

Universidade Federal do Rio de Janeiro

**Diferentes abordagens
na avaliação da estrutura genética populacional**

Catarina da Fonseca Lira de Medeiros

2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



Diferentes abordagens na avaliação da estrutura genética populacional

Catarina da Fonseca Lira de Medeiros

Tese de Doutorado apresentada ao Programa de
Pós-graduação em Química Biológica,
Instituto de Bioquímica Médica,
da Universidade Federal do Rio de Janeiro,
como parte dos requisitos necessários à
obtenção do título de Doutor em Química Biológica.

Orientadores: Paulo Cavalcanti Gomes Ferreira
Mônica Aires Cardoso

Rio de Janeiro
Setembro 2008

Diferentes abordagens na avaliação da estrutura genética populacional

Catarina da Fonseca Lira de Medeiros

Orientadores: Paulo Cavalcanti Gomes Ferreira
Mônica Aires Cardoso

Tese de Doutorado submetida ao Programa de Pós-graduação em Química Biológica, Instituto de Bioquímica Médica, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Doutor em Química Biológica.

Aprovada por:

Presidente, Prof. Paulo Cavalcanti Gomes Ferreira

Prof. Rogerio Margis

Profa. Elizabeth Atalla Mansur de Oliveira

Prof. Franklin David Rumjanek

Rio de Janeiro
Setembro 2008

Lira-Medeiros, Catarina da Fonseca

Diferentes abordagens na avaliação da estrutura genética populacional/Catarina da Fonseca Lira de Medeiros. Rio de Janeiro: UFRJ/IBqM, 2008.

xiv, 243f., il.; 31 cm.

Orientador: Paulo Cavalcanti Gomes Ferreira

Tese (doutorado) – UFRJ/Instituto de Bioquímica Médica/Programa de Pós-graduação em Química Biológica, 2008.

Referências Bibliográficas: f. 15

1. Análise genética. 2. Manguezal. 3. Pau-Brasil. 4. Simulação. 5. Marcador molecular. 6. Estrutura epigenética. I. Ferreira, Paulo Cavalcanti Gomes. II. Universidade Federal do Rio de Janeiro, Instituto de Bioquímica Médica, Programa de Pós- graduação em Química Biológica. III. Diferentes abordagens na avaliação da estrutura genética populacional.

Agradecimentos

Agradeço a todos que participaram comigo desta jornada de quatro anos do doutorado, em especial minha família, Rodrigo e Aninha, meus pais Helena e Lélío e minhas irmãs Adriana e Silvia.

Agradeço também aos amigos do laboratório pelos vários momentos de seriedade e discussão, mas principalmente pelos momentos de descontração.

Agradeço às meninas, minhas amigas de colégio, sempre muito próximas de mim dividindo os bons momentos e ajudando a superar os momentos ruins.

Como não poderia ser, agradeço aos amigos da faculdade que, apesar das dificuldades de nos encontrar, mantemos nossa lista firme e forte com muitas emoções e novidades todos os dias.

Especialmente, agradeço aos meus orientadores, Paulo e Mônica, porque sem eles não estaria realizando este projeto. Eles sempre me dão o suporte necessário para o desenvolvimento do trabalho, mas agüentando também as minhas enrolações.

RESUMO

Diferentes abordagens na avaliação da estrutura genética populacional
Catarina da Fonseca Lira de Medeiros
Orientador(es): Paulo Cavalcanti Gomes Ferreira
Mônica Aires Cardoso

Resumo da Tese de Doutorado submetida ao Programa de Pós-graduação em Química Biológica, Instituto de Bioquímica Médica, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Doutor em Química Biológica

Estudos populacionais com espécies vegetais podem integrar o conhecimento genético, demográfico e ecológico graças a uma série de técnicas e abordagens diferenciadas. O componente genético, por exemplo, pode ser estudado com marcadores moleculares neutros que fornecem dados sobre a diferenciação genética e isolamento entre as populações. Mas, por outro lado, a modelagem computacional é uma nova abordagem que permite avaliar dados genéticos, demográficos e ecológicos de uma espécie e suas populações. O complexo Mata Atlântica destaca-se pelo alto grau de biodiversidade e endemismo encontrado em seus ecossistemas, deixando-a entre os cinco principais *hotspots* do mundo. Espécies como *Caesalpinia echinata* (Pau-Brasil) já foram bastante estudadas geneticamente, porém dados demográficos e ecológicos ainda são desconhecidos. Por outro lado, ecossistemas inteiros como os manguezais, apesar da grande importância econômica e ecológica, ainda não foram desvendados no Brasil, principalmente por estudos genéticos. Neste trabalho, foram utilizadas três abordagens: simulações computacionais com *C. echinata*, estudo da estrutura genética e epigenética com duas espécies de mangue, e desenvolvimento de marcadores microssatélites específicos para estudos de fluxo gênico entre manguezais do Brasil. A primeira etapa do trabalho, mostrou que plantas como *C. echinata* têm sua estrutura genética bastante afetada por algumas características ecológicas, como tipo de cruzamento, dispersão e longevidade, mas de forma independente para marcadores nucleares ou citoplasmáticos. Com isso, foi possível inferir sobre a biologia reprodutiva de *C. echinata*, que ainda não é totalmente conhecida. A segunda abordagem consistiu em estudos genéticos através de marcadores AFLP com duas populações próximas de *Laguncularia racemosa* e *Avicennia schaueriana*, mas que possuem características morfológicas bastante distintas. Os resultados mostraram alguma diferenciação genética entre populações de

