

LIVIA MOURA DE SOUZA

**“MAPA GENÉTICO-MOLECULAR PARA *HEVEA*
BRASILIENSIS E MAPEAMENTO DE QTL’S PARA
CARACTERÍSTICAS DE IMPORTÂNCIA ECONÔMICA”**

CAMPINAS

2012



UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

LIVIA MOURA DE SOUZA

“Mapa genético-molecular para *Hevea brasiliensis* e mapeamento de QTL's para características de importância econômica”

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato (a)
LIVIA MOURA DE SOUZA
[Assinatura]
e aprovada pela Comissão Julgadora.

Tese apresentada ao Instituto de
Biologia para obtenção do Título de
Doutor em GENÉTICA E BIOLOGIA
MOLECULAR na área de GENÉTICA
VEGETAL E MELHORAMENTO.

Orientadora: Profa. Dra Anete Pereira de Souza

Co-Orientador: Prof. Dr. Antônio Augusto Franco Garcia

Campinas, 2012

FICHA CATALOGRÁFICA ELABORADA POR
ROBERTA CRISTINA DAL' EVEDOVE TARTAROTTI – CRB8/7430
BIBLIOTECA DO INSTITUTO DE BIOLOGIA - UNICAMP

So89m Souza, Livia Moura de, 1980-
Mapa genético-molecular para *Hevea brasiliensis* e
mapeamento de QTL's para características de
importância econômica / Livia Moura de Souza. –
Campinas, SP: [s.n.], 2011.

Orientador: Anete Pereira de Souza.
Coorientador: Antônio Augusto Franco Garcia.
Tese (doutorado) – Universidade Estadual de
Campinas, Instituto de Biologia.

1. Hevea. 2. Microssatélites (Genética). 3.
Mapeamento cromossômico. 4. Mapeamento de QTL.
I. Souza, Anete Pereira de, 1962-. II. Garcia, Antônio
Augusto Franco. III. Universidade Estadual de
Campinas. Instituto de Biologia. IV. Título.

Informações para Biblioteca Digital

Título em Inglês: Genetic linkage map for *H. brasiliensis* and QTL's mapping for
important economic traits

Palavras-chave em Inglês:

Hevea

Microsatellites (Genetics)

Chromosome mapping

QTL mapping

Área de concentração: Genética Vegetal e Melhoramento

Titulação: Doutor em Genética e Biologia Molecular

Banca examinadora:

Anete Pereira de Souza [Orientador]

Mario Luiz Teixeira de Moraes

Mirian Perez Maluf

Tatiana Teixeira de Souza Chies

Maria Imaculada Zucchi

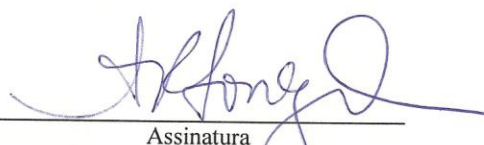
Data da defesa: 27-04-2012

Programa de Pós Graduação: Genética e Biologia Molecular

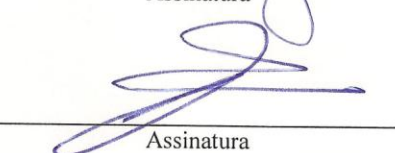
Campinas, 27 de abril de 2012

BANCA EXAMINADORA

Prof(a). Dr(a) Anete Pereira de Souza (Orientador(a))


Assinatura

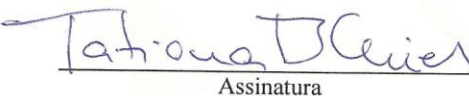
Prof. Dr. Mario Luis Teixeira de Moraes


Assinatura

Prof(a). Dr(a) . Mirian Perez Maluf


Assinatura

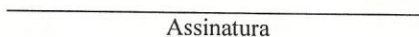
Prof(a). Dr(a) . Tatiana Teixeira de Souza Chies


Assinatura

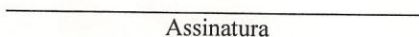
Prof(a). Dr(a) . Maria Imaculada Zucchi


Assinatura

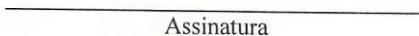
Prof(a). Dr(a) . Michel Georges Albert Vincentz


Assinatura

Prof(a). Dr(a) . José Baldin Pinheiro


Assinatura

Prof(a). Dr(a) . Ananda Aguiar


Assinatura



Cidade Universitária "Zeferino Vaz",
17 de maio de 2012.

CIBio: 03/2012

Projeto CIBio: 14/2003

Identificação:

Doutorado: Livia Moura de Souza, CPG-GBM UNICAMP

Projeto: Mapa genético-molecular para *Hevea brasiliensis* e mapeamento de QTL's para características de importância econômica.

Parecer:

Projeto aprovado pela CIBio/CBMEG em 03/02/2003 sob número 14/2003
(em andamento)

Coordenador: Profa. Dra. Anete Pereira de Souza

Profa. Dra. Edi Lúcia Sartorato
Presidente da CIBio/CBMEG - UNICAMP

"A mente que se abre a uma nova idéia nunca mais volta ao
seu tamanho original."
(Albert Einstein)

Agradecimentos

À Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp), pelo apoio financeiro no projeto e pela bolsa concedida.

À professora Anete Pereira de Souza, pela orientação e por ter me encontrado. Seu exemplo de garra e de luta por justiça e pela ciência sem conflitos de interesse vão me acompanhar sempre. Obrigada por ter confiado em mim para a realização desse trabalho e por sempre me mostrar que no final tudo dá certo.

Ao professor A. Augusto Franco Garcia pela co-orientação e por toda gentileza e disponibilidade no esclarecimento das minhas dúvidas, por me apresentar de forma clara a estatística como aliada.

Ao professor Paulo Souza Gonçalves que sempre foi uma grande fonte de inspiração para continuar a estudar a seringueira, obrigada pelo entusiasmo e por ter me passado seu conhecimento.

Aos pesquisadores do AGAP-CIRAD, Vincent Le Guen e ao Dominique Garcia, colaboradores desse trabalho, por toda ajuda científica e até mesmo com minhas dificuldades com o francês.

À empresa Michelin, pela população de mapeamento e aos seus funcionários pela ajuda com coleta de dados fenotípicos.

À Embrapa pelo material genético utilizado nas análises.

Aos membros da pré-banca, pela disponibilidade e ajuda na pré-avaliação da tese.

Aos membros da banca de defesa, pela disponibilidade em contribuir para que esse trabalho fosse melhorado.

Às secretárias Sandra e Tânia pela atenção, profissionalismo e carinho em me atender em todos os momentos que precisei.

Aos colegas e amigos do laboratório de Análises Genética e Molecular / CBMEG e Barracão, pelos conselhos e ajuda nos momentos de dúvidas, pelos momentos de descontração no nosso “espaço gourmet”, pelo carinho durante estes anos de convivência. Não citarei nomes, pois, com certeza, me esquecerei de alguém. Em quase cinco anos vocês foram muito importantes para que eu ainda tenha a impressão de ter começado ontem, a convivência no laboratório é tão agradável que o tempo parece não ter passado.

Aos colegas do Laboratório de Genética Estatística, ESALQ/USP, essa parceria tem muito futuro!

Ao colega Rodrigo Gazaffi, pelo enorme auxílio na condução das análises estatísticas, pela paciência, pela disponibilidade e pela imensa ajuda com a análise dos dados.

Aos funcionários, Carlão e Zildinha, aos técnicos que passaram pelo laboratório, principalmente à Patty, Alininha e Valéria por serem sempre prestativas e gentis, pela participação essencial no bom andamento dos experimentos.

Ao Juverlandi, um grande amigo sempre disposto a ajudar, obrigada por todas as horas de bate papo no nosso cafezinho e por sempre cuidar do laboratório como sua segunda casa e dos alunos como seus filhos.

À Camila, com quem divido este trabalho desde o começo, passamos por todas as experiências juntas, inúmeras coletas, viagens, acrilamidas... Obrigada pela enorme e fundamental ajuda no desenvolvimento desse projeto, ele também é seu!

Ao Suzuki, por ter alegrado nossa bancada por um tempo e nos ajudado tanto.

A grande amiga Melissa, por ter me recebido no laboratório quando cheguei sem nem saber usar pipeta, por toda a paciência no início. Mesmo com um mundo de distância, continua sempre sendo muito querida.

A toda a minha família, a de sangue (Moura e Souza) e aquela que escolhi para fazer parte da minha vida (Brito e Muszynski), por todo apoio.

Ao meu marido João que sempre acreditou em mim, me apoiando e incentivando com gestos de amor e carinho, muita paciência com meus desesperos e compreensão com as viagens de congressos, coletas... Obrigada por sempre estar ao meu lado!

Às minhas irmãs Cintia e Nadia, sempre me dando apoio e aos meus amores Lara e Joaquim, queridinhos da titia, que nem sempre entendiam que eu tinha que estudar.

Aos meus pais Fátima e Gilberto, pois não tenho palavras que justifiquem todo meu agradecimento, obrigada pelo exemplo de vida e por sempre acreditarem em mim.

Aos que acreditaram sempre, agradeço pela força e apoio, e aos que questionaram também agradeço por me darem ainda mais força para continuar. Também agradeço aos maus momentos e aos meus erros, pois me fizeram entender e dar valor exato as coisas boas.



“Não é o mais forte da espécie que sobrevive, nem o mais inteligente. É aquele que melhor se adapta as mudanças”.

Charles Darwin

Dedico

Aos meus avós
pelo exemplo e inspiração.

Sumário

Resumo.....	xi
Abstract.....	xiii
Prefácio.....	xv
Introdução	1
Capítulo I.....	3
Revisão bibliográfica	5
O gênero <i>Hevea</i>	5
Marcadores moleculares e mapeamento com ênfase em seringueira	8
Marcadores moleculares	8
Mapeamento genético	11
Mapeamento de QTLs	14
Capítulo II.....	17
Objetivos	19
Objetivo Geral	19
Objetivos Específicos	19
Capítulo III.....	21
"Microsatellites from rubber tree (<i>Hevea brasiliensis</i>) for genetic diversity analysis and cross-amplification in six <i>Hevea</i> wild species"	
Capítulo VI.....	33
"Development and characterization of 296 new polymorphic microsatellite markers for rubber tree (<i>Hevea brasiliensis</i>)"	
Capítulo V.....	55
"An integrated molecular genetics linkage map in a full-sib family and QTL mapping of traits related to the growth of the rubber tree (<i>Hevea brasiliensis</i>) in summer and winter in tropical regions"	
Capítulo VI.....	91
Considerações Finais	95
Capítulo VII.....	101
Conclusões	103
Perspectivas	104
Referências.....	105

Resumo

A seringueira é uma espécie de cruzamento misto e com um longo ciclo de vida, o que dificulta a geração de linhagens endogâmicas e, por consequência, a construção e integração de mapas de ligação usando as metodologias convencionais. A exemplo do que tem sido feito em outras espécies vegetais que apresentam limitações para a obtenção de linhagens endogâmicas, os mapas genéticos existentes para seringueira foram construídos utilizando populações F_1 (com diferentes tipos de segregação). A integração dos mapas obtidos para cada um dos genitores é possível com base em marcadores bi-parental, onde ambos os genitores são heterozigóticos podendo segregar nas proporções 1:1:1:1, 1:2:1 e 3:1 na progênie F_1 . Além disso, o uso de marcadores codominantes e multialélicos, como os microssatélites, acrescentam muito à construção do mapa, uma vez que permite a obtenção de estimativas de frequência de recombinação e da fase de ligação com um menor viés. Os microssatélites são marcadores moleculares consagrados, sendo eles a ferramenta de escolha no estudo de diversos organismos pela simplicidade de uso e de análise. Entretanto, para espécies que ainda não têm grande parte de seu genoma sequenciado, a obtenção de marcadores desse tipo passa pela necessidade de desenvolvimento via bibliotecas genômicas.

Esse trabalho objetivou desenvolver, caracterizar e mapear microssatélites em *Hevea brasiliensis* para possibilitar estudos genético-moleculares da variação morfológica, identificação de regiões genômicas associadas a fenótipos de interesse, bem como análises genético-populacionais.

Uma biblioteca genômica enriquecida em motivos repetitivos foi construída, a partir da qual foram desenvolvidos os microssatélites utilizados nesse trabalho. A utilização dos microssatélites de *Hevea brasiliensis* em amplificações heterólogas envolvendo seis espécies do gênero *Hevea* resultou em aproximadamente 98% de sucesso nas amplificações, sugerindo a existência de um complexo formado pelas diferentes espécies do gênero.

Um mapa genético-molecular foi construído utilizando-se 284 marcadores microssatélites na análise de uma população segregante F_1 com 270 indivíduos, a partir do cruzamento entre os genitores PB217 e PR255. Por meio da utilização do programa ONEMAP foi possível a construção de um mapa de ligação integrado que revelou 2840 cM de extensão, distribuídos em 23 grupos de ligação. O mapeamento de QTLs realizado utilizando metodologia de mapeamento por intervalo

composto (CIM) detectou 24 QTLs para altura e circunferência das plantas entre as estações de verão e inverno.

Este trabalho é pioneiro na construção de um mapa integrado para seringueira pelo fato dele ter sido construído unicamente com marcadores microssatélites, os quais são altamente informativos. Além disso, é também o primeiro estudo envolvendo análise de QTL para características relacionadas ao crescimento de plantas em seringueira, bem como é inédita a aplicação da metodologia CIM para população F_1 segregante.

Abstract

The rubber tree is an mixed crossing species with a long life cycle, which makes it difficult for the generation of inbred lines, and, therefore, the construction and integration of linkage maps by using conventional methodologies. Similar to what has been reported from other plant species that have limitations to obtain inbred lines, the genetic maps for rubber tree have been constructed by using F_1 populations (with different types of segregation). The integration of maps obtained for each of the genitors is possible based on bi-parental markers, where both genitors are heterozygous, being able to segregate in ratios 1:1:1:1, 1:2:1 and 3:1 in the F_1 progeny. Moreover, the use of co-dominant and multi-allele markers, such as microsatellites, adds a lot to the construction of maps, since it allows the obtainment of estimates of recombination frequency and linkage phase with less bias. The microsatellites are established molecular markers which are the tool of choice in the study of various organisms for their simplicity of use and analysis. Nevertheless, their application in species whose genomes have not yet been sequenced requires a prior development phase.

The present study intended to develop, characterize and map microsatellites for the species *Hevea brasiliensis* so that initiatives concerning morphological variation, identification of genomic regions linked to phenotypes of interest, as well as population genetic analysis.

A repetitive DNA-enriched library was constructed from which it was developed the microsatellites used in this work. The use of *Hevea brasiliensis* microsatellites for heterologous amplification in other six *Hevea* species was done with approximately 98% of success ratio, suggesting the existence of a complex formed by different species.

A molecular genetic map was constructed using 284 microsatellite markers in the analysis of an F_1 segregating population with 270 individuals from a cross between the parents PB217 and PR255. By using the program ONEMAP it was possible the construction of an integrated linkage map that revealed 2840 cM in length, divided into 23 linkage groups. The QTL mapping methodology was performed using composite interval mapping (CIM) and detected 24 QTLs for plant height and circumference between summer and winter.

This work is a pioneer in building an integrated map for rubber because it was built only with microsatellite markers, which are highly informative. Furthermore, it is also the first one involving the analysis of QTL for characteristics related to the growth of rubber tree plants, as well as novel application of the CIM methodology for this type of population.

Prefácio

Este manuscrito de Tese está apresentado no formato de capítulos. Os resultados obtidos durante as diferentes etapas do desenvolvimento deste trabalho estão apresentados no formato de artigos científicos.

Inicialmente apresenta-se uma revisão bibliográfica, que relata os avanços na pesquisa realizada com *Hevea brasiliensis*, bem como reporta as características dos marcadores microssatélites utilizados como ferramenta molecular em estudos genéticos.

O artigo apresentado no Capítulo IV, intitulado “Microsatellites from rubber tree (*Hevea brasiliensis*) for genetic diversity analysis and cross-amplification in six *Hevea* wild species”, publicado na Conservation Genetics Resources (1, 75-79, 2009), descreve o desenvolvimento de 27 microssatélites a partir de uma biblioteca enriquecida, bem como a avaliação de 31 genótipos de *Hevea brasiliensis* e de outras seis espécies de *Hevea* utilizando-se os microssatélites desenvolvidos.

No Capítulo V, apresenta-se o artigo resultante da participação da autora desta tese em co-autoria no artigo intitulado “Development and characterization of 296 new polymorphic microsatellite markers for rubber tree (*Hevea brasiliensis*)”, o qual foi publicado na revista Plant Breeding (130, 294-296, 2011). Foi feita a genotipagem de 296 marcadores microssatélites (desenvolvidos pelo grupo de pesquisa do CIRAD-França), em 10 genótipos de *Hevea brasiliensis*. O trabalho de colaboração foi realizado durante estágio de dois meses desenvolvido no laboratório do CIRAD em Montpellier. Os genitores da população de mapeamento (que é objeto central de estudo dessa tese) estavam entre os 10 genótipos escolhidos para caracterização dos 296 locos microssatélites. Os resultados do trabalho de caracterização dos 296 marcadores determinaram o polimorfismo existente entre os genitores da população de mapeamento em estudo para os marcadores analisados. Dessa forma, foi possível genotipar a população de mapeamento empregando-se também os marcadores polimórficos desenvolvidos pelos pesquisadores do CIRAD-França.

No Capítulo VI está apresentado o artigo intitulado “An integrated molecular genetics linkage map in a full-sib family and QTL mapping of traits related to the growth of the rubber tree (*Hevea brasiliensis*) in summer and winter in tropical regions” descreve a obtenção do

mapa genético com marcadores microssatélites (desenvolvidos e também aqueles disponíveis na literatura), bem como o mapeamento de características quantitativas associadas ao crescimento e vigor das plantas durante o verão e o inverno.

Finalmente, apresenta-se uma discussão envolvendo todos os resultados obtidos ao longo do desenvolvimento deste trabalho a conclusão geral e perspectivas para a continuidade do trabalho.

Introdução

A seringueira, *Hevea brasiliensis*, ao contrário da maioria das plantas cultivadas, é uma espécie que ainda está sendo domesticada, em razão de ser uma das maiores fontes produtoras de borracha natural. Sua história é tão curta que quase não mudou o seu aspecto de árvore nativa encontrada na Amazônia. No Brasil, a história da produção da borracha vegetal mostra que o país desfrutou da condição de principal produtor e exportador mundial até a metade do século passado, tornando-se importador dessa matéria-prima a partir de 1951.

O maior problema encontrado no cultivo de seringueira em regiões de clima tropical úmido, como é o caso da região Amazônica, é que a condição climática de alta temperatura e alta umidade relativa do ar favorece o desenvolvimento da doença conhecida como "mal das folhas" ou SALB (South American Leaf Blight) em inglês, causada pelo ascomiceto *Microcyclos ulei*. A necessidade de novos cultivares de seringueira adaptados a diferentes regiões ecológicas, denominadas de regiões de escape, constitui um ponto fundamental para o sucesso da heveicultura no Brasil e no mundo.

Outro fator a ser considerado é o tempo necessário para cada ciclo de melhoramento genético da seringueira. Um ciclo de melhoramento vai do cruzamento para obtenção das progênes recombinações até a avaliação final da produção das plantas da progênie podendo demorar aproximadamente 30 anos. Desta forma, torna-se fundamental o desenvolvimento de novas técnicas de avaliação precoce, que possibilitem diminuir e otimizar as avaliações da produção da seringueira.

A biotecnologia e as técnicas de biologia molecular podem ser muito úteis para tornar mais eficiente e rápida a obtenção de novos cultivares. O uso de marcadores moleculares para auxiliar o melhoramento genético é uma estratégia inovadora e eficiente. Eles possibilitam a análise de regiões genômicas que possam, por exemplo, controlar diferentes fenótipos ou que sejam indicativas da dinâmica populacional. Até o momento esta técnica ainda tem um desenvolvimento muito tímido e com poucos avanços em termos de aplicação para o conhecimento do genoma e melhoramento da seringueira.

Neste contexto, o presente trabalho teve como principal objetivo o desenvolvimento de um mapa-genético molecular a partir de microssatélites e o mapeamento de locos associados a características relacionadas ao desenvolvimento e crescimento das plantas durante as estações de verão e inverno. A localização de regiões relacionadas ao crescimento das plantas durante o inverno poderá auxiliar em programas de melhoramento genético que buscam genótipos apropriados para áreas de escape.

Para o desenvolvimento do mapa genético baseado em marcadores microssatélites foi necessária a construção de bibliotecas enriquecidas, a partir das quais regiões contendo microssatélites foram isolados e caracterizados.

A construção de um mapa genético-molecular contribuirá para o mapeamento de muitos outros genes/características importantes, bem como pode resultar também em possibilidade de seleção assistida por marcadores. Uma vez construído um mapa de ligação saturado para *H. brasiliensis*, espera-se poder estender o uso desta ferramenta a diferentes cruzamentos e germoplasmas de seringueira. Além disso, os marcadores desenvolvidos poderão ser utilizados em estudos sobre variabilidade genética, estrutura populacional, fluxo gênico e sistema reprodutivo em *H. brasiliensis*.

Este trabalho foi desenvolvido em colaboração com o Centro de Seringueira do Instituto Agrônomo de Campinas (IAC), Departamento de Genética da ESALQ/USP, Unidade de Pesquisa “Amélioration Génétique et Adaptation des Plantes” do CIRAD em Montpellier-França e, a empresa Michelin do Brasil.

Capítulo I

Revisão bibliográfica

O gênero *Hevea*

A grande importância do gênero *Hevea* decorre da influência que a borracha veio a exercer sobre a civilização do homem, chegando mesmo a caracterizar o ciclo da borracha, cuja origem data de logo após a descoberta da América. Os produtos de borracha provêm de vários grupos de plantas nem sempre muito aparentadas botanicamente, podendo incluir representantes das famílias Euphorbiaceae, Apocynaceae, Asteraceae, etc., dentre as quais se destaca o gênero *Hevea*.

O centro primário de diversidade genética do gênero *Hevea* é o rio Negro, na confluência com o rio Amazonas. Já o centro secundário abrange uma vasta área nas proximidades do município de Borba, no baixo rio Madeira (Wycherley, 1977). Ao redor do centro de origem, espécies de *Hevea* ocorrem naturalmente na Bolívia, Brasil, Colômbia, Guiana Francesa, Guiana, Peru, Suriname e Venezuela. Todas as espécies ocorrem no Brasil exceto *H. microphylla* (Priyadarshan e Clement-Demange, 2004).

A seringueira *Hevea brasiliensis* (Willd. ex A.D.R. de Juss.) Muell.-Arg. é a espécie mais importante desse gênero (Gonçalves et al. 1989) por ser a principal fonte de borracha natural existente.

A literatura científica sobre *Hevea* teve início em 1775 com a descrição de *Hevea guianensis* Aublet (Priyadarshan et al., 2009). É um taxon natural, isto é, um grupo, uma entidade taxonômica definida, muito bem delimitada e de fácil reconhecimento (Pires et al., 2002). Pertence à família Euphorbiaceae que inclui outros importantes gêneros de culturas tropicais, tais como *Ricinus* (mamona) e *Manihot* (mandioca).

Segundo Pires et al. (2002), Richard propôs em 1791 a substituição do nome genérico *Hevea* por *Siphonia*, uma latinização de *Heve* e, em 1824, Jussieu publicou o binômio *Siphonia brasiliensis* para seringueira verdadeira da Amazônia. Como o nome *Siphonia* não estava de acordo com o código Internacional de Nomenclatura, o gênero *Hevea* foi restabelecido em 1865 por Müller, que propôs para seringueira a combinação *Hevea brasiliensis*. Depois dessa segunda espécie classificada, outras espécies foram sendo descritas e a literatura foi se

avolumando em vista do grande interesse pela borracha, de maneira que muitas impropriedades foram surgindo, o que pode ser avaliado pelo considerável número de nomes publicados: 40 espécies e ao todo, 96 incluindo as entidades de categoria sub-específica.

Recentemente foi estabelecida uma relativa ordem quanto à classificação das espécies, alguns estudos taxonômicos foram conduzidos por Baldwin Junior (1947), Seibert (1947), Pires (1973), Schultes (1977), e mais recentemente, Priyadarshan e Gonçalves (2003) e Pires et al. (2002). Baldwin Junior (1947), baseado em levantamento acurado de espécies do gênero *Hevea* nativas da Amazônia e, combinando essas informações com observações citológicas, chegou à conclusão que o gênero *Hevea* possui nove espécies. No Peru, Seibert (1947) reconheceu oito espécies. Os estudos de Schultes (1977) levaram-no a reconhecer nove espécies e quatro variedades, e Pires (1973) reconheceu 11 espécies, incluindo no grupo a mais nova, *Hevea camargoana*, encontrada na Ilha de Marajó.

A classificação atual das espécies do gênero *Hevea* é baseada nos estudos conduzidos por Bailon e Mueller-Argovienensis, e por Hubber, Pax e Ducke (Pires et al., 2002). No Brasil, são conhecidas 11 espécies: *Hevea benthamiana* Müll.Arg., *Hevea brasiliensis* (Willd. ex A.Juss.) Müll.Arg., *Hevea camargoana* Pires, *Hevea camporum* Ducke, *Hevea guianensis* Aubl., *Hevea microphylla* Ule, *Hevea nitida* Mart. ex Müll.Arg., *Hevea paludosa* Ule, *Hevea pauciflora* (Spruce ex Benth.) Müll.Arg., *Hevea rigidifolia* (Spruce ex Benth.) Müll.Arg., *Hevea spruceana* (Benth.) Müll.Arg (Cordeiro et al., 2012).

Segundo Pires et al. (2002), tantas as confusões, interpretações pouco precisas e desacordos entre autores sobre o número e a classificação das espécies do gênero *Hevea* são devido ao fato que apesar desse gênero ser um táxon bem definido, bem limitado e de fácil reconhecimento, o mesmo não se pode dizer das suas espécies, que são dificilmente separáveis uma das outras. O gênero *Hevea* não pode ser separado em espécies naturais por causa das variações ecológicas, das formas de transição e dos muito frequentes híbridos naturais. Não há ainda, entre as espécies, barreiras de reprodução bem individualizadas, podendo-se dizer que se elas tratam de espécies “incipientes”. Poderíamos dizer que elas representam um complexo de espécies em evolução para a especiação.

A falta de clareza nas definições para esse gênero não ocorre somente na classificação taxonômica. Ela também é observada na definição de sua origem ancestral. Alguns estudos citogenéticos apontam que a seringueira seria um anfidiplóide ($2n=4x=36$) que se estabilizou durante o curso da evolução (Priyadarshan et al., 2009). Esta afirmação é apoiada na observação de tetravalentes durante a meiose (Ong, 1975). No entanto, na prática, *Hevea* é considerado como um gênero diplóide ($2n=2x=36$). Estudos de hibridação *in situ* revelaram dois distintos locos 18S-25S rDNA e um loco 5SrDNA, sugerindo uma possível origem alotetraplóide com a perda de um loco 5S rDNA durante o curso da evolução (Leitch et al., 1998). Entretanto, Lespinnasse et al. (2000a) consideraram que embora algumas observações possam sugerir a origem a partir de duas espécies ancestrais com $2n=18$, nenhum ancestral diplóide foi encontrado e as observações feitas com marcadores moleculares revelam que *Hevea* comporta-se como um diplóide $2n=36$.

Alguns trabalhos com marcadores moleculares têm avaliado a transferibilidade de marcadores desenvolvidos para *Hevea brasiliensis* em outras espécies. Souza et al. (2009) observou que todos os microssatélites analisados apresentaram amplificação com variação no tamanho dos alelos para as espécies *H. guianensis*, *H. rigidifolia*, *H. benthamiana*, *H. camargoana*, *H. nítida*, *H. pauciflora*. Saha et al. (2005) obteve transferibilidade dos microssatélites de *H. brasiliensis* para *H. benthamiana* e *H. spruceana*. Essa amplificação “entre-especies” indica claramente que as regiões flanqueadoras dos microssatélites são bem conservadas entre as espécies de *Hevea*. A hibridação e introgressão entre as espécies ocorrem frequentemente em cruzamentos experimentais, mostrando ausência de barreiras genéticas para a reprodução. Wyrcherley (1976) sugere a hipótese de que a especiação de *Hevea* ocorreu em razão da autonomia reduzida do voo dos insetos polinizadores e do sincronismo no florescimento dessas espécies.

Gonçalves e Fontes (em Borém et al., 2009) concluem que a seringueira, ao contrário da maioria das plantas cultivadas, é uma espécie que está sendo domesticada no mundo moderno em razão de ser uma das melhores fontes produtoras de borracha natural. Sua história é tão curta que quase não mudou o seu aspecto de árvore nativa encontrada na Amazônia.

Marcadores moleculares e mapeamento com ênfase em seringueira

Marcadores moleculares

Marcadores moleculares surgiram devido à necessidade de detecção de polimorfismo diretamente no DNA. Eles são definidos como características genéticas que diferenciam dois ou mais indivíduos e possuem herança mendeliana. Os diversos tipos de marcadores moleculares disponíveis diferenciam-se pela tecnologia utilizada, pela sua capacidade de detectar diferenças entre indivíduos, custo, facilidade de uso, consistência e reprodutibilidade. Segundo Milach (1998), as metodologias utilizadas para identificar os tipos de marcadores moleculares dividem-se em dois grupos: hibridização ou amplificação de DNA.

