

**CARLOS ALEXANDRE GOMES RIBEIRO**

**CARACTERIZAÇÃO MOLECULAR DE CULTIVARES DE SOJA POR  
MEIO DE UM SISTEMA DE GENOTIPAGEM FLUORESCENTE  
UTILIZANDO MARCADORES MICROSSATÉLITES COM CAUDA  
ESTENDIDA UNIVERSAL**

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

**VIÇOSA  
MINAS GERAIS - BRASIL  
2011**

Ficha catalográfica preparada pela Seção de Catalogação e  
Classificação da Biblioteca Central da UFRV

T

R484c  
2011

Ribeiro, Carlos Alexandre Gomes, 1985-  
Caracterização molecular de cultivares de soja por meio de  
um sistema de genotipagem fluorescente utilizando marcadores  
microssatélites com cauda estendida universal / Carlos  
Alexandre Gomes Ribeiro. – Viçosa, MG, 2011.  
x, 43f. : il. (algumas col.) ; 29cm.

Orientador: Everaldo Gonçalves de Barros.  
Dissertação (mestrado) - Universidade Federal de Viçosa.  
Referências bibliográficas: f. 36-43.

1. Soja - Melhoramento genético. 2. Microssatélites  
(Genética). 3. Genética vegetal. 3. Polimorfismo (Genética).  
I. Universidade Federal de Viçosa. II. Título.

CDD 22. ed. 635.6552

**CARLOS ALEXANDRE GOMES RIBEIRO**

**CARACTERIZAÇÃO MOLECULAR DE CULTIVARES DE SOJA POR  
MEIO DE UM SISTEMA DE GENOTIPAGEM FLUORESCENTE  
UTILIZANDO MARCADORES MICROSSATÉLITES COM CAUDA  
ESTENDIDA UNIVERSAL**

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

APROVADA: 4 de julho de 2011.

---

**Pesq. Newton Deniz Piovesan**

---

**Prof. Pedro Ivo Vieira Good God  
(Coorientador)**

---

**Prof. Everaldo Gonçalves de Barros  
(Orientador)**

## AGRADECIMENTOS

Agradeço primeiramente a Deus por me proporcionar a oportunidade de crescimento pessoal e profissional através da experiência adquirida em mais uma importante etapa em minha vida.

À Universidade Federal de Viçosa, especialmente ao Programa de Pós-Graduação em Genética e Melhoramento pela oportunidade de realizar o curso e à FINEP pela concessão do financiamento dos estudos.

À AgroGenética que contribuiu diretamente aos estudos efetuados através da logística disponibilizada.

Ao Professor Everaldo Gonçalves de Barros, não só pela orientação, mas pela amizade e pelo exemplo de profissionalismo e dedicação. Muito obrigado pelos conselhos e pela confiança depositada durante todos esses anos.

Aos conselheiros, professores e pesquisadores Francismar Correa Marcelino, Pedro Ivo Vieira Good God, Ivan Schuster, Newton Deniz Piovesan, Eunize Maciel Zambolim e Eveline Teixeira Caixeta pela ajuda e apoio durante o desenvolvimento do trabalho.

Aos amigos da AgroGenética: Janaína, Rosinéa, Luciano, Leandro, Bruna, Ronaldo, Talles, James, Michelle, Sibelle, Darlan, Dahyana, Lílian, Joesse, Carla, Wilton, Flávia, Paulo e Marta. Principalmente aqueles diretamente envolvidos com o trabalho, como a Janaína, que muito me ajudaram e foram fundamentais para o desenvolvimento da pesquisa. Obrigado a todos vocês pela convivência gratificante durante todo esse tempo.

Aos amigos do Bioagro: Demerson, Bruno, Marcão, Luís, Rafael, Arlindo, Klever, Cássio, Rejane, Kátia, Cristina, Samuel, Daniela, Felipe, e Rodrigo, que sempre se mostraram à disposição quando precisei de ajuda e muito contribuíram para a minha formação.

A todos os funcionários da UFV, principalmente aos do BIOAGRO.

Aos meus colegas e amigos de mestrado que me acompanharam nessa jornada, participando direta ou indiretamente deste trabalho.

Aos professores da Universidade Federal de Viçosa, pela valiosa contribuição científica e profissional.

Aos meus amigos de república Hudson e Luciano pelo convívio e pela grande amizade ao longo de todos esses anos.

À Patrícia pelo companheirismo, amor e compreensão e pelo apoio em todos os momentos de dificuldades.

Aos meus pais Duílio e Rosa, à minha irmã Maria Virgínia, e ao meu avô Carmo que nunca deixaram de me apoiar e me incentivar em todas as etapas da minha vida.

## ÍNDICE

	Página
LISTA DE ILUSTRAÇÕES.....	v
LISTA DE TABELAS.....	vi
RESUMO .....	vii
ABSTRACT .....	ix
1. INTRODUÇÃO .....	01
2. MATERIAL E MÉTODOS .....	06
2.1 Material vegetal, extração e quantificação de DNA.....	06
2.2 Seleção dos <i>primers</i> .....	07
2.3 Condições da PCR e eletroforese .....	09
2.4 Índice de polimorfismo, probabilidade de identidade ao acaso e heterozigosidade observada.....	11
2.5 Relações genéticas entre os cultivares .....	12
2.6 Correlação entre os métodos de genotipagem.....	13
3. RESULTADOS E DISCUSSÃO.....	13
3.1 Verificação do perfil de amplificação .....	13
3.2 Estimativa do tamanho dos fragmentos .....	14
3.3 Informatividade dos locos .....	25
3.4 Frequência alélica e probabilidade de identidade ao acaso .....	28
3.5 Análise de correlação de matrizes de dissimilaridade e agrupamento entre os cultivares .....	31
4. CONSIDERAÇÕES FINAIS .....	35
5. REFERÊNCIAS BIBLIOGRÁFICAS .....	36

## LISTA DE ILUSTRAÇÕES

	Página
Figura 1 – Sistema de <i>primers</i> com cauda estendida universal (PCEU).....	8
Figura 2 – Comparação do perfil de amplificação do sistema de genotipagem convencional com o PCEU.....	14
Figura 3 – Representação do perfil de amplificação de genótipos avaliados em sequenciador semi-automático ABI 3130 xl.....	15
Figura 4 – Tamanho estimado e frequência dos alelos detectados nos métodos de genotipagem PCEU e convencional.....	22
Figura 5 – Agrupamento dos cultivares pelo método UPGMA, através da genotipagem pela metodologia PCEU.....	32
Figura 6 – Agrupamento dos cultivares pelo método UPGMA, através da genotipagem pela metodologia convencional.....	33

## LISTA DE TABELAS

	Página
Tabela 1 – Marcadores microssatélites selecionados para análise dos cultivares de soja, sequência senso com cauda estendida universal, antisense, estrutura repetitiva e grupo de ligação.....	10
Tabela 2 – Tamanho estimado dos fragmentos amplificados nos 30 cultivares para cada marcador nas metodologias.....	16
Tabela 3 – Número de alelos, conteúdo de informação polimórfica e heterozigosidade observada nas diferentes metodologias de genotipagem.....	26
Tabela 4 – Frequência alélica e probabilidade de identidade ao acaso dos 30 cultivares em 13 marcadores microssatélites pelo método de genotipagem PCEU.....	29
Tabela 5 – Frequência alélica e probabilidade de identidade ao acaso dos 30 cultivares em 13 marcadores microssatélites pelo método de genotipagem convencional.....	30
Tabela 6 – Formação dos grupos provenientes do agrupamento pelo método UPGMA.....	34

## RESUMO

RIBEIRO, Carlos Alexandre Gomes, M.Sc., Universidade Federal de Viçosa, julho de 2011. **Caracterização molecular de cultivares de soja por meio de um sistema de genotipagem fluorescente utilizando marcadores microssatélites com cauda estendida universal.** Orientador: Everaldo Gonçalves de Barros. Co-orientadores: Pedro Ivo Vieira Good God e Francismar Correa Marcelino

O desenvolvimento de sistemas de identificação molecular para determinar e rastrear a identidade genética de um cultivar representa um grande desafio ao sistema de proteção utilizado no país, até então fundamentado em descritores fenotípicos. A identificação de cultivares de soja é feita atualmente no Brasil com base em 30 descritores morfológicos. Marcadores moleculares do tipo microssatélite são largamente utilizados para fins de determinação da diversidade genética, paternidade, análises de pureza varietal e mapeamento genético. Algumas variações da técnica de genotipagem convencional por microssatélite vêm sendo desenvolvidas com o objetivo de adequar a velocidade e automatização de técnicas atuais, com qualidade de análise e baixo custo. Dentre elas, destaca-se um sistema multiplex de genotipagem em sequenciador semi-automático de DNA, que utiliza *primers* com cauda estendida universal (PCEU). O presente trabalho teve como objetivos padronizar um sistema reprodutível de genotipagem molecular semi-automatizado baseado na metodologia PCEU para a cultura da soja, avaliar diferenças em relação ao sistema de genotipagem convencional, em gel de poliacrilamida e caracterizar com o conjunto de marcadores selecionados, 30 cultivares de soja estimando suas frequências alélicas. Estes descritores genotípicos poderão ser utilizados como ferramenta para a identificação de cultivares e fornecer subsídios para a proteção da propriedade intelectual, determinação e comprovação de pureza varietal. Foram utilizados 30 cultivares comerciais de soja avaliados com um grupo de 22 marcadores microssatélites já descritos na literatura. A genotipagem foi realizada em sequenciador semi-automático (PCEU) e em gel de poliacrilamida (convencional). Dentre os locos analisados, 13 foram selecionados por apresentarem melhor perfil de amplificação, menos produtos *stutter*, e maior número de alelos. O perfil alélico de cada cultivar foi definido primeiramente em sistema tradicional de

genotipagem por eletroforese em gel de poliacrilamida, e depois no sequenciador semi-automatizado. Para a metodologia de genotipagem PCEU, o número total de alelos encontrados foi de 50, variando de 2 (Satt045) a 7 (Satt005). O conteúdo de informação polimórfica (PIC) variou de 0,40 (Satt045) a 0,74 (Satt005) com média de 0,62. Para a metodologia convencional o número de alelos encontrados foi de 38 com uma variação de 2 (Satt045, Satt070 e AF162283) a 5 (Satt005) alelos. O PIC apresentou uma faixa 0,39 (Satt045) a 0,67 (Satt079), com média de 0,56. Os valores de probabilidade de identidade ao acaso diferiram entre os métodos, sendo  $<10^{-5}$  (PCEU) e  $<10^{-4}$  (Convencional). Apesar de diferenças pontuais, observou-se alta correlação entre as matrizes de dissimilaridade genética entre os métodos (0,8026), e os grupos formados apresentaram também alta similaridade e foram concordantes com dados fenotípicos de registro varietal e de genealogia. Além da alta sensibilidade na detecção de alelos de tamanhos muito próximos, o método PCEU permitiu uma identificação mais criteriosa dos artefatos de amplificação. Os grupos de marcadores selecionados em ambas as metodologias permitiram distinguir todas as cultivares analisadas. O método PCEU possui baixo custo quando comparado a técnicas comuns de marcação por fluorescência, além disso, sua alta precisão faz deste um método vantajoso para a caracterização de cultivares e determinação da pureza genética. O grupo de marcadores analisados pode ser usado como um conjunto de descritores genotípicos complementar aos descritores fenotípicos obrigatórios utilizados para fim de proteção de cultivares em testes de distinguibilidade.

## ABSTRACT

RIBEIRO, Carlos Alexandre Gomes, M.Sc., Universidade Federal de Viçosa, July 2011. **Molecular characterization of soybean cultivars by a fluorescent genotyping system using microsatellite markers with a universal tail extension.** Advisor: Everaldo Gonçalves de Barros. Co-advisors: Pedro Ivo Vieira Good God and Francismar Correa Marcelino

The development of molecular identification systems to determine and trace the genetic identity of a cultivar is a major challenge for the cultivar protection system used in Brazil, which is based on phenotypic descriptors. Identification of soybean cultivars in Brazil is currently based on 30 morphological descriptors. Microsatellite markers are widely used for determination of genetic diversity, paternity tests, varietal purity and genetic mapping. Some variations of the conventional technique for microsatellite genotyping have been developed in order to match the speed and automation of current techniques with quality and low cost. Among them, there is a multiplex genotyping system in semi-automatic DNA sequencers, using universal tail extended primers (UTEP). This study aimed to standardize a robust system of semi-automatic molecular genotyping based on the UTEP methodology for soybean identification, to evaluate differences between this method and the conventional genotyping system using polyacrylamide gel electrophoresis and to characterize 30 soybean cultivars with a set of selected markers, estimating their allelic frequencies. These descriptors can be used as genotypic tool for cultivar identification and provide subsidies for the protection of intellectual property, determination and testing of varietal purity. Thirty commercial soybean cultivars were evaluated with a group of 22 microsatellite markers described in the literature. Genotyping was performed in a semi-automatic sequencer (UTEP) or by polyacrylamide gel electrophoresis (conventional). Among the loci analyzed, 13 were selected because they presented good amplification profiles, with few stutter products, and a large number of alleles. The allelic profile of each cultivar was first defined by the genotyping conventional system, and then by the semi-automated sequencer. For the UTEP genotyping methodology, the total number of

alleles found was 50, ranging from 2 (Satt045) to 7 (Satt005). The polymorphic information content (PIC) ranged from 0.40 (Satt045) to 0.74 (Satt005) with an average of 0.62. For the conventional method the number of alleles found was 38 with a range of 2 (Satt045, Satt070 and AF162283) to 5 (Satt005) alleles. The PIC had a range 0.39 (Satt045) to 0.67 (Satt079), with an average of 0.56. The values of probability for random identity differed between the methods being  $<10^{-5}$  (UTEF) and  $<10^{-4}$  (conventional). Despite slight differences, there was a high correlation between genetic dissimilarity matrices obtained by the two methods (0.8026), and the groups formed also showed high similarity among them and they were coherent with the phenotypic data used for varietal registration and with the genealogies of the cultivars. In addition to the high sensitivity for detection of alleles of very close sizes, the UTEF method allowed a more accurate identification of amplification artifacts. The marker selected for both methods allowed the distinction of all cultivars analyzed. The UTEF method has a low cost when compared to common techniques of fluorescence tagging, in addition, its high accuracy makes this an advantageous method for cultivar characterization and determination of genetic purity. The set of markers tested in this work can be used as a set of genotype descriptors complementary to the phenotypic descriptors already used for cultivar protection in distinctness tests.

## INTRODUÇÃO

A capacidade de determinar a identidade genética de um cultivar é fundamental para garantir a rastreabilidade de materiais ao longo da cadeia produtiva e comercial, a pureza varietal nas etapas de certificação e, ainda, para gerar descritores que permitem a proteção de cultivares.

A identificação genética pode ser conduzida com base em descritores fenotípicos ou genotípicos. O uso de descritores fenotípicos pode estar sujeito a limitações, dentre elas a identificação baseada em características de herança genética quantitativa e, portanto, influenciadas pelas variações ambientais e interação genótipo com o ambiente (Collard et al., 2005; Brondani & Brondani, 2006). Além destes aspectos, marcadores morfológicos apresentam a desvantagem de serem somente identificados, em sua maioria, na planta adulta. Na identificação e quantificação da diversidade genética, o uso de características morfológicas, proteínas e isoenzimas apresentam limitações, principalmente no que se refere ao número insuficiente de marcadores disponíveis (Priolli et al., 2002).

