**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Desenvolvimento de marcadores SSR-EST e construção de mapas genéticos
em feijão-comum (*Phaseolus vulgaris* L.)**

Luiz Ricardo Hanai

**Tese apresentada para obtenção do título de doutor em
Agronomia. Área de concentração: Genética e
Melhoramento de Plantas**

**Piracicaba
2008**

Luiz Ricardo Hanai
Bacharel em Ciências Moleculares

**Desenvolvimento de marcadores SSR-EST e construção de mapas genéticos
em feijão-comum (*Phaseolus vulgaris* L.)**

Orientadora:
Prof^a Dra. MARIA LUCIA CARNEIRO VIEIRA

Tese apresentada para obtenção do título de doutor em
Agronomia. Área de concentração: Genética e Melhoramento de
Plantas

Piracicaba
2008

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Hanai, Luiz Ricardo

Desenvolvimento de marcadores SSR-EST e construção de mapas genéticos em feijão-comum (*Phaseolus vulgaris* L.) / Luiz Ricardo Hanai. - - Piracicaba, 2008.
120 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2008.
Bibliografia.

1. Feijão 2. Mapeamento genético 3. Marcador molecular 4. Resistência genética vegetal I. Título

CDD 635.652
H233d

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

DEDICATÓRIA

À minha esposa Camila pelo amor e companheirismo durante a realização do doutorado, e à nossa filha Helena que já está nos enchendo de felicidades, com amor dedico.

Aos meus pais Joaquim e Cleusa por, incondicionalmente, me apoiarem sempre, ofereço.

AGRADECIMENTOS

À Escola Superior de Agricultura “Luiz de Queiroz” – ESALQ/USP, em especial ao Departamento de Genética, pela oportunidade;

À Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, pela bolsa concedida;

À Professora Dra. Maria Lucia Carneiro Vieira, pela orientação, ensinamentos e confiança a mim dispensada desde a Iniciação Científica;

Ao Professor Dr. Luis Eduardo Aranha Camargo, por tornar disponível o banco de EST de feijão - BEST, pelo entusiasmo e pelas ótimas sugestões e idéias;

À Dra. Siu Mui Tsai, por disponibilizar parte do material vegetal usado neste trabalho, e pela amizade e entusiasmo;

Ao Professor Dr. João Bosco dos Santos da Universidade Federal de Lavras, pela atenção dispensada e por ceder parte do material vegetal usado neste trabalho;

Aos Professores do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, pelos conhecimentos compartilhados;

Ao Msci. Marcos Vinícius B. M. Siqueira, pela contribuição no desenho de *primers*;

À Bióloga Carla de Freitas Munhoz, pela ajuda na genotipagem dos acessos de feijão;

À Engenheira Agrônoma Luciane Santini, pelo apoio nos trabalhos de laboratório, sobretudo na genotipagem das populações de mapeamento;

À Engenheira Agrônoma Camila Motta Borgonove e ao Biólogo Luciano Lucena pelas leituras críticas de meus relatórios e manuscritos;

À Dra. Christine P. Stock, pela revisão do abstract;

Aos amigos do laboratório de Biologia Celular e Molecular de Plantas, Ricardo Lopes, Mariza, Maria Rita, Adriano, Juliano, Michel, Frederico, Eder, Francisco, Sheila, Carla, Luciano Consoli, Luciane, Renato, Endson, Helen, Alessandra, Marcelo, Larissa, Thiago, Anselmo, Bruno e Luciano Lucena, pela contribuição no meu desenvolvimento científico e pessoal;

Aos colegas do curso e amigos Marcelo Oliveira, Andréa, Matheus Figueiredo, Walter, Edgar, Marcos Barone, Sidney, Jair e José Manoel, pelos momentos de alegria e descontração;

Aos amigos da República Zona Rural;

A todos que diretamente ou indiretamente contribuíram para a realização deste trabalho.

SUMÁRIO

RESUMO	7
ABSTRACT	8
1 INTRODUÇÃO	9
2 REVISÃO DE LITERATURA	12
2.1 A cultura do feijoeiro	12
2.1.1 Importância econômica e social	12
2.1.2 Pragas e doenças da cultura	13
2.1.3 Recursos genéticos	14
2.1.4 Melhoramento genético da cultura	15
2.2 Marcadores moleculares	17
2.3 Mapas genéticos	22
2.3.1 Conceito e breve histórico	22
2.3.2 Construção de mapas genéticos	24
2.3.3 Aplicações dos mapas genéticos	26
2.3.4 Mapas genéticos em feijão	29
3 MATERIAL E MÉTODOS	33
3.1 Material Vegetal	33
3.2 Extração e quantificação de DNA	34
3.2.1 Isolamento do DNA genômico	34
3.2.2 Quantificação do DNA	35
3.3 Marcadores SSR-EST	36
3.3.1 Desenvolvimento de marcadores SSR-EST	36
3.3.1.1 Caracterização das EST contendo SSR	36
3.3.1.2 Classificação das SSR	36
3.3.1.3 Seleção e desenho dos primers a partir das EST	37
3.3.2 Condições de reação	38
3.3.3 Eletroforese e visualização dos locos	40
3.3.4 Análise dos locos amplificados	41
3.3.4.1 Caracterização dos marcadores SSR-EST	41
3.3.4.2 Seleção de locos polimórficos nas populações BJ e CFM	41

3.4 Marcadores AFLP	42
3.4.1 Geração de marcadores AFLP	42
3.4.2 Análise em gel de poliacrilamida	43
3.5 Análises de mapeamento	44
3.5.1 Integração dos marcadores desenvolvidos no mapa referência BJ	44
3.5.2 Construção do mapa CFM	45
3.5.3 Alinhamento do mapa CFM ao mapa núcleo BJ	45
4 RESULTADOS E DISCUSSÃO	46
4.1 Frequência e localização dos SSR nas EST	46
4.2 Caracterização dos marcadores SSR-EST	49
4.3 Mapeamento genético da população BJ	54
4.3.1 Genotipagem dos marcadores SSR-EST	54
4.3.2 Genotipagem dos marcadores AFLP	55
4.3.3 Mapa de ligação BJ	57
4.4 Mapeamento genético da população CFM	64
4.4.1 Genotipagem dos marcadores SSR-EST	64
4.4.2 Genotipagem dos marcadores AFLP	65
4.4.3 Mapa de ligação CFM	67
4.4.4 Alinhamento do mapa CFM ao mapa consenso BJ	71
5 CONCLUSÕES	74
REFERÊNCIAS	75
ANEXOS	95

RESUMO

Desenvolvimento de marcadores SSR-EST e construção de mapas genéticos em feijão-comum (*Phaseolus vulgaris* L.)

O uso de marcadores moleculares tem contribuído para o estudo da domesticação das espécies de *Phaseolus*, origem e diversidade das cultivares atuais de feijão-comum (*P. vulgaris*), e do controle genético da resistência a diversas doenças. Mapas de ligação têm sido estabelecidos em feijão-comum com base em marcas de RFLP, RAPD, SCAR, isoenzimas e locos fenotípicos, sendo que alguns deles foram reunidos em um mapa mais denso e completo do genoma do feijão, chamado mapa núcleo. No entanto, para o uso efetivo de mapas de ligação em programas de melhoramento estes devem ser suficientemente saturados. O presente trabalho se insere neste contexto, visando saturar o mapa núcleo de *P. vulgaris* a partir do mapeamento de novos marcadores moleculares como AFLP e SSR baseados em EST. Assim, buscou-se desenvolver e caracterizar marcadores SSR oriundos de uma biblioteca de EST, testar a transferibilidade destes para outros cultivares e espécies relacionadas, gerar marcadores AFLP e integrá-los ao mapa consenso da espécie. Também, construir um mapa genético para uma população de interesse no Brasil ('Carioca' x 'Flor de Mayo'), e promover o seu alinhamento com o mapa consenso, complementando desta forma, a caracterização genômica da espécie. Pares de *primers* foram desenhados para 156 SSR-EST. Destes, 138 SSR-EST amplificaram locos claros e reprodutíveis e foram caracterizados usando um conjunto de 26 genótipos de feijões cultivados. Dos locos analisados, 85 se mostraram polimórficos entre os genótipos estudados e apresentaram em média 2,96 alelos por loco e um PIC médio de 0,38. Entre todos os SSR-EST analisados, 50 locos segregaram na população de mapeamento 'Bat 93' x 'Jalo EEP558', 20 locos segregaram na população 'Carioca' x 'Flor de Mayo' e 12 locos foram polimórficos nas duas populações. Marcadores AFLP foram gerados e genotipados nas duas populações. Os 262 locos microssatélites e AFLP genotipados na população 'Bat 93' x 'Jalo EEP558' foram integrados ao mapa núcleo da espécie. Foi obtido um mapa de 1353 cM de comprimento total, contendo 357 marcas, incluindo 9 SSR-genômico, 47 SSR-EST e 190 AFLP. Além disso, outro mapa foi gerado a partir da análise de segregação de 252 marcadores na população 'Carioca' x 'Flor de Mayo'. Este mapa teve 807,5 cM de comprimento, com uma distância média de 5,3 cM entre marcas. Os marcadores microssatélites comuns foram usados como ponte para alinhar e comparar os mapas das duas populações estudadas.

Palavras-Chave: *Expressed Sequence Tag* (EST); Marcador Molecular; Mapa Genético

ABSTRACT

Development of EST-SSR markers and construction of genetic maps in common bean (*Phaseolus vulgaris*)

The use of molecular markers has contributed to the studies regarding domestication of *Phaseolus* species, origin and diversity of current common bean cultivars (*P. vulgaris*), and the genetic control of resistance to several diseases. Linkage maps have been constructed for common bean by using RFLP, RADP, SCAR, isoenzymes and phenotypic markers, some of which were meeting in a more dense and complete map of bean genome, called core map. However, for the effective use of linkage maps in breeding programs they must be sufficiently saturated. This work fits in this context, aiming to saturate the core map of *P. vulgaris* from the mapping of new molecular markers as AFLP and SSR EST-based. Thus, it was tried to develop and characterize SSR markers from a EST library, to test the transferability of these markers to other cultivars and related species, to generate AFLP markers and integrate them to the consensus map of the species. Also, build a genetic map for a population of interest in Brazil ('Carioca' x 'Flor de Mayo'), and promoting its alignment with the consensus map, thus complementing the genomic characterization of the species. Pairs of primers were designed for 156 EST-SSR. Of these, 138 EST-SSR amplified clear and reproducible loci and were characterized using a set of 26 genotypes of cultivated beans. Of the loci tested, 85 were polymorphic between the genotypes studied and showed an average 2.96 alleles per locus and a PIC average of 0.38. Among all examined EST-SSR, 50 loci segregated in 'Bat 93' x 'Jalo EEP558' mapping population, 20 loci segregated in 'Carioca' x 'Flor de Mayo' population and 12 loci were polymorphic in both two populations. AFLP markers were generated and genotyped in the two populations. The 262 microsatellites and AFLP loci genotyped in 'Bat 93' x 'Jalo EEP558' population were integrated onto the core map of the species. A map of 1353 cM total length was obtained, containing 357 markers, including 9 genomic-SSR, 47 EST-SSR and 190 AFLP. Moreover, another map was generated from the analysis of segregation of 252 markers in 'Carioca' x 'Flor de Mayo' population. This map was 807.5 cM long, with an average distance of 5.3 cM between markers. The common microsatellites markers were used as a bridge to align and compare the maps of the two studied populations.

Keywords: Expressed Sequence Tag (EST); Molecular Marker; Genetic Map

1 INTRODUÇÃO

O feijoeiro é a espécie mais cultivada entre as outras do gênero *Phaseolus*, contribuindo com, aproximadamente, 95% da produção mundial de feijões. É cultivado, basicamente, na África tropical (10% da produção mundial) e na América Latina (47% da produção mundial) (FAO, 2008), sendo o Brasil o maior produtor, com uma produção de 3,5 milhões de toneladas por ano, colhida em uma área de 4,2 milhões ha (safra 2005/06; CONAB, 2008). A importância da cultura se deve ao fato de o feijão se constituir em um dos alimentos básicos da população brasileira (cerca de 16 Kg/hab/ano), mas também por representar uma das principais fontes de proteínas na dieta alimentar, principalmente para as camadas de renda menos favorecidas (BROUGHTON et al., 2003).

Por isso, um esforço crescente tem sido dedicado à melhoria e expansão do cultivo de feijões. Os principais objetivos dos programas de melhoramento genético são aumentar a produtividade da cultura, aumentar a qualidade e a quantidade de proteínas, vitaminas e minerais e reduzir alguns fatores antinutricionais presentes no grão, assim como conferir maior resistência a doenças (EVANS, 1986; NAGL; IGNACIMUTHU; BECKER, 1997). Vários estudos genéticos têm sido realizados e as abordagens se baseiam, principalmente, no uso de marcadores moleculares, visando ao esclarecimento sobre a origem e a diversidade das cultivares atuais (GEPTS, 1998), os fatores determinantes da domesticação da espécie (KONINIGE et al., 1996; SONNANTE et al., 1994) e o controle genético da resistência a diversas doenças (ADAM-BLONDON et al., 1994; JUNG et al., 1999; MIKLAS, P. N. et al., 2002; NODARI et al., 1993b; VALLEJOS et al., 2000; YU, Z. H.; STALL; VALLEJOS, 1998).

Os mapas genéticos gerados a partir de marcadores moleculares possibilitam uma ampla cobertura e a análise completa do genoma da espécie. A partir destes mapas, é possível a decomposição de caracteres de herança complexa em seus componentes mendelianos, o mapeamento de regiões genômicas que controlam caracteres de importância econômica, a quantificação do efeito destas regiões para o caráter estudado e a subsequente utilização destas informações no desenvolvimento de cultivares melhoradas (FERREIRA; GRATTAPAGLIA, 1998).

Para o feijão-comum, três principais mapas de ligação foram estabelecidos, utilizando diferentes tipos de marcadores moleculares e fenotípicos (ADAM-BLONDON et al., 1994; NODARI et al., 1993a; VALLEJOS; SAKIYAMA; CHASE, 1992). Mais tarde, estas marcas

foram integradas em um mapa mais denso e completo do genoma do feijão, chamado mapa núcleo (FREYRE et al., 1998). Porém, este mapa de ligação ainda é limitado quanto a informações genéticas sobre características úteis para os melhoristas (resistência a doenças, por exemplo) e não é suficientemente saturado para os geneticistas interessados no isolamento de genes a partir de diversas técnicas de clonagem. Outra limitação refere-se ao fato de as distâncias genéticas variarem consideravelmente entre os mapas de ligação construídos a partir de cruzamentos entre diferentes cultivares de feijão (JUNG et al., 1999; PAREDES; GEPTS, 1995). Assim, um maior grau de saturação do mapa núcleo é de grande interesse, pois se pode aumentar, por exemplo, a chance de encontrar marcadores associados a características de interesse em outras populações.

A saturação de mapas genéticos tem sido buscada usando diferentes técnicas que revelam polimorfismos moleculares. Neste sentido, os marcadores AFLP (*Amplified Fragment Length Polymorphism*) têm se mostrado adequados, uma vez que é possível visualizar até uma centena de locos em uma única corrida eletroforética (VOS et al., 1995). Outras vantagens deste marcador são a não necessidade de conhecimento prévio de seqüências, o baixo custo por loco analisado, além da sensibilidade da técnica a pequenas mutações pontuais nos genomas (YOUNG, W. P.; SCHUPP; KEIM, 1999).

Uma nova abordagem adotada para a saturação de mapas genéticos é a geração de marcadores derivados de genes, principalmente SNP (*Single Nucleotide Polymorphisms*; RAFALSKI, 2002) e SSR (*Single Sequence Repeats*; SHAROPOVA et al., 2002).

Os SSR ou microssatélites têm se tornado a fonte mais popular de marcadores genéticos (SCHLÖTTERER, 2004). Genomas de eucariotos são densamente povoados por microssatélites. Geralmente, cada microssatélite está localizado em um loco simples com grande variação no tamanho dos alelos entre os indivíduos. Além do alto nível de polimorfismo, as seqüências de microssatélites possuem a maioria dos atributos desejáveis para um marcador genético, incluindo um alto índice de informação, fácil distinção entre os alelos, alta reprodutibilidade experimental, codominância, genotipagem rápida e fácil. Em face do desenvolvimento dos SSR como marcadores muitos avanços têm sido obtidos, especialmente no mapeamento de locos quantitativos (QTL - *Quantitative Trait Loci*), na seleção assistida por marcadores e no *screening* de bibliotecas para a clonagem de genes.

Atualmente, sabe-se que boa parte dos microssatélites está contida em regiões expressas do genoma (GAO et al., 2003; VARSHNEY, R. K.; GRANER; SORRELLS, 2005). Isto, juntamente com a crescente disponibilidade de seqüências nos bancos de dados, principalmente EST, tem possibilitado o desenvolvimento de marcadores microssatélites derivados de genes, ou SSR-EST. Neste contexto, o uso destes marcadores tem sido relatado em várias espécies, incluindo o feijão-comum (BLAIR et al., 2003; YU, K. et al., 2000; YU, K. F.; PARK; POYSA, 1999). O fato desses marcadores serem derivados de genes e, conseqüentemente, suas seqüências serem mais conservadas, lhes é atribuído potencial para identificar polimorfismos em espécies relacionadas àquela da qual foi gerado a biblioteca de EST (THIEL et al., 2003; VARSHNEY, R. K. et al., 2005), o que facilita o mapeamento comparativo e a clonagem de genes homólogos.

Recentemente, um esforço no sentido de desenvolver ferramentas genômicas para os feijões foi iniciado com a criação do consórcio internacional 'Phaseomics', e várias bibliotecas de cDNA para feijão-comum e espécies relacionadas começaram a ser construídas (BROUGHTON et al., 2003). Na ESALQ/USP mais de 5.000 EST já foram seqüenciadas e, numa busca prévia, foram identificadas aproximadamente 240 seqüências candidatas a SSR (MELOTTO et al., 2005). O presente trabalho buscou desenvolver marcadores microssatélites derivados de EST e mapeá-los, juntamente com marcas AFLP, em duas populações de *P. vulgaris*: uma derivada do cruzamento 'Bat 93' x 'Jalo EEP558' (chamada de população núcleo de mapeamento ou BJ) e outra oriunda do cruzamento 'Carioca' x 'Flor de Mayo' (população CFM), visando à saturação do mapa núcleo BJ, o estabelecimento de um mapa de ligação para a população CFM, assim como, o alinhamento dos dois mapas.

Para tanto, os seguintes objetivos específicos foram buscados:

- (a) Gerar marcadores AFLP em *P. vulgaris*;
- (b) Desenvolver marcadores moleculares do tipo SSR-EST e caracterizá-los em um conjunto de acessos de feijão-comum e espécies relacionadas;
- (c) Mapear este conjunto de marcadores (AFLP e SSR-EST) no mapa de ligação integrado da espécie BJ;
- (d) Gerar um mapa de ligação para a população CFM, a partir dos marcadores AFLP e SSR-EST;
- (e) Usar os marcadores comuns para alinhar o mapa CFM ao mapa referência BJ.

2 REVISÃO DE LITERATURA

2.1 A cultura do feijoeiro

Além do feijão-comum (*P. vulgaris*), três outras espécies possuem maior importância econômica: *P. coccineus*, *P. lunatus* e *P. acutifolius* var *latifolius*. Os feijões (*Phaseolus* ssp.) são culturas extremamente diversas em termos de morfologia, métodos de cultivo, utilização e ambientes aos quais têm sido adaptados. São encontrados em plantios desde ao nível do mar a até 3000 m de altitude. No Brasil, o feijoeiro é cultivado por pequenos e grandes produtores, em diversificados sistemas de produção e em todas as regiões. Dependendo da cultivar e da temperatura ambiente, pode apresentar ciclos de vida variando de 65 a 100 dias o que possibilita sua produção em sistemas de monocultivo altamente tecnificado, em rotação de cultura com gramíneas ou em associação com outras culturas (principalmente o milho) em nível de subsistência (EMBRAPA, 2003). Dos feijões, são consumidos tanto os grãos maduros quanto imaturos e mesmo folhas e brotos (BROUGHTON et al., 2003).

2.1.1 Importância econômica e social

O feijão-comum é uma das mais antigas culturas agrícolas do novo mundo. Juntamente com o milho e a mandioca, têm sido a base da dieta em baixas e médias latitudes das Américas por milênios. Graças às suas comprovadas propriedades nutracêuticas, o feijão possui grande importância como componente em dietas de combate à fome e à desnutrição. De fato, o feijão supre cerca de 12% da necessidade de um adulto em proteínas, 4% em amido, 8% em cálcio, 20% em fósforo, 16% em magnésio, 27% em ferro, 8% em zinco, 12% em cobre, 18% em manganês, e 12% em potássio, baseado num consumo anual de 15 kg *per capita* (BROUGHTON et al., 2003). Quando combinado com outros alimentos, especialmente cereais, proporciona uma interessante complementação protéica, fornecendo os oito aminoácidos essenciais ao organismo.

Os feijões são as mais importantes leguminosas para o consumo humano direto. Sua produção total é quase duas vezes maior que a de grão de bico, a segunda mais importante leguminosa no mundo (FAO, 2008). Segundo dados estatísticos da FAO (2008), em 2006, a produção mundial de feijões situou-se em torno de 19,6 milhões de toneladas, ocupando uma área de 26,5 milhões de hectares. Cerca de 65,2% da produção foram oriundos de apenas seis países, sendo o Brasil o maior produtor e responsável por 17,6% de toda a produção mundial.

O feijoeiro comum se destaca entre as outras espécies do gênero *Phaseolus*, contribuindo com mais de 95% da produção mundial de feijões. É cultivado, principalmente, na África tropical (10% da produção mundial) e na América Latina (47% da produção mundial). O Brasil é o país que mais produz feijão-comum. Na safra 2005/06 foram colhidos 3,5 milhões de toneladas em uma área de 4,2 milhões de ha (CONAB, 2008). Por outro lado, o Brasil é, também, o maior consumidor do grão no mundo. O consumo médio *per capita* no Brasil é de cerca de 17 kg/ano, com maior consumo nas regiões Sudeste e Nordeste, com 18,2 e 20,8 kg/ano, respectivamente. No entanto, o país ainda apresenta sérios problemas que limitam a produtividade das lavouras e a qualidade dos grãos obtidos, os quais são, principalmente, de ordem fitossanitária (PAULA JÚNIOR; ZAMBOLIM, 1998).

2.1.2 Pragas e doenças da cultura

As doenças estão entre os principais fatores limitantes de produtividade do feijoeiro; tais doenças podem ser fúngicas, bacterianas ou viróticas, sendo a maioria delas ocasionadas por fungos (HALL, 1991). Entre os patógenos fúngicos, destacam-se: *Phaeoisariopsis griseola* (mancha-angular), *Colletotrichum lindemuthianum* (antracnose), *Sclerotinia sclerotiorum* (mofo-branco), *Fusarium oxysporum* f. sp. *phaseoli* (murcha-de-fusarium), *Uromyces appendiculatus* (ferrugem), entre outros (PAULA JÚNIOR; ZAMBOLIM, 1998).

A mancha-angular é uma doença disseminada por toda a região produtora de feijão do Brasil, e pode provocar perdas significativas, alcançando 70% dependendo da cultivar, da época de cultivo e do momento em que ocorre a infecção pelo patógeno (SARTORATO; RAVA, 1992). Com relação à antracnose, os acessos que se mostram resistentes apresentam grãos fora do padrão comercial. Existe uma grande dificuldade em associar em uma cultivar os genes de resistência destes acessos com grãos do tipo carioca, o preferido no Brasil. Outro problema é a grande variabilidade apresentada pelo patógeno. Apesar de já terem sido obtidas muitas linhagens resistentes, algumas das quais são recomendadas para o cultivo, a vida útil dessas cultivares tem sido pequena devido ao grande número de raças existentes do fungo (RAVA et al., 1993; RAVA; PRUCHIO; SARTORATO, 1993).

No que diz respeito às doenças bacterianas, o patógeno de maior relevância para a cultura no Brasil é *Xanthomonas axonopodis* pv. *phaseoli*, causador do crestamento-bacteriano-comum. Em condições favoráveis ao patógeno (temperatura e umidade elevadas), as perdas causadas pelo

crestamento-bacteriano-comum podem ultrapassar 40% (WALLEN; JACKSON, 1975; YOSHII; GALVEZ; ALVAREZ, 1976). A redução da taxa fotossintética nos tecidos verdes remanescentes das folhas infectadas é a principal causa dos danos (DÍAZ et al., 2001).

Em relação às doenças virais, merece destaque o mosaico-dourado, causado pelo BGMV (*Bean Golden Mosaic Virus*) e transmitido pela mosca-branca (*Bemisia tabaci*) (KIMATI et al., 2005; VIEIRA et al., 2005). A doença é largamente disseminada nas regiões produtoras de feijão no Brasil e na América Latina, causando perdas severas - entre 40 e 100% - dependendo da incidência, da época de plantio e da cultivar (MELO et al., 2005).

Frente às elevadas perdas de qualidade e produtividade do feijão, causadas pela incidência de doenças, é imprescindível a obtenção de variedades geneticamente resistentes. A incorporação de resistência a doenças tem efeito estabilizador nas cultivares, evitando a oscilação de rendimento de um plantio para outro, ocasionada pelo ataque de patógenos. E, dessa forma, possibilita à cultivar expressar seu potencial produtivo (VIEIRA et al., 2005).

2.1.3 Recursos genéticos

Dois grandes grupos gênicos são conhecidos entre os feijões cultivados, o Mesoamericano e o Andino. Tomando por base o hábito de crescimento, as características morfológicas, as regiões ecológicas de adaptação e outras informações de 18300 acessos de *P. vulgaris*, Singh (1989) verificou grandes diferenças entre os acessos dos centros primários de domesticação e dividiu o germoplasma de feijão em doze conjuntos gênicos, seis em cada um dos centros primários de domesticação. Segundo Singh, Gepts e Debouck (1991), estes 12 conjuntos gênicos representam seis raças de feijão: a raça Mesoamérica que é formada pelos conjuntos gênicos 1, 2, 3 e 4; a raça Durango, formada pelo grupo gênico 5; a raça Jalisco, formada pelo conjunto gênico 6; a raça Nueva Granada, formada pelos conjuntos gênicos 7, 8 e 9; a raça Chile, formada pelo conjunto gênico 10; e a raça Peru que é formada pelo conjunto gênico 11 e 12. No Brasil, são mais comuns cultivares de sementes pequenas e faseolina do tipo ‘S’ da raça Mesoamérica e cultivares de sementes grandes e faseolina do tipo ‘T’ da raça Nueva Granada (VIEIRA et al., 2005). Assim, as outras 4 raças menos exploradas no Brasil constituem importantes recursos genéticos dentro das formas domesticadas da espécie.

Segundo Singh (2001), o conhecimento sobre, a diversidade disponível no germoplasma domesticado e selvagem, o acesso a ela, e o seu uso são essenciais para a ampliação da base genética das cultivares para um melhoramento sustentável.

Segundo Chacón, Pickersgill e Debuouck (2005), os recursos genéticos de *P. vulgaris* existem em uma complexa composição de grupos gênicos (principais e secundários), raças e tipos intermediários, com ocasional introgressão entre tipos selvagens e domesticados. O tipo selvagem possui uma ampla distribuição geográfica nas Américas do Sul e Central e tanto o tipo selvagem quanto o domesticado de feijão mostram uma maior variação nas populações Mesoamericanas que nas populações Andinas (CHACÓN; PICKERSGILL; DEBOUCK, 2005; GEPTS et al., 1988). Acredita-se que a maioria da variabilidade genética da espécie se originou, provavelmente, durante o processo de dispersão a partir de suas formas ancestrais (ACOSTA-GALLEGOS; KELLY; GEPTS, 2007).

2.1.4 Melhoramento genético da cultura

Os métodos empregados no melhoramento de feijão utilizam a variabilidade natural (como a introdução de linhagens ou cultivares) ou a variabilidade gerada pelo cruzamento entre duas ou mais linhagens ou cultivares diferentes. A variabilidade genética gerada pelos cruzamentos é explorada conforme a herdabilidade do caráter de interesse. Caracteres de alta herdabilidade, como resistências monogênicas, cor e peso dos grãos podem ser selecionados, precocemente, a partir de F₂. Já caracteres de menor herdabilidade, como produtividade de grãos, requerem o avanço de gerações até a homozigose, quando as linhagens recombinantes são avaliadas em experimentos com repetições (RAMALHO; ABREU; SANTOS, 2001).

As populações segregantes podem ser conduzidas por diferentes métodos, no entanto, o mais utilizado tem sido o *bulk* dentro de famílias (PARRELLA, 2006). Neste método, a abertura das famílias é realizada mais cedo que no *bulk* convencional, quando estas ainda estão segregando. Assim, estas famílias são conduzidas por mais algumas gerações, sendo, posteriormente, selecionadas as melhores famílias e as melhores linhagens dentro dessas famílias, com base no seu desempenho durante as gerações. A vantagem na adoção deste método em relação ao *bulk* convencional é que a avaliação das famílias é realizada com base em resultados experimentais e não em análise visual.

Os objetivos específicos dos programas de melhoramento do feijoeiro são aumentar a produtividade da cultura, aumentar a qualidade e a quantidade de proteínas, vitaminas e minerais e reduzir alguns fatores antinutricionais presente no grão, assim como conferir maior resistência a doenças (EVANS, 1986; NAGL; IGNACIMUTHU; BECKER, 1997).

Cultivares mais produtivas têm sido produzidas com sucesso através de melhoramento clássico por muitos anos (SINGH, P. et al., 2007). Abreu e colaboradores (1994) avaliaram o progresso genético durante 20 anos e indicaram um ganho da ordem de 14,5 kg/ha/ano ou 1,9% ao ano. No entanto, é cada vez maior a dificuldade de se continuar obtendo ganhos, já que as diferenças a serem detectadas são cada vez menores. Além disso, existe no Brasil uma preferência por grãos do tipo carioca, o que tem restringido a variabilidade utilizada, devido à necessidade de selecionar apenas indivíduos ou famílias com este tipo de grão. Como a herança deste caráter, assim como, do rendimento de grão é controlada por muitos genes (BASSETT, 1996; LEAKEY, 1998), Ramalho, Abreu e dos Santos (2005) propõem que a seleção recorrente é uma boa opção para aumentar a frequência de alelos favoráveis pela recombinação em sucessivos ciclos de seleção. E, desta forma, continuar obtendo progresso no rendimento de grãos em linhagens com cor de grão comercialmente aceitável. Os autores avaliaram o progresso genético obtido após quatro ciclos de seleção recorrente e estimaram o ganho por ciclo em 7,4% para a produção de grão em feijoeiro.

Com relação à resistência a doenças, dada a grande variabilidade dos patógenos e o aparecimento de novas raças virulentas, o principal desafio é a obtenção de cultivares com classes de grãos comerciais e adequados níveis de resistência. Em função da pequena durabilidade das cultivares resistentes com apenas um alelo de resistência, uma alternativa para aumentar a vida útil dos genes de resistência e, conseqüentemente, das cultivares é a piramidação de diferentes genes em uma única cultivar (MUNDT, 1990; YOUNG, R. A.; KELLY, 1997). Uma dificuldade associada à construção de pirâmides de genes é a identificação da presença simultânea de dois ou mais genes de resistência em um único genótipo, utilizando os procedimentos usuais de inoculação (YOUNG, R. A. et al., 1998). A alternativa tem sido usar marcadores moleculares associados à expressão dos diferentes alelos de resistência. A integração da seleção assistida por marcadores com as abordagens clássicas no melhoramento para resistência a doença está avançando rapidamente para as principais doenças do feijoeiro (MIKLAS, P. N. et al., 2006).

Marcadores moleculares fortemente ligados aos principais genes e QTL de resistência foram relatados para importantes doenças do feijoeiro, como ferrugem (CORREA et al., 2000; MIKLAS, P. N. et al., 2002), antracnose (KELLY; VALLEJO, 2004), mosaico-dourado (STAVELY et al., 2001), crestamento bacteriano (MIKLAS, P. N. et al., 2003) e mofo branco (ENDER; TERPSTRA; KELLY, 2008).

Outra estratégia para acumular em uma ou mais linhagens a maioria dos alelos de resistência é a seleção recorrente. Lyons, Dickson e Hunter (1987), visando à obtenção de cultivares resistentes ao mofo branco, realizaram três ciclos de seleção que resultaram no aumento da resistência com um ganho genético de 31% desde o ciclo inicial ao terceiro ciclo. Como o objetivo de selecionar linhagens resistentes a doenças de solo causadas por *Pythium* spp, *Rhizoctonia solani*, *Macrophomina phaseolina* e *Fusarium* spp, Garcia e colaboradores (2003) avaliaram a taxa de sobrevivência e a produtividade das melhores linhagens provenientes de 6 ciclos de seleção recorrente. Os autores observaram um aumento progressivo na produção e na porcentagem de sobreviventes à medida que se avançaram os ciclos seletivos.

Além da ênfase na obtenção de cultivares mais produtivas e com resistência as principais doenças, as características de processamento ou cozimento necessitam serem avaliadas também, já que a grande maioria dos feijões é consumida diretamente. Estas características incluem razão de hidratação, coloração após o processamento e textura do grão, dentre outras. Posa-Macalincag e colaboradores (2002) sugerem a possibilidade do desenvolvimento de marcadores e de seleção assistida em gerações precoces para melhorar a qualidade de processamento, devido à herança complexa destas características e à necessidade do avanço de gerações para a avaliação fenotípica.

2.2 Marcadores moleculares

Marcadores genéticos são marcas que diferenciam genótipos e são herdados de acordo com as leis de herança Mendeliana (LEFEBVRE; CHÈVRE, 1995). Os marcadores genéticos podem ser morfológicos, bioquímicos ou moleculares. Apesar de os dois primeiros tipos terem tido, inicialmente, grande importância nos estudos genéticos, eles caíram em desuso com o passar do tempo, sendo substituídos pelos marcadores moleculares (SCHLÖTTERER, 2004). Isto aconteceu, principalmente, devido à limitação da disponibilidade destes marcadores e à influência do ambiente ou de outros genes no seu fenótipo. Por outro lado, os marcadores moleculares são

neutros e abundantes no genoma, permitindo maior precisão e detalhamento nos estudos genéticos (LEFEBVRE; CHÈVRE, 1995).

Marcadores moleculares são, basicamente, um conjunto de métodos de detecção de variações nas seqüências de DNA; são formas de explorar os polimorfismos naturais existentes entre os indivíduos e torná-los úteis nas diversas investigações em genética (CAIXETA et al., 2006). Estas diferenças podem ser devido à adição, deleção ou substituição de pares de bases, assim como, devido a rearranjos dos segmentos de DNA por efeito de deleções, inserções, inversões ou translocações ocorridas nas fitas de DNA. A forma de detecção deste polimorfismo é o que determina os diferentes marcadores moleculares. Estes podem ser classificados em duas categorias, dependendo de como o polimorfismo é revelado: por hibridização ou por PCR.

Os marcadores cujo polimorfismo é revelado por hibridização são basicamente o RFLP (*Restriction Fragment Length Polymorphism* - BOTSTEIN et al., 1980) e o VNTR (*Variable Number Tandem Repeats* - NAKAMURA et al., 1987). Nestas técnicas, os locos polimórficos são detectados pela diferença no comprimento dos fragmentos gerados por restrição e revelados a partir da hibridização com sondas marcadas. Estas sondas podem ser insertos clonados a partir da construção de bibliotecas genômicas ou de cDNA para os locos RFLP ou micro- ou minissatélites no caso dos locos VNTR. Estes marcadores possuem algumas vantagens como herança codominante e grande reproducibilidade; no entanto, devido à complexidade técnica envolvida na geração destes marcadores, a necessidade de marcação radioativa e a impossibilidade de automação, estes também foram sendo substituídos (SCHLÖTTERER, 2004).

