

JANAÍNA PAULA MARQUES TANURE

**DESENVOLVIMENTO E VALIDAÇÃO DE MARCADORES
MICROSSATÉLITES PARA O FEIJÃO-COMUM**

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

**VIÇOSA
MINAS GERAIS - BRASIL
2009**

JANAÍNA PAULA MARQUES TANURE

**DESENVOLVIMENTO E VALIDAÇÃO DE MARCADORES
MICROSSATÉLITES PARA O FEIJÃO-COMUM**

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

APROVADA: 03 de Agosto de 2009

**Prof^a. Márcia Regina Costa
(Co-orientador)**

**Prof. Maurílio Alves Moreira
(Co-orientador)**

Pesq^a. Eveline Teixeira Caixeta

**Prof^a. Tânia Maria Fernandes
Salomão**

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

Aos meus amados pais **Paulo** (*in memoriam*) e **Ondina** por todo o amor,
sacrifício, apoio, incentivo e confiança

À minha querida tia **Cecília** por todo amor, carinho, dedicação
e orações

À minha querida irmã **Natália** pelo apoio, companheirismo
e amizade sempre constante

Ao amado **Matheus** por todo amor, confiança,
apoio e companheirismo incondicionais

DEDICO

AGRADECIMENTOS

A Deus, pela proteção e orientação constante, pelas oportunidades, vitórias e bênçãos concedidas e por ser sempre meu porto seguro.

Aos meus pais Paulo (*in memoriam*) e Ondina pelo apoio, confiança e incentivo sempre constantes, por todo o sacrifício e amor incondicional em todos esses anos, essa minha grande vitória a vocês pertence.

À minha querida tia Cecília, minha segunda mãe, por todo o amor, preocupação, sacrifício e pelas orações, pensamentos positivos e de incentivo sempre tão essenciais em minha vida.

À minha querida irmã Natália, pela amizade, apoio, compreensão, companheirismo e cumplicidade, e pela convivência tão agradável.

Ao meu namorado Matheus, pela presença e companheirismo tão constantes em minha vida, por ter abraçado meus sonhos como se fossem os seus próprios, pela confiança, incentivo, amizade, carinho e amor sempre incondicionais, pela compreensão e apoio nos momentos difíceis e pelas boas risadas e companhia tão agradável.

Ao Professor Everaldo Gonçalves de Barros, pela orientação durante todos esses anos, pelos valiosos conselhos e imenso aprendizado, pela compreensão, apoio e confiança em mim depositada, pela disponibilidade e ajuda mesmo quando dispunha de muitos compromissos, e principalmente por colaborar sobremaneira para a minha formação acadêmica e pessoal.

Ao Professor Maurílio Alves Moreira, pelo apoio recebido, sugestões e convivência no laboratório tão agradável.

À Professora Márcia Regina Costa, por ser a grande responsável pelo meu treinamento científico, pela ajuda e disponibilidade sempre incondicionais e presentes, pelos anos de orientação, de aprendizado, de amizade, e de carinho, e por ter aceitado tão prontamente me co-orientar nesse trabalho.

Aos membros da banca examinadora, Professora Tânia Maria Fernandes Salomão e Doutora Eveline Teixeira Caixeta, pelas críticas sempre muito construtivas, contribuições e sugestões para a melhoria desse trabalho.

À Universidade Federal de Viçosa e ao Programa de Pós Graduação em Genética e Melhoramento pela sólida formação acadêmica e pela oportunidade da realização do curso de Mestrado.

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela concessão da bolsa.

À Agrogenética pela possibilidade de cursar uma parte do curso de mestrado com vínculo empregatício.

Aos amigos e companheiros do Laboratório de Genética Molecular de Plantas e Proteína pela convivência tão agradável e pelo crescimento e ajuda compartilhada, em especial ao Klever, Lôreta, Fernanda, Newton, Bruno, Thiago, Giselle, Márcia Costa, Márcia Flores, Polyana, Denise, Danielle, Demerson, Martha, Léo e Pedro Ivo.

Aos funcionários Cássio e Gláucia pela disponibilidade e ajuda.

À Suelen, pela amizade e companheirismo sempre presentes, pelas boas risadas e por compartilhar comigo as dificuldades no laboratório todos esses anos.

À Valéria e Jeziel pela amizade, pela convivência tão agradável, pela ajuda, e pelo crescimento científico que compartilhamos no laboratório nesses anos de brincarmos de cientistas.

Ao estagiário e amigo Luiz, por toda ajuda e disponibilidade constante que foram imprescindíveis para a realização desse trabalho, pela amizade e convivência tão agradável.

Aos amigos da Biologia 2003, pela amizade, companheirismo e apoio sempre constante durante todos esses anos.

Ao amigo Leandro, pela presença constante, amizade incondicional, conselhos, apoio, e por ser o irmão que a vida me deu de presente.

Às amigas Camis, Ju, Laris, Lu e Ní, por toda a amizade, companheirismo, apoio e pelos momentos inesquecíveis de alegria e também de lágrimas que compartilhamos, pelo apoio incondicional.

Aos amigos do Bioprédio, em especial à Camis, Ju, Ní, Natália, Betão, Bernardo, Diego, Diogo e Lu, pela amizade, pelos momentos inesquecíveis vividos nesses anos, pelos sonhos e vitórias compartilhados, e por terem sido minha família em Viçosa.

Aos amigos e companheiros de trabalho da Agrogenética, pela grande receptividade, por todo o apoio e torcida nesse momento final.

E a todos que contribuíram de forma direta ou indireta para a realização desse trabalho.

MUITO OBRIGADA!

BIOGRAFIA

JANAÍNA PAULA MARQUES TANURE, filha de Paulo Roberto Tanure e Ondina Maria Marques Tanure, nasceu em Divinópolis, Minas Gerais, em 07 de Janeiro de 1985. Residente em Divinópolis, concluiu o ensino médio no Colégio Júlio Verne, em 2002.

Em Março de 2003, iniciou o curso de Ciências Biológicas na Universidade Federal de Viçosa (UFV). Graduou-se bacharel em Ciências Biológicas, com ênfase em Genética e Biologia Molecular, em 10 de agosto de 2007.

Em Agosto de 2007, iniciou o curso de mestrado em Genética e Melhoramento pela Universidade Federal de Viçosa (UFV), sob a orientação do Profº Everaldo Gonçalves de Barros, submetendo-se à defesa de dissertação em 03 de agosto de 2009.

Em Abril de 2009, foi admitida no cargo de Analista de Laboratório na empresa Agrogenética - Laboratório de Análises Genéticas Ltda, sendo posteriormente convidada para o cargo de Responsável Técnica.

SUMÁRIO

LISTA DE FIGURAS E TABELAS	VIII
LISTA DE ABREVIações	X
RESUMO.....	XI
ABSTRACT.....	XIII
1. INTRODUÇÃO GERAL	1
2. OBJETIVOS.....	6
2.1. OBJETIVO GERAL.....	6
2.2. OBJETIVOS ESPECÍFICOS	6
3. REFERÊNCIAS BIBLIOGRÁFICAS.....	7
CAPÍTULO I	
DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES PARA O FEIJÃO-COMUM A PARTIR DE BIBLIOTECAS GENÔMICAS ENRIQUECIDAS	
RESUMO.....	10
1. INTRODUÇÃO	11
2. MATERIAL E MÉTODOS.....	16
2.1. MATERIAL GENÉTICO E EXTRAÇÃO DE DNA.....	16
2.2. DIGESTÃO DO DNA GENÔMICO, SEPARAÇÃO E PURIFICAÇÃO DOS FRAGMENTOS DE DNA	17
2.3. REMOÇÃO DAS EXTREMIDADES COESIVAS E DESFOSFORILAÇÃO DOS FRAGMENTOS DE DNA	17
2.4. LIGAÇÃO DO ADAPTADOR UNIVERSAL SNX AOS FRAGMENTOS DE DNA	18
2.4.1. <i>Adaptador SNX</i>	18
2.4.2. <i>Ligação dos fragmentos ao adaptador SNX</i>	19
2.4.3. <i>Confirmação da eficiência da reação de ligação</i>	19
2.5. REAÇÃO DE <i>NICK-LIGATION</i>	19
2.6. REAÇÃO PÓS <i>NICK-LIGATION</i>	20
2.7. HIBRIDIZAÇÃO DOS FRAGMENTOS PÓS <i>NICK-LIGATION</i> COM OLIGONUCLEOTÍDEOS CONTENDO MICROSSATÉLITES	20
2.7.1. <i>Preparação das partículas magnéticas</i>	21
2.7.2. <i>Captura dos fragmentos de DNA hibridizados</i>	21
2.7.3. <i>Síntese de fragmentos de DNA fita dupla</i>	22
2.8. PREPARAÇÃO DOS FRAGMENTOS DE DNA E DO PLASMÍDEO	22
2.9. LIGAÇÃO DOS FRAGMENTOS DE DNA AO VETOR E TRANSFORMAÇÃO DE CÉLULAS COMPETENTES	23
2.10. SELEÇÃO DE CLONES POSITIVOS E SEQÜENCIAMENTO.....	24
2.11. IDENTIFICAÇÃO DE REGIÕES SSR E DESENHO DE <i>PRIMERS</i>	24
2.12. DETERMINAÇÃO DAS CONDIÇÕES DE AMPLIFICAÇÃO E IDENTIFICAÇÃO DE MARCADORES MICROSSATÉLITES POLIMÓRFICOS ENTRE OS GENÓTIPOS ANDINOS E MESOAMERICANOS.....	26
2.13. CARACTERIZAÇÃO DOS <i>PRIMERS</i> POLIMÓRFICOS QUANTO AO ÍNDICE DE INFORMAÇÃO DE POLIMORFISMO (PIC) E NÚMERO DE ALELOS DETECTADOS	27
2.14. ANÁLISES DE SIMILARIDADE	28
3. RESULTADOS E DISCUSSÃO.....	30
3.1. OBTENÇÃO DAS BIBLIOTECAS GENÔMICAS ENRIQUECIDAS	30
3.2. SEQÜENCIAMENTO E DESENHO DE <i>PRIMERS</i> SSR.....	30

3.3. IDENTIFICAÇÃO DE MARCADORES MICROSSATÉLITES POLIMÓRFICOS ENTRE OS GENÓTIPOS ANDINOS E MESOAMERICANOS	38
3.4. ANÁLISE DE SIMILARIDADE	41
4. REFERÊNCIAS BIBLIOGRÁFICAS.....	48

CAPÍTULO II

DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES PARA O FEIJÃO-COMUM A PARTIR DE SEQUÊNCIAS INTERNAS SIMPLES REPETIDAS (ISSR)

RESUMO.....	53
1. INTRODUÇÃO	54
2. MATERIAL E MÉTODOS	58
2.1. MATERIAL GENÉTICO E EXTRAÇÃO DE DNA	58
2.2. AMPLIFICAÇÃO POR PCR DE FRAGMENTOS RICOS EM SEQUÊNCIAS MICROSSATÉLITES, UTILIZANDO MARCADORES ISSR.....	58
2.3. LIGAÇÃO DOS FRAGMENTOS DE DNA AO VETOR E TRANSFORMAÇÃO DE CÉLULAS COMPETENTES	60
2.4. SELEÇÃO DE CLONES POSITIVOS E SEQUENCIAMENTO	60
2.5. IDENTIFICAÇÃO DE REGIÕES SSR E DESENHO DE <i>PRIMERS</i>	61
2.6. DETERMINAÇÃO DAS CONDIÇÕES DE AMPLIFICAÇÃO E IDENTIFICAÇÃO DE MARCADORES MICROSSATÉLITES POLIMÓRFICOS ENTRE OS GENÓTIPOS ANDINOS E MESOAMERICANOS.....	63
2.7. ANÁLISES DE SIMILARIDADE	64
3. RESULTADOS E DISCUSSÃO.....	66
3.1. ANÁLISE DA AMPLIFICAÇÃO POR PCR UTILIZANDO MARCADORES ISSR	66
3.2. SEQUENCIAMENTO E DESENHO DE <i>PRIMERS</i> SSR.....	68
3.3. IDENTIFICAÇÃO DE MARCADORES MICROSSATÉLITES POLIMÓRFICOS ENTRE OS GENÓTIPOS ANDINOS E MESOAMERICANOS ANALISADOS.....	74
3.4. ANÁLISE DE SIMILARIDADE	77
4. REFERÊNCIAS BIBLIOGRÁFICAS.....	79
RESUMO DOS RESULTADOS, CONCLUSÕES E PERSPECTIVAS.....	83

LISTA DE FIGURAS E TABELAS

CAPÍTULO I

- FIGURA 1:** PERFIL ELETROFORÉTICO DO DIAGNÓSTICO DE COLÔNIAS UTILIZANDO OS *PRIMERS* T3 E T7. M: MARCADOR DE TAMANHO 100 PB; 1 A 28: CLONES DIAGNOSTICADOS. A SETA INDICA A POSIÇÃO CORRESPONDENTE A 400 PB.....30
- FIGURA 2:** SEQUÊNCIAS DE FRAGMENTOS CONTENDO DIFERENTES NÚMEROS E TIPOS DE REPETIÇÕES MICROSSATÉLITES. (A) UMA REPETIÇÃO IMPERFEITA E SIMPLES – (TC)₈A(CT)₁₉. (B) DUAS REPETIÇÕES PERFEITAS NO MESMO FRAGMENTO – (GT)₅ E (AG)₈. (C) UMA REPETIÇÃO COMPOSTA – (GA)₃(GGA)₂. A SIMBOLOGIA >>>>> REPRESENTA A REGIÃO DO FRAGMENTO QUE ORIGINOU A SEQUÊNCIA DO *PRIMER* F; <<<<<< A QUE ORIGINOU SEQUÊNCIA DO *PRIMER* R. 31
- FIGURA 3:** PERFIL ELETROFORÉTICO POLIMÓRFICO APRESENTADO PELO *PRIMER* PVAG003 EM GEL DE POLIACRILAMIDA 10%, CORADO COM SOLUÇÃO DE NITRATO DE PRATA . A NUMERAÇÃO 1-18 CORRESPONDE À ORDEM DOS GENÓTIPOS ANDINOS E MESOAMERICANOS DISPOSTA NA TABELA 1, DO ITEM 2.1, COM EXCEÇÃO DOS DOIS ÚLTIMOS GENÓTIPOS QUE NÃO FORAM TESTADOS NO PRESENTE GEL. A LETRA M CORRESPONDE AO MARCADOR DE TAMANHO DE 25 PB.40
- TABELA 1:** GENÓTIPOS ANDINOS E MESOAMERICANOS UTILIZADOS NA ANÁLISE DE POLIMORFISMO DOS *PRIMERS* MICROSSATÉLITES DESENVOLVIDOS16
- TABELA 2:** MOTIVOS MICROSSATÉLITES IDENTIFICADOS, SEQUÊNCIAS DOS *PRIMERS* CORRESPONDENTES, VALORES DE T_M, TIPO DE REPETIÇÃO E TAMANHO DOS FRAGMENTOS AMPLIFICADOS. O SÍMBOLO // ENTRE ALGUNS MOTIVOS DENOTA A PRESENÇA DE OUTROS MOTIVOS SSR RELATIVAMENTE DISTANTES NO MESMO FRAGMENTO.32
- TABELA 3:** TEMPERATURA DE AMPLIFICAÇÃO PADRONIZADA PARA OS 34 PARES DE *PRIMERS* ANALISADOS E IDENTIFICAÇÃO DE QUAIS DETECTARAM POLIMORFISMO ENTRE OS GENÓTIPOS TESTADOS.....38
- TABELA 4:** CARACTERIZAÇÃO DOS LOCOS MICROSSATÉLITES POLIMÓRFICOS QUANTO AO NÚMERO DE ALELOS DETECTADOS, VARIAÇÃO DE TAMANHO (PB) DOS ALELOS, E O CONTEÚDO DE INFORMAÇÃO DE POLIMORFISMO (PIC).39
- TABELA 5:** RESULTADO DA ANÁLISE DE SIMILARIDADE. A PRIMEIRA COLUNA CORRESPONDE AO LOCO MICROSSATÉLITE ANALISADO, A SEGUNDA AO *HIT* ENCONTRADO, SEGUIDA DA PORCENTAGEM DE SIMILARIDADE ENTRE AS DUAS SEQUÊNCIAS, SCORE DE ALINHAMENTO E *E-VALUE*. O NÚMERO DE ACESSO, DO BANCO DO NCBI, DAS SEQUÊNCIAS IDENTIFICADAS COMO SIMILARES ENCONTRA-SE NA ÚLTIMA COLUNA.....44
- TABELA 6:** CATEGORIZAÇÃO DOS PRODUTOS GÊNICOS SEGUNDO CRITÉRIOS DE CLASSIFICAÇÃO DO *GENE ONTOLOGY*.45

CAPÍTULO II

- FIGURA 1:** PERFIL ELETROFORÉTICO EM GEL DE AGAROSE 1,8%/SB 1X DE PRODUTOS DE PCR OBTIDOS A PARTIR DA AMPLIFICAÇÃO POR MARCADORES ISSR. CANALETAS 1-4 CORRESPONDE AO USO DE APENAS UM TIPO DE *PRIMER* ISSR NA REAÇÃO, CANALETAS 5-9 À COMBINAÇÃO DE DOIS *PRIMERS* ISSR DISTINTOS E AS CANALETAS 10-14 À COMBINAÇÃO DE TRÊS *PRIMERS* ISSR. M- MARCADOR DE TAMANHO DE 100 PB.....666
- FIGURA 2:** PERFIL ELETROFORÉTICO EM GEL DE AGAROSE 1,8%/SB 1X DOS PRODUTOS DE AMPLIFICAÇÃO OBTIDOS POR ISSR-PCR, A PARTIR DA UTILIZAÇÃO DE UM *PRIMER* DISTINTO (1-5) E DA COMBINAÇÃO DE DOIS *PRIMERS* (6-15). M- MARCADOR DE TAMANHO DE 100 PB.....677
- FIGURA 3:** PERFIL ELETROFORÉTICO, EM GEL DE AGAROSE 1,8%/SB 1X, DO PCR DE COLÔNIA REALIZADO COM OS *PRIMERS* M13 FOWARD E REVERSE. A LETRA M INDICA O MARCADOR DE TAMANHO DE 100 PB E AS DEMAIS CANALETAS, 29 CLONES ANALISADOS. A SETA INDICA FRAGMENTOS COM APROXIMADAMENTE 400 PB.....68
- FIGURA 4:** SEQÜÊNCIAS DE FRAGMENTOS CONTENDO DIFERENTES NÚMEROS E TIPOS DE REPETIÇÕES MICROSSATÉLITES. (A) UMA REPETIÇÃO IMPERFEITA E SIMPLES – (AG)₈T(CT)₁₉. (B) DUAS REPETIÇÕES PERFEITAS NO MESMO FRAGMENTO – (GT)₅ E (AG)₈. (C) UMA REPETIÇÃO COMPOSTA – (GA)₃(GGA)₂. A SIMBOLOGIA >>>>> REPRESENTA A REGIÃO DO FRAGMENTO QUE ORIGINOU A SEQÜÊNCIA DO *PRIMER* F; <<<<<< A QUE ORIGINOU SEQÜÊNCIA DO *PRIMER* R. 69
- FIGURA 5:** PERFIL ELETROFORÉTICO POLIMÓRFICO APRESENTADO PELO MARCADOR DOMINANTE PVAC001 EM GEL DE POLIACRILAMIDA 10%, CORADO COM SOLUÇÃO DE NITRATO DE PRATA . A NUMERAÇÃO 1-20 CORRESPONDE À ORDEM DOS GENÓTIPOS ANDINOS E MESOAMERICANOS DISPOSTA NA TABELA 1.....766
- TABELA 1:** GENÓTIPOS ANDINOS E MESOAMERICANOS UTILIZADOS NO TESTE DE POLIMORFISMO DOS *PRIMERS* MICROSSATÉLITES DESENVOLVIDOS.588
- TABELA 2:** SEQÜÊNCIAS DOS *PRIMERS* ISSR USADOS NA AMPLIFICAÇÃO DO DNA DE TO E SUAS RESPECTIVAS TEMPERATURAS DE AMPLIFICAÇÃO.....59
- TABELA 3:** RELAÇÃO DOS *PRIMERS* ISSR USADOS NAS REAÇÕES DE AMPLIFICAÇÃO INDIVIDUAIS, E AS COMBINAÇÕES DE DOIS *PRIMERS* UTILIZADAS.....677
- TABELA 4:** MOTIVOS MICROSSATÉLITES IDENTIFICADOS, SEQÜÊNCIAS DOS *PRIMERS* CORRESPONDENTES, VALORES DE T_M, TIPO DE REPETIÇÃO E TAMANHO DOS FRAGMENTOS AMPLIFICADOS. O SÍMBOLO // ENTRE ALGUNS MOTIVOS DENOTA A PRESENÇA DE OUTROS MOTIVOS SSR RELATIVAMENTE DISTANTES NO MESMO FRAGMENTO.70
- TABELA 5:** TEMPERATURA DE AMPLIFICAÇÃO PADRONIZADA PARA OS 10 PARES DE *PRIMERS* ANALISADOS E IDENTIFICAÇÃO DE QUAIS APRESENTARAM PADRÃO DE CODOMINÂNCIA E DE DOMINÂNCIA.755
- TABELA 6:** RESULTADO DA ANÁLISE DE SIMILARIDADE. A PRIMEIRA COLUNA CORRESPONDE AO LOCO MICROSSATÉLITE ANALISADO; A SEGUNDA AO *HIT* ENCONTRADO; SEGUIDA DA PORCENTAGEM DE SIMILARIDADE ENTRE AS DUAS SEQÜÊNCIAS, *SCORE* DE ALINHAMENTO E *E-VALUE*. O NÚMERO DE ACESSO, DO BANCO DO NCBI, DAS SEQÜÊNCIAS IDENTIFICADAS COMO SIMILARES ENCONTRA-SE NA ÚLTIMA COLUNA.....78
- TABELA 7:** CATEGORIZAÇÃO DOS PRODUTOS GÊNICOS SEGUNDO CRITÉRIOS DE CLASSIFICAÇÃO DO *GENE ONTOLOGY*.....78

LISTA DE ABREVIações

AFLP <i>Amplified Fragment Length Polymorphism</i>	PIC <i>Polymorphic Information Content</i>
ATP Adenosina trifosfato	QTL <i>Quantitative Trait Loci</i>
BLAST <i>Basic Local Alignment Search Tool</i>	RAPD <i>Randomly Amplified Polymorphic</i>
BSA Albumina de soro bovina	RFLP <i>Restriction Fragment Length Polymorphism</i>
CAP3 <i>Sequence Assembly Program</i>	RIL <i>Recombinant Inbred Line</i>
CIP <i>Calf Intestinal Alkaline Phosphatase</i>	RNA Ácido ribonucleico
cpSSR <i>Chloroplast Simple Sequence Repeats</i>	SB Borato de sódio
CTAB Brometo de cetiltrimetilamônio	SCAR <i>Sequence Characterized Amplified Region</i>
DNA Ácido desoxirribonucleico	SDS Dodecil sulfato de sódio
DDJ <i>DNA Data Bank of Japan</i>	SSC Citrato salino de sódio
DMSO Dimetilsulfoxido	SSR <i>Simple Sequence Repeats</i>
dNTP Desoxiribonucleosídeo trifosfato	SNP <i>Single Nucleotide Polymorphism</i>
EDTA Ácido etileno diamino tetracético	TAE Tris-base, Ácido acético, EDTA
EMBL <i>European Molecular Biology Laboratory</i>	Tampão NEB <i>New England Biolabs</i>
hsp <i>Heat Shock Protein</i>	Taq <i>Thermus aquaticus</i>
ISSR <i>Internal Simple Sequence Repeats</i>	TBE Tris-base, EDTA
MgCl₂ Cloreto de Magnésio	TE Tris-EDTA
MBN <i>Mung Bean Nuclease</i>	TEMED N,N,N',N' Tetrametiletileno diamina
NADH Nicotinamida adenina dinucleotídeo	Tm Temperatura média de fusão
NCBI <i>National Center for Biotechnology Information</i>	TRIS Tris-(hidroximetil)-aminometano
pb Pares de base	
PCR <i>Polimerase Chain Reaction</i>	
pH Potencial Hidrogeniônico	

RESUMO

TANURE, Janaína Paula Marques, M.Sc., Universidade Federal de Viçosa, agosto de 2009. **Desenvolvimento e validação de marcadores microssatélites para o feijão-comum.** Orientador: Everaldo Gonçalves de Barros. Co-orientadores: Márcia Regina Costa e Maurílio Alves Moreira.