Entre os marcadores mais conhecidos identificados por hibridização estão os marcadores RFLP (*Restriction Fragment Length Polymorphism*) e minissatélites ou locos VNTR (*Variable Number of Tandem Repeats*; Jeffreys et al., 1985). A partir do surgimento da metodologia de amplificação de DNA via reação de polimerase em cadeia (Polimerase Chain Reaction), diversos marcadores moleculares foram desenvolvidos: AFLPs (*Amplified Fragment Length Polymorphisms*) (Vos et al., 1995), RAPDs (*Random Amplified Polymorphic DNA*) (Welsh e McClelland, 1990), SSR (*Simple Sequence Repeats*) (Litt e Luty, 1989), ISSR (*Inter Simple Sequence Repeat*; Zietkiewicz et al., 1994), SNP (*Single Nucleotide Polymorphism*) (Rafalsky, 2002), entre outros. Todos os sistemas de marcadores já foram utilizados no gênero *Hevea* para seleção assistida, facilitando a identificação e caracterização de genes, com exceção apenas dos marcadores SNPs (Priyadarshan e Gonçalves, 2003). Apesar de todas as vantagens deste marcador, na literatura somente um único recente trabalho sobre o desenvolvimento desse marcador foi realizado, onde um total de 5.883 supostas posições de SNP foram descobertas, e 10 marcadores SNP bialélicos foram validados a partir de sequências derivadas de ESTs. Os autores concluem que estes marcadores podem ser úteis para a construção de mapas genéticos, identificação de locos de características quantitativas ligadas a características desejáveis, e o estudo da estrutura genética em *H. brasiliensis* (Pootakham et al., 2011). Os marcadores baseados em PCR abriram novas e inúmeras possibilidades de detecção e avaliação do polimorfismo encontrado na molécula de DNA.

Os marcadores microssatélites ou SSRs (*Simple Sequence Repeats*) têm representado um importante avanço para estudos genéticos. Este tipo de marcador apresenta alto nível de polimorfismo, codominância, elevada heterozigosidade e padrão de herança Mendeliana. Além disso, a sua detecção é possível a partir de qualquer tipo de tecido em qualquer estágio de desenvolvimento. São marcadores baseados na amplificação via PCR (Mullis e Faloona, 1987), facilitando a transferência da técnica entre diferentes grupos de pesquisa.

Microssatélites estão presentes nos genomas de todos os organismos estudados até o momento, sendo encontrados de forma dispersa e em elevada frequência (Hamada *et al.* 1982, Litt e Luty 1989). São sequências de 2-6 pares de bases de comprimento repetidas em *tandem*.

Tais sequências estão distribuídas ao acaso no genoma eucariótico constituindo os locos polimórficos, os quais surgem devido à variação no número de repetições do motivo de DNA em cada loco microssatélite. Dessa forma, cada loco microssatélite pode apresentar diferentes alelos entre os indivíduos da mesma espécie na natureza. As sequências de DNA que flanqueiam estas regiões são conservadas e estáveis em uma dada espécie e, muitas vezes, entre espécies diferentes de um mesmo gênero, permitindo a utilização de *primers* flanqueadores dos microssatélites de uma espécie para as amplificações de microssatélites de diferentes espécies.

O número de repetições do microssatélite pode variar, sendo que sua sequência adjacente costuma ser conservada entre indivíduos diferentes da mesma espécie ou de espécies próximas. Tal fato que permite que o polimorfismo no número de repetições possa ser detectado via PCR, através da utilização de *primers* específicos complementares às regiões adjacentes. Sua principal desvantagem é a necessidade do conhecimento prévio da sequência de DNA da espécie a ser estudada. Porém com a redução dos custos de sequenciamento, esses marcadores se tornaram populares e vêm sendo desenvolvidos a partir de sequências genômicas ou expressas. Sua utilização em plantas é ampla, sendo utilizados tanto em estudos de populações naturais como em estudos mais aplicados em espécies cultivadas (Temnykh *et al.*, 2000; Peleg *et al.*, 2008; Kahrood *et al.*, 2008).

A hipervariabilidade de microssatélites não é totalmente entendida, podendo ser originada através de *crossing-over* desigual ou deslizamento de sequências durante a replicação

do DNA. Estes eventos podem resultar em variação do número de cópias da unidade básica de repetição (Schlotterer e Tautz, 1992). Os motivos repetidos mais frequentes encontrados pela primeira vez foram repetições (AC)_n e (AG)_n (Condit e Hubbell, 1991). Uma busca em bancos de dados de sequências de DNA de eucariotos publicadas revelou que os sítios de microssatélites são distribuídos em uma frequência de 1 a cada 50 mil pares de bases.

A presença dessas características faz com que os microssatélites sejam marcadores indicados para mapeamento genético, para identificação e discriminação de genótipos e estudos de genética de populações.

No gênero *Hevea*, Low et al. (1996) detectaram pela primeira vez, marcadores do tipo microssatélites através de pesquisas em bancos de dados em algumas sequências de genes de *Hevea*. Utilizando esses marcadores, os autores encontraram elevado nível de polimorfismo em *H. pauciflora*, *H. guinensis*, *H. camargoana*, *H. benthamiana* e *H. brasiliensis*. Bibliotecas enriquecidas em sequências repetidas foram produzidas por Atan et al. (1996), Seguin et al. (2003), Roy et al. (2004), Saha et al. (2005) conduzindo à identificação de um grande número de marcadores microssatélites. Souza et al. (2009) a partir do desenvolvimento de biblioteca enriquecida em microssatélites, desenvolveram 27 novos marcadores microssatélites para *H. brasiliensis* e avaliaram a transferibilidade em seis outras espécies do gênero. Outro trabalho recente relacionado com o desenvolvimento de marcadores microssatélites a partir de biblioteca enriquecida foi publicado por Le Guen et al. (2010), onde foram descritos e caracterizados em 10 genótipos de *H. brasiliensis* 296 novos marcadores microssatélites.

A partir de bases de dados de domínio público, Feng et al. (2009) desenvolveram marcadores EST-SSR, detectaram transferibilidade interespecífica e intergenérica desses marcadores e analisaram diversidade genética em 12 genótipos de *H. brasiliensis* e de outras quatro espécies do gênero *Hevea*. Yu et al. (2011) desenvolveram marcadores microssatélites a partir de DNA genômico de *H. brasiliensis* e verificaram a transferibilidade desses marcadores em quatro outras espécies do gênero *Hevea* (*Hevea spruceana*, *Hevea nitida*, *Hevea benthamiana* e *Hevea pauciflora*) e três espécies da família Euphorbiaceae (*Ricinus communis*, *Manihot utilissima* e *Phyllanthus emblica*).

Mapeamento genético

Os fundamentos teóricos sobre a ligação gênica, base para a construção de mapas genéticos (Sturtevant, 1913), conduziu à ideia de fazer um mapa de ligação utilizando a frequência de recombinantes como medida da distância entre dois genes (Coelho e Silva, 2002). Esta técnica clássica foi rejuvenescida a partir do desenvolvimento da tecnologia de marcadores moleculares na década de 80.

A disponibilidade de marcadores, aliada a métodos estatísticos complexos, tem permitido a construção de mapas genéticos para a maioria das espécies vegetais de interesse econômico, até mesmo para aquelas de longo ciclo de vida, como as florestais e as frutíferas (Carneiro e Vieira, 2002).

O desenvolvimento de mapas genéticos é considerado uma das aplicações de maior impacto da tecnologia de marcadores moleculares na análise genética de espécies, e potencialmente, no melhoramento de plantas. No contexto do melhoramento de plantas, mapas genéticos possibilitam a cobertura e análise de genomas; a decomposição de características genéticas complexas nos seus componentes Mendelianos; a localização das regiões genômicas que controlam caracteres de importância; a quantificação do efeito destas regiões na característica estudada e a canalização de toda esta informação para o uso em programas de melhoramento (Cruz e Silva, 2009).

A construção de um mapa de ligação envolve basicamente a aplicação de técnicas de biologia molecular aos conceitos originais de herança genética demonstrada por Mendel e pode ser dividido em três etapas principais: população de mapeamento, seleção de marcadores polimórficos e análise de ligação entre os marcadores.

A escolha de cruzamentos de forma a maximizar a probabilidade de detectar polimorfismo ao nível de DNA é o primeiro passo para a formação de uma população de mapeamento. Um aspecto importante a este respeito é a distância genética que separa os genitores a serem examinados; quanto mais próximos geneticamente, mais difícil será a obtenção de polimorfismos (Paterson et al., 1991). Assim, a escolha de genitores sem ancestral comum, ou mesmo o uso de cruzamentos interespecíficos para espécies de autofecundação, permite otimizar o processo de análise com marcadores moleculares (Tanksley et al., 1993).

Em mapeamento, torna-se importante a utilização de populações segregantes com o máximo de desequilíbrio de ligação, o qual pode ser resultante de diversos fatores (entre os quais seleção e deriva genética), mas o principal é a ligação física dos locos. O desequilíbrio de ligação devido à ligação física dos locos é bastante elevado em populações de cruzamentos controlados; consequentemente, a capacidade de se detectar a ligação entre os dois locos fisicamente ligados também é elevada (Tanksley, 1993; Falconer e Mackay, 1996; Lynch e Walsh 1998).

A população segregante tradicionalmente utilizada para a construção de mapas genéticos é derivada do cruzamento entre linhagens endogâmicas. As gerações F_2 ou aquelas oriundas de retrocruzamentos, cruzamentos entre heterozigotos e linhagens duplo-haplóides são as populações mais usadas na construção de mapas de ligação (Tanksley, 1993). Em espécies de ciclo longo, como é o caso da seringueira, a obtenção de linhagens é impraticável, o que impede o desenvolvimento de população de mapeamento a partir do cruzamento entre linhagens.

Algumas estratégias foram criadas para viabilizar o mapeamento nessas espécies, conhecidas como “outcrossing species”, sendo que uma delas é o duplo pseudo-testecross (Ritter et al., 1990). Essa estratégia consiste na construção de dois mapas individuais, um para cada genitor, no qual cada um apresenta marcadores que segregam em cada genitor (Grattapaglia e Sederoff, 1994).

A construção do mapa genético em *Hevea* requer metodologia específica devido à elevada heterozigosidade nessa espécie. O cruzamento entre dois genitores heterozigotos em *Hevea* pode produzir informação sobre quatro alelos, que segregarão futuramente. Um mapa genético molecular de *Hevea brasiliensis* foi formulado com a ajuda de marcadores RFLPs, AFLPs, microssatélites e isoenzimas por Lespinnasse et al. (2000). Foi realizado um duplo pseudo test-cross e um mapa foi construído separadamente para cada genitor. Os marcadores que segregavam em ambos os genitores foram usados para construir um único mapa. Os genitores usados foram PB 260 (Pb 5/51 x PB 49) e RO 38 (F4542 x AVROS 363). O mapa sintético de F_1 continha 717 marcas que foram distribuídas entre 18 grupos de ligação. Este trabalho foi composto de 301 RFLP, 388 AFLP, 18 microssatélite e 10 isoenzimas. O tamanho genético dos

18 cromossomos foi razoavelmente homogêneo, com média de 120 cM por cromossomo. Muitos marcadores AFLP foram observados em grupo, o que pode ser atribuído à reduzida frequência de regiões de recombinação. Contudo, os marcadores RFLP foram bem distribuídos entre os 18 grupos de ligação, mas isso é insuficiente para um mapa saturado. Todavia, esse trabalho foi o primeiro passo para a construção de um mapa genético em *Hevea*.

Entretanto, tanto do ponto de vista biológico como estatístico, é desejável a integração das informações contidas nesses mapas individuais em um único mapa integrado. Isso só pode ser feito com a presença de marcadores em heterozigose em ambos os genitores, que são utilizados para estabelecer relações de ligação entre os marcadores segregando individualmente em cada genitor (Wu et al., 2000; Garcia et al., 2006; Oliveira et al., 2007).

A construção de mapas genéticos integrados, com a utilização de diferentes tipos de marcadores moleculares que apresentam diferentes segregações traz grandes vantagens, pois permite aumentar a saturação do mapa de ligação e estender a caracterização da variação polimórfica em todo o genoma. A localização de QTLs com maior precisão é facilitada com a disponibilidade de um mapa genético integrado (Maliepaard et al., 1997).

Wu et al. (2002) propuseram a metodologia para a obtenção das fases de ligação e frações de recombinação entre marcadores com diversos tipos de segregação, que foi implementado mais tarde pelo programa ONEMAP (Margarido et al., 2007). Apesar de marcadores dominantes contribuírem para a saturação dos grupos de ligação, marcadores biparentais dominantes com segregação 3:1 e codominantes com segregação 1:2:1 e 1:1:1:1 devem ser utilizados como ponte, permitindo a ligação entre os dois mapas. Esse mesmo autor afirma que é possível a ocorrência de até 18 grupos de segregação, com até quatro alelos informativos para espécies diplóides. As segregações mais informativas, tipo 1:2:1 e 1:1:1:1 só podem ser obtidas por marcadores codominantes tipo microssatélites e SNPs.

Mapeamento de QTLs

Caracteres de importância econômica, em sua maioria, podem ser classificados como quantitativos. “Quantitative trait loci” designados por QTLs são regiões do genoma responsáveis pela expressão de caracteres fenotípicos, que possuem distribuição contínua tais como, altura e peso de plantas e de animais; produção de grãos; teor de óleo; etc.

Até recentemente, análises genéticas e o melhoramento de caracteres quantitativos eram baseados na biometria (Falconer, 1981). Essa estratégia fundamentou-se na caracterização coletiva dos múltiplos fatores afetando uma característica quantitativa possibilitando uma participação da variância fenotípica total nos seus componentes genético e ambiental. A caracterização individual dos locos envolvidos numa característica quantitativa com o auxílio de marcadores genéticos foi feita inicialmente por Sax (1923), em feijão (*Phaseolus vulgaris* L.) e, posteriormente por Thoday (1961), em *Drosophila*. Entretanto, os métodos biométricos não são capazes de permitir a caracterização ou a manipulação dos locos específicos de um grande número de caracteres para melhoramento (Paterson, 1988). Os principais fatores limitantes nesses estudos são o número e a natureza dos marcadores genéticos disponíveis.

Com o advento dos marcadores moleculares (Lander e Bolstein, 1989) tornou-se possível mapear regiões cromossômicas (QTLs) que afetam esses caracteres quantitativos. Uma das aplicações mais importantes dos mapas genéticos é a localização de genes que controlam características de importância econômica que resultam da ação cumulativa de um conjunto de genes. Essas características são denominadas poligênicas, quantitativas ou de herança complexa, e os locos que as controlam são chamados de QTLs (Tanksley, 1993).

Mapear um QTL significa identificar sua posição no genoma e estimar seus efeitos genéticos, tais como: o efeito aditivo, efeito de dominância e outros efeitos presentes no modelo adotado. Três métodos têm sido muito utilizados para mapeamento de QTLs: análise de marcas simples, mapeamento por intervalo (Lander e Bolstein, 1989) e mapeamento por intervalo composto (Zeng, 1993; Jansen 1994; Zeng, 1994).

O método de marcas simples baseia-se na comparação entre as médias fenotípicas obtidas para as diferentes classes genotípicas de um determinado marcador, individualmente.

Os métodos estatísticos geralmente utilizados para verificar se as diferenças observadas são significativas incluem o teste t, análise de variância, regressão linear simples e o método de máxima verossimilhança (Coelho e Silva, 2002). Esse método é o mais simples e não requer a construção de mapas. No entanto a posição do QTL não pode ser precisamente determinada e não é possível diferenciar um QTL de um pequeno efeito situado próximo ao marcador de um QTL de grande efeito (Liu, 1998).

O mapeamento por intervalo utiliza como base as frequências de um par de marcadores adjacentes e um QTL localizado entre esses dois marcadores e, portanto, requer a existência de um mapa de ligação. De acordo com Schuster e Cruz (2004) esse método consiste das seguintes etapas: como a distância entre cada par de marcador é conhecida, o método investiga o intervalo usando a função de mapeamento e estimando os parâmetros do modelo a cada ponto de análise, em seguida a função de verossimilhança ou a análise de regressão (Haley e Knott, 1992) é usada para os cálculos na forma de LOD score para cada ponto, o máximo valor de LOD indica a presença de um QTL.

Este método apresenta vantagens quando comparado ao de marcas simples, pois permite que sejam estimados os parâmetros envolvidos na análise como a média, a variância, o efeito do QTL e os marcadores adjacentes (Lander e Botstein, 1989; Liu, 1998). Embora o mapeamento por intervalo seja mais eficiente que a análise de marcas simples, ele não permite analisar a interação entre QTLs presentes em diferentes regiões do genoma, podendo, portanto mapear “QTLs fantasmas” (Coelho e Silva, 2002).

Com o mapeamento por intervalo composto pode-se resolver parte desses problemas (Schuster e Cruz, 2004), pois ele combina o mapeamento por intervalo com análise de regressão múltipla, que tem o objetivo de eliminar a interferência de QTLs adjacentes ao intervalo que está sendo mapeado, reduzindo a identificação de QTLs inexistentes (Coelho e Silva, 2002). Os marcadores selecionados por regressão são utilizados como cofatores, permitindo controlar os efeitos de outros QTLs. Embora esse método esteja sendo muito utilizado atualmente, ele considera um único QTL de cada vez. Uma desvantagem decorrente do fato de ser um método uni-dimensional tem o risco de se colocar muitas marcas como cofatores.

Para uma melhor capacidade de análise um quarto método foi desenvolvido, chamado de mapeamento por intervalos múltiplos (Kao et al., 1999). Esse método utiliza simultaneamente intervalos múltiplos de marcadores para a identificação de múltiplos possíveis QTLs. Dessa forma, podem ser estimados os efeitos de interação entre QTLs, ou seja, a epistasia, os valores genotípicos dos indivíduos e as herdabilidades das características quantitativas (Kao et al., 1999). O Mapeamento por intervalos múltiplos permite um aumento na precisão e no poder de detecção de QTLs (Kao et al., 1999; Coelho e Silva, 2002).

Em seringueira alguns QTLs já foram mapeados com o objetivo de detectar tolerância a SALB (*Microcyclus ulei*). Lespinasse et al., 2000b utilizaram 195 genótipos provenientes de um cruzamento entre clones de PB 260 (suscetível) e RO 38 (resistente), dando continuidade às análises genéticas realizadas anteriormente (Seguin et al., 1996; Lespinasse et al., 2000a). Oito QTLs foram identificados para tolerância em seringueira. O mapa consenso da população F₁ confirmou os resultados obtidos com o mapa dos genitores.

Le Guen *et al.* (2003), estudando genes que conferem resistência a SALB, encontraram um grande QTL (*M13-lbn*), localizado no g13 no mapa de RO 38, responsável por 36 a 89% da variância fenotípica para resistência.

Mais recentemente Le Guen et al. (2011) desenvolveram um mapa genético para o cruzamento entre PB260 e MDF180 com o objetivo de mapear QTLs para a resistência ao fungo *Microcyclus ulei*, MDF180 já havia sido descrito como um clone resistente ao fungo. O mapa genético foi estabelecido usando marcadores microssatélites e AFLPs. Não foi encontrada nenhuma resistência no parental suscetível, já para o parental resistente um gene qualitativo denominado M15md foi localizado no grupo de ligação 15, quatro QTLs menores também foram identificados, e dois deles apresentaram interação epistática com M15md.

Embora alguns trabalhos com mapeamento genético e mapeamento de QTLs já tenham sido desenvolvidos, ainda não foi possível obter um mapa altamente saturado com marcadores informativos, e o mapeamento de QTLs só foi conduzido, até o presente momento visando obter um maior entendimento da resistência ao *Microcyclus*, nada foi realizado para outras características de interesse econômico.

Capítulo II

Objetivos

Objetivo Geral

- Construção de um mapa genético-molecular e mapeamento de genes associados a características de importância econômica, em seringueira, através de marcadores moleculares do tipo microssatélites.

Objetivos Específicos

- Desenvolvimento de microssatélites para *Hevea brasiliensis*
 - Análise de sequências;
 - Desenho de iniciadores;
 - Caracterização dos microssatélites;
 - Otimização de condições de amplificação;
 - Avaliação do índice polimorfismo.
- Construção do mapa genético-molecular utilizando marcadores microssatélites:
 - Genotipagem da população de mapeamento;
 - Localização dos microssatélites em grupos de ligação.
- Mapeamento de características fenotípicas de interesse:
 - Mensurações fenotípicas de características qualitativas e quantitativas;
 - Mapeamento dos marcadores morfológicos;
 - identificação de QTLs responsáveis pelas características em estudo.

Capítulo III

Capítulo III

“Microsatellites from rubber tree (*Hevea brasiliensis*) for genetic diversity analysis and cross-amplification in six *Hevea* wild species”

Livia Moura de Souza

Camila Campos Mantello

Melissa Oliveira Santos

Paulo de Souza Gonçalves

Anete Pereira de Souza

Conservation Genetics Resources (2009) 1: 75-79.

Microsatellites from rubber tree (*Hevea brasiliensis*) for genetic diversity analysis and cross-amplification in six *Hevea* wild species

L. M. Souza ¹, C. C. Mantello ¹, M. O. Santos¹, P. de S. Gonçalves^{3,4} and A. P. Souza^{1,2}

¹Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas (Unicamp), CP 6010, Campinas, SP CEP 13083-970, Brazil,

²Departamento de Biologia Vegetal.

³Empresa de Pesquisa Agropecuária Brasileira, (Embrapa), ⁴Instituto Agronômico de Campinas (IAC), CP 28, Campinas, SP 13012-970, Brazil;

* Both authors contributed equally to this work

Corresponding author: Anete Pereira de Souza, e-mail: anete@unicamp.br;
PHONE: +55 19 3521 1132, FAX Number: +55 19 3521-1089

Running Title: Microsatellites in *Hevea brasiliensis*

Abstract

Hevea brasiliensis is native to the Amazonian rain forest and an important source of natural rubber. Twenty seven polymorphic microsatellite loci were isolated and characterized from a GA–CA enriched genomic library of *H. brasiliensis*. The number of alleles ranged from two to 20. The observed and expected heterozygosity ranged from 0.13 to 0.88 and from 0.00 to 0.89, respectively. Cross-species amplification of the markers developed for *H. brasiliensis* was successful in the wild *Hevea* species *H. guianensis*, *H. rigidifolia*, *H. nitida*, *H. pauciflora*, *H. benthamiana* and *H. camargoana*. The data indicated a high degree of sequence homology in the microsatellite flanking regions of these species. The developed SSR loci are a potential powerful tool for studies of population genetics, genetic diversity and gene flow among *Hevea* species.

Keywords: *Hevea spp.*, *Hevea brasiliensis*, Microsatellite, Genetic Diversity, Transferability, Cross-Species Amplification.

Hevea brasiliensis is native to the Amazonian rainforest. The genus *Hevea* belongs to the Euphorbiaceae family and comprises 11 inter-crossable species (Pires et al. 2002), which have evolved in the Amazonian rainforest over 100 thousand years ago (Clement-Demange et al. 2000). The species *H. brasiliensis* is the prominent source of natural rubber. In many of its most significant applications, natural rubber cannot be replaced by synthetic alternatives and thus it is of high strategic importance. This singularity is due to the natural rubber unique properties, as resilience, elasticity, impact and abrasion resistance, efficient heat dispersion and malleability at cold temperatures (Cornish 2001; Cataldo 2000).

The genetic variability of *H. brasiliensis* is high at the center of origin and knowledge on such variability is fundamental for conservation, breeding and commercial production of this species. Most often, specific phenotypes of discrete variation have been used as morphological markers. In addition, molecular markers could be highly beneficial as a tool in assisting genetic characterization and breeding (Nodari et al. 1997; Brondani et al. 1998). Moreover, the

conservation of genetic variability and utilization of allied gene resources is critical for the improvement of crop species, including *H. brasiliensis*.

Microsatellites are considered suitable markers for genetic studies. They have the advantage of combining co-dominance and high polymorphism plus abundance and uniform dispersion in plant genomes, so that they are capable of discriminating closely related individuals (Brondani et al. 1998).

Cross-species amplification is highly useful and depends on the extent of the DNA sequence conservation in the primer sites flanking the microsatellite loci and the stability of those sequences during the species evolution (Choumane et al. 2000; Decroocq et al. 2003; Zucchi et al. 2003). Closely related species are more likely to share similar microsatellite priming sites, but also transfer of functional microsatellite primers is occasionally possible from more distantly related species (Lorieux et al. 2000).

In this work, we report the development of *Hevea brasiliensis* primers for amplification of polymorphic microsatellite loci, which are suitable to cross-amplify in six wild *Hevea* species and useful for genetic diversity studies of *Hevea* spp. germplasm.

The studied germplasm comprised 31 genotypes of the rubber tree *H. brasiliensis*, which were gently provided by the Agronomic Institute of Campinas (IAC), as well as the wild congeneric species *H. guianensis*, *H. rigidifolia*, *H. nitida*, *H. pauciflora*, *H. benthamiana* and *H. camargoana*, gently provided by the Brazilian Agricultural Research Corporation (EMBRAPA - Amazônia Ocidental) in Manaus, AM.

Genomic DNA samples were extracted from lyophilized leaf tissues using a modified CTAB method (Doyle and Doyle 1987). A genomic enriched library for *H. brasiliensis* was constructed according to the methodology described by Billote et al. (1999). The DNA samples were digested with *RS*AI and enriched using (CT)₈ and (GT)₈ biotinylated microsatellite probes. Selected DNA fragments were amplified by PCR and then cloned into the pGEM-T vector (Promega). Competent XL1-blue *Escherichia coli* cells were transformed with the recombinant plasmids and cultivated on agar medium containing ampicillin and 100 µg/ml of X-galactosidase. Clones containing the insert were sequenced using the Big Dye Terminator v3.1

Cycle Sequencing Kit (Applied Biosystems) and an automated ABI 377 sequencer (Applied Biosystems).

Of the sequenced *E. coli* clones, 41.2% contained simple sequence repeats. From the 79 sequenced clones with microsatellite motifs, a total of 27 primer pairs complementary to sequences flanking the repeat motifs were designed using PrimerSelect (DNASar). Polymorphism was evaluated in 31 *Hevea brasiliensis* genotypes from the IAC germoplasm collection. The microsatellite fragments were PCR amplified in 25 µL reaction volume, consisting of 20 ng of template DNA, 0.8 µM of each primer, 200 µM of each dNTP, 1.5 mM MgCl₂, 20 mM Tris-HCl, 50 mM KCl, and 0.5 U *Taq* DNA Polymerase. Immediately after an initial denaturation at 94°C for 4 min, PCR amplifications were performed in 30 cycles of 1 min at 94°C, 45 s at the specific annealing temperature of each primer pair and 1 min at 72 °C, followed by a final extension at 72° for 10 min. PCR products were separated by electrophoresis on denaturing acrylamide gels and silver stained (Creste et al. 2001). Molecular size of the DNA fragments was estimated by comparison with standard 10-bp DNA Ladder (Invitrogen).

The 27 evaluated microsatellite loci were polymorphic and showed single-locus amplification (Table 1) in the 31 *Hevea brasiliensis* genotypes. The number of alleles per locus varied from 2 to 16. The observed (*Ho*) and expected (*He*) heterozygosity ranged from 0.13 to 0.88 and from 0.0001 to 0.89, respectively. Descriptive statistics and the test for Hardy-Weinberg Equilibrium were performed using Tools for Genetic Population Analysis (TFPGA) (Miller 1997). Thirteen loci depart significantly from Hardy-Weinberg Equilibrium ($P < 0.05$) (Table 1). The linkage disequilibrium was tested using the Popgene 1.32 (Yeh and Boyle, 1997) and no disequilibrium was detected among all loci. The Polymorph Information Content (PIC) ranged from 0.12 to 0.85 (Table 1).

Cross-species amplifications of the 27 loci in the six wild species *H. guianensis*, *H. rigidifolia*, *H. nitida*, *H. pauciflora* (112 CNSG and 116 CNSG), *H. benthamiana* and *H. camargoana* was investigated using equal PCR conditions. The 27 loci were successfully amplified on the evaluated *Hevea* species (Table 2). The species *H. guianensis*, *H. rigidifolia* and *H. pauciflora* showed a duplicated locus for the HB30, HB17 and HB18 microsatellites, respectively. Locus duplication in *Hevea brasiliensis* was previously reported by Seguin et al.

(1996) and Lespinasse et al. (2000), however this is the first report of loci duplication in other *Hevea* species.

The data presented herein demonstrated a high potential applicability of these isolated microsatellites markers for analyses of intra- and interspecific genetic diversity and gene flow, as well as phylogenetic relationships among the cultivated rubber tree *H. brasiliensis* and wild *Hevea* species. These studies will be of considerable value for the conservation and use of *Hevea* genetic resources.

Acknowledgements

The authors thank the Brazilian Agricultural Research Corporation (EMBRAPA - Amazônia Ocidental) and the Agronomic Institute of Campinas for donating the analyzed *Hevea* germoplasm. The present research was financed by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, 2007/50392-1, 2007/59804-0); L.M.S. and C.M.C. received respectively a undergraduate and graduate scholarship from FAPESP. A.P.S. and P.S.G. received a research fellowship grant from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

- Billotte N, Lagoda PJR, Risterucci AM, Baurens FC (1999) Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits* 54: 277-288
- Brondani RPV, Brondani C, Tarchini R and Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theor Appl Genet* 97:816-827.
- Cataldo, F. (2000) Guayule rubber: a new possible world scenario for the production of natural rubber. *Progr. Rubber Plastics Technol.* 16, 31–59.
- Choumane W, Winter P, Weigand F and Kahl G (2000) Conservation and variability of sequence-tagged microsatellite sites (STMSs) from chickpea (*Cicer arietinum* L.) within the genera *Cicer*. *Theor Appl Genet* 101:269-278
- Clement-Demange A., Legnate H., Seguin M., Carron M.P., Guen V. Le., Chapuset T. et al. (2000). Rubber Tree. In: Charrier A., Jacquot M., Hamon S. and Nicolas D. (eds), *Tropical Plant Breeding*. Collection Reperes, CIRAD-ORSTOM, Montpellier, France, pp. 455–480.
- Cornish, K. (2001) Similarities and differences in rubber biochemistry among plant species. *Phytochem.* 57, 1123–1134
- Creste, S.A et al. (2001). Detection of Single Sequence Repeat Polymorphism in denaturing Polyacrylamide Sequencing Gels by Silver Staining. *P. Mol. Biology Reporter* 19: 299-306.
- Decroocq V, Fave MG, Hagen L, Bordenave L and Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. *Theor Appl Genet* 106:912-922.
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin* 19: 11–15
- Pires, J.M.; Secco, R.de S.; Gomes, J.I. (2002) *Taxonomia e fitogeografia das seringueiras Hevea spp.* Belém: Embrapa Amazônia Oriental, 103p.

