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

**Desenvolvimento e uso de marcadores microssatélites para construção e
integração de mapas genéticos de maracujá-amarelo (*Passiflora edulis* Sims f.
flavicarpa Deg.)**

Eder Jorge de Oliveira

Tese apresentada para obtenção do título de Doutor em
Agronomia, Área de Concentração: Genética e Melhoramento
de Plantas

PIRACICABA

2006

Eder Jorge de Oliveira

Engenheiro Agrônomo

Desenvolvimento e uso de marcadores microssatélites para construção e integração de mapas genéticos de maracujá-amarelo (*Passiflora edulis* Sims f. *flavicarpa* Deg.)

Orientadora:

Prof^ª Dr^ª **MARIA LUCIA CARNEIRO VIEIRA**

Co-orientador:

Prof. Dr. **ANTONIO AUGUSTO FRANCO GARCIA**

Tese apresentada para obtenção do título de Doutor em Agronomia, Área de Concentração: Genética e Melhoramento de Plantas

PIRACICABA

2006

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

Oliveira, Eder Jorge de

Desenvolvimento e uso de marcadores microssatélites para construção e integração de mapas genéticos de maracujá-amarelo (*Passiflora edulis* Sims f. *flavicarpa* Deg.) / Eder Jorge de Oliveira. - - Piracicaba, 2006.
152 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2006.
Bibliografia.

1. Genética molecular vegetal 2. Ligação gênica 3. Mapeamento genético
4. Maracujá-amarelo 5. Marcador molecular I. Título

CDD 634.425

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

DEDICATÓRIA

Dedico à memória de meu pai, Adahil, e à presença de minha mãe Maria, pela educação ofertada, pelo amor dedicado e pela liberdade concedida. Cada um à sua maneira soube trilhar meus caminhos nessa longa jornada da vida.

À Gilmara, minha esposa, que sempre soube compreender e administrar os momentos difíceis que passamos afastados durante a realização do doutorado. Pelo carinho, amor e paciência.

AGRADECIMENTOS

À Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP) e ao Departamento de Genética por propiciarem um bom ambiente de trabalho e geração de conhecimento;

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa concedida;

À Profa. Dra. Maria Lucia Carneiro Vieira que me orientou na execução deste trabalho e em minha formação profissional;

Ao Prof. Dr. Antonio Augusto Franco Garcia, pelo companheirismo, entusiasmo e prontidão nos momentos de dúvidas;

Ao Prof. Dr. Luís Eduardo Aranha Camargo, por colocar o laboratório à disposição no início dos trabalhos de seqüenciamento da biblioteca genômica enriquecida;

À Profa. Dra. Maria Helena Pelegrinelli Fungaro e ao Dr. Laurival Antonio Vilas-Boas da Universidade Estadual de Londrina (UEL) pelo valioso apoio no seqüenciamento da biblioteca genômica enriquecida;

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

À Dra. Maria I. Zucchi do Instituto Agrônomo (IAC) pela valiosa contribuição na construção da biblioteca genômica enriquecida com microssatélites;

Ao Dr. Luciano Consoli da ESALQ/USP, pela grande contribuição na leitura e interpretação das seqüências de DNA, bem como na genotipagem da população segregante;

À bióloga Carla de Freitas Munhoz, pelo grande apoio nos trabalhos de laboratório, sobretudo na genotipagem da população segregante;

Ao graduando Gabriel A. Margarido, pelos conhecimentos compartilhados na programação e utilização do programa *R*;

Ao técnico de laboratório Carlos Alberto de Oliveira, pelo companheirismo e suporte técnico oferecido durante a realização deste trabalho;

A todos os funcionários do Departamento de Genética pelo convívio amigável e pela colaboração na execução dos experimentos;

Aos amigos do Laboratório de Biologia Celular e Molecular de Plantas, Juliano, Mariza, Carla, Michel, Maria Rita, XV, Frederico, Adriano, Hanai, Luciana, Sheila e Renato;

Aos meus amigos da República, Tassiano, Ebert P.O. Flor, Fernando Cárdenaz, Caio, Frederico, Emiliano, André e Douglas;

Aos meus grandes amigos do G8, Tassiano, Farias, Karem, Sybele, Ana, Carol e Juliano, que sempre proporcionaram momentos de alegria e descontração;

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

SUMÁRIO

RESUMO.....	8
ABSTRACT.....	10
1 INTRODUÇÃO.....	12
2 DESENVOLVIMENTO.....	15
2.1 O gênero <i>Passiflora</i>	15
2.2 A cultura do maracujazeiro: aspectos fitotécnicos e de melhoramento genético no Brasil..	17
2.3 Marcadores moleculares e mapas genéticos.....	20
2.4 Marcadores microssatélites: aspectos gerais e metodologias de obtenção.....	26
2.5 Construção de mapas genéticos em populações F ₁ com ênfase em <i>P. edulis</i> f. <i>flavicarpa</i> ...	31
3 Material e métodos.....	41
3.1 Material genético.....	41
3.2 Extração e quantificação do DNA.....	41
3.3 Construção da biblioteca genômica enriquecida com microssatélites.....	42
3.4 Otimização e genotipagem dos genitores da população F ₁ com os marcadores microssatélites.....	46
3.5 Genotipagem com os marcadores AFLP.....	48
3.6 Construção dos mapas de ligação.....	50
3.6.1 Tipos de segregações com base em marcas dominantes e codominantes.....	51
3.6.2 Análises de distorções da segregação esperada.....	53
3.6.3 Codificação dos cromossomos dos genitores.....	54
3.6.4 Algoritmos utilizados na determinação das fases de ligação e frequência de recombinação.....	55
3.6.5 Formação dos grupos de ligação e ordenação das marcas.....	60
3.6.6 Estimativas multiponto.....	61
4 Resultados e discussão.....	63
4.1 Construção da biblioteca genômica enriquecida com microssatélites.....	63
4.1.1 Correção das seqüências e redundância da biblioteca.....	64
4.1.2 Enriquecimento da biblioteca genômica com microssatélites.....	68
4.1.2.1 Motivos e número de repetições dos microssatélites.....	70

4.1.2.2 Eficiência de enriquecimento da biblioteca e desenho dos <i>primers</i>	75
4.1.3 Otimização das condições de PCR para detecção e análise do polimorfismo.....	76
4.2 Análise de ligação.....	88
4.2.1 Segregação dos marcadores microssatélites na população F ₁	88
4.2.2 Análise das segregações dos marcadores AFLP.....	96
4.2.3 Distorções de segregação.....	96
4.2.4 Construção dos mapas de ligação	98
4.2.4.1 Comparações com mapas de maracujá-amarelo previamente construídos.....	103
5. CONCLUSÕES.....	120
REFERÊNCIAS.....	121

RESUMO

Desenvolvimento e uso de marcadores microsatélites para construção e integração de mapas genéticos de maracujá-amarelo (*Passiflora edulis* Sims f. *flavicarpa* Deg.)

O maracujá-amarelo é uma espécie alógama, auto-incompatível, o que prejudica a geração de linhagens endogâmicas e, por consequência, a construção e integração de mapas de ligação pelas metodologias convencionais. A exemplo do que tem sido feito em outras espécies vegetais, populações F_1 (com diferentes tipos de segregação) foram utilizadas para a construção dos mapas genéticos de maracujá. Mapas individuais para cada genitor do cruzamento foram gerados fazendo uso apenas de marcas com segregação 1:1. A integração desses mapas é possível com base em marcadores bi-parentais, quando ambos os genitores são heterozigóticos para os mesmos alelos que segregam na proporção 3:1 em F_1 . Apesar disso, o uso de marcadores codominantes e multialélicos, como os microsatélites, possuem maior valor, uma vez que permite a obtenção de estimativas de frequência de recombinação e da fase de ligação com um menor viés. O objetivo deste estudo foi o desenvolvimento de marcadores microsatélites utilizando bibliotecas genômicas enriquecidas, visando à construção e a integração de mapas genéticos de maracujá-amarelo. Foram usadas 160 plantas oriundas do cruzamento entre os acessos IAPAR-123 x IAPAR-06, marcadores AFLP com segregação 1:1 e 3:1 e marcadores microsatélites desenvolvidos no presente estudo. Para a construção dos mapas utilizou-se um algoritmo que estima, simultaneamente, a fase de ligação e a frequência de recombinação. A biblioteca genômica enriquecida forneceu 107 clones contendo microsatélites, com uma eficiência de enriquecimento da ordem de 11%. Todas as seqüências contendo microsatélites foram analisadas e 26 locos mostraram-se polimórficos: 16 deles segregaram na proporção de 1:1; 1 na proporção de 3:1; 2 na proporção de 1:2:1 e 7 na proporção de 1:1:1:1. Com relação aos dados de AFLP, 253 locos mono-parentais (em 114 deles, o alelo dominante estava presente no genitor 06 e 139, no genitor 123) mostraram-se polimórficos com segregação 1:1. Outros 116 locos mostraram-se bi-parentais, segregando na proporção de 3:1. Foram obtidos 11 grupos de ligação (GL) sendo 8 grupos integrados e um GL com apenas dois marcadores. Também, um GL para cada genitor foi obtido, cuja integração não foi possível por falta de marcadores bi-parentais. Foi possível alocar no *framework* dos mapas genéticos, além dos microsatélites, as marcas de AFLP

com segregação do tipo 3:1, que contribuíram para uma maior saturação e integração de oito dos nove grupos de ligação de maracujá-amarelo ($2n = 18$). Foram obtidos, em média, 25 marcadores por GL, que forneceram um comprimento de mapa de 1.765,6 cM. Este é o primeiro relato em *Passiflora* sobre o uso de marcadores microssatélites e do algoritmo proposto por Wu et al. (2002) visando à integração dos mapas anteriormente construídos com base na estratégia duplo pseudocruzamento teste. Este mapa integrado será útil na alocação de QTL (*Quantitative Trait Loci*) relacionados à resistência a doenças e a características de interesse agrônomo, assim como para estudos genéticos e evolutivos.

Palavras-chave: *Passiflora edulis* f. *flavicarpa*; maracujá-amarelo; ligação gênica; mapas genéticos; marcadores microssatélites; AFLP; integração de mapas

ABSTRACT

Development of microsatellite markers and their use for the generation and integration of genetic maps of yellow passion fruit (*Passiflora edulis* Sims f. *flavicarpa* Deg.)

The yellow passion fruit is an out-breeding, self-incompatible species, but those features impose severe constraints on the generation of inbred lines, consequently impairing the building and integration of linkage maps using conventional methodologies. Similar to the reports from other plant species, F₁ populations (with distinct segregation modes) were used to build passion fruit genetic maps. Individual maps, one for each parent involved in the cross were generated solely employing genetic markers with 1:1 segregation. However, the integration of those maps is feasible using bi-parental markers, i.e. when both parents are heterozygous for the same alleles that show 3:1 segregation in F₁. However, the use of co-dominant multi-allele markers, such as the microsatellites is even more important, since it allows recombination frequency and linkage phase estimates to be obtained less effected by bias. The goal of this research was the development of microsatellite markers using enriched genomic libraries for constructing and integrating genetic maps of yellow passion fruit. We have used 160 plants from a cross between the accessions IAPAR-123 x IAPAR-06, AFLP markers showing 1:1 and 3:1 segregations and microsatellite markers developed in the present work. The maps were built employing an algorithm that simultaneously estimates the linkage phase and the recombination frequency. The enriched library allowed us to obtain 107 microsatellite-containing clones, at an enrichment efficiency of 11%. All microsatellite-containing sequences were analyzed and 26 loci were shown to be polymorphic: 16 of them with a 1:1 segregation; 1 with a 3:1 segregation, 2 showing a 1:2:1 proportion and 7 with a 1:1:1:1 segregation pattern. In relation to the AFLP data, 253 mono-parental loci were shown to be polymorphic with a 1:1 segregation proportion (in 114 loci the dominant allele was present in the parent 06, and in 139 loci, the dominant allele was found in the parent 123). The remaining 116 loci were shown to be bi-parental, segregating in a 3:1 proportion. Eleven linkage groups (LG) were obtained, 8 of them being integrated and one LG with only two markers. Moreover, one LG for each parent was obtained, but due to the absence of bi-parental markers they were not integrated. It was possible to locate on the framework of the genetic maps, along with the microsatellites those AFLP markers showing a 3:1-segregation type,

which contributed to a greater level of map saturation and to the integration of eight of the nine linkage groups of yellow passion fruit ($2n=18$). On average, we have obtained 25 markers per LG, providing a map distance of 1,765.6 cM. This is the first report in *Passiflora* on the use of microsatellite markers and the algorithm proposed by Wu et al. (2002) to integrate maps previously constructed using the double pseudo-testcross strategy. The integrated map will be useful for locating QTL (*Quantitative Trait Loci*) related to disease resistance and agriculturally interesting traits, as well as to provide tools for genetic and evolutionary studies.

Keywords: *Passiflora edulis* f. *flavicarpa*; yellow passion fruit; linkage; genetic maps; microsatellite markers; AFLP; map integration

1 INTRODUÇÃO

O maracujá-amarelo (*Passiflora edulis* Sims f. *flavicarpa* Deg.) é a espécie mais importante do gênero *Passiflora*, sendo o Brasil o centro de diversidade do gênero. Devido à qualidade e ao valor comercial de seus frutos, o maracujá-amarelo está presente em 95% dos pomares comerciais do Brasil.

As estatísticas de produção da cultura são relativamente expressivas, sendo que em 2003 a área plantada foi de 35.000 ha e a produção nacional da ordem de 480 mil toneladas (FNP CONSULTORIA & COMÉRCIO, 2006). Nossa produção média está entre 10 a 18 t/ha/ano, que é muito baixa comparada à produção de até 65 t/ha/ano quando se dispõe de tecnologia e manejo adequados e de variedades melhoradas (MELETTI, SANTOS e MINAMI, 2000).

A espécie apresenta ampla variabilidade natural para ser explorada no melhoramento genético. No entanto, ainda é preciso conhecer e resolver muitos dos problemas enfrentados pela cultura, dentre eles: (i) a insuficiência de estudos básicos na área de genética que fundamentam o melhoramento e, por conseqüente, (ii) a falta de variedades homogêneas e produtivas, tolerantes às principais pragas e doenças, (iii) o uso inadequado de práticas culturais, e a (iii) ausência de tecnologias de produção mais adequadas aos diferentes mercados.

Em termos de genética molecular, é necessário a geração de informações sobre conteúdo de DNA, tamanho do genoma, isolamento e clonagem de genes que controlam caracteres de interesse agrônômico, mecanismo genético que leva à incompatibilidade, mapas de ligação saturados para a detecção e localização de QTL (*Quantitative Trait Loci*).

Os avanços das Biologia Molecular e da Genômica, sobretudo o desenvolvimento dos marcadores de DNA, têm permitido estudar e caracterizar a variabilidade genética em diferentes níveis. O desenvolvimento de mapas genéticos, por exemplo, é uma das aplicações de maior impacto da tecnologia de marcadores moleculares, tornando possível localizar genes que controlam caracteres cuja expressão fenotípica depende da ação de um conjunto de locos, ou QTL (CARNEIRO e VIEIRA, 2002).

Para a construção de mapas genéticos de espécies não endogâmicas, a estratégia denominada duplo pseudocruzamento teste (do inglês, *two-way pseudo-testcross*) tem sido utilizada em virtude da forte depressão endogâmica e dos problemas relativos à auto-incompatibilidade que dificultam a obtenção de linhagens e, por conseguinte, de populações

convencionais de mapeamento (GRATTAPAGLIA e SEDEROFF, 1994; HEMMAT et al., 1994; CROUZILLAT et al., 1996; MARQUES et al., 1998; CERVERA et al., 2001).

A estratégia duplo pseudocruzamento teste baseia-se na utilização de populações F_1 e de locos moleculares que segregam 1:1, como em um retrocruzamento, e têm-se alelos que segregam ora na meiose de um, ora do outro genitor ($Aa \times aa$ ou $aa \times Aa$). Por isso, são construídos mapas individuais, um para cada genitor. Baseado nesta metodologia, Carneiro et al. (2002) construíram o primeiro mapa genético de *Passiflora* utilizando marcadores dominantes RAPD. A população de mapeamento foi composta por 90 plantas F_1 derivadas do cruzamento simples entre os acessos IAPAR-123 (genitor feminino) e IAPAR-06 (genitor masculino). Outro mapa, utilizando 117 plantas desta mesma população, foi construído com base em marcadores AFLP (LOPES et al., 2006).

Contudo, o uso de marcadores ponte, ou seja, de locos cujos alelos segregam em ambos os genitores, permite que essa estratégia seja aperfeiçoada, pois possibilita a união dos mapas genéticos dos genitores, implicando na integração dos mapas individuais. Assim, marcadores dominantes com segregação 3:1 e codominantes com segregação 1:2:1 e 1:1:1:1 podem ser utilizados como ponte. Porém, nesta situação, cada loco pode apresentar diferentes números de alelos com segregações típicas, o que dificulta as análises de ligação já que as fases de ligação dos pares de marcas de cada genitor não são conhecidas a princípio, tornando difícil a detecção dos eventos de recombinação (MALIEPAARD, JANSEN e VAN OOIJEN, 1997; MALIEPAARD et al., 1998; WU et al., 2002).

Apesar dos marcadores com segregação do tipo 3:1 serem úteis na integração dos mapas de ligação individuais, quanto mais informativos forem os marcadores, melhores serão as estimativas das fases de ligação e da frequência de recombinação e, conseqüentemente, mais confiável será a integração dos mapas. Marcadores com segregação informativa são aqueles que conseguem distinguir o maior número de genótipos com base no fenótipo molecular, sendo que os tipos que revelam segregações do tipo 1:1:1:1 e 1:2:1 são os mais informativas, nesta ordem (WU et al., 2002). Estes tipos de configurações só podem ser revelados com o uso de marcadores codominantes, como os microsatélites, que são especialmente importantes, pois são multialélicos e baseados na PCR, demandando menos trabalho no laboratório e custo menor de implementação (BRONDANI et al., 1998).

Os microssatélites podem ser identificados por um *screening* na base de dados de DNA (*Genbank*) da espécie a ser estudada. Contudo, não há registro de descrição e depósito de seqüências de microssatélites para o gênero *Passiflora*, havendo a necessidade de construção e *screening* de uma biblioteca genômica enriquecida, seqüenciamento de clones positivos para microssatélites, desenho de *primers*, seleção dos locos informativos e caracterização da variação alélica.

A hipótese na qual se fundamenta esta pesquisa é que seria possível a construção de um mapa de ligação de maracujá-amarelo, integrado com base em locos ponte e usando-se o método estatístico proposto por Wu et al. (2002), que estima, simultaneamente, a fase de ligação e a frequência de recombinação por meio da análise de verossimilhança. Este seria, porquanto, o primeiro estudo a empregar esta metodologia para análise de dados em espécies frutíferas tropicais não endogâmicas.

Sendo assim, foram perseguidos os seguintes objetivos:

1. a construção de bibliotecas genômicas de maracujá-amarelo enriquecidas com seqüências microssatélites e posterior desenvolvimento destes marcadores genéticos;
2. a construção e integração dos mapas de ligação, utilizando marcadores AFLP e microssatélites, para a mesma população segregante oriunda do cruzamento entre os genitores IAPAR-123 e IAPAR-06.

2 DESENVOLVIMENTO

2.1 O gênero *Passiflora*

Segundo Vanderplank (1996) o maracujazeiro pertence à família *Passifloraceae* que compreende 18 gêneros e cerca de 630 espécies, distribuídas principalmente nas regiões tropicais da América, Ásia e África. O gênero *Passiflora*, originário da América do Sul e cujo centro de distribuição geográfica é o centro-norte do Brasil, possui o maior número de espécies da família, cerca de 465 (VANDERPLANK, 1996). Entretanto, apenas em torno de 60 espécies produzem frutos com valor comercial (SHULTZ, 1968; MANICA, 1997), seja para consumo ou uso medicinal.

Grande parte dessas espécies está dispersa no território nacional, o que confere ao nosso País a condição de um dos principais centros de diversidade genética do gênero (d'EECKENBRUGGE, 2003).

Como principal uso das espécies do gênero *Passiflora*, tem-se a alimentação humana, na forma de sucos, doces, sorvetes, licores e consumo *in natura* da fruta. Entretanto, existe também um forte apelo ornamental em função dos diferentes tipos e formas de flores, folhas e caules, além do valor medicinal condicionado pelas propriedades calmantes da passiflorina encontrada nas flores e folhas.

Dentre as principais espécies produtoras de frutos comestíveis estão: *Passiflora edulis* Sims f. *flavicarpa* Deg. (maracujá-amarelo), *P. edulis* f. *edulis* Sims (maracujá-roxo), *P. alata* Dryand (maracujá doce), *P. ligularis* Juss., *P. nítida* (Kunth.), *P. molinissima* (HBK) Bailey, *P. maliformis* L., *P. caerulea*, *P. laurifolia* Linn. e *P. quadrangularis*. Vieira e Carneiro (2004) fazem uma revisão a respeito do gênero *Passiflora* e apresentam uma compilação das espécies comerciais, bem como suas origens e usos.

O maracujazeiro é uma planta essencialmente alógama, condicionada pelo fenômeno da auto-incompatibilidade, que impede a autofecundação e até mesmo o cruzamento de diferentes plantas com os mesmos alelos de incompatibilidade. Neste caso, a autofecundação leva à perda de variabilidade alélica nos locos de incompatibilidade, o que dificulta a obtenção de linhagens.

Segundo Bruckner et al. (1995) a auto-incompatibilidade do maracujazeiro é do tipo esporofítica. Entretanto, o mesmo grupo de pesquisadores relata que existe um gene de efeito gametofítico associado ao sistema esporofítico (SUASSUNA et al., 2003). As proteínas responsáveis pela auto-incompatibilidade são encontradas no estigma e na parte superior do estilete a partir de dois dias antes da antese, aumentando gradativamente até esse dia (RÊGO, 1997).

Devido ao fenômeno da auto-incompatibilidade, a fertilização do maracujazeiro requer a presença de diferentes genótipos (RUGGIERO, 1987; BRUCKNER et al., 1995), de insetos polinizadores (AKAMINE e GIROLAMI, 1959; CAMILLO, 1987) ou de polinização manual (GRISI Jr., 1973). Graças ao seu tamanho, as mamangavas (*Xylocopa* spp.) são os principais agentes polinizadores, uma vez que podem fazer a coleta de pólen e a polinização ao mesmo tempo (MELETTI, 2003). Outros insetos somente coletam o pólen, não conseguindo promover a polinização cruzada, e, por conseguinte, a geração de frutos. A polinização realizada pelos ventos é praticamente nula, tendo em vista que o pólen do maracujá é bastante pegajoso e pesado (RUGGIERO, LAM-SANCHEZ e BANZATTO, 1976).

O número de cromossomos do gênero é conhecido para muitas espécies, sendo que *P. edulis* possui $2n=18$ cromossomos. Dentro das espécies com este número de cromossomos, é possível a obtenção de cruzamentos interespecíficos viáveis. Contudo, Oliveira (1996) relata que vários híbridos interespecíficos de *P. edulis* x *P. giberti* são estéreis. Algumas espécies possuem $2n=12$, $2n=14$ (*P. holosericea* L.), $2n=20$ e 22 (*P. foetida* L.), $2n=24$ e 36 (*P. suberosa* L.) e $2n=84$ (*P. lutea* L.). Todavia, as principais espécies comerciais de maracujá possuem $2n=18$ (para revisão veja MELO, CERVI e GUERRA, 2001; CUCO et al., 2005).

Estudos sobre o conteúdo de DNA também estão disponíveis para algumas espécies do gênero. Souza et al. (2004), trabalhando com citometria de fluxo, mostraram que o conteúdo $2C$ de DNA em picogramas (pg), variou bastante entre as espécies estudadas. O valor de 1,83 foi encontrado para *P. suberosa*; 3,16 para *P. edulis* f. *edulis*; 3,19 para *P. edulis* f. *flavicarpa* de origem brasileira; 3,21 para *P. edulis* f. *flavicarpa* de origem mexicana; 3,40 para *P. mucronata*; 3,43 para *P. edmundoi*; 3,88 para *P. laurifolia*; 3,92 para *P. giberti* e 5,36 para *P. quadrangularis*, sendo este o maior valor, cerca de 190% maior que o de *P. suberosa*.

2.2 A cultura do maracujazeiro: aspectos fitotécnicos e de melhoramento genético no Brasil

Apesar de existirem várias espécies com frutos comestíveis, o maracujá-amarelo e o maracujá-roxo são os mais difundidos e cultivados no Brasil e no mundo. A planta de maracujá-amarelo assemelha-se muito ao maracujá-roxo, tendo como diferença básica a pigmentação da casca dos frutos por ocasião da maturação, o sabor e a resistência a doenças (SILVA e SÃO JOSÉ, 1994; BRUCKNER, 1997). Inclusive ocorre o cruzamento entre as duas formas. O híbrido chega a produzir plantas férteis com características intermediárias de coloração dos frutos.

O maracujá-amarelo é o mais utilizado em plantios comerciais no Brasil, abrangendo cerca de 95% dos pomares, devido à qualidade dos seus frutos, vigor e produtividade (RUGGIERO et al., 1996; MELETTI e MAIA, 1999).

Nos últimos anos, a produção brasileira de maracujá-amarelo tem sido destinada em partes relativamente iguais entre a indústria de suco concentrado e ao comércio de frutas *in natura* (ROSSI, 1998). Esta proporção está sujeita a alterações, uma vez que a disputa entre esses dois segmentos de mercado leva os produtores a optar pelo preço no curto prazo, em detrimento da constância da oferta no longo prazo. O grande mercado para o suco de maracujá brasileiro tem sido a Comunidade Européia, que consumiu cerca de 60% das exportações no ano de 1996. Os Estados Unidos, o Canadá e o Japão são mercados altamente promissores (JÚNIOR, ESTALISLAU e PAIVA, 2000). Dados mais recentes sobre as exportações brasileiras de suco de maracujá são escassos já que após 1996 as exportações deste suco não foram mais computadas separadamente, tendo sido inseridas na categoria de sucos em geral (MATTA, 2005). Contudo, o mercado internacional de maracujá é considerado emergente, necessitando apenas de uma garantia de continuidade de fornecimento ao longo dos anos (CARRARO e CUNHA, 1994; RUGGIERO et al., 1996).

Apesar da expressão em nível nacional, a cultura ainda enfrenta grandes problemas em termos de produção e comercialização. Do ponto de vista do sistema de produção, a falta de variedades melhoradas constitui um entrave à produção de frutos homogêneos e de qualidade, haja vista que são utilizados para plantio, mudas de variedades locais sem garantia de origem. Estas mudas são de sementes coletadas em plantios comerciais de polinização aberta, onde existe uma grande heterogeneidade entre plantas com relação à qualidade de frutos, potencial produtivo e resistência a pragas e a doenças (STENZEL e SERA, 1999).

Com o uso desse tipo de material propagativo nos plantios comerciais, a produção e, por conseguinte, a comercialização fica comprometida, uma vez que são produzidos frutos desuniformes em termos de tamanho, peso, coloração, acidez, cor do suco, dentre outras características. Isso obriga o produtor a fazer uma seleção de seus frutos para alcançar melhores preços, o que muita vezes onera em muito o custo de produção. O uso de um conjunto de clones de plantas geneticamente superiores poderia amenizar este problema, porém isso não é utilizado pelos produtores em virtude do custo e, também, por problemas relacionados à contaminação por vírus, que podem afetar a qualidade e a sanidade das mudas assim obtidas.

Apesar da utilização de material propagativo de baixa qualidade, existem variedades disponíveis com potencial de produção bastante elevado, em torno de 65 t/ha em condições experimentais, segundo Meletti, Santos e Minami (2000).

O uso de sementes ou mudas produzidas sem controle fitossanitário podem agravar a incidência de pragas e doenças, diminuir em muito o potencial produtivo e onerar os custos de produção, fazendo com que muitos pomares sejam abandonados após o primeiro ano de colheita, obrigando os produtores a buscarem novas áreas de plantio. De acordo com Meletti (2003), esses fatos favoreceram a formação dos ciclos de crescimento e declínio no Brasil, tornando a cultura itinerante.

Todos esses fatores têm prejudicado uma maior expansão da cultura no Brasil ao longo da última década. Em 1992 a produção nacional foi de cerca de 420.000 t/ano em uma área plantada de 32.000 ha/ano, chegando a 485.000 t, plantados em 35.000 ha em 2003 (FNP CONSULTORIA & COMÉRCIO, 2006).

Medidas que visem ao desenvolvimento de variedades com bom potencial de produção, que possuam uniformidade de frutos, menor espessura de casca, alto teor de sólidos solúveis e resistência/tolerância às principais pragas e doenças, poderão contribuir, certamente, para garantir uma oferta constante do produto nos mercados e maior interesse pela cultura por parte dos vários setores do agronegócio brasileiro.

Vários são os métodos de melhoramento aplicáveis ao maracujazeiro, objetivando o aumento da frequência de alelos favoráveis ou a exploração do vigor híbrido ou heterose (MELETTI, 1998; MELETTI, SANTOS e MINAMI, 2000). Neste caso, deve-se formar uma população-base com genótipos superiores para diversas características de interesse como, por exemplo, produtividade, boa qualidade de frutos e resistência a pragas e doenças, sem se esquecer

da necessidade de manutenção da variabilidade alélica para os locos de incompatibilidade. Estas populações seriam utilizadas como fontes de germoplasma em programas de seleção recorrente visando à obtenção de cultivares.

Algumas variedades identificadas com algum tipo de seleção dirigida já estão disponíveis aos produtores brasileiros, dentre elas têm-se os híbridos intra-varietais F₂ IAC (IAC-273, IAC-275 e IAC-277), lançados em 1999, resultantes de um programa de melhoramento de nove anos, baseado em seleção massal, retrocruzamentos e teste de progênies (MELETTI, SANTOS e MINAMI, 2000). Existem, também, três seleções disponíveis no mercado e conhecidas como ‘Maguary’, ‘Sul-Brasil’ e ‘Golden Star’. Outros trabalhos desta natureza foram feitos por Maluf et al. (1989), que por meio de seleção clonal, destacaram um grupo de 24 clones que permitiram a obtenção de ganhos de 29,1 e 89,2% para as características produção total e produção precoce, respectivamente. Viana et al. (2004) iniciaram um trabalho de avaliação de um conjunto de populações de maracujá-amarelo visando o desenvolvimento de variedades. Foi observada alta variabilidade e herdabilidade para número e comprimento de frutos, indicando a possibilidade de utilização da seleção massal para tais características. Já para percentagem de suco e espessura de casca, observou-se pequena variabilidade e baixa herdabilidade, justificando a introdução de variabilidade e uso de métodos de seleção mais elaborados. Moraes et al. (2005) avaliaram a população derivada do cruzamento entre os acessos IAPAR-123 e IAPAR-06, destacando um conjunto de plantas com potencial agrônomo para utilização em ciclos de seleção recorrente.

Nascimento et al. (2003) citam que o melhoramento dessa *Passifloraceae* deve ser feito visando o atendimento de mercados distintos. Assim, nesse mesmo trabalho, os autores avaliaram uma série de progênies de maracujazeiro-amarelo utilizando seleção massal, e destacaram cinco progênies com características desejáveis para o consumo *in natura* e quatro para a indústria de suco concentrado.

No melhoramento visando resistência a doenças, atenção especial deve ser dada ao vírus do endurecimento dos frutos (*passionfruit woodness virus* – *PWV*), à mancha-de-alternaria (*Alternaria* spp.), à verrugose (*Cladosporium herbarium*), à antracnose (*Glomerella cingulata*) e à bacteriose (*Xanthomonas axonopodis* pv. *passiflorae*), todas doenças da parte aérea. Quanto àquelas do sistema radicular, merecem destaque a fusariose, causada por *Fusarium oxysporum* f. sp. *passiflora*, e a podridão-do-pé causada por *Phytophthora* spp. (NAKASONE e PAULL, 1998; LIBERATO, 2002; SANTANA e LAU, 2002). Estas doenças causam grandes perdas de

produção e depreciação dos frutos, chegando a inviabilizar o plantio em determinadas regiões. Alguns estudos sobre transgenia para resistência a vírus tem perspectiva de aplicação comercial (ALFENAS et al., 2005; TREVISAN et al. 2006).

Trabalhos envolvendo o uso de marcadores moleculares e mapeamento de regiões genômicas associadas à tolerância à bacteriose do maracujazeiro têm sido desenvolvidos, fornecendo subsídios aos programas de melhoramento (LOPES, 2003; MATTA, 2005). Um dos grandes desafios dos melhoristas dessa cultura é a utilização de todas essas informações no desenvolvimento de cultivares para atender às demandas do setor produtivo.

2.3 Marcadores moleculares e mapas genéticos

O conhecimento a respeito da Genética iniciou-se com os trabalhos de Mendel em 1865. Trabalhando com ervilhas, Mendel postulou a existência de fatores, hoje conhecidos como genes, que controlavam a manifestação de determinadas características. Ao cruzar plantas puras e contrastantes para a cor das sementes, Mendel concluiu que as plantas heterozigóticas (F_1) produzem dois gametas distintos na mesma proporção e que ambos segregam para a formação dos zigotos. Neste caso, $\frac{1}{4}$ dos indivíduos da geração F_2 é igual ao fenótipo de um dos genitores, $\frac{1}{2}$ dos indivíduos são iguais à geração F_1 e $\frac{1}{4}$ é igual ao outro genitor, resultando no que se conhece por primeira lei de Mendel.

Mendel ainda realizou experimentos para descobrir como a forma e a cor da semente eram herdadas em conjunto, concluindo que cada genitor possuía duas cópias de um gene, e que transmitia cada uma das cópias para seus descendentes de forma independente. Desse modo, elaborou a segunda lei da herança que é a lei da segregação independente, a qual postula que durante a formação dos gametas, a segregação de um par de genes ocorre de forma independente do outro par.

Desvios da segunda lei de Mendel foram observados em ervilha de cheiro (BATESON, SAUNDERS e PUNNETT, 1904) e em *Drosophila* (MORGAN et al., 1915). Muitos genótipos apareciam mais frequentemente do que outros, indicando que pares de genes não segregavam de forma independente. Este fenômeno foi interpretado como ligação gênica, que é a base do mapeamento genético.

Sturtevant em 1913 sugeriu o uso da frequência de aparecimento dos indivíduos recombinantes, ou seja, aqueles cujas combinações de alelos diferem dos genitores, como uma medida da frequência de recombinação entre os genes. A partir desse momento, os trabalhos relacionados com a construção de mapas genéticos se intensificaram numa variedade de estudos e espécies. Em seguida descobriu-se que não havia aditividade nas medidas de distâncias genéticas quando frequências de recombinação maiores que 10% eram utilizadas, devido à presença de *crossing-over* duplo.

Haldane (1919) definiu a distância entre dois genes como uma medida do número de *crossing-overs* entre eles, considerando haver aditividade para formação dos grupos de ligação. A função de Haldane admite que a ocorrência de permuta se dá de modo independente (ausência de interferência), seguindo uma distribuição de Poisson, sendo que a unidade de distância ficou conhecida como Morgan (M). Mais tarde, Kosambi (1994), propôs uma outra função em que é admitida a ocorrência de permutas próximas como eventos não independentes (presença de interferência). Assim, esta função assume a interferência completa entre regiões arbitrariamente próximas, sendo decrescente para locos mais distantes, e igual a zero para locos independentes.

Os primeiros mapas genéticos começaram a ser construídos com base em marcadores morfológicos e citológicos, em espécies vegetais bastante estudadas como o milho, o tomate e a ervilha, em função da disponibilidade de estoques cromossômicos com aberrações, tais como aneuploidias, translocações, deleções e inversões. No caso do uso dos marcadores morfológicos, o nanismo, a deficiência de clorofila, a cor das pétalas e a morfologia foliar foram caracteres bastante utilizados como marcadores (COE, NEUFFER e HOISINGTON, 1988; RICK e YODER, 1988).

Apesar da grande contribuição desses marcadores na construção de mapas genéticos, sua disponibilidade se restringia basicamente a espécies modelo. Além disso, esses marcadores não são neutros em relação a possíveis efeitos fenotípicos, epistáticos ou pleiotrópicos e, em geral, são dominantes ou recessivos. Sobretudo, o número de marcadores morfológicos é muito baixo, restringindo a cobertura dos genomas (STUBER, 1992).

Na década de 50, surgiram os marcadores isoenzimáticos, que são isoformas moleculares da mesma enzima. São o resultado da presença de mais de um gene codificando cada uma das isoformas. As isoenzimas foram utilizadas na construção de diversos mapas (TANKSLEY e

RICK, 1980; VALLEJOS, SAKIYAMA e CHASE, 1992; LANDRY et al., 1991; PHILIPP, WEHLING e WRICKE, 1994).

Com a descoberta da estrutura do DNA por Watson e Crick (1953) e do código genético por Nirenberg e Matthaei (1961), várias descobertas subsequentes permitiram o desenvolvimento de metodologias para se conhecer a sequência de bases de uma determinada região do DNA (SANGER, 1975), a visualização de fragmentos específicos gerados por restrição enzimática (SOUTHERN, 1975) e, posteriormente, a amplificação de fragmentos específicos da molécula (MULLIS e FALOONA, 1987).

O RFLP (*Restriction Fragment Length Polymorphisms*) foi o primeiro marcador de DNA a ser utilizado na construção de mapas genéticos na espécie humana (BOTSTEIN et al., 1980)¹.