A distinguibilidade, requisito fundamental que compõe os testes de DHE (Distinguibilidade, Homogeneidade e Estabilidade) para a proteção de um cultivar, tem sido dificultada devido à alta similaridade fenotípica, cada vez mais evidente, proporcionada pelo estreitamento da base genética de espécies cultivadas, sobretudo em espécies autógamas, como a soja (Priolli et al., 2004; Vieira et al., 2009).

Nos processos de certificação de sementes genéticas, a presença de sementes de outros cultivares em quantidade superior à tolerada, pode condenar todo o lote de sementes (Moreira et al., 1999). Em muitos casos, a identificação destas sementes é determinada por métodos subjetivos, que podem levar a erros e dúvidas na distinção entre os cultivares. Sendo assim, tanto para a proteção de cultivares, quanto para a identificação correta de mistura varietal na produção de sementes, descritores mais informativos e estáveis são necessários, a fim de diferenciar claramente os genótipos (Schuster et al., 2004; Rabel et al., 2010).

Por outro lado, marcadores moleculares podem ser utilizados para caracterizar o genótipo de um indivíduo a partir de amostras de células ou tecidos, portanto, independente do estágio fisiológico da planta (Caixeta et al., 2010), além

de possuírem um número virtualmente ilimitado de polimorfismos distribuídos aleatoriamente ao longo de todo o genoma (Borém & Caixeta, 2010).

No Brasil, desde a aprovação da Lei de Proteção de Cultivares - Nº 9.456 - 25/04/1997, o país passou a contar com um novo mecanismo legal para proteger seus cultivares e, dessa forma, proteger o alto investimento destinado ao seu desenvolvimento, de modo a evitar fraudes e apropriação indevida da propriedade intelectual (Carvalho & Pessanha, 2001; Rodrigues et al., 2008).

A cultura da soja é de grande expressão para o agronegócio nacional, sendo a cultura agrícola brasileira que mais cresceu nas últimas três décadas e corresponde a 49% da área plantada de grãos do país (BRASIL, 2011). O aumento da produtividade é atribuído, além de novas tecnologias de manejo e eficiência dos produtores, aos grandes investimentos e ao desenvolvimento tecnológico, que são transferidos aos produtores via sementes de alta qualidade genética.

O Brasil é atualmente o segundo maior produtor de soja. Na safra 2009/10 foram produzidos 68,69 milhões de toneladas do grão em uma área de 23,47 milhões de hectares. Para a safra 2010/11 a produção aumentou 6,35 milhões de toneladas (9,2%), atingindo o volume de 75,04 milhões de toneladas, com um incremento de área de 690 mil hectares (3%), superando os 23,47 milhões de hectares cultivados com soja no país até o presente momento (CONAB, 2011).

Mais de 800 cultivares de soja estão hoje registrados junto ao Registro Nacional de Cultivares (RNC), sendo 23 somente nos três primeiros meses de 2011 e mais de 450 cultivares estão protegidos no Serviço Nacional de proteção de cultivares (SNPC) (BRASIL, 2011). Embora esse número possa ser expressivo, verifica-se pouca variabilidade genética, devido principalmente ao fato de serem originários de um mesmo conjunto de poucos ancestrais. WYSMIERSKI (2010), analisando um grupo de 444 cultivares brasileiros a partir do coeficiente de parentesco, confirmou que a base genética desses cultivares é constituída por 60 ancestrais, dos quais apenas quatro (CNS, S-100, Nanking e Tokyo) correspondem a 55,26% de toda a base genética. Apesar do número de ancestrais no conjunto de cultivares produzidos em empresas públicas ser maior do que no de cultivares de empresas privadas, ambas as instituições utilizam uma base genética de alta similaridade.

A identificação de genótipos de soja é feita atualmente com base em 30 descritores morfológicos obrigatórios: coloração de flor, hipocótilo, hilo,

tegumento, vagem e pubescência; densidade da pubescência na haste principal; formato, intensidade da cor e rugosidade da folha; hábito e tipo de crescimento; grupo de maturidade relativa; transgenia; forma e tamanho predominante da semente; cor e intensidade do brilho do tegumento; peso de 100 sementes; altura da planta; ciclo vegetativo e ciclo total; reação a peroxidase; reação a pústula bacteriana; reação à mancha “olho de rã”; reação ao cancro da haste; reação à necrose da haste; reação aos nematóides de galhas (*Meloidogyne javanica* e *Meloidogyne incognita*); e reação ao nematóide de cisto (raça 3) (Vieira et al., 2009; MAPA, 2011).

Além dos descritores obrigatórios, a identificação conta com características adicionais que envolvem, em sua maioria, reações a doenças como: podridão parda da haste, oídio, mancha alvo, vírus do mosaico, nematóides de galhas, nematóides de cisto, entre outras (MAPA, 2011). A utilização de técnicas moleculares na obtenção de *fingerprintings* para a caracterização de cultivares, sobretudo perfis alélicos provenientes de microssatélites, vem sendo aplicadas para fins de proteção pelo departamento de agricultura dos Estados Unidos (USDA) como evidência da singularidade de um novo cultivar (Diwan & Cregan, 1997). No Brasil, embora seja utilizada, a análise molecular ainda não é reconhecida como metodologia oficial no processo de proteção de cultivares, sendo este realizado apenas com base em características fenotípicas.

O uso de marcadores moleculares no melhoramento de plantas vem se tornando cada vez mais frequente. Marcadores moleculares do tipo microssatélite são largamente utilizados para fins de determinação da diversidade genética, paternidade, análises de pureza varietal e mapeamento genético (Song et al., 2004; Schuster et al., 2004; Varshney et al., 2005; Garcia et al., 2007). Embora seja moroso e oneroso seu desenvolvimento inicial, sua utilização é justificada por apresentarem características vantajosas para identificação de genótipos, tais como sua natureza multialélica, co-dominante, alto polimorfismo e ampla distribuição no genoma, por serem extremamente conservados dentro da espécie, além da alta reprodutibilidade, característica fundamental para a sua aplicação em diagnósticos moleculares (Powell et al., 1996).

Existem disponíveis atualmente na literatura uma série de microssatélites desenvolvidos para a cultura da soja, o que justifica a sua ampla utilização em

diversos tipos de estudos (Cregan et al., 1994; Akkaya et al., 1995; Diwan & Cregan, 1997; Cregan et al., 1999; Song et al., 2004).

As primeiras técnicas de análise de fragmentos de tamanhos relativamente próximos foi aplicada em géis de poliacrilamida desnaturante ou não desnaturante, corados com nitrato de prata ou brometo de etídio (Caixeta et al., 2010). Porém, a técnica de eletroforese, apesar da grande utilidade, vem gradativamente entrando em desuso para procedimentos de genotipagem, sobretudo quando está envolvido um grande número de marcadores e genótipos. Esse fato se deve principalmente à exigência de considerável prática de laboratório, além de tempo, tanto nos procedimentos laboratoriais quanto na análise dos dados. Metodologias semi-automáticas, em sequenciadores, por outro lado, permitem a avaliação de um maior número de *primers* simultaneamente, uma vez que estes podem ser marcados com diferentes fluorescências, com um menor tempo de corrida, além da análise automatizada dos dados realizada por softwares específicos (Dresler-Nurmi, 2000).

Técnicas semi-automáticas, por detectarem os dados em análise com alta precisão e maior praticidade (Lazaruk et al., 1998), proporcionam uma avaliação mais criteriosa de artefatos de amplificação como *stutter*, picos (+A/-A), *dropout* de alelos, e desta forma reduzindo a probabilidade de determinação incorreta do perfil de um genótipo.

Com o advento de técnicas semi-automáticas para análise de fragmentos, tornou-se bastante comum o uso de *primers* marcados com fluorescência (6-FAM, HEX, TET, NED, entre outros), que são detectados em sequenciadores automáticos de DNA que utilizam a eletroforese capilar. Apesar da notável praticidade desta técnica, a síntese de *primers* marcados com fluorescência ainda possui um custo bastante elevado (5 a 10 vezes) quando comparado ao dos *primers* não marcados (Hayden et al., 2008; Arruda et al., 2010), embora o seu preço venha sendo gradativamente reduzido. Desta forma, novas metodologias vêm sendo desenvolvidas com o objetivo de adequar a velocidade e automatização do processo com qualidade de análise e baixo custo.

Algumas variações da técnica de genotipagem convencional por microssatélite vêm sendo desenvolvidas e dentre elas, merece destaque uma metodologia descrita inicialmente por OETTING et al. (1995), que envolve a utilização conjunta de três *primers* (*primer forward*, *primer reverse* e *primer*

inespecífico marcado com fluorescência). Em um dos pares específicos, além da sequência complementar, é também sintetizada uma cauda estendida em sua extremidade 5', com cerca de 20 nucleotídeos, sem homologia com a região de anelamento do *primer*. O *primer* marcado com fluorescência possui sequência igual a da cauda estendida, e, a partir dos primeiros ciclos de termociclagem, os fragmentos amplificados da região genômica de interesse, tornam-se detectáveis devido à fluorescência incorporada pelo *primer* inespecífico.

A grande vantagem desta metodologia é a necessidade de sintetizar apenas um *primer* marcado para cada uma das fluorescências (6-FAM, HEX, NED, TET), que serão usados nas combinações tríplice, o que resulta em uma redução significativa no custo. Segundo HAYDEN et al. (2008), o custo de um *primer* marcado pode ser de cinco a dez vezes o valor de *primers* não marcados. Além disso, em determinadas análises genéticas, a avaliação de um loco pode não passar de algumas dezenas de vezes, o que inviabiliza a síntese de um *primer* marcado com fluorescência para cada loco, já que quando estocado por muito tempo, o fluoróforo tende a perder intensidade de fluorescência (Missiaggia & Grattapaglia, 2006).

Esta técnica foi inicialmente descrita em estudos para avaliação de genótipos em humanos (Oetting et al., 1995; Neilan et al., 1997; Schuelke, 2000). Há também relatos de sua utilização com sucesso em trabalhos aplicados à genotipagem animal, como é o caso da espécie de lagosta *Panulirus argus* (Diniz et al., 2009) e dos anfíbios *Rhinella schneideri* e *Hypsiboas ranicepse* (Arruda et al., 2010), na genotipagem vegetal, para cultura do feijoeiro (Oblessuc et al., 2009), damasco, cereja, cevada, trigo (Hayden et al., 2008) e eucalipto (Missiaggia & Grattapaglia, 2006).

O presente trabalho teve como objetivo padronizar um sistema reprodutível de genotipagem molecular semi-automatizada baseado na metodologia PCEU (*Primers* com cauda estendida universal) para a cultura da soja, avaliar diferenças em relação ao sistema de genotipagem convencional em gel de poliacrilamida e caracterizar 30 cultivares de soja com o conjunto de marcadores selecionados, estimando suas frequências alélicas. Estes descritores genotípicos poderão ser utilizados como ferramenta para a identificação de cultivares e ampliar o rol de características que diferenciam novos cultivares dos

demais existentes no mercado, além de fornecer subsídios para a proteção da propriedade intelectual, determinação e comprovação de pureza varietal.

## **2. MATERIAL E MÉTODOS**

### **2.1. Material vegetal, extração e quantificação de DNA**

Foi utilizado um grupo de 30 cultivares comerciais de soja, desenvolvidos por diferentes instituições públicas e privadas, são eles: BMX Magna RR; BRS 133; BRS 184; BRS 213; BRS 214; BRS 255 RR; BRS 256 RR; BRS 262; BRS 270 RR; BRS Charrua RR; BRS Gisele RR; BRS MG 850G RR; BRS Pampa RR; BRS MT Pintado; BRS Valiosa RR; CD 202; CD 206; CD 213 RR; CD 214 RR; CD 215; CD 217; CD 219 RR; CD 226 RR; CD 228; CD 229 RR; CD 237 RR; CD 240 RR; MG/BR 46 Conquista; MSoy 7501; Vmax (NK 7059 RR). Os materiais foram cedidos pelo banco ativo de germoplasma da Embrapa Soja Londrina, e por empresas privadas como a Brasmax, Coodetec, Monsanto e Syngenta.

A extração do DNA genômico das variedades de soja foi realizada conforme protocolo descrito por MARCELINO et al. (2008), com pequenas modificações, utilizando a resina Wizard® DNA Purification (Promega). Cinquenta sementes de cada variedade foram trituradas para realização da análise em *bulk*. Foi retirado cerca de 200 mg da amostra e adicionado 1290 µL do tampão de extração (NaCl 150 mM; Tris-HCl 10 mM pH 8,0; EDTA 2 mM pH 8,0; SDS 1 %), 150 µL de hidrocloreto de guanidina 5 M e 60 µL de Proteinase K (20 mg/mL). Após essa etapa, as amostras foram incubadas por aproximadamente 12 h a 55 °C sob agitação constante. Cerca de 500 µL do sobrenadante foram misturados a 1 mL da resina Wizard (Promega), a solução foi aplicada a uma mini coluna Wizard, e com o auxílio de uma bomba de vácuo promoveu-se uma leve sucção facilitando a passagem do líquido através da coluna. Logo após, foram aplicados 2 mL de isopropanol 80 % às colunas, que em seguida, foram centrifugadas a 10.000 g por 2 min para eliminar o isopropanol residual. Os ácidos nucleicos foram eluídos com 70 µL de água MilliQ a 70 °C.

A quantificação do DNA extraído foi realizada em espectrofotômetro, para ajuste de sua concentração (Sambrook et al., 1989) e verificação de sua qualidade de acordo com parâmetros como razão da absorbância nos comprimentos de onda

260/280 nanômetros, degradação e presença de RNA. Após a avaliação e quantificação, o DNA foi diluído em TE 0,1 M para a concentração de 10 ng/μL e armazenado a -20 °C até sua utilização.

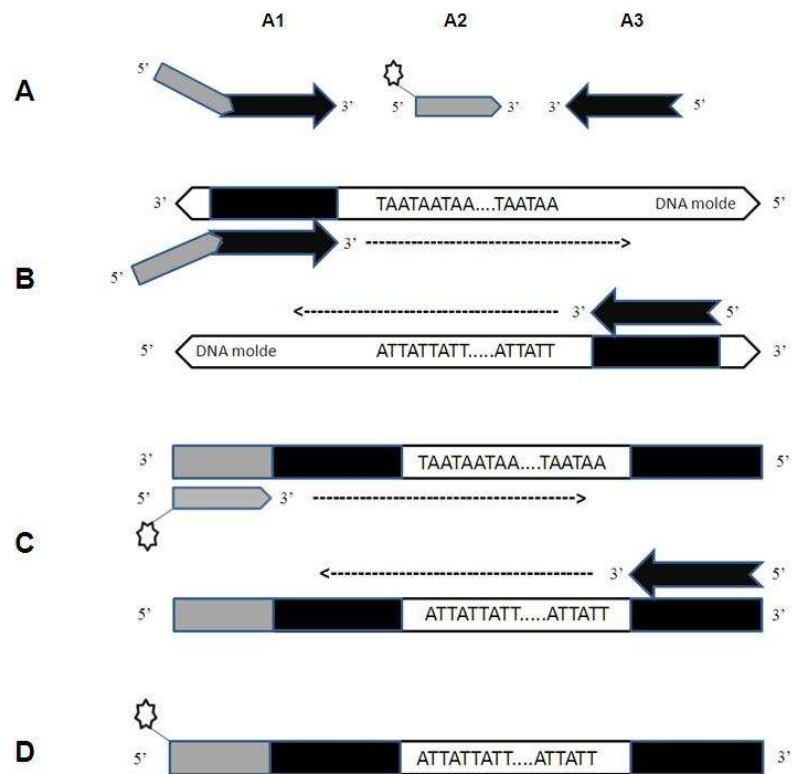
## 2.2. Seleção dos *primers*

Foram selecionados 22 marcadores microssatélites cujas sequências estão disponíveis na literatura, de acordo com o número de alelos por loco apresentado, distribuição nos cromossomos, natureza da repetição e valor de PIC (Conteúdo de informação polimórfica) (Diwan & Cregan, 1997; Song et al., 1999; Priolli et al., 2002; Song et al., 2004; Schuster et al., 2004; Garcia et al., 2007; Mimura et al., 2007). Os marcadores pré-selecionados foram: AF162283, GMABAB, Sat\_038, Sat\_085, Sat\_105, Satt002, Satt005, Satt009, Satt038, Satt042, Satt045, Satt046, Satt070, Satt079, Satt100, Satt114, Satt181, Satt216, Satt233, Satt431, Satt540, Satt586. Aqueles que apresentaram melhor perfil de amplificação, menos produtos *stutter*, maior número de alelos e melhor se adaptaram às condições comuns da PCR, foram conduzidos para as etapas posteriores.