- Lespinasse, D.; et al. (2000) Identification of QTLs involved in the resistance to South American Leaf Blight (*Microcyclus ulei*) in the rubber tree. Theor Appl Genet 100: 975-984.
- Lorieux M, Ndjiondjop M-N and Ghesquière A (2000) A first interspecific *Oryza sativa* e *Oryza glaberrima* microsatellite-based genetic linkage map. Theor Appl Genet 100:593-601.
- Miller MP (1997) Tools for population genetic analysis (TFPGA) 1.3. A windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author.
- Nodari RO, Ducroquet JP, Guerra MP and Meler K (1997) Genetic variability of *Feijoa sellowiana* germplasm. Acta Hort 452:41-46.
- Seguin, M.; et al. (1996) *Hevea* molecular genetics. Plantations, recherche, development 3: 77-88.
- Zucchi MI, Brondani RPV, Pinheiro JB, Chaves LJ, Coelho ASG and Vencovsky R (2003) Genetic structure and gene flow in *Eugenia dysenterica* DC in the Brazilian Cerrado utilizing SSR markers. Genet Mol Biol 26:449-457.
- Yeh FC, Rong-Cai Y, Boyle T (1998) POPGENE version 1.31. Edmonton, Alberta, Canada: University of Alberta, Center for International Forestry Research.

Table 1 – Characteristics of 27 microsatellite loci from *Hevea brasiliensis*. Shown are the loci names, the GeneBank accession number, the forward (F) and reverse (R) primer sequence, repeat motif, number of alleles (N), product size range in base pairs, observed (Ho) and expected (He) heterozygosity and P value HWE. *departs significantly from HWE at $p < 0.005$ after Bonferroni correction.

Locus	Gene bank accession n°	Primer sequence (5'-3')	Repeat motif	N	TA (°C)	Size range(bp)	Ho	He	PIC	P value HWE
HB-1	FJ919780	CTGATGCTGCCAAGCAATAC	(tg)10	9	60	186 - 165	0.645	0.732	0.687	0.623
		CAAACATCGCACTCTCCTCA								
HB-2	FJ919781	TGGGCACAGTTTACAAATAGC	(ca)8	4	60	240 - 232	0.258	0.562	0.512	0.000*
		CACCATGAAATGAATGCCTCT								
HB-3	FJ919782	CACCCACATTAAACAGGACA	(tg)8	4	60	255 - 240	0.200	0.445	0.412	0.007*
		GGTTGCCTTGCTGCTCAT								
HB-4	FJ919783	GGAAGAAAAGGGAATAGAGAGC	(ac)10(aa)(ac)5	5	60	177 - 150	0.645	0.611	0.520	0.092
		ACCCTGCTTAGGCTTGAATG								
HB-6	FJ919785	TGAAAGAAAATGGTGGCTCA	(ac)6	2	60	237 - 235	0.413	0.406	0.319	0.847
		AAAATTTCGGCAAGTCAAGA								
HB-7	FJ919786	AACTTCGCATGAGAGCACAA	(ac)6	2	60	165 - 162	0.896	0.508	0.374	0.000*
		ACTGTGACTTTGCCTCCAC								
HB-8	FJ919787	GGAGGTATTGCTCATGTATGCT	(ca)9	6	60	217 - 210	0.133	0.329	0.312	0.000*
		AGGGCTTTTACAATGTGGA								
HB-9	FJ919788	CACATCATCCATGGTTTTAAATGG	(gt)8	16	60	198 - 168	0.666	0.883	0.858	0.595
		CCATGGAGGCTCTGAAGTG								
HB-10	FJ919789	CACATCATCCATGGTTTTAAATGG	(gt)7(atgt)(ga)17	10	60	203 - 168	0.633	0.831	0.798	0.584
		CTCCATGAGGCTCTGAAGTG								
HB-11	FJ919790	AAAATTTCGGCAAGTCAAGA	(gt)6	4	60	252 - 249	0.333	0.568	0.511	0.005*
		TGAAAGAAAATGGTGGCTCA								
HB-12	FJ919791	AGAATAAAGCCTCCGTGTCTG	(ag)20	9	60	246 - 224	0.548	0.719	0.662	0.474
		GAAGCCAAATGTGTCTGCTGC								
HB-14	FJ919792	TGGGAGAAAGTGGAGGAGAA	(ta)5	2	57	234 - 232	0.064	0.228	0.199	0.000*
		CACCACCTTGTTCACATTG								
HB-15	FJ919793	CGTTGAACGATTTGGAGATG	(ag)5 e (ag)17	11	60	218 - 194	0.677	0.858	0.827	0.032*
		GCTTGGTTTTGAAGACTGC								
HB-16	FJ919794	CGGGTATAGTGGGAGCAGAG	(aag)6	6	60	198 - 188	0.387	0.650	0.576	0.048*
		ACACAACGCCACTGCTTATG								
HB-17	FJ919795	AGGGCTTCGGGACAATCA	(at)4(gt)6	10	60	212 - 201	0.666	0.742	0.700	0.956
		GACATATGCCCAACAAGTG								
HB-18	FJ919796	AAGGGGAAAAGAGAAGAAA	(ag)10	2	57	148 - 146	0.290	0.252	0.217	0.344
		AAGCCCAAGAAAAGAAAGT								
HB-19	FJ919797	AATGGTGTGCAATGTTTCA	(ag)16	7	57	212 - 197	0.551	0.770	0.719	0.335
		CGGGGTCTTCTCTTCA								
HB-20	FJ919798	TATTGGCTTGCTTCTAACC	(gt)13(ga)14	6	60	150 - 124	0.322	0.343	0.322	0.852
		CCTTGACCCACTATCATCT								
HB-21	FJ919799	TTAATTTTATCAGCTTTT	(ca)3 e (at)5	8	47	164 - 145	0.400	0.545	0.519	0.171
		GCATTTACAGTATTTGAA								
HB-22	FJ919800	ATGATCGCTATTCGCTATGG	(ct)17	3	60	178 - 167	0.500	0.406	0.331	0.290
		GATTTGCCCTTCGTCTCC								
HB-24	FJ919801	GCCAAATCAATCACTCATCC	(ca)11	9	60	264 - 232	0.269	0.435	0.414	0.013*
		GTGCTTCGGCTCAATCTGTT								
HB-25	FJ919802	ATGACGCAACTTCTTCTCC	(ga)5 e (tg)9	4	60	226 - 217	0.222	0.608	0.523	0.000*
		TGTTGTATCTTGCTTTCT								
HB-26	FJ919803	TTGCCGAAAGGTAACCTCA	(tg)7	2	56	163 - 162	0.000	0.130	0.120	0.000*
		TGTAATTTGCTGTCCACTTC								
HB-27	FJ919804	AGTCGTGTGTGTCAACATCA	(ttc)6	6	60	177 - 153	0.225	0.607	0.528	0.000*
		TATTCTCCTTGTCGTTGTG								
HB-28	FJ919804	ACAACGACAAGGAGAATAGG	(gtg)5	3	60	160 - 154	0.344	0.305	0.279	0.262
		CAATCGAATATCACAGACG								
HB-29	FJ919805	AGAAACATGCAAGGAACAC	(tc)16	9	56	183 - 163	0.733	0.736	0.692	0.428
		AGCGGACATACGAAGGAA								
HB-30	FJ919806	ATAAACACAAGTGAGTCCTG	(tc)16	9	51	234 - 208	0.193	0.648	0.620	0.000*
		GTTTAGTATATCCCATCCAC								

Table 2. Cross-amplification of developed *H. brasiliensis* microsatellite loci in others species of *Havea*

Primer	<i>H. guianensis</i>	<i>H.rigidifolia</i>	<i>H.benthamiana</i>	<i>H.camargoana</i>	<i>H.nitida</i>	<i>H. pauciflora</i> (112 CNSG)	<i>H.pauciflora</i> (116CNSG)
HB-1	+	+	+	+	+	+	+
HB-2	-	+	+	-	+	+	+
HB-3	+	+	+	+	+	+	+
HB-4	-	+	-	+	+	+	+
HB-6	+	-	+	+	-	+	+
HB-7	+	+	+	+	+	-	+
HB-8	+	-	+	+	+	+	-
HB-9	+	-	+	+	-	-	+
HB-10	+	+	-	+	+	+	+
HB-11	+	+	+	+	+	+	+
HB-12	+	+	+	+	+	+	+
HB-13	-	-	+	+	+	+	+
HB-14	+	+	+	+	+	+	+
HB-15	+	+	+	+	+	+	+
HB-16	+	+	+	+	+	+	+
HB-17	+	+	+	+	+	+	+
HB-18	+	+	+	+	+	+	+
HB-19	+	+	+	+	+	+	+
HB-20	+	+	+	+	+	-	+
HB-21	-	-	-	+	+	-	-
HB-22	+	+	+	+	+	+	+
HB-23	+	+	+	+	+	+	+
HB-24	+	+	+	+	+	-	+
HB-25	-	+	+	+	+	+	+
HB-26	-	+	+	+	+	+	+
HB-27	+	-	+	+	+	+	+
HB-28	+	-	+	+	+	+	+
HB-29	+	+	+	+	+	+	+
HB-30	+	+	+	+	+	+	+

Capítulo IV

Capítulo IV

“Development and characterization of 296 new polymorphic microsatellite markers for rubber tree (*Hevea brasiliensis*)”

Vincent Le Guen
Céline Gay
Tou-Cheu Xiong
Livia Moura de Souza
Maguy Rodier-Goud
Marc Seguin

Plant Breeding 130:294-296,2011

Short Communication

“Development and characterization of two 296 new polymorphic microsatellite markers for rubber tree (*Hevea brasiliensis*)”

V. Le Guen⁽¹⁾, C. Gay⁽²⁾, T.C. Xion⁽³⁾, L.M. Souza⁽⁴⁾, M. Rodier-Goud⁽¹⁾, M. Seguin⁽¹⁾

(1) : CIRAD, Joint Research Unit DAP, 34398 Montpellier Cedex 5, France

(2) : INRA, Joint Research Unit DAP, 34398 Montpellier Cedex 5, France

(3) : INRA, Joint Research Unit BPMP-IBIP, Campus INRA-SupAgro, 2 place Viala,
34060 Montpellier Cedex 1, France

(4) : Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual
de Campinas (Unicamp), CP6010, Campinas, SP, CEP 13083-970, Brazil

Abstract

Enriched genomic libraries in GA and GT tandem repeat motifs were generated from rubber tree (*Hevea brasiliensis*). Design of appropriate primer pairs in the flanking regions enables to produce amplified fragments for 310 of these sequences. The test of these genomic microsatellites on a set of 10 rubber tree genotypes allows the characterization of 296 new polymorphic markers. Expected heterozygosity ranged from 0.10 to 0.93 and allele numbers from 2 to 11. Microsatellite loci with GA motifs were more abundant, longer and more polymorphic than GT motifs. Such a high variability suggests the utility of these markers for the establishment of integrated linkage maps, as well as the characterization of linkage disequilibrium in rubber tree populations.

Keywords

Rubber tree – *Hevea brasiliensis* – Microsatellite – SSR

Rubber tree breeding, like for many other tree crops, is a long and tedious process. Large scale cultivation of a new cultivar can only be reached after 20 to 25 years of field experiments on large areas (Webster and Baulkwill, 1989). Breeders are thus interested in developing alternative strategies to improve and speed up the breeding scheme, using genetic markers. Molecular markers are being used for more than fifteen years in *Hevea* breeding, mostly for diversity studies (Besse et al., 1994), assessment of genetic distance among cultivars (Leconte et al., 1994; Varghese et al., 1997), genetic mapping (Lespinnasse et al., 2000a) and identification of genetic loci implied in the expression of agronomical traits (Le Guen et al., 2003; Le Guen et al., 2007; Lespinnasse et al., 2000b). However, markers that were used until now were predominantly isozymes, RAPD, RFLP or AFLP™. Only few microsatellite markers have been already involved in linkage map (Lespinnasse et al., 2000a; Seguin et al., 2001), in characterization of cultivated clones (Saha et al., 2005) or in study of genetic structure of wild populations (Le Guen et al., 2009; Lekawipat et al., 2003). Because of their high level of polymorphism, microsatellite markers are very useful in plant breeding. As they are locus specific markers, a set of microsatellite can be used to construct various linkage maps and

provides bridge markers to merge them in an integrated and denser map. A good coverage of the genome can be further used to assess the linkage disequilibrium and to determine what fraction of it can be attributed to the genetic structure of the studied populations or to the physical proximity between loci in the genome. Such information is fundamental to detect new loci controlling desirable characters in a genome wide association study, or to locate with more precision known genes.

Recently, twenty seven polymorphic microsatellite markers and their use for the evaluation of genetic diversity have been described (Souza et al., 2009). In the present work we describe the enrichment of *Hevea* genomic DNA in GT and GA microsatellite sequences, the development of corresponding loci into molecular markers, and the genetic characterization of these markers.

Total DNA was extracted from fresh leaves of the cultivar PB260, following standard cetyltrimethyl ammonium bromide extraction procedure (Richards et al., 1994). PB260 is a high-yielding Asiatic cultivar which has been used as progenitor in a previous mapping population (Lespinnasse et al., 2000a). A genomic library enriched for CA and GA –repeated microsatellite sequences was produced according to the protocol previously described (Billotte et al., 1999) with the following modifications. Four independent experiments were carried out to produce this enriched library. In the first three experiments, total DNA was restricted with the endonuclease *Rsa*I. The selection of microsatellite sequences was performed using two biotin-labelled oligoprobes, 5'-I*IIITCTCTCTCTCTCTC-3' or 5'-I*IIIIACACACACACACAC-3' where I* is a biotinylated Inosine, either separately (to generate libraries enriched in respectively GA or GT tandem repeats) or jointly to simultaneously enrich in GA and GT repeats. In a fourth experiment, total DNA was restricted with the endonuclease *Sau*3A and the same two oligoprobes were used to simultaneously enrich in GA and GT. After ligation, transformation was made in Epicurian XL1-blue *Escherichia coli* strain. The clones were transferred to microplates containing 150 µl of LB/Ampicilin solution. Insert sizes were estimated by agarose gel electrophoresis of insert PCR products using M13 universal primers. Gels were alkaline Southern transferred to Hybond N+ nylon membranes (Amersham) and hybridized with a $\gamma^{32}\text{P}$ radiolabeled (CT)₁₅ and (CA)₁₅ probes. A total of 1128 clones were selected following this protocol and sequenced by the Centre National de Séquençage (CNS,

Génoscope, Evry, France). Primer pairs were designed using the online Primer 3 package (Whitehead Institute for Biomedical Research), and a total of 399 sequence data were of sufficient quality for a correct primer definition. Oligonucleotides were commercially synthesised by Sigma-Aldrich, with a M13 19 base extension added to forward primers. Polymorphism of the markers was assessed on ten rubber tree genotypes, among which six were Asiatic cultivars, three were issued from (Asiatic cultivars x wild Amazonian genotypes) crosses, and one from a (*H. brasiliensis* x *H. benthamiana*) inter-specific cross. The PCR amplification mixtures (10 µL) contained 25 ng of template DNA, 0.2 µM of each primer, 200µM of dNTP, 2mM of MgCl₂, 1X PCR buffer and 1U *Taq* DNA polymerase. Touch down cycling conditions consisted of an initial denaturing step of 4 min at 94 °C followed by 10 cycles of denaturation, annealing and elongation with a 0.5 °C annealing decrease per cycle starting from 55 °C to 50 °C (94 °C for 45 s, 55 °C for 1min, 72 °C for 1min 15s), 25 cycles with annealing at 50 °C (94 °C for 45s, 50 °C for 1min, 72 °C for 1min), and a final elongation step at 72 °C for 5 min. PCR products were denaturated for 3 min at 94 °C before separation on a 6.5% polyacrylamide gel using a Li-Cor 4300 DNA analyzer. Products were visualized with a fluorescently labelled M13 primer at 700 or 800 nm wavelength.

A total of 396 sequences were tested for their polymorphism, from which 310 (79%) gave detectable amplified products. Numbers of markers according to the type of repeat motif are summarized in Table 1. GA repeats were more abundant than GT repeats, as already characterized in various other plant genomes (Morgante and Olivieri, 1993). The overall percentage of polymorphic markers was 95%.

Characteristics of the 296 polymorphic loci are described in table S1 in supplementary electronic material. Expected and observed heterozygosity, as well as inbreeding coefficient were computed with GDA software (Lewis and Zaykin, 2001). The number of detected alleles on the 10 genotypes varies from 2 to 11 with an average of 5.5. The overall observed heterozygosity was high ($H_o = 0.66$), and positively correlated with the repeat number ($R^2 = 0.64$). Expected heterozygosity (H_e) ranged from 0.10 to 0.93. Marker frequencies were plotted separately for markers with simple GT or GA perfect motifs according to the repeat number (Figure 1) and the allele number (Figure 2). In average, the length of GA microsatellites was

clearly greater (mode = 16 repeats) than that of GT microsatellites (mode = 8 repeats). This is reflected in a greater number of alleles for GA markers (mean = 5.9) than for GT markers (mean = 4.7).

A high proportion of these 296 markers therefore provide sufficient allelic forms to be employed in various linkage maps and to merge these maps for a more precise location of loci. Linkage disequilibrium assessment among wild or improved rubber tree populations can also be considered using this large set of microsatellite markers.

References

- Besse, P., M. Seguin, P. Lebrun, M.H. Chevallier, D. Nicolas, and C. Lanaud, 1994: Genetic diversity among wild and cultivated populations of *Hevea brasiliensis* assessed by nuclear RFLP analysis. *Theoretical and Applied Genetics* **88**, 199-207.
- Billotte, N., P. Lagoda, A.M. Risterucci, and F.C. Baurens, 1999: Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits* **54**, 277-288.
- Le Guen, V., F. Doaré, C. Weber, and M. Seguin, 2009: Genetic structure of Amazonian populations of *Hevea brasiliensis* is shaped by hydrographical network and isolation by distance. *Tree Genetics and Genomes* **5**, 673-683.
- Le Guen, V., D. Lespinasse, G. Oliver, M. Rodier-Goud, F. Pinard, and M. Seguin, 2003: Molecular mapping of genes conferring field resistance to South American Leaf Blight (*Microcyclus ulei*) in rubber tree. *Theoretical and Applied Genetics* **108**, 160-167.
- Le Guen, V., D. Garcia, C.R.R. Mattos, F. Doaré, D. Lespinasse, and M. Seguin, 2007: Bypassing of a polygenic *Microcyclus ulei* resistance in rubber tree, analyzed by QTL detection. *New Phytologist* **173**, 335-345.
- Leconte, A., P. Lebrun, D. Nicolas, and M. Seguin, 1994: Electrophoresis. Application to Hevea clone identification. *Plantations, recherche, développement* **1**, 28-36.
- Lekawipat, N., K. Teerawatanasuk, M. Rodier-Goud, M. Seguin, A. Vanavichit, T. Toojinda, and S. Tragoonrung, 2003: Genetic diversity analysis of wild germplasm and cultivated clones of *Hevea brasiliensis* Muell. Arg. by using microsatellite markers. *Journal of Rubber Research* **6**, 36-47.
- Lespinasse, D., M. Rodier-Goud, L. Grivet, A. Leconte, H. Legnate, and M. Seguin, 2000a: A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. *Theoretical and Applied Genetics* **100**, 127-138.
- Lespinasse, D., L. Grivet, V. Troispoux, M. Rodier-Goud, F. Pinard, and M. Seguin, 2000b: Identification of QTLs involved in the resistance to South American leaf blight (*Microcyclus ulei*) in the rubber tree. *Theoretical and Applied Genetics* **100**, 975-984.

Lewis, P.O., and D. Zaykin, 2001: Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from <http://lewis.eeb.uconn.edu/lewishome/software.html>.

Morgante, M., and A.M. Olivieri, 1993: PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal* **3**, 175-182.

Richards, E., M. Reichardt, and S. Rogers, 1994: Preparation of genomic DNA from plant tissue, In: V. Benson Chanda, (ed.) *Current protocols in molecular biology*, Vol. 1. John Wiley e Sons, Inc., Harvard.

Saha, T., C.B. Roy, and M.A. Nazeer, 2005: Microsatellite variability and its use in the characterization of cultivated clones of *Hevea brasiliensis*. *Plant Breeding* **124**, 86-92.

Seguin, M., C. Gay, T.-C. Xiong, and M. Rodier-Goud, 2001: Microsatellite markers for genome analysis of rubber tree (*Hevea spp.*), In: IRRDB, (ed.) *IRRDB Symposium 2001*, 1-6, Vol. 1, Montpellier (France).

Souza, L.M., C.C. Mantello, M.O. Santos, P.S. Gonçalves, and A.P. Souza, 2009: Microsatellites from rubber tree (*Hevea brasiliensis*) for genetic diversity analysis and cross-amplification in six *Hevea* wild species. *Conservation Genetics Resources* **1**, 75-79.

Varghese, Y.A., C. Knaak, M.R. Sethuraj, and W. Ecke, 1997: Evaluation of random amplified polymorphic DNA (RAPD) markers in *Hevea brasiliensis*. *Plant Breeding* **116**, 47-52.

Webster, C.C., and W.J. Baulkwill, 1989: *Rubber*. Longman ed. Longman Singapore Publishers Ltd.

Acknowledgements

This work was mainly supported by funding from the French government (SCSP-Cirad). DNA sequencing of genomic clones was carried out under a grant from Génoscope (Evry, France).

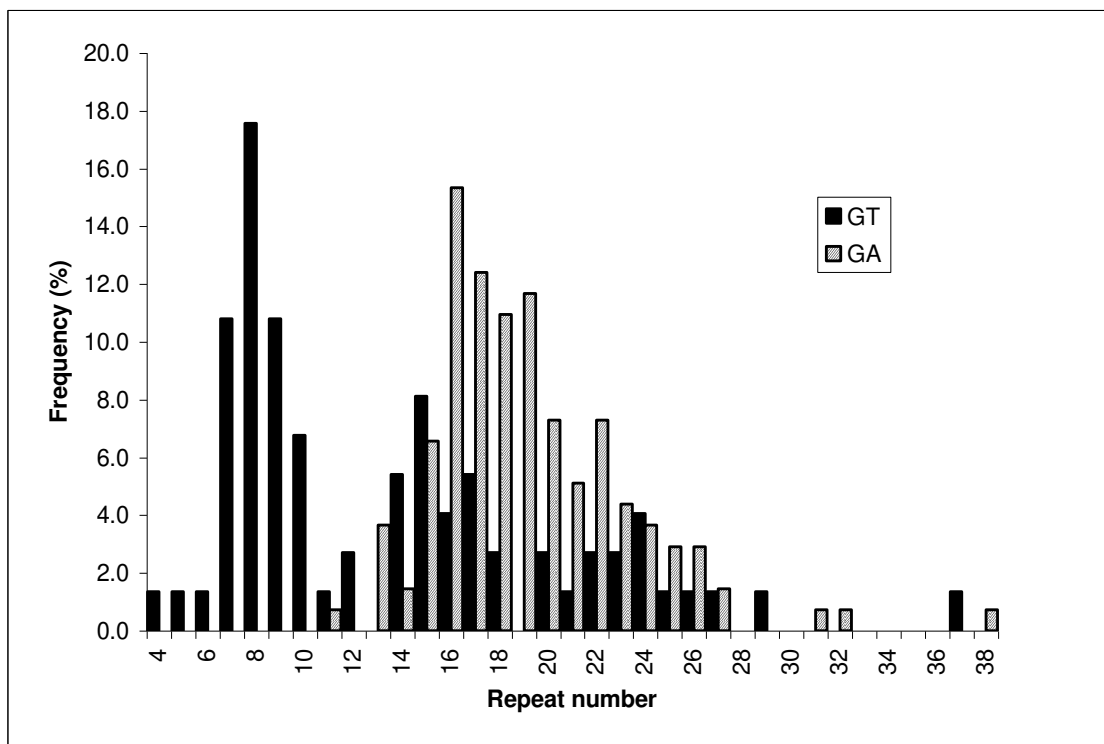


Figure 1. Frequency of markers according to the repeat number

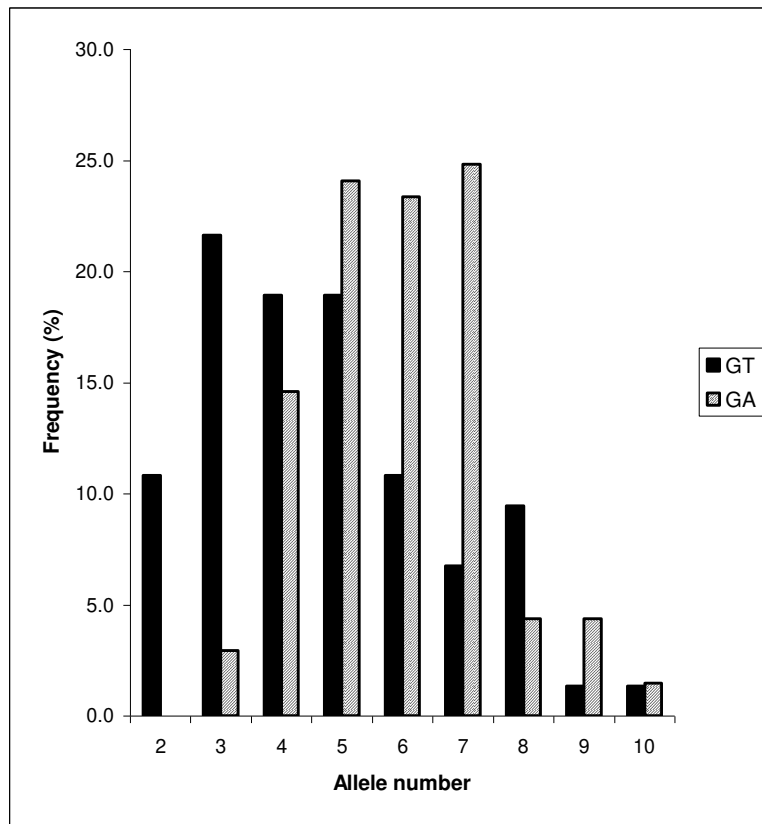


Figure 2. Frequency of markers according to the allele number

Table 1. Polymorphism and range of repeat number in the 310 tested microsatellite markers according to the nature of repeat number

Repeat motif	Number of markers		Range of repeat number
	Tested	Polymorphic (%)	
GA	140	137 (98)	11-38
imperfect GA	11	11	
GT	84	74 (88)	4-37
imperfect GT	9	8 (89)	
compound (GT) _n (GA) _p	29	29	16-44
compound (GT) _n (CT) _p	6	6	9-41
other compounds	20	20	13-40
Multiple	9	9	21-40
Tri (GAA)	1	1	5
Tetra (GAAA)	1	1	5

Imperfects are repeat motifs with one or two nucleotide substitutions

Compounds are adjacent repeat motifs of different types

Multiples are non adjacent repeat motifs

For compound markers, the repeat number is the sum of repeats of each motif

Table S1. Characterization of 296 polymorphic microsatellite loci of rubber tree (*Hevea brasiliensis*)

Columns correspond to locus name (Locus), accession number in EMBL databank (Accession), forward and reverse primers, tandem repeat motif (Repeat), number of alleles per locus (N), number of successfully amplified versus tested accessions (A), allele size range in base pairs (SR), expected (H_e) and observed (H_o) heterozygosity

* significant deviation from Hardy-Weinberg equilibrium (P < 0.01)

