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

**Desenvolvimento de marcadores SSR e SNP em maracujá-doce a partir de
uma biblioteca enriquecida com genes de resposta à *Xanthomonas*
*axonopodis***

Zirlane Portugal da Costa

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Genética e
Melhoramento de Plantas

**Piracicaba
2014**

Zirlane Portugal da Costa
Engenheira Agrônoma

**Desenvolvimento de marcadores SSR e SNP em maracujá-doce a partir de
uma biblioteca enriquecida com genes de resposta à *Xanthomonas*
*axonopodis***

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientadora
Prof^a. Dra. **MARIA LUCIA CARNEIRO VIEIRA**

Dissertação apresentada para obtenção do título de Mestra em
Ciências. Área de concentração: Genética e Melhoramento de
Plantas

Piracicaba
2014

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Costa, Zirlane Portugal da

Desenvolvimento de marcadores SSR e SNP em maracujá-doce a partir de uma biblioteca enriquecida com genes de resposta à *Xanthomonas axonopodis* / Zirlane Portugal da Costa. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2014.

117 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2014.
Bibliografia.

1. *Passiflora alata* 2. Maracujá-doce 3. Microsatélites 4. Marcadores SNP
5. Polimorfismo molecular 6. Genes diferencialmente expressos I. Título

CDD 634.425
C837d

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

A minha mãe Zilda

Ao meu pai Osias

Meus irmãos Ozéas, Otoniel e Oziel

Ao meu sobrinho Gabriel

Ao meu namorado Ricardo Gusmão

A minha melhor amiga Jamille

Aos meus amigos

Com amor, dedico este trabalho.

A Deus, pela vida, saúde e força

Ofereço

AGRADECIMENTOS

À Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ / USP) e ao Departamento de Genética, pela oportunidade de realizar o curso de Mestrado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelas bolsas concedidas.

À Professora Dra. Maria Lucia Carneiro Vieira pela orientação, apoio, confiança e oportunidades concedidas durante o desenvolvimento deste trabalho.

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

À Dra. Carla de Freitas Munhoz, pela amizade, atenção, auxílio e esclarecimentos durante as etapas do trabalho.

Ao Carlos Alberto de Oliveira, pela amizade e auxílio técnico durante a realização do trabalho.

Aos funcionários do Departamento de Genética pelo convívio amigável e pela colaboração durante a execução do trabalho.

Aos amigos do Laboratório de Genética Molecular de Plantas Cultivadas, Alessandra, Anselmo, Augusto, Bruno, Carmelice, Cassia, Carla, Guilherme, Helen, Luciane, Lourdes, Luiz Augusto e Rafael pelos auxílios, sugestões e companheirismo.

A todos os amigos que contribuíram para a realização deste trabalho.

SUMÁRIO

RESUMO.....	9
ABSTRACT.....	11
LISTA DE FIGURAS.....	13
LISTA DE TABELAS.....	15
1 INTRODUÇÃO.....	17
2 REVISÃO BIBLIOGRÁFICA.....	19
2.1 A cultura do maracujazeiro.....	19
2.2 Marcadores SSR.....	20
2.3 Marcadores SNP.....	24
2.4 Polimorfismo molecular em <i>Passiflora edulis</i> e <i>P. alata</i>	27
3 MATERIAL E MÉTODOS.....	29
3.1 Material Biológico.....	29
3.2 Análise das sequências de transcritos para o desenvolvimento dos marcadores SSR e SNP.....	29
3.3 Marcadores SSR.....	30
3.3.1 Prospeção e classificação dos locos SSR.....	30
3.3.2 Desenho dos <i>primers</i> e testes de amplificação.....	30
3.3.3 Transferibilidade de <i>primers</i> de <i>P. edulis</i> f. <i>flavicarpa</i> para <i>P. alata</i>	31
3.3.4 Eletroforese e avaliação do polimorfismo.....	32
3.4 Marcadores SNP.....	35
3.4.1 Seleção das sequências e desenho de <i>primers</i>	35
3.4.2 Condições de amplificação e transferibilidade dos <i>primers</i> de <i>P. edulis</i> f. <i>flavicarpa</i> para <i>P. alata</i>	39
3.4.3 Avaliação, purificação e sequenciamento dos produtos de PCR.....	40
3.4.4 Processamento, alinhamento das sequências e identificação dos potenciais SNPs.....	40
4 RESULTADOS.....	43
4.1 Marcadores SSR.....	43
4.1.1 Caracterização das regiões contendo SSR.....	43
4.1.2 Testes e seleção de <i>primers</i> de SSR.....	51
4.1.3 Análise do polimorfismo em gel de poliacrilamida desnaturante.....	55
4.2 Marcadores SNP.....	62

5 DISCUSSÃO.....	73
5.1 Marcadores SSR.....	73
5.1.1 Caracterização das regiões contendo SSR.....	73
5.1.2 Testes e seleção de <i>primers</i> de SSR.....	75
5.1.3 Análise do polimorfismo em gel de poliacrilamida desnaturante.....	76
5.2 Marcadores SNP.....	77
5.3 Polimorfismo molecular em <i>Passiflora</i>	81
6 CONCLUSÕES.....	83
REFERÊNCIAS.....	85
ANEXOS.....	99

RESUMO

Desenvolvimento de marcadores SSR e SNP em maracujá-doce a partir de uma biblioteca enriquecida com genes de resposta à *Xanthomonas axonopodis*

Um dos desafios atuais da pesquisa em frutíferas tropicais é incorporar abordagens baseadas em marcadores moleculares nos programas convencionais de melhoramento. O maracujá-doce (*Passiflora alata*) é uma espécie diploide, de fecundação cruzada e pouco explorada. Recentemente, nosso grupo construiu um mapa de ligação de *P. alata* composto de diferentes tipos de marcadores moleculares. Além disso, dispõe-se de um conjunto de transcritos de *Passiflora edulis*, obtidos a partir de duas bibliotecas de expressão: *forward* e *reverse* onde foram isolados transcritos diferencialmente expressos na planta inoculada com *Xanthomonas axonopodis* (*Xap*) (672) e na planta controle, não inoculada (310), respectivamente. Assim, neste estudo, este conjunto de transcritos foi explorado visando ao desenvolvimento de marcadores SSR e SNP com o intuito de enriquecer, posteriormente, o mapa de ligação de *P. alata* com marcadores funcionais putativos. Para o desenvolvimento dos marcadores SSRs, as 672 sequências da biblioteca *forward* foram investigadas e em 91 delas foram encontrados 115 SSRs. Como esperado, a classe de repetições trinucleotídicas foi a mais abundante, sendo o motivo (AG)_n o mais comum entre as repetições dinucleotídicas. Desenhou-se *primers* para amplificar 42 desses SSRs. Dois acessos de *P. edulis* e seis indivíduos da população de mapeamento de *P. alata* foram usados nos testes de transferibilidade e avaliação do polimorfismo. Trinta e quatro pares de *primers* apresentaram bom padrão de amplificação, porém apenas 10 deles revelaram polimorfismo em *P. alata*. Para o desenvolvimento dos marcadores SNPs, 118 sequências selecionadas das bibliotecas de expressão *forward* e *reverse* foram usadas para o desenho de *primers*; 37 delas foram usadas para avaliar o polimorfismo no mesmo *set* de indivíduos de *P. alata*. Foram encontrados 34 locos contendo SNPs bialélicos em 16 fragmentos gênicos, cujas sequências variaram em tamanho de 332 a 872 pb. Considerando todos os fragmentos gênicos (16), foi analisado um total de 10.003 pb; a frequência de SNPs foi estimada como sendo 1 a cada 294 pb. Observou-se a mesma ocorrência de SNPs (50%, 17/34) em regiões codantes e não-codantes. Uma função putativa pôde ser atribuída a todos os fragmentos gênicos de *P. alata*, sendo que 82% mostraram homologia com as mesmas proteínas das sequências de origem, isoladas de *P. edulis*. No geral, os locos marcadores apresentaram baixo nível de polimorfismo molecular. Este é o primeiro trabalho sobre o desenvolvimento de locos marcadores funcionais putativos em *Passiflora* usando transcritos expressos em resposta à *Xap*.

Palavras-chave: *Passiflora alata*; Maracujá-doce; Microssatélites; Marcadores SNP; Polimorfismo molecular; Genes diferencialmente expressos

ABSTRACT

Development of SSR and SNP markers in sweet passion fruit from a library enriched for genes induced in response to *Xanthomonas axonopodis*

One of the current challenges of tropical fruit research is to incorporate molecular marker-based approaches into conventional breeding programs. The sweet passion fruit (*Passiflora alata*) is a diploid, outcrossing and underexploited species. Recently, our group has constructed a *P. alata* linkage map consisted of different types of molecular markers. Moreover, we have a set of transcripts of *Passiflora edulis* obtained from two expression libraries: the forward and the reverse where differentially expressed transcripts were isolated from a plant inoculated with *Xanthomonas axonopodis* (*Xap*) (672), and from the control plant, uninoculated (310), respectively. Thus, in this study, this set of transcripts were exploited aiming at the development of SNP and SSR markers for future enrichment of the *P. alata* linkage map with putative functional markers. For the development of SSR markers, the 672 sequences from the forward library were investigated and 91 of them were found to have 115 SSRs. As expected, the trinucleotide class of repeats was the most abundant, and the (AG)_n motif was the most common among the dinucleotide repeats. Primers were designed to amplify 42 of these SSRs. Transferability tests and polymorphism investigation were carried out using two accessions of *P. edulis* and six individuals of the mapping population of *P. alata*. Thirty-four primer pairs showed a good amplification pattern but only 10 loci revealed polymorphism in *P. alata*. For the development of SNP markers, 118 sequences selected from forward and reverse expression libraries were used for designing primers; 37 were used to assess the polymorphism in the same set of individuals of *P. alata*. Thirty-four biallelic SNPs were found in 16 gene fragment sequences that ranged in size from 332 to 872 bp. Considering all gene fragments, a total of 10,003 bp was obtained; the frequency of SNPs was estimated to be 1 every 294 bp. The same prevalence of SNPs (50%, 17/34) was observed within coding and non-coding regions. A putative function was assigned to all gene fragments of *P. alata*; 82% of them have shown homology to the original protein sequences isolated from *P. edulis*. Overall, the marker loci showed a low level of molecular polymorphism. This is the first report on the development of putative functional marker loci in *Passiflora* using transcripts induced in response to *Xap*.

Keywords: *Passiflora alata*; Sweet passion fruit; Microsatellites; SNP markers; Molecular polymorphism; Differentially expressed genes

LISTA DE FIGURAS

- Figura 1 - Parte da saída do software *Sciroko* utilizado na busca de microssatélites nos 672 transcritos da biblioteca de expressão *forward*.....43
- Figura 2 - Identificação e classificação dos motivos SSRs nas sequências dos transcritos PeA2P4 (Putative chromatin remodelling complex ATPase chain ISWI), PeC3P4 (Cysteine proteinase inhibitor 3), PeD9P3 (1-aminocyclopropane-1-carboxylate oxidase 4) e PeH4P6 (Peroxisomal biogenesis factor 6-like, partial).....44
- Figura 3 - Distribuição dos microssatélites quanto ao motivo [mononucleotídeos (1), dinucleotídeos (2), trinucleotídeos (3), tetranucleotídeos (4), pentanucleotídeos (5) e hexanucleotídeos (6)] e a classe identificada.....45
- Figura 4 - Saída do aplicativo *web PRIMER3* utilizado para o desenho dos pares de *primers* flanqueadores da região do microssatélite (CA)₈ do transcrito A1P4, neste exemplo.....51
- Figura 5 - Eletroforese em gel de agarose (1,8%) corado com Sybr Green mostrando o padrão de amplificação de 14 microssatélites. Canaletas 1 a 5 (para cada *primer*): ‘IAPAR-06’, ‘IAPAR-123’, ‘2(12)’, ‘SV3’ e Controle negativo, sem DNA *template*. L é o ladder 100 pb.....54
- Figura 6 - Eletroforese em gel de poliacrilamida desnaturante mostrando o padrão de amplificação de 9 locos de microssatélites. Canaletas 1 a 10 (para cada *primer*): acessos ‘IAPAR-06’, ‘IAPAR-123’, ‘2(12)’ em duplicata, ‘SV3’ em duplicata e 4 indivíduos da F₁. L é o ladder 10 pb.....57
- Figura 7 - Polimorfismo e segregação dos locos polimórficos. Nas canaletas 1 e 2 estão os acessos ‘2(12)’ e ‘SV3’, genitores da população F₁, e nas demais canaletas está a progênie. As setas indicam os locos polimórficos para os marcadores que segregam 1:1.....59
- Figura 8 - Eletroforese em gel de agarose (1,8%) corado com Sybr Green mostrando o padrão de amplificação de 6 fragmentos gênicos. Canaletas 1 a 5 (para cada par de *primers*): acessos ‘IAPAR-06’, ‘IAPAR-123’, ‘2(12)’, ‘SV3’ e Controle negativo. L é o ladder 100 pb.....62
- Figura 9 - Sequências e eletroferogramas dos acessos ‘2(12)’, ‘SV3’ e dos indivíduos da geração filial F₁-67 e F₁-100 relativos ao fragmento gênico PaD5P10 de *P. alata*: há um SNP na posição 350.....65

Figura 10 - BLASTN das sequências dos *contigs* de *P. alata* revelados pela amplificação dos *primers* desenhados a partir de sequências de cDNA dos genes PeB8P9 (Protein linase 2B), PeP3RA2 (Lactoylglutathione lyase) e PeContig21 (30S Ribosomal protein 3-1) de *P. edulis*. As linhas horizontais (em cinza) representam os possíveis íntrons presentes do DNA genômico de *P. alata*.....67

Figura 11 - Número de polimorfismos encontrados nas sequências dos genes putativos (ou *contigs*) de *P. alata*. A/G e C/T correspondem a transições e os demais a transversões.....68

LISTA DE TABELAS

Tabela 1 - Acessos de maracujá-doce e de maracujá-azedo utilizados nos testes de amplificação e polimorfismo para os marcadores SSR e SNP.....	29
Tabela 2 - Concentração dos reagentes e relação dos <i>mix</i> utilizados para a otimização da PCR para cada par de <i>primers</i> usados nos testes de transferibilidade.....	32
Tabela 3 - Configurações genéticas dos locos marcadores, segundo o tipo de cruzamento, e os respectivos padrões de bandas observadas (alelos) nos genitores e nas progênes oriundas de cruzamentos entre genitores não-endogâmicos.....	35
Tabela 4 - Lista dos transcritos selecionados das bibliotecas de expressão para o desenho de <i>primers</i> e proteína codificada pelo respectivo transcrito.....	36
Tabela 5 - Tipo de repetição, motivo e número encontrado entre os 115 microssatélites distribuídos nas 91 sequências expressas de <i>Passiflora edulis</i>	46
Tabela 6 - Lista dos 42 transcritos contendo SSRs selecionados da biblioteca de expressão <i>forward</i> para o desenho de <i>primers</i> e proteína respectiva codificada.....	47
Tabela 7 - Resultado da análise <i>in silico</i> para 42 regiões contendo microssatélites: sequências dos pares de <i>primers</i> (5'-3'), motivo, classe e tamanho esperado do respectivo alelo (AE, em pb).....	48
Tabela 8 - Resultado das amplificações com os 42 pares de <i>primers</i> de microssatélites em <i>Passiflora alata</i> em gel de agarose.....	52
Tabela 9 - Resultados das amplificações dos 42 pares de <i>primers</i> de microssatélites de <i>Passiflora alata</i> em gel de poliacrilamida desnaturante: qualidade da amplificação (nº de locos e nitidez das bandas), presença de bandas <i>stutter</i> , análise genética (locos P, polimórficos, e M, monomórficos) e tamanhos observados dos alelos (AO, em pb).....	55
Tabela 10 - Classificação dos locos SSR quanto ao tipo de configuração genética, segundo a terminologia de Wu et al. (2002).....	61
Tabela 11 - Resultado da análise das sequências dos 37 genes putativos de <i>Passiflora alata</i> selecionados.....	63
Tabela 12 - SNPs putativos e respectiva localização, se intrônico (I) ou exônico (E), genótipo dos genitores e de dois indivíduos da F ₁ de <i>Passiflora alata</i>	66
Tabela 13 - Classificação dos SNPs identificados em cada gene putativo (ou <i>contig</i>) de <i>P. alata</i> quanto ao tipo de substituição.....	68

Tabela 14 - Resumo das consultas via BLASTX dos <i>contigs</i> obtidos para os 16 fragmentos gênicos com presença de SNP para <i>Passiflora alata</i> em relação ao acesso com maiores <i>scores</i> de alinhamento.....	70
Tabela 15 - Tamanho esperado e observado das sequências (em pb) dos 16 fragmentos gênicos de <i>P. alata</i> e ocorrência de SNPs em cada fragmento gênico estudado...	71

1 INTRODUÇÃO

Boa parte das fruteiras tropicais é pouco estudada do ponto de vista genético, embora várias delas sejam usadas como alimento, ou tenham potencial econômico para consumo *in natura* e para serem exploradas em processos industriais. O maracujá-doce (*Passiflora alata* Curtis) é uma espécie diploide ($2n=18$), de fecundação cruzada (FERREIRA et al., 2010) e muito pouco explorada comercialmente. A espécie de maracujá mais cultivada no país é a *Passiflora edulis* Sims, também conhecida como maracujá-azedo, representando 95% dos pomares comerciais. Entretanto, devido aos preços diferenciados, o maracujá-doce vem ganhando importância dentro do mercado de frutas *in natura* e para compor *blends* em sucos industrializados.

A inexistência de uma cultivar homogênea e produtiva tem sido o fator limitante para elevar o valor e a expansão dessa cultura. Sem dúvida, as técnicas de melhoramento genético devem contribuir para o desenvolvimento e a competitividade da cultura, viabilizando a obtenção de cultivares que atendam aos padrões referenciais (KAVATI; PIZA JR, 2002). Além disso, as tecnologias advindas do conhecimento de genética molecular, especificamente o desenvolvimento de marcadores de DNA, têm contribuído com os programas de melhoramento, mesmo para as espécies pouco estudadas sob o ponto de vista genético (COLLARD; MACKILL, 2008).

Microssatélites, também conhecidos como sequências simples repetidas (SSR), são regiões compostas por pequenos motivos de 1 a 6 nucleotídeos repetidos em *tandem*, que são dispersos no genoma de eucariotos e procariotos (veja OLIVEIRA et al., 2006). Muito usados como marcadores genéticos, os SSRs são baseados em PCR, revelam polimorfismos e por isso tem sido aplicados em mapeamento, identificação e proteção de cultivares e em diagnose molecular (KING; THOROGOOD; EDWARDS, 2008).

Variações em um único nucleotídeo em sequências de DNA são conhecidas como SNPs. O potencial dos SNPs como marcadores tem sido claramente demonstrado nas análises do genoma humano, pois vários milhões de SNPs têm sido identificados e analisados (FRAZER et al., 2007). Com um número tão grande de marcadores disponível e densidade extremamente alta, é possível o exame de todo o genoma, permitindo a associação do perfil de marcas individuais com o de caracteres de herança complexa (MCCARTHY, 2008). Da mesma forma, isso pode ser útil em estudos com plantas cultivadas.

O desenvolvimento de mapas genéticos é considerado uma das aplicações de maior impacto da tecnologia de marcadores moleculares. Procedimentos como a localização e o

mapeamento de locos quantitativos (QTLs), estudos de sintenia e clonagem de genes com base em mapas de ligação passaram a representar complementos importantes a serem considerados pelos geneticistas (CARNEIRO; VIEIRA, 2002). Nas espécies em que ocorre autoincompatibilidade e o cruzamento entre indivíduos é obrigatório, como o maracujazeiro, não é possível a obtenção de populações convencionais de mapeamento. Para contornar esse problema, inicialmente, foi adotada a estratégia *2-way pseudo-testcross* (RITTER; GEBHARDT; SALAMINI, 1990) para a geração de mapas, usando uma população F_1 segregante e marcas dominantes, o que resulta na geração de mapas individuais, sendo um por genitor. A inconveniência desta estratégia tem sido superada pelo uso de marcadores codominantes e estatísticas mais robustas. Neste contexto, um mapa genético integrado de maracujá-doce foi construído por Pereira et al. (2013) e consiste basicamente em uma compilação de marcadores AFLP, M-AFLP, SSR, RGA, TRAP e SNP.

De posse de um mapa genético, é possível a alocação de novos marcadores, visando a sua associação com caracteres fenotípicos de interesse. Nosso grupo de pesquisa construiu uma biblioteca de expressão enriquecida com genes de resposta de maracujá à inoculação com *Xanthomonas axonopodis*, principal patógeno bacteriano que ataca os pomares (MUNHOZ, 2013). Sendo assim, neste trabalho se deu continuidade a esses estudos desenvolvendo marcadores SSRs e SNPs gênicos, a partir da biblioteca de expressão, para futura alocação desses locos no mapa de ligação de maracujá-doce desenvolvido por Pereira et al. (2013).

2 REVISÃO BIBLIOGRÁFICA

2.1 A cultura do maracujazeiro

O maracujazeiro pertence à família *Passifloraceae*, que é amplamente distribuída nos trópicos e regiões temperadas, sendo composta por 18 gêneros e cerca de 630 espécies. O gênero *Passiflora* é o mais importante economicamente e possui 129 espécies conhecidas, nativas do Brasil, das quais 83 são endêmicas, podendo ser utilizadas como alimento, remédios e ornamento (CERVI et al., 2011).

De acordo com Faleiro et al. (2008), *Passiflora edulis* (maracujá-azedo) e *Passiflora alata* (maracujá-doce) são as espécies mais cultivadas, sendo estimado que essas espécies ocupem mais de 90% da área cultivada no mundo. O Brasil se destaca como maior produtor mundial de maracujá, com produção da ordem de 920 mil toneladas, produzidas em 62 mil hectares, sendo que 76% da produção são oriundos da região Nordeste. Neste cenário, o Estado da Bahia ocupa a posição de liderança, sendo responsável por 52% da área plantada e 50% da produção brasileira (IBGE, 2010). Há um registro, em 2001, que 95% da produção são oriundos do maracujazeiro-azedo que é destinada para o mercado de frutas e para a indústria de sucos (MELETTI; BRÜCKNER, 2001).

Entretanto, devido aos preços diferenciados, a cultura do maracujá-doce vem ganhando importância dentro do mercado de frutas *in natura* e para compor *blends* em sucos industrializados. Além disso, suas folhas possuem propriedades farmacológicas associadas à produção de passiflorina, um calmante natural. Seu valor ornamental está associado às flores, coloridas e perfumadas. O maracujá-doce é uma espécie nativa da América do Sul, especialmente do Brasil, mas encontrada também no Peru, Paraguai e Argentina, ainda desconhecida da maioria da população. O aumento da área plantada, embora pequeno, e o crescente interesse pela cultura criam novas demandas de geração de cultivares e tecnologias de produção.

Com a considerável demanda por frutos de maracujá e consequente uso da espécie em monocultura, houve maior incidência de pragas e doenças nos pomares. As doenças são causadas por fungos, vírus e bactérias. Tais problemas acarretam redução do tempo de exploração comercial dos pomares e, até mesmo, impossibilitam seu plantio em algumas regiões, o que exige esforços para que as perdas diminuam e se impeça a disseminação das doenças (SANTOS FILHO et al., 2004).

Uma das doenças do maracujazeiro que se tem maior dificuldade em controlar é a bacteriose causada por *Xanthomonas axonopodis* pv. *passiflorae*, por não se dispor de um controle químico eficiente.

Essa doença provoca perdas expressivas em maracujazeiro-doce e azedo durante os períodos mais quentes e úmidos do ano. Ocorre em todas as regiões onde se cultiva o maracujazeiro, sendo mais severa nas regiões mais quentes e úmidas. Nas folhas, o patógeno induz pequenas lesões encharcadas e translúcidas que necrosam. Com o desenvolvimento da doença, ocorre seca das folhas e, posteriormente, a desfolha, reduzindo consideravelmente a produtividade. A bactéria pode atingir os feixes vasculares e se propagar por todos os tecidos da planta, provocando a seca e até a morte de plantas em cultivares muito suscetíveis (JUNQUEIRA; JUNQUEIRA, 2007; VIANA et al. 2003). Nos frutos, as manchas são grandes, bem delimitadas, inicialmente esverdeadas e oleosas, podendo coalescer e causar lesões de grandes extensões. Essas manchas são superficiais, entretanto, em condições favoráveis, o patógeno pode penetrar a polpa, contaminando as sementes e promovendo a sua fermentação (HALFELD-VIEIRA; NECHET, 2006; JUNQUEIRA; JUNQUEIRA, 2007; VIANA et al., 2003). Em maracujá doce, pode ocorrer a deterioração do mesocarpo, caracterizada por podridão mole, com anasarca, descoloração e consequente separação da casca, com o aparecimento ou não de cavidades (MALAVOLTA JÚNIOR et al., 2001)

Por isso, se torna importante o uso de variedades resistentes, o que realça a necessidade de pesquisas que visem buscar essas novas variedades (veja MUNHOZ et al., 2011). Atualmente, a geração de variedades melhoradas tende a incorporar metodologias de genética molecular. Por exemplo, a escolha de genitores com base em fenótipos moleculares ou a identificação de marcadores moleculares ligados fisicamente a genes que controlam caracteres de importância agrônômica podem auxiliar no processo de seleção genética (COLLARD; MACKILL, 2008), inclusive visando à resistência à doenças.