Atualmente, existem novos marcadores que exploram a hibridização de ácidos nucléicos, como é o caso dos recém-desenvolvidos DArT (*Diversity Arrays Technology* - JACCOUD et al., 2001), SFP (*Single Feature Polimorphism* - BOREVITZ et al., 2003) e GEM (*Gene Expression Marker* - RONALD et al., 2005). Estes marcadores são independentes de eletroforese em gel, utilizam microarranjos de oligonucleotídeos nos quais são hibridizadas sondas para sua revelação. De forma geral, estes três marcadores possibilitam uma boa cobertura genômica e são automatizáveis, por outro lado, eles possuem caráter dominante.

Após a invenção da PCR (*Polymerase Chain Reaction* - MULLIS; FALOONA, 1987) no início da década de 1990, a capacidade de visualização de locos genômicos, sem a necessidade de marcação e hibridização possibilitou o desenvolvimento de uma série de marcadores moleculares, sendo os mais importantes o RAPD (*Random Amplified Polymorphic DNA* - WILLIAMS et al.,

1990), os microssatélites e o AFLP (*Amplified Fragment Length Polymorphism* - VOS et al., 1995). Os marcadores RAPD são facilmente gerados, pois se baseiam na amplificação de segmentos arbitrários de DNA ao longo do genoma, a partir de iniciadores curtos (10 pb). Sempre que houver o anelamento destes iniciadores em sentidos opostos ocorre a amplificação do segmento entre eles, gerando, assim, vários locos ao mesmo tempo; a eletroforese em gel de agarose (1,4% p/v) permite a sua rápida detecção. Os marcadores RAPD possibilitaram a construção de mapas genéticos para diversas espécies, incluindo as leguminosas, feijão (ADAM-BLONDON et al., 1994; CARNEIRO et al., 2002; GRATTAPAGLIA; SEDEROFF, 1994), *Vigna* (KAGA et al., 1996), grão de bico (SIMON; MUEHIBAUER, 1997) e lentilha (EUJAYL et al., 1998). No entanto, estes marcadores possuem herança dominante e os indivíduos heterozigóticos não podem ser distinguidos dos homozigóticos. Além disso, devido ao uso de iniciadores curtos, a estringência da PCR deve ser baixa para que haja o seu anelamento, o que torna a amplificação muito sensível a variações, dificultando a sua reprodutibilidade e, conseqüentemente, as comparações entre *profiles* gerados por diferentes laboratórios.

O AFLP (VOS et al., 1995) explora a variação do comprimento de fragmentos de restrição, como no RFLP. No entanto, os fragmentos são amplificados seletivamente via PCR, reduzindo o trabalho exigido na detecção do RFLP. Os *profiles* de AFLP são oriundos dos fragmentos gerados pela digestão do DNA genômico por duas enzimas de restrição, uma de corte raro e outra de corte freqüente. Após a digestão, os fragmentos de restrição recebem seqüências específicas, denominadas adaptadores, que se ligam às suas extremidades. Subconjuntos de fragmentos são amplificados, seletivamente, via PCR, utilizando como *primers* seqüências complementares às seqüências dos adaptadores acrescidas de 1 a 3 nucleotídeos arbitrários na posição 3', permitindo a seleção dos fragmentos que serão amplificados. A separação dos fragmentos se dá em gel de poliacrilamida (gel de seqüenciamento), de modo que os fragmentos de corte raro-freqüente sejam prioritariamente visualizados. Diferentes subconjuntos de fragmentos podem ser obtidos modificando-se a combinação dos nucleotídeos adicionais nos *primers* complementares aos adaptadores. O polimorfismo detectado pelo AFLP é observado, pela ou ausência do fragmento amplificado, conferindo a estas marcas um caráter dominante (LOPES, 2003). Cada gel, em geral, é capaz de revelar a variação em até uma centena de fragmentos (locos) o que se constitui em uma vantagem deste marcador em relação aos demais. A técnica permite a detecção de um grande número de pontos de mutação por reação, pois uma

simples mudança de base pode ocasionar a perda ou o ganho de um sítio de restrição e, por isso, estes marcadores são importantes para a saturação de mapas genéticos (VUYLSTEKE et al., 1999).

Quanto aos microssatélites (LITT; LUTY, 1989), também conhecidos como SSR (*Simple Sequence Repeats* - JACOB et al., 1991) ou STR (*Short Tandem Repeats* - EDWARDS et al., 1991), estes surgiram a partir da sugestão de Jeffreys e colaboradores (1988) de combinar a especificidade e a rapidez da PCR com a informatividade dos marcadores VNTR em humanos. Loci microssatélites são seqüências de um a seis pares de bases, chamadas de motivos, repetidas em tandem. A variação no número de repetições do motivo é a base do seu polimorfismo. Por outro lado, as seqüências que flanqueiam o microssatélite são conservadas, sendo utilizadas para o ancoramento dos *primers* e a amplificação do loco. Os produtos da PCR possuem mobilidade eletroforética que difere de acordo com o número de unidades repetidas no alelo. Estes marcadores possuem natureza codominante, são PCR-específicos, passíveis de transferência entre espécies próximas e mostram-se reproduzíveis entre diferentes laboratórios, o que os tem tornado importantes na construção e na integração de mapas genéticos (AKKAYA et al., 1995; BLAIR et al., 2003; MORETZSOHN et al., 2005; SONG et al., 2004).

Uma questão limitante para o uso dos microssatélites está relacionada ao seu desenvolvimento. A estratégia tradicional de desenvolvimento é bastante complexa e cara, envolvendo a construção de uma biblioteca genômica, hibridização com sondas repetitivas, seqüenciamento, busca de microssatélites, desenho e síntese de *primers* (OLIVEIRA, 2006).

No entanto, o aumento da disponibilidade de seqüências de DNA nos bancos de dados tem possibilitado a mineração de seqüências contendo microssatélites (LI et al., 2004; MORGANTE; HANAFEY; POWELL, 2002). Hoje, se sabe que boa parte dos microssatélites está contida em regiões expressas do genoma (revisado por LI et al., 2004) o que permitiu uma alternativa para o desenvolvimento de marcadores microssatélites a partir da análise de seqüências depositadas nos bancos de dados. Estas seqüências, em sua grande maioria são seqüências expressas (ESTs) e podem ser analisadas usando programas computacionais específicos para a busca de motivos repetidos em *tandem*. Depois de identificados os microssatélites nas ESTs, *primers* flanqueando as repetições podem ser desenhados para a amplificação do loco microssatélite.

Os marcadores microssatélites derivados de EST (SSR-EST) possuem vantagens adicionais por serem desenvolvidos a partir de regiões transcritas do genoma. Um marcador derivado de EST que esteja associado a um caráter de interesse pode ser parte do próprio gene controlando o caráter (THIEL et al., 2003). Outra vantagem está no fato de que as seqüências gênicas serem mais conservadas se comparadas às seqüências anônimas. Assim, estes marcadores derivados de genes têm grandes chances de serem transferidas para espécies relacionadas com aquela da qual foi gerada a biblioteca de EST (THIEL et al., 2003; VARSHNEY, R. K. et al., 2005).

O uso de SSR derivados de c-DNA ou EST tem sido relatado em espécies de diferentes famílias, incluindo uva (SCOTT et al., 2000), cana-de-açúcar (CORDEIRO et al., 2001b), trigo (EUJAYL et al., 2002), cevada (THIEL et al., 2003), batata (FEINGOLD et al., 2005), pimentão (SANWEN et al., 2001), soja (TIAN et al., 2004) e feijão-comum (BLAIR et al., 2003; YU, K. et al., 2000).

Visando estudar a transferibilidade de SSR-EST, Gao e colaboradores (GAO et al., 2003) avaliaram a utilidade de 597 pares de *primers* desenhados a partir de seqüências ESTs de trigo em outras espécies. Os autores encontraram 478 *primers* funcionais dos quais 255 amplificaram fragmentos em arroz, milho e soja, confirmando o alto nível de transferibilidade entre espécies para este tipo de marcadores.

Yu e colaboradores (YU, K. et al., 2000) atribuíram 15 marcadores SSR, identificados a partir seqüências de DNA de feijão depositadas no banco de dados *GenBank*, no mapa consenso de ligação de *P. vulgaris*. Depois disto, Blair e colaboradores (BLAIR et al., 2003) mapearam microssatélites derivados de genes em uma população oriunda do cruzamento DOR364 x G9833, e atribuíram estes marcadores ao mapa consenso de Davis (FREYRE et al., 1998) e ao mapa de Paris (VALLEJOS; SAKIYAMA; CHASE, 1992). Estes marcadores forneceram a base para ancorar e alinhar os mapas genéticos.

A maior limitação dos SSR-EST reside no número de seqüências depositadas nos bancos de dados. Em feijão-comum, a quantidade de informação disponível para este tipo de abordagem, até pouco tempo atrás, era bastante limitada. Melotto e colaboradores (2005) relataram que em setembro de 2004 haviam apenas 1.828 entradas de *P. vulgaris* no NCBI (*National Center for Biotechnology Information*). No entanto, um esforço no sentido de desenvolver ferramentas genômicas para a espécie foi iniciado com a criação do consórcio internacional 'Phaseomics', e

várias bibliotecas de cDNA para feijão-comum e espécies relacionadas estão sendo construídas (BROUGHTON et al., 2003). Na ESALQ/USP, cerca de 5.000 EST já foram seqüenciadas e, numa busca prévia, foram identificadas mais de 240 seqüências candidatas a SSR (MELOTTO et al., 2005). Outras bibliotecas de EST em feijoeiro foram também disponibilizadas (RAMIREZ et al., 2005). Neste último trabalho foram seqüenciadas 21 mil EST a partir de cinco diferentes bibliotecas de cDNA de feijoeiro que geraram quase oito mil transcritos. Atualmente, existem seqüências de 129 genes e mais de 83 mil EST de feijão depositadas no banco de dados (NCBI, 2008), uma considerável fonte de marcadores moleculares para a espécie.

2.3 Mapas genéticos

2.3.1 Conceito e breve histórico

Dois locos que não segregam de forma independente geram o que se conhece por desequilíbrio de ligação, significando que estão ligados fisicamente. A ligação é a base do mapeamento genético. Este fenômeno foi descoberto por Bateson e Punnett (1906), no início do século passado, quando verificaram a falta de independência na segregação de dois genes, em ervilhas. Em seguida, Sturtevant (1913) sugeriu que se poderia usar a frequência de aparecimento dos indivíduos recombinantes, ou seja, com fenótipos diferentes dos pais, como uma medida da fração de recombinação ($=r$) entre os genes sob estudo. Usando a premissa de que as permutações ou *crossing overs* ocorrem ao acaso ao longo do cromossomo, a possibilidade de recombinação é maior para genes mais distantes do que para aqueles mais proximamente ligados. Esta premissa se constituiu na base do mapeamento genético, e a taxa de recombinação entre os locos é usada para ordenar e posicionar os genes nos grupos de ligação.

Os primeiros mapas genéticos foram construídos observando-se a segregação de marcadores morfológicos e citológicos. No entanto, estes marcadores dependem da geração de estoques cromossômicos por aberrações cromossômicas, como aneuploidias, translocações, deleções e inverções (CARNEIRO; VIEIRA, 2002), os quais são restritos a espécies de amplo conhecimento citológico. Em trigo, por exemplo, os estoques de substituição cromossômica permitiram que genes que controlam as principais proteínas de armazenamento fossem mapeados nos cromossomos dos grupos homeólogos 1 e 6, como mostrado na revisão de Gupta, Kulwal e Rustgi (2005). Já, os marcadores morfológicos contribuíram grandemente para a construção dos primeiros mapas genéticos, incluindo o de feijão (BASSETT, 1991).

Entre a décadas de 1950 e 1960 foram descobertas as isoenzimas (MARKERT; MOLLER, 1959), que resultam da presença de mais de um gene codificando isoformas de uma mesma proteína. São classificadas como marcadores bioquímicos e se mostraram interessantes para a construção de mapas por terem um controle genético simples e, na maioria das vezes, possuírem herança codominante (CARNEIRO; VIEIRA, 2002). Assim, diversos mapas de ligação foram desenvolvidos com marcadores isoenzimáticos (PHILIPP; WEHLING; WRICKE, 1994; TANKSLEY; RICK, 1980; VALLEJOS; SAKIYAMA; CHASE, 1992). Estes locos marcadores estão disponíveis de forma limitada no genoma (MURPHY et al., 1990), de modo que são insuficientes para fins de mapeamento, principalmente quando se pretende cobrir amplamente o genoma (CAIXETA et al., 2006; FERREIRA; GRATTAPAGLIA, 1998).

A partir da década de 1980, com o surgimento dos marcadores de DNA, o mapeamento genético tornou-se ilimitado a todas as espécies vegetais e animais. Os marcadores de DNA (ou marcadores moleculares) são baseados em polimorfismos nas seqüências de DNA. Essencialmente, os marcadores moleculares são livres de efeitos pleiotrópicos e não são influenciados pelo estágio de desenvolvimento do organismo, nem pelo ambiente. Além disso, geralmente, não têm efeito no fenótipo porque eles são reflexos de variações naturais presentes nas seqüências de DNA. Assim, um mapa de ligação detalhado pode ser construído utilizando apenas um cruzamento e uma população de mapeamento (KUMAR, 1999). Estas propriedades dos marcadores moleculares fizeram deles candidatos ideais para uso em mapeamento genético.

A partir dos marcadores moleculares, espécies que já possuíam mapas genéticos desenvolvidos tiveram o número de marcadores mapeados rapidamente ampliado. Em feijão, por exemplo, até 1991 haviam 46 marcadores mapeados em 13 grupos de ligação (BASSETT, 1991). Já nos três anos seguintes, 443 marcadores RFLP e 102 marcadores RAPD foram mapeados (ADAM-BLONDON et al., 1994; NODARI et al., 1993a; VALLEJOS; SAKIYAMA; CHASE, 1992). Mesmo espécies tropicais, para as quais não havia descrição de ligação entre qualquer tipo de marcador, tiveram mapas genéticos rapidamente desenvolvidos a partir desses marcadores, como o eucalipto (GRATTAPAGLIA; SEDEROFF, 1994) e o maracujá (CARNEIRO et al., 2002).

2.3.2 Construção de mapas genéticos

Mapa de ligação é a representação linear da posição dos genes ou marcadores dentro dos grupos de ligação que compõem o complemento haplóide de uma espécie. As posições no mapa são inferidas a partir das estimativas das frequências de recombinação entre os locos marcadores. Quando um grande número de marcadores está segregando em uma população de mapeamento, a análise desta segregação através de procedimentos estatísticos adequados permite a construção de um mapa de ligação. A metodologia de construção de um mapa integra uma série de procedimentos que incluem: (i) o desenvolvimento de populações segregantes adequadas; (ii) a genotipagem dos locos na população de mapeamento e nos seus progenitores (iii) a verificação da segregação individual dos locos; e (iv) a análise do conjunto de dados gerados para a verificação de ligação, a estimativa da distância e a ordenação dos marcadores nos grupos de ligação (TANKSLEY, 1993).

Existem variados tipos de populações segregantes que podem ser utilizadas para a construção de mapas. As comumente usadas são aquelas derivadas do cruzamento de linhagens puras e contrastantes, que podem derivar de endogamia ou de haplóides duplos (TANKSLEY, 1993). A primeira geração filial (F_1) é autofecundada ou retrocruzada com um dos progenitores dando origem às populações segregantes F_2 , RC_1 ou RIL. A escolha do tipo de população mais apropriada ao mapeamento está relacionada com o tipo de marcador, a finalidade do mapa a ser construído e as características da espécie, como tipo de reprodução e duração do ciclo de vida. O máximo de informação genética é alcançado quando se usa uma população F_2 genotipada com marcadores codominantes (STAUB; SERQUEN; GUPTA, 1996). Populações formadas por linhagens endogâmicas recombinantes (RIL – *Recombinant Inbred Lines*) possuem a vantagem de poderem ser “eternizadas”, pois, novos marcadores podem sempre ser acrescentados, gerando mapas de alta resolução, e caracteres de interesse podem ser avaliados em diferentes locais e épocas, requisitos importantes no mapeamento de QTL, principalmente quando o objetivo final é a seleção assistida por marcadores (KELLY et al., 2003).

Uma vez tendo os dados de genotipagem dos locos marcadores na progênie segregante e respectivos progenitores, seguem-se, então, as análises estatísticas. Como, geralmente, um elevado número de marcadores é usado na construção dos mapas de ligação, as análises são realizadas com o auxílio de programas computacionais. Para isso, métodos estatísticos foram

desenvolvidos para as principais populações de mapeamento e implementados em programas computacionais, como Mapmaker/Exp (LANDER, E. et al., 1987), Crimap (GREEN; FALLS; CROOKS, 1990), GMendel (LIU, B. H.; KNAPP, 1992), JoinMap (VAN OOIJEN; VOORRIPS, 2001), Multimap (MATISE; PERLIN; CHAKRAVARTI, 1994), Outmap (BUTCHER et al., 2002), AntMap (IWATA; NINOMIYA, 2006) e OneMap (MARGARIDO; SOUZA; GARCIA, 2007).

São necessários, basicamente, três procedimentos estatísticos: (i) verificação do padrão Mendeliano de segregação de cada loco marcador; (ii) análise da ligação entre todos os pares de marcadores para a formação dos grupos de ligação; e (iii) estimação da ordem e distância dos locos marcadores dentro dos grupos de ligação.

Primeiramente, é realizada uma análise da segregação dos marcadores individualmente. Espera-se que cada loco marcador se comporte de acordo com o padrão Mendeliano típico para a população e o tipo de marcador. A estatística mais usada é o qui-quadrado (χ^2), no qual a hipótese H_0 não admite a ocorrência de distorção. Bearzoti (2000) recomenda o descarte dos locos apresentando distorção de segregação, visando a não comprometer a qualidade do mapa, pois os locos com distorção afetam os testes estatísticos usados para detectar a ligação.

Diferentes tratamentos podem ser realizados para a formação dos grupos de ligação. Estatisticamente, grupo de ligação pode ser definido como um conjunto de locos herdados juntos (LIU, B.H., 1998). As estatísticas usadas para esta análise podem ser o χ^2 , a razão de verossimilhança, um limiar para a frequência de recombinação, dentre outras (MOLLINARI, 2007). Nesta fase, todos os pares de marcadores são testados; ao final, a propriedade transitiva é usada para o estabelecimento dos grupos de ligação, ou seja, se os marcadores A e B não estão ligados, mas o marcador B está ligado ao C e C ligado ao A, então os marcadores B, C e A pertencem ao mesmo grupo de ligação.

Uma vez formados os grupos de ligação e de posse das estimativas das frequências de recombinação entre os locos marcadores, é realizada a análise visando à ordenação e à estimação das distâncias entre os marcadores nos grupos de ligação. A ordenação dos locos marcadores tem sido tratada como um processo que visa a minimizar o número de *crossing-overs* (CARNEIRO; VIEIRA, 2002). Diversos critérios foram desenvolvidos a fim de se comparar e julgar qual a melhor ordem. Alguns métodos usam estatísticas a partir da análise de dois pontos, por exemplo, a soma mínima dos coeficientes de recombinação adjacentes (SARF, *Sum of Adjacent*

Recombination Fractions - FALK, 1989), o produto mínimo dos coeficientes de recombinação adjacentes (PARF, *Product of Adjacent Recombination Fractions* - WILSON, 1988) ou a soma mínima dos LODs adjacentes (SALOD, *Sum of Adjacent LOD Scores* - WEEKS; LANGE, 1987). Por outro lado, tem-se o método da verossimilhança. Lander e Green (LANDER, E. S.; GREEN, 1987) usaram a verossimilhança como critério para a comparação entre diferentes ordens. Neste caso, são realizadas análises multiponto, onde a verossimilhança é obtida via cadeia de Markov oculta, a qual usa a informação de todos os marcadores, simultaneamente, para determinar a ordem mais verossímil e estimar as frações de recombinação entre eles. Algoritmos baseados neste método foram implementados em vários programas de mapeamento genético, por exemplo, Mapmaker/Exp, Crimap e Multimap.

Finalmente, as frequências de recombinação são transformadas em distância genética de mapa através de funções de mapeamento, para que haja aditividade entre as distâncias dos marcadores no mapa. Uma unidade de medida de distância de mapa é definida como a distância entre dois locos em que se espera a ocorrência de uma permuta por gameta, por geração (WEIR, 1996). Denominada Morgan e expressa em centiMorgans (cM), ela é uma função da fração de recombinação e possui caráter aditivo. Assim, as funções de mapeamento transformam as frequências de recombinação em distâncias aditivas (LYNCH; WALSH, 1998). As principais são a função de Haldane (HALDANE, 1919), que admite que a ocorrência de permuta se dá na ausência de interferência, seguindo uma distribuição de Poisson, e a função de Kosambi (KOSAMBI, 1944), que admite a ocorrência de permutas próximas como eventos não independentes. Função de Kosambi admite interferência completa entre regiões arbitrariamente próximas, que decresce com o aumento da distância, sendo nula para locos independentes.

2.3.3 Aplicações dos mapas genéticos

Os mapas genéticos fornecem referenciais genéticos ao longo do genoma e suas principais aplicações são: (i) contribuir para o conhecimento básico da organização e evolução do genoma; (ii) permitir estudos de mapeamento comparativo; e (iii) permitir a localização de regiões controlando caracteres tanto de herança simples quanto complexa.

Através da comparação de mapas genéticos é possível se estudar os processos evolutivos do genoma, verificando-se a sintonia entre espécies e as alterações ocorridas durante a evolução, por exemplo, rearranjos, duplicações e deleções de segmentos cromossômicos. Em meados da

década de 1980, quando os marcadores RFLP foram usados, pela primeira vez, em plantas, ficou claro que sondas RFLP poderiam ser transferidas entre espécies para fornecerem âncoras que permitiriam a comparação entre os genomas. Bonierbale, Plaisted e Tanksley (1988) mostraram que os mapas de tomate e batata eram muito similares. Da mesma forma, Chao, Sharp e Gale (1988) mostraram que três genomas diplóides que formam o trigo hexaplóide tinham mantido quase idêntica a ordenação de seus genes. Estes estudos deram os primeiros sinais de que o arranjo linear dos genes em plantas pode ter permanecido conservado durante um longo período evolutivo. Assim, a extrapolação de informações obtidas a partir de espécies-modelo (arabidopsis e arroz, por exemplo) para as espécies cultivadas menos estudadas pode servir de base para aumentar o conhecimento sobre elas (GALE; DEVOS, 1998). Estudos a cerca da fixação biológica de nitrogênio em leguminosas demonstraram que genes funcionalmente conservados ocupam posições sintênicas em diversas espécies (RIELY et al., 2004). Mesmo grandes famílias gênicas, e que evoluem rapidamente, como os genes de resistência do tipo NBS-LLR (*Nucleotide-Binding Site-Leucine-Rich Repeat*) podem apresentar sintenia nas leguminosas. Por exemplo, Zhu e colaboradores (ZHU et al., 2002) mostraram através de mapeamento comparativo entre *Medicago truncatula* e *Glycine max* e entre *M. truncatula* e *Pisum sativum* que vários blocos de locos gênicos de resistência se mostraram sintênicos entre essas espécies.

Uma importante utilidade dos mapas genéticos é, sem dúvida, o mapeamento de genes e QTL. A partir do mapeamento de QTL é possível estudar a base genética dos caracteres complexos (PATERSON, 1995). Do ponto de vista do melhoramento, é interessante que os mapas possam ser completamente anotados indicando a localização de genes e QTL, sejam eles para resistência a pragas e doenças, performance agronômica, adaptabilidade, síndrome da domesticação, qualidade, ou qualquer outro caráter importante. Estas informações podem ser usadas tanto para a obtenção de novas cultivares pela seleção assistida por marcadores, quanto no auxílio aos melhoristas pela melhor compreensão dos efeitos e modo de ação dos locos controlando os caracteres de interesse.

Em feijão, tem havido um grande interesse no mapeamento de genes e de caracteres quantitativos de importância agronômica, como resistência a pragas e doenças (MIKLAS, P. N. et al., 2006), tolerância a estresses abióticos (BEEBE et al., 2006), síndrome da domesticação (KOINANGE et al., 1996; PARK et al., 2000), nodulação (NODARI et al., 1993a) e caracteres relacionadas com a produtividade (TAR'AN; MICHAELS; PAULS, 2002). Estes trabalhos têm

aumentado o conhecimento sobre o controle genético de diversos caracteres, fornecendo informações a respeito do número de QTL envolvidos, seu efeito e sua localização nos grupos de ligação.

A seleção assistida por marcadores é a aplicação mais direta destes conhecimentos. O potencial de se mapear diferentes QTL afetando um caráter de interesse abre a possibilidade de se transferir simultaneamente vários QTL e combiná-los em uma cultivar melhorada (DITA et al., 2006).

Com relação à resistência a doenças, Mutlu e colaboradores (2005) usaram a seleção fenotípica combinada com a seleção assistida para desenvolver linhagens de feijão das classes pinto, a mais consumida no Estados Unidos, resistentes ao cretamento bacteriano comum. Ender, Terpstra e Kelly (2008) avaliaram a eficiência da seleção assistida para aumentar a resistência ao mofo branco em linhagens de feijão. Dois cruzamentos diferentes com a mesma fonte doadora de resistência foram avaliados. Este trabalho mostrou que a seleção realizada com base em QLT nos grupos de ligação B2 e B7 se mostrou mais eficiente que a seleção através de avaliações fenotípicas, suportando a eficácia da seleção assistida para locos de herança complexa. Além disso, os QTL introgrididos se mostraram estáveis nos dois *backgrounds* genéticos analisados.

Concibido e colaboradores (2003) mapearam um loco quantitativo associado ao aumento da produtividade a partir de uma introdução de *Glycine soja* e o introgridiram em seis *backgrounds* genéticos de *G. max*, usando retrocruzamento assistido por marcadores. A disponibilidade de um mapa genético possibilitou o isolamento de regiões específicas do genoma e a sua transferência para variedades comerciais com um mínimo de *linkage drag*. Esta pesquisa demonstrou o potencial de se usar germoplasma exótico com o auxílio de mapas genéticos para aumentar a produtividade em soja, uma cultura com limitada diversidade genética no germoplasma elite.

Outra aplicação dos mapas genéticos está no uso de informações a cerca da posição dos genes envolvidos na síndrome da domesticação. Estas informações podem facilitar a introgressão de diversidade genética a partir de linhagens silvestres, importantes fontes de novos alelos, no germoplasma adaptado (KELLY et al., 2003). Kelly e colaboradores (2003) apontam que o conhecimento prévio da ligação entre os genes de domesticação e aqueles controlando uma determinada característica agrônômica pode orientar os melhoristas a cerca da dificuldade de se

obter linhagens recombinantes carregando a característica de interesse sem a perda do fenótipo domesticado para arquitetura da planta, indeiscência, etc. Da mesma forma, é de grande ajuda o conhecimento da ligação genética entre quaisquer caracteres de interesse. Por exemplo, a ligação em repulsão entre QTL afetando a resposta ao vírus do mosaico dourado do feijoeiro e a resistência ao crestamento bacteriano foi observada no grupo de ligação B7 (MIKLAS, PHILLIP N. et al., 1996). Isto explica a consistente correlação fenotípica negativa observada entre estes dois caracteres no cruzamento ‘Dourado’ x ‘Xan 176’. Concluindo, esses conhecimentos podem colaborar no planejamento de cruzamentos a serem realizados, a fim de se obter mais rapidamente os resultados esperados pelos melhoristas.

2.3.4 Mapas genéticos em feijão

O feijão-comum é uma espécie diplóide, com $2n=2x=22$; seus cromossomos são extremamente pequenos quando comparados aos de outras espécies vegetais (MOSCONE et al., 1999), possuindo um genoma de tamanho estimado entre 686 e 750 milhões de bases (Mb) por conjunto haplóide (BENNETT; LEITCH, 2005).

O primeiro mapa de ligação para o feijão-comum foi publicado há 47 anos (LAMPRECHT, 1961). Este mapa inicial foi baseado em poucas dezenas de marcadores morfológicos. Durante as décadas de 70 e 80, diversos artigos reportando ligação entre diferentes marcadores morfológicos e bioquímicos foram publicados em feijão. Bassett (1991) publicou um mapa a partir da revisão destes trabalhos, incluindo alguns contestando as interpretações e a validação dos dados de Lamprecht (1961). Estas contradições a cerca da ligação entre alguns locos ocorriam devido aos marcadores usados não serem neutros. Assim, alguns marcadores só se expressavam em determinados genótipos, outras vezes possuíam efeitos epistáticos ou mesmo pleiotrópicos. Após o esforço de Bassett (1991) em revisar o mapa genético de feijão, 13 grupos de ligação foram formados. A nomenclatura de alguns genes marcadores foi uniformizada, assim como, outros marcadores foram retirados do mapa, por não haver evidências de *crossing over* suficientes para suportar sua ligação genética.

Mais tarde, com o desenvolvimento dos marcadores RFLP, Vallejos, Sakiyama e Chase (1992) publicaram o primeiro mapa usando marcadores moleculares em feijão. Os autores analisaram 224 locos RFLP, além de alguns marcadores morfológicos e protéicos em uma população de retrocruzamento. A geração F_1 oriunda do cruzamento entre a linhagem ‘XR-235-

1', de origem Mesoamericana, e a cultivar 'Calima', de origem Andina, foi retrocruzada com a linhagem 'XR-235-1' para originar a população de mapeamento. Uma biblioteca genômica de fragmentos *Pst I* enriquecida com seqüências de cópia única foi utilizada para o desenvolvimento das sondas (CHASE; ORTEGA; VALLEJOS, 1991). Ao final, 216 RFLP foram ordenados em 11 grupos de ligação, cobrindo 960 cM do genoma de feijão.

Em seguida, Nodari e colaboradores (1993a) usaram 15 clones de genes conhecidos e 115 clones gerados a partir de duas bibliotecas genômicas para revelar marcadores RFLP. Estes, juntamente com mais 8 RAPD e alguns marcadores morfológicos e bioquímicos, foram usados para construir um mapa de ligação para uma população F₂ derivada do cruzamento entre as cultivares 'Bat 93', de origem Mesoamericana, e 'Jalo EEP558', de origem Andina. O mapa final possuiu 143 locos distribuídos em 15 grupos de ligação, e um comprimento total de 827 cM.

Um terceiro mapa baseado em marcadores moleculares foi construído a partir de uma população de retrocruzamento ('Ms8EO2' x 'Corel') (ADAM-BLONDON et al., 1994). Baseado em 51 locos RFLP, 100 RADP, dois SCARs (*Sequence-Characterized Amplified Regions*), e 4 marcadores morfológicos, um mapa de 567,5 cM foi gerado. Além destes, outros mapas genéticos foram, rapidamente, desenvolvidos para o feijão-comum (JUNG et al., 1996; JUNG et al., 1997; KOINANGE et al., 1996; MIKLAS, PHILLIP N. et al., 1996).

A construção de mapas a partir de diferentes populações de mapeamento se deve, principalmente, à multiplicidade dos objetivos de melhoramento em feijão, pois nenhuma população, por mais divergente que seja, segrega para todos os caracteres de interesse (KELLY et al., 2003).

Beebe e colaboradores (2006), por exemplo, utilizaram uma população de RIL oriunda do cruzamento entre a cultivar 'DOR 364' e a landrácea 'G19833' para construir um mapa com a finalidade de mapear QTL relacionados à aquisição de fósforo e à arquitetura de raiz. O genótipo 'G19833' possui alta capacidade de acumular fósforo em solos deficientes neste nutriente e por isso foi escolhido para este cruzamento. Já Tar'an, Michaels e Pauls (2002) usaram a população derivada do cruzamento 'OAC Seaforth' x 'OAC 95-4' para gerar um mapa genético e mapear 14 caracteres quantitativos relacionados com a produtividade e seus componentes e com a arquitetura da planta em feijão.

Da mesma forma, outras populações de mapeamento foram usadas para a construção de mapas com a finalidade de se mapear caracteres específicos aos objetivos de cada grupo, como

por exemplo, resistência ao cretamento bacteriano (ARIYARATHNE et al., 1999; NODARI et al., 1993b; VALLEJOS; SAKIYAMA; CHASE, 1992), *Fusarium* (SCHNEIDER; GRAFTON; KELLY, 2001), mofo branco (KOLKMAN; KELLY, 2003; MIKLAS, P. N. et al., 2001), à antracnose (ADAM-BLONDON et al., 1994; RODRIGUEZ-SUAREZ et al., 2007) e mancha angular (MELO; SANTOS; FERREIRA, 2002b; NODARI et al., 1993b).

Visando permitir a comparação entre diferentes mapas, Freyre e colaboradores (1998) iniciaram um esforço para a criação de um mapa consenso para a espécie. Eles analisaram a segregação de um conjunto de marcadores comuns aos diferentes mapas em uma população núcleo de mapeamento. A população escolhida por Freyre e seus colaboradores (1998) foi uma população de RIL resultante do cruzamento entre os genótipos ‘Bat 93’ e ‘Jalo EEP558’ (BJ). Esta foi escolhida por mostrar uma grande diversidade para vários tipos de marcadores (NODARI et al., 1993a) e por segregar para múltiplas interações microorganismo-hospedeiro (NODARI et al., 1993b). O mapa consenso gerado integrou ao mapa BJ (NODARI et al., 1993a) marcadores dos mapas da Florida (VALLEJOS; SAKIYAMA; CHASE, 1992) e de Paris (ADAM-BLONDON et al., 1994). O mapa integrado foi gerado a partir de 563 marcadores e possuiu um comprimento final de 1226 cM. Os marcadores comuns permitiram estabelecer a correspondência entre os grupos de ligação dos três mapas. Estudos subsequentes continuam a mapear novos marcadores nesta população, fazendo deste mapa um mapa referência para a espécie (BLAIR et al., 2003; GRISI et al., 2007; PAPA et al., 2005; YU, K. et al., 2000).

Pedrosa e colaboradores (2003), usando hibridização fluorescente *in situ* de *pools* de sondas RFLP geneticamente ligadas, integraram o mapa de ligação da Florida (VALLEJOS et al., 1992) ao mapa cromossômico mitótico de *P. vulgaris*. Os *pools* de sondas usados permitiram a atribuição de todos os grupos de ligação aos cromossomos de dois cultivares de feijão, ‘Saxa’ e ‘Calima’.

Apesar da nomenclatura dos cromossomos de feijão ter sido primeiramente proposta por Moscone e colaboradores (1999), devido à alta variação no tamanho dos cromossomos dentro da espécie (PEDROSA-HARAND et al., 2006) e da grande utilização pela comunidade científica da nomenclatura dos grupos de ligação proposta por Freyre e colaboradores (1998), a designação B1 a B11, usada para denominar os grupos de ligação do mapa consenso, passou a ser a nomenclatura para os cromossomos de feijão-comum (Figura 1) (PEDROSA-HARAND;

PORCH; GEPTS, 2008), evidenciando, mais uma vez, o mapa BJ como mapa referência da espécie.

Figura 1 - Idiograma dos cromossomos da variedade 'Calima' de *P. vulgaris* enumerada de acordo com a nomenclatura dos grupos de ligação do mapa consenso de Freyre et al. (1998). A posição dos clones RFLP usados para a atribuição dos grupos de ligação aos cromossomos é indicada em preto; a localização aproximada no mapa consenso de ligação dos *pools* de clones RFLP usados nas hibridizações *in situ* é indicada por setas; as marcas nas extremidades dos grupos de ligação representam os marcadores RFLP distais, segundo o mapa de Freyre et al. (1998). Fonte: Pedrosa-Harand; Porch; Gepts (2008)

3 MATERIAL E MÉTODOS

3.1 Material Vegetal

Para a caracterização dos SSR-EST foram usados 23 genótipos de *Phaseolus vulgaris*, sendo 17 linhagens de origem Mesoamericana, 6 linhagens de origem Andina (Tabela 1), e três espécies relacionadas *P. acutifolius*, *P. lunatus* e *Vigna sp.* Este conjunto de genótipos compreende uma amostra representativa dos genótipos cultivados de feijão-comum, além de incluir as linhagens parentais das duas populações de mapeamento usadas neste trabalho ('Bat 93' x 'Jalo EEP558' e 'Carioca' x 'Flor de Mayo').