Os iniciadores foram sintetizados pela Prodimol Biotecnologia (<http://prodimol.com.br>) com exceção dos *primers* senso marcados com fluorescência, que foram sintetizadas pela Applied Biosystems (<http://appliedbiosystems.com.br>). As sequências dos marcadores estão disponíveis no portal do Departamento de Agricultura dos Estados Unidos (<http://bldg6.arsusda.gov/~pooley/soy/cregan/soymap.htm>).

Fez-se o uso de um sistema multiplex, conforme descrito inicialmente por OETTING et al. (1995), que contém três tipos de *primers* distintos: um *primer* senso microssatélite que possui, além de sua sequência específica, uma cauda de 17 pares de bases (pb) em sua extremidade 5', com sequência idêntica à dos *primers* universais comumente usados em biologia molecular, tais como M13 (GTAAAACGACGGCCAGT), T3 (ATTAACCCTCACTAAAG) e T7 (AATACGACTCACTATAG); um *primer* anti-senso microssatélite regular; e um terceiro *primer* senso, marcado com fluoróforo, com a mesma sequência da cauda do *primer* senso (Figura 1). Durante as etapas iniciais da PCR o *primer* senso microssatélite com a cauda estendida universal irá ancorar na região genômica flanqueadora do motivo repetitivo, promovendo a extensão da fita. De modo similar, ocorrerá ao *primer* microssatélite anti-senso, e ao final de alguns ciclos

iniciais, o *primer* microssatélite anti-senso poderá ancorar em fitas recém produzidas e sintetizar sequências complementares ao molde da cauda estendida universal. Após mais alguns ciclos da reação, o *primer* marcado com fluorescência irá ancorar nessas regiões complementares e sintetizar assim fragmentos que serão fluorescentes (Figura 1).



**Figura 1.** Sistema de *primers* com cauda estendida universal (PCEU). *Primers* utilizados na amplificação do fragmento fluorescente (A). *Primer* senso microssatélite com a cauda estendida universal (A1); *Primer* marcado com fluorescência com a mesma sequência da cauda estendida universal (A2). *Primer* anti-senso microssatélite (A3). Primeiras etapas do ciclo de amplificação (B), as regiões específicas do genoma complementar aos *primers* microssatélites estão demonstradas pelas caixas pretas. Pareamento do *primer* marcado com fluorescência, após a síntese de fragmentos nas etapas iniciais da PCR (C). A caixa cinza representa a sequência da cauda estendida universal incorporada ao fragmento amplificado devido ao anelamento do *primer* demonstrado em B. Amplicons marcados com fluorescência em etapas finais da PCR (D).  
Fonte: adaptado de Schuelke (2000).

Foi verificada a possibilidade de pareamento das sequências dos *primers* universais com o genoma da soja através do uso do algoritmo BLASTn no banco de dados do Genbank.

Para verificar a eficiência da combinação da sequência da cauda, com a sequência específica do *primer*, foi realizado um teste com o objetivo de confirmar a diferença de tamanho esperado, em pares de bases, entre os produtos amplificados em PCR convencional e por PCR no método de *primers* com cauda estendida universal (PCEU). Os cultivares BMX Magna RR e CD 202 foram avaliados em duplicata utilizando os 14 marcadores selecionados apresentados na Tabela 1.

### 2.3. Condições da PCR e eletroforese

As reações de PCR foram preparadas para um volume final de 20  $\mu$ L, contendo 30 ng de DNA de soja, tampão de PCR 1X (Tris-HCl 10 mM, pH 8,8, KCl 50 mM, Nonidet P40 0,08%), MgCl<sub>2</sub> 1,5 mM, dNTP 0,2 mM, 0,2  $\mu$ M de cada *primer* (senso e anti-senso) e 1,0 U de Taq DNA polimerase (Fermentas). Para as reações de PCR com PCEU, foi adicionado um *primer* senso com aproximadamente o dobro da extensão dos demais *primers*. Desta forma foi necessária a redução de sua concentração em 5 vezes (0,04  $\mu$ M) em relação aos outros dois *primers* presentes na reação (anti-senso e senso marcado com fluorescência) que foram mantidos na concentração do procedimento convencional (0,2  $\mu$ M). As reações de PCR foram conduzidas separadamente para cada loco microssatélite. O programa de amplificação consistiu de uma etapa inicial de desnaturação de 94 °C por 5 min, seguida de 35 ciclos com desnaturação a 94 °C por 40 s, anelamento a 50 °C por 40 s e extensão a 72 °C por 40 s. A etapa de extensão final foi programada para 72 °C por 25 min.

Os produtos da PCR convencional foram separados por eletroforese em gel de poliacrilamida 10%, usando-se tampão TBE 1X (Tris 89 mM e EDTA 20 mM ácido bórico 89mM, pH 8,0), por aproximadamente 3 h a 120 volts. Foram utilizadas cubas verticais com placas de gel de 15 cm de altura com capacidade para 60 amostras em gel de 1,0 mm de espessura. Após a eletroforese, os géis foram corados em solução de brometo de etídio (0,2  $\mu$ g/ml) por 10 min. A variação do tamanho dos fragmentos de cada loco foi estimada por meio de comparação visual com o marcador de tamanho de 50 pb (Promega).

Para a detecção simultânea dos alelos amplificados por fluorescência, fez-se o agrupamento dos marcadores com diferentes fluóforos. Foi possível o agrupamento de marcadores com o mesmo tipo de fluorescência, respeitando a

**Tabela 1.** Marcadores microssatélites selecionados para análise dos cultivares de soja, sequência senso com cauda estendida universal (CEU) destacada em preto, antisenso, estrutura repetitiva e grupo de ligação (G.L.).

Marcador	C.E.U	Senso 5'→3'	Antisenso 5'→3'	Estrutura repetitiva	G.L.
AF162283/T3	<b>ATTAACCCCTCACTAAAG</b>	GCGAGTTCTGGATGTAGG	GCGTGGCGGCTTTGGTAG	(CT) <sub>11</sub>	G
GMABAB/T3	<b>ATTAACCCCTCACTAAAG</b>	CAAAACATAAAAAAGGTGAGA	AAGAACCACACTAATATTATT	(ATT) <sub>25</sub>	N1
Satt002/T7	<b>AATACGACTCACTATAG</b>	TGTGGGTAAAATAGATAAAAAAT	TCATTTTGAATCGTTGAA	(ATT) <sub>25</sub>	D2
Satt005/T7	<b>AATACGACTCACTATAG</b>	TATCCTAGAGAAGAACTAAAAAA	GTCGATTAGGCTTGAAATA	(ATT) <sub>19</sub>	D1b
Satt045/T7	<b>AATACGACTCACTATAG</b>	TGGTTTCTACTTTCTATAATTATTT	ATGCCTCTCCCTCCT	(ATT) <sub>18</sub>	E
Satt070/M13	<b>GTAAAACGACGGCCAGT</b>	TAAAAATTTAAAATACTAGAAGACAAC	TGGCATTAGAAAATGATATG	(ATT) <sub>24</sub>	B2
Satt079/T3	<b>ATTAACCCCTCACTAAAG</b>	AGTCGAAGATACACAATTAGAT	CTTTTAGACACAAATTTATCACT	(ATT) <sub>13</sub>	C2
Sat_085/T3	<b>ATTAACCCCTCACTAAAG</b>	GGTTTTAGATCCTTAAATTTGT	GGGAAGCAAGTAGCT	(AT) <sub>23</sub>	C1
Satt114/T3	<b>ATTAACCCCTCACTAAAG</b>	GGGTTATCCTCCCAATA	ATATGGGATGATAAGGTGAAA	(ATT) <sub>17</sub>	F
Satt233/M13	<b>GTAAAACGACGGCCAGT</b>	AAGCATACTCGTCGTAAC	GCGGTGCAAAGATATTAGAAA	(ATT) <sub>16</sub>	A2
Satt431/T7	<b>AATACGACTCACTATAG</b>	GCGTGGCACCCTTGATAAATAA	GCGCACGAAAGTTTTTCTGTAACA	(ATT) <sub>21</sub>	J
Satt540/M13	<b>GTAAAACGACGGCCAGT</b>	CTGGCGAATCAAGCTTTGTAAC	CCGTGATTGCGAAGAGGATATT	(ATT) <sub>12</sub>	M
Satt586/M13	<b>GTAAAACGACGGCCAGT</b>	GCGGCCTCCAACTCCAAGTAT	GCGCCCAAATGATTAATCACTCA	(ATT) <sub>19</sub>	F
Satt612/T7	<b>AATACGACTCACTATAG</b>	GTCATACTGGGTGTTTCATTTATGAC	GCGCCTTTTAGTCTCTGAAAGTATTT	(ATT) <sub>10</sub>	G

diferença mínima de 50 pb entre o limite da faixa amplificável de cada marcador e a quantidade aceitável de artefatos de amplificação de modo a garantir a nitidez na interpretação dos alelos.

A eletroforese capilar foi realizada no sequenciador ABI 3130xl (Applied Biosystems) utilizando o polímero POP7 (Applied Biosystems). A amostra utilizada para a corrida eletroforética constitui-se de 0,4 µL de cada produto de PCR, 8,55 µL de formamida Hi-Di e 0,25 µL do padrão de tamanho molecular 500 ROX de coloração vermelha (Applied Biosystems). Essa solução foi aquecida por 3 min a 95 °C, para desnaturação e imediatamente colocada em gelo de modo a evitar o reanelamento das fitas do DNA. Após esse procedimento, foi realizada a corrida eletroforética. Os dados foram coletados automaticamente e codificados com auxílio do software Gene Mapper 4.0 (Applied Biosystems).

O grupo de cultivares foi genotipado por ambas as técnicas, PCR convencional avaliado em gel de poli-acrilamida e PCR com PCEU avaliados em eletroforese capilar. A genotipagem fluorescente foi realizada em um sistema de prova e contra-prova, desde a reação até a análise dos fragmentos, realizados em dias distintos, com o intuito de avaliar a reprodutibilidade dos dados e determinar o tamanho dos alelos gerados levando em consideração o motivo de cada marcador (dinucleotídeo ou trinucleotídeo).

#### **2.4. Índice de polimorfismo, probabilidade de identidade ao acaso e heterozigosidade observada**

O índice de polimorfismo foi calculado para cada marcador com base na frequência de cada alelo, segundo a expressão matemática que indica o valor de PIC (Conteúdo de Informação Polimórfica), mostrada abaixo:

$$PIC = 1 - \sum_{j=1}^n p_{ij}^2$$

Onde  $p_{ij}$  é a frequência do  $j$ -ésimo alelo para o  $i$ -ésimo marcador (Anderson et al., 1993). Segundo KEIM et al. (1992) o índice de polimorfismo refere-se a diversidade genética de um loco. A estimativa da diversidade genética é um indicador do poder de discriminação genética do loco em um teste de identidade (Weir, 1996).

Foi calculada a probabilidade de identidade ao acaso através do grupo de marcadores utilizados na discriminação inequívoca dos cultivares, conforme descrito por SCHUSTER et al. (2009). Onde:

$$\text{PROBABILIDADE DE IDENTIDADE AO ACASO} = \prod_{i=1}^n P_{ii}$$

Em que  $P_{ii}$  é a frequência do alelo  $A_i$ , e  $n$  o número dos marcadores avaliados.

A partir do perfil alélico molecular apresentado pelos cultivares gerado em cada marcador foi calculada a heterozigosidade observada ( $H_o$ ) como demonstrado abaixo:

$$H_o = 1 - \sum_{i=1}^n P_{ii}$$

Onde  $P_{ii}$  é a frequência observada de genótipos homozigotos para o  $i$ -ésimo alelo.

## 2.5. Relações genéticas entre os cultivares

As relações genéticas entre os genótipos foram avaliadas por meio do complemento do índice de similaridade para variáveis codominantes e multialélicas, a partir do índice não-ponderado utilizado para o cálculo da matriz de similaridade, foi obtida uma matriz de dissimilaridade para avaliação da distância genética entre os cultivares. O índice não-ponderado é dado por:

$$S_{ii'} = \frac{1}{2L} \sum_{j=1}^L c_j$$

em que:

L: número total de locos analisados; e

$C_j$ : número de alelos comuns entre os pares de acessos  $i$  e  $i'$ .

A matriz foi estabelecida por meio de escores com base no número de alelos detectados em cada marcador. A partir da matriz de dissimilaridade foi obtida a correlação cofenética e realizado o agrupamento dos cultivares pelo

método hierárquico UPGMA (*unweighted pair-group using an arithmetic average*) para a construção do dendograma. As análises foram realizadas através do programa Genes versão 2009 (Cruz, 2006).

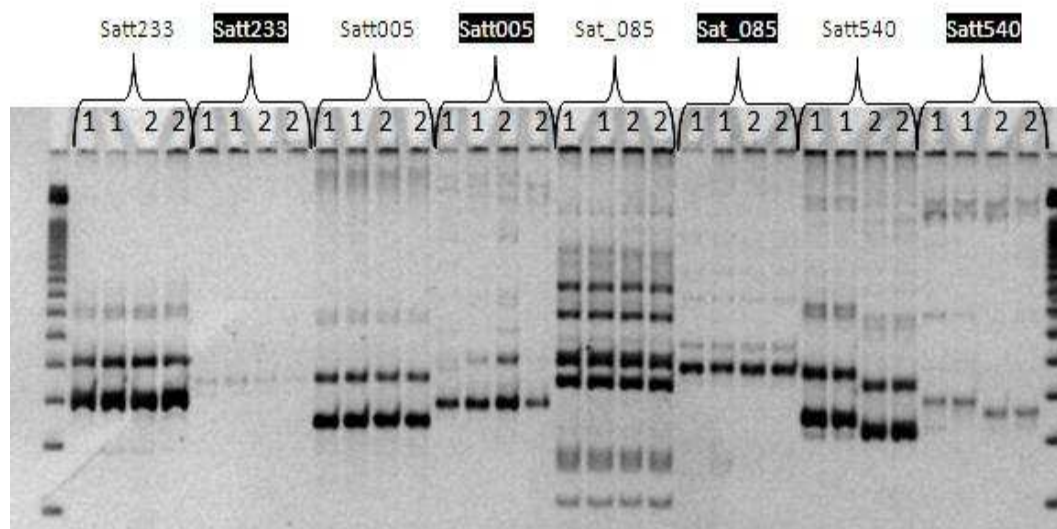
## **2.6. Correlação entre os métodos de genotipagem**

Foi realizada uma análise de correlação entre as matrizes de dissimilaridade com o objetivo de avaliar a similaridade entre a distância genética gerada por ambos os métodos de genotipagem. Foi realizado o teste de Mantel com 10.000 simulações através do programa Genes versão 2009 (Cruz, 2006). Os dendrogramas obtidos a partir das matrizes também foram comparados e avaliados quanto à formação dos grupos de genótipos.