A. schaueriana. Mas *L. racemosa*, por outro lado, não tinha nenhuma diferenciação e foi em seguida investigada por marcadores MSAP. A estrutura genética e epigenética reveladas pelo MSAP mostraram que a diferenciação epigenética era maior do que a genética entre as populações de *L. racemosa*, indicando que as diferenças morfológicas encontradas nesta espécie seriam resultado de fenômenos epigenéticos, como metilação do DNA, alterados pelo estresse salino ao qual uma das populações é submetida diariamente. Por último, foram isolados marcadores microssatélites nucleares aleatórios e de regiões expressas do genoma para as duas espécies de mangue: *L. racemosa* e *A. schaueriana*. Os microssatélites aleatórios são difíceis de serem isolados, pois necessitam da criação de uma biblioteca de DNA enriquecida com microssatélites a fim de obter clones dos quais podem ser isolados os marcadores. Isso gera um empecilho devido ao alto dispêndio de tempo e dinheiro, muitas vezes sem sucesso, no isolamento de marcadores que sejam amplificados via PCR, mas também polimórficos. Por outro lado, os microssatélites de regiões gênicas são mais fáceis de isolar porque podem ser usados bancos de EST já seqüenciados. Apesar destes bancos geralmente não serem encontrados para espécies selvagens, este problema ainda pode ser contornado devido à alta transferabilidade destes marcadores entre espécies correlatas. Durante todo este trabalho, foi possível abordar diferentes técnicas usadas em estudos genéticos com dois modelos diferentes, o Pau-Brasil e os manguezais. Todas as metodologias se mostraram eficientes em cumprir os objetivos de cada etapa do trabalho.

Palavras-chave: Análise genética, manguezal, Pau-Brasil, simulação, marcador molecular, estrutura epigenética

Rio de Janeiro
Setembro 2008

ABSTRACT

Different approaches to assess population genetic structure
Catarina da Fonseca Lira de Medeiros
Orientador(es): Paulo Cavalcanti Gomes Ferreira
Mônica Aires Cardoso

Abstract da Tese de Doutorado submetida ao Programa de Pós-graduação em Química Biológica, Instituto de Bioquímica Médica, da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários à obtenção do título de Doutor em Química Biológica.

Plant species population study can integrate the genetic, demographic and ecological characteristics using different techniques and approaches. The genetic component, for example, can be studied with neutral molecular markers that provide data on genetic differentiation and isolation among populations. On the other hand, computer modeling is a new approach for assessing genetic, demographic and ecological data of a species and their populations. The complex Atlantic Rainforest has one of the highest degrees of biodiversity and endemism and it is among the top five hotspots of the world. Species such as *Caesalpinia echinata* (Brazil-wood) have been fairly genetically studied, but ecological and demographic data are still missing. On the other hand, whole ecosystems such as mangroves, despite the great economic and ecological importance, have not yet been fully studied in Brazil mainly using genetic approaches. In this work, three approaches were used: (i) computer simulations for *C. echinata* populations; (ii) assess of genetic and epigenetic structure within two mangrove species; and (iii) development of microsatellite specific markers for gene flow studies between Brazilian mangroves. The first part of this work showed that plants like *C. echinata* have its genetic structure significantly affected by some ecological characteristics, such as selfing rate, fragmentation and life span, but in an independently way for nuclear and cytoplasmic markers. It was possible to infer some reproductive biology characteristics of *C. echinata*, which is not yet fully understood. Another approach assessed the genetic studies by AFLP markers of two nearby populations, which have distinct morphological characteristics, of *Laguncularia racemosa* and *Avicennia schaueriana*. Some genetic differentiation was found among *A. schaueriana* populations but *L. racemosa* populations had no differentiation and, consequently, were investigated by MSAP markers. The genetic and epigenetic structures revealed by MSAP showed that

the epigenetic differentiation was greater than the genetic one between *L. racemosa* populations. This suggested that morphological differences found within populations are results from epigenetic phenomena such as DNA methylation, caused by salt stress to which one of the populations is daily submitted. Finally, nuclear microsatellite markers were isolated from random and expressed regions of the genome for both mangrove species: *L. racemosa* and *A. schaueriana*. The random microsatellites are difficult to be isolated; as they need the creation of an enriched DNA library to obtain clones of which the markers can be isolated. This is an obstacle due to high expenditure of time and money, often without success for the isolation of PCR amplifiable and polymorphic microsatellite markers. Moreover, genic microsatellites are easier to isolate because EST databases can be used. Although these databases are restricted to some species, generally not found for wild species. This problem can be minimized with the high transferability of these markers between related species. Throughout this work, it was possible to address different techniques used in genetic studies with two different models, Brazil-wood and mangroves. All methodologies have efficiently fulfilled the goals of each study.