Tabela 1 - Conjunto de genótipos representantes dos dois grupos gênicos de *Phaseolus vulgaris*

Genótipo	Grupo gênico	Genótipo	Grupo gênico
Bat 93	Mesoamericano	Baetão	Mesoamericano
Brasil 2	Mesoamericano	Tu	Mesoamericano
Porrillo	Mesoamericano	IAC UNA	Mesoamericano
Rio Tibaji	Mesoamericano	Sanilac	Mesoamericano
Mulatinho	Mesoamericano	Jamapa	Mesoamericano
Carioca	Mesoamericano	Pompadour	Andino
Puebla 152	Mesoamericano	Antioquia 8	Andino
Great Northern	Mesoamericano	Jalo EEP558	Andino
Flor de Mayo	Mesoamericano	Jabona	Andino
Garbancillo	Mesoamericano	Red Kidney	Andino
Apetito Blanco	Mesoamericano	CAL 143	Andino
Barbunya	Mesoamericano		

Para os estudos de mapeamento foram utilizadas duas populações de feijoeiro conforme descrito a seguir:

(i) População BJ: População derivada do cruzamento entre 'Bat 93' e 'Jalo EEP558', a qual foi primeiramente usada em F₂ para a construção do mapa de Davis, CA (NODARI et al., 1993a). Após o avanço de gerações, as linhagens recombinantes (RILs) foram usadas na integração dos principais mapas genéticos de feijão-comum (FREYRE et al., 1998). No presente trabalho, foram utilizados os progenitores 'Bat 93' e 'Jalo EEP558' e 74 RILs na geração F₈. O material foi gentilmente cedido pela Dra. Siu Mui Tsai (CENA/USP) e pela Dra. Rosana Pereira Vianello Brondani (EMBRAPA Arroz e Feijão). Os genitores desta população foram selecionados por

terem tido uma origem evolutiva divergente, ou seja, ‘Bat 93’ é uma linhagem de origem Mesoamerica desenvolvida no CIAT (Centro Internacional de Agricultura Tropical /Cali, Colombia) e ‘Jalo EEP558’ é uma seleção da landrácea ‘Jalo’, de origem Andina, obtida na Estação Experimental de Patos de Minas, Brasil, que segrega para diversas características, incluindo resistência ao vírus do mosaico comum do feijoeiro, ao cretamento bacteriano, à mancha angular, ferrugem e antracnose (NODARI et al., 1993a). De agora em diante esta população será chamada de BJ.

(ii) População CFM: População gerada a partir do cruzamento entre os genitores Carioca e Flor de Mayo, de agora em diante chamada população CFM. A partir da geração F_1 foram obtidas, por autofecundação, as gerações de F_2 até F_5 , sendo a população conduzida através do “método da população”. Na geração F_5 foram coletadas sementes de 196 plantas individuais, tomadas ao acaso, originando 196 famílias $F_{5:6}$ que multiplicadas originaram as linhagens recombinantes utilizadas no primeiro trabalho de mapeamento desta população (MELO; SANTOS; FERREIRA, 2002b). As sementes dos genitores e das linhagens foram gentilmente cedidas pelo Dr. João Bosco dos Santos (Universidade Federal de Lavras). No presente trabalho, foram usadas 94 das 196 RILs disponíveis. Esta população segrega para diversas características, incluindo reação ao oídio e à mancha-angular (MELO; SANTOS; FERREIRA, 2002b).

3.2 Extração e quantificação de DNA

3.2.1 Isolamento do DNA genômico

Para o isolamento do DNA genômico, todos os genótipos foram cultivados em casa de vegetação. O tecido foliar foi colhido dos primeiros trifólios abertos e mantido em nitrogênio líquido até a maceração. A extração foi feita de acordo com Doyle & Doyle (1987), com algumas modificações. Foram utilizados 2 g de tecido foliar por genótipo. Para o rompimento das paredes e membranas celulares, os tecidos foliares foram macerados em um almofariz com nitrogênio líquido usando um pistilo. O tecido pulverizado foi adicionado a um tubo de 15 ml contendo 8 ml de Tampão de Extração de DNA [100 mM de Tris HCl (pH 8,0); 1,4 M de NaCl; 2% (p/v) de CTAB; 20 mM de Na_2EDTA pH 8,0; 1% (p/v) de PVP 40.000] previamente aquecido a 65 °C. Após a completa homogeneização das amostras na solução, elas foram incubadas a 65 °C por uma hora em banho maria, com leve agitação a cada 15 min. Em seguida,

foi realizada a extração da parte orgânica das amostras pela adição de 6 ml de clorofórmio-álcool isoamílico (24:1) seguida de agitação manual por 5 min. e centrifugação por 15 min. a 3.000 g. A fase aquosa foi, então, transferida para novos tubos de 15 ml que foram tratados com RNase (25 ng/ml) e incubados por 2 h a 37 °C. Para a precipitação do DNA, foram adicionados 5 ml de isopropanol absoluto e os tubos mantidos por duas horas a 4 °C. O DNA precipitado foi enrolado em bastão de vidro e transferido para tubo de 2,0 ml. Foram realizadas duas lavagens: uma com uma Solução de Lavagem (7,5 M de AcNH_3 e 70% de etanol) e outra com etanol 70%. A ressuspensão do DNA foi realizada em 1,0 ml de TE [10 mM de Tris HCl (pH 8,0); 1 mM de Na_2EDTA pH 8,0], deixando a solução *overnight* a temperatura ambiente. A solução contendo DNA passou, em seguida, por duas extrações orgânicas: foi adicionado 1,0 ml de fenol saturado e o tubo gentilmente agitado e centrifugado por 10 min. a 10.000 g. A fase aquosa foi recuperada e adicionada a um novo tubo de 2,0 ml contendo 750 μl de clorofórmio. A suspensão foi gentilmente agitada e centrifugada por mais 10 min. a 10.000 g. A fração contendo DNA (cerca de 500 μl) foi transferida para novo tubo de 2,0 ml e o DNA foi precipitado pela adição de 1.250 μl de etanol absoluto e 250 μl de acetato de amônio (7,5 M). Os tubos foram mantidos por 20 min. a -20 °C para favorecer a precipitação. Em seguida, os tubos foram centrifugados por 5 min. a 8.000 g e o *pellet* lavado com 500 μl de álcool 70%. O álcool foi descartado e os tubos foram colocados para secar por 2 h à temperatura ambiente. Por fim, o DNA foi ressuspensionado em 200 μl de TE e estocado a -20 °C para uso posterior.

3.2.2 Quantificação do DNA

A concentração do DNA foi estimada em gel de agarose 1,0% (p/v) por comparações de sua fluorescência com aquelas de concentrações conhecidas de DNA do fago lambda, após coloração com brometo de etídeo (10 mg/ml) durante 15 min. O resultado foi documentado por fotografia sob luz UV usando equipamento de fotodocumentação (Geldoc – Bio-Rad). A partir desta quantificação, duas alíquotas de trabalho foram feitas, uma para o uso na amplificação dos marcadores SSR-EST (10 ng/ μl) e outra para digestão enzimática para uso nas etapas posteriores da técnica de AFLP (50 ng/ μl).

3.3 Marcadores SSR-EST

3.3.1 Desenvolvimento de marcadores SSR-EST

Uma coleção de 3.126 seqüências de *P. vulgaris*, isoladas a partir de *seedlings* inoculados e não inoculados com *Colletotrichum lindemuthianum*, foi usada como fonte de SSR-EST (MELOTTO et al., 2005). Uma busca por microssatélites perfeitos foi realizada para os motivos di-, tri- e tetranucleotídicos, repetidos pelo menos 5 vezes em tandem (MELOTTO et al., 2005).

3.3.1.1 Caracterização das EST contendo SSR

Todas as seqüências haviam sido originalmente anotadas por Melotto e colaboradores (2005), no entanto, aquelas contendo microssatélites foram novamente alinhadas contra o banco de seqüências de aminoácidos para serem melhor caracterizadas. Estas análises foram realizadas através da ferramenta BLASTx (ALTSCHUL et al., 1990) diretamente no sítio do NCBI. Ainda, as 219 seqüências foram caracterizadas de acordo com a posição do SSR dentro de cada EST, isto é, se o microssatélite estava localizado na região 3' não traduzida (3' UTR), 5' não traduzida (5' UTR) ou codante (CDS). Esta análise foi realizada através da busca das trincas putativas de início e fim da tradução em cada EST, assim como dos quadros de leitura que possuíam o melhor alinhamento de aminoácidos com seqüências similares depositadas no banco de dados *GenBank* não redundante, usando o algoritmo BLASTx (ALTSCHUL et al., 1990).

3.3.1.2 Classificação das SSR

Os microssatélites presentes nas 219 EST foram classificados quanto aos motivos repetidos. Os microssatélites dinucleotídicos foram agrupados em quatro motivos (AT, AG, AC e CG) e os trinucleotídicos em 10 motivos (GAA, AAT, CAT, AAC, TGG, AGG, TAG, TGC, CGG e ACG). Já os tetranucleotídicos foram somente sete, e foram agrupados em cinco motivos (AAAG, AATC, AATG, AATT e TGCA). Cada motivo incluiu todas as seqüências complementares e permutadas para as seqüências acima apresentadas entre parênteses, por exemplo, o motivo dinucleotídico AG representa as seqüências AG, GA, e os complementos CT e TC e o motivo trinucleotídico GAA representa as seqüências GAA, AAG, AGA e seus complementos TTC, TCT e CTT.

3.3.1.3 Seleção e desenho dos primers a partir das EST

As seqüências flanqueando os SSR nas EST foram usadas para desenhar *primers* entre 16 e 24 pares de bases. Para aumentar a chance de desenhar bons *primers*, primeiramente foi selecionada, entre as seqüências contendo SSR, aquelas em que o SSR esteve distante pelo menos 50 bases de ambas as extremidades 5' e 3' da seqüência. O programa Primer-3 (ROZEN; SKALETSKY, 2000) foi usado para gerar os *primers*. Para reduzir a ocorrência de bandas espúrias geradas durante a amplificação e permitir o uso de vários pares de *primers* sob as mesmas condições de PCR, dois critérios de estringência foram empregados no desenho: (1) T_m variando de 57 a 63 °C e (2) diferença máxima de 3 °C entre as T_m dos *primers* de cada par. A Figura 2 mostra as seqüências de um par de *primers* desenhado a partir da EST PVEPLE1002A08 contendo um SSR trinucleotídico. Antes da síntese, cada *primer* foi checado quanto à possibilidade de auto anelamento ou de formação de estruturas secundárias. Ainda, cada par de *primers* foi testado quanto à possibilidade de anelamento cruzado. Para tanto, foi utilizado o programa Gene Runner (SPRUYT; BUQUICCHIO, 1994).

Os *primers* desenhados foram sintetizados pela IDT[®] (Integrated DNA Technologies, Inc.) na escala de 25 nmoles cada.

PRIMER PICKING RESULTS FOR PVEPLB1002A08.g

```

OLIGO          start      len      tm      gc%      any      3' seg
LEFT PRIMER    37        20       60.66   45.00   4.00   0.00   TCCCTGGAACGAAAAATCGT
RIGHT PRIMER   232       20       59.17   50.00   2.00   0.00   CGAAGTAACAGCGGCTGAA
SEQUENCE SIZE: 499
INCLUDED REGION SIZE: 499

PRODUCT SIZE: 196, PAIR ANY COMPL: 4.00, PAIR 3' COMPL: 1.00
TARGETS {start, len}*: 101,21

1  TCAATGCTAAAGTGCAATGTAAATCTCATCTCTTTCTCTCCCTGGAACGAAAAATCGTGTC A
                                     >>>>>>>>>>>>>>>>>>
61 ACATATGCACAAGAAGACGGAACCTCACCAAACATCATCATCCCTTCTTCTCTTCTTCTTCT
                                     *****
121 TCCCCATCCCCATCCCCATCCATCACAAGCCCACACGCTCATCCACAGTAGGTCCCACTCC
    *
181 TTTGTGGGCCCTCTCCCGCGCTTCAATCCCTCTTGCACCCCTCTTTACTTCTGTCACGAG
                                     <<<<<<<<<<<<<<<<<<
241 GACGACGACGAAGACGAAGAAAAGGAACCTCTTGTTTTCTGACCCTGTTAAGATTGGACCT

301 TCCGATTTTTCATATTCTCCGAGTTGTAGGGCAGGCTGCCCTTTGCCAAAGTCTTTCTTCTT

361 AGAAAAAAAGCAAATTGCTTCGATGCTGACGCGGCTCTTCGCCATGAAGCTCATGACAAAG

421 GACACCCGTCATAAACAAAAACCATCTTGATTACATGAAGCGGAGAGCGCATATTCTCACT

481 AAACTTCTTTCACCCCTTTCA

```

Figura 2 - Arquivo de saída do programa *Primer 3*, mostrando dados dos *primers forward* (*left primer*) e *reverse* (*right primer*) desenhados para o microssatélite (CTT)₇ contido na sequência PVEPLE1002A08. São mostrados os sítios iniciais de ancoramento dos *primers* na EST, as Tm em °C, as porcentagens de GC, as sequências 5'-3' dos *primers*, o tamanho esperado do fragmento amplificado, assim como a sequência alvo (*) e a posição dos *primers forward* (>) e *reverse* (<) na EST

3.3.2 Condições de reação

As reações de amplificação dos SSR-EST para a caracterização dos marcadores em um conjunto de genótipos foram realizadas em um volume final de 20 µl em placas de PCR de 96 poços, usando um termociclador Gene Amp® PCR System 9700 (Applied Biosystems). Aproximadamente 20 ng de DNA molde foram adicionados a uma solução contendo 0,25 µM de ambos os *primer forward e reverse*, 200 µM de cada desoxinucleotídeo trifosfatado (dNTP), 10 mM de Tris-HCl (pH 8,3), 50 mM de KCl e 1,5 mM de MgCl₂ e 0,5 unidade de Taq DNA polimerase (Invitrogen). Os programas usados na PCR são mostrados na Tabela 2.

Já as reações para fins de mapeamento foram realizadas em um volume final de 16 µl de *mix*, contendo 20 ng de DNA; 8,0 µl de Master Mix da Promega [0,4 unidade de Taq DNA polimerase, 0,2 mM de cada dNTP, 10 mM de Tris-HCl (pH 8,3), 50 mM de KCl e 1,5 mM de

MgCl₂] e 0,2 µM de cada *primer*. A amplificação foi efetuada em termociclador Gene Amp[®] PCR System 9700 (Applied Biosystems), a partir de um dos quatro programas descritos na Tabela 2, dependendo do par de *primers*.

Um conjunto de 15 *primers* SSR-genômicos, previamente selecionados por apresentar polimorfismo em ambas as populações (CAMPOS, 2007; comunicação pessoal), também foi usado na genotipagem das duas populações de mapeamento. As condições de amplificação seguiram conforme descrito nas publicações originais (CAMPOS et al., 2007; BENCHIMOL et al., 2007).

Tabela 2 - Programas usados para amplificar os locos SSR-EST de feijão. Estão mostradas as condições de desnaturação da dupla fita de DNA, anelamento dos *primers*, extensão da nova fita de DNA, e o número de ciclos (Nº ciclos) usados em cada programa

Programa*	Desnaturação	Anelamento	Extensão	Nº ciclos
1	94 °C / 30 seg	65 °C / 45 seg (-1°C por ciclo)	72 °C / 30 seg	13 X
	94 °C / 30 seg	52 °C / 30 seg	72 °C / 45 seg	25 X
2	94 °C / 30 seg	60 °C / 45 seg (-1°C por ciclo)	72 °C / 30 seg	10 X
	94 °C / 30 seg	50 °C / 30 seg	72 °C / 45 seg	25 X
3	94 °C / 30 seg	55 °C / 45 seg (-1°C por ciclo)	72 °C / 30 seg	10 X
	94 °C / 30 seg	45 °C / 30 seg	72 °C / 45 seg	25 X
4	94 °C / 30 seg	45 °C / 45 seg (+1°C por ciclo)	72 °C / 30 seg	10 X
	94 °C / 30 seg	55 °C / 30 seg	72 °C / 45 seg	25 X

* Cada programa foi constituído por dois conjuntos de ciclos, um primeiro no qual a temperatura de anelamento variou a cada ciclo e um segundo no qual a temperatura de anelamento foi mantida constante. Em todos os programas houve uma desnaturação inicial a 94°C por 3 min. e uma extensão final de 5 min. a 72°C; os ciclos foram corridos em um termociclador GenAmp[®] PCR-System 9700 (Applied Biosystems).

3.3.3 Eletroforese e visualização dos locos

Os produtos resultantes das reações de amplificação de uma parte das amostras foram primeiramente separados em gel de agarose (1,5%) e corados em banho de brometo de etídeo (50 ng/l) para a visualização das bandas antes da corrida em gel de poliacrilamida desnaturante. Aqueles locos que estavam claramente visíveis em gel de agarose foram, então, separados em gel de sequenciamento [acrilamida/bisacrilamida (19:1) 6%, uréia 7,5 M, tampão TBE 1X (89 mM Tris base; 89 mM de ácido bórico; 2,5 mM EDTA 2Na)] de 0,4 mm de espessura.

A eletroforese se deu a uma potência constante de 80 W por cerca de 3 h em um aparato de eletroforese Sequi-Gen[®] (Bio-Rad). *Ladders* de 10 e 100 pb foram usados como marcadores de peso molecular.

A coloração foi realizada com nitrato de prata em cinco etapas de acordo com Creste, Tulmann-Neto e Figueira (2001), com algumas modificações como descrito abaixo.

- (i) Fixação: a fixação do gel foi feita por imersão em solução de etanol 10% e ácido acético 1%, sob lenta agitação por 10 min., seguida de uma lavagem com água destilada e deionizada, durante 1 min., sob agitação.
- (ii) Oxidação: o procedimento de pré-tratamento com uma solução de ácido nítrico 1,5% foi realizado durante 3 min., sob lenta agitação, após a qual seguiu-se uma nova lavagem do gel com água destilada deionizada durante 1 min., sob agitação.
- (iii) Impregnação com prata: a impregnação do gel foi realizada com a solução de nitrato de prata (AgNO_3) 0,2% durante 20 min., sob agitação, e protegida da luz. Em seguida, foram realizadas duas novas lavagens com água destilada deionizada, com duração de 30 seg. cada.
- (iv) Revelação: para revelação dos géis foram utilizados 2,0 l de solução contendo 60 g de Na_2CO_3 e 2,0 ml de formaldeído (37%). Nesta etapa, 1,0 l da solução foi aplicado sobre o gel até começarem a surgir bandas. Após este passo, a solução foi descartada e outro litro foi adicionado para finalizar o processo de revelação. Após a revelação da placa, a solução foi descartada e o bloqueio foi realizado em seguida.
- (v) Bloqueio: o bloqueio da revelação se deu em solução de ácido acético 5% durante 5 min., seguindo-se uma nova lavagem com água destilada deionizada durante 1 min. Após este passo, a placa foi colocada em posição vertical em um local seco, arejado e à temperatura ambiente para a secagem do gel, possibilitando posterior análise das bandas.

Os géis foram digitalizados por um scanner Color Page-HR3200 (Genius) e as imagens armazenadas em computador para posterior leitura (genotipagem).

3.3.4 Análise dos locos amplificados

3.3.4.1 Caracterização dos marcadores SSR-EST

Os SSR-EST desenvolvidos foram amplificados em um conjunto de genótipos para a caracterização dos marcadores. Após a separação dos produtos de amplificação em gel de sequenciamento, os locos foram manualmente lidos e os dados dos SSR-EST foram organizados em uma planilha.

O PIC (*Polymorphic Information Content*) foi usado para medir a informatividade de cada um dos locos gerados, de acordo com a eq. (1) (ANDERSON et al., 1993).

$$PIC = 1 - \sum_{i=1}^k P_i^2 \quad (1)$$

Na eq. (1), k é o número total de alelos detectado por um SSR-EST e P_i a frequência do i -ésimo alelo detectado na amostra estudada.

Os dados de todos os locos foram usados para a análise de agrupamento dos genótipos avaliados. Uma matriz de distâncias genéticas foi gerada usando a distância de Roger's modificada (WRIGHT, 1978) através do pacote TFPGA (MILLER, 1997). Para o agrupamento usou-se o método conhecido por UPGMA (*Unweighted Pair-Group Method with Arithmetical Averages*). Os dados de agrupamento foram, então, usados para a geração de um dendograma pelo programa *Drawgram* do pacote computacional PHYLIP (FELSENSTEIN, 2004).

3.3.4.2 Seleção de locos polimórficos nas populações BJ e CFM

Os locos SSR-EST amplificados a partir dos genitores das duas populações de mapeamento tiveram o tamanho de seus alelos anotados. Aqueles locos que se prestaram para mapeamento, apresentando polimorfismo em pelo menos uma das populações utilizadas, foram, então, amplificados em toda a população. Os dados de segregação foram anotados em uma planilha e usados na construção dos mapas genéticos.

3.4 Marcadores AFLP

3.4.1 Geração de marcadores AFLP

O desenvolvimento de marcadores AFLP para *Phaseolus vulgaris* envolveu os passos propostos por Vos e colaboradores (VOS et al., 1995) com algumas modificações propostas por Lopes e colaboradores (2006), conforme descrito a seguir:

(i) Restrição do DNA: cerca de 200 ng de DNA foram digeridos pela combinação de duas enzimas, uma de corte raro (*Eco R-I* ou *Pst-I*) e uma de corte freqüente (*Mse-I*). A digestão se deu em um volume final de 20 µl contendo 5 U de *Eco-RI* ou *Pst-I* (Promega), 5 U de *Mse-I* (New England Biolabs), 2 µl de tampão de digestão universal *One Phor All* 10X (Amersham/GE) e 2 µg de BSA (New England Biolabs). A solução de digestão foi incubada em termociclador Gen Amp® PCR System 9700 Thermocycler (Applied Biosystems) durante 4 h a 37 °C para a restrição do DNA e, em seguida, por 20 min. a 65 °C para a inativação das enzimas. A fixação do gel foi feita por imersão em solução de etanol 10% e ácido acético 1%, sob lenta agitação, por 10 min., seguida de uma lavagem com água destilada e deionizada, durante 1 min., sob agitação.

(ii) Ligação de adaptadores: adaptadores específicos (Tabela 3) foram ligados às extremidades dos fragmentos de DNA gerados pela restrição. Os adaptadores de fita dupla foram previamente preparados misturando-se quantidades equimolares dos dois oligos de fita simples (50 µM para adaptadores *Mse-I*, 5 µM para adaptadores *Eco-RI* e *Pst-I*). A solução foi aquecida a uma temperatura de 90°C por 10 min. e deixada resfriar lentamente. A ligação dos adaptadores às extremidades dos fragmentos foi obtida adicionando-se 20 µl de uma solução de ligação [1X de tampão de ligação T4 DNA ligase (New England Biolabs), 1 µM de adaptador *Mse-I*, 0,1 µM de adaptador *Eco-RI* ou *Pst-I* e 1 U de T4 DNA ligase (New England Biolabs)] aos 20 µl da solução de digestão do passo anterior. A ligação foi realizada por 14 h a 17°C; os tubos foram incubados em termociclador Gen Amp® PCR System 9700 Thermocycler (Applied Biosystems).

(iii) Reação de pré-amplificação: para a amplificação dos marcadores AFLP, primeiramente, foi realizada uma pré-amplificação com *primers* específicos, acrescidos de uma base seletiva. A reação foi realizada em um volume final de 20 µl contendo 2,5 U de Taq DNA polimerase (Promega), 1X tampão da Taq DNA polimerase (Promega), 1,5 µM de MgCl₂, 200 µM de dNTPs, 125 nM de cada *primer* (E+A com M+C ou P+A com M+G, mostrados na Tabela 3), e 3 µl de DNA digerido e ligado. Este coquetel foi incubado a 94 °C por 2 min. sendo, em seguida, submetido a 26 ciclos nas seguintes condições: 94 °C por 1 min., 56 °C por 1 min., 72 °C por 1

min.; ao final, foi mantido por mais 5 min. a 72 °C para a extensão final das fitas de DNA. Os produtos da reação foram diluídos 10X em água ultrapura e armazenados a -20 °C.

(iv) Amplificação seletiva: combinações de *primers* específicos para os adaptadores acrescidos de 3 bases seletivas na extremidade 3' foram utilizadas na amplificação seletiva. A reação foi realizada em um volume final de 15 µl contendo 1,0 U de Taq DNA polimerase (Promega), 1X tampão da Taq DNA polimerase (Promega), 1,5 µM de MgCl₂, 200 µM de dNTPs, 135 nM do iniciador Eann ou Pann (Tabela 3), 165 nM do iniciador Menn ou Mgnn (Tabela 3) e 5 µl de DNA pré-amplificado. O coquetel foi incubado a 94 °C por 2 min.; em seguida, foi submetido a 12 ciclos nas seguintes condições: 94 °C por 30 seg., 65 °C por 30 seg., 72 °C por 1 min., e a mais 23 ciclos nas seguintes condições: 94 °C por 30 seg., 56 °C por 30 seg., 72 °C por 1 min.; ao final, foi mantido por mais 5 min. a 72 °C para a extensão final das fitas de DNA.

Tabela 3 - Sítios de restrição, seqüências de adaptadores e de iniciadores usados para as análises de AFLP

Enzima	Sítio de restrição	Especificação	Seqüência (5' – 3')*
<i>Eco</i> -RI	5' GAATTC 3' 3' CTTAAG 5'	Adaptador	5'-CTCGTAGACTGCGTACC-3' 5'-AATTGGTACGCAGTCTAC-3'
		<i>Primer</i> E+N	5'-GACTGCGTACCAATTCN-3'
		<i>Primer</i> Ennn	5'-GACTGCGTACCAATTCNNN-3'
<i>Pst</i> -I	5' CTGCAG 3' 3' GACGTC 5'	Adaptador	5'-CTCGTAGACTGCGTACATGCA-3' 5'-TGTACGCAGTCTAC-3'
		<i>Primer</i> P+N	5'-GACTGCGTACATGCAGN-3'
		<i>Primer</i> Pnnn	5'-GACTGCGTACATGCAGNNN-3'
<i>Mse</i> -I	5' TTAA 3' 3' AATT 5'	Adaptador	5'-GACGATGAGTCCTGAG-3' 5'TACTCAGGACTCAT-3'
		<i>Primer</i> M+N	5'-GATGAGTCCTGAGTAAN-3'
		<i>Primer</i> Mnnn	5'-GATGAGTCCTGAGTAANNN-3'

* N: nucleotídeo arbitrário usado nas amplificações seletivas.

3.4.2 Análise em gel de poliacrilamida

Os produtos resultantes da PCR foram separados em gel de poliacrilamida desnaturante [acrilamida/bisacrilamida (19:1) 5%, uréia 7,0 M, tampão TBE 1X (89 mM Tris base; 89 mM de

ácido bórico; 2,5 mM EDTA 2Na)] de 0,4 mm de espessura. O gel foi submetido à eletroforese a uma potência constante de 70 W por 4 h e 30 min., usando um aparato de eletroforese Sequi-Gen[®] (Bio-Rad). Para evitar a perda de locos de baixo peso molecular, um gradiente foi gerado, adicionando-se 0,5 M de acetato de sódio ao tampão TBE da parte inferior do aparato de eletroforese. Os marcadores de peso molecular *Ladder* 25 pb (Invitrogen) e *Ladder* 100 bp (Invitrogen) foram misturados em igual proporção e corridos nas canaletas das extremidades do gel.

A revelação dos locos se deu por coloração com nitrato de prata pelo método proposto por Creste e colaboradores (CRESTE; TULMANN-NETO; FIGUEIRA, 2001), conforme detalhado no item 1.3.3. Os géis foram digitalizados em *scanner* Color Page-HR3200 (Genius) e as imagens armazenadas em computador para posteriores análises.

3.5 Análises de mapeamento

3.5.1 Integração dos marcadores desenvolvidos no mapa referência BJ

Um teste de aderência χ^2 ($\alpha=0,05$) foi realizado para todos os locos, utilizando-se as funções da planilha Excel do Windows (Microsoft Windows).

O conjunto de dados dos marcadores presentes no *framework* do mapa integrado de Freyre e colaboradores (FREYRE et al., 1998) nos foi gentilmente cedido pelo Dr. Paul Gepts (*University of Califórnia, USA*). Estes dados foram usados para integrar os marcadores gerados no presente trabalho ao mapa referência de feijão. O mapa de ligação foi construído com o programa Mapmaker/Exp versão 3.0 para Windows (LANDER, E. et al., 1987). A opção ‘ri-self’ foi usada para o mapeamento da população BJ. Inicialmente, para a determinação dos grupos de ligação, três marcas do *framework* do mapa consenso (FREYRE et al., 1998), sendo uma em cada extremidade e uma no centro do grupo de ligação, foram ancorados em cada um dos 11 grupos de *P. vulgaris* previamente nomeados de B1 a B11. Os marcadores SSR-genômicos, SSR-EST, AFLP e o restante dos marcadores do *framework* foram, então, atribuídos aos grupos de ligação com valor de LOD > 3,0 através do comando ‘assign’. O comando ‘suggest subset’ foi usado para a identificação de marcadores altamente informativos e, ao mesmo tempo, relativamente bem distribuídos nos grupos de ligação. Estes marcadores foram ordenados usando o comando ‘compare’, com 10 marcadores por grupo de ligação. Os locos restantes foram adicionados

usando o comando 'try' e a ordem verificada pelo comando 'ripple'. Aqueles marcadores que não puderam ser ordenados com um LOD > 2,0 foram atribuídos aos intervalos mais verossímeis.

As frequências de recombinação foram transformadas em distâncias de mapa utilizando a função de mapeamento de Kosambi (1944). O mapa final foi desenhado com o programa MapChart 2.2 (VOORRIPS, 2002).

3.5.2 Construção do mapa CFM

O teste de aderência χ^2 ($\alpha=0,05$) foi realizado para todos os locos utilizando-se as funções da planilha Excel do Windows (MICROSOFT, 1997). Os marcadores foram, então, classificados de acordo com o número de dados perdidos e aderência à segregação esperada.

O programa Mapmaker/Exp versão 3.0 (LANDER et al., 1987) foi usado para as análises de agrupamento e ordenação dos locos no mapa. A determinação dos grupos de ligação foi realizada através do comando 'group' com LOD *score* entre 4,0 e 8,0 e distância menor que 50 cM (função de Kosambi). Os grupos com até 10 marcadores tiveram a ordem determinada pelo comando 'compare'. Para os grupos maiores, os 10 marcadores mais informativos - com poucos dados perdidos e apresentando a segregação esperada - foram ordenados pelo comando 'compare'. Os locos restantes foram adicionados usando o comando 'try' e a ordem verificada pelo comando 'ripple'. Aqueles marcadores que não puderam ser ordenados com um LOD > 2,0 foram atribuídos, como acessórios, aos intervalos mais verossímeis.

As frequências de recombinação foram transformadas em distâncias de mapa utilizando a função de mapeamento de Kosambi (KOSAMBI, 1944). O mapa final foi desenhado com o programa MapChart 2.2 (VOORRIPS, 2002).

3.5.3 Alinhamento do mapa CFM ao mapa núcleo BJ.

Os marcadores microsatélites que segregaram nas duas populações foram utilizados para o alinhamento entre os mapas BJ e CFM. O programa MapChart 2.2 (VOORRIPS, 2002) foi utilizado para a visualização dos mapas alinhados.

4 RESULTADOS E DISCUSSÃO

4.1 Frequência e localização dos SSR nas EST

Mais de 5.200 EST foram seqüenciadas a partir de tecidos de feijão e, numa busca prévia, foram identificadas 240 SSR (MELOTTO et al., 2005). Hu e colaboradores (2004) encontraram 394 SSR dentre mais de 4.000 EST de batata doce e, recentemente, Peng e Lapitam (2005) identificaram mais de 40.000 SSR dentro de quase meio milhão de EST de trigo. Isto indica que existem microssatélites com frequência relativamente alta na parte expressa do genoma de espécies cultivadas.

A análise da função putativa das EST seqüenciadas por Melotto e colaboradores (2005) distribuiu estas EST em 14 categorias pelo alinhamento com seqüências protéicas da base de dados não-redundante do *GenBank* (BLASTx). A mesma análise foi realizada nas EST contendo SSR e uma proporção similar foi encontrada entre as 14 categorias. Este resultado mostra não haver distribuição preferencial dos SSR em categoria específica de genes (Tabela 4).

Tabela 4 - Distribuição das frequências dos *unigenes* detectados nas 5.255 EST e daqueles contendo SSR em 14 categorias

Categoria	3.126 <i>unigenes</i> (Melotto et al., 2005)	219 <i>unigenes</i> contendo SSR
Fotossistema	4,2%	6,4%
Metabolismo geral	12,7%	9,1%
Metabolismo de macromoléculas	12,9%	12,3%
Processos celulares	7,6%	8,2%
Vias de ubiquitinização	0,9%	0,9%
Estresses abióticos	2,8%	3,2%
Estresses bióticos	3,1%	1,4%
Domínios relacionados com resistência a doenças	1,2%	2,3%
Retrotransposons	0,3%	0,0%
Transdução de sinais	2,3%	1,8%
<i>No hit</i>	11,4%	12,3%
<i>Low hit</i>	9,5%	11,0%
Genes conservados de função desconhecida	27,7%	28,3%
Não classificados	3,5%	2,7%

Os 240 SSR contidos em 219 *Unigenes* puderam ser divididos em três tipos: 138 (57%) foram compostos de motivos dinucleotídicos, 97 (40%) de motivos trinucleotídicos e sete (3%) de motivos tetranucleotídicos. O mais abundante de todos os motivos SSR foi o AG presente em 33,5% do total (Tabela 5). Estudos iniciais da abundância dos diferentes motivos SSR em bases de dados de plantas indicaram o motivo AT como a repetição dinucleotídica mais comum (LAGERCRANTZ; ELLEGREN; ANDERSSON, 1993). No entanto, em feijão encontra-se, mais freqüentemente, a repetição de dinucleotídeos AG. De fato, seqüências expressas têm mostrado uma maior freqüência da repetição AG do que AT tanto em *Arabidopsis thaliana* quanto em soja (MORGANTE; HANAFEY; POWELL, 2002). A presença preferencial de motivos SSR em regiões codantes, isto é, mostrando um viés para composições nucleotídicas específicas (LI et al., 2004) pode estar relacionada com as altas freqüências de certos aminoácidos (TÓTH; GASPARI; JURKA, 2000).

Entre os SSR trinucleotídicos, os motivos GAA foram os mais abundantes, contando por 14,2% do total (Tabela 5) ou 35,4% dos trinucleotídicos. Wang e colaboradores (WANG et al., 1994) apontaram os motivos AAT, AAC e AGC como os mais comuns em plantas, diferentemente dos dados encontrados no presente trabalho, em feijão. Diferenças também existem entre espécies de plantas para a prevalência de motivos particulares de SSR. Em cereais, por exemplo, o motivo CGG é o mais comum (CORDEIRO et al., 2001a; VARSHNEY, R.K. et al., 2002), enquanto em soja (GAO et al., 2003) e *arabidopsis* (CARDLE et al., 2000) o motivo GAA prevalece, corroborando as freqüências encontradas no presente trabalho.

Poucos SSR tetranucleotídicos foram aqui encontrados, por isso, quase nada se pode inferir a respeito deles. A baixa freqüência de SSR com motivos tetranucleotídicos se deve, principalmente, à alta estringência usada na sua busca. O critério usado foi o número de repetições, fixada em no mínimo 5 para qualquer tamanho de motivo. Por isso, é mais provável se encontrar SSR com motivos dinucleotídicos com 10 pb do que motivos trinucleotídicos com 15 pb ou tetranucleotídicos de, no mínimo, 20 pb de comprimento.