O feijão-comum (*Phaseolus vulgaris* L.) é uma cultura de destacada importância nutricional, econômica e social. Programas de melhoramento genético do feijoeiro têm utilizado marcadores moleculares como importantes ferramentas auxiliares em diversos tipos de estudos genéticos. Diferentes classes de marcadores têm sido desenvolvidas, dentre as quais se destacam os microssatélites. Os microssatélites (SSR) são seqüências simples repetidas de DNA, que se repetem em *tandem* ao longo do genoma, formando sítios altamente polimórficos, o que possibilita o seu uso como marcas moleculares. Como marcadores, são codominantes, multialélicos, e aplicáveis em diversos tipos de estudos, principalmente no mapeamento genético. *Primers* que flanqueiam seqüências SSR geralmente são desenhados a partir da construção de bibliotecas genômicas, bibliotecas genômicas enriquecidas, seqüências depositadas em bancos de dados e, alternativamente, a partir de seqüências internas simples repetidas (ISSR). Geneticistas moleculares têm desenvolvido marcadores SSR com o intuito de mapear genes que codificam determinadas características de interesse. Entretanto, não existe um mapa consenso saturado para o feijão que sirva como referência para auxiliar na construção de mapas específicos. Nesta perspectiva, o Programa de Melhoramento Genético do Feijoeiro do BIOAGRO/UFV desenvolveu uma população de RIL's que poderá ser usada para integrar, em um único mapa, todos os marcadores SSR já desenvolvidos. No entanto, para a saturação do mapa, há necessidade de um grande número de marcadores. O presente trabalho teve o objetivo de desenvolver e validar *primers* que amplifiquem regiões contendo microssatélites a partir da metodologia da construção de bibliotecas genômicas enriquecidas e a partir de ISSR. Na primeira metodologia, em trabalho anterior, foram construídas duas bibliotecas genômicas enriquecidas para seqüências SSR. No presente trabalho, a partir das bibliotecas genômicas desenvolvidas, foram selecionados 207 clones contendo insertos de tamanho desejado. Destes, foram seqüenciados 196 (94,68%), dos quais 184 (88,89%) puderam

ser analisados, sendo 133 clones da biblioteca 1 e 51 da biblioteca 2. Foram detectados 48 (26,09%) clones redundantes. A análise dos clones permitiu identificar 66 (49,62%) motivos SSR na biblioteca 1 e 20 (39,22%) na biblioteca 2, a partir dos quais foram desenhados 56 pares de *primers*. Destes, 34 tiveram suas condições de amplificação otimizadas e padronizadas e foram testados quanto ao polimorfismo detectado entre 20 genótipos andinos e mesoamericanos, incluindo os genitores AND277 e Rudá. Todos os *primers* geraram produtos de amplificação e seis (17,65%) amplificaram produtos polimórficos entre os genótipos testados. Em relação à metodologia de enriquecimento por ISSR-PCR foram selecionados 250 clones contendo insertos com tamanho desejado, obtidos a partir da amplificação por ISSR-PCR, clonagem dos fragmentos e transformação de células competentes. Dos 250 clones, 168 (67,2%) foram sequenciados e 103 (41,2%) puderam ser analisados. Foram detectados 30 clones redundantes (29,13%). A análise das sequências permitiu identificar 58 motivos microssatélites (56,31%) e foi possível o desenho de 32 pares de *primers*. Destes, 10 tiveram suas condições de amplificação padronizadas e foram analisados quanto ao polimorfismo detectado entre os mesmos 20 genótipos andinos e mesoamericanos utilizados na metodologia de bibliotecas genômicas enriquecidas. Dos 10 pares de *primers* testados, seis comportaram-se como marcadores codominantes e quatro como dominantes. Dos codominantes nenhum mostrou-se polimórfico dentre os genótipos testados. Adicionalmente, as sequências contendo motivos microssatélites, obtidas a partir das duas metodologias utilizadas, foram submetidas à busca por similaridade com sequências já caracterizadas em bancos públicos de sequências. Foi identificada similaridade com regiões transcritas e não traduzidas, e com regiões codificadoras de proteínas, a partir do genoma nuclear, mitocondrial e do cloroplasto, e também a partir de sequências advindas de retrotransposons. As duas metodologias utilizadas foram eficazes para a seleção, no feijoeiro, de seqüências contendo microssatélites. Estes resultados representam um primeiro esforço no sentido de selecionar marcadores moleculares que serão futuramente mapeados na população consenso de RILs, além de fornecer marcadores que poderão ser usados nos mais variados tipos de estudos genéticos, contribuindo de fundamental maneira para o aprimoramento dos programas de melhoramento do feijoeiro comum que utilizam marcadores moleculares.

ABSTRACT

TANURE, Janaína Paula Marques, M.Sc., Universidade Federal de Viçosa, August 2009. **Development and validation of microsatellite markers for the common bean.** Adviser: Everaldo Gonçalves de Barros. Co-advisers: Márcia Regina Costa and Maurílio Alves Moreira.

Common bean (*Phaseolus vulgaris* L.) is a crop of great nutritional, economical and social importance. Breeding programs use molecular markers as important auxiliary tools for various types of genetic studies. Different classes of molecular markers have been developed, and among them microsatellites highlight. Microsatellites are DNA simple sequence repeats (SSR) distributed in tandem along the genome, forming highly variable polymorphic sites, enabling their use as molecular markers. SSRs are codominant, multiallelic, and thus can be used in several types of studies, mainly for genetic mapping. Primers flanking microsatellite sequences are commonly developed from genomic libraries, enriched genomic libraries, sequences obtained from databases and, alternatively, from internal simple sequence repeats (ISSR-PCR). Common bean breeding molecular geneticists have developed SSR markers for mapping specific traits of interest. However, a saturated consensus genetic map for common bean has not been established so far that can be used as a reference for the development of specific maps. Therefore, the BIOAGRO/UFV common bean breeding program developed a population of recombinant inbred lines (RILs) which is suggested to be used to integrate, in one single map, all microsatellite markers that have been developed so far. However, to saturate the map, there must be a great number of markers available. The objective of the present study was to develop and validate primers that amplify microsatellites sequences obtained from enriched genomic libraries and from ISSR sequences. In the first case, two enriched genomic libraries for microsatellite sequences, that had been developed in a previous study, were used. In the present study, 207 clones were selected from these two genomic libraries. One hundred and ninety six of these clones (94.68%) were sequenced and 184 (88.89%) of them could be analyzed, 133 clones from library 1 and 51 from library 2. Forty eight redundant clones (26.09%) were detected. Clone analysis led to the identification of 66 (49.62%) microsatellite motifs in library 1 and 20 (39.22%) in library 2, and 56 primer pairs

were designed. From the 56 primer pairs developed, 34 were characterized and tested in 20 Mesoamerican and Andean genotypes, including AND277 and Rudá. All the primer pairs were able to generate PCR products and six (17.65%) generated polymorphic DNA bands among the tested genotypes. In the ISSR enrichment methodology, 250 clones were selected with sizes over 400 bp. From these 250 clones, 168 (67.2%) were sequenced and 103 (41.2%) could be analyzed. Thirty redundant clones (29.13%) were detected. Clone analyses led to the identification of 58 microsatellite motifs (56.31%) and 32 primer pairs were developed. Out of these, 10 were characterized and tested in the same genotypes used in the previous methodology. Out of the 10 primer pairs tested, six were identified as codominant markers and the other four as dominant. The codominant markers revealed no polymorphisms among the tested genotypes. Additionally, microsatellite containing sequences obtained from both methodologies were submitted to BLAST analysis against sequences deposited in public databases. Similarity was identified between the SSR sequences and transcribed and non-transcribed regions, from nuclear, mitochondrial and chloroplast genomes, and also with retrotransposon sequences. Both methodologies were effective for selecting, in common bean, sequences that contain microsatellites. The results obtained represent an initial effort to select molecular markers that will be mapped in the future RIL's consensus population, contributing for the construction of a saturated genetic map for the species. In addition, these primers can be used in different types of genetic studies which are important for common bean breeding programs that use molecular markers.

1. INTRODUÇÃO GERAL

O feijão-comum (*Phaseolus vulgaris* L.) é um dos alimentos básicos da população mundial, sendo a leguminosa mais consumida mundialmente. O feijão é uma importante fonte de proteínas (~22% de sua constituição), vitaminas (folatos) e minerais (Ca, Cu, Fe, Mg, Mn, Zn) na dieta humana, especialmente em países em desenvolvimento (MIKLAS, 2006). O gênero *Phaseolus* compreende aproximadamente 55 espécies, das quais apenas cinco são cultivadas para o consumo humano: o feijoeiro comum (*Phaseolus vulgaris*); o feijão de lima (*P. lunatus*); o feijão Ayocote (*P. coccineus*); o feijão tepari (*P. acutifolius*); e o *P. polyanthus*. Desses, o feijão comum é mais cultivado mundialmente (EMBRAPA ARROZ E FEIJÃO, 2009) e consiste de basicamente dois *pools* gênicos, o Mesoamericano e o Andino, caracterizados por atributos morfológicos e padrão da proteína faseolina nas sementes.

No Brasil, a média anual de consumo de feijão por pessoa, atualmente, é de 12,7 kg. (EMBRAPA ARROZ E FEIJÃO, 2009). O feijão-comum é cultivado ao longo de todo o ano, na maioria dos estados brasileiros, proporcionando constante oferta do produto no mercado, e caracteriza-se como uma cultura de destacada importância nutricional, econômica e social. No âmbito econômico e social, sua relevância se dá pelo fato de ser cultivado tanto por pequenos agricultores quanto por agricultores altamente tecnificados. Pelas razões acima, o país se destaca como maior consumidor e produtor mundial desta leguminosa (EMBRAPA ARROZ E FEIJÃO, 2009). No entanto, embora a produção seja alta, ela está abaixo das potencialidades da cultura.

Para uma cultura cujo potencial produtivo pode chegar a 4.000 Kg por hectare, a produtividade média nacional ainda é pouco expressiva, com rendimento de apenas 725 kg/ha na segunda safra de 2008/2009 e de 967 kg/ha na terceira safra (CONAB, 2009). Diversos fatores abióticos e bióticos limitam a produtividade do feijão, tais como a seca, solos com baixa fertilidade e o acometimento da cultura por diversos tipos de pragas e doenças. Além disso, outros fatores como o cultivo em consórcio, o plantio com baixo nível tecnológico, a desorganização do mercado e a falta de integração dos diversos elos na cadeia produtiva contribuem muito para a baixa produtividade obtida (COSTA, 2007).

Dentro deste cenário, programas de melhoramento envolvendo a cultura estão bem estabelecidos, tanto em países desenvolvidos quanto naqueles em desenvolvimento, e visam à busca por cultivares produtivos, resistentes a pragas e doenças e tolerantes a estresses abióticos. Diversas ferramentas assistem a estes programas, dentre elas, os marcadores moleculares, os quais têm sido utilizados com êxito em estudos de diversidade genética, genotipagem molecular, reconstrução filogenética, mapeamento genético, seleção assistida por marcadores, dentre diversos outros. No Brasil, várias instituições desenvolvem importantes programas de melhoramento do feijoeiro, a exemplo da Universidade Federal de Viçosa (UFV), da Embrapa Arroz e Feijão, da Empresa de Pesquisa Agropecuária de Minas Gerais (Epmig), da Universidade Federal de Lavras (UFLA), da Universidade de São Paulo (Escola Superior de Agricultura Luiz de Queiroz) e do Instituto Agrônomo de Campinas (IAC). No entanto, no que diz respeito ao uso de marcadores moleculares no melhoramento, a falta de troca de informações tem dificultado grandes avanços, principalmente na consolidação de um mapa consenso do feijoeiro, com elevado grau de saturação, a partir de marcadores de grande poder informativo e reprodutíveis, como os microssatélites.

Além dos marcadores microssatélites, outras classes de marcadores foram extensivamente usadas em estudos de mapeamento genético do feijoeiro, como os marcadores RFLPs (NODARI et al. 1993;), RAPDs (KELLY, 2003) e AFLPs (PALLOTTINI et al., 2004; TA'RAN et al., 2002). No entanto, tais tipos de marcadores apresentam muitas desvantagens quando comparados aos marcadores microssatélites.

Microssatélites ou SSR (*Simple Sequence Repeats*) são seqüências simples repetidas de DNA, encontradas dispersas ao longo do genoma de organismos eucarióticos, formando sítios com alto grau de variabilidade (LITT e LUTY, 1989). A constatação de microssatélites em diferentes espécies de plantas e o grande polimorfismo apresentado sugeriu o potencial uso dos mesmos como marcadores. Tais marcadores apresentam várias vantagens em relação a outras classes de marcadores. Além da natureza multialélica, são codominantes, altamente reprodutíveis e permitem ampla cobertura do genoma (CAIXETA et al., 2009).

Devido à grande importância dessa classe de marcadores, vários *primers* microssatélites estão sendo desenvolvidos para o feijão,

tradicionalmente, a partir de sequências depositadas em bancos de dados públicos e a partir de bibliotecas genômicas enriquecidas. Os primeiros *primers* microsatélites para o feijão, baseados em sequências obtidas de bancos públicos, foram desenvolvidos por Yu et al. (1999, 2000), somando um total de 38 pares de *primers*; seguidos por Blair et al. (2003) e Guerra-Sanz (2004), que desenvolveram 57 e 20 *primers*, respectivamente. Masi et al. (2003) desenvolveram o primeiro sistema multiplex SSR-PCR para o feijão, com *primers* desenvolvidos a partir de sequências de bancos públicos, e padronizaram as condições de um sistema semi-automatizado baseado em fluorescência, para genotipagem em seqüenciador automático. Já os primeiros *primers* desenvolvidos a partir de bibliotecas genômicas enriquecidas foram primeiramente obtidos por Gaitán et al. (2002), Métais et al. (2002) e Yaish e Pérez De La Vega (2003), resultando no desenvolvimento de 68, 15 e 21 pares de *primers*, respectivamente. Desde então, Caixeta et al. (2005) usaram bibliotecas de cromossomos artificiais de bactéria (BACs) para desenvolver 28 pares de *primers* SSR, e Buso et al. (2006) e Benchimol et al. (2007) desenvolveram um grupo de 20 e 123 marcadores, respectivamente, a partir de bibliotecas enriquecidas. Mais recentemente, Blair et al. (2008) obtiveram 85 pares de *primers* a partir de bibliotecas enriquecidas para motivos AT e ATA; L'taief et al. (2008) desenvolveram 18 *primers* por meio de seleção de colônias positivas para motivos SSR, via hibridização com sondas SSR; e Hanai et al. (2007) desenvolveram 120 marcadores a partir de bibliotecas enriquecidas e 40 a partir de bibliotecas de cDNA obtidas a partir do *Bean EST Project* (BEST, 2007). Tais marcadores estão sendo utilizados para os mais diversos tipos de estudos tais como a identificação de genótipos, caracterização de germoplasma comercialmente cultivado e de germoplasma silvestre, análises filogenéticas e, principalmente, mapeamento genético.

Os mapas genéticos obtidos 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 e GRATTAPAGLIA,1998). Sendo, portanto, de fundamental importância para os programas de melhoramento do feijoeiro.

A primeira população usada para construção de um mapa núcleo para o feijoeiro foi obtida por FREYRE et al. (1998), sendo constituída de 79 RIL's resultantes do cruzamento entre os genótipos BAT93 e Jalo EEP558 (BJ). A população BJ apresenta alto nível de diversidade, sendo contrastante para várias características de interesse, pois seus genitores originaram-se dos *pools* gênicos andino e mesoamericano, respectivamente. Porém, tal população não é ideal para estudos de mapeamento, devido ao seu baixo número de indivíduos. Estudos de simulações computacionais concluíram que mapas com sérias distorções são obtidos com o uso de populações de tamanhos insuficientes, mesmo com grande quantidade de marcas. O tamanho mínimo de uma população de RIL ideal para a obtenção de mapas com boa confiabilidade deveria ser de 200, 300 e 500 indivíduos nos casos de saturação alta, mediana e baixa, respectivamente (SILVA, 2007). Desde a publicação dos trabalhos de FREYRE et al. (1998), estudos de mapeamento genético foram realizados na população BJ e estão disponíveis na literatura (YU et al., 2000; BLAIR et al., 2003; PEDROSA, 2003; BLAIR et al., 2006; OCHOA et al., 2006; GRISI et al., 2007; BLAIR et al., 2008). Tradicionalmente tal população tem sido usada como população núcleo para estudos de mapeamento, por ter sido a primeira desenvolvida para o feijão-comum, visando tais tipos de estudos, e principalmente pela indisponibilidade de outra de tamanho e características adequadas que pudesse servir como uma população base para a construção de um mapa consenso saturado para a cultura do feijoeiro. O atual mapa de ligação de referência da cultura ainda é muito 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 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 (HANAI et al., 2007).

Por esse motivo, o Programa de Melhoramento do Feijoeiro do BIOAGRO-UFV desenvolveu uma população de 600 RIL's derivada do cruzamento entre o genótipo andino AND277 e o mesoamericano Rudá, conduzida pelo método SSD (dados não publicados). Tal população constitui-

se, na atualidade, na maior população de RIL's já desenvolvida para o feijoeiro comum, para fins de mapeamento. O Programa sugere que a nova população de RIL's seja utilizada como uma população base para a construção de um mapa genético consenso para a espécie de *Phaseolus vulgaris*. Para a saturação do mapa genético da cultura, o que é fundamental para o aprimoramento dos programas de melhoramento, é essencial que cada vez mais marcadores reprodutíveis e informativos como os microssatélites sejam desenvolvidos. E, apesar de todos os estudos realizados, o número de marcadores microssatélites disponíveis para a cultura ainda é pequeno, quando comparado com outras culturas de grande importância social e econômica, como a soja e o arroz (BENCHIMOL et al., 2007; BUSO, 2006).

Assim sendo, o presente trabalho teve o objetivo de desenvolver e validar marcadores baseados em microssatélites, com o intuito de fornecer ferramentas que possam contribuir para a consolidação de um mapa consenso saturado para o feijoeiro, além de prover marcadores que poderão ser úteis para os mais diversos tipos de estudos genéticos, e assim auxiliar os programas de melhoramento dessa cultura.

2. OBJETIVOS

2.1. Objetivo Geral

O presente trabalho teve o objetivo de desenvolver e validar *primers* que amplifiquem regiões contendo motivos microssatélites em *Phaseolus vulgaris*, a partir de bibliotecas genômicas enriquecidas e a partir de sequências internas simples repetidas (ISSR).

2.2. Objetivos Específicos

1. Identificar motivos microssatélites e desenhar *primers* flanqueadores em sequências de DNA obtidas a partir da:
 - Análise de duas bibliotecas genômicas enriquecidas para os motivos microssatélites (TC)₁₅, (GT)₁₅, (CTT)₁₀ e para os motivos (AGG)₁₀, (CCA)₁₀, (GGC)₁₀, construídas em trabalho anterior;
 - Análise de biblioteca enriquecida com fragmentos contendo motivos microssatélites obtida a partir de sequências internas simples repetidas (ISSR), construída no presente trabalho;
2. Identificar possíveis funções celulares relacionadas às sequências que apresentaram motivos SSR, por meio de busca de similaridade com sequências já caracterizadas em bancos públicos de sequências;
3. Selecionar *primers* polimórficos entre 20 genótipos andinos e mesoamericanos de *Phaseolus vulgaris*.

3. REFERÊNCIAS BIBLIOGRÁFICAS

- BENCHIMOL L. L., CAMPOS, T., CARBONELL, S. A. M., COLOMBO, C. A., CHIORATTO, A. F., FORMIGHIERE, E. F., GOUVÊA, L. R. L., SOUZA, A. P. (2007). **Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers.** Genetic Resource and Crop Evolution, 54: 1747-1762.
- BLAIR, M. W., PEDRAZA, F., BUENDIA, H. F., GAITÁN-SOLÍS, E., BEEBE, S. E., GEPTS, P., TOHME, J. (2003). **Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.).** Theoretical and Applied Genetics, 107: 1362-1374.
- BLAIR, M.W., IRIARTE, G., BEEBE, S. (2006). **QTL analysis of yield traits in a advanced backcross population population derived from a cultivated Andean x wild common bean (*Phaseolus vulgaris* L.) cross.** Theoretical and Applied Genetics, 112: 1149-1163.
- BLAIR, M. W., BUENDÍA, H. F., GIRALDO, M. C., MÉTAIS, I., PELTIER, D. (2008). **Characterization of AT-rich microsatélites in common bean (*Phaseolus vulgaris* L.).** Theoretical and Applied Genetics, 118: 91-103.
- BEST. Bean EST Project. Disponível em <http://lgm.esalq.usp.br/BEST>. Acesso em Julho/2009
- CAIXETA, E. T., OLIVEIRA, A. C. B., BRITO, G. G., SAKIYAMA, N. S. (2009). **Tipos de marcadores moleculares. In: Marcadores Moleculares.** BORÉM, A., CAIXETA, E. T. 2º Edição. Viçosa/MG: Editora UFV. 532 p.
- BUSO G. S. C., AMARAL, Z. P. S., BRONDANI, R. P. V, FERREIRA, M. E. (2006). **Microsatellite markers for the common bean *Phaseolus vulgaris*.** Molecular Ecology Notes, 6: 252–254.
- CAIXETA, E.T., BORÉM, A., KELLY, J.D. (2005). **Development of microsatellite markers based on BAC common bean clones.** Crop Breeding and Applied Biotechnology, 5(2):125-133.
- CONAB. Companhia Nacional de Abastecimento. Indicadores da Agropecuária, Safra 2008/09. Disponível em: http://www.conab.gov.br/conabweb/download/safra/9_levantamento_jun2009.pdf Acesso em Junho/2009.
- COSTA, M.R. (2007). **Melhoramento de feijões preto e vermelho visando a resistência à antracnose, mancha-angular e ferrugem, com o auxílio de marcadores moleculares.** Viçosa: UFV. 88 p. (Tese-Doutorado).
- EMBRAPA ARROZ E FEIJÃO. Disponível em: <http://www.cnpaf.embrapa.br/>. Acesso em Junho/2009

- FERREIRA, M.E., GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética.** Brasília/DF: EMBRAPA/CENARGEM, 1998. 220 p.
- FREYRE, R.P., SKROCH, V., GEFFROY, A.F., *et al.* (1998). **Towards an integrated linkage map of common bean: 4. Development of a core linkage map and alignment of RFLP maps.** Theoretical and Applied Genetics, 97: 847-856.
- GAITÁN-SOLÍS, E., DUQUE, M.C., EDWARDS, K.J., TOHME, J (2002). **Microsatellite repeats in common bean (*Phaseolus vulgaris*): Isolation, Characterization, and Cross-Species Amplification in *Phaseolus* spp.** Crop Science, 42: 2128-2136.
- GRISI, M.C.M., BLAIR, M.W., GEPTS P., BRONDANI, C., PEREIRA, P.A.A., BRONDANI, R.P.V. (2007). **Genetic map of a new set of microsatellite markers in a reference common bean (*Phaseolus vulgaris*) population BAT93 X Jalo EEP558.** Genetic and Molecular Research, 6: 691-706.
- GUERRA-SANZ, J. M. (2004). **New SSR markers of *Phaseolus vulgaris* from sequence databases.** Plant Breeding, 123: 87-89.
- HANAI, L. R., CAMPOS, T., CAMARGO, L. E. A., BENCHIMOL, L. L., SOUZA, A. P., MELOTTO, M., CARBONELL, S. A. M., CHIORATTO, A. F., CONSOLI, L., FORMIGHIERI, E. F., SIQUEIRA, M. V. B. M., TSAI, S. M., VIEIRA, M. L. C. (2007). **Development, characterization, and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources.** Genome, 50: 266-277.
- KELLY, J. D., GEPTS, P., MIKLAS, P.N., COYNE, D. P. (2003). **Tagging and mapping of genes and QTL and molecular marker-assisted selection for traits of economic importance in bean and cowpea.** Field Crops Research, 32: 135-154.
- LITT, M., LUTY, J.A (1989). **A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene.** American Journal of Human Genetics, 44: 398-401.
- L'TAIEF, B., HORRES, R., JUNGSMANN, R., MOLINA, C., SIFI, B., LACHAËL, M., WINTER, P., KAHL, G. (2008). **Locus-specific microsatellite markers in common bean (*Phaseolus vulgaris* L.): isolation and characterization.** Euphytica, 162: 301-310.
- MASI, P., ZEULI, P. L. S., DONINI, P. (2003). **Development and analysis of multiplex microsatellite markers set in common bean (*Phaseolus vulgaris* L.).** Molecular Breeding, 11: 303-313.
- MÉTAIS, I., HAMON, B., JALOUZOT, R., PELTIER, D. (2002). **Structure and level of genetic diversity in various bean types evidenced with microsatellites markers isolated from a genomic enriched library.** Theoretical and Applied Genetics, 104: 1346-1352.

- MIKLAS, P.N., KELLY, J.D., BEEBE, S.E., BLAIR, M.W. (2006). **Common bean breeding for resistance against biotic and abiotic stresses: From classical to MAS**. *Euphytica*, 147: 234-230
- NODARI, R.O., TSAI, S.M., GILBERTSON, R.L., GEPTS, P. (1993). **Towards an integrated linkage map of common bean 2. Development of an RFLP-based linkage map**. *Theoretical and Applied Genetics*, 85: 513-520.
- OCHOA, I. E., BLAIR, M.W., LYNCH, J.P. (2006). **QTL analysis of adventitious root formation in common bean (*Phaseolus vulgaris* L.) under contrasting phosphorus availability**. *Crop Science*, 46: 1609-1621.
- PALLOTTINI, L., GARCIA, E., KAMI, J., BARCACCIA, G., GEPTS, P. (2004). **The genetic anatomy of a patented yellow bean**. *Crop Science*, 44: 968-977.
- PEDROSA, A., VALLEJOS, C.E., BACHMAIR, A., SCHWEIZER, D. (2003). **Integration of common bean (*Phaseolus vulgaris* L.) linkage and chromosomal maps**. *Theoretical Applied Genetics*, 106: 205-212.
- SILVA, L. C., CRUZ, C. D., MOREIRA, M. A., BARROS, E. G. (2007). **Simulation of population size and genome saturation level for genetic mapping of recombinant inbred lines (RILs)**. *Genetics and Molecular Biology*, 30: 1101-1108.
- TA'RAN, B., MICHAELS, THOMAS E., PAULS, K.P. (2002). **Genetic mapping of agronomic traits in common bean**. *Crop Science*, 42: 544-556.
- YAISH M.W.F., VEGA, M. P. D. L. (2003). **Isolation of (GA)_n microsatellite sequences and description of a predicted MADS-box sequence isolated from common bean (*Phaseolus vulgaris* L.)**. *Genetics and Molecular Biology*, 26: 337-342
- YU, K., PARK, S.J., POYSA, V. (1999). **Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*)**. *Genome*, 42: 27-34.
- YU, K., PARK, S. J., POYSA, V., GEPTS, P. (2000). **Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.)**. *The Journal of Heredity*, 91: 429-434.

CAPÍTULO I

DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES PARA O FEIJÃO-COMUM A PARTIR DE BIBLIOTECAS GENÔMICAS ENRIQUECIDAS

RESUMO

Microsatélites (SSR) são repetições em *tandem* de 1 a 6 pb que se estendem ao longo do genoma, formando sítios altamente polimórficos. Como marcadores apresentam várias vantagens e são utilizados em diversos estudos. *Primers* SSR geralmente são desenhados a partir de sequências depositadas em bancos de dados, do desenvolvimento de bibliotecas genômicas e de bibliotecas genômicas enriquecidas. Para o feijoeiro, a última é a metodologia mais utilizada. O presente trabalho teve o objetivo de desenvolver e validar marcadores SSR, a partir de bibliotecas genômicas enriquecidas. No presente trabalho, a partir de duas bibliotecas enriquecidas desenvolvidas, foram selecionados 207 clones, dos quais 184 (88,89%) foram analisados, sendo 133 clones da biblioteca 1 e 51 clones da biblioteca 2. A análise dos clones permitiu identificar 66 (49,62%) motivos SSR na biblioteca 1 e 20 (39,22%) motivos SSR na biblioteca 2, a partir dos quais foram desenhados 56 pares de *primers*. Destes, 34 tiveram suas condições de amplificação padronizadas, e seis (17,65%) amplificaram fragmentos polimórficos entre 20 genótipos andinos e mesoamericanos. Adicionalmente, as sequências contendo motivos microsatélites foram submetidas à busca por similaridade com sequências já caracterizadas em bancos públicos de sequências. Foi identificada similaridade com regiões transcritas e não traduzidas, e com regiões codificadoras de proteínas do genoma nuclear, mitocondrial e do cloroplasto, e com regiões de retrotransposons. Os resultados permitiram a constatação da presença de SSR em importantes regiões de funções conhecidas no genoma do feijão e o desenvolvimento de *primers* SSR para genoma de cloroplasto, estudo ainda pouco descrito na literatura para a cultura. Os marcadores microsatélites desenvolvidos no presente trabalho poderão ser utilizados nos mais diversos tipos de estudos genéticos como importantes ferramentas auxiliares dos programas de melhoramento do feijoeiro.