2.2 Marcadores SSR

Microssatélites, também conhecidos como sequências simples repetidas (SSR) são regiões de DNA repetitivo compostos por pequenos motivos de 1 a 6 nucleotídeos repetidos em *tandem*, que são dispersos no genoma de eucariotos e procariotos (TÓTH et al., 2000). Repetições de mono-, di-, tri- e tetranucleotídeos caracterizam os principais tipos de SSRs, mas repetições de cinco (penta-) ou seis (hexa-) nucleotídeos também são usualmente encontradas (ELLEGEN, 2004).

Marcadores SSR são classificados de acordo com o tipo de sequência repetida, como perfeita, imperfeita, interrompida ou composta. Em um SSR perfeito a sequência repetida não é interrompida por nenhuma base que não pertença ao motivo (por exemplo, TATATATATATATATA), enquanto em um SSR imperfeito há uma base entre os motivos repetidos que não corresponde ao motivo (por exemplo, TATATATACTTATATA). No caso de SSRs interrompidos há uma pequena sequência intercalada que não corresponde ao motivo repetido (por exemplo, TATATACGTGTTATATATATA), enquanto em um SSR composto há duas sequências repetidas adjacentes (por exemplo, TATATATATAGTGTGTGTGTGT) (veja revisão de OLIVEIRA et al., 2006).

SSRs são regiões do genoma que sofrem altas taxas de mutação. Embora a dinâmica de mutação dessas regiões genômicas não seja bem entendida (SCHLÖTTERER, 2000), sabe-se que a taxa de mutação é bem maior do que em outras partes do genoma, variando de 10^{-2} a 10^{-6} nucleotídeos por loco, por geração (SIA et al., 2000). Como dito, as causas das variações não são completamente entendidas, mas são de grande interesse para a Medicina, por exemplo, porque a instabilidade dos SSRs tem relação com o câncer (ODA et al., 2005) ou expansões de locos SSRs são responsáveis por mais de 40 desordens neurológicas (PEARSON; EDAMURA; CLEARY, 2005). Têm sido sugeridos mecanismos para explicar a alta taxa de mutação das SSRs, incluindo erros durante a recombinação, crossing-over desigual e *slippage*, ou deslizamento, da polimerase durante a replicação ou reparo do DNA (STRAND et al., 1993). Segundo Schlötterer (2000), o número e o tipo de repetição, a sequência flanqueadora e a recombinação são os fatores que influenciam as taxas de mutação das SSRs.

A densidade das SSRs tende a ser positivamente correlacionada com o tamanho do genoma (KATTI; RANJEKAR; GUPTA, 2001). Entre os genomas de eucariotos completamente sequenciados, a densidade de SSRs é maior em mamíferos. No entanto, em plantas, a frequência de SSRs parece estar negativamente correlacionada com o tamanho do genoma. Morgante, Hanafey e Powell (2002) encontraram que a frequência de SSRs difere entre *Arabidopsis*, milho, soja, algodão e arroz. Esta frequência é maior em *Arabidopsis* e menor em espécies com genomas comparativamente maiores, como o milho e o algodão. Santos (2013) explorou uma biblioteca genômica inserida em BACs (*Bacterial Artificial Chromosomes*) por meio da técnica de *BAC-end sequencing* visando promover os primeiros *insights* sobre a composição e organização do genoma de *Passiflora edulis*, além de gerar novos candidatos a marcadores moleculares. Assim, identificou 610 potenciais novos

marcadores SSR nos 5.704.725 pb analisados, com densidade média estimada em 10,7 SSRs a cada 100 kpb.

Há uma relação linear positiva entre a frequência de SSRs e a presença de cópias únicas de DNA, sugerindo que SSRs parecem ser mais frequentes em cópias únicas de DNA do que em DNAs repetitivos. Outra característica peculiar da maioria dos genomas de plantas é que $(AT)_n$ é o motivo mais comum entre os dinucleotídeos (LAGERCRANTZ; ELLEGREN; ANDERSSON, 1993).

Várias linhas de pesquisa têm demonstrado que a distribuição genômica das SSRs não é ao acaso, por conta de seus efeitos na organização da cromatina, regulação da atividade gênica, recombinação, replicação do DNA, ciclo celular e sistema de reparo (*mismatch*). Estas SSRs com funções putativas podem estar localizadas em genes ou regiões regulatórias do genoma (LI et al., 2004).

Muitos relatos demonstraram que um grande número de SSRs está localizada nas regiões transcritas do genoma, incluindo genes codificadores de proteínas e em sequências expressas (ESTs, de *expressed sequence tags*) (MORGANTE; HANAFEY; POWELL, 2002). Por exemplo, verificou-se que 10% das SSRs identificadas em primatas (JURKA; PETHIYAGODA, 1995), 15% em coelho (VAN LITH; VAN ZUTPHEN, 1996), e 9,1% e 10,6% em porco e frango (MORAN, 1993), respectivamente, estão localizadas em genes codificadores de proteínas ou em quadros de leitura abertos (ORFs). Em cereais (milho, trigo, cevada, sorgo e arroz) 1,5% a 7,5% das ESTs contêm SSRs (KANTETY et al., 2002; THIEL; MICHALEK; GRANER, 2003). Essas ESTs têm uma série de funções, pois são enzimas do metabolismo, proteínas estruturais e de armazenamento, de sinalização e fatores de transcrição, sugerindo haver algum envolvimento das SSRs no metabolismo da planta e no processo evolutivo (MARCOTTE et al., 1999). Debates sobre se as SSRs, de fato, desempenham papel funcional no desenvolvimento do organismo, adaptação, sobrevivência e evolução, são extensos na literatura.

SSRs são usadas como marcadores genéticos para mapeamento, identificação de cultivares, proteção de germoplasma, determinação de hibridismo e análises de variabilidade em *pools* gênicos e como marcas para diagnóstico molecular (KING; THOROGOOD; EDWARDS, 2008). Marcadores SSR são comumente empregados para a análise da estrutura genética de populações de espécies silvestres e cultivadas por conta de sua natureza codominante e alta informatividade (ZUCCHI et al., 2002; PINTO et al., 2003). Em adição, marcadores SSR são baseados em PCR, extremamente polimórficos, e altamente informativos

como resultado do número e da frequência de alelos detectados e sua habilidade de distinguir indivíduos intimamente relacionados (KING; THOROGOOD; EDWARDS, 2008).

As sequências que flanqueiam as repetições de SSRs são conservadas e, por isso, são usadas para o desenho de *primers* específicos, possibilitando a amplificação, via PCR, dos fragmentos que contêm o DNA repetitivo em indivíduos da mesma espécie, e, em alguns casos, entre diferentes espécies ou até gêneros da mesma família (CHOUMANE et al., 2004).

A detecção dos alelos SSRs é feita em gel de poliacrilamida, ou agarose especial de alta resolução, uma vez que é necessária uma matriz adequada para a separação eletroforética de alelos que diferem por poucos pares de bases. Podem ainda ser detectados pelo uso de *primers* fluorescentes por eletroforese em plataformas de sequenciamento. Cada loco de SSR é analisado a partir do par de *primers* construído especificamente para a sua amplificação. Mais do que um loco pode ser analisado simultaneamente quando os alelos de cada loco têm tamanhos suficientemente diferentes e migram para zonas separadas do gel. Nestes métodos de genotipagem, denominados “multiplex”, mais de um par de *primers* específicos são utilizados na mesma PCR, otimizando, portanto, o processo de genotipagem (CHAMBERLAIN et al., 1988; EDWARDS; GIBBS, 1994).

Como método de isolamento de SSRs, têm-se as bibliotecas genômicas enriquecidas, que consistem em um conjunto de clones de DNA entre os quais são selecionados os de interesse, ou seja, cujos insertos contêm SSRs complementares a uma sonda repetitiva (BILLOTE et al., 1999; KIJAS et al., 1994). SSRs também podem ser isoladas a partir de bibliotecas de cDNA. Este processo requer o isolamento e sequenciamento dos clones contendo SSRs, desenho dos *primers*, otimização da amplificação via PCR para cada par de *primers*, e um teste de polimorfismo em alguns indivíduos não relacionados. Desde então, centenas de espécies vegetais foram estudadas com base nesta tecnologia, com vistas a analisar a sua diversidade.

Com o estabelecimento dos projetos de sequenciamento de ESTs, visando à descoberta de transcritos e genes em várias espécies de plantas, uma riqueza de informações de sequências tem sido obtida e depositada em bancos de dados (RUDD, 2003). Além disso, dados de sequência para muitos genes totalmente caracterizados e clones de cDNA têm sido gerados para várias plantas, tal como o arroz (KIKUCHI et al., 2003). Pelo uso de programas computacionais, os dados de ESTs, genes e clones de cDNA podem ser “baixados” a partir do *GenBank* e verificados quanto à presença de SSRs, que são referidos como EST-SSRs ou SSRs gênicos. Neste caso, a geração de EST-SSRs é relativamente fácil e de baixo custo; no entanto, a geração desses SSRs é limitada às espécies ou parentes próximos dos quais existe

um grande número de ESTs disponíveis. SSRs gênicos têm algumas vantagens intrínsecas em relação aos SSRs genômicos porque eles são rapidamente obtidos por triagem eletrônica (*in silico*), além de estarem presentes em regiões expressas do genoma. A utilidade desses SSRs também reside na sua esperada transmissibilidade porque os *primers* são projetados a partir das regiões codantes e, portanto, mais conservadas do genoma. Por conta dessas vantagens e da disponibilidade pública de grandes quantidades de dados de sequências, SSRs gênicos têm sido identificados, desenvolvidos e utilizados numa variedade de estudos em várias espécies vegetais (VARSHNEY; GRANER; SORRELLS, 2005).

Atualmente, as tecnologias de sequenciamento de última geração (NGS - *Next-Generation Sequencing*) permitem o desenvolvimento de um grande número de marcadores de DNA dentro de um curto espaço de tempo (YANG et al., 2012). Obviamente, o sequenciamento do DNA genômico em larga escala é uma estratégia útil para identificar todos os tipos de DNA repetitivo, incluindo as SSRs (NYBOM; WEISING; ROTTER, 2014).

Utilizando 6.216 ESTs de *Salvia fruticosa*, *S. miltiorrhiza*, *S. sclarea* e *Stenogyne rugosa*, Karaca et al. (2012) desenvolveram 75 marcadores EST-SSRs. Estes foram testados em 12 espécies da família *Lamiaceae*, sendo 52 transmissíveis através dos gêneros. Adicionalmente, foi verificado que esta transmissibilidade diminuía com o aumento da distância evolutiva das espécies. Em *Arabidopsis*, Cosson, Decroocq e Revers (2014), desenvolveram 96 marcadores SSR altamente polimórficos, a partir de sequências de *Arabidopsis* inseridas em BACs, permitindo o mapeamento de QTLs entre os 24 acessos da coleção de *Versailles* (*Versailles Arabidopsis Stock Center*), que contem uma alta parte da diversidade genética da espécie. Adicionalmente, mostrou-se que apenas três desses marcadores permitem diferenciar todos os 24 acessos, sendo uma poderosa ferramenta para controlar a identidade ou o cruzamento entre quaisquer dos acessos. Silva et al. (2013), trabalhando com *Brachiaria ruziziensis*, uma das espécies forrageiras mais importantes cultivadas nos trópicos, obtiveram 198 SSRs polimórficos em 11 acessos de *B. ruziziensis*. Esses marcadores foram obtidos a partir da montagem parcial do genoma da espécie através do sequenciamento de *single-ends*. O uso desta abordagem para o desenvolvimento de marcadores SSR é promissor para as espécies com limitada informação genômica, cujos programas de melhoramento se beneficiariam a partir do uso de ferramentas genômicas. Este é o primeiro conjunto de marcadores SSR desenvolvidos para esta espécie.

Uma vantagem adicional do uso do sequenciamento de última geração para o desenvolvimento de marcadores SSR é a grande quantidade de dados de sequência nuclear e de organelas gerado (NYBOM; WEISING; ROTTER, 2014). Por exemplo, Krapp et al.

(2012), agruparam aproximadamente 84% do genoma cloroplastidial de *Dyckia marnier-lapostollei* (Bromeliaceae), desenvolvendo um total de 34 SSRs cloroplastidiais (cpSSRs), que produziram produtos de amplificação altamente polimórficos em várias espécies de *Dyckia*.

Como novas metodologias de sequenciamento estão sendo descobertas constantemente, em ritmo rápido, uma redução ainda maior dos custos pode ser esperada. Além disso, segundo Nybom, Weising e Rotter (2014), a grande quantidade de dados de sequências obtida em um experimento, em pequena escala, permite a seleção de SSR candidatos mais promissores, por exemplo, apenas com repetições de trinucleotídeos, ou com repetições perfeitas, aumentando a chance de sucesso na geração de SSRs para uso como marcadores genético/funcionais.

2.3 Marcadores SNP

Variações em um único nucleotídeo em uma dada sequência genômica de indivíduos de uma população são conhecidas como SNPs. Eles ocorrem em frequências mais elevadas que aquelas esperadas para as mutações de ponto. O potencial dos SNPs tem sido claramente demonstrado nas análises do genoma humano, pois vários milhões de marcadores SNP foram identificados e analisados (FRAZER et al., 2007). Com um número tão grande de marcadores disponível, é possível o exame de todo o genoma; como a densidade de marcadores é extremamente alta, isto permite a associação do perfil dos locos de SNPs com os fenótipos de caracteres de herança complexa, o que é chamado de *genome-wide association studies* (GWAS), ou genética associativa (MCCARTHY, 2008).

Apesar dos SNPs serem amplamente distribuídos ao longo dos genomas, sua ocorrência e distribuição variam entre as espécies. No milho, por exemplo, há 1 SNP a cada 60-120 pb (CHING et al., 2002), enquanto é estimado que em humanos há 1 SNP a cada 1.000 pb (SACHIDANANDAM et al., 2001), sendo mais predominantes em regiões não codantes do genoma. Nas regiões codantes, um SNP é dito “não sinônimo” quando resulta em mudança na sequência de aminoácidos (SUNYAEV et al., 1999), ou é dito “sinônimo”, quando não altera a sequência de aminoácidos da proteína correspondente.

Com o advento do sequenciamento em larga escala e os diferentes projetos genoma em execução, a disponibilidade de sequências para diferentes organismos tornou-se uma das mais importantes fontes de marcadores moleculares. Em plantas, essas fontes estão sendo exploradas para o desenvolvimento de marcadores informativos, tais como SSRs e SNPs e de

marcadores direcionados a sequências alvos, como os RGAs (*Resistance Gene Analogs*) (LOPEZ et al., 2003).

Métodos de genotipagem de alto rendimento, incluindo abordagens como *chips* de DNA, PCR alelo-específico e abordagens com extensão de *primer*, que envolvem o sequenciamento de *amplicons* na busca de variantes, tornam os SNPs especialmente atrativos como marcadores genéticos. Eles são adequados para a automação e são usados para uma variedade de propósitos, incluindo a identificação de cultivares e construção de mapas genéticos altamente densos (AGARWAL; SHRIVASTAVA; PADH, 2008). Locos do SNPs também proporcionam informações a cerca da sentença entre o genoma da espécie alvo e o de uma espécie modelo. Além disso, os SNPs também são considerados estáveis durante os processos evolutivos das espécies, o que permite o seu uso em estudos de genética de populações e filogenia (JIN et al., 2003).

Diversas técnicas são usadas para a identificação de vários números de SNPs em uma determinada espécie. Para muitas plantas cultivadas, um grande número de ESTs está disponível, sendo fonte para a busca de SNPs. Outra abordagem que é também baseada em ESTs envolve o uso de *arrays* contendo oligonucleotídeos derivados de um grande número de genes. Em resumo, este método envolve o desenvolvimento de *primers* para a amplificação de fragmentos de DNA derivados de genes, ESTs ou outras sequências genômicas de cópia única. Os produtos de amplificação de um número de genótipos representativos são totalmente sequenciados e as sequências obtidas são posteriormente comparadas umas com as outras usando ferramentas de alinhamento de sequências e *pipelines* de bioinformática. A vantagem desta abordagem é que a sequência de cada indivíduo investigado é determinada através do sequenciamento das duas fitas de DNA e os SNPs podem ser identificados de maneira muito confiável, com uma taxa de erro usualmente significativa abaixo de 5%. Outra vantagem é que, com o sequenciamento de fragmentos grandes, é possível identificar não somente SNPs, mas também haplótipos (GANAL; ALTMANN; RÖDER, 2009). Muitos espécies foram investigadas com esta abordagem, como milho, *Arabidopsis*, arroz, tomate, beterraba-sacarina, cevada e abeto.

Em arroz, utilizando seis diferentes cultivares, foi possível a identificação de 2.800 SNPs a partir de 417 sequências anotadas, sendo possível a obtenção de 213 marcadores codominantes (NASU et al., 2002). Em *Arabidopsis*, Nordborg et al. (2005) encontraram um total de mais de 17.000 SNPs em uma população de 96 indivíduos, sendo possível estudar a estrutura genética desta população. Em tomate, Van Deynze et al. (2007) realizaram um *screening* em 967 transcritos, pertencentes a um grupo de ESTs conservadas, de cópia única,

ortólogas entre tomate e *Arabidopsis*, para a busca de polimorfismo em 12 linhagens, sendo possível a identificação de 1.487 SNPs e 282 *indels*. Em beterraba-sacarina, Schneider et al. (2007) identificaram 538 SNPs, a partir de 315 ESTs e 43 locos não codantes derivados de RFLPs, utilizando-os para a construção de mapas de ligação em três populações. Em cevada, 437 sequências foram obtidas a partir de ESTs, sendo possível a identificação de 1.257 SNPs, dos quais 216 foram utilizados para a construção de um mapa genético consenso desenvolvido a partir de dados de segregação de marcas de três populações (KOTA et al., 2008). Em abeto, Pavy et al. (2008) utilizaram um *array* de 768 SNPs identificados no DNA genômico de genitores de duas populações, uma de abeto branco (*Picea glauca*) e outra de abeto preto (*Picea mariana*), sendo possível a construção de dois mapas de ligação individuais para cada espécie, bem como a construção de uma mapa composto para os genomas de *Picea*. No milho, usando mapeamento associativo e escaneamento de todo o genoma, foi possível a identificação de locos com maior efeito no conteúdo de ácido oleico na semente (BELÓ et al., 2008).

2.4 Polimorfismo molecular em *Passiflora edulis* e *P. alata*

P. edulis e *P. alata* são espécies diploides e auto-incompatíveis (BRUCKNER et al., 1995; FERREIRA et al., 2010; REGO et al., 2000), altamente dependentes da polinização cruzada para a frutificação devido a presença da autoincompatibilidade, mecanismo que impede a autofertilização, favorecendo o cruzamento entre plantas, e assim promove variabilidade genética dentro da espécie (BRUCKNER et al., 1995; HIGASHIYAMA, 2010). Espera-se que a autoincompatibilidade mantenha altos níveis de heterozigosidade dentro das populações. No entanto, pouco se sabe sobre o grau de variabilidade e heterozigosidade dentro das populações naturais e artificiais de *P. edulis* e, mormente, de *P. alata* (veja PENHA et al., 2013).

Estudos recentes sobre o polimorfismo molecular em acessos de maracujá-doce e maracujá-azedo, realizados com base em vários tipos de marcadores, sugerem haver um baixo nível de polimorfismo, o que não seria esperado para espécies de fecundação cruzada. Exemplificando, Penha et al. (2013), trabalhando com marcadores SSR, encontraram 35% de divergência entre os dois acessos de maracujá-doce, genitores de uma população de mapeamento; Pereira et al. (2013), trabalhando com marcadores SSR, AFLPs, RGAs, TRAPs e SNPs, e com os mesmos acessos, encontraram 20,3% de polimorfismo. Em maracujá-azedo, Oliveira et al. (2008) encontraram 25% de polimorfismo entre dois acessos genitores

fenotipicamente contrastantes de uma população de mapeamento. Nestes trabalhos, os acessos investigados eram oriundos de regiões geográficas bem distintas.

Nesses estudos, o polimorfismo molecular foi medido usando dois acessos fenotipicamente distintos de diferentes origens geográficas. Como todas as plantas são produzidas naturalmente por reprodução sexual alógama, seria esperado encontrar maiores níveis de polimorfismo molecular. Além disso, espécies cultivadas de *Passiflora* são caracterizadas como culturas incipientes, pouco selecionadas. Também, é válido notar que há muito pouca informação de sequências de DNA sobre as passifloras nos bancos de dados públicos, limitando a análise da variabilidade (PEREIRA et al., 2013).

No que se refere a estudos populacionais, mais recentemente, o polimorfismo de microssatélites foi investigado usando 114 genótipos de uma coleção de passifloras da Empresa Brasileira de Pesquisa Agropecuária (Cruz das Almas, Bahia) e da UESB (Itapetinga, Bahia). A percentagem de locos polimórficos foi de 15% em *Passiflora edulis*, 29% em *P. setacea* e 20% em *P. cincinnata* (CERQUEIRA-SILVA et al., 2014). Este nível de polimorfismo, de 15 a 30%, parece ser característico do gênero, como demonstrado em vários trabalhos do nosso grupo (CARNEIRO; VIEIRA, 2002; LOPES et al. 2006; OLIVEIRA et al., 2008; PEREIRA et al., 2013; PENHA et al., 2013;) e sugerido na literatura (CERQUEIRA-SILVA et al., 2012).

3 MATERIAL E MÉTODOS

3.1 Material Biológico

Nos testes de amplificação dos marcadores SSR e SNP, foram avaliados os genitores da população de mapeamento estudada por Pereira et al. (2013), a saber, os acessos ‘SV3’ e ‘2(12)’, e, a título de controle, também foram usados os acessos ‘IAPAR-123’ e ‘IAPAR-06’, genitores de uma população de mapeamento de maracujá-azedo (OLIVEIRA et al., 2008). Nos testes de polimorfismo para ambos os marcadores foram avaliados os mesmos genitores da população de mapeamento estudada por Pereira et al. (2013) e parte da sua progênie (Tabela 1). O DNA de todos os acessos foi extraído por Nunes (2010), utilizando o protocolo de Murray e Thompson (1980) adaptado, e armazenado a -20 °C.

Tabela 1 - Acessos de maracujá-doce e de maracujá-azedo utilizados nos testes de amplificação e polimorfismo para os marcadores SSR e SNP

	SSR	SNP
Testes de amplificação	‘2(12)’, ‘SV3’, ‘IAPAR-123’ e ‘IAPAR-06’	‘2(12)’, ‘SV3’, ‘IAPAR-123’ e ‘IAPAR-06’
Testes de polimorfismo	‘2(12)’, ‘SV3’, 30 indivíduos da população F ₁ segregante	‘2(12)’, ‘SV3’, ‘F ₁ -67’ e ‘F ₁ -100’

3.2 Análise das sequências de transcritos para o desenvolvimento dos marcadores SSR e SNP

Para o desenvolvimento dos marcadores SSR e SNP, foi usado um conjunto de sequências expressas de maracujá-azedo (genótipo ‘IAPAR123’). Estas sequências foram obtidas por Munhoz (2013), cujo objetivo foi identificar transcritos de maracujá-azedo diferencialmente expressos em resposta à inoculação por *Xanthomonas axonopodis* pv. *passiflorae* (Xap). Para isso, a autora construiu duas bibliotecas de expressão usando a metodologia de hibridização subtrativa por supressão (SSH) (DIATCHENKO et al., 1996), sendo que na biblioteca *forward* foram isolados os transcritos diferencialmente expressos na planta inoculada com Xap e na biblioteca *reverse* os transcritos diferencialmente expressos na planta controle, não inoculada.

Especificamente para o desenvolvimento dos marcadores SSR foi usado o conjunto de sequências da biblioteca *forward* (42 transcritos, ao final da análise das sequências) enquanto para o desenvolvimento dos marcadores SNP foram usados os transcritos de ambas as bibliotecas (118 transcritos, ao final da análise das sequências).

3.3 Marcadores SSR

3.3.1 Prospeção e classificação dos locos SSR

Utilizaram-se 672 sequências oriundas da biblioteca de expressão *forward* para a busca de SSRs, com o auxílio do software Sciroko ver 3.4 (KOFER; SCHÖTTERER; LELLEY, 2007). Foram considerados insertos positivos aqueles que apresentaram repetições maiores ou iguais a 10 para mono-, 5 para di-, e 3 para tri-, tetra-, penta- ou hexanucleotídeos. A classificação, quanto ao número de repetições, foi baseada em cada motivo, incluindo todas as sequências complementares e as permutadas. Por exemplo, o motivo dinucleotídico AG representa as sequências AG, GA, e os complementos CT e TC e o motivo trinucleotídico GAA representa as sequências GAA, AAG, AGA e seus complementos TTC, TCT e CTT (Penha, 2007).