## **3. RESULTADOS E DISCUSSÃO**

### **3.1. Verificação do perfil de amplificação**

A diferença de 17 pb entre os amplicons gerados nas diferentes metodologias foi confirmada (Figura 2), levando em consideração a inclusão da cauda estendida universal, nos fragmentos provenientes da genotipagem no sequenciador semi-automático de DNA. Dos 22 marcadores testados, 14 foram selecionados por apresentarem melhor perfil de amplificação quando genotipados em gel de poliacrilamida, e foram adaptados para a utilização no sistema de genotipagem PCEU. Os 14 marcadores selecionados apresentaram produtos de amplificação mensuráveis, com exceção do marcador Satt586, possivelmente devido à formação de estruturas secundárias quando combinado à cauda estendida. Na Figura 2, são mostrados alguns produtos de amplificação gerados com os *primers* avaliados. Os marcadores foram testados nos dois métodos utilizando os genótipos BMX Magna RR e CD202, os quais foram avaliados em duplicata. Foi observado que amplicons obtidos a partir da metodologia PCEU, embora um pouco menos intensos, apresentaram uma menor quantidade de bandas inespecíficas, sendo o marcador Sat\_085 o que apresentou esta diferença de forma mais significativa.



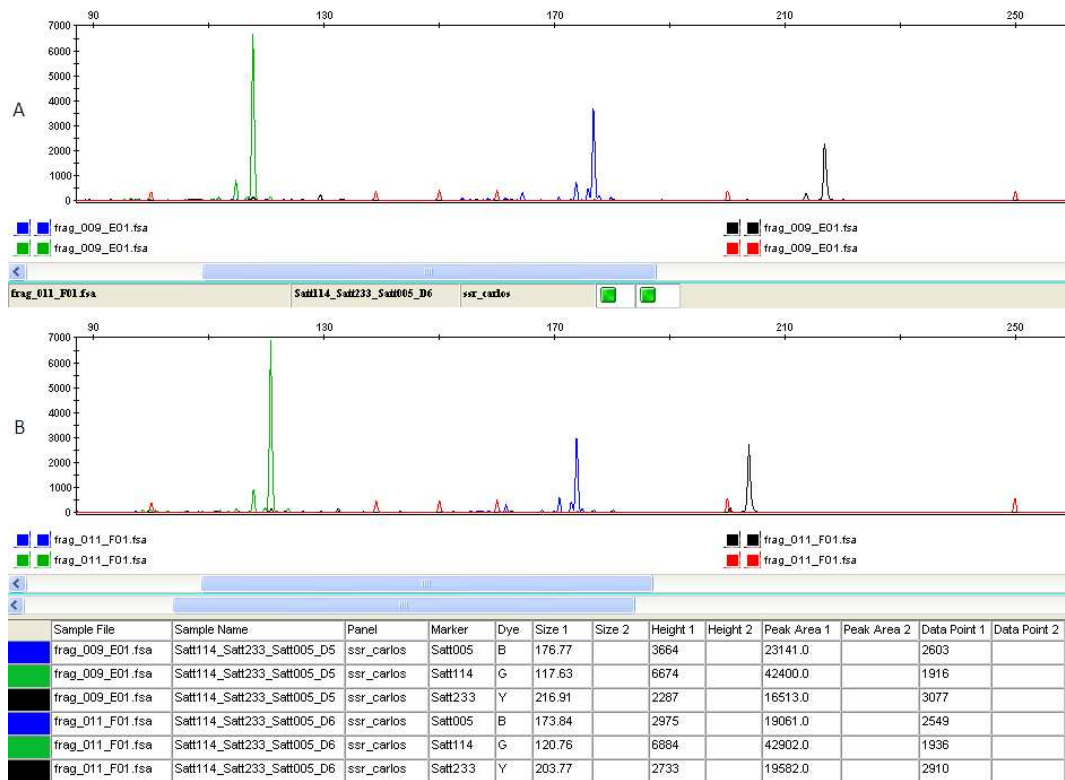
**Figura 2.** Comparação do perfil de amplificação do sistema de genotipagem convencional com o PCEU. Os marcadores estão descritos na porção superior da figura, sendo aqueles destacados em preto os genotipados pelo sistema PCEU. Foram avaliados os cultivares BMX Magna RR, representado pelo número 1, e o CD202, representado pelo número 2, ambos avaliados em duplicata.

A representação da avaliação simultânea com marcadores de diferentes fluorescências e do perfil de amplificação dos genótipos CD202 e CD228 estão mostrados na Figura 3. Esses cultivares foram avaliados com os marcadores Satt005, Satt114 e Satt233, sendo os três *primers* polimórficos entre os genótipos, cujos tamanhos específicos dos fragmentos amplificados estão descritos na figura.

### 3.2. Estimativa do tamanho dos fragmentos

Os tamanhos dos fragmentos estimados para cada marcador em ambos os métodos de genotipagem para os 30 cultivares avaliados estão mostrados na Tabela 2. Na metodologia PCEU os fragmentos foram estimados com base em valores de genotipagem de prova e contra-prova, considerando-se o motivo repetitivo de cada marcador. Foi observado que fragmentos com diferenças de tamanho menor que quatro pares de bases, tiveram sua identificação dificultada quando genotipados em gel de poliacrilamida, resultando em uma alteração no número de alelos detectados nos métodos bem como uma possível alteração da frequência alélica encontrada.

Fragmentos cuja diferença de tamanho de apenas uma unidade de repetição, foram considerados como um mesmo alelo, de acordo com a genotipagem convencional.



**Figura 3.** Representação do perfil de amplificação de genótipos avaliados em sequenciador semi-automático ABI 3130 xl. Os genótipos: CD202 (A) e CD228 (B). Foram avaliados com os marcadores Satt005, Satt114 e Satt233. Os picos verdes representam os alelos 118 (A) e 121 (B), os picos azuis representam os alelos 177 (A) e 174 (B), e os picos pretos representam os alelos 217 (A) e 205 (B).

A título de exemplo, o marcador Satt070 apresentou 3 alelos de acordo com o método PCEU (165pb, 189pb e 192pb), e 2 alelos (150pb e 175pb) quando genotipado pelo método convencional. Os fragmentos de tamanho 189pb e 192pb (PCEU) foram visualizados como sendo um mesmo alelo (175pb) pelo método convencional, e desta forma sua frequência alélica (0,38) correspondeu à soma das frequências dos alelos 189pb (0,23) e 192pb (0,15). O mesmo não ocorreu com marcador Satt233 que apresentou 3 alelos, independente da metodologia de genotipagem, que é explicado pela maior diferença entre os tamanhos dos alelos. O tamanho e frequência dos alelos dos marcadores foram comparados e estão mostrados na Figura 4.

Artefatos de amplificação também podem ter influenciado no surgimento ou desaparecimento de alelos detectáveis nas diferentes metodologias. A *Taq* DNA polimerase, tem a propriedade de adicionar um nucleotídeo (em geral adenosina) à extremidade 3' do fragmento recém sintetizado.

**Tabela 2.** Tamanho estimado dos fragmentos amplificados nos 30 cultivares em cada marcador nas metodologias PCEU e convencional

Genótipos	AF162283		GMABAB		Sat_085	
	PCEU	*Conv.	PCEU	*Conv.	PCEU	*Conv.
BMX Magna RR	249	230	187	170	221	205
BRS133	241	220	166	150	229	215
BRS184	241	220	178	160	219	205
BRS213	249	230	166	150	191	180
BRS214	249	230	166	150	221	205
BRS255RR	249	230	166	150	219	205
BRS256RR	241	220	166	150	219	205
BRS262	241	220	166	150	191	180
BRS270RR	247	230	178	160	191	180
BRScharruaRR	241	220	166	150	229	215
BRSGiseleRR	241	220	178	160	219	205
BRSMG850GRR	241	220	178	160	219	205
BRSPampaRR	241	220	166	150	229	215
BRSMTPintado	249	230	166	150	191	180
BRSValiosaRR	241	220	178	160	219	205
CD202	245	230	175	160	219	205
CD206	249	230	175	160	219	205
CD213RR	241	220	178	160	221	205
CD214RR	249	230	178	160	221	205
CD215	245	220	178	160	219	205
CD217	241	220	178	-	191	180
CD219RR	247	230	178	160	191	180
CD226	245	230	166	150	191	180
CD228	247	230	182	170	191/219	180
CD229RR	247	230	166	150	191	180
CD237RR	241	220	178	160	191	180
CD240RR	245	230	178	160	-	205
MG/BR Conquista	241	220	178	160	219	205
MSoy 7501	241	220	178	160	219	205
Vmax	249	230	166	150	219	205

\* Convencional  
 “-“ Dado perdido

Continua...

**Tabela 2** (cont.). Tamanho estimado dos fragmentos amplificados nos 30 cultivares em cada marcador nas metodologias PCEU e convencional

Genótipos	Satt002		Satt005		Satt045	
	PCEU	*Conv.	PCEU	*Conv.	PCEU	*Conv.
BMX Magna RR	148	140	177	165	151	140
BRS133	142	135	198	190	157	145
BRS184	142/148	135	174	165	157	145
BRS213	148	140	174/177	165	151	140
BRS214	154	145	177	165	151	140
BRS255RR	148/154	135	201	190	151	140
BRS256RR	142/154	135	174	165	157	145
BRS262	142	135	177	165	151	140
BRS270RR	142	135	177	165	151	140
BRSharruaRR	142	135	198	190	-	145
BRSGiseleRR	154	145	174	165	151	140
BRSMG850GRR	154	-	165	155	151/157	140
BRSPampaRR	142	135	198	190	157	145
BRSMTPintado	148	-	177	165	151	140
BRSValiosaRR	154	145	165	155	151/157	140
CD202	142	135	177	165	151	140
CD206	142	135	198	190	151	140
CD213RR	142	135	177	165	157	145
CD214RR	142	135	177	165	151	140
CD215	142	135	177	165	151	140
CD217	148	140	165	155	151	140
CD219RR	142	135	174	165	151	140
CD226	148	140	177	165	151	140
CD228	142	-	174	165	151	140
CD229RR	142	-	156	145	151	140
CD237RR	148	-	177	165	151	140
CD240RR	142	-	177	165	157	145
MG/BR Conquista	154	145	165	155	151	140
MSoy 7501	148	140	183/198	170	157	145
Vmax	142	135	165	155	151	140

\* Convencional  
 “-“ Dado perdido

Continua...

**Tabela 2** (cont.). Tamanho estimado dos fragmentos amplificados nos 30 cultivares em cada marcador nas metodologias PCEU e convencional

Genótipos	Satt070		Satt079		Satt114	
	PCEU	*Conv.	PCEU	*Conv.	PCEU	*Conv.
BMX Magna RR	165	150	166	-	121	110
BRS133	165	150	142	130	109	100
BRS184	165	150	160	145	109	100
BRS213	165/189	150/175	166	150	109	100
BRS214	192	175	166	150	109	100
BRS255RR	165/189	150/175	166	-	94	80
BRS256RR	165	150	160	145	94	80
BRS262	192	175	142	130	109	100
BRS270RR	192	175	166	150	109	100
BRScharruaRR	165	150	160	145	109	100
BRSGiseleRR	165	150	160	145	94	80
BRSMG850GRR	165	150	160	145	94	80
BRSPampaRR	165	150	142	130	109	100
BRSMTPintado	192	175	166	150	109	100
BRSValiosaRR	165	150	160	145	94	80
CD202	165	150	142	130	118	110
CD206	165	150	160	145	94	80
CD213RR	165	150	166	150	109	100
CD214RR	189	175	166	150	121	110
CD215	165	150	142	130	109	100
CD217	189	175	142	130	94	80
CD219RR	189	175	-	130	109	100
CD226	189	175	166	150	109	100
CD228	165/192	150/175	163	150	121	110
CD229RR	165	150	142	130	94	80
CD237RR	189	175	142	130	94	80
CD240RR	165	150	163	150	118	110
MG/BR 46 Conquista	165	150	160	145	94	80
MSoy 7501	189	175	160	145	109	100
Vmax	165	150	166	150	121	110

\* Convencional

“-“ Dado perdido

Continua...

**Tabela 2** (cont.). Tamanho estimado dos fragmentos amplificados nos 30 cultivares em cada marcador nas metodologias PCEU e convencional

Genótipos	Satt233		Satt431		Satt540	
	PCEU	*Conv.	PCEU	*Conv.	PCEU	*Conv.
BMX Magna RR	217	200	251	240	183	170
BRS133	205	185	251	240	186	170
BRS184	205	185	251	240	186	170
BRS213	217/226	200/205	206	190	165	150
BRS214	217	200	251	240	171	160
BRS255RR	205/226	185/205	218	210	165	150
BRS256RR	226	205	206/218	210	186	170
BRS262	217	200	251	240	183	170
BRS270RR	205	185	194/206	190/240	183	170
BRScharruaRR	205	185	251	240	171	170
BRSGiseleRR	205	185	206/218	210/240	171	160
BRSMG850GRR	205/226	185/205	218	210	171	160
BRSPampaRR	217	200	206/251	190/240	165	150
BRSMTPintado	205	185	251	240	183	-
BRSValiosaRR	205	185	218	210	171	160
CD202	217	200	218	210	171	160
CD206	205	185	251	240	186	170
CD213RR	205/226	185/205	206/251	190/240	165	150
CD214RR	217	200	206/218	190/210	183	170
CD215	217	200	251	240	171	160
CD217	217	200	251	240	183	-
CD219RR	217	200	251	240	186	-
CD226	217/226	200	194/206	190	165	150
CD228	205	185	251	240	171	-
CD229RR	205	185	251	240	171	160
CD237RR	217	200	251	240	183	-
CD240RR	217	200	218	210	171	-
MG/BR Conquista	205	185	218	210	171	150
MSoy 7501	217	200	251	210/240	186	-
Vmax	217	200	251	210	171	-

\* Convencional  
 “-“ Dado perdido

Continua...