1. INTRODUÇÃO

Marcadores genéticos são quaisquer características, processos bioquímicos ou fragmentos de DNA que permitam a distinção de indivíduos geneticamente diferentes. Com o desenvolvimento das técnicas de biologia molecular, surgiram diversos métodos de detecção de polimorfismo genético diretamente em nível de DNA e, assim, uma nova classe de marcadores. Classificados como marcadores de DNA, estes se baseiam na variação natural da seqüência de bases do DNA, apresentam nível de polimorfismo alto para cada loco estudado e não são afetados pela variação nas condições ambientais ou fisiológicas do indivíduo, sendo amplamente usados em estudos de variados organismos e para diversos tipos de análises genéticas. São vários os tipos de marcadores de DNA, cada qual com a sua particularidade.

Os primeiros marcadores de DNA desenvolvidos foram os RFLPs. Entretanto, com o desenvolvimento da técnica de PCR na década de 80 pelo pesquisador britânico Kary Mullis (MULLIS e FALOONA, 1987), uma nova geração de marcadores moleculares ganhou grande importância, sendo classificados como marcadores baseados em PCR, podendo ser citados, dentre eles, o RAPD, o AFLP e os microssatélites.

Microssatélites ou *Simple Sequence Repeats* (SSR) são repetições de aproximadamente 1 a 6 pb que se estendem ao longo do DNA, sendo encontrados tanto em genomas de eucariotos quanto em procaríotos (FIELD e WILLS, 1996; HANCOCK et al.,1996). Morgante e Oliveiri (1993) foram os primeiros pesquisadores a identificar seqüências microssatélites localizadas em genes de plantas. Quanto à distribuição, Varshney et al. (2005) e Morgante et al. (2002) estimaram a densidade de microssatélites em regiões expressas de várias espécies de cereais e encontraram, em média, um SSR a cada 6 Kb. Em uma pesquisa realizada no GenBank, Yu et al. (1999) examinaram 326 seqüências de DNA de feijoeiro comum, e identificaram um microssatélite, em média, a cada 7 kb de DNA.

Os microssatélites estão distribuídos ao longo de todo o genoma, sendo encontrados tanto em regiões codificadoras como em não-codificadoras. Tóth et al. (2000) avaliaram a abundância de microssatélites em vários grupos taxonômicos e compararam a sua distribuição em éxons, íntrons e regiões intergênicas. Eles observaram que, em todos os táxons, repetições de

trinucleotídeos e de hexanucleotídeos foram encontradas em maior quantidade em éxons, resultado semelhante ao encontrado por Morgante et al. (2002) e Gao et al. (2003). No feijoeiro comum, uma pesquisa realizada por Yu et al. (1999) revelou que os microssatélites mais comuns são aqueles constituídos por repetições dos dinucleotídeos AT/TA e CT/GA, sendo o primeiro predominante. Foram também encontradas repetições de tri e tetranucleotídeos, porém em menor abundância. A maioria das repetições de dinucleotídeos foram encontradas em regiões intergênicas ou de íntrons, ao passo que as de trinucleotídeos foram encontradas, em grande parte, em regiões codificadoras e as de tetranucleotídeos, em regiões intergênicas. Hanai et al. (2007), analisando 3126 sequências expressas do *Bean Est Project* (BEST, 2009), constataram que o dinucleotídeo mais freqüente nas sequências expressas também foi AG/TC, assim como observado por Morgante et al. (2002) para *Arabidopsis thaliana* e *Glycine max*. Em relação aos trinucleotídeos, GAA/CTT, CAT/GTA e TGG/ACC foram os mais encontrados.

A constatação de seqüências microssatélites em diferentes espécies de plantas sugeriu o potencial uso dos mesmos como marcadores. As seqüências de DNA que flanqueiam os microssatélites são geralmente conservadas entre os indivíduos de uma mesma espécie, permitindo o desenho de *primers* específicos que amplificam, via PCR, fragmentos contendo o DNA repetitivo. A técnica de PCR utilizando marcadores SSR permite a detecção das variações nesses locos repetitivos, revelando polimorfismo devido às diferenças no número de vezes (n) em que, por exemplo, um dinucleotídeo (AG)_n se repete. Essas variações no número de repetições constituem-se, em última análise, em variações no comprimento do segmento gerado por PCR e na separação desses fragmentos amplificados em gel de eletroforese. Deste modo, cada segmento amplificado com diferente tamanho representa um alelo diferente do mesmo loco.

Como marcadores, os SSR são multialélicos, codominantes, apresentam alto poder de resolução, e o desenho de *primers* específicos garante a alta repetibilidade dos resultados. A técnica associa a simplicidade e rapidez da PCR, além de permitir uma ampla cobertura do genoma (CAIXETA et al., 2009). Todas essas características fazem dos SSR ferramentas eficazes, não só para o mapeamento genético, mas também para estudos de ligação gênica, identificação e caracterização molecular de genótipos, proteção de variedades,

avaliação da pureza de sementes, conservação de germoplasma, reconstrução filogenética, estudos de diversidade genética, análise gênica e de locos para caracteres quantitativos (QTL), análise de *pedigree* e seleção assistida por marcadores. O alto nível de variação detectado reduz, inclusive, o número de marcadores requeridos para distinguir genótipos (CAIXETA et al., 2009).

Devido às diversas vantagens oferecidas pelos marcadores microssatélites, vários *primers* SSR têm sido desenvolvidos para um grande número de espécies de interesse comercial. Para esse fim, várias metodologias estão propostas na literatura.

Tradicionalmente, a estratégia de desenvolvimento de *primers* SSR envolve a construção de uma biblioteca genômica, hibridização com sondas contendo microssatélites, seqüenciamento dos clones hibridizados e desenho de *primers* flanqueando os microssatélites encontrados. No entanto, a etapa de seleção é trabalhosa e demorada, devido à baixa freqüência de clones contendo repetições microssatélites (CAIXETA et al. 2009), que varia, normalmente, de 12% a menos de 0,04% do total de clones analisados (ZANE et al., 2002). Outras metodologias alternativas vêm sendo sugeridas de modo a superar as limitações impostas pela técnica tradicional.

Uma das metodologias alternativas envolve a construção e utilização de bibliotecas genômicas enriquecidas para motivos SSR, inicialmente proposta por Ostrander et al. (1992). Esta estratégia consiste na digestão do DNA genômico usando uma combinação de enzimas de restrição, a seleção de fragmentos de tamanho desejáveis e a ligação dos fragmentos digeridos a um adaptador (HAMILTON et al., 1999). Em seguida, ocorre a hibridização seletiva dos fragmentos com sondas marcadas com biotina ou sondas aderidas a uma membrana de náilon, e a recuperação dos fragmentos hibridizados usando partículas magnéticas cobertas com estreptoavidina ou por várias lavagens para eluir o DNA associado à sonda. Essa metodologia envolve um número de etapas relativamente grande, no entanto, a eficiência de seleção dos clones contendo microssatélites é bem maior que no método tradicional, sendo a eficiência de clones positivos obtidos de 40 a 50% do total.

É possível também identificar microssatélites por meio da análise de seqüências de DNA depositadas em bancos públicos, como a associação dos bancos GenBank (EUA), EMBL (Europa), DDJ (Japão), acessível no endereço

eletrônico do NCBI (NCBI, 2009); ou a partir de bancos desenvolvidos a partir de consórcios públicos e privados específicos para cada cultura, tais como o Projeto Genolyptus e o Projeto Genoma Café, para o eucalipto e o café, respectivamente. (GENOLYPTUS, 2009; PROJETO GENOMA CAFÉ, 2009). A pesquisa a bancos de dados é menos laboriosa e demanda menos recurso financeiro do que a construção de bibliotecas genômicas. No entanto, a maior limitação do uso dessa estratégia, para a identificação e desenvolvimento de marcadores SSR, reside na baixa disponibilidade de seqüências depositadas, para algumas espécies, em bancos de dados, limitando, assim, o estudo apenas a espécies cujo genoma esteja bem caracterizado (CAIXETA et al., 2009).

Devido à sua grande aplicabilidade nos mais diversos tipos de estudos, diversos trabalhos são encontrados na literatura que descrevem o desenvolvimento de marcadores microssatélites para o feijão, tradicionalmente, por meio da busca em bancos de dados (YU et al., 1999; YU et al., 2000; BLAIR et al., 2003; MASI, 2003; GUERRA-SANZ et al., 2004), e da construção de bibliotecas genômicas enriquecidas (GAÍTAN-SOLÍS et al., 2002; MÉTAIS et al., 2002; YAISH e PÉREZ DE LA VEGA., 2003; BUSO et al., 2006; BENCHIMOL et al., 2007; HANAI et al., 2007; BLAIR et al., 2008; L'TAIEF et al., 2008), sendo a metodologia de construção de bibliotecas a mais frequentemente utilizada para a cultura do feijoeiro.

Microssatélites são usualmente considerados como marcadores de DNA neutros, do ponto de vista evolutivo, e sua localização é tido como randomicamente distribuída ao longo do genoma eucariótico (SCHLOTTERER e WIEHE, 1999). No entanto, vários estudos têm demonstrado a frequente presença de microssatélites em seqüências funcionais e gênicas (LI et al. 2004). Na tentativa de caracterização de locos microssatélites do genoma do feijoeiro, desenvolvidos a partir de bibliotecas genômicas enriquecidas, quanto à sua localização e possível função celular relacionada, trabalhos foram conduzidos primeiramente por Gaitán et al. (2002) que identificaram similaridade com vários genes MAD-box relacionados à resposta a diversos tipos de estresse e envolvidos no desenvolvimento de raiz do feijoeiro comum. Yaish and Perez de la Vega (2003) identificaram a presença de um SSR localizado imediatamente *upstream* ao códon da metionina de uma seqüência MAD-box da subfamília AGL2 de *Arabidopsis thaliana*. Mais recentemente, Benchimol et

al. (2007) identificaram e caracterizaram microssatélites que apresentaram similaridade com genes que codificam proteínas de ligação ao DNA e quinases, e com regiões gênicas do genoma do cloroplasto (plastoma) de *Phaseolus vulgaris*, algo não identificado anteriormente para a cultura do feijoeiro, caracterizando um estudo inédito.

Assim sendo, o presente trabalho teve o objetivo de desenvolver e validar marcadores microssatélites para o feijão, a partir de bibliotecas genômicas enriquecidas, e de caracterizar putativas funções relacionadas às regiões repetitivas identificadas, com o intuito de fornecer ferramentas que possam ser úteis em diversos tipos de estudos genéticos, contribuindo para a melhoria dos programas de melhoramento dessa cultura.

2. MATERIAL E MÉTODOS

Os procedimentos descritos nos tópicos 2.1. ao 2.9. foram desenvolvidos por MAFRA (2007), dentro do Programa de Melhoramento do Feijoeiro do BIOAGRO/UFV, e serão descritos por se tratar da metodologia empregada nas etapas iniciais do desenvolvimento do presente trabalho.

2.1. Material genético e Extração de DNA

O cultivar de feijão TO foi usado como fonte de DNA genômico para a construção das bibliotecas enriquecidas. Este cultivar é do *pool* gênico mesoamericano e teve origem no cruzamento de “Tanderette” (suscetível a antracnose) com a linhagem mexicana “Acapulco” (resistente a antracnose). Além disso, este cultivar possui o gene dominante *Co-4* de resistência à antracnose (FOUILLLOUX, 1976). Os cultivares andinos e mesoamericanos descritos na Tabela 1 foram utilizados para a análise de polimorfismo dos *primers* SSR desenvolvidos no presente trabalho. O DNA foi extraído de folhas jovens dos cultivares com base no método do CTAB descrito por Doyle e Doyle (1990), com modificações (COSTA, 2007).

Tabela 1: Genótipos andinos e mesoamericanos utilizados na análise de polimorfismo dos *primers* microssatélites desenvolvidos

Genótipo	Pool gênico	Genótipo	Pool gênico
AND 277	Andino	G 5686	Mesoamericano
RUDÁ	Mesoamericano	TO	Mesoamericano
JALO EEP 558	Andino	VC3	Mesoamericano
CAL 143	Andino	MÉXICO 54	Mesoamericano
BJ8	Andino	MAR 2	Mesoamericano
BOLON BAYO	Andino	BAT 332	Mesoamericano
BJ3	Andino	TALISMÃ	Mesoamericano
BJ6	Andino	DOM TIMÓTEO	Mesoamericano
JALO M665	Andino	AMENDOIM	Mesoamericano
G 11796	Andino	OURO NEGRO	Mesoamericano

2.2. Digestão do DNA genômico, Separação e Purificação dos fragmentos de DNA

Parte da metodologia utilizada por Mafra (2007) foi descrita por Queiroz (2004). Para a digestão do DNA genômico, foram testadas várias combinações de enzimas de corte raro e freqüente. A combinação escolhida foi aquela que produziu, preferencialmente, fragmentos de 200 a 1000 pb. As enzimas utilizadas foram *Rsal*, *AluI*, *EcoRI* e *NheI*. Foram usados na reação de digestão aproximadamente 50 µg de DNA genômico, 50 unidades (U) de cada enzima, tampão NEB 2 1X, 100 µg/mL de BSA (albumina sérico bovina) e água para completar o volume para 200 µL. O conteúdo da reação foi mantido a 37 °C por 12 h.

A separação e purificação dos fragmentos de interesse foram realizadas de acordo com Wang et al. (2004). Esta metodologia consiste em realizar três etapas de eletroforese com o intuito de, a cada corrida eletroforética, concentrar os fragmentos de tamanho desejado em uma área menor no gel de agarose, permitindo uma otimização da etapa posterior de purificação do DNA a partir do gel. Em seguida, os fragmentos foram então purificados com o kit Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, EUA), segundo recomendações do fabricante. A eficiência da reação de digestão foi confirmada por análise eletroforética de uma alíquota de 3 µL de amostra, em gel de agarose 1,2% / SB (hidróxido de sódio 10 mM, pH 8,5 ajustado com ácido bórico). Após a corrida, o gel foi fotodigitalizado, sob luz ultravioleta, utilizando o sistema Eagle Eye II (Stratagene, La Jolla, CA, EUA).

2.3. Remoção das extremidades coesivas e Desfosforilação dos fragmentos de DNA

A remoção das extremidades coesivas foi necessária porque nem todas as enzimas usadas para digerir o DNA produzem extremidades abruptas, que são necessárias para a ligação do adaptador universal SNX. A etapa de desfosforilação da extremidade 5' impede que os fragmentos se liguem uns aos outros.

A reação para a remoção das extremidades coesivas foi realizada com os fragmentos de DNA digeridos e purificados, 15 U da enzima MBN (*Mung*

Bean Nuclease), tampão MBN 1X e água deionizada para completar o volume final de 240 µL. Os reagentes foram transferidos para um microtubo, que foi mantido a 30 °C por 30 min, e, posteriormente, os fragmentos de DNA com extremidades abruptas foram purificados com o kit Wizard® SV Gel and PCR Clean-Up System (Promega), segundo recomendações do fabricante. A reação de desfosforilação foi realizada com os fragmentos de DNA com extremidades abruptas, 30 U da enzima CIP (*Calf Intestinal Alkaline Phosphatase*), tampão NEB 3 1X e água deionizada para completar o volume final. A reação foi preparada com a metade do volume de CIP (15 U) em um primeiro momento e permaneceu por 30 min a 37 °C e em seguida, por 15 min a 50 °C. Após esta etapa, foi adicionado o restante do volume de CIP (15 U) e a reação foi novamente incubada por 30 min a 37 °C e depois 15 min a 50 °C. Os fragmentos de DNA desfosforilados foram então purificados com o kit Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, EUA), segundo recomendações do fabricante.

2.4. Ligação do adaptador universal SNX aos fragmentos de DNA

2.4.1. Adaptador SNX

O adaptador SNX foi desenvolvido por Hamilton et al. (1999). Este adaptador possui um sítio de restrição para as enzimas *StuI* e *NheI*. A restrição do adaptador com as enzimas *StuI* e *NheI* origina sítios de clonagem com extremidades abruptas e coesivas, respectivamente, sendo que as extremidades abruptas são aquelas que se ligarão aos fragmentos de interesse. A possível formação de dímeros entre os adaptadores cria um sítio para a enzima *XmnI*, no entanto, a adição desta enzima na reação de ligação favorece a ligação dos adaptadores preferencialmente aos fragmentos de interesse. Outra importante característica destes adaptadores é a presença de uma cauda poli-A em uma das extremidades 3'OH. A cauda poli-A polariza o adaptador de maneira que apenas uma extremidade possa servir como extremidade para ligação.

2.4.2. Ligação dos fragmentos ao adaptador SNX

A reação de ligação do adaptador aos fragmentos de DNA foi realizada em mistura contendo os fragmentos desfosforilados, adaptador SNX fita dupla 2 μ M, tampão NEB 2 1X, ATP 1mM, BSA (albumina sérico bovina) 100 μ g/mL, 10 U da enzima *Xmnl* e 30 unidades *Weiss* de T₄ DNA ligase. A mistura de ligação foi transferida para um termociclador modelo 9600 (Perkin Elmer-Cetus, Norwalk, CT, EUA), programado para 18 ciclos a 16 °C por 30 min e 37 °C por 10 min. Uma etapa final a 65 °C por 20 min foi realizada, para inativar as enzimas *Xmnl* e T₄ DNA ligase. O DNA foi armazenado a – 20 °C.

2.4.3. Confirmação da eficiência da reação de ligação

A eficiência da reação de ligação foi confirmada por PCR, utilizando o oligonucleotídeo SNX_{forward} como *primer*. A reação de amplificação foi feita com a mistura de ligação (10 ng/ μ L), tampão Phoneutria 1X, MgCl₂ 2 mM, *primer* SNX_f (0,8 μ M), cada um dos desoxirribonucleotídeos (dATP, dTTP, dCTP e dGTP) em uma concentração de 150 μ M e 1,5 U da enzima *Taq* Phoneutria. Os ciclos de amplificação foram programados para: uma etapa a 96 °C por 5 min; seguido de 40 ciclos a 96 °C por 45 s, 62 °C por 60 s e 72 °C por 90 s. Após os 40 ciclos, foi realizada uma etapa final a 72 °C por 7 min. A eficiência da reação de ligação foi confirmada por eletroforese dos produtos da reação em minigel de agarose 1,2% / SB 1X.

2.5. Reação de *nick-ligation*

A realização de reações de *nick-ligation* foi necessária para evitar a perda dos adaptadores em etapas posteriores do processo. A partir da etapa de desfosforilação dos fragmentos de DNA, a única possibilidade de formação da ligação fosfodiéster seria entre a extremidade 3' (-OH) do fragmento e a 5' (-PO₄⁻) do adaptador, assim, os fragmentos ficariam ligados em ambas as extremidades em apenas uma das fitas. O *nick* formado poderia resultar, nas etapas posteriores, na perda do adaptador em uma das extremidades dos fragmentos. Assim, foi necessário reparar este *nick* utilizando uma *Taq* DNA

polimerase (Promega) que possui atividade exonucleásica 5'→3', capaz de repará-lo.

Para sua realização, o conteúdo da reação de ligação foi dividido em alíquotas. Para cada alíquota, foi realizada a reação de *nick-ligation*, em mistura contendo tampão (Promega, Madison, WI, EUA) 1X, MgCl₂ 2 mM, BSA 25 µg/mL, cada um dos desoxirribonucleotídeos (dATP, dTTP, dCTP e dGTP) em uma concentração de 150 µM e 1 U da enzima *GoTaq* DNA polimerase (Promega, Madison, WI, EUA). As reações foram realizadas em um termociclador modelo 9600 (Perkin Elmer-Cetus, Norwalk, CT, EUA) programado para 72 °C por 30 min. O conteúdo de todas as reações foi reunido e purificado com o kit Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, EUA), segundo recomendações do fabricante.

2.6. Reação pós *nick-ligation*

O conteúdo da reação de *nick-ligation* foi dividido em alíquotas. Para cada alíquota, foi realizada a reação pós *nick-ligation* com tampão 1X (Promega, Madison, WI, EUA), MgCl₂ 2 mM, *primer* SNX_f (0,8 µM), cada um dos desoxirribonucleotídeos (dATP, dTTP, dCTP e dGTP) 150 µM e 2 U da enzima *GoTaq* polimerase (Promega, Madison, WI, EUA). Os ciclos de amplificação foram realizados em um termociclador modelo 9600 (Perkin Elmer-Cetus, Norwalk, CT, EUA) programado para: uma etapa a 96 °C por 5 min; seguido de 15 ciclos a 96 °C por 45 s, 62 °C por 60 s e 72 °C por 90 s. Após os 15 ciclos, foi realizada uma etapa final a 72 °C por 7 min. Os produtos da reação pós *nick-ligation* foram purificados com o Kit de purificação Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, EUA).

2.7. Hibridização dos fragmentos pós *nick-ligation* com oligonucleotídeos contendo microssatélites

Os fragmentos digeridos, purificados e ligados ao adaptador foram selecionados quanto à presença de microssatélites, por meio da metodologia de hibridização usando oligonucleotídeos biotinizados. Os oligonucleotídeos biotinizados usados foram (TC)₁₅, (GT)₁₅, (AGG)₁₀, (CCA)₁₀, (CTT)₁₀ e (GGC)₁₀.

Foram preparadas duas reações de hibridização, uma contendo as sondas (TC)₁₅, (GT)₁₅ e (CTT)₁₀ e outra com as sondas trinucleotídicas (AGG)₁₀, (CCA)₁₀ e (GGC)₁₀. Cada reação de hibridização continha 6 µL (600 ng) dos produtos da reação de amplificação, 50 µL do tampão de hibridização (SSC 12 X; SDS 0,1%), 2 pmols de cada oligonucleotídeo biotilado específico (1 µM) e água deionizada para completar o volume final de 100 µL. A mistura foi aquecida a 95 °C por 15 min. Em seguida, os tubos foram incubados por 14 h a 65 °C.

2.7.1. Preparação das partículas magnéticas

A solução contendo as partículas *Dynabeads* M-270[®] (DynaBiotech, Lake Success, NY, EUA) foi cuidadosamente agitada, a fim de obter uma suspensão homogênea. O volume de 6 µL da solução contendo as partículas (50 µg/µL) foi transferido para um microtubo de 1,5 mL. Os microtubos foram mantidos próximos a um ímã e, após 2 min, o sobrenadante da mistura foi removido. Aos mesmos microtubos, foram adicionados 100 µL do tampão de lavagem (Tris-HCl 10 mM, pH 7,5; EDTA 1 mM; NaCl 1 M). Em seguida, os microtubos foram afastados do magneto e as partículas foram cuidadosamente ressuspendidas. Os microtubos foram novamente aproximados do magneto e, após 2 min, o tampão de lavagem foi novamente removido. Este procedimento de lavagem das partículas foi repetido por três vezes. As partículas ressuspendidas foram divididas em alíquotas antes da última etapa de remoção do sobrenadante. Essas alíquotas foram transferidas para microtubos de 1,5 mL. Os microtubos foram aproximados do ímã e, após 2 min, o sobrenadante foi removido. A quantidade de partícula magnética coberta com estreptoavidina transferida para cada microtubo foi de 300 µg.

2.7.2. Captura dos fragmentos de DNA hibridizados

Os fragmentos de DNA hibridizados com os oligonucleotídeos biotilados foram transferidos para cada microtubo, contendo as partículas magnéticas. Os microtubos foram incubados a 37 °C por 1 hora, sendo cuidadosamente agitados a cada 15 min. Em seguida, os microtubos foram aproximados de um magneto e, após 2 min, o sobrenadante da mistura foi

descartado. Inicialmente, o conteúdo de todos os microtubos foi lavado duas vezes com 200 μ L de SSC 2X (NaCl 3M, Citrato de sódio 0,3M, pH 7,0) e SDS 0,1%, em temperatura ambiente em um intervalo de 5 min. Posteriormente, o conteúdo dos microtubos foram lavados duas vezes a 65 °C com 200 μ L de SSC 1X e SDS 0,1% em um intervalo de 15 min. No final de cada etapa de lavagem, o microtubo foi aproximado do magneto e, após 2 min, o sobrenadante da mistura foi removido. Após a remoção do sobrenadante da última etapa, foram adicionados 100 μ L de tampão TE 0,1M (Tris-HCl 10 mM, pH 8,0; EDTA 0,1 mM) pré-aquecido a 95 °C em todos os microtubos. Os microtubos foram incubados a 95 °C por 10 min e então aproximados ao magneto. O sobrenadante de cada um foi transferido para um microtubo novo e armazenado a – 20 °C.

2.7.3. Síntese de fragmentos de DNA fita dupla

Os fragmentos de DNA fita simples capturados por hibridização com as sondas foram convertidos em fita dupla utilizando o adaptador SNX *forward* como *primer*. Os ciclos de amplificação foram realizados em um termociclador modelo 9600 (Perkin Elmer-Cetus, Norwalk, CT, EUA), programado para: uma etapa a 96 °C por 5 min; seguida de 40 ciclos a 96 °C por 45 s, 62 °C por 60 s e 72 °C por 90 s. Após os 40 ciclos, foi realizada uma etapa final a 72 °C por 7 min. A eficiência da reação de amplificação foi confirmada por meio de análise eletroforética de uma alíquota em minigel de agarose 1,2% / SB 1X. Os fragmentos de DNA fita dupla foram purificados com o kit Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, EUA), segundo recomendações do fabricante. A concentração dos fragmentos de DNA foi estimada por espectrometria.

2.8. Preparação dos fragmentos de DNA e do plasmídeo

Os fragmentos de DNA enriquecidos com os diferentes tipos de microssatélites foram digeridos com 2 U de *NheI* em tampão NEB 2 1X., por 14 h a 37 °C no termociclador modelo 9600 (Perkin Elmer-Cetus, Norwalk, CT, EUA). A reação continha cerca de 400 ng dos fragmentos de DNA dupla fita, 2

U de *NheI*, tampão NEB 2 1X, BSA 100 µg/mL e água deionizada em um volume final de reação de 30 µL. Após a digestão, os fragmentos foram purificados com o Kit Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, EUA) e eluídos em 30 µL.

A reação de digestão de 3 µg do plasmídeo pBluescript SK (+) foi realizada com 10 U de *XbaI* em tampão NEB 2 1X. A mistura de reação permaneceu por 14 h a 37 °C no termociclador. O conteúdo de todas as reações foi purificado com o kit PCR Purification Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, EUA), segundo recomendações do fabricante. Após a purificação, o plasmídeo foi desfosforilado com 10 U da enzima CIP em tampão NEB 3 1X em um volume final de 50 µL a 37 °C, por 10 h no termociclador. Após a reação, o plasmídeo foi purificado, conforme descrito anteriormente. O plasmídeo desfosforilado foi ressuspendido em tampão EB (Tris-HCl 10 mM, pH 8,5) na concentração de 100 ng/µL. A eficiência da reação de digestão do plasmídeo foi confirmada por meio de análise eletroforética de uma alíquota de 3 µL de amostra em minigel de agarose 1,2% / SB 1X. Após a corrida, o gel foi fotodigitalizado, sob luz ultravioleta, utilizando o sistema Eagle Eye II (Stratagene, La Jolla, CA, EUA).