3.3.2 Desenho dos *primers* e testes de amplificação

As sequências contendo SSRs foram utilizadas para o desenho dos pares de *primers* flanqueadores da região repetida usando o aplicativo *web PRIMER3* (ROZEN; SKALETISKY, 2000; disponível em: <<http://frodo.wi.mit.edu/primer3/>>). A fim de se obter um bom padrão de amplificação para vários pares de *primers*, utilizaram-se os seguintes critérios: (i) temperatura de anelamento (TA) variando de 52 a 60 °C, (ii) proporção de GC entre 40 e 60%, (iii) diferença máxima de 3 °C na TA entre os *primers* de cada par. Adicionalmente, cada par de *primers* foi checado quanto às possibilidades de auto anelamento ou anelamento cruzado com auxílio do *software* GENE RUNNER (SPRUYT; BUQUICCHIO, 1994).

A síntese dos *primers* foi realizada pela empresa Sigma-Aldrich na escala de 25 nmol. A nomenclatura de cada par de *primers* contém as iniciais da espécie, o nome do clone que originou cada sequência, segundo o esquema adotado por Munhoz (2013), anotando-se o número da placa da biblioteca (que varia de 1 a 10) e a respectiva coordenada do clone (as letras A até H identificam a linha, e os números 1 a 12 identificam a coluna). Por exemplo,

PeA10P10 corresponde ao clone de *P. edulis* cuja sequência possibilitou o desenho de um par de *primers*, estando presente na “Placa 10”, sob a coordenada “linha A-coluna 10”. Clones com duas ou mais corridas originaram *contigs* (sequências consenso). *Primers* desenhados a partir dessas sequências foram nomeados de acordo com o número do *contig*.

Os testes de amplificação dos pares de *primers* foram realizados com os DNAs dos genótipos ‘2(12)’ e ‘SV3’ de maracujá-doce e, como controle, foram usados os DNAs dos genótipos ‘IAPAR-123’ e ‘IAPAR-06’, de maracujá-azedo. As reações foram realizadas contendo cerca de 30 ng de DNA genômico, 1X de tampão de PCR, 2 mM de MgCl₂, 1,2 U de enzima Taq DNA polimerase, 0,25 mM de cada dNTP, 0,3 µM de cada *primer*, e água ultrapura para um volume final de 20 µL. As amplificações foram realizadas em termociclador utilizando o seguinte programa: desnaturação inicial a 95 °C por 5 min, seguida de 35 ciclos de 95 °C por 40 s, 55 °C por 40 s e 72 °C por 1 min, com extensão final de 72 °C por 8 min.

Os produtos das reações foram submetidos à eletroforese (3V/cm) em gel de agarose 1,2% em TBE 1 X (Tris-base 89 mM, ácido bórico 89 mM, Na₂EDTA 2,5 mM) com corante fluorescente 0,5 X (SYBR® Safe DNA Gel Stain, Promega). O gel foi visualizado e digitalizado sob luz ultravioleta (UV) em equipamento de fotodocumentação (Multi Doc-It Digital Imaging System, UVP).

3.3.3 Transferibilidade de *primers* de *P. edulis* f. *flavicarpa* para *P. alata*

No intuito de testar a transferibilidade dos *primers*, os produtos de PCR obtidos pela amplificação com os acessos de maracujá-doce e de maracujá-azedo, espécie da qual as sequências foram obtidas, foram avaliados em géis de agarose. De acordo com a visualização do padrão de bandas, eventualmente, foi necessário realizar a otimização das condições de amplificação e da concentração dos reagentes. Para os pares de *primers* em que não houve amplificação, ou que houve amplificação de bandas inespecíficas em *P. alata*, foi feita a otimização das condições de PCR usando quatro diferentes *mix* de reação (Tabela 2), como proposto por Penha (2007).

Tabela 2 - Concentração dos reagentes e relação dos *mix* utilizados para a otimização da PCR para cada par de *primers* usados nos testes de transferibilidade

Reagentes	Mix 1	Mix 2	Mix 3	Mix 4
Tampão	1X	1X	1X	1X
<i>Primer F+R*</i>	0,3µM	0,3 µM	0,3 µM	0,3 µM
DNA	30 ng	30 ng	30 ng	30 ng
MgCl ₂	1 mM	2 mM	2 mM	2,5 mM
dNTP	0,25 mM	0,25 mM	0,25 mM	0,25 mM
<i>Taq polimerase</i>	1,2 U	1,2 U	1,2 U	1,2 U
Água deionizada	9,16 µL	8,76 µL	8,36 µL	7,96 µL

* F e R – *primers forward* e *reverse*, respectivamente.

3.3.4 Eletroforese e avaliação do polimorfismo

Os produtos resultantes da PCR foram separados em gel de poliacrilamida [acrilamida/bisacrilamida (19:1) 5%, uréia 7,0 M, tampão TBE 1X] de 0,4 mm de espessura. O gel foi corrido a uma potência constante de 80 W por cerca de 2 a 3 h, de acordo com o fragmento amplificado, usando um aparato de eletroforese Sequi-Gen® (Bio-Rad). O *Ladder* de 10 pb (Invitrogen) foi usado como marcador de peso molecular.

A coloração foi realizada com nitrato de prata (AgNO₃) em cinco etapas, de acordo com Creste, Tulmann Neto e Figueira (2001), com algumas modificações, como descrito a seguir:

Etapa 1 (fixação): a fixação do gel foi feita por imersão em solução de etanol 10% e ácido acético 1%, agitando por 15 min, seguida de uma lavagem com água destilada e deionizada, durante 1 min, com agitação.

Etapa 2 (oxidação): o pré-tratamento foi realizado com solução de ácido nítrico 1,5% durante 5 min, seguindo-se uma nova lavagem do gel com água destilada deionizada durante 1 min.

Etapa 3 (impregnação com prata): a impregnação do gel foi feita em solução de nitrato de prata 0,2% durante 20 min, sob agitação e ao abrigo de luz, seguida de duas novas lavagens com água destilada e deionizada, com duração de 30 segundos cada.

Etapa 4 (revelação): para a revelação dos géis, foi utilizada uma solução de 3L contendo 180 g de Na₂CO₃ e 1,8 mL de formaldeído (37%). Nesta etapa, 1 litro da solução foi aplicada sobre o gel até começarem a surgir bandas. Após este passo, a solução foi descartada e os outros dois litros foram adicionados um após o outro para finalizar o processo de revelação. Após a revelação da placa, a solução foi descartada.

Etapa 5 (bloqueio): o bloqueio da revelação foi feito em solução de ácido acético 5% durante 5 min, seguindo-se de uma nova lavagem com água destilada deionizada durante 1 min. Após este passo, para a secagem do gel, a placa foi colocada em posição vertical em um local seco e arejado, sob temperatura ambiente.

Os géis foram analisados colocando-se a placa sobre uma fonte de luz branca, e digitalizados por um scanner (Genius Color Page 3200).

Foram considerados polimórficos os locos que apresentaram alelos diferentes entre os acessos ‘SV3’ e ‘2(12)’ e segregaram na progênie F_1 . Adicionalmente, os marcadores polimórficos foram nominados segundo o tipo de segregação tal como proposto por Wu et al. (2002) para populações F_1 segregantes. Cada um dos alelos do marcador, simbolizados por a , b , c ou d , são codominantes em relação aos demais, mas dominantes em relação ao alelo nulo, simbolizado por o . Dependendo de como alelos diferentes são combinados nos dois genitores da população F_1 , existe um total de 18 possíveis tipos de cruzamento (configurações genéticas) para cada loco do marcador. Caracterizam-se, portanto, quatro tipos de segregações possíveis decorrentes do número de fenótipos moleculares na população, que varia de quatro até dois. Baseado no padrão de bandas dos marcadores em ambos os genitores e na progênie, esses 18 tipos de cruzamentos são classificados em sete grupos (MALIEPAARD; JANSEN; VAN OOIJEN, 1997; WU et al., 2002) (Tabela 3):

A) locos que são heterozigóticos em ambos os genitores e segregam na proporção de 1:1:1:1, envolvendo quatro alelos ($ab \times cd$), três alelos não-nulos ($ab \times ac$), três alelos não-nulos e um nulo ($ab \times cc$), ou dois alelos nulos e dois não-nulos ($ao \times bo$);

B) locos que são heterozigóticos em ambos os genitores e segregam na proporção de 1:2:1 os quais incluem três subgrupos:

B_1) um genitor com dois diferentes alelos dominantes e o outro com um alelo dominante e um nulo ($ab \times ao$);

B_2) o recíproco de B_1 ;

B_3) ambos os genitores têm o mesmo genótipo para dois alelos codominantes ($ab \times ab$);

C) locos que são heterozigóticos em ambos os genitores e segregam na proporção de 3:1 ($ao \times ao$);

D) locos que apresentam a configuração de cruzamento teste entre os genitores e segregam na proporção de 1:1, incluindo dois subgrupos:

D_1) heterozigótico em um genitor e homozigótico no outro, incluindo três alelos $ab \times cc$, dois alelos $ab \times aa$, $ab \times oo$ e $bo \times aa$, e um alelo, com três alelos nulos, $ao \times oo$.

D_2) o recíproco de D_1 .

Cabe esclarecer que as segregações apresentadas na tabela 3 não foram aqui testadas, apenas serviram como referencia para o reconhecimento do tipo de configuração genética e dos respectivos alelos.

Tabela 3 - Configurações genéticas dos locos marcadores, segundo o tipo de cruzamento, e os respectivos padrões de bandas observadas (alelos) nos genitores e nas progênes oriundas de cruzamentos entre genitores não-endogâmicos

Tipo de cruzamento	Genitores			Progenie			
		Cruzamento	Bandas observadas	Observação	Bandas observadas	Segregação	Nº de fenótipos
A		1 $ab \times cd$	$ab \times cd$	Assimétrico	ac, ad, bc, bd	1:1:1:1	4
		2 $ab \times ac$	$ab \times ac$	Assimétrico	a, ac, ba, bc	1:1:1:1	4
		3 $ab \times co$	$ab \times c$	Assimétrico	ac, a, bc, b	1:1:1:1	4
		4 $ao \times bo$	$a \times b$	Assimétrico	ab, a, b, o	1:1:1:1	4
B	B1	5 $ab \times ao$	$ab \times a$	Assimétrico	$ab, 2a, b$	1:2:1	3
	B2	6 $ao \times ab$	$a \times ab$	Assimétrico	$ab, 2a, b$	1:2:1	3
	B3	7 $ab \times ab$	$ab \times ab$	Simétrico	$a, 2ab, b$	1:2:1	3
C		8 $ao \times ao$	$a \times a$	Simétrico	$3a, o$	3:1	2
D	D1	9 $ab \times cc$	$ab \times cc$	Assimétrico	ac, bc	1:1	2
		10 $ab \times aa$	$ab \times aa$	Assimétrico	a, ab	1:1	2
		11 $ab \times oo$	$ab \times o$	Assimétrico	a, b	1:1	2
		12 $bo \times aa$	$b \times aa$	Assimétrico	ab, a	1:1	2
		13 $ao \times oo$	$a \times o$	Assimétrico	a, o	1:1	2
	D2	14 $cc \times ab$	$cc \times ab$	Assimétrico	ac, bc	1:1	2
		15 $aa \times ab$	$aa \times ab$	Assimétrico	a, ab	1:1	2
		16 $oo \times ab$	$o \times ab$	Assimétrico	a, b	1:1	2
		17 $aa \times bo$	$aa \times b$	Assimétrico	ab, a	1:1	2
		18 $oo \times ao$	$o \times a$	Assimétrico	a, o	1:1	2

Fonte: Adaptado de Wu et al. (2002)

3.4 Marcadores SNP

3.4.1 Seleção das sequências e desenho de *primers*

Dos 998 transcritos obtidos nas bibliotecas de expressão por Munhoz (2013), 118 (Tabela 4) foram selecionados para o desenho de *primers*. As sequências correspondentes aos transcritos apresentados na tabela 4 foram analisadas visando desenhar os *primers forward* e *reverse*. A fim de se obter um bom padrão de amplificação, utilizaram-se os seguintes critérios: (i) temperatura de anelamento (TA) variando de 52 a 60 °C, (ii) proporção de G+C entre 40 e 60%, (iii) Tm (*melting temperature*) dos membros do par de *primers* não diferir mais de 5 °C, (iiii) menor possibilidade de formação de estruturas secundárias, como *hairpins*, dímeros e *loops*. Para o desenho e análise das sequências dos *primers* foi utilizado o programa GENE RUNNER (SPRUYT; BUQUICCHIO, 1994). A síntese dos *primers* foi realizada pela

empresa Sigma-Aldrich na escala de 25 nmol. A nomenclatura adotada foi a mesma descrita para os SSRs gênicos.

Tabela 4 - Lista dos transcritos selecionados das bibliotecas de expressão para o desenho de *primers* e proteína codificada pelo respectivo transcrito

(continua)

Nome da sequência	Proteína codificada
PeA4P8	Putative serine proteinase
PeA5P9	ABC transporter F family member 5
PeA6P9	Lipoxygenase family
PeA12P4	ABC transporter G family member 15
PeA12P5	ABC transporter G family member 35
PeB8P9	Protein kinase 2B
PeC1P6	Proteinase-like protein
PeC1P9	DNA-binding family protein
PeC1P10	Protein phosphatase 2C 7
PeC3P4	Cysteine proteinase inhibitor 3
PeC6P4	Putative polygalacturonase
PeC9P4	Nudix hydrolase
PeC10P2	SUMO-conjugating enzyme UBC9
PeC11P10	Probable serine/threonine-protein kinase
PeContig1	Trypsin and protease inhibitor
PeContig10	Putative cathepsin B-like protease family protein
PeContig11	Putative hydroxypyruvate reductase
PeContig13R	Ribulose biphosphate carboxylase oxygenase activase chloroplast
PeContig14	Peptidase C1A subfamily
PeContig15	Light harvesting-like protein 3
PeContig16	Basic endochitinase B
PeContig17	Cytochrome c oxidase subunit 2
PeContig19	Putative ATP citrate lyase
PeContig20	Oxygen-evolving enhancer protein 1-1
PeContig20R	MPBQ MSBQ methyltransferase 2
PeContig21	30S ribosomal protein 3-1
PeContig22R	Pdi-like protein
PeContig27	Hevein-like protein
PeContig29	Cationic amino acid transporter
PeContig33	Mitochondrial outer membrane protein porin 1
PeContig35	DEAD-box ATP-dependent RNA helicase 3
PeContig39	RNA-dependent RNA polymerase 1
PeContig40	Serine/threonine-protein phosphatase 2A
PeContig41	Elongation factor Tu
PeContig45	NAD(P)H-quinone oxidoreductase subunit J

Tabela 4 - Lista dos transcritos selecionados das bibliotecas de expressão para o desenho de *primers* e proteína codificada pelo respectivo transcrito

(continuação)

Nome da sequência	Proteína codificada
PeContig50	NADP-dependent malate dehydrogenase
PeContig5	Basic endochitinase B
PeContig52	31 kDa ribonucleoprotein, chloroplastic
PeContig53	predicted protein
PeContig57	Glycine dehydrogenase [decarboxylating] 2
PeContig6	ATP synthase gamma chain 1
PeContig8R	Chlorophyll a b binding
PeD1P8	Glutamate receptor 3.7
PeD4P6	Guanylylate cyclase
PeD5P4	box/LRR-repeat protein 4
PeD5P10	Sugar transport protein 13
PeD7P9	Putative uncharacterized protein
PeD8P6	Aldehyde dehydrogenase family 3 member I1
PeD10P6	global transcription factor group
PeE1P7	Dihydrolipoamide S-acetyltransferase
PeE8P7	Nucleobase-ascorbate transporter 6
PeE12P7	Protein transport inhibitor response 1
PeF1P8	Syntaxin-32
PeF1P9	Auxin response factor 2
PeF7P5	Serine/threonine-protein kinase CTR1
PeF12P4	GATA transcription factor 8
PeF12P5	36kDa-peroxisomal membrane protein (PMP36)
PeG1P9	ABC transporter I family member 10
PeG2P6	Probable carboxylesterase 15
PeG5P9	Calcium-dependent protein kinase 33
PeG6P4	Glutamate--cysteine ligase, chloroplastic
PeG6P7	Lipoxygenase 2
PeG10P5	Leucine-rich repeat receptor-like protein kinase
PeG12P5	Receptor-like protein kinase
PeG12P6	Putative monooxygenase
PeG12P7	predicted protein
PeH3P9	Alpha-L-arabinofuranosidase 1
PeH5P4	Putative beta-carotene hydroxylase
PeH5P10	Inactive beta-amylase 9
PeH5P8	Late embryogenesis abundant protein
PeH6P10	Nuclear transcription factor Y subunit C-9

Tabela 4 - Lista dos transcritos selecionados das bibliotecas de expressão para o desenho de *primers* e proteína codificada pelo respectivo transcrito

(continuação)

Nome da sequência	Proteína codificada
PeH7P6	Beta-glucosidase-like SFR2
PeH8P3	Homocysteine S-methyltransferase 3
PeH10P4	Alpha-L-fucosidase 1
PeH10P7	Probable WRKY transcription factor 33
PeH11P9	Putative type 2 ribosome-inactivating protein
PeH12P9	Putative casein kinase I
PeP1RA5	Harpin-induced protein
PeP1RB6	GTP-binding protein gb2
PeP1RD10	Mechano sensitive ion channel domain-containing protein
PeP1RE8	Calmodulin-like protein 1-like
PeP1RG8	box family protein
PeP1RH6	Exosome complex component mtr3-like
PeP2RA12	Ferredoxin i precursor
PeP2RB8	Leucine-rich repeat family protein
PeP2RF11	Auxin-responsive protein
PeP2RF12	Mitotic checkpoint protein
PeP2RF4	Stearoyl-acp desaturase
PeP2RF6	Methyltransferase pmt20
PeP2RG11	Heat shock transcription factor hsf1
PeP2RH10	RNA helicase
PeP3RA2	Lactoylglutathione lyase
PeP3RB4	Cc-nbs-Irr resistance protein
PeP3RB6	Kinesin-like protein
PeP3RC1	Strubbelig-receptor family 7 protein
PeP3RC2	Receptor protein kinase clavata1
PeP3RC9	Uncharacterized protein
PeP3RF11	Vacuolar h ⁺ -translocating inorganic pyrophosphatase
PeP3RF2	Proteasome subunit beta type-1
PeP3RF8	Threonine dehydratase deaminase mutant
PeP3RG12	Shaggy-related protein kinase eta
PeP3RH11	Peroxisomal acyl- oxidase 1 ^a
PeP3RH3	Receptor protein
PeP4RA10	Ribonuclease t2
PeP4RB1	ribulose bisphosphate small subunit
PeP4RB2	Pheophorbide a oxygenase
PeP4RB3	Cellulose synthase

Tabela 4 - Lista dos transcritos selecionados das bibliotecas de expressão para o desenho de *primers* e proteína codificada pelo respectivo transcrito

Nome da sequência	Proteína codificada	(conclusão)
PeP4RB5	Glycolate oxidase	
PeP4RB7	Superoxide dismutase	
PeP4RB8	Xyloglucan endotransglycosylase hydrolase	
PeP4RD10	E3 ubiquitin-protein ligase	
PeP4RD5	Loxc homologue	
PeP4RD8	Ubiquitin-conjugating enzyme	
PeP4RF3	Homeobox-leucine zipper protein athb-6	
PeP4RG8	cdp-diacylglycerol--glycerol-3-phosphate3phosphatidyltransferase	
PeP4RH8	glyceraldehyde-3-phosphate dehydrogenase	
PeP4RH10	RNA-binding protein	

3.4.2 Condições de amplificação e transferibilidade dos *primers* de *P. edulis* f. *flavicarpa* para *P. alata*

Como as sequências que foram usadas para o desenho dos *primers* (Tabela 4) são correspondentes a transcritos de *P. edulis*, testou-se a transferibilidade dos *primers* para uso em *P. alata*. Assim, as reações de PCR foram realizadas com os acessos ‘2(12)’ e ‘SV3’ de maracujá-doce e ‘IAPAR-123’ e ‘IAPAR-06’ de maracujá-azedo. De acordo com o resultado das amplificações no maracujá-doce, ou seja, a visualização do padrão de bandas em agarose, eventualmente, foi necessário otimizar as condições de amplificação. A partir desta avaliação, para alguns *primers* que não resultaram em amplificação, ou que resultaram bandas inespecíficas em *P. alata*, foi alterada a temperatura de anelamento, de 55 a 65 °C, de acordo com o *primer*, a fim de se obter uma amplificação específica, ou seja, uma única banda no gel.

As PCRs foram realizadas contendo cerca de 30 ng de DNA genômico, 1X de tampão de PCR, 2 mM de MgCl₂, 1,2 U de enzima Taq DNA polimerase, 0,25 mM de cada dNTP, 0,3 µM de cada *primer*, e água ultrapura para um volume final de 20 µL. As amplificações foram realizadas em termociclador usando o seguinte programa: desnaturação inicial a 95 °C por 5 min, seguida de 35 ciclos de 95 °C por 40 s, 55 °C por 40 s e 72 °C por 1 min, com extensão final de 72 °C por 8 min.

O produto das amplificações foi separado por eletroforese em gel de agarose 1,2%.

3.4.3 Avaliação, purificação e sequenciamento dos produtos de PCR

As amplificações que resultaram em bandas específicas foram usadas para a busca de SNPs através do sequenciamento direto dos *amplicons*. Para cada amplificação que resultou em banda específica, foram sequenciados os *amplicons* dos acessos ‘2(12)’ e ‘SV3’ em duplicata, e o *amplicon* de dois indivíduos da progênie, escolhidos aleatoriamente, quais sejam, F₁-67 e F₁-100, para se verificar a presença (ou não) de SNPs. Para isso, estes produtos de PCR foram purificados com o *kit* Illustra GFX 96 PCR Purification Kiq® (GE Healthcare), conforme as especificações do fabricante. As concentrações dos produtos purificados foram estimadas por espectrofotometria, usando-se o equipamento NanoDrop 2000® (Thermo Scientific), e o sucesso do procedimento de purificação foi verificado a partir da eletroforese de alíquotas de 5 µL desses produtos em gel de agarose (1,2%), corado com SYBR SAFE® (Invitrogen). Neste caso, a quantificação das amostras foi estimada a partir da comparação com alíquotas de DNA do fago λ de concentrações conhecidas. Os géis foram visualizados sob luz UV e fotodocumentados pelo sistema de imagem digital MultiDoc-It - UVP.

Nas reações de sequenciamento foi usado o BigDye® Terminator v3.1 Cycle Sequencing Kit (ABI, CA, USA). Foram realizadas normalizações das concentrações dos produtos de PCR purificados a fim de garantir uma relação ótima entre o tamanho do *amplicon* correspondente e a quantidade (em ng) a ser usada nas reações de sequenciamento. As condições de reação foram modificadas em relação àquelas sugeridas pelo fabricante, a fim de melhorar a relação custo benefício e garantir um bom nível de qualidade e tamanho de sequência. Cada reação de sequenciamento continha 50 ng de produto de PCR purificado, 2 µL de tampão 5x, 0,6 µL de BigDye Terminator v3.1, 1,0 µL de *primer* 5 µM, sendo completada com água ultrapura até o volume de 10 µL. As amplificações foram realizadas em termociclador da Applied Biosystems, utilizando o seguinte programa: desnaturação inicial a 96 °C por 1 min, seguida de 25 ciclos de 96 °C por 10 s, 50 °C por 5 s e 60 °C por 4 min. As reações de sequenciamento bem como a eletroforese (ABI 3730 DNA Analyser®) foram realizadas no Centro de Estudos do Genoma Humano da Universidade de São Paulo.

3.4.4 Processamento, alinhamento das sequências e identificação dos potenciais SNPs

O software CODONCODEALIGNER v. 3.5.7 (CodonCode Corporation) foi usado para a leitura dos arquivos de saída dos sequenciadores automáticos. Esta suíte permite a visualização dos eletroferogramas, realiza as análises de qualidade das bases lidas (função

‘*Call bases with phred*’) e elimina as bases iniciais e finais das sequências que possuem baixa qualidade (função ‘*Clip ends*’). O critério utilizado no “*end clipping*” foi o *default* do software, que adota índice phred de qualidade “20” como limite para considerar a base como verdadeira, ou seja, a probabilidade de a base estar incorreta é de 1 em 100.

As sequências foram alinhadas utilizando a função ‘*Assemble*’ e os SNPs candidatos foram identificados por meio de inspeção visual dos eletroferogramas, corrigindo-se, manualmente, eventuais equívocos de leitura.

Foram considerados polimórficos os sítios em que havia variação em um mesmo ponto (base) nas sequências dos progenitores ‘2(12)’ e ‘SV3’.

4 RESULTADOS

4.1 Marcadores SSR

4.1.1 Caracterização das regiões contendo SSR

Dos 672 transcritos da biblioteca de expressão *forward*, foram encontrados 91 (13,5%) contendo 115 SSRs. O software *Sciroko* ver 3.4 (KOFLEK; SCHÖTTERER; LELLEY, 2007) utilizado na busca dos SSRs mostrou-se eficiente na busca dos motivos perfeitos (Figura 1), sendo a classificação dos motivos resultante da análise do avaliador (Figura 2), através da identificação do motivo na sequência.