Tabela 5 - Frequência dos motivos SSR entre os 242 SSR presentes nos 219 *unigenes* de *P. vulgaris*

	Motivo	Número de repetições						Total/motivo	
		5	6	7	8	9	>10	Nº	(%)
Dinucleotídeos	AG	42	15	6	4	5	8	80	(33,3)
	AT	20	8	3		1	5	37	(15,3)
	AC	17	1	1		1		20	(8,3)
	CG							0	(0,0)
Trinucleotídeos	GAA	18	11	4			1	34	(14,2)
	CAT	11	6	1				18	(7,5)
	TGG	6	3	1				10	(4,2)
	AAC	6	2					8	(3,3)
	TGC	6	1					7	(2,9)
	AAT	4	1	1				6	(2,5)
	CGG	4						4	(1,7)
	TAG	2	1					3	(1,3)
	AGG	2		1				3	(1,3)
	ACG	3						3	(1,3)
Tetranucleotídeos	AAAG	1		1				2	(0,8)
	AATC	1						1	(0,4)
	AATG	1	1					2	(0,8)
	AATT	1						1	(0,4)
	TGCA	1						1	(0,4)
Total		147	50	19	4	7	14	240	(100)

Os SSR trinucleotídicos são relatados como a classe mais abundante de microssatélites em EST de várias espécies incluindo as dicotiledôneas *A. thaliana* e soja (GAO et al., 2003; MORGANTE; HANAFEY; POWELL, 2002). No entanto, de forma geral, os SSR dinucleotídicos foram encontrados como os mais abundantes neste trabalho. Isto, provavelmente, também se deve aos critérios usados na busca por SSR, como discutido anteriormente. Quando foram classificados segundo a posição em que se encontraram no gene, os SSR dinucleotídicos estiveram preferencialmente localizados na porção não-codante das seqüências enquanto, interessantemente, SSR trinucleotídicos foram mais abundantes em regiões codantes das seqüências (Figura 3). A ocorrência preferencial deste tipo de SSR em regiões codantes é devido ao fato deles não causarem mudança na fase de leitura de gene (LI et al., 2004) e assim, não provocar uma forte alteração na proteína final.

Figura 3 - Distribuição das frequências dos SSR-EST baseada em suas posições no gene (5' UTR: região 5' não traduzida; CDs: região codificadora de proteína ;UTR: região 3' não traduzida)

4.2 Caracterização dos marcadores SSR-EST

No total, 137 pares de *primers*, dos 156 desenhados, amplificaram 138 locos. O par de *primers* PvM13 amplificou 2 locos e 19 pares de *primers* não amplificaram bandas em nenhuma das 4 condições testadas. Inicialmente, 40 locos SSR-EST foram avaliados em um conjunto de 25 genótipos, sendo 23 de *P. vulgaris*, um de *P. lunatus* e um de *Vigna* sp (HANAI et al., 2007; anexo B). Posteriormente, o restante dos SSR-EST (98 locos) foi avaliado usando o mesmo conjunto de genótipos de *P. vulgaris* e um acesso de *P. acutifolius*. Do total de 138 locos amplificados, 9 não foram usados nas análises por não permitirem uma leitura clara do gel. Assim, 130 locos foram analisados, dos quais 85 se mostraram polimórficos entre os genótipos estudados e 44 amplificaram alelos únicos.

Os *primers* desenvolvidos para SSR-EST de *P. vulgaris* foram testados em três espécies relacionadas. Cento e sete dos 138 locos microssatélites amplificados em *P. vulgaris* foram transferidos com sucesso para, pelo menos, uma das três espécies avaliadas. Em *P. acutifolius*, foram testados 106 pares de *primers* e nas espécies *P. lunatus* e *Vigna* sp foram testados 52 pares de *primers*. As frequências de amplificação foram de 60%, 87% e 83%, para *P. acutifolius*, *P. lunatus* e *Vigna* sp., respectivamente (Tabela A1; Anexo A). No entanto, desconsiderando-se os *primers* que não amplificaram nem mesmo em *P. vulgaris*, estas taxas sobem para 97,7% para *P. lunatus*, 93,0% para *Vigna* sp. e 76% para *P. acutifolius*. A taxa de transferência para *P.*

acutifolius foi menor que para *P. lunatus* ou mesmo para *Vigna* sp., o que está em desacordo com a literatura que tem mostrado maior sucesso na transferência dentro do mesmo táxon (PEAKALL et al., 1998). No entanto, não se pode afirmar que a transferabilidade de *P. vulgaris* para *Vigna* é maior que para *P. acutifolius*, já que neste trabalho a grande maioria dos *primers* testados foi diferente para estas espécies (Tabela A1; Anexo A). Apesar da amplificação não garantir que estes locos sejam polimórficos nas espécies para as quais os *primers* foram transferidos, fica evidente que as regiões flanqueando os microssatélites nas EST são conservadas entre estas espécies. Além do mais, o presente estudo serve de ponto de partida para a validação destes marcadores em tais espécies, o que poderá contribuir para estudos de sintenia e facilitar a transferência, a partir de espécies relacionadas, de genes importantes para melhoramento do feijão-comum.

No total, 85 SSR-EST apresentaram-se polimórficos entre os 26 genótipos. Para medir a quantidade de informação dos marcadores SSR-EST, o PIC foi calculado para cada um dos locos polimórficos. Os valores variaram entre 0,08 e 0,91 com uma média de 0,38 por loco. Uma média de 2,96 alelos por loco polimórfico foi observada, e os locos com maior número de alelos foram o PvM21 e o PvM115, com 12 e 13 alelos, respectivamente. Nossos dados são similares àqueles de microssatélites gênicos: Yu e colaboradores (1999), usando 12 SSR, desenvolvidos a partir de seqüências do banco de dados *Genbank*, para acessar a variação alélica em 12 linhagens de feijão, encontraram uma média de 2,9 alelos por loco; Guerra-Sanz (2004) detectou 3,1 alelos por loco quando analisou 20 SSR usando 12 acessos de *P. vulgaris*. Por outro lado, microssatélites genômicos têm se mostrado mais informativos: Gaitán-Solis e colaboradores (2002), por exemplo, encontraram uma média de 6,0 alelos por loco em 68 locos SSR desenvolvidos a partir de bibliotecas genômicas, resultado similar ao encontrado por Buso e colaboradores (2006) que, a partir de 10 locos, também derivados de biblioteca genômica, encontraram uma média de 7 alelos por loco. Estas diferenças podem ser devido ao fato dos SSR-EST estarem associados a regiões expressas do genoma com uma função biológica e, por isso, serem mais conservadas que a porção não-codante. Outra possível explicação pode estar na diversidade dos genótipos utilizados na caracterização dos microssatélites, pois Benchimol e colaboradores (2007), usando um grupo de 20 genótipos, sendo 12 comuns aos utilizados neste trabalho, encontraram em 87 microssatélites genômicos 3 alelos por loco, em média.

A análise de locos SSR mostrou uma relação inversamente proporcional entre a frequência de polimorfismo e o tamanho do microsatélite (Figura 4A). Este resultado está de acordo com a literatura que afirma que a estrutura e o tamanho do SSR são considerados como os principais fatores que afetam a variabilidade dos microsatélites e, dessa forma, locos SSR com mais repetições tendem a ser mais polimórficos e a ter uma maior amplitude de variação (CHO et al., 2000; INNAN; TERAUCHI; MIYASHITA, 1997). De fato, os dois locos mais polimórficos e com maiores amplitudes de variação foram aqueles com mais repetições, o PvM21 e o PvM115 com os motivos (AT)₁₄ e (CT)₂₅, respectivamente (Figura 5). No entanto, o tamanho do microsatélite não parece ter uma relação direta com o PIC, pelo menos para aqueles menores que 25 pb. Na Figura 4B pode-se observar microsatélites de apenas 10 pb e altamente informativos. As informações detalhadas de cada um dos locos estão apresentadas na Tabela A1 (Anexo A).

Análises cofenéticas foram realizadas com base nos 85 locos polimórficos. O dendrograma obtido dividiu os genótipos em dois grandes grupos, um contendo as 17 linhagens de origem Mesoamericana e outro contendo as 6 linhagens de origem Andina, revelando assim uma clara diferença entre os feijões mesoamericanos e andinos. O acesso de *P. acutifolius* não se agrupou a nenhum dos dois grandes grupos, se posicionando como um *out-group* no dendrograma (Figura 6). A maioria dos genótipos se agrupou da mesma forma que quando foram usados apenas 31 locos polimórficos (HANAI et al., 2007). Os genótipos Bat 93, Garbancillo, Flor de Mayo, Apetito Blanco e Great Northern se agruparam em um dos clusters mesoamericanos e o restante dos genótipos formou outro cluster, com pequenas alterações em relação as análises prévias (HANAI et al., 2007). Já os genótipos de origem Andina permaneceram como um grupo único, distante dos demais.

Figura 4 - Distribuição dos SSR-EST baseada no polimorfismo e nos seus valores de PIC. Em **A** é mostrada a relação entre o tamanho do microsatélite e a frequência de polimorfismo; em **B** pode ser observada a distribuição do PIC dos locos SSR-EST em relação ao tamanho dos microsatélites

Figura 5 - Padrão de amplificação de marcadores SSR-EST em 23 linhagens de *P. vulgaris* em gel de poliácridamida 5%, corado com nitrato de prata. Em **A** e **B** são mostrados os locos PvM21 e PvM115, respectivamente

Figura 6 – Dendrograma resultante da análise de agrupamento de 23 genótipos de *P. vulgaris* e um acesso de *P. acutifolius*, com base em 85 marcadores SSR-EST. A árvore foi construída pelo método UPGMA a partir das distâncias calculadas pelo método de Rogers' modificado (WRIGHT, 1978). M indica o grupo dos genótipos de origem Mesoamericana e A indica os genótipos de origem Andina

Os marcadores SSR-EST mostraram haver uma grande diferença entre os dois grupos gênicos, Andino e Mesoamericano. No entanto, a variabilidade dentro de cada um dos grupos foi consideravelmente menor, da mesma forma que nos resultados anteriores, usando marcadores RFLP (CHASE; ORTEGA; VALLEJOS, 1991; NODARI et al., 1992).

Nossos resultados comprovam a utilidade dos microssatélites derivados de genes para estudos de diversidade genética em feijão. Além disso, alguns SSR-EST aqui desenvolvidos têm grande potencial para serem utilizados na identificação de cultivares por possuírem um alto valor de PIC, como por exemplo, os SSR-EST PvM21 e PvM115 (Figura 5).

4.3 Mapeamento genético da população BJ

4.3.1 Genotipagem dos marcadores SSR-EST

Os 50 locos SSR-EST que se mostraram polimórficos entre os genitores da população BJ foram amplificados usando as 75 RIL e foram usados na construção do mapa genético. A segregação de 3 locos SSR-EST pode ser visualizada na Figura 7. Cerca de 28% dos locos SSR-EST apresentaram distorção de segregação com relação ao esperado (1:1) quando aplicado o teste de Qui-quadrado ($\alpha=0,05$). Oito locos apresentaram excesso de alelos do genitor ‘Bat 93’ e 6 locos amplificaram mais alelos do genitor ‘Jalo EEP558’.

Também, 9 locos microssatélites derivados de bibliotecas genômicas foram genotipados na população BJ. Estes locos foram selecionados por mostrarem-se polimórficos nas duas populações de mapeamento usadas neste trabalho e porque poderiam ajudar no posterior alinhamento dos mapas.

Figura 7 - Padrão de segregação de três marcadores SSR-EST. De cima para baixo estão mostrados os locos PvM120, PvM126, PvM93. Nas canaletas 1 e 2 estão os genitores ‘Bat 93’ e ‘Jalo EEP558’, respectivamente. Nas restantes, estão as 75 linhagens da população BJ, exceto que na canaleta M está o marcador de peso molecular ladder 10 pb (Invitrogen)

4.3.2 Genotipagem dos marcadores AFLP

Inicialmente, um teste de combinações foi realizado a fim de identificar as combinações que amplificaram mais locos polimórficos. Foram testadas, no total, 28 combinações de *primers* seletivos, sendo 12 combinações a partir da digestão com *Eco*-RI e *Mse*-I e 16 combinações *Pst*-I e *Mse*-I para o cruzamento BJ. As 28 combinações produziram um total de 1252 locos, dos quais 361 (29%) se mostraram polimórficos entre os genitores da população BJ. Quinze combinações, das previamente avaliadas quanto ao polimorfismo entre os genitores, foram selecionadas e genotipadas na população de 75 RIL, produzindo um total de 203 locos segregantes (Tabela 6). A segregação de alguns locos gerados por uma das combinações selecionadas (*EacgMctc*) pode ser visualizada na Figura 8.

Tabela 6 - Combinações de *primers* seletivos genotipadas, mostrando o número de locos de AFLP amplificados e os polimorfismos gerados por combinação

Combinação	Combinação de <i>primers</i>	Nº de locos amplificados	Nº de locos polimórficos
C01	PaaaMgct	55	18
C02	PaacMgca	56	19
C03	PacgMgct	45	5
C04	PagaMgga	54	8
C05	PagaMgca	48	11
C06	PactMgga	37	4
C07	PaacMgct	25	6
C08	EaaaMcac	62	20
C09	EagaMcag	65	10
C10	EagcMcgc	77	23
C11	EacgMctc	67	23
C12	EaaaMctc	57	18
C13	EagcMcag	47	8
C14	EaaaMcag	57	18
C15	EagaMcac	48	12
	Total	1296	203

Vinte e oito dos 203 locos, ou seja, aproximadamente 14% apresentaram alguma distorção de segregação, metade da frequência encontrada usando marcadores microssatélites. É provável que esta diferença seja devido aos SSR-EST serem derivados de genes, os quais estão mais sujeitos a algum tipo de seleção do que as seqüências anônimas. Já entre as enzimas usadas, nenhuma diferença foi detectada. Os locos gerados pela combinação de enzimas *Eco*-RI e *Mse*-I apresentaram 14% de distorção de segregação e os gerados pela combinação *Pst*-I e *Mse*-I apresentaram 13% de distorção. Dos 28 locos que desviaram da segregação esperada, 13 apresentaram excesso do alelo de ‘Jalo EEP558’ e 15 apresentaram excesso do alelo de ‘Bat 93’.

Figura 8 - Padrão de AFLP gerado a partir da combinação de *primers* seletivos *EacgMctc*. Nas canaletas 1 e 2 estão os genitores ‘Bat 93’ e ‘Jalo EEP558’, respectivamente. Na canaleta M está o marcador de peso molecular ladder 100+25 pb (Invitrogen). Nas restantes, estão as 75 linhagens da população BJ. As setas na esquerda indicam exemplos de marcadores AFLP mapeados

4.3.3 Mapa de ligação BJ

Para a população BJ, um total de 510 locos foi inicialmente usado, incluindo 248 marcadores presentes no mapa integrado da espécie (FREYRE et al., 1998), 9 SSR-genômicos, 50 SSR-EST e 203 AFLPs. Após as primeiras análises, 118 locos RAPD do mapa integrado (FREYRE et al., 1998) foram retirados, pois grande parte destes locos possuíam dados de segregação para apenas 44 das 75 RIL usadas neste trabalho. Por isso, foram excluídas nas análises seguintes. Aproximadamente 17% dos locos SSR e AFLP genotipados apresentaram significativa distorção de segregação ($\alpha=0,05$). No entanto, apesar de não terem sido descartados, eles não foram usados no primeiro passo de ordenação dos marcadores, quando o comando 'compare' foi utilizado, sendo adicionados posteriormente.

Um conjunto de dados de 392 marcadores foi finalmente usado para a integração dos SSR e AFLP no mapa núcleo. Dentre os marcadores previamente mapeados na população BJ, foram escolhidos três RFLP posicionados, respectivamente, nas extremidades e no centro de cada grupo ligação. Estes marcadores foram atribuídos aos seus respectivos grupos, de B1 até B11, e serviram de âncora para a formação dos 11 grupos de ligação. Do total, 357 marcadores (111 marcadores do mapa integrado, 56 SSR e 190 AFLPs) formaram os 11 grupos de ligação e 35 permaneceram não ligados. O *framework* do mapa foi gerado a partir da ordenação de 293 locos, que produziram um mapa de comprimento total de 1.353 cM. Os 64 locos restantes não puderam ser ordenados com um LOD mínimo de 2 e foram excluídos das análises multiponto. Estes foram atribuídos aos intervalos mais verossímeis, sem influenciar nas distâncias do mapa, e aparecem como marcadores acessórios ao lado do *framework* (entre parênteses) (Figura 9).

Figura 9 - Mapa genético da população BJ. Os marcadores alocados no *framework* foram ordenados com LOD > 2,0 e as suas posições estão indicadas à esquerda. Os marcadores não ordenados estão entre parênteses e foram alocados nos intervalos mais verossímeis. Os marcadores do *framework* do mapa integrado da espécie estão destacados em *itálico*. Os locos realçados em **negrito** são microsatélites e os demais AFLP. Os marcadores mostrando distorção de segregação são indicados com * ($P < 0,05$), ** ($P < 0,01$) e *** ($P < 0,001$)

Figura 9 - Mapa genético da população BJ. Os marcadores alocados no *framework* foram ordenados com LOD > 2,0 e as suas posições estão indicadas à esquerda. Os marcadores não ordenados estão entre parênteses e foram alocados nos intervalos mais verossímeis. Os marcadores do *framework* do mapa integrado da espécie estão destacados em itálico. Os locos realçados em negrito são microsatélites e os demais AFLP. Os marcadores mostrando distorção de segregação são indicados com * ($P < 0,05$), ** ($P < 0,01$) e *** ($P < 0,001$)

O mapa aqui gerado se mostrou consistente com relação à ordenação das marcas nos grupos de ligação. Dentre os marcadores presentes no mapa gerado por Freyre e colaboradores (1998) mapeados neste trabalho, nenhum par de marcas mostrou discrepância na sua ordem em relação ao mapa núcleo. As diferenças encontradas foram relacionadas às distâncias entre marcadores, as quais podem ser explicadas por diferenças no tamanho da população e na qualidade dos dados usados. Enquanto Freyre e colaboradores usaram 91 RIL, o mapa aqui apresentado foi construído com base em apenas 75 linhagens. Por outro lado, dentre os 620 marcadores usados por Freyre et al. (1998), 429 locos RAPD provenientes do mapa de Wisconsin tinham 48% de dados perdidos (Dr. Paul Gepts, comunicação pessoal). Já no presente trabalho, apesar do número menor de linhagens (75 RILs), foram usados locos com apenas 6,4% de dados perdidos, em média.

O comprimento total do mapa gerado no presente trabalho foi um pouco superior ao mapa integrado de Freyre e colaboradores (1998), que teve um tamanho final de 1226 cM. No entanto, está de acordo com o mapa recém publicado por Grisi e colaboradores (2007), que produziram um mapa com 1358 cM de comprimento a partir da integração de 104 marcadores microsatélites

nesta mesma população de mapeamento. Como a população e mesmo a geração das RIL foram iguais nos três trabalhos, esta variação no tamanho dos mapas se deve, provavelmente, às diferenças no número e tipos de marcadores utilizados na construção de cada mapa, pois diferentes classes de marcadores podem se distribuir diferentemente no genoma. No grupo B1, por exemplo, foram mapeados seis marcadores AFLP e três microssatélites na sua extremidade inferior (Figura 9). Estes podem ter amostrado uma região diferente daquela amostrada no mapa de Freyre et al. (1998) e, desta forma, ter contribuído para o aumento no tamanho deste grupo de ligação. A mesma inferência vale para o grupo B4 que teve três marcadores AFLP mapeados na extremidade inferior. Juntos, estes dois grupos foram responsáveis por mais de 60 cM do aumento do comprimento deste mapa em relação ao mapa integrado de Freyre et al. (1998).

Os grupos de ligação tiveram tamanhos entre 92 e 175 cM, com número de locos variando de 24 a 39 por grupo (Tabela 7). De forma geral, o número de marcadores mapeados em um dado grupo de ligação foi correlacionado com o respectivo comprimento do grupo (cM). No entanto, os grupos de ligação B3 e B5 tiveram os maiores números de marcadores, 39 cada um, enquanto o grupo de maior comprimento genético foi o B2. Além disso, o grupo de ligação B10 foi o mais curto, o que não coincidiu com o grupo com o menor número de marcadores (B9) (Tabela 7). A densidade do mapa foi de 4,6 cM entre marcas, em média, com marcadores relativamente bem distribuídos ao longo dos 11 grupos. Contudo, diferenças na densidade de marcas foram observadas em alguns intervalos. A densidade de marcas nos intervalos 58-68 cM no grupo B4, 31-41 cM no grupo B5 e 47-57 cM no grupo B8 foi superior a de outras regiões (Figura 9). Estes clusters presentes nos grupos B4, B5 e B8 foram formados quase que exclusivamente por marcadores AFLP. A distribuição não randômica dos marcadores AFLP gerados pela combinação de enzimas *Eco*-RI e *Mse*-I foi previamente reportada em mapas genéticos de outras espécies, incluindo soja (YOUNG et al., 1999), feijão-azuki (HAN et al., 2005) e maracujá (LOPES et al., 2006). No presente trabalho, a combinação de enzimas *Pst*-I e *Mse*-I foi usada em adição à combinação *Eco*-RI/*Mse*-I, conforme sugestão de Young e colaboradores (1999), que propuseram que a enzima sensível a metilação *Pst*-I seria mais eficiente que a *Eco*-RI no desenvolvimento de marcadores AFLP randomicamente distribuídos nos mapas de ligação. De fato, dos 71 locos *Pst*-I/*Mse*-I aqui usados, apenas 4 locos co-segregaram, dois no grupo de ligação B3 e dois no cluster central do grupo B4 (Figura 9).

Tabela 7 - Distribuição dos marcadores SSR e AFLP nos 11 grupos de ligação do mapa consenso de *Phaseolus vulgaris* gerado no presente estudo

Grupo de Ligação	Comprimento do mapa (cM)	Densidade ¹	Número de marcadores total		Marcadores adicionados	
			Total	Frame ²	SSR	AFLP
B1	129,5	5,0	36	26	6	19
B2	174,9	5,8	34	30	7	14
B3	166,1	6,2	39	27	7	21
B4	128,2	4,7	30	27	1	22
B5	107,2	3,7	39	29	8	18
B6	110,8	4,0	32	28	3	16
B7	105,0	4,2	26	25	2	13
B8	113,3	3,9	38	29	7	24
B9	120,4	5,5	24	22	6	11
B10	92,1	4,0	26	23	4	13
B11	106,1	3,9	33	27	5	19
Total	1.353,6	4,6	357	293	56	190

¹ Distância média entre marcadores no *framework*; ² Frame: marcadores ordenados no *framework*.

Com relação aos marcadores apresentando distorção de segregação, houve uma localização preferencial em certas regiões genômicas nos grupos de ligação B1, B2, B3, B6, B7, B8 e B9. Os marcadores com segregação distorcida, alocados nos grupos de ligação B1, B3 e B6, mostraram-se enviesados para o progenitor ‘Jalo EEP558’. Já os marcadores com desvio de segregação, alocados nos grupos de ligação B2, B7 e B8, mostraram-se enviesados para o progenitor ‘Bat 93’. Blair e colaboradores (2003) constataram em feijão que locos mostrando distorção de segregação foram alocados nos grupos de ligação B1, B2, B3 e B8 com excesso de alelos do genitor ‘Jalo EEP558’, e nos grupos B5, B9, B10 e B11 com predominância do alelo do genitor ‘Bat 93’. A predominante transmissão de alelos andinos nos grupos de ligação B1 e B3, em duas populações de mapeamento, sugere que as regiões genômicas são as mesmas (Bat x Jalo e DOR364 x G19833) e que, possivelmente, isto reflita a localização de genes para incompatibilidade ou adaptabilidade que influenciaram no processo de seleção em ambas as populações (BLAIR et al., 2003).

A saturação de mapas genéticos é de grande importância para estudos genéticos e desenvolvimento de cultivares melhoradas. Quanto maior o número de marcas alocadas no mapa,

maiores são as chances de saturar os *gaps* existentes no mapa consenso, o que é importante para a seleção assistida por marcadores e para o mapeamento de QTL, pois a eficiência tanto do mapeamento de QTL quanto da seleção é aumentada pela presença de marcadores regularmente espaçados ao longo dos grupos de ligação (LANDER, E. S.; BOTSTEIN, 1989). De forma geral, o mapa final apresentou uma boa cobertura genômica. Os *gaps* presentes nos grupos B4 e B6 do mapa integrado de Freyre e colaboradores (1998) foram aqui saturados por marcadores AFLP. Apenas o grupo de ligação B9 permaneceu com um *gap* maior que 30 cM entre duas marcas. Por outro lado, no grupo B5, o maior *gap* foi de 8,3 cM, com uma densidade média de 3,7 cM entre marcas. Além disso, algumas regiões contendo QTL de resistência à antracnose e ao crestamento bacteriano (MIKLAS et al., 2006) foram saturadas por marcadores SSR e AFLP. No grupo de ligação B1, a região onde fora mapeado um QTL para antracnose (76-95 cM) foi saturada com oito marcadores adicionais, sendo seis AFLPs e dois SSR-EST, e a região 58-72 cM do grupo B2 contendo outro QTL de resistência à antracnose foi saturada com três AFLPs e um SSR-EST. O mesmo foi observado nos intervalos 18-33 cM do grupo B7 e 6-24 cM do grupo B9, regiões contendo QTL de resistência ao crestamento bacteriano, onde foram mapeados três e oito marcadores adicionais, respectivamente.

Dentre os marcadores microssatélites, 56 deles foram aqui ordenados em 50 posições diferentes ao longo dos 11 grupos de ligação, cinco co-segregaram com outro microssatélite e um foi deixado como marca acessória (Figura 9). Apesar desta categoria de marcador ter sido mapeada em todos os grupos de ligação, sua distribuição não foi homogênea entre os grupos, variando de 8 marcadores no grupo B5 a apenas um microssatélite no grupo B4 (Tabela 7). Interessantemente, o mesmo grupo B4, no qual menos SSR-EST foram mapeados no presente trabalho, foi aquele onde 5 dos 16 microssatélites mapeados por Yu e colaboradores (2000) se integraram. Recentemente, Grisi e colaboradores (2007) integraram um total de 104 microssatélites no mapa BJ, os quais também não se distribuíram de forma homogênea entre os grupos, variando de 4 a 13 locos por grupo.

Neste trabalho, foram mapeados diferentes SSR-EST apresentando similaridade com proteínas conhecidas (Tabela 8). Por exemplo, o marcador PvM18 presente no grupo de ligação B5 deriva de uma EST apresentando similaridade com uma proteína da família LRR (*leucine-rich repeat*) de *Arabidopsis thaliana*. O loco PvM97 mapeado no grupo B1 foi originado de uma EST similar a uma proteína de *Glycine latifolia* (AAO91809) induzida por estresse hídrico. E, o SSR-

EST PvM95 mapeado no grupo B3 é altamente similar a uma hidroximetiltransferase (NP_195506) que está relacionada com a resposta de hipersensibilidade em *A. thaliana*.

Tabela 8 - SSR-EST integrados no mapa consenso BJ apresentando similaridade com seqüências de proteínas depositadas no GenBank

(continua)					
SSR-EST	G.L. ¹	No. GenBank ²	E-value	Proteína apresentado similaridade	Organismo
PvM001	B8	P25699	4e-67	Ferritina	<i>Phaseolus vulgaris</i>
PvM002	B10	ABE77731	4e-16	Proteína hipotética conservada	<i>Medicago truncatula</i>
PvM002	B10	ABE77731	4e-16	Proteína hipotética conservada	<i>M. truncatula</i>
PvM007	B5	ABA40448	5e-71	Proteína parecida com frutoquinase 2	<i>Solanum tuberosum</i>
PvM010	B1	NP_194345	1e-8	Proteína da família Zinc finger (GATA type)	<i>A. thaliana</i>
PvM012	B4	ABE81465	5e-34	C2	<i>M. truncatula</i>
PvM013	B10	NP_565320	3e-46	Proteína ribossomal 40S S17 (RPS17B)	<i>A. thaliana</i>
PvM017	B8	CAA42617	8e-41	Ribulose bisfosfato carboxilase	<i>P. vulgaris</i>
PvM018	B5	NP_188718	2e-29	Proteína da família LRR	<i>A. thaliana</i>
PvM022	B9	AAA33039	4e-25	Proteína de ligação a RNA	<i>Mesembryanthemum crystallinum</i>
PvM034	B9	BAA19769	1e-37	SRC2	<i>Glycine max</i>
PvM036	B7	NP_194937	8e-17	Galactosiltransferase	<i>A. thaliana</i>
PvM046	B2	Q94A43	4e-21	Proteína homóloga à BES1/BZR1	<i>A. thaliana</i>
PvM052	B5	NP_196885.1	7e-05	Proteína da família glutaredoxina	<i>A. thaliana</i>
PvM059		CAN61408	1e-44	Proteína hipotética	<i>Vitis vinifera</i>
PvM061	B9	ABE85172	1e-49	Proteína no apical meristem (NAM)	<i>M. truncatula</i>
PvM062	B5	XP_968694	0.66	Proteína similar a CG13503-PF, isoforma F	<i>Tribolium castaneum</i>
PvM066	B6	NP_001077728	3e-12	Proteína associada à dormência/auxina	<i>A. thaliana</i>
PvM073		AAB00554	3e-21	Dehidrina	<i>P. vulgaris</i>
PvM075	B9	CAN73178	2e-22	Proteína hipotética	<i>V. vinifera</i>
PvM078	B3	NP_974919	3e-20	Proteína desconhecida	<i>A. thaliana</i>
PvM087	B3	AAY54007.1	1e-76	Protease parecida com subtilisina	<i>Arachis hypogaea</i>
PvM088	B2	CAN65462	0.074	Proteína hipotética	<i>V. vinifera</i>
PvM093	B2	ABO83187	2e-18	UDP-glucuronosil/UDP-glucosiltransferase	<i>M. truncatula</i>
PvM095	B3	P49357	2e-66	Serina hidroximetiltransferase	<i>Flaveria pringlei</i>

Tabela 8 - SSR-EST integrados no mapa consenso BJ apresentando similaridade com seqüências de proteínas depositadas no GenBank

(conclusão)					
SSR-EST	G.L. ¹	No. GenBank ²	E-value	Proteína apresentado similaridade	Organismo
PvM097	B1	AAO91809	1e-38	Proteína induzida por estresse hídrico	<i>G. latifolia</i>
PvM098	B11	NP_200264	5e-35	Proteína hipotética	<i>A. thaliana</i>
PvM103	B8	AAN03469	6e-71	Enzima de conjugação com ubiquitina	<i>G. max</i>
PvM118	B8	ABF70090.1	3e-60	Proteína quinase putativa	<i>Musa balbisiana</i>
PvM120	B1	ABO78563	7e-06	Proteína P do sistema de clivagem de glicina	<i>M. truncatula</i>
PvM123	B1	ABN08632	0.65	SBP	<i>M. truncatula</i>
PvM124	B3	AAM63014.1	5e-27	Proteína desconhecida	<i>A. thaliana</i>
PvM127	B10	CAN64365	2e-07	Proteína hipotética	<i>V. vinifera</i>
PvM145	B5	NP_001064383	5.9	Os10g0340800	<i>O. sativa</i>
PvM148	B3	AAW28572.1	1e-38	<i>Putative oxygen evolving enhancer protein 3</i> , idêntica	<i>Solanum demissum</i>
PvM150	B11	NP_191557.1	3e-47	Proteína contendo domínio de ligação a tRNA	<i>A. thaliana</i>
PvM153	B2	AAL50980	5e-13	Proteína parecida com inibidor de bax	<i>Brassica oleracea</i>

¹ G. L. Grupo de ligação do mapa consenso em que o SSR-EST foi mapeado; ² No. Número de acesso no banco de dados GenBank.

4.4 Mapeamento genético da população CFM

4.4.1 Genotipagem dos marcadores SSR-EST

Dos 138 locos SSR-EST, apenas 21 se mostraram polimórficos entre os genitores da população CFM. Estes foram genotipados nas 94 RIL, porém dois deles não segregaram na população e um terceiro amplificou um loco de difícil leitura. A baixa frequência de polimorfismo observada para este cruzamento reflete a menor variabilidade genética existente dentro do grupo Mesoamericano em relação àquela existente entre os grupos Mesoamericano e Andino (Figura 6). Outros 9 locos SSR genômicos que segregaram em ambas as populações foram genotipados.

Os locos foram classificados em relação ao tamanho do fragmento amplificado para permitir a eletroforese simultânea de mais de um loco. As PCRs foram realizadas independentemente para cada par de *primers* e, em seguida, “multiplexadas” a fim de se otimizar as corridas. Na Figura 10, por exemplo, pode-se observar a segregação de 5 locos microssatélites em um único gel.

Figura 10 - Padrão de segregação de cinco marcadores SSR-EST, sendo que a eletroforese dos locos foi “multiplexada”. De cima para baixo estão mostrados os locos PvM118, PvM145, PvM79, PVM97 e PvM103. Nas canaletas 1 e 2 estão os genitores ‘Carioca’ e ‘Flor de Mayo’, respectivamente. Nas restantes, estão as 94 linhagens da população CFM, exceto que na canaleta M está o marcador de peso molecular ladder 10 pb (Invitrogen)

4.4.2 Genotipagem dos marcadores AFLP

Inicialmente, 41 combinações de *primers* seletivos foram testadas, 25 combinações a partir da digestão com as enzimas *Eco*-RI e *Mse*-I e 16 a partir da digestão com as enzimas *Pst*-I e *Mse*-I. Para as combinações a partir da digestão com *Eco*-RI e *Mse*-I foi observada uma média de 58 locos por combinação de *primers*, com uma frequência de polimorfismo entre 5% e 27% (17%, em média). No entanto, um número bem menor de locos foi visualizado nas amplificações a partir da digestão com *Pst*-I e *Mse*-I, 45 locos, em média, com 11% de polimorfismo, em média. Assim, 27 combinações de *primers* seletivos foram utilizadas para a obtenção de 225

locos polimórficos (Tabela 9). O perfil de amplificação da combinação C08 (*EaaaMcac*) é apresentado na Figura 11. Todos os géis obtidos mostraram um excelente padrão de amplificação, permitindo leituras confiáveis para a genotipagem.

Tabela 9 - Combinações de *primers* seletivos genotipadas na população CFM. O número total de locos amplificados e o número de locos polimórficos são mostrados

Combinação	Combinação de <i>primers</i>	Nº de locos amplificados	Nº de locos polimórficos
C01	EacgMcat	74	17
C02	EaggMcac	60	12
C03	EaacMcac	85	4
C04	EagaMcag	65	11
C05	EaaaMcat	57	11
C06	EaccMcat	72	11
C07	EaaaMcaa	53	11
C08	EaaaMcac	68	17
C09	EaacMcaa	48	13
C10	EacaMcaa	57	11
C11	EacgMcaa	78	9
C12	EtaaMgcc	48	7
C14	EttaMgcg	39	2
C15	EtttMgcc	53	8
C16	EtttMgcg	57	10
C17	PaaaMcaa	52	6
C18	PaagMcaa	46	6
C19	PaagMcac	42	5
C20	PaagMctc	39	4
C21	PactMcaa	48	6
C22	PactMcac	41	5
C23	PactMctc	45	4
C24	PagaMcac	38	3
C25	EaaaMcag	50	10
C26	EaaaMctc	48	12
C27	EacgMctc	45	8
C28	EagcMcgc	43	2
Total		1451	225

Figura 11 - Padrão de AFLP gerado a partir da combinação de *primers* seletivos *EaaaMcac*. Nas canaletas 1 e 2 estão os genitores ‘Carioca’ e ‘Flor de Mayo’, respectivamente. Na canaleta M está o ladder 100 +25 pb (Invitrogen) e nas restantes estão as 94 RIL da população CFM. As setas de ponta curta indicam exemplos de loco dominantes e as setas de ponta longa indicam os locos codominantes

4.4.3 Mapa de ligação CFM

Dados de segregação de 252 locos (9 locos SSR-genômicos, 18 SSR-EST e 225 AFLPs) foram analisados para a construção do mapa da população CFM. Apesar de a maioria dos locos terem segregado conforme o esperado (1:1), quando foi aplicado o teste de χ^2 ($\alpha=0,05$), o número de locos apresentando distorção foi significativamente maior nesta população (43%) comparativamente à população BJ (17%). Nem todas as marcas que desviaram do esperado foram descartadas das análises. Ressalte-se que o teste de χ^2 foi realizado, por loco, como se fossem testes únicos. Porém, quando se trata de múltiplos testes independentes, como é o caso, pois se trabalhou com 157 marcadores e, dentre eles, muitos são independentes, deve-se realizar uma correção. Isto elevará o valor do χ^2 para os níveis de significância previamente determinados ou, por outro lado, reduzirá o valor de α . Por exemplo, quando se trata de teste único, o valor de α para um corte a 5% de probabilidade é de 0,05, já o valor de α , após aplicada a correção de

Bonferroni para 12 testes independentes, com o mesmo nível de probabilidade, cai para 0,0043. Apesar de não se ter realizado esta correção com a finalidade de gerar um valor de corte de χ^2 , os marcadores apresentando forte distorção não foram usados na determinação do *framework*, sendo adicionados posteriormente pelo comando ‘try’ do Mapmaker e mantidos apenas quando não alteraram a ordem das marcas vizinhas nos grupos de ligação.