Nesta técnica, o polimorfismo ocorre devido à perda ou surgimento de sítios de restrição, inserções, deleções ou outros rearranjos que ocorrem entre dois sítios de restrição (LANZA, GUIMARÃES e SCHUSTER, 2000). Os marcadores RFLP apresentam algumas vantagens, a principal delas é a herança codominante. Devido ao fato de representarem locos únicos em cada genoma, possibilitam a utilização de sondas heterólogas, permitindo o mapeamento comparativo entre espécies. A técnica de RFLP é bastante laboriosa, com grande complexidade técnica, requerendo laboratórios bem equipados e mão-de-obra especializada. Mesmo assim, na época de sua descrição, diversos mapas foram construídos com base nestes marcadores. Em vegetais, destacam-se os mapas de brássicas (FERREIRA, WILLIAMS e OSBORN, 1994; SONG et al., 1991), algodão (REINISCH et al., 1994), milho (BURR et al., 1988; WEBER e HELENTJARIS, 1989), arroz (CAUSSE et al., 1994), soja (KEIM et al., 1990; DIERS et al., 1992) e de tomate (TANKSLEY et al., 1992).

Com a descoberta da técnica de PCR (*Polymerase Chain Reaction*) por Mullis e Faloona (1987), utilizando uma DNA polimerase termoestável e termocicladores programáveis com elevada capacidade de processamento, houve uma grande automatização na síntese enzimática *in vitro* de DNA. A partir daí, surgiram uma ampla gama de marcadores moleculares em base em PCR, sendo o RAPD (*Random Amplified Polymorphic DNA*) o primeiro deles (WILLIAMS et al., 1990)¹.

¹ Para maiores informações sobre a técnica, consultar Ferreira e Grattapaglia (1998); Vieira, Vello e Silva Filho (2004); Fungaro e Vieira (2001)

O polimorfismo deste tipo de marcador se deve a variações nas seqüências complementares aos sítios de anelamento dos *primers* resultando em apenas dois fenótipos moleculares, presença ou ausência de bandas, daí porque a herança deste marcador é dominante.

Segundo alguns autores a técnica de RAPD popularizou o uso de marcadores moleculares para uma variedade de espécies animais e vegetais. Dentre os vários estudos feitos com esta técnica, a construção de mapas genéticos apresentou grande destaque. Foram construídos mapas de *Pinus* (GRATTAPAGLIA et al., 1991), *Arabidopsis* (REITTER et al., 1992), eucalipto (GRATTAPAGLIA e SEDEROFF, 1994), maçã (HEMMAT et al., 1994), cana-de-açúcar (SOBRAL e HONEYCUTT, 1993) e maracujá-amarelo (CARNEIRO et al., 2002), dentre outros.

Os marcadores microssatélites vieram em seguida. Os microssatélites ou SSR (*Single Sequence Repeats*) são seqüências de 1 a 6 nucleotídeos repetidas em *tandem* (LITT e LUTY, 1989). As seqüências que flanqueiam os microssatélites são bastante conservadas até mesmo entre espécies, sendo utilizadas para o desenho de *primers* e amplificação dos microssatélites. O polimorfismo é detectado devido à diferença de tamanho entre os fragmentos amplificados, que se deve ao número diferenciado de repetições dentro dos microssatélites. Estes marcadores são muito interessantes para uma série de estudos genéticos devido à sua natureza codominante e multialélica. Os microssatélites foram utilizados para a saturação de mapas genéticos de cacau (PUGH et al., 2004), eucalipto (BRONDANI, BRONDANI e GRATTAPAGLIA, 2002), seringueira (LESPINASSE et al., 2000) e lentilha (TEMNYKH et al., 2000; HAMWIEH et al., 2005). Devido à sua importância no contexto desta tese, este marcador será melhor apresentado e discutido no item 2.4.

O AFLP (*Amplified Fragment Length Polymorphisms*) é uma classe de marcadores que aliam a especificidade dos sítios de restrição do RFLP à praticidade de amplificação da PCR (VOS et al., 1995).

A técnica baseia-se na digestão do DNA genômica com duas enzimas de restrição e ligação de adaptadores específicos com terminais complementares às extremidades coesivas dos sítios de restrição. É utilizado um par de adaptadores para cada sítio de restrição. Os fragmentos digeridos e ligados aos adaptadores são amplificados por PCR com *primers* pré-seletivos, cuja seqüência é complementar a um dos adaptadores acrescido de um nucleotídeo na extremidade 3'. Esta pré-amplificação é feita com o objetivo de aumentar a proporção dos fragmentos de interesse com os sítios de restrição utilizados inicialmente. Estes fragmentos são amplificados

novamente com *primers* seletivos com sequência igual ao dos pré-seletivos acrescidos de mais dois nucleotídeos na extremidade 3'. A detecção do polimorfismo é feita em géis de poliacrilamida, corados com prata, radioatividade ou fluorescência.

Essa técnica mostra boa reprodutibilidade dos resultados. Como vantagem tem-se: a exploração de regiões genômicas arbitrárias sem conhecimento prévio da sequência de DNA; a especificidade da PCR e a grande quantidade de informação gerada em um único gel de poliacrilamida, uma vez que são explorados polimorfismos de restrição e amplificação seletiva. Entretanto, são marcadores dominantes, apresentando baixo conteúdo informativo por loco, além de o procedimento ser trabalhoso e exigir um DNA de alta qualidade.

Devido ao seu grande poder de saturação de mapas genéticos em pouco tempo, os marcadores AFLP foram utilizados para uma série de espécies vegetais como cevada (QI, STAM e LINHOUT, 1998), alfafa (JULIER et al., 2003); arroz (MAHESWARAN et al., 1997), milho (VUYLSTEKE et al., 1999), pêssego (LU et al., 1998), sorgo (KLEIM et al., 2000) e maracujá-amarelo (MATTA, 2005; MORAES, 2005; LOPES et al., 2006).

Nos últimos anos, em função do acúmulo de informações a respeito dos diferentes tipos de marcadores moleculares e das populações segregantes ou de mapeamento como são conhecidas, o uso dos mapas genéticos no mapeamento de características de herança complexa, os chamados QTL (*Quantitative Trait Loci*) ganharam grande impulso. A necessidade de obtenção de mapas altamente saturados levou os pesquisadores a utilizarem as informações de todos os tipos de marcadores disponíveis. Assim, diversos mapas foram construídos ou integrados para muitas espécies como *Populus* (CERVERA et al., 2001), *Eucalyptus camaudulensis* (AGRAMA, GEORGE e SALAH, 2002), *Medicago truncatula* (THOQUET et al., 2002), cacau (PUGH et al., 2004), soja (SONG et al., 2004) e cana-de-açúcar (GARCIA et al., 2006).

Recentemente, certas mutações pontuais foram transformados numa nova classe de marcadores moleculares, caracterizando os SNP (*Single Nucleotide Polymorphism*). Estes vêm ganhando espaço, principalmente para espécies com grande volume de informações geradas a partir dos projetos genomas. Os SNPs são bastante abundantes no genoma humano, ocorrendo um SNP a cada 1000 pb, ou seja, um SNP por centimorgan (MARTIN et al., 2000). Mais de 8051 SNP foram identificados em 12 acessos de *Arabidopsis*, e estão disponíveis para acesso público no banco de dados (SCHMID et al., 2003). A partir do momento que as sequências de DNA de

uma dada região genômica é conhecida, *primers* podem ser sintetizados para amplificar os alelos de interesse revelando polimorfismos existentes. Nos últimos anos, uma ampla variedade de técnicas de genotipagem de SNP foi desenvolvida (KWOK, 2001; GUT, 2001).

Devido ao fato de serem marcadores altamente abundantes no genoma, de ação codominante e de efeito predominantemente neutro, diversos mapas de ligação também foram construídos com esse tipo de marcador, como exemplo, têm-se os mapas de *Drosophila* (BERGER et al., 2001), *Arabidopsis* (PONCE, ROBLES e MICOL, 1999; TORJEK et al., 2003), *Caenorhabditis elegans* (SWAN et al., 2002) e de bovino (SNELLING et al., 2005).

Os estudos de mapeamento genético são importantes, mormente quando se objetiva contribuir para o melhoramento genético de características de interesse via seleção assistida por marcadores (SAM) (HOSPITAL, CHEVALET e MULSANT, 1992; SPELMAN e BOVENHIS, 1998; KELLY et al., 2003). Os mapas genéticos contribuem também para: i) acessar a diversidade de populações (FERGUSON, 1989); ii) estudos evolutivos de herança entre espécies relacionadas (BROAD e HILL, 1994), e entre diferentes populações de mapeamento de uma mesma espécie (SONG et al., 2004; ROMÁN et al. 2004) e; iii) estudos de sintenia entre espécies afins (DOERGE, ZENG e WEIR, 1997) e identificação de genes ortólogos (KALO et al., 2000).

Os marcadores moleculares também podem ajudar a entender a evolução do cariótipo das espécies. No gênero *Helianthus*, por exemplo, foi demonstrada a presença de 27 segmentos colineares resultantes de no mínimo oito translocações e três inversões. Juntos, estes rearranjos requerem no mínimo 20 quebras e fusões cromossômicas (BURKE et al., 2004).

O mapeamento de regiões genômicas associadas à herança de características quantitativas se constitui em uma aplicação importante dos mapas genéticos construídos com base em marcadores moleculares. Diversos QTL têm sido mapeados para resistência a doenças em frutíferas como: pêssego (DIRLEWANGER et al., 1999; FOULONGNE et al., 2003), maçã (STANKIEWICZ-KOSYL, PITERA e GAWRONSKI, 2005), cacau (FLAMENT et al., 2001; RISTERUCCI et al., 2003) e maracujá-amarelo (MATTA, 2005; LOPES et al., 2006). Outros QTL relacionado à produtividade e qualidade de frutos também foram detectados (CONNER, BROWN e WEEDEN, 1998; DIRLEWANGER et al., 1999; GARCIA et al., 2000; FANIZZA et al., 2005, MORAES, 2005).

2.4 Marcadores microssatélites: aspectos gerais e metodologias de obtenção

Os genomas eucariotos são densamente povoados por seqüências simples repetidas, as quais consistem de 1 a 6 nucleotídeos repetidos em *tandem*. Essas regiões repetidas são denominadas de microssatélites, SSR (*Simple Sequence Repeats*) ou STR (*Short Tandem Repeats*).

De acordo com Weber (1990), os microssatélites podem ser classificados em: i) Perfeitos: quando a seqüência repetida não é interrompida por qualquer base que não pertença ao motivo, por exemplo: TATATATATATATATA; ii) Imperfeitos: quando existe entre os motivos, pares de bases que não correspondem ao mesmo, por exemplo: TATATATAcTATATA e iii) Compostos: quando a seqüência contém duas seqüências repetidas distintas adjacentes, por exemplo: TATATATAGTGTGTGTGTGT.

Os microssatélites estão presentes em regiões codantes e não codantes do genoma e são usualmente caracterizados pelo alto polimorfismo que é decorrente das altas taxas de mutação presente nesses locos. Diferentes hipóteses têm sido formuladas para explicar os mecanismos que levam às altas taxas de mutação que variam de loco para loco, sendo esta variação de 10^{-2} a 10^{-6} nucleotídeos por loco por geração (SCHLÖTTERER, 2000). Dentre elas, podem ser citados: os erros que acontecem nos processos de recombinação, o *crossing-over* desigual e o processo de *slippage* da polimerase no processo de replicação ou reparo do DNA (STRAND et al., 1993). Uma ampla revisão sobre esse assunto pode ser consultada em Oliveira et al. (2006).

Os marcadores microssatélites foram primeiramente desenvolvidos em humanos (LITT e LUTY, 1989; WEBER e MAY, 1989; TAUTZ, 1989) e logo receberam a atenção dos melhoristas e geneticistas de plantas, pois vários estudos demonstraram que os microssatélites são amplamente distribuídos no genoma das espécies (BRUNEL, 1994).

As seqüências de DNA que flanqueiam os microssatélites são geralmente conservadas dentro de uma mesma espécie, permitindo a seleção de *primers* específicos que amplificam, via PCR, fragmentos contendo o DNA repetitivo em todos os genótipos. Quando os microssatélites são individualmente amplificados, usando o par de *primers* complementar às seqüências únicas que os flanqueiam, eles quase que invariavelmente mostram extensivo polimorfismo para tamanho. Esta variação no tamanho dos produtos de PCR é consequência da ocorrência de diferentes números de unidades repetidas dentro da estrutura do microssatélite (MORGANTE e

OLIVIERI, 1993; McCOUCH et al., 1997; CREGAN et al., 1999; ASHKENAZI et al., 2001). Desse modo, vários alelos podem ser detectados para uma determinada população. Indivíduos homozigóticos possuem o mesmo número de repetições nos cromossomos homólogos, enquanto indivíduos heterozigóticos possuem número de repetições diferentes nos dois homólogos. Portanto, o loco é definido pelo par de *primers* e os vários alelos, pelo tamanho das bandas amplificadas.

Variações no número das repetições em *tandem* se acumulam mais rapidamente na população do que mutações de ponto e do que os eventos de inserções/deleções responsáveis pelo polimorfismo detectado por outras técnicas moleculares (BROWN et al., 1996). Segundo alguns autores, esses marcadores estão substituindo os demais em vários tipos de estudos genéticos, devido a sua reprodutibilidade, simplicidade e rapidez da técnica, pequena quantidade de DNA requerida, baixo custo de utilização, grande poder de resolução e altos níveis de polimorfismo (CHEN et al., 1997; BRONDANI et al., 1998; GIANFRANCESCHI et al., 1998; RALLO, DORADO e MARTÍN, 2000). Entretanto, observações feitas em maracujá-amarelo, demonstram que o desenvolvimento e a otimização dos locos de microssatélites a partir de bibliotecas genômicas enriquecidas não é uma tarefa muito simples, haja vista que problemas decorrentes da amplificação desses locos podem diminuir a eficiência e a rapidez dessa técnica.

Segundo Ferreira e Grattapaglia (1998), os microssatélites constituem uma das classes de marcadores moleculares mais polimórficos. Entretanto, atualmente, os SNP vêm merecendo destaque, em função de revelar quase 80% de todo o polimorfismo conhecido em humanos, sendo possível encontrar um SNP a cada 1000 pares de bases nesta espécie. Apesar dos SNPs serem bastante freqüentes e mutacionalmente mais estáveis do que os microssatélites, eles são predominantemente bialélicos o que os torna menos informativo do que os microssatélites.

Outra vantagem sobre os demais marcadores baseados em PCR, como o RAPD, é a codominância dos microssatélites (MUDGE et al., 1997). A codominância torna-os mais informativos para análises de ligação (YU, PARK e POYSA, 1999). Além disso, os microssatélites parecem ter uma distribuição freqüente e aleatória, permitindo uma cobertura completa do genoma (RALLO, DORADO e MARTÍN, 2000).

Essas características tornam os microssatélites úteis na construção de mapas de ligação, na análise gênica e de locos quantitativos (QTL) (SCHULER et al., 1996; KNAPIK et al., 1998; TEMNYKH et al., 2000; HAMWIEH et al., 2005), *fingerprinting* (WRIGHT e BENTZEN, 1994;

BOWERS et al., 1999; SCHLÖTTERER, 2000), proteção de variedades (PRIOLLI et al., 2002), caracterização e conservação de germoplasma (MASONA et al., 2005), estudos de diversidade genética (COLLEVATTI, BRONDANI e GRATTAPAGLIA, 1999; ZUCCHI et al., 2002), análise de *pedigree* (DREISIGACKER et al., 2004) e análise de bibliotecas genômicas para clonagem de genes (WANCHANA et al., 2005).

Apesar das vantagens apresentadas, os marcadores microsatélites estão disponíveis para uso apenas em plantas modelo (CHEN et al., 1997; McCOUCH et al., 1997). A grande limitação da tecnologia está no seu desenvolvimento, pois envolve um processo demorado, trabalhoso e com alto custo (McCOUCH et al., 1997; KÖLLIKER et al., 2001). No entanto, essa desvantagem é compensada pela ampla gama de potencialidades de pesquisa e tecnologias que se abrem ao se dispor de tal marcador. Uma vez desenvolvidos, eles podem ser utilizados com a facilidade, eficiência e a rapidez típica da técnica de PCR.

Os microsatélites podem ser identificados por um *screening* na base de dados de DNA (*Genbank*) da espécie que se tem interesse (ZANE, BARGELLONI e PATARNELLO, 2002). Atualmente, esta é uma estratégia interessante em função dos grandes programas de seqüenciamento do genoma (como um todo ou de suas partes), oferecendo novas oportunidades para o desenvolvimento desse tipo de marcador de forma a contribuir principalmente na geração de mapas de ligação altamente saturados.

McCouch et al. (2002) utilizaram essa estratégia para o mapeamento de 2.240 novos microsatélites em arroz. Ainda, Mahalakshmi et al. (2002) encontraram 7.325 microsatélites de *Medicago* spp, numa busca no banco de dados. Contudo, estas informações não estão disponíveis para a maioria das espécies, como é o caso do maracujá-amarelo, havendo a necessidade de se recorrer à construção de bibliotecas genômicas.

As estratégias clássicas de isolamento de microsatélites envolvem a construção de bibliotecas genômicas selecionadas para um certo tamanho de inserto, e o *screening* de milhares de clones com sondas apropriadas (RASSMANN, SCHLÖTTERER e TAUTZ, 1991). Embora relativamente simples, em especial para genomas ricos em microsatélites, esta metodologia pode ser bastante tediosa e ineficiente para espécies com baixa frequência de microsatélites, o que pode não ser interessante quando se pretende obter um grande número de seqüências.

No trabalho de Toth, Gáspari e Jurka (2000), é apresentada a frequência de microsatélites em diferentes organismos. Em *Rodentia*, são esperados cerca de 429 microsatélites por Mb,

enquanto que em *Embryophyta* são esperados cerca de 154. Assim, numa biblioteca genômica de 2.500 clones, com tamanho médio de inserto de 400 pb, sem sobreposição, representando 1 Mb, é esperado um enriquecimento em seqüências de microssatélites de 17,16 e 6,16% em *Rodentia* e *Embryophyta*, respectivamente. Tal fato demonstra o diferencial de enriquecimento em função da espécie.

Outro aspecto interessante é que as frequências dos motivos (tipo de repetição) dos microssatélites variam bastante entre espécies e muitas das seqüências homólogas às sondas comumente utilizadas no *screening* de algumas espécies podem ser sub-representadas em outras (ZANE, BARGELLONI e PATARNELLO, 2002).

As classes de microssatélites contendo repetições de dinucleotídeos mais encontradas no genoma de plantas são AT/TA, AG/TC, seguidas de AC/GT; as de trinucleotídeos são TAT e TCT (MORGANTE e OLIVIERI, 1993; WANG et al., 1994). Todavia, existe uma série de dificuldades técnicas no *screening* de microssatélites constituídos por AT/TA, devido à sua auto-complementariedade, o que prejudica o processo de hibridização. Assim, as sondas CA e GA são as mais usadas.

A percentagem de clones positivos (contendo o inserto de interesse) com as metodologias convencionais de construção de bibliotecas genômicas é muito baixa, variando de 0,13 a 4,5% em mamíferos, de 0,066 a 8,92% em peixes, 0,059 a 5,8% em plantas e 0,0025 a 1,7% em pássaros (ZANE, BARGELLONI e PATARNELLO, 2002). Por isso, essa metodologia de isolamento e desenvolvimento de microssatélites pode se mostrar bastante ineficiente.

Uma alternativa à construção de bibliotecas genômicas baseia-se na amplificação de microssatélites desconhecidos utilizando *primers* de RAPD e posterior hibridização com sonda de microssatélites (CIFARELLI, GALLITELLI e CELLINI, 1995) ou clonagem de todos os produtos de RAPD e posterior *screening* por *Southern blot* (LUNT, HUTCHINSON e CARVALHO, 1999). Esta metodologia é baseada na observação de que há regiões repetitivas nos *amplicons* detectados por RAPD. Estratégias baseadas na extensão de *primers* também têm sido propostas para a construção de bibliotecas genômicas enriquecidas. Os trabalhos de Paetkau (1999) e Ostrander et al. (1992) descrevem esta metodologia como sendo bastante eficiente na construção de bibliotecas enriquecidas com repetições AC. As estratégias dos autores citados anteriormente, envolvem a construção de uma biblioteca genômica primária em que o DNA genômico fragmentado é inserido em vetores de fagos ou fagmídeos para a obtenção de uma

biblioteca de DNA fita simples. Este DNA fita simples é usado como molde para uma reação de extensão do *primer*, constituído por uma sequência repetitiva, que gera DNA de fita dupla apenas nos vetores contendo as repetições desejadas. Os dois processos de enriquecimento divergem na metodologia utilizada para recuperar os produtos amplificados. Entretanto, devido ao fato de serem estratégias bastante trabalhosas, elas não são comumente empregadas.

Os métodos mais empregados, atualmente, para o isolamento de microssatélites são baseadas em hibridização seletiva. Os protocolos básicos foram propostos por Karagyzov, Kalchieva e Chapman (1993), Armour et al. (1994) e Kijas et al. (1994), sendo que muitas modificações foram feitas ao longo do tempo. Segundo Zane, Bargelloni e Patarnello (2002), estes protocolos são bastante eficientes e aplicáveis para basicamente todas as espécies.

A primeira etapa na construção dessas bibliotecas enriquecidas é fragmentar o DNA em pequenos fragmentos para serem ligados a vetores ou adaptadores específicos. Após esta etapa, e dependendo da quantidade de DNA, este é hibridizado com sondas contendo as repetições desejadas, sendo a sonda ligada a uma membrana de nylon (KARAGYZOV, KALCHIEVA e CHAPMAN 1993; ARMOUR et al., 1994). A sonda também pode ser 5' biotinilada e aderida a contas magnéticas ligadas a estreptavidina (KIJAS et al., 1994). Assim, o DNA hibridizado pode ser seletivamente removido. Após a hibridização e muitas etapas de lavagem para remover o DNA não-ligado, o DNA é eluído e recuperado por PCR. Em seguida, o DNA enriquecido é clonado em um vetor adequado. Dependendo da eficiência do procedimento, os clones podem ser sequenciados diretamente ou passarem por um *screening* por *Southern blot* ou por PCR.

Outra abordagem denominada FIASCO (*Fast Isolation by AFLP of Sequences Containing Repeats*) tem sido proposta na literatura. Esta se baseia na técnica de AFLP, na qual o DNA é simultaneamente digerido com as enzimas *MseI* e ligado a adaptadores *MseI* AFLP (5'-TACTCAGGACTCAT3'/5'-GACGATGAGTCCTGAG-3'). Em seguida, é feita uma pré-amplificação com uma mistura de *primers* contendo as quatro possíveis bases seletivas (*MseI*-N), permitindo a amplificação de todos os fragmentos flanqueados por sítios *MseI*. O DNA é então, hibridizado com sondas marcadas e biotiniladas, sendo recuperado com o uso das contas magnéticas ligadas à estreptavidina. O restante dos passos é feito de forma similar às outras metodologias descritas anteriormente (ZANE, BARGELLONI e PATARNELLO, 2002).

Apesar das dificuldades, o desenvolvimento de microssatélites tem-se tornado cada vez mais acessível, principalmente devido às novas técnicas de enriquecimento de bibliotecas

genômicas e a tecnologias rápidas de sequenciamento automático baseadas em fluorescência (RAFALSKI et al., 1996). Diversos pesquisadores têm utilizado esta nova abordagem no desenvolvimento de microssatélites tais como em sorgo (TARAMINO et al., 1997), *Populus* (VAN DER SCHOOT et al., 2000), batata (ASHKENAZI et al., 2001), palmito (GAIOTTO, BRONDANI e GRATTAPAGLIA, 2001), azevém (JONES et al., 2001), trevo branco (KÖLLIKER et al., 2001) e pêssago (ARANZANA et al., 2002).

Butcher e Moran (2000) desenvolveram uma série de microssatélites em *Acacia mangium*. Estes autores encontraram de 1 a 40% de seqüências contendo microssatélites utilizando bibliotecas genômicas convencionais e enriquecidas, respectivamente. Valores de 10 a 40% de enriquecimento também foram observados em eucalipto (BRONDANI et al., 1998).

Atualmente, há uma quantidade pequena de seqüências de DNA de maracujá-amarelo depositadas no banco de dados (*GenBank*). Por isso, metodologias eficientes de isolamento de seqüências contendo microssatélites devem ser buscadas. Sendo assim, o uso de estratégias baseadas em hibridização seletiva é de especial importância para a cultura do maracujá-amarelo, pois possibilitam com grande eficiência a obtenção e desenvolvimento dos marcadores microssatélites.

2.5 Construção de mapas genéticos em populações F₁ com ênfase em *P. edulis* f. *flavicarpa*

A construção de mapas de ligação envolve basicamente a aplicação das técnicas de genética molecular aos conceitos originais da herança apresentados por Mendel. Depois do aparecimento do primeiro mapa genético baseado em marcadores moleculares, desenvolvido por Botstein et al. (1980), vários mapas vêm sendo construídos para diversos organismos.

A metodologia de construção de um mapa integra um certo número de técnicas que incluem o desenvolvimento de linhagens genitoras e populações segregantes adequadas, a identificação dos genótipos nos locos marcadores e a utilização de técnicas de análises estatísticas e computacionais para a estimativa de ligação e da distância entre marcadores, bem como sua ordem nos cromossomos.

Como existem centenas de marcadores segregando, as análises são amplamente computadorizadas. Métodos estatísticos para as análises de ligação e construção dos mapas foram

desenvolvidos para populações convencionais de mapeamento e implementados em programas computacionais como Mapmaker/Exp (LANDER et al., 1987), Crimap (GREEN, FALLS e CROOKS, 1990), GMendel (LIU e KNAPP, 1992), JoinMap (VON OOIJEN e VOORRIPS, 2001), Multimap (MATISE, PERLIN e CHAKRAVARTI, 1994) e Outmap (BUTCHER et al., 2002).

As populações convencionais de mapeamento referem-se àquelas populações segregantes derivadas do cruzamento entre linhagens puras (homozigóticas). A geração F_1 é autofecundada ou retrocruzada de forma a produzir gerações F_2 , F_n ou RC_1 , segregantes tanto para os marcadores moleculares como para as características fenotípicas de interesse. Nestes casos, a escolha apropriada da população a ser mapeada deve levar em consideração o tipo de marcador e a espécie (TANKSLEY et al., 1988). O máximo de informação genética é obtida quando se utiliza populações F_2 e marcadores codominantes, uma vez que estes marcadores permitem o conhecimento completo do genótipo em questão.

Para a construção dos mapas de ligação, várias adaptações do método da verossimilhança foram desenvolvidas, merecendo destaque os trabalhos de Lander e Green (1987) cujos algoritmos foram implementados nos programas Mapmaker/Exp, Crimap e Multimap. O método de análise multiponto usa, simultaneamente, informações de todos os marcadores em um grupo de ligação para determinar a sua ordem, bem como estimar a frequência de recombinação entre eles. Contudo, esses programas pressupõem que as fases de ligação entre os alelos do marcador sejam conhecidas nos genitores. Pacotes computacionais baseados em verossimilhança não podem ser usados para o mapeamento de algumas espécies, quando se dispõe de marcadores com diferentes tipos de segregação e a fase de ligação não é conhecida, como é o caso dos maracujás.

Para culturas em que é possível obter linhagens via sucessivas autofecundações, a análise de ligação com marcadores moleculares é feita facilmente, pois as fases de ligação dos diferentes locos marcadores são conhecidas, o que permite a distinção entre os genótipos recombinantes e os não-recombinantes (LANDER et al., 1987). Nas populações de mapeamento mais comuns (F_2 , retrocruzamento e linhas puras recombinantes), marcadores moleculares bi-alelicos, como RFLP, RAPD e AFLP são suficientes para caracterizar o polimorfismo existente em todos os locos do genoma. Contudo, estes marcadores não são ideais nas espécies de polinização aberta, em que populações tradicionais de mapeamento não são obtidas, pois os problemas de depressão por endogamia ou de auto-incompatibilidade não permitem a obtenção de linhagens

(GRATTAPAGLIA e SEDEROFF, 1994; HEMMAT et al., 1994; LOPES et al., 2006), ou quando existe um longo período entre gerações que impede, na prática, a obtenção de linhagens homozigóticas. Nestes casos, ou seja, em cruzamentos entre plantas não endogâmicas podem haver até quatro alelos segregando, os quais só podem ser visualizados com marcadores apropriados como os microsatélites (OLSON et al., 1989).

Esforços no sentido de desenvolver ferramentas robustas de análises de dados moleculares em populações não endogâmicas foram feitos nos últimos anos (RITTER, GEBHARDT e SALAMINI, 1990; ARUS et al., 1994; RITTER e SALAMINI, 1996; MALIEPAARD, JANSEN e VAN OOIJEN, 1997; WU et al., 2002). Entretanto, a análise genética e os métodos estatísticos neste tipo de população são mais complicados do que nas espécies em que é possível a obtenção de linhagens, devido ao número de alelos e padrão de segregação que varia para cada loco e as fases de ligação entre os diferentes marcadores que não são conhecidas *a priori*.

Visando contornar esses problemas, tem-se utilizado a verossimilhança e a estratégia dita duplo pseudocruzamento teste na qual a análise de ligação é feita para cada genitor separadamente. Esta estratégia baseia-se no fato de que se um indivíduo for heterozigótico para um marcador e outro indivíduo tiver genótipo nulo, isto é, outras formas alélicas não detectadas (ausência de banda no gel), a progênie F_1 segrega na proporção 1:1 para presença e ausência de bandas. Após a genotipagem de uma amostra de várias dezenas de indivíduos F_1 para os marcadores selecionados, dois conjuntos de marcadores são gerados, resultando em um mapa de ligação para cada genitor, ou seja, os mapas são indivíduo-específicos (GRATTAPAGLIA e SEDEROFF, 1994). Alguns autores utilizaram esta abordagem na construção de mapas de espécies altamente heterozigóticas como eucalipto (GRATTAPAGLIA e SEDEROFF, 1994, MARQUES et al., 1998), maçã (HEMMAT et al., 1994), cacau (CROUZILLAT et al., 1996) e *Populus* (CERVERA et al., 2001).

Para a construção do primeiro mapa de ligação de maracujá-amarelo, utilizou-se a configuração de duplo pseudocruzamento teste e marcadores RAPD em 90 plantas de uma população F_1 oriundas do cruzamento entre os acessos IAPAR-06 e IAPAR-123 (CARNEIRO et al., 2002). Para os dois genitores, foram encontrados 9 grupos de ligação. O mapa do genitor IAPAR-123 foi de 728 cM, com uma distância média de 11,2 cM entre locos, sendo que o tamanho dos grupos de ligação variou de 56 a 144,6 cM. O mapa do genitor IAPAR-06 foi de 783,5 cM, com distância média entre locos de 12,2 cM e comprimento dos grupos variando de

20,6 a 144,2 cM (Tabela 1). Nesta mesma população, outro mapa baseado em AFLP foi construído, utilizando 117 plantas F₁, também obtendo 9 grupos de ligação para cada genitor (LOPES et al., 2006). Para o acesso IAPAR-123, o comprimento dos grupos de ligação variou de 8,1 a 140,4 cM, com média de 54,3 cM, comprimento total de 488,9 cM e 9,6 cM de distância média. O mapa do genitor IAPAR-06 foi de 790,2 cM, com 12,6 cM de distância média entre os locos e comprimento dos grupos de ligação variando de 12,9 a 208,2 cM (Tabela 1).

Alternativamente, Moraes (2005) estabeleceu a homologia entre os grupos de ligação dos genitores, utilizando marcas de AFLP com segregação 1:1 e 3:1 num conjunto de 160 plantas desta mesma população (Figura 1). Este autor utilizou o programa TreeMap (COELHO, 2005), que trabalha com mistura de segregações, para construção dos mapas e posicionamento das marcas 3:1. Em seguida um arquivo de dados, somente com marcas 1:1, foram analisados no programa Mapmaker/EXP versão 3.0 (LANDER et al., 1987), para obtenção das estimativas multiponto. As marcas 3:1 foram colocadas como acessórias servindo apenas para o estabelecimento da homologia dos mapas. Neste caso, foram encontrados 9 grupos de ligação para o genitor IAPAR-123, abrangendo um comprimento de mapa de 471,5 cM, com uma distância média de 9,8 cM entre locos e variação no tamanho dos grupos de ligação de 11,1 a 93,5 cM. Para o genitor IAPAR-06 foram encontrados 10 grupos de ligação com comprimento de mapa de 727,8 cM e distância média consecutiva entre locos de 13,7 cM. Além disso o comprimento dos grupos de ligação variou de 3,9 e 212,1 cM. Na Tabela 1 é apresentada uma comparação entre os 3 mapas de maracujá-amarelo, utilizando a estratégia duplo pseudocruzamento teste.

Entretanto, a configuração de duplo pseudocruzamento teste apresenta algumas limitações, uma vez que os mapas são indivíduo-específicos e apenas uma parte dos marcadores é usada, não podendo ser alinhados ou integrados como tais, a não ser que sejam utilizados outros marcadores comuns a ambos os mapas (GRATTAPAGLIA e SEDEROFF, 1994; DALBÓ et al., 2000; WU et al., 2002). Além disso, a integração dos mapas é problemática (MALIEPAARD, JANSEN e VAN OOIJEN, 1997), sob o ponto de vista estatístico e computacional, uma vez que poucos programas estão disponíveis para este tipo de população, bem como, do ponto de vista biológico, ser necessário o desenvolvimento e mapeamento de um grande número de marcadores informativos em ambos os genitores.

Tabela 1 - Comparação entre os mapas de ligação de maracujá-amarelo construídos por Carneiro et al. (2002), Moraes (2005) e Lopes et al. (2006). Distâncias medidas em cM de acordo com a função de Kosambi

Característica	Carneiro et al. (2002)		Moraes (2005)		Lopes et al. (2006)	
	IAPAR-123	IAPAR-06	IAPAR-123	IAPAR-06	IAPAR-123	IAPAR-06
Número de locos	148	121	179	154	140	115
Grupos de ligação	9	9	9	10	9	9
Marcas ordenadas no <i>framework</i>	65	64	57	63	57	72
Marcas acessórias	70	32	122	90	78	38
Comprimento do mapa	727,7	783,5	471,5	727,8	488,9	790,2
Comprimento do maior grupo	144,6	144,2	93,5	212,1	140,4	208,2
Comprimento do menor grupo	56,0	20,6	11,1	3,9	8,1	11,9
Comprimento médio dos grupos	80,9	87,1	52,4	72,8	54,3	87,8
Distância média entre marcas	11,2	12,2	9,8	13,7	9,6	12,6

Nos últimos anos, diversos trabalhos demonstraram métodos para determinar as fases de ligação entre locos marcadores com um número variável de alelos (RITTER, GEBHARDT e SALAMINI, 1990; ARUS et al., 1994; RITTER e SALAMINI 1996; MALIEPAARD, JANSEN e VAN OOIJEN, 1997). Entretanto, uma das maiores limitações destes estudos é que uma fórmula geral não foi derivada para estimar simultaneamente a ligação e a fase de ligação para um conjunto misto de dados de marcadores, incluindo todos os possíveis padrões de segregação. Em muitos casos, os dois possíveis critérios empregados para a caracterização da fase de ligação, isto é, a estimativa de frequência mínima de recombinação e o de *LOD score*, podem não ser consistentes tal como demonstrado por Maliepaard, Jansen e Van Ooijen (1997).

Van Ooijen e Voorrips (2001) desenvolveram o programa computacional JoinMap para resolver os problemas de integração desses mapas de ligação. Este programa combina dados de diferentes experimentos ou *pedigrees*, utilizando apenas frequências de recombinação entre parte dos marcadores para estimar sua ordem nos grupos de ligação. Assim, não é esperado que este programa seja tão preciso quanto os pacotes baseados em verossimilhança multiloco.

A construção de mapas de ligação em populações não endogâmicas ganhou grande impulso nos últimos anos com os avanços das metodologias de determinação das fases e frequências de recombinação para locos com diferentes tipos de segregação. Esta integração dos mapas pode ser feita com o uso de qualquer marca que segregue em ambos os genitores, inclusive

marcas com segregação do tipo 3:1, geradas por marcadores dominantes como o RAPD e o AFLP.

Figura 1 - Homologia dos mapas de ligação dos acessos IAPAR-123 e IAPAR-06, construídos por Moraes (2005): grupos de ligação I, II, III, IV e V

Figura 2 - Homologia dos mapas de ligação dos acessos IAPAR-123 e IAPAR-06, construídos por Moraes (2005): grupos de ligação VI, VII, VIII, IX e X

Matta (2005) utilizou marcas geradas pela técnica de AFLP que segregam na proporção 3:1 para integrar os mapas dos genitores IAPAR-06 e IAPAR-123, em maracujá-amarelo, utilizando o programa JoinMap. Neste estudo, foram utilizadas 114 (30,9%) marcas informativas oriundas do genitor IAPAR-06, 139 (37,7%) marcas informativas do genitor IAPAR-123 e 116 (31,4%) marcadores bi-parentais (Figura 3).

De acordo com Maliepaard, Jansen e Van Ooijen (1997), existe um grande viés das estimativas quando são utilizados marcadores com segregação 3:1 para a integração dos mapas. A detecção de ligação significativa para um determinado par de marcas não implica, necessariamente que a medida da estimativa da frequência de recombinação seja acurada. A obtenção de estimativas confiáveis é um aspecto fundamental do mapeamento para se determinar a correta ordem e distância entre os marcadores. Embora as marcas 3:1 possam ser empregadas

na integração das marcas 1:1, seu uso para estabelecer a correta ordem entre esses dois grupos será limitado considerando populações pequenas (MATTA, 2005).

Figura 3 - Mapa integrado dos acessos IAPAR-123 e IAPAR-06, obtidos por Matta (2005) com o uso do programa JoinMap: grupos de ligação I, II e III

Essas considerações fazem os marcadores moleculares multialélicos e codominantes, como os microssatélites, serem extremamente úteis para promover essa integração, pois permitem a obtenção de marcas com segregação informativa (1:1:1:1 e 1:2:1), possibilitando a obtenção das fases de ligação, frequência de recombinação e ordenação com um menor viés.