Kew-words: Genetic analysis, mangroves, Brazil-wood, simulation, molecular marker, epigenetic structure

Rio de Janeiro
Setembro 2008

Lista das Figuras

Figura 1.1	24
Figura 1.2	25
Figura 1.3	27
Figura 1.4	28
Figura 1.5	30
Figura 1.6	33
Figura 1.7	44
Figura 1.8	44
Figura 1.9	46
Figura 1.10	47
Figura 1.11	49
Figura 2.1	60
Figura 2.2	61
Figura 2.3	64
Figura 2.4	69
Figura 2.5	72
Figura 2.6	76
Figura 2.7	88
Figura 2.8	89
Figura 2.9	92
Figura 2.10	94
Figura 2.11	99
Figura 3.1	114
Figura 3.2	115
Figura 3.3	116
Figura 3.4	119
Figura 3.5	137
Figura 3.6	138
Figura 3.7	138
Figura 3.8	140
Figura 3.9	141
Figura 3.10	143
Figura 3.11	143
Figura 3.12	144
Figura 3.13	145
Figura 3.14	146
Figura 3.15	146

Lista das Tabelas

Tabela 1.1	34
Tabela 1.2	37
Tabela 1.3	40
Tabela 2.1	83
Tabela 2.2	84
Tabela 2.3	85
Tabela 2.4	85
Tabela 2.5	90
Tabela 3.1	110
Tabela 3.2	112
Tabela 3.3	117
Tabela 3.4	124
Tabela 3.5	133
Tabela 3.6	133
Tabela 3.7	135
Tabela 3.8	137
Tabela 3.9	145

Lista de Abreviaturas

AFLP – Amplified Fragment Length Polymorphism

BPCA – Bayesian Principal Component Analysis

EST – Expressed Sequence Tag

MCMC - Monte-Carlo Markov Chain

MSAP – Methyl-Sensitive Amplified Polymorphism

NCBI - National Center for Biotechnology Information

pb – Pares de base

PCA – Principal Component Analysis

PCR – Polymerase Chain Reaction

QTL – Quantitative Trait *Loci*

RAPD – Random Amplified Polymorphic DNA

RFLP – Restriction Fragment Length Polymorphism

SSR – Simple Sequence Repeat

TIGR – The Institute for Genomic Research

UTR – Untranslated Region

Sumário

Agradecimentos	v
Lista das Figuras	x
Lista das Tabelas	xi
Lista de Abreviaturas	xii
Sumário	xiii
Introdução Geral	15
Capítulo 1	19
1.1 Introdução	20
1.2 Objetivos	31
1.3 Materiais e Métodos	32
1.3.1 Parâmetros demográficos e ciclo de vida	34
1.3.2 Parâmetros genéticos e sistema de cruzamento	36
1.3.3 Dispersão do pólen e da semente	38
1.3.4 Simulações	39
1.3.5 Análise dos resultados	41
1.4 Resultados	42
1.4.1 Sistema de cruzamento	42
1.4.2 Taxa de dispersão e fragmentação	45
1.4.3 Ciclo de vida	46
1.4.4 Fragmentação recente	48
1.5 Discussão	50
1.5.1 Sistema de cruzamento	51
1.5.2 Taxa de dispersão e fragmentação	52
1.5.3 Ciclo de vida	54
1.5.4 Fragmentação recente	54
1.5.5 Questões conservacionistas	55
Capítulo 2	57
2.1 Introdução	58
2.2 Objetivos	67
2.3 Materiais e Métodos	68
2.3.1 Material Vegetal	68
2.3.2 Extração do DNA	70
2.3.3 AFLP	70
2.3.3.1 Protocolo de amplificação	71
2.3.3.2 Análise dos dados	74
2.3.4 MSAP	75
2.3.4.1 Protocolo de amplificação	75
2.3.4.2 Análise dos dados	79
2.4 Resultados	82
2.4.1 Marcadores AFLP	82
2.4.1.1 Diversidade genética	82
2.4.1.2 Estrutura populacional	84
2.4.2 Marcadores MSAP	86
2.4.2.1 Níveis de polimorfismo	87
2.4.2.2 Estrutura genética populacional	90