Os grupos de ligação foram primeiramente formados usando o comando ‘group’ com um LOD > 6,0 e distância menor que 50 cM (Kosambi). A partir deste critério, os marcadores foram separados em 23 grupos de ligação. A distribuição dos marcadores previamente mapeados na população BJ foi verificada e revelou algumas discrepâncias. Em um mesmo grupo de ligação, foram agrupados marcadores de dois diferentes grupos do mapa núcleo BJ, os quais somente foram separados por um LOD > 8,0. Por outro lado, foi observado que marcadores de um mesmo grupo de ligação do mapa BJ estavam separados em mais de um grupo em CFM. Assim, as informações prévias dos marcadores comuns ao mapa BJ foram usadas para a formação final dos grupos de ligação. Finalmente, 244 marcadores foram agrupados em 12 grupos de ligação e 8 permaneceram não ligados. Um total de 151 marcadores, incluindo SSR e AFLPs, foram ordenados no *framework* do mapa CFM com LOD > 2,0 (Tabela 10).

Tabela 10 - Distribuição dos marcadores SSR e AFLP nos 12 grupos de ligação do mapa CFM

Grupo de Ligação ¹	Comprimento do mapa (cM)	Densidade ²	Número de marcadores			
			Total	Frame ³	SSR	AFLP
CFM1	147,0	5,1	61	29	3	58
CFM2	115,8	9,7	16	12	3	13
CFM3	112,3	5,3	36	21	1	35
CFM4	93,8	7,8	13	12	0	13
CFM5	93,3	5,8	39	16	2	37
CFM6	69,8	4,4	29	16	4	25
CFM7	65,8	7,3	9	9	4	5
CFM8	33,5	6,7	5	5	1	4
CFM9	31,7	4,0	8	8	3	5
CFM 10	24,1	2,7	14	9	1	13
CFM 11	10,8	1,4	8	8	2	6
CFM12	9,6	1,6	6	6	1	5
Total	824,7	6,9	244	151	25	219

¹ Os grupos de ligação foram nomeados de acordo com as iniciais dos genitores ‘Carioca’ e ‘Flor de Mayo’, seguido pelos números 1 a 12 desde o maior até o menor grupo de ligação; ² Distância média entre marcadores no *framework*;

³ Frame: marcadores ordenados no *framework* com LOD > 2,0.

O mapa gerado teve um comprimento total de 807,5 cM, cobrindo aproximadamente 67% do genoma da espécie que foi estimado em cerca de 1.200 cM (VALLEJOS; SAKIYAMA; CHASE, 1992). A variação do comprimento dos grupos de ligação foi de 9,6 a 147,0 cM, enquanto o número e a porcentagem (entre parênteses) de marcadores por grupo de ligação variou de 5 (2,0%) a 61 (25,0%). A densidade média do *framework* foi de 5,3 cM entre marcas, variando de 1,4 cM no grupo CFM11 a 9,7 cM no grupo CFM2 (Tabela 10). Apesar de relativamente pouco saturado, o mapa apresentou apenas um *gap* com mais de 30 cM de comprimento e 9 com mais de 20 cM (Figura 12). Neste mapa a aglomeração de marcadores em algumas regiões foi maior que no mapa BJ, apresentado anteriormente. No grupo CFM1, foi observado um *cluster* contendo 14 marcadores. Os grupos CFM3, CFM5 e CFM6 também apresentaram *clusters* (Figura 12). Sakiyama e Chase (1992) detectaram que aproximadamente 20% dos marcadores usados na construção do mapa de ligação para a população ‘XR-235-1’ x ‘Calima’ estiveram agrupados em sete *clusters*. Dentre as possíveis explicações para a formação destes aglomerados de marcadores estão o pequeno tamanho da população amostrada somado ao fato de a maioria dos marcadores analisados serem AFLPs gerados pelo uso da enzima *Eco*-RI, que tem sido relacionada à distribuição não randômica de marcadores, em diversos trabalhos.

Cerca de 38% dos marcadores não puderam ser ordenados com $LOD > 2,0$, ou seja, a sua posição no mapa não foi pelo menos 100 vezes mais provável que a segunda melhor posição. Assim, 93 marcadores foram adicionados como acessórios no mapa. No grupo CFM5, o número de marcadores acessórios foi maior que o de marcadores ordenados no *framework* (Tabela 10). A dificuldade de ordenação dos marcadores, que foram colocados como acessórios no mapa, se deve, em parte, à presença de *clusters* em certas regiões genômicas e ao tamanho relativamente pequeno da população usada neste estudo.

Os marcadores microssatélites estiveram distribuídos em todos os grupos de ligação, exceto no grupo CFM4 (Tabela 10). Em ambos os grupos CFM6 e CFM7 foram observados quatro marcadores microssatélites, enquanto no restante foram alocadas de 1 a 3 marcas. Apesar do baixo número de marcadores microssatélites mapeados, eles foram de grande importância, pois 20 deles foram também mapeados na população BJ e permitiram o alinhamento de parte dos dois mapas.

Figura 12 - Mapa de ligação da população CFM. Os marcadores alocados no *framework* foram ordenados com LOD > 2,0 e as suas posições estão indicadas à esquerda. Os marcadores não ordenados estão ao lado do *framework* e foram alocados nos intervalos mais verossímeis. Os locos realçados em negrito são microssatélites e os demais AFLP. Os marcadores mostrando distorção de segregação são indicados com * ($P < 0,05$), ** ($P < 0,01$) e *** ($P < 0,001$)

Figura 12 - Mapa de ligação da população CFM. Os marcadores alocados no *framework* foram ordenados com LOD > 2,0 e as suas posições estão indicadas à esquerda. Os marcadores não ordenados estão ao lado do *framework* e foram alocados nos intervalos mais verossímeis. Os locos realçados em negrito são microssatélites e os demais AFLP. Os marcadores mostrando distorção de segregação são indicados com $*(P<0,05)$, $**(P<0,01)$ e $*** (P<0,001)$

Neste trabalho foi construído o primeiro mapa de ligação para a população CFM usando marcadores AFLP e microssatélites. Os primeiros mapas publicados para outras populações de feijão-comum apresentaram resultados semelhantes aos apresentados neste trabalho. O presente mapa teve 807,5 cM de comprimento total, enquanto os primeiros mapas para diversas populações tiveram tamanhos variando de 426 a 960 cM de comprimento: ‘Bat 93’ x ‘Jalo EEP558’ com 960 cM (NODARI et al. 1993), ‘XR-235-1’ x ‘Calima’ (VALLEJOS; SAKIYAMA; CHASE, 1992) com 827 cM e ‘PC-50’ x ‘XAN-159’ com 426 cM (JUNG et al. 1997). A integração de diferentes classes de marcadores, somada à amostragem de um número maior de linhagens desta população poderão fornecer uma maior consistência ao mapa CFM.

4.4.4 Alinhamento do mapa CFM ao mapa consenso BJ

Onze marcadores SSR-EST e outros 9 SSR genômicos foram mapeados tanto na população BJ quanto na população CFM. Os marcadores microssatélites comuns estiveram distribuídos em 9 grupos de ligação do mapa núcleo BJ. O mapeamento comparativo dos mapas das duas populações permitiu identificar, no mapa CFM, oito dos 11 grupos de ligação do mapa consenso BJ e determinar a orientação de cinco deles (Figura 13). Nos grupos B4 e B9 não foram alocados marcadores comuns ao mapa da população CFM e o marcador PvM95, mapeado no grupo B3, permaneceu não ligado no mapa CFM. Como apresentado no item 3.4.3, foram formados 12 grupos de ligação para a população CFM. Através dos marcadores PvM118, PvM103 e JF53, os grupos de ligação CFM1 e CFM12 foram alinhados ao grupo B8 de BJ.

Figura 13 – Alinhamento do mapa genético da população CFM (grupos de ligação designados de CFM1 a CFM12) ao mapa referência BJ (B1-B11) enriquecido neste trabalho com marcadores SSR e AFLP. Os marcadores microssatélites comuns aos dois mapas estão destacados em negrito e as linhas indicam os marcadores equivalentes. Para facilitar a visualização e a comparação entre os mapas, todos os marcadores redundantes e acessórios foram retirados.

A comparação entre os mapas mostrou que as distorções de segregação ocorreram em algumas posições equivalentes nos dois mapas BJ e CFM. Todavia, os marcadores comuns apresentaram distorções de segregação mais fortes na população CFM que quando mapeados na população BJ, a exemplo dos marcadores FJ_037 e PvM115 mapeados no grupo B2/CFM3, e de PvM118 no grupo B8/CFM12 (Figura 13). Além disso, os marcadores PvM14 e IAC001 mapeados no grupo B6/CFM11 e PvM 103 e FJ_53 nos grupos B8/CFM11 se mostraram distorcidos apenas na população CFM. A maior diferença entre as duas populações está no grupo CFM9 (correspondente ao grupo B10 no mapa BJ) onde foi observada distorção na segregação de todos os marcadores, enquanto que no mapa BJ não foi observada nenhuma marca com segregação distorcida (Figura 13).

Aqui, 8 dos grupos de ligação do mapa CFM foram alinhados aos respectivos grupos do mapa BJ. O mapeamento de outros marcadores do mapa BJ na população CFM completará este alinhamento, permitindo não apenas a correspondência entre os grupos de ligação, mas a transferência de informações entre os mapas. QTL relacionados ao florescimento e de resistência à mancha angular e ao oídio foram previamente mapeados na população CFM (MELO; SANTOS; FERREIRA, 2002a, b). O alinhamento do mapa CFM ao mapa consenso BJ permitirá que marcadores próximos às regiões contendo os QTL sejam identificados no mapa BJ e usados para saturar pontualmente as regiões de interesse na população CFM. Além disso, QTL para estes caracteres mapeados na população BJ, ou mesmo em outras populações cujos mapas foram integrados, permitirão o acúmulo de informações a cerca do controle de tais caracteres.

5 CONCLUSÕES

- O banco de EST de feijão é uma importante fonte de marcadores microssatélites para *Phaseolus*.
- Marcadores SSR-EST são úteis para estudos de diversidade genética e para a construção de mapas moleculares e particularmente importantes para o alinhamento de diferentes mapas genéticos de feijão-comum.
- Marcadores AFLP podem ser usados para a construção e a saturação de mapas de ligação em feijão-comum.
- Marcadores AFLP gerados com a enzima *Pst*-I estão mais uniformemente distribuídos no genoma do feijoeiro que aqueles gerados com a enzima *Eco*-RI.
- O uso combinado de marcadores SSR-EST e AFLP resulta em considerável saturação de mapas genéticos, como exemplificado, neste trabalho, para o feijão-comum.

REFERÊNCIAS

ABREU, A. D. B.; RAMALHO, M. A. P.; SANTOS, J. B.; MARTINS, L. A. Progresso do melhoramento genético do feijoeiro nas décadas de setenta e oitenta nas regiões Sul e Alto Paranaíba em Minas Gerais. **Pesquisa Agropecuária Brasileira**, Brasília, v. 29, p. 105-112, 1994.

ACOSTA-GALLEGOS, J. A.; KELLY, J. D.; GEPTS, P. Prebreeding in common bean and use of genetic diversity from wild germplasm. **Crop Science**, Madison, v. 47, p. S44-S59, 2007.

ADAM-BLONDON, A. F.; SEVIGNAC, M.; BANNEROT, H.; DRON, M. SCAR, RAPD and RFLP markers linked to a dominant gene (Are) conferring resistance to anthracnose in common bean. **Theoretical and Applied Genetics**, Berlin, v. 88, p. 865-870, 1994.

AKKAYA, M. S.; SHOEMAKER, R. C.; SPECHT, J. E.; BHAGWAT, A. A.; CREGAN, P. B. INTEGRATION OF SIMPLE SEQUENCE REPEAT DNA MARKERS INTO A SOYBEAN LINKAGE MAP. **Crop Science**, Madison, v. 35, p. 1439-1445, 1995.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic local alignment search tool. **Journal of Molecular Biology**, London, v. 215, p. 403-410, 1990.

ANDERSON, J. A.; CHURCHILL, G. A.; AUTRIQUE, J. E.; TANKSLEY, S. D.; SORRELLS, M. E. Optimizing parental selection for genetic linkage maps. **Genome**, Ottawa, v. 36, p. 181-186, 1993.

ARIYARATHNE, H. M.; COYNE, D. P.; JUNG, G.; SKROCH, P. W.; VIDAVER, A. K.; STEADMAN, J. R.; MIKLAS, P. N.; BASSETT, M. J. Molecular mapping of disease resistance genes for halo blight, common bacterial blight, and bean common mosaic virus in a segregating population of common bean. **Journal of the American Society for Horticultural Science**, Alexandria, v. 124, p. 654-662, 1999.

BASSETT, M. J. List of genes – *Phaseolus vulgaris* L. **Annual Report of the Bean Improvement Cooperative**, East Lansing, v. 39, p. 1-19, 1996.

BASSETT, M. J. A revised linkage map of common bean. **Hortscience**, Alexandria, v. 26, p. 834-836, 1991.

BATESON, W.; PUNNETT, R. C. Experimental studies in the physiology of heredity. **Reports to the Evolution Committee of the Royal Society**, London, v. 3, p. 11-30, 1906.

BEARZOTI, E. Mapeamento de QTL. In: PINHEIRO, J. B.; CARNEIRO, I. F. **Análise de QTL no melhoramento de plantas**. Goiânia: FUNAPE, 2000. p. 63-223.

BEEBE, S. E.; ROJAS-PIERCE, M.; YAN, X.; BLAIR, M. W.; PEDRAZA, F.; MUÑOZ, F.; TOHME, J.; LYNCH, J. P. Quantitative trait loci for root architecture traits correlated with phosphorus acquisition in common bean. **Crop Science**, Madison, v. 46, p. 413-423, 2006.

BENCHIMOL, L. L.; DE CAMPOS, T.; CARBONELL, S. A. M.; COLOMBO, C. A.; CHIORATTO, A. F.; FORMIGHIERI, E. F.; GOUVEA, L. R. L.; DE SOUZA, A. P. Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. **Genetic Resources and Crop Evolution**, Dordrecht, v. 54, p. 1747-1762, 2007.

BENNETT, M. D.; LEITCH, I. J. Nuclear DNA amounts in angiosperms: Progress, problems and prospects. **Annals of Botany**, London, v. 95, p. 45-90, 2005.

BLAIR, M. W.; PEDRAZA, F.; BUENDIA, H. F.; GAITAN-SOLIS, E.; BEEBE, S. E.; GEPTS, P.; TOHME, J. Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). **Theoretical and Applied Genetics**, Berlin, v. 107, p. 1362-1374, 2003.

BONIERBALE, M. W.; PLAISTED, R. L.; TANKSLEY, S. D. RFLP Maps Based on a Common Set of Clones Reveal Modes of Chromosomal Evolution in Potato and Tomato. **Genetics**, Austin, v. 120, p. 1095-1103, 1988.

BOREVITZ, J. O.; LIANG, D.; PLOUFFE, D.; CHANG, H. S.; ZHU, T.; WEIGEL, D.; BERRY, C. C.; WINZELER, E.; CHORY, J. Large-scale identification of single-feature polymorphisms in complex genomes. **Genome Research**, Woodbury, v. 13, p. 513-523, 2003.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of a genetic-linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, Chicago, v. 32, p. 314-331, 1980.

BROUGHTON, W. J.; HERNANDEZ, G.; BLAIR, M.; BEEBE, S.; GEPTS, P.; VANDERLEYDEN, J. Beans (*Phaseolus* spp.) - model food legumes. **Plant and Soil**, The Hague, v. 252, p. 55-128, 2003.

BUSO, G. S. C.; AMARAL, Z. P. S.; BRONDANI, R. P. V.; FERREIRA, M. E. Microsatellite markers for the common bean *Phaseolus vulgaris*. **Molecular Ecology Notes**, Oxford, v. 6, p. 252-254, 2006.

- BUTCHER, P. A.; WILLIAMS, E. R.; WHITAKER, D.; LING, S.; SPEED, T. P.; MORAN, G. F. Improving linkage analysis in outcrossed forest trees: an example from *Acacia mangium*. **Theoretical and Applied Genetics**, Berlin, v. 104, p. 1185-1191, 2002.
- CAIXETA, E. T.; OLIVEIRA, A. C. B.; BRITO, G. G.; SAKIYAMA, N. S. Tipos de Marcadores Moleculares. In: BORÉM, A.; CAIXETA, E. T. **Marcadores Moleculares**. Viçosa: UFV, 2006. cap. 1, p. 9-78.
- CAMPOS, T.; BENCHIMOL, L. L.; CARBONELL, S. A. M.; CHIORATTO, A. F.; FORMIGHIERI, E. F.; DE SOUZA, A. P. Microsatellites for genetic studies and breeding programs in common bean. **Pesquisa Agropecuaria Brasileira**, Brasília, v. 42, p. 589-592, 2007.
- CARDLE, L.; RAMSAY, L.; MILBOURNE, D.; MACAULAY, M.; MARSHALL, D.; WAUGH, R. Computational and experimental characterization of physically clustered simple sequence repeats in plants. **Genetics**, Austin, v. 156, p. 847-854, 2000.
- CARNEIRO, M. S.; VIEIRA, M. L. C. Mapas genéticos em plantas. **Bragantia**, Campinas, v. 61, p. 89-100, 2002.
- CARNEIRO, M. S.; CAMARGO, L. E.; COELHO, A. S.; VENCOVSKY, R.; RUI, P. L.; STENZEL, N. M.; VIEIRA, M. L. RAPD-based genetic linkage maps of yellow passion fruit (*Passiflora edulis* Sims. f. *flavicarpa* Deg.). **Genome**, Ottawa, v. 45, p. 470-478, 2002.
- CHACÓN, M. I.; PICKERSGILL, B.; DEBOUCK, D. G. Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. **Theoretical and Applied Genetics**, Berlin, v. 110, p. 432-444, 2005.
- CHAO, S.; SHARP, P. J.; GALE, M. D. A linkage map of wheat homoeologous group 7 chromosomes using RFLP markers. In: INTERNATIONAL WHEAT GENETICS SYMPOSIUM, 7., 1988, Cambridge. **Proceedings...** Cambridge: Bath Press, 1988. p. 493-498.
- CHASE, C. D.; ORTEGA, V. M.; VALLEJOS, C. E. DNA restriction fragment length polymorphisms correlate with isozyme diversity in *Phaseolus vulgaris* L. **Theoretical and Applied Genetics**, Berlin, v. 81, p. 806-811, 1991.

CHO, Y. G.; ISHII, T.; TEMNYKH, S.; CHEN, X.; LIPOVICH, L.; PARK, W. D.; AYRES, N.; CARTINHO, S.; MCCOUCH, S. R. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). **Theoretical and Applied Genetics**, Berlin, v. 100, p. 713-722, 2000.

CONAB. **Central de informações agropecuárias**. Disponível em: <<http://www.conab.gov.br/conabweb/index.php?PAG=101>>. Acesso em: 05 mai. 2008

CONCIBIDO, V.; VALLEE, B.; MCCLAIRD, P.; PINEDA, N.; MEYER, J.; HUMMEL, L.; YANG, J.; WU, K.; DELANNAY, X. Introgression of a quantitative trait locus for yield from Glycine soja into commercial soybean cultivars. **Theoretical and Applied Genetics**, Berlin, v. 106, p. 575-582, 2003.

CORDEIRO, G. M.; CASU, R.; MCINTYRE, C. L.; MANNERS, J. M.; HENRY, R. J. Microsatellite markers from sugarcane (*Saccharum* spp) ESTs across transferable to erianthus and sorghum. **Plant Science**, Limerick, v. 160, p. 1115-1123, 2001a.

CORDEIRO, G. M.; CASU, R.; MCINTYRE, C. L.; MANNERS, J. M.; HENRY, R. J.; 160(6):1115-1123, M. M. F. S. S. S. E. C. T. T. E. A. S. P. S. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. **Plant Science**, Limerick, v. 160, p. 1115-1123, 2001b.

CORREA, R. X.; COSTA, M. R.; GOOD-GOD, P. I.; RAGAGNIN, V. A.; FALEIRO, F. G.; MOREIRA, M. A.; DE BARROS, E. G. Sequence Characterized Amplified Regions Linked to Rust Resistance Genes in the Common Bean. **Crop Science**, Madison, v. 40, p. 804-807, 2000.

CRESTE, S.; TULMANN-NETO, A.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reporter**, Athens, v. 19, p. 1-8, 2001.

DÍAZ, C. G.; BASSANEZI, R. B.; GODOY, C. V.; LOPES, D. B.; BERGAMIN FILHO, A. Quantificação do efeito do cretamento bacteriano comum na eficiência fotossintética e na produção do feijoeiro. **Fitopatologia Brasileira**, Brasília, v. 26, p. 71-76, 2001.

DITA, M. A.; RISPAIL, N.; PRATS, E.; RUBIALES, D.; SINGH, A. B. Biotechnology approaches to overcome biotic and abiotic stress in legumes. **Euphytica**, Wageningen, v. 147, p. 1-24, 2006.

DOYLE, J. J.; DOYLE, J. L. A rapid DNA isolation procedure to small amounts of fresh leaf tissue. **Phytochemical Bulletin**, [s.l.], v. 19, p. 11-15, 1987.

EDWARDS, A.; CIVITELLO, A.; HAMMOND, H. A.; AND CASKEY, C. T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. **American Journal of Human Genetics**, Chicago, v. 49, p. 746-756, 1991.

EMBRAPA. Cultivo do Feijoeiro Comum. Disponível em: <<http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Feijao/CultivodoFeijoeiro>>. Acesso em: 05 mai de 2008.

ENDER, M.; TERPSTRA, K.; KELLY, J. D. Marker-assisted selection for white mold resistance in common bean. **Molecular Breeding**, Dordrecht, v. 21, p. 149-157, 2008.

EUJAYL, I.; BAUM, M.; POWELL, W.; ERSKINE, W.; PEHU, E. A genetic linkage map of lentil (*Lens sp.*) based on RAPD and AFLP markers using recombinant inbred lines. **Theoretical and Applied Genetics**, Berlin, v. 97, p. 83-89, 1998.

EUJAYL, I.; SORRELLS, M. E.; BAUM, M.; WOLTERS, P.; POWELL, W. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. **Theoretical and Applied Genetics**, Berlin, v. 104, p. 399-407, 2002.

EVANS, A. M. Beans. In: SIMMONDS, N. W. **Evolution of crop plants**. Hong Kong: Longman Scientific and Technical, 1986. p. 168-172.

FALK, C. T. A simple scheme for preliminary ordering of multiple loci: application 10 45 CF families. In: ELSTON, R. C.; A., S. M.; RODGE, S. E.; MCCLUE, J. W. **Multipoint mapping and linkage based upon affected pedigree members**. New York: Alan R Liss, 1989. chap. 6, p. 17-22.

FAO. **Food and Agriculture Organization of the United Nations**: Statistical database. Disponível em: <<http://faostat.fao.org/>>. Acesso em: 05 mai. 2008.

FEINGOLD, S.; LLOYD, J.; NORERO, N.; BONIERBALE, M.; LORENZEN, J. Mapping and characterization of new EST-derived microsatellites for potato (*Solanum tuberosum* L.). **Theoretical and Applied Genetics**, Berlin, v. 111, p. 456-466, 2005.

FELSENSTEIN, J. **PHYLIP (Phylogeny Inference Package) 3.6**. Seattle: Department of Genetics, University of Washington. 2004. Disponível em: <<http://evolution.genetics.washington.edu/phylip.html>>.

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

FREYRE, R.; SKROCH, P. W.; GEFFROY, V.; ADAM-BLONDOND, A. F.; SHIRMOHAMADALI, A.; JOHNSON, W. C.; LLACA, V.; NODARI, R. O.; PEREIRA, P. A.; TSAI, S. M.; TOHME, J.; DRON, M.; NIENHUIS, J.; VALLEJOS, V. A.; GEPTS, P. Towards an integrated linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP. **Theoretical and Applied Genetics**, Berlin, v. 97, p. 847-856, 1998.

GAITÁN-SOLÍS, E.; DUQUE, M. C.; EDWARDS, K. J.; TOHME, J. Microsatellite repeats in common bean (*Phaseolus vulgaris*): Isolation, characterization, and cross-species amplification in *Phaseolus* ssp. **Crop Science**, Madison, v. 42, p. 2128-2136, 2002.

GALE, M. D.; DEVOS, K. M. Plant Comparative Genetics after 10 Years. **Science**, Washington, v. 282, p. 656-659, 1998.

GAO, L. F.; TANG, J. F.; LI, H. W.; JIA, J. Z. Analysis of microsatellites in major crops assessed by computational and experimental approaches. **Molecular Breeding**, Dordrecht, v. 12, p. 245-261, 2003.

GARCIA, R.; ROBINSON, R. A.; AGUILAR, J. A.; SANDOVAL, S.; GUZMAN, R. Recurrent selection for quantitative resistance to soil-borne diseases in beans in the Mixteca region, Mexico. **Euphytica**, Wageningen, v. 130, p. 241-247, 2003.

GEPTS, P. Origin and evolution of common bean: past events and recent trends. **Hortscience**, Alexandria, v. 33, p. 1124-1130, 1998.

GEPTS, P.; KMIETIK, K.; PEREIRA, P.; BLISS, F. A. Dissemination pathways of common bean (*Phaseolus vulgaris*, Fabaceae) deduced from phaseolin electrophoretic variability. **Economic Botany**, Bronx, v. 40, p. 451-468, 1988.

GRATTAPAGLIA, D.; SEDEROFF, R. Genetic Linkage Maps of *Eucalyptus grandis* and *Eucalyptus urophylla* Using a Pseudo-Testcross: Mapping Strategy and RAPD Markers. **Genetics**, Austin, v. 137, p. 1121-1137, 1994.

GREEN, P.; FALLS, K.; CROOKS, S. **Documentation for CRI-MAP**. Saint Louis: Washington University School of Medicine. 1990. Disponível em:
<<http://linkage.rockefeller.edu/soft/crimap/>>

- GRISI, M. C. M.; BLAIR, M. W.; GEPTS, P.; BRONDANI, C.; PEREIRA, P. A. A.; BRONDANI, R. P. V. Genetic mapping of a new set of microsatellite markers in a reference common bean (*Phaseolus vulgaris*) population BAT93 x Jalo EEP558. **Genetics and Molecular Research**, Ribeirão Preto, v. 6, p. 691-706, 2007.
- GUERRA-SANZ, J. M. New SSR markers of *Phaseolus vulgaris* from sequence databases. **Plant Breeding**, Berlin, v. 123, p. 87-89, 2004.
- GUPTA, P. K.; KULWAL, P. L.; RUSTGI, S. Wheat cytogenetics in the genomics era and its relevance to breeding. **Cytogenetic and Genome Research**, Basel, v. 109, p. 315-327, 2005.
- HALDANE, J. B. S. The combination of linkage values, and the calculation of distance between the loci of linked factors. **Journal of Genetics**, Bangarole, v. 8, p. 299-309, 1919.
- HALL, R. **Compendium of Bean Disease**. St. Paul: APS Press, 1991. 73 p.
- HANAI, L. R.; CAMPOS, T.; CAMARGO, L. E. A.; BENCHIMOL, L. L.; DE SOUZA, A. P.; MELOTTO, M.; CARBONELL, S. A. M.; CHIORATTO, A. F.; CONSOLI, L.; FORMIGHIERI, E. F.; SIQUEIRA, M.; TSAI, S. M.; VIEIRA, M. L. C. Development, characterization, and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources. **Genome**, Ottawa, v. 50, p. 266-277, 2007.
- HU, J.; NAKATANI, M.; MIZUNO, K.; FUJIMURA, T. Development and characterization of microsatellite markers in sweetpotato. **Breeding Science**, Tokyo, v. 54, p. 177-188, 2004.
- INNAN, H.; TERAUCHI, R.; MIYASHITA, N. T. Microsatellite Polymorphism in Natural Populations of the Wild Plant *Arabidopsis thaliana*. **Genetics**, Austin, v. 146, p. 1441-1452, 1997.
- IWATA, H.; NINOMIYA, S. AntMap: Constructing Genetic Linkage Maps Using an Ant Colony Optimization Algorithm. **Breeding Science**, Tokyo, v. 56, p. 371-377, 2006.
- JACCOUD, D.; PENG, K.; FEINSTEIN, D.; KILIAN, A. Diversity Arrays: a solid state technology for sequence information independent genotyping. **Nucleic Acids Research**, Oxford, v. 29, n. e25, 2001. Disponível em: <<http://nar.oxfordjournals.org/cgi/reprint/29/4/e25>>
- JACOB, H. J.; LINDPAINTNER, K.; LINCOLN, S. E.; KUSUMI, K.; BUNKER, R. K.; MAO, Y.-P.; GANTEN, D.; DZAU, V. J.; LANDER, E. S. Genetic mapping of a gene causing hypertensive rat. **Cell**, Cambridge, v. 67, p. 213-224, 1991.

JEFFREYS, A. J.; WILSON, V.; NEWUMANN, R.; KEYTE, J. Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. **Nucleic Acids Research**, Oxford, v. 16, p. 10953-10971, 1988.

JUNG, G.; SKROCH, P. W.; COYNE, D. P.; NIENHUIS, J.; ARNAUD-SANTANA, E.; ARIYARATHNE; KAEPLER, S. M. Molecular-marker-based genetic analysis of tepary bean-derived common bacterial blight resistance in different developmental stages of common bean. **Journal of the American Society for Horticultural Science**, Alexandria, v. 122, p. 329-337, 1997.

JUNG, G.; SKROCH, P. W.; NIENHUIS, J.; COYNE, D. P.; ARNAUD-SANTANA, E.; ARIYARATHNE, H. M.; MARITA, J. M. Confirmation of QTL associated with common bacterial blight resistance in four different genetic backgrounds in common bean. **Crop Science**, Madison, v. 39, p. 1448-1455, 1999.

JUNG, G.; COYNE, D. P.; SKROCH, P. W.; NIENHUIS, J.; ARNAUD-SANTANA, E.; BOKOSI, J.; ARIYARATHNE, H. M.; STEADMAN, J. R.; BEAVER, J. S.; KAEPLER, S. M. Molecular markers associated with plant architecture and resistance to common blight, web blight, and rust in common beans. **Journal of the American Society for Horticultural Science**, Alexandria, v. 121, p. 794-803, 1996.

KAGA, A.; OHNISHI, M.; ISHII, T.; KAMIJIMA, O. A genetic linkage map of azuki bean constructed with molecular and morphological markers using an interspecific population (*Vigna angularis* x *V-nakashimae*). **Theoretical and Applied Genetics**, Berlin, v. 93, p. 658-663, 1996.

KELLY, J. D.; VALLEJO, V. A. A comprehensive review of the major genes conditioning resistance to anthracnose in common bean. **Hortscience**, Alexandria, v. 39, p. 1196-1207, 2004.

KELLY, J. D.; GEPTS, P.; MIKLAS, P. N.; COYNE, D. P. Tagging and mapping of genes and QTL and molecular marker-assisted selection for traits of economic importance in bean and cowpea. **Field Crops Research**, Amsterdam, v. 82, p. 135-154, 2003.

KIMATI, H.; RESENDE, J. A. M.; BERDAMIN FILHO, A.; CAMARGO, L. E. A. **Doenças das Plantas Cultivadas**. 4.ed. São Paulo: Agronômica Ceres, 2005. p.

KOINANGE, E. M. K.; SINGH, S. P.; GEPTS, P.; 36:1037-1045, G. C. O. T. D. S. I. C. B. C. S. Genetic control of the domestication syndrome in common bean. **Crop Science**, Madison, v. 36, p. 1037-1045, 1996.

KOLKMAN, J. M.; KELLY, J. D. QTL conferring resistance and avoidance to white mold in common bean. **Crop Science**, Madison, v. 43, p. 539-548, 2003.

KOSAMBI, D. D. The estimation of map distance from recombination values. **Annuaire of Eugenetics**, Washington, v. 12, p. 172-175, 1944.

KUMAR, L. S. DNA markers in plant improvement: An overview. **Biotechnology Advances**, New York, v. 17, p. 143-182, 1999.

LAGERCRANTZ, U.; ELLEGREN, H.; ANDERSSON, L. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. **Nucleic Acids Research**, Oxford, v. 21, p. 1111-1115, 1993.

LAMPRECHT, H. Weitere kopplungsstudien an *Phaseolus vulgaris* mitt einer übersicht über die Koppelungsgruppen. **Agri Hortique Genetica**, Landskrona, v. 19, p. 319-332, 1961.

LANDER, E.; GREEN, P.; ABRAHAMSON, J.; BARLON, A.; DALEY, M.; LINCOLN, S.; NEWBURG, L. MAPMAKER: an interactive computer package for constructing primary genetic linkages maps of experimental and natural populations. **Genomics**, San Diego, v. 1, p. 174-181, 1987.

LANDER, E. S.; GREEN, P. Construction of multilocus genetic linkage maps in humans. **Proceeding of the National Academic Science of the USA**, Washington, v. 84, p. 2363-2367, 1987.

LANDER, E. S.; BOTSTEIN, D. Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. **Genetics**, Austin, v. 121, p. 185-199, 1989.

LEAKEY, C. L. A. Genotypic and phenotypic markers in common bean. In: GEPTS, P. **Genetic resources of *Phaseolus* beans**. Dordrecht: Kluwer Academic, 1998. p. 245-327.

LEFEBVRE, V.; CHÈVRE, A. M. Tools for marking plant disease and pest resistance genes: a review. **Agronomie**, Paris , v. 15, p. 3-19, 1995.

LI, Y. C.; KOROL, A. B.; FAHIMA, T.; NEVO, E. Microsatellites within genes: Structure, function, and evolution. **Molecular Biology and Evolution**, Chicago, v. 21, p. 991-1007, 2004.

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

LIU, B. H. **Statistical genomics**. New York: CRC, 1998. 610 p.

LIU, B. H.; KNAPP, S. J. GMENDEL, a program for Mendelian segregation and linkage analysis of individual or multiple progeny population using log-likelihood ratios. **Journal of Heredity**, Washington, v. 8, p. 407-418, 1992.

LOPES, R. **Mapas de ligação AFLP e identificação de genes de resistencia à *Xanthomonas campestris* pv. *passiflorae* em maracujá-amarelo**. 2003. 126 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2003.

LOPES, R.; LOPES, M. T. G.; CARNEIRO, M. S.; MATTA, F. P.; CAMARGO, L. E. A.; VIEIRA, M. L. C. Linkage and mapping of resistance genes to *Xanthomonas axonopodis* pv. *passiflorae* in yellow passion fruit. **Genome**, Ottawa, v. 49, p. 17-29, 2006.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, 1998. 978 p.

LYONS, M. E.; DICKSON, M. H.; HUNTER, J. E. RECURRENT SELECTION FOR RESISTANCE TO WHITE MOLD IN PHASEOLUS SPECIES. **Journal of the American Society for Horticultural Science**, Alexandria, v. 112, p. 149-152, 1987.