Sequence_Name	Total_Length	Score	Motif	Motif_Standardized (Total_sequence)	SSR_Start	SSR_End			
A01_PL4_01			CA	AC	365	374	10	10	0
A01_PLACA7_01			TTTC	AAAG	194	205	12	12	0
A02_PL4_02			AT	AT	127	142	16	16	0
A02_PL9_02			TTG	AAC	327	338	12	12	0
A02_PLACA2_02			TA	AT	305	317	13	13	0
A02_PLACA2_02			GT	AC	328	342	15	15	0
A03_PL6_01			AGA	AAG	122	131	10	10	0
A03_PLACA8_01			TG	AC	488	498	11	11	0
A04_PL9_02			CAG	AGC	202	212	11	11	0
A05_PLACA2_01			AAG	AAG	275	286	12	12	0
A06_PLACA8_02			TC	AG	64	78	15	15	0
A08_PLACA2_02			GGCGTC	ACGCCG	81	99	19	19	0
A09_PLACA7_01			GGGA	AGGG	5	17	13	13	0
A10_PL4_02			CAG	AGC	121	131	11	11	0
A10_PL9_02			GA	AG	260	269	10	10	0
A10_PLACA10_02			TTTG	AAAC	7	27	21	21	0
A10_PLACA10_02			T	A	125	134	10	10	0
A10_PLACA10_02			GGT	ACC	370	381	12	12	0
A10_PLACA8_02			CTT	AAG	208	217	10	10	0
A12_PL9_02			CTG	AGC	233	242	10	10	0
A12_PLACA3_02			TGC	AGC	448	457	10	10	0
B01_PL1_03			AGTTTC	AAACTG	232	255	24	24	0
B01_PL5_03			AGA	AAG	181	192	12	12	0
B01_PL6_03			TTG	AAC	54	64	11	11	0
B01_PLACA8_03			A	A	492	501	10	10	0
B02_PLACA2_04			T	A	535	545	11	11	0
B05_PLACA3_03			TCC	AGG	241	250	10	10	0
B07_PL6_03			AAG	AAG	199	209	11	11	0

Copyright:
DI Dr. Robert Kofler
Vienna
robert@kofler.or.at

Figura 1 - Parte da saída do software *Sciroko* utilizado na busca de microssatélites nos 672 transcritos da biblioteca de expressão *forward*

> PeA2P4

GGTAAGAGAGGACGGATAACTCCCCTTCAGAGTTCAGGCAGCCAATATATTCCTTGGAGGCCTTAATAAATAA
 AGGCAACAAGTCATATATGCCATGGTCACATAAGTTGATTCTACGTCTATACATATATATATATATATCTA
 GTAGATCTACGAGTATTGACTCTCCAGGCGAGATCATGGGTCTCAACAGTTACAAAGATAACAAAATTAATGG
 GATTCAGGTTTCAGGCCATAATCACCCTTATATCTCTGCTACATTCTTAGATGCAAGAAAAGTTTATCCCCTC
 CACCAACTCTGAGCTGCACCTACATTACTTCTTCCCTCTTTCCCGTCACGTAGTCATCCATGGATAATTGCTTC
 CGTTTCTTCTGAAGAGTGGGACTCTCAGTTTGCCTTCCCGTTCCACGCTTGGATGGTGTCAAACCTCTTGGCAA
 GCTTCTTTTCTTTACGAGCCTGTCTTTCCCTTTCATCATACTCTTGGTTCTCTTTCTCC

Motivo: (AT)₁₀ - Classificação: Imperfeito

>PeC3P4

ATGGGGGTAGCAACATTGGCATTGTTGGTTTCGAATACTAGCTAGTGAGAAATATTACGAGAAATTCACTTCAAA
 TAGAGATGGCAACAAACGGAGTTGTTGGTGGCATTGAGGTCAATCCCATCACCCATGAAATTGAGGAGCTTGC
 TCGATTTGCTGTGGAGAATCAGAAAACGTATCCCAACCTGGTGTTTCAAGAAAGGTGGTGAGCGCAAGGCAACAG
 GTGGTATCAGGCATACTCTACTACATCACCTGGAGGCAGCGGAGGGTGAACATCAATTCCTATGAGACCA
 AAGTCTTGGTGAAGGAATGGCTTCACTACAAGGAAGTTGTTGAATTCAAGCGTCTTCTAACATCTCCTCATCA
 GGAGCCCCAAGTCGACGGTAAGCTTGGTGGCATTAAAGTCATTCCCGACGCGCCCAACCATGAGATCATT
 GAACTGGCTCGTTTTGCCGTGGACGAGCACAACAAGAAAGAGGTTACGTATATATATATTTGCTTAATAATAA
 TAAACAGTATCTTGATTAATTAATTTGACTAATGCAGCATCTTAATTTTCGTCCACTAATTATCATTATTCTT
 GAGCAGAAATCACATCTGGT

Motivo: (TA)₅(N)₅(TAA)₄ - Classificação: Composto

>PeD9P3

ATGGGGAAGGAAGCACAAAGCCTTACATAAGAGAGAGAGAGAGAGAGAGAGAAACAGTGAGAACATAGAGAGA
 TGGAGTTTTCATTGATTAACATGGAGAAGCTCAATGGTGAGGAGAGAGCAGCCATCATGGAAAAGATCAAAGA
 TGCTTGCGAGAAGTGGGGCTTCTTTGAGGTGCTGAACCATGGCATACCCCATGACTTTTTGGACACCGTGGAG
 AGTATGACCAAGGCACACTACAAGAAATGCATGGAGCAAAGATTCAATGAATTGGTGGCGAGCAAAGCCCTGG
 AGGATGTCCAAACAGAACTCAAAGATATGGACTGGGAGAGCACCTTCTATTTACGCCATCTGCCTGAGTCAA
 TATTGCCGAGATCCCTGATCTGGATGATGA

Motivo: (AG)₂₄ - Classificação: Interrompido

>PeH4P6

TTAAAAGAGAGAAAAAAGGAAAAAGGGTTTGCTAGCAGCGGTGTGAAGAACTGAAAAGAGAAACCTATATATG
 ACACAATAGGAATAGGGTCAGCAGGAGGATATAAAAAATAAAAACAGTGGTTCGTTTCGCCTCCTCGAGACGCG
 ACCAAGAATGGAATAAACAAAAGGAATGGCTGGCCATAATATAATTATATATCATATCTATACACAAAGAGAAG
 AGAAGAGAAGAGAAGAGAAACCACTGCTCTACATGAAGTAGCTGAGAGATTTCTTCTTTCTGGATCCACCTTCT
 CCGTATAAAATCATTCCATTGGAGGAGTTTCGTTTCATATTTGTGGACTCCGATGACACACTCGCACACACCTGCT
 CGTGTGCAAATCTAAAATCCTCCATCTTCAGAGGACGGACGTCGGTGCTGGTATAAAGTGCTGGGGAAGGTCT
 GTTCTCTTCCAATGCCAGGGCCTTCTCCTTTTTTCTTCTTCTCAAATTTCTCGGATGGCGCAATGTGCTGCAG
 TCACGCACAGGTTTCTTAGATCGCTGCCAGAATAACCATCAGTCATACTAGC

Motivo: (AAGAG)₅ - Classificação: Perfeito

Figura 2 - Identificação e classificação dos motivos SSRs nas sequências dos transcritos PeA2P4 (Putative chromatin remodelling complex ATPase chain ISWI), PeC3P4 (Cysteine proteinase inhibitor 3), PeD9P3 (1-aminocyclopropane-1-carboxylate oxidase 4) e PeH4P6 (Peroxisomal biogenesis factor 6-like, partial)

Dos 115 microssatélites encontrados nas sequências, 18 foram mononucleotídicos (15,70%), 32 dinucleotídicos (27,82%), 47 trinucleotídicos (40,86%), 11 tetranucleotídicos (9,56%), 4 pentanucleotídicos (3,47%) e 3 hexanucleotídicos (2,60%). Dentre as classes de microssatélites, os perfeitos representaram 57,09%, seguidos dos interrompidos (27,19%), dos imperfeitos (13,15%) e dos compostos (1,76%) (Figura 3).

Figura 3 - Distribuição dos microssatélites quanto ao motivo [mononucleotídeos (1), dinucleotídeos (2), trinucleotídeos (3), tetranucleotídeos (4), pentanucleotídeos (5) e hexanucleotídeos (6)] e a classe identificada

Quanto à composição nucleotídica dos motivos (Tabela 5), verificou-se que, dentre os dinucleotídicos, o motivo AG representou 62,5% do total (20/32), ficando os demais motivos AC e AT representados por 15,6% (5/32) e 21,9% (7/32), nesta ordem. Já para os trinucleotídeos, repetições de maior ocorrência na biblioteca, observou-se que os motivos AAG, ACC e AGC foram os mais frequentes, com 21,27% (10/47), 19,14% (9/47) e 17,02% (8/47) do total, respectivamente. Dos 91 transcritos contendo SSRs, 42 foram selecionados para o desenho de *primers* (Tabela 6).

Tabela 5 - Tipo de repetição, motivo e número encontrado entre os 115 microssatélites distribuídos nas 91 sequências expressas de *Passiflora edulis*

Tipo de repetição	Motivo*	Número
Mononucleotídeos	A	18
Dinucleotídeos	AC	5
	AT	7
	AG	20
Trinucleotídeos	ACC	9
	AAG	10
	AAC	2
	ATC	6
	AGC	8
	CCG	2
	AAT	2
	AGG	6
	ACG	2
Tetranucleotídeos	TTTC	1
	TTTG	1
	ATTA	1
	GTTT	1
	CATA	1
	CCTT	1
	AACC	1
	TATC	1
	ATGG	1
	AATT	1
	AGAT	1
Pentanucleotídeos	AAAAG	1
	GAGAG	1
	TAAAA	1
	AAGAG	1
Hexanucleotídeos	GGCGTC	1
	AGTTTC	1
	CCACCT	1
Total	-	115

Nota: * Os motivos mononucleotídicos incluem os complementos, e os di- e trinucleotídicos, incluem os complementos e as permutações das bases.

Tabela 6 - Lista dos 42 transcritos contendo SSRs selecionados da biblioteca de expressão *forward* para o desenho de *primers* e proteína respectiva codificada

(continua)

Nome da sequência	Proteína codificada
PeA1P4	Phosphoglycolate phosphatase-like
PeA1P7	Probable plastid-lipid-associated protein 4, chloroplastic-like
PeA10P10	Putative casein kinase
PeA10P9	Phosphoglycolate phosphatase 1B, chloroplastic-like
PeB1P1	Photosystem II reaction center PSB28 protein, chloroplastic
PeB1P5	Phosphatidylinositol kinase
PeB11P9	Soluble inorganic pyrophosphatase 1, chloroplastic
PeC3P4	Cysteine proteinase inhibitor 3
PeC9P4	Nudix hydrolase 27, chloroplastic
PeContig46	Cytochrome P450 76C2
PeContig48	Basic endochitinase B
PeContig49	E3 ubiquitin-protein ligase UPL6
PeContig52	31 kDa ribonucleoprotein, chloroplastic
PeContig57	Glycine dehydrogenase [decarboxylating] 2, mitochondrial
PeContig59	Replication factor C subunit 1-like
PeContig64	ATP-dependent Clp protease proteolytic subunit-related protein 4, chloroplastic
PeD3P4	Histone deacetylase 19
PeD4P5	40S ribosomal protein S13-2
PeD4P6	Guanylyl cyclase-related family protein
PeD8P6	Aldehyde dehydrogenase family 3 member I1, chloroplastic
PeD9P3	1-aminocyclopropane-1-carboxylate oxidase 4
PeE1P7	Dihydrolipoamide S-acetyltransferase, putative
PeE5P4	Disease resistance protein
PeE5P6	Transcription factor PIF7
PeE5P8	Serine/threonine-protein kinase BRI1-like 2 = receptor de brassinosteróide
PeE10P7	Long chain acyl-CoA synthetase 2
PeF1P8	Syntaxin-32
PeF6P8	40S ribosomal protein S8-1

Tabela 6 - Lista dos 42 transcritos contendo SSRs selecionados da biblioteca de expressão *forward* para o desenho de *primers* e proteína respectiva codificada

(continuação)

Nome da sequência	Proteína codificada
PeF8P2	Telomere repeat-binding protein 5
PeF12P4	GATA transcription factor 8
PeF12P5	36kDa-peroxisomal membrane protein (PMP36)
PeG2P10	Putative uncharacterized protein A
PeG2P7	Triosephosphate isomerase, cytosolic
PeG5P5	Glycogenin-1 isoform 2
PeG6P3	Hypothetical protein
PeG9P9	Leucine aminopeptidase 1
PeH3P10	Guanine deaminase-like, partial
PeH4P6	Peroxisomal biogenesis factor 6-like, partial
PeH5P4	Putative beta-carotene hydroxylase
PeH6P6	29 kDa ribonucleoprotein, chloroplastic
PeH8P3	Homocysteine S-methyltransferase 3
PeH11P6	Phospholipid hydroperoxide glutathione peroxidase 1, chloroplastic

A seguir, é apresentado o resultado da análise *in silico* dos transcritos selecionados visando desenhar as sequências dos pares de *primers* com o auxílio do programa *web PRIMER3* (ROZEN; SKALETSKY, 2000) e amplificar os 42 locos de microssatélites (Tabela 7; Figura 4).

Tabela 7 - Resultado da análise *in silico* para 42 regiões contendo microssatélites: sequências dos pares de *primers* (5'-3'), motivo, classe e tamanho esperado do respectivo alelo (AE, em pb)

(continua)

Código	<i>Primers forward</i> (linha superior) e <i>reverse</i> (linha inferior)	Motivo	Classe	AE
PeA1P4	TGGTGCCTCTGTACATTC CCCCACACTTCATCCTTG	(CA) ₈	Perfeito	298
PeA1P7	GGGACCACAAGCGTTATC CTTAGCATCCTTGCCCTTG	(TTTC) ₃	Perfeito	234
PeA10P10	TGGAATGTGATTTGCATGG TAGGTATTGCTGGCGAAG	(GGT) ₄	Perfeito	219
PeA10P9	AATGCTTGATAACACATGCTC GGCTGCAAACTCTTCTTG	(GA) ₅	Perfeito	154
PeB1P1	ATGTCTCCTCGGAAATGC CACCACCTCCCTTTTGTG	(AGTTTC) ₄	Perfeito	155
PeB1P5	GCGATGAGTGTGCTAGAGG TTCAGACAGCCGAGGAAG	(AGA) ₅	Perfeito	183

Tabela 7 - Resultado da análise *in silico* para 42 regiões contendo microssatélites: sequências dos pares de *primers* (5'-3'), motivo, classe e tamanho esperado do respectivo alelo (AE, em pb)

Código	<i>Primers forward</i> (linha superior) <i>e reverse</i> (linha inferior)	Motivo	(continuação)	
			Classe	AE
PeB11P9	CCGAACAGCAAACCTCTG CCGTAGCAACCTCCATCT	(TTC) ₇	Perfeito	283
PeC3P4	TTGAACTGGCTCGTTTTG GGACGAAAATTAAGATGCTG	(TA) ₅ (N) ₅ (TAA) ₄	Composto	129
PeC9P4	GTCCGCTCCTTATTGTCC AGTTCACGAGGCAAATCC	(AGG) ₆	Perfeito	227
PeContig46	GATCAATCGTCATCCTCCT TCTCGATGCACAGGACAG	(CA) ₉	Perfeito	227
PeContig48	GGGGACACAGAGACACATC ATGCAATCAATGGGGAAG	(CTC) ₇	Perfeito	168
PeContig49	AAGCCAAAGTCGGTGTG TAGCCCAGACAAGGAAGG	(GAA) ₆	Perfeito	186
PeContig52	CTTGCCTCTCTCCCTCTC GGGTTTTGGGTTTGACAG	(CT) ₂₆	Perfeito	298
PeContig57	AAGATCGGCGGTAACAAC GGAATCGAGCTTGTCTGG	(CAG) ₄	Perfeito	290
PeContig59	CAGAAGGAAAACCAGCAAG CCCCATCATCATCACTTTC	(TGA) ₇	Perfeito	298
PeContig64	TCCCCCATTTTCTCTTCG CGATTCCCGAACAAGTAAG	(TC) ₅	Perfeito	259
PeD3P4	GGCTCCACCTCCTATTCC CGAAGTGTTGCATCTTTCC	(CCT) ₉	Perfeito	216
PeD4P5	CCTCAAATCGGGAACCTCC AAGCCTTTCTTCGCAAC	(TTC) ₅	Perfeito	243
PeD4P6	CATTTTCCCAACACCAGC CTTAATCCACGGACTCGG	(AACC) ₃	Perfeito	150
PeD8P6	CTTTGACGGTCGCATTG GGAAGCATACCCCACATC	(ATGG) ₃	Perfeito	299
PeD9P3	AGGAAGCACAAGCCTTACA ATGGTTTCAGCACCTCAAAG	(AG) ₂₄	Interrompi do	179
PeE1P7	CTCAGCTCCACCATGACTG TTCCTCGACCATGATTGC	(TGT) ₄	Perfeito	164
PeE5P4	CAAATGACTCTGGCTTCTTC TGATCCAAGGGATAGAGCTG	(TCA) ₄	Perfeito	219
PeE5P6	CAGCCTCCAACCTCTCCAAC CATGTTAGCTCTGCAACTCC	(ATC) ₂₃	Perfeito	263
PeE5P8	GGAAGCAAACACCAAAATC TCGTGTCATCGAACCTC	(TCG) ₇	Perfeito	245
PeE10P7	TTAATGCCACAGCCCAAC TGACCCAAAATCAAACACC	(AC) ₇	Perfeito	235
PeF1P8	ATGCTTCACAGGCTCCTC TCAGCTCTGCTCTGCATG	(GCA) ₁₀	Interrompi do	169
PeF6P8	TGACATTGGTCGCAAGAAG GACCACCACCAAATTGCTC	(CTG) ₅	Perfeito	176
PeF8P2	GGTTTTAGGGGGTTTTGC CAAGAAGGGCACATCACAC	(AG) ₈	Perfeito	249
PeF12P4	GAAAACACCGAGTCGAAAG AACACCATGATTGGACAGC	(GTG) ₇	Perfeito	271
PeF12P5	GGAGTCAGAGATGGTTTCG AGTGATGAGCTGAGCGATG	(TC) ₁₃	Interrompi do	230
PeG2P10	TCTTCATCACCGCAAAAAG GGGGTCTGGAGAGAGAAAG	(GAA) ₄	Perfeito	300
PeG2P7	CCCACAGATCCATCAAACC TGGTTCCATTGCATTTCC	(CAC) ₉	Perfeito	217

Tabela 7 - Resultado da análise *in silico* para 42 regiões contendo microssatélites: sequências dos pares de *primers* (5'-3'), motivo, classe e tamanho esperado do respectivo alelo (AE, em pb)

Código	<i>Primers forward</i> (linha superior) e <i>reverse</i> (linha inferior)	Motivo	(conclusão)	
			Classe	AE
PeG5P5	GACGGGGGCAAACTTAC GACATCCATCATAGCGACC	(AT) ₅	Perfeito	273
PeG6P3	TGTTCAAGCCCATCTTCG GCTCTCGCTCTTGTCGTAG	(GTG) ₄	Perfeito	362
PeG9P9	CTTGTTGAGTGGGTGTTGC CACTGCTTACAGCCCCAG	(AGAT) ₃	Perfeito	156
PeH3P10	CAATGCCAATTCAGCTCAG AACATCTCACATGCCCTTG	(TAAAA) ₃	Perfeito	252
PeH4P6	GAATAGGGTCAGCAGGAGG GAGTGTGTCATCGGAGTCC	(AAGAG) ₅	Perfeito	270
PeH5P4	TAACCCCGTAATCCTCTCC CATGGGCAGTCTGTTTCG	(CT) ₁₀	Interrompido	261
PeH6P6	CTTTCAATCCAAGCCATG TTCTGATGAGCCGAGAGG	(CTT) ₆	Perfeito	217
PeH8P3	AATAGCGAGAGGCATTGG ATTTGGCACTCCAGAGAGG	(CT) ₅	Perfeito	234
PeH11P6	AAACTGCCTGCTTGGAAG ACCCAGAAAGACCAGACG	(TCT) ₄	Perfeito	265

06' de *P. edulis*. A qualidade da amplificação foi avaliada em relação à amplificação de fragmentos únicos ou múltiplos (Tabela 8, Figura 5).

Tabela 8 - Resultado das amplificações com os 42 pares de *primers* de microssatélites em *Passiflora alata* em gel de agarose

(continua)

Código	Resultado da amplificação
PeA1P4	Fragmento único
PeA1P7	Fragmento único
PeA10P10	Fragmento único
PeA10P9	Fragmento único
PeB1P1	Fragmento único
PeB1P5	Fragmento único
PeB11P9	Fragmento único
PeC3P4	Fragmento único
PeC9P4	Fragmentos múltiplos
PeContig46	Fragmentos múltiplos
PeContig48	Fragmento único
PeContig49	Fragmentos múltiplos
PeContig52	Fragmentos múltiplos
PeContig57	Fragmentos múltiplos
PeContig59	Fragmento único
PeContig64	Fragmento único
PeD3P4	Fragmentos múltiplos
PeD4P5	Fragmentos múltiplos
PeD4P6	Fragmentos múltiplos
PeD8P6	Fragmentos múltiplos
PeD9P3	Fragmentos múltiplos
PeE1P7	Fragmento único
PeE5P4	NA
PeE5P6	Fragmentos múltiplos
PeE5P8	Fragmentos múltiplos
PeE10P7	Fragmento único
PeF1P8	Fragmentos múltiplos
PeF6P8	Fragmentos múltiplos

Tabela 8 - Resultado das amplificações com os 42 pares de *primers* de microssatélites em *Passiflora alata* em gel de agarose

(conclusão)	
Código	Resultado da amplificação
PeF8P2	Fragmento único
PeF12P4	Fragmentos múltiplos
PeF12P5	Fragmentos múltiplos
PeG2P10	Fragmento único
PeG2P7	Fragmentos múltiplos
PeG5P5	Fragmento único
PeG6P3	Fragmentos múltiplos
PeG9P9	NA
PeH3P10	Fragmentos múltiplos
PeH4P6	Fragmentos múltiplos
PeH5P4	Fragmentos múltiplos
PeH6P6	Fragmento único
PeH8P3	Fragmento único
PeH11P6	Fragmentos múltiplos

Notas:

NA - Não houve amplificação ou não apareceu a banda na revelação do gel

Figura 5 - Eletroforese em gel de agarose (1,8%) corado com Sybr Green mostrando o padrão de amplificação de 14 microssatélites. Canaletas 1 a 5 (para cada *primer*): 'IAPAR-06', 'IAPAR-123', '2(12)', 'SV3' e Controle negativo, sem DNA *template*. L é o ladder 100 pb

Dos 42 pares de *primers* avaliados, 40 resultaram em algum tipo de amplificação. Desses, 18 mostraram amplificação de um único fragmento e 22 de fragmentos múltiplos, fazendo-se necessária a otimização das reações. A otimização foi realizada testando diferentes concentrações dos reagentes, tal como descrito na tabela 1, e diferentes temperaturas de anelamento. Desses 22 *primers*, foi possível a otimização da reação para 16. Assim, um bom padrão de amplificação foi obtido para um total de 34 *primers*. Portanto, observou-se ~83% (34/41) de transferibilidade entre as espécies, ou seja, dos *primers* desenhados com base em sequências de *P. edulis* a maioria resultou em amplificações em *P. alata*.

Esses 34 *primers*, então, foram avaliados quanto ao polimorfismo de alelos entre acessos de maracujá em gel de poliacrilamida desnaturante.

4.1.3 Análise do polimorfismo em gel de poliacrilamida desnaturante

Para verificar a qualidade da amplificação foram observados dois aspectos após a eletroforese em géis de poliacrilamida desnaturante: número de locos originados, uma vez que um par de *primers* pode amplificar locos únicos ou múltiplos; a presença de bandas *stutter* ou “fantasmas”. Já em relação ao polimorfismo, os locos foram caracterizados como sendo monomórficos ou polimórficos, a depender da existência de alelos distintos entre os genitores ou nos representantes da F₁ utilizados. Adicionalmente, foram caracterizados os tamanhos observados dos alelos amplificados por cada *primer* de SSR (Tabela 9, Figura 6).