2.4.2.3	Análise de componente principal (PCA)	93
2.5	Discussão	95
2.5.1	Marcadores AFLP	95
2.5.2	Marcadores MSAP	98
Capítulo 3		103
3.1	Introdução	104
3.1.1	Marcadores Moleculares	104
3.1.2	Aplicações ecológicas dos microssatélites	111
3.1.3	Publicações atuais sobre conservação	113
3.1.4	Espécies de mangue	120
3.2	Objetivos	121
3.3	Materiais e Métodos	122
3.3.1	SSR-EST	122
3.3.1.1	Isolamento dos marcadores SSR-EST	122
3.3.1.2	Caracterização dos marcadores SSR-EST	124
3.3.2	SSR nucleares aleatórios	125
3.3.2.1	Criação da biblioteca de DNA enriquecida com microssatélites	126
3.3.2.2	Busca de clones positivos	129
3.3.2.3	Seqüenciamento da biblioteca de DNA	130
3.3.2.4	Análise das seqüências e desenho dos <i>primers</i>	131
3.3.2.5	Caracterização dos <i>primers</i> SSR nucleares	132
3.4	Resultados	134
3.4.1	Isolamento dos marcadores microssatélites do tipo SSR-EST	134
3.4.2	Isolamento dos marcadores microssatélites nucleares aleatórios	139
3.5	Discussão	147
	Referências bibliográficas	157
	Glossário	171
	Anexo I	173
	Anexo II	205
	Anexo III	206
	Anexo IV	231

Introdução Geral

A biodiversidade é abordada em três níveis: a diversidade genética, que se refere à variedade de informação genética contida tanto nos níveis intra e interpopulacional como interespecífico; a diversidade de espécies, que se refere à grande variedade de seres vivos; e a diversidade de ecossistemas, que se refere à variedade de habitats, comunidades bióticas e processos ecológicos (Lira, 2003). A diversidade genética tem papel importante na conservação de espécies ameaçadas e na manutenção de suas populações ao longo do tempo, pois espécies com baixa variação genética têm geralmente uma redução na capacidade de sobreviver a mudanças ambientais durante o curso da evolução (Frankham, 1995). A integração entre o conhecimento genético, demográfico e ecológico de uma espécie é importante para a formulação de estratégias de conservação de espécies que levem em conta estes três aspectos ao mesmo tempo (Escudero *et al.*, 2003).

A Mata Atlântica é um complexo de formações vegetais, englobando florestas ombrófilas, florestas estacionais, além de formações costeiras como restingas, manguezais, praias e dunas, e ainda campos de altitude e encaves de cerrado (Lira, 2003). Sua importância está centrada no alto grau de biodiversidade e endemismo encontrado em seus ecossistemas, o que a torna um dos cinco principais “hotspots” de biodiversidade do mundo (Myers *et al.*, 2000). Algumas espécies de animais e vegetais se destacam dentro deste ecossistema devido à sua importância histórica como o Pau-Brasil (*Caesalpinia echinata* Lam.). Mas há também ecossistemas inteiros com grande importância ecológica e econômica como restingas e manguezais. Ocorrendo em baixas latitudes na interface entre terra e mar, as florestas de manguezais são bastante robustas e altamente adaptáveis, ocupando um ambiente submetido às alterações diárias de

marés, causando variações na temperatura, quantidade de água, salinidade e níveis de anoxia (Alongi, 2008).

O componente genético de uma espécie pode ser estudado através de marcadores moleculares que são neutros em relação a efeitos fenotípicos, com mínimo ou nulo efeito epistático¹ ou pleiotrófico² e estão aleatoriamente inseridos no genoma, e, por conseguinte, fornecem resultados sobre a diversidade genética dentro de cada população e a diferenciação entre elas (Escudero *et al.*, 2003). Tradicionalmente, a diferenciação genética e isolamento entre as populações são estudadas pela estimativa dos parâmetros da estrutura populacional como F_{ST} ³ e fluxo gênico (Bensch and Åkesson, 2005). Atualmente, uma nova abordagem que vem sendo utilizada é a modelagem de populações que permite analisar resultados tanto genéticos quanto ecológicos sobre uma espécie através de simulações computacionais (Topping *et al.*, 2003).

O uso de simulações em estudos populacionais possibilita avaliar de forma rápida e eficiente a viabilidade das populações através de parâmetros genéticos e demográficos, o que sempre foi dificultado devido à necessidade de formular estudos demográficos de longo prazo em campo, de conhecer a biologia reprodutiva da espécie e, por último, de avaliar a estrutura genética e os efeitos ambientais sobre as populações (Volis *et al.*, 2005).

O uso de marcadores moleculares é recorrente em estudos de diversidade genética com ênfase na conservação de espécies, sendo os marcadores mais utilizados atualmente RAPD (*Random Amplified Polymorphic DNA*), AFLP (*Amplified Fragment Length Polymorphism*) e SSR (*Simple Sequence Repeat*) ou microssatélites (Lira *et al.*, 2006). Estes marcadores são bastante informativos, porém o RAPD e o AFLP são

¹ Fenômeno em que vários pares de genes interagem entre si para influenciar uma única característica fenotípica.