Locus	Accession	Forward primer	Reverse primer	Repeat	N	A	SR	H _e	H _o
A31	AF383940	TCCTGCCATCCTTATCCT	TTTTGTATTGCCCCAGCCGTGAGT	(GA)22	9	10/10	174-214	0.900	0.900
a36	AY486855	TGTATTTGTTCTAGAATGCTTTCA	AACAACAACAGCATTTTTTCCC	(GT)8(GA)20	5	10/10	262-282	0.511	0.600
a58	AY486857	TTTTGTGGTTTCATTTGGCTTT	GGTCTTGATGATGTTGTCCAGGT	(CA)10(GA)18	5	10/10	159-177	0.726	0.800
a66	AF383941	CCAACGCTACCTTCC	TACACCACAACACTTCC	(CT)18	9	10/10	245-269	0.900	0.900
a87	AY486860	AGAGAGGGAATATGAGCAAGTGA	ATTTCTGGACCATGTTCTTGT	(GA)18	6	10/10	159-185	0.789	0.800
a90	AY486861	TGAACCGACAACAAGAAACAAAC	CCCTCAAAACCCAACTCAA	(GA)12 (GA)7 (GA)8	9	10/10	219-237	0.832	0.600
a95	AY486862	ACAAGACCTGATACCAAGAAACC	GTCTCCAGAAACATGCCTTA	(GA)5 CAGA(CA)12 (GA)14	7	10/10	226-252	0.805	0.800
a104	AY486863	CTGCGCCCTTGCTGGATA	ATTCGCGGCGAGATCGAG	(CT)18	9	10/10	163-197	0.847	0.800
a105	AY486890	CTTTTTGGACCTTGTGTTTATTG	GGTTTTAGTTATGGCTTTGCTGA	(GA)17(GT)9	7	10/10	145-191	0.874	1.000
a120	AY486865	CTTCCCATTAGGCTCTGTCTC	ATTTCTTGATCTCTGGAGGGTG	(GT)8(GA)12	7	10/10	145-191	0.816	0.700
a123	AY486862	CTTCACACTTCGCTTCCAAAC	TGGCCTCAGTAATCATGCAA	(CT)16	5	10/10	177-199	0.795	0.900
a129	AY486867	AGCCAATCATCAGGAATAGTGC	GATACAACCTGGGGGATGGT	(GA)20	6	9/10	273-289	0.745	0.778
a131	AY486868	CCTATATTGAATTGGGCTTTTGA	AAGATGAAGTGATGGAAGACAACA	(CT)25	6	10/10	273-289	0.795	0.900
a137	AY486871	GGCTGTGGCAATAGGATAAC	CTTATCTTTCTCACTAAACCAGCC	(GT)17 (GA)15	7	10/10	185-205	0.737	0.800
a144	AY486873	AAGCAGCGAAAGTTTATACAGCA	TTCTACTACCCAGGCCAGGG	(CT)24	4	10/10	135-155	0.574	0.600
a148	AY486874	TGTGCTTTTGAATAGGGAATCA	TATGGTTTGAGCTGTGTGTGT	(GA)19	5	9/10	249-283	0.791	0.556*
a156	AY486875	GCCATTTGCAGTTTATGTAGA	AACAGATTTGAAGGCAGCTAGAC	(CT)8 (CT)11	7	10/10	232-252	0.774	1.000
a157	AY486876	TCCAGGTAAAGTGACAGCAGTAA	AACCAAACCTCATAGAAAGCAGCAG	(CT)23	3	10/10	180-196	0.468	0.300
a170	AY486878	CTTCACCTATTTTGGGCTGTCT	GGTGATTGCTTCCATTACTTTT	(CT)21	7	10/10	180-196	0.858	0.800
a173	AY486879	ATCAGCAAACTCCAAGAAGC	CGCAGCTACAATGTCCATAAA	(CT)19	7	10/10	247-271	0.821	0.900
a179	AY486881	CAAAGTTGAAAATGTTAAAGGGAA	CATATCAAGAAACACAGCAAAAAA	(CT)19	7	10/10	159-185	0.832	0.900
a183	AY486882	AGAATGGGGCTACTTGTGTTG	AGAACCGCTAGAAGAACCTCC	(CT)25	4	8/10	231-245	0.642	0.625
a200	AY486884	GCTTTCACCTCTCTCGCATCA	ACAAAACAACCGAACACAAACA	(CT)17	5	10/10	204-226	0.632	0.700
a214	AY486886	GAAAAGATGGAAGGAGTGAATGT	TTCTCCCATTTTGTGCTTATCTC	(GA)24	8	10/10	131-167	0.747	0.800
a218	FN552726	TGCTGTAGGTTTGAGTTATTTTG	TTCTTGGTCGGAAGCGT	(CT)19	4	10/10	160-178	0.674	0.600
a220	AY486888	GCACAAAATACTTCCTACACACA	CATAGCACAGGCAGGTCTCTT	(GA)15	5	10/10	165-191	0.758	0.800
a235	AY486559	CTTTATGAGGTTGTAGTTTTAGG	AAAAGCAGGTGGTTCTCAGT	(CA)11 (CA)4 (GA)4 (GA)21	10	10/10	253-279	0.921	1.000
a238	AY486560	TTGCATGAAGAGTTTGAAGTCC	CAGGTTTCGTTTCTTAGTTTCTCA	(GC)8 (GT)5 (GA)14	6	9/10	184-212	0.758	0.667
a241	AY486561	CAGACCACACCAACCTCAAC	CAGGTTTCGTTTCTTAGTTTCTCA	(CT)18	7	10/10	143-181	0.816	0.700
a251	AY486563	TAGACATTCACCCCTTTTACG	TAGTCCTTTTGTGCGGTGC	(GA)20	6	10/10	235-267	0.826	0.900
a262	AY486564	CCAACCTTTTCTCACTCTTTTC	ATTTCTTTGGATTGGATTTGGAT	(CT)21 (CT)11	11	10/10	228-286	0.916	0.900
a265	AY486566	GCAGAAATGGCAACTCAATAGAC	GAAACAATCAAATCCGCATAGG	(GA)15	4	8/10	177-217	0.642	0.750
a268	AY486565	CAAAAAGAAAGGAAAATGGAGA	CAACCATACAGCAAGGCAAG	(GA)25	4	10/10	224-242	0.595	0.667

Table S1. (Continued)

Locus	Accession	Forward primer	Reverse primer	Repeat	N	A	SR	H _e	H _o
a274	AY486602	AACAAGCGATAACTAATACACCA	ACATCCTCATTTTCCAAGTTCA	(GT)17(GC)6 nn (CT)17	5	8/10	129-137	0.767	0.875
a275	AY486567	TCCATCTCTCCTCCATTACCC	GCATTATTGTCATCAAAGAGCCA	(CAT)7(CT)29	8	10/10	145-219	0.884	0.800
a282	AY486569	CACTCCACCCCTTCTTTCTC	AAGTCCATAACATAAAGTCTTCG	(CT)19	4	10/10	98-116	0.721	0.500
a283	AY486570	TCTGTGGTTGTGCAATTTTGC	TGAAGGGTGGAGGGTTTAGG	(CT)18	7	10/10	169-195	0.800	0.800
a288	AY486572	AGTTTTGGGACCTCCTTATTACA	TTTTAGGCTACACTCATTCCACA	(GA)27	6	10/10	203-229	0.832	0.900
a289	AY486573	GGATTGGATATAAGCAGGCTGT	TGTTCCAAAATCCATAGGCAT	(CT)25	7	8/10	153-179	0.867	0.750
a296	AY486575	TCACGCAAGGGGTAATCAC	CTTGAAAGAACTTGTAGTGATG	(CT)26	10	10/10	135-163	0.921	0.900
a297	AY486576	CGAGCACTGAAAAAGATAGA	TGAAAAGAAAGTTATGGATTGGA	(GA)16	6	10/10	187-215	0.763	0.700
a299	AY486577	AAAATTGATGAGTTGGCAGAGAG	TTTTCCATTTCAGTCCGCATA	(GA)19	6	10/10	187-215	0.684	0.700
a301	AY486578	AAAACAAGAAATAAGAATGCAA	GCGTGGACTACTCAGAAAGAAA	(CT)20	3	10/10	237-241	0.584	0.600
a302	AY486579	AAGATTGTGCTTCTCCTC	AGCCTTCAAACCTAACC	(CT)15	5	10/10	152-166	0.442	0.400
A309	AY486580	CCAACGCTACCTTCC	TACACCACAACAACCTCC	(GA)16	8	10/10	174-214	0.858	1.000
a312	AY486892	GGGGTCTGTATTCTTTCA	TGCCTCTTGCTGGATGCTAT	(GA)26	10	10/10	122-176	0.932	0.900
A320	AY486583	GCAATAAGAGCACAGAAGATAG	TTTACTCCAACCATCCAACAT	(CA)3CG(CA)5(GA)19	8	10/10	217-241	0.858	0.800
A324	AY486584	CCCTAGAAATGAAGAGAAAAGA	GAGGAGAGTTGAGAAGGAGTTG	(GA)4 (GA)7 (GA)8	2	10/10	168-180	0.100	0.100
a330	AY486585	GGCAGTAGCACAATCATTTTTAGTA	TTTCCTCACTGTTTTGTCAATTCC	(GA)24	5	9/10	168-180	0.771	0.444
a337	AY486586	ACCAAGTCTAAATCCAACACAA	AACACAAAAGCTAATGAACAA	(GA)22	8	10/10	97-125	0.805	0.700
a365	AY486589	AATATGCAGACACTTGACTCCC	GATTGCACTCTTCTATTGCTCTC	(GA)16 (GA)12	6	10/10	215-237	0.574	0.700
a367	AY486590	CGGTATTCATTTGCTAGATTTCA	CCCATTGAGTATTTAAAGAGAGC	(CT)26 (CA)8	4	10/10	123-171	0.363	0.400
a385	AY486591	AACTCGCCGATAAGGTTTCT	GAATAAAACAACACGACAAGTCA	(CT)22	5	9/10	92-118	0.693	0.778
a394	AY486593	CTTTCTATACTCAATGCCCCAC	TCATGTCCATTTCAATCTTTCC	(CT)17	4	10/10	208-218	0.437	0.100*
a403	AY486594	TCCACTGTCTTGCTCCTGCT	GTGCCCTCTCCCTATTGGTTCT	(GA)23 nn (GGT)4 (GTT)5	6	10/10	133-161	0.579	0.600
a438	AY486596	GACATCTCTTCATCCCTCAAAA	GGGTTGACCACTTCCATACTG	(CT)24 (CA)14	6	10/10	228-250	0.726	0.900
a460	AY486598	ATAGTGAAAGCCCCAACCC	TCACATCAGATCAGTCGTAGA	(CT)22	6	10/10	93-129	0.826	0.800
a463	AY486599	ATCACTTGGAGAGGGAGGG	GCTGCGTTTTAATTGCTGG	(GA)16	7	10/10	178-210	0.800	0.700
a491	AY486601	CTTGACGTTGCGATTCCCT	CATACCCATTTACATACACACACC	(CT)5 AT(CT)14 CAAA(CA)4(CG)4 nn (CA)5 nn (GT)4	6	10/10	136-154	0.684	0.700
A2298	AY486676	TACCTCTGCTAAAATACCC	TACAAAGAGGAGAGCGCA	(CT)15	5	10/10	255-271	0.737	0.700
A2306	AY486662	GTGCCAATCATCTCTTTT	GAGAACAAATTAACAGCCTC	(GA)21	4	10/10	159-187	0.695	0.800
A2316	AY486667	AACACAGAAGAGAAAAGAGG	GACTCCCAACACAGCA	(GA)26	7	10/10	171-193	0.816	0.700
A2317	AY486672	CTCTTTGTTTTCTTCCC	AACCACAAAACCCCA	(CT)19	5	10/10	140-150	0.668	0.800
A2320	AY486663	AAAGAACAAGGGGAAA	CATGATATGGACGATGAA	(GA)18	7	9/10	287-309	0.843	0.667
A2326	AY486674	AGCTTGCACGGGTAT	ACTGGTATGTTGGTGTITT	(GA)15 (GAA)4	6	9/10	301-315	0.758	0.444*
A2329	AY486670	ATTCAAAGAACACACCCCT	GTTCTTTTCTCTTCTTT	(GA)19	6	10/10	196-210	0.779	0.700
A2336	AY486664	TGAGAAGAGTGGAAGATAGA	ATTGAAGCAGCACAAAC	(GA)19	7	10/10	76-96	0.737	0.700
A2348	AY486665	ACCGACTAACTAATTCC	TAAATGCCCTTCAACA	(GA)17	7	10/10	301-338	0.789	0.400
A2355	AY486675	CCAGCCTCATCAAAAG	AGCAGTTGAAGTTGGATT	(GAA)5	2	10/10	177-179	0.189	0.200
A2365	AY486666	GCTATCCATCAGGCAA	CTCCTTCTTTCTCCCC	(CT)16	6	10/10	177-209	0.811	1.000
A2368	AY486668	CCTTCTAATTTCTCAAGGT	CTTATTTTACTGGGCTAATG	(CT)14	5	9/10	173-231	0.739	0.667
A2374	AY486671	ACCGAATCCAGTGCTT	TTTGTGGCTTGATATGTG	(GA)15	7	10/10	259-283	0.805	0.700

Table S1. (Continued)

Locus	Accession	Forward primer	Reverse primer	Repeat	N	A	SR	H _e	H _o
A2381	AY486669	TGGTTGCTTTGTTTCTGT	CAGCTTGCACGGTATG	(CT)18	9	10/10	253-275	0.853	0.800
A2386	AY486677	AGCAAACACCCCTTCT	CGCTTCTCTTCATTCT	(CT)18	5	10/10	162-202	0.758	0.800
A2387	AY486690	GTTCCCTCTTGCTATCTT	GAGGACCCAGGACTTT	(CT)17	4	9/10	176-194	0.627	0.444
A2388	AY486678	CCTTTCTGTGATTGGTG	TGCAAAATTGACCTTCT	(GA)5 (GA)6 (GA)5	3	10/10	103-127	0.279	0.100
A2389	AY486679	CTTTCTTTTGGTCTTTCTC	GCAACTCATCCACCAC	(GA)13 (GAA)5	4	9/10	183-191	0.706	0.778
A2391	AY486680	CTATGGATGTGATTGTGTG	CCCAGAATATTGGCTTAT	(CT)18	6	10/10	239-269	0.811	0.600
A2394	AY486692	ACTTTTGTTTGTCTTCCAC	CGGTAGATTTTGTCCA	(CT)9 (CT)17	5	10/10	199-219	0.653	0.400
A2396	AY486693	GAAGGAAAGTGAAGGGA	CCAACCAATAAAGGAAGA	(CT)6 (CT)16	8	10/10	221-255	0.795	0.800
A2398	AY486694	GGACTATTTCTTGGATGG	GTGGATAGCAAGTGAGGT	(CT)22	5	10/10	188-218	0.442	0.500
A2401	AY486880	TTTCGGGAGATTACAGATTGAG	CACACGGTTCAGGCACAC	(CT)18 (CA)7	7	9/10	93-123	0.784	0.778
A2402	AY486864	GGTATCATGTTATGCCTTACGG	ATGCTCCTATCACTTTCCCTAT	(CT)19 (CA)9	7	10/10	104-136	0.874	0.600
A2404	AY486695	AATGGAAGTCAAAAGGA	GATAAGAGTGAAAGAGACAAC	(CT)13 (CT)7	6	10/10	321-343	0.811	0.900
A2405	AY486696	AAGTTGATGTTAGGAGGAA	GAACTACGTCTGCATTGA	(GA)21	5	10/10	147-173	0.663	1.000
A2406	AY486697	GTCCACAGAAAATAAATCA	AGCCATTTTCTCACCTC	(GA)20	4	9/10	112-146	0.477	0.444
A2410	AY486877	TGCTTTCTTCTTGTCTTTCT	GGTCTTGCAAGTATTGAGGATT	(CT)20	7	10/10	225-247	0.811	0.800
A2412	AY486699	TCCATGAGCAACTCTTC	GTCTTCTACATTCTCTACTGACT	(GA)17	7	10/10	281-297	0.842	0.900
A2413	AY486681	ATCCAAACCTGCTCATAC	GACCCCTATCCAAAGA	(GA)17	7	10/10	287-311	0.637	0.500
A2415	AY486720	GCTTGTCAGGGAGAAA	CAGAGCATAGCAAAACAAC	(CT)19	6	10/10	257-271	0.826	1.000
A2416	AY486682	AAACGACAGGGAAACTAC	GCAACTCTGTCTATCCAA	(CT)18	6	10/10	243-269	0.805	1.000
A2418	AY486701	TTTCTTCTGGCTTTGG	GAACTATGGAAGGGAATG	(CT)15	4	9/10	120-128	0.699	0.444
A2419	AY486702	CTATTGTTGATGTCCTTCC	CATGTCACCATATACCATTT	(CT)17	7	10/10	274-288	0.874	0.900
A2420	AY486703	TGTTAGGACTATGGTTTGAG	GGAGGTGTTTTCTTTGAG	(CT)23	6	10/10	173-203	0.758	0.800
A2422	FN552727	AGTTATCGCCAGTTTCG	CAACCAACAGAAACCAG	(GA)18	4	10/10	238-250	0.489	0.400
A2423	AY486704	CGAGTTCTTGTGCTTTC	GCAGGATTCTCGTTAGTT	(CT)11	5	10/10	269-289	0.742	0.500
A2428	AY486714	CATTCTCTTTGCTTAGGG	GTTGTCCTTCTTATTGG	(GA)13	6	7/10	63-75	0.680	0.889
A2429	AY486715	AATGGCTTCCCAGATAC	AGTAACAGCTTACCAACAAA	(GA)23	7	10/10	217-247	0.882	1.000
A2431	AY486716	TGTTTTCGTTGCTCTTC	TCTGCGTTTCCTCTTC	(CT)27	6	10/10	217-253	0.795	0.700
A2432	AY486717	CACAAGGGACAGATGG	GTAGGGGTGATTAGGGA	(GA)18	5	10/10	81-99	0.632	0.700
A2433	AY486683	CAAAATGGTGGACGCT	TATGTCTGTTGTTGCTTG	(GA)20	7	10/10	86-100	0.800	0.600
A2435	AY486718	ACAGGGAGTCAGGGAG	CGGGGAGATTACAGATAG	(CT)17	6	9/10	234-250	0.719	0.778
A2436	AY486719	GTCCATTCTCATCCCTT	CTATCCTCCTATCGCCT	(GA)10 (GA)9	7	10/10	184-214	0.689	0.400
A2438	AY486558	TTGCTGTTTTCTCTTGG	TAGGCTGGAAACTTGG	(CT)8 CC(CT)8 (CA)5	6	10/10	246-278	0.721	0.900
A2439	FN552728	GCAGACAAGGCGTAGA	CTTGCTTTTGACTGGTG	(GA)16	7	10/10	301-347	0.811	0.700
A2440	AY486707	TTCTTGACCATACTCGG	TGTTTCATGGCTTACA	(GA)16	6	10/10	267-293	0.847	0.900
A2442	AY486708	CTCGCTCAGTTTTACTTCT	GATGAGGGTAGAGTTTGG	(CT)16	6	10/10	147-171	0.684	0.700
A2481	AY486844	TCGATTTCCCTAGATGAG	TTTTGAAGAGTGAAGGAAG	(CT)16	3	9/10	138-144	0.386	0.222
A2482	AY486728	ATGTGAGGATGTTGTTTCT	ATTTGATTATGCTACTGGG	(CT)20	9	10/10	239-285	0.874	0.800
A2483	AY486729	TACATAAATCCAGCCCA	ACTCATCTATTTGTCTACCC	(CT)21	6	10/10	182-210	0.863	1.000
A2484	AY486722	TCTAAACTTCATTGGTGCT	TGCTTGTTCAATTCTAACCT	(CT)17	3	10/10	188-198	0.563	0.600

Table S1. (Continued)

Locus	Accession	Forward primer	Reverse primer	Repeat	N	A	SR	H _e	H _o
A2485	FN552729	TCCCTTGCTCTGGTG	AGGAAAGTAAAGGAGTGAA	(CT)22	5	9/10	317-335	0.614	0.556
A2486	AY486723	CCTCAAACCTAAGCAG	AGTCGGTTTCTTCATT	(GA)18	5	10/10	173-189	0.789	0.700
A2489	AY486730	CGGTATGCTTACGA	GAAATTGAGTTGAAAGGTG	(GA)22	5	10/10	121-139	0.742	0.700
A2492	AY486732	TCTCTGCTAATACCCTCC	GAAGAGTAAAGATGCGGT	(CT)23 (AT)6	6	10/10	170-210	0.737	0.600
A2497	AY486724	TCAGGGCATTAGTCAAG	TGAATGGGGAAGAGG	(CT)16	8	10/10	251-283	0.868	1.000
A2498	AY486734	GCGTTGTGCTATTGTG	TCCATGTTTCTTTTGA	(GA)24	6	9/10	267-305	0.680	0.778
A2500	AY486735	TACCAATCACTCTCCCTT	GTTGTTTGGGTTTGTG	(CT)15 (GT)5	8	10/10	87-105	0.847	1.000
A2503	AY486736	TAGGAATATCATGTCCAAAC	TAACCTGTGTCCCATTC	(GA)23	9	10/10	253-287	0.853	0.700
A2504	AY486737	AACAGGAAACAGAGGAAA	ATGAGTCCATCAAACAATC	(GA)23	7	10/10	111-165	0.711	0.800
A2505	AY486726	AACTCGACCAGGTAAT	GGTAGGATAAGTGTGAAG	(CT)16	7	10/10	149-169	0.816	0.800
A2506	AY486738	TCATACTTTCTGCTCTTTACC	CTCCTCCATCCAGAGATT	(CT)17	4	9/10	199-221	0.654	0.222
A2507	AY486739	GCAGGATTAGGGATAAGA	TAACACCATGCAGAAGAG	(GT)6 (GA)22	7	10/10	288-314	0.742	0.800
A2508	AY486740	ATTAGTGGTGTATTGTGG	ATACCTTGAATGGCTTGT	(CT)21	6	10/10	235-267	0.737	0.600
A2510	AY486742	TGTTTCTCTCTTCATTTCAG	CCTTTACACTCGGGG	(CT)28 (CA)16	5	10/10	209-261	0.700	0.700
A2511	AY486743	CGGAGAAAGAGTATGGT	ACTGTAGACGGATAAGGG	(GA)26	7	10/10	94-128	0.711	0.800
A2518	AY486745	TGGCATCAAATCCTTC	ATGCTTCCCTTGTAAATG	(GA)16	4	10/10	295-319	0.442	0.500
A2522	AY486747	ACCAACTACACAAGGGAC	GGAACACAATCACCAAA	(GA)18	5	10/10	74-94	0.632	0.700
A2523	AY486748	GACCAAACTCACAATCCT	GAACGCAACTTCTTTTG	(GA)38	5	7/10	193-219	0.736	0.571
A2525	AY486582	GAGTTAGAGGTAGAAGCCA	CCTGTATGAAATCAAGAGAA	(CT)13	7	10/10	147-179	0.863	0.800
A2528	AY486749	CTTGTGTCGTGGTATGAA	CAAAACATAAGGGCACA	(GA)17	4	10/10	296-312	0.684	1.000
A2530	AY486750	ACCCCTCTTTAGGTTTGT	CTATTTTCATCGCTTATCTG	(CT)23	6	10/10	269-285	0.768	0.600
A2532	AY486597	CAAAATACCCCTCAGTTT	ATGGAGAAGTGATGAAGAG	(CT)15	6	10/10	90-120	0.632	0.800
A2534	AY486751	AGTTCAAGTTTGTCCATC	TTTTCCCTCCCTTCA	(GA)21	7	9/10	215-241	0.824	0.889
A2535	AY486752	CCTGCCCTTGATTTTCT	CACAGCAAATAGGGTTACT	(GA)13	7	10/10	318-360	0.843	1.000
A2536	AY486753	AATCAGCAGCATCTACAA	CAAGGGATTTTACCTCAC	(CT)21	7	10/10	133-163	0.850	1.000
A2673	FN552730	CTGGAGACGATATGAGGT	CTTGCTGGCTTGTTTAG	(GA)31	7	10/10	147-187	0.695	0.500
A2676	AY486818	AATTGCGAAAGGATACA	CCCTCACAATCACAAC	(GA)17	5	10/10	121-137	0.695	0.500
A2680	AY486819	TCAACCTCGCTTTCAC	CCACAAAATAGCATAGCA	(CT)19	7	10/10	208-224	0.795	0.900
A2682	AY486820	CCCAACAAAGAAAAGAAC	GGAGGACAGTGATAGTGAA	(GA)32	5	10/10	288-318	0.774	0.700
A2683	AY486807	AGCCTTTTGGATGGTT	TGGAAAGTGTGCTC	(GA)17	6	9/10	183-213	0.771	0.778
A2684	AY486821	TATGCGAACAAGGAAAG	ACAGGGATTACACATACAA	(GA)20	6	10/10	243-267	0.637	0.800
A2685	AY486817	ATCTAAAGTGAGGGCATT	TGCACAAGAAGAATTAGG	(CT)16	5	10/10	301-341	0.784	0.600
A2686	AY486808	CTTCCATCTCCTTCCC	GCCTATCCAAAACACATT	(CT)17	6	10/10	118-144	0.747	0.600
A2689	AY486822	CCATTCTTGATACTTCCAC	CAAAACCCACAAAGAAA	(CT)22	7	10/10	69-97	0.642	0.700
A2695	AY486823	GAACAAGCAGCAAGTAAA	AACGAGGAGGAAGCA	(GA)20	7	10/10	296-316	0.842	0.700
A2697	FN552731	ATTTGGTTCGCTGTATG	CTTTGTGGCTCTGTATTTT	(CT)17	5	10/10	201-237	0.742	0.500
A2698	AY486824	GGAAATGGGTCTGTTG	CCTGACTTGATTGTATCCT	(GA)16	5	10/10	124-144	0.824	1.000
A2699	AY486825	TTGTAGTTTGCCTCCT	CATTTCATCCTCCTTCT	(CT)16	7	10/10	122-142	0.853	0.600
A2702	AY486826	CAGTGACCTGGAATGTG	GAAAACAGAGACTTTTGA	(CT)14	5	10/10	228-258	0.700	0.900

Table S1. (Continued)

Locus	Accession	Forward primer	Reverse primer	Repeat	N	A	SR	H _e	H _o
A2705	AY486828	GCTGACCAAGAATCAAA	GTAGACCAGAGGCAGAAC	(GA)17	5	10/10	282-306	0.663	0.800
A2706	AY486829	TAACCTCTTTTCATCATCAC	CTTCGATTCTTTGTCACTT	(GA)24	6	10/10	89-109	0.797	1.000
A2723	AY486835	GCAATAAATGAGTGAGCA	AGAAACGGATAGTGGAAA	(CT)15	4	10/10	269-279	0.647	0.400
A2728	AY486811	CCTCCTTTTCCCCTT	TCGGTTTCTCGCTCT	(CT)22	5	10/10	203-225	0.779	0.700
A2730	AY486838	TGAATAACAGCATCAAGG	CTAATCCCAAGCGAA	(GA)19	7	10/10	282-302	0.805	1.000
A2732	AY486839	CCATAGATGCAAGGCA	AAACACCAAGGAGGAAC	(CT)13	4	8/10	308-318	0.634	0.778
A2734	AY486840	TTACACTCATCTCTACACTCC	CTTGTGATTGAGACACATAA	(CT)16	6	9/10	94-126	0.765	0.667
A2736	AY486841	GCAACCTGATGAATAAAGA	AAATGAGAAACAAGAAGACC	(CT)15	6	10/10	95-113	0.784	0.700
A2737	AY486812	AACATCAAAGGGTTTCAG	CAGAGCGAGGAAGAGTAG	(GA)16	5	10/10	254-276	0.774	0.600
A2739	AY486842	AATAAAGAGCGAGCGA	CAGGGATGCACATAAAC	(GA)16	4	10/10	129-137	0.647	0.400
A2742	AY486845	GAAATGGTCAGTAGGGG	AATCAGTGGCTTCATCTT	(CT)19	6	10/10	217-247	0.747	0.700
A2744	AY486846	ACAGAAGAAGGAAATTAAGG	GCTGAGTTGAGTGGAGTT	(GA)17	5	10/10	272-302	0.763	0.900
A2745	AY486847	AGTCTCACTCTTACAACGTC	TTTGTGCTGTTTCTCCT	(GA)19	4	9/10	137-157	0.542	0.667
A2746	AY486848	TCAGGGCATTAGTCAAG	GAAATCAGGCCACAAG	(CT)19	7	9/10	167-199	0.869	0.889
A2747	AY486849	GTGTGAGCGTGTGTTGG	CATGTAAGCATTGTTGGAG	(CT)18	8	10/10	197-223	0.863	1.000
A2749	AY486851	CTGTAAAGCATCAAGTCA	TTTCTTTTGGTTCTCTCTC	(GA)19	6	10/10	269-301	0.853	0.700
A2757	AY486813	GTGCCATTTTGTGATGT	TAGCCAACCACTCATTT	(CT)16	8	10/10	279-309	0.847	1.000
A2758	AY486814	CCAAACCCAAAGAAA	AAACAAAGGGAGAGGAG	(CT)8 (CT)20	6	10/10	205-225	0.795	0.900
A2762	AY486852	CATGAGCACAAGTTACCA	CAGGGGAATTTAACAAAC	(CT)16	4	10/10	123-139	0.700	0.500
A2763	AY486853	AGTCCTATTCCTTCCCTT	CATGACAATTCCCCAC	(CT)22	7	10/10	241-269	0.853	0.700
A2765	AY486815	TCAAGGTTGATTCTTCT	CACTCATTCTTTTCTCC	(GA)13	5	7/10	203-219	0.758	0.429
A2766	AY486816	AGTTCCCTTGAGTGCTT	CACCCTGGCCTATTC	(GA)16	5	10/10	158-172	0.574	0.600
A2767	AY486854	CACCTTATCTGGTTTTCGT	TGCATGTTGTGTTGCT	(CT)20	5	10/10	187-201	0.800	0.900
t45	AY486791	AGCAACATCCCAATAATCAGAG	AGCATTCACTTTCCATTGAGGTA	(GA)18 nn (GT)19	7	10/10	179-191	0.779	0.700*
t73	AY486792	CCCCCATCTCTTATATAGACACA	GAGATCAGAGACAACTTGACCTG	(CA)21	5	10/10	193-255	0.753	0.700
t93	AY486793	TCTCAATTGCTTTGCTCATTTTA	CAGTTGTCTGTCTGAAGATGCTGT	(CA)27	6	10/10	123-143	0.800	0.800
t97	AY486794	AACCAGCGCAATAACATAAACCC	GTCTCAACTAAGTTTCTGAAAAGC	(GT)24	6	10/10	123-143	0.832	1.000
t140	AY486795	TGAACAATCATTCCTTTCTTTTCC	GGGTTTCATACACAAGTCATTTCC	(GT)10 (GT)9	5	10/10	234-248	0.805	0.800
t182	AY486797	GCCACATACAAATAAAACCAATCT	CCATGAAAAAGTAGTGCTTACATT	(CA)23	8	10/10	137-173	0.884	0.800
t283	AY486798	CCTCCCTCATTTTCTCTCTCTCT	TATTCAGCATTTTCTGTCATTTGTG	(CA)27 (TA)5	9	10/10	207-255	0.905	0.900
t319	AY486799	ATGGGGAGAAAAATAGAGAGAT	CAAACACTGAAAACAACAATGG	(GT)12	6	10/10	185-229	0.805	0.900
T373	AY486800	AAAAGAACAGATTGCCAAAC	ACCAAAGCCAACTCCAAATG	(GT)7 (GT)33	4	10/10	186-222	0.679	0.200*
T415	AY486806	GTGTTAATGCTTGCTTTCT	GCTCCCTTAGTTGTTTGT	(CT)10 (CA)12 CG (CA)6 nn (CA)9	5	9/10	179-191	0.752	0.556
t595	AY486801	GGGCAGGGATAGTGAATTAAG	TAAGAGGAGGAGACGTATTGGA	(GT)24	6	10/10	225-273	0.747	0.600
t686	AY486803	GAAAACTTGCTGCGGAAAC	CAATAATGTCCCAACCAAAA	(GT)37	10	10/10	193-255	0.863	0.900
T1067	AY486889	TATGGCTATGTTCTGCTATGC	GGACTCTTGATGATTTTGGTT	(CA)18	8	10/10	244-264	0.784	0.700
T2085	AY486905	ACTACAAGTCTCATTGCTTT	TTTAGAGGGTGCGGT	(CA)9	3	8/10	187-211	0.451	0.556
T2086	AY486906	AACCAGATAAGGATTTTAGG	ACTCTTGGAATTTCTTTGG	(CA)15	6	8/10	175-195	0.830	0.889
T2091	AY486907	GAAAAGCCAAGCAAAA	CAGCGGTAACAAAGATG	(CA)9	8	10/10	277-315	0.853	0.700