2.9. Ligação dos fragmentos de DNA ao vetor e Transformação de células competentes

A reação de ligação foi realizada com 15 µL dos fragmentos de DNA digeridos (cerca de 180 ng), tampão NEB 2 1X, ATP 0,8 mM, BSA 100 µg/mL, 150 ng do plasmídeo pBluescript SK (+) digerido e desfosforilado, 7 U de *NheI* e 6 U Weiss de *T₄* DNA ligase. O procedimento foi realizado em um termociclador modelo 9600 (Perkin Elmer-Cetus, Norwalk, CT, EUA), programado para 19 ciclos a 16 °C, por 30 min, e 37 °C por 10 min. Uma etapa final a 65 °C por 20 min foi realizada para inativar as enzimas *NheI* e *T₄* DNA ligase. O conteúdo da reação permaneceu à temperatura ambiente por 30 min e, em seguida, foi utilizado para transformação de células ultra-competentes da cepa DH5α de *Escherichia coli*. A transformação foi realizada por meio de choque térmico, de acordo com protocolo-padrão descrito por Ausubel (1998).

2.10. Seleção de clones positivos e Seqüenciamento

Após transformação por choque térmico, foi realizada a seleção de colônias brancas transformantes-recombinantes. A confirmação da presença do inserto foi feita por meio da técnica de PCR de colônia, utilizando-se os *primers* T3 e T7. Tais *primers* amplificam um fragmento do vetor de 170 pb e, por esse motivo, apenas as colônias que apresentaram um fragmento de no mínimo 400pb foram selecionadas para o seqüenciamento. O DNA plasmidial das colônias de interesse foi extraído com o Kit de extração Wizard® Plus SV Minipreps DNA Purification System, segundo recomendações do fabricante (Promega, Madison, WI, EUA). Após a extração e quantificação do DNA plasmidial, as amostras foram seqüenciadas em seqüenciador automático Applied Biosystems modelo 3730xl, na empresa Macrogen (Seoul, Coréia do Sul), pelo método de terminação da cadeia por didesoxirribonucleotídeos (SANGER et al., 1977).

2.11. Identificação de regiões SSR e Desenho de *primers*

As sequências que apresentaram valor PHRED de qualidade maior que 20, tiveram seu perfil cromatográfico analisados no programa CHROMAS (<http://www.technelysium.com.au/chromas>). O valor PHRED=20 está relacionado à probabilidade estimada de 1 erro a cada 100 bases seqüenciadas, ou seja, corresponde a 99% de certeza de que um nucleotídeo apresentado em uma dada posição no perfil cromatográfico corresponde à realidade. Esse valor é comumente padronizado na literatura como o valor PHRED mínimo para que uma sequência seja considerada de boa qualidade para análise. Adicionalmente, foram consideradas sequências de boa qualidade aquelas cujo cromatograma apresentou picos de fluorescência agudos, regularmente espaçados e com pouco ou nenhum ruído de fundo. Em seguida, foi utilizada a ferramenta VecScreem para identificar quais os segmentos das sequências de DNA eram correspondentes ao vetor plasmidial utilizado durante a etapa de clonagem. O VecScreem realiza uma análise de similaridade, por meio do algoritmo BLAST, entre a sequência de entrada e um banco de dados especializado e não-redundante de sequências de vetores usualmente utilizados em experimentos de clonagem, o UniVec, identificando,

na sequência de interesse, as regiões correspondentes ao vetor plasmidial. As três ferramentas descritas encontram-se disponíveis na página do NCBI (NCBI, 2009). Logo após, foi utilizado o programa MEGA4 (<http://www.megasoftware.net/mega.html>), para alinhamento global das sequências *forward* e *reverse* advindas do mesmo clone, e obtenção da sequência consenso. Quando o fragmento de DNA genômico clonado era grande, não era possível um alinhamento de boa qualidade das sequências *forward* e *reverse*, pois a fração das sequências que apresentavam sobreposição era geralmente pequena, o que dificultava a obtenção de uma sequência consenso. Nesses casos foi utilizado o programa CAP3 (<http://pbil.univ-lyon1.fr/cap3.php>), que realiza o agrupamento de sequências contíguas, ou seja, sequenciais, gerando *contigs* e fornecendo a sequência consenso. O programa CAP3 foi utilizado também para a identificação de clones redundantes. Seguiu-se a análise das sequências quanto à presença de motivos microssatélites, por meio da ferramenta SSRIT (<http://www.gramene.org/>). Não existe um critério padrão estabelecido na literatura para a definição de um potencial microssatélite, assim sendo, foi seguido o seguinte critério mínimo: A) $(NN)_5$ – 5 repetições de dinucleotídeos; B) $(NNN)_4$ - 4 repetições de trinucleotídeos e C) $(NNNN)_3$ – 3 repetições de tetranucleotídeos, segundo Caixeta et al. (2005). No entanto, alguns motivos menores foram igualmente selecionados por estarem inseridos em um motivo maior, porém composto.

Os *primers* microssatélites foram desenhados com auxílio do programa Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Para o desenho dos *primers* foram seguidos os critérios recomendados por Yu et al. (1999), com modificações: A) *Tamanho dos primers*: de 18 pb a 30 pb; 2) *Tm*: mínimo de 48 °C e máximo de 60 °C, com ótimo de 52 °C e *Tm* máxima de 5 °C entre os *primers forward* e *reverse* da mesma sequência ; 3) *Concentração de sal*: 50 mM; 4) % GC: mínimo de 20% e máximo de 80%; 5) *Produtos de amplificação*: entre 100 – 400 pb. Após o desenho dos *primers* foi verificada a estabilidade dos mesmos por meio do programa *Primerselect* 3.1 (DNASTAR, Madison, WI, EUA).

O *Primerselect* 3.1 analisa a possibilidade de formação de heterodímeros entre os *primers forward* e *reverse*, identifica a existência de auto-complementariedade e formação de homodímeros, e ainda identifica a possível

formação de estruturas secundárias (*hairpin*) durante a futura reação de PCR, por meio de padrões termodinâmicos pré-estabelecidos. O programa *Primerselect* considera que os seguintes fatores são indispensáveis para um *primer* ser considerado estável: 1) *Estabilidade do pentâmero 3'* - é um valor limite de estabilidade para os cinco últimos nucleotídeos do terminal 3'. *Primers* com o valor de energia livre (G) do pentâmero menor do que -8,5 kCal/mol apresentam uma tendência de pareamento em regiões inespecíficas e são mais prováveis de formar estruturas em grampo e dímeros com ele mesmo. O *primer* ideal deve possuir um gradiente decrescente de G que vai de alta estabilidade na extremidade 5' (G < -8,5 kCal/mol) até baixa estabilidade na extremidade 3' (G > -8,5 kCal/mol). 2) *Seqüência 3' única de 7 bases* - este critério limita o pareamento inespecífico, usando este critério, qualquer *primer* que não apresente uma seqüência específica de 7 bases ou superior que seja exatamente complementar ao sítio de anelamento é eliminado. 3) *Formação de dímeros* - este critério limita no valor 2 a 3 o número de pares de bases que o dímero pode conter. O valor 3 raramente altera o desempenho do *primer*, a menos que o pareamento seja próximo à extremidade 3'. 4) *Formação de estruturas em grampo* - este critério limita no valor 2 o número de pares de bases que a estrutura pode conter. O valor 3 raramente altera o desempenho do *primer*, a menos que o pareamento seja próximo à extremidade 3'. 5) *Formação de fita dupla próximo à extremidade 3'* - este critério elimina *primers* candidatos que apresentem qualquer dímero ou estrutura em grampo a uma distância de 8 pares de base da extremidade 3'.

2.12. Determinação das condições de amplificação e Identificação de marcadores microssatélites polimórficos entre os genótipos andinos e mesoamericanos

Os *primers* analisados no programa *PrimerSelect* e considerados estáveis termodinamicamente foram sintetizados por empresa especializada (Sinapse, São Paulo/SP, Brasil). Para a determinação e validação das condições ideais de amplificação de cada *primer* desenvolvido, os cultivares Rudá e AND277 foram utilizados como fonte de DNA para a realização de reações SSR-PCR, nas quais os *primers* foram submetidos a um gradiente de temperaturas de amplificação.

As reações de amplificação foram feitas em um volume de 15 µL, contendo 30 ng de DNA, Tris-HCl 12 mM (pH 8,3), KCl 60 mM, MgCl₂ 2 mM, cada um dos desoxirribonucleotídeos (dATP, dTTP, dGTP e dCTP) numa concentração de 100 µM, 0,3 µM de cada *primer* e 1U da enzima *Taq* DNA polimerase. O procedimento foi realizado em um termociclador Eppendorf modelo Mastercycler EP Gradiente, programado para uma etapa inicial a 94 °C por 4 min, seguida de 40 ciclos a 94 °C por 40 s, 48-62°C por 40 s e 72 °C por 60 s, e uma etapa final a 72 °C por 7 min. Os produtos de amplificação foram submetidos à eletroforese em gel de agarose 1,8%/SB 1X, corados com brometo de etídeo e fotodigitalizados utilizando o sistema L-PIX (Loccus Biotecnologia, São Paulo, Brasil).

Após a determinação da melhor temperatura de amplificação, os *primers* foram testados quanto ao polimorfismo detectado entre os cultivares descritos na Tabela 1 do item 2.1, genitores usados no Programa de Melhoramento do Feijoeiro do BIOAGRO/UFV. As reações de amplificação para cada um dos *primers* padronizados foram realizadas sob as mesmas condições usadas para as reações de padronização. As temperaturas de amplificação definidas e utilizadas no teste de polimorfismo, para cada *primer*, estão indicadas na Tabela 3 do item 3.1, na sessão “Resultados e Discussão”. Os produtos de amplificação foram submetidos à eletroforese em gel de poliacrilamida nativo 10%, contendo acrilamida/bis-acrilamida na proporção de 29:1; tampão TAE 10X (Tris-base 400 mM, ácido acético 200 mM e EDTA 10 mM); na presença de 167,5 µl de persulfato de amônio 10% e 167,5 µl do catalisador TEMED, para um volume final de 25ml de gel. Os géis foram submetidos a uma voltagem de 120 volts por 4 h e corados com solução de nitrato de prata, de acordo com o protocolo proposto por Sanguinetti et al. (1994), com modificação do tempo de imersão dos géis em cada uma das quatro soluções usadas no procedimento para 10 minutos.

2.13. Caracterização dos *primers* polimórficos quanto ao Índice de Informação de Polimorfismo (PIC) e número de alelos detectados

Para os *primers* que amplificaram produtos polimórficos entre os genótipos testados, foi calculado o Índice de Informação de Polimorfismo (PIC). O PIC foi calculado de acordo com a equação de Anderson et al. (1993):

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

Onde k é o número total de alelos detectados para um loco microssatélite e P_i é a frequência do i -ésimo alelo detectado nos vinte genótipos de feijão analisados.

O número de alelos em cada loco foi determinado por análise visual, comparando-se a mobilidade no gel dos fragmentos gerados com a dos marcadores de tamanho DNA Step Ladder 25 pb (Fermentas International Inc., Ontário, Canadá), em gel nativo de poliacrilamida 10%, nas mesmas condições descritas no item 2.12.

2.14. Análises de similaridade

As regiões microssatélites identificadas, e para as quais puderam ser desenhados *primers* flanqueadores, foram submetidas à busca por similaridade com sequências já caracterizadas e depositadas no consórcio de bancos públicos de sequências GenBank (EUA), EMBL (Europa) e DDJ (Japão), disponível no endereço eletrônico do NCBI (NCBI, 2009).

Primeiramente, foi utilizada a ferramenta *PrimerBLAST* para analisar se os *primers* desenvolvidos amplificavam sequências do genoma de *Phaseolus vulgaris* e de outras espécies que já haviam sido caracterizadas e depositadas no banco de sequências. Para tornar os parâmetros menos rigorosos selecionou-se a opção *nr (non-redundant)*, na seção *database*, que permite uma busca mais ampla, pois engloba todos os bancos de dados não redundantes disponíveis no NCBI (bancos curados e não curados, de sequências expressas e não-expressas, para todas as espécies e não somente para *Phaseolus vulgaris*).

Logo em seguida, as sequências amplificadas pelos *primers* SSR desenvolvidos, obtidas no item 2.10., foram submetidas a uma busca de similaridade contra o banco de dados completo do NCBI, via BLASTN, algoritmo que identifica similaridade entre a sequência de nucleotídeos de entrada (*query*) com as demais sequências de nucleotídeos depositadas. Quando a busca identificou similaridade com múltiplas sequências disponíveis no banco de dados, foi considerado o primeiro *hit*, ou seja, a sequência que

apresentou, em conjunto, a maior porcentagem de similaridade com a sequência de entrada (85%-100%), associada ao maior *score* de alinhamento e ao menor *e-value* (*e-value* máximo de e^{-10}). Todas as ferramentas acima descritas encontram-se disponíveis na página do NCBI (NCBI, 2009).

Para identificação e categorização das funções celulares associadas às sequências gênicas identificadas pela análise de similaridade, foi realizada uma pesquisa ao banco SwissProt (<http://www.expasy.ch/sprot>) e ao banco GO (*Gene Ontology*) (<http://www.geneontology.org/>), respectivamente.

3. RESULTADOS E DISCUSSÃO

3.1. Obtenção das bibliotecas genômicas enriquecidas

Como resultado dos procedimentos descritos nos tópicos 2.1 ao 2.9 da sessão “Material e Métodos”, realizados por MAFRA (2007), foram desenvolvidas duas bibliotecas genômicas enriquecidas para motivos microssatélites, a primeira para os motivos (TC)₁₅, (GT)₁₅ e (CTT)₁₀ (Biblioteca 1) e a segunda para os motivos (AGG)₁₀, (CCA)₁₀ e (GGC)₁₀ (Biblioteca 2). No presente trabalho, as duas bibliotecas desenvolvidas foram analisadas em busca de motivos microssatélites para os quais pudessem ser desenhados *primers* flanqueadores.

3.2. Seqüenciamento e Desenho de *primers* SSR

Após a etapa de transformação de células ultra-competentes, 400 colônias transformantes-recombinantes foram obtidas e 207 foram selecionadas por apresentarem fragmentos clonados acima de 400 pb, identificadas por meio do procedimento de PCR de colônias (Figura 1).

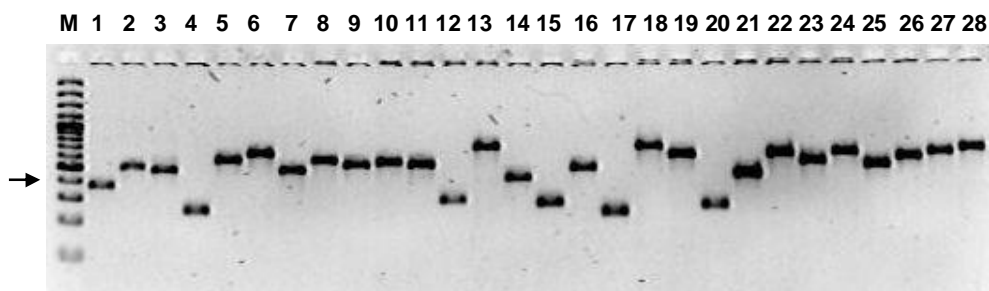


Figura 1: Perfil eletroforético do diagnóstico de colônias utilizando os *primers* T3 e T7. M: marcador de tamanho 100 pb; 1 a 28: clones diagnosticados. A seta indica a posição correspondente a 400 pb.

Das 207 colônias selecionadas, 196 (94,68%) foram sequenciados, dos quais 184 (88,89%) apresentaram perfil de sequenciamento de boa qualidade e puderam ser analisados, sendo 133 clones da biblioteca 1 e 51 clones da biblioteca 2. Foram detectados 48 (26,09%) clones redundantes, dentre o total de clones analisados nas duas bibliotecas. Foi possível identificar repetições perfeitas, imperfeitas e compostas, que variaram em tipo e em número (Figura

repetições e o número de alelos em um microsatélite. No entanto, Yu et al. (1999) encontraram microsatélites curtos que se mostraram polimórficos entre os genótipos testados, enquanto que outros mais longos, como (AT)₂₂, foram monomórficos. Os resultados obtidos por Caixeta et al. (2005) também mostraram que é possível encontrar polimorfismo mesmo entre motivos SSR relativamente pequenos como (TAA)₃. Benchimol et al. (2007) desenvolveram *primers* que amplificavam motivos microsatélites perfeitos pequenos como (CA)₅ e (GT)₅ e também motivos compostos de repetições menores tais como (TG)₄TT(TG)₄(AT)₃, (TG)₄T(TG) e (AC)₃(TC)₂, que se mostraram polimórficos entre os 20 genótipos analisados. Hanai et al., (2007) também chegaram a resultados semelhantes ao detectarem polimorfismo em motivos como (TC)₅, (CA)₅ e (CT)₂(GT)₃(AAG)₂(GT)₄. Por esta razão, no presente trabalho, motivos compostos por repetições menores foram igualmente considerados para o desenho de *primers*. Motivos repetitivos de 7-9 pb, definidos por Tautz e Renz (1984) como minissatélites, que se repetiam de duas a quatro vezes, foram igualmente considerados.

A denominação dos *primers* SSR foi feita de acordo com Queiroz (2004), com modificações. Consistiu primeiramente das letras PV de *Phaseolus vulgaris*, seguida da unidade repetitiva presente em maior número no motivo SSR e um número arbitrário a partir de 001 para sequências que possuam a mesma unidade repetitiva em maior número. Quando a sequência possuía mais de uma unidade com o mesmo número de repetições, foi considerada a primeira unidade repetitiva para efeito de nomenclatura. As designações F e R foram utilizadas para diferenciar os *primers forward* e *reverse*, respectivamente. Por exemplo, para o *primer* PVga001F, F designa o *primer* SSR *forward* flanqueando repetições do motivo GA. A Tabela 2 mostra os motivos microsatélites identificados e as sequências dos *primers* que foram desenhados.

Tabela 2: Motivos microsatélites identificados, sequências dos *primers* correspondentes, valores de T_m, tipo de repetição e tamanho dos fragmentos amplificados. O símbolo // entre alguns motivos denota a presença de outros motivos SSR relativamente distantes no mesmo fragmento.

Motivo microsatélite	Primer	Seqüência	T _m (°C)	Tam. frag. (pb)
(GA) ₃ (GGA) ₂ (GAA) ₂	PVga001 F	AAAGATAGAAGTGAAGAGAA	50,10	208

	PVga001 R	ACTTGTC AAGAGTCAAAGTG	50,00	
(GGC) ₂ (AG) ₃	PVag001 F	TCGATGACAGTCAGCAACAG	57,20	150
	PVag001 R	TGACACCGCATCACTATTCT	57,60	
(AG) ₃ ATA ₃ (GT) ₅ //(AG) ₈	PVag003 F	AGTGGACTCAGCTGCGAGAG	60,90	163
	PVag003 R	CACGCCAAATAAACCCATTC	60,10	
(AG) ₄ A(AG)(GA) ₃ // (AG) ₅	PVag004 F	AGGTA AAAATTGAGAGAACC	50,20	250
	PVag004 R	TTGAAACAGAAGAAAAACAT	50,00	
(TC) ₃ CG ₂ (GTT) ₂ T (GAG) ₂ A(AGG) ₂	PVtc001 F	CTAGTTTGAATGCTTTTGAC	50,20	372
	PVtc001 R	ACTGTAAA ACTGACCAACTG	50,10	
(TTG) ₂ TC(CT) ₂ TA(CT) ₂ T ₃ A(GCCT) ₂ (TG) ₂ GAG ₂ (TC) ₃	PVtc002 F	AGTTCAAGGGAAGATAAAAG	50,40	218
	PVtc002 R	CCTCCCTATTACTAAACCAT	50,20	
(TC) ₄ //(GT) ₃ // (GA) ₂ G(GA) ₃	PVtc003 F	CCACCACTAACATTAAGAA	50,40	168
	PVtc003 R	ATTCACACAGAACAATGATC	49,80	
(TC) ₃ AG(ATC) ₂ (AAAG) ₂	PVtc004 F	CCGGTGTGCCTGATAGTTCT	60,10	268
	PVtc004 R	ACAGTTGGAGGCTTGTGGAC	60,10	
(TC) ₂ A(TC)(CT) ₂ C ₆ (TC) C(TC) ₄	PVtc005 F	GTTCTCCTAATCATAAGCAA	49,70	310
	PVtc005 R	ATGTTAAGTGTTC AAGTGC	50,30	
(TC) ₈ A(CT) ₁₉	PVct001 F	GGATCTTGGAGGGCAACCTA	61,30	183
	PVct001 R	CTCATTGCGTCTACCAGTGC	59,40	
(CCA) ₂ (AAGA) ₂ (ACT) ₂	PVcca001 F	TAACATTCCTTTCATTTAC	50,20	195
	PVcca001 R	TCTGAAGTTTATGTGATGGT	50,30	
(CCT) ₂ (CTC) ₂ // (GATC) ₂	PVcct001 F	GGAGTTTCTGTGTTTCGTCC	58,30	179
	PVcct001 R	CAGATCGCCAAGGTGATTGA	57,10	
(CTTT)CT(CTTT) ₂	PVcttt001 F	GATTCTAAGCACAACTG	50,80	162
	PVcttt001 R	GATAGGTGCTGCTGGAGA	54,60	
(CTTT)TG(CTTT) ₂ TC (CT) ₂ G ₂ (CTA) ₂	PVcttt002 F	TGTATAGCTCTCCCACAAC	50,80	209
	PVcttt002 R	GAGTTGAAAATAAGCGTAGA	50,00	
(TGG) ₅ (TGA) ₂	PVtgg001 F	CATTGTTGGAGTTGTGGTTG	54,40	87
	PVtgg001 R	TTCTGATGATGCTTGTAGGT	54,80	
(AT) ₂ (CA)A(CA) ₄ (TTC) ₂	PVca001 F	ATAAATCATAAATGAACCAAC	48,60	

	PVca001 R	ATCAGGTTGAGAGCAGTAG	50,00	165
(TAG) ₂ TA(ACA) ₃	PVaca001 F	AGACAGAAGGCAAACTTAT	50,60	250
	PVaca001 R	AGTGTAGTTGTTGATGGAAC	49,70	
(CG)A(CG) ₃ CT(GA) ₂ GT (GA)	PVcg001 F	GCAGATGATGAGGTCCAC	54,50	341
	PVcg001 R	ATGCTCAGGCTCATCTAT	50,20	
(CAA) ₂ CAC(CAA) ₃ // (AAC) ₃ (GGAG) ₂	PVcaa001 F	TCAAACCCTACAACATACTT	49,70	185
	PVcaa001 R	CTCTACAAAGCTTCCACTT	49,50	
(CATTG) ₃ C(TG) ₂ // (GT) ₃ (TTC) ₂	PVcattg001 F	CACTCAAAGGTACGATTTT	50,20	270
	PVcattg001 R	ACGTAAGAAAAGAAGAACG	49,70	
(GAA) ₃ (GA) ₂ G ₃ (GA)T (GA)AT(TG) ₃ (TGAAA) ₃ // (GT) ₄	PVgaa001F	CATGTATGAGATGTGAAGAT	52,36	219
	PVgaa001R	AGATCTGGCAGCGACAAC	57,36	
(TTGATC) ₂ //(AGA) ₂ // (ATCA) ₂	PVttgatc001F	CAACACACCAACACTAAGA	49,84	193
	PVttgatc001R	GCTGAGATTGCTGAAAGTC	53,43	
(CTTTGC) ₂ //(TTG) ₂ // (TGCCC) ₂	PVctttgc001F	GCTTCTACACAGTGCTTCTT	52,33	162
	PVctttgc001R	CGGTTCAACGGGCTAAATA	58,64	
(AC) ₃ A(AC) ₂ //(AC) ₃ (AAAC) ₂ //(TTG) ₂	PVac001F	CCAACACTAAGAATGCTGA	52,09	150
	PVac001R	GTGTGAGATAGGATTTCTGCTT	55,23	
(CTC) ₂ //(TCT) ₂ CC (CTT) ₂ TA(TTCCG) ₂ A (AG) ₂	PVctc001F	TGCCACTCTACAGAACCG	55,66	152
	PVctc001R	TGACTGAATTGGTCCTAGTT	52,15	
(TTGAA) ₂ //(TGTC) ₃ // (TCTTC) ₂	PVtgtc001F	ATTCTAAGCACAACAACACTGG	52,95	162
	PVtgtc001R	CGATAGGTGCTGCTGGAG	58,02	
(AG) ₃ //(GT) ₅ T ₂ CT(GT) ₃ A(AG) ₈ //(AATGG) ₂	PVag006F	GCTGCGAGAGCTTCGTTT	59,42	203
	PVag006R	GTGGCTAATACCCTCCTT	54,42	
(GAGTTTTT) ₂ //(TGAT) ₂ //(TCT) ₂ //(CTG) ₂ // (TGTT) ₂ //(GT) ₃ // (GATCAA) ₂	PVgt001F	TTGAGAATTAATCTTTGACA	48,28	248
	PVgt001R	CCAACACACCAACACTAAG	52,16	
(GAA) ₄ (GA) ₄ //(GA) ₂ A(GA)A(GA)	PVgaa002F	ACAGGGCATGAGGAGAAG	56,12	213
	PVgaa002R	GCTGTCACCAACAAAGAC	51,99	
(GA) ₃ //(TG) ₃ //(GAAA) ₂ // (GT) ₂ G(AAGA) ₂ // (AAGCA) ₂ //(TAT) ₂ A (TGTT) ₂	PVga003F	CGGCTGAGTAATGTTTAC	52,27	338
	PVga003R	CATGACTGATTATGCTTTGA	52,10	
(ATTAA) ₂ //(CTC) ₂ // (CAA) ₂ //(CAA) ₂ //(AC) ₃	PVac003F	TAACCACTATCATCAAGCA	49,71	194