Neste sentido, os marcadores microssatélites tendem a substituir, como ponte, os marcadores do tipo 3:1 e mesmo outros codominantes como as isoenzimas e os RFLPs, pois permitem obter mapas mais confiáveis para os trabalhos de mapeamento de QTL, dado que a maioria dos locos controladores de características quantitativas nestas populações são multialélicos. Marcadores multialélicos e codominantes permitirão a localização, entendimento e manipulação adequada de toda a variação alélica que segrega para os QTL (BRONDANI, BRONDANI e GRATTAPAGLIA, 2002).

Associado ao desenvolvimento de marcadores informativos, o desenvolvimento de metodologias que melhorem as estimativas de fase de ligação, frequência de recombinação e ordenação dos locos ao longo dos grupos de ligação é de fundamental importância para o mapeamento em populações formadas a partir de parentais não endogâmicos, uma vez que o JoinMap, que é disponível atualmente, não utiliza os melhores algoritmos para tal.

O programa Outmap, desenvolvido por Butcher et al. (2002), é bastante robusto na análise de dados para esse tipo de população. Entretanto, utiliza apenas informações de marcadores codominantes, não sendo possível a utilização de marcas com segregação 3:1.

Mais recentemente Wu et al. (2002) desenvolveram um algoritmo geral baseado em verossimilhança para estimar, simultaneamente, a frequência de recombinação e as fases de ligação de todos os marcadores ligados segregando nesse tipo de população F_1 . A fase de ligação descreve a configuração dos alelos de cada par de locos heterozigóticos nos cromossomos homólogos em um único genitor, e pode ser determinada examinando-se como os alelos alternativos em um cromossomo homólogo são carregados em um genitor, para quaisquer dois marcadores ligados.

Esta metodologia é mais sofisticada para a construção de mapas de ligação em espécies não endogâmicas do que aquela implementada no programa JoinMap (VON OOIJEN e VOORRIPS, 2001) que utiliza a proposta de Maliepaard, Jansen e Van Ooijen (1997).

A metodologia proposta por Wu et al. (2002) distingue as fases de ligação entre os pares de marcadores com base na probabilidade *a posteriori* de cada fase de ligação, que permite estimar, simultaneamente, a frequência de recombinação e a fase de ligação, solucionando alguns problemas estatísticos apontados por Maliepaard, Jansen e Van Ooijen (1997) sendo, portanto, mais interessante para obtenção de mapas mais consistentes.

Segundo Garcia et al. (2006), outra vantagem desta metodologia é a possibilidade de se fazer uma análise multiponto utilizando a verossimilhança, que permite uma melhoria na obtenção de estimativas de frequência de recombinação entre os marcadores, contribuindo para a geração de mapas mais confiáveis. Estes autores empregaram a proposta de Wu et al. (2002) para construção de mapas de ligação numa população oriunda do cruzamento entre duas cultivares pré-comerciais de cana-de-açúcar ('SP80-180' e 'SP80-4966'). Dos 1118 marcadores de dose única empregados (RFLP, SSR e AFLP), 39% segregaram na proporção de 1:1 e os outros 61% na proporção de 3:1. Estes últimos foram utilizados para estabelecer a ligação entre os marcadores *testcross* (1:1). Foram encontrados 131 grupos de ligação com comprimento total 2.602,4 cM e densidade de marcadores de 7,3 cM. O programa JoinMap foi empregado para a construção de mapas de ligação nessa mesma população, sendo encontrados 98 grupos de ligação com comprimento total de 1.340 cM e 6,2 cM de densidade dos marcadores. Em suma, a metodologia de Wu et al. (2002), permitiu a formação de um maior número de grupos de ligação, construção de grupos de ligação com um maior número de marcadores e, por conseguinte, uma maior cobertura do genoma (duas vezes mais, em termos de comprimento de mapas), em relação ao JoinMap.

Diante dessas considerações, é perfeitamente justificável o uso da metodologia de Wu et al. (2002) na integração dos mapas de ligação anteriormente construídos com base na estratégia duplo pseudocruzamento teste e com marcadores dominantes para maracujá-amarelo (CARNEIRO et al., 2002; LOPES et al., 2006). Em síntese, esta é a hipótese na qual se baseia o presente trabalho, ou seja, de que o desenvolvimento dos marcadores microsatélites servirá como ponte para a geração de um mapa de ligação integrado de *P. edulis* f. *flavicarpa*.

3 Material e métodos

3.1 Material genético

Os acessos de maracujá-amarelo IAPAR-123 e IAPAR-06, cedidos pela Estação Experimental de Londrina do Instituto Agrônomo do Paraná (IAPAR), Londrina/PR, foram utilizados para o desenvolvimento da biblioteca genômica enriquecida para dinucleotídios (GA)_n e (CA)_n.

A população de mapeamento foi constituída por 160 plantas F₁ oriundas do cruzamento dos acessos IAPAR-123 e IAPAR-06, utilizados como genitores feminino e masculino, respectivamente. O acesso IAPAR-123 apresenta importância comercial com características agronômicas desejáveis, entre elas, certo grau de resistência à *Xanthomonas axonopodis* pv. *passiflorae*, agente causal da bacteriose do maracujazeiro, enquanto o acesso IAPAR-06, introduzido de Marrocos, apresenta suscetibilidade a esse patógeno e características agronômicas inferiores. A escolha dos genitores foi feita em função do seu contraste fenotípico e molecular, gerando uma população segregante.

De acordo com Moraes et al. (2005), esta população segregante também possui grande diversidade genética para produtividade, número de frutos por planta, peso médio dos frutos, percentagem de polpa e teor de sólidos solúveis, podendo ser usada em ciclos de seleção recorrente visando à melhoria destas características.

3.2 Extração e quantificação do DNA

O DNA genômico foi extraído de folhas dos acessos IAPAR-123 e IAPAR-06 utilizando o protocolo com CTAB descrito por Murray e Thompson (1980). Em seguida, o DNA foi submetido à eletroforese (3 V/cm) em gel de agarose 1% (p/v), sendo corado com brometo de etídeo (10 µg/ml). A quantificação foi feita por meio de comparações visuais da intensidade de fluorescência das bandas com àquelas de concentrações conhecidas de DNA do fago λ (Gibco) sob luz ultravioleta.

3.3 Construção da biblioteca genômica enriquecida com microssatélites

A construção da biblioteca genômica foi feita utilizando o protocolo desenvolvido por Kijas et al. (1994), com algumas modificações conforme descrito em Oliveira et al. (2005), o qual será descrito brevemente a seguir.

- Restrição do DNA

O DNA genômico (5µg) foi digerido por hidrólise enzimática, adicionando-se na reação 10,0 µL de tampão de reação 10X, 6,0 µL da enzima de restrição *RsaI* (10u/µL, Invitrogen), 10,0 µL de espermidina 40 mM e água Mili-Q para completar 100,0 µL. O material foi incubado a 37 °C por 3 horas para restrição e depois por 15 minutos a 70 °C para inativar as endonucleases de restrição. Os fragmentos foram ligados a adaptadores específicos.

- Ligação dos adaptadores

Este protocolo requer duas etapas de PCR para aumentar a quantidade de fragmentos na reação, aumentando as chances de sucesso na seleção de seqüências com microssatélites. Assim, para atender a esta etapa, os fragmentos digeridos foram ligados a adaptadores específicos.

Para a ligação dos adaptadores *Rsa21* (5' CTCTTGCTTACGCGTGGACTA 3') e *Rsa25* (5' TAGTCCACGCGTAAGCAAGAGCACA 3'), a reação foi feita com: 1µg do DNA digerido na etapa anterior, 10 µL de tampão 5X [Promega - (150 mM Tris-HCl (pH 7,8), 50 mM de MgCl₂, 50 mM de DTT e 5 mM de ATP)], 2 µL de cada adaptador (10 µM), 4 unidades de T4 DNA ligase (Promega) e água para completar 50 µL. A reação foi incubada por 2 horas a 20 °C.

- Pré-amplificação

Os fragmentos ligados aos adaptadores foram pré-amplificados usando-se o *primer Rsa21*. A reação foi feita adicionando 3 µL do produto da ligação, 2 µL de *primer* (10 µM), 5 µL de tampão 10X [Fermentas - (100 mM Tris-HCl (pH 8,8), 500 mM de KCl e 0,8% de Nonidet P40)], 1,5 µL de MgCl₂ (50 mM), 4 µL de dNTP (2,5 mM), 3 unidades de Taq DNA polimerase (Fermentas) e água Mili-Q para completar 50 µL. Os ciclos de amplificação por PCR foram feitos conforme descrito a seguir: 95 °C por 4 min, seguindo-se 20 ciclos de 94 °C por 30 s, 60 °C por 1 min, 72 °C por 2 min e extensão final de 72 °C por 8 min.

Após a pré-amplificação, fragmentos entre 300 e 800 pb (pares de bases) foram selecionados de um gel de agarose a 0,8 %, preparado com TBE 1X (100 mM Tris, 100 mM ácido bórico e 2 mM EDTA), utilizando-se como marcador de peso molecular 100 ng de *Ladder* 100 pb DNA (Invitrogen). Para recuperar o DNA digerido do gel de agarose, foi utilizado o kit Wizard PCR Preps DNA Purification System (Promega).

- Seleção dos fragmentos contendo microssatélites

Após a ligação do DNA aos adaptadores, foi feita uma seleção dos fragmentos contendo seqüências de microssatélites. Os fragmentos contendo seqüências ricas em GA e CA foram selecionados por meio da hibridização com oligonucleotídeos biotinilados biotina-III(CT)_8 e biotina-III(GT)_8 complementares às seqüências GA e CA, respectivamente, os quais foram recuperados com estreptavidinas ligadas a contas magnéticas.

Como as contas magnéticas marcadas com estreptavidina precisam de uma preparação prévia, homogeneizou-se um tubo do produto Streptavidin Magnesfere Paramagnetic Particle (Promega) usando vortex. Em seguida, os 600 μL do produto foram magnetizados (deixados aderir à parede) por 30 segundos, sendo descartado o sobrenadante. Após este processo, foram feitas três lavagens com 300 μL de SSC 0,5X. A cada lavagem, descartou-se o sobrenadante, segurando as esferas magnéticas na parede do tubo com auxílio de um rack Magna Separator Magnetic Particle (Promega). No final, as esferas foram ressuspensas em 100 μL de SSC 0,5X.

Para o preparo do DNA, juntou-se 400 μL de água aos 100 μL de DNA purificado. Esta solução foi incubada a 95°C por 15 min. Em seguida, adicionaram-se 13 μL de SSC 20X e 3 μL de cada sonda de microssatélite biotinilado - $\text{Biotina-III(CT)}_8 + \text{Biotina-III(GT)}_8$ - (50 μM), sendo hibridizados à temperatura de 60 °C por 20 minutos, com agitação suave a cada 2 minutos.

Os 100 μL de esferas pré-lavadas foram misturadas com os 519 μL de mistura de hibridização, sendo incubados por 10 minutos à temperatura ambiente agitando suavemente o tempo todo. Após este procedimento, a solução foi magnetizada por 30 segundos, o sobrenadante descartado e a solução ressuspensa em 300 μL de SSC 0,1X. Esta última etapa foi repetida por três vezes.

Na última etapa, a solução foi ressuspensa em 100 μL de água e magnetizada por 30 segundos. O sobrenadante foi reservado em um tubo *ependorf* e a solução novamente

ressuspendida e magnetizada com 150 μL de água. As duas partes foram misturadas num volume final de 250 μL e conservadas a $-20\text{ }^{\circ}\text{C}$.

- Amplificação dos fragmentos selecionados, clonagem e transformação

A seguir, os fragmentos foram amplificados pela técnica de PCR utilizando o *primer Rsa21*, nas seguintes condições: 10 μL dos fragmentos selecionados pelas contas magnéticas, 4 μL de *primer* (10 μM), 10 μL de tampão 10X (Fermentas), 3 μL de MgCl_2 (50 mM), 8 μL de dNTP (2,5 mM), 1 unidade de Taq DNA polimerase (Fermentas) e água Mili-Q para completar 100 μL . A reação de PCR foi feita conforme descrito a seguir: $95\text{ }^{\circ}\text{C}$ por 1 min, seguindo-se 20 ciclos de $94\text{ }^{\circ}\text{C}$ por 40 s, $60\text{ }^{\circ}\text{C}$ por 1 min, $72\text{ }^{\circ}\text{C}$ por 2 min e extensão final de $72\text{ }^{\circ}\text{C}$ por 5 min.

Em seguida, os insertos foram clonados no vetor pGEMT (Promega), utilizando-se 1 μL do vetor, 10 μL de tampão 2X [Promega - (60 mM Tris-HCl (pH 7,8), 20 mM de MgCl_2 , 20 mM de DTT, 2 mM de ATP e 10% de polietilenoglicol)], 1 μL de T4 DNA ligase e 8 μL do produto da amplificação. Sendo a reação incubada *overnight* por $4\text{ }^{\circ}\text{C}$.

A transformação foi feita em células super-competentes XL1-Blue. Cerca de 100 μL do produto da transformação foram plaqueados em placas de Petri contendo meio LB, X-Gal e IPTG, e foram incubadas a $37\text{ }^{\circ}\text{C}$ por 14 horas.

- Amplificação dos insertos clonados contendo microssatélites

Para verificar a presença dos insertos nas colônias transformadas (colônias brancas), foram feitas reações de PCR com o *primer Rsa21*. Para isso, as colônias foram coletadas com auxílio de um palito, sendo mergulhado em placas de PCR contendo 15 μL de água Mili-Q. Uma réplica desta placa de PCR foi feita em placas contendo meio LB com 50 $\mu\text{g/mL}$ de ampicilina. As reações de PCR foram feitas com: 15 μL de água Mili-Q autoclavada (onde as colônias foram inoculadas), 2,5 μL de tampão 10X (Fermentas), 1 μL de MgCl_2 (50 mM), 2 μL de dNTP (2,5 mM), 0,625 μL do *primer Rsa21* (10 μM), 1 unidade de Taq DNA polimerase (Fermentas) e água para completar 25 μL . A reação de PCR foi feita com desnaturação inicial de $95\text{ }^{\circ}\text{C}$ por 4 min, seguida de 30 ciclos a $94\text{ }^{\circ}\text{C}$ por 30 s, $52\text{ }^{\circ}\text{C}$ por 45 s, $72\text{ }^{\circ}\text{C}$ por 90 s e uma extensão final de $72\text{ }^{\circ}\text{C}$ por 8 min.

- *Seleção e seqüenciamento dos clones positivos*

Após a seleção das colônias contendo os fragmentos clonados na etapa anterior, foi feita a extração dos plasmídeos (miniprep) pelo método de lise alcalina. Em seguida, os clones foram seqüenciados utilizando o kit DYEnamic ET Terminator (Amersham Biosciences) com os *primers* universais M13 *forward* e/ou *reverse*.

As amostras foram seqüenciadas utilizando o seguinte programa: 25 ciclos de 95 °C por 20 s, 50 °C por 15 s e 60 °C por 60 s. A reação foi purificada e as amostras foram submetidas a corrida em seqüenciador MegaBACE 1000 (Amersham Biosciences) ou ABI 377 (Applied Biosystems). Quando se utilizou o MegaBace 1000, três corridas de seqüenciamento foram feitas com diferentes condições de injeção das amostras para a obtenção de um maior número de seqüências com boa qualidade².

- *Análise das seqüências*

Para a obtenção de seqüências adequadas para o desenho dos *primers*, foi utilizado o programa computacional *Staden Package*³. A análise das seqüências foi feita a partir de diferentes corridas de um mesmo clone, obtido com um *primer* M13 *forward*, verificando a qualidade das bases e corrigindo possíveis problemas decorrentes do seqüenciamento. Quando se dispôs dessa mesma seqüência, obtida com o *primer* M13 *reverse*, foi possível a formação dos *contigs*.

Ao final desse processo, esse mesmo *software* foi utilizado para verificar a existência de redundância na biblioteca, sendo que seqüências com similaridade maior do que 90% foram consideradas redundantes.

Para a identificação das seqüências contendo microssatélites, utilizou-se o programa *Tandem Repeat Finder* – TRF (BENSON, 1999). Os *primers* foram desenhados para seqüências com no mínimo 15, 6, 4 e 3 repetições do motivo do microssatélite para mono, di, tri e tetranucleotídeos (e acima), respectivamente.

² Esta etapa foi realizada no Laboratório de Genoma da Universidade Estadual de Londrina, sob supervisão da Profa. Maria Helena Pelegrinelli Fungaro

³ Fonte: <<http://staden.sourceforge.net>>. Acesso em: 10 fev. 2005.

- Desenho dos primers

Os pares de *primers* complementares às seqüências flanqueando os elementos de repetição foram desenhados utilizando o programa PRIMER3 (ROZEN e SKALETSKY, 2000). Foram utilizados os seguintes critérios para este desenho: i) temperatura de anelamento (TA) variando de 52 a 60 °C e (2) diferença máxima de 2 °C na TA entre pares de *primers*; iii) conteúdo GC entre 40 e 60%. Os *primers* foram sintetizados pelas empresas Invitrogen, MWG Oligo Synthesis Report ou IDT (Promega).

3.4 Otimização e genotipagem dos genitores da população F₁ com marcadores microssatélites

Cada reação de PCR foi otimizada para amplificar o alelo esperado de cada seqüência identificada. Para isso, foi usado o DNA dos dois genitores do cruzamento. O tamanho dos fragmentos foi determinado com o uso do marcador de peso molecular *Ladder* 10 pb (Invitrogen). Foram preparados quatro diferentes *mix* de reação para a otimização das reações de PCR conforme o *primer* (Tabela 2) e duas diferentes condições de amplificação em esquema de *touch-down*, com base nos programas TD60 e TD56 (Tabela 3).

Tabela 2 - Concentração dos reagentes e relação dos *mix* usados para a otimização das reações de PCR de acordo com os *primers* de microssatélites de maracujá-amarelo

Reagentes	Mix 2	Mix 4	Mix 5	Mix 7
Tampão*	1X	2X	1X	1X
Primer F + R**	0,3 µM	0,3 µM	0,3 µM	0,3 µM
DNA	20 ng	20 ng	20 ng	20 ng
MgCl ₂	1,5 mM	1,5 mM	2,5 mM	2,5 mM
dNTP	200 µM	200 µM	200 µM	350 µM
Taq Polimerase	0,5 U	0,5 U	0,5 U	0,5 U
Água Mili-Q	9,7 µL	7,7 µL	8,9 µL	7,7 µL

* O tampão dos *mix* para a concentração 1X continha 10mM de Tris-HCl pH 8,8; 50 mM de KCl e 0,8% de Nonidet P40;

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

Os reagentes que variaram entre os *mix* foram a concentração do tampão, o $MgCl_2$ e o dNTP.

Para alguns pares de *primers*, a utilização de uma temperatura de anelamento específica produziu melhores resultados, sendo utilizadas as temperaturas de 60, 56 e 52 °C. Assim, o programa básico foi constituído por uma desnaturação inicial de 94 °C por 5 min e 30 ciclos de 94 °C por 40 s (60, 56 ou 52 °C, de acordo com o *primer*), 72 °C por 50 s e uma extensão final de 5 min a 72 °C.

Tabela 3 - Condições de amplificação usadas para a otimização das reações de PCR de acordo com os *primers* de microssatélites de maracujá-amarelo

Programa TD60			Programa TD56		
Nº de ciclos	Temperatura	Tempo	Nº de ciclos	Temperatura	Tempo
1	94 °C	5 min	1	94 °C	5 min
8	94 °C	40 s	12	94 °C	40 s
	60°C - 0,5 °C por ciclo	40 s		56°C - 0,5 °C por ciclo	40 s
	72 °C	50 s		72 °C	50 s
24	94 °C	40 s	20	94 °C	40 s
	56 °C	40 s		56 °C	40 s
	72 °C	50 s		72 °C	50 s
1	72 °C	5 min	1	72 °C	5 min
1	8 °C	∞	1	8 °C	∞

As amplificações foram feitas em termociclador PCR Express Hybaid (Amersham Pharmacia Biotech), PTC-100 (MJ Research) ou em GeneAmp PCR System 9700 (Applied Biosystems). A visualização dos produtos de PCR foi feita em gel de agarose 3% (p/v) para a escolha do melhor *mix* de reação, bem como da condição de amplificação.

Após a escolha desses parâmetros, foram feitas novas reações com o DNA dos dois genitores do cruzamento e mais seis plantas da população segregante para verificar o possível padrão de segregação dos locos de microssatélites. Os produtos resultantes destas reações foram

separados em gel desnaturante 6% de poliacrilamida + 7 M de uréia em tampão TBE 1X e visualizados por coloração com prata (CRESTE, TULMANN NETO e FIGUEIRA, 2001). A genotipagem da população segregante F₁ foi feita seguindo os mesmos procedimentos.

3.5 Genotipagem com marcadores AFLP

Os dados de AFLP foram gerados por Lopes et al. (2006) e Matta (2005). Brevemente, o DNA genômico foi extraído de folhas novas e sua digestão foi feita com duas combinações de enzimas, *EcoRI/MseI* e *PstI/MseI*. Para a digestão com *EcoRI/MseI* foram usados 300 ng de DNA, 5,0 µL de tampão 10X *One Phor All* (Amersham Biosciences), 0,5 µL de BSA (10 µg/µL), 0,5 µL da enzima *MseI* (10 u/µL, New England Biolabs) e 0,5 µL da enzima *EcoRI* (10 u/µL; Gibco). A reação foi ajustada para um volume de 50 µL. O mesmo procedimento foi realizado para a digestão do DNA com *PstI/MseI* com exceção do tampão *React1* 10X (Gibco) e da enzima *PstI* (10 u/µL; Gibco). As reações de digestão foram incubadas a 37 °C por 3 h e inibidas a 70 °C por 15 min.

Em seguida, foi feita a ligação dos adaptadores utilizando-se 2,0 µL de tampão 5X da T4 DNA ligase (Invitrogen), 1,0 µL dos adaptadores das enzimas de corte raro (*EcoRI* ou *PstI*), 1,0 µL dos adaptadores da enzima de corte freqüente (*MseI*), 1,0 µL da enzima T4 DNA ligase (1 u/µL, Invitrogen), 5,0 µL de água Mili-Q e 40 µL da solução de digestão. As reações de ligação foram incubadas a 20 °C por 3 horas.

A amplificação pré-seletiva foi feita usando-se um nucleotídeo 3' seletivo e três combinações (E+A/M+C, E+T/M+G e P+A/M+C)⁴. A reação foi feita com 3,0 µL do DNA digerido e ligado, 0,5 µL dos *primers* complementares aos adaptadores para as enzimas de corte raro (E+A, E+C ou P+A) (50 ng/µL), 0,5 µL dos *primers* complementares aos adaptadores para a enzima de corte freqüente (M+C ou M+G) (50 ng/µL), 4,0 µL de dNTP (2,5 mM), 2,0 µL de tampão 10X [Promega - (500 mM KCl, 100 mM de Tris-HCl (pH 9,0) e 1% de Triton[®] X-100)], 1,2 µL de MgCl₂ (25 mM), 3 unidade de Taq DNA polimerase (Promega) e 8,2 µL de água Mili-Q. As reações foram feitas nas seguintes condições: 94 °C por 2 min, seguindo-se 26 ciclos de 94

⁴ E+A/M+C = *EcoRI* + adenina/*MseI* + citosina; E+T/M+G = *EcoRI* + timina/*MseI* + guanina e P+A/M+C = *PstI* + adenina/*MseI* + citosina;

°C por 1 min, 56 °C por 1 min e 72 °C por 1 min, e uma extensão final de 72 °C por 5 min. Os produtos pré-amplificados foram diluídos em 80 µL de água Mili-Q.

A amplificação seletiva foi realizada com 1,5 µL dos produtos pré-amplificados e diluídos, 0,5 µL dos *primers* (E+NNN ou P+NNN) (50 ng/µL), 0,6 µL dos *primers* (M+NNN) (50 ng/µL), 1,6 µL de dNTP (2,5 mM), 2,0 µL de tampão 10X (Promega), 1,2 µL de MgCl₂ (25 mM), 1,6 unidade de DNA polimerase (Promega) e 12,28 µL de água Mili-Q. As bases seletivas, bem como os códigos de identificação das 34 combinações de AFLP são apresentadas na Tabela 4.

Tabela 4 - Códigos atribuídos às 34 combinações de iniciadores AFLP usadas para as amplificações seletivas. Em negrito, estão os 3 nucleotídeos seletivos adicionados à extremidade 3' dos iniciadores, de acordo com Matta (2005)

Códigos	Combinações
EM01	5'- GACTGCGTACCAATTC AAA -3' + 5'- GATGAGTCCTGAGTAA CGC -3'
EM02	5'- GACTGCGTACCAATTC AAA -3' + 5'- GATGAGTCCTGAGTAA CTA -3'
EM03	5'- GACTGCGTACCAATTC AAC -3' + 5'- GATGAGTCCTGAGTAA CAA -3'
EM04	5'- GACTGCGTACCAATTC AAC -3' + 5'- GATGAGTCCTGAGTAA CGT -3'
EM05	5'- GACTGCGTACCAATTC AAG -3' + 5'- GATGAGTCCTGAGTAA CAA -3'
EM06	5'- GACTGCGTACCAATTC AAG -3' + 5'- GATGAGTCCTGAGTAA CAC -3'
EM07	5'- GACTGCGTACCAATTC ACC -3' + 5'- GATGAGTCCTGAGTAA CTC -3'
EM08	5'- GACTGCGTACCAATTC ACC -3' + 5'- GATGAGTCCTGAGTAA CTG -3'
EM09	5'- GACTGCGTACCAATTC ACC -3' + 5'- GATGAGTCCTGAGTAA CTT -3'
EM10	5'- GACTGCGTACCAATTC ACG -3' + 5'- GATGAGTCCTGAGTAA CGC -3'
EM11	5'- GACTGCGTACCAATTC ACG -3' + 5'- GATGAGTCCTGAGTAA CTT -3'
EM12	5'- GACTGCGTACCAATTC ACT -3' + 5'- GATGAGTCCTGAGTAA CGC -3'
EM13	5'- GACTGCGTACCAATTC AGG -3' + 5'- GATGAGTCCTGAGTAA CAA -3'
EM14	5'- GACTGCGTACCAATTC AGG -3' + 5'- GATGAGTCCTGAGTAA CAG -3'
EM15	5'- GACTGCGTACCAATTC AGG -3' + 5'- GATGAGTCCTGAGTAA CTT -3'
EM16	5'- GACTGCGTACCAATTC ACA -3' + 5'- GATGAGTCCTGAGTAA CAA -3'
EM17	5'- GACTGCGTACCAATTC ACA -3' + 5'- GATGAGTCCTGAGTAA CAC -3'
EM18	5'- GACTGCGTACCAATTC ACA -3' + 5'- GATGAGTCCTGAGTAA CTC -3'
EM19	5'- GACTGCGTACCAATTC AGA -3' + 5'- GATGAGTCCTGAGTAA CAG -3'
EM20	5'- GACTGCGTACCAATTC TAC -3' + 5'- GATGAGTCCTGAGTAA GGT -3'
EM21	5'- GACTGCGTACCAATTC TAG -3' + 5'- GATGAGTCCTGAGTAA GCT -3'
EM22	5'- GACTGCGTACCAATTC TAT -3' + 5'- GATGAGTCCTGAGTAA GGC -3'
EM23	5'- GACTGCGTACCAATTC TTA -3' + 5'- GATGAGTCCTGAGTAA GGT -3'
EM24	5'- GACTGCGTACCAATTC AGA -3' + 5'- GATGAGTCCTGAGTAA CAA -3'
PM01	5'- GACTGCGTACATGCAG ACA -3' + 5'- GATGAGTCCTGAGTAA CAC -3'
PM02	5'- GACTGCGTACATGCAG ACA -3' + 5'- GATGAGTCCTGAGTAA CAG -3'
PM03	5'- GACTGCGTACATGCAG ACC -3' + 5'- GATGAGTCCTGAGTAA CAC -3'
PM04	5'- GACTGCGTACATGCAG ACT -3' + 5'- GATGAGTCCTGAGTAA CAG -3'
PM05	5'- GACTGCGTACATGCAG AGC -3' + 5'- GATGAGTCCTGAGTAA CGT -3'
PM06	5'- GACTGCGTACATGCAG AAC -3' + 5'- GATGAGTCCTGAGTAA ACTA -3'
PM07	5'- GACTGCGTACATGCAG ACG -3' + 5'- GATGAGTCCTGAGTAA CTG -3'
PM08	5'- GACTGCGTACATGCAG AGC -3' + 5'- GATGAGTCCTGAGTAA CTA -3'
PM09	5'- GACTGCGTACATGCAG AGA -3' + 5'- GATGAGTCCTGAGTAA CAC -3'
PM10	5'- GACTGCGTACATGCAG AAA -3' + 5'- GATGAGTCCTGAGTAA CAC -3'

Nesta codificação, os dois primeiros caracteres são as letras que correspondem às enzimas de restrição (*EcoR*1 = E, *Pst*1 = P e *Mse*1 = M), o terceiro e o quarto caracteres são números que identificam as combinações dos nucleotídeos seletivos. O código de identificação dos locos AFLP é composto pelos quatro caracteres que identificam a combinação seguidos do tamanho do fragmento do loco em pares de base (veja Matta, 2005).

Foram avaliadas 174 combinações de *primers* na amplificação seletiva, sendo 90 E+ANN/M+CNN, 28 E+TNN/M+GNN e 56 P+ANN/M+CNN. Foi feita uma avaliação preliminar das combinações dos *primers* com os dois genitores e duas plantas da população segregante. Assim, foram analisadas 34 combinações para o mapeamento, sendo 20 combinações E+ANN/M+CNN, 4 E+TNN/M+GNN e 10 P+ANN/M+CNN. As reações de PCR foram realizadas sob as seguintes condições: 94 °C por 2 min, seguido de 12 ciclos a 94 °C por 30 s, 65 °C por 30 s (-0,7 s por ciclo), 72 °C por 1 min, mais 23 ciclos nas mesmas condições, modificando-se somente a temperatura de anelamento para 56 °C, mais uma extensão final a 72 °C por 2 min.

Os produtos das reações de AFLP (20 µL) foram misturados com 10 µL de uma solução de formamida (98% de formamida; 10 mM de EDTA pH 8,0; 0,002% de azul de bromofenol e 0,002% de xileno cianol). Esta mistura foi desnaturada a 95 °C por 5 min sendo aplicados 14 µL nos poços dos géis. O tampão TBE 1X foi preparado para a pré-corrida dos géis de poliacrilamida 6%, (acrilamida/bisacrilamida (19:1) e 7,5 M de uréia) no equipamento Sequi-Gen GT (BioRad) durante 1h a 80 W. Após esse período, as amostras foram aplicadas, e os géis foram submetidos a 80 W por 4 h e 30 min, sendo em seguida corados com nitrato de prata (CRESTE, TULMANN NETO e FIGUEIRA, 2001).

3.6 Construção dos mapas de ligação

Para construção dos mapas, foi usada a metodologia de Wu et al. (2002), que será apresentada em detalhes. Algumas modificações foram implementadas, principalmente no que se refere aos métodos de ordenação.

3.6.1 Tipos de segregações com base em marcas dominantes e codominantes

Nesse item, será apresentada a notação, bem como a classificação que foi adotada para os marcadores. Cada alelo marcador foi simbolizado por a , b , c e d , sendo codominante em relação aos outros e dominantes para os alelos nulos, simbolizados por o . Os cruzamentos foram denotados de forma que sempre o primeiro genitor indicado é o feminino. Considerando um cruzamento entre dois acessos heterozigóticos e marcadores dominantes AFLP, podem ser observados três tipos de segregação: a) Segregação do tipo $ao \times oo$, ou seja, loco informativo para o genitor feminino, cuja segregação na progênie é do tipo 1:1; b) Segregação do tipo $oo \times ao$, ou seja, informativo para o genitor masculino, cuja segregação na progênie é do tipo 1:1 e; c) Segregação do tipo $ao \times ao$, ou seja, loco heterozigótico em ambos genitores, cuja segregação esperada na progênie é do tipo 3:1.

No caso dos microssatélites, é possível identificar locos com até quatro alelos segregando, além de alelos nulos, todos segregando em um único loco. Além disso, o número de alelos pode variar entre os locos. Dependendo de como os diferentes alelos são arranjados nos dois genitores do cruzamento, existe um total de 18 possíveis tipos de cruzamentos para os locos marcadores, conforme apresentado na Tabela 5.

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 e VAN OOIJEN, 1997; WU et al., 2002):

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 que incluem três subgrupos:

B_1) um genitor com dois diferentes alelos dominantes e o outro com um loco 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 .

Tabela 5 - Possíveis genótipos dos marcadores microsatélites, padrões de alelos observados nos genitores e em suas progênie, oriundos do cruzamento de genitores não endogâmicas (Modificado de WU et al., 2002 e MALIEPAARD, JANSEN e VAN OOIJEN, 1997)

Tipo de cruzamento		Genitores			Progênie		
		Cruzamento	Alelos observados	Obs.: ¹	Alelos observados	Segregação	Nº de fenótipos moleculares
<i>A</i>	1	$ab \times cd$	$ab \times cd$	A	ac, ad, bc, bd	1:1:1:1	4
	2	$ab \times ac$	$ab \times ac$	A	a, ac, ba, bc	1:1:1:1	4
	3	$ab \times co$	$ab \times c$	A	ac, a, bc, b	1:1:1:1	4
	4	$ao \times bo$	$a \times b$	A	ab, a, b, o	1:1:1:1	4
<i>B</i>	B_1 5	$ab \times ao$	$ab \times a$	A	$ab, 2a, b$	1:2:1	3
	B_2 6	$ao \times ab$	$a \times ab$	A	$ab, 2a, b$	1:2:1	3
	B_3 7	$ab \times ab$	$ab \times ab$	S	$a, 2ab, b$	1:2:1	3
<i>C</i>	8	$ao \times ao$	$a \times a$	S	$3a, o$	3:1	2
<i>D</i>	D_1 9	$ab \times cc$	$ab \times cc$	A	ac, bc	1:1	2
	10	$ab \times aa$	$ab \times a$	A	a, ab	1:1	2
	11	$ab \times oo$	$ab \times o$	A	a, b	1:1	2
	12	$bo \times aa$	$b \times a$	A	ab, a	1:1	2
	13	$ao \times oo$	$a \times o$	A	a, o	1:1	2
	D_2 14	$cc \times ab$	$c \times ab$	A	ac, bc	1:1	2
	15	$aa \times ab$	$a \times ab$	A	a, ab	1:1	2
	16	$oo \times ab$	$o \times ab$	A	a, b	1:1	2
	17	$aa \times bo$	$a \times b$	A	ab, a	1:1	2
	18	$oo \times ao$	$o \times a$	A	a, o	1:1	2

¹ A- marcador assimétrico e S marcador simétrico

Os marcadores do tipo B_3 e C têm o mesmo genótipo em ambos os genitores e por isso são chamados de marcadores simétricos, enquanto os outros possuem genótipo específico para cada genitor, sendo chamados de assimétricos (Tabela 5).

De acordo com Wu et al. (2002), denominando um genitor por P o outro por Q , e os dois cromossomos homólogos de cada genitor por P_1 e P_2 e por Q_1 e Q_2 , respectivamente, então, para uma marca particular, independente do tipo de marcador, são possíveis quatro pareamentos entre os cromossomos dos genitores (PCG) na progênie: P_1Q_1 , P_1Q_2 , P_2Q_1 e P_2Q_2 . Por isso, os quatro grupos classificados na Tabela 5 são definidos em termos da correspondência entre os PCG e os fenótipos moleculares observados.

Para proceder à análise entre dois marcadores arbitrários, todas as possíveis combinações dos sete grupos de marcas foram consideradas. Teoricamente, são 28 possíveis combinações: A vs. A , A vs. B_1 , A vs. B_2 , A vs. B_3 , A vs. C , A vs. D_1 , A vs. D_2 , B_1 vs. B_1 , B_1 vs. B_2 , B_1 vs. B_3 , B_1 vs. C , B_1 vs. D_1 , B_1 vs. D_2 , B_2 vs. B_2 , B_2 vs. B_3 , B_2 vs. C , B_2 vs. D_1 , B_2 vs. D_2 , B_3 vs. B_3 , B_3 vs. C , B_3 vs. D_1 , B_3 vs. D_2 , C vs. C , C vs. D_1 , C vs. D_2 , D_1 vs. D_1 , D_1 vs. D_2 e D_2 vs. D_2 .

3.6.2 Análises de distorções da segregação esperada

Os locos genotipados foram submetidos ao teste de aderência de “qui-quadrado” (χ^2) para a verificação do padrão de segregação, sendo obtidos os p-valores de cada teste, com base na hipótese H_0 de se obter um valor igual ou superior ao χ^2 calculado tomando-se por base a amostra dos dados.

Como são feitos múltiplos testes, é necessário considerar um nível de significância conjunto. Assim, foi utilizado o critério da “razão de falsas descobertas” (FDR - *False Discovery Rate*) proposto por Storey e Tibshirani (2003), para estabelecer o nível de significância conjunto dos testes. Neste caso, foram obtidos para cada loco, individualmente, o valor da estatística χ^2 e o correspondente valor de P , sendo que estes valores foram ordenados em ordem crescente e a posição do teste individual do *rank* de probabilidades representado pela letra i . Foi adotado como ponto de corte o nível $\alpha=0,05$ para o conjunto dos t testes. O valor de i a partir do qual foi rejeitada a hipótese H_0 foi determinado a partir da equação:

$$\alpha \geq \frac{tP_i}{i} \quad (1)$$

As marcas que apresentaram distorção das proporções mendelianas no teste de aderência não foram utilizadas na construção dos mapas de ligação. Foram descartados também locos com mais de 10% de observações perdidas. Tais critérios também foram adotados por Moraes (2005) na construção de mapas de ligação de maracujá-amarelo usando marcas de AFLP.