Table S1. (Continued)

Locus	Accession	Forward primer	Reverse primer	Repeat	N	A	SR	H _e	H _o
T2093	AY486908	CATCTCCACCACCTACA	AGGAGAGCAATAACCAAC	(CA)15	7	10/10	277-315	0.842	0.900
T2094	AY486909	CATCTCACTCCCCGT	ATCCCATCACAGAGTTTT	(GT)8 C(GT)7 (GA)12	8	8/10	309-333	0.892	0.750
T2096	AY486910	GAATCCATTAAAGCCCA	GCTCTTTTCACACATTTTAC	(CA)10	4	9/10	198-208	0.601	0.556
T2099	AY486603	CCAATGCTTAGAGGTGA	TGTGGGTGGATAGGAG	(CT)15 (CA)18	3	9/10	244-248	0.601	0.444
T2101	AY486604	TCCCATTCTTTTGTG	GGAACGATGTGCATTAG	(CA)13 (CA)10	3	10/10	104-146	0.674	0.400
T2103	AY486605	ATCCTACAATGCTCATCC	TGGTGTGCTTCATTTCT	(CA)8 (TA)5	5	10/10	111-123	0.600	0.300
T2106	AY486606	GCCATCCTCCTCTACTC	CTCACCCCATCTTATCAC	(GT)14	3	9/10	306-326	0.601	0.778
T2113	AY486607	ATCTTCAACCTGTTC	CAAGACCAGCAACTTACA	(GT)26	8	10/10	121-159	0.868	1.000
T2115	AY486609	CCACACTTTCCTCTTC	TCGCTGTATCTTGTGCT	(TG)10 (TA)6	3	10/10	165-193	0.674	0.200*
T2127	AY486610	CACCATTGAATTGCTAA	AGCATAGAAACAGCAAAG	(CA)24	3	8/10	180-194	0.508	0.375
T2140	AY486611	GCCATACCTTGCGTC	CACAAAGTCAAAGCACC	(GT)17	5	10/10	93-117	0.653	0.900
T2447	AY486709	ACACTAGGTATGTTGAATGG	CTGTGTCAAGTGGAGGA	(CA)14	5	9/10	243-257	0.667	0.444
T2449	AY486684	TCAAGAAATGCGAACA	GTGAAGTGGGAGCAAC	(CA)10	4	10/10	96-104	0.753	0.900
T2538	AY486802	ATTCCACCCTAGACCAC	CTCCTCAAATCCAGCA	(CA)25	8	10/10	190-226	0.874	0.900
T2542	AY486754	GTCTTCTAACTCCGCTCT	TTCTCTCCCTCCTTCTT	(CA)20	8	10/10	294-340	0.868	1.000
T2587	AY486763	CTTTGGATGACTGAAGG	GAAAGGAAAGAAAGGAGAA	(GT)6	3	9/10	321-331	0.451	0.556
T2588	AY486764	GACATCATCCCAGAAGG	AAGCAAATCATCCAAGAG	(GT)8	4	10/10	153-163	0.647	0.500
T2591	AY486765	CTCAAAGAGGAAGGATTT	GGAGTGTCTGCTTAGTGT	(CT)20 (CA)19	6	10/10	294-314	0.768	0.800
T2595	AY486766	TCCCACATCTATTGACAC	CATCAACGGGCTTTT	(CA)16 (TA)4	5	9/10	260-274	0.739	0.333*
T2597	AY486767	GGAATACAGCCCAAGA	TGATAAATGTGGTGTGGT	(GT)8 (GC)4 (GT)6 GCGT(GA) 17	5	10/10	273-293	0.679	1.000
T2603	AY486768	TAGCAGAAGCAGTTATGG	TTATCTATTGGACTGAAGGA	(CT)19 (CA)18	9	10/10	282-322	0.895	0.700
T2605	AY486769	GATTGAGCAGGAAGGA	AAACCTTGAGGATGTAATG	(CT)15 (CA)13	5	10/10	145-165	0.721	0.500
T2607	AY486770	CATTTGTAACTTTTCCC	CGACAACCATATCCTCA	(GT)8	3	9/10	72-78	0.529	0.333
T2611	AY486771	GAGCTACTATGGTTTGAGG	ATGGGTGTGATTTGCTA	(CA)7	6	10/10	277-301	0.811	1.000
T2612	FN552732	TGTGCTTGCTGTAGGT	TTTTCTGATGGTGAGTTTAG	(CT)15 (CA)16	8	10/10	141-171	0.879	0.800
T2617	AY486772	CAACCTGGCTCAATACA	GAGGACGGGGAGTAA	(GT)8 (GC)6 nn (GA)14	6	10/10	280-302	0.789	0.500
T2619	AY486773	GGTTTTCAGGTTCCATT	AGGTGTTCTTTGGAGGT	(GT)7 +(T)13	3	9/10	289-301	0.386	0.444
T2620	FN552733	TCCAAGCAAAGAAACAC	GAAACTGATTTGGTCATCT	(CA)7	3	10/10	245-253	0.532	0.300
T2625	AY486775	TGTGAGATTTAACTAATTT	AATGACGGAATCACAG	(GT)8	2	10/10	80-82	0.268	0.300
T2628	AY486776	AAGCAAACAATCAGCAC	GGAAGCAGAACGACC	(CA)15	4	10/10	120-130	0.437	0.500
T2635	AY486777	AGAAGAATAGAAATGGGT	GTATTCACTTTTGGCATT	(GT)5 (GC)6 (GT)7	6	10/10	287-349	0.747	0.400*
T2636	AY486778	GGCTCCCATCAAATAAG	GCAGCATTCCATTAATA	(GT)14	3	10/10	133-137	0.611	0.400
T2639	AY486779	GAATCACTTTACCATAAC	TGCATCCTCTAGTAGC	(CATA)5 (CA)9 (TA)4	4	7/10	317-323	0.442	0.375
T2645	FN552735	GAAAATAACTTCTTGGTGG	AATCAGTGGAGAATATGG	(GT)29	5	9/10	177-233	0.673	0.444*
T2647	AY486781	CAATCCTCATTATTCCAAC	GTTCAATTCCTTGTATGCT	(CA)18	3	9/10	219-227	0.503	0.333
T2648	AY486782	GCCTTTACTTTTCCCC	TGATGCTGTCTTTTGTCT	(GT)22	7	8/10	282-322	0.817	0.875
T2650	AY486784	ATCGTGACTTAGGTAATAATG	ACGAACATAAAGAAAGGAG	(GT)20	5	10/10	215-269	0.668	0.700
T2656	AY486785	GATTATGGCCTTTTACAGA	TTTAATCGCTTGAATG	(CA)15	5	10/10	141-155	0.511	0.500
T2659	AY486787	GCAAAGAAGACAACCA	CAAGAGAAAAGCCGAG	(CA)14	4	8/10	207-215	0.725	0.375

Table S1. (Continued)

Locus	Accession	Forward primer	Reverse primer	Repeat	N	A	SR	H _e	H _o
T2668	AY486789	GTCAAACCTCTAGCACATTTTC	TTCAATCCATTAACCCA	(CA)12	4	10/10	316-324	0.558	0.500
TA2155	AY486626	ATGTTAGTCTTTGCATCGT	CTTTTAGGTTCTTTGTGCT	(CA)9	7	10/10	280-316	0.800	0.800
TA2156	AY486613	TTGGCCTCAAATCAAA	TGCTGTCGGTAAGAATG	(GT)5 (CT)4	2	10/10	301-305	0.100	0.100
TA2158	AY486615	TAAGTCAGCAAAACCAAG	TAACCTCATTGGTGGA	(CA)7 (CA)14	4	9/10	271-305	0.471	0.556
TA2161	AY486616	TGTAGTTTGTGGGACTTG	CCATTTCTTTGCTTTTG	(CA)7	3	10/10	124-128	0.695	0.900
TA2163	AY486617	ATGCAACAGAGTAGGAGA	TCAAAGCAAATGAAGTG	(CT)11CACAAA (CA)7	7	10/10	188-222	0.853	0.800
TA2197	AY486631	CTCCTTTACACAGCCA	CGTCAAAAGATTCCAAA	(CA)5	2	10/10	147-149	0.268	0.100
TA2203	AY486632	GCAGTATAATTCTTTGTTCTTTTC	TGAAGAAGGCTCAACATC	(GT)9 (GT)8	3	8/10	203-215	0.542	0.556
TA2205	AY486651	GGACTAAGCAGAGCCA	TATGGGAGTTAGAAGCAA	(CA)8 C8	3	10/10	87-95	0.584	0.200
TA2252	AY486641	AGCACAAAACCTCTCTTC	ATGTCTGATTGTGTATCCC	(CA)17	6	9/10	222-236	0.719	0.778
TA2284	AY486660	TTTATATCAGTGGAGTTTGG	TCTCTTGCTTAGTTGTAGG	(CA)8	3	9/10	278-288	0.216	0.111
TA2288	AY486650	CCATTGTAGTTGAATAGGAG	ACATTAGCCATCACCAG	(GT)8	3	8/10	225-229	0.608	0.875
TA2452	AY486711	CTTCTACATTTCTTTGC	GCTATTCCAGCTTTT	(CA)23	8	10/10	275-301	0.853	1.000
TA2545	AY486755	AGATCAAACCTACCTGAA	TCGTATGTGAAGGACAA	(GT)8 (GA)22	4	9/10	320-330	0.686	0.667
TA2550	AY486756	GCCTTATATTCTGAGCTTG	GACCTTGTTGAATGGA	(CA)16	4	10/10	200-206	0.721	0.800
TA2553	AY486757	CATCTATAACCCCTCGAAAA	TTGTTGAAAGGCACATC	(CA)6 (CA)8	3	9/10	179-217	0.523	0.222
TAs2170	AY486618	CTCCTTGAACCTAACAAA	CCCACAAATCTGCAC	(GA)16 nn (GT)8	8	10/10	143-165	0.789	0.600
TAs2172	AY486619	AGGAATGCTAAGGGTATG	AGGAGATTGTGGAAGAAA	(CA)17	5	10/10	100-128	0.800	0.400
TAs2173	AY486627	TGAAAAGTGTAGTGGGC	TGGATTCAATTATGTACCTC	(CT)19 (CA)4	3	10/10	175-185	0.679	0.500
TAs2175	AY486620	GAATTGATGCAGGAGG	TACAAGGCAAATGAAGG	(CA)17	5	10/10	186-200	0.742	0.800
TAs2178	AY486628	CTTGATGGATGAAAGCTAC	AGAAGAAAAGGTGGGTG	(GT)8 (GA)14	5	10/10	217-269	0.711	0.900
TAs2186	AY486629	TTTTGTCGCATTTCC	CTTGCTTCGCCAC	(CT)22 (CA)17	4	9/10	193-221	0.699	0.444
TAs2188	AY486623	TCAAATCAGTCCCACC	CCAGCGAGAACTTTTA	(CA)10 nn (GA)17	7	10/10	127-143	0.842	0.700
TAs2191	AY486624	GATGGTGAGGTAAAGTGT	AGACTCAGCAAGAAAGGT	(CA)10	7	10/10	121-151	0.695	0.800
TAs2192	AY486625	TTAGAGAGATGAGGAAATGA	GCAGTTAGGGTTTGA	(CA)9 (CA)9	4	10/10	142-284	0.605	0.400*
TAs2196	AY486630	GGCTGAAATCTTTTGGA	CTTATGCCCATGAACG	(CT)13 (CA)20	6	7/10	96-138	0.783	0.625
TAs2207	AY486652	TCATTCTTACCACCAC	ACCTACAACATTTACCAAGA	(A)13 (CA)8	4	10/10	193-211	0.363	0.400
TAs2212	AY486633	TACAACCAGCGAGATTAC	TCTTTATCTTCTTCAGCC	(CA)8	4	10/10	271-277	0.363	0.400
TAs2216	AY486653	GAACCAAAATACACATAGGG	ATCTCAAAGAAGTGGTGAA	(CA)9	3	10/10	161-179	0.353	0.200
TAs2222	AY486636	TCTTGCTTGCTGTATTTGT	TCTAACTCCCTTATTCCTGT	(CT)18 (CA)12	4	10/10	221-247	0.658	0.700
TAs2224	AY486637	CATCCTTTCTGTTCTGTTTC	AGTTGGTGATTGTGGTGT	(CA)15	2	10/10	187-197	0.337	0.400
TAs2225	AY486638	GATCTCGATGTTTTTGGT	AATTTGAGTGTGTGCTTG	(CA)10	3	10/10	177-191	0.568	0.800
TAs2227	AY486639	ATCAATCCCGTGTCTCTC	CTAATCCTTGCCCTTGT	(GT)7	2	10/10	195-201	0.505	0.400
TAs2230	AY486654	CTATCTCTGCTGGAACATAA	CAATAAAAGCAACCGAC	(CA)7	3	9/10	198-202	0.216	0.222
TAs2251	AY486640	AGATCACCAAGCATTACTC	GAAGGCAAGTCTCAACA	(GT)9	9	10/10	85-127	0.905	0.600
TAs2254	AY486661	TTTGTGAGATGGCAGTAG	CCCCTTGTTCTTTATTTG	(GT)8 (GA)9	3	10/10	222-240	0.416	0.300
TAs2258	AY486642	TTGTTGAGTTTTGTAGAAGG	CACCAGAAGAGGCGA	(GT)22	3	10/10	246-254	0.279	0.300
TAs2260	AY486656	TTTCATGGCCTAAAGTG	CTTGCTAATGTAGTGTGGT	(GT)11 (GA)10 nn (GAAA)4	7	9/10	177-211	0.830	0.333*
TAs2262	AY486643	ATCACCTCCCCATTTTC	TGAAATTCTGTAGCATAGC	(GT)9	5	10/10	221-235	0.511	0.500

Table S1. (Continued)

Locus	Accession	Forward primer	Reverse primer	Repeat	N	A	SR	H _e	H _o
TAs2264	AY486644	CCTATCACCAAAACCAA	AGACCAATCAGAACCTTAC	(CA)11	5	10/10	120-138	0.774	1.000
TAs2265	AY486657	TTCAACAAGTACACATTTTAG	ATTTTCATTTTCTGGG	(CA)8	5	9/10	238-258	0.601	0.556
TAs2270	AY486645	CCCTCTCAAAAGTCAAAA	TAATGGACCCCTACAAAG	(GT)7	4	10/10	92-100	0.574	0.800
TAs2271	AY486646	CAACAAATATGGCTTCTAC	AAGTACATGGCAGCTTAAC	(CA)16	5	10/10	209-217	0.774	0.800
TAs2276	AY486659	CATTTGGTGCATTGATT	GTCCGACAAAAGGAGTT	(CT)18 (CA)15	3	10/10	221-239	0.416	0.300
TAs2277	FN552736	TAACATGCTACTCATCCAAC	CTCATCCTTTGCTATTCTC	(CA)9	3	10/10	139-143	0.279	0.300
TAs2282	AY486648	ATTAAAGTACACAGCAGGTT	ACAGCTAAACTATATGAGGC	(CA)8 (CA)7	4	10/10	296-348	0.616	0.300
TAs2283	AY486649	TGTCATGTTGGGTCTTC	TTGTAAGAAAATCAGGCTC	(GT)4	2	10/10	78-82	0.268	0.300
TAs2455	AY486685	TTTCTCTTCTCTTCTCCTTT	AATGGAACCTCTCTTGCTC	(CT)11 (GT)18	2	10/10	161-165	0.268	0.100
TAs2456	AY486712	GCCAGAAATAAGAAAAGG	GGACCGCTAGTTTGAG	(CT)19 (CA)8	7	8/10	308-330	0.817	0.875
TAs2457	AY486686	AGATTTTGGTTTGCTCC	CACTCCCTTTACATTTCC	(GT)7 (GA)11	3	10/10	280-304	0.195	0.200
TAs2464	AY486687	AGCTTTCTCTTGCTTC	AGGATTTTCAGTAATGTCTACC	(GT)16	6	10/10	199-267	0.789	0.900
TAs2465	AY486688	ATCTCTTCTGTGCAATTCT	TATTGCGTGTGGCTT	(CT)19 (CA)19	4	10/10	183-189	0.563	0.600
TAs2547	AY486758	ACACTTTTCTATATGTCGGT	ACTTACAGCAGGTTTGG	(GT)8	2	10/10	155-157	0.395	0.300
TAs2548	AY486887	ATAATTTGTGCCCTC	TTTGGGTTCCCTTGATT	(CT)13 (CA)8 (CT)9	6	10/10	259-319	0.789	1.000
TAs2557	AY486759	TCTCCATTCTCACTCCA	GTCAATTGTTTTGATGCC	(CT)20 (GT)21	4	8/10	259-319	0.675	0.250
TAs2558	AY486760	ACTCAATACAAGGAAGGT	AGTTCCAAAGGTCGTG	(GT)9 GAATGT (GA)19	8	10/10	208-224	0.847	0.900
TAs2559	AY486761	GCTTGTCTTCTCTCCAC	AGAACCAAAACGCTAATC	(GT)7	4	10/10	258-276	0.363	0.400
TAs2688	FN552737	TGACCACAAACACCAA	TGTATCGTCAAAATCCC	(GAAA)5	4	10/10	298-314	0.537	0.600
TAs2694	FN552738	TTGGTAGCTTAACACGG	AGACACCTAGAAACCCAA	(GT)9 nn (GA)12	3	8/10	274-294	0.307	0.333
TAs2698	AY486894	ATCAAGCAAAGAAACCTC	GTAAGTGCTCAACCTCTGT	(CA)7	5	10/10	314-358	0.711	0.700
TAs2706	AY486895	CCATAGTCAATTAAGCTCA	GATAATTTCAAAGAAAAGGC	(CA)8	4	10/10	187-215	0.647	0.600
TAs2709	AY486896	TTGGTAAGAATGTTGGTG	CCATAAGAGACTCAATCCA	(GT)10	4	10/10	323-337	0.553	0.400
TAs2724	AY486898	ACCAAATCAAGCATC	ACATCATTTTGAGCC	(GT)9	5	9/10	139-157	0.667	0.556
TAs2730	AY486899	GTATGGATTTACCTCATTTT	AGCTGGCTCATGTCTC	(GT)8	2	10/10	69-71	0.505	0.400
TAs2743	AY486900	AGCATTGAAGAAGCCA	GCAGGGAGGTCAAAA	(GT)8	4	10/10	112-122	0.647	0.800
TAs2744	AY486901	CATCCATTCACCCAAC	TTGTCCCGGTTACATC	(GT)8 (GA)8	3	9/10	149-153	0.392	0.444
TAs2746	AY486902	GCCACTCAATTTCTCTTC	AGTAAATCAATGTACAGG	(GT)15	7	10/10	159-193	0.784	0.800
TAs2748	AY486903	AAATTGGGAAGCTGTG	AGTAATCAGCAAAGGGAA	(GT)8	2	10/10	197-199	0.505	0.200

Capítulo V

Capítulo V

“An integrated molecular genetics linkage map in a full-sib family and QTL mapping of traits related to the growth of the rubber tree (*Hevea brasiliensis*) in summer and winter in tropical regions”

Livia Moura Souza

Rodrigo Gazaffi

Camila Campos Mantello

Dominique Garcia

Vincent Le Guen

Antonio Augusto Franco Garcia

Anete Pereira de Souza

An integrated molecular genetics linkage map in a full-sib family and QTL mapping of traits related to growth of rubber tree (*Hevea brasiliensis*) in winter and summer in tropical regions.

Livia M. Souza^{1,*}, Rodrigo Gazaffi^{2,*}, Camila C. Mantello¹, Dominique Garcia³, Vincent Le Guen³, Antonio Augusto Franco Garcia², Anete Pereira de Souza^{1, 4, §}

¹Molecular Biology Center and Genetic Engineering, UNICAMP, Campinas, SP, 13083-875, Brazil

²Department of Genetics "Luiz de Queiroz" College of Agriculture University of São Paulo, Piracicaba, SP, 13400-970, Brazil

³ CIRAD, UMR AGAP, 34398 Montpellier, France

⁴Department of Plant Biology – Biology Institute and Molecular Biology Center and Genetic Engineering, UNICAMP, Campinas, SP, 13083-875, Brazil

*Both authors contributed equally in this work

§Corresponding author

Email Addresses:

LMS: liviamoura31@gmail.com

CCM: camila.mantello@gmail.com

RG: rgazaffi@gmail.com

DG: dominique.garcia@cirad.fr

VLG: vincent.le_guen@cirad.fr

AAFG: aafgarci@gmail.com

APS: anetepsouza@gmail.com

Abstract

Background

The rubber tree (*Hevea* spp.) is the main plant for natural rubber production. It is cultivated in numerous equatorial and tropical countries. Due to genetic and physiological issues, inbred lines of this species are not available, which requires alternative approaches for its characterization, such as performing genetic mapping in full-sib crosses derived from outbred parents. In the present work, an integrated genetic map for the rubber tree was obtained for a full-sib cross saturated with simple sequence repeats (SSRs) and expressed sequence tag (EST-SSR) markers. To study the genetic architecture of traits related to growth under two different conditions (winter and summer), QTL mapping was also performed using the integrated map for height and girth growth *via* composite interval mapping.

Results

The obtained molecular genetic map consists of 284 markers distributed in 23 linkage groups (LG), with a total length of 2793 cM and an average genetic distance of 9.83 cM between adjacent markers. The largest group covers a distance of 228.7 cM (29 markers) and the smallest 2.72 cM (2 markers). In the QTL mapping analysis, a total of 24 QTLs were detected for growth traits in the summer and winter seasons. A comparison between the seasons of the year was also possible. In regard to the height, the QTLs detected during summer were different from those detected during winter, which was also observed for the circumference. Only the QTLs located in linkage group 16 associated to height were stable during both seasons of the year.

Conclusions

Integrated maps are important for outbreed species because they represent the polymorphisms observed in the genitors in a more realistic way. The QTL mapping performed in the present scenario revealed interesting information, such as a dominance effect and the unique segregation patterns that each QTL could exhibit, independent of the flanking markers. The QTLs identified in this work, especially those related to phenotypic variation during winter, can help to reveal factors that may facilitate the comprehension of these traits in future studies of marker-assisted selection, particularly when the objective of a breeding program is to obtain phenotypes adapted to sub-optimal regions.

Keywords

Rubber tree, *Hevea brasiliensis*, microsatellite markers, genetic linkage, F1 population, QTL.

Background

Hevea brasiliensis (Wild. Ex. Adr. de. Juss. Muell. Arg.) is a commercial rubber producing species belonging to the *Euphorbiaceae* family. The genus *Hevea* encompasses 11 species native to the Amazon region (Pires et al., 2002), and although the Amazon provides optimal weather for the rubber tree, attempts to cultivate it in marginal environments have been prompted for several reasons. For instance, in South America, the yield of rubber is limited by a fungus, *Microcyclus ulei*, which causes the devastating South American Leaf Blight (SALB). A promising strategy to thwart *M. ulei* is to develop *Hevea* cultivars in dryer and colder areas (Jacob et al. 1999). Thus, rubber production was extended to many sub-optimal environments

worldwide during the late 1970s, such as northeast India, highlands and coastal areas of Vietnam, southern China and the southern plateau of Brazil (Raj et al. 2005), all of which exhibit various climatic constraints. Although these areas satisfy mostly the basic requirements for rubber, they also present stressful situations, such as low temperature and dry periods (Priyadarshan and Goncalves, 2003). In addition to the damage and the low growth rate caused by the cold temperatures, the production of latex is halted for a period of 1–3 months each year in these areas (Rao et al. 1998; Jacob et al. 1999).

Rubber breeding has long been based on techniques that involve statistics and quantitative genetic approaches to reveal the best genotypes to be used as new cultivars. Molecular genetic markers have provided new possibilities for characterizing genotypes for the purpose of identifying cultivars, analyzing genetic diversity, establishing relationships between agricultural traits and genetic factors (QTLs), and identifying genes of interest (Besse et al., 1994; Low et al., 1996; Varghese et al., 1997; Lespinasse et al., 2000a; Le Guen et al., 2003; Roy et al., 2004; Saha et al., 2005; Souza et al., 2009; Le Guen et al., 2010; Yu et al., 2011, Mantello et al., in preparation). Although the genetic basis of disease resistance has been studied, less is known about the genetic basis of cold tolerance and characteristics related to growth and plant development in the rubber tree.

Genetic maps are of great use in breeding programs for many species. However, for perennial trees, such as *Hevea* spp., it is difficult to obtain homozygous inbred lines due to their allogamous nature, long juvenile period and inbreeding depression. Therefore, the traditional approaches used to construct linkage maps for F₂ and backcross populations cannot be used in these species (Grattapaglia and Sederoff, 1994). Traditionally, the construction of genetic maps

from a full-sib cross (or F1 populations) has been achieved using the double “pseudo-testcross” strategy, where all markers are considered separately for their heterozygosity in each parent and analyzed as 1:1 segregation pattern. Two parental maps are established and homologous markers segregating in both parents are used to merge the two parental maps in a consensus map. Due to the lack of information from these markers in one of the genitors, two individual maps are constructed (i.e., one for each genitor genotype). In the case of the rubber tree, the first genetic map was generated using a double “pseudo-testcross” strategy (Lespinasse et al., 2000a) and was developed with several markers. This was an important step toward obtaining a saturated genetic linkage for the rubber tree.

In this kind of population, markers that segregate in 3:1 (dominant), 1:2:1 (codominant), and 1:1:1:1 (codominant) ratios can also be found, which can be used to integrate individual linkage map, although only the 1:1 was used for this strategy. Dominant markers provide less information in linkage analyses than codominant markers, and loci that segregate in a 1:1:1:1 or 1:2:1 fashion are highly informative (Wu et al., 2002a). Different numbers of alleles can segregate at each F_1 locus, and different locus configurations should occur. Wu et al. (2002a) proposed an alternative strategy that applies maximum likelihood methods for the simultaneous estimation of linkage and linkage phases, resolving several of the difficulties pointed out by Maliepaard et al. (1997). This method was successfully used in a sugarcane population (Garcia et al., 2006; Oliveira et al., 2007) and in yellow passion fruit (Oliveira et al., 2008), applying a new method for incorporating information from markers that segregate in different ratios.

OneMap software (Margarido et al., 2007) was first developed to facilitate linkage analyses in outcrossing species using the methodology proposed by Wu et al. (2002a). The software was updated to perform multipoint analysis based on the Hidden Markov Method (Margarido et al., 2009), as described by Wu et al. (2002b). This methodology allows the analysis of a mixed set of different marker types exhibiting various segregation patterns, resulting in a genetic map that represents the polymorphisms of the cross in a more realistic way.

While many statistical methods have been specifically developed to map QTLs in outcrossing species, the general double pseudo-testcross method has been widely used to study QTLs in the rubber tree (Lespinasse et al., 2000; Le Guen et al., 2007; Le Guen 2011). Gazaffi et al. (in preparation) presented an approach to perform QTL mapping in an integrated genetic map through Composite Interval Mapping (CIM). Using this methodology, it is possible to infer the segregation of QTLs as well the linkage phase between markers. The method is based on a mixture of models, and multipoint probabilities are also used in this approach.

Here, we present an integrated rubber tree map saturated with simple sequence repeats (SSR) and expressed sequence tags (EST-SSR) and perform QTL mapping using CIM to determine the rate of development during the summer and winter periods to identify growth characteristics in *H. brasiliensis* (height and circumference).

Results and discussion

Marker polymorphism and segregation analyses

Among the 603 microsatellite markers evaluated, 51% (308) showed polymorphism between the parents of the mapping population. The level of polymorphism observed in the present study was expected, as a similar level of polymorphism has been observed in studies by Lespinasse et al. (2000a). Although EST-derived SSRs have been shown to be less polymorphic (39%) than those derived from genomic sequences (54.5%), they are present in gene-rich regions of the genome and are normally abundant (Scott, 2000). All of the 284 identified polymorphic markers were used for genotyping of the 270 F1 offspring in the mapping population.

A chi-squared test was carried out on the genotyped polymorphic loci and showed that 138 (48.6%) loci exhibited a segregation ratio of 1:1; 31 (10.9%) presented a ratio of 1:2:1; and 115 (40.5%) showed a ratio of 1:1:1:1. Of the markers evaluated, 3% (9) showed distorted segregation ($P \leq 0.005$, chi-squared test). The results of the linkage analysis revealed that the markers exhibiting distorted segregation were distributed throughout the genome and were not the cause of distortion in the LGs. The low number of genomic regions showing a skewed segregation in this map is similar to what is observed in other crops, such as melon (Diaz et al., 2011), but differs from what is observed in lettuce (Truco et al., 2007). The detected deviations from Mendelian segregation ratios may be due to various processes, possibly including the existence of gametophyte selection for sub-lethal genes, i.e., genes controlling the viability of

pollen, zygotes or seedlings, putatively located within one or more of these linkage groups (Yan et al., 2005).

Linkage map

Microsatellite markers were used to construct a new genetic map with other genetic populations and statistical parameters. The genome coverage and LGs in the original maps were considerably improved by the new method of analysis and the use of codominant markers. As a result, the map included 284 markers assigned to 23 linkage groups and spanned a total of 2793 cM and the previous map developed to rubber tree is 689 cM smaller than that (Lespinasse et al., 2000a), more informative genetic linkage was determined because it was possible to use makers with segregation ratios of 1:1:1:1, 1:2:1 and 1:1 to constructed a integrated map. Le Guen et al. (2011) and Triwitayakorn et al. (2011) developed a synthetic map for the rubber tree using markers with segregation in both genitors, but it was analyzed as a double pseudo-testcross.