MARGARIDO, G. R. A.; SOUZA, A. P.; GARCIA, A. A. F. OneMap: software for genetic mapping in outcrossing species. **Hereditas**, Lund, v. 144, p. 78-79, 2007.

MARKERT, C. L.; MOLLER, F. Multiple forms of enzymes: tissue, ontogenetic and species specific patterns. **Proceedings of the National Academic Science of the USA**, Washington, v. 45, p. 753-763, 1959.

MATISE, T. C.; PERLIN, M.; CHAKRAVARTI, A. Automated construction of genetic-linkage maps using an expert-system (MULTIMAP) - a human genome linkage map. **Nature Genetics**, New York, v. 6, p. 384-390, 1994.

MELO, L. C.; SANTOS, J. B.; FERREIRA, D. F. Mapeamento de QTLs para florescimento do feijoeiro com marcadores RAPD em diferentes ambientes. **Ciência e Agrotecnologia**, Lavras, v. 26, p. 768-779, 2002a.

MELO, L. C.; SANTOS, J. B.; FERREIRA, D. F. Mapeamento de QTLs para reação ao oídio e mancha angular do feijoeiro comum em diferentes locais. **Pesquisa Agropecuária Brasileira**, Brasília, v. 37, p. 1115-1126, 2002b.

MELO, L. C.; DEL PELOSO, M. J.; FARIA, C. J. D.; YOKOYAMA, M.; ROSARIA, L.; BRONDANI, R. P. V.; BRONDANI, C.; FARIA, L. C. D. **Controle Genético da Reação do Feijoeiro Comum ao Vírus do Mosaico Dourado**. Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2005. 16 p.

MELOTTO, M.; MONTEIRO-VITORELLO, C. B.; BRUSCHI, A. G.; CAMARGO, L. E. A. Comparative bioinformatic analysis of genes expressed in common bean (*Phaseolus vulgaris* L.) seedlings. **Genome**, Ottawa, v. 48, p. 562-570, 2005.

MICROSOFT. **Microsoft Office 97**. São Paulo: 1997. Disponível em: <1 CD-ROM>.

MIKLAS, P. N.; JOHNSON, W. C.; DELORME, R.; GEPTS, P. QTL conditioning physiological resistance and avoidance to white mold in dry bean. **Crop Science**, Madison, v. 41, p. 309-315, 2001.

MIKLAS, P. N.; HU, J. G.; GRUNWALD, N. J.; LARSEN, K. M. Potential application of TRAP (targeted region amplified polymorphism) markers for mapping and tagging disease resistance traits in common bean. **Crop Science**, Madison, v. 46, p. 910-916, 2006.

MIKLAS, P. N.; JOHNSON, E.; STONE, V.; BEAVER, J. S.; MONTROYA, C.; ZAPATA, M. Selective Mapping of QTL Conditioning Disease Resistance in Common Bean. **Crop Science**, Madison, v. 36, p. 1344-1351, 1996.

MIKLAS, P. N.; PASTOR-CORRALES, M. A.; JUNG, G.; COYNE, D. P.; KELLY, J. D.; MCCLEAN, P. E.; GEPTS, P. Comprehensive linkage map of bean rust resistance genes. **Annual Report of the Bean Improvement Cooperative**, East Lansing, v. 45, p. 125-129, 2002.

MIKLAS, P. N.; COYNE, D. P.; GRAFTON, K. F.; MUTLU, N.; REISER, J.; LINDGREN, D. T.; SINGH, S. P. A major QTL for common bacterial blight resistance derives from the common bean great northern landrace cultivar Montana No. 5. **Euphytica**, Wageningen, v. 131, p. 137-146, 2003.

MILLER, M. P. **Tools for population genetics analyses (TFPGA) 1.3:** a Windows program for the analysis of allozyme and molecular population genetic data. 1997. Disponível em: <<http://herb.bio.nau.edu/~miller/tfpga.htm>>.

MOLLINARI, M. **Comparação de algoritmos usados na construção de mapas genéticos.** 2007. 73 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2007.

MORETZSOHN, M. C.; LEOI, L.; PROITE, K.; GUIMARÃES, P. M.; LEAL-BERTIOLI, S. C.; GIMENES, M. A.; MARTINS, W. S.; F., V. J.; GRATTAPAGLIA, D.; BERTIOLI, D. J. A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). **Theoretical and Applied Genetics**, Berlin, v. 111, p. 1060-1071, 2005.

MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. **Nature Genetics**, New York, v. 30, p. 194-200, 2002.

MOSCONE, E. A.; KLEIN, F.; LAMBROU, M.; FUCHS, J.; SCHWEIZER, D. Quantitative karyotyping and dual-color FISH mapping of 5S and 18S-25S rDNA probes in the cultivated *Phaseolus* species (Leguminosae). **Genome**, Ottawa, v. 42, p. 1224-1233, 1999.

MULLIS, K. B.; FALOONA, F. A. Specific synthesis of DNA invitro via a polymerase-catalyzed chain-reaction. **Methods in Enzymology**, New York, v. 155, p. 335-350, 1987.

MUNDT, C. C. PROBABILITY OF MUTATION TO MULTIPLE VIRULENCE AND DURABILITY OF RESISTANCE GENE PYRAMIDS. **Phytopathology**, Saint Paul, v. 80, p. 221-223, 1990.

MURPHY, R. W.; SITES, J. W. J.; BUTH, D. G.; HAUFLE, C. H. Proteins I: Isozyme electrophoresis. In: HILLIS, D. M.; MORITZ, C. **Molecular Systematics**. Sunderland: Sinauer, 1990. p. 45-126.

MUTLU, N.; MIKLAS, P. N.; REISER, J.; COYNE, D. P. Backcross breeding for improved resistance to common bacterial blight in pinto bean (*Phaseolus vulgaris* L.). **Plant Breeding**, Berlin, v. 124, p. 282-287, 2005.

NAGL, W.; IGNACIMUTHU, S.; BECKER, J. Genetic engineering and regeneration of *Phaseolus* and *Vigna*. State of the art and new attempts. **Journal of Plant Physiology**, Stuttgart, v. 150, p. 625-644, 1997.

NAKAMURA, Y.; LEPPERT, M.; OCONNELL, P.; WOLFF, R.; HOLM, T.; CULVER, M.; MARTIN, C.; FUJIMOTO, E.; HOFF, M.; KUMLIN, E.; WHITE, R. Variable number of tandem repeat (VNTR) markers for human-gene mapping. **Science**, Washington, v. 235, p. 1616-1622, 1987.

NCBI. **National Center for Biotechnology Information**. Disponível em: <<http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucleotide>>. Acesso em: 15 dez. 2007.

NODARI, R. O.; TSAI, S. M.; GILBERTSON, R. L.; GEPTS, P. Towards an integrated linkage map of common bean. 2. Development of an RFLP-based linkage map. **Theoretical and Applied Genetics**, Berlin, v. 85, p. 513-520, 1993a.

NODARI, R. O.; TSAI, S. M.; GUZMÁN, P.; GILBERTSON, R. L.; GEPTS, P. Towards an integrated linkage map of common bean. 3. Mapping genetic factors controlling host-bacteria interactions. **Genetics**, Austin, v. 134, p. 341-350, 1993b.

OLIVEIRA, E. J. **Desenvolvimento e uso de marcadores microssatélites para construção e integração de mapas genéticos de maracujá-amarelo (*Passiflora edulis* Sims f. *flavicarpa* Deg.)**. 2006. 152 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2006.

OOIJEN, J. W. van; VOORRIPS, R. E. **JoinMap 3.0 Software for the calculations of genetic linkage maps**. Wageningen: Plant Research International. 2001. 1 CD ROM.

PAPA, R.; ACOSTA, J.; DELGADO-SALINAS, A.; GEPTS, P. A genome-wide analysis of differentiation between wild and domesticated *Phaseolus vulgaris* from Mesoamerica. **Theoretical and Applied Genetics**, Berlin, v. 111, p. 1147-1158, 2005.

PAREDES, O. M.; GEPTS, P. Segregation and recombination in inter-gene pool crosses of *Phaseolus vulgaris* L. **Journal of Heredity**, Washington, v. 86, p. 98-106, 1995.

PARK, S. O.; COYNE, D. P.; JUNG, G.; WSKROCH, P.; ARNAUD-SANTANA, E.; STEADMAN, J. R.; ARIYARATHNE, H. M.; NIENHUIS, J. Mapping of QTL for seed size and shape traits in common bean. **Journal of the American Society for Horticultural Science**, Alexandria, v. 125, p. 466-475, 2000.

PARRELLA, N. N. L. D. **Seleção de famílias de feijão com resistência à antracnose, produtividade e tipo de grão carioca**. 2006. 50 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Universidade Federal de Lavras, Lavras, 2006.

PATERSON, A. H. MOLECULAR DISSECTION OF QUANTITATIVE TRAITS - PROGRESS AND PROSPECTS. **Genome Research**, Woodbury, v. 5, p. 321-333, 1995.

PAULA JÚNIOR, T. J.; ZAMBOLIM, L. Doenças. In: VIEIRA, C.; PAULA JÚNIOR, T. J.; BORÉM, A. **Feijão: Aspéctos gerais e cultura no Estado de Minas**. Viçosa: UFV, 1998. p. 375-433.

PEAKALL, R.; GILMORE, S.; KEYS, W.; MORGANTE, M.; RAFALSKI, A. Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. **Molecular Biology and Evolution**, Chicago, v. 15, p. 1275-1287, 1998.

PEDROSA-HARAND, A.; PORCH, T.; GEPTS, P. Standard nomenclature for common bean chromosomes and linkage groups. **Annual Reptort of the Bean Improvement Cooperative**, East Lansing, v. 51, p. 106-107, 2008.

PEDROSA-HARAND, A.; DE ALMEIDA, C. C. S.; MOSIOLEK, M.; BLAIR, M.; SCHWEIZER, D.; GUERRA, M. Extensive ribosomal DNA amplification during Andean common bean (*Phaseolus vulgaris* L.) evolution. **Theoretical and Applied Genetics**, Berlin, v. 112, p. 924-933, 2006.

PEDROSA, A.; VALLEJOS, C. E.; BACHMAIR, A.; SCHWEIZER, D. Integration of common bean (*Phaseolus vulgaris* L.) linkage and chromosomal maps. **Theoretical and Applied Genetics**, Berlin, v. 106, p. 205-212, 2003.

PENG, J. H.; LAPITAN, N. L. V. Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. **Functional and Integrative Genomics**, Berlin , v. 5, p. 80 - 96, 2005.

PHILIPP, U.; WEHLING, P.; WRICKE, G. A. A linkage map of rye. **Theoretical and Applied Genetics**, Berlin, v. 88, p. 234-248, 1994.

POSA-MACALINCAG, M. C. T.; HOSFIELD, G. L.; GRAFTON, K. F.; UEBERSAX, M. A.; KELLY, J. D. Quantitative trait loci (QTL) analysis of canning quality traits in kidney bean (*Phaseolus vulgaris* L.). **Journal of the American Society for Horticultural Science**, Alexandria, v. 127, p. 608-615, 2002.

PRIOLLI, R. H. G.; MENDES-JUNIOR, C. T.; ARANTES, N. E.; CONTEL, E. P. B. Characterization of Brazilian soybean cultivars using microsatellite markers. **Genetics and Molecular Biology**, Ribeirão Preto, v. 25, p. 185-193, 2002.

RAFALSKI, A. Applications of single nucleotide polymorphisms in crop genetics. **Current Opinion in Plant Biology**, London, v. 5, p. 94-100, 2002.

RAMALHO, M. A. P.; ABREU, A. D. F. B.; SANTOS, J. B. Melhoramento de espécies autógamas. In: NASS, L. L. E. A. **Recursos Genéticos e melhoramento**. Rondonópolis: Fundação MT, 2001. 1183 p.

RAMALHO, M. A. P.; ABREU, A. D. B.; DOS SANTOS, J. B. Genetic progress after four cycles of recurrent selection for yield and grain traits in common bean. **Euphytica**, Wageningen, v. 144, p. 23-29, 2005.

RAMIREZ, M.; GRAHAM, M. A.; BLANCO-LOPEZ, L.; SILVENTE, S.; MEDRANO-SOTO, A.; BLAIR, M. W.; HERNANDEZ, G.; VANCE, C. P.; LARA, M. Sequencing and Analysis of Common Bean ESTs. Building a Foundation for Functional Genomics. **Plant Physiology**, Washington, v. 137, p. 1211-1227, 2005.

RAVA, C. A.; PRUCHIO, A. F.; SARTORATO, A. Caracterização de patótipos de *Colletotrichum lindemuthianum* que ocorrem em algumas regiões produtoras de feijão comum. **Fitopatologia Brasileira**, Brasília, v. 19, p. 167-172, 1993.

RAVA, C. A.; MOLINA, J.; KAUFFMANN, M.; BRIONES, I. Determinación de razas fisiológicas de *Colletotrichum lindemuthianum* en Nicaragua. **Fitopatologia Brasileira**, Brasília, v. 18, p. 388-391, 1993.

RIELY, B. K.; ANÉ, J.-M.; PENMETS, R. V.; COOK, D. R. Genetic and genomic analysis in model legumes bring Nod-factor signaling to center stage. **Current Opinion in Plant Biology**, London, v. 7, p. 408-413, 2004.

RODRIGUEZ-SUAREZ, C.; MENDEZ-VIGO, B.; PANEDA, A.; FERREIRA, J. J.; GIRALDEZ, R. A genetic linkage map of *Phaseolus vulgaris* L. and localization of genes for specific resistance to six races of anthracnose (*Colletotrichum lindemuthianum*). **Theoretical and Applied Genetics**, Berlin, v. 114, p. 713-722, 2007.

RONALD, J.; AKEY, J. M.; WHITTLE, J.; SMITH, E. N.; YVERT, G.; KRUGLYAK, L. Simultaneous genotyping, gene-expression measurement, and detection of allele-specific expression with oligonucleotide arrays. **Genome Research**, Woodbury, v. 15, p. 284-291, 2005.

ROZEN, S.; SKALETSKY, H. J. Primer3 on the WWW for general users and for biologist programmers. In: KRAWETZ, S.; MISENER, S. **Bioinformatics Methods and Protocols: Methods in Molecular Biology**. Totowa: Humana Press, 2000. p. 365-386.

SANWEN, H.; BAOXI, Z.; MILBOURNE, D.; CARDLE, L.; GUIMEI, Y.; JIAZHEN, G. Development of pepper SSR markers from sequence databases. **Euphytica**, Wageningen, v. 117, p. 163-167, 2001.

SARTORATO, A.; RAVA, C. A. Influência da cultivar e do número de inoculações na severidade da mancha-angular (*Isariopsis griseola*) e nas perdas na população do feijoeiro comum (*Phaseolus vulgaris*). **Fitopatologia Brasileira**, Brasília, v. 17, p. 247-251, 1992.

SCHLÖTTERER, C. The evolution of molecular markers - just a matter of fashion? **Nature Reviews Genetics**, London, v. 5, p. 63-69, 2004.

SCHNEIDER, K. A.; GRAFTON, K. F.; KELLY, J. D. QTL analysis of resistance to fusarium root rot in bean. **Crop Science**, Madison, v. 41, p. 535-542, 2001.

SCOTT, K. D.; EGGLER, P.; SEATON, G.; ROSSETTO, M.; ABLETT, E. M.; LEE, L. S.; HENRY, R. J. Analysis of SSRs derived from grape ESTs. **Theoretical and Applied Genetics**, Berlin, v. 100, p. 723-726, 2000.

SHAROPOVA, N.; MCMULLEN, M. D.; SCHULTZ, L.; SCHROEDER, S.; SANCHEZ-VILLEDA, H.; GARDINER, J.; BERGSTROM, D.; HOUCHINS, K.; MELIA-HANCOCK, S.; MUSKET, T.; DURU, N.; POLACCO, M.; EDWARDS, K.; RUFF, T.; REGISTER, J. C.; BROUWER, C.; THOMPSON, R.; VELASCO, R.; CHIN, E.; LEE, M.; WOODMAN-CLIKEMAN, W.; LONG, M. J.; LISCUM, E.; CONE, K.; DAVIS, G.; COE, E. H. J. Development and mapping of SSR markers for maize. **Plant Molecular Biology**, Dordrecht, v. 48, p. 463-481, 2002.

SIMON, C. J.; MUEHIBAUER, F. J. Construction of a Chickpea Linkage Map and Its Comparison With Maps of Pea and Lentil. **The Journal of Heredity**, Oxford, v. 88, p. 115-119, 1997.

SINGH, P.; TERAN, H.; LEMA, M.; WEBSTER, D. M.; STRAUSBAUGH, C. A.; MIKLAS, P. N.; SCHWARTZ, H. F.; BRICK, M. A. Seventy-five years of breeding dry bean of the western USA. **Crop Science**, Madison, v. 47, p. 981-989, 2007.

SINGH, S. P. Patterns of variation in cultivated common bean (*Phaseolus vulgaris*, Fabaceae). **Economic Botany**, Bronx, v. 43, p. 39-57, 1989.

SINGH, S. P. Broadening the genetic base of common bean cultivars: A review. **Crop Science**, Madison, v. 41, p. 1659-1675, 2001.

SINGH, S. P.; GEPTS, P.; DEBOUCK, D. G. Races of common bean (*Phaseolus vulgaris*). **Economic Botany**, Bronx, v. 45, p. 379-396, 1991.

SONG, Q. J.; MAREK, L. F.; SHOEMAKER, R. C.; LARK, K. G.; CONCIBIDO, V. C.; DELANNAY, X.; SPECHT, J. E. A new integrated genetic linkage map of the soybean. **Theoretical and Applied Genetics**, Berlin, v. 109, p. 122-128, 2004.

SONNANTE, G.; STOCKTON, T.; NODARI, R. O.; BECERRA VELÁSQUEZ, V. L.; GEPTS, P. Evolution of genetic diversity during the domestication of common-bean (*Phaseolus vulgaris* L.). **Theoretical and Applied Genetics**, Berlin, v. 89, p. 629-635, 1994.

SPRUYT, M.; BUQUICCHIO, F. **Gene Runner 3.01**. Hastings Software. 1994. Disponível em: <<http://www.generunner.net/>>.

STAUB, J. E.; SERQUEN, F. C.; GUPTA, M. Genetic markers, map construction, and their application in plant breeding. **HortScience**, Alexandria, v. 31, p. 729-740, 1996.

STAVELY, J. R.; MCMILLAN, R. T.; BEAVER, J. S.; MIKLAS, P. N. Release of three McCaslan type, indeterminate, rust and golden mosaic resistant snap bean germplasm lines BelDade RGM 4,5 and 6. **Annual Report of the Bean Improvement Cooperative**, East Lansing, v. 44, p. 197-198, 2001.

STURTEVANT, A. H. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. **Journal of Experimental Zoology**, New York, v. 14, p. 43-59, 1913.

TANKSLEY, S. D. Mapping polygenes. **Annual Review of Genetics**, Palo Alto, v. 27, p. 205-233, 1993.

TANKSLEY, S. D.; RICK, C. M. Isozymic linkage map of the tomato; application in genetics and breeding. **Theoretical and Applied Genetics**, Berlin, v. 57, p. 161-170, 1980.

TAR'AN, B.; MICHAELS, T. E.; PAULS, K. P. Genetic mapping of agronomic traits in common bean. **Crop Science**, Madison, v. 42, p. 544-556, 2002.

THIEL, T.; MICHALEK, W.; VARSHNEY, R. K.; GRANER, A.; 106(3):411-422., E. E. D. F. T. D. A. C. O. G.-D. S.-M. I. B. H. V. L. T. A. G. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). **Theoretical and Applied Genetics**, Berlin, v. 106, p. 411-422, 2003.

TIAN, A.-G.; WANG, J.; CUI, P.; HAN, Y.-J.; XU, H.; CONG, L.-J.; HUANG, X.-G.; WANG, X.-L.; JIAO, Y.-Z.; WANG, B.-J.; WANG, Y.-J.; ZHANG, J.-S.; CHEN, S.-Y. Characterization of soybean genomic features by analysis of its expressed sequence tags. **Theoretical and Applied Genetics**, Berlin, v. 108, p. 003-913, 2004.

TÓTH, G.; GASPARI, Z.; JURKA, J. Microsatellites in different eukaryotic genomes: survey and analysis. **Genome Research**, Woodbury, v. 10, p. 967-81., 2000.

VALLEJOS, C. E.; SAKIYAMA, N. S.; CHASE, C. D. A molecular marker-based linkage map of *Phaseolus-vulgaris* L. **Genetics**, Austin, v. 131, p. 733-740, 1992.

VALLEJOS, C. E.; MALANDRO, J. J.; SHEEHY, K.; ZIMMERMANN, M. J. Detection and cloning of expressed sequences linked to a target gene. **Theoretical and Applied Genetics**, Berlin, v. 101, p. 1109-1113, 2000.

VARSHNEY, R. K.; GRANER, A.; SORRELLS, M. E. Genic microsatellite markers in plants: features and applications. **Trends in Biotechnology**, Amsterdam, v. 23, p. 48-55, 2005.

VARSHNEY, R. K.; THIEL, T.; STEIN, N.; LANGRIDGE, P.; GRANER, A. *In silico* analysis on frequency and distribution of microsatellites in ESTs of some cereal species. **Cellular and Molecular Biology Letters**, Warsaw, v. 7, p. 537-46, 2002.

VARSHNEY, R. K.; SIGMUND, R.; BORNER, A.; KORZUN, V.; STEIN, N.; SORRELLS, M. E.; LANGRIDGE, P.; GRANER, A. Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. **Plant Science**, Limerick, v. 168, p. 195-202, 2005.

VIEIRA, C.; BORÉM, A.; RAMALHO, M. A. P.; CARNEIRO, J. E. S. Melhoramento do feijão. In: BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: UFV, 2005. cap. 9, p. 301-391.

VOORRIPS, R. E. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. **Journal of Heredity**, Washington, v. 93, p. 77-78, 2002.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VANDELEE, T.; HORNES, M.; FRIJTERS, A.; POT, J.; PELEMAN, J.; KUIPER, M.; ZABEAU, M. AFLP - A new technique for DNA-fingerprinting. **Nucleic Acids Research**, Oxford, v. 23, p. 4407-4414, 1995.

VUYLSTEKE, M.; MANK, R.; ANTONISE, R.; BASTIAANS, E.; SENIOR, M. L.; STUBER, C. W.; MELCHINGER, A. E.; LUBBERSTEDT, T.; XIA, X. C.; STAM, P.; ZABEAU, M.; KUIPER, M. Two high-density AFLP (R) linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. **Theoretical and Applied Genetics**, Berlin, v. 99, p. 921-935, 1999.

WALLEN, V. R.; JACKSON, H. R. Model for yield loss determination of bacterial blight of field beans utilizing infrared photography combined with field plot studies. **Phytopathology**, Saint Paul, v. 65, p. 942-948, 1975.

WANG, Z.; WEBER, J. L.; ZHONG, G.; TANKSLEY, S. D. Survey of plant short tandem repeats. **Theoretical and Applied Genetics**, Berlin, v. 88, p. 1-6, 1994.

WEEKS, D.; LANGE, K. Preliminary ranking procedure for multilocus ordering. **Genomics**, San Diego, v. 1, p. 236-242, 1987.

WEIR, B. **Genetic data analysis**. 2.ed. Sunderland: Sinauer Associates, 1996. 447 p.

WILLIAMS, J. G. K.; KUBELIK, A. R.; LIVAK, K. J.; RAFALSKI, J. A.; TINGEY, S. V. DNA polymorphisms amplified by arbitrary primers are useful as genetic-markers. **Nucleic Acids Research**, Oxford, v. 18, p. 6531-6535, 1990.

WILSON, S. R. A major simplification in preliminary ordering of linked loci. **Genetic Epidemiology**, New York, v. 5, p. 75-80, 1988.

WRIGHT, S. **Evolution and the genetics of populations**. Chicago: Univ. Chicago Press, 1978. 511 p.

YOSHII, K.; GALVEZ, G. E.; ALVAREZ, G. Estimation of yield losses in beans caused by common blight. **Proceeding of the American Phytopathology Society**, Saint Paul, v. 3, p. 298-299, 1976.

YOUNG, R. A.; KELLY, J. D. RAPD markers linked to three major anthracnose resistance genes in common bean. **Crop Science**, Madison, v. 37, p. 940-946, 1997.

YOUNG, R. A.; MELOTTO, M.; NODARI, R. O.; KELLY, J. D. Marker-assisted dissection of the oligogenic anthracnose resistance in the common bean cultivar, 'G 2333'. **Theoretical and Applied Genetics**, Berlin, v. 96, p. 87-94, 1998.

YOUNG, W. P.; SCHUPP, J. M.; KEIM, P. DNA methylation and AFLP marker distribution in the soybean genome. **Theoretical and Applied Genetics**, Berlin, v. 99, p. 785-792, 1999.

YU, K.; PARK, S. J.; POYSA, V.; GEPTS, P. Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). **Journal of Heredity**, Washington, v. 91, p. 429-434, 2000.

YU, K. F.; PARK, S. J.; POYSA, V. Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). **Genome**, Ottawa, v. 42, p. 27-34, 1999.

YU, Z. H.; STALL, R. E.; VALLEJOS, C. E. Detection of genes for resistance to common bacterial blight of beans. **Crop Science**, Madison, v. 38, p. 1290-1296, 1998.

ZHU, H.; CANNON, S. B.; YOUNG, N. D.; COOK, D. R. Phylogeny and Genomic Organization of the TIR and Non-TIR NBS-LRR Resistance Gene Family in *Medicago truncatula*. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 15, p. 529-539, 2002.

ANEXOS

ANEXO A – Tabela A1

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primer</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continua)	
						PIC ⁴	Loco transf. ⁵
PvM001/TC34	TTC7	GCTCTAGCCTCAATCCTTGG GGTTCCTGCTTCCCACCTT	240/240	1	4	0,60	V, PI
PvM002/TC116	CTT6	CGCCATTTGGATTGGATT AGGCGTGGAAGTGGAGTG	191/191	1	6	0,73	V, PI
PvM003/TC116	TTC6	CCGCCTTCTTCTTCTTCTTC CGGCGAGTCATCTTTTCC	167/167	1	4	0,62	V, PI
PvM004/TC323	TTC10	GGTTCCTCCTCCTTCTGCT GCGCCGTCTTTTGGTAGT	210/210	2	3	0,65	V, PI
PvM005/TC559	TA19	AAGATCCACGGTGAAGGTG GCCCTCACTGCTATGGTCA	204/	NO			
PvM006/TC103	TAC6	TTGCCAACAAATGGAAGTGA TTGCACAATGAGAAGGGAGA	190/170	2	3	0,15	V, PI
PvM007/TC127	ATG6	CTGGAATGGCTTACAGATGC CCAAGATCAGGCACAGTGAA	213/213	2	2	0,46	PI
PvM008/TC134	AGT5	TGGAAAACAGGGCTGAATG AGCGCAGTTGGAGTAAGCA	171/	NO			
PvM009/PVEPSE3030D08	AAG5	GCTTGACGCCTTTGAACC CGTGGGCATTGATTCTCTC	216/216	1	1	Mono	V, PI
PvM010/PVEPSE3030I07	ACT5	GAAGGGCAATGGCTGAAG TCCCGTGGAACACTTGG	256/256	1	NA		
PvM011/PVEPSE3030J03	AGA6	TGGGATTTTGAGCGTGTG TCCACCGACTTACCGAACA	154/154	2	3	0,46	V, PI
PvM012/PVEPSE3030L20	TGC5	CCAGTGTGTCAACAAATGGA AGGCTTAGGGAATGGCTCA	191/191	2	2	0,38	V, PI
PvM013.1/PVEPSE3030M09	GAA6	GAGAAGCCGCAGAGAGGA AGATGCCGCGAACAGAAC	216/216	2	4	0,71	V, PI

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM013.2/PVEPSE3030M09	GAA6	GAGAAGCCGCAGAGAGGA AGATGCCCGCGAACAGAAC	216/240	2	4	0,73	V, PI
PvM014/PVEPLE1004G06	AATC5	AAAGCCACAAAAACCATAAACC GCCTCCGAAAGATTCAACG	204/204	2	3	0,61	PI
PvM015/PVEPSE2010E01	TGCA5	GGTGGAGGCAGTGGTGT TTCACCTTCCCAAGACCTGA	185/185	2	2	0,46	V, PI
PvM016/PVEPSE2022E07	TTC7	ATATCGCCACCCACTCCA AACCGAAAGGGCAAAACG	179/179	2	2	0,08	V, PI
PvM017/PVEPSE2030E02	ATGA5	AGGATGGGTTCCTGTGCTT GGCAGCTTCCACATCGTC	213/213	2	2	0,40	V, PI
PvM018/PVEPSE3003E05	AATT5	CAACCATCTCTGTGGCAAG TGAGTTTGGGAGAGTGAGTTTT	202/202	2	3	0,58	V, PI
PvM019/PVEPSE3022E08	AGC5	ACCAGGTGAAAGCCTTGAG AGCTGCTGCTGCTGCTTC	169/	NO			
PvM020/PVEPLE1004B07	CT7	CGCTTTGATAAAACGCTTGC TGTTGGTGCTAACGATGGAA	222/340	2	1	Mono	V, PI
PvM021/TC449	AT14	CCCATCTTCCATTTCCACAG AAGCCACAGAAAACCTGAGCA	251/251	2	12	0,86	V, PI
PvM022/PVEPSE2004H10	TC5	ACTCTCACAATGGCGGAATC GGCGTTTTCTCCCTCTTCTT	222/222	2	3	0,40	V, PI
PvM023/PVEPSE2004G02	AG5	CCTGTTCCCTCTGTGTGAGA AATCAACCCCATCAAACCAA	219/700	2	1	Mono	V, PI
PvM024/PVEPSE2009B07	TG5	CAGAAAAATGGGCACAAGG AAAGTCAACACGCACACACC	345/345	2	1	Mono	V, PI
PvM025/TC54	TGA5	GAAGACGTGAAGGAGCCACT CACTTTCCAGAGAGAACGCAAT	183/183	2	2	0,08	V, PI
PvM026/TC80	TCA7	TTCTCCACCCCATTAAGTGC CTCGGAACAAAGTCCTCCAC	206/206	2	2	0,08	V, PI

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM027/TC527	GTT5	GCAGACCCTGTCAACAACAA GCCATGAAGCAAGGTGAAGA	192/192	2	1	Mono	V, PI
PvM028/TC390	TCA5	GGAGAGGAAACACGTAATGGA TCGACAACAAAAGAAAGAGAA	189/189	2	2	0,37	V, PI
PvM029/TC487	AAG5	CAGCAGAACAGCAAGTGATTC GGTGGCAGGTACGAAAGAGT	275/275	2	2	0,08	V, PI
PvM030/TC159	TGC5	GGAGTTGAAAAGGGTGGCTA TGATACAGAAGCAGGGAGAACA	222/222	2	5	0,55	V, PI
PvM031/PVEPSE3030F06	ATC5	GAGAAGCCAGGAACCTTA GCATCCCTTTTGTGTC	159/159	3	NA		
PvM032/PVEPSE3029G08	TTG5	GCCTTCTTTTCCAGCAG GCTTATTGTTGTTTTCTTTCTC	220/220	3	1	Mono	V, PI
PvM033/PVEPSE3027C10	GAG5	CTTCTCCCTCTTCCTCCT GGACAGTGATAGTTCAGATCC	273/273	3	2	0,08	V, PI
PvM034/PVEPSE3022E01	TCT6	GCCGCTGTCTTTGTCCAC GCATTCAAGGTTTTTCACCAA	240/	3	3	0,58	V, PI
PvM035/PVEPSE3022B11	AGA5	TCTGAAGCAGTCTCTCTGG ATCGTCTTACCACAAACG	102/	NO			
PvM036/PVEPSE3011F10	ATC5	GCCCCCTATTTCCCTACC CATCTTCTTCCTCACCTTA	201/201	3	2	0,47	V, PI
PvM037/PVEPSE3011A04	GAG7	GGACAGGGGAATCTCACTA CGTGGTCGTGGTGGTG	158/158	3	2	0,37	V, PI
PvM038/PVEPSE3006H03	AAT7	TTAGGGTCTCATAACAAA CAGTAAACTCACAAATATAC	162/	NO			
PvM039/PVEPSE3006G01	GAT6	ATCTCTTTCGGGCAGA ATTGTGCTTTCTCAACG	133/133	3	1	Mono	V, PI
PvM040/PVEPSE3001F10	CTG6	CTCTCCATTTTGTCTAG GAACCATCAGAACTAC	202/202	3	3	0,55	V, PI

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM041/PVEPSE2029D10	CT5	AAATGTCCCCAGCCACAC TTCCCTTCACCTTACCCTTC	174/300	3	1	Mono	V, Pl
PvM042/PVEPSE2027C10	AG5	CCATCGTCCTTTTCTTCAACA CACTGTCCTTGGGGCTATC	255/255	3	2	0,08	V, Pl
PvM043/PVEPSE2026A06	AG10	CGAGCAAAGACATCGGTTT CATTGTAAGCCAGGGAAGG	112/	4	1	Mono	
PvM044/PVEPSE2024H04	TC8	CTCTTACCAGCAGCATCA CTCCTTCTCTCTTTTCTCT	182/600	3	1	Mono	V, Pl
PvM045/PVEPSE2024D09	CT24	GAGGGTAGTTTCTGGA GACCAAGTTCTCTTTC	232/232	3	4	0,66	V, Pl
PvM046/PVEPSE2024B07	GA5	CGCTTATGTTGGGTGGAG CGTGGAAGGAGAGGGAGA	246/246	3	2	0,47	V, Pl
PvM047/PVEPSE2022E03	CA5	AGCACTATGATTTTGAAT CCTGGGTTGAAGGAAT	205/700	3	NA		
PvM048/PVEPSE2021H06	CT6	GCAAATGGTGGGGTGATG AGGGAGGTGAGGTGGTGA	122/122	3	1	Mono	V, Pl
PvM049/PVEPSE2027H04	TC6	ATCTCTGCGTCGGAGTGTTT TATCCTTCGCCCAATCCATA	197/197	3	2	0,27	V, Pl
PvM050/PVEPSE2025C02	TC5	GCGTGGGGTCTTTTCTCTCT AACGCATGGAAGTGCAAGTT	206/206	4	1	Mono	V, Pl
PvM051/PVEPSE2019C07	CA5	CCAACCACCCTTCTTCTCTCT ATCCGTGCCAACCATCTTT	199/199	4	2	0,22	V, Pl, Pa
PvM052/PVEPSE2018H03	CA7	TCCTCTGCAAACAAACCCTA TTCGCCAAATTCGACAATCT	204/500	4	3	0,62	V, Pl, Pa
PvM053/PVEPSE2016A02	TC6	TGGGGGAATTGAGTGAACAT TCTTTCTTGCGCTTTTCTCTTC	200/200	3	4	0,51	Pa
PvM054/PVEPSE2015G11	TC5	GCATCGTTGTTTCCTTCACA GATTGGAGAGCGGATTGGTA	193/193	3	2	0,09	Pa