² Quando uma variação devida a um só gene não se traduz num, mas em vários efeitos fenotípicos.

³ Índice de diferenciação populacional baseado em polimorfismo genético.

dominantes e por isso não permitem o estudo de heterozigotos, enquanto que os microssatélites são co-dominantes, o que possibilita a visualização de cada alelo por indivíduo.

Através de uma técnica baseada no AFLP, o MSAP (*Methyl-Sensitive Amplified Polymorphism*), é possível avaliar os dois componentes: genético e epigenético do genoma de uma espécie. As alterações epigenéticas são baseadas numa série de processos moleculares que podem induzir, reduzir ou silenciar a atividade de genes específicos como: (1) metilação de citosinas no DNA; (2) remodelagem da estrutura da cromatina através de modificações químicas, em particular acetilação e metilação de histonas; e (3) processos regulatórios mediados por RNA de interferência (Bossdorf *et al.*, 2008). A técnica do MSAP permite a análise de padrões de metilação de citosinas através do uso de duas enzimas isoesquisômeras que reconhecem o mesmo sítio de clivagem, porém possuem diferentes sensibilidades à metilação de citosinas. Torna-se interessante o uso desta técnica em populações fenotipicamente diferentes, mas geneticamente similares, porque os fenômenos epigenéticos são responsáveis por alterações de fenótipos sem que ocorra alteração de genótipo, ou seja, sem alterações nas seqüências de nucleotídeos (Jablonka and Lamb, 1998). Isso pode, portanto, explicar diferenças morfológicas em espécies que possuem indivíduos sem muita variação genética.

Os microssatélites ou SSR são regiões hipervariáveis do genoma e se destacam por permitir a análise de homozigotos e heterozigotos numa população, fenômeno chamado de co-dominância. Porém o alto custo inicial para o desenvolvimento desta técnica, aliado à dificuldade de se obter marcadores eficientes, inibe o uso desta técnica em espécies selvagens ou de baixa importância econômica que nunca foram estudadas geneticamente (Maguire *et al.*, 2000). Por outro lado, uma nova modalidade de

microssatélites foi encontrada dentro de seqüências expressas do genoma, chamados de SSR-EST. O custo inicial para se obter os marcadores SSR-EST é mais baixo sendo facilitado pela existência de bancos de EST⁴ disponíveis na internet e programas de computador gratuitos usados especificamente para isolar estes marcadores (Thiel *et al.*, 2003), porém são poucas as espécies que já possuem bancos de EST disponibilizados.

Ao longo desta tese foram utilizadas diferentes abordagens visando o estudo populacional de espécies vegetais. Primeiramente, a espécie *Caesalpinia echinata* Lam. (Pau-Brasil) foi utilizada como modelo de simulações que objetivaram aprofundar o conhecimento dos efeitos de características ecológicas de uma determinada espécie sobre a diferenciação genética entre suas populações. Posteriormente, os alvos de estudo foram duas espécies de mangue, *Laguncularia racemosa* e *Avicennia schaueriana*, ambas presentes no município do Rio de Janeiro, onde foi feito um estudo mais aprofundado nas populações de Guaratiba usando a técnica do AFLP. Com resultados bastante interessantes de baixa diferenciação genética entre as populações de *L. racemosa*, apesar do alto grau de distinção morfológica entre elas, a técnica do MSAP acessou o componente genético e epigenético do genoma de indivíduos das duas populações de Guaratiba ao mesmo tempo. Por último, foram isolados marcadores microssatélites do tipo SSR para ambas as espécies de mangue e SSR-EST para o gênero *Avicennia*.

⁴ Um segmento de cDNA (DNA complementar ao RNA) com seqüência idêntica à parte da região transcrita de um gene.

Capítulo 1

Simulação da evolução genética de uma metapopulação usando o Pau-Brasil (*Caesalpinia echinata* Lam.) como espécie-modelo

Manuscrito submetido ao periódico *Biological Conservation* (Anexo I):

“How life-cycle, mating system and fragmentation processes can influence nuclear and cytoplasmic genetic differentiation? A modeling study based on *Caesalpinia echinata* Lam. (Leguminosae)”

1.1 Introdução

Três características são importantes para impedir que ocorra diferenciação genética entre as populações de plantas: localização geográfica, sistema de cruzamento e fluxo gênico (Hamrick e Godt, 1990). A localização geográfica diz respeito ao ambiente, pois as plantas estarão mais bem adaptadas e o seu desenvolvimento será impulsionado quando fatores bióticos ou abióticos à sua volta forem favoráveis ao seu crescimento e reprodução.