3.6.3 Codificação dos cromossomos dos genitores

Considerando um grupo de ligação com m marcadores, $M^1, M^2, \dots, M^m$, os cromossomos dos genitores podem ser arbitrariamente codificados usando seus respectivos alelos. Cada par de cromossomo para o primeiro marcador foi identificado como $P_1||P_2$ e $Q_1||Q_2$, sendo que $||$ representa as duas cromátides dos cromossomos homólogos. A fase de ligação entre os alelos do primeiro (M^1) e segundo marcador (M^2) foi determinada pelo arranjo dos alelos alternativos do segundo marcador. Têm-se, então, quatro arranjos no total ($A_\omega, \omega = 1, \dots, 4$) para os dois genitores, como mostrado a seguir:

$$\begin{array}{ccc}
 P & \times & Q \\
 \\
 A_1 & \begin{array}{c} P_1^1 \parallel P_2^1 \\ P_1^2 \parallel P_2^2 \end{array} & \begin{array}{c} Q_1^1 \parallel Q_2^1 \\ Q_1^2 \parallel Q_2^2 \end{array} \\
 \\
 A_2 & \begin{array}{c} P_1^1 \parallel P_2^1 \\ P_1^2 \parallel P_2^2 \end{array} & \begin{array}{c} Q_1^1 \parallel Q_2^1 \\ Q_2^2 \parallel Q_1^2 \end{array} \\
 \\
 A_3 & \begin{array}{c} P_1^1 \parallel P_2^1 \\ P_2^2 \parallel P_1^2 \end{array} & \begin{array}{c} Q_1^1 \parallel Q_2^1 \\ Q_1^2 \parallel Q_2^2 \end{array} \\
 \\
 A_4 & \begin{array}{c} P_1^1 \parallel P_2^1 \\ P_2^2 \parallel P_1^2 \end{array} & \begin{array}{c} Q_1^1 \parallel Q_2^1 \\ Q_2^2 \parallel Q_1^2 \end{array}
 \end{array}$$

em que A_1 representa os dois marcadores ligados em associação/associação, A_2 ligados em associação/repulsão, A_3 em repulsão/associação e A_4 em repulsão/repulsão.

3.6.4 Algoritmos utilizados na determinação das fases de ligação e frequência de recombinação

Assumiu-se que os dois genitores têm a mesma frequência de recombinação para os marcadores correspondentes, sendo $\mathfrak{g}^{12}$ a frequência de recombinação entre o primeiro e o segundo marcador, M^1 e M^2 , por exemplo. A probabilidade condicional dos genótipos PCG de M^2 na progênie, dado o genótipo PCG de M^1 é a probabilidade de transição dos eventos de recombinação (h_{ω}^{12}) entre os dois marcadores num dado arranjo. A probabilidade de transição e o número de eventos de recombinação (d_{ω}^{12}) que ocorrem entre os dois marcadores num dado ω -ésimo arranjo A_{ω} depende de cada arranjo. A matriz h_{ω}^{12} foi então construída para cada par de marcas assumindo-se um determinado arranjo (A_1, A_2, A_3 e A_4). Assim, h_{ω}^{12} e d_{ω}^{12} num dado ω -ésimo arranjo A_{ω} são expressos em forma matricial em H_{ω}^{12} e D_{ω}^{12} :

$\Rightarrow$ Arranjo A_1

$$H_{\omega}^{12} = \begin{matrix} & \begin{matrix} P_1^2 Q_1^2 & P_1^2 Q_2^2 & P_2^2 Q_1^2 & P_2^2 Q_2^2 \end{matrix} \\ \begin{matrix} P_1^1 Q_1^1 \\ P_1^1 Q_2^1 \\ P_2^1 Q_1^1 \\ P_2^1 Q_2^1 \end{matrix} & \begin{bmatrix} (1 - \mathfrak{g}^{12})^2 & \mathfrak{g}^{12}(1 - \mathfrak{g}^{12}) & \mathfrak{g}^{12}(1 - \mathfrak{g}^{12}) & (\mathfrak{g}^{12})^2 \\ \mathfrak{g}^{12}(1 - \mathfrak{g}^{12}) & (1 - \mathfrak{g}^{12})^2 & (\mathfrak{g}^{12})^2 & \mathfrak{g}^{12}(1 - \mathfrak{g}^{12}) \\ \mathfrak{g}^{12}(1 - \mathfrak{g}^{12}) & (\mathfrak{g}^{12})^2 & (1 - \mathfrak{g}^{12})^2 & \mathfrak{g}^{12}(1 - \mathfrak{g}^{12}) \\ (\mathfrak{g}^{12})^2 & \mathfrak{g}^{12}(1 - \mathfrak{g}^{12}) & \mathfrak{g}^{12}(1 - \mathfrak{g}^{12}) & (1 - \mathfrak{g}^{12})^2 \end{bmatrix} \end{matrix}$$

$$D_{\omega}^{12} = \begin{matrix} & \begin{matrix} P_1^2 Q_1^2 & P_1^2 Q_2^2 & P_2^2 Q_1^2 & P_2^2 Q_2^2 \end{matrix} \\ \begin{matrix} P_1^1 Q_1^1 \\ P_1^1 Q_2^1 \\ P_2^1 Q_1^1 \\ P_2^1 Q_2^1 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

$\Rightarrow$ Arranjo A_2

$$H_{\omega}^{12} = \begin{matrix} & P_1^2 Q_1^2 & P_1^2 Q_2^2 & P_2^2 Q_1^2 & P_2^2 Q_2^2 \\ \begin{matrix} P_1^1 Q_1^1 \\ P_1^1 Q_2^1 \\ P_2^1 Q_1^1 \\ P_2^1 Q_2^1 \end{matrix} & \begin{bmatrix} g^{12}(1-g^{12}) & (1-g^{12})^2 & (g^{12})^2 & g^{12}(1-g^{12}) \\ (1-g^{12})^2 & g^{12}(1-g^{12}) & g^{12}(1-g^{12}) & (g^{12})^2 \\ (g^{12})^2 & g^{12}(1-g^{12}) & g^{12}(1-g^{12}) & (1-g^{12})^2 \\ g^{12}(1-g^{12}) & (g^{12})^2 & (1-g^{12})^2 & g^{12}(1-g^{12}) \end{bmatrix} \end{matrix}$$

$$D_{\omega}^{12} = \begin{matrix} & P_1^2 Q_1^2 & P_1^2 Q_2^2 & P_2^2 Q_1^2 & P_2^2 Q_2^2 \\ \begin{matrix} P_1^1 Q_1^1 \\ P_1^1 Q_2^1 \\ P_2^1 Q_1^1 \\ P_2^1 Q_2^1 \end{matrix} & \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix} \end{matrix}$$

$\Rightarrow$ Arranjo A_3

$$H_{\omega}^{12} = \begin{matrix} & P_1^2 Q_1^2 & P_1^2 Q_2^2 & P_2^2 Q_1^2 & P_2^2 Q_2^2 \\ \begin{matrix} P_1^1 Q_1^1 \\ P_1^1 Q_2^1 \\ P_2^1 Q_1^1 \\ P_2^1 Q_2^1 \end{matrix} & \begin{bmatrix} g^{12}(1-g^{12}) & (g^{12})^2 & (1-g^{12})^2 & g^{12}(1-g^{12}) \\ (g^{12})^2 & g^{12}(1-g^{12}) & g^{12}(1-g^{12}) & (1-g^{12})^2 \\ (1-g^{12})^2 & g^{12}(1-g^{12}) & g^{12}(1-g^{12}) & (g^{12})^2 \\ g^{12}(1-g^{12}) & (1-g^{12})^2 & (g^{12})^2 & g^{12}(1-g^{12}) \end{bmatrix} \end{matrix}$$

$$D_{\omega}^{12} = \begin{matrix} & P_1^2 Q_1^2 & P_1^2 Q_2^2 & P_2^2 Q_1^2 & P_2^2 Q_2^2 \\ \begin{matrix} P_1^1 Q_1^1 \\ P_1^1 Q_2^1 \\ P_2^1 Q_1^1 \\ P_2^1 Q_2^1 \end{matrix} & \begin{bmatrix} 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \end{bmatrix} \end{matrix}$$

$\Rightarrow$ Arranjo A_4

$$H_{\omega}^{12} = \begin{matrix} & P_1^2 Q_1^2 & P_1^2 Q_2^2 & P_2^2 Q_1^2 & P_2^2 Q_2^2 \\ \begin{matrix} P_1^1 Q_1^1 \\ P_1^1 Q_2^1 \\ P_2^1 Q_1^1 \\ P_2^1 Q_2^1 \end{matrix} & \begin{bmatrix} (g^{12})^2 & g^{12}(1-g^{12}) & g^{12}(1-g^{12}) & (1-g^{12})^2 \\ g^{12}(1-g^{12}) & (g^{12})^2 & (1-g^{12})^2 & g^{12}(1-g^{12}) \\ g^{12}(1-g^{12}) & (1-g^{12})^2 & (g^{12})^2 & g^{12}(1-g^{12}) \\ (1-g^{12})^2 & g^{12}(1-g^{12}) & g^{12}(1-g^{12}) & (g^{12})^2 \end{bmatrix} \end{matrix}$$

$$D_{\omega}^{12} = \begin{matrix} & P_1^2 Q_1^2 & P_1^2 Q_2^2 & P_2^2 Q_1^2 & P_2^2 Q_2^2 \\ \begin{matrix} P_1^1 Q_1^1 \\ P_1^1 Q_2^1 \\ P_2^1 Q_1^1 \\ P_2^1 Q_2^1 \end{matrix} & \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix} \end{matrix}$$

Sendo A_{ω}^2 o ω -ésimo arranjo dos alelos no segundo marcador M^2 , dado que M^l é fixo, de acordo com o teorema de Bayes, a probabilidade *a posteriori* do arranjo A_{ω}^2 dado M pode ser expressa como:

$$P(A_{\omega}^2 | M) = \frac{P(A_{\omega}^2)P(M|A_{\omega}^2)}{\sum_{\omega=1}^4 P(A_{\omega}^2)P(M|A_{\omega}^2)} = \frac{P(M|A_{\omega}^2)}{\sum_{\omega=1}^4 P(M|A_{\omega}^2)} \quad (2)$$

sendo $P(A_{\omega}^2)$ a distribuição de probabilidades *a priori* do A_{ω}^2 , que é assumida como uniforme, e $P(M | A_{\omega}^2)$ a verossimilhança de M dado A_{ω}^2 . Assumindo que os dados dos marcadores são independentes entre os N descendentes, têm-se:

$$P(M|A_{\omega}^2) = \prod_{j=1}^N P(M_j|A_{\omega}^2) \quad (3)$$

$$P(M|A_{\omega}^2) = \prod_{j=1}^N \sum_{i_1=1}^{p_1} \sum_{i_2=1}^{p_2} P(M_{1j} | (PQ)_{i1}) P((PQ)_{i2} | (PQ)_{i1}) P(M_{2j} | (PQ)_{i2})$$

sendo p_k , o número de genótipos distinguíveis (fenótipos) na progênie do marcador M^k , que é 4, 3, 3, 3, 2, 2 e 2 para os tipos A , B_1 , B_2 , B_3 , C , D_1 e D_2 (Tabela 5), respectivamente; $P(M_{kj} | (PQ)_{ik})$ é a variável indicadora que descreve o i_k ésimo fenótipo PCG do marcador M^k para a progênie j , que é 1 se o fenótipo da progênie j no marcador M^k (M_{kj}) é consistente com $(PQ)_{ik}$, e 0 se não o for; e $P((PQ)_{i2} | (PQ)_{i1})$, a probabilidade condicional (transição) do i_2 ésimo fenótipo PCG do marcador M^2 , dado o i_1 ésimo PCG do marcador M^1 , como descrito para os diferentes arranjos.

A equação (3) pode ser escrita na forma matricial como mostrado a seguir:

$$P(M | A_{\omega}^2) = \prod_{j=1}^N m_{i_1j}^T I_{p_1}^T H_{\omega}^{12} I_{p_2}^T m_{i_2j} = \prod_{j=1}^N m_{i_1j}^T P_{\omega}^{12} m_{i_2j} \quad (4)$$

sendo m_{i_kj} o vetor com dimensão p_k da variável indicadora $P(M_{kj} | (PQ)_{i_k})$ para o marcador M^k ; I_{p_k} a matriz de incidência com dimensão $(4 \times p_k)$ que relaciona o genótipo PCG com o seu fenótipo, feito para todos os quatro arranjos alélicos possíveis, em função dos diferentes tipos de combinações de marcadores utilizados, conforme apresentado a seguir:

$$I_{pA}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

para o cruzamento A ;

$$I_{pB1}^T = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

para o cruzamento B_1 ;

$$I_{pB2}^T = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

para o cruzamento B_2 ;

$$I_{pB3}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

para o cruzamento B_3 ;

$$I_{pC}^T = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

para o cruzamento C ;

$$I_{pD1}^T = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

para o cruzamento D_1 ;

$$I_{pD2}^T = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

e para o cruzamento D_2 .

H_{ω}^{12} é a matriz (4 x 4) de probabilidade de transição do genótipo PCG do marcador M^l , para o M^2 , $P_{\omega}^{12} = I_{p_1}^T H_{\omega}^{12} I_{p_2}$ é a matriz ($p_1 \times p_2$) de probabilidade de transição do marcador M^l para o M^2 ao nível fenotípico, e T indica o operador de transposição do vetor ou matriz.

Para obter a estimativa de máxima verossimilhança (EMV) para $\mathcal{G}^{12}$, o algoritmo EM (esperança e maximização) foi usado, conforme proposto por Dempster, Laird e Rubin (1977) e Lander e Green (1987) e adaptado por Wu et al. (2002) a essa situação. Assim, na etapa de esperança (E) utilizou-se a matrix $H_{\omega}^{12\{\tau\}}$ baseada na estimativa atual de $\mathcal{G}_{\omega}^{12\{\tau\}}$ e calculou-se o número esperado de eventos de recombinação entre as marcas M^l e M^2 para a descendência j sob o arranjo A_{ω}^2 , de acordo com:

$$c_{i_1 i_2 j}^{\omega\{\tau+1\}} = \frac{m_{i_1 j}^T [I_{p_1}^T (D_{\omega}^{12} \circ H_{\omega}^{12\{\tau\}}) I_{p_2}] m_{i_2 j}}{m_{i_1 j}^T I_{p_1}^T H_{\omega}^{12\{\tau\}} I_{p_2} m_{i_2 j}} \quad (5)$$

Na etapa de maximização (M), foi calculado o $\mathcal{G}_{\omega}^{12\{\tau+1\}}$ sob o arranjo A_{ω}^2 , utilizando a equação:

$$\mathcal{G}_{\omega}^{12\{\tau+1\}} = \frac{1}{2N} \sum_{j=1}^N \sum_{i_1=1}^{p_1} \sum_{i_2=1}^{p_2} c_{i_1 i_2 j}^{\omega\{\tau+1\}} \quad (6)$$

Estes procedimentos iterativos foram repetidos até haver convergência. Após a obtenção das estimativas de verossimilhança da frequência de recombinação, a fase de ligação mais provável entre os marcadores M^l e M^2 foi inferida pelo cálculo da probabilidade *a posteriori* dos quatro diferentes arranjos A_{ω}^2 . O arranjo que possuía a maior probabilidade *a posteriori*, com frequência de recombinação $< 0,50$ representava a mais provável fase de ligação entre os dois marcadores, sendo, então, adotada nas análises.

3.6.5 Formação dos grupos de ligação e ordenação das marcas

Todas estas análises foram implementadas no pacote *R* (R DEVELOPMENT CORE TEAM, 2004). Os dados dos genitores foram agrupados com intuito de se construir o primeiro mapa genético integrado de maracujá-amarelo, utilizando marcas com segregação 1:1 e 3:1 geradas pela técnica de AFLP e com segregação 1:1, 3:1, 1:2:1 e 1:1:1:1 geradas pelos marcadores microssatélites.

Nessa primeira etapa, foi possível a determinação das fases de ligação e da frequência de recombinação entre todos os pares de locos, permitindo a atribuição dos locos aos grupos de ligação (GL). Na construção dos grupos de ligação, é utilizada a propriedade transitiva, ou seja, se o loco 1 estiver ligado ao loco 2 e este, por sua vez, estiver ligado ao loco 3, então os locos 1 e 3 pertencem ao mesmo grupo de ligação. Como critérios para determinar a ligação foram usados $LOD\ score > 5,0$ e frequência de recombinação inferior a 0,35.

Após a formação dos grupos de ligação, iniciou-se o processo de ordenação das marcas dentro de cada grupo com o uso do algoritmo *Rapid Chain Delineation* (RCD) (DOERGE, 1996). De posse das frequências de recombinação entre todos os pares de marcas, o algoritmo do RCD inicia a construção da cadeia (ordenação do grupo de ligação) com o par de marcadores que possui a menor estimativa de frequência de recombinação. Em seguida, verifica-se dentro do grupo de marcadores restantes, aquele que possui a menor frequência de recombinação em relação ao par que iniciou a cadeia, sendo adicionado à esquerda ou à direita, em função de sua menor distância em relação a um ou outro marcador e, assim, sucessivamente, até que todos os marcadores sejam acrescentados.

Em seguida, a ordem sugerida pelo RCD foi verificada com o uso do algoritmo *Ripple* (LANDER et al., 1987), com base na menor SARF (*Sum of Adjacent Recombination Fraction*) e na função de verossimilhança de cada sub-ordem (WU et al., 2002). Para isso, o número de marcas incluídas no comando *Ripple* foi de 6 para o RCD e de 3 para o uso da verossimilhança.

A programação no *R* permitiu a obtenção das 10 melhores ordens via comando *Ripple*. Por isso, marcas foram retiradas dos GL quando prejudicavam a sua ordenação, a qual foi considerada correta quando não havia diferença entre as 10 melhores ordens. Após esta etapa, foi utilizada a verossimilhança de cada sub-ordem, com 3 marcas, para confirmação da ordenação sugerida pelo RCD.

3.6.6 Estimativas multiponto

As análises de três pontos foram realizadas visando-se obter as melhores estimativas da frequência de recombinação, especialmente, nos casos em que marcadores parcialmente informativos estavam envolvidos, além de contribuir para a verificação da ordenação sugerida pelo RCD.

De acordo com Wu et al. (2002), considerando três marcadores em ordem M^1 , M^2 e M^3 , e assumindo que o arranjo dos alelos nos dois genitores foi fixado para o M^1 , existem $4 \times 4 = 16$ possíveis arranjos ($A_{\omega 2 \omega 3}^{23}$) para os marcadores M^2 e M^3 . Denominando $\mathcal{G}^{12}$ como sendo a frequência de recombinação entre M^1 e M^2 , $\mathcal{G}^{23}$ e $\mathcal{G}^{13}$ são definidos similarmente. Estas frequências de recombinação estão associadas com as probabilidades de um *crossing-over* ocorrer entre os marcadores M^1 e M^2 , e entre M^2 e M^3 . O evento de ocorrência ou não de *crossing-over* em cada intervalo é denotado por G^{11} e G^{00} , respectivamente, enquanto a ocorrência de um *crossing-over* apenas no primeiro intervalo ou no segundo é denotada por G^{10} e G^{01} , respectivamente. As probabilidades de ocorrência destes eventos é denotada por g_{00} , g_{01} , g_{10} e g_{11} , respectivamente, cuja soma é 1. Assim, $\mathcal{G}^{12} = g_{00} + g_{11}$, $\mathcal{G}^{23} = g_{01} + g_{11}$ e $\mathcal{G}^{13} = g_{01} + g_{10}$; $g_{11} = \frac{1}{2} (\mathcal{G}^{12} + \mathcal{G}^{23} - \mathcal{G}^{13})$, $g_{10} = \frac{1}{2} (\mathcal{G}^{12} + \mathcal{G}^{13} - \mathcal{G}^{23})$, $g_{01} = \frac{1}{2} (\mathcal{G}^{23} + \mathcal{G}^{13} - \mathcal{G}^{12})$ e $g_{00} = 1 - \frac{1}{2} (\mathcal{G}^{12} + \mathcal{G}^{13} + \mathcal{G}^{23})$.

Para a análise de três pontos existem 16 (16×4) matrizes de probabilidade de transição para os marcadores M^1 , M^2 e M^3 , chamada de $H_{\omega 2 \omega 3}^{123}$, cada uma correspondendo a um arranjo particular $A_{\omega 2 \omega 3}^{23}$ condicionada ao *status* fixo do marcador M^1 . Do mesmo modo, existem 16 (16×4) matrizes para o número de *crossing-overs* que ocorrem para G^{00} , G^{01} , G^{10} e G^{11} entre os três marcadores sob um determinado arranjo, sendo $G_{\omega 2 \omega 3}^{00}$, $G_{\omega 2 \omega 3}^{01}$, $G_{\omega 2 \omega 3}^{10}$ e $G_{\omega 2 \omega 3}^{11}$, respectivamente.

Para um dado arranjo $A_{\omega 2 \omega 3}^{23}$, a verossimilhança é:

$$P(M | A_{\omega 2 \omega 3}^{23}) = \prod_{j=1}^N [m_{i_1 j}^T \otimes m_{i_2 j}^T] [I_{p_1}^T \otimes I_{p_2}^T] H_{\omega 2 \omega 3}^{123} I_{p_3}^T m_{i_3 j} \quad (7)$$

onde $\otimes$ é o produto de Kronecker.

Da mesma forma que no teste de dois pontos, o algoritmo EM foi utilizado para obtenção das estimativas de máxima verossimilhança (EMV) para a frequência de recombinação entre os três marcadores. O que difere na análise de três pontos é que são obtidas EMV para as probabilidades de ocorrência de *crossing-over* considerando cada arranjo. Assim, as EMV para os *g*'s são transformadas em EMV para as frequências de recombinação.

Com a análise de três pontos, são obtidas duas estimativas diferentes de um mesmo intervalo para a frequência de recombinação (exceto para os intervalos situados nas extremidades dos grupos de ligação). Para a construção do mapa integrado, foram utilizadas estimativas ponderadas pelo recíproco da variância das duas estimativas como sugerido por Ridout et al. (1998) e Wu et al. (2002).

Para converter as frequências de recombinação em distâncias de mapa, expressas em centiMorgans, foi utilizada a função de mapeamento de Kosambi (KOSAMBI, 1944).

4 Resultados e discussão

4.1 Construção da biblioteca genômica enriquecida com microssatélites

O protocolo de construção da biblioteca genômica enriquecida com microssatélites envolve uma série de etapas. Neste caso, existe a necessidade de um cuidado especial durante os procedimentos, haja vista que o sucesso da técnica somente se verificou nas etapas finais de *screening* por PCR ou pelo seqüenciamento direto dos clones positivos, de modo que erros cometidos nas etapas anteriores podem prejudicar o andamento do trabalho. Cabe, então, mencionar os principais aspectos que devem ser observados durante sua construção.

O uso de controles, principalmente nas etapas de: i) digestão do DNA; ii) amplificação por PCR dos fragmentos digeridos; iii) seleção e purificação do DNA com tamanho adequado para o seqüenciamento; iv) hibridização e posterior amplificação por PCR, são de grande importância, pois permitem a detecção de erros que, porventura, possam ocorrer.

A seleção dos insertos é um passo importante e tem como objetivo facilitar os futuros trabalhos de seqüenciamento (fragmentos com tamanho entre 300 e 600 pb são ideais). Tais fragmentos foram eluídos do gel de agarose, onde foram, previamente, separados por tamanho. A eluição pode ser feita após a digestão do DNA ou após a ligação dos adaptadores aos fragmentos digeridos. O último procedimento oferece a vantagem de remover os adaptadores não ligados e outros reagentes da PCR, que podem dificultar os procedimentos seguintes na construção da biblioteca enriquecida. Nesta etapa, a principal cautela refere-se à eluição, principalmente porque pode ocorrer grandes perdas de DNA.

Como o princípio do método é a recuperação de fragmentos contendo repetições de microssatélites a partir de um *pool* de DNA, o processo de hibridização com as sondas repetitivas deve ser feito com bastante atenção, principalmente em relação à temperatura de hibridização, que deve ser de 60 °C.

Outra observação de interesse refere-se ao número dos ciclos de amplificação. É preciso tomar cuidado para que não haja amplificação excessiva dos fragmentos selecionados após a hibridização, de modo a aumentar a redundância e por consequência a eficiência da biblioteca genômica (vinte ciclos de amplificação produzem resultados satisfatórios). A seguir são

apresentados os resultados obtidos com a construção da biblioteca genômica de maracujá-amarelo.

4.1.1 Correção das seqüências e redundância da biblioteca

Foram construídas bibliotecas genômicas enriquecidas com repetições GT e CT para os genitores IAPAR-123 e IAPAR-06. Todas as seqüências obtidas foram editadas no programa computacional *Staden Package*. Combinações de diferentes corridas de uma mesma seqüência foram analisadas nesse programa com objetivo de verificar sua qualidade e, eventualmente, corrigir possíveis erros de seqüenciamento. Tais erros são devidos à: i) inibição da reação por contaminantes, ii) baixa quantidade de DNA e/ou *primer*, iii) presença de resíduos de *primers* ou de dNTP após a purificação da reação, iv) contaminação das seqüências por diferentes vetores, com ou sem insertos, v) presença de estrutura secundária, principalmente em regiões ricas em GC.

Um outro problema que, usualmente, acontece no seqüenciamento é o término gradual da seqüência em função da presença de seqüências repetitivas, como os microssatélites, obtendo-se apenas um dos lados da seqüência flanqueadora. Tal fato pode ser resolvido promovendo o seqüenciamento no outro sentido da fita de DNA, possibilitando o desenho dos *primers* do microssatélite. Contudo, o tamanho da seqüência não é conhecido previamente, tendo em vista a falta de informação do número de repetições do microssatélite.

Para cada clone obtido na biblioteca genômica enriquecida, dispunha-se de no mínimo duas corridas eletroforéticas do seqüenciador. Assim, foi possível promover o alinhamento e as devidas correções em cada seqüência. A Figura 4 representa duas corridas de uma seqüência em que foram observados erros em algumas bases (em verde). Foram obtidas 947 seqüências, sendo 501 do genitor IAPAR-06 e 446 do IAPAR-123. Após a edição e correção, todas as seqüências foram comparadas para a identificação de seqüências redundantes. Das 947 seqüências, 134 (14,2%) foram consideradas redundantes por apresentarem uma similaridade de até 90% em número de bases. Foram observadas 35 seqüências diferentes representadas duas vezes. Outras 10, 5, 4, 5, 1 e 1 seqüências diferentes foram representadas por 3, 4, 5, 6, 11 e 14 vezes, respectivamente (Figura 5).

Figura 4 - Exemplo acerca de erros de seqüenciamento em duas corridas eletroforéticas de uma mesma seqüência

Figura 5 - Distribuição e número de seqüências redundantes encontrados nas bibliotecas enriquecidas dos acessos IAPAR-123 (123) e IAPAR-06 (06) de maracujá-amarelo

De acordo com essas análises de redundância, foram encontradas 134 seqüências redundantes, representando 61 diferentes seqüências. A Figura 6 ilustra um exemplo de dois clones redundantes, o C02P16 e o D01P16.

A redundância pode ser atribuída à presença de clones duplicados pelo próprio procedimento da PCR durante as etapas de enriquecimento, à presença de diferentes alelos, uma vez que os genitores utilizados na construção da biblioteca são não endogâmicos, além da presença de locos duplicados no genoma. Sequências com mais de 90% de similaridade foram consideradas redundantes, sendo que muitas delas apresentaram variações na composição de nucleotídeos. Paniego et al. (2002), adotaram critérios menos restritivos para analisar uma biblioteca enriquecida com microssatélites de girassol, utilizando similaridade $\geq 85\%$ para considerar as sequências como redundantes.

Dentro das 134 sequências redundantes, foram observadas sequências dos dois genitores utilizados para a construção da biblioteca. Descartando a possibilidade de eventuais erros de sequenciamento, tendo em vista a prévia correção das sequências, estas sequências redundantes possivelmente correspondem a diferentes alelos de ambos os genitores ou mesmo locos duplicados. Entretanto, para a confirmação desta hipótese, seria necessário o desenho de *primers* para amplificar estas sequências por PCR, posterior clonagem e sequenciamento.

Também foram encontradas 13 sequências similares contendo microssatélites, representando 6 sequências únicas. Interessantemente, é que uma dessas sequências únicas corresponde a uma sequência redundante de cada genitor (Figura 7).

Em maracujá-amarelo, a redundância dentro da classe de microssatélites foi de 5,2% (7 sequências em 134 microssatélites encontrados), que pode ser considerada baixa em função dos relatos em outras espécies.

Todo esse trabalho foi feito visando a obtenção de sequências com boa qualidade, que é de fundamental importância para a análise dos microssatélites, uma vez que os *primers* são desenhados nas regiões flanqueadoras dos microssatélites para promover uma amplificação específica.

Observa-se a presença de 21 nucleotídeos diferentes nas regiões flanqueadoras e 2 dentro do motivo do microssatélite, respectivamente, nas sequências D03P16 do genitor IAPAR-06 e E01P09 do genitor IAPAR-123, havendo uma dissimilaridade entre as duas sequências de 6,0% (Figura 7). Um aspecto importante a ser observado é a não conservação completa das regiões flanqueadoras entre os genitores neste microssatélite. Nesta situação, apenas a sequência D03P16 do genitor IAPAR-06 foi utilizada para o desenho de *primers*, levando-se em consideração as regiões similares em ambas as sequências. Assim, diferenças entre alelos de ambos os genitores

podem ser devidas às variações das repetições do microssatélite, além de inserções ou deleções presentes em ambos os genitores.

Figura 6 - Análise de redundância das seqüências obtidas a partir de uma biblioteca genômica enriquecida com microssatélites de maracujá-amarelo

```

cap_D03P16 (06) 121 agtgtaggcgcgaagaaggaagaatagaaaaaataatttttgtt-gaaacggttatctttg
cap_E01P09 (123) 1 ..... *****t*****c*****

cap_D03P16 (06) 180 aagcgcagattctgttgcagatatagattcggatgcggaatcctctcccggtatccac
cap_E01P09 (123) 24 *****t*****

cap_D03P16 (06) 240 atgaacacttcgcatggctagattgctggagttggtctctctctctctgggtttctcacg
cap_E01P09 (123) 84 *c*****c**t*g**c*****gg-----**g*****

cap_D03P16 (06) 300 gcaaaaacctgatgaaacacagagaaaaaataatttt--cttttgttctgcttttctgtt
cap_E01P09 (123) 142 *****t*****t*****

cap_D03P16 (06) 358 tttaacgcacgaatgatcgtgtcagagagatatgcctatttctctgttttagtcgcttgact
cap_E01P09 (123) 201 *cc*****

cap_D03P16 (06) 418 aaacggtatgtaataattttgcagattacatctgacatcacatcctaataat
cap_E01P09 (123) 261 *****

cap_D03P16 (06) 478 ctaattagattagatattatagttacggtttga-ttcggaatcaaacgggtccgatcca
cap_E01P09 (123) 321 *****-*****-**a*****

cap_D03P16 (06) 537 gtcgg
cap_E01P09 (123) 379 ***-aagatagtgacgcgtaagcaaga

```

Figura 7 - Similaridade entre duas seqüências contendo microssatélites, presentes em ambos os genitores

Deleções/inserções de bases únicas ou de longos fragmentos de DNA nas regiões flanqueadoras foram reportadas como fonte de variação dos microssatélites (GIANFRANCESCHI et al., 1998; ASHWORTH et al., 2004). Além do mais, o polimorfismo nas regiões flanqueadoras podem afetar os sítios de anelamento dos *primers*, promovendo o aparecimento de alelos nulos (RALLO, DORADO e MARTÍN, 2000).

Para as outras seqüências contendo microssatélites, três são redundantes para o genitor IAPAR-06, sendo a seqüência A12P01 e G09P04, com 8,4%, H01P16 e B03P16 com 0,6%, D07P16 e E06P16 com 0,9% de dissimilaridade de bases. Já para o IAPAR-123, as seqüências D05P07, B01P10 e B03P10 apresentam uma dissimilaridade de 10% enquanto as seqüências F04P02 e F01P09 apresentam uma dissimilaridade de 3,9%.

A redundância das seqüências reduz a eficiência de construção da biblioteca, exigindo a construção de bibliotecas com maior número de clones para se garantir maior número de locos de microssatélites. He et al. (2003) construíram uma biblioteca genômica de amendoim e encontraram uma redundância de 67% das seqüências de microssatélites. Já em *Lolium perenne* L., Jones et al. (2001) encontraram uma redundância de 16,4%.

4.1.2 Enriquecimento da biblioteca genômica com microssatélites

A biblioteca foi constituída por um tamanho médio de inserto de 450 bp, o que facilitou o trabalho de seqüenciamento. Por outro lado, não foi possível o seqüenciamento completo dos insertos para a maioria dos clones, em função da dificuldade de seqüenciar regiões repetitivas havendo um término gradual das mesmas.

Dos 947 clones seqüenciados, foram encontrados 127 contendo microssatélites não redundantes, o que demonstra uma eficiência de enriquecimento de 13,4%. Não foi possível desenhar *primers* para 20 seqüências, em função da falta de uma das regiões flanqueadoras dos microssatélites. Por isso, a eficiência de construção da biblioteca reduziu para 11,3%.

Nesse sentido, foram encontrados 66 microssatélites perfeitos (quando a seqüência repetida não é interrompida por qualquer base que não pertença ao motivo), 55 imperfeitos (quando existem bases entre os motivos que não correspondem ao mesmo) e 6 compostos (quando a seqüência contém duas seqüências repetidas distintas adjacentes). Dentro da classe de

microsatélites perfeitos, existem quatro seqüências de mono, 52 de di, 3 de tri, 1 de tetra, 3 de penta e 3 de hexanucleotídeos. A classe de microsatélites imperfeitos possui 2 seqüências de mono, 41 de di, 2 de tri, 8 de tetra, 1 de hexa e 1 de heptanucleotídeos. Já na classe de microsatélites compostos, todas as 6 seqüências são de dinucleotídeos (Figura 8).

Em uma destas seqüências, foram encontrados dois microsatélites imperfeitos. Neste caso, foi desenhado um par de *primers* para amplificação de ambos. O desenho de *primers* para amplificação desses dois microsatélites, separadamente, não apresenta interesse prático em estudos de mapeamento, uma vez que estariam completamente ligados.

Resultados semelhantes foram encontrados por Brondani et al. (1998) na construção de bibliotecas enriquecidas para as repetições AG e AC em eucalipto, relatando que a maioria dos microsatélites encontrados, foi de microsatélites perfeitos, 81 e 48% para as bibliotecas AG e AC, respectivamente.

Figura 8 - Classes de microsatélites encontrados em maracujá-amarelo em função dos motivos

Esses resultados são interessantes, haja vista que microsatélites perfeitos são tidos, em média, como mais polimórficos do que aqueles com interrupções dentro dos motivos (WEBER, 1990), fazendo com que se tenha preferência no isolamento e desenho de *primers* para este tipo de microsatélite (ELLEGREN, PRIMMER e SHELDON, 1995).

Segundo Harr, Zanger e Schlötterer (2000), o processo de *slippage* do DNA além de promover a geração de novos alelos para os microssatélites, também é importante para promover a remoção das interrupções dos microssatélites imperfeitos. A “purificação” das interrupções é feita com o seguinte mecanismo: durante a replicação do DNA, ocorre a formação de uma alça na fita molde, sendo que a interrupção está contida dentro da mesma. Em seguida, ocorre a síntese da nova fita de DNA sem a interrupção, de tal sorte que esta seqüência tornar-se-á um microssatélite perfeito. Ainda segundo estes autores, as interrupções dos microssatélites são consideradas como um estado transitório de evolução e não de degeneração dos locos de microssatélites.

4.1.2.1 Motivos e número de repetições dos microssatélites

Do total de microssatélites, 3,9% são de mononucleotídeos, 78,7% são repetições de dinucleotídeos, 3,9% são de trinucleotídeos, 7,1% de tetranucleotídeos, 2,4% são de pentanucleotídeos, 3,2% são de hexanucleotídeos e 0,8% de heptanucleotídeos.

Os motivos dos microssatélites encontrados, bem como o número de seqüências em cada um deles, estão apresentados na Tabela 6. Como esperado, a classe de dinucleotídeos foi a mais freqüente em virtude do uso de sondas CT e GT. Ainda dentro dessa classe, microssatélites com motivos AC/GT e CA/TG foram mais freqüentes do que aqueles com motivos AG/CT e GA/TC, correspondendo a 52,8 e 25,2% do total, respectivamente.

O tamanho dos microssatélites foi aqui definido em função do número de repetições não interrompidas. Analisando somente os motivos de dinucleotídeos GT e CT, bem como os microssatélites compostos por estes motivos, verifica-se uma maior freqüência de motivos do tipo GT, porém com menor número de repetições em relação aos motivos CT. Motivos do tipo GT e CT apresentaram, em média, $10,3 \pm 4,3$ (erro padrão) e $23,2 \pm 8,7$ repetições, respectivamente (Figura 9).

Os microssatélites compostos com motivos GT, apresentaram maior número de repetições em relação aos perfeitos e imperfeitos. Contudo, este maior número de repetições foi devido ao outro motivo que acompanha este tipo de microssatélite.

Tabela 6 - Distribuição dos tipos de repetições e motivos dos microssatélites, bem como o número de seqüências em cada classe (Nº)

Tipo de repetição	Motivo	Nº
Mononucleotídeos	T	3
	G	2
Dinucleotídeos	AC/GT	23
	CA/TG	38
	AG/CT	15
	GA/TC	17
	CT/AG+AG/CT	1
Trinucleotídeos	AGA	1
	TGT	1
	AGC	1
	CTT	1
	CTT	1
Tetranucleotídeos	GAAA	1
	TTAA	5
	AAAC	1
	ACCG	1
	AAAG	1
Pentanucleotídeos	ATCGT	1
	ATCTA	1
	CATAT	1
Hexanucleotídeos	TTATGG	1
	GTTGTG	1
	TTTTTG	1
	ATCACA	1
Heptanucleotídeos	CTTTAGC	1
Compostos	(TA) _n (TG) _n	2
	(AT) _n (TG) _n	1
	(AT) _n (GT) _n	1
	(AT) _n (AC) _n	1
	(CT) _n (GT) _n	1

Resultados contraditórios foram observados em abacate por Ashworth et al. (2004), pois motivos do tipo CT foram mais frequentes e com maior número de repetições, quando comparados com motivos do tipo GT.