(CAT) ₂	PVac003R	TTCTTCACATCTTACACTTCTA	49,51	
(GCT) ₂ /(TCGC) ₂ CG (CTA) ₂ /(CTTTT) ₂ TC (CT) ₂ GG(CTA) ₂ /(TC) ₂ T (TGCC) ₂ /(CAG) ₃	PVcag001F	ACTCGGGTCAGAGCGTAA	56,86	280
	PVcag001R	ATTCCCGATATAGGTCAAC	51,41	
(TC) ₂ C(ACA) ₂ /(CTT) ₃ // (TCC) ₂ C ₂ AG(AAACT) ₂	PVctt001F	ATTCCTGCCTTCATAACAT	52,17	233
	PVctt001R	CACGGTTTGATAGAGTTG	49,43	
(GA) ₂ (AT) ₂ (GGT) ₂ (CT) ₂ /(AGTGT) ₂ (GTT) ₂ T (TAA) ₂ /(GT) ₂ (GAT) ₂	PVga004F	AGATTTGATTGTATCCTGAGT	50,42	357
	PVga004R	CCTGGCGTGAGTTTAGAA	54,81	
(GGT) ₃ /(GAG) ₂ /(GA) (GAAA) ₂ /(GAGG) ₂ C (GAG) ₂ /(GCG) ₂ (GAA) ₂ /(AGGGAA) ₂	PVggt001F	GCGTAGTCATGTTGAGGA	52,67	387
	PVggt001R	GCATACGTTGTGGGCTGT	58,05	
(GGGTC) ₂ /(AAGT) ₂ // (TGG) ₂ A(AG) ₂ /(GGT) ₃ (GT) ₂ /(TAA) ₂ /(TGA) ₃	PVtga001F	GAACCATGGATAATCTATTT	48,31	317
	PVtga001R	AGACCAGGTTTCCACTATTC	53,22	
(AG) ₃ /(GGT) ₂ /(AC)C (AC)(CAG) ₂ /(AG) ₂ (CG) ₂ /(TC) ₂ (GT) ₂ T (AC) ₂	PVag007F	CAGGTGGCAGAGAGGTAG	54,10	400
	PVag007R	GTGTGCTCATGCGAGGTC	58,85	
(GCA) ₂ /(GCT) ₂ // (AGC) ₂ (ACG)(AC) ₂ T ₂ G (TA) ₂ G(CT) ₂ C ₂ (CA) ₂ // (TCGC) ₂ CG(CTA) ₂ // (CTTT) ₂ (TC)(CT) ₂ G ₂ (CTA) ₂ /(TC) ₂ T(TGCC) ₂ /(CAG) ₃	PVcag002F	TACTCGGGTCAGAGCGTA	55,29	399
	PVcag002R	ATTCCCGATATAGGTCAAC	51,41	
(TGCG) ₂ /(TGT) ₂ /(GT) (GTT) ₃ G(GT) ₂ G(GT)// (GA) ₂	PVggt001F	TTTGTCGTAGTAGATGCTGA	51,93	245
	PVggt001R	CTCATATTCTGGCTGGTAGT	51,92	
(TTGATGATA)(GTGAC GATA)(TTGATGATA) (GTGACGATA) ₂ /(AT) ₂ (GAGCT)G(GAGCT)//(AT)A(AT) ₂ /(TC) ₃ /(TG) ₂	PVtc007F	CCTGGATAACAAGCCTAAG	52,57	387
	PVtc007R	TATGAGTCGAAGCAATAACA	51,94	
(GATG) ₂ /(AGCTTTCA) ₂ /(GT) ₂ /(GGT) ₂ (GA) ₂	PVgatg001F	TGTGTGGAATAGATGCTTTA	51,75	174
	PVgatg001R	ATTATAGTGTTCGCTGCTC	52,37	
(TCT) ₃ /(CA) ₂ (CT) ₂ // (CT) ₅ (TCT) ₃ (TCC) ₂	PVct002 F	GCTGTCACCAACAAAGAC	51,99	213
	PVct002 R	ACAGGGCATGAGGAGAAG	56,12	
(GA) ₂ (AC) ₂ /(AC) ₂ A (GCAAT) ₃ G(GA) ₂ AG (GA) ₂	PVgcaat001F	AAGAAACACACGAACATCA	52,16	144
	PVgcaat001R	TCCATTCTCCCTTACTT	51,90	
(TTC) ₄ C(TCT) ₂₀ (TC) ₂	PVtct002 F	ACCATGGTTCAGTCCCTT	55,10	179

	PVtct002 R	CTTCATGTACGCATTCTCT	51,09	
(CA) ₂ (AGCA) ₂ GTC ₂ (CT) ₂ (TTA) ₂ (GGTAA) ₂	PVca002 F	GTTCTGTCCCTTTACCATTTT	50,99	180
	PVca002 R	TGATTTTCGTGCATAACTTAC	51,40	
(AACC) ₂ (AT) ₂ A(AT)// (TTTGC) ₂ //(TCTT) ₂ (CA) C ₂ (TG)(CA) ₂ (TG) (CAC) ₂ A ₃ (TG)(TTTC) ₂ // (GA) ₄	PVga005F	CATGACTGATTATGCTTTGA	52,10	391
	PVga005R	TCTTTGATTTGTGCGCTT	51,72	
(CA)(CCA) ₃ CA ₃ (CCA) ₂ CT(AC)(CAA) ₂ //(TC) ₂ A (ACC) ₅ //(CTT) ₂ C(CA) ₂	PVacc001F	CTTATGTACACCTTTTCCT	51,81	224
	PVacc001R	CTCTCCAAGTCTTCATCAG	52,39	
(TGT) ₃ (GT)(TG)/(GT) ₅ T C ₂ (GT) ₂ A(GT)T (CATGAG) ₂ (GA) ₂ (GA) ₂	PVgt002F	TGAATCTCATGTTTTTGTG	51,98	241
	PVgt002R	CCTCCTCAGTGTTTTGATAC	52,15	
(CCTCCA) ₂ //(TCT) ₃ (TC)//(TCT) ₄	PVtct003 F	GGATAACCTTCACCTTCTAC	50,41	187
	PVtct003 R	GATCATCTATGTCCACAAC	52,52	
[(TGTATAGA)(GGTTC TTC)] ₂ (TGTATACA) (GGTCTTC) (TGTATAGA) (GGTCTTC)	PVx001 F	TACTATTTGTACTTTGTGTTGTAA	49,39	160
	PVx001 R	TTTGTAATAATAGACTCTTATGAT	49,40	
(GT) ₄ //(GTTTA) ₂ G (GTTT) ₂ G(AGAGGA) ₂	PVgt003F	GTGACAGGATAAGTAGAATGA	49,27	140
	PVgt003R	TTAAATACAACCAGGTAGCC	52,15	
(CTTT) ₂ (GAT) ₂ (CT) (GA) ₃ //(TTTT) ₂ (CA) (TTAC) ₂	PVga006F	TCAGTTGTGTATGTTTATTCTG	50,57	220
	PVga006R	GGTGTGCCTGATAGTTCTG	52,63	
(GAA) ₁₆ //(AGG) ₂ (TG) ₂ (TC)(TG)(TTGCAA) ₂	PVgaa005F	GTCATGGGAAACAGTAGAAG	52,15	236
	PVgaa005R	CCGAAGGAATATACAGTGTC	52,20	
(GT) ₂ T(GT) ₂ TC ₂ (GT) ₂ A (TGGG) ₂	PVgt004F	ATCCAAATGATGAGAATGAT	51,71	217
	PVgt004R	CCTCACACAAACAACTGTAA	51,95	
(AC) ₂ T(CACCCA) ₂ // (TCA) ₂ C ₂ A(GGAACA) ₂ G(GT) ₂	PVac004F	ATGAAGAGCAAAGAACACT	51,65	235
	PVac004R	GTAGGAGGCAAGGTATTATG	51,61	
(TG) ₂ (CA)(TG) ₂ //(AC) ₂ (GT)(TC) ₂ C(TC)	PVtg001F	GCCTTCCAATGTGAGTAA	51,73	286
	PVtg001R	GGCTTGGAAGTTAGACAAA	52,97	

Como pôde ser observado na Tabela 2, os motivos microssatélites mais freqüentes foram da classe AG/TC. Yu et al. (1999), Melotto et al. (2005), Hanai et al. (2007), Benchimol et al. (2007) também verificaram uma alta freqüência do motivo GA/CT no genoma do feijoeiro. O resultado encontrado

no presente trabalho confirma a alta ocorrência deste motivo dinucleotídico no genoma da espécie.

Dentre os locos microssatélites identificados, o número máximo de repetições encontrado de um dinucleotídeo foi (CT)₁₉, para o loco PVct001, e de um trinucleotídeo foi (TCT)₂₀, para o loco PVtct002. Resultados semelhantes foram obtidos por Benchimol et al. (2007), que constataram que o motivo GA/CT não apenas estava entre os mais freqüentes dentre os locos analisados, como também foi o que constituiu os motivos mais longos no genoma do feijão, assim como observado no presente trabalho para os locos PVag003, PVct001 e PVag006, que apresentaram os motivos mais longos.

Foi também observado que alguns motivos não foram compatíveis com as sondas utilizadas na etapa de enriquecimento. Levando-se em consideração apenas as maiores repetições presentes nos motivos encontrados, foi possível identificar 51 (77,27%) motivos SSR compatíveis com as sondas utilizadas na biblioteca 1, enquanto que na biblioteca 2 apenas 8 (40%) motivos microssatélites foram encontrados. O baixo percentual de enriquecimento da biblioteca 2 sugere que os motivos SSR usados são pouco freqüentes no genoma do feijão ou que as condições utilizadas nas etapas de hibridização e lavagem não favoreceram ao enriquecimento para estas repetições. .

Gaitán et al. (2002) construíram três bibliotecas enriquecidas e seqüenciaram 856 (18,3%) clones para obterem 213 (24,88%) pares de *primers*, dos quais 68 geraram produtos de amplificação. Yaish e Pérez De La Vega (2003) identificaram 21 (35%) locos microssatélites ao analisarem 60 seqüências obtidas a partir de uma biblioteca genômica enriquecida para motivos GA. Benchimol et al. (2007) construíram uma biblioteca enriquecida para os motivos (CT) e (GT) e seqüenciaram 1209 (48,8%) clones para obterem 714 (59,06%) seqüências contendo microssatélites, das quais 540 foram utilizadas para o desenho de *primers*. No presente trabalho, foram seqüenciados 196 (94,68%) clones, dos quais 184 (88,89%) apresentaram seqüências com boa qualidade para serem analisados, sendo 133 clones da biblioteca 1 e 51 clones da biblioteca 2. A análise dos clones permitiu identificar 66 (49,62%) motivos SSR na biblioteca 1 e 20 (39,22%) na biblioteca 2, no entanto, a partir dos critérios para desenho de *primers* estabelecidos no item 2.11 da sessão “Material e Métodos”, foi possível desenhar 56 pares de *primers* flanqueadores.

A alta eficiência obtida na seleção de clones positivos para motivos microssatélites, 49,62% para a biblioteca 1 e 39,22% para a biblioteca 2, e no desenho de *primers*, 56 (30,43%) *primers* SSR desenvolvidos a partir de 184 clones analisados, indicam que o método utilizado foi eficaz na identificação e desenvolvimento de marcadores microssatélites para o feijão-comum.

3.3. Identificação de marcadores microssatélites polimórficos entre os genótipos andinos e mesoamericanos

Dos 56 pares de *primers* desenvolvidos no presente trabalho, 34 tiveram suas condições de amplificação padronizadas e foram analisados quanto ao polimorfismo detectado entre os 20 genótipos andinos e mesoamericanos descritos na Tabela 1, do item 2.1 da sessão “Material e Métodos”. A tabela 3 apresenta a temperatura de amplificação padronizada para cada um dos 34 pares de *primers* analisados e quais detectaram polimorfismo entre os genótipos testados. Os demais 22 pares de *primers* não analisados ainda serão padronizados e testados em próxima oportunidade.

Tabela 3: Temperatura de amplificação padronizada para os 34 pares de *primers* analisados e identificação de quais detectaram polimorfismo entre os genótipos testados.

Primer	Temperatura de amplificação (°C)	Análise de Polimorfismo
PVga001	59 °C	Monomórfico
PVcttt002	59 °C	Monomórfico
PVtgg001	59 °C	Polimórfico
PVca001	59 °C	Monomórfico
PVaca001	59 °C	Monomórfico
PVcg001	59 °C	Monomórfico
PVcattg001	59 °C	Monomórfico
PVag001	54 °C	Monomórfico
PVag003	54 °C	Polimórfico
PVag004	54 °C	Monomórfico
PVtc001	54 °C	Monomórfico
PVtc002	54 °C	Monomórfico
PVtc003	54 °C	Monomórfico
PVtc004	54 °C	Monomórfico
PVtc005	54 °C	Monomórfico
PVct001	54 °C	Polimórfico
PVcca001	54 °C	Polimórfico

PVcct001	54 °C	Monomórfico
PVcttt001	54 °C	Monomórfico
PVcaa001	51 °C	Monomórfico
PVgaa001	54 °C	Monomórfico
PVttgatc001	54 °C	Monomórfico
PVctttgc001	54 °C	Monomórfico
PVac001	54 °C	Monomórfico
PVctc001	54 °C	Monomórfico
PVtgtc001	54 °C	Monomórfico
PVag006	54 °C	Polimórfico
PVgt001	54 °C	Monomórfico
PVgaa002	54 °C	Monomórfico
PVga003	54 °C	Polimórfico
PVac003	54 °C	Monomórfico
PVcag001	54 °C	Monomórfico
PVctt001	54 °C	Monomórfico
PVga004	54 °C	Monomórfico

Todos os 34 pares de *primers* testados geraram produtos de amplificação e seis (17,64%) amplificaram produtos polimórficos entre os genitores testados. Os *primers* que se mostraram polimórficos foram caracterizados quanto ao número de alelos que apresentaram e quanto ao Conteúdo de Informação de Polimorfismo (PIC). Os resultados obtidos encontram-se descritos na Tabela 4.

Tabela 4: Caracterização dos locos microssatélites polimórficos quanto ao número de alelos detectados, variação de tamanho (pb) dos alelos, e o Conteúdo de Informação de Polimorfismo (PIC).

SSR	Motivo SSR e número de repetições	Número de alelos detectados	PIC
PVag003	(AG) ₃ ATA ₃ (GT) ₅ //(AG) ₈	2 alelos	0,4950
PVct001	(TC) ₈ A(CT) ₁₉	4 alelos	0,5700
PVtgg001	(TGG) ₅ (TGA) ₂	2 alelos	0,3200
PVcca001	(CCA) ₂ (AAGA) ₂ (ACT) ₂	2 alelos	0,4800
PVag006	(AG) ₃ //(GT) ₅ T ₂ CT(GT) ₃ A (AG) ₈ //(AATGG) ₂	2 alelos	0,4986
PVga003	(GA) ₃ //(TG) ₃ //(GAAA) ₂ //(GT) ₂ G(AAGA) ₂ //(AAGCA) ₂ //(TAT) ₂ A(TGGT) ₂	2 alelos	0,5000

O número de alelos detectados variou de dois a quatro, e o valor de PIC variou de 0,3200 a 0,5700, com uma média de 0,4773. O maior valor de PIC (0,5700) foi associado ao loco PVct001, que apresentou também o maior motivo dinucleotídico perfeito identificado no presente trabalho (CT)₁₉. O valor médio de PIC encontrado para os *primers* desenvolvidos foi próximo ao obtido por Métails et al. (2002), que estimaram um PIC médio de 0,44 entre 15 locos microssatélites, desenvolvidos a partir de duas bibliotecas genômicas enriquecidas de *Phaseolus vulgaris*, e por Hanai et al. (2007) que obtiveram um PIC médio de 0,45, e uma variação de dois a três alelos para os locos desenvolvidos e testados em 25 acessos andinos e mesoamericanos de *Phaseolus vulgaris*. Os *primers* PVtgg001 e PVcca001 não amplificaram produtos polimórficos entre os genitores AND277 e Rudá, apesar de apresentarem polimorfismo entre outros genótipos analisados. Para exemplificação, a Figura 3 ilustra o perfil polimórfico apresentado pelo *primer* PVag003. A ordem dos genótipos dispostos no gel corresponde à mesma descrita na Tabela 1, do item 2.1 da sessão “Material e Métodos”, com exceção dos dois últimos genótipos que foram testados em outro gel não apresentado no presente trabalho.

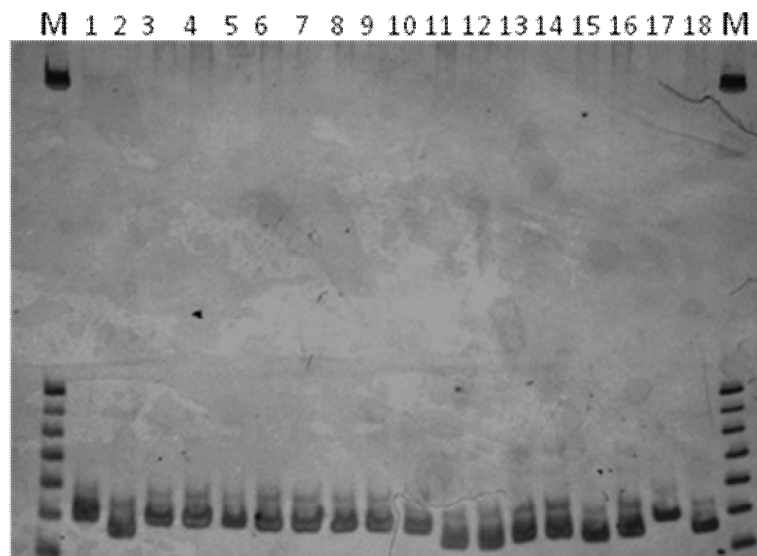


Figura 3: Perfil eletroforético polimórfico apresentado pelo *primer* PVag003 em gel de poli-acrilamida 10%, corado com solução de nitrato de prata. A numeração 1-18 corresponde à ordem dos genótipos andinos e mesoamericanos disposta na Tabela 1, do item 2.1, com exceção dos dois últimos genótipos que não foram testados no presente gel. A letra M corresponde ao marcador de tamanho de 25 pb.

Assim como reportado por Gaitán et al. (2002), o fato de um dado motivo microssatélite ser composto apenas de pequenas unidades repetitivas não limita a possibilidade desse *primer* poder detectar polimorfismo entre diferentes genótipos analisados. Adicionalmente, resultados encontrados por Benchimol et al. (2007) comprovaram que motivos compostos podem prover um padrão polimórfico informativo. Os autores identificaram uma alta frequência de pequenas unidades repetitivas, usualmente de três e quatro unidades, constituindo muitos motivos compostos, que se mostraram polimórficos entre os genótipos analisados e que levou à conclusão de que tal padrão microssatélite é comumente encontrado no genoma do feijoeiro. Esse padrão foi também confirmado pelos resultados obtidos no presente trabalho, a exemplo do conteúdo informativo apresentado pelos *primers* PVcca001 e PVga003, que flanqueiam microssatélites compostos de motivos perfeitos pequenos e ainda assim mostraram-se polimórficos entre os genótipos analisados (Tabela 5).

3.4. Análise de similaridade

Foi identificada similaridade entre as sequências obtidas neste trabalho com regiões transcritas e não traduzidas, e com regiões codificadoras de proteínas presentes no genoma nuclear, mitocondrial e do cloroplasto. Foi também identificada similaridade com sequências de retrotransposons e com sequências microssatélites de *P. vulgaris* já publicadas na literatura (BENCHIMOL et al., 2007). Os resultados de similaridade encontrados para os locos desenvolvidos estão descritos na Tabela 5, assim como a porcentagem de similaridade entre as sequências de entrada e o primeiro *hit*, o *score* de alinhamento entre elas e o respectivo *e-value*. A categorização das sequências, segundo critérios de classificação do *Gene Ontology*, está descrita na Tabela 6. Dentre todos os locos microssatélites analisados no presente trabalho, somente os descritos na Tabela 5 apresentaram similaridade com alguma sequência já depositada e caracterizada do banco de dados. Da mesma forma, dentre os locos dispostos na Tabela 5, somente aqueles que apresentaram sequências para as quais foi identificada alguma função celular já caracterizada estão categorizados na Tabela 6.

Em relação ao genoma nuclear, foi identificada similaridade das regiões microssatélites analisadas com regiões codificadoras do fator de transcrição GT-2 e smart00415 de *Glycine Max* (soja). O fator de transcrição GT-2 é componente de uma família de fatores de transcrição com domínios de ligação ao DNA bastante conservados, e estão relacionados à ativação de genes responsivos ao estímulo luminoso. Já o fator de transcrição smart00415 é um elemento regulatório de resposta ao estresse por calor, característico de *Glycine max*.

Em relação ao genoma mitocondrial, foi detectada alta similaridade com a região codificadora da subunidade 5 da enzima NADH desidrogenase de *Vitis vinifera*. Tal subunidade é componente do complexo I NADH desidrogenase da cadeia respiratória e sua função é a transferência de elétrons do NADH para a cadeia respiratória. O complexo protéico está localizado na membrana interna mitocondrial.

Em relação ao genoma do cloroplasto, houve similaridade com a região codificadora da ribonucleoproteína S19, e com os genes *ycf2* e *rrn23* do RNA ribossomal 23S, todos em *Phaseolus vulgaris*. Assim como com o gene *hsp22* para proteína de baixo peso molecular de resposta ao estresse por calor de *Glycine max*. A proteína S19 é uma ribonucleoproteína que forma um complexo com a proteína S13 que se liga ao rRNA 16S, constituinte da subunidade 30S do ribossomo de cloroplastos. Já o gene *rrn23* codifica o rRNA 23S, constituinte da subunidade 50S também de ribossomos de cloroplastos. A proteína *hsp22* atua como uma chaperona que impede a desnaturação de outras proteínas durante o estresse celular provocado pelo estímulo externo de calor. Vários trabalhos foram relatados na literatura de identificação de microssatélites na região promotora de proteínas *hsp* (*heat shock protein*) de *Drosophila* (SANDALTZOPOULOS et al., 1995), *Aspergillus* (PUNT et al., 1990), *Phytophythora* (CHEN e ROXBY, 1997) e de mamíferos (LISOWSKA et al. 1997). Em relação ao gene *ycf2*, de *P. vulgaris*, este ainda não teve sua função precisamente caracterizada. Sabe-se que se trata de um gene codificante de uma provável ATPase, e sua presença na planta não-fotossintetizante *Epifagus virginiana* e resultados de experimentos mais elaborados em tabaco indicam que a proteína tem uma função putativa essencial não relacionada à fotossíntese (WOLFE, 1994)

Foi também identificada alta similaridade com duas regiões microssatélites do feijão, publicadas por Benchimol et al. (2007), e com o retrotransposon Tpv2-1c da família Tpv2 de retrotransposons exclusivos de *Phaseolus vulgaris*. Esses elementos são preferencialmente transcritos em condições de estresse, principalmente relacionados à resposta da planta ao ataque de patógenos (GARBER et al., 1999). Microssatélites em famílias de retrotransposons exclusivos de *Phaseolus vulgaris* foram previamente descritos por Melotto et al. (2005) e por Benchimol et al. (2008). A caracterização e uso de elementos transponíveis é uma importante ferramenta para estudos funcionais e de *screening* de mutantes para o feijão-comum, uma vez que essa cultura tem se mostrado recalcitrante à transformação genética (MELOTTO et al., 2005).

Foi ainda identificada similaridade com cDNA de *Glycine max*, ainda não caracterizado. Interessantemente o loco microssatélite sob análise, PVct001, corresponde ao motivo dinucleotídico perfeito mais longo detectado no presente trabalho, (TC)₈A(CT)₁₉. Segundo Li et al. (2002), é baixa a ocorrência de motivos microssatélites longos em regiões de éxons, principalmente de dinucleotídeos, por se constituírem em regiões do genoma com alto índice mutacional, o que poderia acarretar mudanças consideráveis no fenótipo protéico correspondente a essas regiões gênicas, implicando em danos para o organismo. No entanto, os dados encontrados nesse trabalho contestam esse fato, uma vez que a similaridade detectada entre o loco microssatélite e o *hit* cDNA de *Glycine max* foi consistente, pois apresentou alta porcentagem de similaridade, alto *score* de alinhamento e baixo *e-value*, indicando uma conservação de região gênica entre as leguminosas feijão e soja. Os motivos repetitivos presentes nos genes *ycf2* e no gene que codifica a ribonucleoproteína S19 constituem-se em minissatélites, com unidades repetitivas de 8 e 9 pb, respectivamente.

Tabela 5: Resultado da análise de similaridade. A primeira coluna corresponde ao loco microsatélite analisado, a segunda ao *hit* encontrado, seguida da porcentagem de similaridade entre as duas sequências, score de alinhamento e *e-value*. O número de acesso, do banco do NCBI, das sequências identificadas como similares encontra-se na última coluna.

SSR	RESULTADO DE SIMILARIDADE PELO BLASTN	% DE SIMILARIDADE	SCORE	e-VALUE	NÚMERO DE ACESSO
PVga001	Gene hsp22 do genoma do cloroplasto. Produto: proteína de baixo peso molecular de resposta ao stress (calor), de <i>Glycine max</i>	92%	556	2e-155	X63198
PVtc002	Gene mitocondrial para a subunidade 5 da enzima NADH desidrogenase, de <i>Vitis vinifera</i>	91%	516	4e-143	FM179380
PVct001	cDNA de função não caracterizada, de <i>Glycine max</i>	98%	520	4e-120	AK286979
PVcttt002	Gene rrn23 do genoma do cloroplasto. Produto: RNA ribossomal 23S de cloroplasto	99%	309	3e-81	EU196765
PVca001	Sequência microsatélite de <i>Phaseolus vulgaris</i> , já publicada na literatura	86%	191	3e-45	EF136465
PVaca001	Gene para fator de transcrição GT-2, de <i>Glycine max</i>	89%	243	9e-61	AF372499
PVx001	Gene do genoma do cloroplasto. Produto: proteína ribossomal S19 de cloroplasto, de <i>Phaseolus vulgaris</i>	99%	936	7e-129	EU196765
PVtg001	cDNA de retrotransposon Tpv2-1c, de <i>Phaseolus vulgaris</i>	99%	577	1e-161	AJ005761
PVtc009	Gene para fator de transcrição smart00415 de resposta ao estresse por calor, de <i>Glycine max</i>	86%	294	2e-76	ABC47851
PVct004	Sequência microsatélite de <i>Phaseolus vulgaris</i> , já publicada na literatura	98%	889	0	DQ469500
PVtc007F	Gene ycf2 do genoma do cloroplasto de <i>Phaseolus vulgaris</i> , de função hipotética.	100%	715	0	EU196765

Tabela 6: Categorização dos produtos gênicos segundo critérios de classificação do *Gene Ontology*.