Existem diferentes sistemas de cruzamento nas plantas. Além de ter a possibilidade de fazer reprodução assexuada, chamada de crescimento vegetativo, as plantas podem ser autógamias ou alógamas, ou seja, podem ter predominância de autofecundação ou de fecundação cruzada, respectivamente. Sistemas de cruzamento baseados na autofecundação geralmente resultam em maior diferenciação entre as populações, pois a informação genética não é trocada com indivíduos de outras populações. Por outro lado, em espécies alógamas, ocorre naturalmente troca de informação genética entre suas populações devido à preferência de cruzamento, o que geralmente diminui a diferenciação entre elas.

A terceira característica é o fluxo gênico, dependente dos mecanismos de dispersão do pólen e da semente. O sucesso do cruzamento entre indivíduos e populações está totalmente relacionado com o sistema de dispersão. Diferentes mecanismos de dispersão possibilitam que o pólen e a semente alcancem longas distâncias e, conseqüentemente, outras populações, principalmente no caso de espécies que vivem em ambientes fragmentados. Em decorrência da fragmentação de habitat, populações naturais podem se tornar pequenas e isoladas, ou até se extinguirem. Se o fluxo gênico entre os fragmentos é limitado, devido ao tipo de sistema de cruzamento ou de dispersão da semente e do pólen, as populações estarão mais propensas ao

efeito da deriva gênica⁵, resultando na estruturação da diversidade genética entre as populações, ou seja, em sua diferenciação (Hamrick, 2004). Quanto menor for o número e o tamanho das populações, maior a probabilidade de sofrerem uma diminuição na diversidade genética dentro das populações e um aumento na diferenciação genética entre as populações (Young *et al.*, 1996). Entretanto, espécies que sobreviveram à fragmentação de seu habitat, ou seja, que se mantiveram ao longo das gerações nestes ambientes isolados, podem se recuperar quanto ao nível de diversidade genética em suas populações naturais, desde que aquelas três características citadas acima (localização geográfica, sistema de cruzamento e fluxo gênico) estejam atuando de forma eficiente no ambiente fragmentado.

A princípio, qualquer tipo de diferenciação entre populações ou estruturação da diversidade genética é prejudicial para a sobrevivência da espécie. A diversidade genética de uma determinada espécie é muito importante para que ela mantenha o seu potencial evolutivo, ou seja, ela contribui na sua adaptação em um ambiente em alteração, e por isso a perda de diversidade genética está geralmente associada à redução do *fitness*⁶ (Yao *et al.*, 2007). Isso acontece porque a variabilidade genética é imprescindível para que ocorra seleção natural sobre uma espécie, e esta seleção estará atuando sobre os genótipos melhor adaptados ao ambiente.

A manutenção da diversidade genética de espécies submetidas à fragmentação em seu habitat natural depende de vários fatores ecológicos (Yao *et al.*, 2007) dentre eles a sua diversidade genética antes da fragmentação, o fluxo gênico (Bossuyt, 2007), o sistema de cruzamento e a duração do ciclo de vida (Honnay e Jacquemyn, 2007). Mas pouco se sabe sobre o efeito direto de cada um destes fatores ecológicos sobre a diferenciação genética entre

⁵ Flutuações estocásticas na frequência alélica dentro de uma população, mais evidente em populações pequenas.

⁶ A medida da habilidade do indivíduo de uma espécie em sobreviver e se reproduzir.

as populações.

A espécie *Caesalpinia echinata* Lam. (Familia: Leguminosae), endêmica do Brasil, ocorre na Mata Atlântica, um ecossistema muito fragmentado atualmente, mas que contém um dos mais altos índices de diversidade biológica (*hotspot*⁷) do mundo (Myers *et al.*, 2000). Esta fragmentação ocorreu devido aos séculos de exploração da espécie seguida da destruição de seu habitat causada pelo desenvolvimento e crescimento desordenado do país. Como resultado, o Pau-Brasil foi uma das espécies mais exploradas nos últimos 500 anos e esta entre as espécies ditas como ameaçadas de extinção em lista nacional e internacional (Lira, 2003).

Da madeira do Pau-Brasil era extraído um pigmento vermelho, a brasileína, muito utilizado para tingir tecidos na época imperial (Carvalho, 1994). Após o descobrimento do corante sintético – anilina – em meados do século XIX, o Pau-Brasil deixou de ser explorado de forma tão avassaladora. Atualmente, a sua principal importância é na fabricação de arcos de violino, devido à grande flexibilidade e resistência da madeira do Pau-Brasil, características ainda não encontradas em nenhuma outra espécie vegetal (Carvalho, 1994). Mesmo após muitos anos de exploração desta espécie e com o estado atual da Mata Atlântica no Brasil, algumas populações naturais ainda resistem ao longo da costa. Porém, as principais ameaças ao Pau-Brasil continuam sendo o isolamento de suas populações naturais e a destruição de seu habitat, já que restam apenas em torno de 7% da cobertura original da Mata Atlântica no Brasil (SOS Mata Atlântica INPE, 2002).

