por Diwan e Cregan (TAG 95:723-733:1997) em uma análise de 35 genótipos representativos da variabilidade alélica do germoplasma norte-americano. É interessante observar, entretanto, que no conjunto de 125 amostras foram observados diversos alelos não reportados naquele trabalho o que confirma o fato de que material genético exclusivo do Brasil está presente no pool gênico estudado neste trabalho.

As estimativas de poder informativo dependem do universo de cultivares amostradas e tendem a ser maiores quanto maior o número e a origem diversificada das cultivares analisadas, o que maximiza a probabilidade de novos alelos serem detectados. Entretanto, para o conjunto de cultivares registradas e protegidas no Brasil, este banco de dados de frequências alélicas constitui a referência fundamental para o embasamento dos cálculos de probabilidade de identidade genética entre amostras teste e amostras padrão. Vale lembrar que as estimativas de frequência de alelos raros, como por exemplo o alelo 111 para o loco Satt038 (f=0,008) devem ser utilizadas de forma conservadora elevando a frequência para um valor equivalente a 5/2n onde n é igual ao número de indivíduos analisados para a construção do banco ("ceiling principle" = 5/2n onde n=125) (National Research Council 1996). No caso deste banco, para alelos raros deverá ser adotada uma estimativa conservadora utilizando como teto mínimo de frequência o valor de 0,02.

O dendrograma das distâncias genéticas entre as amostras (Figura 2), permite visualizar os relacionamentos de distâncias genéticas entre as 125 amostras. No dendrograma pode-se observar que a grande maioria das amostras apresentou um perfil genético claramente distinto das demais amostras analisadas, geralmente com mais de três diferenças genotípicas. Entretanto das (125 x 124)/2 = 7750 comparações dois a dois, foi verificado que as amostras 67 e 75 apresentaram-se geneticamente indistinguíveis para os 15 locos analisados.

Além da identidade das amostras 67 e 75, foram observados outros três casos nos quais apenas um marcador apresentou genótipo distinto entre as duas amostras. Estes casos foram: 76 e 112, 45 e 56A e 73 e 95 (Figura 2). Uma vez que todos os microsatélites são sujeitos a mutações a uma taxa aproximada de 0,001 por geração existe uma probabilidade de 1 em 1.000 de que esta diferença não seja real e sim fruto de uma mutação. Neste sentido, é importante adotar uma abordagem conservadora e exigir no mínimo duas diferenças para declarar a distigibilidade com elevada confiabilidade. Com isso a declaração de que as duas amostras são geneticamente distintas quando na verdade elas são iguais passa a ser da ordem de 0,000001 (erro do tipo I), ou seja 1 em 1.000.000. Portanto, para melhor resolver a distigibilidade genética destes três pares de amostras, sugere-se uma análise adicional com pelo menos outros seis locos microsatélites.

Os dados genéticos e a análise de agrupamento mostram que as amostras com a mesma numeração seguida pelas letras A e B (14 e 14A, 31 e 31A; 56, 56A e 56B; 86, 86A e 86B; 106 e 106A) são geneticamente distintas, ou seja, estas cultivares são geneticamente diferentes.

Finalmente, é importante relatar que algumas amostras apresentaram fortes indícios de mistura de linhas e uma amostra apresentou evidências de heterozigosidade residual. As amostras 55 e 86B, apresentaram indícios de serem compostas por uma mistura de linhas, mistura esta evidenciada em dois dos 15 locos em sucessivas análises. Na análise de amostras de 50 sementes foram detectados dois alelos simultaneamente, e na análise de sementes individuais destas amostras os resultados foram diferentes entre sementes da mesma amostra. Este resultado indica mistura e não heterozigosidade uma vez que a detecção de dois alelos somente ocorreu na amostra de sementes em "bulk". Este fato foi observado em dois locos, o que sugere a hipótese de mistura de linhas para as amostras 55 e 86B. O mesmo foi observado na análise da amostra 9 porém em apenas um loco. A amostra 91, por sua vez, apresentou indícios de heterozigose residual, confirmada em pelo menos um loco em amostras de sementes individuais. Seja no caso da mistura de linhas bem como no caso da suspeita heterozigosidade residual, uma análise complementar com maior número de sementes individuais destas amostras deverá ser realizada para confirmar definitivamente estas hipóteses.