Produto gênico	Anotação <i>Gene Ontology</i> Categorias: Processo, Função e Componente
Fator de transcrição smart00415, em <i>Glycine max</i>	Processo: "Regulação da transcrição, DNA dependente"; "Proteína de resposta à estresse" Função: "Atividade de fator de transcrição"; "Ligação ao DNA sequência-específico"; Componente: "Nuclear"
Fator de transcrição GT-2, em <i>Glycine max</i>	Função: "Ligação ao DNA" Componente: "Nuclear"
Subunidade 5 da enzima NADH desidrogenase	Processo: "Síntese de ATP acoplada ao transporte de elétrons"; "Redução oxidativa" Função: "Atividade de NADH desidrogenase associada à ubiquinona"; "Atividade de oxiredutase" Componente: "Mitocondrial"; "Membrana"; "Proteína integral de membrana"
Ribonucleoproteína S19	Processo: "Processo de tradução" Função: "Constituinte estrutural do ribossomo"; "Proteína de ligação ao rRNA" Componente: "Intracelular"; "Ribossomo"; "Pequena unidade ribossomal"; "Cloroplasto"; "Complexo ribonucleoproteico"; "Ribossomo"
Gene ycf2 de função hipotética	Processo: "Atividade nucleosídeo-trifosfatase"; "Ligação à nucleotídeo"; "Ligação ao ATP"; Componente: "Cloroplasto"; "Estroma do cloroplasto"
rRNA 23S	Processo: "Tradução" Função: "Constituinte estrutural do ribossomo" Componente: "Ribossomo"; "Cloroplasto"; "Complexo ribonucleoproteico"; "Subunidade maior ribossomal"; "Intracelular"; "Ribossomal"
Gene hsp22	Processo: "Resposta à estresse" Componente: "Cloroplasto"

A identificação e o desenvolvimento de marcadores microssatélites presentes no genoma do cloroplasto é ainda pouco explorado para a cultura do feijão. Microssatélites de cloroplasto (cpSSRs) já estão sendo extensivamente desenvolvidos e utilizados para várias culturas por apresentarem certas vantagens, a exemplo, permitem geralmente uma alta detecção de polimorfismo entre os genótipos analisados e apresentam herança uniparental materna, sendo bastante utilizados em estudos de pureza e diversidade genética, identificação molecular de genótipos e em estudos evolutivos. Estudos utilizando cpSSRs no feijoeiro já foram realizados principalmente em análises de diversidade genética entre cultivares locais de *Phaseolus vulgaris* (ANGIOI et al., 2009a) e entre diferentes espécies do gênero *Phaseolus* (SICARD et al., 2005). No entanto, os locos microssatélites utilizados nesses estudos foram desenvolvidos para outros grupos taxonômicos, *Glycine max* e gimnospermas, respectivamente, e não especificamente para a cultura do feijão. Para o feijão, tais marcadores foram identificados até o momento apenas por Benchimol et al. (2007) (três locos). No entanto, um estudo mais direcionado especificamente ao desenvolvimento de cpSSRs em feijão está em andamento (ANGIOI et al., 2009b) no *Center for Biotechnology Development and Biodiversity Research*, localizado na Università degli Studi di Sassari, Itália (dados não publicados).

A publicação de estudos científicos que envolvam a identificação de cpSSRs no genoma do feijoeiro é recente. O seqüenciamento do genoma completo do cloroplasto de *Phaseolus vulgaris* e sua conseqüente caracterização (GUO et al., 2007) representou um importante avanço para os estudos genômicos na espécie, pois, anteriormente à isso, a quantidade de sequências de origem cloroplastidial já caracterizadas e disponíveis em bancos públicos de sequências era muito pequena. Segundo GUO et al. (2007), os padrões de motivos repetitivos identificados em *Phaseolus vulgaris* indica uma importante função na diversificação entre os genomas de cloroplastos das leguminosas, além de apresentarem altas taxas evolucionárias de mudanças tanto a nível genômico, quanto gênico, podendo ser, portanto, bastante úteis em estudos evolutivos e de diversificação genética.

Estudos genômicos desenvolvidos para a cultura do feijão são ainda muito limitados. Nesse sentido, análises comparativas e de similaridade com outras espécies de plantas, com genomas mais bem caracterizados tais como

Glycine max e *Arabidopsis thaliana*, são de crucial importância. A identificação de regiões microssatélites em regiões gênicas é particularmente interessante, do ponto de vista da aplicação desses marcadores. Tais marcadores, apesar de se apresentarem menos polimórficos em comparação aos genômicos (EUJAYL, 2001), são mais informativos e vantajosos em estudos de seleção assistida por marcadores moleculares e em estudos de evolução. Adicionalmente, caracterizam-se como uma rica fonte de sequências comumente usadas no *screening* por SNPs, além de contribuírem para a construção de mapas comparativos baseados em sequências expressas, resultando no desenvolvimento de mapas transcricionais, que constituem uma valiosa fonte para futuros estudos de desequilíbrio de ligação na cultura. Além disso, tais tipos de estudos podem impulsionar o ancoramento de mapas genéticos a mapas físicos, caracterizando um primeiro passo para a clonagem de genes de grande importância para programas de melhoramento da cultura do feijoeiro (MELOTTO et al., 2005).

Os resultados obtidos no presente trabalho permitiram a constatação da presença de SSRs em regiões relacionadas a importantes funções no genoma do feijão e o desenvolvimento de *primers* cpSSRs, ainda pouco descritos na literatura científica para a cultura. Os marcadores microssatélites desenvolvidos poderão ser usados como importantes ferramentas para os mais variados tipos de estudos genéticos para a espécie, inclusive em estudos de mapeamento genético.

4. REFERÊNCIAS BIBLIOGRÁFICAS

- ANDERSON, J.A., CHURCHILL, G.A., AUTRIQUE, J.E., TANKSLEY, S.D., SORRELLS, M.E. (1993). **Optimizing parental selection for genetic linkage maps**. *Genome*, 36: 181-186.
- ANGIOI, S.A., RAU, D., RODRIGUEZ, M., LAGOZZO, G., DESIDERIO, F., PAPA, R., ATTENE, G. (2009a). **Nuclear and chloroplast microsatellite diversity in *Phaseolus vulgaris* L. from Sardinia (Italy)**. *Molecular Breeding*, 23:413-429.
- ANGIOI, S.A., DESIDERIO, F., RAU, D. (2009b- em preparação, dados não publicados). **Development and use of chloroplast microsatellites in *Phaseolus* spp. and other legumes**. (trabalho referenciado em ANGIOI 2009a)
- AUSUBEL, F (1998). **Current Protocols in Molecular Biology**. New York: John Wiley & Sons, Tenton Data Systems.
- BENCHIMOL L. L., CAMPOS, T., CARBONELL, S. A. M., COLOMBO, C. A., CHIORATTO, A. F., FORMIGHIERE, E. F., GOUVÊA, L. R. L., SOUZA, A. P. (2007). **Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers**. *Genetic Resource and Crop Evolution*, 54: 1747-1762.
- BEST. Bean EST Project. Disponível em <http://lqm.esalq.usp.br/BEST>. Acesso em Julho/2009
- BLAIR, M. W., BUENDÍA, H. F., GIRALDO, M. C., MÉTAIS, I., PELTIER, D. (2008). **Characterization of AT-rich microsatélites in common bean (*Phaseolus vulgaris* L.)**. *Theoretical and Applied Genetics*, 118: 91-103.
- BLAIR, M. W., PEDRAZA, F., BUENDIA, H. F., GAITÁN-SOLÍS, E., BEEBE, S. E., GEPTS, P., TOHME, J. (2003). **Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.)**. *Theoretical and Applied Genetics*, 107: 1362-1374.
- CAIXETA, E. T., OLIVEIRA, A. C. B., BRITO, G. G., SAKIYAMA, N. S. (2009). **Tipos de marcadores moleculares. In: Marcadores Moleculares**. BORÉM, A., CAIXETA, E. T. 2º Edição. Viçosa/MG: Editora UFV. 532 p.
- BUSO G. S. C., AMARAL, Z. P. S., BRONDANI, R. P. V, FERREIRA, M. E. (2006). **Microsatellite markers for the common bean *Phaseolus vulgaris***. *Molecular Ecology Notes*, 6: 252–254.
- CAIXETA, E. T., BORÉM, A., KELLY, J. D. (2005). **Development of microsatellite markers based on BAC common bean clones**. *Crop Breeding and Applied Biotechnology*, 5(2): 125-133.
- CHEN, Y., ROXBY, R. (1997). **Identification of a functional CT-element in the *Phytophthora infestans* piypt1 gene promoter**. *Gene*, 198: 159-164.

- COSTA, M.R. (2007). **Melhoramento de feijões preto e vermelho visando a resistência à antracnose, mancha-angular e ferrugem, com o auxílio de marcadores moleculares.** Viçosa: UFV. 88 p. (Tese-Doutorado).
- DOYLE, J. J., DOYLE, J. L. (1990). **Isolation of plant DNA from fresh tissue.** Focus, 12: 13-15.
- EIJAYL, I., SORRELLS, M.E., BAUM, M., WOLTERS, P., POWELL, W. (2001). **Assessment of genotype variation among cultivated durum wheat based on EST-SSRs and genomic SSRs.** Euphytica, 119: 39-43.
- FIELD, D., WILLS, C. (1996). **Long, polymorphic microsatellites in simple organisms.** Proceedings. Biological Sciences/ The Royal Society, 263: 209–215.
- FOUILLOUX, G. (1976). **Bean anthracnose: new genes of resistance.** Annual Report of the Bean Improvement Cooperative, 19: 36-37.
- GAITÁN-SOLÍS, E., DUQUE, M.C., EDWARDS, K.J., TOHME, J (2002). **Microsatellite repeats in common bean (*Phaseolus vulgaris*): Isolation, Characterization, and Cross-Species Amplification in *Phaseolus* spp.** Crop Science, 42: 2128-2136.
- GAO, L., TANG, J., LI, H., JIA, J. (2003). **Analysis of microsatellites in major crops assessed by computational and experimental approaches.** Molecular Breeding, 12: 245-261.
- GARBER, K., BILIC, I., PUSCH, O., TOHME, J., BACHMAIR, A., SCHWEIZER, D., JANTSCH, V. (1999). **The Tpv2 family of retrotransposons of *Phaseolus vulgaris*: structure, integration, characteristics, and use for genotype classification.** Plant Molecular Biology, 39: 797-807.
- GENOLYPTUS. Rede Brasileira de Pesquisa do Genoma do Eucalipto. Disponível em <http://www.genolyptus.org.br/tiki-index.php>. Acesso em Julho/2009.
- GUERRA-SANZ, J. M. (2004). **New SSR markers of *Phaseolus vulgaris* from sequence databases.** Plant Breeding, 123: 87-89.
- GUO, X., CASTILLO-RAMIREZ, S., GONZÁLEZ, V., BUSTOS, P., FERNÁNDEZ-VÁZQUEZ, J.L., SANTAMARÍA, R.I., ARELLANO, J., CEVALLOS, M.A., DÁVILA, G. (2007). **Rapid evolutionary change of common bean (*Phaseolus vulgaris* L.) plastome, and the genomic diversification of legume chloroplasts.** BMC Genomics, 8: 228-244.
- HAMILTON, M. B., PINCUS, E. L., Di FIORE, A., FLEISCHER, R. C. (1999). **Universal linker and ligations procedures for construction of genomic DNA libraries enriched for microsatellites.** Biotechniques, 27: 500-505.

- HANAI, L. R., CAMPOS, T., CAMARGO, L. E. A., BENCHIMOL, L. L., SOUZA, A. P., MELOTTO, M., CARBONELL, S. A. M., CHIORATTO, A. F., CONSOLI, L., FORMIGHIERI, E. F., SIQUEIRA, M. V. B. M., TSAI, S. M., VIEIRA, M. L. C. (2007). **Development, characterization, and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources.** *Genome*, 50: 266-277.
- HANCOCK, J. M. (1996). **Simple sequences in a 'minimal' genome.** *Nature Genetics*, 14: 14-15.
- L'TAIEF, B., HORRES, R., JUNGSMANN, R., MOLINA, C., SIFI, B., LACHAËL, M., WINTER, P., KAHL, G. (2008). **Locus-specific microsatellite markers in common bean (*Phaseolus vulgaris* L.): isolation and characterization.** *Euphytica*, 162: 301-310.
- LI, Y., KOROL, A.B., FAHIMA, T., BEILES, A., NEVO, E. (2002). **Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review.** *Molecular Ecology*, 11:2453-2465.
- LISOWSKA, K., LOCH, T., FISZER-KIERZKOWSKA A., SCIEGLINSKA, D., KRAWCZYK, A. (1997). **Identification of a microsatellite region composed of a long homopurine/homopyrimidine tract surrounded by AT-rich sequences upstream of the rat stress-inducible hsp 70.1 gene.** *Acta Biochimica Polonica*, 44: 147-152.
- MAFRA V. S. (2007). **Desenvolvimento de marcadores microssatélites para o feijão-comum (*Phaseolus vulgaris*).** Viçosa: UFV. 52 p. (Monografia de final de curso)
- MASI, P., ZEULI, P. L. S., DONINI, P. (2003). **Development and analysis of multiplex microsatellite markers set in common bean (*Phaseolus vulgaris* L.).** *Molecular Breeding*, 11: 303-313.
- MELOTTO, M., MONTEIRO-VITORETTO, C.B., BRUSCHI, A.G., CAMARGO, L.E.A. (2005). **Comparative bioinformatic analysis of genes expressed in common bean (*Phaseolus vulgaris* L.).** *Genome*, 48: 562-570.
- MÉTAIS, I., HAMON, B., JALOUZOT, R., PELTIER, D. (2002). **Structure and level of genetic diversity in various bean types evidenced with microssatellites markers isolated from a genomic enriched library.** *Theoretical and Applied Genetics*, 104: 1346-1352.
- MORGANTE M., HANAFEY, M., POWELL, W. (2002). **Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes.** *Nature Genetics*, 30: 195-200.
- MORGANTE, M., OLIVIERI, A. M. (1993). **PCR-amplified microsatellites as markers in plant genetics.** *Plant Journal*, 3: 175-182.

MULLIS, K., FALOONA, F. (1987). **Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction.** *Methods in Enzymology*, 55:335-350.

NCBI. National Center for Biotechnology Information. Disponível em <http://www.ncbi.nlm.nih.gov/>. Acesso em Julho/2009

OSTERLUND, M. T., PATERSON, A. H. (2002). **Applied plant genomics: the secret is integration.** *Current Opinion in Plant Biology*, 5: 141-145.

OSTRANDER, E. A., JONG, P. M., RINE, J., DUYK, G. (1992). **Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences.** *Proceedings of the National Academy of Science*, 89: 3419-3423.

PROJETO GENOMA CAFÉ. Disponível em <http://www.lge.ibi.unicamp.br/cafe/>. Acesso em Julho/2009.

PUNT, P.J., DINGEMANSE, M.A., KUYVENHOVEN, A. (1990). **Functional elements in the promoter region of the *Aspergillus nidulans* gpd A gene coding glyceradehyde-3-phosphate dehydrogenase.** *Gene*, 93: 101-109.

QUEIROZ, V. T. de (2004). **Obtenção de primers microssatélites e desenvolvimento, validação e mapeamento de marcadores SCAR em feijoeiro-comum.** Viçosa: Universidade Federal de Viçosa. 155p (Tese de Doutorado)

SANDALZOPOULOS, R., MITCHELMORE, C., BONTE, E., WALL, G., BECKER, P.B. (1995). **Dual regulation of the *Drosophila* hsp26 promoter in vitro.** *Nucleic Acids Research*, 23: 2479-2487.

SANGER, F., NICKLEN, S., COULSON, A. R. (1977). **DNA sequencing with chain-terminating inhibitors.** *Proceedings of the National Academy of Science*, 74: 5463-5467.

SANGUINETTI, C.J., DIAS NETO, E., SIMPSON, A.J.G. (1994). **Rapid silver staining and recovery of PCR products separated on polyacrylamide gels.** *Biotechniques*, 17: 915-919.

SCHLOTTERER, C., WIEHE, T. (1999). **Microsatellites, a neutral marker to infer selective sweeps.** In: *Microsatellites: Evolution and Applications* (eds Goldstein, D.B., Schlotterer, C.), 238-247. Oxford University Press, Oxford.

SICARD, D., NANNI, L., PORFIRI, O., BULFON, D., PAPA, R. (2005). **Genetic diversity of *Phaseolus vulgaris* L. and *P. coccineus* L. landraces in central Italy.** *Plant Breeding*, 124: 464-472.

TAUTZ, D., RENZ, M. (1984). **Simple sequence repeats are ubiquitous repetitive components of eukaryotic genomes.** *Nucleic Acids Research*, 12: 4127-4137.

- TÓTH, G., GÁSPÁRI, Z., JURKA, J. (2000). **Microsatellites in different eukaryotic genomes: Survey and analysis.** *Genome Research*, 10: 967-981.
- WANG, X. W., KAGA, A., TOMOOKA, N., VAUGHAN, D. A. (2004). **The development of SSR markers by a new method in plants and their application to gene flow studies in azuki bean [*Vigna angularis* (Willd.) Ohwi & Ohashi].** *Theoretical and Applied Genetics*, 109: 352-360.
- WEBER, J. L. (1990). **Informativeness of human (dC- dA)_n(dG- dT)_n polymorphisms.** *Genomics*, 7: 524-530.
- WOLFE, K.H. (1994). **Similarity between putative ATP-binding sites in land plant plastid ORF2280 proteins and the FTSH/CDCH8 family of ATPases.** *Current Genetics*, 25: 379-383.
- VARSHNEY, R. K., GRANER, A., SORRELLS, M. E. (2005). **Genic microsatellite markers in plants: features and applications.** *Trends in Biotechnology*, 23: 48-55.
- YAISH M.W.F., VEGA, M. P. D. L. (2003). **Isolation of (GA)_n microsatellite sequences and description of a predicted MADS-box sequence isolated from common bean (*Phaseolus vulgaris* L.).** *Genetics and Molecular Biology*, 26: 337-342
- YU, K., PARK, S.J., POYSA, V. (1999). **Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*).** *Genome*, 42: 27-34.
- YU, K., PARK, S. J., POYSA, V., GEPTS, P. (2000). **Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.).** *The Journal of Heredity*, 91: 429-434.
- ZANE, L., BARGELLONI L., PATARNELLO, T. (2002). **Strategies for microsatellite isolation: a review.** *Molecular Ecology*, 11: 1-16.

CAPÍTULO II

DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES PARA O FEIJÃO-COMUM A PARTIR DE SEQUÊNCIAS INTERNAS SIMPLES REPETIDAS (ISSR)

RESUMO

Microssatélites ou *Simple Sequence Repeats* (SSR) são sequências de 1 a 6 pb que se repetem em série no genoma, formando sítios polimórficos. *Primers* flanqueadores são geralmente desenvolvidos a partir de bibliotecas genômicas, bibliotecas genômicas enriquecidas, e de sequências em bancos de dados. Anteriormente a essas metodologias tradicionais, técnicas de enriquecimento mais simples eram utilizadas, a exemplo do enriquecimento por RAPD-PCR, e, mais recentemente, por ISSR-PCR. Esse último vem se mostrando uma alternativa rápida, eficaz e de menor custo, que vem acrescentar-se às técnicas já utilizadas. O presente trabalho teve o objetivo de desenvolver *primers* SSR por meio da metodologia de enriquecimento por ISSR-PCR. Para tal, foi realizada amplificação por ISSR-PCR de fragmentos do genoma do feijão, clonagem dos produtos de PCR e transformação em células competentes. Foram selecionados 250 clones, dos quais 103 (41,2%) foram analisados. A análise das sequências permitiu identificar 58 motivos microssatélites (56,31%) e foi possível o desenho de 32 pares de *primers*. Destes, 10 tiveram suas condições de amplificação padronizadas, sendo que seis comportaram-se como marcadores codominantes e quatro como dominantes. Dos codominantes nenhum mostrou-se polimórfico entre 20 genótipos andinos e mesoamericanos testados. As sequências contendo motivos SSR foram submetidas à busca por similaridade com sequências em bancos de dados. Foi identificada similaridade com regiões transcritas com função não caracterizada, e com a região codificadora de uma proteína inibidora de cinase ciclina-dependente. A metodologia utilizada foi eficaz para a seleção no feijoeiro de seqüências contendo motivos SSR, a partir das quais foi possível desenhar *primers* flanqueadores. Os marcadores microssatélites desenvolvidos no presente trabalho poderão ser usados como ferramentas eficazes em diversos tipos de estudos genéticos para o feijoeiro, dentre os quais o mapeamento genético, visando contribuir de forma significativa para o incremento de programas de melhoramento da cultura.

1. INTRODUÇÃO

Microsatélites ou *Simple Sequence Repeats* (SSR) são seqüências simples repetidas de DNA distribuídas ao longo do genoma de organismos, formando sítios com alto grau de variabilidade (LITT e LUTY, 1989). Como marcadores, apresentam várias vantagens em relação a outras classes de marcadores, são multialélicos, codominantes, altamente reprodutíveis e permitem ampla cobertura do genoma (CAIXETA et al., 2009).

O desenvolvimento de *primers* que flaqueiam regiões microsatélites tradicionalmente envolve o desenvolvimento e *screening* de bibliotecas genômicas enriquecidas e não-enriquecidas. É a metodologia tradicionalmente mais utilizada para tal fim, no entanto, constitui-se de uma técnica bastante laboriosa. Anteriormente à metodologia de bibliotecas genômicas, uma técnica de enriquecimento relativamente mais simples e que demanda menores recursos financeiros, usando marcadores RAPD (*Random Amplified Polymorphic DNA Polymerase Chain Reaction*), foi descrita. Essa metodologia envolve as etapas de amplificação por PCR usando marcadores RAPD, clonagem dos produtos de PCR em vetores comerciais, transformação em células competentes e posterior *screening* por clones que possuam insertos com seqüências SSR. No entanto, o polimorfismo detectado por marcadores RAPD reflete tanto a variação nas regiões microsatélites quanto mutações pontuais, inserções ou deleções de seqüências em outras regiões do genoma. Tal fato caracteriza a maior limitação dessa metodologia, uma vez que contribui para a baixa eficiência na seleção de clones que contenham seqüências com motivos microsatélites. Mais recentemente, uma técnica alternativa que utiliza marcadores ISSR (*Internal Simple Sequence Repeats*) tem sido usada e sugerida como alternativa ao enriquecimento a partir de amplificação por marcadores RAPD, uma vez que tem como alvo apenas regiões ricas em motivos microsatélites, refletindo assim, somente o polimorfismo gerado nesses locos (NEST et al., 2000).

Devido à grande aplicabilidade nos mais diversos tipos de estudos, diversos trabalhos são encontrados na literatura científica que descrevem o desenvolvimento de marcadores microsatélites para o feijão, tradicionalmente, por meio da busca em bancos de dados (YU et al., 1999; YU et al., 2000; BLAIR et al., 2003; MASI, 2003; GUERRA-SANZ et al., 2004), e da construção de

bibliotecas genômicas enriquecidas (GAÍTAN-SOLÍS et al., 2002; MÉTAIS et al., 2002; YAISH e PÉREZ DE LA VEGA., 2003; BUSO et al., 2006; BENCHIMOL et al., 2007; HANAI et al., 2007; BLAIR et al., 2008; L'TAIEF et al., 2008), sendo a metodologia de desenvolvimento e análise de bibliotecas genômicas a mais frequentemente utilizada para a cultura do feijoeiro. Ainda não foi descrito na literatura o desenvolvimento de marcadores microssatélites para o feijão, a partir da metodologia de enriquecimento por ISSR-PCR.

Marcadores ISSR ou Sequências Internas Simples Repetidas surgiram como um sistema alternativo de análise molecular, com a confiabilidade e vantagens dos marcadores microssatélites. São marcadores baseados em PCR e a técnica envolve a amplificação de seqüências situadas entre motivos microssatélites localizados próximos no genoma. Os *primers* ISSR são constituídos de motivos repetitivos, caracterizando-se verdadeiras seqüências microssatélites, que geralmente apresentam ancorados na sua extremidade 5' ou 3', resíduos de 1-4 purinas ou pirimidinas. As seqüências repetidas no genoma são localizadas arbitrariamente pelos *primers* ISSR e o padrão de amplificação de fragmentos observados em géis de eletroforese revela um número bem maior de fragmentos por *primer* do que, por exemplo, marcadores RAPD (WANG, 2003). Os ISSR são marcadores multialélicos, com herança mendeliana, dominantes, apresentam alta reprodutibilidade devido ao comprimento longo do *primer* e à alta temperatura de anelamento geralmente requerida, fornecem *fingerprint* altamente polimórfico, além de apresentarem a facilidade de poderem ser desenvolvidos sem a necessidade de conhecimento prévio de alguma seqüência alvo específica. A técnica de amplificação utilizando marcadores ISSR fornece um sistema rápido, confiável e altamente informativo para o *fingerprinting* de cultivares, para a análise de diversidade, variabilidade genética e mapeamento genômico (WANG, 2003).

Em plantas, vários trabalhos envolvendo o uso de tais marcadores já foram realizados para diferentes culturas, em estudos de diversidade genética, identificação de genótipos, mapeamento genômico e estudos de ligação de genes específicos. Para a cultura do feijoeiro, Galvan et al. (2003) realizaram estudos de diversidade genética e observaram que, em relação aos marcadores RAPD, esses marcadores se mostraram ótimas ferramentas para identificar molecularmente feijões de acordo com seu *pool* gênico de origem, distinguindo-os em feijões andinos e mesoamericanos. Métais et al. (2000)

compararam as técnicas de RFLP, ISSR e RAPD na detecção de polimorfismos entre 24 cultivares comerciais de feijão e concluíram também a efetividade desses marcadores. O programa de Melhoramento do Feijoeiro do BIOAGRO/UFV tem atualmente desenvolvido bibliotecas genômicas enriquecidas para o desenvolvimento de marcadores microssatélites. Entretanto, tem-se buscado também metodologias alternativas. A amplificação utilizando marcadores ISSR pode ser usada como ferramenta eficaz para a obtenção de sequências enriquecidas para motivos microssatélites, e assim permitir o desenvolvimento de *primers* SSR flanqueadores. É descrito que o genoma de plantas possui regiões ricas em motivos microssatélites, nas quais geralmente se concentram SSR distintos (LI et al., 2002), assim sendo, os fragmentos amplificados por ISSR-PCR têm alta probabilidade de possuírem microssatélites, principalmente quando se combinam mais de um *primer* ISSR diferente em uma mesma reação PCR, caracterizando o enriquecimento (GALVAN et al., 2003). Os fragmentos obtidos por PCR são então purificados, clonados em vetor apropriado e transformados em células competentes. Os clones transformantes-recombinantes são seqüenciados e submetidos ao *screening* por motivos SSR e desenho de *primers* flanqueadores.

NEST et al. (2000) trabalhando com eucalipto, desenvolveram e validaram *primers* SSR, a partir da metodologia de enriquecimento por ISSR-PCR. Para tal, foram utilizados os genótipos TAG 5 e ZG 14, de *Eucalyptus grandis*. Vinte clones transformantes-recombinantes foram obtidos e todos continham sequências com motivos microssatélites. A partir dos SSR identificados, treze *primers* flanqueadores foram desenhados e testados quanto ao polimorfismo detectado entre os dois genótipos analisados. Desses, dez se mostraram polimórficos.