**Tabela 2** (cont.). Tamanho estimado dos fragmentos amplificados nos 30 cultivares em cada marcador nas metodologias PCEU e convencional

Genótipos	Satt612	
	PCEU	*Conv.
BMX Magna RR	256	245
BRS133	256	245
BRS184	256	245
BRS213	256	245
BRS214	256	245
BRS255RR	256	245
BRS256RR	256	245
BRS262	256	245
BRS270RR	250	240
BRSCharruaRR	250/256	245
BRSGiseleRR	250	240
BRSMG850GRR	256	245
BRSPampaRR	250	240
BRSMTPintado	256	245
BRSValiosaRR	256	245
CD202	265	250
CD206	256	245
CD213RR	250	240
CD214RR	256	245
CD215	265	250
CD217	256	245
CD219RR	250	240
CD226	265	250
CD228	250	240
CD229RR	250	240
CD237RR	256	245
CD240RR	265	250
MG/BR Conquista	256	245
MSoy 7501	256	245
Vmax	265	250

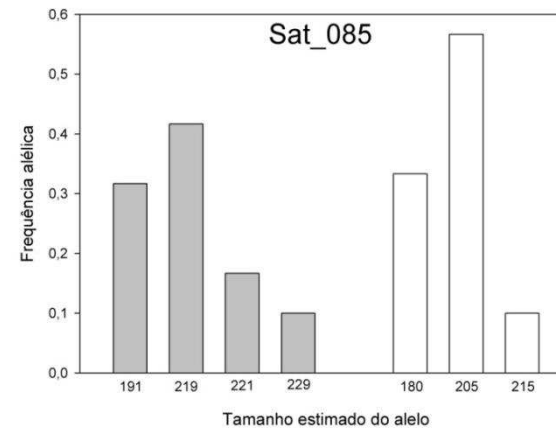
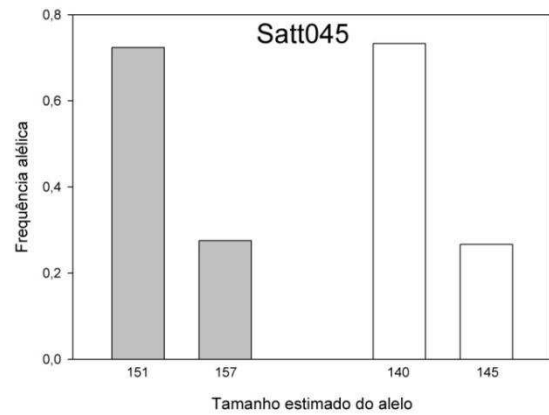
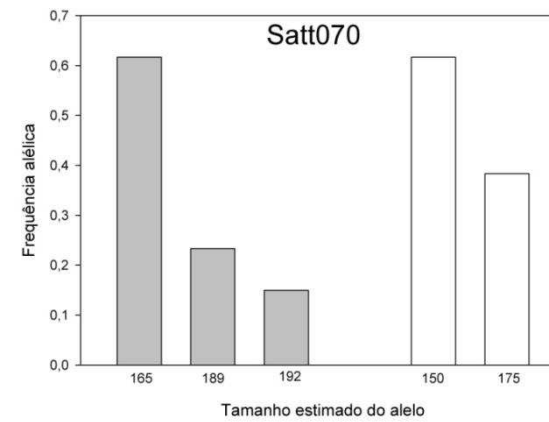
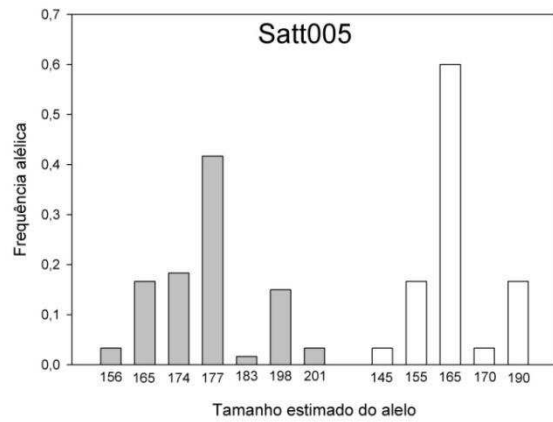
\* Convencional  
 “-“ Dado perdido

Essa adenilação tende a influenciar na determinação exata dos tamanhos dos fragmentos amplificados (Pompanon et al., 2005). Para contornar tal situação um aumento no tempo de extensão final (mais de 20 minutos) no ciclo da PCR proporciona uma maior chance de que todos os fragmentos sejam adenilados, uniformizando desta forma o tamanho dos produtos de PCR gerados, e com isso diminuindo erros provenientes dessa propriedade da *Taq* DNA polimerase.

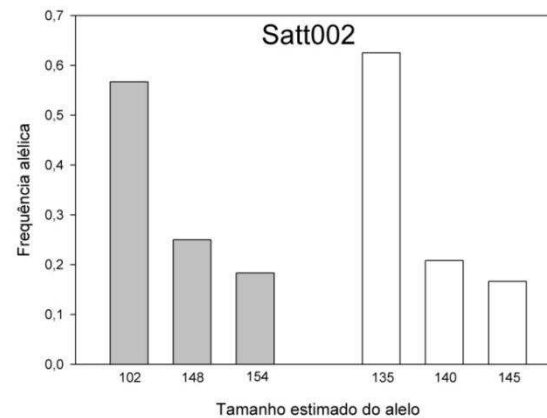
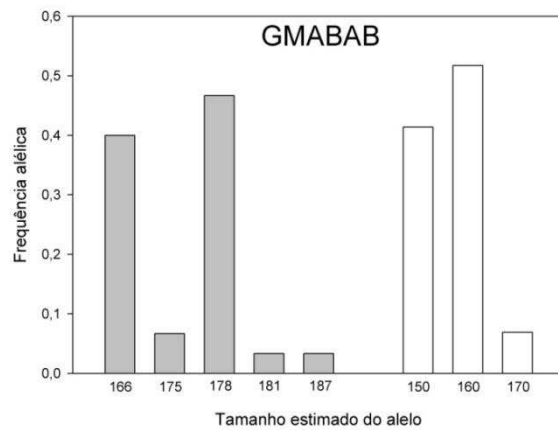
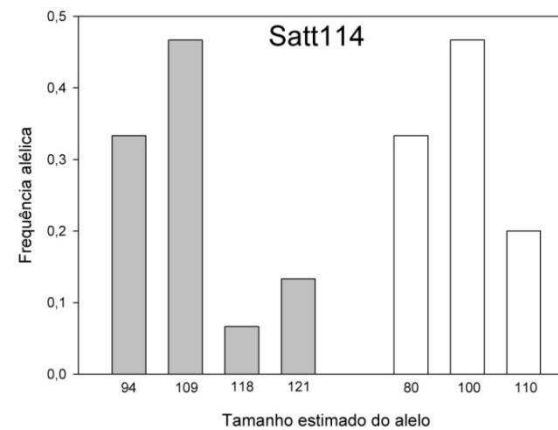
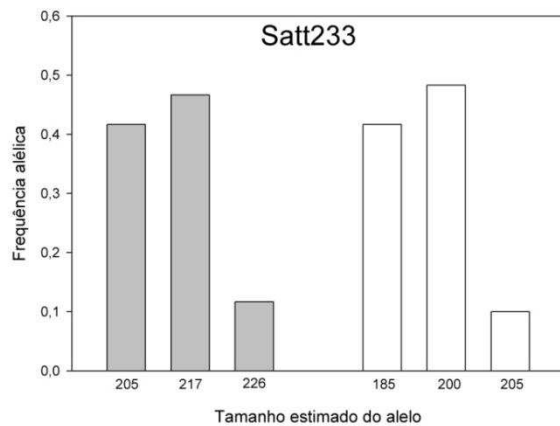
A observação de variações de um par de bases pode também indicar a presença de micro variações devido à ocorrência de mutações pontuais. Entretanto, apesar do acréscimo de tempo na extensão final nas amplificações, alguns locos ainda poderão apresentar variações devido à adição de dATP (+A/-A), (Medeiros, 2006). Além disso, a temperatura ambiente do local onde o sequenciador se encontra afeta significativamente o tamanho estimado do alelo. De acordo com DAVISON & CHIBA (2003), com uma alteração na temperatura do laboratório de 5 °C, o tamanho estimado do alelo pode sofrer variações em até 0,7 pb.

Outro artefato comum que pode influenciar na interpretação dos dados é o *stutter*, que é explicado pelo “escorregamento” (*slippage*) da DNA polimerase durante a replicação produzindo fragmentos de tamanhos próximos ao tamanho do alelo verdadeiro (Hauge & Litt, 1993). Sua magnitude é variável e dependente do loco em avaliação, podendo apresentar dois ou mais picos *stutter* (Dewoody et al., 2006). O *stutter* pode sobrepor alelos heterozigotos e interferir na interpretação do padrão de amplificação gerado (Miller & Yuan, 1997), ou seja, um heterozigoto pode ser classificado como homozigoto, em casos de sobreposição, ou um homozigoto pode ser classificado como heterozigoto nas situações em que o *stutter* se apresente em alta intensidade.

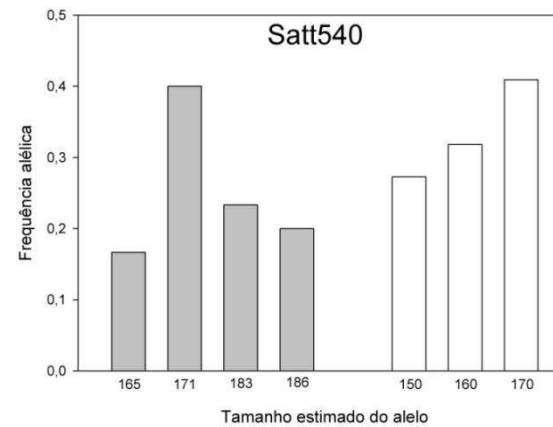
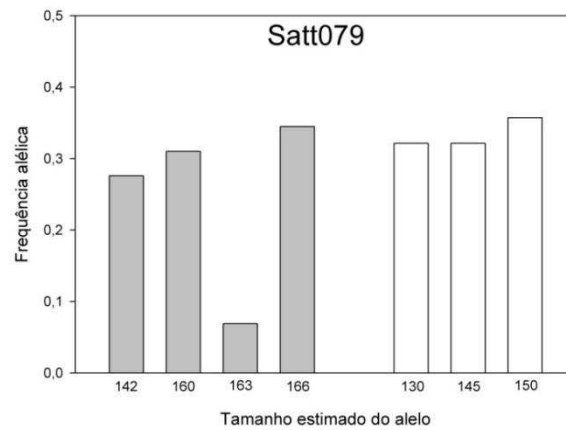
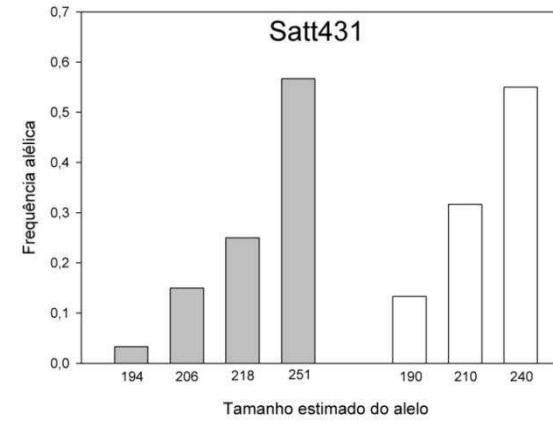
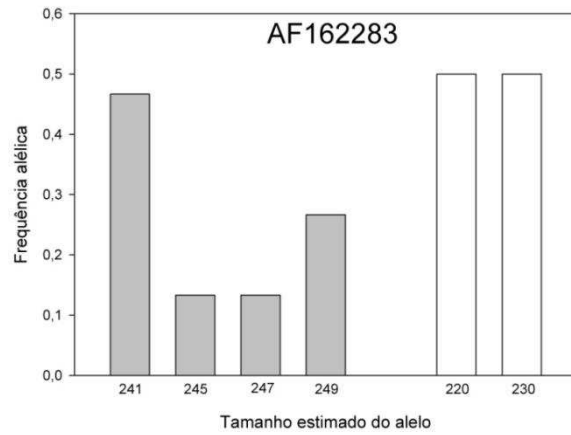
Erros de genotipagem também podem ocorrer devido à amplificação preferencial de alelos menores (Van Oosterhout et al., 2004). Esse artefato, também conhecido como “*dropout*” de alelos, está na maioria das vezes associado a baixa qualidade e/ou quantidade de DNA da amostra e pode levar a uma falsa classificação dos alelos (Pompanon et al., 2005; Hoffman & Amos, 2005; Johnson & Haydon, 2006), isto é, em alguns casos ao se avaliar um indivíduo heterozigoto, na presença desse artefato, o indivíduo poderá ser considerado erroneamente como homozigoto.



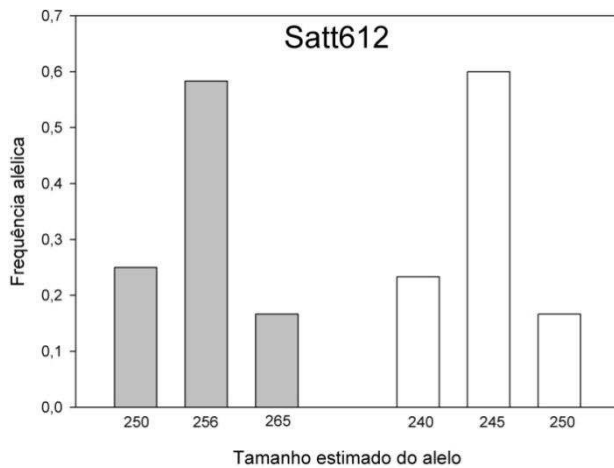
**Figura 4.** Tamanho estimado e frequência dos alelos detectados nos métodos: PCEU (cinza) e convencional (branco) em cada marcador.



**Figura 4** (cont.). Tamanho estimado e frequência dos alelos detectados nos métodos: PCEU (cinza) e convencional (branco) em cada marcador.



**Figura 4** (cont.). Tamanho estimado e frequência dos alelos detectados nos métodos: PCEU (cinza) e convencional (branco) em cada marcador.



**Figura 4** (cont.). Tamanho estimado e frequência dos alelos detectados nos métodos: PCEU (cinza) e convencional (branco) em cada marcador.

### 3.3. Informatividade dos locos

Para a metodologia de genotipagem PCEU, o número total de alelos encontrado foi de 50, variando de 2 a 7 entre os marcadores, sendo o marcador Satt005 o que apresentou maior número de alelos polimórficos. O PIC variou de 0,40 (Satt045) a 0,74 (Satt005), com média de 0,62. Para a metodologia convencional o número de alelos encontrados foi de 38 com uma variação de 2 (Satt045, Satt070 e AF162283) a 5 (Satt005) alelos. O PIC apresentou uma faixa 0,39 (Satt045) a 0,67 (Satt079), com média de 0,56 (Tabela 3).

A diferença apresentada com relação ao número de alelos, valor de PIC e heterozigosidade observada entre as metodologias é explicada pela maior precisão da metodologia aplicada no sequenciador, já que esta possui uma alta sensibilidade na detecção de alelos de tamanhos estreitamente similares (Wenz et al., 1998; Medeiros, 2006). Foi possível a detecção de até dois alelos a mais para um mesmo marcador quando se utilizou o método PCEU (Satt005, GMABAB e AF162283). Da mesma forma, por estar relacionado ao número de alelos detectados, o valor de PIC foi maior no conjunto de marcadores quando genotipados pelo método PCEU.

**Tabela 3.** Número de alelos, conteúdo de informação polimórfica (PIC) e heterozigose observada (H. Obs.) nas diferentes metodologias.