Em conclusão, este trabalho estabelece um novo sistema de identificação genética de cultivares de soja baseado em uma tecnologia rápida, econômica e, mais importante, de alto poder de resolução para testes de DHE. Das 125 amostras enviadas para análise, 123 foram declaradas geneticamente distintas e apenas duas foram geneticamente indistinguíveis com os 15 marcadores utilizados. Outros três pares de amostras apresentaram apenas uma diferença, que, embora seja muito provavelmente real, merece uma confirmação com análises de locos genéticos adicionais.

Este trabalho estabelece um banco de dados de perfis genéticos e frequências alélicas de marcadores microsatélites para as cultivares de soja registradas no Brasil e representa uma ferramenta fundamental para a condução de testes de identidade genética entre amostras questionadas e amostras padrão em procedimentos de controle de qualidade e fiscalização de identidade genética. Este banco é público e é disponibilizado pela BRASPOV aos interessados. Durante os anos de 2002 e 2003 análises de identidade de variedades de soja tem sido realizadas no nosso laboratório com o objetivo de defender a propriedade intelectual de cultivares de soja protegidas no país.

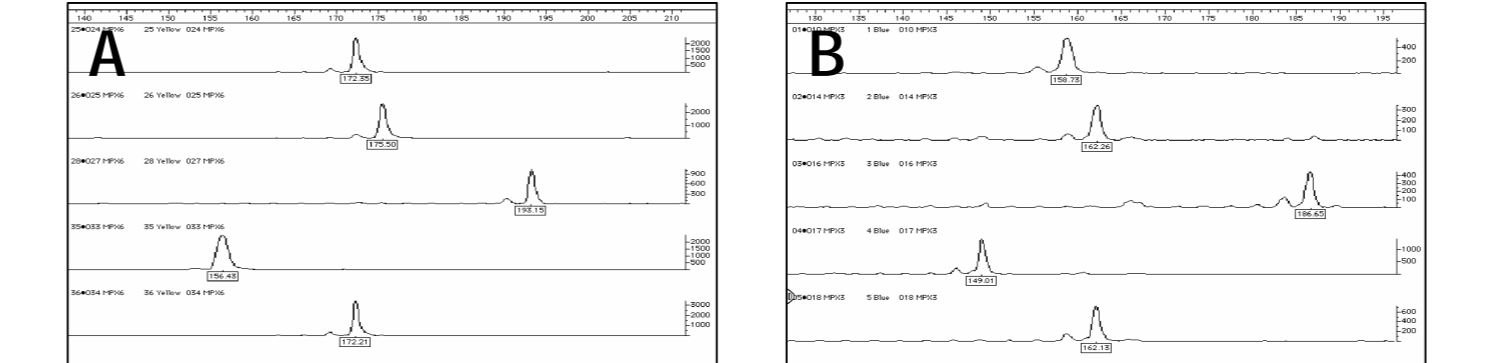


Figura 1. Exemplos de eletroferogramas das análises genéticas com dois marcadores moleculares e 10 amostras. Parâmetros de desempenho fornecido e de investigação de paternidade dados sistemas multiplex. Painel A - Marcador Satt005 – amostras 24, 25, 27, 33 e 34; Painel B - Marcador Satt070 – amostras 10, 14, 16, 17 e 18

Tabela 2. Banco de dados de perfis genéticos multiloco de 30 das 125 amostras de cultivares de soja para 15 locos microsatélites polimórficos. Os locos microsatélites foram designados de acordo com a nomenclatura vigente para soja e as amostras numeradas de acordo com a codificação atribuída pelo SNPC/MA. Os alelos são identificados em número de pares de bases de DNA de acordo com a sequência de A ou T trinucleotídeos do loco microsatélite.