A alta eficiência alcançada na seleção de clones contendo motivos SSR, obtida por NEST et al., sugere a utilização dessa metodologia como uma boa alternativa para o desenvolvimento de marcadores microssatélites principalmente quando há poucas seqüências disponíveis em bancos de dados para uma determinada espécie de interesse, como no caso do feijoeiro, ou quando não há o acesso a bibliotecas genômicas ou a bibliotecas genômicas enriquecidas. O presente trabalho teve o objetivo de desenvolver marcadores

microsatélites a partir da metodologia de enriquecimento por ISSR-PCR e caracterizar putativas funções relacionadas com os motivos identificados, com o propósito de fornecer marcadores informativos que possam ser utilizados como ferramentas eficazes para os mais diversos tipos de estudos genéticos na espécie *Phaseolus vulgaris*, contribuindo para o incremento de programas de melhoramento do feijoeiro que utilizam marcadores moleculares.

2. MATERIAL E MÉTODOS

2.1. Material genético e Extração de DNA

O cultivar de feijão TO foi utilizado como fonte de DNA genômico para a construção da biblioteca. Este cultivar é do *pool* gênico mesoamericano e teve origem no cruzamento de “Tanderette” (suscetível a antracnose) com a linhagem mexicana “Acapulco” (resistente a antracnose). Além disso, possui o gene dominante *Co-4* de resistência à antracnose (FOUILLLOUX, 1976). Os cultivares andinos e mesoamericanos descritos na Tabela 1 foram utilizados para teste de polimorfismo dos *primers* SSR desenvolvidos no presente trabalho. O DNA foi extraído de folhas jovens, pelo método do CTAB descrito por DOYLE e DOYLE (1990), com modificações (COSTA, 2007).

Tabela 1: Genótipos andinos e mesoamericanos utilizados no teste de polimorfismo dos primers microssatélites desenvolvidos.

Genótipo	Pool gênico	Genótipo	Pool gênico
AND 277	Andino	G 5686	Mesoamericano
RUDÁ	Mesoamericano	TO	Mesoamericano
JALO EEP 558	Andino	VC3	Mesoamericano
CAL 143	Andino	MÉXICO 54	Mesoamericano
BJ8	Andino	MAR 2	Mesoamericano
BOLON BAYO	Andino	BAT 332	Mesoamericano
BJ3	Andino	TALISMÃ	Mesoamericano
BJ6	Andino	DOM TIMÓTEO	Mesoamericano
JALO M665	Andino	AMENDOIM	Mesoamericano
G 11796	Andino	OURO NEGRO	Mesoamericano

2.2. Amplificação por PCR de fragmentos ricos em seqüências microssatélites, utilizando marcadores ISSR

Em cada reação de amplificação com os *primers* ISSR foram utilizados 30 ng de DNA, Tris-HCl 12 mM (pH 8,3), KCl 60 mM, MgCl₂ 2 mM, cada um dos desoxinucleotídios (dATP, dTTP, dGTP e dCTP) em uma concentração de

100 µM, 0,3 µM de cada *primer*, 1U da enzima Taq DNA polimerase, em um volume de 15 µL. Foram realizadas diferentes reações utilizando combinações de um, dois e três *primers* ISSR, com o objetivo de definir qual a melhor combinação a ser usada em procedimentos posteriores. Na Tabela 2 estão indicados os *primers* ISSR utilizados. O procedimento foi realizado em um termociclador Perkin Elmer-Cetus, modelo 9600, programado para um ciclo inicial a 94 °C por 4 min, seguido de 35 ciclos a 94 °C por 1 min, 50 °C por 1 min e 72 °C por 2 min, e uma etapa de extensão final a 72 °C por 7 min. E posterior redução da temperatura a 4 °C. Os produtos de amplificação foram submetidos à uma eletroforese em gel de agarose 1,8%, corados com brometo de etídeo e fotodigitalizados utilizando o sistema L-PIX (Loccus Biotecnologia, São Paulo, Brasil).

Tabela 2: Seqüências dos *primers* ISSR usados na amplificação do DNA de TO e suas respectivas temperaturas de amplificação.

Seqüência do <i>Primer</i>	Temperatura de amplificação
(AG) ₈ C	50 °C
(AGG) ₆	50 °C
(GACA) ₆	50 °C
(AC) ₈ G	50 °C
(AC) ₈ YT	50 °C

Abreviatura: Y= (C,T)

Após a definição das melhores combinações a serem utilizadas, novas reações de amplificação por ISSR foram realizadas, em um volume final de 100 µL, utilizando 30 ng de DNA, Tris-HCl 12 mM (pH 8,3), KCl 60 mM, MgCl₂ 2 mM, cada um dos desoxinucleotídios (dATP, dTTP, dGTP e dCTP) em uma concentração final de 2,5mM, 0,3 µM de cada *primer* ISSR, e 1U da enzima GoTaq DNA polimerase (Promega, Madison, WI, EUA). As condições de termociclagem mantiveram-se as mesmas descritas no parágrafo anterior. Para a verificação da qualidade dos produtos de amplificação gerados, uma alíquota de 5 µl foi submetida à eletroforese em gel de agarose 1,8%, corado com brometo de etídeo e fotodigitalizado utilizando o sistema L-PIX (Loccus Biotecnologia, São Paulo, Brasil). O restante do conteúdo de todas as reações

foi purificado com o kit Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, EUA), segundo recomendações do fabricante.

2.3. Ligação dos fragmentos de DNA ao vetor e Transformação de células competentes

Os fragmentos purificados foram clonados em plasmídeo pGEM T-Easy Vector (Promega, Madison, WI, EUA), segundo recomendações do fabricante. Para a ligação dos insertos ao plasmídeo, foi utilizado 15 ng de DNA, 50 ng de plasmídeo, 1X Rapid Ligation Buffer – T4 DNA Ligase e 3U da enzima T4 DNA Ligase, em um volume final de reação de 10 µl. O plasmídeo utilizado é um vetor linearizado que possui um resíduo de timidina no terminal 3' de ambas as extremidades do sítio de clonagem. Essa característica aumenta a eficiência da ligação dos produtos de PCR ao sítio de clonagem, por evitar a recircularização do vetor e por fornecer uma extremidade coesiva favorável à ligação de fragmentos amplificados por DNA polimerases termoestáveis com a propriedade de acrescentar um resíduo de adenina à extremidade dos fragmentos, a exemplo da GoTaq DNA polimerase, utilizada nas reações de ISSR-PCR descritas no item 2.2.

Após a clonagem, foi realizada a transformação de células ultra-competentes da cepa DH5α de *Escherichia coli*, pelo protocolo padrão de transformação por choque térmico descrito por Ausubel (1998). As células transformadas foram plaqueadas em placas de petri contendo 20 ml de meio LB sólido, suplementado com 100mg/ml de antibiótico ampicilina, 10 µl de IPTG 0,1 M e 40 µl de X-gal 2X.

2.4. Seleção de clones positivos e Sequenciamento

Após transformação por choque térmico, foi feita a seleção de colônias brancas, e verificação de transformantes-recombinantes por meio da técnica de PCR de colônia, utilizando os *primers* M13 *Foward* e *Reverse*. Os clones que apresentaram insertos com tamanho de no mínimo 400 pb foram selecionados para o sequenciamento. O DNA plasmidial das colônias de interesse foi extraído, utilizando o Kit de extração Wizard® *Plus* SV Minipreps DNA Purification System, segundo recomendações do fabricante (Promega,

Madison, WI, EUA). Após a extração e quantificação do DNA plasmidial, as amostras foram seqüenciadas em um sequenciador automático Applied Biosystems modelo 3730xl, na empresa Macrogen (Seoul, Coréia do Sul), pelo método de terminação da cadeia por didesoxirribonucleotídeos (SANGER *et al.*, 1977).

2.5. Identificação de regiões SSR e Desenho de *primers*

As seqüências que apresentaram valor PHRED de qualidade maior que 20, tiveram seu perfil cromatográfico analisados no programa CHROMAS (<http://www.technelysium.com.au/chromas>). O valor PHRED=20 está relacionado à probabilidade estimada de 1 erro a cada 100 bases seqüenciadas, ou seja, corresponde a 99% de certeza de que um nucleotídeo apresentado em uma dada posição no perfil cromatográfico corresponde à realidade. Esse valor é comumente padronizado na literatura como o valor PHRED mínimo para que uma seqüência seja considerada de boa qualidade para análise. Adicionalmente, foram consideradas seqüências de boa qualidade aquelas cujo cromatograma apresentou picos de fluorescência agudos, regularmente espaçados e com pouco ou nenhum ruído de fundo. Em seguida, foi utilizada a ferramenta VecScreem para identificar quais os segmentos das seqüências de DNA eram correspondentes ao vetor plasmidial utilizado durante a etapa de clonagem. O VecScreem realiza uma análise de similaridade, por meio do algoritmo BLAST, entre a seqüência de entrada e um banco de dados especializado e não-redundante de seqüências de vetores usualmente utilizados em experimentos de clonagem, o UniVec, identificando, na seqüência de interesse, as regiões correspondentes ao vetor plasmidial. As três ferramentas descritas encontram-se disponíveis na página do NCBI (NCBI, 2009). Logo após, foi utilizado o programa MEGA4 (<http://www.megasoftware.net/mega.html>), para alinhamento global das seqüências *forward* e *reverse* advindas do mesmo clone, e obtenção da seqüência consenso. Quando o fragmento de DNA genômico clonado era grande, não era possível um alinhamento de boa qualidade das seqüências *forward* e *reverse*, pois a fração das seqüências que apresentavam sobreposição era geralmente pequena, o que dificultava a obtenção de uma seqüência consenso. Nesses casos foi utilizado o programa CAP3

(<http://pbil.univ-lyon1.fr/cap3.php>), que realiza o agrupamento de sequências contíguas, ou seja, sequenciais, gerando *contigs* e fornecendo a sequência consenso. O programa CAP3 foi utilizado também para a identificação de clones redundantes. Seguiu-se a análise das sequências quanto à presença de motivos microssatélites, por meio da ferramenta SSRIT (<http://www.gramene.org/>). Não existe um critério padrão estabelecido na literatura para a definição de um potencial microssatélite, assim sendo, foi seguido o seguinte critério mínimo: A) (NN)₅ – 5 repetições de dinucleotídeos; B) (NNN)₄ - 4 repetições de trinucleotídeos e C) (NNNN)₃ – 3 repetições de tetranucleotídeos, segundo Caixeta et al. (2005). No entanto, alguns motivos menores foram igualmente selecionados por estarem inseridos em um motivo maior, porém composto.

Os *primers* microssatélites foram desenhados com auxílio do programa Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Para o desenho dos *primers* foram seguidos os critérios recomendados por Yu et al. (1999), com modificações: A) *Tamanho dos primers*: de 18 pb a 30 pb; 2) *T_m*: mínimo de 48 °C e máximo de 60 °C, com ótimo de 52 °C e *T_m* máxima de 5 °C entre os *primers forward* e *reverse* da mesma sequência ; 3) *Concentração de sal*: 50 mM; 4) % GC: mínimo de 20% e máximo de 80%; 5) *Produtos de amplificação*: entre 100 – 400 pb. Após o desenho dos *primers* foi verificada a estabilidade dos mesmos por meio do programa *Primerselect* 3.1 (DNASTAR, Madison, WI, EUA).

O *Primerselect* 3.1 analisa a possibilidade de formação de heterodímeros entre os *primers forward* e *reverse*, identifica a existência de auto-complementariedade e formação de homodímeros, e ainda identifica a possível formação de estruturas secundárias (*hairpin*) durante a futura reação de PCR, por meio de padrões termodinâmicos pré-estabelecidos. O programa *Primerselect* considera que os seguintes fatores são indispensáveis para um *primer* ser considerado estável: 1) *Estabilidade do pentâmero 3'* - é um valor limite de estabilidade para os cinco últimos nucleotídeos do terminal 3'. *Primers* com o valor de energia livre (G) do pentâmero menor do que -8,5 kCal/mol apresentam uma tendência de pareamento em regiões inespecíficas e são mais prováveis de formar estruturas em grampo e dímeros com ele mesmo. O *primer* ideal deve possuir um gradiente decrescente de G que vai de alta estabilidade na extremidade 5' (G < -8,5 kCal/mol) até baixa estabilidade na

extremidade 3' ($G > -8,5$ kcal/mol). 2) *Seqüência 3' única de 7 bases* – este critério limita o pareamento inespecífico, usando este critério, qualquer *primer* que não apresente uma seqüência específica de 7 bases ou superior que seja exatamente complementar ao sítio de anelamento é eliminado. 3) *Formação de dímeros* – este critério limita no valor 2 a 3 o número de pares de bases que o dímero pode conter. O valor 3 raramente altera o desempenho do *primer*, a menos que o pareamento seja próximo à extremidade 3'. 4) *Formação de estruturas em grampo* - este critério limita no valor 2 o número de pares de bases que a estrutura pode conter. O valor 3 raramente altera o desempenho do *primer*, a menos que o pareamento seja próximo à extremidade 3'. 5) *Formação de fita dupla próximo à extremidade 3'* – este critério elimina *primers* candidatos que apresentem qualquer dímero ou estrutura em grampo a uma distância de 8 pares de base da extremidade 3'.

2.6. Determinação das condições de amplificação e Identificação de marcadores microssatélites polimórficos entre os genótipos andinos e mesoamericanos

Os *primers* analisados no programa *PrimerSelect* e considerados estáveis termodinamicamente foram sintetizados por empresa especializada (Sinapse, São Paulo/SP, Brasil). Para a determinação e validação das condições ótimas de amplificação de cada *primer* desenvolvido, os cultivares Rudá e AND277 foram utilizados como fonte de DNA para a realização de reações SSR-PCR, nas quais os *primers* foram submetidos a um gradiente de temperaturas de amplificação.

As reações de amplificação no teste de gradiente foram realizadas em um volume de 15 μ L, contendo 30 ng de DNA, Tris-HCl 12 mM (pH 8,3), KCl 60 mM, $MgCl_2$ 2 mM, cada um dos desoxinucleotídios (dATP, dTTP, dGTP e dCTP) numa concentração de 100 μ M, 0,3 μ M de cada *primer* e 1U da enzima *Taq* DNA polimerase. O procedimento foi realizado em um termociclador Eppendorf modelo Mastercycler EP Gradiente, programado para uma etapa inicial a 94 °C por 4 min, seguida de 40 ciclos a 94 °C por 40 s, 48-62°C por 40 s e 72 °C por 60 s, e uma etapa final a 72 °C por 7 min. Os produtos de amplificação foram submetidos à eletroforese em gel de agarose 1,8%/SB 1X,

corados com brometo de etídeo e fotodigitalizados utilizando o sistema L-PIX (Loccus Biotecnologia, São Paulo, Brasil).

Após a determinação da melhor temperatura de amplificação, os *primers* foram testados quanto ao polimorfismo detectado entre os cultivares descritos na Tabela 1 do item 2.1, que são genitores usados no Programa de Melhoramento do Feijoeiro do BIOAGRO/UFV. As reações de amplificação para cada um dos *primers* padronizados foram realizadas sob as mesmas condições usadas para as reações de padronização. As temperaturas de amplificação definidas e utilizadas no teste de polimorfismo, para cada *primer*, estão indicadas na Tabela 5 do item 3.5, na sessão “Resultados e Discussão”. Os produtos de amplificação foram submetidos à eletroforese em gel de poliacrilamida nativo 10%, contendo acrilamida/bis-acrilamida na proporção de 29:1; tampão TAE 10X (Tris-base 400 mM, ácido acético 200 mM e EDTA 10 mM); na presença de 167,5 µl de persulfato de amônio 10% e 167,5 µl do catalisador TEMED, para um volume final de 25ml de gel. Os géis foram submetidos a uma voltagem de 120 volts por 4 h e corados com solução de nitrato de prata, de acordo com o protocolo proposto por Sanguinetti et al. (1994), com modificação do tempo de imersão dos géis em cada uma das quatro soluções usadas no procedimento para 10 minutos.

2.7. Análises de similaridade

As regiões microsatélites identificadas, e para as quais puderam ser desenhados *primers* flanqueadores, foram submetidas à busca por similaridade com sequências já caracterizadas e depositadas no consórcio de bancos públicos de sequências GenBank (EUA), EMBL (Europa) e DDJ (Japão), disponível no endereço eletrônico do NCBI (NCBI, 2009).

Primeiramente, foi utilizada a ferramenta *PrimerBLAST* para analisar se os *primers* desenvolvidos amplificavam sequências do genoma de *Phaseolus vulgaris* e de outras espécies que já haviam sido caracterizadas e depositadas no banco de sequências. Para tornar os parâmetros menos rigorosos selecionou-se a opção *nr (non-redundant)*, na seção *database*, que permite uma busca mais ampla, pois engloba todos os bancos de dados não redundantes disponíveis no NCBI (bancos curados e não curados, de

sequências expressas e não-expressas, para todas as espécies e não somente para *Phaseolus vulgaris*).

Logo em seguida, as sequências amplificadas pelos *primers* SSR desenvolvidos, obtidas no item 2.10., foram submetidas a uma busca de similaridade contra o banco de dados completo do NCBI, via BLASTN, algoritmo que identifica similaridade entre a sequência de nucleotídeos de entrada (*query*) com as demais sequências de nucleotídeos depositadas. Quando a busca identificou similiaridade com múltiplas sequências disponíveis no banco de dados, foi considerado o primeiro *hit*, ou seja, a sequência que apresentou, em conjunto, a maior porcentagem de similaridade com a sequência de entrada (85%-100%), associada ao maior *score* de alinhamento e ao menor *e-value* (*e-value* máximo de e^{-10}). Todas as ferramentas acima descritas encontram-se disponíveis na página do NCBI (NCBI, 2009).

Para identificação e categorização das funções celulares associadas às sequências gênicas identificadas pela análise de similaridade, foi realizada uma pesquisa ao banco SwissProt (<http://www.expasy.ch/sprot>) e ao banco GO (*Gene Ontology*) (<http://www.geneontology.org/>), respectivamente.

3. RESULTADOS E DISCUSSÃO

3.1. Análise da amplificação por PCR utilizando marcadores ISSR

O perfil eletroforético em gel de agarose dos produtos de PCR obtidos a partir da amplificação por marcadores ISSR foi analisado, de modo a definir qual a quantidade de *primers* ISSR diferentes utilizada em cada reação (uso de apenas um *primer* distinto, da combinação de dois *primers*, e de três *primers*) seria a mais adequada para ser adotada. Para tal, as combinações que originaram produtos de amplificação com padrão mais definido de bandas foram selecionadas para serem utilizadas nas etapas posteriores de enriquecimento por ISSR-PCR (Figura 1).

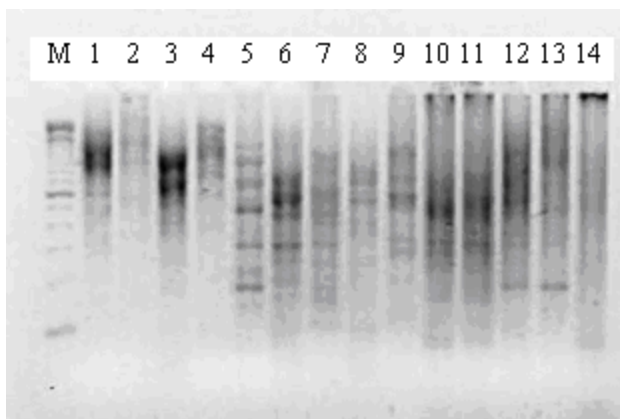


Figura 1: Perfil eletroforético em gel de agarose 1,8%/SB 1X de produtos de PCR obtidos a partir da amplificação por marcadores ISSR. Canaletas 1-4 corresponde ao uso de apenas um tipo de *primer* ISSR na reação, canaletas 5-9 à combinação de dois *primers* ISSR distintos e as canaletas 10-14 à combinação de três *primers* ISSR. M- Marcador de tamanho de 100 pb.

O uso de apenas um tipo de *primer* e da combinação de dois *primers* foram as condições selecionadas. A combinação de três *primers* distintos não foi adotada por apresentar um perfil de arraste e de não definição de bandas dos produtos de PCR gerados. O perfil eletroforético observado ocorreu provavelmente devido à amplificação de fragmentos de tamanhos muito próximos que não se distinguiram visualmente no gel.

Após a definição de quantos *primers* distintos seria adequado utilizar, novas reações de amplificação foram realizadas com as condições selecionadas de um *primer* e combinação de dois *primers*, em um volume final

maior, de modo a gerar produtos de PCR em volume adequado para ser utilizado em etapa posterior de clonagem em vetor plasmidial. Na Tabela 3 encontra-se descrita a relação dos *primers* utilizados e das combinações adotadas.

Tabela 3: Relação dos *primers* ISSR usados nas reações de amplificação individuais, e as combinações de dois *primers* utilizadas.

Primers ISSR individuais	Combinações de dois primers ISSR
(AG) ₈ C	(AG) ₈ C e (AGG) ₆
(AGG) ₆	(AG) ₈ C e (GACA) ₄
(GACA) ₆	(AG) ₈ C e (AC) ₈ G
(AC) ₈ G	(AG) ₈ C e (AC) ₈ YT
(AC) ₈ YT	(AGG) ₆ e (GACA) ₄
	(AGG) ₆ e (AC) ₈ G
	(AGG) ₆ e (AC) ₈ YT
	(GACA) ₄ e (AC) ₈ G
	(GACA) ₄ e (AC) ₈ YT
	(AC) ₈ G e (AC) ₈ YT

Uma alíquota de 5 ul dos produtos de PCR obtidos foram submetidos à eletroforese em gel de agarose 1,8% para verificação da qualidade dos produtos gerados (Figura 2), na ordem correspondente à indicada na Tabela 2. Após a verificação do padrão de amplificação, o restante da reação foi purificado para posterior etapa de clonagem.

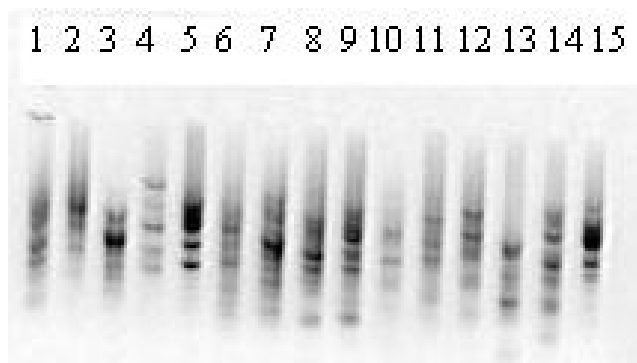


Figura 2: Perfil eletroforético em gel de agarose 1,8%/SB 1X dos produtos de amplificação obtidos por ISSR-PCR, a partir da utilização de um *primer* distinto (1-5) e da combinação de dois *primers* (6-15). M- Marcador de tamanho de 100 pb.

SSR e um número arbitrário a partir de 001 para sequências que possuam a mesma unidade repetitiva em maior número. Quando a sequência possuía mais de uma unidade com o mesmo número de repetições, foi considerada a primeira unidade repetitiva para efeito de nomenclatura. As designações F e R foram utilizadas para diferenciar os *primers forward* e *reverse*, respectivamente. Por exemplo, para o *primer* PVga001F, F designa o *primer* SSR *forward* flanqueando repetições do motivo GA. A Tabela 4 mostra os motivos microsatélites identificados e as sequências dos *primers* que foram desenhados.

Tabela 4: Motivos microsatélites identificados, sequências dos *primers* correspondentes, valores de T_m, tipo de repetição e tamanho dos fragmentos amplificados. O símbolo // entre alguns motivos denota a presença de outros motivos SSR relativamente distantes no mesmo fragmento.