Marcadores	Número de alelos		PIC		Heterozigosidade observada	
	PCEU	*Conv.	PCEU	*Conv.	PCEU	*Conv.
Satt005	7	5	0,74	0,58	0,0667	0,0
Satt045	2	2	0,40	0,39	0,069	0,0
Satt070	3	2	0,54	0,47	0,1	0,1
Sat_085	4	3	0,69	0,56	0,0333	0,0
Satt233	3	3	0,60	0,58	0,1667	0,1333
GMABAB	5	3	0,62	0,56	0,0	0,0
Satt114	4	3	0,65	0,63	0,0	0,0
Satt002	3	3	0,58	0,54	0,0667	0,0
AF162283	4	2	0,68	0,50	0,0	0,0
Satt079	4	3	0,70	0,67	0,0	0,0
Satt431	4	3	0,59	0,58	0,2333	0,2
Satt540	4	3	0,72	0,66	0,0	0,0
Satt612	3	3	0,57	0,56	0,0333	0,0
<b>Média</b>	<b>3,846</b>	<b>2,923</b>	<b>0,62</b>	<b>0,56</b>	<b>0,0592</b>	<b>0,0333</b>

\*Conv. – convencional

SAYAMA et al. (2011) desenvolveram um painel formado por 304 SSR para ampla genotipagem do genoma da soja afim de facilitar o mapeamento genético e avaliaram um grupo de 87 genótipos de soja com 4 marcadores re-selecionados para cada grupo de ligação. Nesse grupo de 80 locos eles obtiveram uma média de 14,8 alelos e um PIC médio de 0,78. SINGH et al. (2010), avaliando 44 genótipos de soja com diferentes respostas ao fotoperíodo com 40 SSR e 5 AFLP, obtiveram um PIC médio de 0,507 e 0,120, respectivamente, e 120 alelos foram detectados para os 34 SSR polimórficos, com média de 3,53 alelos por loco. WANG et al. (2008) avaliaram a diversidade genética de 1383 acessos de germoplasma de soja com 60 SSR e encontraram 1.111 alelos com média de 18,52 por loco. O PIC variou de 0,456 a 0,928 com média de 0,815. MIAN et al. (2008) determinaram a diversidade genética de 48 genótipos de soja

diferindo em sintomas do vírus do mosqueado do feijoeiro em soja (*Bean Pod Mottle Virus*) com 84 SSR, encontrando 271 alelos, com média de 3,23 e com o valor médio de PIC de 0,53. MULATO et al. (2009) estudaram a diversidade genética de 79 acessos de soja de diferentes regiões do mundo com 30 marcadores SSR, sendo 20 genômicos e 10 EST-SSR. Encontraram um total de 259 alelos com um número de alelos por loco que variou de 2 a 21, com média de 8,63. O valor de PIC variou de 0,166 a 0,921, com média de 0,626 e a heterozigosidade observada foi de 0,025.

A maior quantidade de alelos e valor de PIC encontrados por alguns desses trabalhos é justificada por, além do maior número de genótipos avaliados, pela inclusão de acessos de banco de germoplasma nas análises de diversidade (Wang et al. 2008) resultando em uma maior variabilidade genética do material e conseqüentemente aumentando a probabilidade de se encontrar diferentes alelos, o que influencia diretamente no aumento do valor de PIC.

Valores superiores também foram observados ao se comparar a heterozigosidade observada entre os métodos, obtendo-se média de 0,0592 (PCEU) e 0,0333 (convencional). Apesar de serem avaliados os mesmos genótipos com os mesmos marcadores em ambas as metodologias, a diferença na presença de heterozigotos observada está intrinsecamente relacionada com o sistema de detecção utilizado. Devido à soja ser uma espécie autógama o esperado é que todos os indivíduos de um cultivar sejam homozigotos. Entretanto, em alguns cultivares dois alelos puderam ser detectados em alguns locos, o que pode caracterizar, portanto, uma mistura de linhas puras (Oliveira, 2010).

Segundo SCHUSTER et al. (2009), ao avaliarem o poder de detecção de um alelo pouco frequente em um *bulk* de DNA, confirmaram que a medida que se aumenta o número de diferentes indivíduos para a detecção simultânea, a chance de amplificação do alelo menos frequente é diminuída. Em uma proporção de até 1:7 SCHUSTER et al. (2004) identificaram, sem dificuldades, em análises de determinação de pureza varietal, dois alelos presentes em um *bulk* de DNA.

CRYER et al. (2005), utilizando um sistema de detecção de locos microssatélites, baseado na ligação de amplicons em um vetor de clonagem para re-amplificação com *primers* fluorescentes, demonstraram que um homozigoto possuindo um alelo raro pode ser detectado com confiabilidade em um conjunto de até 12 indivíduos.

GIANCOLA et al. (2002) sugerem que 5 indivíduos são suficientes para representação de um genótipo com poucas chances de erro na determinação do perfil alélico. Desta forma, ao realizar a análise simultânea de 50 indivíduos para representar o genótipo, para um alelo ser detectável em um *bulk* em análise através da metodologia convencional, segundo SCHUSTER et al. (2004), pelo menos oito indivíduos deverão apresentar esse alelo em seu genoma. Já quando avaliado em sequenciador, devido à sensibilidade do sistema de detecção, mesmo uma pequena quantidade de produto de PCR amplificado pode ser mais facilmente detectável pelo equipamento, e além disso, sua melhor separação eletroforética dos amplicons permite uma identificação mais precisa de fragmentos com pequenas diferenças de tamanho, justificando o maior número de alelos e heterozigotos encontrados pelo método PCEU.

#### **3.4. Frequência alélica e probabilidade de identidade ao acaso**

A partir da genotipagem da população, foi calculada a frequência alélica para cada marcador e a probabilidade de identidade ao acaso (Tabela 4 e 5). Os valores de probabilidade de identidade ao acaso diferiram entre os métodos, sendo <0,001% (PCEU) e <0,01% (convencional). Essa diferença está relacionada com precisão do método PCEU em detectar diferentes alelos.

Segundo SCHUSTER et al. (2009) pode-se calcular, através do produto das frequências alélicas de um indivíduo em cada marcador, a probabilidade de identidade ao acaso, que remete a probabilidade de dois indivíduos distintos, tomados ao acaso na população, serem iguais. O cálculo para situações em que os genótipos apresentam-se como heterozigotos foi realizado através da soma das frequências dos dois alelos encontrados no loco em questão, de modo a considerar a possível presença de um ou do outro alelo em uma análise de determinação de identidade. Sendo assim, para o conjunto dos 13 marcadores microssatélites utilizados, a probabilidade de um cultivar tomado aleatoriamente na população ser igual a outro, é menor que 0,00001 (PCEU), 0,0001 (convencional), ou seja, ao se comparar duas amostras com os 13 marcadores, e elas apresentarem o mesmo perfil alélico, pode-se dizer que essas amostras são idênticas a uma probabilidade de maior que 99,999% ( $1 - 0,00001$ ) pelo método PCEU e 99,99% ( $1 - 0,0001$ ) pelo método convencional. Esses valores obtidos (99,999% e 99,99%) são o poder de exclusão da técnica.

**Tabela 4.** Frequência alélica e probabilidade de identidade ao acaso (PI) dos 30 cultivares em 13 marcadores microssatélites pelo método PCEU

Genótipos	Satt005	Satt045	Satt070	Sat_085	Satt233	GMABAB	Satt114	Satt002	AF162283	Satt079	Satt431	Satt540	Satt612	PI
BMX Magna RR	0,417	0,724	0,617	0,167	0,467	0,033	0,133	0,250	0,267	0,345	0,567	0,241	0,583	<0,001%
BRS133	0,150	0,276	0,617	0,100	0,417	0,400	0,467	0,567	0,467	0,276	0,567	0,207	0,583	<0,001%
BRS184	0,183	0,276	0,617	0,417	0,417	0,467	0,467	0,817	0,467	0,310	0,567	0,207	0,583	<0,001%
BRS213	0,600	0,724	0,850	0,317	0,583	0,400	0,467	0,250	0,267	0,345	0,150	0,172	0,583	<0,001%
BRS214	0,417	0,724	0,150	0,167	0,467	0,400	0,467	0,183	0,267	0,345	0,567	0,379	0,583	<0,001%
BRS255RR	0,033	0,724	0,850	0,417	0,533	0,400	0,333	0,567	0,267	0,345	0,250	0,172	0,583	<0,001%
BRS256RR	0,183	0,276	0,617	0,417	0,117	0,400	0,333	0,750	0,467	0,310	0,400	0,207	0,583	<0,001%
BRS262	0,417	0,724	0,150	0,317	0,467	0,400	0,467	0,567	0,467	0,276	0,567	0,241	0,583	<0,001%
BRS270RR	0,417	0,724	0,150	0,317	0,417	0,467	0,467	0,567	0,133	0,345	0,183	0,241	0,250	<0,001%
BRSCharruaRR	0,150	0,724	0,617	0,100	0,417	0,400	0,467	0,567	0,467	0,310	0,567	0,379	0,833	<0,001%
BRSGiseleRR	0,183	0,724	0,617	0,417	0,417	0,467	0,333	0,183	0,467	0,310	0,400	0,379	0,250	<0,001%
BRSMG850GRR	0,167	1,000	0,617	0,417	0,533	0,467	0,333	0,183	0,467	0,310	0,250	0,379	0,583	<0,001%
BRSPampaRR	0,150	0,276	0,617	0,100	0,467	0,400	0,467	0,567	0,467	0,276	0,717	0,172	0,250	<0,001%
BRSMTPintado	0,417	0,724	0,150	0,317	0,417	0,400	0,467	0,250	0,267	0,345	0,567	0,241	0,583	<0,001%
BRSValiosaRR	0,167	1,000	0,617	0,417	0,417	0,467	0,333	0,183	0,467	0,310	0,250	0,379	0,583	<0,001%
CD202	0,417	0,724	0,617	0,417	0,467	0,067	0,067	0,567	0,133	0,276	0,250	0,379	0,167	<0,001%
CD206	0,150	0,724	0,617	0,417	0,417	0,067	0,333	0,567	0,267	0,310	0,567	0,207	0,583	<0,001%
CD213RR	0,417	0,276	0,617	0,167	0,533	0,467	0,467	0,567	0,467	0,345	0,717	0,172	0,250	<0,001%
CD214RR	0,417	0,724	0,233	0,167	0,467	0,467	0,133	0,567	0,267	0,345	0,400	0,241	0,583	<0,001%
CD215	0,417	0,724	0,617	0,417	0,467	0,467	0,467	0,567	0,133	0,276	0,567	0,379	0,167	<0,001%
CD217	0,167	0,724	0,233	0,317	0,467	0,467	0,333	0,250	0,467	0,276	0,567	0,241	0,583	<0,001%
CD219RR	0,183	0,724	0,233	0,317	0,467	0,467	0,467	0,567	0,133	1,000	0,567	0,207	0,250	<0,001%
CD226	0,417	0,724	0,233	0,317	0,583	0,400	0,467	0,250	0,133	0,345	0,183	0,172	0,167	<0,001%
CD228	0,183	0,724	0,767	0,733	0,417	0,033	0,133	0,567	0,133	0,069	0,567	0,379	0,250	<0,001%
CD229RR	0,033	0,724	0,617	0,317	0,417	0,400	0,333	0,567	0,133	0,276	0,567	0,379	0,250	<0,001%
CD237RR	0,417	0,724	0,233	0,317	0,467	0,467	0,333	0,250	0,467	0,276	0,567	0,241	0,583	<0,001%
CD240RR	0,417	0,276	0,617	0,167	0,467	0,467	0,067	0,567	0,133	0,069	0,250	0,379	0,167	<0,001%
MG/BR Conquista	0,167	0,724	0,617	0,417	0,417	0,467	0,333	0,183	0,467	0,310	0,250	0,379	0,583	<0,001%
MSoy 7501	0,167	0,276	0,233	0,417	0,467	0,467	0,467	0,250	0,467	0,310	0,567	0,207	0,583	<0,001%
Vmax	0,167	0,724	0,617	0,417	0,467	0,400	0,133	0,567	0,267	0,345	0,567	0,379	0,167	<0,001%

**Tabela 5.** Frequência alélica e probabilidade de identidade ao acaso (PI) dos 30 cultivares em 13 marcadores microssatélites pelo método convencional

Genótipos	Satt005	Satt045	Satt070	Sat_085	Satt233	GMABAB	Satt114	Satt002	AF162283	Satt079	Satt431	Satt540	Satt612	PI
BMX Magna RR	0,600	0,733	0,617	0,567	0,483	0,069	0,200	0,208	0,500	1,000	0,550	0,409	0,600	<0,01%
BRS133	0,167	0,267	0,617	0,100	0,417	0,414	0,467	0,625	0,500	0,321	0,550	0,409	0,600	<0,01%
BRS184	0,600	0,267	0,617	0,567	0,417	0,517	0,467	0,625	0,500	0,321	0,550	0,409	0,600	<0,01%
BRS213	0,600	0,733	1,000	0,333	0,583	0,414	0,467	0,208	0,500	0,357	0,133	0,273	0,600	<0,01%
BRS214	0,600	0,733	0,383	0,567	0,483	0,414	0,467	0,167	0,500	0,357	0,550	0,318	0,600	<0,01%
BRS255RR	0,167	0,733	1,000	0,567	0,517	0,414	0,333	0,625	0,500	1,000	0,317	0,273	0,600	<0,01%
BRS256RR	0,600	0,267	0,617	0,567	0,100	0,414	0,333	0,625	0,500	0,321	0,317	0,409	0,600	<0,01%
BRS262	0,600	0,733	0,383	0,333	0,483	0,414	0,467	0,625	0,500	0,321	0,550	0,409	0,600	<0,01%
BRS270RR	0,600	0,733	0,383	0,333	0,417	0,517	0,467	0,625	0,500	0,357	0,683	0,409	0,233	<0,01%
BRSCharruaRR	0,167	0,267	0,617	0,100	0,417	0,414	0,467	0,625	0,500	0,321	0,550	0,409	0,600	<0,01%
BRSGiseleRR	0,600	0,733	0,617	0,567	0,417	0,517	0,333	0,167	0,500	0,321	0,867	0,318	0,233	<0,01%
BRSMG850GRR	0,167	0,733	0,617	0,567	0,517	0,517	0,333	1,000	0,500	0,321	0,317	0,318	0,600	<0,01%
BRSPampaRR	0,167	0,267	0,617	0,100	0,483	0,414	0,467	0,625	0,500	0,321	0,683	0,273	0,233	<0,01%
BRSMTPintado	0,600	0,733	0,383	0,333	0,417	0,414	0,467	0,625	0,500	0,357	0,550	0,410	0,600	<0,01%
BRSValiosaRR	0,167	0,733	0,617	0,567	0,417	0,517	0,333	0,167	0,500	0,321	0,317	0,318	0,600	<0,01%
CD202	0,600	0,733	0,617	0,567	0,483	0,517	0,200	0,625	0,500	0,321	0,317	0,318	0,167	<0,01%
CD206	0,167	0,733	0,617	0,567	0,417	0,517	0,333	0,625	0,500	0,321	0,550	0,409	0,600	<0,01%
CD213RR	0,600	0,267	0,617	0,567	0,517	0,517	0,467	0,625	0,500	0,357	0,683	0,273	0,233	<0,01%
CD214RR	0,600	0,733	0,383	0,567	0,483	0,517	0,200	0,625	0,500	0,357	0,450	0,409	0,600	<0,01%
CD215	0,600	0,733	0,617	0,567	0,483	0,517	0,467	0,625	0,500	0,321	0,550	0,318	0,167	<0,01%
CD217	0,167	0,733	0,383	0,333	0,483	1,000	0,333	0,208	0,500	0,321	0,550	1,000	0,600	<0,01%
CD219RR	0,600	0,733	0,383	0,333	0,483	0,517	0,467	0,625	0,500	0,321	0,550	1,000	0,233	<0,01%
CD226	0,600	0,733	0,383	0,333	0,483	0,414	0,467	0,208	0,500	0,357	0,133	0,273	0,167	<0,01%
CD228	0,600	0,733	1,000	0,333	0,417	0,069	0,200	1,000	0,500	0,357	0,550	1,000	0,233	<0,01%
CD229RR	0,033	0,733	0,617	0,333	0,417	0,414	0,333	1,000	0,500	0,321	0,550	0,318	0,233	<0,01%
CD237RR	0,600	0,733	0,383	0,333	0,483	0,517	0,333	0,625	0,500	0,321	0,550	0,410	0,600	<0,01%
CD240RR	0,600	0,267	0,617	0,567	0,483	0,517	0,200	1,000	0,500	0,357	0,317	1,000	0,167	<0,01%
MG/BR Conquista	0,167	0,733	0,617	0,567	0,417	0,517	0,333	0,167	0,500	0,321	0,317	0,273	0,600	<0,01%
MSoy 7501	0,033	0,267	0,383	0,567	0,483	0,517	0,467	0,208	0,500	0,321	0,867	1,000	0,600	<0,01%
Vmax	0,167	0,733	0,617	0,567	0,483	0,414	0,200	0,625	0,500	0,357	0,317	1,000	0,167	<0,01%