Tabela 9 - Resultados das amplificações dos 34 pares de *primers* de microssatélites de *Passiflora alata* em gel de poliacrilamida desnaturante: qualidade da amplificação (nº de locos e nitidez das bandas), presença de bandas *stutter*, análise genética (locos P, polimórficos, e M, monomórficos) e tamanhos observados dos alelos (AO, em pb)
(continua)

Código	Nº de locos	<i>Stutter</i>	Nitidez	Rastro	Análise genética	AO (*)
PeA1P4	Único	Não	Nítidas	Sim	M	>330*
PeA1P7	Múltiplos	Não	Nítidas	Sim	M	220/235
PeA10P10*	Múltiplos	Não	Nítidas	Sim	P	120
PeA10P9	Múltiplos	Não	Nítidas	Sim	M	255
PeB1P1	Único	Não	Nítidas	Sim	M	130
PeB1P5	Múltiplos	Sim	Nítidas	Sim	P	155-165/195-205
PeB11P9	Único	Sim	Nítidas	Sim	M	>330*
PeC3P4	Múltiplos	Sim	Nítidas	Sim	M	120/142
PeC9P4	Único	Sim	Nítidas	Não	M	240*
PeContig46	Único	Não	Nítidas	Não	M	>330*
PeContig48	Múltiplos	Não	Nítidas	Não	M	132/162
PeContig49*	Múltiplos	Sim	Nítidas	Não	P	250
PeContig52*	Múltiplos	Sim	Nítidas	Não	P	250
PeContig59*	Múltiplos	Sim	Rastro	Sim	P	250-260
PeContig64*	Múltiplos	Sim	Nítidas	Sim	P	>330
PeD3P4	Múltiplos	Sim	Nítidas	Não	M	>330

Tabela 9 - Resultados das amplificações dos 34 pares de *primers* de microssatélites de *Passiflora alata* em gel de poliacrilamida desnaturante: qualidade da amplificação (nº de locos e nitidez das bandas), presença de bandas *stutter*, análise genética (locos P, polimórficos, e M, monomórficos) e tamanhos observados dos alelos (AO, em pb)
(continuação)

Código	Nº de locos	<i>Stutter</i>	Nitidez	Rastro	Análise genética	AO (*)
PeD4P5	Múltiplos	Sim	Nítidas	Sim	M	>330
PeD4P6	Múltiplos	Sim	Nítidas	Sim	M	>330
PeD8P6	Múltiplos	Sim	Nítidas	Sim	M	300/>330
PeD9P3	Múltiplos	Sim	Nítidas	Não	M	>330
PeE1P7	Único	Não	Nítidas	Sim	M	164
PeE5P8*	Múltiplos	Sim	Nítidas	Sim	P	162/248-252/230
PeE10P7	Múltiplos	Não	Nítidas	Sim	P	242-244
PeF1P8	Múltiplos	Sim	Nítidas	Sim	M	>330
PeF6P8	Múltiplos	Sim	Nítidas	Sim	M	>330
PeF8P2	Único	Sim	Nítidas	Sim	M	>330*
PeF12P4	Múltiplos	Sim	Nítidas	Sim	M	130/270
PeG2P10	Múltiplos	Não	Nítidas	Sim	M	310/312
PeG5P5	Múltiplos	Não	Nítidas	Sim	M	273/>330
PeG6P3*	Múltiplos	Sim	Nítidas	Sim	P	150-160
PeH3P10	Múltiplos	Não	Nítidas	Não	M	128/250/252
PeH4P6*	Múltiplos	Sim	Nítidas	Sim	P	265/266/276
PeH8P3	Múltiplos	Não	Nítidas	Não	M	255/230
PeH11P6	Múltiplos	Não	Nítidas	Não	M	205/270/273/290

Notas:

No código:

* Indica *primers* que apresentaram amplificações inespecíficas, mas observou-se polimorfismo

No alelo observado (AO):

* Tamanho dos alelos observados diferiram do tamanho esperado

/ Separa os alelos observado por tamanho

- Separa os alelos dentro de um intervalo de tamanho

Figura 6 - Eletroforese em gel de poliacrilamida desnaturante mostrando o padrão de amplificação de 9 locos de microssatélites. Canaletas 1 a 10 (para cada *primer*): acessos 'IAPAR-06', 'IAPAR-123', '2(12)' em duplicata, 'SV3' em duplicata e 4 indivíduos da F₁. L é o ladder 10 pb

Nos géis de poliacrilamida desnaturante foi possível verificar que a maioria dos pares de *primers* (67,5%) levou à amplificação de inúmeras bandas, além daquelas esperadas na ocasião do desenho dos *primers*. Quando utilizaram-se os *primers* PeA1P4, PeB11P9, PeC9P4, PeContig46 e PeF8P2, que revelaram um único *amplicon*, observou-se que o tamanho observado dos alelos foi maior que o esperado, por conta da provável presença de íntros entre os éxons, já que os *primers* foram desenhados a partir de sequências de cDNA. Para os *primers* que revelaram fragmentos múltiplos não foi possível essa análise, pela inviabilidade de se detectar o loco esperado.

Cada SSR pôde ser analisado quanto ao polimorfismo. Os 10 SSRs que revelaram polimorfismo foram reamplificados em parte da população F₁ (Figura 7), para se observar a segregação e daí classifica-los segundo o modelo proposto por Wu et al. (2002) para

populações F_1 oriundas de cruzamentos bi-parentais. Analisando a tabela 3 e a figura 7, realizou-se a classificação quanto ao tipo de configuração genética (Tabela 10), embora as segregações tenham sido observadas (e não testadas) em uma pequena amostra.

Figura 7 - Polimorfismo e segregação dos locos polimórficos. Nas canaletas 1 e 2 estão os acessos '2(12)' e 'SV3', genitores da população F₁, e nas demais canaletas está a progênie. As setas indicam os locos polimórficos para os marcadores que segregam 1:1

PeContig64

PeE5P8

PeE10P7

PeG06P3

PeH04P6

Figura 7 - Polimorfismo e segregação dos locos polimórficos. Nas canaletas 1 e 2 estão os acessos '2(12)' e 'SV3', genitores da população F₁, e nas demais canaletas está a progênie. As setas indicam os locos polimórficos para os marcadores que segregam 1:1

Tabela 10 - Classificação dos locos SSR quanto ao tipo de configuração genética, segundo a terminologia de Wu et al. (2002)

Marcador	Motivo microssatélite	Locos revelados	Configuração genética do loco	Segregação esperada na população F ₁	Nº de alelos
PeA10P10	(GGT) ₄	1	D2.18	1:1	2
PeB1P5	(AGA) ₅	1	A2	1:1:1:1	4
PeContig49	(GAA) ₆	1	D2.18	1:1	2
		2	D2.18	1:1	2
PeContig52	(CT) ₂₆	1	A4	1:1:1:1	4
PeContig59	(TGA) ₇	1	A4	1:1:1:1	4
PeE10P7	(AC) ₇	1	D2.18	1:1	2
PeE5P8	(TCG) ₇	1	A4	1:1:1:1	4
PeG6P3	(GTG) ₄	1	D1.13	1:1	2
		2	D1.13	1:1	2
		3	D2.18	1:1	2
		4	D2.18	1:1	2
PeH4P6	(AAGAG) ₅	1	D2.18	1:1	2
		2	D1.13	1:1	2

Notas: * Baseado na classificação de Wu et al. (2002)

Observa-se que os marcadores PeB1P5, PeContig52, PeContig59 e PeE5P8 apresentam a configuração A, o tipo mais informativo de cruzamento, com segregação 1:1:1:1, sendo A2 para o marcador PeB1P5 e A4 para os marcadores PeContig52, PeContig59 e PeE5P8. Nesses marcadores, os quatro alelos estão segregando já que os genótipos dos genitores são heterozigóticos para alelos distintos (*ab* x *cd*) em um único loco, sendo que na configuração A4 há um alelo nulo (*ao* x *bo*).

Os marcadores PeA10P10, PeE10P7, PeContig64, PeContig49, PeH04P6 e PeG06P4 apresentam a configuração do tipo D (Tabela 10) a qual representa locos que estão na configuração de cruzamento-teste e portanto segregam na proporção 1:1, sendo incluídos em dois grupos: D1, quando um dos pais é heterozigótico e o outro é homozigótico e D2, o recíproco de D1.

O padrão de amplificação de alguns locos microssatélites estudados neste trabalho mostra que alguns alelos específicos são acompanhadas por bandas *stutter* no gel, como é o caso do SSR PeContig52 (Figura 7). Adicionalmente, nos locos do tipo A4 (Tabela 10; Figura 7), observou-se a presença de alelos nulos.

4.2 Marcadores SNP

Dos 118 *primers* desenhados a partir das sequências de transcritos selecionados das bibliotecas de expressão, 97 (82,2%) revelaram algum padrão de amplificação.

Observou-se que o uso do programa computacional GENNE RUNNER (SPRUYT; BUQUICCHIO, 1994) foi bastante adequado para o desenho e a análise das sequências dos *primers*.

Dos *amplicons* obtidos para os 97 transcritos, a maioria (67%, 65/97) apresentou amplificação inespecífica. Esses *primers*, tal como o PeP3RF8, foram submetidos a testes de otimização, com aumento da temperatura de anelamento para obtenção de um único *amplicon*, como ocorreu quando se analisou o padrão revelado pelo par de *primers* PeP3RC2 (Figura 8). Assim, foi possível a otimização da reação para 74 desses, com a obtenção de um *amplicon* único. Como esperado, observou-se alta transferibilidade (~63%, 74/97) dos pares de *primers* de uma espécie para outra. Adicionalmente, 43 pares de *primers* revelaram *amplicons* de tamanho maior do que o esperado.

Figura 8 - Eletroforese em gel de agarose (1,8%) corado com Sybr Green mostrando o padrão de amplificação de 6 fragmentos gênicos. Canaletas 1 a 5 (para cada par de *primers*): acessos 'IAPAR-06', 'IAPAR-123', '2(12)', 'SV3' e controle negativo. L é o ladder 100 pb

Após as PCRs e a obtenção dos *amplicons*, os respectivos fragmentos gênicos foram purificados e sequenciados diretamente. Foi possível o sequenciamento de 37 deles. Foram obtidas sequências dos acessos '2(12)' e 'SV3' para cada fragmento gênico, permitindo a análise quanto à presença de SNPs nas sequências desses acessos. Para os fragmentos gênicos onde detectou-se um potencial SNP, foram obtidas, adicionalmente, sequências dos indivíduos F₁-67 e F₁-100, da população filial, para a confirmação do SNP. Optou-se pelo método de sequenciamento direto (sem clonagem) dos fragmentos por permitir a identificação de heterozigotos na mesma sequência.

O sequenciamento dos fragmentos amplificados a partir da análise inicial dos 37 transcritos resultou em:

- (i) presença de no mínimo um SNP nas sequências dos genitores ‘2(12)’ e ‘SV3’, e nos indivíduos F₁- 67 e F₁-100 em 16 fragmentos gênicos (14 transcritos resultaram em um fragmento gênico e um transcrito resultou em dois fragmentos gênicos);
- (ii) ausência de SNPs nas sequências de 13 fragmentos gênicos;
- (iii) sequências de baixa qualidade inviabilizando a análise e detecção de SNPs em 9 fragmentos gênicos.

A tabela 11 sintetiza os resultados da análise das sequências geradas a partir do desenho de *primers* tomando por base os transcritos de *P. edulis*. Portanto, os fragmentos aqui gerados são correspondentes a genes putativos de *P. alata*.

Tabela 11 - Resultado da análise das sequências dos 37 genes putativos de *Passiflora alata* selecionados

(continua)	
Gene putativo	Resultado do sequenciamento
PaA4P8	Sequência de baixa qualidade
PaA5P9	Ausência de SNP
PaA6P9	Presença de SNP
PaA12P4	Sequência de baixa qualidade
PaA12P5	Presença de SNP
PaB8P9	Presença de SNP
PaC11P10	Presença de SNP
PaContig13R	Ausência de SNP
PaContig15	Presença de SNP
PaContig17	Ausência de SNP
PaContig21	Presença de SNP
PaContig22R	Presença de SNP
PaContig39	Sequência de baixa qualidade
PaContig45	Ausência de SNP
PaContig52	Sequência de baixa qualidade
PaContig53	Presença de SNP
PaD1P8	Ausência de SNP
PaD5P4	Sequência de baixa qualidade
PaD5P10	Presença de SNP
PaE1P7	Presença de SNP
PaE12P7	Sequência de baixa qualidade
PaF1P8	Presença de SNP
PaF1P9	Ausência de SNP
PaG2P6	Ausência de SNP
PaG6P4	Presença de SNP
PaG10P5	Ausência de SNP
PaG12P5	Presença de SNP
PaH3P9	Ausência de SNP

Tabela 11 - Resultado da análise das sequências dos 37 genes putativos de *Passiflora alata* selecionados

(continuação)	
Gene putativo	Resultado do sequenciamento
PaH5P8	Sequência de baixa qualidade
PaH8P3	Sequência de baixa qualidade
PaH10P4	Ausência de SNP
PaH12P9	Presença de SNP
PaP1RA5	Ausência de SNP
PaP3RA2	Presença de SNP
PaP3RC9	Ausência de SNP
PaP4RA10	Ausência de SNP
PaP4RB1	Sequência de baixa qualidade

As sequências dos 13 fragmentos gênicos nas quais não foi detectado polimorfismo entre os genitores foram descartadas.

Para os 16 fragmentos gênicos putativos de *P. alata* nos quais foram detectados SNPs, os *amplicons* dos 4 acessos foram resequenciados para confirmação desses SNPs. Como o método de sequenciamento utilizado gera sequências de até 700 pb, para os *amplicons* com tamanho de no mínimo 1.000 pb, foram obtidas sequências nos dois sentidos, *forward* e *reverse*, no intuito de obter a sequência completa. No total, após a retirada das duplicatas, foram selecionadas as 85 sequências de melhor qualidade, com representação dos genitores e dos indivíduos da progênie para cada um dos 15 genes.

Ao submeter essas 85 sequências ao ‘*Clip ends*’, no *software* CODONCODE ALIGNER foram excluídas 16 e 115 bases das extremidades 5’ e 3’, respectivamente, do conjunto total de sequências. Essas 85 sequências foram alinhadas por meio da função ‘*Assemble*’. Os *contigs* foram obtidos para os 16 fragmentos gênicos putativos que apresentaram polimorfismo. Esses *contigs* variaram de 332 a 872 pb, com tamanho médio de 625 pb. Para os 16 fragmentos gênicos, foram obtidas leituras a partir da extremidade 5’, resultante do sequenciamento *forward*, e para um gene, adicionalmente, foi obtida uma leitura a partir da extremidade 3’, resultante do sequenciamento *reverse*; assim, adicionou-se nos códigos a letra F ou R para representar a direção do sequenciamento. Foram identificados 34 SNPs nos 16 fragmentos gênicos putativos. Na figura 8, são apresentadas as sequências e os eletroferogramas correspondentes aos genitores ‘2(12)’, ‘SV3’, F₁-67 e dos indivíduos F₁-100 relativos ao fragmento gênico PaD5P10 (*Major facilitator superfamily protein*). De acordo

com a nomenclatura da *International Union of Pure and Applied Chemistry*, outras letras, além de A, C, T e G, podem aparecer nas sequências. Essas representam ambiguidades, ou SNPs bialélicos, como se segue: R (G/A); Y (T/C); K (G/T); M (A/C), S (G/C) e W (A/T).

As sequências dos demais genes se encontram nos anexos desta dissertação. Adicionalmente, esses genes foram validados por algoritmo de alinhamento local BLASTN, disponível no site do NCBI (<http://www.ncbi.nlm.nih.gov>), contra a sequência de cDNA de origem, sendo possível a identificação de éxons e íntrons, bem como a classificação dos SNPs de acordo com sua posição no gene (Tabela 12, Figura 9).

Figura 9 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', e dos indivíduos da geração filial F₁-67 e F₁-100 relativos ao fragmento gênico PaD5P10 de *P. alata*: há um SNP na posição 350

Tabela 12 - SNPs putativos e respectiva localização, se intrônico (I) ou exônico (E), genótipo dos genitores e de dois indivíduos da F₁ de *Passiflora alata*

Gene	Posições no <i>contig</i>	Localização do SNP	Genótipos dos genitores		Genótipos F ₁	
			'2(12)'	'SV3'	F ₁ - 67	F ₁ - 100
PaA6P9	532	I	G/C	G/G	G/C	C/C
PaA12P5_F	127	E	A/T	T/T	A/T	T/T
	235	I	T/C	T/C	T/C	T/C
PaA12P5_R	160	I	G/T	G/T	G/T	G/T
	169	I	G/A	G/G	G/A	G/G
PaB8P9	333	I	A/A	A/C	A/A	A/A
	342	I	T/T	A/T	T/T	T/T
	344	I	C/C	A/C	C/C	C/C
	347	I	T/T	G/T	T/T	T/T
PaC11P10	44	E	??	C/T	C/T	C/T
PaContig15	158	E	G/C	G/G	G/C	G/C
PaContig21	309	E	G/T	T/T	T/T	G/T
	323	E	C/C	G/C	G/C	G/C
PaContig22R	108	E	G/C	C/C	G/C	??
	336	I	G/A	G/G	G/A	??
	364	I	A/C	A/A	A/C	??
PaContig53	104	E	T/C	T/C	T/C	C/C
	303	E	C/C	T/C	C/C	C/C
	343	E	T/C	C/C	??	T/T
	347	E	T/C	C/C	??	C/C
	349	E	T/C	C/C	??	T/T
PaD5P10	181	E	A/G	A/A	G/G	A/A
	350	E	A/T	T/T	A/A	T/T
	371	E	A/G	A/A	G/G	A/A
	381	E	A/G	G/G	A/A	G/G
PaE1P7	212	E	G/C	C/C	G/C	G/C
PaF1P8	163	I	T/C	T/C	T/C	T/T
	246	I	T/G	T/G	T/G	T/T
	247	I	T/G	T/G	T/G	T/T
PaG6P4	287	I	G/C	C/C	C/C	C/C
PaG12P5	264	E	T/T	A/T	A/T	T/T
PaH12P9	286	I	G/A	G/A	G/A	G/A
PaP3RA2	186	I	T/C	C/C	??	T/C
	225	I	T/C	C/C	??	T/C

Nota: Os símbolos ?? representam genótipos não identificados.

Figura 10 - BLASTN das sequências dos *contigs* de *P. alata* revelados pela amplificação dos *primers* desenhados a partir de sequências de cDNA dos genes PeB8P9 (Protein linase 2B), PeP3RA2 (Lactoylglutathione lyase) e PeContig21 (30S Ribosomal protein 3-1) de *P. edulis*. As linhas horizontais (em cinza) representam os possíveis íntrons presentes do DNA genômico de *P. alata*

Observou-se, no presente trabalho, igual ocorrência de SNPs em regiões codantes e não-codantes, ou seja, 50% (17/34). Os SNPs encontrados em regiões codantes foram analisados quanto à tradução para aminoácidos. Dos 17 SNPs encontrados em éxons, 13 (76,4%), resultaram em mudança no aminoácido codificado, totalizando uma relação de 1,3:1 de mutação silenciosa versus não silenciosa. Adicionalmente, pela divisão dos aminoácidos em classes (VOET; VOET, 2013), cinco dessas mutações não silenciosas foram não conservativas, ou seja, a mudança em um nucleotídeo resultou na mudança de um aminoácido com diferentes características, podendo gerar uma mudança na função da proteína.

Do total de substituições encontradas, todas foram bialélicas, ou seja, havia apenas duas alternativas para cada posição. Verificou-se que das substituições encontradas, as transversões, que é a substituição de uma purina por uma pirimidina ou vice-versa, foram as mais frequentes (~53%, 18/34). Também, entre as substituições encontradas, observou-se que as transições do tipo C/T foram as mais frequentes (29,4%, 10/34) (Tabela 13, Figura 11).

Tabela 13 - Classificação dos SNPs identificados em cada gene putativo (ou *contig*) de *P. alata* quanto ao tipo de substituição

Gene	Nº de SNPs	Transição	Transversão
PaA6P9	1		1
PaA12P5_F	2	1	1
PaA12P5_R	2	1	1
PaB8P9	4		4
PaC11P10	1	1	
PaContig15	1		1
PaContig21	2	1	1
PaContig22R	3	1	2
PaContig53	5	5	
PaD5P10	4	3	1
PaE1P7	1		1
PaF1P8	3	1	2
PaG6P4	1		1
PaG12P5	1		1
PaH12P9	1	1	
PaP3RA2	2	2	
Total	34	16	18

Figura 11 - Número de polimorfismos encontrados nas sequências dos genes putativos (ou *contigs*) de *P. alata*. A/G e C/T correspondem a transições e os demais SNPs a transversões

Os *contigs* dos 16 fragmentos gênicos de *Passiflora alata* foram, adicionalmente, usados na busca de sequências similares na base de dados do NCBI por BLASTX para verificar a sua identidade. A cobertura dos fragmentos foi determinada tomando-se os acessos

com maiores *scores* e identidades, já que os maiores valores indicam os melhores alinhamentos. Após análise no BLASTX, uma função putativa pôde ser atribuída a todos os 16 fragmentos gênicos de *P. alata*. Dentre esses, 81,2% (13/16) mostraram homologia com as mesmas proteínas das sequências de origem e 18,8% tiveram similaridade com proteínas não nomeadas (Tabela 14).

Tabela 14 - Resumo das consultas via BLASTX dos *contigs* obtidos para os 16 fragmentos gênicos com presença de SNP para *Passiflora alata* em relação ao acesso com maiores *scores* de alinhamento

Gene/ <i>Contig</i>	Proteína similar e espécie vegetal	Cobertura	Identidade	<i>Scores</i>	<i>E-value</i>
PaA6P9	Linoleate 13S-lipoxygenase 2-1, chloroplastic [Vitis vinifera]	78%	68%	338	2e-105
PaA12P5_F	ABC transporter G family member 29-like [Vitis vinifera]	29%	92%	87.8	1e-09
PaA12P5_R	ABC transporter G family member 36 [Arabidopsis thaliana]	60%	61%	120	6e-28
PaB8P9	Kinase 2B family protein [Populus trichocarpa]	24%	92%	134	2e-18
PaC11P10	Putative serine/threonine-protein kinase [Morus notabilis]	71%	89%	147	3e-39
PaContig15	Hypothetical protein [Morus notabilis]	10%	87%	60,1	2e-07
PaContig21	30S ribosomal protein 3, chloroplastic-like [Cucumis sativus]	40%	88%	134	6e-36
PaContig22R	Disulfide isomerase-like protein isoform 1 [Theobroma cacao]	57%	75%	206	2e-26
PaContig53	Unnamed protein product [Vitis vinifera]	17%	80%	71,6	4e-12
PaD5P10	Major facilitator superfamily protein [Theobroma cacao]	80%	90%	303	7e-97
PaE1P7	Hypothetical protein [Amborella trichopoda]	36%	93%	114	5e-27
PaF1P8	Syntaxin-32 like [Citrus sinensis]	26%	93%	89,7	9e-18
PaG6P4	Glutamate-cysteine ligase, chloroplastic-like [Glycine max]	11%	96%	47.8	0,001
PaG12P5	Hypothetical protein [Populus trichocarpa]	62%	82%	174	4e-48
PaH12P9	Kinase family protein [Populus trichocarpa]	61%	98%	265	5e-29
PaP3RA2	Lactoylglutathione lyase, putative [Ricinus communis]	58%	77%	211	4e-34

Os tamanhos observados das sequências diferiram dos tamanhos esperados para todos os 16 fragmentos gênicos, com 13 genes com tamanho da sequência observado maior do que o esperado, e 3 deles com tamanho menor que o esperado.

Foi calculada também a frequência de SNP em cada gene. Adicionalmente, para os 16 fragmentos gênicos foi obtido um total de 10.003 pb, assim, a frequência de SNP foi estimada em 1 SNP a cada 294 pb, aproximadamente (Tabela 15).

Tabela 15 - Tamanho esperado e observado das sequências (em pb) dos 16 fragmentos gênicos de *P. alata* e ocorrência de SNPs em cada fragmento gênico estudado

Gene/ <i>Contig</i>	Tamanho esperado	Tamanho observado	Nº de SNPs	Ocorrência de SNPs
PaA6P9	541	872	1	1 SNP a cada 872 pb
PaA12P5_F	459	473	2	1 SNP a cada 237 pb
PaA12P5_R	459	533	2	1 SNP a cada 267 pb
PaB8P9	507	803	4	1 SNP a cada 201 pb
PaC11P10	505	332	1	1 SNP a cada 332 pb
PaContig15	677	857	1	1 SNP a cada 857 pb
PaContig21	468	506	2	1 SNP a cada 253 pb
PaContig22R	570	747	3	1 SNP a cada 249 pb
PaContig53	523	469	5	1 SNP a cada 94 pb
PaD5P10	566	614	4	1 SNP a cada 154 pb
PaE1P7	513	465	1	1 SNP a cada 465 pb
PaF1P8	545	768	3	1 SNP a cada 256 pb
PaG6P4	584	593	1	1 SNP a cada 593 pb
PaG12P5	448	491	1	1 SNP a cada 491 pb
PaH12P9	550	841	1	1 SNP a cada 841 pb
PaP3RA2	416	639	2	1 SNP a cada 320 pb

5 DISCUSSÃO

5.1 Marcadores SSR

5.1.1 Caracterização das regiões contendo SSR

Os SSRs encontrados no conjunto de sequências expressas analisadas neste trabalho totalizaram 13,5% (91/672). Em café (*Coffea sp*), um conjunto de ESTs foi analisado, encontrando-se 13,5% delas contendo SSRs (AGARRWAL et al., 2006). Resultados aproximados foram encontrados por Zheng et al. (2013), trabalhando com sequenciamento em larga escala de duas espécies de *Amorphophallus*. Estes autores relataram que 11,8% das sequências transcritas continham SSRs. Taxas menores foram descritas em ESTs de cevada (VARSHNEY et al., 2006), trigo (PENG; LAPITAN, 2005) e batata-doce (WANG et al., 2011), com 7,41%, 2,8% e 8,2% respectivamente. Vale ressaltar que a frequência de SSRs em um determinado grupo de sequências é dependente dos parâmetros para a busca de candidatos a SSRs, como *threshold* do comprimento da repetição e o número de unidades de repetição.