Nº amostra	Satt008	Satt011	Satt014	Satt016	Satt019	Satt022	Satt025	Satt028	Satt031	Satt034	Satt037	Satt040	Satt043	Satt046	Satt049	Satt052
1	131	137	90	206	134	168	176	172	162	146	126	169	160	220	178	
2	139	134	95	206	134	177	176	172	162	164	126	169	160	184	178	
3	113	146	95	212	134	168	176	172	159	150	132	169	160	205	178	
4	113	146	119	236	134	174	164	172	162	164	132	169	160	184	178	
5	119	146	80	212	134	174	164	172	159	150	126	169	160	220	178	
6	119	140	80	206	128	174	148	172	159	160	126	169	160	220	178	
7	113	146	95	203	134	177	176	172	162	164	126	151	160	184	178	
8	119	134	95	206	134	168	148	172	162	150	126	157	160	184	178	
9	119	134	95	236	134	168	176	172	162	164	126	169	160	220	178	
10	119	140	95	218	134	174	148	157	159	150	126	139	160	220	178	
11	119	140	119	218	134	171	173	157	162	150	126	139	160	220	178	
12	113	146	95	206	140	168	148	172	183	162	132	160	160	205	178	
13	117	146	95	200	140	171	148	172	162	150	126	154	160	205	178	
14	119	146	116	203	149	171	148	175	162	150	132	151	160	208	178	
14A	119	146	95	200	140	168	173	172	159	150	126	154	166	220	178	
15	113	146	95	236	134	168	164	172	162	150	126	169	160	184	178	
16	115	146	80	218	140	171	148	172	186	162	132	139	160	220	178	
17	115	143	80	203	134	171	164	172	150	160	132	151	160	220	178	
18	119	146	80	212	134	171	148	175	162	150	132	160	160	163	178	
19	119	140	95	206	140	174	148	172	159	150	132	169	160	220	178	
20	119	140	80	200	140	171	164	178	162	164	126	154	160	208	178	
21	117	146	95	203	140	174	148	172	141	150	132	151	160	205	178	
22	113	146	80	212	140	174	148	193	162	150	126	169	160	183	223	
23	119	146	116	236	134	171	148	175	162	166	126	169	160	184	178	
24	119	146	95	212	140	171	148	172	159	148	126	169	160	208	178	
25	117	137	95	200	134	171	176	175	162	150	132	154	160	163	178	
26	113	146	95	236	134	168	176	172	162	146	126	169	160	184	178	
27	119	146	80	203	128	171	148	193	162	160	132	151	160	220	178	
28	117	140	80	200	134	171	164	172	159	164	126	154	164	163	178	
29	117	146	95	203	134	168	148	172	159	160	126	151	160	238	229	
30	119	146	80	203	140	171	148	172	162	150	126	151	164	205	178	

Código do loco microsatélite	Elemento repetido	Grupo de ligação	Sequência do iniciador direto 5'-3'	Sequência do iniciador reverso 5'-3'
Sat_038	DinucL	O	cttccaattgagaccctta	gtcttlltaaacaccacactt
Satt586	TrinucL	-	ggcgctccaacccaagat	gcgcccaaatgataaacctca
Satt045	TrinucL	E	tggtttctacttctaataatt	atgctcttccctct
Satt042	TrinucL	A1	gacttaattgctcttagta	gttgctgaccactactt
Satt070	TrinucL	B2	taaaaataaataactagagaacac	tgccattgaaagaatgatg
Satt038	TrinucL	G	gggaactctttttcttcaitaagt	ggccattgaaatggtttgactca
Satt009	TrinucL	F	aaaaaaggatccaagcc	cttcaaatctttagttgctgac
Satt005	TrinucL	D1	tatcttagagaagactcaaaa	cttcaattgagcttagta
GMABA/B	DinucL	N1	caaacataaaaagggggga	aagacacccaactaataatt
Satt002	TrinucL	D2	tggggaalaatgataaaaat	ctatttga