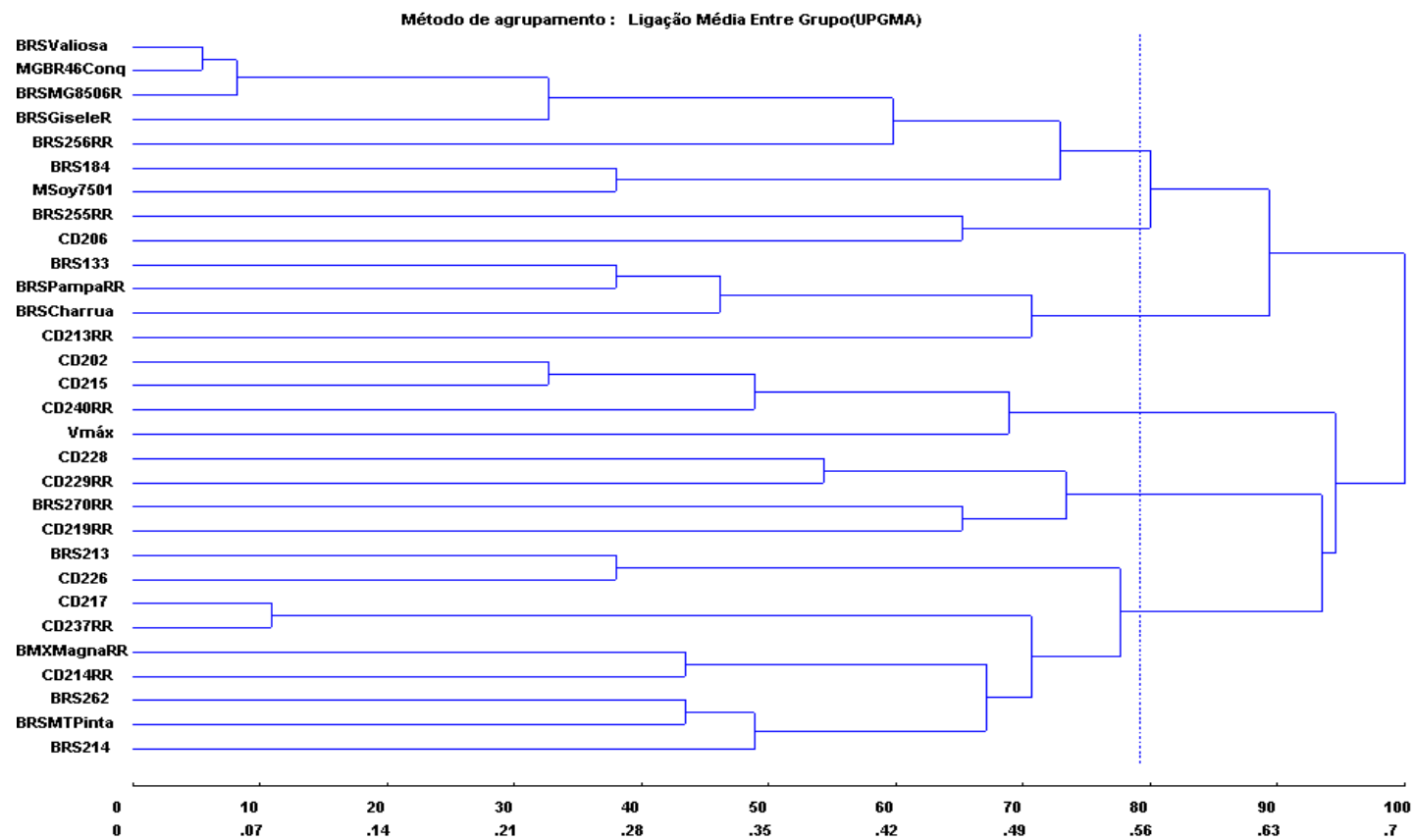
### **3.5. Análise de correlação de matrizes de dissimilaridade e agrupamento entre os cultivares**

A análise de correlação entre as matrizes de dissimilaridade dos métodos avaliados foi significativa pelo Teste de Mantel e pelo Teste t a 1% de significância com correlação de 0,8026. A alta similaridade era esperada já que foram avaliados os mesmos indivíduos com o mesmo grupo de marcadores. Os valores de correlação não foram superiores ao encontrado devido as diferenças atribuídas às características de cada método de genotipagem, como a detecção do número de alelos, e heterozigosidade observada.

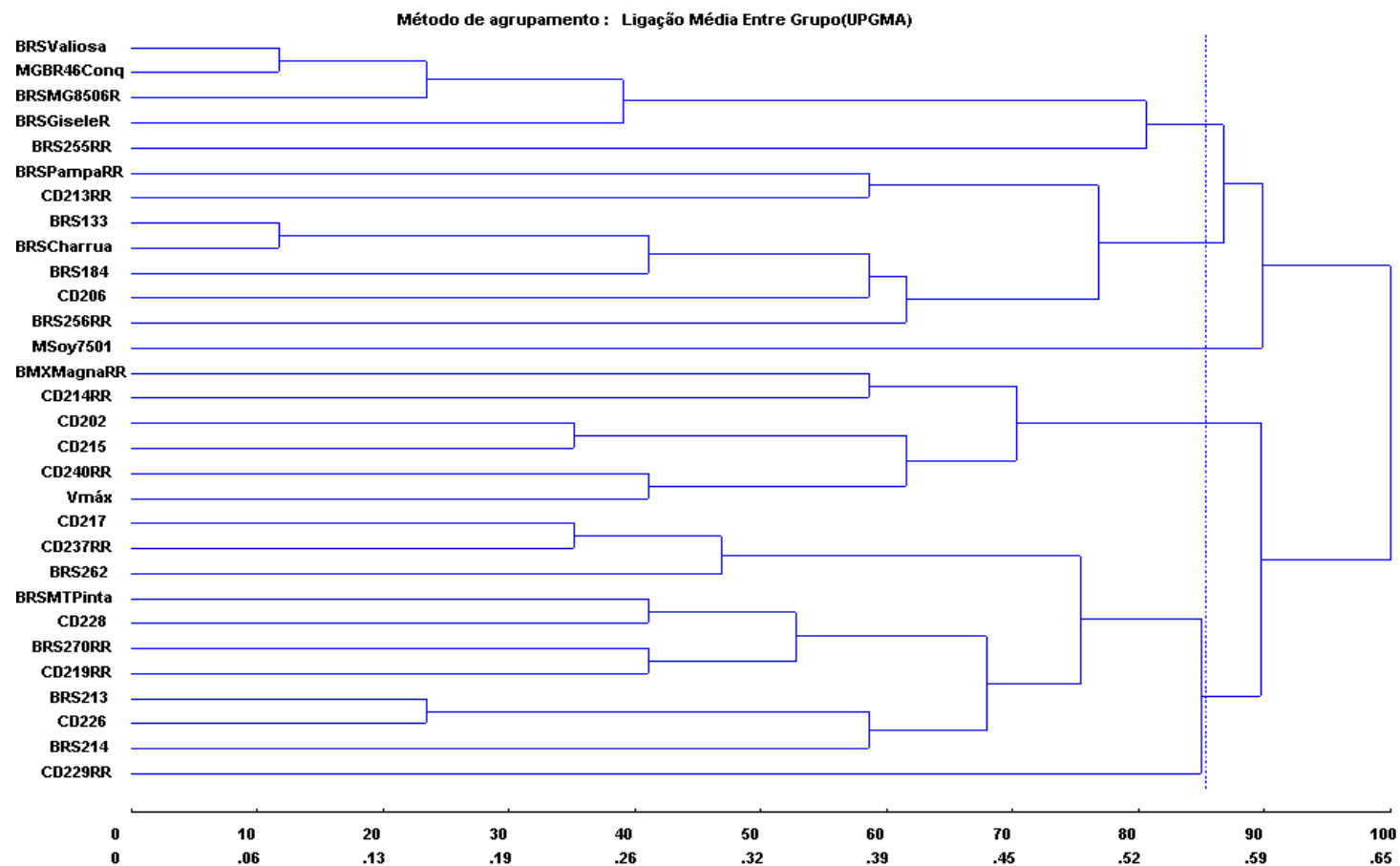
Os dendogramas obtidos a partir das matrizes também foram comparados e avaliados quanto a formação dos grupos de genótipos (Figura 5 e 6). A partir do conjunto de marcadores utilizados, todos os genótipos puderam ser diferenciados, até aqueles estreitamente próximos como os cultivares MGBR46 Conquista e BRS Valiosa RR.

Com uma correlação cofenética de 0,7431, distorção de 2,3497% e estresse de 15,3286% em um ponto de corte de 79,06% no métodos PCEU, cinco grupos foram discriminados. No método convencional a correlação cofenética foi de 0,7071 com valores de distorção e estresse de 3,0424% e 17,4424% respectivamente, cinco grupos foram discriminados no ponto de corte de 85,36%. Os grupos formados, apesar de diferenças pontuais, tiveram alta similaridade entre os métodos e foram concordantes com dados fenotípicos de registro varietal e genealogia.

O cultivar BRS Valiosa RR é essencialmente derivado do cultivar MGBR46 Conquista. O cultivar BRSMG 850G RR foi protegido como sendo mais parecido com o cultivar BRS Valiosa RR. O cultivar BRS Pampa RR foi obtido a partir de cruzamento triplo entre Embrapa 61, E96-246 e BRS 133, seguido de retrocruzamento com BRS 133. O cultivar BRS Charrua RR foi protegido como sendo mais parecido com BRS133. O cultivar CD 215 RR foi protegido como sendo mais parecido com CD 202. O cultivar Vmax (NK 7059 RR) foi protegido também como sendo mais parecido com CD 202. Os cultivares BRS 213 e CD 226 foram protegidos como sendo mais parecidos com o cultivar CD 201. O cultivar CD 237 RR é essencialmente derivado do cultivar CD 217 (MAPA, 2011).



**Figura 5.** Agrupamento dos cultivares pelo método UPGMA, através da genotipagem pela metodologia PCEU.



**Figura 6.** Agrupamento dos cultivares pelo método UPGMA, através da genotipagem pela metodologia convencional.

Na tabela 6 está representada a formação dos grupos em ambos os métodos de genotipagem. Apesar de pequena, a existência de diferenças no arranjo do agrupamento entre os métodos (PCEU e convencional), pode influenciar no estudo da diversidade genética. A escolha dos genitores e a caracterização da variabilidade genética existente são fundamentais para o incremento da eficiência em programas de melhoramento (Brammer, 2002). Para um aumento no ganho genético é preciso variabilidade genética na população proveniente do cruzamento e a informação proveniente do estudo da diversidade genética gera um indicativo de grande utilidade na escolha adequada dos genitores. Desse modo, a classificação dos genitores e a realização de cruzamento entre tipos geneticamente distintos podem contribuir para a ampliação da variância genética em populações segregantes (Messmer et al.,1993).

**Tabela 6.** Formação dos grupos provenientes do agrupamento pelo método UPGMA conforme representado na figura 5, a partir do ponto de corte de 79% (PCEU) e 85% (convencional). Genótipos destacados em preto são os que se mantiveram em grupos similares em ambos os métodos.

PCEU				
Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
BRSValiosaRR MG/BRConquista BRSMG850GRR BRSGiseleRR BRS255RR BRS184 CD206 BRS256RR Msoy7501	BRS133 BRSPampaRR BRSharruaRR CD213RR	CD202 CD215 CD240RR V <sub>max</sub>	CD228 CD229RR BRS270RR CD219RR	BRS213 CD226 CD217 CD237RR BRS262 BRSMTPintado BRS214 BMXMagnaRR CD214RR
Convencional				
Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
BRSValiosaRR MG/BRConquista BRSMG850GRR BRSGiseleRR BRS255RR	BRS133 BRSPampaRR BRSharruaRR CD213RR BRS184 CD206 BRS256RR	CD202 CD215 CD240RR V <sub>max</sub> BMXMagnaRR CD214RR	Msoy7501	BRS213 CD226 CD217 CD237RR BRS262 BRSMTPintado BRS214 CD228 CD229RR BRS270RR CD219RR

Devido à alta similaridade fenotípica, cada vez mais evidente, proporcionada pelo estreitamento da base genética de espécies superiores, marcadores moleculares

vêm se tornando importantes ferramentas de caracterização por serem de rápida e fácil avaliação e por contribuírem para uma distinção mais acurada entre cultivares.

GIANCOLA et al. (2002), avaliaram a contribuição relativa de marcadores moleculares RAPD, AFLP, SSR e marcadores morfológicos para gerar informações sobre as relações genéticas entre cultivares de soja visando sua caracterização. O agrupamento de cultivares através dos dados provenientes de microssatélites mostraram uma maior similaridade com o perfil gerado com base em informações de pedigree de qualquer outro marcador. Os resultados por eles encontrados sugerem que as quatro técnicas descrevem a variabilidade genética de forma diferente e específica, e que a combinação da informação dos microssatélites com os descritores morfológicos resultaram em um agrupamento altamente similar comparado ao agrupamento gerado por dados de pedigree, verificando que os dois sistemas juntos (SSR e morfológico) fornecem informações diferentes e complementares. Portanto esses marcadores podem ser efetivamente aplicados para a classificação de germoplasma e proteção de propriedade intelectual.

#### **4. CONSIDERAÇÕES FINAIS**

Foi possível distinguir, por análise molecular, através do grupo de 13 marcadores microssatélites selecionados por ambos os métodos de genotipagem, todos os cultivares em análise, inclusive aqueles estreitamente relacionados, sendo estes dados concordantes com a genealogia e proteção de cultivar.

A metodologia PCEU se mostrou mais eficiente em detectar fragmentos de tamanhos próximos, comparada ao sistema de genotipagem em gel de poliacrilamida, resultando em uma avaliação mais precisa e confiável dos alelos presentes na população e dos artefatos de amplificação.

O baixo custo quando comparado a técnicas de marcação por fluorescência em cada marcador e sua alta precisão faz do método PCEU, um método vantajoso para a caracterização de cultivares e determinação da pureza genética.

O grupo de marcadores apresentados neste trabalho pode ser usado para a obtenção de descritores genotípicos complementares aos descritores fenotípicos obrigatórios para fim de proteção de cultivares em testes de distinguibilidade.