The length of each group varied from 2.7 cM (LG2b) to 228.7 cM (LG10) (Figure 1). The distribution of the markers for each chromosome is indicated in Table 1. Distorted markers are indicated with an asterisk in Figure 1 ($P>0.05$) and the color of the markers indicate the origem of them (in black Le Guen et al., 2010; in red Mantello et al., 2012; in blue Feng at al., 2009 and in green Souza et al., 2009).

The linkage groups were organized according to the numbers from the map developed previously by Lespinasse et al. (2000) and a map developed by CIRAD (unpublished) for the identification of the syntenic markers. In this case, the 23 groups were classified from 1 to 19,

with groups 2, 6, 11 and 17 being further divided into subgroups “a” and “b”. The density of the markers in our map is lower (one marker for every 10 cM) than the density of the map developed by Lespinasse et al. (2000a) (one marker per 3 cM) but this map has been saturated with many AFLP markers. Additionally, the marker distribution among the linkage groups (LG) was not uniform, which generated a number of gaps (LG 11a, 12, 13, 15 and 16). The largest gap, 66.3 cM, was observed in LG 12 (Figure 1).

One possible explanation for the observed gaps, especially for the groups that could not be clustered, may be the lower presence of SSR and EST-SSR polymorphisms in these regions. For example, in LG 6, the published map indicates clusters of markers, primarily AFLP markers, found in some linkage group regions, but it is impossible to join LG 6a and 6b. AFLP clustering has commonly been reported in saturated maps for melons (Diaz et al., 2011), lettuce (Truco et al., 2007), and potatoes (van Os, et al., 2006) and is usually associated with the heterochromatic regions near the centromeres. Although the regions showing AFLP clustering are likely indicative of centromere positions, comprehensive cytogenetic analyses would be necessary to demonstrate this association in the rubber tree.

Another explanation for the gaps found in the map could be that either the recombination events or mapped loci were not evenly distributed throughout the genome. The low density of markers in some of the linkage groups might also correspond to highly homozygous regions that are subject to lower recombination frequencies (Castiglioni et al., 1999).

The expected number of eighteen linkage groups for the linkage map of the rubber tree ($2n = 36$) was exceeded by five linkage groups, all of which showed a low number of markers

per linkage group. Because some linkage groups were small, it is safe to conclude that the apparent excess of linkage groups may be due to incomplete coverage of the genome by the marker loci. Comparison with the maps published by Lespinasse et al. (2000a) and Le Guen et al. (2011), which used several markers, was shown that it is possible to obtain 18 linkage groups. Additionally, in the genetic linkage map developed using only SSR and EST-SSR markers recently published by Triwitayakorn et al. (2011), the number of linkage groups exceeds the haploid number, providing the same results as obtained in the present study, with 23 linkage groups, suggesting that different markers are required to fill the gaps between adjacent markers.

QTL analysis of growth traits in summer and winter

QTL analysis of growth during different seasons of the year was performed using composite interval mapping (CIM). The analyses were conducted separately for each season. The results are summarized in Figures 2, 3, 4 and 5 and in Table 2.

The number of QTLs detected for growth traits in summer was lower than for winter traits. Using a LOD Score of 3 as a threshold, 24 QTLs were identified, with 11 being detected during summer and 13 during winter. In general, the trait exhibiting the greatest number of QTLs was the girth growth (a total of 14 QTLs, including 6 during summer and 8 QTLs in winter). For height, 10 QTLs were detected (5 QTLs in summer and 5 QTLs in winter).

With respect to the girth growth during summer, 6 QTLs were detected, located in linkage groups 5 (21 and 71 cM), 8 (12 cM), 10 (119 cM), 13 (119 cM) and 16 (43.3 cM), together explaining 25.7% of the variation in this characteristic (Figure 2). Among these 6 QTLs,

four (SC.5_i, SC.5_ii, SC10 and SC16) showed a 1:1 segregation pattern, and SC.8 and SC.13 segregated in a 1:1:1:1 and 1:2:1 fashion, respectively. It should be noted that a dominance effect was only observed for the QTLs segregating at a 1:1:1:1 ratio in this case.

During winter, 8 QTLs were found for the girth growth, located in linkage groups 1 (68.16 cM), 2a (45 cM), 6b (7 cM), 8 (15 cM and 57 cM), 16 (43.34 cM and 75 cM), and 18 (23 cM), explaining 37.1% of the variation of this phenotypic characteristic (Figure 3). There was a predominance of QTLs segregating in a 1:1:1:1 fashion (WC.1, WC.2a, WC.8_i, WC.8_ii, and WC.16_ii) and 2 QTLs that showed a 1:2:1 segregation (WC.6b and WC.18). Of the 8 mapped QTLs, 6 exhibited a dominance effect. It should be noted that QTL WC.6b explains 14.3% of this trait, which is the highest R² value observed among all of the QTLs detected in this study.

As for height, 5 QTLs were identified during summer (Figure 4), in linkage groups 5 (194 cM), 8 (164.5 cM), 10 (151cM), 13 (23 cM) and 16 (163.3cM), together explaining 23.15% of the phenotype variation, and their segregation patterns were 1:1, 1:2:1, 1:2:1, 1:1:1:1 and 1:1:1:1, respectively. Observing the QTL located in linkage group 5, which showed a 1:1 segregation pattern, because only the dominance effect was significant. For this trait, 4 QTLs showed a dominance effect among the 5 QTLs detected. It is worth noting that the 2 QTLs with the highest LOD scores (LG 8, at 5.44, and LG 13, at 6.31) each explained approximately 11% of the trait.

In winter, 5 other QTLs for height were identified, in linkage groups 6a (68.4 cM) and 6b (21 cM), 14 (155 cM), 16 (172) and 19 (7 cM). Among these QTLs, a dominance effect was only observed for QTLs WH.6a and WH.14. It can be noted that genitor PR 255 contributed a significant effect for 4 of the 5 detected QTLs. Figure 5 suggests that the QTLs mapped to

linkage groups 6a and 6b may be the same QTL because other studies indicate that these groups are connected. However, in the present dataset it was not possible to link them due to the lack of markers in this region.

Conner et al. (1998) has studied an apple population and its growth during the period juvenile and has identified ten genomic regions associated with growth in terms of height; 4 were detected in two years, linked to 6 QTLs detected in only one year.

A comparison between the different seasons of the year was also performed. In regard to height, the QTLs detected during summer were different from those detected during winter, which can also be observed for the girth growth, for which only the QTLs located in the linkage group 16 were stable during both seasons of the year.

When comparing height and girth growth during summer, we noted that the linkage groups with significant QTLs were the same (5, 8, 10, 13, and 16), indicating a high relation between these characteristics, which was expected because they are both related to plant growth. When the same comparison was performed for winter, we identified some linkage groups that were mapped to QTLs for both characteristics, whereas some QTLs were located in other linkage groups (for height: QTLs WH6a, WH14e and WH19, and for circumference: WC1, WC2a, WC8i, WC8ii and WC18).

A priori, it is not possible to state whether the mapped QTLs for the height and CBH of the plants are, in fact, associated with the same genes or adjacent genes. A study like this requires the utilization of a large number of progenies and saturation regions where QTLs are mapped with molecular markers. However, when the mode of action of coincident QTLs is similar, a pleiotropic effect is a plausible explanation (Conner et al., 1998)

Making a comparison of the regions mapped between different studies on rubber trees is not trivial because the investigated markers and populations of rubber trees are not the same, nor are the traits considered. The QTL mapping studies for rubber trees developed to date have been performed for the purpose of characterizing *Microcyclus ulei* resistance (Le Guen et al., 2003, 2008, 2011). To our knowledge, this is the first QTL mapping study carried out using an integrated genetic map for the rubber tree, which allowed identification of QTLs with different segregation patterns and included the possibility of identifying QTLs with a dominance effect, which was not possible using the other approaches applied previously. In this work, 16 QTLs were detected with different segregation patterns (1:1:1:1 and 1:2:1), which may have contributed to the high R² values obtained.

Genitors PR 255 and PB 217 showed contrasting and complementary phenotypes. Considering the significant additive effects detected, it is possible to verify that both parents contributed equally to the polymorphism in the investigated traits. A total of 29 marginal effects were identified: 13 were associated with PB 217 and 16 with PR 255. The results obtained through QTL mapping provide a first glance into the genetic control of growth characteristics in the rubber tree. Of the 24 detected QTLs, among which 13 (54.1%) are due to dominance effects. It should be noted that additive effects are the main reason for polymorphism in the progeny, but dominance effects play an important role in the genetic architecture of these traits, especially for height in summer and CBH in winter.

QTL mapping provides general confirmation that complex traits in trees are indeed controlled by many genes and that the individual effects of these genes on a phenotype are small. The practical implication of these results is that broad, genome-wide searches will be

needed to account for all of the genes that control complex traits and for all of the segregating variation in populations (Neale et al., 2011).

The QTLs identified in this work, particularly those related to phenotypic variation during winter, can be used as a starting point for investigations of marker-assisted selection, especially when the objective of a breeding program is to obtain phenotypes adapted to regions of escape. These investigations are possible because winter (April until October), in the local where these experiments would be performed, is dry and cold compared to the summer period (Figure 6), which makes it difficult for the fungus to develop. Further investigations in other crosses are required to identify other QTLs affecting frost tolerance in *H. brasiliensis*. It is too early to say that the regions identified here are in fact related to cold resistance, but our findings represent a starting point.

Mapping QTLs in rubber trees is a challenge due to the long developmental cycle of these plants, and the size of the experiment is a limiting factor for the establishment of mapping populations in different environments. Therefore, it is necessary to perform further studies addressing characteristics associated with the development of these plants to select the genotypes that have adapted to escape areas. This was the first work approaching this type of characteristic in the rubber tree.

Methods

Mapping Population

The mapping population consisted of 270 individuals derived from a cross between genotypes PB 217 and PR 255. The first genotype (PB 217) has a high yield potential, which is

expressed throughout the lifetime of its rubber production, and exhibits low metabolic activity and high level of sucrose in its latex vessels. The second (PR 255) shows tolerance to injury and cold. Progeny individuals were obtained through controlled pollination of the two parental genotypes and multiplied by bud grafting onto rootstock (seedlings of GT1 genotype) to generate replicates for the field experiment.

DNA extraction and Molecular marker analysis

Genomic DNA was extracted from 300 mg of lyophilized leaf tissue using the cetyl trimethyl ammonium bromide (CTAB) method, as described by Murray and Thompson (1980). The obtained DNA concentrations were estimated through electrophoresis on ethidium bromide-stained agarose gels using appropriate molecular weight standards.

The polymorphism of a total 603 microsatellite markers from different sources was screened on both PB217 and PR255 genotype parents. For this total, 425 SSR markers were from a enriched genomic library (Souza et al., 2009; Le Guen et al., 2010 and Mantello et al., in preparation) and 178 sequence tag-derived simple sequence repeats (EST-SSRs) markers (Feng et al., 2009).

Two alternative techniques were used for vizualisation of the PCR products, both of then being used separation by electrophoresis on denaturing acrylamide gels. For vizualisation by silver staining, polymerase chain reaction (PCR) was performed as described by Souza et al. (2009). For vizualisation by fluorescence using a LI-COR 4300 DNA analyzer, oligonucleotides with with an M13-19 base extension were added to the forward primers. The PCR amplification mixtures followed the methodology described by Le Guen et al. (2010).

Linkage Map

The linkage analysis was performed using OneMap software (Margarido et al., 2007; Margarido et al., 2009), version 2.0-1, in which the multipoint approach is implemented using HMM models (Wu et al., 2002b). First, cosegregation groups were established using a LOD score of 4.5 and a recombination fraction of 0.4. The order of the markers was obtained through comparison (for groups with up to six markers), and an order algorithm (more than six) was applied (Lander et al., 1987; Mollinari et al., 2009). For the final order, the recombination fraction between the markers was converted into centimorgans using the Kosambi map function (Kosambi, 1944).

Experimental Design and Statistical analysis

The mapping population was planted in the Edouard Michelin Plantation (Mato Grosso state, Brazil, 17° 23' 59.60'' S and 54° 44' 53.93'' W, altitude 519 m) characterized by a subtropical climate with a cold and dry period of four months from May to October (Figure 6). The planting was realized in 2006 using a randomized block study design with four replicates. Each block included was made of 280 elementary plots: 1 plot for each of the 270 progeny individuals, for the 2 parental genotypes and for 8 other cultivars as comparative controls. Each elementary plot consisted of four grafted plants of a same genotype. The total experiment covered 7.2 hectares, including 80 rows and 56 columns, which allowed better control of the environmental conditions. The silvicultural management procedures were the same as those used for commercial purposes.

The circumference at breast height (CBH) and total height were measured to evaluate the development of the individual rubber trees. These growth characteristics were measured in April and October from October 2007 up until October 2009. The traits included in the analysis were the ratio girth growth and the height during the summer (Oct to Apr) and winter periods (Apr to Oct) over two years. The statistical model used for the analysis was as follows:

$$y_{ijuv} = \mu + G_i + B_j + R_{uj} + C_{vj} + e_{ijuv}$$

where y_{ijuv} is the growth characteristic measured on the basis of one season over two years for the i^{th} genotype in the j^{th} block, the u^{th} row and the v^{th} column; μ is the intercept; G_i is the random effect of the i^{th} genotype $\sim N(0, V_g)$; B_j is the fixed effect of the j^{th} block; R_{uj} is the random effect of the u^{th} row within the j^{th} block $\sim N(0, V_{\text{row}})$; C_{vj} is the random effect of the v^{th} column within the j^{th} block $\sim N(0, V_{\text{col}})$; and e_{ij} is the error term $\sim N(0, V)$.

The estimation of the genotypic values was done to perform the QTL mapping in the following step. This analysis was done using the Genstat Software (Paine et al., 2010).

QTL mapping

QTL mapping was performed using the integrated genetic map employing the approach described by Gazaffi et al. (in preparation). The two traits (girth growth and height) were analyzed separately for each season. Briefly, the applied methodology extends composite interval mapping (CIM), as presented by Zeng (1994), for an outbred scenario. The model

consists of three main effects: two additive effects (one for each parent) and one dominance effect. The model considers mixture models and maximum likelihood estimates. Multipoint probabilities were used to determine the QTL genotype. The mapping strategy is based on two steps: first, a scan of the genome is performed to detect QTLs (each 1 cM in this work). In the second step, this mapped region is fully characterized; i.e., significant effects are identified, along with the linkage phase between the markers and QTLs as well as the segregation pattern. The linkage phase between the QTLs and markers reveals the location of alleles that increase or reduce a trait. The segregation pattern is identified based on the significant effects and their magnitude. In diploid progeny, a QTL can exhibit a segregation ratio of 1:1 (only one significant effect), 1:2:1 (two significant effects, which are statistically similar), 3:1 (three significant effects, with all being similar) or 1:1:1:1 (other cases). The details of the methodology are presented by Gazaffi et al. (in preparation). To perform the CIM analysis, the cofactors were included in the model to control the QTLs located outside the mapping region. For this purpose, multiple linear regressions using stepwise selection based on the Akaike Information Criterion were considered, and the window size was 20 cM. To declare a QTL, an LOD score of 3.0 was assumed.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LMS carried out the molecular genetic studies, helped to perform the statistical analysis and drafted the manuscript. **CCM** helped in the field and molecular genetics studies. **DG** coordinated and supervised the field studies. **VLG** participated in the evaluations of the molecular data. **RG** participated in the molecular genetic map construction, performed statistical analysis of the phenotypic data and QTL mapping and helped to draft the manuscript. **AAFG** and **APS** conceived the study, participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors gratefully acknowledge the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) (2007/50392-1) for financial support and scholarships to LMS, CCM and RG, the Conselho Nacional de Desenvolvimento Científico e Tecnológico for research fellowships to APS, AAFG and PSG, and Michelin Ltda, Itiquira-MS, Brazil for producing the biparental population studied in this work available.

References

Besse P, Seguin M, Lebrun P, Chevallier MH, Nicolas D, Lanaud C: **Genetic diversity among wild and cultivated populations of *Hevea brasiliensis* assessed by nuclear RFLP analysis.** *Theoretical Applied Genetics* 1994, 88, 199-207.

Castiglioni P, Ajmone-Marsan P, Van Wijk R, Motto M: **AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution.** *Theor Appl Genet* 1999, **99**: 425–431.

Conner PJ, Brown SK, Weeden NF: **Molecular marker analysis of quantitative traits for growth and development in juvenile apple trees.** *Theor. Appl. Genet.* 1998, **96**: 1027 –1035.

Creste S, Tulmann Neto A, Figueira A: **Detection of Single Sequence Repeat Polymorphism in denaturing Polyacrylamide Sequencing Gels by Silver Staining.** *P. Mol. Biology Reporter* 2001, **19**, 299-306.

Díaz A, Fergani M, Formisano G, Ziarsolo P, Blanca J, Fei Z, Staub JE, Zalapa JE, Cuevas HE, Dace G, Oliver M, Boissot N, Dogimont C, Pitrat M, Hofstede R, van Koert P, Harel-Beja R, Tzuri G, Portnoy V, Cohen S, Schaffer A, Katzir N, Xu Y, Zhang H, Fukino N, Matsumoto S, Garcia-Mas J, Monforte AJ: **A consensus linkage map for molecular markers and Quantitative Trait Loci associated with economically important traits in melon (*Cucumis melo* L.).** *BMC Plant Biol.* 2011, doi:10.1186/1471-2229-11-111.

Feng SP, Li EWG, Huang EHS, Wang JY, Wu YET: **Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*).** *Molecular Breeding* 2009, **23**, 85-97.

Garcia AAF, Kido EA, Meza AN, Souza HMB, Pinto LR, Pastina MM, Leite CS, Silva JAG, Ulian EC, Figueira AVO, Souza AP: **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* 2006, **112**, 298-314.

Grattapaglia D, Sederoff R: **Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers.** *Genetics* 1994, **137**, 1121-1137.

Jacob J, Annmalainathan K, Alam BM, Sathick MB, Thapaliyal AP, Devakumar AS: **Physiological constraints for cultivation of *Hevea brasiliensis* in certain unfavorable agro climatic regions of India.** *Indian J Nat Rub Res* 1999, **12**:1–16.

Kosambi DD: **The estimation of map distances from recombinant values.** *Annals of Eugenics* 1944, **12**, 172-175.

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln Newburg EE: **MapMaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations.** *Genomics* 1987, **1**, 174–181.

Le Guen M, Lespinasse D, Oliver G, Rodier-Goud M, Pinard F, Seguin M: **Molecular mapping of genes conferring field resistance to South American Leaf Blight (*Microcyclus ulei*) in rubber tree.** *Theor Appl Genet* 2003, **108**, 160-167.

Le Guen, V., Garcia, D., Doare, F., Mattos CR, Condina V, Couturier C, Chambon A, Weber C, Espéout S, Seguin M: **A rubber tree's durable resistance to *Microcyclus ulei* is conferred by a qualitative gene and a major quantitative resistance factor.** *Tree Genet, Genomes* 2011, doi:10.1007/s11295-11011-11038-11297.

Le Guen V, Gay C, Xiong TC, Souza LM, Rodier-Goud M, Seguin M: **Development and characterization of 296 new polymorphic microsatellite markers for rubber tree (*Hevea brasiliensis*).** *Plant Breeding* 2010. DOI: 10.1111/j.1439-0523.2010.01774.x

Lespinasse D, Grivet L, Troispoux V, Rodier-Goud M, Pinard F, Seguin M: **Identification of QTLs involved in the resistance to South American Leaf Blight (*Microcyclus ulei*) in the rubber tree.** *Theor. Appl. Genet* 2000b, **100**, 975–984.

Lespinasse D, Rodier-Goud M, Grivet L, Leconte A, Legnat'e H, Seguin M: **A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite and isozyme markers.** *Theor. Appl. Genet.* 2000a, **100**, 127–138.

Low FC, Atan S, Jaafar H, Tan H: **Recent advances the development of molecular markers for *Hevea* studies.** *J. Nat. Rubber Res* 1996, **11**, 32-44.

Maliepaard C, Jansen J, Van Ooijen JW: **Linkage analysis in a full-sib family of an outbreeding plant species: Overview and consequences for applications.** *Genet. Res* 1997, **70**, 237-250.

Margarido GRA, Souza AP, Garcia AAF: **OneMap: software for genetic mapping in outcrossing species.** *Hereditas* 2007, **144**, 78-79.

Mollinari M, Margarido GRA, Vencovsky R, Garcia AAF: **Evaluation of algorithms used to order markers on genetic maps.** *Heredity* 2009, **103**, 494–502.

Murray MG, Thomson WF: **Rapid isolation of high molecular weight plant DNA.** *Nucl Acids* 1980, **8**: 4321-4325.

Neale DB, Kremer A: **Forest tree genomics: growing resources and applications.** *Nature reviews* 2011, **12**, 111-122.

Oliveira EJ, Vieira MLC, Garcia AAF, Munhoz CF, Margarido GRA, Consoli L, Matta FP, Moraes MC, Zucchi MI, Fungaro MHP: **An Integrated Molecular Map of Yellow Passion Fruit Based on Simultaneous Maximum-Likelihood Estimation of Linkage and Linkage Phases.** *Journal of the American Society for Horticultural Science* 2008, **133**, 35-4.

Oliveira KM, Pinto LR, Marconi TG, Margarido GRA, Pastina MM, Teixeira LHM, Figueira AVO, Ulian EC, Garcia AAF, Souza AP: **Functional genetic linkage map based on EST-markers for a sugarcane (*Saccharum spp.*) commercial cross.** *Molecular Breeding* 2007, **20**, 189-208.

Payne RW, Murray DA, Harding SA, Baird DB, Soutar DM: *GenStat for Windows (13th Edition) Introduction.* VSN International, Hemel Hempstead, 2010.

Pires JM, Secco RS, Gomes JI: *Taxonomia e fitogeografia das seringueiras (Hevea spp.).* Belém: Embrapa Amazônia Oriental 2002.

Priyadarshan PM, Gonçalves PS: **Hevea gene pool for breeding.** *Genet. Resour. Crop Evol.* 2003, **50**, 101–114.

Raj S, Das G, Pothan J, Dey S: **Relationship between latexyield of *Hevea brasiliensis* and antecedent environmental parameters.** *Int J Biometeorol* 2005, **49**:189–196.

Rao PS, Saraswathyamma CK, Sethuraj MR: **Studies on the relationship between yield and meteorological parameters of para rubber tree (*Hevea brasiliensis*).** *Agr Forest Meteorol* 1998, **90**:235–245.

Roy CB, Nazeer MA, Saha T: **Identification of simple séquence repeats in rubber (*Hevea brasiliensis*).** *Currente Science* 2004, **87**, 807-811.

Saha T, Bindu Roy C, Nazeer MA: **Microsatellite variability and its use in the characterization of cultivated clones of *Hevea brasiliensis*.** *Plant Breeding* 2005, **124**, 86-92.

Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, Henry RJ: **Analysis of SSRs derived from grape ESTs.** *Theoretical Applied Genetics* 2000, **100**, 723–726.

Souza LM, Mantello CC, Santos MO, Gonçalves PS, Souza AP: **Microsatellites from rubber tree (*Hevea brasiliensis*) for genetic diversity analysis and cross-amplification in six *Hevea* wild species.** *Conservation Genet Resour* 2009, **1**, 75–79.

Triwitayakorn K, Chatkulkawin P, Kanjanawattanawong S, Sraphet S, Yoocha T, Sangsrakru D, Chanprasert J, Ngamphiw C, Jomchai N, Therawattanasuk K, Tangphatsornruang S: **Transcriptome Sequencing of *Hevea brasiliensis* for Development of Microsatellite Markers and Construction of a Genetic Linkage Map.** *DNA Research* 2011, doi:10.1093/dnares/dsr034.

Truco MJ, Antonise R, Lavelle D, Ochoa O, Kozik A, Witsenboer H, Fort SB, Jeuken MJW, Kesseli RV, Lindhout P, Michelmore RW, Peleman J: **A highdensity, integrated genetic linkage map of lettuce (*Lactuca spp.*).** *Theor Appl Genet* 2007, **115**:735-746.

van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ, Caromel B, Ghareeb B, Isidore E, de Jong W, van Koert P, Lefebvre V, Milbourne D, Ritter E, van der Voort JNAMR, Rousselle Bourgeois F, van Vliet J, Waugh R, Visser RGF, Bakker J, van Eck HJ: **Construction of a 10,000 Marker Ultradense Genetic Recombination Map of Potato: Providing a framework for accelerated gene isolation and a genomewide physical map.** *Genetics* 2006, **173**:1075-1087.

Varghese YA, Knaak C, Sethuraj MR, Ecke W: **Evaluation of random amplified polymorphic DNA (RAPD) markers in *Hevea brasiliensis*.** *Plant Breeding* 1997, **116**, 47-52.

Wu R, MA CX, Painter I, Zeng ZB: **Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species.** *Theoretical Population Biology* 2002a, **61**, 349-363.

Wu R, MA CX, WU SS, Zeng ZB: **Linkage mapping of sex-specific differences.** *Genetical Research* 2002b, **79**, 85-96.

Yan Z, Denneboom C, Hattendorf A, Dolstra O, Debener T, Stam P, Visser PB: **Construction of an integrated map of rose with AFLP, SSR, PK, RGA, RFLP, SCAR and morphological markers.** *Theor Appl Genet* 2005, **110**, 766–777

Yu F, Wang BH, Feng SP, Wang JY, Li WG, Wu YT: **Development, characterization, and cross-species/genera transferability of SSR markers for rubber tree (*Hevea brasiliensis*).** *Plant Cell Reports* 2011, **30**, 335-344.

Zeng ZB: **Precision mapping of quantitative trait loci.** *Genetics* 1994, **136**, 1457–1468.

Figures

Figure 1 - Genetic linkage

Genetic linkage map of the rubber tree F1 population (PB 217 x PR 255) developed from EST-SSR and SSR markers distributed among 23 linkage groups.

Figure 2 - CBH during summer

The linkage groups that contained mapped QTLs for the circumference at breast height trait during summer.

Figure 3 - CBH during winter

The linkage groups that contained mapped QTLs for the circumference at breast height trait during winter.

Figure 4 - Height during summer

The linkage groups that contained mapped QTLs for the height trait during summer.

Figure 5 - Height during winter

The linkage groups that contained mapped QTLs for the height trait during winter.

Figure 6 - Temperature and precipitation

Data concerning the average temperature (maximum and minimum) and precipitation during the years 2006 to 2009 in Itiquira-MS.

Tables

Table 1 - Distribution of the markers

The distribution and type of the SSR and EST-SSR markers in the linkage groups.

Table 2 - QTLs for the CBH and height

The QTLs detected by Composite Interval Mapping (CIM) for the height and circumference in summer and winter in a full-sib family of rubber trees.

Detailed legend: Significant (above 3.0) LOD scores are shown in bold.

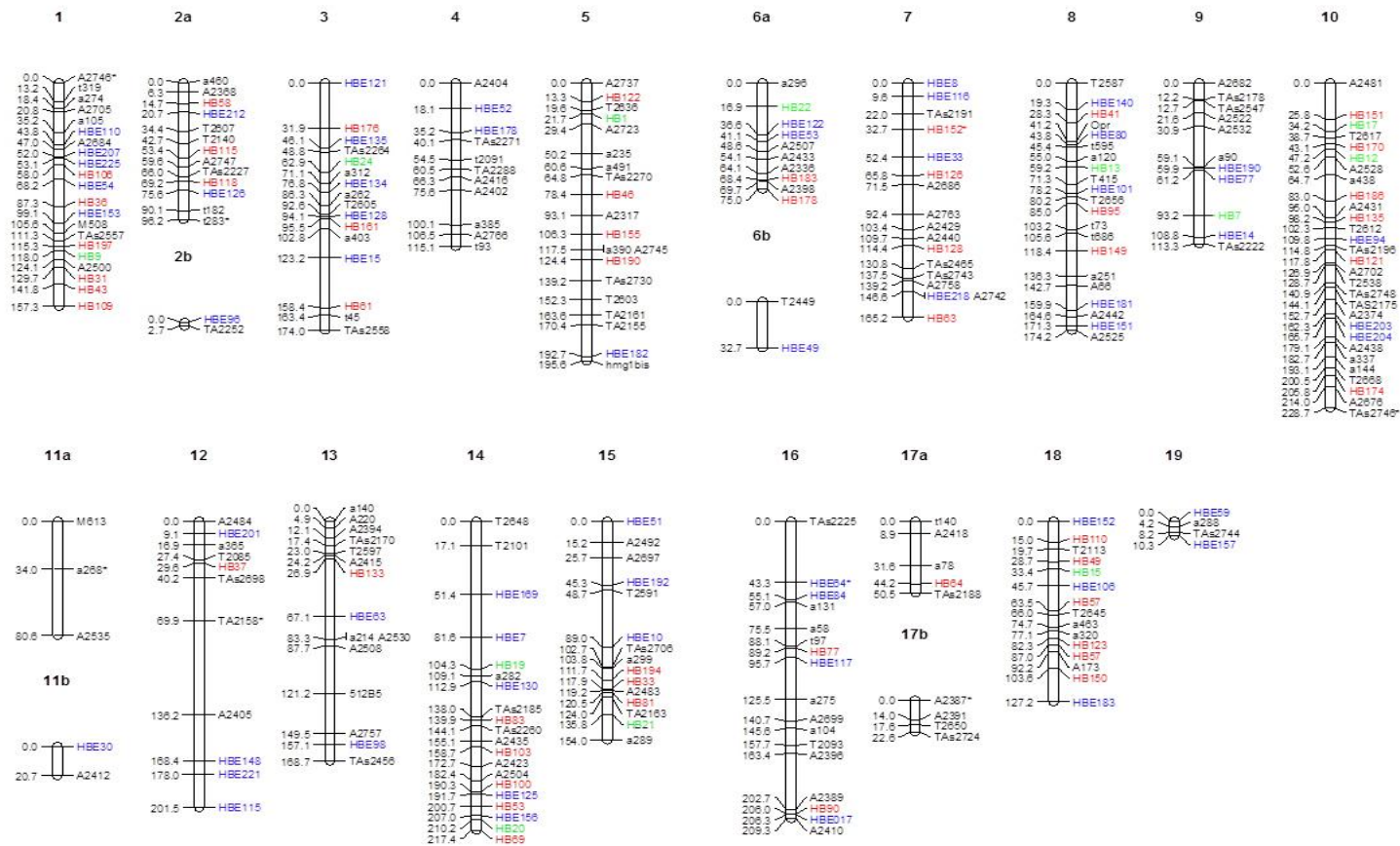
(1) Linkage Group

(2) Distance in relation to the first marker to the left of the respective linkage group.