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM055/PVEPSE2014C04	AT6	TGGTCTTCCTGGAGATGAAAG GCACAAAGTGCCCAAATG	209/209	1	1	Mono	
PvM056/PVEPSE2013B09	TA6	AGCTGGCAGAACTTGGA TTGAAACCCAGGTCTATTCTCTTC	214/214	3	4	0,62	Pa
PvM057/PVEPSE2030D11	CT5	TCCTCAAATGCTTGGTTTCA CCGGTTCTCACCAGATTGTT	199/199	1	1	Mono	
PvM058/PVEPSE2031B01	TC5	ATTGGGACGGGGATTG ACTTTTCCTTCTCTTCTCCTCT	151/151	3	3	0,38	Pa
PvM059/PVEPSE2031C12	CT5	TGAGGCACACTTCGACTACAA ATGGCAAGAACGTGAAAAGG	158/158	2	2	0,38	Pa
PvM060/PVEPSE2032B08	AC5	TTCACATTCTTCACTCACA ATCCCTTGGACGATGAC	161/161	3	2	0,08	Pa
PvM061/PVEPSE2033A05	AC6	GAAAAAGCTATGGGCCACAC CGGTAGAGAATCCAGCGAGA	231/231	2	2	0,39	
PvM062/PVEPSE2033F05	TC5	GCACCACTTTCTCAGATTTGG CTGTCATGGCGATGGTGTTA	172/172	3	3	0,48	Pa
PvM063/PVEPSE2034F05	AT5	TAACCACCCACAAGCATTGA TCCAAATAATTCCCCTGCTG	165/	NO			
PvM064/PVEPSE2035G05	GA6	AGTACCCTACACGTATCACTG GCCACTAGCCCATCTTCT	163/163	1	NA		
PvM065/PVEPSE2129G03	TA5	GAGCGCACAAATGCAATAAT GGGAGGATAACGACAATAGG	241/241	1	2	0,09	Pa
PvM066/PVEPSE3001H04	AT5	CCTCAGCAAGCTCCGTAAAC CATGCAACCTGAAACACTCC	196/196	1	2	0,51	Pa
PvM067/PVEPSE3002D09	GT5	CTGGTGCAGCAAGGTTTTTC TGCGCCAGAATTTACAATACAC	216/216	1	2	0,09	Pa
PvM068/PVEPSE3007B09	AT6	CCATTACGTGTCCAAGGTCTC CAAAAAGGAATCGGACTCTAGG	269/269	3	3	0,51	Pa

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM069/PVEPSE3008F01	TC7	CAATAAAAACCGCAGAGAATCC GCGTAAGCAGGTGGGTAAGT	217/217	1	2	0,09	
PvM070/PVEPSE3010C09	AT6	GAAGGGTCATTGTGGAC TCTTTTGTCACTCCTAATCC	161/161	1	NA		
PvM071/PVEPSE3011E03	CT5	ACGGATGGCAAGGACTCA AACACTACCGCAGCACCA	188/	NO			
PvM072/TC144	TC9	GAAACACAATGGCTCAAGCA CAAACCAGCAGCAACAAGAG	200/200	3	1	Mono	Pa
PvM073/TC154	AAG6	CTTCACCGATCTGACAGCAG TTCTCCAGGAACCCCTTCTTT	194/600	4	3	0,44	Pa
PvM074/TC241	GCG5	TGAACACCCTTTCGTCAACTC CCTTCTCCTCCCTGATTTT	200/450	4	2	0,08	Pa
PvM075/PVEPLE1001E04	GAT5	ATTGGAAGGGGGATGAACCT TAGGAGAGTGCCCAGTGCTT	222/222	4	3	0,44	Pa
PvM076/PVEPLE1002A08	CTT7	TCCCTGGAAGGAAAAATGGT CGAAGTAAAGAGGGGTGGAA	196/	NO			
PvM077/PVEPLE1005D03	TGA6	AGGGGTGAGTATGGTTAT CACTGGCATTGATTAC	207/	NO			
PvM078/PVEPSE2003D01	TCT6	CATTACTCCCTTCCCTTCCA ACTTCACTGGCACTCACAGG	174/174	1	NA		
PvM079/PVEPSE2003F03	GAA5	TCCGAAACTAACAGAACAAGGA TTATGGACAGTGGACGAACG	191/191	1	2	0,50	Pa
PvM080/PVEPSE2004B12	CAC5	ACTATCCCCAACCAACCACCA TCATAGCCTCCACCAGAAGG	204/204	4	1	Mono	Pa
PvM081/PVEPSE2004F01	ATG5	TCCCAGTGAAATCTGCTACAT CCAACAATAAAATGGCACAAA	176/176	4	2	0,08	Pa
PvM082/PVEPSE2004G10	TTA5	TTTCATCTTCAAACTCTCATCAA ATGGCAGACCGAACTCTCTA	183/183	4	2	0,16	Pa

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM083/PVEPSE2005C02	TTC5	AGAAATGAGTGAGCGTGTCG ACGGCTGATAGAGCGGGTA	152/152	1	1	Mono	Pa
PvM084/PVEPSE3013B07	TG5	AGAAGGTGAAGTATGATTTA TACATTGATTTTGGGTCA	244/244	1	1	Mono	
PvM085/PVEPSE3015B05	TA6	ATTCACGGTTAGGATAGAT TGGCTGTCTTGTTTCT	111/111	1	1	Mono	
PvM086/PVEPSE3018D04	AT5	ACCGGGAGAAAAGGAAGAAC GCCAAAAAGTAAGAAGCAGCA	210/250	1	2	0,09	Pa
PvM087/PVEPSE3021G12	TA5	CATCACGACAGTTTTCC CATTACATTACACCTATTCA	278/278	1	NA		
PvM088/PVEPSE3022A06	AG5	AGGTGCAAAAGCAAGAATGC TCCCCAGCAGTATGATGACA	208/208	1	1	Mono	Pa
PvM089/PVEPSE3023B04	TC5	CGAAGTAGCACGCAGAGAGA GCAAGACGGGAGTGATGG	185/290	1	2	0,08	Pa
PvM090/PVEPSE3023H10	TC6	CCCTTCACCGCTCCATACT CGGAAACAGTGGCTGGTCTA	250/250	1	2	0,08	Pa
PvM091/PVEPSE3025H02	AT7	CTCCATTTCCCCTTTG GTTACCCCAAGTTACAAGAA	208/	NO			
PvM092/PVEPSE3027F14	TA6	CGTGGGCTACACTGGAAGAC CATCTGCAAATGCTTGCTGT	240/240	1	3	0,35	Pa
PvM093/PVEPSE3029J06	GA5	TTATATGCGGAGCAAAGGATG GACCGACAAAGCCTTCATTC	216/216	1	2	0,38	Pa
PvM094/PVEPSE3030F08	CT6	AGTCACCGGATGCTCTGAAG GGGGTTTAGGGAAAGCAAAA	185/185	1	1	Mono	Pa
PvM095/TC17	AC9	GGGCGAGTTCATCTTCACAA AGGATCAACAGCCTCCAGTG	243/400	1	4	0,69	Pa
PvM096/TC18	CT5	GCCAGTTTGTTGCTTGGAAT GAAGAATGGTTGCTCCCTCA	190/190	1	1	Mono	Pa

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM111/TC427	TA5	TGGCTGGTAACAACAACTTCA TACGACATGGACCAAAACGA	200/200	1	1	Mono	
PvM112/TC471	AG5	GGTTCCTGCCCCAGAATC TTGCTGCTTCCGTTTCAGTA	251/251	1	1	Mono	Pa
PvM113/TC514	CT5	CTGGTGGAGTTGTGGCTGTA GGATTCTGCTCAAAGGGAGA	180/180	1	1	Mono	Pa
PvM114/TC529	CT5	TCTTCTTCTTCCAAGCCTTCC TATGGCAGCAAACGGGTAT	201/1000	1	1	Mono	Pa
PvM115/PVEPLE1003G09	CT25	AAATGTAAAGTGCTCCAT CTGAGAGAAAGAAAGAGACA	143/143	2	13	0,91	Pa
PvM116/PVEPSE2002D05	GA5	TATTGCCCAAAGATGCTTCC ATACCCAAAATGCACCCACT	179/179	1	1	Mono	Pa
PvM117/PVEPSE2004H03	TC7	CACACATCTCACCTTGTCTCTCTC GCGGCATTCAACACTGCT	165/700	1	1	Mono	Pa
PvM118/PVEPSE2006F05	TC8	GCGAACACATTACACAACA AATCCCGACCAAGTCTCCTT	215/215	1	5	0,54	Pa
PvM119/PVEPSE2006H01	TC9	AGAAGAAGAAGTAGAACAGAATC GCTGGAGTTGGGGTCAGT	226/1100	1	NA		
PvM120/PVEPSE2009C06	TC6	CCCCATCCTCACTCACAAAC TCATTCTTCTGGTGCTGCTTT	211/211	1	2	0,39	Pa
PvM121/PVEPSE2010B07	TC5	CTGCTTTTGAACCCTACTTCT TTGGGCGTTGTTTAGTGAAA	224/	NO			
PvM122/PVEPSE2010B07	CA5	CATTTGCTTGCTTATCTGC TTTGGATTTTCCCTCACA	201/	NO			
PvM123/PVEPSE2013B07	CT9	CCCACTACCACTCCTTTG AGCGTCTGAACCCATTG	194/194	2	4	0,47	Pa
PvM124/PVEPSE2013E06	TA5	TCCCCTACTGATGTGTTC AAATGTCTGGTTTTGC	193/193	2	2	0,39	

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM097/TC19	TA5	CAAGAGTGAAGGGGCAGTTT CGGCCAACCCTACTTTTAG	155/155	1	3	0,66	
PvM098/TC70	TC5	CTCCCTCCTTCACACTCTTTG GGGGTAGTCCTTTGGAACG	114/114	1	3	0,46	Pa
PvM099/TC70	TC5	CTCTTTGTGTGGCTCCCTGT GGTGGCATCTTTTCTACCTCA	203/2000	1	1	Mono	Pa
PvM100/TC147	AT7	GCCAATAACATCACCTCAT GGGTAGTCAGTTTCTAAGCA	290/290	1	4	0,36	Pa
PvM101/TC158	TC6	TCACATACAAGCACATCAACCA CCAGCAACTCCAGACACAGT	176/660	1	1	Mono	Pa
PvM102/TC161	CA5	CGGAAATAGAACTCAAACCTT AATGGAGAGAGAGCAAGG	259/259	1	1	Mono	Pa
PvM103/TC168	AG8	CGAGAAAGAGAGAGAAGAGTTT GTGTTGGTGTGATGCTGAG	145/145	1	3	0,58	Pa
PvM104/TC187	AC5	CCAATCGAAAGACCCAAATC GATCGCAGAGCCTTCAAATC	191/191	1	2	0,08	Pa
PvM105/TC188	AT5	CCGTATGAACATCAACCATT CACAAGAGGATTTCCAACCTCT	232/500	1	2	0,29	
PvM106/TC195	AT5	GCTGGTGTGGAGATGATACT TGTTGATGTTGAAAGGAGTG	219/219	1	1	Mono	Pa
PvM107/TC198	AT5	GGAGCCCCTCACCAATAGAC CGCCCTGGATCGTTAGATAG	347/347	1	1	Mono	Pa
PvM108/TC202	CT5	CATCTTGAGTTTATCCCTGCTC AATTCCAAGGCCATTGACAC	238/238	1	NA		
PvM109/TC306	TC5	CATTCTAAACAACAGACCCCAAT CGTGGCAGTTCCCTCTTT	116/	NO			
PvM110/TC325	CT5	CGCACACGCTCTCTCTCT CTTCTGCTGCTCCTCAATCTT	202/	2	NA		

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM125/PVEPSE2014B09	AC5	CAAACCTCAACACCCTCTCA GGAAACAGCAGCACCAC	198/198	1	1	Mono	
PvM126/PVEPSE2014F08	TC7	AAATCCTCTTCCACCTTTG AACACGCACACACAGACA	132/132	1	3	0,49	
PvM127/PVEPSE2015H07	TC5	AACTTTCTTTGACCCTCTC GCTTTGTCTTGTCTTCCA	155/155	2	6	0,72	Pa
PvM128/PVEPSE2016C01	GA11	GTTTCTGCGGTGAGAGTG AGCCTTCGTCGTTCTTCT	130/130	1	3	0,43	
PvM129/PVEPSE2017D09	GA5	GGTAGGACAACAAGGTAACG AAAACAACGCAGGAAATG	167/167	2	1	Mono	
PvM130/PVEPSE2018H02	TA5	TGGACAACAACACTACACA AAGACATTACATAATAAGG	265/265	2	2	0,18	
PvM131/PVEPSE2018H09	GA5	TGATTGTTCTTGCCTGGT TTCCATCCTCATCCTTG	142/	NO			
PvM132/PVEPSE2018H10	AG5	GCCGAAGCAGATAAAAGG TGTTCTCGTCTCCAATGC	191/191	2	4	0,51	Pa
PvM133/PVEPSE2019C09	AT5	GAAGTCTCTTGGGTTTCAGA CCCTATTTTTCACTCTCACC	193/193	2	1	Mono	
PvM134/PVEPSE2020C04	CA5	TCATACTACCTTGTTCTTACATTCA CACGCTGTTCCACCTG	182/182	2	1	Mono	Pa
PvM135/PVEPSE2005C06	TTG5	TCTCTCTGTTTCTTTTCTGGGTTT CGTTGTCTTTTTCTCCGCTTT	151/225	1	1	Mono	
PvM136/PVEPSE2006E08	TGG5	GCAACTATGCTGGGAAAGGA GGGCATTCTTCAGATTCTCG	217/217	1	2	0,08	Pa
PvM137/PVEPSE2007B06	TTC6	TTCTCCTGTCCCTCCTTGTTG GCCTTCCCACCAGGTTTAG	197/197	1	2	0,16	Pa
PvM138/PVEPSE2009H08	AGA5	TGATGAGGCTGAGGAGAAGG TCGAACACAATGTCCGAATC	216/	NO			

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(continuação)	
						PIC ⁴	Loco transf. ⁵
PvM139/PVEPSE2011H08	AGC5	AGAGGACGATGGATTGGT TCAGAACAAAGGGTTTTTCG	288/288	2	1	Mono	
PvM140/PVEPSE2013G02	GGT5	ATATTCTCGGCTCCGTTGGT TCCTCTTTCTTTGCCTCAGC	214/	NO			
PvM141/PVEPSE2013G03	CCA5	CTTCGTCATCTCCCACAA CCATCATCTCCTTTGGAA	214/214	2	1	Mono	
PvM142/PVEPSE2014H11	CAG5	CAGATAAATAGCCCTGTGC GCATTCAAACGATGACTCTT	188/188	1	2	0,16	
PvM143/PVEPSE2016B08	ATG5	TCTTCCACTCCTCCACCT CCTAAGCACCCCAAAAAT	183/170	1	1	Mono	
PvM144/PVEPSE2017C07	TTC5	AACCACGCTCTCTTCTCC ATCCCTGAAAAACCA	163/	NO			
PvM145/PVEPSE2019C10	TCC5	TTTCAGTTCGGGATTGTTCC ATTGGTGGAGGTGGGAGAG	203/203	1	3	0,16	Pa
PvM146/PVEPSE2020D07	CGC5	ACCATTGACTCCGATGT ACGATAAACCCAATCCACCA	187/400	1	1	Mono	Pa
PvM147/PVEPSE2020F11	TCA5	ATTTGCTGTTTTCTGCTG GAAGTTGAGGTGGGAGGT	151/151	1	1	Mono	
PvM148/PVEPSE2023B04	CCA7	ACCTCAAAACCCACCACAAA GAAGTGCTCCCAGATGAAGG	191/191	1	3	0,48	Pa
PvM149/PVEPSE2025B02	ACG5	ATGACGAGGAACCAGACACC GTCGCTTCAACGAGGATGTT	213/213	1	2	0,08	Pa
PvM150/PVEPSE2030D10	TCT5	CTCCCAAGTTCCCCTTTTTTC ATGCGTCCAACCCTTATGTC	199/199	1	2	0,43	Pa
PvM151/PVEPSE2031D01	TTA6	GGGCTGCAAAGGTGACATTA ACCGAAAACCCATGCTACTC	204/215	1	3	0,29	Pa
PvM152/PVEPSE2032F06	TTG6	ATTTTGGAGCGAAACAGCAT GAGAACCTCGTCGTCGTCTT	200/200	1	4	0,36	Pa

Tabela A1 - Análise dos 157 locos SSR-EST, mostrando o motivo, a sequência de *primers*, os tamanhos observado e esperado dos alelos, o programa de PCR usado, o número de alelos, o PIC e a espécie para a qual o marcador foi transferido

Código do loco/código da EST	Motivo	Sequência (5' - 3') dos <i>primers</i> <i>Forward/Reverse</i>	Tamanho esperado/ observado (pb)	Programa de PCR ^{1,2}	Número de alelos ³	(conclusão)	
						PIC ⁴	Loco transf. ⁵
PvM153/PVEPSE3002B08	AGA5	ATGGGCGATCTGGACTATGT TTGAAGACCAAGCCAAGAGAA	183/270	1	2	0,41	Pa
PvM154/PVEPSE3021F08	TCA5	ATTCACCTCCTTCTGCTTCG GCTCCTTCTCTCCAGTCTCG	216/490	1	1	Mono	Pa
PvM155/PVEPSE3029G07	AAC5	AAGTATGGAAACTGGGGCTGT GTGCGTGGAATAATGTGGTG	206/206	2	2	0,08	Pa
PvM156/PVEPSE2021C05	TTC6	CACACTTCAACTCCAAAGG CCAACCCTCGCAAAAT	219/1100	2	NA		

¹ Os programas de PCR são apresentados na Tabela 2; ² NO: PCR não otimizada; ³ NA: Loco não analisado no conjunto de genótipos; ⁴ Mono: Loco monomórfico no conjunto de genótipos; ⁵ Loco transf.: Loco transferido para uma ou mais espécies: V: *Vigna* sp.; Pl: *Phaseolus lunatus*; Pa: *Phaseolus acutifolius*.

ANEXO B

Artigo científico publicado no periódico *Genome* (Ottawa):

HANAI, L. R.; CAMPOS, T.; CAMARGO, L. E. A.; BENCHIMOL, L. L.; DE SOUZA, A. P.; MELOTTO, M.; CARBONELL, S. A. M.; CHIORATTO, A. F.; CONSOLI, L.; FORMIGHIERI, E. F.; SIQUEIRA, M.; TSAI, S. M.; VIEIRA, M. L. C. Development, characterization, and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources. **Genome**, Ottawa, v. 50, p. 266-277, 2007.

Development, characterization, and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources

Luiz Ricardo Hanai, Tatiana de Campos, Luis Eduardo Aranha Camargo, Luciana Lasry Benchimol, Anete Pereira de Souza, Maeli Melotto, Sérgio Augusto Moraes Carbonell, Alisson Fernando Chioratto, Luciano Consoli, Eduardo Fernando Formighieri, Marcos Vinícius Bohrer Monteiro Siqueira, Siu Mui Tsai, and Maria Lucia Carneiro Vieira

Abstract: Microsatellites or SSRs (single sequence repeats) have been used to construct and integrate genetic maps in crop species, including *Phaseolus vulgaris*. In the present study, 3 cDNA libraries generated by the Bean EST project (<http://lgm.esalq.usp.br/BEST/>), comprising a unigene collection of 3126 sequences and a genomic microsatellite-enriched library, were analyzed for the presence of SSRs. A total of 219 expressed sequence tags (ESTs) were found to carry 240 SSRs (named EST-SSR), whereas 714 genomic sequences contained 471 SSRs (named genomic-SSR). A subset of 80 SSRs, 40 EST-SSRs, and 40 genomic-SSRs were evaluated for molecular polymorphism in 23 genotypes of cultivated beans from the Mesoamerican and Andean genetic pools, including Brazilian cultivars and 2 related species. Of the common bean genotypes, 31 EST-SSR loci were polymorphic, yielding 2–12 alleles as compared with 26 polymorphic genomic-SSRs, accounting for 2–7 alleles. Cluster analysis from data using both genic and genomic-SSR revealed a clear separation between Andean and Mesoamerican beans. The usefulness of these loci for distinguishing bean genotypes and genetic mapping is discussed.

Key words: *Phaseolus vulgaris*, microsatellites, genetic markers, expressed sequence tags (EST).

Résumé : Des microsatellites ou SSRs (« simple sequence repeats ») ont été employés pour produire et intégrer des cartes génétiques chez plusieurs espèces cultivées, incluant le *Phaseolus vulgaris*. Dans le présent travail, trois banques d'ADNc générées par le « Bean EST project » (<http://lgm.esalq.usp.br/BEST/>) totalisant 3 126 unigènes et une banque génomique enrichie en microsatellites ont été analysées pour la présence de microsatellites. Au total, 219 ESTs comprenaient 240 SSRs (EST-SSR) tandis que 714 séquences génomiques comptaient 471 SSRs (SSR génomiques). Un sous-ensemble de 80 SSRs, 40 EST-SSRs et 40 SSRs génomiques, ont été évalués pour leur polymorphisme chez 23 génotypes de haricots cultivés provenant des ressources génétiques mésoaméricaines et andéennes, incluant des cultivars brésiliens, et deux espèces apparentées. Chez les génotypes du haricot, 31 EST-SSRs étaient polymorphes et révélaient entre 2 et 12 allèles tandis que 26 des SSR génomiques étaient polymorphes avec entre 2 et 7 allèles. Une analyse de groupement employant les données des deux types de SSR a montré une séparation claire des génotypes andéens et mésoaméricains. L'utilité de ces loci pour la discrimination de ces génotypes et pour la cartographie génétique est discutée.

Mots-clés : *Phaseolus vulgaris*, microsatellites, marqueurs génétiques, étiquettes de gènes exprimés (EST).

Introduction

Microsatellites or single sequence repeats (SSRs) are DNA regions composed of small motifs of 1 to 6 nucleotides repeated in tandem, which are widespread in both eukaryotic and prokaryotic genomes (Field and Wills 1996;

Tóth et al. 2000). Microsatellite markers were developed for a wide range of plant species, including many crops. The advantage of microsatellites as molecular markers is that they detect polymorphisms at highly variable loci (Morgante and Olivieri 1993; Powell et al. 1996; Hedrick 2001). Furthermore, they can be amplified by PCR, are codominant,

Received 9 August 2006. Accepted 6 December 2006. Published on the NRC Research Press Web site at genome.nrc.ca on 9 May 2007.

Corresponding Editor: F. Belzile.

L.R. Hanai, L.E.A. Camargo, M. Melotto, L. Consoli, M.V.B.M. Siqueira, and M.L.C. Vieira.¹ Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz", Avenida Pádua Dias, 11, 13418-900, Piracicaba, Brazil.

T. de Campos, A.P. de Souza, and E.F. Formighieri. Universidade Estadual de Campinas, Instituto de Biologia, 13083-970, Campinas, Brazil.

L.L. Benchimol, S.A.M. Carbonell, and A.F. Chioratto. Instituto Agronômico de Campinas, Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais, Campinas, 13001-970, Brazil.

S.M. Tsai. Universidade de São Paulo, Centro de Energia Nuclear na Agricultura, Piracicaba, Brazil.

¹Corresponding author (e-mail: mlcvieir@esalq.usp.br).

and are distributed throughout the genome. Together, these characteristics make microsatellite loci one of the best genetic markers for mapping purposes (Oliveira et al. 2006).

Microsatellites were developed from both coding and non-coding regions of plant genomes (Brown et al. 1996; Scott 2001). Several resources can be used to search for SSR, including a variety of DNA libraries such as genomic, genomic-enriched for SSR, bacterial artificial chromosome and cDNA libraries, as well as public databases, including expressed sequence tags (ESTs) databases (Scott 2001; Broughton et al. 2003). EST-SSRs have some intrinsic advantages over anonymous markers because they can be easily found by in-silico analysis of EST databases and are associated with a coding gene. In addition, as primers are designed from highly conserved coding regions, i.e., from genes, across-species amplification is expected to be more successful than that obtained from genomic DNA (Varshney et al. 2005). Nevertheless, EST-SSRs are less polymorphic than genomic-SSR (Eujayl et al. 2001).

Beans are the most important grain legumes for human consumption in the developing world. Molecular markers have assisted common bean breeding programs in various ways (Jarne and Lagoda 1996): to investigate the origin and diversity of current cultivars (Gepts 1998), the domestication of *Phaseolus vulgaris* (Koinange et al. 1996), and the genetic control of resistance to important diseases (Nodari et al. 1993; Yu et al. 1998; Miklas et al. 2003). Several *Phaseolus* linkage maps have been developed in recent years (Freyre et al. 1998; Yu et al. 2000), and SSRs have been used mainly for constructing and integrating them (Yu et al. 2000; Tarlan et al. 2001; Blair et al. 2003).

In this study, we explored the Bean EST (BEST) database (Bean EST project; <http://lgm.esalq.usp.br/BEST/>) and a genomic library highly enriched for microsatellites screened for CT and GT sequences (Billotte et al. 1999), used to develop, characterize, and compare the SSR loci generated from both sources. Furthermore, we assessed 80 SSR markers to evaluate the diversity of a collection of *P. vulgaris* accessions, including South American genotypes.

Materials and methods

Plant material and DNA extraction

A set of 23 common bean inbred lines and 2 other related species were used to evaluate allelic variations at 80 SSR loci identified in both kinds of sequences, i.e., genomic and genic: 17 lines of Mesoamerican origin ('Brasil 2', 'Porcillo', 'Rio Tibagi', 'Mulatinho', 'Carioca Comum', 'Puebla 152', 'Great Northern', 'Flor de Mayo', 'Garbancillo', 'Apetito Blanco', 'Jamapa', 'Barbunya', 'IAC-UNA', 'Baetão', 'Sanilac', 'Tu', and 'Bat 93'), 6 lines of South American origin ('Pompadour', 'Antioquia 8', 'Jalo EEP 558', 'Jabola', 'CAL 143', and 'Red Kidney'), and 2 related species, *P. lunatus* ('87 JP 12') and *Vigna* sp. ('Fradinho Cruzeiro'). This set comprises a representative sample of cultivated dry beans and includes the parental lines of 3 mapping populations: 'Bat 93' × 'Jalo EEP558' (Nodari et al. 1993), 'Carioca' × 'Flor de Mayo' (Melo et al. 2002), and 'CAL 143' × 'IAC UNA' (Instituto Agronômico, Campinas, Brazil).

Total genomic DNA was extracted from leaf tissues using

the cetyltrimethylammonium bromide extraction method as described by Doyle and Doyle (1987). DNA was resuspended in Tris-EDTA buffer (10 mmol/L Tris-HCl, pH 8.0; 1 mmol/L EDTA, pH 8.0), and DNA concentrations were estimated by electrophoresis on ethidium-bromide-stained agarose gels using appropriate molecular mass standards.

EST-SSR and genomic-SSR analyses, and primer designing

A unigene collection of 3126 sequences of *P. vulgaris* EST isolated from seedlings inoculated and not inoculated with the fungal pathogen *Colletotrichum lindemuthianum* was used as a source of EST-SSR (Melotto et al. 2005). We searched the collection for the presence of perfect SSR using SSRIT software (Temnykh et al. 2001). The minimum length, i.e., 5 repeat units for di-, tri-, and tetranucleotides, was used as a criterion for SSR mining. SSR motifs were found in 219 unigene sequences. We grouped the dinucleotide SSRs into 4 motifs (AT, AG, AC, and CG) and the trinucleotides into 10 motifs (GAA, AAT, CAT, AAC, TGG, AGG, TAG, TGC, CGG, and ACG). The tetranucleotides were as few as 7 and were grouped in 5 motifs (AAAG, AATC, AATG, AATT, and TGCA). Each motif included all the complementary and permuted sequences for a given simple sequence. These sequences were further classified according to the position of the SSRs within the gene, that is, in the 3' untranslated region (UTR), 5'UTR, or coding sequence (CDS). This analysis was performed by searching the putative start and stop codons in each EST after determining the frame of the transcribed EST. For this purpose, the BLASTx algorithm was used to obtain the best amino acid alignment of the EST with respect to similar sequences deposited in the nonredundant GenBank database.

For primer design, only the EST-SSRs that contained enough nonrepetitive flanking regions was used. PCR primers between 16 and 24 bp in length were designed to flank the SSRs using Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). To standardize the amplification conditions and to decrease the probability of generating spurious PCR fragments, 2 criteria were used to design the primers: a melting temperature ranging from 57 to 63 °C and a maximum difference of 3 °C between the estimated optimal melting temperature for the forward and reverse primer annealing.

An enriched-microsatellite library was screened for 2 repeated sequences, CT and GT. It was constructed using the Brazilian variety 'IAC-UNA', following the protocol of Billotte et al. (1999). Sequencing was performed with T₇ or SP₆ primer sites (all clones were sequenced in both directions) using the Bigdye Terminator Cycle Sequencing Kit (Applied Biosystems, ABI 377 sequencer). Readings were processed by the Phred v. 0.000925.c base-calling program (<http://www.phrap.org>; Ewing et al. 1998), and poly-A/T tails, vector, and adapter sequences were trimmed using cross-match analyses. Clustering was performed using CAP3 software with default parameters (Huang and Madan 1999). Primer designing for genomic-SSR was carried out by means of PrimerSelect (DNASar Inc.) according to the following conditions: expected size of the amplified product ranging from 150 to 350 pb; GC content between 40% and

60%; melting temperature (T_m) from 45 to 60 °C; primer length ranging from 18 to 22 bp; and no hairpin or dimer formation.

EST-SSR and genomic-SSR assays

The PCR for EST-SSR assays was performed in a final volume of 20 µL in 96-well plates using a Gen Amp[®] PCR System 9700 thermocycler (Applied Biosystems Inc.). Approximately 20 ng of template DNA was mixed in a solution containing 0.25 µmol/L of both forward and reverse primer, 200 µmol/L of each deoxynucleoside triphosphate (dNTP), 2.5 mmol/L $MgCl_2$, 10 mmol/L Tris-HCl, 50 mmol/L KCl, and 0.5 U Taq DNA polymerase (Invitrogen). Amplification profiles are shown in Table 1.

For the genomic SSR assay, the annealing temperature (T_a) of each microsatellite sequence was identified using a temperature gradient on a PTC-200 thermocycler (MJ Research Inc.). PCRs were performed in a 25 µL final volume containing 50 ng of template DNA, 0.2 µmol/L of each forward and reverse primer, 100 µmol/L of each dNTP, 2.0 mmol/L $MgCl_2$, 10 mmol/L Tris-HCl, 50 mmol/L KCl, and 0.5 U Taq DNA polymerase (Invitrogen). For the reactions the following conditions were used: 1 min at 94 °C; then, 30 cycles of 1 min at 94 °C, 1 min at T_a , and 1 min at 72 °C, ending with a final extension for 5 min at 72 °C.

Five microlitres of the amplification products, previously denatured with 1× loading buffer (0.2% each bromophenol blue and xylene cyanol, 10 mmol/L EDTA, pH 8.0, and 95% formamide) at 95 °C for 5 min, was loaded on 6% w/v denatured polyacrylamide gels using a 10 bp ladder as a size standard. Silver staining was performed according to Creste et al. (2001).

Data analysis

The polymorphic information content (PIC) was used to measure the informativeness of the 80 SSR loci, according to the equation of Anderson et al. (1993)

$$PIC = 1 - \sum_{i=1}^k P_i^2$$

where k is the total number of alleles detected for a SSR and P_i the frequency of the i th allele detected in the 25 bean genotypes.

SSR data were scored visually, and genetic distances among the genotypes were estimated separately for EST-SSRs and genomic-SSRs by using Wright's distance (Wright 1978) modified from Roger's distance. The genetic distance matrix was estimated using TFPGA v. 1.3 (Miller 1997). Cluster analyses were performed according to the UPGMA method (unweighted pair-group method with arithmetical averages). Trees were drawn using Drawgram with PHYLIP computer package v. 3.6 (Felsenstein 2004). The stability of the clusters was tested by a resampling procedure with 10 000 bootstraps using the BooD program (Coelho 2002).

Results

Frequency and distribution of EST-SSR

Previous studies published by our group (Melotto et al. 2005) identified a set of 240 potential SSR markers in 219

Table 1. Number of cycles and annealing temperature (T_a) of PCR programs used to amplify common bean EST-SSR loci.

	PCR program		
	1	2	3
Stage 1			
Cycles	12	10	10
T_a (°C)	65–53 ^a	55–50 ^b	55–45 ^a
Stage 2			
Cycles	25	25	25
T_a (°C)	53	50	45

Note: An initial denaturing step was conducted at 94 °C for 2 min before stage 1 in all 3 programs. At both stages 1 and 2, denaturing and elongation conditions were held constant at 94 °C for 30 s and 72 °C for 1 min, respectively. The annealing duration in all programs was 45 s and 1 min at stages 1 and 2, respectively.

^aAnnealing temperature was decreased by 1 °C after each cycle.

^bAnnealing temperature was decreased by 0.5 °C after each cycle.

unigene sequences: 137 (57%) were composed of dinucleotide motifs, 96 (40%) of trinucleotide motifs, and 7 (3%) of tetranucleotide motifs. The average length of the repeats was 12.6 bp for dinucleotides, 16.6 bp for trinucleotides, and 21.7 bp for tetranucleotides. The dinucleotide motifs were the most abundant: AG with 80 occurrences (33.3%) followed by AT with 37 occurrences (15.4%). The CG motif never occurred in the sequences (Table 2). The most frequent trinucleotide motif was GAA, with 34 occurrences (14.2%), followed by CAT and TGG with 18 (7.5%) and 10 occurrences (4.2%), respectively.

EST-SSR loci were mostly found in UTR (33% and 23% at 5' and 3' ends, respectively). Dinucleotides were found predominantly in UTR (73%), and trinucleotides were preferentially found in coding regions (71%) (Fig. 1).

Here, we designed primers for 156 SSRs. The remaining 86 either had an unsuitable GC content or were not flanked by nonrepeated sequences because they were located at the EST ends. PCR assays were conducted for a subset of 46 loci. Four primer pairs did not yield a PCR product, and 2 primer pairs produced large fragments, probably because of the presence of introns in the amplicon causing problems, as large fragments cannot be accurately scored for small size differences. The remaining 40 EST-SSR markers, which produced fragments of suitable size, were used for further analyses.

A total of 2479 clones were screened from the 'IAC-UNA' library, and 1453 (58.6%) putative positive colonies were isolated, of which 1209 (48.8%) were sequenced. A total of 714 sequences contained microsatellites. The screening of the library showed that 471 sequences contained simple motifs (perfect and imperfect) and 243 composite ones. From Phred/Cap3 analyses, 478 contigs and 648 singlets were obtained. Around 540 sequences were suitable for designing primers as a result of their sequence flanking repeats or complexity. SSR containing GT/CA (38%) or

Table 2. Frequency of individual SSR motifs among 240 microsatellites isolated from the BEST database.

SSR type	Motif	No. of repeats							Total/motif	
		5	6	7	8	9	10–15	16–25	No.	(%)
Dinucleotides	AG	42	15	6	4	5	6	2	80	(33.3)
	AT	20	8	3		1	4	1	37	(15.3)
	AC	17	1	1		1			20	(8.3)
	CG								0	(0.0)
Trinucleotides	GAA	18	11	4			1		34	(14.2)
	CAT	11	6	1					18	(7.5)
	TGG	6	3	1					10	(4.2)
	AAC	6	2						8	(3.3)
	TGC	6	1						7	(2.9)
	AAT	4	1	1					6	(2.5)
	CGG	4							4	(1.7)
	TAG	2	1						3	(1.3)
	AGG	2		1					3	(1.3)
	ACG	3							3	(1.3)
	AAAG	1		1					2	(0.8)
	AATC	1							1	(0.4)
	AATG	1	1						2	(0.8)
Tetranucleotides	AATT	1							1	(0.4)
	TGCA	1							1	(0.4)
Total		146	50	19	4	7	11	3	240	(100)

Fig. 1. Distribution of bean EST-SSR loci according to their gene position.

GA/CT (22.7%) motifs were the most frequent among simple repeats.

Polymorphism of SSR loci in *P. vulgaris* genotypes

A set of 80 SSR loci, 40 EST-SSRs (Table 3), and 40 genomic-SSRs (Table 4, Genbank accession numbers EF136464 to EF136503) was used to study the allelic diversity in the selected common bean genotypes. Out of those 40 EST-SSR markers, 11 had dinucleotide repeats, 23 had trinucleotide, and 6 had tetranucleotide repeats. A total of 31 EST-SSRs were polymorphic among the 25 common bean genotypes (Table 3). The average number of alleles was 2.7 for the 40 EST-derived loci but 3.1 for the polymorphic ones. Of all the EST-SSR markers, the locus PvM21 revealed the most alleles, i.e., 12 (Fig. 2). For the genomic SSRs, 15 loci had simple dinucleotide repeats and 25 were composite ones. A total of 26 loci were polymorphic among the genomic-SSRs. The greatest number of alleles per locus (7) was detected at the locus FJ55, which amplified 3 dis-

tinct microsatellites: 1 composite between 2 simple ones ((TC)₅–(TC)₁₀(CT)₆–(CT)₁₅). The average number of alleles was 2.4 for the 40 genomic-SSR-derived markers (2 to 7) but 3.2 when only polymorphic markers were considered.