Até hoje, a biologia reprodutiva desta espécie ainda não foi completamente resolvida, dado de extrema relevância para se formular estratégias eficientes não só para a conservação desta espécie, mas também de todo o seu ecossistema. Existem algumas hipóteses quanto ao

⁷ Áreas consideradas prioritárias para conservação, que possuem altos índices de biodiversidade.

ciclo de vida do Pau-Brasil, sendo provável que ele viva em torno de 300 anos (H.C. Lima – comunicação pessoal). Por outro lado, pouco se sabe sobre a biologia floral da espécie, porque suas flores são de difícil manipulação devido à grande fragilidade delas (M.A. Cardoso – comunicação pessoal). Porém, um estudo atual mostra que a polinização ocorre por abelhas dos gêneros *Xylocopa* e *Centris*, mas também pela abelha *Apis mellifera*, sendo a dispersão da semente autocórica, ou seja, feita pela gravidade (Borges *et al.*, 2008). A figura 1.1 mostra detalhes de folhas e flores de um indivíduo de Pau-Brasil.

Sete populações naturais de Pau-Brasil distribuídas ao longo da costa brasileira foram analisadas usando dois marcadores moleculares dominantes (RAPD e AFLP; Cardoso *et al.*, 1998; Cardoso *et al.*, 2005), mas também com marcadores microssatélites de cloroplasto (Lira *et al.*, 2003). Ambos marcadores dominantes revelaram uma forte diferenciação genética diretamente correlacionada com as distâncias geográficas entre as populações, caracterizando o isolamento por distância (Lira 2003). Foram verificados os níveis de diversidade genética: $H_e = 0,636$ e $0,125$; e de diferenciação genética: $F_{ST} = 0,58$ e $0,62$; com estes marcadores (Cardoso *et al.*, 1998 e 2005). Este isolamento entre as populações foi ainda mais evidente através de marcadores citoplasmáticos que mostraram diversidade genética ainda mais baixa ($H_e = 0,085$), porém diferenciação genética bastante alta ($F_{ST} = 0,91$) (Lira *et al.*, 2003). Estes índices foram um reflexo direto do número de haplótipos observados nas populações naturais de Pau-Brasil. Lira *et al.* (2003) mostraram que, do total de oito haplótipos observados, sete eram específicos de uma única população, e além disso, só duas populações eram compostas por mais de um haplótipo. Este nível de conhecimento genético é extremamente importante para que seja feito um planejamento visando a conservação e sobrevivência da espécie (Rajora e Mosseler, 2001).



Figura 1.1 Fotos ilustrativas da árvore, folhas e flores da espécie *Caesalpinia echinata* Lam., mais conhecida como Pau-Brasil.

Com base nos dados moleculares obtidos com os três marcadores moleculares, foi formulada a hipótese de que as populações naturais de Pau-Brasil estariam isoladas, formando três principais grupos, desde o último período glacial ou até antes disso (Lira, 2003). Das sete populações investigadas, Mamanguape (PA) e Mata da Estrela (RN) formariam o primeiro grupo no nordeste do Brasil; Eunápolis (BA) e Aracruz (ES) formariam o segundo grupo mais ao centro; enquanto que Cabo Frio (RJ), Saquarema (RJ) e Guaratiba (RJ) formariam o último grupo no sudeste do país (Figura 1.2; Lira *et al.*, 2003; Cardoso *et al.*, 2005).



Figura 1.2 Mapa de localização das populações naturais de *Caesalpinia echinata* Lam. estudadas geneticamente com marcadores AFLP e microssatélites de cloroplasto. Fonte: Cardoso *et al.* (2005).

A figura 1.3 mostra duas árvores de distância genética calculadas a partir dos resultados obtidos com os marcadores AFLP (Cardoso *et al.*, 2005) e microssatélites de cloroplasto (Lira *et al.*, 2003). O cladograma obtido com dados dos marcadores dominantes mostra uma clara separação dos grupos nordeste (Mamanguape e Mata da Estrela) e sudeste (Cabo Frio, Saquarema e Guaratiba). As populações Eunápolis e Aracruz ficaram distantes dos outros dois grupos, mas não se agruparam entre si (Figura 1.3 a). Por outro lado, o cladograma obtido com dados dos marcadores citoplasmáticos mostra nitidamente a formação dos três grupos principais: nordeste, centro e sudeste (Figura 1.3 b). Ambas as árvores corroboraram com a hipótese do isolamento antigo das populações desde as últimas glaciações. Estas regiões podem ter sido fragmentadas não apenas recentemente com o desmatamento cada vez mais amplo da Mata Atlântica, mas milhares ou até milhões de anos atrás (Lira *et al.*, 2003).