(3) α_p : additive effect in relation to genitor PB 217; α_q : additive effect in relation to genitor

PR 255 and δ_{pq} : dominant effect.

Figure1. The genetic linkage map of the rubber tree F1 population (PB 217 x PR 255) developed from the EST-SSR (markers in blue) and SSR (markers in red, green and black) markers distributed among the 23 linkage groups.



*markers with distorted segregation

Markers in blue is from Feng et al., 2009, in red is from Mantello et al., in preparation, in green is from Souza et al., 2009 and in black is from Le Guen et al., 2010.

Figure 2. The linkage groups that contained mapped QTLs for the girth growth trait (CBH) during summer.

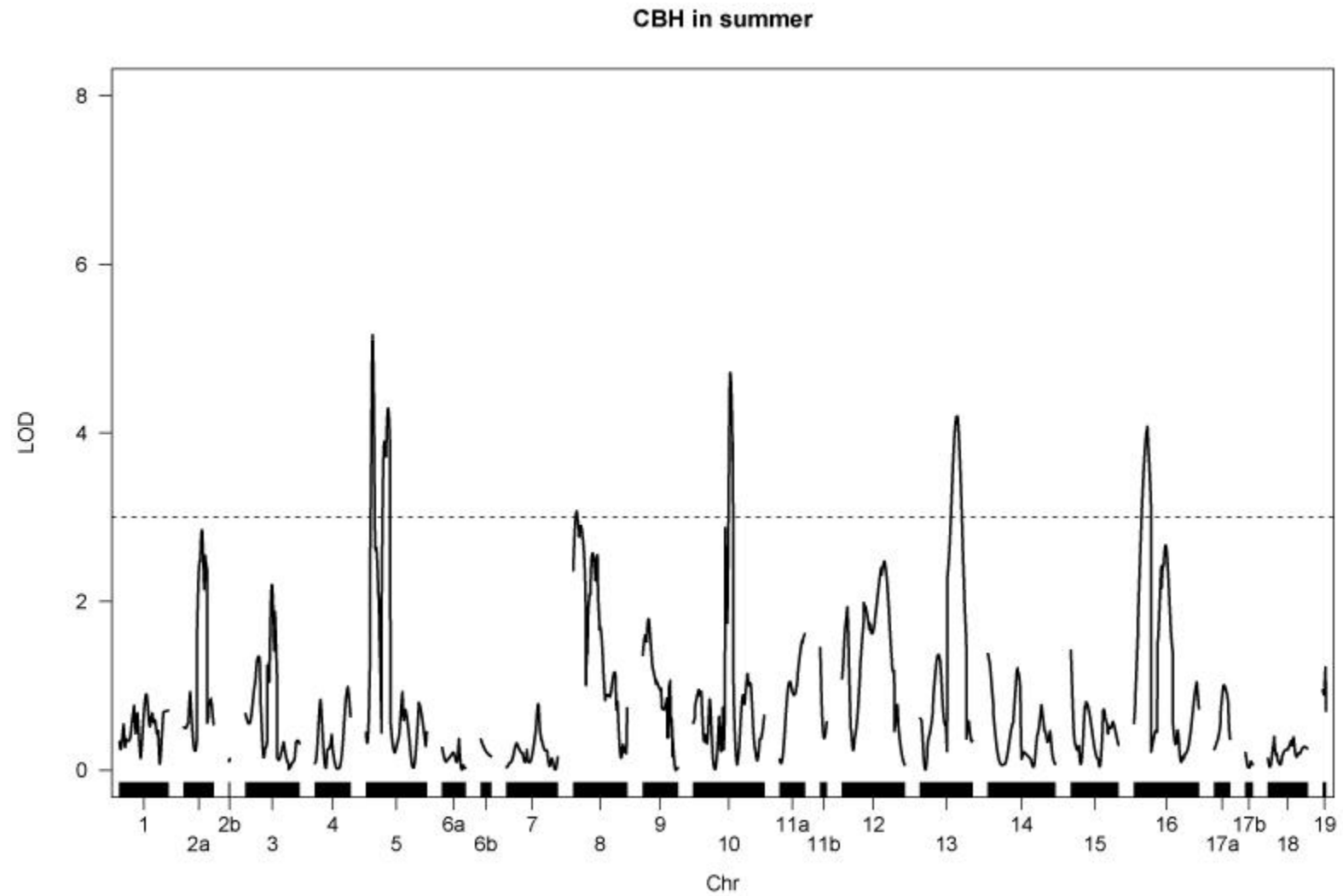


Figure 3. The linkage groups that contained mapped QTLs for the girth growth trait (CBH) during winter.

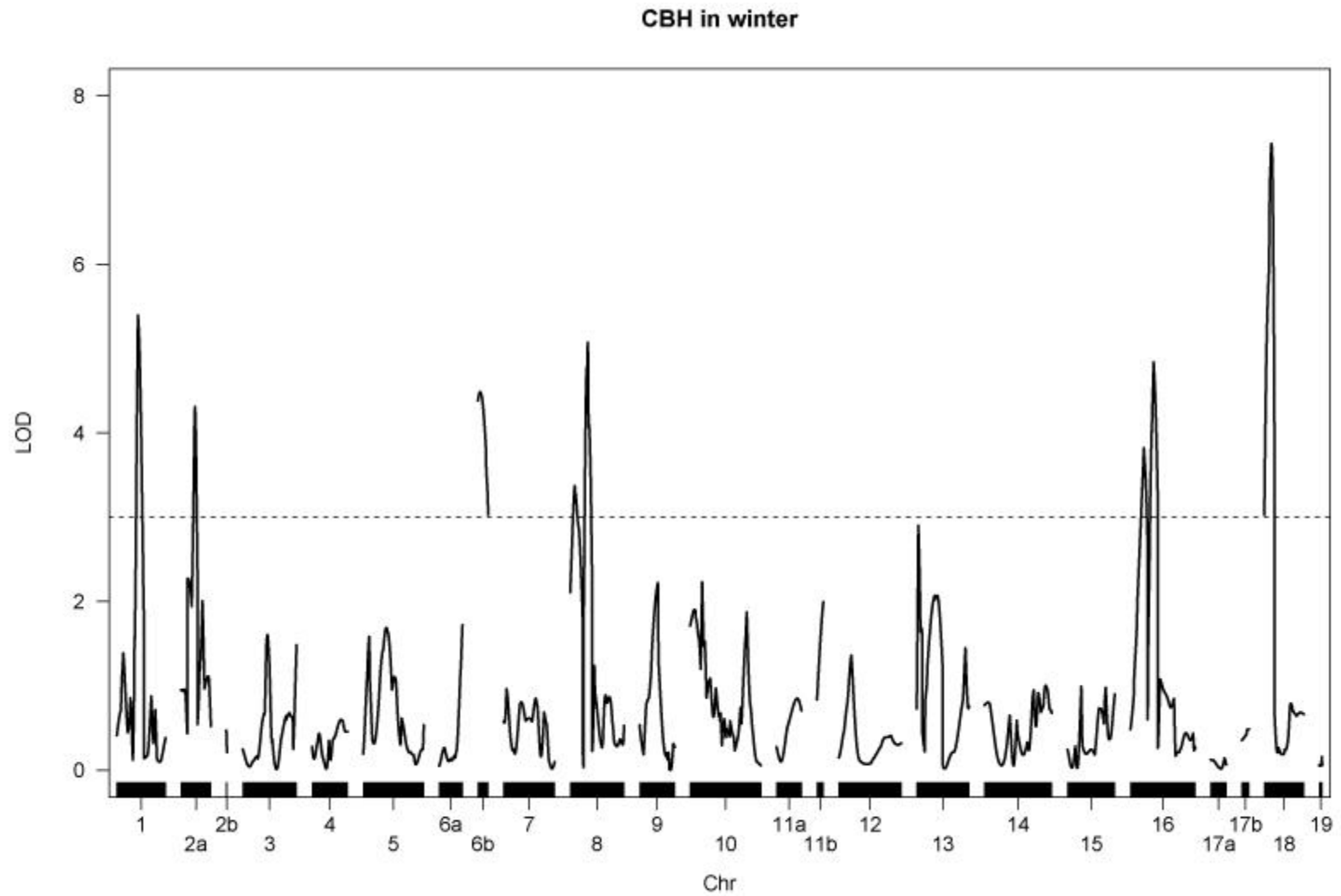


Figure 4. The linkage groups that contained mapped QTLs for the height trait during summer.

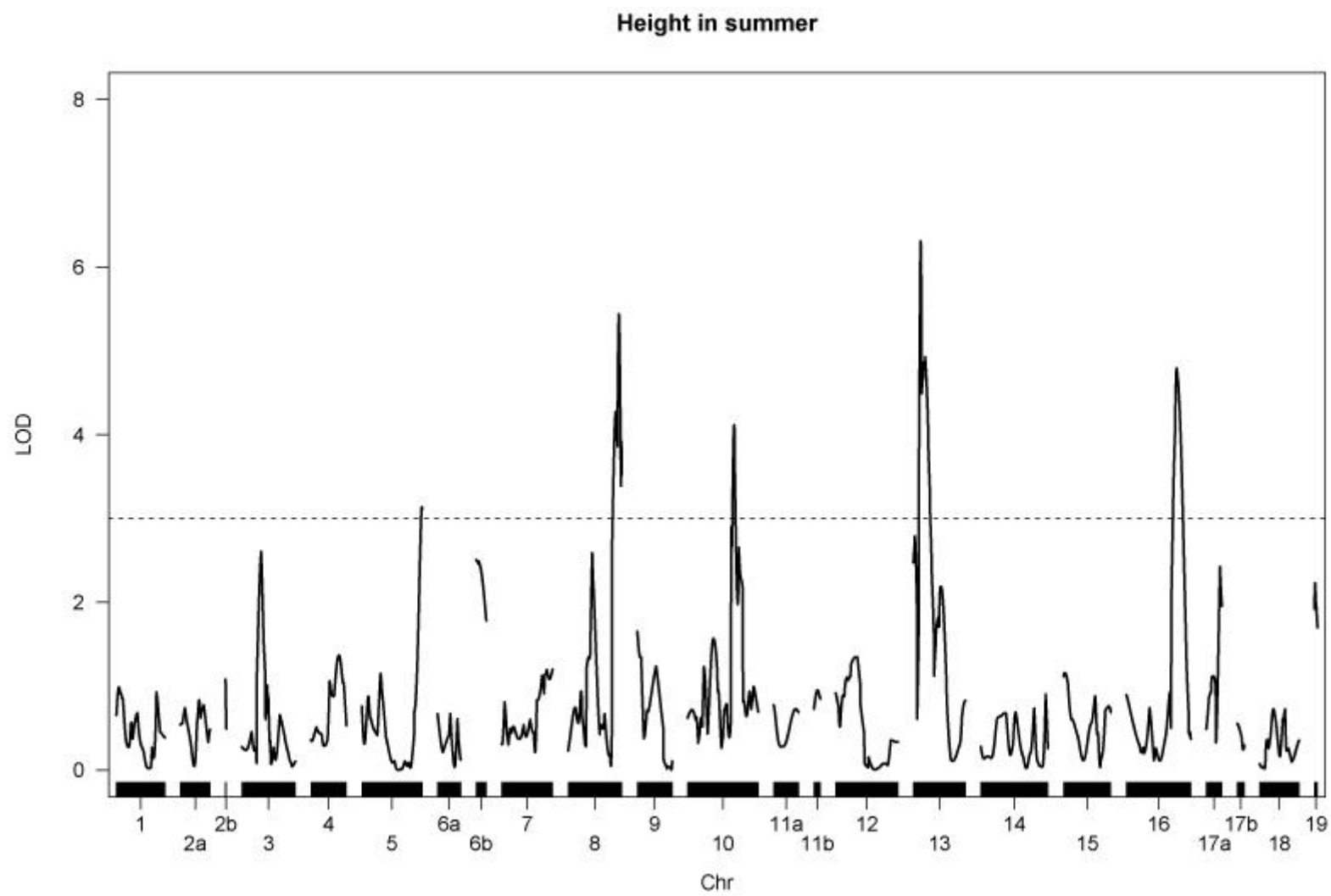


Figure 5. The linkage groups that contained mapped QTLs for the height trait during winter.

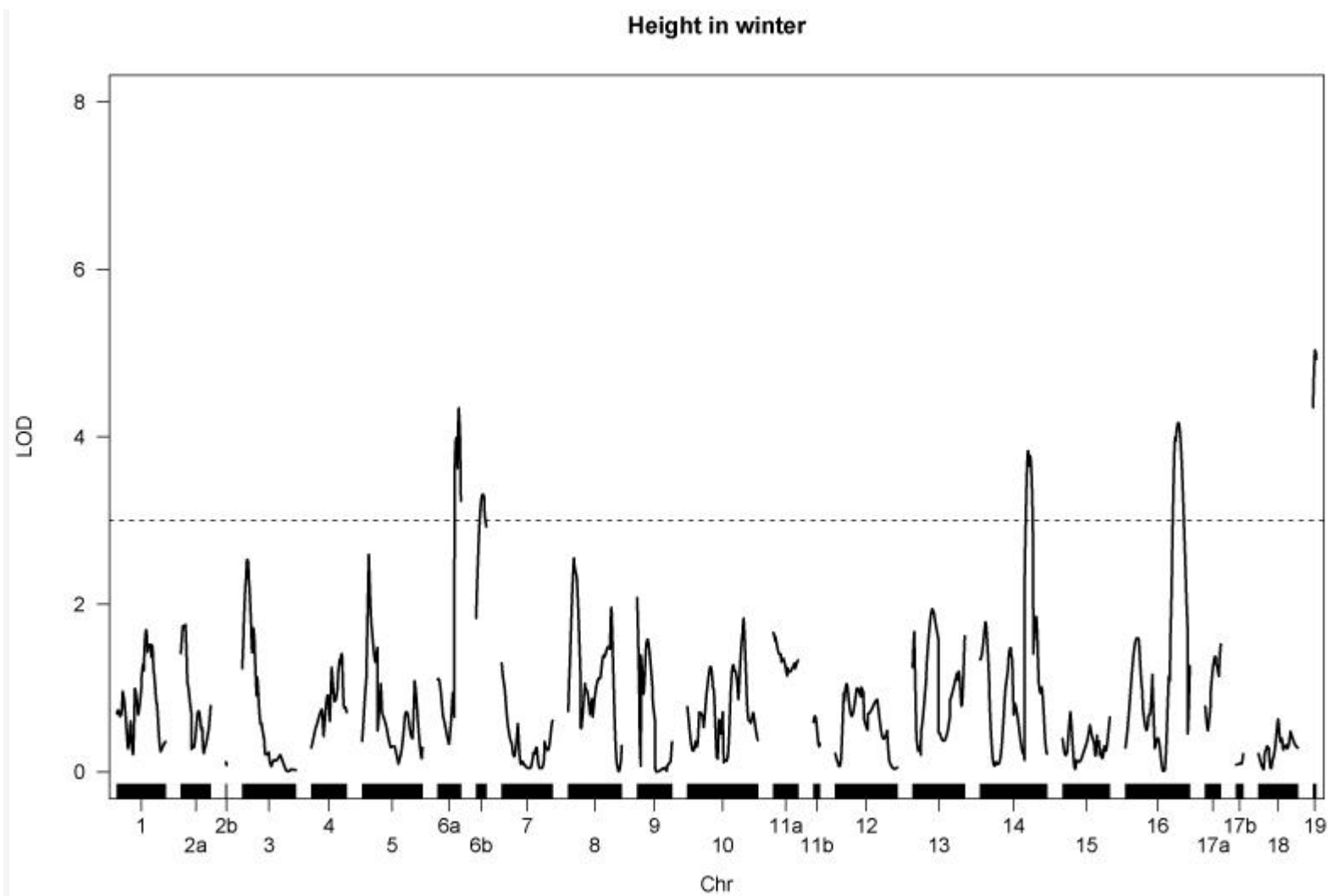


Figure 6: Data concerning the average temperature (maximum and minimum) and precipitation during the years 2006 to 2009 in Itiquira-MS.

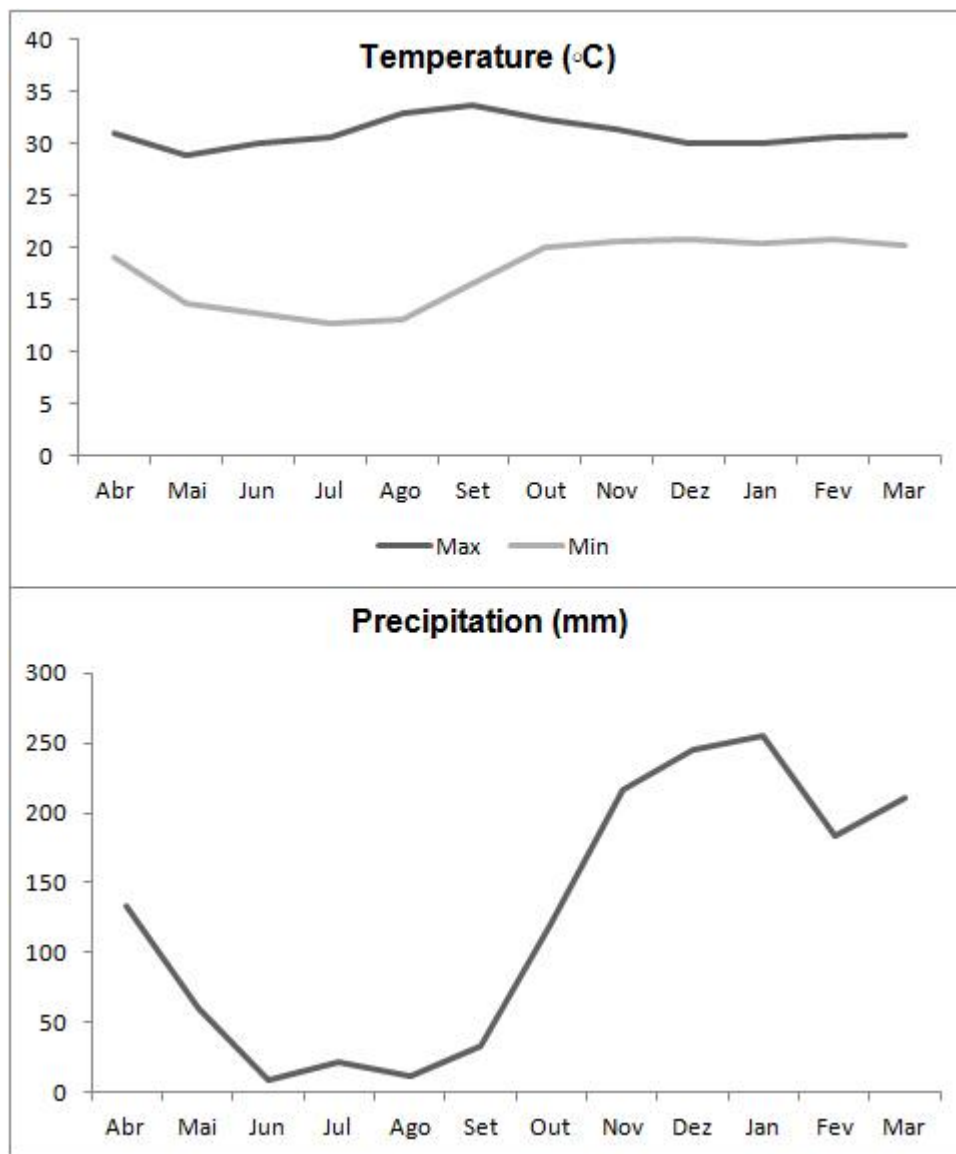


Table 1. The distribution and type of SSR and EST-SSR markers in the linkage groups.

Linkage Group	Biparental		Uniparental	TOTAL	Length	Markers density
	1:1:1:1	1:2:1	1:1			
LG 1	10	1	10	21	157.30	7.49
LG2a	5	1	7	13	96.24	7.40
LG2b	-	2	-	2	2.72	1.36
LG3	5	1	10	16	174.00	10.88
LG4	7	2	2	11	115.10	10.46
LG5	10	1	9	20	195.60	9.78
LG6a	2	1	7	10	74.95	7.50
LG6b	-	1	1	2	32.65	16.33
LG7	6	2	9	17	165.20	9.72
LG8	7	3	11	21	174.20	8.30
LG9	1	1	9	11	113.30	10.30
LG10	14	6	9	29	228.70	7.89
LG11a	3	-	-	3	80.63	26.88
LG11b	2	-	-	2	20.70	10.35
LG12	2	2	7	11	201.50	18.32
LG13	7	-	8	15	168.70	11.25
LG14	8	-	12	20	217.40	10.87
LG15	6	4	5	15	154.00	10.27
LG16	9	2	6	17	209.30	12.31
LG17a	2	-	3	5	50.46	10.09
LG17b	-	1	3	4	22.64	5.66
LG18	7	-	8	15	172.20	11.48
LG19	2	-	2	4	10.25	2.56
Total	115	31	138	284	2837.74	9.99

Table 2. The QTLs detected by Composite Interval Mapping (CIM) for the height and girth growth (circumference) in summer and winter in a full-sib family of the rubber tree.

Trait	Flanking markers	QTL name	LG ⁽¹⁾	cM ⁽²⁾	LOD Global	R ²	α_p ⁽³⁾	LOD	α_q ⁽³⁾	LOD	δ_{pq} ⁽³⁾	LOD	Segregation
Summer Circunference (SC)	T2636-HB1	SC.5_i	5	21.00	5.16	9.70	0.5991	3.62	-0.1385	0.13	0.2451	0.48	1:1
	TAs2270-HV46	SC.5_ii	5	71.00	4.29	9.00	0.0848	0.10	0.4891	3.87	0.1717	0.50	1:1
	T2587-HBE140	SC.8	8	12.00	3.07	8.70	0.0868	0.03	0.2703	1.21	-0.3903	1.88	1:1:1:1
	HB121-A2702	SC.10	10	119.00	4.71	8.75	0.2020	0.81	0.4720	4.00	0.0420	0.04	1:1
	A2508-512B5	SC.13	13	119.00	4.20	9.90	-0.3456	1.80	-0.3867	2.48	0.0161	-0.01	1:2:1
	HBE64	SC.16	16	43.34	4.08	7.34	0.4935	3.55	0.0990	0.09	0.0641	0.00	1:1
						25.74							
Winter Circunference (WC)	HBE54	WC.1	1	68.16	5.39	9.43	0.2716	0.25	-0.9940	3.20	0.7396	1.89	1:1:1:1
	T2140-HB115	WC.2a	2a	45.00	4.31	8.78	-0.7896	2.01	-0.0864	0.03	0.9097	2.53	1:1:1:1
	T2449-HBE49	WC.6b	6b	7.00	4.48	14.35	0.9811	2.39	-1.0657	3.21	0.4064	0.32	1:2:1
	T2587-HBE140	WC.8_i	8	15.00	3.37	10.14	0.8182	1.36	0.3749	0.46	-0.9158	1.78	1:1:1:1
	a120-HB13	WC.8_ii	8	57.00	5.07	9.79	-0.3518	0.33	-1.0823	3.35	0.8400	2.22	1:1:1:1
	HBE64	WC.16_i	16	43.34	3.82	6.94	1.3951	3.73	-0.2642	0.24	0.0677	0.00	1:1
	a131-a58	WC.16_ii	16	75.00	4.84	8.69	-1.0193	2.00	0.9588	2.57	-0.6207	1.04	1:1:1:1
	T2113-HB49	WC.18	18	23.00	7.43	14.38	0.3243	0.33	-0.8905	2.49	1.2459	4.24	1:2:1
						37.15							
Summer Height (SH)	HBE182-hmg1bis	SH.5	5	194.00	3.14	5.67	-0.1179	0.05	-0.3678	0.60	-0.7339	2.37	1:1
	A2442	SH.8	8	164.57	5.44	11.00	-0.5764	1.32	0.6119	1.49	-0.8690	2.79	1:2:1
	TAS2175-A2374	SH.10	10	151.00	4.12	8.29	-1.0157	2.61	-0.4524	0.64	0.6110	0.84	1:2:1
	T2597	SH.13	13	23.04	6.31	10.84	-0.3800	0.43	-0.9922	2.78	-1.1088	4.34	1:1:1:1
	A2396	SH.16	16	163.36	4.79	8.31	0.5888	1.65	-0.8398	3.32	0.1172	0.06	1:1:1:1
							23.15						
Winter Height (WH)	HB183	WH.6a	6a	68.40	4.34	9.02	0.7353	3.14	0.3003	0.56	0.3906	0.99	1:2:1
	T2449-HBE49	WH.6b	6b	21.00	3.31	9.91	0.3231	0.64	-0.7508	2.96	0.3314	0.67	1:1
	A2435	WH.14	14	155.09	3.83	7.86	-0.1872	0.28	-0.5754	2.02	0.5427	1.78	1:1:1:1
	A2396-A2389	WH.16	16	172.00	4.17	9.54	-0.6764	2.45	0.5671	1.83	-0.1521	0.07	1:1:1:1
	a288-TAs2744	WH.19	19	7.00	5.03	9.01	-0.1165	0.08	0.8246	4.45	0.2677	0.47	1:1
							20.75						

The significant (above 3) LOD scores are shown in bold.

(1) Linkage Group

(2) Distance in relation to the first marker to the left of the respective linkage group.

(3) α_p : Additive effect in relation to the genitor PB 217, α_q : Additive effect in relation to the genitor PR 255 and δ_{pq} : dominant effect.

Capítulo VI

Considerações Finais

A biotecnologia e as técnicas de biologia molecular são muito úteis para incrementar a eficiência e rapidez dos programas de melhoramento genético. Até o momento elas ainda têm um desenvolvimento muito tímido e com poucos avanços em termos de aplicação para o conhecimento do genoma e melhoramento de *Hevea brasiliensis*. Neste contexto, nosso grupo de pesquisa iniciou os estudos com seringueira em 2007, em parceria com o grupo de pesquisa do Cirad e do Instituto Agronômico de Campinas, com o objetivo de desenvolver um instrumental molecular que pudesse ser utilizado em estudos genéticos, bem como nos programas de melhoramento da espécie. Foi então iniciado o desenvolvimento de marcadores moleculares do tipo microssatélites, a construção de um mapa genético molecular para a população F₁ deste trabalho, e o mapeamento de caracteres quantitativos associados ao crescimento das plantas no verão e inverno, nesta população de *Hevea brasiliensis*. Esta foi a primeira iniciativa rumo ao desenvolvimento de ferramentas moleculares e estudos de mapeamento genético-molecular para a espécie *H. brasiliensis* no Brasil.

Devido a características como herança mendeliana codominante, multialelismo e distribuição ampla pelo genoma os marcadores microssatélites tornaram-se muito populares na Genética. Entretanto, o desenvolvimento dos marcadores microssatélites para espécies sem genoma sequenciado requer a construção de bibliotecas genômicas. Sendo assim, para o desenvolvimento de marcadores microssatélites para *H. brasiliensis* foi adotada a estratégia de enriquecimento de biblioteca genômica por sondas repetitivas (Billotte et al.,1999). Os resultados obtidos referentes a esta etapa do trabalho foram apresentados no artigo presente no Capítulo IV. A seguir apresentamos resultados complementares deste trabalho.

Foram sequenciados 576 clones de *H. brasiliensis* da biblioteca construída, enriquecida em microssatélites. O tamanho médio dos insertos foi de 650 pb. As amplificações mostraram que houve sucesso na clonagem (Figura 1).



Figura 1 – Padrão de amplificação dos clones de microssatélites com RSA 21.

Para o desenvolvimento de microssatélites adotou-se como critério a seleção de regiões com motivos que apresentassem pelo menos cinco repetições para os dinucleotídeos, quatro para os tri e três para os tetra, penta e hexanucleotídeos. O número médio de repetições variou de três nos pentanucleotídeos a sete nos dinucleotídeos, para seleção dos microssatélites.

Devido ao enriquecimento da biblioteca construída para os dinucleotídeos AG/TC e AC/TG, estas foram as sequências mais abundantes, perfazendo 45% dos microssatélites encontrados. Dentre os diferentes tipos de repetições, os dinucleotídeos foram também os mais abundantes (68%), seguidos pelos tri (13%), compostos imperfeitos (10%) compostos perfeitos (7%) e tetra (1,4%). A literatura disponível sobre microssatélites em *Hevea* cita que as repetições GA no genoma são muito frequentes no genoma de *Hevea*, sendo mais abundantes que CA (Saha et al., 2005). Nossos resultados concordaram com essas afirmações.

Para a caracterização dos marcadores microssatélites foi realizada uma coleta de folhas em fazenda pertencente à APTA/IAC em Votuporanga/SP, cujo objetivo foi obter genótipos contrastantes para a avaliação do polimorfismo dos microssatélites desenvolvidos (Tabela 1).

Em março de 2008 foi feita uma visita a EMBRAPA Ocidental situada em Manaus (AM) com o objetivo de trazer para o Instituto Agrônomo de Campinas (IAC) uma cópia em forma de estacas das espécies de *Hevea* que estão sendo mantidas pela EMBRAPA, esse material foi entregue ao pesquisador Dr. Paulo Gonçalves de Souza (IAC / EMBRAPA) que reproduziu esse material. Foram também coletadas folhas dessas espécies para analisar a transferibilidade dos marcadores microssatélites desenvolvidos em *H. brasiliensis* para as outras espécies (Tabela 1).