To measure the informativeness of SSR markers, PIC was calculated for each of the 80 loci using the 25 genotypes. PIC values ranged from 0.08 to 0.86, with an average of 0.44 for EST-SSR and from 0.08 to 0.81 with an average of 0.45 for genomic-SSR. Data from genic and genomic SSRs were used for phenetic analysis among the 25 bean lines. The dendrogram (Fig. 3A) obtained from data on 31 polymorphic EST-SSRs first clustered a Mesoamerican *Vigna* species and 18 bean lines, 17 of them belonging to the gene pool of Mesoamerican origin and 'Fradinho Cruzeiro'. The second cluster in the dendrogram comprised 6 lines of Andean common beans and the cultivar '87 JP 12' of *P. lunatus*. Cluster analyses from data on the 26 genomic SSRs and genic SSRs also revealed a clear difference between Andean and Mesoamerican beans (Fig. 3).

Discussion

SSR in expressed sequences

Microsatellites are usually considered as evolutionary neutral DNA markers and selectively neutral sequences randomly or nearly randomly distributed over the euchromatic genome (Schlotterer and Wiehe 1999). However, reports have repeatedly shown that SSRs also exist within genes or functional sequences (Li et al. 2004). Yaish and Pérez de la Vega (2003) isolated a GA SSR that was immediately upstream of the methionine codon of the MADS-box in common bean. Hu et al. (2004) found 394 SSRs in more than 4000 sweet potato ESTs, and recently Peng and Lapitan (2005) identified more than 40 000 SSRs in nearly half a million wheat ESTs. Melotto et al. (2005) identified 240

Table 3. PvM loci derived from common bean expressed sequence tags (ESTs).

SSR code	Bean unigene	Forward and reverse primer sequences (5'→3')	Repeat motif	Gene location	Predicted/ observed length (bp) ^a	PCR program	PIC	Na
PvM01	TC34	GCTCTAGCCTCAATCCTTGG GGTTCCTGCTTCCCACTTT	(TTC) ₇	5'UTR	240/240	1	0.60	4
PvM02	TC116	CGCCATTGGATTGGATT AGCGTGGAAGTGGAGTG	(CTT) ₆	CDS	191/191	1	0.73	6
PvM03	TC116	CGCCCTTCTTCTTCTCTTC CGGCGAGTCATCTTTTC	(TTC) ₆	CDS	167/167	1	0.62	4
PvM04	TC323	GGTTCCTCTCTCTCTCTGCT GCGCCGTCTTTTGGTAGT	(TTC) ₁₀	5'UTR	210/210	2	0.65	3
PvM06	TC103	TTGCCAACAAATGGAAGTGA TTGCAACAATGAGAAGGGAGA	(TAC) ₆	3'UTR	190/170	2	0.15	3
PvM07	TC127	CTGGAATGGCTTACAGATGC CCAAGATCAGGCACAGTGAA	(ATG) ₆	CDS	213/213	2	0.46	2
PvM09	PVEPSE3030D08	GCTTGACGCCCTTTGAACC CGTGGCATTGATTCCTC	(AAG) ₅	CDS	216/216	1	mono	1
PvM11	PVEPSE3030J03	TGGGATTTGAGCGGTGTG TCCACCGACTTACCGAACA	(AGA) ₆	n.d.	154/154	2	0.46	3
PvM12	PVEPSE3030L20	CCAGTGTGTCAACAAATGGA AGGCTTAGGGAATGGCTCA	(TGC) ₅	CDS	191/191	2	0.38	3
PvM13a	PVEPSE3030M09	GAGAAAGCCGACAGAGGA AGATGCCGGAACAGAAC	(GAA) ₆	CDS	216/216	2	0.71	4
PvM13b	PVEPSE3030M09	GAGAAAGCCGACAGAGGA AGATGCCGGAACAGAAC	(GAA) ₆	CDS	216/240	2	0.73	4
PvM14	PVEPLE1004G06	AAAGCCACAAAACATAAACCC GCCTCCGAAAGATTCAACG	(AATC) ₅	n.d.	204/204	2	0.61	3
PvM15	PVEPSE2010E01	GGTGGAGCACTGGTGTT TTCACCTTCCCAAGACCTGA	(TGA) ₅	n.d.	185/185	2	0.46	2
PvM16	PVEPSE2022E07	ATATGCCACCCACTCCA AACCAGAAAGGCAAAACG	(TTTC) ₇	CDS	179/179	2	0.08	2
PvM17	PVEPSE2030E02	AGGATGGGTCCGTGCTT GGCAGCTTCCACATCGTC	(ATGA) ₅	CDS	213/213	2	0.40	2
PvM18	PVEPSE3003E05	CAACCATCTCTGTGGCAAG TGAGTTTGGGAGAGTAGTTT	(AATT) ₅	3'UTR	202/102	2	0.58	3
PvM20	PVEPSE1004B07	CGCTTTGATAAAACGCTTC TGTTGGTGCTAACGATGGAA	(CT) ₇	3'UTR	222/340	2	mono	1
PvM21	TC449	CCCATCTTCCATTTCACAG AAGCCACAGAAAACCTGAGCA	(AT) ₁₄	CDS	251/251	2	0.86	12
PvM22	PVEPSE2004H10	ACTCTCAATGGCGGATC GGCGTTTCTCCCTCTCTT	(TC) ₅	CDS	222/222	2	0.40	3
PvM24	PVEPSE2009B07	CAGAAAAATGGGCACAAGG AAAGTCAACACGCACACACC	(TG) ₅	3'UTR	345/ 345	2	mono	1
PvM25	TC54	GAAGACGTGAAGGAGCCACT	(TGA) ₇	3'UTR	183/183	2	0.08	2

Table 3 (concluded).

SSR code	Bean unigene	Forward and reverse primer sequences (5'→3')	Repeat motif	Gene location	Predicted/ observed length (bp) ^a	PCR program	PIC	Na ^a
PvM26	TC80	CAC TTTCAGAGAGAACGCAAT TTCTCCACCCCAATTAAC TGC	(TCA) ₇	CDS	206/206	2	0.08	2
PvM27	TC159	CTCGGAACAAAAGTCTCCAC GGAGTTGAAAAGGTGGCTA	(TGC) ₅	CDS	192/192	2	mono	1
PvM28	TC390	TGATACAGAAAGCAGGGAGAACAA GGAGAGGAAACACGTAATGGA	(TCA) ₅	n.d.	189/189	2	0.37	2
PvM29	TC487	TCGACAAACAAAAGAAAGAAAGAA CAGCAGAACAGCAAGTGATTC	(AAG) ₅	5'UTR	275/275	2	0.08	2
PvM30	TC527	GGTGGCAGGTACGAAAGAGT GCAGACCTGTCAACAACAA	(GTT) ₅	3'UTR	222/222	2	0.55	5
PvM32	PVEPSE3030F06	GCCATGAAAGCAAGGTGAAGA GAGAAGCCAGAACTTA	(ATC) ₅	n.d.	159/159	3	mono	1
PvM33	PVEPSE3029G08	GCATCCCTTTTGTGTC GCCTTCTTTTCCAGCAG	(TTG) ₅	CDS	220/220	3	0.08	2
PvM34	PVEPSE3027C10	GTCATATGTTGTTTCTTTCTC CTTCTCCCTCTCTCTCT	(GAG) ₅	n.d.	273/273	3	0.58	3
PvM36	PVEPSE3011F10	GGACAGTGATAGTTCAGATTC GCCCTATTTCCCTACC	(ATC) ₅	CDS	201/201	3	0.47	2
PvM37	PVEPSE3011A04	CATCTCTCTCTCACCCCTTA GGACAGGGAATCTCACTA	(GAG) ₇	CDS	158/158	3	0.37	2
PvM39	PVEPSE3006G01	CGTGGTCGTGGTGGTG ATCTCTTCGGGCAGA	(GAT) ₆	CDS	133/133	3	mono	1
PvM40	PVEPSE3001F10	ATTGTGCTTCTCAACG CTCTCCATTTGTCTAG	(CTG) ₆	CDS	202/202	3	0.55	3
PvM41	PVEPSE2029D10	GAACCATCAGAAAACCTAC AAATGTCCCAGCCACAC	(CT) ₅	5'UTR	174/300	3	mono	1
PvM42	PVEPSE2027C10	TTCCTTCACCTTACCCTTC CCATCGTCTTTTCTTCAACA	(AG) ₅	3'UTR	255/255	3	0.08	2
PvM44	PVEPSE2024H04	CACGTCTCTGGGGCTATC CTCTACCAGCAGCATCA	(TC) ₈	5'UTR	182/600	3	mono	1
PvM45	PVEPSE2024D09	CTCTCTCTCTCTTCTCT GAGGGTAGTTCTGGA	(CT) ₂₄	5'UTR	232/232	3	0.66	4
PvM46	PVEPSE2024B07	GACCAAGTCTCTTC CGCTATGTTGGTGGAG	(GA) ₅	CDS	246/246	3	0.47	2
PvM48	PVEPSE2021H06	CGTGAAGAGAGAGGGGAGA GCAAAATGGTGGGTGATG	(CT) ₆	n.d.	122/122	3	mono	1
PvM49	PVEPSE2027H04	AGGGAGGTGAGGTGGTGA ATCTCTGCGTCGAGTGTTT TATCCTTCGCCCAATCCATA	(TC) ₆	5'UTR	197/197	3	0.27	2

Note: Na, number of alleles amplified using 25 bean genotypes; n.d., not determined, because either no hit or low hit in BLASTx.

^aAt least 1 genotype amplified the observed size.

Table 4. FJ loci derived from the enriched-microsatellite library.

SSR code	Forward and reverse primer sequence (5'→3')	Repeat motif	Predicted length (bp)	Range (bp)	Ta (°C)	PIC	Na
FJ 21/EF136464	TAATACTTGTATCTCCCTTGTG GGCCATGGAAAACTTGTA	(GT) ₄ G (GT) ₃	293	310	45.0	mono	1
FJ22/EF136465	CACCGCTTCTACCTCCGTCC GGGGCCGCATAGTGATTCTTCT	(CT) ₃ (CA) ₂	276	293	64.8	mono	1
FJ 23/EF136466	TAAATAGCCCTAAAACAAATCC ATATAGTGGCTTCCTCAAG	(AG) ₃ (TA) ₂	200	200	60.0	mono	1
FJ 24/EF136467	TGCAATCCAAACGAGTGAAG AAATACAGTTGAAATCCTCCAC	(GT) ₃ AT (GT) ₃	290	310	45.0	mono	1
FJ 25/EF136468	AAGCTTTTGGGTTTCTCA TTATCCCAAGGGCTACATTC	(AC) ₇	291	310	45.0	mono	1
FJ 26/EF136469	ATACCCAATTTAATGTTCTTTC GAGGTATACGTAGTTTGGAGTT	(CT) ₆	168	190–193	60.0	0.08	2
FJ 27/EF136470	ATGAAAGCGATAGTGATGAGAT CCAGTGGGTAATGAGTTG	(GA) ₁₇	246	230–270	60.0	0.66	4
FJ 28/EF136471	GCGCATTATCTCATCCATTAT CCAGGCAGAGCAAGCAAGTT	(AC) ₁₁	192	199–207	60.0	0.53	3
FJ 29/EF136472	CATGTTGGCTTAGTGATGTTT CATGTTGGCTTAGTGATGTTT	(TA) ₅ (TG) ₈	237	240–269	60.0	0.80	6
FJ 30/EF136473	TTCTCTTCTTCTTCCTTCCTC TTTTCTCTTCCACCTTCTCTA	(GT) ₆ ; (AG) ₅	251	255	60.0	mono	1
FJ 31/EF136474	TTTTTAACATGCTCCCTCTAC GGTCCACAATCAAGCAGTCAAG	(AC) ₉ (AT) ₄	289	308–312	60.0	0.21	2
FJ 32/EF136475	AGGGCTAGGAAACACCGAAAAC CCCGCCCCATCAGAAAAAG	(TG) ₆	159	159	60.0	mono	1
FJ 33/EF136476	GTGTGATGAGCCAAAGGAAGT GAAACAAAGCTCAAAACAGTAA	(GT) ₆	268	290	60.0	mono	1
FJ 34/EF136477	CTTCTGGGCTTGTTTCA CTATATCAGCCTAGTTATTTTCA	(GT) ₆	271	287–290	60.0	0.44	2
FJ 35/EF136478	TAATCTTAAACCAAAAACCACA GGGGATCGAATGAAAAGACT	(CA) ₁₀ (TA) ₁₁	264	300–340	60.0	0.44	2
FJ 36/EF136479	TGAATTTTAAAGTGCAAGAT CCAAACACAATAGAAATGATAG	(GT) ₅	178	188	45.0	0	1
FJ 37/EF136480	TTATCGCTAGATTCTTGTGTTA GGATGCATCTTTAGGAGTGA	(TA) ₂ (GT) ₆	281	140–152	60.0	0.73	4
FJ 38/EF136481	TTCTCTTCTTCTTCCTTCCTC TTTTCTCTTCCACCTTCTCTA	(GT) ₆	251	260	60.0	mono	1
FJ 39/EF136482	TGCTTACGCTCCAGTCATT TAGCAAAACACAGAAACGAAACC	(CA) ₃ (CT) ₂	280	262	60.0	mono	1
FJ 40/EF136483	TCAACATTCTTGGTGTGGTT AATGTTTATATTGGTTCACTCA	(GT) ₅	288	260	45.0	mono	1
FJ 41/EF136484	CATTCTTAGTGGTGTGTTGTC AATCGGTTGTTTCGCATAA	(CT) ₂ (GT) ₃ ; (AAG) ₂ (GT) ₄	268	278–280	60.0	0.21	2
FJ 42/EF136485	GGAGCGTAAAACCACCTT TCGAAAGCTTGAAGACCTTG	(AC) ₅ ; (AC) ₇	202	215–231	60.0	0.15	3
FJ 43/EF136486	GTGTGATGAGCCAAAGGAAGT GAAACAAAGCTCAAAACAGTAA	(TG) ₅	268	300	52.7	mono	1
FJ 44/EF136487	TATCAGCCTAGTTATTTTCAAG CATACTTTTCTTATTTTCTGGA	(CA) ₆	255	260–263	53.7	0.37	2
FJ 45/EF136488	TTGATCATGAGTTCAGAGTCCA GATTGATAGAGATTGCCATAAGTC	(CA) ₆	226	239–240	60.0	0.21	2
FJ 46/EF136489	AGTGGACCACTTCTTCTTC CAATTTACCAGCTGGACTTTAG	(CT) ₉	263	290–310	60.0	0.29	4
FJ 47/EF136490	ATGCTGGAATGACTGTGATGGA AATGGGTGAAATAAGTGGGTTG	(TC) ₂₀	248	278–312	60.0	0.76	6
FJ 48/EF136491	ATTGAAGTGAGTTATTTGTGAAG	(TG) ₅	229	250–260	60.0	0.37	2

Table 4 (concluded).

SSR code	Forward and reverse primer sequence (5'→3')	Repeat motif	Predicted length (bp)	Range (bp)	Ta (°C)	PIC	Na
FJ 49/EF136492	CATTCTTTATTGGTTTTCTTAG ATGAGAAAATAACGTGGAGGAA GACGGGATCGGTTTTACTGAG	(AG) ₁₀	264	275–280	60.0	0.40	2
FJ 50/EF136493	CATTTCATGTGATTTCTCTGT ATTGTGCTCTCCTTTGATAA	(CA) ₅	226	228–230	45.0	0.40	2
FJ 51/EF136494	AAGATTTTCCAGCGACCTC TTCACGGATCCAAAGTATTCTC	(CA) ₅	212	213–238	60.0	0.74	6
FJ 52/EF136495	GCTCGTGAATTCAAGGGTAA CCGAATCGGTAATAGTAAATCC	(CA) ₅	158	158–162	60.0	0.08	2
FJ 53/EF136496	CCGCGGACTACTGTAATCTATG GCAGCAAAAACAAAATCAAATC	(CT) ₁₀	174	195–218	45.0	0.65	5
FJ 54/EF136497	AGCATGCCCTGTTCTTGCTA TGTGAATTATGGATGTGTTTGT	(TC) ₁₀	248	282–292	60.0	0.57	3
FJ 55/EF136498	CAAACACTGACACACAAACGA	(TC) ₅ ; (TC) ₁₀ (CT) ₆ ; (CT) ₁₅	180	144–230	60.0	0.77	7
FJ 56/EF136499	AGAGAAATCAAAGCCGAAGGA ATTATGCTCTTCTTTGTTT GAAAATATTAGAAAAATGCCGC	(CA) ₅	263	266–282	60.0	0.42	3
FJ 57/EF136500	CCTAGGTTCTGCGCTGCTTTTG TCCCTCTCATCATCCCATAA	(AG) ₅ ; (AGG) ₅	263	227	45.0	mono	1
FJ 58/EF136501	GAGCTTCGGTTCTTTCTTTG GCAGCCTCCACATAATAGTAA	(TC) ₈	186	195–198	60.0	0.48	2
FJ 59/EF136502	GAGCGTGTCCGAAGCGAGTT ACCCAGCCGAGACCAGCAG	(AG) ₅ (GT) ₈ ; (GA) ₅	201	237–240	60.0	0.61	3
FJ 60/EF136503	GGAGGGCATGAGAAGACGACT ACAAGACACGGGCAAAATAAAT	(CT) ₅	134	147–150	45.0	0.40	2

Note: Na, number of alleles amplified using 25 bean genotypes.

SSRs in 5255 common bean ESTs, thus indicating that SSRs also exist in the expressed genome of this crop species.

In the Bean EST database, SSRs were found in 7% of the total unigene sequences (Melotto et al. 2005). The SSR frequency in the Bean EST database is similar to that found in soybeans (5.2%) but lower than in sweet potatoes (9.1%) (Cardle et al. 2000). However, this comparison should be interpreted with caution, as the method for identifying microsatellites may well be different.

Earlier studies on the abundance of different SSR motifs in plant databases indicated that AT was the most common dinucleotide repeat (Lagercrantz et al. 1993). However, our data showed that in the common bean the most frequent dinucleotide repeat was AG. In *Arabidopsis thaliana* and soybean, in fact, expressed sequences had shown a higher frequency of AG repeats than AT repeats (Morgante et al. 2002).

Among the trinucleotidic microsatellites, GAA, CAT, and TGG motifs were the most abundant, accounting for 35.4%, 18.8%, and 10.4% of occurrences, respectively. Wang et al. (1994) identified AAT, AAC, and TGC as the most common in plants, in contrast to our data. However, differences also exist among plant species in terms of the prevalence of particular microsatellite motifs. In cereals, for example, the CGG motif is the most common (Cordeiro et al. 2001; Varshney et al. 2002), whereas in soybean (Gao et al. 2003) and *Arabidopsis* (Cardle et al. 2000) the GAA motif prevails. Our research supports the proposal that the motif

GAA is more prevalent among dicots, as previously reported.

The trinucleotide repeats are seen as the most abundant class of microsatellites in the EST sequences of several species, including the dicots *A. thaliana* and soybean (Morgante et al. 2002; Gao et al. 2003). We found dinucleotide microsatellites as the most abundant in our EST libraries. This can be explained by the different stringencies used for the SSR detection. We used a minimum of 10 bp for the dinucleotides and 15 bp for the trinucleotides. Furthermore, the dinucleotide microsatellites were found mainly in noncoding regions 5'UTR and 3'UTR (Fig. 1). Trinucleotides, in turn, were most abundant in coding portions of the EST (Fig. 1). The preferential occurrence of this type of SSR in coding regions may result from the fact that they do not cause a disruption in the triplet reading frame.

Of the ESTs further characterized in this study, 27 were homologous to known genes, as indicated by previous BLASTx alignments carried out by our group (Melotto et al. 2005). The PVEPSE3003E05 EST is similar to a disease resistance protein family (LRR) from *A. thaliana*. This EST contains a (AATT)₅ microsatellite in the 3'UTR. Another example is the TC127 consensus sequence of a cluster of ESTs, which carries the microsatellite (ATG)₆ and is similar to a salt-tolerance protein from *A. thaliana*.

The presence of SSR sequences in known gene sequences provides an opportunity to study the correlation between repeat numbers and the functional aspects of genes them-

Fig. 2. Display of a *Phaseolus vulgaris* microsatellite marker (PvM21) among 25 bean genotypes detected on silver-stained 6% polyacrylamide gels.

selves. EST-SSR combines the advantages of microsatellite variability with the information potentially carried by expressed sequences. Then, once EST-SSR loci have been mapped, they automatically yield information as to the location of the genes that carry them, thus representing an efficient way of mapping cDNA sequences and increasing the density of gene markers on linkage maps.

SSR polymorphism

The SSR markers derived from EST are considered less powerful in the discrimination of genotypes than other sources. Eujayl et al. (2001) compared EST- and genomic-SSR markers to investigate genotypic variation of 64 durum wheat lines, landraces, and varieties, obtaining 25% polymorphic loci among 137 EST-SSR markers and 50% among 108 genomic-SSR markers, with an average of 4.1 and 5.0 alleles per locus, respectively. Interestingly, we found similar results at both EST-SSR and genomic-SSR loci. For the EST-SSR markers we found an average of 2.7 alleles per locus and an average PIC value of 0.44, and for the genomic-SSR 2.4 alleles per locus were found, with a mean PIC value of 0.45.

Our EST-SSR data agree with previous findings from common-bean-sequence databases: by using 12 SSR loci de-

veloped from the GenBank database to assess the allelic variation in 12 common bean inbred lines, Yu et al. (1999) found 2.9 alleles per locus. Guerra-Sanz (2004) detected an average of 3.1 alleles per locus at 20 loci derived from public databases assessed in 12 *Phaseolus* accessions. However, the genomic SSRs had been more informative than in the present study. Gaitán-Solís et al. (2002) used 68 primer pairs to investigate the polymorphism among 21 *P. vulgaris* genotypes and found 6 alleles, on average, per primer pair. A similar result was described by Buso et al. (2006), who studied 10 loci isolated from a genomic library enriched for AG in 85 bean accessions. Average values of 7 alleles per locus and PIC of 0.59 were found.

In general, cluster analyses based on SSR data grouped our samples in accordance with the seed size (Fig. 3A and B) in the same way as was previously reported for grouping of common bean races (Singh 2001). The analysis based on EST-derived genetic distances (Fig. 3A) showed a clear distinction between Mesoamerican and Andean materials, both grouped separately with bootstrap support values of 69% and 80%, respectively. 'Flor de Mayo' appeared as an outlier, externally linked to other Mesoamerican beans. 'Bat 93', one of the parents of the bean core map population grouped together with 'Carioca Comum', 'IAC-UNA', 'Ja-

Fig. 3. Relationships among 25 bean genotypes. The unrooted trees ((A) from EST-SSR markers and (B) from genomic SSR markers) were constructed from the modified Rogers' distance matrix using the UPGMA (unweighted paired group method using arithmetic averages) and Drawgram programs of the PHYLIP package (Felsenstein 2004).

mapa', and 'Baetão'. With the exception of 'Tu', which is Andean, the first group consisted of Mesoamerican small-seeded beans, and the second group included Andean large-seeded beans. The Andean 'CAL-143' grouped together with 'J. EEP552', the other parent of the core map population with 98% bootstrap support. 'Fradinho Cruzeiro', a *Vigna* spp. accession, was internally linked to other Mesoamerican beans, as indicated by its small-seeded feature. In addition, the '87 JP 12', which is a *P. lunatus* L., was externally linked to Andean beans.

When the analysis based on genomic-derived genetic distances was performed, the 2-group pattern was maintained (Fig. 3B), though some modifications were noted, such as the clustering of 'Flor the Mayo' with 'Garbancillo', another 'Jalisco' genotype. Interesting groups were formed, such as that between 'Rio Tibagi' and 'IAC-UNA', improved cultivars that have black tegument and medium sized seeds. Some of the white Mesoamerican seeded beans tended to cluster together, e.g., 'Apetito Blanco', 'Garbancillo', and 'Great Northern'. 'Fradinho Cruzeiro' once more grouped with other Mesoamerican small-seeded beans.

Overall, microsatellites were useful for distinguishing Andean and Mesoamerican genotypes, for uncovering seed-size relations, and for separating cultivars from wild accessions. The ability to differentiate Andean and Mesoamerican gene pools was shared by both types of marker developed here, and dendrogram comparison showed only slight differences. 'Jalo EEP558' and 'Bat 93', common beans used as parents of the core mapping population (Nodari et al. 1993), were well differentiated by both sources of microsatellite loci.

In conclusion, we have developed a new set of microsatellite markers based on EST and genomic sequences and showed that the proportion of polymorphic SSRs in expressed regions of the bean genome is high, although the average number of alleles from EST-SSR and genomic-SSR were practically the same. Moreover, microsatellites developed from enriched-genomic libraries seemed to be more informative than EST-derived ones. The increasing availability of bean sequences from EST projects, as the one coordinated by the international consortium *Phaseomics* (Broughton et al. 2003), are valuable sources for SSR marker development. The markers developed here will certainly be useful for genetic mapping and marker-assisted selection programs. They are currently being mapped in an important recombinant inbred line population, 'Bat 93' × 'Jalo EEP558', and 'IAC-UNA' × 'CAL-143', which is part of the genetic material developed in Brazil, the world's main bean producer.

Acknowledgments

The EST data were based on projects supported by FAPESP, Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil (grants 00/09049-2 and 02/09049-8). This study was supported by FAPESP (Grants 04/13547-9; 02/03225-9; 02/00752-8) and CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil (140310/2005-3). We sincerely thank C.A. Oliveira for his technical support, and Ricardo Vieira Cesar for proofreading the manuscript.

References

- Anderson, J.A., Churchill, G.A., Autrique, J.E., Tanksley, S.D., and Sorrells, M.E. 1993. Optimizing parental selection for genetic linkage maps. *Genome*, **36**: 181–186.
- Billotte, N., Lagoda, P.J.L., Risterucci, A.M., and Baurens, C. 1999. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits*, **54**: 277–288.
- Blair, M.W., Pedraza, F., Buendia, H.F., Gaitán-Solís, E., Beebe, S.E., Gepts, P., and Tohme, J. 2003. Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* **107**: 1362–1374. doi:10.1007/s00122-003-1398-6. PMID:14504741.
- Broughton, W.J., Hernández, G., Blair, M., Beebe, S., Gepts, P., and Vanderleyden, J. 2003. Beans (*Phaseolus* spp.) – model food legumes. *Plant Soil*, **252**: 55–128. doi:10.1023/A:1024146710611.
- Brown, S.M., Hopkins, M.S., Mitchell, S.E., Senior, M.L., Wang, T.Y., Duncan, R.R., et al. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench.]. *Theor. Appl. Genet.* **93**: 190–198. doi:10.1007/s001220050265.
- Buso, G.S.C., Amaral, Z.P.S., Brondani, R.P.V., and Ferreira, M.E. 2006. Microsatellite markers for the common bean *Phaseolus vulgaris*. *Mol. Ecol. Notes*, **6**: 252. doi:10.1111/j.1471-8286.2006.01210.x.
- Cardle, L., Ramsay, L., Milbourne, D., Macaulay, M., Marshall, D., and Waugh, R. 2000. Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics*, **156**: 847–854. PMID:11014830.
- Coelho, A.S.G. 2002. BooD: avaliação dos erros associados a estimativas de distâncias/similaridades genéticas através do procedimento de bootstrap com número variável de marcadores. Computer software distributed by the author. Laboratório de Genética Vegetal, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, Brasil. [In Portuguese.]
- Cordeiro, G.M., Casu, R., McIntyre, C.L., Manners, J.M., and Henry, R.J. 2001. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. *Plant Sci.* **160**: 1115–1123. doi:10.1016/S0168-9452(01)00365-X. PMID:11337068.
- Creste, S., Tulmann-Neto, A., and Figueira, A. 2001. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. *Plant Mol. Biol. Rep.* **19**: 299–306.
- Doyle, J.J., and Doyle, J.L. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**: 11–15.
- Eujayl, I., Sorrells, M.E., Baum, M., Wolters, P., and Powell, W. 2001. Assessment of genotype variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. *Euphytica*, **119**: 39–43. doi:10.1023/A:1017537720475.
- Ewing, B., Hillier, L., Wendl, M.C., and Green, P. 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res.* **8**: 175–185. PMID:9521921.
- Felsenstein, J. 2004. PHYLIP (Phylogeny Inference Package) v. 3.6 [computer program]. Department of Genetics, University of Washington, Seattle. [Distributed by the author.]
- Field, D., and Wills, C. 1996. Long, polymorphic microsatellites in simple organisms. *Proc. Biol. Sci.* **263**: 209–215. doi:10.1098/rspb.1996.0033. PMID:8728984.
- Freyre, R., Skroch, P., Geffroy, V., Adam-Blondon, A.-F., Shirmohamadi, A., Johnson, W., et al. 1998. Towards an integrated linkage map of common bean. 4. Development of a core map

- and alignment of RFLP maps. *Theor. Appl. Genet.* **97**: 847–856. doi:10.1007/s001220050964.
- Gaitán-Solís, E., Duque, M.C., Edwards, K.J., and Tohme, J. 2002. Microsatellite Repeats in Common Bean (*Phaseolus vulgaris*): isolation, characterization, and cross-species amplification in *Phaseolus* spp. *Crop Sci.* **42**: 2128–2136.
- Gao, L., Tang, J., Li, H., and Jia, J. 2003. Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Mol. Breed.* **12**: 245–261. doi:10.1023/A:1026346121217.
- Gepts, P. 1998. Origin and evolution of common bean: past events and recent trends. *HortScience*, **33**: 1124–1130.
- Guerra-Sanz, J.M. 2004. New SSR markers of *Phaseolus vulgaris* from sequence databases. *Plant Breed.* **123**: 87–89. doi:10.1046/j.0179-9541.2003.00929.x.
- Hedrick, P.W. 2001. Conservation genetics: where are we now? *Trends Ecol. Evol.* **16**: 629–636.
- Hu, J., Nakatani, M., Mizuno, K., and Fujimura, T. 2004. Development and characterization of microsatellite markers in sweetpotato. *Breed. Sci.* **54**: 177–188. doi:10.1270/jsbbs.54.177.
- Huang, X., and Madan, A. 1999. CAP3: A DNA sequence assembly program. *Genome Res.* **9**: 868–877. doi:10.1101/gr.9.9.868. PMID:10508846.
- Jarne, P., and Lagoda, P.J.L. 1996. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **11**: 424–429. doi:10.1016/0169-5347(96)10049-5.
- Koinange, E.M.K., Singh, S.P., and Gepts, P. 1996. Genetic control of the domestication syndrome in common-bean. *Crop Sci.* **36**: 1037–1045.
- Lagercrantz, U., Ellegren, H., and Andersson, L. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Res.* **21**: 1111–1115. doi:10.1093/nar/21.5.1111. PMID:8464696.
- Li, Y.-C., Korol, A.B., Fahima, T., and Nevo, E. 2004. Microsatellites within genes: structure, function, and evolution. *Mol. Biol. Evol.* **21**: 991–1007. PMID:14963101.
- Melo, L.C., Santos, J.B., and Ferreira, D.F. 2002. Mapeamento de QTLs para reação ao oídio e mancha-angular do feijoeiro-comum em diferentes locais. *Pesqui. Agropecu. Bras.* **37**: 1115–1126.
- Melotto, M., Monteiro-Vitorello, C.B., Bruschi, A.G., and Camargo, L.E.A. 2005. Comparative bioinformatic analysis of genes expressed in common bean (*Phaseolus vulgaris* L.) seedlings. *Genome*, **48**: 562–570. doi:10.1139/g05-010. PMID:16121253.
- Miklas, P.N., Coyne, D.P., Grafton, K.F., Mutlu, N., Reiser, J., Lindgren, D.T., and Singh, S.P. 2003. A major QTL for common bacterial blight resistance derives from the common bean great northern landrace cultivar Montana No. 5. *Euphytica*, **131**: 137–146. doi:10.1023/A:1023064814531.
- Miller, M.P. 1997. Tools for population genetic analysis (TFPGA), version 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author. Available from <http://herb.bio.nau.edu/~miller> [accessed January 2006].
- Morgante, M., Hanafey, M., and Powell, W. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat. Genet.* **30**: 194–200. doi:10.1038/ng822. PMID:11799393.
- Morgante, M., and Olivieri, A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. *Plant J.* **3**: 175–182. doi:10.1046/j.1365-3113.1993.0119-00999.x. PMID:8401603.
- Nodari, R.O., Tsai, S.M., Yu, K.F., Park, S.J., Zhang, B.L., Haffner, M., et al. 1993. Towards an integrated linkage map of common bean. 3. Mapping genetic factors controlling host-bacteria interactions. *Genetics*, **134**: 341–350. PMID:8514141.
- Oliveira, E.J., Pádua, J.G., Zucchi, M.I., Vencovsky, R., and Vieira, M.L.C. 2006. Origin, evolution and genome distribution of microsatellites. *Genet. Mol. Biol.* **29**: 294–307.
- Peng, J.H., and Lapitan, N.L.V. 2005. Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. *Funct. Integr. Genomics*, **5**: 80–96. doi:10.1007/s10142-004-0128-8. PMID:15650880.
- Powell, W., Machray, G.C., and Provan, J. 1996. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **1**: 215–222.
- Schlotterer, C., and Wiehe, T. 1999. Microsatellite, a neutral marker to infer selective sweeps. In *Microsatellites: evolution and applications*. Edited by D.B. Goldstein and C. Schlotterer. Oxford University Press, Oxford. pp. 238–247.
- Scott, K.D. 2001. Microsatellites derived from ESTs, and their comparison with those derived by other methods. In *Plant genotyping: the DNA fingerprinting of plants*. Edited by R.J. Henry. CAB International, Wallingford. pp. 255–237.
- Singh, S.P. 2001. Broadening the genetic base of common bean cultivars: a review. *Crop Sci.* **41**: 1659–1675.
- Tarlan, B., Michaels, T.E., and Pauls, K.P. 2001. Mapping genetic factors affecting the reaction to *Xanthomonas axonopodis* pv. *phaseoli* in *Phaseolus vulgaris* L. under field conditions. *Genome*, **44**: 1046–1056. PMID:11768208.
- Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., and McCouch, S. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Res.* **11**: 1441–1452. doi:10.1101/gr.184001. PMID:11483586.
- Tóth, G., Gaspari, Z., and Jurka, J. 2000. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* **10**: 967–981. doi:10.1101/gr.10.7.967. PMID:10899146.
- Varshney, R.K., Thiel, T., Stein, N., Langridge, P., and Graner, A. 2002. In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cell. Mol. Biol. Lett.* **7**: 537–546. PMID:12378259.
- Varshney, R.K., Graner, A., and Sorrells, M.E. 2005. Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* **23**: 48–55. doi:10.1016/j.tibtech.2004.11.005. PMID:15629858.
- Yaish, M.W.F., and Pérez de la Vega, M. 2003. Isolation of (GA)_n microsatellite sequences and description of a predicted MADS-box sequence isolated from common bean (*Phaseolus vulgaris* L.). *Genet. Mol. Biol.* **26**: 337–342.
- Yu, Z.H., Stall, R.E., and Vallejos, C.E. 1998. Detection of genes for resistance to common bacterial blight of beans. *Crop Sci.* **38**: 1290–1296.
- Yu, K., Park, S.J., and Poysa, V. 1999. Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). *Genome*, **42**: 27–34. doi:10.1139/gen-42-1-27.
- Yu, K., Park, S.J., Poysa, V., and Gepts, P. 2000. Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *J. Hered.* **91**: 429–434. doi:10.1093/jhered/91.6.429. PMID:11218079.
- Wang, Z., Weber, J.L., Zhong, G., and Tanksley, S.D. 1994. Survey of plant short tandem repeats. *Theor. Appl. Genet.* **88**: 1–6.
- Wright, S. 1978. Evolution and the genetics of populations, Vol. 4. Variability within and among natural populations. University of Chicago Press, Chicago.