## 5. REFERENCIAS

- Anderson, J. A.; Churchill, G. A.; Autrique, J. E.; Tanksley, S. D.; Sorrells M. E. Optimizing parental selection for genetic linkage maps. **Genome**, v.36, p.181-186, 1993.
- Arruda, M. P.; Gonçalves, E. C.; Schneider, M. P. C.; Silva, A. L. C.; Morielle-Versute, E. An alternative genotyping method using dye-labeled universal primer to reduce unspecific amplifications. **Molecular Biology Reports**, v.37, p.2031-2036, 2010.
- Akkaya, M. S.; Shoemaker, R. C.; Specht, J. E.; Bhagwat, A. A.; Cregan, P. B. Integration of simple sequence repeat (SSR) DNA markers into a soybean linkage map. **Crop Science**, v.35, p.1439-1445, 1995.
- Borém, A. & Caixeta, E. T. **Marcadores Moleculares**. 2 ed. Viçosa/MG: UFV, 2010.
- Brammer, S. P. Variabilidade e diversidade genética vegetal: requisito fundamental em um programa de melhoramento. **Documento online EMBRAPA Trigo**, Passo Fundo, Disponível em: <[http://www.cnpt.embrapa.br/biblio/p\\_do29.htm](http://www.cnpt.embrapa.br/biblio/p_do29.htm)>. 6 p., no.29, 2006.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento – MAPA. **CultivarWeb**. Disponível em: <<http://www.agricultura.gov.br>>. Acesso em: abril/2011.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento – MAPA. **Instrução para execução dos ensaios de distinguibilidade, homogeneidade e estabilidade de cultivares de soja**. Disponível em: <<http://www.agricultura.gov.br>>. Acesso em: abril/2011.
- Brondani R. P. V. & Brondani, C. Aplicação de tecnologias genômicas baseadas em marcadores microssatélites para discriminação de cultivares e análise de pureza genética em feijoeiro comum. **Comunicado técnico EMBRAPA Arroz e feijão**, Santo Antônio de Goiás, 4 p. no.132, 2006.

- Caixeta, E. T.; Oliveira, A. C. B.; Brito, G. G.; Sakiyama, N. S. Tipos de Marcadores Moleculares. In: **Marcadores Moleculares**. 2 ed. Viçosa-MG, p.011-094, 2010.
- Carvalho, S. M. P. & Pessanha, L. D. R. Propriedade intelectual, estratégias empresariais e mecanismos de apropriação econômica do esforço de inovação no mercado brasileiro de sementes. **Revista de Economia Contemporânea**, v.5, n.1, p. 151-182, 2001.
- Collard, B. C. Y.; Jahufer, M. Z. Z.; Brouwer, J. B.; Pang, E. C. K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. **Euphytica** v.142, p.169-196, 2005.
- COMPANHIA NACIONAL DE ABASTECIMENTO - CONAB. Acompanhamento da Safra Brasileira: Grãos - Safra 2010/2011 - Oitavo Levantamento Disponível em: <<http://www.conab.gov.br/>>. Acesso em Maio/2011.
- Cregan, P. B.; Bhagwat, A. A.; Akkaya, M. S.; Rongwen, J. Microsatellite fingerprinting and mapping of soybean. **Methods in Molecular and Cellular Biology**, v.5 p.49-61, 1994.
- Cregan, P. B.; Jarvik, T.; Bush, A. L.; Shoemaker, R. C.; Lark, K. G.; Kahler, A. L.; Vantoai, T. T.; Lohnes, D. G.; Chung, J.; Specht, J. E. An integrated genetic linkage map of the soybean. **Crop Science**, v.39, p.1464-1490, 1999.
- Cruz, C. D. **Programa GENES - Versão Windows. Aplicativo Computacional em Genética e Estatística**. Viçosa: UFV, 648p., 2001.
- Cryer, N. C.; Butler, D. R.; Wilkinson, M. J. High throughput, high resolution selection of polymorphic microsatellite loci for multiplex analysis. **Plant Methods**, v.1:3, 2005.
- Davison, A. & Chiba, S. Laboratory temperature variation is a previously unrecognized source of genotyping error during capillary electrophoresis. **Molecular Ecology Notes**, v.3, p.321-323, 2003.

- Dewoody, J.; Nason, J. D.; HIPKINS, V. D. Mitigating scoring errors in microsatellite data from wild populations. **Molecular Ecology Notes**, v.6, p.951-957, 2006.
- Diniz, F. M.; Iyengar, A.; Lima, P. S. C.; Maclean, N.; Bentzen, P. Application of a double-enrichment procedure for microsatellite isolation and the use of tailed primers for high throughput genotyping. **Genetics and Molecular Biology**, v.30, p.380-384, 2007.
- Diwan, N. & Cregan, P. B. Automated sizing of fluorescent labeled simple sequence repeat markers to assay genetic variation in soybean. **Theoretical and Applied Genetics**. v.95, p.723-733, 1997.
- Dresler-Nurmi, A.; Terefework, Z.; Kaijalainen, S.; Lindström, K.; Hatakka, A. Silver stained polyacrylamide gels and fluorescence-based automated capillary electrophoresis for detection of amplified fragment length polymorphism patterns obtained from white-rot fungi in the genus *Trametes*. **Journal of Microbiological Methods**, v.41, p.161-172, 2000.
- Garcia, A.F.; Alberini, J.L.; Zucchi, M.I.; De Souza, A.P. Microsatellite molecular markers in the cultivar identification of Brazilian soybean for human consumption. **Crop Breeding and Applied Biotechnology**, v.7, p.155-164, 2007.
- Giancola, S.; Poltri, S. M.; Lacaze, P.; Hopp, H. E. Feasibility of integration of molecular markers and morphological descriptors in a real case study of a plant variety protection system for soybean. **Euphytica**, v.127, p.95-113, 2002.
- Hauge X. Y. & Litt, M. A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. **Human Molecular Genetics**, v.4, p.411-415, 1993.
- Hayden, M. J.; Nguyen, T. M.; Waterman, A.; Chalmers, K. J. Multiplex-Ready PCR: A new method for multiplexed SSR and SNP genotyping. **BMC Genomics**, v.9:80, 2008.

- Hoffman, J. I. & Amos, W. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. **Molecular Ecology**, v.14, p.599-612, 2005.
- Johnson, P. C. & Haydon, D. T. Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. **Genetics**, v.175, p.827-842, 2007.
- Keim, P.; Beavis, W.; Schupp, J.; Freestone, R. Evaluation of soybean RFLP markers diversity in adapted germplasm. **Theoretical and Applied Genetics**, v.85, p.205-212, 1992.
- Lazaruk, K.; Walsh, P. S.; Oaks, F.; Gilbert, D.; Rosenblum, B. B.; Menchen, S.; Scheibler, D.; Wenz, H. M.; Holt, C.; Wallin, J. Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. **Electrophoresis**, v.19, p.86-93, 1998.
- Marcelino, F. C.; Guimarães, M. F. M.; Barros E. G.; Detection and quantification of Roundup Ready® soybean residues in sausage samples by conventional and real-time PCR, **Ciência e Tecnologia de Alimentos**, v.28, p38-45, 2008.
- Medeiros, A. C. B. Desenvolvimento, caracterização e estimativas de erros de genotipagem de locos microssatélites de *Tabebuia aurea* (Bignoniaceae). 111f. Dissertação (Mestrado Ciências Genômicas e Biotecnologia) - Universidade Católica de Brasília, Brasília, DF, 2006.
- Messmer, M. M.; Melchinger, A. E.; Hermann, R. G.; Boppenmaier, J. Relationships among early European maize inbreds: II. Comparison of pedigree and RFLP data. **Crop Science**. v.33, p. 944-950, 1993.
- Mian, M. A. R.; Kang, S. T.; Redinbaugh, M. G. Microsatellite diversity of soybean genotypes differing in bean pod mottle virus leaf symptom. **Canadian Journal of Plant Science**, v.87, p.359-367, 2008.
- Miller, M. J. & Yuan B. Z. Semiautomated Resolution of overlapping stutter patterns in genomic microsatellite analysis. **Analytical Biochemistry**, v.251, p.050-056, 1997.

- Mimura, M.; Coyne, C. J.; Bambuck, M. W.; Lumpkin, T. A. SSR diversity of vegetable soybean [*Glycine max* (L.) Merr.], **Genetic Resources and Crop Evolution**, v.54, p.497-508, 2006.
- Missiaggia, A. & Grattapaglia, D. Plant microsatellite genotyping with 4-color fluorescent detection using multiple-tailed primers. **Genetics and Molecular Research**, v.5, p.72-78, 2006.
- Moreira, C. T.; Souza, P. I. M.; Farias Neto, A. L.; Almeida, L. A. Ocorrência de variações na coloração do hilo de sementes de cultivares de soja [*Glycine max* (L.)Merrill]. **Comunicado técnico EMBRAPA Cerrados**, Planaltina, 4p. no. 5, 1999.
- Mulato, B. M.; Möller, M.; Zucchi, M. I.; Quecini, V.; Pinheiro, J. B. Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers. **Pesquisa Agropecuária Brasileira**, v.45, p.276-283, 2010.
- Neilan, B. A.; Wilton, A. N.; Jacobs, D. A universal procedure for primer labelling of amplicons. **Nucleic Acids Research**, v.25, n.14, p.2938-2939, 1997.
- Oblessuc, P. R.; Campos, T.; Cardoso, J. M. K.; Sforça, D. A.; Baroni, R. M.; Souza, A. P.; Benchimol, L. L. Adaptation of fluorescent technique for genotyping with new microsatellite markers in common bean. **Pesquisa Agropecuária Brasileira**, v.44, n.6, p.638-644, 2009.
- Oeting, W. S.; Lee, H. K.; Flanders, D. J.; Wiesner, G. L.; Selers, T. A.; King, R. A. Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. **Genomics**, v.30, p.450-458, 1995.
- Oliveira, M. B.; Vieira, E.S.N.; Schuster, I. Construction of a molecular database for soybean cultivar identification in Brazil. **Genetics and Molecular Research** v.9, p.705-720, 2010.
- Pompanon, F.; Bonin, A.; Bellemain, E.; Taberlet, P. Genotyping errors: causes consequences, and solutions. **Nature Reviews Genetics**, v.6, p.847-859, 2005.

- Powell, W.; Machray, G. C.; Provan, J. Polymorphism revealed by simple sequence repeats. **Trends in Plant Science**. v.1, p.215-222, 1996.
- Priolli, R. H. G.; Mendes-Junior, C. T.; Arantes, C. E.; Contel, E. P. B. Characterization of Brazilian soybean cultivars using microsatellite markers. **Genetics and Molecular Biology**, v.25, p.185-193, 2002.
- Priolli, R. H. G.; Mendes-Junior, C. T.; Sousa, S. M. B.; Sousa, N. E. A.; Contel, E. P. B. Diversidade genética da soja entre períodos e entre programas de melhoramento no Brasil. **Pesquisa Agropecuária Brasileira**, v.39, p.967-975, 2004.
- Rabel M.; Vieira, E. S. N.; Lana, U. G. P.; Paiva, E.; Sehnem, A. P. A.; Schuster, I. Marcadores moleculares microssatélites na avaliação de sementes de soja com variação na coloração do hilo. **Revista Brasileira de Sementes**, v.32, n.2, p.019-025, 2010.
- Rodrigues, D. H.; Alcântara Neto, F.; Schuster, I. Identification of essentially derived soybean cultivars using microsatellite markers, **Crop Breeding and Applied Biotechnology**, v.8, p.74-78, 2008.
- Sambrook, J.; Fritsch, E. F.; Maniatis, T. **Molecular cloning: a laboratory manual**. Cold Spring Harbor University Press, Cold Spring Harbor, NY. 1989.
- Sayama, T.; Hwang, T. Y.; Komatsu, K.; Takada, Y.; Takahashi, M.; Kato, S.; Sasama, H.; Higashi, A.; Nakamoto, Y.; Funatsuki, H.; Ishimoto, M. Development and application of a whole-genome simple sequence repeat panel for high-throughput genotyping in soybean. **DNA Research**, v.18, p.107-115, 2011.
- Schuelke, M. An economic method for the fluorescent labeling of PCR fragments. **Nature Biotechnology**, v.18, p.233-234, 2000.
- Schuster, I.; Queiroz, V. T.; Teixeira, A. I.; Barros, E. G.; Moreira, M. A. Determinação da pureza varietal de sementes de soja com o auxílio de marcadores moleculares microssatélites. **Pesquisa Agropecuária Brasileira**, v.39, n.3, p.247-253, 2004.

- Schuster, I.; Vieira, E. S. N.; Padilha, L. Marcadores moleculares no pós-melhoramento. In: **Marcadores moleculares**. 2 ed. Viçosa-MG, p.177-208, 2009.
- Singh, R. K.; Bhatia, V. S.; Bhat, K. V.; Mohapatra, T.; Singh, N. K.; Bansal, K. C.; Koundal, K. R. SSR and AFLP based genetic diversity of soybean germplasm differing in photoperiod sensitivity. **Genetics and Molecular Biology**, v.33, p.319-324, 2010.
- Song, Q. J.; Marek L. F.; Shoemaker R. C.; Lark K. G.; Concibido V. C.; Delannay X.; Specht J. E.; and Cregan P. B. A new integrated genetic linkage map of the soybean. **Theoretical and Applied Genetics**, v.109, p.122-128, 2004.
- Song, Q. J.; Quigley, C. V.; Nelson, R. L.; Carter, T. E.; Boerma, H. R.; Strachan, J.; Cregan, P. B. A selected set of trinucleotide simple sequence repeat markers for soybean variety identification. **Plant Varieties and Seeds**, v.12, p.207-220, 1999.
- United States Department of Agriculture – USDA. **Mapped Soybean SSR Loci**  
Disponível em: <<http://bldg6.arsusda.gov/~pooley/soy/cregan/soymap.htm>>  
Acesso em: fev/2010.
- Van Oosterhout, C.; Hutchinson, W. F.; Wills, D. P. M.; Shipley, P. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. **Molecular Ecology Notes**, v.4, p.535-538, 2004.
- Varshney R.K.; Graner A.; and Sorrells M. E. Genic microsatellite markers in plants: features and applications. **Trends in Biotechnology**, v.23, p.48-55, 2005.
- Vieira, E. S. N.; Pinho, E. V. R. V.; Carvalho, M. G. G.; Silva, P. A. Caracterização de cultivares de soja por descritores morfológicos e marcadores bioquímicos de proteínas e isoenzimas. **Revista Brasileira de Sementes**, v.31, n.01, p.086-094, 2009.
- Wang, L. X.; Guan, R. X.; Li, Y. H.; Lin, F. Y.; Luan, W. J.; Li, W.; Ma, Y. S.; Liu, Z. X.; Chang, R. Z.; Qiu, L. J. Genetic diversity of Chinese Spring Soybean Germplasm Revealed by SSR Markers. **Plant Breeding**, v.127, p.56-61, 2008.

Weir, B. **Genetics data analysis II**. Sinauer Associates, Inc., Sunderland, Massachusetts. 376p. 1996.

Wenz, H. M.; Robertson, J. M.; Menchen, S.; Oaks, F.; Demorest, D. M.; Scheibler, D.; Rosenblum, B. B.; Wike, C.; Gilbert, D. A.; Efcavitch J. W. High-precision genotyping by denaturing capillary electrophoresis. **Genome Research**, v.8, p.69-80, 1998.

Wysmierski, P. T. Contribuição genética dos ancestrais da soja às cultivares brasileiras. 100f. Dissertação (Mestrado em Ciências: Genética e melhoramento de plantas) - Universidade de São Paulo - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, SP, 2010.