Figura 9 - Distribuição do comprimento em função do número e dos motivos dos microssatélites de maracujá-amarelo

Observou-se nesse conjunto de seqüências que quanto maior o tamanho do motivo, menor é o seu número de repetições. Os hexa, penta e tetranucleotídeos apresentaram, em média, $3,8 \pm 1,0$; $5,0 \pm 2,0$ e $5,9 \pm 2,8$ repetições, respectivamente. Os trinucleotídeos apresentaram, em média, $6,4 \pm 2,8$ repetições. No caso dos dinucleotídeos, o número médio de repetições foi de $14,8 \pm 8,9$, mostrando uma alta variância associada, uma vez que o número de repetições variou de 6 até 42. Já para os mononucleotídeos e microssatélites compostos, a média foi de $23,6 \pm 7,0$ e $24,0 \pm 11,3$ repetições, também com amplitude bastante elevada, de 16 a 35 e de 13 a 44 repetições, respectivamente (Figura 10).

Figura 10 - Distribuição do comprimento das seqüências repetitivas em função do número e das classes de microssatélites encontrados em maracujá-amarelo

A frequência das classes de microssatélites é bastante variável em plantas (WANG et al., 1994). Numa revisão feita por Powell, Machray e Provan (1996), estes autores relataram que os motivos AG são mais abundantes no genoma de plantas do que os motivos AC, diferentemente do que ocorre no genoma de mamíferos (WEISSENBAACH et al., 1992). Entretanto, ainda existe uma discordância sobre o assunto, uma vez que em algumas espécies os motivos AG são mais abundantes do que os AC (BRONDANI et al., 1998; VARHNEY et al., 2000) e, em outras, os motivos AC são mais frequentes (ECHT e MAY-MARQUARDT, 1997; BRYAN et al., 1997; BUTCHER et al., 2000). Os dados aqui apresentados indicam haver uma maior frequência de motivos GT no genoma de *P. edulis* f. *flavicarpa*. Contudo, para estabelecer melhor esta relação, seria necessário o *screening* de um conjunto maior de seqüências.

A abundância de motivos GT (mais do que o dobro do que CT) e a caracterização dos microssatélites perfeitos em *P. edulis* f. *flavicarpa* foi bastante similar aos relatos de eucalipto

(BRONDANI et al., 1998) e arroz (CHEN et al., 1997), duas espécies muito distantes filogeneticamente de *Passiflora*.

Os fatores responsáveis pela diferença em abundância dos motivos dos microssatélites entre os vários gêneros em plantas ainda não são muito bem esclarecidos (ECHT e MAY-MARQUARDT, 1997). Contudo, a escolha das enzimas de restrição pode influenciar no número e no tipo de repetição encontrados na construção das bibliotecas genômicas (HAMILTON e FLEISCHER, 1999). A construção de bibliotecas genômicas enriquecidas com microssatélites, também, tornou-se uma prática rotineira nos últimos anos em plantas (ZANE, BARGELLONI e PATARNELLO, 2002), sendo que a maioria delas utiliza sondas de dinucleotídeos, principalmente AG e AC, embora, alguns autores afirmem que os dinucleotídeos AT sejam os mais abundantes nos genomas vegetais (PRIMMER et al., 1997). Contudo, formações de estruturas secundárias podem ocorrer devido ao pareamento interno de bases nas sondas (AT)_n, dificultando os procedimentos de hibridização com as seqüências alvo, fazendo com que as informações sobre a abundância deste tipo de seqüência não esteja disponível para muitas espécies.

Um aspecto interessante observado nesse trabalho, foi a detecção de motivos do tipo AT, preferencialmente associados com motivos GT/AC, formando os microssatélites compostos. Resultados semelhantes foram obtidos em *Acacia mangium*, para a qual foi construída uma biblioteca genômica enriquecida com repetições do tipo AT dentre outras repetições. Contudo, repetições AT somente foram isoladas quando associadas a outras, ou seja, formando os microssatélites compostos (BUTCHER et al., 2000). Resultados semelhantes também foram encontrados em girassol (PANIEGO et al., 2002).

É comum, serem encontrados na literatura, relatos sobre a presença de outros tipos de microssatélites, senão os dinucleotídeos para os quais as bibliotecas genômicas enriquecidas foram desenvolvidas (BUTCHER et al., 2000; ASHKENAZI et al., 2001). Este fato pode ser explicado pela baixa estringência adotada nos procedimentos de hibridização que faz com que as sondas se liguem de uma forma não completamente pareada com as seqüências alvo.

Um fato interessante foi o aparecimento de 7,1% de tetranucleotídeos em *P. edulis* f. *flavicarpa*, pois a biblioteca foi construída para detecção de dinucleotídeos. Em trigo, este tipo de motivo é raro, como relatado por Varshney et al. (2000) e Ma, Roder e Sorrels (1996), mesmo

construindo-se bibliotecas enriquecidas especificamente para este tipo de sequência. Também, foi aqui constatada a porcentagem de 3,9% de trinucleotídeos.

4.1.2.2 Eficiência de enriquecimento da biblioteca e desenho de *primers*

O enriquecimento de bibliotecas genômicas é uma metodologia simples para o isolamento de microssatélites, sendo amplamente aplicado para uma grande variedade de espécies, porém com uma eficiência variável de recuperação de sequências contendo microssatélites (BUTCHER et al., 2000; HE et al., 2003). Alguns relatos demonstram uma baixa eficiência, de 1,7% em batata (ASHKENAZI et al., 2001) e 2,1% em mamão (SANTOS et al., 2003); outros relatos apontam para alta eficiência de enriquecimento, de 46% em *Lolium perenne* L. (JONES et al., 2001) e de 28% em melão (RISTSCHEL et al., 2004). Já bibliotecas genômicas convencionais (sem enriquecimento) possuem uma baixa eficiência de recuperação de sequências contendo microssatélites. Em batata, por exemplo, Ashkenazi et al. (2001) relatam uma eficiência de 0,37% na recuperação de microssatélites utilizando esta abordagem.

Neste estudo, 127 sequências apresentaram microssatélites, sendo disponíveis para o desenho de *primers*. Contudo, foram desenhados *primers* para 107 sequências (84,25%) que incluíam 55 microssatélites perfeitos, 46 imperfeitos e 6 compostos. Para 15,8% restantes não foi possível desenhar *primers*, principalmente, devido ao fato dos microssatélites estarem próximos ao sítio de clonagem.

Em sorgo, Taramino et al. (1997) utilizaram procedimentos de enriquecimento semelhantes aos aqui apresentados, e encontraram 70% das sequências contendo microssatélites, com as repetições próximas ao sítio de clonagem. Neste trabalho, foram selecionados fragmentos de tamanho relativamente pequeno (300 a 600 pb), sendo que a escolha das enzimas de restrição utilizadas na construção da biblioteca pode influenciar no número e tipo de repetição a ser encontrada, bem como no processo de sequenciamento (HAMILTON e FLEISCHER, 1999).

Trabalhos de literatura revelam haver uma baixa relação de microssatélites detectados por *primers* desenhados, devido aos motivos já citados, bem como pela baixa qualidade das sequências. Em trigo Varshney et al. (2000), desenharam *primers* para apenas 39% dos microssatélites detectados. Em amendoim, He et al. (2003) desenharam *primers* para 22,67% das sequências. A alta relação microssatélites detectados/desenhados neste trabalho (84,25%) foi

obtida em função da edição das seqüências no programa computacional *Staden Package*. Esta abordagem promove um melhor aproveitamento das seqüências, uma vez que diferentes corridas de uma mesma seqüência permitem uma complementariedade na qualidade das seqüências em diferentes posições. Esta correção é trabalhosa, pois as seqüências devem ser corrigidas base a base. Contudo, os resultados são interessantes quando se deseja obter seqüências de boa qualidade para desenho de *primers*. Também, se disponibiliza para os bancos de dados, seqüências altamente confiáveis.

A partir deste trabalho, é possível afirmar que a presença dos microssatélites em maracujá-amarelo é tão abundante quanto em outras espécies frutíferas, como mamão (SANTOS et al., 2003), melão (RISTCHEL et al., 2004) e abacate (ASHWORTH et al., 2004). Nossos resultados sugerem que o sistema aqui adotado pode ser promissor no isolamento e caracterização de outros microssatélites em maracujá-amarelo.

4.1.3 Otimização das condições de PCR para detecção e análise do polimorfismo

As análises genéticas com marcadores microssatélites envolvem a amplificação de DNA por PCR utilizando *primers* complementares às regiões flanqueadoras desses locos. Contudo, a PCR é bastante sensível a uma série de fatores, principalmente, em relação aos reagentes utilizados na reação e à temperatura de anelamento dos *primers* para haver amplificação. Assim, variações na concentração de magnésio, tampão e dNTP, além do uso diferentes programas de amplificação foram analisados.

Foram desenhados 107 pares de *primers* para os marcadores microssatélites desenvolvidos para *P. edulis* f. *flavicarpa*. As seqüências dos *primers*, o tipo de motivo encontrado, bem como sua classificação, o tamanho do alelo seqüenciado, os *mix* e as temperaturas de anelamento (TA), além da qualidade da amplificação são apresentadas na Tabela 7. Para estes locos de microssatélites, o tamanho dos alelos esperados variou de 126 até 392 pb.

O tamanho do alelo esperado para o loco *PE57* (o código adotado para os *primers*, refere-se também ao nome do loco em *P. edulis* f. *flavicarpa*) não era conhecido *a priori*, tendo em vista que o microssatélite estava no meio da seqüência. Mesmo seqüenciando ambos os lados do inserto, não foi possível completar o *contig*. Assim, foram desenhados *primers* na seqüência *forward* e outro na *reverse*. No processo de otimização desse loco, verificou-se que havia a

amplificação de uma banda mais forte por volta de 720 pb, sendo considerado o tamanho esperado para o alelo.

Para cada marcador, foi escolhido o *mix* e o programa de amplificação que rendeu maior quantidade de produto com menor quantidade de bandas inespecíficas, características desejáveis para a facilitar a leitura e a genotipagem dos locos nos géis de poliacrilamida.

Com relação aos *mix* utilizados, 8,4% das reações⁵ foram otimizadas com o *mix* 2, 21,5% com o *mix* 4, 30,8% com *mix* 5, 18,7% com o *mix* 7 e 20,6% das reações não foram otimizadas com nenhum dos *mix* testados, em virtude de não apresentarem nenhum produto amplificado ou por formarem um *smear* (rastro) no gel. Estas reações que não produziram amplificação foram também avaliados sob temperaturas de anelamento menores, porém sem grande sucesso.

Observou-se que a maior parte das reações dos *primers* (52,3%) foi otimizada com os *mix* 4 e 5, na qual o que variou foi a concentração de tampão, 2X e 1X no *mix* 4 e 5, respectivamente, bem como na quantidade de MgCl₂, que foi de 1,5 e 2,5 mM no *mix* 4 e 5, respectivamente. Diferenças entre os *mix* 2 e 7, foram devidas à concentração de 1,5 e 2,5 mM de MgCl₂, e de 200 e 350 µM de dNTP, respectivamente (Tabela 2).

Segundo Rahman, Jaquish e Khasa (2000), a concentração ótima de magnésio para amplificação da maioria dos locos de algumas espécies de coníferas (*Picea* spp e *Larix* spp) foi de 1,0 mM, sendo que concentrações de 0,5 mM não produziram bandas visíveis, enquanto sob concentrações de 3,5 mM não houve produção de bandas visíveis.

Para amplificação dos locos de microssatélites de *P. edulis* f. *flavicarpa*, não houve grandes restrições à amplificação em altas concentrações de magnésio – 4,0 mM (dados não mostrados), sendo que as concentrações de 1,5 a 2,5 mM são ideais.

Com relação aos programas de amplificação para definição da temperatura de anelamento de cada *primer*, verificou-se que 17,6% das reações foram otimizadas com o programa TD56, 71,8% com o programa TD60, 1,2% com temperatura de anelamento específica de 52 e 60 °C e 8,2% delas com temperatura de 60 °C (Tabela 7).

Como as condições de PCR precisam ser otimizadas para cada loco, em virtude da grande variação nos *primers* e nos produtos amplificados, Senior et al. (1996) propuseram o uso da estratégia conhecida como *touch-down*, na qual a temperatura de anelamento diminui com o

⁵ Neste contexto, entenda-se otimização das reações como otimização para cada *primer* específico que correspondia a um loco de microssatélite

avanço nos ciclos de amplificação da PCR. Quando muitas reações precisam ser otimizadas, esta metodologia oferece bons resultados, em curto espaço de tempo. Excluindo-se os *primers* que não produziram amplificação, aproximadamente 90% das reações foram otimizadas utilizando esta metodologia em maracujá-amarelo.

Resultados semelhantes foram observados por Rahman, Jaquish e Khasa (2000) em *Picea* spp e *Larix* spp, pois o protocolo de *touch-down* produziu bons resultados, com amplificação de bandas intensas em géis de poliacrilamida. Esta metodologia amplifica produtos específicos sob temperaturas maiores, seguindo-se de uma ligação forte dos *primers* no sítio alvo com a progressiva diminuição da temperatura. Temperaturas de anelamento mais estridentes, especialmente durante os primeiros ciclos, são recomendadas para aumentar a especificidade da reação.

Das 79,4% das reações que apresentaram boa amplificação (bandas bem visíveis e sem *stutter*), 25,2% apresentaram uma amplificação específica dos alelos esperados, 39,3% apresentaram amplificação dos alelos desenhados, porém com amplificação de outros alelos de outros locos, e 14,9% amplificaram outros alelos de tamanhos diferentes comparativamente aos desenhados (Tabela 7).

Os *primers* que amplificaram alelos de outros locos foram testados em temperaturas de anelamento maiores do que aquelas para o qual as respectivas reações foram otimizadas, com intuito de eliminar estas amplificações indesejáveis. Contudo, este procedimento não produziu melhorias, pois ocorreu uma inibição total da amplificação de todos os locos.

PE48 e *PE30* são exemplos de locos para os quais não se obteve nenhum produto de amplificação, enquanto os locos *PE32* e *PE85* formaram um rastro no gel (Figura 11 e 12). Os locos *PE38* (Figura 11) e *PE54* e *PE15* (Figura 12), apresentaram boa amplificação dos alelos esperados, apesar da grande influência das bandas *stutter*, que dificultam a leitura dos géis. Os *primers* relativos aos locos *PE69*, *PE57*, *PE52*, *PE67* (Figura 11) e *PE25*, *PE60*, *PE68*, *PE16*, *PE77* e *PE22* (Figura 12), amplificaram os alelos esperados, bem como alelos de outros locos. Os *primers* correspondentes ao loco *PE51*, apresentaram amplificação muito fraca do alelo esperado (206 pb) (Figura 11). Contudo, outros fragmentos (locos) foram amplificados com formação de bandas fortes. Informações relativas a todos os *primers* são apresentadas na Tabela 7.

Figura 11 - Amplificação do DNA de um grupo de plantas da população segregante de maracujá-amarelo. Os números 1 e 2 correspondem aos genitores IAPAR-06 e IAPAR-123, respectivamente, 3 a 8, plantas F_1 do cruzamento. Esta mesma sequência de plantas foi utilizada em reações com todos os *primers*. As setas mostram os alelos esperados segundo a leitura das sequências

Desconsiderando os *primers* que não produziram produtos de amplificação, 73% deles produziram bandas *stutter*. O aparecimento de tais bandas ocorre, principalmente, em repetições de dinucleotídeos (CREGAN et al., 1994). Este tipo de motivo constitui 78,7% da biblioteca genômica de *P. edulis* f. *flavicarpa* aqui construída. Segundo Ashworth et al. (2004), locos que revelam muitas bandas residem em regiões duplicadas no genoma ou resultam de desenho inadequado dos *primers*. O aparecimento de bandas *stutter* se deve à ocorrência do processo de *slippage* durante a reação de PCR (HAUGE e LITT, 1993), sendo que estas bandas tendem a diminuir com o aumento do tamanho do motivo (EDWARDS et al., 1991).

Figura 12 - Amplificação do DNA de um grupo de plantas da população segregante de maracujá-amarelo. Os números 1 e 2 correspondem aos genitores IAPAR-06 e IAPAR-123, respectivamente, 3 a 8, plantas F_1 do cruzamento. Esta mesma sequência de plantas foi utilizada em reações com todos os *primers*. As setas mostram os alelos esperados segundo a leitura das seqüências

Em girassol, locos de microssatélites com motivos $(ATT)_n$, $(TGG)_n$ e $(GA)_n$ produzem poucas bandas *stutter*, enquanto locos contendo os motivos $(GT)_n$, $(CATA)_n$ e $(ATC)_n$, apresentam um perfil de baixa qualidade, com muitas bandas *stutter*, porém permitindo a leitura dos géis (PANIEGO et al., 2002).

Em abacate, 49% dos *primers* que amplificam locos de di e trinucleotídeos mostraram-se promissores para estudos de diversidade genética, com padrões de bandas interpretáveis (ASHWORTH et al., 2004). Ainda segundo estes autores, modificações nas condições da PCR e novo desenho de *primers* para os locos “problemáticos” raramente levam à uma melhoria

significativa no padrão de bandas, sugerindo haver influência do genoma (*background*) da espécie no padrão de bandas.

Numa biblioteca genômica enriquecida com microssatélites de girassol, 22% dos *primers* falharam na amplificação dos locos ou mostraram um padrão de bandas complexo, com pouca reprodutibilidade, não sendo utilizados na caracterização dos locos de um conjunto de 16 linhagens (PANIEGO et al., 2002). Resultados semelhantes foram obtidos por Yu et al. (2000), a partir de uma biblioteca enriquecida para di e tetranucleotídeos, também de girassol, sendo que apenas 75,5% dos *primers* produziram um bom padrão de bandas.

Considerando apenas os *primers* que produziram amplificação específica, a taxa de amplificação dos microssatélites de *P. edulis* f. *flavicarpa* isolados neste trabalho foi de 25,2%. Se incluirmos os *primers* que produziram amplificações de outros locos, esta taxa chega a 64,5%. Este potencial foi maior, por exemplo, do que os 30% de *primers* funcionais relatados por Röder et al. (1995; 1998) em trigo. A inespecificidade de amplificação dos locos de microssatélites tem sido verificada em outras espécies, principalmente devido a processos de duplicação de regiões do genoma, além da evolução particular das famílias de seqüências repetitivas (AKAGI et al., 1998). Entretanto, para saber se estes outros locos amplificados contêm as repetições dos microssatélites para o qual os *primers* foram desenhados, é necessário o isolamento, a clonagem e o seqüenciamento de cada alelo.

Múltiplos produtos são resultado de vários sítios de anelamento complementares ao *primer* ao longo do genoma da espécie sob estudo. A amplificação de múltiplos locos é comum em espécies com origem aloploplóide, pois muito provavelmente ocorreram eventos de fusão e duplicação cromossômica durante a evolução (BUTELER et al., 1999). Contudo, se nesta amplificação inespecífica houver a possibilidade de identificação dos diferentes locos, como acontece quando se trabalha com a construção de mapas de ligação, abre-se a possibilidade de fazer uma genotipagem multiloco. Tal abordagem foi adotada durante o desenvolvimento desta tese e será discutida adiante, no item 4.2.4.

Tabela 7 - Código do loco e sequência de bases dos 107 pares de *primers* de microssatélites desenhados para *P. edulis* f. *flavicarpa*, tipo de motivo e sua classificação segundo Weber (1990), tamanho do alelo sequenciado, *mix* e temperatura de anelamento (TA) ideais, além da qualidade da amplificação encontrada e dos locos com presença de bandas *stutter*

(continua)

Loco	Primer forward	Primer reverse	Motivo	Classes	Alelo (pb)	Mix	TA	Qualidade da amplificação	Bandas <i>stutter</i>
PE01	caggatagcagcagcaatga	agccaaatgtcaaactgaac	(GT) ₇	Perfeito	171	7	TD60	AOL	S
PE02	gggacgacaatcaagtgagg	cccaaactatgcaacaccaa	(TG) ₇	Perfeito	215	4	TD60	LBA	S
PE03	gcagcgagggaagaaaaa	tgagacatcgctgcgtgaa	(GA) ₁₀	Interrompido	156	2	TD60	LBA	S
PE04	atgcttttgaaatccgttt	tgctcatgcaaagtcactgg	(TG) ₉	Imperfeito	235	5	TD60	LBA	S
PE06	agcggggaggagagtagc	gcctgatgtcaaaaacacag	(CA) ₇	Interrompido	187	4	TD60	LBA+OL	N
PE07	tgctcattgatggtgcttg	tcgtctcttctcctctca	(GA) ₂₃	Interrompido	138	4	TD60	LBA	S
PE08	ccggatacccacgcatta	tctaagagcggaggaaagc	(GTTGTG) ₄	Perfeito	282	4	TD56	LBA+OL	S
PE09	ggaaatccgaaaactggtg	gggcctttatccatgttga	(AT) ₅ (AC) ₈	Composto	268	4	TD56	LBA+OL	S
PE10	aaccttgatctccagcctat	gttttcgccgcgtatt	(GA) ₃₄	Interrompido	234	4	TD60	LBA	S
PE11	gcataagttgtcggcttgg	cctcgaaccttatcatcca	(GT) ₁₁	Perfeito	178	4	TD60	LBA	S
PE12	cgtaatatgtttgggcact	atcatgggcgaactcattt	(TG) ₈	Perfeito	150	7	TD60	LBA+OL	S
PE13	aagcacccaatcggtga	cccctgcacactgagta	(GT) ₆	Interrompido	172	4	TD60	LBA+OL	N
PE14	tggtgttgctgaatttcattt	tacgcgcctagcgtattctt	(GA) ₂₁	Imperfeito	245	4	TD60	LBA	S
PE15	accgttaaatccaagcaagt	aaatgcaaagaatgatatgtta	(CTTTAGC) ₅	Imperfeito	204	7	60	LBA	N
PE16	cgcattgtgtttcctctg	cagtcaaagctcgttctcc	(TG) ₂₃	Perfeito	240	7	60	LBA+OL	S
PE17	acttcgttggtatgcacttg	catagaatgcaagggttcaca	(AC) ₂₂	Perfeito	295	4	TD56	LBA+OL	N
PE18	ccgtgaaccaaccatttctc	ttgcagcacaacaagtaa	(TG) ₉	Perfeito	220	7	TD60	LBA+OL	S
PE19	ttaacaggacttagcacttga	ctcatccttcttccatttgg	(CA) ₁₄	Perfeito	245	5	52	LBA	N
PE20	aggatcaccatagaaaacat	gttaggttgccattgctctt	(AAAC) ₄	Imperfeito	242	7	TD60	LBA+OL	S

Tabela 7 - Código do loco e sequência de bases dos 107 pares de *primers* de microssatélites desenhados para *P. edulis* f. *flavicarpa*, tipo de motivo e sua classificação segundo Weber (1990), tamanho do alelo sequenciado, *mix* e temperatura de anelamento (TA) ideais, além da qualidade da amplificação encontrada e dos locos com presença de bandas *stutter*

(continuação)

Loco	Primer forward	Primer reverse	Motivo	Classes	Alelo (pb)	Mix	TA	Qualidade da amplificação	Bandas <i>stutter</i>
PE21	cccgaagattggtcgta	atccaatggcaggaaggct	(GT) ₁₅	Perfeito	237	4	TD60	LBA+OL	S
PE22	cctcaatcaggcacgcttat	ctgggcataccgaagtc	(ACCG) ₄	Perfeito	241	5	TD60	LBA+OL	N
PE23	caatcccttgaccataga	cgtccatccttctccttt	(GA) ₁₉	Perfeito	206	7	TD56	LBA+OL	S
PE24	tcaaactgaactcgtaaagg	gtgctgggagactgatgtt	(CA) ₁₅	Interrompido	294	5	TD60	LBA+OL	S
PE25	tagagaaaagcacacacaca	cgaggtcgagtttacagaaa	(CA) ₉	Perfeito	246	4	56	LBA+OL	S
PE26	gctttcatattcgggttg	ttgcttgagttggaggaag	(TG) ₁₇	Perfeito	277	4	TD60	LBA+OL	S
PE27	ttgctcattgcactcatct	gcagacatttctggagca	(GT) ₇	Perfeito	139	5	TD60	LBA	S
PE28	gccactaacgttaactgtgct	caagctcttattagcatcca	(CT) ₁₆	Perfeito	243	4	TD60	LBA	N
PE29	cggatgaaggctcgtcttt	cggcacactcacctctcc	(AG) ₁₃	Interrompido	314	7	TD60	LBA	S
PE30	tagccagcaactgacaccaa	ccagaataccaatcccagggt	(TTATGG) ₃	Perfeito	199	-	-	SA	-
PE31	ccctcaacactctgctatct	catcacctctaaccacacaaa	(TG) ₇	Perfeito	307	7	TD60	LBA+OL	S
PE32	ggccaaccattcaaccaata	caagcacatgaatcaaaatcg	(GA) ₁₉	Interrompido	359	-	-	Rastro	-
PE33	cgtgcttgggcgagagtagt	cgtgattctgcggcttc	(ATCGT) ₇	Perfeito	200	2	TD60	LBA+OL	S
PE34	actgcgtgtgattgtgt	gaagattgcctactcgcc	(TG) ₉	Imperfeito	368	-	-	Rastro	-
PE35	attatgcctaaaaacccaaa	tgatccagaggttgagagg	(CA) ₉	Imperfeito	225	7	TD60	LBA	S
PE36	taagctgccttcccctcagt	tgaacctgctgctcatgtc	(AG) ₃₃	Interrompido	226	5	TD60	AOL	S
PE37	caaaaggataggcctgatgtc	tgcttggtcatccactgaag	(TG) ₈	Perfeito	232	2	TD60	LBA	N
PE38	gatcggtcctcggttagac	agtcacacagcatgagaaatc	(TG) ₈	Perfeito	215	5	TD56	LBA	S
PE39	ctttcaggtcccgatcattt	tatcttcaagttggccctgt	(CT) ₉ +(AG) ₉	Interrompido	187	-	-	SA	-

Tabela 7 - Código do loco e sequência de bases dos 107 pares de *primers* de microssatélites desenhados para *P. edulis* f. *flavicarpa*, tipo de motivo e sua classificação segundo Weber (1990), tamanho do alelo sequenciado, *mix* e temperatura de anelamento (TA) ideais, além da qualidade da amplificação encontrada e dos locos com presença de bandas *stutter*

(continuação)

Loco	Primer forward	Primer reverse	Motivo	Classes	Alelo (pb)	Mix	TA	Qualidade da amplificação	Bandas <i>stutter</i>
PE40	acaacaaagactgcgggaac	ctcttttgggttgctctctg	(GAAA) ₅	Interrompido	223	5	TD60	LBA+OL	S
PE41	atcgggggttcgcttatttg	cgttcacaccttagtgggcta	(TTAA) ₅	Interrompido	220	5	TD60	LBA	S
PE42	gtcacttcattcttctttcc	ttagcccaactcaaacacaa	(GT) ₈	Perfeito	216	7	TD60	LBA	N
PE43	ggaatcactcttgcgcttct	tctgtcttatgccactgttg	(TG) ₁₃	Perfeito	241	2	TD60	LBA+OL	N
PE44	tggtgtgccctttctctct	gcacacacacgacgtatgga	(TA) ₆ (TG) ₁₂	Composto	165	4	TD60	LBA+OL	S
PE45	gcatttaccgcgaattacc	caagttccaggcctcctc	(CA) ₂₆	Interrompido	249	5	TD60	AOL	N
PE46	catttctcagtcaccgata	gtcaatgcagtcattcacaa	(AT) ₂₈ (TG) ₁₆	Composto	294	7	TD60	LBA+OL	S
PE47	ttcactcacgcattctatcac	accgaaacatcccagaag	(GT) ₈	Perfeito	364	7	TD56	LBA+OL	S
PE48	aatactgacaagaaagtggct	acggagagcaaatgactaca	(GT) ₁₁	Interrompido	259	-	-	SA	-
PE49	ggccatttgcctgattta	ggagaacattgctaagtgc	(TG) ₇	Perfeito	283	5	TD60	LBA+OL	S
PE50	ctcgattgccattagtagtca	aaaacttccgcttggttag	(AC) ₁₆	Perfeito	212	-	-	SA	-
PE51	tcaagggttgattcggttcag	acacaatcggtggaaggat	(AT) ₉ (GT) ₁₃	Composto	206	5	TD60	AOL	S
PE52	tgagcatttgggtggttt	ttctttcttcttcttctcc	(AGA) ₁₁	Interrompido	245	5	TD60	LBA+OL	S
PE53	cccttgccttgtttccttt	gcagtttcggttggttctc	(GT) ₈	Perfeito	296	5	TD60	AOL	S
PE54	tggtgtgtgtgggtgattag	cattctcctgccactgagt	(TG) ₇	Perfeito	176	2	60	LBA	N
PE55	gacctagaactgccttga	atctgtgctgctactggct	(TG) ₈	Perfeito	221	-	-	SA	-
PE56	ccttgttggttttcgactaca	tgaggacaatcaatcggaca	(GT) ₁₅	Interrompido	207	7	TD60	AOL	N
PE57	gtggagttctgcagtgag	aaggaaatccgtctcttgc	(TG) ₉	Perfeito	?	5	TD56	LBA+OL	N
PE58	gcaatttcaccattctctgct	ccacggctatggatgttc	(AC) ₁₁	Perfeito	243	2	TD60	LBA	S

Tabela 7 - Código do loco e sequência de bases dos 107 pares de *primers* de microssatélites desenhados para *P. edulis* f. *flavicarpa*, tipo de motivo e sua classificação segundo Weber (1990), tamanho do alelo seqüenciado, *mix* e temperatura de anelamento (TA) ideais, além da qualidade da amplificação encontrada e dos locos com presença de bandas *stutter*

(continuação)

Loco	Primer forward	Primer reverse	Motivo	Classes	Alelo (pb)	Mix	TA	Qualidade da amplificação	Bandas <i>stutter</i>
PE59	gaacacttcgcatggctaga	ttccgaatcaaaccgtaact	(ATCTA) ₃	Perfeito	276	7	TD56	LBA	S
PE60	tcctcaccttgtttatgct	aatgacctattgaacctgga	(TG) ₁₂	Perfeito	139	4	TD56	LBA	S
PE61	ctgtagccctgcaacaacat	agacgcgagccagatcatatc	(T) ₃₅	Interrompido	242	-	-	Rastro	-
PE62	accgagacaacgctagagga	caacgtattctacccataggc	(AG) ₁₀	Interrompido	228	5	TD60	LBA+OL	S
PE63	tgtagctccctcatgtcggtga	tcagtgttcacccatgtagcagg	(TG) ₉	Perfeito	150	-	-	Rastro	-
PE64	atcaattacgcaccccaaac	ggaacgtcaatcaagttagga	(AC) ₈	Perfeito	228	4	TD56	LBA	N
PE65	ccttgctgatttctgtgtg	acgctcatggaggtaaagc	(CT) ₁₆ (GT) ₁₄	Composto	247	4	TD60	AOL	S
PE66	ccatagtccaacaagcatc	gctgtggaccctaactcagtc	(AC) ₉	Perfeito	165	5	TD60	LBA	N
PE67	aacctccaaagcaactctc	aaaccgaatagcacatcatt	(GT) ₁₅	Interrompido	249	5	TD60	LBA+OL	N
PE68	aatcccagcagccacaac	cgaacaagtgaagaagaaga	(CT) ₁₂	Interrompido	294	4	60	LBA+OL	S
PE69	cgacacgggacgggtgtt	acgtgtccgggtagtcag	(TG) ₈	Perfeito	364	5	TD56	LBA+OL	S
PE70	cggagcttgggtgagtg	ggacacacatggacaccaga	(GT) ₁₂	Interrompido	259	5	TD60	LBA+OL	S
PE71	attcctatccaatctgttcc	cctgtttcttcttcttctt	(AAAG) ₁₃	Imperfeito	283	-	-	SA	-
PE72	accacatcacacaatatca	tgagaggagagaaaaatcgt	(CATAT) ₅	Perfeito	212	5	TD60	AOL	S
PE73	tagtgtggggagggttaca	cccaatggctcaaaatgact	(TTTTTG) ₃	Perfeito	324	5	60	LBA+OL	S
PE74	ccctcttatcaatagcgttgg	gcacgagcacgagtattatt	(ATCACA) ₅	Interrompido	215	2	TD56	LBA	N
PE75	cacaatcggtgggaaagata	gtagtttgggcagtttgc	(TG) ₁₇	Perfeito	178	7	TD60	LBA	S
PE76	actgttcttgattccgataa	ctacttaccgctgacacac	(TG) ₈	Perfeito	392	-	-	Rastro	-
PE77	ttggcctcgcataagaactc	gatcggtctgtgtgtgtgtt	(CA) ₈	Perfeito	126	7	60	LBA+OL	S

Tabela 7 - Código do loco e sequência de bases dos 107 pares de *primers* de microssatélites desenhados para *P. edulis* f. *flavicarpa*, tipo de motivo e sua classificação segundo Weber (1990), tamanho do alelo seqüenciado, *mix* e temperatura de anelamento (TA) ideais, além da qualidade da amplificação encontrada e dos locos com presença de bandas *stutter*

(continuação)									
Loco	Primer forward	Primer reverse	Motivo	Classes	Alelo (pb)	Mix	TA	Qualidade da amplificação	Bandas <i>stutter</i>
PE78	ctattcacaacctcaaacca	aagactgacctaatactgactgcc	(GA) ₂₉	Imperfeito	238	-	-	SA	-
PE79	tccagtgtgattgctcgcttc	tgcttgcatagattgctga	(T) ₂₂	Imperfeito	204	-	-	SA	-
PE80	tgagcctttgtgttagtttc	agactctcctttcttctct	(GA) ₁₂	Interrompido	321	-	-	SA	-
PE81	acagaggttcagtcagctcat	ttgcatcaatcgtaacaccttc	(CA) ₇	Perfeito	299	4	TD60	AOL	S
PE82	ccttgactcttgaccctga	tatccatccaacctccgaac	(TG) ₉	Perfeito	244	7	TD56	LBA+OL	N
PE83	gtcttctttgtttgtttgg	tgggagatgagagatgaaca	(T) ₁₆	Perfeito	246	-	-	SA	-
PE84	gctctctcttcttctctc	caaccattcaaccagtaaga	(CT) ₂₂	Interrompido	245	-	-	Rastro	-
PE85	gacgcagtcacagatgagga	atgtgcgagcatcagttgac	(TGT) ₅	Perfeito	239	-	-	Rastro	-
PE86	tcattgctgttttcagcaca	tgaataacccaaatgcaatg	(TG) ₁₂	Interrompido	161	-	-	SA	-
PE87	tcttcgtttcactggctcaa	attcactgtcggcacaaaca	(CT) ₂₂	Interrompido	232	-	-	SA	-
PE88	cttcagggtcacacacatt	gttcaccttttagtgggct	(TTAA) ₆	Interrompido	293	7	TD60	LBA	S
PE89	cgctgttatctctgcttacg	ggacgaaggctcgactttta	(TC) ₂₃	Interrompido	327	-	-	SA	-
PE90	tcaggaagattgcatgttagt	ctgggtttgtttatgttgc	(AGC) ₅	Perfeito	245	5	TD60	LBA	N
PE91	gaaggattggccttgcataa	tgatttggggctaaggttg	(CA) ₇	Perfeito	209	5	TD60	LBA+OL	S
PE96	ggttttgcttgattacttttc	ggatcttcccaaatgactc	(CA) ₁₀	Perfeito	184	4	TD60	LBA+OL	S
PE97	gccttcccctcagtataaaa	ctcttctctctctctctcc	(GA) ₃₁	Interrompido	176	7	60	AOL	N
PE98	gaatcgagcctgttatgttt	cgactcgtagttcgtttaagt	(TTAA) ₆	Interrompido	154	2	TD60	LBA+OL	N
PE99	gccacacaacaacaaacaa	aaagccaccgatccctct	(CTT) ₄	Perfeito	289	5	TD60	AOL	S
PE100	aactcctgcacgctggat	ccgtatcggtttgtgaaa	(GT) ₈	Interrompido	246	5	TD60	LBA+OL	S

Tabela 7 - Código do loco e sequência de bases dos 107 pares de *primers* de microssatélites desenhados para *P. edulis* f. *flavicarpa*, tipo de motivo e sua classificação segundo Weber (1990), tamanho do alelo seqüenciado, *mix* e temperatura de anelamento (TA) ideais, além da qualidade da amplificação encontrada e dos locos com presença de bandas *stutter*

(conclusão)									
<i>Loco</i>	<i>Primer forward</i>	<i>Primer reverse</i>	<i>Motivo</i>	<i>Classes</i>	<i>Alelo (pb)</i>	<i>Mix</i>	<i>TA</i>	<i>Qualidade da amplificação</i>	<i>Bandas stutter</i>
<i>PE101</i>	aaaggaggcaaggcaaag	tagatggagagagatgtggaa	(CT) ₁₆	Perfeito	288	2	TD56	LBA+OL	S
<i>PE102</i>	cggaaaccctagacatgaaaa	gagtgagcgtggagagaaaaa	(TC) ₁₄	Interrompido	166	5	TD60	AOL	S
<i>PE103</i>	gatcttcgcgttcagag	ccggtcgaacactaattt	(CT) ₁₄	Interrompido	231	5	TD60	AOL	S
<i>PE104</i>	ctctgaacttcgatcgagca	aggaggggagagatgtgaaa	(AC) ₁₀	Interrompido	176	-	-	Rastro	-
<i>PE105</i>	atggcatgacacaagagt	ccccgatatgtgtcatct	(TA) ₈ (TG) ₉	Composto	160	4	TD60	AOL	S
<i>PE106</i>	tgggttggtgtattgtatg	atcagtcggacagctctttt	(CA) ₆	Interrompido	211	5	TD60	AOL	S
<i>PE107</i>	aaccagacaattccagca	taagatgcttctgtaagggtg	(CT) ₄₀	Interrompido	191	-	-	SA	-
<i>PE108</i>	aacaacgctagacaacgaat	cgagagagagagagagagaga	(CT) ₄₂	Interrompido	165	4	TD56	AOL	S

AOL: sem amplificação dos alelos seqüenciados, mas com amplificação de alelos de outros locos; LBA: loco amplificado, com presença do alelo esperado; LBA + OL: loco amplificado, com presença do alelo esperado, mais outros locos amplificados; AS: sem amplificação; S e N: presença e ausência de bandas *stutter*, respectivamente

4.2 Análise de ligação

4.2.1 Segregação dos marcadores microssatélites na população F₁

Do total de 107 *primers* de microssatélites, 85 apresentaram uma boa amplificação, sendo que 27 apresentaram amplificação específica dos alelos esperados; 42 apresentaram boa amplificação dos alelos desenhados, porém com amplificação de outros alelos de outros locos; e 16 amplificaram outros alelos de tamanhos diferenciados em relação aos desenhados (Tabela 8). Inicialmente, estes *primers* foram utilizados para amplificação dos alelos dos dois genitores juntamente com seis plantas da população segregante, com intuito de verificar a existência de polimorfismo para a construção dos mapas de ligação. Com base nisso, foi possível inferir a provável segregação dos marcadores. Como resultado final, 21 *primers* mostraram-se polimórficos entre os genitores e as seis plantas F₁ (Tabela 8).