Com relação ao número de repetições, 40,8% dessas foram trinucleotídicas. Resultados similares relativamente à predominância de repetições trinucleotídicas em sequências expressas foram encontrados em outras espécies. Em trigo 59,7% dos EST-SSRs continham repetições trinucleotídicas, seguido de 19,7% mono- e 16,7% dinucleotídicas (PENG; LAPITAN, 2005). Em *Medicago truncatula*, 39,3% dos SSRs encontrados em ESTs eram trinucleotídicos (CHANDRA, 2011). Em batata-doce, Wang et al. (2011) encontraram 42,8% e 41,2% de SSRs tri- e dinucleotídicos respectivamente. Predominância de SSRs trinucleotídicos foram encontrados também em milheto (*Setaria italica*, 60,1%; KUMARI et al., 2013), e cevada (52,6%; VARSHNEY et al., 2006). No entanto, alguns trabalhos encontram predominância de repetições dinucleotídicas. Por exemplo, em gergelim a frequência de SSRs dinucleotídicos foi de aproximadamente 40% e 48% para SSRs com tamanho maior que 15 e maior que 18 pb respectivamente (ZHANG et al., 2012); já em café, Agarrwal et al. (2006) encontraram 46,3%.

Nossos resultados estão de acordo com estudos anteriores. Varshney, Korzun e Böner (2004) citam que, em ESTs, as repetições trinucleotídicas são as mais comuns, seguidas pelas repetições di- e tetranucleotídicas, dependendo do estudo. Por exemplo, entre as espécies de cereais, os trinucleotídicos são os mais frequentes (54 a 78%) seguido dos dinucleotídicos (17,1 a 40,4%) e tetranucleotídicos (3 a 6%) (VARSHNEY et al., 2002).

A variação considerável na frequência e distribuição de EST-SSRs entre os diferentes estudos é plausível devido à variação no tamanho das amostras, nos critérios de pesquisa, no tamanho da base de dados, e ferramentas usadas para o desenvolvimento dos SSRs bem como a diferenças intrínsecas às espécies estudadas (KUMARI et al., 2013) .

A dominância dos SSRs triméricos pode ser explicada pela supressão dos SSRs não triméricos nas regiões codantes (éxons), devido ao risco de mutações *frameshift* que podem ocorrer quando esses SSRs se alteram em uma ou duas bases (MERTZGAR; BYTOF; WILLS, 2000). As expansões e deleções em regiões codantes podem ser toleradas em repetições trinucleotídicas, que não perturbam as ORFs (*Open Reading Frames*) (KATTI et al., 2000; ASP et al., 2007). Repetições de tri- e hexanucleotídeos são muito mais encontradas relativamente às demais em ambas as sequências, codantes e não codantes. Porém, outros tipos de repetição são bem menos frequentes em regiões codantes do que em regiões não codantes. Os excessos de repetições tri- e hexanucleotídicas são principalmente controlados por pressão de mutação (MERTZGAR; BYTOF; WILLS, 2000). Os SSRs trinucleotídicos são relatados como a classe mais frequente em ESTs em plantas cultivadas, incluindo o trigo (KANTETY et al., 2002; MORGANTE; HANAFEY; POWEL, 2002; VARSHNEY et al., 2002; GAO et al., 2003).

Quanto à composição dos motivos nos transcritos de *P. alata*, verificou-se que **AG** está presente em 62,5% entre os dinucleotídicos. Já para os trinucleotídeos observou-se que os motivos **AAG**, **ACC** e **AGC** foram os mais frequentes, com 21,3%, 19,2% e 17,0% de representação respectivamente. Esses resultados estão de acordo com estudos anteriores, por exemplo, em batata-doce, os motivos di- e trinucleotídicos dominantes foram AG (26,9%) e AAG (13,5%), respectivamente (WANG et al., 2011). Em café, os motivos AG e AAG também foram os mais representativos, com 52,8% entre os di- e 28,8% entre os trinucleotídicos respectivamente (AGGARWAL et al., 2007). Em *M. truncatula*, 60,0% dos dinucleotídicos eram do tipo AG e 30,5% dos trinucleotídicos eram AAG (CHANDRA, 2011).

Kumpatla e Mukhopadhyay (2005) examinaram mais de 1,54 milhões de EST pertencentes a 55 espécies de dicotiledôneas, nas quais a frequência de ESTs com SSRs variou entre 2,65% a 16,82% nas espécies investigadas. Os motivos AG e AAG, seus complementos e permutações, foram os mais abundantes entre os di- e trinucleotídeos respectivamente, sugerindo haver um padrão entre as dicotiledôneas.

É válido lembrar que, em geral, os motivos ricos em GC são os menos frequentes em dicotiledôneas, ocorrendo o contrário em algumas monocotiledôneas. Neste trabalho, observa-

se esta tendência, pois o motivo **CCG** foi o menos frequente entre os trinucleotídicos, apenas 4,3%. Varshney et al. (2002) observaram que o motivo CCG, seus complementos e permutações, foi o mais predominante nas espécies de cereais. O motivo CCG constituiu metade dos trinucleotídeos em arroz, enquanto foi raro em dicotiledôneas (*Arabidopsis* e soja) e moderadamente abundante em outras monocotiledôneas, como o milho e o trigo. Estes resultados apoiam fortemente e estendem a noção da raridade da repetição CCG em um grande número de dicotiledôneas. Diferenças em relação à frequência dos motivos de SSRs ricos em GC foram relatadas em ESTs de outras espécies, e sua raridade parece ser uma característica comum em dicotiledôneas (CARDLE et al., 2000; GAO et al., 2003). Uma hipótese para explicar essas diferenças seria o alto conteúdo de GC e as consequentes vias de uso dos códons em ESTs de monocotiledôneas. Isso pode contar para a ocorrência diferencial da repetição CCG em mono e dicotiledôneas (MORGANTE; HANAFEY; POWELL, 2002).

5.1.2 Testes e seleção de *primers* de SSR

Houve 95,2% de amplificação, o que revela a boa qualidade do desenho dos *primers* aqui usados. Somente para 2 pares de *primers* não foi observada amplificação de SSRs no maracujá-doce: a amplificação não ocorreu em nenhum dos genótipos quando se usou o par PeE5P4, isso pode ter ocorrido devido a algum problema com a sequências dos *primers*, como a formação de estruturas secundárias ou auto anelamento, ou a sequência complementar a sequência do *primer* no genoma pode ser interrompida por um íntron. Já para PeG9P9, a amplificação ocorreu somente em *P. edulis*, mostrando que este par não é transferível para *P. alata*.

Após a otimização da reação para os *primers* que amplificaram múltiplos fragmentos, um padrão de qualidade foi obtido para 34 pares de *primers*. Portanto, 85% (34/40) deles mostraram transferibilidade de *P. edulis* para *P. alata*. Essa porcentagem é satisfatoriamente alta. Oliveira et al. (2013) avaliaram 41 SSRs desenvolvidos para *P. edulis* em relação a sua transferibilidade para 11 espécies de *Passiflora*. Todos os marcadores foram amplificados em pelo menos uma espécie, com uma média de 68,8% de transferibilidade. Os resultados indicaram um alto nível de conservação nucleotídica nas regiões para os quais os *primers* foram desenhados nas espécies avaliadas. Cerqueira-Silva et al. (2014) realizaram ensaios de transferibilidade para 69 pares de *primers* desenvolvidos a partir de duas bibliotecas genômicas enriquecidas produzidas para as espécies *P. edulis* e *P. setacea*. Todos os 69 *primers* mostraram alta porcentagem de amplificação entre espécies de *Passiflora*: 88% em *P.*

edulis, 70% em *P. setacea* e 80% em *P. cincinnata*. Adicionalmente, foi demonstrada a transferibilidade de 25 locos previamente caracterizados para *P. cincinnata* (CERQUEIRA-SILVA et al., 2012) já que mostraram 48% de amplificação em *P. edulis* e 28% em *P. setacea*.

5.1.3 Análise do polimorfismo em gel de poliacrilamida desnaturante

A maioria dos pares de *primers* (67,5%, 23/34) resultaram em várias bandas, além daquelas esperadas por ocasião do seu desenho. Essas amplificações inespecíficas podem ser devido à ocorrência de regiões genômicas similares às sequências dos *primers* de SSR, ou à repetição da própria região ao longo do genoma, já que os *primers* são de sequências de cDNA, podendo anelarem-se em regiões codantes de genes que podem estar em um *cluster*, dificultando a otimização da reação. Os PeA1P4, PeB11P9, PeC9P4, PeContig46 e PeF8P2 mostraram amplificação de um único produto (*amplicon*), porém de tamanho maior que o esperado, sendo isso justificado pela ocorrência de íntrons entre os éxons, já que os *primers* foram desenhados, como dito, a partir de cDNA. Para os *primers* que mostraram amplificação de fragmentos múltiplos (23) não foi possível essa análise, pela inviabilidade de se detectar o loco esperado.

Poczai et al. (2013) citam que, dependendo da região onde está o gene, o padrão de bandas pode ser variável, podendo ser difícil prever o tamanho exato dos produtos de PCR para famílias gênicas que resultam em múltiplos produtos. Isso pode resultar pela ocorrência de pseudogenes, ou de erros durante a PCR que geram artefatos. As condições de PCR devem ser otimizadas cuidadosamente, porque produtos de amplificação não específicos, como heterodúplex, podem ocorrer. Polimorfismos baseados em regiões conservadas ou em famílias gênicas resultam de inserções/deleções nos fragmentos amplificados, representando diferentes alelos do gene alvo.

Dos 34 SSRs aqui avaliados em géis de poliacrilamida, 10 foram polimórficos. Esses foram classificados quanto ao tipo de cruzamento, segundo Wu et al (2002). Observou-se que os locos marcadores PeB1P5, PeContig52, PeContig59 e PeE5P8 são do tipo A, o mais informativo geneticamente, com segregação esperada 1:1:1:1; o tipo A2 foi caracterizado para o loco PeB1P5 e A4 para os locos PeContig52, PeContig59 e PeE5P8. Nesses locos, foram observados quatro alelos, sendo que nos locos com configuração A4, há um alelo nulo.

O padrão de amplificação de alguns locos microssatélites estudados neste trabalho mostra que há bandas específicas acompanhadas por bandas *stutter*, como é o caso do loco

PeContig52 (Figura 7). As bandas *stutter* são uma ocorrência comum em regiões microssatélites, uma vez que, ao se caracterizarem por repetições de bases, a Taq DNA polimerase pode cometer erros de incorporação de dNTPs mais frequentemente. A genotipagem por microssatélites pode ser dificultada pela presença dessas bandas *stutter* (GINOT et al., 1996; HARKER, 2001). Por exemplo, um padrão potencialmente ambíguo pode ser produzido quando dois alelos de um heterozigoto diferem em tamanho por uma unidade, causando a sobreposição de bandas *stutter* e produção de bandas adjacentes de intensidade similar (HOFFMAN; AMOS, 2005).

Nos locos de configuração A4 (Tabela 10; Figura 7), há alelos nulos como consequência de uma mutação próxima ao local de ancoragem do *primer*. Geralmente, apenas substituições próximas a extremidade 3' do *primer* ou inserções ou deleções na região genômica complementar causam problemas de amplificação (POMPANON et al., 2005).

Bonin et al. (2004) citam alguns passos para minimizar os erros de genotipagem, entre elas: utilizar amostras controle, incluir uma amostra previamente amplificada em cada PCR como referência, replicar as amplificações, executar testes repetidamente independentes, eliminar marcadores de qualidade duvidosa, etc. Esses cuidados foram tomados neste trabalho no intuito de evitar equívocos na presente análise e em análises posteriores, de mapeamento de locos SSR, por exemplo.

5.2 Marcadores SNP

Dos 118 *primers* desenhados para amplificar os transcritos selecionados das bibliotecas de expressão, 97 (82,2%) apresentaram algum padrão de amplificação e apenas 21 não mostraram amplificação, demonstrando que os *primers* foram de boa qualidade. Dos *amplicons* obtidos para os 97 transcritos, a maioria apresentou amplificação inespecífica, sendo possível a otimização da reação para 74 desses, com a obtenção de um *amplicon* único. Como esperado, observou-se alta transferibilidade (62,7%) dos *primers* de *P. edulis* para *P. alata*. Sugere-se que as amplificações inespecíficas são resultado do anelamento dos *primers* em regiões duplicadas do genoma ou que os transcritos estudados pertencem a famílias de genes. Adicionalmente, 43 apresentaram *amplicons* de tamanho maior do que o esperado, devido à presença de íntron(s) separando os éxons utilizados como base para o desenho dos *primers*.

Foi possível o sequenciamento de 37 transcritos. Optou-se pelo método de sequenciamento direto (sem clonagem) dos fragmentos, pois os marcadores serão

posteriormente alocados no mapa de ligação de *P. alata*, sendo a população segregante composta por 180 indivíduos, ficando inviável a sua clonagem; além disso, as chances são altas de ambas fitas de DNA de cada cromossomo não serem representadas nos sequenciamentos de insertos.

O sequenciamento resultou em: presença de no mínimo um SNP nas sequências de 16 fragmentos gênicos; ausência de SNPs nas sequências de 13 fragmentos gênicos e sequências de baixa qualidade inviabilizando a análise e detecção de SNPs em 9 fragmentos gênicos. Para esses 9, observou-se nos cromatogramas, a sobreposição de picos em alguns pontos. Isto pode ser devido aos *primers* anelarem-se em duas ou mais regiões do DNA genômico, gerando mais de um *template* na amostra. O fato de serem sequências de genes que podem estar organizadas em *cluster* de resistência corrobora este resultado.

No total, 34 SNPs foram detectados em 16 fragmentos gênicos, cuja frequência variou significativamente. A frequência média foi de 1 SNP a cada 294 pb. Essa frequência é mais baixa, quando comparada àquela encontrada em um estudo anterior, realizado por Pereira et al. (2013), que relataram 1 SNP a cada 177 pb, analisando 7 genes dos acessos ‘2(12)’ e ‘SV3’ de *P. alata*. Os SNPs aqui encontrados foram detectados pelo alinhamento de 4 genótipos; assim, eventualmente, alguns SNPs foram excluídos, acreditando se tratarem de falso-positivos. Essa diferença é mais possivelmente devida ao conjunto maior de fragmentos gênicos aqui usados para a detecção de SNPs.

A frequência de SNPs no genoma das plantas varia significativamente, dependendo de alguns fatores. Assim, a seleção dos genes candidatos, a diversidade da população sob avaliação, a característica (codante ou não-codante) das regiões analisadas e o sistema de cruzamento podem afetar estas estimativas (GRATTAPAGLIA et al, 2011; LEPOITTEVIN et al, 2010; LEONFONTE et al., 2013).

Em beterraba-sacarina (*Beta vulgaris*), uma espécie de fecundação cruzada e parcialmente auto-incompatível (LANGE, BRANDENBURG; BOCK, 1999), 1.246 SNPs foram identificados em 89.941 pb originados de 315 ESTs: a frequência foi de 1 SNP a cada 72,2 pb (SCHNEIDER et al., 2007). Em milho, uma espécie de fecundação cruzada, frequências de 1 SNP ou INDEL a cada 41 pb e 1 SNP a cada 60,8 pb foram observadas em fragmentos de 18 genes, analisando-se 38 linhagens (CHING et al., 2002). Para o azevém, uma gramínea forrageira (*Lolium perene*), a frequência de 1 SNP a cada 54,1 pb foi observada em 100 genes (COGAN et al., 2006). Em eucalipto (*Eucalyptus camaldulensis*), a frequência média foi de 1 SNP a cada 83,9 pb analisando-se 41 genes (HENDRE; KAMALAKANNAN; VARGHESE, 2012). Taxas semelhantes foram encontradas em espécies altamente

heterozigóticas, como o pinheiro (1/102,6 pb; DANTEC et al., 2004), a uva (1/64 pb; LIJAVETZKY et al., 2007) e o centeio (1/52 pb; LI et al., 2011). Já em espécies de autofecundação, o polimorfismo encontrado teve frequências muito menores. As taxas de SNP nas regiões codantes de espécies autógamas foram: 1/333 pb em arroz (YU et al., 2005); 1/191 pb em soja (VAN et al., 2005) e 1/336 pb em *A. thaliana* (SCHMID et al., 2003).

Pelo exposto acima, observa-se que a frequência do polimorfismo determinada neste trabalho e em Pereira et al. (2013) para *P. alata* não está em concordância com o relatado para outras espécies de fecundação cruzada. Interessantemente, a frequência de SNPs em *P. alata* se aproxima mais das taxas encontradas em espécies de autofecundação. Isto corrobora a ideia de que o polimorfismo molecular não é alto nas passifloras investigadas.

A frequência dos SNPs aqui encontrados foi de 50% em regiões codantes e em não-codantes. Do total de substituições encontradas, todas foram bialélicas. Estudos sugerem que mutações sejam mais frequentes em regiões não codantes do genoma, pois nessas sequências espera-se maior tolerância a eventos mutacionais, levando a maiores níveis de polimorfismo (SCHNEIDER et al., 2007). Assim, em eucalipto, os SNPs foram mais frequentes em regiões intrônicas e UTRs em comparação com os éxons, pois estes são submetidos à seleção evolutiva (HENDRE; KAMALAKANNAN; VARGHESE, 2012); da mesma forma, em uva, a frequência das variações foi maior nas regiões não codantes (1/47 pb) comparativamente às codantes (1/69 pb) (LIJAVETZKY et al., 2007).

Verificou-se que as substituições de bases encontradas nas sequências de *P. alata* foram preferencialmente transversões (~53%), com uma razão transição/transversão de 0,89. Em plantas, há registros que as substituições do tipo transição são mais frequentes. Em beterraba-sacarina, a razão transição/transversão foi 1,64. (SCHNEIDER et al., 2007), em ervilha 2:1 (LEONFONTE et al., 2013) e, em eucalipto, a razão foi de 2,76 (HENDRE; KAMALAKANNAN; VARGHESE, 2012). Em uva, das substituições, 59,3% foram devidas a transições e 40,7%, a transversões (LIJAVETZKY et al., 2007). Em concordância com a frequência encontrada para *P. alata*, essa razão foi de 0,92 em soja (ZHU et al., 2003).

Os resultados observados em relação à frequência de SNPs em regiões codantes e predominância das substituições do tipo transversão em *P. alata* sugerem baixa pressão de seleção nos fragmentos gênicos analisados, ou que o tamanho das sequências foi relativamente pequeno, em média 625 pb.

Das substituições encontradas, observou-se que as transições do tipo **C/T** foram as mais frequentes (29,4%, 10/34), seguido pelas transições do tipo **A/G** e transversões do tipo **G/C**, com 17,6% (6/34) de ocorrência. As transversões **C/T** e **A/G** foram as mais encontradas

em plantas. Em ervilha, as duas variações de SNP mais comuns foram A/G e C/T , representando 36% e 32% de todas as variações, respectivamente. As outras variações (T/G, C/G, A/C e A/T) representaram menos de 10% do total (LEONFONTE et al., 2013). Estudos em leguminosas, como o trevo branco (HAND et al., 2008) e o grão-de-bico (GAUR et al., 2012) também relatam substituições A/G e C/T como as mais comuns. A alta proporção de transições C/T tende a ser devido a reações de desaminação das 5-metilcitosinas, que ocorre ao longo do tempo evolutivo, particularmente nos nucleotídeos CpG (HOLLIDAY; GRIGG, 1993). A alta mutabilidade das ilhas de CpG é a principal razão para a maior proporção de transições do que transversões nos éxons de plantas, bem como de animais (DUNCAN; MILLER, 1980; JIANG et al., 2008).

Dos 17 SNPs encontrados em éxons, 13 (76,4%), resultaram em mudança no aminoácido codificado, havendo uma relação de 1,3:1 de mutação silenciosa versus não silenciosa. Adicionalmente, pela divisão dos aminoácidos em classes (VOET; VOET, 2013), 5 (38,5%) dessas mutações não silenciosas foram não conservativas, ou seja, a mudança em um nucleotídeo resultou na mudança de um aminoácido com diferentes características, podendo gerar uma mudança na função da proteína. Em uva, nas regiões codantes, foi observada uma razão de 1:1 de mutação silenciosa versus não silenciosa, com 16% das mutações não silenciosas originando mudanças de aminoácidos não conservativos (LIJAVETZKY et al., 2007). A razão de mutação silenciosa versus não silenciosa foi de 1,5:1 (PAVY et al., 2006) em abeto, e de 2:1 em *Arabidopsis* 2:1 (SCHMID et al., 2003). Os maiores valores das mutações não silenciosas sugere a existência de uma pressão de seleção reduzida, resultando em uma maior diversidade proteica que pode ser a base da variação fenotípica (LIJAVETZKY et al., 2007). Em eucalipto, observou-se que as transições resultaram em mais mutações sinônimas, ao passo que as transversões resultaram em mais mutações não sinônimas, como esperado (DAGAN; TALMOR; GRAUR, 2002).

Uma função putativa foi atribuída a todos os 15 genes avaliados neste estudo, 80% das funções mostrando homologia com as mesmas proteínas codificadas pelas sequências de *P. edulis*.

Os locos SNP são de grande importância em estudos genéticos por serem codominantes e, nesse caso, gênicos. Vale destacar o loco PaA6P9, cuja sequência tem similaridade com uma proteína da família das lipoxigenases. O gene da lipoxigenase foi o mais expresso em resposta à infecção por *Xap* em *P. edulis* (MUNHOZ, 2013), sendo de grande valor mapeá-lo em um dos grupos de ligação do mapa de *P. alata*. Sanier et al. (2012) estudaram o gene GhLOX2 em algodão e confirmaram que este gene codifica uma

lipoxigenase-13 e apresenta um peptídeo de trânsito para o cloroplasto. Os autores observaram que o gene GhLOX2 está altamente relacionado com a resposta de hipersensibilidade, pois foi altamente expresso nas primeiras horas após inoculação com *Xanthomonas campestris* pv. *malvacearum*.

5.3 Polimorfismo molecular em *Passiflora*

A taxa média de polimorfismo nos genótipos aqui investigados considerando todos os marcadores (SSRs e SNPs) foi 36,7%. Essa taxa é maior do que a encontrada em estudos anteriores em *P. alata*. As taxas foram de 18%, 20% e 35% utilizando AFLP (NUNES, 2010), SSR (PENHA et al., 2013) e SSR, SNP, AFLP e RGA (PEREIRA et al., 2013), respectivamente. Para *P. edulis* foram de 26,4%, 13,0% e 24,7% com base em RAPD (CARNEIRO; VIEIRA, 2002), AFLP (LOPES et al., 2006) e SSR (OLIVEIRA et al., 2008), respectivamente. No entanto, ainda pode ser considerada baixa para espécies de cruzamento essencialmente aberto. Conforme determinado pela análise de SSRs, níveis mais elevados de polimorfismo foram relatados para outras espécies autoincompatíveis, incluindo a chicória (69,3%; CADALEN et al., 2010), o cacau (52,0%; PUGH et al., 2004) e a maçã (50,5%; CELTON et al. 2009).

O baixo polimorfismo parece ser característico do gênero *Passiflora*, ao menos para o grupo de espécies que vem sendo analisadas. Em um estudo visando ao desenvolvimento de marcadores microssatélites para *P. cincinnata*, apenas 21% deles foram polimórficos (CERQUEIRA-SILVA et al., 2012). Em outro estudo, 69 SSRs foram desenvolvidos a partir de uma biblioteca genômica enriquecida desenvolvida com genótipos de *P. edulis* e *P. setacea*. Entre os 69 locos, 15% foram polimórficos em *P. edulis*, 29% em *P. setacea* e 20% em *P. cincinnata* (CERQUEIRA-SILVA et al., 2014). Em maracujá-roxo (*Passiflora edulis* Sims), foram observados 7% de locos AFLP polimórficos e nenhum polimorfismo em relação a SSR (ORTIZ et al., 2012).

Vale salientar que, nesse estudo, o polimorfismo molecular foi medido usando dois acessos fenotipicamente distintos oriundos de diferentes regiões geográficas. Notavelmente, relativamente aos locos SNP, a contribuição do genitor feminino, que pertence a uma progênie de uma planta-mãe de origem amazônica, foi maior do que a do masculino, que foi coletado em um pomar comercial na região de Piracicaba, SP.

Completando, os polimorfismos aqui relatados entre os genitores da população de mapeamento de *P. alata*, todos em regiões expressas, devem contribuir para o enriquecimento

do mapa de ligação e facilitar o mapeamento de locos quantitativos associados à resposta da população à infecção por *Xap*.

6 CONCLUSÕES

O uso de sequências de bibliotecas de expressão de *Passiflora edulis* permite, tal como realizado neste trabalho, o desenvolvimento de marcadores genéticos em *Passiflora alata*, com alta transferibilidade de locos de uma espécie para outra.

O desenvolvimento de marcadores SSR (ou microssatélites) e SNPs para *Passiflora alata* a partir dos transcritos de *Passiflora edulis* proporciona a detecção de locos polimórficos.