Motivo Microsatélite	Primer	Seqüência	T _m (°C)	T. frag. (pb)
(AT) ₂ (TCT) ₂ (GA) ₂ //(ATT) ₂ (CATTT) ₂ //(TC) ₅ (TA) ₄ //(AT) ₅ (AC)(AT) ₄	PVtc008F	TGAGGGACATCAAACACTACTA	50,56	352
	PVtc008R	AAATGACACTGAGACACTG	48,47	
(GA) ₄ (TC) ₅	PVga002F	GTGACCTAAAAGTGAAAACA	55,29	250
	PVga002R	CTAATCCTCATCATCACATC	51,41	
(AC) ₅ //(AC) ₅	PVac001F	CTAATCCTCATCATCACATC	51,93	332
	PVac001R	ACCACACTCGCTGAAAAA	51,92	
(GAA) ₄ (AT) ₂ //(AT) ₄	PVgaa004F	TCAAGAATCATAAGGAAAGA	52,27	238
	PVgaa004R	GATAAACCTGTCACTTGGTA	52,10	
(AC) ₅ //(AC) ₅	PVac002F	ATTGCATAAACTTCGTTCT	49,71	333
	PVac002R	ACCACACTCGCTGAAAAA	49,51	
(TTC) ₄ (AT) ₃	PVttc001F	GATAAACCTGTCACTTGGTA	56,86	239
	PVttc001R	TCAAGAATCATAAGGAAAGA	51,41	
(GAA) ₄ (GA) ₄	PVgaa003F	TCAAGAATCATAAGGAAAGA	50,4	260
	PVgaa003R	GATAAACCTGTCACTTGGTA	50,2	
(GAA) ₄ (CG) ₄	PVgaa004F	TCAAGAATCATAAGGAAAGA	50,4	238
	PVgaa004R	GATAAACCTGTCACTTGGTA	49,8	
(AG) ₄ //(AAG) ₄ //(CG) ₄	PVag001F	CAAAACAGAAAACCTAACAC	48,6	229

	PVag001R	CGAACATTCTCTTCTTATTC	50	
(TC) ₅ //(AG) ₅ (AT) ₂	PVtc006F	GTTAGGTGAGATGTGTCTTG	50,6	159
	PVtc006R	CATTTCTACCTTTCTAAGCA	49,7	
(AG) ₅ (AT) ₃	PVag002F	TACAACATAACTGCGTCCT	54,5	210
	PVag002R	GTCATTATCACAAATTCGTC	50,2	
(AAAT) ₂ A ₂ C(AGTA) ₂ //(GTA) ₃ (TTA)A(TTA))/(AGGT) ₂	PVgta001F	TCTCACTCATCAGACTCAC	48.18	306
	PVgta001R	ATATTAAGCTATGGGACTGC	51.33	
(AT) ₃ (TA) ₂ (AT) ₂ (AG) ₂ //(AAT) ₂ (GT)(CA) ₃ (AG) ₂ (GT) ₂	PVat001F	CAGTTTTGCGTGTGCTAC	53.64	173
	PVat001R	TGATTCTCATGTCGTGAA	50.99	
(AT) ₃ //(AACAG) ₂ //(G CA) ₂ (CA) ₂	PVat001 F	GACAGACACTAGTAAGCCTCA	52.18	205
	PVat001 R	GGGAATTCGATTAGAAAG	49.64	
(TAAA) ₂ (TAC) ₃ //(TC) ₄ C(TCT) ₂ //(CA) ₂ C(C A) ₂ C(CA) ₂ (CT) ₃ T ₂ (CT) ₂ C(CT) ₂	PVtc009F	ACAACCCAAATTAATACATCT	50.60	250
	PVtc009R	AGAGTAGAATGCCTTCACAG	52.06	
(GA) ₄ (CT)(GA) ₂	PVga007F	GTGTCTACTGATCTCCCAGA	52.40	280
	PVga007R	AAGCCTAGAAGAAGTTGAG	48.82	
(GAGT) ₃ //(GAA) ₃ //(A GA) ₂ TA ₂ (GAA) ₂ //(GA) ₆ A(GA)	PVga008F	GACATACCCATCTCACAAAC	52.63	212
	PVga008R	ATACATTAGCCCAACAACAT	52.15	
(TGGAAG) ₂ T ₂ (GC) ₂ (CG)G(CG)//(CCGG GC) ₂	PVcg002 F	AAATAATAACGACCAAGCAC	52.13	244
	PVcg002 R	CTGATCTCAAGGGATGATT	52.27	
(CT) ₂ (TGA) ₂ //(CA) ₄ (GA)(CA)(GAG) ₂	PVca003 F	GTAAACGGAAGGCAGTAG	52.65	181
	PVca003 R	AAGATTTTGAGGATAGGTT	52.11	
(GT) ₂ (TTGCA) ₂ A(CA AAC) ₂ (ACC) ₂ //(AT) ₄ (GT)(AT) ₃	PVat002 F	AAGATCTCTTTTGGGGTTAT	52.11	294
	PVat002 R	AACGTACAATCATAGTTTTTCA	51.92	
(TA) ₂ T(TA) ₂ T(TA)AT (TA)	PVta001F	AGAAAATTGAACACGGTCTA	52.44	203
	PVta001R	GCCTTGATATATTGACGAAA	52.98	
(GTAA) ₂ TA(TTG) ₂ //(AG) ₃ (TG)G(TG)//(TC) ₃ T(TC)	PVtc010F	ATAATTCACCGGATTCAAA	52.98	203
	PVtc010R	TAACACAAGTATGGAAGCAA	51.87	
(GAA) ₂ T(GAA) ₂ (GA) G(GA)A(GA)	PVga009F	ACAACTCACCCAACAAC	52.75	196
	PVga009R	ATAATACCTACCAACGCAAC	52.00	
(TC) ₂ T(TG) ₂ T(TG) ₂	PVgga001F	GGTTAGAGGAGGGGAGTAT	52.19	207

(TG) ₃ G(GGA) ₄	PVgga001R	CAAACAAAGAACAAGTGACA	52.10	
(GATAA)G(GATAA) G(GAT) ₃ (GA)G(GA)	PVgat001F	AAGTGACATTGTGAAGGTT	50.03	158
	PVgat001R	ACAAATGGAACCTGGTAAA	50.49	
(CCA) ₂ T(CT) ₃ C(CT) T(CT) ₃ T ₂ (CT) ₂ //(AG) G ₂ (AG) ₃	PVct004 F	TCTGTTGCTGTTCCCTGAG	52.86	272
	PVct004 R	GTTTTTCATGGCAAGATACT	51.97	
(AT) ₃ A(AT) ₂ //(CTT) ₂ A(AT) ₂ T(AT)	PVat003 F	CGTGAATAAAATTAGTGGAAC	49.98	172
	PVat003 R	ATGGATAAAAAGTTTGATTGG	51.83	
(GT) ₉	PVgt005F	TTGGCCTATTATTTGAACT	50.03	126
	PVgt005R	TTCCCTGATACCTATGTAGTG	51.90	
(GT)G(GT) ₄	PVgt006F	CTTTGATAAGTGGAGGAA	48.11	270
	PVgt006R	AAAGCATTGCAGATTAGAGT	51.88	
(GAC) ₂ (GAG)(GAC) GA(GGT) ₂ //(CGTTG G) ₂	PVgac001F	ATCCACAGAATAAAAGAAAG	48.74	214
	PVgac001R	GGAAGGAAAAGGTAAAAGT	52.03	
(CTT)TCG(CTT)(CT TC) ₃	PVcttc001F	TGGATCAAATAAAACTGA	50.87	196
	PVcttc001R	CCAGAAAGATCACAACGA	52.64	
(AG) ₃ T(GAA) ₂ //(GAA)) ₂	PVag007F	TTAAGCCTTTGTTTGTATC	48.04	299
	PVag007R	ACATTGTACGAATCTGTGTT	50.28	

Como pôde ser observado na Tabela 4, não houve predominância de nenhuma classe específica de motivos microssatélites, dentre os identificados no presente trabalho. No entanto, Yu et al. (1999), Melotto et al. (2005), Hanai et al. (2007) e Benchimol et al. (2007) observaram em seus trabalhos uma maior freqüência de microssatélites constituídos do motivo AG/TC, no genoma do feijoeiro. Entretanto, essa correlação não foi encontrada no presente trabalho. Foi observado também que, dos 58 motivos microssatélites indentificados, 31 (53,45%) foram obtidos a partir da amplificação por PCR que utilizou a combinação de dois *primers* ISSR e 27 (46,55%) obtidos a partir da amplificação por apenas um tipo de *primer* distinto. O resultado obtido demonstrou que ambas as estratégias foram eficazes na seleção de regiões microssatélites, a partir do genoma do feijão.

A partir da análise dos locos microssatélites descritos na Tabela 4, foi identificado que o número máximo de repetições encontrado para dinucleotídeos foi (GT)₉, para o loco PVgt005F, e para trinucleotídeos foi quatro repetições, de diferentes motivos, para os locos PVgaa004F, PVttc001F, PVgaa003F, PVgaa004F, PVag001F, PVgga001F. O número máximo de repetições encontrados foi baixo, quando comparado aos resultados obtidos pela metodologia de bibliotecas genômicas enriquecidas (HANAI et al., 2007; BENCHIMOL et al., 2007; BLAIR et al., 2008; L'TAIEF et al., 2008). No entanto, resultados encontrados por Caixeta et al. (2005), Hanai et al., (2007) e Benchimol et al., (2007) demonstraram que *primers* que amplificam regiões microssatélites constituídas de motivos compostos e com número pequeno de repetições podem efetivamente serem informativos quanto ao polimorfismo detectado entre diferentes genótipos.

Na literatura científica já foram descritas diferentes estratégias de desenvolvimento de marcadores microssatélites por meio do enriquecimento utilizando marcadores ISSR, a exemplo das metodologias desenvolvidas por Nest et al. (2000), já descrita anteriormente, e por Lian et al. (2001). No último trabalho, os resultados obtidos indicaram que a metodologia adotada apresenta alta eficiência na seleção de motivos microssatélites constituídos de alto número de repetições e no desenvolvimento de *primers* que apresentaram um alto índice de conteúdo polimórfico. O método envolveu: (1) amplificação por ISSR-PCR, de modo a gerar produtos de amplificação com motivos microssatélites em suas extremidades; (2) clonagem e seqüenciamento dos fragmentos gerados e desenho de um primeiro *primer* na região entre os motivos microssatélites das extremidades, e de um segundo entre a sequência do primeiro e o motivo repetitivo, para serem usados em etapa posterior de nested-PCR; (3) construção de bibliotecas de fragmentos digeridos e ligados a adaptadores especiais que permitam via nested-PCR, juntamente com os *primers* anteriormente obtidos, que ocorra o procedimento de PCR de supressão, descrito por Siebert et al. (1995). A partir dos produtos gerados é possível identificar a outra região flangeadora do motivo microssatélite e o desenho de *primers*, de modo a permitir a amplificação. A metodologia descrita para o desenvolvimento de marcadores microssatélites foi utilizada por Guo e Luo (2006), para a fruta japonesa *Diospyros kaki* Thumb., e os resultados obtidos confirmaram a alta eficiência de enriquecimento proporcionada por

essa estratégia, além de permitir que se determine logo nas etapas iniciais do projeto qual o motivo microssatélite que se deseja posteriormente selecionar e a partir dele desenvolver *primers flanqueadores*.

Em relação à eficiência de enriquecimento das metodologias tradicionais de *screening* em bibliotecas genômicas, estudos realizados apresentaram porcentagem de seleção de clones positivos para motivos microssatélites que variaram de valores menores tais como 24,88%, obtido por Gaitán et al. (2002), até altos índices de enriquecimento como 59,06%, obtido por Benchimol et al. (2007) e 56,31% alcançado pela metodologia descrita no primeiro capítulo do presente trabalho.

Em relação à metodologia de enriquecimento por amplificação utilizando marcadores ISSR, Nest et al. (2000) obtiveram 65% de eficiência de seleção de clones positivos para motivos microssatélites. Treze *primers* foram desenhados, a partir de 20 motivos SSR identificados e, destes, dez foram polimórficos entre os genótipos de eucalipto analisados. Os resultados obtidos nesse trabalho indicam uma alta eficiência não só na seleção de clones como também no desenvolvimento de marcadores SSR informativos, a partir da citada metodologia

No presente trabalho, a análise das sequências permitiu identificar 58 motivos microssatélites, apresentando uma eficiência de 56,31% e foi possível o desenho de 32 pares de *primers*. A alta eficiência obtida na seleção de clones positivos para motivos microssatélites indica que o método utilizado foi eficaz na identificação e desenvolvimento de marcadores microssatélites para o feijão-comum.

3.3. Identificação de marcadores microssatélites polimórficos entre os genótipos andinos e mesoamericanos analisados

Dos 32 pares de *primers* desenvolvidos, 10 tiveram suas condições de amplificação padronizadas e foram analisados quanto ao polimorfismo detectado entre os 20 genótipos andinos e mesoamericanos descritos na Tabela 1 do item 2.1.

A análise do perfil eletroforético do gradiente de temperatura realizado para os 10 pares de *primers* permitiu observar que quatro *primers* comportaram-se como marcadores dominantes, apresentando padrão de presença e ausência de bandas, em vez do perfil de codominância

comummente apresentado pelos marcadores microssatélites. Os demais 6 *primers* analisados comportaram-se como codominantes.

A Tabela 5 apresenta a temperatura de amplificação padronizada para cada um dos 10 pares de *primers* analisados e quais apresentaram padrão de dominância e de codominância. Os demais 22 pares de *primers* não analisados ainda serão padronizados e testados em próxima oportunidade.

Tabela 5: Temperatura de amplificação padronizada para os 10 pares de *primers* analisados e identificação de quais apresentaram padrão de codominância e de dominância.

Primer	Temperatura de amplificação (°C)	Padrão apresentado
PVtc008	54 °C	Codominante
PVga002	54 °C	Codominante
PVac001	54 °C	Dominante
PVgaa004	54 °C	Codominante
PVac002	55 °C	Codominante
PVttc001	55 °C	Dominante
PVgaa003	55 °C	Dominante
PVgaa004	55 °C	Dominante
PVag001	55 °C	Codominante
PVtc006	55 °C	Codominante

Os 10 marcadores que tiveram as condições de amplificação padronizadas foram testados quanto ao polimorfismo detectado entre os 20 genótipos analisados, em gel de poliacrilamida 10%. Nenhum dos *primers* codominantes mostraram-se polimórficos. Já os marcadores dominantes detectaram polimorfismo entre os cultivares AND277 e Rudá, na forma de presença e ausência de bandas. Para exemplificação, a Figura 5 ilustra o padrão dominante apresentado pelo *primer* PVac001. A ordem dos genótipos dispostos no gel corresponde à mesma descrita na Tabela 1, do item 2.1 da sessão “Material e Métodos”.

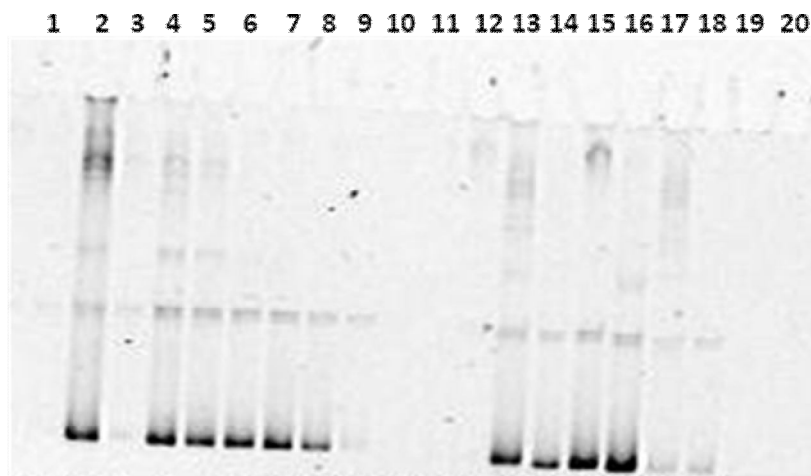


Figura 5: Perfil eletroforético polimórfico apresentado pelo marcador dominante PVac001 em gel de poliacrilamida 10%, corado com solução de nitrato de prata. A numeração 1-20 corresponde à ordem dos genótipos andinos e mesoamericanos disposta na Tabela 1.

Na literatura científica já foram descritos trabalhos utilizando *primers* SSR de carácter dominante para fins de mapeamento, a exemplo de Vilanova et al. (2003) e Karaca et al. (2002), que trabalharam com damasco e algodoeiro, respectivamente. Para o feijão, Yu et al., (2000) detectaram comportamento dominante em um dos *primers* desenvolvidos em seu estudo, e descreve que a ocorrência de marcadores microssatélites dominantes é possível e já havia sido detectado em estudos anteriores realizados pela sua equipe. Além do mapeamento genético, marcadores de carácter dominante, tanto microssatélites quanto outros tipos de marcadores, já foram extensivamente utilizados para outros tipos de estudos genéticos para o feijão, tais como em seleção assistida por marcadores moleculares e estudos de diversidade genética. Em relação aos demais marcadores de carácter dominante, tais como os marcadores SCAR e RAPD, os microssatélites ainda apresentam as vantagens de alta reprodutibilidade de resultados e de serem marcadores multialélicos.

Assim, os marcadores desenvolvidos no presente trabalho, tanto os de carácter codominante quanto os dominantes, poderão ser usados como importantes ferramentas para os mais variados tipos de estudos genéticos para a espécie, inclusive em estudos de mapeamento genético, colaborando para o incremento de programas de melhoramento do feijoeiro que utilizam marcadores moleculares.

3.4. Análise de similaridade

A partir da análise realizada, foi identificada similaridade com região transcrita de função não caracterizada, em *Glycine max* e com a região codificadora do gene para inibidor 1 de cinase ciclina-dependente, também em *Glycine Max* (Tabela 6).

As cinases ciclina-dependente são enzimas que catalizam a transferência de um grupo fosfato, usualmente do ATP, para uma proteína celular com função regulatória do ciclo celular, mais especificamente, relacionada ao trânsito de células durante a divisão celular. Os inibidores tipo 1 são responsáveis pela interrupção, prevenção ou redução da atividade dessas enzimas, com conseqüente efeito de interrupção do ciclo celular (Tabela 7). Dentre todos os locos microssatélites analisados no presente trabalho, somente os descritos na Tabela 6 apresentaram similaridade com alguma sequência já depositada e caracterizada do banco de dados. Da mesma forma, dentre os locos dispostos na Tabela 6, somente aquele que apresentou sequência para a qual foi identificada alguma função celular já caracterizada está categorizado na Tabela 7.

Tabela 6: Resultado da análise de similaridade. A primeira coluna corresponde ao loco microsatélite analisado; a segunda ao *hit* encontrado; seguida da porcentagem de similaridade entre as duas sequências, *score* de alinhamento e *e-value*. O número de acesso, do banco do NCBI, das sequências identificadas como similares encontra-se na última coluna.

SSR	RESULTADO DE SIMILARIDADE PELO BLASTN	% DE SIMILARIDADE	SCORE	e-VALUE	NÚMERO DE ACESSO
PVag001	cDNA - Glycine max	92%	255	1e-64	AK243796
PVag002	Gene para inibidor 1 de cinase ciclina-dependente	93%	87.9	1e-13	AY439101

Tabela 7: Categorização dos produtos gênicos segundo critérios de classificação do *Gene Ontology*.

Produto gênico	Anotação <i>Gene Ontology</i> Categorias: Processo, Função e Componente
Inibidor 1 de cinase ciclina-dependente	Processo biológico: "Interrupção do ciclo celular" Função molecular: "Atividade de cinase"; "Atividade de inibidor de cinase ciclina-dependente" Componente celular: "Nuclear"

4. REFERÊNCIAS BIBLIOGRÁFICAS

- AUSUBEL, F. (1998) **Current protocols in molecular biology**. New York: John Wiley & Sons, Tenton Data Systems
- BENCHIMOL L. L., CAMPOS, T., CARBONELL, S. A. M., COLOMBO, C. A., CHIORATTO, A. F., FORMIGHIERE, E. F., GOUVÊA, L. R. L., SOUZA, A. P. (2007). **Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers**. Genetic Resource and Crop Evolution, 54: 1747-1762.
- BLAIR, M. W., BUENDÍA, H. F., GIRALDO, M. C., MÉTAIS, I., PELTIER, D. (2008). **Characterization of AT-rich microsatellites in common bean (*Phaseolus vulgaris* L.)**. Theoretical and Applied Genetics, 118: 91-103.
- BLAIR, M.W., PEDRAZA, F., BUENDIA, H.F., GAITÁN-SOLÍS, E., BEEBE, S.E., GEPTS, P., TOHME, J (2003). **Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.)**. Theoretical and Applied Genetics, 107:1362-1374.
- CAIXETA, E. T., OLIVEIRA, A. C. B., BRITO, G. G., SAKIYAMA, N. S. (2009). **Tipos de marcadores moleculares**. In: **Marcadores Moleculares**. BORÉM, A., CAIXETA, E. T. 2^o Edição. Viçosa/MG: Editora UFV. 532 p.
- BUSO G. S. C., AMARAL, Z. P. S., BRONDANI, R. P. V, FERREIRA, M. E. (2006). **Microsatellite markers for the common bean *Phaseolus vulgaris***. Molecular Ecology Notes 6, 252–254.
- BUSO, G. S. C. , CIAMPI A. Y., MORETZSOHN M. C., AMARAL, Z. P. S., BRONDANI, R. V. (2003) **Desenvolvimento e caracterização de marcadores microssatélites em espécies vegetais tropicais**. Biotecnologia- Ciência e Desenvolvimento, 30: 113-115.
- CAIXETA, E.T., BORÉM, A., KELLY, J.D. (2005). **Development of microsatellite markers based on BAC common bean clones**. Crop Breeding and Applied Biotechnology, 5(2):125-133.
- COSTA, M.R. (2007). **Melhoramento de feijões preto e vermelho visando a resistência à antracnose, mancha-angular e ferrugem, com o auxílio de marcadores moleculares**. Viçosa: UFV. 88 p. (Tese-Doutorado).
- DOYLE, J.J., DOYLE, J.L. (1990) **Isolation of plant DNA from fresh tissue**. Focus, 12: 13-15.
- FOUILLOUX, G. (1976). **Bean anthracnose: new genes of resistance**. Annual Report of the Bean Improvement Cooperative, 19:36-37.
- GAITÁN-SOLÍS, E., DUQUE, M.C., EDWARDS, K.J., TOHME, J (2002). **Microsatellite repeats in common bean (*Phaseolus vulgaris*): Isolation, characterization, and cross-species amplification in *Phaseolus* spp.** Crop Science, 42:2128-2136.

- GALVAN M.Z, BORNET B., BALATTI P. A., BRANCHARD M..(2003) **Inter simple sequence repeat (ISSR) markers as a tool for the assessment of both genetic diversity and gene pool origin in common bean (*Phaseolus vulgaris* L.)**. *Euphytica*, 132: 297–301
- GUERRA-SANZ, J.M. (2004) **New SSR markers of *Phaseolus vulgaris* from sequence databases**. *Plant Breeding*, 123: 87-89.
- GUO, D.L., LUO, Z.R. (2006). **Development of SSR primers using ISSR-PCR in *Diospyros kaki* Thumb**. *Molecular Ecology Notes*, 6: 886-887
- HANAI, L. R., CAMPOS, T., CAMARGO, L. E. A., BENCHIMOL, L. L., SOUZA, A. P., MELOTTO, M., CARBONELL, S. A. M., CHIORATTO, A. F., CONSOLI, L., FORMIGHIERI, E. F., SIQUEIRA, M. V. B. M., TSAI, S. M., VIEIRA, M. L. C. (2007). **Development, characterization, and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources**. *Genome*, 50: 266-277.
- KARACA M. , SAHA S., JENKINS J. N., ZIPF A., KOHEL R., STELLY D.M. (2002). **Simple Sequence Repeat (SSR) Markers Linked to the *Ligon Lintless* (Li1) Mutant in Cotton**. *The Journal of Heredity*, 93: 150-152
- L'TAIEF, B., HORRES, R., JUNGSMANN, R., MOLINA, C., SIFI, B., LACHAËL, M., WINTER, P., KAHL, G. (2008). **Locus-specific microsatellite markers in common bean (*Phaseolus vulgaris* L.): isolation and characterization**. *Euphytica*, 162: 301-310.
- LIAN, C., ZHOU, Z., HOGETSU, T. (2001). **A simple method for developing microsatellite markers using amplified fragments of inter-simple sequence repeat (ISSR)**. *Journal of Plant Research*, 144: 381-385.
- LI, Y., KOROL, A.B., FAHIMA, T., BEILES, A., NEVO, E. (2002). **Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review**. *Molecular Ecology*, 11:2453-2465.
- LITT, M., LUTY, J.A (1989). **A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene**. *Am. J. Hum. Genet.*, 44: 398-401.
- MASI, P., ZEULI, P. L. S., DONINI, P. (2003). **Development and analysis of multiplex microsatellite marker set in common bean (*Phaseolus vulgaris* L.)**. *Molecular Breeding*, 11: 303-313.
- MELOTTO, M., MONTEIRO-VITORETTO, C.B., BRUSCHI, A.G., CAMARGO, L.E.A. (2005). **Comparative bioinformatic analysis of genes expressed in common bean (*Phaseolus vulgaris* L.)**. *Genome*, 48: 562-570.
- MÉTAIS, I., AUBRY, C., HAMON B., JALOUZOT R., PELTIER D (2000). **Description and analysis of genetic diversity between commercial bean**

lines (*Phaseolus vulgaris* L.). Theoretical Applied Genetics, 101: 1207–1214.

MÉTAIS, I., HAMON, B., JALOUZOT, R., PELTIER, D (2002). **Structure and level of genetic diversity in various bean types evidenced with microsatellites markers isolated from a genomic enriched library.** Theoretical and Applied Genetics 104:1346-1352.

NCBI. National Center for Biotechnology Information. Disponível em <http://www.ncbi.nlm.nih.gov/>. Acesso em Julho/2009

NEST, M. A. V. D., STEENKAMP E. T., WINGFIELD B. D., WINGFIELD M. J (2000). **Development of simple sequence repeat (SSR) markers in *Eucalyptus* from amplified inter-simple sequence repeats (ISSR).** Plant Breeding, 119: 433-436.

QUEIROZ, V.T. de (2004). **Obtenção de primers microsatélites e desenvolvimento, validação e mapeamento de marcadores SCAR em feijoeiro-comum.** Tese de Doutorado. 155p.

SANGER, F., NICKLEN, S., COULSON, A. R. (1977). **DNA sequencing with chain-terminating inhibitors.** Proceedings of the National Academy of Science, 74: 5463-5467.

SANGUINETTI, C.J., DIAS NETO, E., SIMPSON, A.J.G. (1994). **Rapid silver staining and recovery of PCR products separated on polyacrylamide gels.** Biotechniques, 17: 915-919.

SIEBERT, P.D., CHENCHIK, A., KELLOGG, G.E., LUKYANOV, K. A., LUKYANOV, S. A. (1995). **An improved PCR method for walking in uncloned genomic DNA.** Nucleic Acids Research., 23: 1087- 1088.

VILANOVA S., ROMERO C., ABBOTT A. G., LLICER G., BADENES M. L. (2003). **An apricot (*Prunus armeniaca* L.) F2 progeny linkage map based on SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits.** Theoretical Applied Genetics, 107:239–247

YAISH M.W.F., VEGA, M. P. D. L. (2003). **Isolation of (GA)_n Microsatellite Sequences and Description of a Predicted MADS-box Sequence Isolated from Common Bean (*Phaseolus vulgaris* L.).** Genetics and Molecular Biology, 26: 337-342

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

YU, K.; PARK, S.J.; POYSA, V.; GEPTS, P (2000). **Integration of simple sequence repeat (SSR) Markers Into a molecular Linkage Map of Common Bean (*Phaseolus vulgaris* L.).** The Journal of Heredity, 91: 429-434.

WANG JB (2003). **ISSR markers and their applications in plant genetics.** Yi Chuan, 24:613-616.

WEBER, J. L. (1990). **Informativeness of human (dC- dA)_n.(dG- dT)_n polymorphisms.** Genomics, 7: 524-530.

RESUMO DOS RESULTADOS, CONCLUSÕES E PERSPECTIVAS

- A partir do desenvolvimento de duas bibliotecas genômicas enriquecidas, foram analisados 184 clones (88,89%), a partir dos quais foi possível identificar 66 (49,62%) motivos SSR na biblioteca 1 e 20 (39,22%) motivos SSR na biblioteca 2. Foram desenhados 56 pares de *primers*. Destes, 34 tiveram suas condições de amplificação padronizadas e seis (17,65%) amplificaram produtos polimórficos entre os genitores testados.
- A partir do desenvolvimento de uma biblioteca enriquecida com fragmentos microssatélites, obtida a partir da amplificação por marcadores ISSR, foi possível analisar 103 clones (41,2%), a partir dos quais foi possível identificar 58 motivos microssatélites (56,31%), e o desenvolvimento de 32 pares de *primers*. Destes, 10 tiveram suas condições de amplificação padronizadas, sendo que seis comportaram-se como marcadores codominantes e quatro como dominantes. Dentre os codominantes nenhum mostrou-se polimórfico dentre os genótipos testados, já os dominantes detectaram polimorfismo entre os genótipos na forma de presença e ausência de bandas.
- A alta porcentagem de enriquecimento de sequências contendo motivos microssatélites obtida indica que as duas metodologias utilizadas foram eficazes para a seleção, no feijoeiro, de seqüências contendo motivos microssatélites.
- Os resultados obtidos no presente trabalho permitiram a constatação da presença de SSR em regiões relacionadas a importantes funções no genoma do feijão e o desenvolvimento de *primers* cpSSRs, ainda pouco descritos na literatura científica para a cultura. Os marcadores microssatélites desenvolvidos poderão ser usados como ferramentas eficazes em diversos tipos de estudos genéticos, inclusive em estudos de mapeamento genético visando a consolidação de um mapa genético consenso saturado para a espécie, contribuindo para o aprimoramento dos programas de melhoramento do feijoeiro comum que utilizam marcadores moleculares.