Não foi possível promover a amplificação dos alelos dos locos *PE22* e *PE25*, com tamanho esperado de 241 e 246 pb, respectivamente, pois estes *primers* amplificaram outras bandas com pesos moleculares diferentes daquelas para as quais foram desenhados. Algumas destas bandas apresentaram segregação mendeliana e alta reprodutibilidade, sendo utilizadas na construção dos mapas genéticos, mesmo não se tendo a certeza de que sejam seqüências de microssatélites. Estes locos foram codificados da mesma forma que os outros, porém acrescentado o tamanho do alelo que se mostra polimórfico na população, como por exemplo, *PE22_270* (Tabela 8). Sendo assim, foram utilizados 26 locos de microssatélites na construção dos mapas genéticos de maracujá-amarelo.

Kenis e Keulemans (2005) também utilizaram marcadores microssatélites que apresentavam mais de um loco, sendo que os marcadores CH01B12, CH02D12 e CH03g12 mostraram dois conjuntos de bandas oriundas de cada genitor. No trabalho de Hamwich et al. (2005), 28,5% dos locos utilizados na construção dos mapas de ligação são oriundos de fragmentos de outros locos além daqueles para os quais os *primers* foram desenhados. O mesmo foi observado em *Actinidia* spp. (TESTOLIN et al., 2001).

Tabela 8 - Relação dos *primers*/locos que mostraram polimorfismo na população F₁ do cruzamento entre os acessos IAPAR-123 e IAPAR-06 de maracujá-amarelo

<i>Primer/loco</i>	Motivo*	Classes	Alelo (pb)**	Tipo de segregação
<i>PE06</i>	(CA) ₇	Interrompido	187 / 201 / 209	1:1
<i>PE07</i>	(GA) ₂₃	Interrompido	138 / 150	1:1
<i>PE08</i>	(GTTGTG) ₄	Perfeito	273 / 282	1:1
<i>PE09</i>	(AT) ₅ (AC) ₈	Composto	268 / 286	1:1
<i>PE11</i>	(GT) ₁₁	Perfeito	160 / 178 / 186 / n	1:1:1:1
<i>PE12</i>	(TG) ₈	Perfeito	144 / 150	1:1
<i>PE13</i>	(GT) ₆	Interrompido	172 / 280	1:1
<i>PE14</i>	(GA) ₂₁	Imperfeito	245 / 261	1:1
<i>PE15</i>	(CTTTAGC) ₅	Imperfeito	202 / 204 / 210 / n	1:1:1:1
<i>PE16</i>	(TG) ₂₃	Perfeito	232 / 235 / 240 / 248	1:1:1:1
<i>PE17</i>	(AC) ₂₂	Perfeito	273 / 275 / 283 / 295	1:1:1:1
<i>PE18</i>	(TG) ₉	Perfeito	216 / 218 / 220	1:1:1:1
<i>PE19</i>	(CA) ₁₄	Perfeito	239 / 245 / 249	1:1
<i>PE20</i>	(AAAC) ₄	Imperfeito	242 / 258	1:1
<i>PE21</i>	(GT) ₁₅	Perfeito	237 / 233 / 229	1:1:1:1
<i>PE22_270</i>	—	—	270 / n	3:1
<i>PE22_360</i>	—	—	360 / n	1:1
<i>PE22_420</i>	—	—	420 / n	1:1
<i>PE22_640</i>	—	—	640 / n	1:1
<i>PE23</i>	(GA) ₁₉	Perfeito	206 / 210 / n / n	1:1:1:1
<i>PE24</i>	(CA) ₁₅	Interrompido	294 / 306	1:2:1
<i>PE25_206</i>	—	—	206 / n	1:1
<i>PE25_214</i>	—	—	214 / n	1:1
<i>PE25_236</i>	—	—	236 / n	1:1
<i>PE26</i>	(TG) ₁₇	Perfeito	277 / 275	1:2:1
<i>PE27</i>	(GT) ₇	Perfeito	139 / 143	1:1

*“—” loco com segregação mendeliana e tamanhos de alelos diferentes dos esperados

**“n” indica a presença de alelo nulo

No presente trabalho, foram gerados sete conjuntos de marcas que não corresponderam ao tamanho esperado dos alelos dos locos, ou seja, *PE22_270*, *PE22_360*, *PE22_420*, *PE22_640*, *PE25_206*, *PE25_214* e *PE25_236*, mas mesmo assim foram utilizados na construção dos mapas. Neste caso, todos os locos foram alocados nos GL, com exceção do loco *PE22_640*. O restante

foi alocado em 4 grupos de ligação diferentes, com exceção dos locos *PE22_270* e *PE22_360*, que permaneceram no grupo de ligação VI, a uma distância de cerca de 100 cM. Tal fato indica que em maracujá-amarelo é possível o desenvolvimento e utilização de locos de microssatélites cujos *primers* amplificam não somente os alelos para os quais foram desenhados, mas outros alelos de locos distribuídos ao longo do genoma da espécie.

Do total de locos de microssatélites analisados, 75,3% foram monomórficos e 24,7% polimórficos. Do total de locos polimórficos, 61,5% segregaram na proporção de 1:1 (sendo 38,4% informativos para o genitor IAPAR-06, e 23,1% para o genitor IAPAR-123). Os outros 38,5% são locos bi-parentais, sendo que 3,9% segregam na proporção de 3:1; 7,7% na proporção de 1:2:1 e 26,9% na proporção de 1:1:1:1 (Tabela 9).

Em cacau, cerca de 52% dos locos de microssatélites foram polimórficos, sendo que 35,8% foram bi-parentais, 57,7 e 6,4% foram informativos para os genitores UF676 e UPA402, respectivamente (PUGH et al., 2004). Lespinasse et al. (2000) encontraram 46% de polimorfismo em seringueira, sendo 61% bi-parentais e 39% segregando apenas para um genitor. Segundo Testolin et al. (2001), em *Actinidia* spp., 41% dos *primers* de microssatélites foram polimórficos, sendo que 84% segregaram para um dos genitores, e 16% foram bi-parentais. Como se observa, o nível de polimorfismo nestas espécies foi bem maior do que o observado em maracujá-amarelo. Isto pode ser devido à própria diferença do conteúdo genômico destas espécies, ao baixo polimorfismo dos genitores utilizados no cruzamento, bem como ao baixo número de locos utilizados no presente trabalho.

Um fato interessante é que houve amplificação de somente um dos alelos dos locos *PE22_640*, *PE25_206* e *PE25_214* no genitor IAPAR-06, enquanto o genitor IAPAR-123 apresentou alelo nulo (configuração *oo x ao*) (Figura 13 e 14). A mesma configuração (*ao x oo*) foi observada para os locos *PE22_360* e *PE22_420* no genitor IAPAR-123 (Figura 13). Já o loco *PE22_270*, mostrou amplificação de um único alelo em ambos os genitores na configuração *ao x ao* (segregação 3:1). Neste caso, estes alelos levaram à manifestação da dominância, pois os dois genótipos particulares *aa* e *ao* não podem ser distinguidos pelo fenótipo molecular (Figura 13).

Figura 13 - Amplificação do *primer/loco* PE22. (M) Marcador de peso molecular - *Ladder* 10pb. (1) IAPAR-123, (2) IAPAR-06; canaletas 3 a 10: plantas da população segregante F₁. As setas indicam os locos segregantes

Figura 14 - Amplificação do *primer/loco* PE25. (M) Marcador de peso molecular - *Ladder* 10pb. (1) IAPAR-06, (2) IAPAR-123; canaletas 3 a 12: plantas da população segregante F₁. As setas indicam os locos segregantes

Tabela 9 - Tipos de segregação encontrada para os locos de microssatélites amplificados na população F₁ do cruzamento entre IAPAR-123 e IAPAR-06

Segregação (♀ x ♂)	Locos	Alelos da F ₁	Genitor
<i>cc x ab</i>	<i>PE06 e PE19</i>	<i>ac, bc</i>	IAPAR-06
<i>aa x ab</i>	<i>PE08, PE12, PE13, PE14 e PE25_236</i>	<i>a, ab</i>	IAPAR-06
<i>oo x ao</i>	<i>PE22_640, PE25_206 e PE25_214</i>	<i>a, o</i>	IAPAR-06
<i>ao x oo</i>	<i>PE22_360 e PE22_420</i>	<i>a, o</i>	IAPAR-123
<i>ab x aa</i>	<i>PE07, PE09, PE20 e PE27</i>	<i>a, ab</i>	IAPAR-123
<i>ao x ao</i>	<i>PE22_270</i>	<i>a, o</i>	Bi-parentais
<i>ab x ab</i>	<i>PE24 e PE26</i>	<i>a, ab, b</i>	Bi-parentais
<i>ab x co</i>	<i>PE11 e PE15</i>	<i>ac, a, bc, b</i>	Bi-parentais
<i>ab x cd</i>	<i>PE16 e PE17</i>	<i>ac, bc, ad, bd</i>	Bi-parentais
<i>ab x ac</i>	<i>PE18 e PE21</i>	<i>a, ab, ac, bc</i>	Bi-parentais
<i>ao x bo</i>	<i>PE23</i>	<i>ab, b, a, o</i>	Bi-parentais

Nessas situações, os *primers* de microssatélites se anelam a apenas um dos cromossomos homólogos da espécie, de tal sorte que apenas um dos alelos daquele loco é amplificado por PCR em um ou em ambos os genitores. O polimorfismo pode ser decorrente de mutações ou deleções no sítio de anelamento dos *primers*, ou a inserções de bases entre estes sítios. Contudo, os estudos com seqüências de microssatélites mostram que as bases que flanqueiam os microssatélites são bastante conservadas e que o polimorfismo dos microssatélites é devido à diferença no número de repetições. Tal fato conduz à hipótese de que estes locos não sejam de microssatélites, mas sim de outras regiões genômicas aleatórias, uma vez todos estes locos foram obtidos de dois únicos *primers*, *PE22* e *PE25*, em que não foi possível a amplificação dos alelos para os quais foram desenhados. Estes *primers* apresentaram um conjunto de bandas monomórficas, e outras polimórficas.

Os locos *PE11* e *PE15* segregaram na proporção de 1:1:1:1 (fenótipos *ac*, *a*, *bc* e *b*), considerada totalmente informativa, pois todos os possíveis genótipos do loco são identificados por meio dos fenótipos moleculares (Figura 15). Os locos *PE11* e *PE15* apresentam um alelo nulo oriundo do genitor IAPAR-06 (genótipo *co*), enquanto o loco *PE23*, mostra um alelo nulo

oriundo de ambos os genitores (configuração $ao \times bo$) (Figura 16). Tais alelos de microssatélites somente podem ser identificados quando se trabalha com populações de mapeamento, já que é possível a observação do padrão de segregação mendeliana na descendência. Nos estudos de diversidade genética, nos quais são avaliados conjuntos de indivíduos ou de acessos, estes alelos não são detectados, fazendo com que as estimativas de heterozigosidade e de PIC (*Polymorphic Information Content*) sejam subestimadas. De acordo com Powell, Machray e Provan (1996), a ocorrência de alelos nulos pode resultar em viés associado ao uso de microssatélites para análise de espécies não endogâmicas, pois podem levar a subestimativas ou mesmo superestimativas de fluxo gênico.

Os locos *PE16* e *PE17* também mostram segregação totalmente informativa, com quatro alelos ($ab \times cd$) (Figura 17); já os locos *PE18* e *PE21* possuem três diferentes alelos ($ab \times ac$) (Figura 18).

Segregações do tipo 1:2:1, com a configuração $ab \times ab$, foram observadas para os locos *PE24* e *PE26* (Figura 19).

Figura 15 - Amplificação do *primer/loco* *PE11*. (M) Marcador de peso molecular - *Ladder* 10pb. (1) IAPAR-123, (2) IAPAR-06; canaletas 3 a 12: plantas da população segregante F_1 . As setas indicam os locos segregantes

Figura 16 - Amplificação do *primer/loco* PE23. (M) Marcador de peso molecular - *Ladder* 25pb. (1) IAPAR-06, (2) IAPAR-123; canaletas 3 a 12: plantas da população segregante F_1 . As setas indicam os locos segregantes

Figura 17 - Amplificação do *primer/loco* PE17. (M) Marcador de peso molecular - *Ladder* 25pb. (1) IAPAR-123, (2) IAPAR-06; canaletas 3 a 12: plantas da população segregante F_1 . As setas indicam os locos segregantes

Figura 18 - Amplificação do *primer/loco* *PE18*. (M) Marcador de peso molecular - *Ladder* 25pb. (1) IAPAR-06, (2) IAPAR-123, canaletas 3 a 11: plantas da população segregante F_1 . As setas indicam os locos segregantes

Figura 19 - Amplificação do *primer/loco* *PE24*. (M) Marcador de peso molecular - *Ladder* 25pb. (1) IAPAR-06, (2) IAPAR-123; canaletas 3 a 10: plantas da população segregante F_1 . As setas indicam os locos segregantes

Dois alelos oriundos do genitor IAPAR-123 e segregando em F_1 na proporção de 1:1 (*ab x aa*) foram observados para os locos *PE07*, *PE09*, *PE20* e *PE27*, enquanto os alelos locos *PE08*, *PE12*, *PE13*, *PE14* e *PE25_236* segregaram na proporção 1:1 e são oriundos do genitor IAPAR-06 (Tabela 9). De outra maneira, os locos *PE06* e *PE19* apresentaram três diferentes alelos segregando na proporção de 1:1 (genitor IAPAR-06).

4.2.2 Análise das segregações dos marcadores AFLP

Considerando 109 plantas comuns à população F_1 usada no presente estudo, Lopes et al. (2006) mostraram que o nível de polimorfismo com marcadores AFLP foi de aproximadamente 13%, que é quase a metade do polimorfismo encontrado para os locos de microssatélites (24,7%).

Para o mapeamento das marcas, foram utilizados 253 locos polimórficos com padrão de segregação 1:1, sendo que em 114 deles o alelo dominante (presença da banda) estava presente no genitor IAPAR-06, enquanto em 139 deles, o alelo dominante estava presente no genitor IAPAR-123. Outros 116 locos bi-parentais, segregando na proporção de 3:1, foram também incluídos na construção dos mapas genéticos. Este mesmo conjunto de dados foi analisado por Matta (2005) e Moraes (2005) na construção de mapas genéticos de maracujá-amarelo.

Cabe lembrar que a diferença deste estudo em relação aos dois anteriores refere-se à inclusão de marcas de microssatélites, visando à construção de um mapa integrado.

4.2.3 Distorções de segregação

Foi adotado o teste de aderência de χ^2 (qui-quadrado), com nível de significância $\alpha=0,05$, para verificar a hipótese nula (H_0) de segregação mendeliana. Para a verificação da existência de distorções de segregações das marcas utilizou-se o critério FDR (*False Discovery Rate*), definido como sendo a proporção de hipóteses nulas verdadeiras entre as hipóteses nulas rejeitadas, ou seja, a proporção de erros devidos à falsa rejeição de H_0 , também chamada proporção de falsos positivos. Mesmo após a correção de FDR, as marcas EM03174 e EM08348, presentes no genitor IAPAR-06, apresentaram desvio da proporção 1:1. Em relação às marcas bi-parentais, 12 apresentaram desvio da proporção de 3:1. Nenhuma marca de microssatélite apresentou distorção de segregação após a correção usando FDR (Tabela 10).

Distorções de segregação de marcadores têm sido observadas para algumas espécies não endogâmicas como *Populus* (CERVERA et al., 2001), eucalipto (GRATTAPAGLIA e SEDEROFF, 1994) e maracujá-amarelo (LOPES et al., 2006; MATTA, 2005; MORAES, 2005). Os desvios de segregação podem ser devidos a fatores biológicos ou de amostragem da população utilizada para fins de mapeamento. Cervera et al. (2001) observaram que algumas marcas que co-segregavam com genes de resistência a *Melampsora* também mostravam desvios

de segregação devido à morte de genótipos suscetíveis. Segundo Kuang et al. (1999), o agrupamento em *clusters* de marcadores que apresentam distorção de segregação, parece estar associado a regiões genômicas que afetam a viabilidade dos zigotos. A não inclusão dessas marcas pode comprometer a formação dos grupos de ligação e a saturação de regiões com baixa densidade de marcas. Contudo, a entrada dessas marcas pode aumentar a chance de se cometer o erro Tipo I (ou seja, rejeição da hipótese de nulidade), além de incorrer na falta de acurácia na determinação das distâncias de mapa (CLOUTIER, CAPPADOCIA e LANDRY, 1997).

Tabela 10 - Número e tipos de marcas utilizadas na construção dos mapas genéticos de maracujá-amarelo

Nº de locos	Mono-parentais (1:1)		Bi-parentais		
	IAPAR-06	IAPAR-123	3:1	1:2:1	1:1:1:1
AFLP - polimórficos	114	139	116	-	-
Microsatélites - polimórficos	10	6	1	2	7
Com distorção de segregação esperada	2	-	12	-	-
Com mais de 10% de dados perdidos	14	19	40	-	-
Utilizados na construção dos mapas	108	126	65	2	7

Critérios que controlam o nível de significância global são recomendados para evitar que seja cometido o erro tipo I quando são realizados múltiplos testes (BEARZOTI, 2000; SILVA e VENCOSKY, 2002). No caso, isto corresponde a rejeitar erroneamente um loco que apresenta o padrão de segregação testado em H_0 .

Moraes (2005) construiu mapas individuais para os genitores IAPAR-123 e IAPAR-06, utilizando as mesmas marcas de AFLP aqui usadas, com segregação 1:1 e 3:1, e o critério de FDR para verificação de distorções da segregação mendeliana. Segundo este autor, se fosse utilizado apenas o nível de significância individual ($\alpha = 0,05$), teriam sido considerados com desvio de segregação 15 locos do acesso IAPAR-06 (~13%), 17 locos do IAPAR-123 (~12%) e 22 locos bi-parentais (~19%), os quais não seriam incluídos nas análises de ligação.

Alguns autores utilizam a correção de Bonferroni como uma alternativa para contornar o problema de múltiplos testes (PROVINCE, 1999). Supondo t testes e adotando-se um nível de

significância (α) para cada teste, tem-se que o nível de significância conjunto do teste (α^*), considerando os t testes independentes, será:

α^* = probabilidade de rejeitar pelo menos uma hipótese nula verdadeira, ou seja,

$\alpha^* = 1 -$ probabilidade de não rejeitar nenhuma hipótese nula verdadeira, isto é,

$\alpha^* = 1 - (1 - \alpha)^t$. Como o nível de significância conjunto cresce à medida que aumenta o número de testes realizados, a correção de Bonferroni determina o valor do nível de significância individual (α) que proporcionará o nível de significância conjunto (α^*) desejado, dado por:

$$\alpha = -\exp\left(\frac{\ln(1 - \alpha^*)}{t}\right) + 1. \text{ A rigor, os testes não são independentes, uma vez que os}$$

marcadores pertencentes ao mesmo grupo de ligação co-segregam. Isso faz com que o critério de Bonferroni seja um pouco conservador, ou seja, o valor real de α^* é menor do que o nominal, o que pode aumentar a dificuldade em se rejeitar H_0 (BEARZOTI, 2000).

O critério de FDR proposto por Storey e Tibshirani (2003) tem como vantagem o fato de que o nível de significância e o poder do teste independem do número de testes realizados (t).

4.2.4 Construção dos mapas de ligação

Foram utilizadas 308 marcas para construção dos mapas de ligação, todas sem distorção de segregação e com menos de 10% de observações perdidas, sendo 126 informativas para o genitor IAPAR-123, 108 para o IAPAR-06 e 74 bi-parentais (Tabela 10). As análises de dois pontos foram usadas para estimar as frequências de recombinação e as fases de ligação separadamente para cada par de marcas. As fases de ligação entre dois marcadores adjacentes foram determinadas pela comparação das probabilidades *a posteriori* entre os quatro possíveis arranjos de acordo com a equação (2) e pelos valores das frequências de recombinação.

Utilizando a metodologia proposta por Wu et al. (2002) para promover a integração dos mapas dos genitores, 298 marcas (97%) foram ligadas com $LOD \geq 5,0$ e frequência de recombinação $\leq 0,35$, formando 10 grupos de ligação. Contudo, um dos grupos de ligação foi formado por um grande número de marcas (107). Neste grupo, uma das marcas de AFLP estava ligada a apenas uma marca. Ao se retirar esta marca (possivelmente um falso positivo), o grupo foi dividido em dois, sendo formados, assim, 11 grupos de ligação. Foram encontrados dois

grupos de ligação a mais em relação ao número de pares de cromossomos homólogos de maracujá-amarelo ($2n=18$), sendo que um dos grupos (GL-X) foi formado por apenas duas marcas, informativas para o genitor IAPAR-06. Provavelmente, este grupo pertença a algum dos grupos de ligação formados, porém não foi possível promover sua ligação.

Moraes (2005) comenta em seu trabalho que existem indicações de que os GL-IX e X do IAPAR-06, ambos constituídos por dois locos no *framework*, sejam, na verdade, um único grupo, haja vista que esses dois grupos se unem por intermédio de marcas de segregação 3:1 que foram descartadas devido a apresentarem muitos dados perdidos. Entretanto, estas observações não foram confirmadas com a entrada dos locos de microssatélites no mapa único integrado.

Além disso, outros dois grupos de ligação não puderam ser integrados em virtude da inexistência de marcas bi-parentais, formando GL separados, sendo denominados GL-IV(06) e GL-IV(123), pertencentes aos genitores IAPAR-06 e IAPAR-123, respectivamente. Tais grupos foram considerados como pertencentes ao GL-IV, que permaneceu separado como grupos de ligação indivíduo-específico.

Na Tabela 11 é apresentada a distribuição do número e dos tipos de marcas encontradas por grupo de ligação. Das 16 marcas de microssatélites com segregação do tipo 1:1, 12 foram agrupadas e 11 alocadas no *framework* dos mapas. Todas as marcas de microssatélites com segregação 3:1 e 1:2:1 foram alocadas no *framework*. No entanto, das 7 marcas com segregação totalmente informativa (1:1:1:1), apenas 3 permaneceram no *framework*. Mesmo empregando apenas 26 locos de microssatélites, foi possível observar sua distribuição em todos os grupos de ligação, com exceção do GL-V e GL-IX, já que o GL-X provavelmente pertença a um dos outros grupos.

Apesar da utilização de poucos locos de microssatélites (26) de maracujá-amarelo, nossos resultados estão de acordo com a distribuição aleatória e com a boa cobertura do genoma oferecida por estes marcadores. Tal como observado em eucalipto (BRONDANI et al., 1998) e pinus (DEVEY et al., 1999).

Com relação às marcas de AFLP, 83,0 e 87,5% daquelas com segregação do tipo 1:1 e 3:1, respectivamente, foram alocadas no *framework* dos grupos de ligação (Tabela 11).

Tabela 11 - Distribuição do número, tipos de marcas agrupadas e alocadas no *framework* de cada grupo de ligação do mapa integrado dos genitores IAPAR-123 e IAPAR-06

GL	Nº de marcas agrupadas							Nº de marcas alocadas no <i>framework</i>						
	AFLP		Microssatélite				Total	AFLP		Microssatélite				Total
	1:1	3:1	1:1	3:1	1:2:1	1:1:1:1		1:1	3:1	1:1	3:1	1:2:1	1:1:1:1	
I	30	11	2	-	1	-	44	24	10	2	-	1	-	37
II	27	6	1	-	-	2	36	23	4	-	-	-	1	28
III	31	9	1	-	-	-	41	20	9	1	-	-	-	30
IV(06)	8	-	1	-	-	-	9	8	-	1	-	-	-	9
IV(123)	15	-	1	-	-	-	16	15	-	1	-	-	-	16
V	14	3	-	-	-	-	17	13	3	-	-	-	-	16
VI	39	18	2	1	-	3	63	32	16	2	1	-	1	52
VII	25	3	3	-	-	2	33	23	3	3	-	-	1	30
VIII	11	6	1	-	1	-	19	11	6	1	-	1	-	19
IX	13	5	-	-	-	-	18	10	5	-	-	-	-	15
X	2	-	-	-	-	-	2	2	-	-	-	-	-	2
Não ligados	3	3	4	-	-	-	10	-	-	-	-	-	-	-
Total	218	64	16	1	2	7	308	181	56	11	1	2	3	254

Nove grupos de ligação foram obtidos para os genitores IAPAR-123 e IAPAR-06, em mapas de ligação previamente construídos para esta mesma população utilizando a estratégia duplo pseudocruzamento teste com base em marcadores RAPD e AFLP com segregação 1:1 (CARNEIRO et al., 2002; LOPES et al., 2006).

Matta (2005) construiu o primeiro mapa integrado de maracujá-amarelo para 160 plantas desta população, utilizando marcas bi-parentais de AFLP com segregação 3:1. Neste estudo, foram encontrados apenas 8 grupos de ligação. O autor utilizou o programa JoinMap (VAN OOIJEN e VOORRIPS, 2001). As principais diferenças no padrão de agrupamento com relação ao presente trabalho referem-se aos GL-IV(06), IV(123), III, VII, IX e X. Nos trabalhos conduzidos por Matta (2005), os GL-IV(06) e IV(123) foram unidos pela marca bi-parental de AFLP EM1695, a qual foi agrupada, porém aqui não alocada no *framework* do GL-II. Outra diferença é que o GL-II (MATTA, 2005) deu origem aos GL-III e VII (presente trabalho) e, da mesma forma, o GL-VI deu origem aos GL-IX e X.

Moraes (2005) utilizou marcas de AFLP com segregação 1:1 e 3:1, na construção dos mapas de ligação desta espécie, usando o programa computacional TreeMap (COELHO, 2005) e o programa Mapmaker/EXP versão 3.0 (LANDER et al., 1987). Primeiramente, os conjuntos contendo os locos com segregação 1:1 e 3:1 de cada genitor foram analisados no TreeMap, o qual permite a análise de locos com mistura de segregações, sendo anotadas as posições dos locos 3:1 nos grupos de ligação. Em seguida os locos com segregação 1:1 foram duplicados e recodificados, para serem analisados pelo programa Mapmaker/EXP, para obtenção de estimativas multiponto das distâncias entre marcas, haja vista que o TreeMap não possui esta opção. Os locos 3:1 e os 1:1 que apresentavam problemas de ordenação foram colocados como acessórios. Foram encontrados 9 GL para o genitor IAPAR-123 e 10 GL para o IAPAR-06. A homologia entre os grupos de ligação foi estabelecida por meio dos locos bi-parentais (3:1) alocados como acessórios. No trabalho de Moraes (2005) houve um avanço, pois se estabeleceu a homologia dos GL que estavam separados, um para cada genitor.

No presente trabalho, as marcas dentro de cada grupo foram ordenadas pelo algoritmo RCD (*Rapid Chain Delineation*), baseado na menor SARF (*Sum of Adjacent Recombination Fraction*). Diferentemente do trabalho de Moraes (2005), marcas biparentais 3:1, que não “atrapalhavam” a ordenação das marcas foram mantidas no *framework*. Em todas as etapas de ordenação, as 10 melhores ordens fornecidas pelo algoritmo RCD foram analisadas, sempre

verificando os conjuntos de marcas que mantinham suas ordens nos GL e eliminando as que causavam problemas na ordenação. Assim, quando as 10 melhores ordens eram iguais, a ordenação foi considerada satisfatória para todo o conjunto de dados analisados.

Após esta etapa, com uma ordem pré-estabelecida, foram feitas as análises de três pontos. Por exemplo, considerando $M^1, M^2 \dots M^n$ como sendo a ordem de um determinado grupo de ligação, a primeira análise foi feita para as marcas M^1, M^2 e M^3 , estimando, simultaneamente, a fase de ligação e a frequência de recombinação para as marcas M^2 e M^3 . A segunda análise de três pontos foi feita para as marcas M^2, M^3 e M^4 , estimando os mesmos parâmetros, para as marcas M^3 e M^4 , já considerando a fase de ligação de M^2 . Tal procedimento foi feito até a última marca. Após a obtenção das estimativas de três pontos, foi possível promover a confirmação da ordenação fornecida anteriormente pelo RCD. Este procedimento é recente na literatura, mormente em plantas.

Quando os marcadores adjacentes são submetidos a análise de três pontos e a ordem é confirmada, são obtidas duas estimativas diferentes para a frequência de recombinação para um mesmo intervalo, com exceção do intervalo correspondente aos extremos dos grupos de ligação. Neste caso, utilizou-se o recíproco da variância das duas estimativas separadas como uma ponderação na obtenção das frequências de recombinação de cada intervalo (RIDOUT et al., 1998). Esta abordagem multiponto permite a obtenção de estimativas da fração de recombinação com maior acurácia.

Os grupos de ligação deste trabalho foram constituídos por duas (GL-X) até 52 (GL-VI) marcas alocadas no *framework* e por um total de 254 marcas (82,47% das marcas analisadas), com média de 25 marcas por GL. A distância total, considerando a soma dos GL foi de 1.765,6 cM, variando de 3,9 a 307,1 cM, com uma média de 176 cM por GL. A distância média entre marcas consecutivas no *framework* foi de 6,95 cM, sendo que o maior intervalo foi de 35,15 cM (GL VII) (Figuras 20 a 27 e Tabela 12).

Tabela 12 - Comparações entre os mapas de ligação dos acessos IAPAR-06 e IAPAR-123 construídos com base em locos AFLP e usando a estratégia duplo pseudocruzamento teste (MORAES, 2005) e o mapa integrado construído no presente trabalho. O tamanho da população foi de 160 indivíduos em ambos os estudos

Característica	IAPAR-123	IAPAR-06	Mapa Integrado
Número de locos	179	154	298
Grupos de ligação	9	10	10
Marcas ordenadas no <i>framework</i>	57	63	254
Marcas acessórias	122	90	-
Comprimento do mapa (cM – Kosambi)	471,5	727,8	1.765,6
Comprimento do maior grupo (cM – Kosambi)	93,5	212,1	307,1
Comprimento do menor grupo (cM – Kosambi)	11,1	3,9	3,9
Comprimento médio dos grupos (cM – Kosambi)	52,4	72,8	176,5
Distância média entre marcas (cM – Kosambi)	9,8	13,7	6,95

4.2.4.1 Comparações com mapas de maracujá-amarelo previamente construídos

No mapa único integrado não foi atribuída a nenhuma marca a condição de marca acessória, sendo todas alocadas no *framework* (sem prejuízo da ordenação das marcas). No trabalho de Moraes (2005), 59% das marcas em relação às do *framework* foram alocadas como acessórias (Tabela 12). Este procedimento foi necessário para eliminar as marcas com segregação 3:1 e algumas 1:1 que prejudicavam a ordenação para, em seguida, utilizar o mapa genético para a identificação de QTL via QTLCartographer versão 1.17 (BASTEN, WEIR e ZENG, 2004), que não trabalha com misturas de segregações.

Uma observação importante é que o número de marcas no *framework* e o comprimento dos grupos de ligação aumentaram em relação ao trabalho de Moraes (2005) (Tabela 13). Ainda, com a maior saturação dos mapas de ligação, a distância média entre marcas passou de 13,7 e 9,8 cM nos mapas do IAPAR-123 e IAPAR-06, respectivamente, para 6,95 cM no mapa integrado (Tabela 12).

Observou-se que cerca de 70% dos marcadores concentraram-se em cinco grupos de ligação (GL-I, II, III, VI e VII), cobrindo 63,15% do comprimento dos GL (Tabela 13). Resultados semelhantes em maracujá-amarelo também foram observados por Matta (2005), pois

cerca de 70% dos marcadores concentraram-se nos 5 primeiros GL, cobrindo 75% do comprimento total dos grupos de ligação (genoma).

Nas Figuras 20 a 27, são apresentados os 10 GL de maracujá-amarelo integrados, juntamente com os GL obtidos por Moraes (2005) para os genitores IAPAR-123 e IAPAR-06. Os GL do IAPAR-06 e do IAPAR-123 estão denominados com o número do grupo seguido da identificação 06 e 123, respectivamente, enquanto o mapa integrado recebe apenas o número do grupo de ligação, com exceção do grupo de ligação não integrado (GL-IV), cujos mapas receberam a denominação “06” e “123” para diferenciar dos mapas individuais de Moraes (2005).

Tabela 13 - Correspondência entre o número de marcas e comprimento dos grupos de ligação dos mapas individuais de maracujá-amarelo construídos por Moraes (2005) e do mapa integrado obtido pelo presente estudo

Grupos de ligação	IAPAR-123 ^a		IAPAR-06 ^a		Mapa Integrado	
	Nº de marcas	Comprimento (cM)	Nº de marcas	Comprimento (cM)	Nº de marcas	Comprimento (cM)
GL-I	7	43,7	6	17,0	37	180,4
GL-II	7	62,2	11	212,1	28	236,6
GL-III	7	75,0	11	131,9	30	153,9
GL-IV(06)	-	-	6	83,5	9	133,0
GL-IV(123)	7	41,1	-	-	16	101,3
GL-V	5	40,3	5	59,1	16	117,8
GL-VI	9	67,3	10	141,9	52	307,1
GL-VII	2	11,1	5	21,7	30	236,9
GL-VIII	5	37,3	5	45,6	19	150,9
GL-IX	8	93,5	2	11,1	15	147,5
GL-X	-	-	2	3,9	2	3,9

^aMapas obtidos por Moraes (2005)

De modo geral, houve uma grande correspondência entre os mapas individuais e o mapa integrado, com relação à atribuição e à ordenação das marcas nos grupos de ligação. Contudo, algumas diferenças podem ser observadas em alguns grupos de ligação do mapa integrado. No GL-I, a marca *D_I*-EM06273, informativa para o IAPAR-123, que estava ligada a 4,4 cM da

marca D_1 -EM01525 no mapa de Moraes (2005) ficou separada a uma distância de 82,8 cM no mapa integrado. Tal fato pode ter ocorrido devido ao aumento do número de marcas e a não entrada da marca D_1 -EM17230 no mapa integrado. Neste mesmo grupo, ainda houve uma inversão das marcas D_2 -EM19226 e D_2 -EM13340, informativas para o IAPAR-06. Além disso, a marca D_2 -EM24343 não foi alocada no *framework* do mapa integrado (Figura 20).

No GL-II, três marcas presentes no genitor IAPAR-06 não foram alocadas no *framework* do mapa integrado, pois não foi possível obter uma ordenação consistente com as outras marcas. Além disso, algumas inversões das ordens foram observadas para as marcas D_2 -EM12182, D_2 -EM06119 e D_2 -EM1387, informativas para o IAPAR-06, e para a marca D_1 -EM22112, informativa para o IAPAR-123 (Figura 21).

Já no GL-III, a marca D_1 -EM21254 (IAPAR-123) e as marcas D_2 -EM08570, D_2 -EM01130, D_2 -PM01332, D_2 -EM09177, D_2 -EM16512, D_2 -EM01234 e D_2 -EM13620 (IAPAR-06) não foram alocadas no mapa integrado. Ainda duas inversões na ordem foram observadas para duas marcas informativas para o IAPAR-123 (D_1 -EM14379 e D_1 -EM21204) e IAPAR-06 (D_2 -EM13100 e D_2 -EM20326) (Figura 22).

O GL-IV não foi integrado devido à falta de marcas bi-parentais. Entretanto, nove e três outras marcas foram alocadas no *framework* do GL-IV(123) e GL-IV(06), respectivamente. Houve, também, inversão na ordem das marcas D_1 -EM15204, D_1 -PM08137 e D_1 -EM24470, informativas para o genitor IAPAR-123 (Figura 23).

No GL-V, houve uma inversão na ordem das marcas D_1 -EM04114 e D_1 -EM17240, informativas para o genitor IAPAR-123 e a marca D_2 -EM02369 (IAPAR-06) não foi alocada no mapa integrado (Figura 24). Maiores diferenças foram observadas para o GL-VI onde não houve integração de uma e cinco marcas dos genitores IAPAR-123 e IAPAR-06, respectivamente. Inversões também foram observadas em marcas informativas para o IAPAR-123 (D_1 -EM09360 e D_1 -EM24335) e IAPAR-06 (D_2 -EM072363 e D_2 -EM03288) (Figura 25). No GL-VIII, houve apenas uma inversão para as marcas D_2 -EM18124 e D_2 -EM22110, informativas para o IAPAR-06 (Figura 24).