Este trabalho confirma que as espécies já investigadas de *Passiflora* apresentam baixo polimorfismo molecular, comparativamente a outras espécies de fecundação cruzada obrigatória.

Passiflora alata é uma espécie pouco estudada do ponto de vista genético e apresenta potencial para ser explorada, justificando o desenvolvimento de novos marcadores genéticos a partir de bibliotecas de expressão e genômicas. Este estudo e disponibilizou ferramentas para o mapeamento de marcadores funcionais bem como para uso em análises de diversidade em passifloras.

Nossa experiência, utilizando diferentes estratégias para investigar sequências de genes foi satisfatória, e forneceu informações adicionais quanto à ocorrência de SSRs e SNPs em *Passiflora alata*.

REFERÊNCIAS

- AGARWAL, M.; SHRIVASTAVA, N.; PADH, H. Advances in molecular marker techniques and their applications in plant sciences. **Plant Cell Reports**, Berlin, v. 27, p. 617-631, 2008.
- AGGARWAL, R.K.; HENDRE, P.S.; VARSHNEY, R.K.; BHAT, P.R.; KRISHNAKUMAR, V.; SINGH, L. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. **Theoretical and Applied Genetics**, Berlin, v. 114, p. 359-372, 2007.
- ASP, T.; FREI, U.K.; DIDION, T.; NIELSEN, K.K.; LÜBBERSTEDT, T. Frequency, type, and distribution of EST-SSRs from three genotypes of *Lolium perenne*, and their conservation across orthologous sequences of *Festuca arundinacea*, *Brachypodium distachyon*, and *Oryza sativa*. **BMC Plant Biology**, London, v. 7, p. 822-828, 2007.
- BELÓ, A.; ZHENG, P.; LUCK, S.; SHEN, B.; MEYER, D. J.; LI, B.; TINGEY, S.; RAFALSKI, A. Whole genome scan detects an allelic variant of *fad2* associated with increased oleic acid levels in maize. **Molecular Genetics and Genomics**, Berlin, v. 279, p. 110, 2008.
- BILLOTE, N.; LAGODA, P.J.L.; RISTERUCCI, A.M.; BAURENS, F.C. Microsatellite-enriched libraries: applied methodology for the development of SSRs markers in tropical crops. **Fruits**, Paris, v. 54, p. 277-288, 1999.
- BONIN, A.; BELLEMAIN, E.; BRONKEN E.P.; POMPANON, F.; BROCHMANN, C.; TABERLET, P. How to track and assess genotyping errors in population genetics studies. **Molecular Ecology**, Oxford, v. 13, p. 3261-3273, 2004.
- BRUCKNER, C.H.; CASALI, V.W.D.; MORAES, C.F.; REDAZZI, A.J.; SILVA, E.A.M. Self-incompatibility in passion fruit (*Passiflora edulis* Sims). **Acta Horticultrae**, Lima, n. 370, p. 45-57, 1995.
- CADALEN, T.; MÖRCHEN, M.; BLASSIAU, C.; CLABAUT, A.; SCHEER, I.; HILBERT, J. L.; HENDRIKS, T.; QUILLET, M. C. Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.). **Molecular Breeding**, Dordrecht, v. 25, p. 699-722, 2010.
- CARDLE, L.; RAMSAY, L.; MILBOURNE, D.; MACAULAY, M.; MARSHALL, D.; WAUGH, R. Computational and experimental characterization of physically clustered simple sequence repeats in plants. **Genetics**, Austin, v. 156, p. 847-854, 2000.
- CARNEIRO, M.S.; CAMARGO, L.E.A.; COELHO, A.S.G.; VENCOVSKY, R.; LEITER JUNIOR, R.P.; STENZEL, N.M.C.; VIEIRA, M.L.C. RAPD-based genetic linkage maps of yellow passion fruit (*Passiflora edulis* Sims. f. *flavicarpa* Deg.). **Genome**, Ottawa, v. 45, p. 670-678, 2002.
- CARNEIRO, M.S.; VIEIRA, M.L.C. Mapas genéticos em plantas. **Bragantia**, Campinas, v. 61, p. 89-100, 2002.

CELTON, J.M.; TUSTIN, D.S.; CHAGNÉ, D.; GARDINER, S.E. Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. **Tree Genetics and Genomes**, Berlin, v. 5, p. 93-107, 2009.

CERQUEIRA-SILVA, C.B.M.; SANTOS, E.S.L.; CONCEIÇÃO, L.D.; CARDOSO-SILVA, C.B.; PEREIRA, A.S.; OLIVEIRA, A.C.; CORRÊA, R.X. Genetic variation in a wild population of the 'sleep' passion fruit (*Passiflora setacea*) based on molecular markers. **Genetics and Molecular Research**, Ribeirão Preto, v. 11, p. 731-738, 2012a.

CERQUEIRA-SILVA, C.B.M.; SANTOS, E.S.L.; SOUZA, A.M.; MORI, G.M.; OLIVEIRA, E.J.; CORRÊA, R.X.; SOUZA, A.P. Development and characterization of microsatellite markers for the wild South American *Passiflora cincinnata* (Passifloraceae). **American Journal of Botany**, Columbus, v. 99, p. 170-172, 2012b.

CERQUEIRA-SILVA, C.B.M.; SANTOS, E.S.L.; VIEIRA, J.G.P.; MORI, G.M.; JESUS, O.N.; CORRÊA, R.X.; SOUZA, A.P. New microsatellite markers for wild and commercial species of *Passiflora* (Passifloraceae) and cross-amplification. **Applications in Plant Sciences**, St. Louis, v. 2, p. 1-5, 2014.

CERVI, A.C.; MILWARD-DE-AZEVEDO, M.A.; BERNACCI, L.C.; NUNES, T.S. **Passifloraceae**. In: Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro, 2011. Disponível em <http://floradobrasil.jbrj.gov.br/2010/FB000182>. Acesso em: 17 mai. 2013.

CHAMBERLAIN, J.S.; GIBBS, R.A.; RANIER, J.E.; NGUYEN, P.N.; CASKEY, C.T. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. **Nucleic Acids Research**, Oxford, v. 16, p. 11141-11156, 1988.

CHANDRA, A. Use of EST database markers from *M. truncatula* in the transferability to other forage legumes. **Journal of Environmental Biology**, Lucknow, v. 32, p. 347-354, 2011.

CHING, A.; CALDWELL, K.S.; JUNG, M.; DOLAN, M.; SMITH, O.S.H.; TINGEY, S.; MORGANTE, M.; RAFALSKI, A.J. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. **BMC Genetics**, London, v. 3, p. 19-33, 2002.

CHOUMANE, W.; WINTER, P.; BAUM, M.; KAHL, G. Conservation of microstellite flanking sequences in different taxa of Leguminosae. **Euphytica**, Wageningen, v. 138, p. 239-245, 2004.

COGAN, N.O.I.; PONTING, R.C.; VECCHIES, A.C.; DRAYTON, M.C.; GEORGE, J.; DRACATOS, P.M.; DOBROWOLSKI, M.P.; SAWBRIDGE, T.I.; SMITH, K.F.; SPANGENBERG, G.C.; FORSTER, J.W. Gene-associated single nucleotide polymorphism discovery in perennial ryegrass (*Lolium perenne* L.). **Molecular Genetics and Genomics**, Berlin, v. 276, p. 101-112, 2006.

COLLARD, B.C.Y.; MACKILL, D.J. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. **Philosophical Transactions of the Royal Society B**, London, v. 363, p. 557-572, 2008.

COSSON, P.; DECROOCQ, V.; REVERS F. Development and characterization of 96 microsatellite markers suitable for QTL mapping and accession control in an *Arabidopsis* core collection. **Plant Methods**, London, v. 10, p. 1-6, 2014.

CRESTE, S.; TULMANN, A.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reports**, New York, v. 19, p. 229-306, 2001.

DAGAN, T.; TALMOR, Y.; GRAUR, D. Ratios of radical to conservative amino acid replacement are affected by mutational and compositional factors and may not be indicative of positive Darwinian selection. **Molecular Biology and Evolution**, Chicago, v. 19, p. 1022-1025, 2002.

DANTEC, L.L.; CHAGNÉ, D.; POT, D.; CANTIN, O.; GARNIER-GÉRÉ, P.; BEDON, F.; FRIGERIO, J. M.; CHAUMEIL, P.; LÉGER, P.; GARCIA, V.; LAIGRET, F.; DARUVAR, A. de; PLOMION, C. Automated SNP detection in expressed sequence tags: Statistical considerations and application to maritime pine sequences. **Plant Molecular Biology**, Dordrecht, v. 54, p. 461-470, 2004.

DIATCHENKO, L.; LAU, Y.F.; CAMPBELL, A.P.; CHENCHIK, A.; MOQADAM, F.; HUANG, B.; LUKYANOV, S.; LUKYANOV, K.; GURSKAYA, N.; SVERDLOV, E.D.; SIEBERT, P.D. Suppression subtractive hybridization: a method for generating 140 differentially regulated or tissue-specific cDNA probes and libraries. **Proceedings of the National Academy of Sciences**, Washington, v. 93, n. 12, p. 6025–6030, 1996.

DUNCAN, B.K.; MILLER, J.H. Mutagenic deamination of cytosine residues in DNA. **Nature**, New York, v. 287, p. 560-561, 1980.

EDWARDS, M.C.; GIBBS, R.A. Multiplex PCR: Advantages, development, and applications. **Genome Research**, New York, v. 3, p. 65-75, 1994.

ELLEGREN, H. Microsatellites: Simple sequences with complex evolution. **Nature Reviews Genetics**, London, v. 5, p. 435-445, 2004.

FALEIRO, F.G.; JUNQUEIRA, N.T.V.; BRAGA, M.F.; PEIXOTO, J.R. Caracterização de germoplasma e melhoramento genético do maracujazeiro assistidos por marcadores moleculares: resultados de pesquisa 2005-2008. **Boletim de Pesquisa e Desenvolvimento**, Planaltina, v. 1, n. 1, 59 p, 2008.

FERREIRA, T.G.; PENHA, H.A.; ZUCCHI, M.I.; SANTOS, A.A.; HANAI, L.R.; JUNQUEIRA, N.; BRAGA, M.F.; VENCOVSKY, R.V.; VIEIRA, M.L.C. Outcrossing rate in sweet passion fruit based on molecular markers. **Plant Breeding**, Berlin, v. 129, p. 727-730, 2010.

FRAZER, K.A.; BALLINGER, D.G.; COX D. R.; HINDS, D.A.; STUVE, L.L.; GIBBS, R.A.; BELMONT, J.W.; BOUDREAU, A.; HARDENBOL, P.; LEAL, S.M.; PASTERNAK, S.; WHEELER, D.A.; WILLIS, T.D.; YU, F.; YANG, H.; ZENG, C.; GAO, Y.; HU, H.; HU, W.; LI, C.; LIN, W.; LIU, S.; PAN, H.; TANG, X.; WANG, J.; WANG, W.; YU, J.; ZHANG, B.; ZHANG, Q.; ZHAO, H.; ZHAO, H.; ZHOU, J.; GABRIEL, S. B.; BARRY.; R.; BLUMENSTIEL, B.; CAMARGO, A.; DEFELICE, M.; FAGGART, M.; GOYETTE, M.; GUPTA, S.; MOORE, J.; NGUYEN, H.; ONOFRIO, R. C.; PARKIN, M.; ROY, J.; STAHL, E.; WINCHESTER, E.; ZIAUGRA, L.; ALTSHULER, D.; SHEN, Y.; YAO, Z.; HUANG, W.; CHU, X.; HE, Y.; JIN, L.; LIU, Y.; SHEN, Y.; SUN, W.; WANG, H.; WANG, Y.; WANG, Y.; XIONG, X.; XU, L.; WAYE, M. M.; TSUI, S. K.; XUE, H.; WONG, J. T.; GALVER, L. M.; FAN, J. B.; GUNDERSON, K.; MURRAY, S. S.; OLIPHANT, A. R.; CHEE, M. S.; MONTPETIT, A.; CHAGNON, F.; FERRETTI, V.; LEOBOEUF, M.; OLIVIER, J. F.; PHILLIPS, M. S.; ROUMY, S.; SALLÉE, C.; VERNER, A.; HUDSON, T. J.; KWOK, P. Y.; CAI, D.; KOBOLDT, D. C.; MILLER, R. D.; PAWLIKOWSKA, L.; TAILLON-MILLER, P.; XIAO, M.; TSUI, L. C.; MAK, W.; SONG, Y. Q.; TAM, P. K.; NAKAMURA, Y.; KAWAGUCHI, T.; KITAMOTO, T.; MORIZONO, T.; NAGASHIMA, A.; OHNISHI, Y.; SEKINE, A.; TANAKA, T.; TSUNODA. T.; DELOUKAS, P.; BIRD, C. P.; DELGADO, M.; DERMITZAKIS, E. T.; GWILLIAM, R.; HUNT, S.; MORRISON, J.; POWELL, D.; STRANGER, B. E.; WHITTAKER, P.; BENTLEY, DR.; DALY, M. J.; DE BAKKER, P. I.; BARRETT, J.; CHRETIEN, Y. R.; MALLER, J.; MCCARROLL, S.; PATTERSON, N.; PE'ER, I.; PRICE, A.; PURCELL, S.; RICHTER, D. J.; SABETI, P.; SAXENA, R.; SCHAFFNER, S. F.; SHAM, P. C.; VARILLY, P.; ALTSHULER, D.; STEIN, L. D.; KRISHNAN, L.; SMITH, A. V.; TELLO-RUIZ, M. K.; THORISSON, G. A.; CHAKRAVARTI, A.; CHEN, P. E.; CUTLER, D. J.; KASHUK, C. S.; LIN, S.; ABECASIS, G. R.; GUAN, W.; LI, Y.; MUNRO, H. M.; QIN, Z. S.; THOMAS, D. J.; MCVEAN, G.; AUTON, A.; BOTTOLO, L.; CARDIN, N.; EYHERAMENDY, S.; FREEMAN, C.; MARCHINI, J.; MYERS, S.; SPENCER, C.; STEPHENS, M.; DONNELLY, P.; CARDON, L. R.; CLARKE, G.; EVANS, D. M.; MORRIS, A. P.; WEIR, B. S.; TSUNODA, T.; MULLIKIN, J. C.; SHERRY, S. T.; FEOLO, M.; SKOL, A.; ZHANG, H.; ZENG, C.; ZHAO, H.; MATSUDA, I.; FUKUSHIMA, Y.; MACER, D. R.; SUDA, E.; ROTIMI, C. N.; ADEBAMOWO, C. A.; AJAYI, I.; ANIAGWU, T.; MARSHALL, P. A.; NKWODIMMAH, C.; ROYAL, C. D.; LEPPERT, M. F.; DIXON, M.; PEIFFER, A.; QIU, R.; KENT, A.; KATO, K.; NIIKAWA, N.; ADEWOLE, I. F.; KNOPPERS, B. M.; FOSTER, M. W.; CLAYTON, E. W.; WATKIN, J.; GIBBS, R. A.; BELMONT, J. W.; MUZNY, D.; NAZARETH, L.; SODERGREN, E.; WEINSTOCK, G. M.; WHEELER, D. A.; YAKUB, I.; GABRIEL, S. B.; ONOFRIO, R. C.; RICHTER, D. J.; ZIAUGRA, L.; BIRREN, B. W.; DALY, M. J.; ALTSHULER, D.; WILSON, R. K.; FULTON, L. L.; ROGERS, J.; BURTON, J.; CARTER, N. P.; CLEE, C. M.; GRIFFITHS, M.; JONES, M. C.; MCLAY, K.; PLUMB, R. W.; ROSS, M. T.; SIMS, S. K.; WILLEY, D. L.; CHEN, Z.; HAN, H.; KANG, L.; GODBOUT, M.; WALLENBURG, J. C.; L'ARCHEVÊQUE, P.; BELLEMARE, G.; SAEKI, K.; WANG, H.; AN, D.; FU, H.; LI, Q.; WANG, Z.; WANG, R.; HOLDEN, A. L.; BROOKS, L. D.; MCEWEN, J. E.; GUYER, M. S.; WANG, V. O.; PETERSON, J. L.; SHI, M.; SPIEGEL, J.; SUNG, L. M.; ZACHARIA, L. F.; COLLINS, F. S.; KENNEDY, K.; JAMIESON, R.; STEWART, J. A second generation human haplotype map of over 3.1 million SNPs. International HapMap Consortium1. *Nature*, New York, v. 18, p. 851-861, 2007.

GANAL, M.W.; ALTMANN, T.; RÖDER, M.S. SNP identification in crop plants. *Current Opinion in Plant Biology*, London, v. 12, p. 211-217, 2009.

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

GAUR, R.; AZAM, S.; JEENA, G.; KHAN, A.W.; CHOUDHARY, S.; JAIN, M.; YADAV, G.; TYAGI, A.K.; CHATTOPADHYAY, D.; BHATIA, S. High-throughput SNP discovery and genotyping for constructing a saturated linkage map of chickpea (*Cicer arietinum* L.). **DNA Research**, Oxford, v. 19, p. 357-373, 2012.

GINOT, F.; BORDELAIS, I.; NGUYEN, S.; GYAPAY, G. Correction of some genotyping errors in automated fluorescent microsatellite analysis by enzymatic removal of one base overhangs. **Nucleic Acids Research**, Oxford, v. 24, p. 540-541, 1996.

GRATTAPAGLIA, D.; SILVA-JUNIOR, O.B.; KIRST, M.; LIMA, B.M. de; FARIA, D. A.; PAPPAS JUNIOR, G.J. High-throughput SNP genotyping in the highly heterozygous genome of Eucalyptus: Assay success, polymorphism and transferability across species. **BMC Plant Biology**, London, v. 11, p. 1-18, 2011.

HALFELD-VIEIRA, B. A.; NECHET, K. L. **Mancha-bacteriana do maracujá: sintomas, danos e medidas de controle**. Boa Vista: Embrapa Roraima, 2006. 4 p. (Embrapa Roraima. Comunicado técnico, 03).

HAND, M.L.; PONTING, R.C.; DRAYTON, M.C.; LAWLESS, K. A.; COGAN, N.O.I.; CHARLES BRUMMER, E.; SAWBRIDGE, T.I.; SPANGENBERG, G.C.; SMITH, K.F.; FORSTER, J.W. Identification of homologous, homoeologous and paralogous sequence variants in an outbreeding allopolyploid species based on comparison with progenitor taxa. **Molecular Genetics and Genomics**, Berlin, v. 280, p. 293-304, 2008.

HARKER, N. Collection, Reporting and Storage of Microsatellite Genotype Data. In: HENRY, R. J. (Ed.). **Plant Genotyping: the DNA Fingerprinting of Plants**. CAB International, Wallingford; New York, 2001. p. 251-264.

HENDRE, P.S.; KAMALAKANNAN, R.; VARGHESE, M. High-throughput and parallel SNP discovery in selected candidate genes in *Eucalyptus camaldulensis* using Illumina NGS platform. **Plant Biotechnology Journal**, Oxford, v. 10, p. 646-656, 2012.

HIGASHIYAMA, T. Peptide signaling in pollen-pistil interactions. **The Plant Cell Physiology**, Oxford, v. 51, p. 177-189, 2010.

HOFFMAN, J.I.; AMOS, W. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. **Molecular Ecology**, Oxford, v. 14, p. 599-612, 2005.

HOLLIDAY, R.; GRIGG, G.W. DNA methylation and mutation. **Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis**, Amsterdam, v. 285, p. 61-67, 1993.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produção agrícola municipal**. Disponível em: <<http://sidra.ibge.gov.br/2010>>. Acesso em: 12 nov. 2013.

- JIANG, Z.; WU, X.L.; ZHANG, M.; MICHAL, J.J.; WRIGHT JUNIOR, R.W. The complementary neighborhood patterns and methylation-to-mutation likelihood structures of 15,110 single-nucleotide polymorphisms in the bovine genome. **Genetics**, Austin, v. 180, p. 639-647, 2008.
- JIN, H.J.; KWAK, K.D.; HAMMER, M.F.; NAKAHORI, Y.; SHINKA, T.; LEE, J.W.; JIN, F.; JIA, X.; TYLER-SMITH, C.; KIM, W. Y-chromosomal DNA haplogroups and their implications for the dual origins of the Koreans Hum. **Genetics**, Austin, v. 114, p. 27–35, 2003.
- JUNQUEIRA, N. T. V.; JUNQUEIRA, K. P. Manejo das principais doenças do maracujazeiro. In: SUSSEL, A. A. B.; MEDEIROS, F. H. V.; RIBEIRO JÚNIOR, P. M.; UCHOA, C. N.; AMARAL, D. R.; MEDEIROS, F. C. L.; PEREIRA, R. B.; SANTOS, J.; LIMA, L. M.; ROSWALKA, L. C. **Manejo integrado de doenças de fruteiras**. Lavras: Ufla, 2007. 1 CD-ROM.
- JURKA, J.; PETHIYAGODA, C. Simple repetitive DNA sequences from primates: compilation and analysis. **Journal of Molecular Evolution**, New York, v. 40, p. 120–126, 1995.
- KANTETY, R.V.; LA ROTA, M.; MATTHEWS, D.E.; SORRELLS, M.E. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. **Plant Molecular Biology**, Dordrecht, v. 48, p. 501-510, 2002.
- KATTI, M.V.; RANJEKAR, P.K.; GUPTA, V.S. Differential distribution of simple sequence repeats in eukaryotic genome sequences. **Molecular Biology and Evolution**, Oxford, v. 18, p. 1161-1167, 2001.
- KATTI, M.V.; SAMI-SUBBU, R.; RANJEKAR, P.K.; GUPTA, V. S. Amino acid repeat patterns in protein sequences: their diversity and structural-functional implications. **Protein Science**, New York, v. 9, p. 1203-1209, 2000.
- KAVATI, R.; PIZA JUNIOR, C.T. **Cultura do maracujá-doce**. Campinas: CATI, 2002. (CATI. Boletim Técnico, 244).
- KIKUCHI, S.; TOMINAGA, D.; ARITA, M.; TAKAHASHI, K.; TOMITA, M. Dynamic modeling of genetic networks using genetic algorithm and S-system. **Bioinformatics**, Oxford, v. 19, p. 643-650, 2003.
- KIJAS, J.M.; FOWLER, J.C.; GARBETT, C.A.; THOMAS, M.R. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidincoated magnetic particles. **Biotechniques**, Natick, v. 16, p. 656-662, 1994.
- KING, J.; THOROGOOD, D.; EDWARDS, K.J. Development of a genomic microsatellite library in perennial ryegrass (*Lolium perenne* L.). **Annals of Botany**, London, v. 101, p. 845–853, 2008.
- KOFLER, R.; SCHLÖTTERER, C.; LELLEY, T. SciRoKo: A new tool for whole genome microsatellite search and investigation. **Bioinformatics**, Oxford, v. 23, p. 1683-1685, 2007.

KOTA, R.; VARSHNEY, R.K.; PRASAD, M.; ZHANG, H.; STEIN, N.; GRANER, A. EST-derived single nucleotide polymorphism markers for assembling genetic and physical maps of the barley genome. **Functional and Integrative Genomics**, Berlin, v. 8, p. 223-233, 2008.

KRAPP, F.; WÖHRMANN, T.; PINANGÉ, D.S.B.; BENKO-ISEPPON, A.M.; HUETTEL, B.; WEISING, K. A set of plastid microsatellite loci for the genus *Dyckia* (Bromeliaceae) derived from 454 pyrosequencing. **American Journal of Botany**, Columbus, v. 12, p. 470-473, 2012.

KUMARI, K.; MUTHAMILARASAN, M.; MISRA, G.; GUPTA, S.; SUBRAMANIAN, A.; PARIDA, S. K.; CHATTOPADHYAY, D.; PRASAD, M. Development of eSSR-Markers in *Setaria italica* and Their Applicability in Studying Genetic Diversity, Cross-Transferability and Comparative Mapping in Millet and Non-Millet Species. **PLoS ONE**, San Francisco, v. 8, p. 1-15, 2013.

KUMPATLA, S.P.; MUKHOPADHYAY, S. Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. **Genome**, Ottawa, v. 48, p. 985-998, 2005.

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

LANGE, W.; BRANDENBURG, W.A.; BOCK, T.S.M. Taxonomy and cultonony of beet (*Beta vulgaris* L.). **Botanical Journal of the Linnean Society**, London, v. 130, p. 81-96, 1999.

LEONFORTE, A.; SUDHEESH, S.; COGAN, N.O.I.; SALISBURY, P.A.; NICOLAS, M. E.; MATERNE, M.; FORSTER, J.W.; KAUR, S. SNP marker discovery, linkage map construction and identification of QTLs for enhanced salinity tolerance in field pea (*Pisum sativum* L.). **BMC Plant Biology**, London, v. 13, p. 1-14, 2013.

LEPOITTEVIN, C.; FRIGERIO, J.M.; GARNIER-GÉRÉ, P.; SALIN, F.; CERVERA, M.; VORNAM, B.; HARVENGT, L.; PLOMION, C. In Vitro vs in Silico detected SNPs for the development of a genotyping array: What can we learn from a non-model species? **PLoS ONE**, San Francisco, v. 5, p. 1-9, 2010.