Tabela 1. Relação de genótipos coletados em Votuporanga-SP e Manaus-AM em 2008

Genótipos APTA/IAC			Espécies / Manaus (AM)
RRIM 600	PB 233	IRCA 27	<i>H. guianensis</i>
RRIM 606	PB 235	IRCA 209	<i>H. rigidifolia</i>
RRIM 701	PB 311	IRCA 230	<i>H. nítida</i>
RRIM 710	PB 346	IRCA 707	<i>H. pauciflora</i> - 112 CNSG
RRIM 728	PC 140	IRCA 1159	<i>H. benthamiana</i>
RRIM 805	RRIC 100	GT1	<i>H. camargoana</i>
RRIM 809	IAC 306	PB217	<i>H. pauciflora</i> -116 CNSG
RRIM 913	IAC 307	PR255	
RRIM 915	IAC 309		
RRIM 937	IAC 313		
RRII 105	IAC 318		

O desenvolvimento da biblioteca e a caracterização dos marcadores microssatélites tiveram continuidade em um projeto de iniciação científica, processo FAPESP 2007/59804-0, desenvolvido por Camila Campos Mantello.

Para a caracterização dos microssatélites desenvolvidos, foram utilizados na genotipagem 37 diferentes genótipos de seringueira, sendo que 31 deles são de *H. brasiliensis* e seis são de: *H. guianensis*, *H. rigidifolia*, *H. nitida*, *H. pauciflora* (2), *H. benthamiana*, *H. camargoana*, (Tabela 1). Inicialmente a otimização das amplificações foi realizada em PCR com gradiente de temperatura. Porém foi verificado que a maioria dos microssatélites apresentava uma boa amplificação a 60°C, o que levou a um teste inicial a 60°C e, quando este não funcionava, era então utilizado o gradiente de temperatura. Após a verificação da amplificação em géis de agarose 3%, os microssatélites foram analisados em géis de acrilamida 6% (Figura 2).

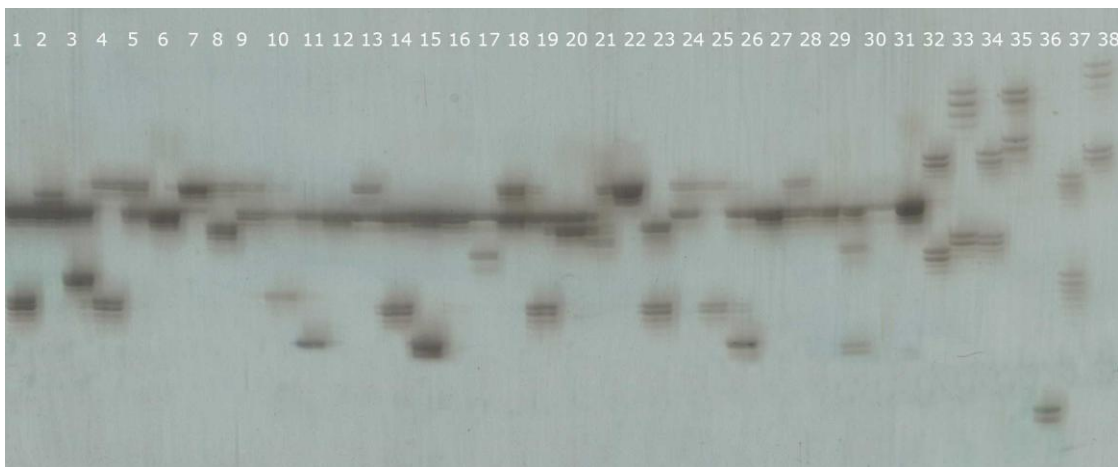


Figura 2 . Microsatélite HB 1 amplificado em gel de acrilamida 6%.

A avaliação da transferibilidade dos microsatélites desenvolvidos para *H. brasiliensis* para outras espécies do gênero não fazia parte de nossos objetivos iniciais. Entretanto, devido à coleta das diferentes espécies de *Hevea*, proporcionada pela parceria com o pesquisador Paulo Gonçalves de Souza e a Embrapa Ocidental, optamos por verificar a transferibilidade desses marcadores. Um indivíduo de cada espécie foi usado nas amplificações com a intenção de estimarmos simultaneamente a qualidade dos produtos de amplificação e o polimorfismo.

A elevada taxa de transferibilidade dos marcadores entre-espécies indica que as regiões flanqueadoras dos microsatélites são bem conservadas entre as espécies de *Hevea*. Segundo Barbará et al. (2009) 40% dos microsatélites são transferíveis para espécies de plantas do mesmo gênero. Em seringueira os marcadores microsatélites caracterizados apresentaram elevada transferibilidade. A espécie *H. carmargoana* obteve 96% de transferibilidade e *H. rigifolia* apresentou a menor porcentagem, sendo esta 79%.

Na figura 2 podemos observar a variabilidade de alelos entre os genótipos analisados. Os indivíduos 1 a 31 representam genótipos de *H. brasiliensis* e do 32 a 38 são genótipos de outras espécies de *Hevea*. A variabilidade observada pode ser explicada pela natureza alógama da espécie. Lidah et al. (2006) verificaram que a taxa de fecundação cruzada é de 96 a 97% variando de acordo com a época do ano.

Segundo Pires et al. (2002), o gênero *Hevea* não pode ser dividido em espécies naturais por causa das variações ecológicas, das formas de transição e dos muito híbridos naturais encontrados. Ainda não existem barreiras reprodutivas bem estabelecidas entre as diferentes espécies de *Hevea*, podendo-se dizer que se tratam de espécies incipientes.

No Capítulo V foi apresentado uma parte dos resultados obtidos durante o estágio desenvolvido no exterior, de maio a julho de 2009, no CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement) em Montpellier – França, no grupo de pesquisa em *Hevea*.

Inicialmente, o objetivo era avaliar o polimorfismo dos marcadores nos genitores da população de mapeamento e genotipar alguns locos nos híbridos. Como tínhamos outros oito genitores de outras populações de mapeamento, esses genótipos foram acrescentados na primeira fase do trabalho, e já foram caracterizados e poderão ser usados em futuros projetos de mapeamento genético.

Os marcadores microsatélites desenvolvidos pelo grupo de pesquisa do CIRAD, não haviam sido publicados e, com essa seleção dos marcadores polimórficos em 10 genótipos contrastantes de *Hevea brasiliensis* foi possível caracterizar esses marcadores e publicá-los, disponibilizando-os para o uso em outros trabalhos científicos.

A porcentagem de marcadores heterozigotos nos genótipos analisados variou de 50,4 a 68,7%. A Figura 3 apresenta a caracterização de seis locos microsatélites em dez genótipos de seringueira. Em destaque estão os genitores PB217 e PR255.

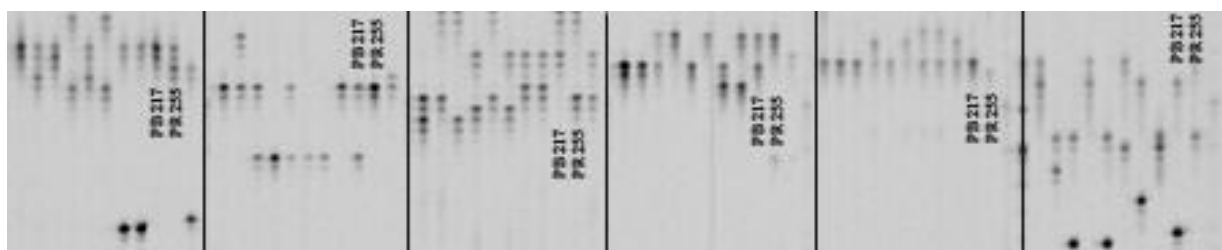


Figura 3. Seleção de marcadores microsatélites nos genitores da população de mapeamento, com destaque para os genótipos PB217 e PR255.

A genotipagem dos híbridos F₁ segregantes foi realizada no sequenciador 4300 LI-COR DNA Analyzer. Foram genotipados 149 locos microssatélites, dos quais 103 apresentaram qualidade de bandas que permitiram boa leitura e interpretação dos géis (Figura 4).

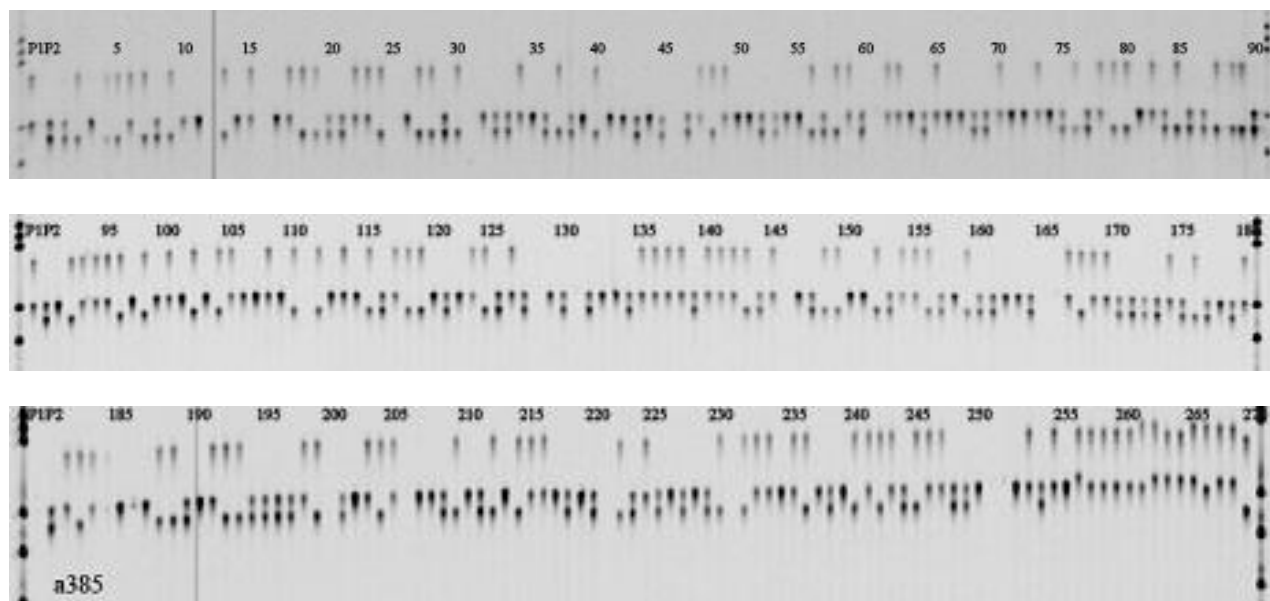


Figura 4. Genotipagem do loco microssatélite a385 na população de mapeamento em gel de acrilamida 6,5% no sequenciador DNA Analyzer – LICOR, onde P1 e P2 correspondem aos genitores PB217 e PR255 respectivamente.

Finalmente, no capítulo VI, com um estudo mais abrangente utilizando os marcadores desenvolvidos pelo nosso grupo de pesquisa, associado a outros marcadores publicados na literatura, foi construído o primeiro mapa genético molecular para seringueira utilizando apenas marcadores microssatélites, bem como realizado o mapeamento de características fenotípicas associadas ao crescimento e desenvolvimento das plantas.

A maioria dos delineamentos genéticos usados para a construção dos mapas genéticos de ligação utiliza populações provenientes do cruzamento entre linhagens endogâmicas (por exemplo, usando retrocruzamentos ou autofecundações). Os métodos genético-estatísticos, nesse caso, já estão bem estabelecidos e implementados em diversos softwares, como por exemplo, o MAPMAKER/EXP. No entanto, para a seringueira, a obtenção de tais linhagens é impraticável,

principalmente em função da grande depressão por endogamia que ocorre quando são efetuadas autofecundações. Nesse caso, as populações de mapeamento são gerações F_1 obtidas a partir de cruzamentos entre indivíduos não endogâmicos.

Para espécies florestais a abordagem utilizando duplo *pseudo-testcross* é amplamente utilizada. Tal estratégia consiste na construção de dois mapas individuais (um para cada genitor), através da identificação de polimorfismos em marcadores de dosagem única para cada parental (Grattapaglia e Sederoff, 1994).

Entretanto, tanto do ponto de vista biológico como estatístico, é desejável a integração das informações contidas nesses mapas individuais em um único mapa integrado. Isso só pode ser feito com a presença de marcadores em heterozigose em ambos os genitores, os quais são utilizados para estabelecer relações de ligação entre os marcadores segregando individualmente em cada parental (Garcia et al., 2006).

Seguindo essa abordagem, e com a disponibilidade do programa Onemap, desenvolvido pelo grupo do professor Antonio Augusto Franco Garcia da Esalq-USP (Margarido et al., 2007) foi possível a construção de um mapa integrado para seringueira utilizando população F_1 segregante.

Para a localização dos QTLs as análises fenotípicas da progênie tiveram início em outubro de 2007 pelos funcionários da antiga Fazenda Experimental Eduard Michelin em Itiquira – MT. As coletas dos dados variaram em número a cada ano, sendo uma coleta em outubro de 2007, duas coletas em 2008, 2009 e 2010 realizadas nos meses de abril e outubro de cada ano. A coleta de dados fenotípicos de outubro de 2009 foi realizada pelos alunos envolvidos nesse projeto, juntamente com os funcionários da fazenda de pesquisa.

O desenvolvimento do mapa de QTL, assim como o mapa genético, utilizou uma abordagem inovadora de análise, baseada no mapeamento por intervalo composto (CIM), desenvolvida pelo Dr. Rodrigo Gazaffi (Gazaffi et al., em preparação para publicação). Nessa metodologia, foi considerado um mapa genético integrado, construído através da abordagem multiponto, com marcadores apresentando diferentes padrões de segregação. O modelo baseou-se em três contrastes ortogonais para estimar os efeitos aditivos e de dominância para os alelos dos QTLs.

Capítulo VII

Conclusões

- As bibliotecas enriquecidas construídas possibilitaram a identificação de novos marcadores microsatélites com elevado polimorfismo para seringueira.
- Os níveis de amplificação heteróloga foram elevados. Sendo assim, os marcadores desenvolvidos poderão ser usados para investigação de espécies aparentadas a *Hevea* e da relação entre elas. Com base nos resultados obtidos é possível investir no cruzamento interespecífico, pois aparentemente o conjunto gênico disponível para o melhoramento da *Hevea* é aparentemente muito maior do que inicialmente imaginado.
- A população de mapeamento apresenta variabilidade genética suficiente para ser explorada em processos de seleção.
- Foi possível construir e integrar o mapa de ligação de seringueira, obtendo-se 23 grupos de ligação. Algumas regiões com baixa frequência de recombinação podem estar relacionadas com regiões de centrômero, e não foram suficientemente saturadas com os marcadores microsatélites. Desse modo fica evidente a necessidade de efetuar a saturação do atual mapa com outros tipos de marcadores.
- Os genitores P217 e PR255 apresentaram em sua maioria QTLs com efeitos favoráveis, com predomínio de ação gênica aditiva. Esse tipo de ação gênica foi detectado pela primeira vez em mapeamento de QTLs para características relacionadas aos desenvolvimento de seringueira.
- O modelo aplicado para o mapeamento permitiu mapear pela primeira vez QTLs com diferentes padrões de segregação.
- Os marcadores e o mapa forneceram simultaneamente novas informações para a espécie e ferramentas para novas abordagens de investigação.

Perspectivas

Apesar dos avanços obtidos neste trabalho de tese, há ainda um imenso caminho para trilhar com relação ao mapeamento de caracteres de interesse agrônômico em *Hevea*, tema ainda muito pouco explorado. Neste contexto, em 2008 foi iniciado um segundo projeto em colaboração com o IAC, que pretende obter duas populações F_1 de *H. brasiliensis* oriundas de genitores que apresentam genes para tolerância ao frio e elevada produtividade de látex (GT1 x PB235 e GT1 x RRIM701) para posterior mapeamento com marcadores moleculares e mapeamento de QTLs para características de importância econômica.

Outras atividades de pesquisa estão sendo estruturadas entre diferentes equipes de pesquisa que visam o estudo da variabilidade genética do germoplasma de seringueira brasileiro *ex situ* e análise do desequilíbrio de ligação. Este projeto apresenta cinco metas principais: (1) coleta de amostras de DNA de seringueiras em coleções *ex situ* existentes no Brasil; (2) Detecção de SNPs no transcriptoma e no genoma de *Hevea*; (3) Mapeamento genético de SNPs nas populações de mapeamento; (4) Genotipagem do germoplasma *ex situ* com os SNPs mapeados e (5) Análise da diversidade genética e caracterização de desequilíbrio de ligação no germoplasma *ex situ*.

A população de mapeamento genético PB217 x PR255 utilizada neste trabalho e disponibilizada pela empresa Michelin Ltda está atualmente ameaçada, pois, após 30 anos de atuação, a Plantações Edouard Michelin de Mato Grosso, do grupo francês Michelin, passou por uma reestruturação radical. Oitenta por cento da área de 10 mil hectares da empresa no município de Itiquira, na região sul do Estado, foram vendidas ao Grupo Maggi. Nesse processo, a Michelin vai parar a exploração agrícola da seringueira na região e manter apenas uma fábrica de beneficiamento de borracha natural e um laboratório de pesquisa na área de seringueira. O anúncio do processo de reestruturação foi feito aos funcionários da empresa em 03/11/2009.

Por enquanto o Grupo Maggi proprietário da área destinada a estudos experimentais, tem um acordo com a empresa Michelin para manter a área experimental com a população de seringueiras avaliadas nesse trabalho por mais alguns anos. Pesquisadores e colaboradores da empresa Michelin informaram que pretendem transferir o experimento da fazenda de Itiquira para uma unidade de pesquisa do grupo Michelin localizada na Bahia, o que possibilitará futuros avanços na pesquisa envolvendo essa população de mapeamento.

Referências

Referências

- Atan, S.; Low, F.C.; Saleh, N.M. 1996. Construction of a microsatellite enriched library from *Hevea brasiliensis*. **J Nat Rubb Res** 11: 247-255.
- Baldwin J.J.T. 1947. *Hevea*: a first interpretation. A cytogenetic survey of a controversial genus, with a discussion of its implications to taxonomy and to rubber production. **J Hered** 38, 54–64.
- Borém, A.; Lopes, M. T. G.; Clement, C. R. 2009. Domesticação e Melhoramento: Espécies Amazônicas. Viçosa, MG.
- Carneiro, M.S.; Vieira, M.L.C. 2002. Mapas genéticos em plantas . **Bragantia**, 61:89-100.
- Cordeiro, I., Secco, R. 2012. *Hevea* in Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro. (<http://floradobrasil.jbrj.gov.br/2012/FB022703>).
- Coelho, A.S.G.; Silva H.D.; Construção de mapas genéticos e mapeamento de QTLs. ESALQ, Piracicaba, 2002, 77p. (apostila).
- Condit, R.; Hubbell, S.P. 1991. Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. **Genome**, 34-1, 66-71.
- Cruz, C.D.; Silva, L.C. 2009. Análise de marcadores moleculares. In: Bórem, A.; Caixeta, E.T. 2009.(Eds.). Marcadores moleculares. Viçosa, 2:371-442.
- Falconer, D.S.; Mackay, T.F.C. 1996. Introduction to Quantitative Genetics. 4.ed. London: Longman Scientific e Technical, 464p.
- Falconer, D.S. 1981. Introduction to quantitative genetics. 2. ed. London: Longman. 340 p.
- Feng, S.P.; Li, E.W.G.; Huang, E.H.S.; Wang, J.Y.; Wu, E.Y.T. 2009. Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*). **Molecular Breeding**, 23:85-97.
- Garcia, A.A.F.; Kido, E.A.; Meza, A.N.; Souza, H.M.B.; Pinto, L.R.; Pastina, M.M.; Leite, C.S.; Silva, J.A.G. Da; Ulian, E.C.; Figueira, A.V.O.; Souza, A.P. 2006. 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** 112: 298-314.
- Gonçalves P.S.; Cardoso M.; Ortolani A.A. 1990. Origin, variability and domestication of *Hevea*. **Pesq. Agrop. Brasileira** 25, 135–156.
- Gonçalves P.S.; Cardoso M.; Boaventura, M.A.M.; Martins, A.L.M.; Lavorentti, C. 1989. Biologia, citogenética e ploidia de espécies do gênero *Hevea*. **O Agrônomo**, 41, 40-64.
- Grattapaglia, D.; Sederoff, R. 1994. Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. **Genetics**, 137, 1121-1137.

- Haley, C.S.; Knott, S. A.; J.; Elsen, M. 1994. Mapping quantitative trait loci in crosses between outbred lines using least squares. **Genetics**, 136:1195–1207.
- Hamada H, Petrino MG, Kakanaga T. 1982. A novel repeated element with ZDNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. **Proceedings of the National Academy of Science USA**. 79: 6465- 6469.
- Jansen, R.C.; Stam, P. 1994. High resolution of quantitative traits into multiple loci via interval mapping. **Genetics**, 136:1447-1455.
- Jeffreys, A.J.; Wilson, V.; Thein, S.L. 1985. Hypervariable ‘minisatellite’ regions in human DNA. **Nature**, 314, 67-73.
- Kahrood HV, Korori SA, Pirseyedi M, Shirvany A, Danehkar A. 2008. Genetic variation of mangrove species *Avicennia marina* in Iran revealed by microsatellite markers. **African Journal of Biotechnology** 7: 3017-3021.
- Kao, C.H.; Zeng, Z.-B.; Teasdale, R. 1999. Multiple interval mapping for quantitative trait loci. **Genetics**, 152:1203-1216
- Lander, E.S.; Botstein, D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, 121:185-199.
- Le Guen V., Gay C., Xiong T.C., Souza L.M., Rodier-Goud M., Seguin M. 2010 Development and characterization of 296 new polymorphic microsatellite markers for rubber tree (*Hevea brasiliensis*). **Plant Breeding**. DOI: 10.1111/j.1439-0523.2010.01774.x
- Le Guen, M.; Lespinasse, D.; Oliver, G.; Rodier-Goud, M.; Pinard, F.; Seguin, M. 2003. Molecular mapping of genes conferring field resistance to South American Leaf Blight (*Microcyclus ulei*) in rubber tree. **Theor Appl Genet** 108:160-167.
- Leitch, A.R.; Lim, K.Y.; Leitch, I.J.; O’Neill, M.; Chye, M.L.; Low, F.C. 1998. Molecular cytogenetic studies in rubber. *Hevea brasiliensis*. Muell. Arg. (Euphorbiaceae). **Genome** 41, 64–467.
- Lespinasse, D.; Grivet, L.; Troispoux, V.; Rodier-Goud, M.; Pinard, F.; Seguin M. 2000b. Identification of QTLs involved in the resistance to South American Leaf Blight (*Microcyclus ulei*) in the rubber tree. **Theor. Appl. Genet.** 100, 975–984.
- Lespinasse, D.; Rodier-Goud, M.; Grivet, L.; Leconte, A.; Legnat’e, H.; Seguin, M. 2000a. A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite and isozyme markers. **Theor. Appl. Genet.** 100, 127–138.
- Litt, M.; Luty, J.A. 1989. A hypervariable microsatellite revealed by in vitro amplification of dinucleotide repeat within the cardiac muscle actin gene. **American Journal of Human Genetics** 44: 398-401.

- Low, F.C.; Atan, S.; Jaafar, H.; Tan, H. 1996. Recent advances the development of molecular markers for *Hevea* studies. **J. Nat. Rubber Res.** 11, 32-44.
- Lynch, M.; Walsh, B. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA. 980p.
- Maliepaard, C.; Jansen, J.; Van Ooijen, J.W. 1997. Linkage analysis in a full-sib family of an outbreeding plant species: Overview and consequences for applications. **Genet. Res.** 70:237-250.
- Margarido, G.R.A.; Souza, A.P.; Garcia, A.A.F. 2007. OneMap: software for genetic mapping in outcrossing species. **Hereditas** 144: 78-79
- Milach, S.C.K. Principais tipos de marcadores moleculares e suas características. In: S. C. K. Milach (Ed.); **Marcadores Moleculares em Plantas**. p.17-28, 1998. Porto Alegre: UFRGS.
- Mullis, K.; Fallona, F. 1987. Specific synthesis of DNA *in vitro* via polymerase catalysed chain reaction. **Methods Enzymology**, 55:335-350.
- Oliveira, K.M.; Pinto, L.R.; Marconi, T.G.; Margarido, G.R.A.; Pastina, M.M.; Teixeira, L.H.M.; Figueira, A.V.; Ulian, E.C.; Garcia, A.A.F.; Souza, A.P. 2007. Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. **Molecular Breeding** 21: 1-20.
- Ong S.H. 1975. Chromosome morphology at the pachytene stage in *Hevea brasiliensis*: a preliminary report. **Proc Intl Rubb Conf.** 2: 3-12.
- Paterson, A.H.; Lander, E.S.; Hewitt, J.D.; Peterson, S.; Lincoln, S.E.; Tanksley, S.D. 1981. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. **Nature**, 335: 721-726.
- Peleg Z., Fahima T., Abbo S., Krugman T., Saranga Y. 2008. Genetic structure of wild emmer wheat populations as reflected by transcribed versus anonymous SSR markers. **Genome** 51: 187-195.
- Pires J.M. 1973. Revisão do gênero *Hevea*: descrição da espécie e distribuição geográfica. Relatório Anual. Belém, Instituto de Pesquisa Agropecuária do Norte (Projeto de Botânica – Subprojeto revisão do gênero *Hevea*. Sudhevea/Dnpea (Ipean), pp 6-66.
- Pires, J.M.; Secco, R. de S.; Gomes, J.I. Taxonomia e fitogeografia das seringueiras (*Hevea* spp.). Belém: Embrapa Amazônia Oriental, 2002, 103p.
- Priyadarshan P.M., Clement-Demange A. 2004. Breeding *Hevea* rubber: formal and molecular genetics. **Adv. Genet.** 52, 51-115.
- Priyadarshan P.M. Gonçalves P.S. 2003. *Hevea* gene pool for breeding. **Genet. Resour. Crop Evol.** 50, 101-114.

Priyadarshan, P.M.; Gonçalves, P.S.; Omokhame, K.O. Breeding *Hevea* rubber. In: Jain, S.M.; Priyadarshan, P.M. (Ed.). **Breeding plantation tree crops**: tropical species. New York: Springer Science, 2009. 469-524.

Rafalski, A. 2002. Applications of single nucleotide polymorphisms in crop genetics. **Current Opinion in Plant Biology**, 5: 94-100.

Ritter E., Gebhardt C., Salamini F. 1990. Estimation of recombination frequencies and construction of RFLP linkage maps in plants. **Genetics** 125: 645-654.

Roy, C. B.; Nazeer, M. A.; Saha, T. 2004: Identification of simple sequence repeats in rubber (*Hevea brasiliensis*). **Current Science** 87: 807-811.

Saha, T.; Bindu Roy, C.; Nazeer, M. A. 2005: Microsatellite variability and its use in the characterization of cultivated clones of *Hevea brasiliensis*. **Plant Breeding** 124, 86-92.

Sax, K. 1923. The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. **Genetics** 8:552-560.

Schultes R.E. 1977. A new infrageneric classification of *Hevea*. **Botanical Museum Leaflets of Harvard University** 25, 243–257.

Schuster, I.; Cruz, C. D. 2004. Estatística genômica aplicada a populações derivadas de cruzamentos controlados. Viçosa , MG: Editora UFV, 568p.

Segun, M.; Besse, P.; Lespinasse, D.; Lebrun, P.; Rodier-Goud, M.; Nicolas, D.; 1996. *Hevea* molecular genetics. **Plantations Rech. Devel.** 3: 77-88.

Seibert, R.J. 1947. A study of *Hevea* in the Republic of Peru. **Annals Missouri Botanic Garden**, 34: 261-352.

Souza, L.M, Mantello, C.C, Santos M.O. , Gonçalves P.S., Souza A.P. 2009. Microsatellites from rubber tree (*Hevea brasiliensis*) for genetic diversity analysis and cross-amplification in six *Hevea* wild species, **Conservation Genet Resour**, 1:75–79.

Sturtevant, A.H. 1913. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. **J. Exper. Zool.** 14: 43-59.

Tanksley, S.D. 1993. Mapping polygenes. **Annual Review of Genetics**, 27:205-233.

Temnykh S., Park W.D., Ayres N., Cartinhour S., Hauck N., Lipovich L., Cho Y.G., Ishii T., McCouch S.R. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). **Theoretical and Applied Genetics** 100: 1432-2242.

Thoday, J.M. 1961. Location of polygenes. **Nature**, 191: 368-370.

Toledo, E.R. Leandro, R.A.; Souza Junior, C.L.; Souza, A.P. 2008. Mapeamento de QTLs : Uma abordagem Bayesiana. **Bras. Biom.** 26: 107-118.

- Vos, P.; Hogers, R.; Bleeker, M.; et al. 1995. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, 23: 4407-4414.
- Welsh, J.; McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. **Nucleic Acids Research**, 18: 7213-7218.
- Wu, R.; Ma, C.X.; Painter, I.; Zeng, Z.-B. 2002. Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. **Theoretical Population Biology** 61: 349-363.
- Wycherley, P.R. The genus *Hevea*. In: WORKSHOP ON INTERNATIONAL COLLABORATION *HEVEA* BREEDING AND THE COLLECTION AND ESTABLISHMENT OF MATERIALS FROM NEOTROPIC, Kuala Lumpur, 1977, 12p.
- Yu, F.; Wang, B H; Feng, S P; Wang, J Y; Li, W G; Wu, Y T. 2011. Development, characterization, and cross-species/genera transferability of SSR markers for rubber tree (*Hevea brasiliensis*). **Plant Cell Reports**, 30:335-344.
- Zeng, Z.-B. 1994. Precision mapping of quantitative trait loci. **Genetics**, 136:1457-1466.
- Zeng, Z.-B. 1993. Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. **Proceedings of the National Academy of Science USA**, 90: 10972-10976.
- Zeng, Z.-B.; Kao, C-H.; Basten, C.J. 1999. Estimating the genetic architecture of quantitative traits. **General Research**, 74:279-289.
- ZIETKIEWICZ, E.; RAFALSKI, A.; LABUDA, D. 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. **Genomics**, 2: 176-183.