Para a integração do GL-IX, não foi possível a alocação das duas únicas marcas do *framework* do IAPAR-06, sendo que esta integração foi feita para todas as marcas do mapa individual do genitor IAPAR-123. Inversões na ordem das marcas foram observadas apenas para as marcas D_1 -PM01254, D_1 -EM03219 e D_1 -EM16920 (Figura 27).

Figura 20 - Mapa de ligação (GL-I) integrado com marcadores bi-parentais, AFLP e microssatélites (negrito e sublinhado). À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-123 e à direita, para o genitor IAPAR-06. No centro, o mapa integrado obtido no presente trabalho. As linhas mostram as marcas comuns aos mapas; as demais são marcas alocadas como acessórias por Moraes (2005)

Figura 21 - Mapa de ligação (GL-II) integrado com marcadores bi-parentais, AFLP e microssatélites (negrito e sublinhado). À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-123 e à direita, para o genitor IAPAR-06. No centro, o mapa integrado obtido no presente trabalho. As linhas mostram as marcas comuns aos mapas; as demais são marcas alocadas como acessórias por Moraes (2005)

Figura 22 - Mapa de ligação (GL-III) integrado com marcadores bi-parentais, AFLP e microssatélites (negrito e sublinhado). À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-123 e à direita, para o genitor IAPAR-06. No centro, o mapa integrado obtido no presente trabalho. As linhas mostram as marcas comuns aos mapas; as demais são marcas alocadas como acessórias por Moraes (2005)

Figura 23 - Mapa de ligação (GL-IV) integrado com marcadores bi-parentais, AFLP e microsatélites (negrito e sublinhado). À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-123 e IAPAR-06, e à esquerda os mapas obtidos no presente estudo. As linhas mostram as marcas comuns aos mapas; as demais são marcas alocadas como acessórias por Moraes (2005)

Figura 24 - Mapas de ligação (GL-V e VIII) integrados com marcadores bi-parentais, AFLP e microsatélites (negrito e sublinhado). À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-123 e à direita, para o genitor IAPAR-06. No centro, o mapa integrado obtido no presente trabalho. As linhas mostram as marcas comuns aos mapas; as demais são marcas alocadas como acessórias por Moraes (2005)

Figura 25 - Mapa de ligação (GL-VI) integrado com marcadores bi-parentais, AFLP e microssatélites (negrito e sublinhado). À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-123 e à direita, para o genitor IAPAR-06. No centro, o mapa integrado obtido no presente trabalho. As linhas mostram as marcas comuns aos mapas; as demais são marcas alocadas como acessórias por Moraes (2005)

Figura 26 - Mapa de ligação (GL-VII) integrado com marcadores bi-parentais, AFLP e microssatélites (negrito e sublinhado). À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-123 e à direita, para o genitor IAPAR-06. No centro, o mapa integrado obtido no presente trabalho. As linhas mostram as marcas comuns aos mapas; as demais são marcas alocadas como acessórias por Moraes (2005)

Figura 27 - Mapa de ligação (GL-IX) integrados com marcadores bi-parentais, AFLP e microssatélites. À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-123 e à direita, para o genitor IAPAR-06. No centro, o mapa integrado obtido no presente trabalho. As linhas mostram as marcas comuns aos mapas; as demais são marcas alocadas como acessórias por Moraes (2005)

Figura 28 - Mapa de ligação (GL-X) integrado com marcadores bi-parentais, AFLP e microssatélites. À direita, *framework* dos grupos de ligação obtido por Moraes (2005) para o genitor IAPAR-06, e à esquerda o mapa obtidos no presente estudo. As linhas mostram as marcas comuns aos mapas

Para muitos grupos de ligação, os mapas integrados foram consistentes com a ordem das marcas nos GL individuais, embora algumas pequenas diferenças com relação à ordem dos marcadores tenha sido observada. Isso não necessariamente representa um viés, uma vez que a ordem obtida após as análises é apenas uma estimativa da ordem real dos marcadores nos cromossomos e, portanto, pequenas diferenças nas estimativas podem sempre ocorrer. O mapa

único integra locos que segregam em F_1 e cujos alelos estão presentes em um ou ambos os genitores.

Para pares de locos heterozigóticos em ambos os genitores, a frequência de recombinação estimada é uma média da frequência de recombinação das meioses de ambos os genitores. Esta estimativa combinada pode diferir das estimativas de frequência de recombinação em cada genitor e pode causar uma alteração na ordem das marcas no mapa integrado em comparação ao mapa único.

Yin et al. (2002) observaram a existência de inversões entre marcadores nos grupos de ligação de *Populus*. Estes autores levantaram a hipótese de que tal fato seria devido a problemas estatísticos de ordenação ou rearranjos cromossômicos entre as duas espécies utilizadas no cruzamento. O tamanho limitado da população de mapeamento, que reflete o número de meioses usadas para estimar as frequências de recombinação, também influencia na ordem das marcas (BRONDANI, BRONDANI e GRATTAPAGLIA, 2002).

As pequenas inversões na ordem das marcas observadas neste trabalho são devidas, provavelmente, à entrada dos locos de microssatélites e marcas de AFLP com segregação do tipo 3:1 e 1:1, que não foram alocadas no trabalho de Moraes (2005). A atribuição destas marcas no mapa único promove uma maior saturação, aumentando o comprimento dos grupos de ligação (Tabela 12 e 13). Em todos os grupos de ligação, houve aumento no número de marcas e no comprimento dos mapas. Esta maior saturação diminui as distâncias entre as marcas, propiciando maiores chances de inversões de ordenação nos grupos de ligação. Além disso, dificulta a permanência de algumas marcas que são, então, retiradas dos grupos de ligação. Porém, inversões de ordem em marcas próximas têm poucas consequências práticas, tanto para os mapas *per se*, como para a localização de QTL.

Segundo Moraes (2005), a formação de *clusters* em determinadas regiões também dificulta a obtenção da ordem correta dos marcadores. Agrupamentos de marcadores AFLP oriundos da combinação *EcoRI/MseI* aparecem em regiões específicas dos cromossomos (BECKER et al., 1995; KEIM et al., 1997; ALONSO-BLANCO et al., 1998), enquanto combinações do tipo *PstI/MseI* produzem marcas com distribuição mais uniforme no genoma (VUYLSTEKE et al., 1999).

Este mesmo conjunto de dados foi analisado por Matta (2005) para a construção de um mapa integrado para o cruzamento, utilizando os locos bi-parentais de AFLP (3:1) como

marcadores ponte no programa JoinMap. O autor relata que houve uma grande dificuldade no agrupamento dos locos. Segundo o procedimento de agrupamento inserido no JoinMap, os grupos de ligação obtidos por Matta (2005) foram formados com razões de verossimilhança (LOD) distintas. O Grupo III foi formado com um LOD 4, o Grupo I com LOD 6, o Grupo II com LOD 7, os Grupos VI e VIII com LOD 9 e os Grupos IV, V e VII com LOD 10. Ou seja, foram utilizados valores de *LOD scores* muito altos para a separação de alguns grupos, resultando em oito GL. Este autor comenta ainda que a ordenação das marcas nesses GL ficou muito prejudicada e pouco confiável.

Em seguida, Moraes (2005) construiu dois mapas com uma ordenação baseada somente em marcadores 1:1; e o uso de locos 3:1 apenas para a identificação e o alinhamento dos GL homólogos dos genitores. Para isto, foi utilizado o programa TreeMap (COELHO, 2005), cujos algoritmos são fundamentados na mesma metodologia proposta por Wu et al. (2002) para determinação das fases de ligação e da frequência de recombinação. Segundo Moraes (2005), mesmo sendo possível a inclusão de alguns locos bi-parentais no *framework*, estes não foram incluídos, pois os mapas foram construídos visando ao mapeamento de QTL pelo programa QTLCartographer que não trabalha com mistura de segregações.

Mesmo com a disponibilidade de marcadores bi-parentais, muitos autores construíram mapas de ligação separadamente para cada genitor (BRONDANI et al., 1998; SEEFELDER et al., 2000; CERVERA et al., 2001; TESTOLIN et al., 2001; AGRAMA, GEORGE e SALAH, 2002; BRONDANI, BRONDANI e GRATTAPAGLIA, 2002; KENIS e KEULEMANS, 2005). Outros preferiram promover a integração dos mapas utilizando o programa JoinMap (RISTERUCCI et al., 2000; WU, COLLINS e SEDGLEY, 2004; HAMWICH et al., 2005; MATTA, 2005). Optamos aqui pela segunda abordagem (WU et al., 2002), em função das vantagens oferecidas de se ter um mapa integrado, que servirá como um mapa de referência de *P. edulis* f. *flavicarpa*. Estudos de organização do genoma de *Passiflora*, evolução das espécies com base em sintenia de mapas saturados são alguns dos aspectos que se buscou colaborar. Além disso, o mapa integrado aqui construído contribuirá com os estudos sobre mapeamento de QTL de interesse (resistência à bacteriose, por exemplo), controlados por locos cujos alelos de ambos os genitores segreguem em F_1 .

O trabalho de Garcia et al. (2006) mostra uma comparação entre a construção de mapas de ligação em cana-de-açúcar utilizando a metodologia proposta por Wu et al. (2002) e o programa

JoinMap, o qual, basicamente, emprega a metodologia de Malliepard et al. (1997). Nesta comparação, a primeira metodologia promoveu a formação de 131 GL, com 2.602,4 cM de comprimento de mapa e densidade dos marcadores de 7,3 cM. Com o uso do programa JoinMap, foram gerados 98 GL, com 1.340 cM de comprimento de mapa e densidade dos marcadores de 6,2 cM. Assim, a metodologia de Wu et al. (2002) permitiu a detecção de ligação de um maior número de marcadores e a geração de um maior número de grupos de ligação, mostrando-se mais eficiente do que o programa JoinMap para gerar e integrar mapas de ligação de cana-de-açúcar.

No presente trabalho, a metodologia proposta por Wu et al. (2002) foi implementada no programa *R* (R DEVELOPMENT CORE TEAM, 2004) para a integração dos mapas dos genitores IAPAR-123 e IAPAR-06, utilizando marcas de AFLP com segregação do tipo 1:1 e 3:1 e microssatélites com segregação 1:1, 3:1, 1:2:1 e 1:1:1:1. Este é o primeiro relato do uso de marcadores microssatélites e desta metodologia em maracujá-amarelo, visando à integração dos mapas, anteriormente construídos a partir da estratégia duplo pseudocruzamento teste (CARNEIRO et al., 2002; LOPES et al., 2006).

Para promover a análise de ligação, é preciso observar às quais classes pertencem as marcas, de acordo com a Tabela 5. A combinação entre as marcas D_1 (informativas para o IAPAR-123) e D_2 (informativas para o IAPAR-06) não fornece informações para a distinção entre indivíduos recombinantes e parentais, de tal forma que a análise de ligação não pode ser feita. Contudo, esta ligação foi detectada indiretamente pelas marcas bi-parentais, pertencentes aos grupos A , B_3 e C (Tabela 5).

Quando se trabalha com populações de tamanho infinito, as estimativas de máxima verossimilhança são não viesadas. Segundo Maliepaard, Jansen e Van Ooijen (1997), a acurácia e o poder de estimação das frequências de recombinação menores do que 0,25 são boas quando se trabalha com progênies compostas por mais de 100 indivíduos, como é o caso do presente estudo. Além disso, a análise de ligação difere quanto à acurácia, para marcas com diferentes níveis de informação genética. Segundo Wu et al. (2002), o viés das estimativas de frequência de recombinação é elevado quando se trabalha com marcadores simétricos (aqueles que possuem a mesma configuração em ambos os genitores) como os do tipo B_3 e C . No entanto, estes problemas podem ser amenizados quando se trabalha com análises do tipo multiponto. Neste caso, as inferências podem ser feitas incluindo marcadores assimétricos.

Neste estudo, as análises de três pontos foram utilizadas para verificar a ordem fornecida pelo RCD, e fornecer melhores estimativas para as frequências de recombinação. Esta metodologia foi comprovada por Wu et al. (2002), isto é, de que melhores estimativas de frequência de recombinação e ordenação são obtidas comparativamente à análise de dois pontos. Estes autores ainda afirmam que esta melhoria é tanto maior quanto mais informativa for a marca utilizada (com exceção da análise de marcas simétricas dispostas consecutivamente no grupo de ligação).

Apesar dos problemas apresentados pelo uso de marcas dominantes com segregação 3:1, para fins de detecção de ligação, cerca de 87,5% delas puderam ser alocadas no *framework* do mapa integrado de maracujá-amarelo. O número médio de marcas 3:1 ao longo dos grupos de ligação foi de 5,6. Resultados semelhantes foram obtidos em mapas de uva (DOUCLEFF et al., 2004).

Quando se pretende promover a integração de mapas em espécies não endogâmicas, o ideal é utilizar marcadores microssatélites com segregação completamente (1:1:1:1) ou parcialmente informativa (1:2:1). Este tipo de marcador permite uma melhor inferência das fases e frequências de recombinação entre os diferentes tipos de segregações encontradas nessas espécies. Pugh et al. (2004) promoveram a integração de mapas de ligação em cacau, utilizando 465 marcadores codominantes (RFLP, microssatélites, isoenzimas e genes análogos de resistência - Rgenes-RFLP) e obtiveram 10 grupos de ligação. Lespinasse et al. (2000) empregaram 128 marcadores RFLP, 29 AFLP, 11 microssatélites e 4 isoenzimas como locos biparentais para promover a integração dos mapas de seringueira.

O mapa genético aqui construído sintetiza um conjunto de informações inicialmente obtidas por Carneiro et al. (2002) que estabeleceram o primeiro mapa de maracujá-amarelo, utilizando uma população segregante de 90 plantas e marcadores moleculares RAPD na estratégia duplo pseudocruzamento teste. Estas marcas de RAPD não foram integradas no presente mapa pelo de fato de haver poucas plantas (indivíduos F_1) em comum. Neste caso, o número de observações perdidas seria de mais de 50%. De acordo com Doerge (1996), para populações relativamente grandes, com cerca de 250 plantas, a performance do algoritmo RCD para estimação da ordenação dos grupos de ligação começa a ser afetada quando mais existem mais de 20% de observações perdidas. No presente caso foram usadas 160 plantas, portanto, mais de 12% de dados perdidos poderiam afetar a ordenação. As marcas de RAPD com segregação do

tipo 1:1 e 3:1 foram utilizadas visando a sua integração nos mapas, porém sem grande sucesso, pois muitas das marcas alocadas no *framework* do mapa integrado com marcadores AFLP e microssatélites foram retiradas e muitos grupos de ligação não puderam ser integrados (dados não apresentados).

Em seguida, Lopes (2003) utilizou 117 plantas dessa mesma população, marcadores AFLP e a mesma estratégia de mapeamento. Matta (2005) trabalhou com 160 plantas dessa mesma população, visando à integração dos mapas usando o programa JoinMap e marcas de AFLP com segregação do tipo 1:1 e 3:1. Contudo, esta integração não foi muito consistente, pois foram obtidos oito grupos de ligação, um a menos do que o esperado. Segundo o autor, a ordenação de algumas marcas não foi muito acurada. Tal fato pode ser devido à metodologia de mapeamento implementada no JoinMap, que utiliza apenas das frequências de recombinação entre pares de marcas para a estimação da melhor ordem.

O trabalho seguinte foi o de Moraes (2005) que utilizou o mesmo grupo de marcas e de plantas, porém fez o uso do programa TreeMap que utiliza os algoritmos propostos por Wu et al. (2002). A estratégia foi diferente, já que não foi feita a integração dos mapas, mas sim a alocação das marcas 3:1 como acessórias nos mapas construídos, estabelecendo, assim, uma homologia entre eles.

Neste estudo, foram introduzidas marcas de microssatélites, sendo obtidos 10 grupos de ligação (sendo 1 deles constituído por apenas duas marcas que provavelmente está ligado a algum dos outros nove). Também, foi possível alocar no *framework*, marcas de AFLP com segregação do tipo 3:1, que contribuíram para uma maior saturação e integração de oito dos nove grupos de ligação de maracujá-amarelo. Tudo isso, só foi possível devido ao acúmulo de dados para uma mesma população de mapeamento. Contudo, não houve integração do GL-IV por isso, o aumento do número de marcas informativas, principalmente de microssatélites, se faz necessário. O uso do algoritmo de estimação da fase de ligação e da frequência de recombinação simultaneamente, proposto por Wu et al. (2002) e implementado no programa *R* foi muito promissor na integração dos mapas de ligação de maracujá-amarelo, utilizando dados de marcadores microssatélites e AFLP.

A integração de mapas de ligação de espécies não endogâmicas com o uso de diferentes tipos de marcadores permite uma melhor caracterização do polimorfismo existente ao longo de todo o genoma. Além disso, os mapas dos genitores podem ser apresentados separadamente para

investigar a presença de QTL segregando para cada genitor. Todavia, se os QTL estão segregando para ambos os genitores e se pretende estudar as diferentes combinações alélicas dos QTL, é melhor utilizar o mapa integrado com todos os marcadores (MALIEPAARD et al., 1998).

Em maracujá-amarelo existe uma série de características de importância econômica que possuem herança complexa, envolvendo múltiplos genes, como a resistência a doenças, a produtividade e a qualidade de frutos. Todas estas características estão sendo estudadas no Departamento de Genética da ESALQ/USP e de posse do mapa único, integrado, as estimativas dos efeitos e localização dos QTL tendem a ser mais acuradas.

O *framework* dos mapas aqui apresentados será útil para alocação de novas marcas de microssatélites que serão progressivamente mapeadas na busca de integração completa dos grupos de ligação e de sua maior saturação, principalmente em regiões genômicas específicas de interesse aplicado. A alocação de novas marcas de microssatélites também contribuirá para o estabelecimento da correspondência entre mapas genéticos de diferentes cruzamentos, em estudos de mapeamento comparativo e evolução (sintenia).

5. CONCLUSÕES

A construção de bibliotecas genômicas enriquecidas é particularmente interessante em maracujá-amarelo, uma vez que não há seqüências de microssatélites depositadas no *GenBank* (<http://www.ncbi.nih.gov>). Com esta estratégia, 11,3% de eficiência de enriquecimento foi atingida, isolando 107 seqüências contendo microssatélites, sendo 55 perfeitos, 46 imperfeitos e 6 compostos;

É possível a construção de mapas de ligação com base em marcas dominantes e codominantes, a partir da análise de uma população F_1 segregante de maracujá-amarelo, assim como alocar marcadores dominantes bi-parentais, com segregação do tipo 3:1, no *framework* dos grupos de ligação, sem prejuízo da ordenação dos locos;

A utilização do algoritmo proposto por Wu et al. (2002), que estima as fases de ligação e a frequência de recombinação simultaneamente, é extremamente promissora para a integração dos mapas de ligação de populações não endogâmicas, como as de maracujá-amarelo, pois esta é uma espécie auto-incompatível;

As abordagens adotadas neste trabalho permitiram a obtenção de oito dos nove grupos de ligação de maracujá-amarelo ($2n = 18$), saturados e integrados;

Existe a necessidade de se desenvolver mais marcadores microssatélites para a geração de todos os grupos de ligação de maracujá-amarelo, com boa acurácia.

REFERÊNCIAS

- AGRAMA, H.A.; GEORGE, T.L.; SALAH, S.F. Construction of genome map for *Eucalyptus camaldulensis* Dehn. **Silvae Genética**, Frankfurt, v.51, p. 5-6, 2002.
- FNP CONSULTORIA & COMÉRCIO. Maracujá. In _____. **Agrianual 2006**: anuário da agricultura brasileira. São Paulo, 2006. p. 370-375.
- AKAGI, H.; YOKZEKI, Y.; INAGAKI, A.; FUJIMURA, T. Origin and evolution of twin microsatellites in the genus *Oryza*. **Heredity**, London, v.81, p.187-197, 1998.
- AKAMINE, E.K.; GIROLAMI, G. Pollination and fruit set in the yellow passion fruit. University of Hawaii, Honolulu. Agricultural Experiment Station. **Technical Bulletin**, Honolulu, v.39, 44 p. 1959.
- ALFENAS, P.F.; BRAZ, A.S.K.; TORRES, L.B.; SANTANA, E.N.; NASCIMENTO, A.V.S.; CARVALHO, M.G.; OTONI, W.C.; ZERBINI, M. Transgenic passionfruit expressing RNA derived from *Cowpea aphid-borne mosaic virus* is resistant to passionfruit woodiness disease. **Fitopatologia Brasileira**, Brasília, v.30, p.33-38, 2005.
- ALONSO-BLANCO, C.; PEETERS, A.J.M.; KOORNNEEF, M.; LISTER, C.; DEAN, C.; VAN DEN BOSCH, N.; POT, J.; KUIPER, M.T.R. Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. **Plant Journal**, Gainesville, v.14, p.259-271, 1998.
- ARANZANA, M.J.; GARCIA, M.A.J.; CARBO, J.; ARUS, P. Development and variability analysis of microsatellite markers in peach. **Plant Breeding**, Berlin, v.121, p.87-92, 2002.
- ARMOUR, J.A.; NEUMANN, R.; GOBERT, S.; JEFFREYS, A.J. Isolation of human simple repeat loci by hybridization selection. **Human Molecular Genetics**, Oxford, v.3, p.599-565, 1994.

ARUS, P.; OLARTE, C.; ROMERO, M.; VARGAS, F. Linkage analysis of 10 isozyme genes in F₁ segregating almond progenies. **Journal of the American Society for Horticultural Science**, Alexandria, v.119, p.339-344, 1994.

ASHKENAZI, V.; CHANI, E.; LAVI, U.; LEVY, D. HILLEL, J.; VEILLEUX, E. Development of microsatellite markers in potato and their use in phylogenetic and fingerprinting analyses. **Genome**, Ottawa, v.44, p.50-62, 2001.

ASHWORTH, V.E.T.M.; KOBAYASHI, M.C; de la CRUZ, M.; CLEGG, M.T. Microsatellite markers in avocado (*Persea americana* Mill.): development of dinucleotide and trinucleotide markers. **Scientia Horticulturae**, Nem York, v.101, p.255-267, 2004.

BASTEN, C.J.; WEIR, B.S.; ZENG, Z-B. **QTL Cartographer, Version 1.17**. Department of Statistics, North Carolina State University, Raleigh, NC. 2004.

BATESON, W., SAUNDERS, E.R.; PUNNETT, R.C. Experimental studies in the physiology of heredity, London. **Reports to the Evolution Committee of the Royal Society**, London, v.2, p.1-131, 1904.

BEARZOTI, E. Mapeamento de QTL. In: PINHEIRO, J.B.; CARNEIRO, I.F. (Ed.). **Análise de QTL no melhoramento de plantas**. Goiânia: FUNAPE, 2000. p.63-224.

BECKER, J.; VOS, P.; KUIPER, M.; SALAMINI, F.; HEUN, M. Combined mapping of AFLP and RFLP markers in barley. **Molecular Genetics and Genomics**, Berlin, v.259, p.65-73, 1995.

BENSON, G. Tandem repeats finder: a program to analyze DNA sequences. **Nucleic Acids Research**, London, v.27, p.573-580, 1999.

BERGER, J.; SUZUKI, T.; SENTI, K.A.; STUBBS, J.; SCHAFFNER, G.; DICKSON, B.J. Genetic mapping whit SNP markers in *Drosophila*. **Nature Genetics**, New York, v.29, p.475-481, 2001.

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

BOWERS, J.; BOURSQUOT, J.M.; THIS, P.; CHU, K.; JOHANSSON, H.; MEREDITH, C. Historical genetics: The parentage of Chardonnay, Gamay, and other wine grapes of northeastern France. **Science**, Washington, v.285, p.1562-1565, 1999.

BROAD, T.E.; HILL, D.F. Mapping the sheep genome, practice, progress and promise. **British Veterinary Journal**, London, v.150, n.3, p.237-252, 1994.

BRONDANI, R.P.V.; BRONDANI, C.; GRATTAPAGLIA, D. Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. **Molecular Genetics and Genomics**, Berlin, v.267, p.338-347, 2002.

BRONDANI, R.P.V.; BRONDANI, C.; TARCHINI, R.; GRATTAPAGLIA, D. Development, characterisation and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. **Theoretical and Applied Genetics**, Berlin, v.97, p.816-827, 1998.

BROWN, A.H.D.; GARVIN, D.F.; BURDON, J.J.; ABBOTT, D.C.; READ, B.J. The effect of combining scald resistance genes on disease levels, yield and quality traits in barley. **Theoretical and Applied Genetics**, Berlin, v.93, p.361-366, 1996.

BRUCKNER, C.H. Perspectives do melhoramento genético do maracujazeiro. In: SÃO JOSÉ, A.R.; BRUCKNER, C.H.; MANICA, I.; HOFMANN, M. (Ed.). **Maracujá: Temas selecionados - melhoramento, morte prematura, polinização, taxonomia**. Porto Alegre: Cinco Continentes, 1997. p.7-24.

BRUCKNER, C.H.; CASALI, V.W.D.; MORAES, C.F. de; REGAZZI, A.J.; SILVA, E.A.M. da. Self-incompatibility in passion fruit (*Passiflora edulis* Sims). **Acta Horticulturae**, Bélgica, v.370, p. 45-57, 1995.

BRUNEL, D. A microsatellite marker in *Helianthus annuus* L. **Plant Molecular Biology**, Oxford, v.24, p.397-400, 1994.

BRYAN, G.J.; COLLINS, A.J.; STEPHENSON, P.; ORRY, A.; SMITH, J.B.; GALE, M.D. Isolation and characterization of microsatellite from hexaploid bread wheat. **Theoretical and Applied Genetics**, Berlin, v.94, p.557-563, 1997.

BURKE, J.M.; LAI, Z.; SALMASO, M.; NAKAZATO, T.; TANG, S.; HEESACKER, A.; KNAPP, S.J.; RIESEBERG, L.H. Comparative mapping and rapid karyotypic evolution in the genus *Helianthus*. **Genetics**, Bethesda, v.167, p.449-457, 2004.

BURR, B.; BURR, F.A.; THOMPSON, K.H.; ALBERTSON, M.; STUBER, C.W. Gene mapping with recombinant inbreds in maize. **Genetics**, Bethesda, v.118, p. 519-526, 1988.

BUTCHER, P.A.; MORAN, G.F. Genetic linkage mapping in *Acacia mangium*. 2. Development of an integrated map from two outbred pedigrees using RFLP and microsatellite loci. **Theoretical and Applied Genetics**, Berlin, v.101, p.594-605, 2000.

BUTCHER, P.A.; DECROOCQ, S.; GRAY, Y.; MORAN, G.F. Development, inheritance and cross-species amplification of microsatellite markers from *Acacia mangium*. **Theoretical and Applied Genetics**, Berlin, v.101, p.1282-1290, 2000.

BUTCHER, P.A.; WILLIAMS, E.R.; WHITAKER, D.; LING, S.; SPEED, T.P.; MORAN, C.F. Improving linkage analysis in outcrossed forest trees - an example from *Acacia mangium*. **Theoretical and Applied Genetics**, Berlin, v.104, p.1185-1191, 2002.

BUTLER, M.I.; JARRET, R.L.; LABONTE, D.R. Sequence characterization of microsatellites in diploid and polyploid *Ipomoea*. **Theoretical and Applied Genetics**, Berlin, v.99, p.123-132, 1999.

CAMILLO, E. Polinização do maracujazeiro. In: RUGGIERO, C. (Ed.). **Maracujá**. Ribeirão Preto: Legis Summa, 1987. p. 97-103.

CARNEIRO, M.S.; VIEIRA, M.L.C. Mapas genéticos em plantas. **Bragantia**, Campinas, v.61, p.89-100, 2002.

CARNEIRO, M.S.; CAMARGO, L.E.A.; COELHO, A.S.G.; VENCOVSKY, R.; LEITE JÚNIOR, 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.

CARRARO, A.F.; CUNHA, M.M. **Manual de exportação de frutas**. Brasília: MAARA-SDR - FRUPEX/IICA, 1994, 254p.

CAUSSE, M.A.; FULTON, T.M.; CHO, Y.G.; AHN, S.N.; CHUNWONGSE, J.; WU, K.; XIAO, J.; YU, Z.; RONALD, P.C.; HARRINGTON, S.E.; SECOND, G.; McCOUCH, S.R.; TANKSLEY, S. Saturated molecular map of the rice genome based on an interespecific backcross population. **Genetics**, Bethesda, v. 138, p.1251-1274, 1994.

CERVERA, M.T.; STORME, V.; IVENS, B.; GUSMÃO, J.; LIU, B.H.; HOSTYN, V.; SLYCKEN, J.V.; MONTAGU, M.V.; BOERJAN, W. Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. **Genetics**, Bethesda, v.158, p.787-809, 2001.

CHEN, X.; TEMNYKH, S.; XU, Y.; CHO, Y.G.; McCOUCH, S.R. Development of a microsatellite *framework* map providing genome-wide coverage in rice (*Oryza sativa* L.). **Theoretical and Applied Genetics**, Berlin, v.95, p.553-567, 1997.

CIFARELLI, R.A.; GALLITELLI, M.; CELLINI, F. Random amplified hybridization microsatellites (RAHM): isolation of a new class of microsatellite-containing DNA clones. **Nucleic Acids Research**, London, v.23, p.3802-3803, 1995.

CLOUTIER, S., CAPPADOCIA, M.; LANDRY, B.S. Analysis of RFLP mapping inaccuracy in *Brassica napus* L. **Theoretical and Applied Genetics**, Berlin, v.95, p.83-91, 1997.

COE, H.E.; NEUFFER, M.G.; HOISINGTON, D.A. The genetics of corn. In: SPRAGUE, G.F.; DUDLEY, J.W. (Ed.). **Corn and corn improvement**. 3rd ed. Madison: American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, 1988. p.81-237.

COELHO, A.S.G. **TreeMap - Software para construção de mapas genéticos em cruzamentos derivados de genótipos não-endogâmicos**. Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Goiás. Goiânia, 2005.

COLLEVATTI, R.G.; BRONDANI, R.V.; GRATTAPAGLIA, D. Development and characterisation of microsatellite markers for genetic analysis of a Brazilian endangered tree species *Caryocar brasiliense*. **Heredity**, London, v.83, p.748-756, 1999.

CONNER, P.J.; BROWN, S.K.; WEEDEN, N.F. Molecular-marker analysis of quantitative traits for growth and development in juvenile apple trees. **Theoretical and Applied Genetics**, Berlin, v.96, p.1027-1035, 1998.

CREGAN, P.B.; BHAGWAT, A.A.; AKKAYA, M.S.; RONGWEN, J. Microsatellite fingerprinting and mapping of soybean. **Methods in Molecular and Cellular Biology**, New York, v.5, p.49-61, 1994.

CREGAN, P.B.; MUDGE, J.; FICKUS, E.W.; MAREK, L.F.; DANESH, D.; DENNY, R.; SHOEMAKER, R.C.; MATTHEWS, B.F.; JARVIK, T.; YOUNG, N.D. Targeted isolation of simple sequence repeat markers through the use of bacterial artificial chromosomes. **Theoretical and Applied Genetics**, Berlin, v.98, p.919-928, 1999.

CRESTE, S.; TULMANN NETO, A.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reporter**, New York, v.19, p.299-306, 2001.

CROUZILLAT, D.; LERCETEAU, E.; PETIARD, V.; MORERA, J.; RODRIGUEZ, H.; WALKER, D.; PHILLIPS, W.; RONNING, C.; SCHNELL, R.; OSEI, J.; FRITZ, P. *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. **Theoretical and Applied Genetics**, Berlin, v.93, p.205-214, 1996.

CUCO, S.M.; VIEIRA, M.L.C.; MONDIN, M.; AGUIAR-PERECIN, M.L.R. Comparative karyotype analysis of three *Passiflora* L. species and cytogenetic characterization of somatic hybrids. **Caryologia**, Florence, v.58, p.220-228, 2005.

D'EECKENBRUGGE, G.C. Exploração da diversidade genética das passifloras (1 CD-ROM). In: SIMPÓSIO BRASILEIRO SOBRE A CULTURA DO MARACUJAZEIRO, 6., Campos dos Goytacazes, 2003. **Palestras**. Campos dos Goytacazes: Cluster Informática, 2003.

DALBÓ, M.A.; YE, G.N.; WEEDEN, N.F.; STEINKELLNER, H.; SEFC, K.M.; REISCH, B.I. A gene controlling sex in grapevines placed on a molecular marker-based genetic map. **Genome**, Ottawa, v.43, p.333-340, 2000.

DEMPSTER, A.P.; LAIRD, N.M.; RUBIN, D.B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society, Series B**, London, v.39, p.1-38, 1977.

DEVEY, M.E.; SEWELL, M.M.; UREN, T.L.; NEALE, D.B. Comparative mapping in loblolly and radiata pine using RFLP and microsatellite markers. **Theoretical and Applied Genetics**, Berlin, v.99, p.656-662, 1999

DIERS, B.W.; MANSUR, L.; IMSANDE, J.; SHOEMAKER, R.C. Mapping *Phytophthora* resistance loci in soybean with restriction fragment length polymorphism markers. **Crop Science**, Madison, v. 32, p. 377-383, 1992.

DIRLEWANGER, E.; MOING, A.; ROTHAN, C.; SVANELLA, L.; PRONIER, V.; GUYE, A.; PLOMION, C.; MONET, R. Mapping QTL controlling fruit quality in peach (*Prunus persica* (L.) Batsch). **Theoretical and Applied Genetics**, Berlin, v.98, p.18-31, 1999.

DOERGE, R.W. Constructing genetic maps by rapid chain delineation. **Journal of Agricultural Genomics**, v.2, article 6, 1996. Disponível em: <<http://probe.nalusda.gov:8000/otherdocs/jqtl>>. Acesso em: 12 mar.2005.

DOERGE, R.W.; ZENG, Z.B.; WEIR, B.S. Statistical issues in the search for genes affecting quantitative traits in experimental populations. **Statistical Science**, Hayward, v.12, n.3, p. 195-219, 1997.

DOUCLEFF, M; JIN, Y.; GAO, F.; RIAZ, S.; KRIVANEK, A.F.; WALKER, M.A. A genetic linkage map of grape utilizing *Vitis rupestris* x *Vitis arizonica*. **Theoretical and Applied Genetics**, Berlin, v.109, p.1178-1187, 2004.

DREISIGACKER, S.; ZHANG, P.; WARBURTON, M.L.; van GINKEL, M.; HOISINGTON, D.; BOHN, M.; MELCHINGER, A.E. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different megaenvironments. **Crop Science**, Madison, v.44, p.381-388, 2004.

ECHE, C.S.; MAY-MARQUARDT, P. Survey of microsatellite DNA in pine. **Genome**, Ottawa, v.40, p.9-17, 1997.

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

ELLEGREN, H.; PRIMMER, C.R.; SHELDON, B.C. Microsatellite 'evolution': directionality or bias? **Nature Genetics**, New York, v.11, p.360-362, 1995.

FANIZZA, G.; LAMAJ, F.; COSTANTINI, L.; CHAABANE, R.; GRANDO, M.S. QTL analysis for fruit yield components in table grapes (*Vitis vinifera*) **Theoretical and Applied Genetics**, Berlin, v.111, p.658-664, 2005.

FERGUSON, A. Genetic differences among brown trout (*Salmo trutta*) stocks and their importance for the conservation and management of the species. **Freshwater Biology**, Oxford, v.21, p.35-46, 1989.

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

FERREIRA, M.E.; WILLIMAS, P.H.; OSBORN, T.C. RFLP mapping of *Brassia napus* using doubled-haploid. **Theoretical and Applied Genetics**, Berlin, v. 89, p.615-621, 1994.

FLAMENT, M.H.; KEBE, I.; CLEMENT, D.; PIERETTI, I.; RISTERUCCI, A.M.; N'GORAN, J.A.K.; CILAS, C.; DESPRAUX, D.; LANAUD, C. Genetic mapping of resistance factors to *Phytophthora palmivora* in cocoa. **Genome**, Ottawa, v.44, p.79-85, 2001.

FOULONGNE, M.; PASCAL, T.; PFEIFFER, F.; KERVELLA, J. QTLs for powdery mildew resistance in peach × *Prunus davidiana* crosses: consistency across generations and environments. **Molecular Breeding**, New York, v.12, p.33-50, 2003

FUNGARO, M.H.P.; VIEIRA, M.L.C. Marcadores Moleculares. In: SERAFINI, L.A.; BARROS, N.M.; AZEVEDO, J.L.A. (Org.). **Biotecnologia na Agricultura e Agroindústria**. Piracicaba: Guaíba, 2001. v.1, p. 153-200.

GAIOTTO, F.A.; BRONDANI, R.P.V.; GRATTAPAGLIA, D. Microsatellite markers for heart of palm - *Euterpe edulis* and *E. oleraceae* Mart. (Arecaceae). **Molecular Ecology Notes**, Oxford, v.1, p.86-88, 2001.

GARCIA, A.A.F.; KIDO, E.A.; MEZA, A.N.; SOUZA, H.M.B.; PINTO, L.R.; PASTINA, M.M.; LEITE, C.S.; DA SILVA, J.A.G.; ULIAN, E.C.; FIGUEIRA, A.; SOUZA, A.P. Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. **Theoretical and Applied Genetics**, Berlin, v.112, p.298-314, 2006.

GARCIA, M.R.; ASINS, M.J.; CARBONELL, E.A. QTL analysis of yield and seed number in *Citrus*. **Theoretical and Applied Genetics**, Berlin, v.101, p.487-493, 2000.

GIANFRANCESCHI, L.; SEGLIAS, N.; TARCHINI, R. KOMJANC, M.; GESSLER, C. Simple sequence repeats for the genetic analysis of apple. **Theoretical and Applied Genetics**, Berlin, v.96, p.1069-1076, 1998.