LI, Y.; HASENEYER, G.; SCHÖN, C.C.; ANKERST, D.; KORZUN, V.; WILDE, P.; BAUER, E. High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (*Secale cereale* L.) genes involved in frost response. **BMC Plant Biology**, London, v. 11, p. 1-14, 2011.

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

LIJAVETZKY, D.; CABEZAS, J.; IBÁÑEZ, A.; RODRÍGUEZ, V.; MARTÍNEZ-ZAPATER, J. M. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. **BMC Genomics**, London, v. 8, p. 1-11, 2007.

- LOPES, R.; LOPES, M.T.G.; CARNEIRO, M.S.; MATTA, F.D.P.; CAMARGO, L.E.A.; VIEIRA, M.L.C. Linkage and mapping of resistance genes to *Xanthomonas axonopodis* pv. *passiflorae* in yellow passion fruit. **Genome**, Ottawa, v. 49, p. 17-29, 2006.
- LÓPEZ, C.E.; ACOSTA, I.F.; JARA, C.; PEDRAZA, F.; GAITAN-SOLIS, E.; GALLEGU, G.; BEEBE, S.; TOHME, J. Identifying resistance gene analogs associated with resistances to diferente pathogens in common bean. **Phytopathology**, Saint Paul, v. 93, p. 88-95, 2003.
- MALAVOLTA JÚNIOR, V. A.; BERIAM, L. O. S.; RODRIGUES NETO, J. **Podridão do fruto, novo sintoma relacionado a *Xanthomonas axonopodis* pv. *Passiflorae***. Arquivos do Instituto Biológico, Campinas, v. 68, n. 2, p. 121-123, jul./dez. 2001.
- MALIEPAARD, C.; JANSEN, J.; VAN OOIJEN, J. W. Linkage analysis in a full-sib family of an outbreeding plant species: Overview and consequences for applications. **Genetical Research**, London, v. 70, p. 237-250, 1997.
- MARCOTTE, E.M.; PELLEGRINI, M.; YEATES, T.O.; EISENBERG, D. A census of protein repeats. **Journal of Molecular Biology**, London, v. 293, p. 151-160, 1999.
- McCARTHY, M.I.; ABECASIS, G.R.; CARDON, L.R.; GOLDSTEIN, D.B.; LITTLE, J.; IOANNIDIS, J.P.; HIRSCHHORN, J.N. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. **Nature Reviews Genetics**, London, v. 5, p. 356-369, 2008.
- MELETTI, L.M.M.; BRÜCKNER, C.H. Melhoramento Genético. In: BRÜCKNER, C.H.; PICANÇO, M.C. **Maracujá: tecnologia de produção, pós-colheita, agroindústria, mercado**. Porto Alegre: Cinco Continentes, 2001. p. 345-385.
- METZGAR, D.; BYTOF, J.; WILLS, C. Selection against frameshift mutations limits microsatellite expansion in coding DNA. **Genome Research**, New York, v. 10, p. 72-80, 2000.
- MORAN, C. Microsatellite repeats in pig (*Sus domestica*) and chicken (*Gallus domesticus*) genomes. **Journal of Heredity**, Oxford, v. 84, p. 274-280, 1993.
- MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. **Nature Genetics**, New York, v. 30, p. 194-200, 2002.
- MUNHOZ, C.F. **Identificação de genes de maracujá azedo diferencialmente expressos durante a interação com *Xanthomonas axonopodis***. 2013. 174p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2013.
- MUNHOZ, C. F.; WEISS, B.; HANAI, L. R.; ZUCCHI, M. I.; FUNGARO, M. H. P.; OLIVEIRA, A. L. M.; MONTEIRO-VITORELLO, C. B.; VIEIRA, M. L. C. Genetic diversity and a PCR-based method for *Xanthomonas axonopodis* detection in passion fruit. **Phytopathology**, Saint Paul, v. 101, p. 416-424, 2011.

MURRAY, M.G.; THOMPSON, W.F. Rapid isolation of high molecular weight plant DNA. **Nucleic Acids Research**, Oxford, v. 8, p. 1134-1137, 1980.

NASU, S.; SUZUKI, J.; OHTA, R.; HASEGAWA, K.; YUI, R.; KITAZAWA, N.; MONNA, L.; MINOBE, Y. Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. **DNA Research**, Oxford, v. 9, p. 163-171, 2002.

NORDBORG, M.; HU, T.T.; ISHINO, Y.; JHAVERI, J.; TOOMAJIAN, C.; ZHENG, H.; BAKKER, E.; CALABRESE, P.; GLADSTONE, J.; GOYAL, R.; JAKOBSSON, M.; KIM, S.; MOROZOV, Y.; PADHUKASAHASRAM, B.; PLAGNOL, V.; ROSENBERG, N.A.; SHAH, C.; WALL, J.D.; WANG, J.; ZHAO, K.; KALBFLEISCH, T.; SCHULZ, V.; KREITMAN, M.; BERGELSON, J. The pattern of polymorphism in *Arabidopsis thaliana*: **PLoS Biology**, San Francisco, v. 3, p. 1289-1299, 2005.

NUNES, E.S. **Caracterização fenotípica e molecular de uma população F1 de maracujá-doce visando à construção de mapas de ligação e identificação de QTL**. 2010. 138p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2010.

NYBOM, H.; WEISING, K.; ROTTER, B. DNA fingerprinting in botany: Past, present, future. **Investigative Genetics**, London, v. 5, p. 1-88, 2014.

ODA, K.; MATSUOKA, Y.; FUNAHASHI, A.; KITANO, H. A comprehensive pathway map of epidermal growth factor receptor signaling. **Molecular Systems Biology**, Germany, v. 1, p. 1-36, 2005.

OLIVEIRA, E.J.; PÁDUA, J.G.; ZUCCHI, M.I.; VENCOVSKY, R.; VIEIRA, M.L.C. Origin, evolution and genome distribution of microsatellites. **Genetics and Molecular Biology**, Ribeirão Preto, v. 29, p. 294-307, 2006.

OLIVEIRA, E.J.; VIEIRA, M.L.C.; GARCIA, A.A F.; MUNHOZ, C.F.; MARGARIDO, G.R.A.; CONSOLI, L.; MATTA, F.P.; MORAES, M.C.; ZUCCHI, M.I.; FUNGARO, M. H.P. An integrated molecular map of yellow passion fruit based on simultaneous maximum-likelihood estimation of linkage and linkage phases. **Journal of the American Society for Horticultural Science**, Alexandria, v. 133, p. 35-41, 2008.

OLIVEIRA, G. A. F.; PÁDUA, J. G.; COSTA, J. L.; JESUS, O. N. de; CARVALHO, F. M. de; OLIVEIRA, E. J. de; OLIVEIRA, J. L. Cross-species amplification of microsatellite loci developed for *Passiflora edulis* Sims. in related passiflora species: **Brazilian Archives of Biology and Technology**, Curitiba, v. 56, p. 785-792, 2013.

ORTIZ, D.C.; BOHÓRQUEZ, A.; DUQUE, M.C.; TOHME, J.; CUÉLLAR, D.; MOSQUERA VÁSQUEZ, T. Evaluating purple passion fruit (*Passiflora edulis* Sims f. *edulis*) genetic variability in individuals from commercial plantations in Colombia. **Genetic Resources and Crop Evolution**, Dordrecht, v. 59, p. 1089-1099, 2012.

PAVY, N.; PARSONS, L.S.; PAULE, C.; MACKAY, J.; BOUSQUET, J. Automated SNP detection from a large collection of white spruce expressed sequences: Contributing factors and approaches for the categorization of SNPs. **BMC Genomics**, London, v. 7, p. 1-14, 2006.

PAVY, N.; PELGAS, B.; BEAUSEIGLE, S.; BLAIS, S.; GAGNON, F.; GOSSELIN, I.; LAMOTHE, M.; ISABEL, N.; BOUSQUET, J. Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: Application to the large and unsequenced genomes of white spruce and black spruce. **BMC Genomics**, London, v. 9, p. 1-17, 2008.

PEARSON, C.E.; EDAMURA, K.N.; CLEARY, J.D. Repeat instability: Mechanisms of dynamic mutations. **Nature Reviews Genetics**, London, v. 6, p. 729-742, 2005.

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

PENHA, H.A. **Desenvolvimento de marcadores microssatélites em *Passiflora alata* Curtis**. 2008. 93p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2008.

PENHA, H.A.; PEREIRA, G.D.S.; ZUCCHI, M.I.; DINIZ, A.L.; VIEIRA, M.L.C. Development of microsatellite markers in sweet passion fruit, and identification of length and conformation polymorphisms within repeat sequences. **Plant Breeding**, Berlin, v. 132, p. 731-735, 2013.

PEREIRA, G.S.; NUNES, E.S.; LAPERUTA, L.D.C.; BRAGA, M.F.; PENHA, H.A.; DINIZ, A.L.; MUNHOZ, C.F.; GAZAFFI, R.; GARCIA, A.A.F.; VIEIRA, M.L.C. Molecular polymorphism and linkage analysis in sweet passion fruit, an outcrossing species. **Annals of Applied Biology**, Warwick, v. 162, p. 347-361, 2013.

PINTO, L.R.; VIEIRA, M.L.C.; DE SOUZA JUNIOR, C.L.; SOUZA, A.P. Genetic diversity assessed by microsatellites in tropical maize population submitted to high-density reciprocal recurrent selection. **Euphytica**, Wageningen, v. 134, p. 277-286, 2003.

POCZAI, P.; VARGA, I.; LAOS, M.; CSEH, A.; BELL, N.; VALKONEN, J.P.T.; HYVÖNEN, J. Advances in plant gene-targeted and functional markers: A review. **Plant Methods**, London, v. 9, p. 1-31, 2013.

POMPANON, F.; BONIN, A.; BELLEMAIN, E.; TABERLET, P. Genotyping errors: causes, consequences and solutions. **Nature Reviews Genetics**, London, v. 6, p. 847-859, 2005.

PUGH, T.; FOUET, O.; RISTERUCCI, A.M.; BROTTIER, P.; ABOULADZE, M.; DELETREZ, C.; COURTOIS, B.; CLEMENT, D.; LARMANDE, P.; N'GORAN, J.A.K.; LANAUD, C. A new cacao linkage map based on codominant markers: Development and integration of 201 new microsatellite markers. **Theoretical and Applied Genetics**, Berlin, v. 108, p. 1151-1161, 2004.

RÊGO, M.M.; RÊGO, E.R.; BRUCKNER, C.H.; SILVA, E.A.M. da; FINGER, F.L.; PEREIRA, K.J.C. Pollen tube behavior in yellow passion fruit following compatible and incompatible crosses. **Theoretical and Applied Genetics**, Berlin, v. 101, p. 685-689, 2000.

RITTER, E.; GEBHARDT, C.; SALAMINI, F. Estimation of recombination frequencies and

construction of linkages maps from crosses between heterozygous parents. **Genetics**, Austin, v.125, p. 645-654, 1990.

ROZEN, S.; SKALETISKY, H.J. **Primer3**: pick *primers* from a DNA sequence. 2000. Disponível em: <<http://frodo.wi.mit.edu/primer3/input.htm>>. Acesso em: 10 nov. 2013.

RUDD, S. Expressed sequence tags: alternative or complement to whole genome sequences? **Trends in Plant Science**, Oxford, v. 8, p. 321–329, 2003.

SACHIDANANDAM, R.; WEISSMAN, D.; SCHMIDT, S.C.; KAKOL, J.M.; STEIN, L. D.; MULLIKIN, J.C.; MORTIMORE, B.J.; WILLEY, D.L.; HUNT, S.E.; COLE, C.G.; COGGILL, P.C.; RICE, C.M.; NING, Z.; ROGERS, J.; BENTLEY, D.R.; KWOK, P.Y.; MARDIS, E.R.; YEH, R.T.; SCHULTZ, B.; COOK, L. DAVENPORT, R.; DANTE, M.; FULTON, L.; HILLIER, L.; WATERSTON, R.H.; MCPHERSON, J. D.; GILMAN, B.; SCHAFFNER, S.; VAN ETEN, W.J.; REICH, D.; HIGGINS, J.; DALY, M.J.; BLUMENSTIEL, B.; BALDWIN, J.; STANGE-THOMANN, N.; ZODY, M.C.; LINTON, L.; LANDER, E.S.; ALTSHULER, D. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. **Nature**, New York, v. 409, p. 928–933, 2001.

SANIER, C.; SAYEGH-ALHAMDI, M.; JALLOUL, A.; CLERIVET, A.; NICOLE, M.; MARMEY, P. A 13-lipoxygenase is Expressed Early in the Hypersensitive Reaction of Cotton Plants to *Xanthomonas campestris* pv. *malvacearum*. **Journal of Phytopathology**, Berlin, v. 160, p. 286-293, 2012.

SANTOS, A.A. **Exploração de uma biblioteca genômica de *Passiflora edulis* f. *flavicarpa* por sequenciamento de BAC-ends**. 2013. 141p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2013.

SANTOS FILHO, H.P.; LARANJEIRA, F.F.; SANTOS, C.C.F.; BARBOSA, C.J. Doenças do maracujazeiro. In: LIMA, A.A; CUNHA, M.A.P. (Ed). **Maracujá: produção e qualidade na passicultura**. Cruz das Almas: Embrapa Mandioca e Fruticultura, 2004. p. 239-280.

SCHLOTTERER, C. Evolutionary dynamics of microsatellite DNA. **Chromosoma**, Berlin, v. 109, p. 365-371, 2000.

SCHMID, K. J.; SÖRENSEN, T.R.; STRACKE, R.; TÖRJÉK, O.; ALTMANN, T.; MITCHELL-OLDS, T.; WEISSHAAR, B. Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana*: **Genome Research**, New York, v. 13, p. 1250-1257, 2003.

SCHNEIDER, K.; KULOSA, D.; SOERENSEN, T.R.; MÖHRING, S.; HEINE, M.; DURSTEWITZ, G.; POLLEY, A.; WEBER, E.; JAMSARI; LEIN, J.; HOHMANN, U.; TAHIRO, E.; WEISSHAAR, B.; SCHULZ, B.; KOCH, G.; JUNG, C.; GANAL, M. Analysis of DNA polymorphisms in sugar beet (*Beta vulgaris* L.) and development of an SNP-based map of expressed genes. **Theoretical and Applied Genetics**, Berlin, v. 115, p. 601-615, 2007.

SIA, E.A.; BUTLER, C.A.; DOMINSKA, M.; GREENWELL, P.; FOX, T.D.; PETES, T. D. Analysis of microsatellite mutations in the mitochondrial DNA of *Saccharomyces cerevisiae*. **Proceedings of the National Academy of Sciences**, Washington, v. 97, n. 1, p. 250-255, 2000.

SILVA, P.I.T.; MARTINS, A.M.; GOUVEA, E.G.; PESSOA-FILHO, M.; FERREIRA, M. E. Development and validation of microsatellite markers for *Brachiaria ruziziensis* obtained by partial genome assembly of Illumina single-end reads. **BMC Genomics**, London, v. 14, p. 1-9, 2013.

SPRUYT, M.; BUQUICCHIO, F. **Gene Runner for Windows**, 1994. Disponível em: <<http://www.generunner.net/>>. Acesso em: 12 jun. 2013.

STRAND, M.; PROLLA, T.A.; LISKAY, R.M.; PETES, T.D. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. **Nature**, London, v. 365, p. 274-276, 1993.

SUNYAEV, S.R.; EISENHABER, F.; RODCHENKOV, I.V.; EISENHABER, B.; TUMANYAN, V.G.; KUZNETSOV, E.N. PSIC: profile extraction from sequence alignments with position-specific counts of independent observations. **Protein engineering**, Oxford, v. 12, p. 387-394, 1999.

THIEL, T.; MICHALEK, V.; GRANER, A. Exploiting EST data-bases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). **Theoretical and Applied Genetics**, Berlin, v. 106, p. 411-422, 2003.

TÓTH, G.; GÁSPÁRI, Z.; JURKA, J. Microsatellites in different eukaryotic genomes: survey and analysis. **Genome Research**, New York, v. 10, p. 967-981, 2000.

VAN DEYNZE, A.E.; STOFFEL, K.; BUELL, R.C.; KOZIK, A.; LIU, J.; VAN DER KNAAP, E.; FRANCIS, D. Diversity in conserved genes in tomato. **BMC Genomics**, London, v. 8, p. 1-9, 2007.

VAN, K.; HWANG, E.Y.; KIM, M.Y.; PARK, H.J.; LEE, S.H.; CREGAN, P.B. Discovery of SNPs in soybean genotypes frequently used as the parents of mapping populations in the United States and Korea. **Journal of Heredity**, Oxford, v. 96, p. 529-535, 2005.

VAN LITH, H.A.; VAN ZUTPHEN, L.F. Characterization of rabbit DNA microsatellites extracted from the EMBL nucleotide sequence database. **Animal Genetics**, Oxford, v. 27, p. 387-395, 1996.

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

VARSHNEY, R.K.; GROSSE, I.; HÄHNEL, U.; SIEFKEN, R.; PRASAD, M.; STEIN, N.; LANGRIDGE, P.; ALTSCHMIED, L.; GRANER, A. Genetic mapping and BAC assignment of EST-derived SSR markers shows non-uniform distribution of genes in the barley genome: **Theoretical and Applied Genetics**, Berlin, v. 113, p. 239-250, 2006.

VARSHNEY, R.K.; KORZUN, V.; BÖRNER, A. Molecular maps in cereals: Methodology and progress. In: GUPTA, P.K.; VARSHNEY, R.K. (Ed.). **Cereal genomics**. The Netherlands: Kluwer Academic Publishers, 2004. p 35-82.

VARSHNEY, R.K.; THIEL, T.; LANGRIDGE, N.S.P.; GRANER, A. In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. **Cellular and Molecular Biology Letters**, Wrocław, v. 7, p. 537-546, 2002.

VIANA, F. M. P.; FREIRE, F. C. O.; CARDOSO, J. E.; VIDAL, J. C. **Principais doenças do maracujazeiro na Região Nordeste e seu controle**. Fortaleza: Embrapa Agroindústria Tropical, 2003. 12 p. (Embrapa Agroindústria Tropical. Comunicado técnico, 86).

VOET, D.; VOET, J. G. **Bioquímica**. 4. ed. Porto Alegre: Artmed, 2013. p 66.

WANG, Z.; LI, J.; LUO, Z.; HUANG, L.; CHEN, X.; FANG, B.; LI, Y.; CHEN, J.; ZHANG, X. Characterization and development of EST-derived SSR markers in cultivated sweetpotato (*Ipomoea batatas*). **BMC Plant Biology**, London, v. 11, p. 1471-2229, 2011.

WU, R.; MA, C.-X.; PAINTER, I.; ZENG, Z.-B. Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. **Theoretical Population Biology**, New York, v. 61, p. 349-363, 2002.

YANG, H.; TAO, Y.; ZHENG, Z.; LI, C.; SWEETINGHAM, M.W.; HOWIESON, J.G. Application of next-generation sequencing for rapid marker development in molecular plant breeding: a case study on anthracnose disease resistance in *Lupinus angustifolius* L. **BMC Genomics**, London, v. 13, p 1-11, 2012.

YU, J.; WANG, J.; LIN, W.; LI, S.; LI, H.; ZHOU, J.; NI, P.; DONG, W.; HU, S.; ZENG, C.; ZHANG, J.; ZHANG, Y.; LI, R.; XU, Z.; LI, X.; ZHENG, H.; CONG, L.; LIN, L.; YIN, J.; GENG, J.; LI, G.; SHI, J.; LIU, J.; LV, H.; LI, J.; DENG, Y.; RAN, L.; SHI, X.; WANG, X.; WU, Q.; LI, C.; REN, X.; LI, D.; LIU, D.; ZHANG, X.; JI, Z.; ZHAO, W.; SUN, Y.; ZHANG, Z.; BAO, J.; HAN, Y.; DONG, L.; JI, J.; CHEN, P.; WU, S.; XIAO, Y.; BU, D.; TAN, J.; YANG, L.; YE, C.; XU, J.; ZHOU, Y.; YU, Y.; ZHANG, B.; ZHUANG, S.; WEI, H.; LIU, B.; LEI, M.; YU, H.; LI, Y.; XU, H.; WEI, S.; HE, X.; FANG, L.; HUANG, X.; SU, Z.; TONG, W.; TONG, Z.; YE, J.; WANG, L.; LEI, T.; CHEN, C.; CHEN, H.; HUANG, H.; ZHANG, F.; LI, N.; ZHAO, C.; HUANG, Y.; LI, L.; XI, Y.; QI, Q.; LI, W.; HU, W.; TIAN, X.; JIAO, Y.; LIANG, X.; JIN, J.; GAO, L.; ZHENG, W.; HAO, B.; LIU, S.; WANG, W.; YUAN, L.; CAO, M.; MCDERMOTT, J.; SAMUDRALA, R.; WONG, G. K. S.; YANG, H. The genomes of *Oryza sativa*: A history of duplications. **PLoS Biology**, San Francisco, v. 3, p. 266-281, 2005.

ZHANG, H.; WEI, L.; MIAO, H.; ZHANG, T.; WANG, C. Development and validation of genic-SSR markers in sesame by RNA-seq. **BMC Genomics**, London, v. 13, p. 1-10, 2012.

ZHENG, X.; PAN, C.; DIAO, Y.; YOU, Y.; YANG, C.; HU, Z. Development of microsatellite markers by transcriptome sequencing in two species of *Amorpha* (Araceae). **BMC Genomics**, London, v. 14, p. 1-11, 2013.

ZHU, Y. L.; SONG, Q.J.; HYTEN, D.L.; VAN TASSELL, C.P.; MATUKUMALLI, L.K.; GRIMM, D.R.; HYATT, S.M.; FICKUS, E.W.; YOUNG, N.D.; CREGAN, P.B. Single-nucleotide polymorphisms in soybean. **Genetics**, Berlin, v. 163, p. 1123-1134, 2003.

ZUCCHI, M.I.; BRONDANI, R.V.; PINHEIRO, J.B.; VENCOVSKY, R. Transferability of microsatellite markers from *Eucalyptus* spp. to *Eugenia dysenterica* (Myrtaceae). **Molecular Ecology Notes**, Oxford, v. 2, p. 512-514, 2002.

ANEXOS

ANEXO A

Figura A1 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaA6P9** (1º conjunto) e **PaA12P5F** (2º conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaA6P9 há um SNP na posição 532 e no fragmento gênico PaA12P5F na posição 127

Figura A2 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaA12P5F** (1º conjunto) e **PaA12P5R** (2º conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaA12P5F há um SNP na posição 235 e no fragmento gênico PaA12P5R na posição 160

Figura A3 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F1-67 e F1-100 relativos aos fragmentos gênicos **PaA12P5R** (1º conjunto) e **PaB8P9** (2º conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaA12P5R há um SNP na posição 169 e no fragmento gênico PaB8P9 na posição 333

Figura A4 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaB8P9** (1º e 2º conjuntos) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No 1º conjunto há um SNP na posição 342 e no 2º conjunto na posição 344

Figura A5 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaB8P9** (1º conjunto) e **PaC11P10** (2º conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaB8P9 há um SNP na posição 347 e no fragmento gênico PaC11P10 na posição 44

Figura A6 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaContig15** (1º conjunto) e **PaContig21** (2º conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaContig15 há um SNP na posição 158 e no fragmento gênico PaContig21 na posição 309

Figura A7 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaContig21** (1º conjunto) e **PaContig22R** (2º conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaContig21 há um SNP na posição 323 e no fragmento gênico PaContig22R na posição 108

Figura A8 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos ao fragmento gênico **PaContig22R** (1° e 2° conjuntos) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No 1° conjunto há um SNP na posição 336 e no 2° conjunto na posição 364

Figura A9 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos ao fragmento gênico **PaContig53** (1° e 2° conjuntos) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No 1° conjunto há um SNP na posição 104 e no 2° conjunto na posição 303

Figura A11 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaContig53** (1º conjunto) e **PaD5P10** (2º conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaContig53 há um SNP na posição 349 e no fragmento gênico PaD5P10 na posição 181

Figura A12 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos ao fragmento gênico **PaD5P10** (1° e 2° conjuntos) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No 1° conjunto há um SNP na posição 350 e no 2° conjunto na posição 371

Figura A13 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaD5P10** (1º conjunto) e **PaE1P7** (2º conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaD5P10 há um SNP na posição 381 e no fragmento gênico PaE1P7 na posição 212

Figura A14 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos ao fragmento gênico **PaF1P8** (1° e 2° conjuntos) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No 1° conjunto há um SNP na posição 163 e no 2° conjunto na posição 246

Figura A15 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos ao fragmento gênico **PaF1P8** (1° e 2° conjuntos) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No 1° conjunto há um SNP na posição 247 e no 2° conjunto na posição 287

Figura A16 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos aos fragmentos gênicos **PaG12P5** (1° conjunto) e **PaH12P9** (2° conjunto) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No fragmento gênico PaG12P5 há um SNP na posição 264 e no fragmento gênico PaH12P9 na posição 286

Figura A17 - Sequências e eletroferogramas dos acessos '2(12)', 'SV3', F₁-67 e F₁-100 relativos ao fragmento gênico **PaP3RA2** (1° e 2° conjuntos) de *P. alata* obtidos a partir de transcritos de *P. edulis*. No 1° conjunto há um SNP na posição 186 e no 2° conjunto na posição 225