GRATTAPAGLIA, D.; SEDEROFF, R. Genetic linkage maps of *Eucalyptus grandis* e *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. **Genetics**, Bethesda, v.137, p.1121-37, 1994.

GRATTAPAGLIA, D.; WILCOX, P.; CHAPARRO, J.X.; O'MALLESY, D.; McCORD, S. A RAPD map of loblolly pine in 60 days. In: INTERNATIONAL CONGRESS OF THE INTERNATIONAL SOCIETY FOR PLANT MOLECULAR BIOLOGY, 3., 1991, Tucson. **Abstract**. Tucson: 1991, p.2224.

GREEN, P.; FALLS, K.; CROOKS, S.: **Documentation for CRIMAP, version 2.4**. Washington: Univ. School of Medicine, St. Louis, MO. 1990.

GRISI JR., C. Método de polinização artificial do maracujazeiro, *Passiflora edulis*. In: CONGRESSO BRASILEIRO DE FRUTICULTURA, 2., 1973, Viçosa. **Anais**. Viçosa: SBF, 1973. v.2, p. 433-436.

GUT, I.G. Automation in genotyping of single nucleotide polymorphism. **Human Mutation**, New York, v.17, p.475-492, 2001.

HALDANE, J.B.S. The combination of linkage values, and the calculation of distances between the loci of linked factors. **Journal of Genetics**, London, v. 8, p. 299-309, 1919.

HAMILTON, M.B.; FLEISCHER, R.C. Cloned microsatellite repeats differ between 4-base restriction endonucleases. **Journal of Heredity**, London, v.90, p.561-563, 1999.

HAMWIEH, A.; UDUPA, S.M.; CHOUMANE, W.; SARKER, A.; DREYER, F.; JUNG, C.; BAUM, M. A genetic linkage map of *Lens* sp. based on microsatellite and AFLP markers and the localization of fusarium vascular wilt resistance. **Theoretical and Applied Genetics**, Berlin, v.110, p.669-677, 2005.

HARR, B.; ZANGER, B.; SCHLÖTTERER, C. Removal of microsatellite interruptions by DNA replication slippage: Phylogenetic evidence from *Drosophila*. **Molecular Biology Evolution**, New York, v.17, p.1001-1009, 2000.

HAUGE, X.Y.; LITT, M. A study of the origin of shadow bands seen when typing dinucleotide repeat polymorphisms by the PCR. **Human Molecular Genetics**, Oxford, v.2, p.411-415, 1993.

HE, G.; MENG, R.; NEWMAN, M.; GAO, G.; PITTMAN, R.N.; PRAKASH, C.S. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). **BMC Plant Biology**, London, v.3, p.1-6, 2003.

HEMMAT, M.; WEEDEN, N.F.; MANGANARIS, A.G.; LAWSON, D.M. Molecular marker linkage map for apple. **Journal of Heredity**, London, v.85, p.4-11, 1994.

HOSPITAL, F.; CHEVALET, C.; MULSANT, P. Using markers in gene introgression breeding programs. **Genetics**, Bethesda, v.132, n.4, p.1199-1210, 1992.

JONES, E.S.; DUPAL, M.P.; KÖLLIKER, R.; DRAYTON, M.C.; FORSTER, J.W. Development and characterisations of simple sequence repeat (SSR) markers for perennial ryegrass (*Lolium perenne* L.), **Theoretical and Applied Genetics**, Berlin, v.102, p.405-415, 2001.

JULIER, B.; FLAJOULOT, S.; BARRE, P.; CARDINET, G.; SANTONI, S.; HUGUET, T.; HUYGHE, C. Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. **BMC Plant Biology**, London, v.3, p.1-19, 2003.

JÚNIOR, F.L.C.; ESTALISLAU, M.L.L.; PAIVA, B.M. Aspectos econômicos da cultura do maracujá. **Informe Agropecuário**, Belo Horizonte, v.21, p.10-17, 2000.

KALO, P.; ENDRE, G.; ZIMANYI, L.; CSANADI, G.; KISS, G.B. Construction of an improved linkage map of diploid alfalfa (*Medicago sativa*). **Theoretical and Applied Genetics**, Berlin, v.100, p. 641-657, 2000.

KARAGYOZOV, L.; KALCHEVA, I.D.; CHAPMAN, V.M. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. **Nucleic Acids Research**, London, v.21, p.3911-3912, 1993.

KEIM, P.; DIERS, B.W.; OLSON, T.C.; SHOEMAKER, R.C. RFLP mapping in soybean: association between marker loci and variation in quantitative traits. **Genetics**, Bethesda, v.126, p.735-742, 1990.

KEIM, P.; SCHUPP, J.M.; TRAVIS, S.E.; CLAYTON, K.; ZHU, T.; SHI, L.; FERREIRA, A.; WEBB, D.M. A high-density soybean genetic map based on AFLP markers. **Crop Science**, Madison, v.37, p.537-543, 1997.

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

KENIS, K.; KEULEMANS, J. Genetic linkage maps of two apple cultivars (*Malus domestica* Borkh.) based on AFLP and microsatellite markers. **Molecular Breeding**, Nova York, v.15, p.205-219, 2005.

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 streptavidin-coated magnetic particles. **Biotechniques**, Natick, v.16, p.656-662, 1994.

KLEIN, P.E.; KLEIN, R.R.; CARTINHO, S.W.; ULANCH, P.E.; DONG, J.; OBERT, C.A.; MORISHIGE, D.T.; SCHLUETER, S.D.; CHILDS, K.L.; ALE, M.; MULLET, J.E. A High-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. **Genome Research**, New York, v.6, p.789-807, 2000.

KNAPIK, E.W.; GOODMAN, A.; EKKER, M.; CHEVRETTE, M.; DELGADO, J.; NEUHAUSS, S.; SHIMODA, N.; DRIEVER, W.; FISHMAN, M.C.; JACOB, H.J. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). **Nature Genetics**, New York, v.18, p.338-343, 1998.

KÖLLIKER, R.; JONES, E.S.; DRAYTON, M.C.; DUPAL, M.P.; FOSTER, J.W. Development and characterisation of simple sequence repeat (SSR) markers for white clover (*Trifolium repens* L.). **Theoretical and Applied Genetics**, Berlin, v.102, p.416-424, 2001.

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

KUANG, H.; RICHARDSON, T.; CARSON, S.; WILCOX, P.; BONGARTEN, B. Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. I. Genetic map with distorted markers. **Theoretical and Applied Genetics**, Berlin, v.98, p.697-703, 1999.

KWOK, P. Methods for genotyping single nucleotide polymorphism. **Annual Reviews of Genomics and Human Genetics**, Palo Alto, v.2, p.235-258, 2001.

LANDER, E. S.; GREEN, P. Construction of multilocus genetic linkage maps in human. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.84, p.2363-2367, 1987.

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

LANDRY, B.; HUBERT, N.; ETOH, T.; HARADA, J.J.; LINCOLN, S.E. A genetic map for *Brassica oleraceae* based on restriction fragment length polymorphism detected with expressed DNA sequences. **Genome**, Ottawa, v.34, p.543-552, 1991.

LANZA, M.A.; GUIMARÃES, C.T.; SCHUSTER, I. Aplicação de marcadores moleculares no melhoramento genético. **Informe Agropecuário**, Belo Horizonte, v.21, p. 97-108, 2000.

LESPINASSE, D.; RODIER, G.M.; GRIVET, L.; LECONTE, A.; LEGNATE, H.; SEGUIN, M. A saturated genetic linkage map of rubber tree (*Hevea* spp) based on RFLP, AFLP, microsatellite and isozymes markers. **Theoretical and Applied Genetics**, Berlin, v.100, p.127-138, 2000.

LIBERATO, J.R. Controle das doenças causadas por fungos, bactérias e nematóides em maracujá. In: ZAMBOLIM, L.; VALE, F.X.R.; MONTEIRO, A.J.A.; COSTA, H. (Ed.). **Controle de doenças de plantas: fruteiras**. Viçosa: UFV, 2002, v.2, p. 699-826.

LITT, M.; LUTY, L.A. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. **American Journal of Human Genetics**, Chicago, v.44, p.398-401, 1989.

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

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

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

LU, Z.X.; SOSINSKI, B.; REIGHARD, G.L.; BAIRD, W.V.; ABBOTT, A.G. Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks. **Genome**, Ottawa, v.41, p.199-207, 1998.

LUNT, D.H.; HUTCHINSON, W.F.; CARVALHO, G.R. An efficient method for PCR-based identification of microsatellite arrays (PIMA). **Molecular Ecology**, Oxford, v.8, p.893-894, 1999.

MA, Z.Q.; Roder, M.; SORRELLS, M.E. Frequencies and sequence characteristics of di-, tri- and tetra-nucleotide microsatellites in wheat. **Genome**, Ottawa, v.39, p.123-130, 1996.

MAHALAKSHMI, V.; APARNA, P.; RAMADEVI, S.; ORTIZ, R. Genomic sequence derived simple sequence repeats markers: a case study with *Medicago* spp. **Electronic Journal of Biotechnology**, Valparaiso, v.5, p.233-242, 2002.

MAHESWARAN, M.; SUBUDHI, P.K.; NANDI, S.; XU, J.C.; PARCO, A.; YANG, D.C.; HUANG, N. Polymorphism, distribution, and segregation of AFLP markers in a doubled-haploid rice population. **Theoretical and Applied Genetics**, Berlin, v.94, p.39-45, 1997.

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.

MALIEPAARD, C.; ALSTON, F.H.; VAN ARKEL, G.; BROWN, L.M.; CHEVREAU, E.; DUNEMANN, F.; EVANS, K.M.; GARDINER, S.; GUILFORD, P.; VAN HEUSDEN, A.W.; JANSE, J.; LAURENS, F.; LYNN, J.R.; MANGANARIS, A.G.; DEN NIJS, A.P.M.; PERIAM, N.; RIKKERINK, E.; ROCHE, P.; RYDER, C.; SANSAVINI, S.; SCHMIDT, H.; TARTARINI, S.; VERHAEGH, J.J.; VRIELINK-VAN GINKEL, M.; KING, G.J. Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. **Theoretical and Applied Genetics**, Berlin, v.97, p.60-73, 1998.

MALUF, W.R.; SILVA, J. A. R.; GRATTAPAGLIA, D.; TOMA-BRAGHINI, M.; CORTE, R. D.; MACHADO, M.; CALDAS, L.S. Genetic gains via clonal selection in passion fruit *Passiflora edulis* Sims. **Revista Brasileira de Genética**, Ribeirão Preto, v. 12, p. 833-841, 1989.

MANICA, I. Maracujá: Taxonomia - anatomia - morfologia: In. SÃO JOSÉ, A.L.; BRUCKNER, C.H.; MANICA, I.; HOFMANN, M. (Ed.). **Maracujá: Temas selecionados - Melhoramento, morte prematura, polinização, taxonomia**. Porto Alegre: Cinco Continentes, 1997. p.7-24.

MARQUES, C.M.; ARAÚJO, J.A.; FERREIRA, J.G.; WHETTEN R.; O'MALLEY, D.M.; LIU, B-H.; SEDEROFF, R. AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. **Theoretical and Applied Genetics**, Berlin, v.96, p.727-737, 1998.

MARTIN, E.R.; LAI, E.H.; GILBERT, J.R.; ROGALA, A.R.; AFSHARI, A.J.; RILEY, L.; FINCH, K.L.; STEVENS, F.; LIVAK, K.J.; SLOTTERBECK, B.D.; SLIFER, S.H.; WARREN, L.L.; CONNEALLY, P.M.; SCHMECHEL, D.E.; PURVIS, I.; PERICAK-VANCE, M.A.; ROSES, A.D.; VANCE, J.M. SNPing away at complex diseases: Analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease. **American Journal of Human Genetics**, Chicago, v.67, n.2, p.383-394, 2000.

MASONA, S.L.; STEVENSA, M.R.; JELLENA, E.N.; BONIFACIOB, A.; FAIRBANKSA, D.J.; COLEMANA, C.E.; MCCARTYA, R.R.; RASMUSSENA, A.G.; MAUGHANA, P.J. Development and use of microsatellite markers for germplasm characterization in quinoa (*Chenopodium quinoa* Willd.). **Crop Science**, Madison, v.45, p.1618-1630, 2005.

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

MATTA, F.P. **Mapeamento de QRL para *Xanthomonas axonopodis* pv. *passiflorae* em maracujá-amarelo (*Passiflora edulis* Sims f. *flavicarpa* Deg.)**. 2005. 231 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luis de Queiroz", Universidade de São Paulo, Piracicaba, 2005.

McCOUCH, S.R.; CHEN, X.; PANAUD, O.; TEMNYKH, S.; XU, Y.; CHO, Y.G.; HUANG, N.; ISHII, T.; BLAIR, M. Microsatellite marker development, mapping and applications in rice genetics and breeding. **Plant Molecular Biology**, Oxford, v.35, p.89-99, 1997.

McCOUCH, R.; TEYTELMAN, L.; XU, Y.; LOBOS, K.B.; CLARE, K.; WALTON, M.; FU, B.; MAGHIRANG, R.; LI, Z.; XING, Y.; ZHANG, Q.; KONO, I.; YANO, M.; FJELLSTROM, R.; DECLERCK, G.; SCHNEIDER, D.; CARTINHO, S.; WARE, D.; STEIN, L. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). **DNA Research**, Oxford, v.9, p.199-207, 2002.

MELETTI, L.M.M. **Caracterização agronômica de progênies de maracujazeiro-amarelo (*Passiflora edulis* Sims f. *flavicarpa* Deg.)**. 1998. 92 p. Tese (Doutorado em Fitotecnia) - Escola Superior de Agricultura “Luis de Queiroz”, Universidade de São Paulo, Piracicaba, 1998.

MELETTI, L.M.M. Comportamento de híbridos e seleções de maracujazeiro (*passifloraceae*) (1 CD-ROM). In: SIMPÓSIO BRASILEIRO SOBRE A CULTURA DO MARACUJAZEIRO, 6., Campos dos Goytacazes, 2003. **Palestras**. Campos dos Goytacazes: Cluster Informática, 2003.

MELETTI, L.M.M.; MAIA, M.L. **Maracujá: produção e comercialização**. Campinas: Instituto Agrônomo, 1999. 62 p. (Boletim Técnico, 181).

MELETTI, L.M.M.; SANTOS, R.R.; MINAMI, K. Melhoramento do maracujazeiro-amarelo: Obtenção do ‘Composto IAC-27’. **Scientia Agricola**, Piracicaba, v. 56, p. 491-498, 2000.

MELO, N.F.; CERVI, A.C.; GUERRA, M. Karyology and cytotaxonomy of the genus *Passiflora* L. (*Passifloraceae*). **Plant Systematics and Evolution**, New York, v.226, p.69-84, 2001.

MORAES, M.C. **Mapas de ligação e mapeamento de QTL (“Quantitative Trait Loci”) em maracujá-amarelo (*Passiflora edulis* Sims f. *flavicarpa* Deg.)**. 2005. 141p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2005.

MORAES, M.C.; GERALDI, I.O.; MATTA, F.P.; VIEIRA, M.L.C. Genetic and phenotypic parameter estimates for yield and fruit quality traits from a single wide cross in yellow passion fruit. **Hortscience**, Alexandria, v.40, p.1978-1981, 2005.

MORGAN, T.H.; STURTEVANT, A.H.; MULLER, H.J.; BRIDGES, C.B. **The Mechanism of Mendelian Heredity**. New York: Henry Holt and Company, 1915, 262p.

MORGANTE, M.; OLIVIERI, A.M. PCR-amplified microsatellites as markers in plant genetics. **The Plant Journal**, Gainesville, v.3, p.175-182, 1993.

MUDGE, J.; CREGAN, P.B.; KENWORTHY, J.P.; KENWORTHY, W.J.; ORF, J.H.; YOUNG, N.D. Two microsatellite markers that flank the major soybean cyst nematode resistance locus. **Crop Science**, Madison, v.37, p.1611-1615, 1997.

MULLIS, K.; FALOONA, F. Specific synthesis of DNA *in vitro* via polymerase-catalyzed chain reaction. **Methods in Enzymology**, New York, v.115, p. 335-350, 1987.

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

NAKASONE, H.Y.; PAULL, R.E. **Tropical fruits**. New York: CAB International, 1998. 445p. (Crop Production Science in Horticulture Series, 7).

NASCIMENTO, W.M.O.; TOMÉ, A.T.; OLIVEIRA, M.S.P.. MULLER, C.H., CARVALHO, J.E.U. Seleção de progênies de maracujazeiro-amarelo (*Passiflora edulis* f. *flavicarpa*) quanto à qualidade de frutos. **Revista Brasileira de Fruticultura**, Jaboticabal, v.25, p.186-188, 2003.

NIRENBERG, M.; MATTHAEI, J.H. Dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.47, p.1588-1602, 1961.

OLIVEIRA, A.M.A. **Reprodução e citogenética de espécies de *Passiflora***. 1996. 145 p. Tese (Doutorado em Genética) - Universidade Estadual Paulista “Júlio de Mesquita”, São José do Rio Preto, 1996.

OLIVEIRA, E.J.; PÁDUA, J.G.; ZUCCHI, M.I.; CAMARGO, L.E.A.; FUNGARO, M.H.P.; VIEIRA, M.L.C. Development and characterization of microsatellite markers from the yellow passion fruit (*Passiflora edulis* f. *flavicarpa*). **Molecular Ecology Notes**, Oxford, v.5, p.331-333, 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, 2006. *In press*.

OLSON, M.; HOOD, L.; CANTOR, C.; BOTSTEIN, D. A common language for physical mapping of the human genome. **Science**, Washington, v.245, p.1434-1435, 1989.

OSTRANDER, E.A.; JONG, P.M.; RINE, J.; DUYK, G. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.89, p.3419-3423, 1992.

PAETKAU, D. Microsatellites obtained using strand extension: An enrichment protocol. **Biotechniques**, Natick, v.26, p.690-697, 1999.

PANIEGO, N.; ECHAIDE, M.; MUÑOZ, M.; FERNÁNDEZ, L.; TORALES, S.; FACCIO, P.; FUXAN, I.; CARRERA, M.; ZANDOMENI, R.; SUÁREZ, E.Y.; HOPP, H.E. Microsatellite isolation and characterization in sunflower (*Helianthus annuus* L.). **Genome**, Ottawa, v.45, p.34-43, 2002.

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

PONCE, M.R.; ROBLES, P.; MICOL, J.L. High throughput genetic mapping in *Arabidopsis thaliana*. **Molecular General Genetics**, Berlin, v.261, p.408-415, 1999.

POWELL, W.; MACHRAY, G.C.; PROVAN, J. Polymorphism revealed by simple sequence repeats. **Trends in Plant Science**, Oxford, v.1, p.209-245, 1996.

PRIMMER, C.R.; RAUDSEPP, T.; CHOWDHARY, B.P.; MOLLER, A.R.; ELLEGREN, H. Low frequency of microsatellites in the avian genome. **Genome Research**, New York, v.7, p.471-482, 1997.

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

PROVINCE, M.A. Sequential methods of analysis for genome scan. In: RAO, D.C.; PROVINCE, M.A. **Dissection of complex traits**. San Diego: Academic Press, 1999. 583 p.

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.

QI, X., STAM, P.; LINDHOUT, P. Use of locus-specific AFLP markers to construct a high-density molecular map in barley. **Theoretical and Applied Genetics**, Berlin, v.96, p.376-384, 1998.

R DEVELOPMENT CORE TEAM: R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, 2004. Disponível em: <<http://www.r-project.org>>. Acesso em: 12 dez. 2004.

RAFALSKI, J.A.; VOGEL, J.M.; MORGANTE, M.; POWEL, W. ANDRE, C.; TINGEY, S. Generating and using DNA markers in plant. In: BIRREN, B.; LAI, E. (Ed.). **Non-mammalian genomic analysis: a practical guide**. San Diego: Academic Press, 1996. p. 75-135.

RAHMAN, M.H.; JAQUISH, B.; KHASA, P.D. Optimization of PCR protocol in microsatellite analysis with silver and SYBR[®] stains. **Plant Molecular Biology Reporter**, New York, v.18, p.339-348, 2000.

RALLO, P.; DORADO, G.; MARTÍN, A. Development of simple sequence repeats (SSRs) in olive tree (*Olea europaea* L.). **Theoretical and Applied Genetics**, Berlin, v.101, p.984-989, 2000.

RASSMANN, K.; SCHLÖTTERER, C.; TAUTZ, D. Isolation of simple sequence loci for use in polymerase chain reaction-based DNA fingerprinting. **Electrophoresis**, Weinheim, v.12, p.113-118, 1991.

RÊGO, M.M. **Genética, interação pólen-pistilo e expressão de proteínas na auto-incompatibilidade do maracujazeiro (*Passiflora edulis* Sims.)**. 1997. 66 p. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa, 1997.

REINISCH, A.J.; DONG, J.; BRUBAKER, C.L.; STELLY, D.M.; WENDEL, J.F.; PARESON, A.H. A detailed RFLP map of cotton, *Gossypium hirsutum* x *Gossypium barbadense*: Chromosome organization and evolution in a disomic polyploid genome. **Genetics**, Bethesda, v.138, p.829-847, 1994.

REITTER, R.S.; WILLIAMS, J.G.K.; FELDMANN, K.A.; RAFALSKI, J.A.; TINGEY, S.V.; SCOLNIK, P.A. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. **Proceedings of the National Academic of Sciences of the United States of America**, Washington, v.89, p.1477-1481, 1992.

RICK, C.M.; YODER, J.I. Classical and molecular genetics of tomato highlights and perspectives. **Annual Review of Genetics**, Palo Alto, v.22, p.281-300, 1988.

RIDOUT, M.S.; TONG, S.; VOWDEN, C.J.; TOBUTT, K.R. Three-point linkage analysis in crosses of allogamous plant species. **Genetical Research**, London, v.72, p.111-112, 1998.

RISTERUCCI, A.M.; GRIVET, L.; N'GORAN, J.A.K.; PIERETTI, I.; FLAMENT, M.H.; LANAUD, C. A high-density linkage map of *Theobroma cacao* L. **Theoretical and Applied Genetics**, Berlin, v.101, p.948-955, 2000.

RISTERUCCI, A.M.; PAULIN, D.; N'GORAN, J.A.K.; LANAUD, C. Identification of QTL related to cocoa resistances to three species of *Phytophthora*. **Theoretical and Applied Genetics**, Berlin, v.108, p.168-174, 2003.

RITSCHER, P.S.; LINS, T.C.L.; TRISTAN, R.L.; BUSO, G.S.C.; BUSO, J.A.; FERREIRA, M.E. Development of microsatellite markers from an enriched genomic library for genetic analysis of melon (*Cucumis melo* L.). **BMC Plant Biology**, London, v.4, p.1-14, 2004.

RITTER, E.; SALAMINI, F. The calculation of recombination frequencies in crosses of allogamous plant species with applications to linkage mapping. **Genetics Research**, New York, v.67, p.55-65, 1996.

RITTER, E.; GEBHARDT, C.; SALAMINI, F. Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. **Genetics**, Bethesda, v.125, p.645-654, 1990.

RÖDER, M.S.; KORZUN, V.; WENDHAKKE, K.; PLASCHKE, J.; TIXER, M.H.; LEROY, P.; GRANAL, M.W. A microsatellite map of wheat. **Genetics**, Bethesda, v.149, p.2007-2023, 1998.

RÖDER, M.S.; PLASCHKE, J.; KONIG, S.V.; BORNER, A.; SORRELS, M.E.; TANKSLEY, S.D.; GRANAL, M.W. Abundance, variability and chromosomal location of microsatellite in wheat. **Molecular General Genetics**, Berlin, v.246, p.327-333, 1995.

ROMÁN, B.; SATOVIC, Z.; POZARKOVA, D.; MACAS, J.; DOLEZEL, J.; CUBERO, J.I.; TORRES, A.M. Development of a composite map in *Vicia faba*, breeding applications and future prospects. **Theoretical and Applied Genetics**, Berlin, v.108, p.1079-1088, 2004.

ROSSI, A.D. Comercialização do maracujá. In: RUGGIERO, C. (Ed.). **Maracujá: do plantio à colheita**. Jaboticabal: FUNEP, 1998. p. 279-287.

ROZEN, S.; SKALETISKY, H.J. PRIMER3. 2000. Disponível em: <http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi>. Acesso em: 15 jul. 2003.

RUGGIERO, C. Alguns fatores que podem influir na frutificação. In:_____. **Maracujá**. Ribeirão Preto: Legis Summa, 1987. p.76-85.

RUGGIERO, C.; LAM-SANCHEZ, A.; BANZATTO, D.A. Studies on natural and controlled pollination in yellow passion fruit (*Passiflora edulis* f. *flavicarpa* Deg.). **Acta Horticulturae**, Bélgica, v.57, p.121-124, 1976.

RUGGIERO, C.; SÃO JOSÉ, A.R.; VOLPE, C.A.; OLIVEIRA, J.C.; DURINGAN, J.F.; BAUMGARTNER, J.G.; SILVA, J.R.; NAKAMURA, K.; FERREIRA, M.E.; KAVATI, R.; PEREIRA, V.P. **Maracujá para exportação: aspectos técnicos da produção**. Brasília: EMBRAPA - SPI, 1996. 64 p. (FRUPEX. Produções Técnicas, 19).

SANGER, F. Croonian lecture, 1975 - nucleotide-sequences in DNA. **Proceedings of the Royal Society of London Series B-Biological Sciences**, Great Britain, v.191, p.317-333, 1975.

SANTANA, E.N.; LAU, D. Controle do vírus que causa endurecimento-dos-frutos-do-maracujazeiro. In: ZAMBOLIM, L., VALE, F.X.R.; MONTEIRO, A.J.A., COSTA, H. (Ed.). **Controle de doenças de plantas: fruteiras**. Viçosa: UFV, 2002. v.2, p. 827-836.

SANTOS, S.C.; RUGGIERO, C.; SILVA, C.L.S.P.; LEMOS, E.G.M. A microsatellite library for *Carica papaya* L. cv. Sunrise Solo. **Revista Brasileira de Fruticultura**, Jaboticabal, v.25, p.263-267, 2003.

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

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

SCHULER, G.D.; BOGUSKI, M.S.; STEWART, E.A.; STEIN, L.D.; GYAPAY, G.; RICE, K.; WHITE, R.E.; RODRIGUEZ-TOME, P.; AGGARWAL, A.; BAJOREK, E.; BENTOLILA, S.; BIRREN, B.B.; BUTLER, A.; CASTLE, A.B.; CHIANNILKULCHAI, N.; CHU, A.; CLEE, C.; COWLES, S.; DAY, P.J.; DIBLING, T.; DROUOT, N.; DUNHAM, I.; DUPRAT, S.; EAST, C.; HUDSON, T.J. A gene map of the human genome. **Science**, New York, v.274, p.540-546, 1996.

SCHULTZ, A. **Introdução ao estudo da botânica sistemática**. 3.ed. Porto Alegre: Globo, 1968. 215p.

SEEFELDER, S.; EHRMAIER, H.; SCHWEIZER, G.; SEIGNER, E. Male and female genetic linkage map of hops, *Humulus lupulus*. **Plant Breeding**, Berlin, v.119, p.249-255, 2000.

SENIOR, M.L.; CHIN, E.C.L.; LEE, M.; SMITH, J.S.C.; STUBER, C.W. Simple sequence repeat markers developed from maize sequences found in the GENBANK Database: map construction. **Crop Science**, Madison, v.36, p.1676-1683, 1996.

SILVA, AC.; SÃO JOSÉ, A.R. Classificação botânica do maracujazeiro. In: SÃO JOSÉ, A.C. (Ed). **Maracujá, produção e mercado**. Vitória da Conquista: Universidade Estadual do Sudoeste da Bahia, 1994. 255p.

SILVA, H.D.; VENCOVSKY, R. Testes múltiplos e associação de marcas associadas a QTL's. **Scientia Agricola**, Piracicaba, v.99, p.755-762, 2002.

SNELLING, W.M.; CASAS, E.; STONE, R.T.; KEELE, J.W.; HARHAY, G.P.; BENNETT, G.L.; SMITH, T.P.L. Linkage mapping bovine EST-based SNP. **BMC Genomics**, London, v.6, p.1-10, 2005.

SOBRAL, B.W.S.; HONEYCUTT, R.J. High output genetic mapping of polyploid using PCR-generated markers. **Theoretical and Applied Genetics**, Berlin, v.86, p.105-112, 1993.

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

SONG, K.M.; SUZUKI, J.Y.; SLOCUM, M.K.; WILLIAMS, P.H.; OSBORN, T.C. A linkage map of *Brassica rapus* (syn. *campestris*) based on restriction fragment length polymorphism loci. **Theoretical and Applied Genetics**, Berlin, v. 82, p.296-304, 1991.

SOUTHERN, E.M. Detection of specific sequence among DNA fragments separated by electrophoresis. **Journal of Molecular Biology**, New York, v.98, p.503-517, 1975.

SOUZA, M.M.; PALOMINO, G.; PEREIRA, T.N.S.; PEREIRA, M.G.; VIANA, A.P. Flow cytometric analysis of genome size variation in some *Passiflora* species. **Hereditas**, Mendelska sällskapet, v.141, p. 31-38, 2004.

SPELMAN, R.; BOVENHUIS, H. Genetic response from marker assisted selection in an outbred population for differing marker bracket sizes and with two identified quantitative trait loci. **Genetics**, Bethesda, v.148, p.1389-1396, 1998.

STANKIEWICZ-KOSYL, M.; PITERA, E.; GAWRONSKI, S.W. Mapping QTL involved in powdery mildew resistance of the apple clone U 211. **Plant Breeding**, Berlin, v.124, p.63-66, 2005.

STENZEL, N.M.C.; SERA, T. Melhoramento genético de maracujá-amarelo (*Passiflora edulis* f. *flavicarpa*) no Paraná. In: REUNIÃO TÉCNICA DE PESQUISA EM MARACUJAZEIRO, 2., 1999, Londrina. **Anais**. Londrina: IAPAR-SBF, 1999, p.81.

STOREY, J.D.; TIBSHIRANI, R. Statistical significance for genomewide studies. **Proceedings of the National Academic of Sciences of the United States of America**, Washington, v.5, p.9440-9445, 2003.

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.

STUBER, C.W. Biochemical and molecular markers in plant breeding. In: DUDLEY, J.W.; HALLAUER, A.R.; RYDER, M. (Eds). **Plant Breeding Reviews**, Westport: John Wiley e Sons, 1992. v.9, p.37-61.

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

SUASSUNA, T.M.F.; BRUCKNER, C.H.; CARVALHO, C.R.; BORÉM, A. Self-incompatibility in passionfruit: evidence of gametophytic-sporophytic control. **Theoretical and Applied Genetics**, Berlin, v. 106, p.298-302; 2003.

SWAN, K.A.; CURTIS, D.E.; McCUSICK, K.B.; VOINOV, A.V.; MAPA, F.A.; CANCELLA, M.R. High throughput gene mapping in *Caenorhabditis elegans*. **Genome Research**, New York, v.12, p.1100-1105, 2002.

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

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

TANKSLEY, S.D.; GANAL, M.W.; PRINCE, J.P.; VICENTE, M.C.; BONIERBALE, M.W.; BROUN, P.; FULTON, T.M.; GIOVANNONI, J.J.; GRANDILLO, S.; MARTIN, G.B.; MESSEGUER, R.; MILLER, J.C.; MILLER, L.; PATERSON, A.H.; PINEDA, O.; RÖDER, M.S.; WING, R.A.; WU, W.; YOUNG, N.D. High density molecular linkage maps of the tomato and potato genomes. **Genetics**, Bethesda, v.132, p.1141-1160, 1992.

TANKSLEY, S.D.; MILLER, J.; PATERSON, A.; BERNATSKY, R. **Molecular mapping of plant chromosomes**. In: GUSTAFSON, J.P.; APPELS, R. (Ed.). Chromosome structure and function. New York: Plenum Press, 1988. p.157-173.

TARAMINO, G.; TARCHINI, R.; FERRARIO, S.; LEE, M.; PE, M.E. Characterisation and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. **Theoretical and Applied Genetics**, Berlin, v.95, p.66-72, 1997.

TAUTZ, D. Hypervariability of simple sequences as a general source of polymorphic DNA markers. **Nucleic Acids Research**, London, v.17, p.6463-6471, 1989.

TEMNYKH, S.; PARK, W.D.; AYRES, N.; CARTINHO, S.; HAUCK, N.; LIPOVICH, L.; CHO, Y.G.; ISHII, T.; MCCOUCH, S.R. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). **Theoretical and Applied Genetics**, Berlin, v.100, p.697-712, 2000.

TESTOLIN, R.; HUANG, W.G.; LAIN, O.; MESSINA, R.; VECCHIONE, A.; CIPRIANI, G. A kiwifruit (*Actinidia* spp.) linkage map based on microsatellites and integrated with AFLP markers. **Theoretical and Applied Genetics**, Berlin, v.103, p.30-36, 2001.

THOQUET, F.; GHÉRARDI, M.; JOURNET, E.P.; KERESZT, A.; ANÉ, J.M.; PROSPERI, J.M.; HUGUET, T. The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. **BMC Plant Biology**, London, v.2, p.1-13, 2002.

TÖRJÉK, O.; MEYER, R.C.; MÜSSIG, C.; SCHMID, K.J.; SÖRENSEN, T.R.; WEISSHAAR, B.; MITCHELL-OLDS, T.; ALTMAN, T. Establishment of a high efficiency SNP-based framework marker set for *Arabidopsis*. **The Plant Journal**, Gainesville, v.36, p.122-140, 2003.

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

TREVISAN, F.; MENDES, B.M.J.; MACIEL, S.C ; VIEIRA, M.L.C.; MELETTI, L.M.M.; REZENDE, J.A.M. Resistance to *Passion Fruit Woodiness Virus* in transgenic passionflower expressing the virus coat protein gene. **Plant Disease**, 2006. *In press*.

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

VAN DER SCHOOT, J. POPÍŠKOVÁ, M.; VOSMAN, B.; SMULDERS, M.J.M. Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.) **Theoretical and Applied Genetics**, Berlin, v.101, p.317-322, 2000.

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

VANDERPLANK, J. **Passion flowers**. London: Cambridge Press, 1996. 224p.

VARSHNEY, R.K.; KUMAR, A.; BALYAN, H.S.; ROY, J.K.; PRASAD, M.; GUPTA, P.K. Characterization of microsatellite and development of chromosome specific STMS markers in bread wheat. **Plant Molecular Biology Reporter**, New York, v.18, p.5-16, 2000.

VIANA, A.P.; PEREIRA, T.N.S.; PEREIRA, M.G.; AMARAL JÚNIOR, A.T.; SOUZA, M.M.; MALDONADO, J.F.M. Parâmetros genéticos em populações de maracujazeiro-amarelo. **Revista Ceres**, Viçosa, v.51, p.545-555, 2004.

VIEIRA, M.L.C.; CARNEIRO, M.S. *Passiflora* spp. Passionfruit. In: LITZ, R.E. (Ed.). **Biotechnology of Fruit and Nut Crops**. Oxford; CABI Publishing, 2004, p. 435-453.

VIEIRA, M.L.C.; VELLO, N.A.; SILVA FILHO, M.C. Genética e Melhoramento Vegetal. In: LUIS MIR. (Org.). **Genômica**. São Paulo, 2004. v.1, p. 679-699.

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

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

WANCHANA, S.; KAMOLSUKYUNYONG, W.; RUENGPHAYAK, S.; TOOJINDA, T.; TRAGOONRUNG, S.; VANAVICHITA, A. A rapid construction of a physical ccontig across a 4.5 cM region for rice grain aroma facilitates marker enrichment for positional cloning. **ScienceAsia**, Bangkok, v.31, p.299-306, 2005.

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

WATSON, J.D.; CRICK, F.H.C. Molecular structure of nucleic acids - a structure for deoxyribose nucleic acid. **Nature**, London, v.171, n.4356, p.737-738, 1953.

WEBER, D.; HELENTJARIS, T. Mapping RFLP loci in maize using B-A translocations. **Genetics**, Bethesda, v.121, p.583-590, 1989.

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

WEBER, J.L.; MAY, P.E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. **American Journal of Human Genetics**, Chicago, v.44, p.388-396, 1989.

WEISSENBACH, J.; GYAPAY, G.; DIB, C.; VIGNAL, A.; MORISETTE, J.; MILLASEAU, P.; VAYSSEIX, G.; LATHROP, M. A second generation linkage map of the human genome. **Nature**, London, v.359, p.794-801, 1992.

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

WRIGHT, J.M.; BENTZEN, P. Microsatellites, genetic markers of the future. **Reviews in Fish Biology and Fisheries**, London, v.4, p.384-388, 1994.

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.

WU, S.B.; COLLINS, G.; SEDGLEY, M. A molecular linkage map of olive (*Olea europaea* L.) based on RAPD, microsatellite, and SCAR markers. **Genome**, Ottawa, v.47, p.26-35, 2004.

YIN, T.; ZHANG, Z.; HUANG, M.; WANG, M.; ZHUGE, Q.; TU, S.; ZHU, L.H.; WU, R. Molecular linkage maps of the *Populus* genome. **Genome**, Ottawa, v.45, p.541-555, 2002.

YU, J.K.; MANGOR, J.; EDWARDS, E.; THOMPSON, L.; KNAPP, S.J. Simple sequence repeats in sunflower: length polymorphisms among elite inbred lines for several motifs. In: PLANT AND ANIMAL GENOME CONFERENCE, 8., 2000, San Diego. **Abstract**. San Diego, 2000. Disponível em: <<http://www.intl-pag.org/pag/8/abstracts/pag8735.html>>. Acesso em: 16 jul. 2005.

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

ZANE, L.; BARGELLONI, L.; PATARNELLO, T. Strategies for microsatellite isolation: a review. **Molecular Ecology**, Oxford, v.11, p.1-16, 2002.

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