A teoria de que as populações naturais do Pau-Brasil teriam sofrido fragmentação desde milhares de anos atrás, é corroborada por estudos que sugerem a ocorrência de expansões e contrações das florestas secas e úmidas durante as glaciações (Pennington *et al.*, 2002). As áreas mais secas do Complexo Mata Atlântica, onde ocorre o Pau-Brasil, sofreram expansões durante os eventos glaciais, quando os ambientes mais secos eram favorecidos, e depois passavam por contrações nos períodos interglaciais (Pennington *et al.*, 2004). A figura 1.4 mostra a distribuição das florestas secas sazonais nas Américas Central e do Sul, que no Brasil está representada por parte do Complexo da Mata Atlântica, pela Caatinga e pelo Cerrado.

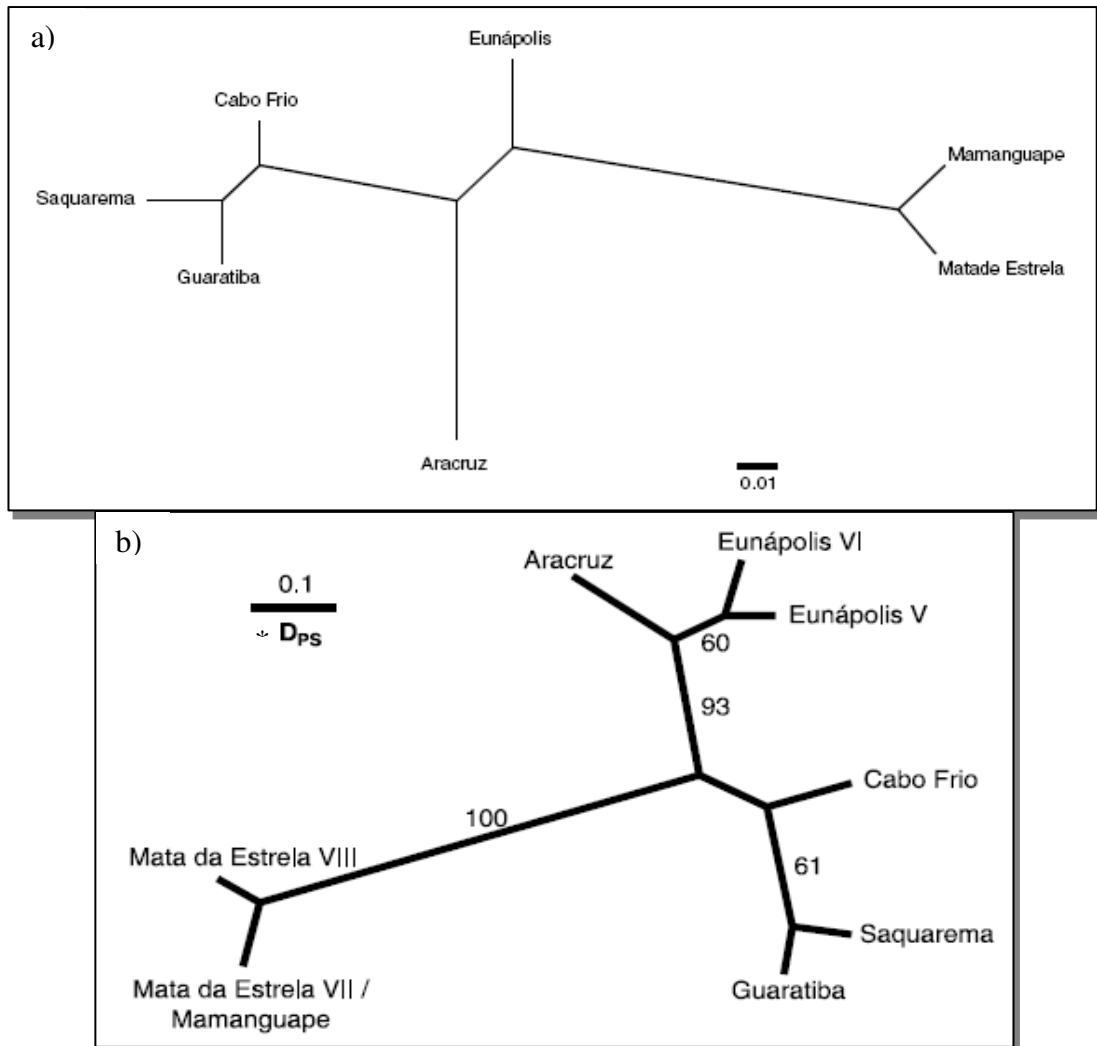


Figure 1.3 Árvores de distância genética calculadas a partir de dados moleculares dos marcadores (a) AFLP e (b) microssatélites de cloroplasto para sete populações naturais de *C. echinata* analisadas por Lira *et al.* (2003) e Cardoso *et al.* (2005). Na figura (b), as populações Mata da Estrela e Eunápolis estão indicadas com os números V, VI, VII e VIII que representam os quatro haplótipos encontrados nestas populações, enquanto que as outras populações tiveram um único haplótipo (detalhes em Lira *et al.*, 2003)
 * D_{PS} é um método que compara a proporção de alelos compartilhados entre haplótipos.

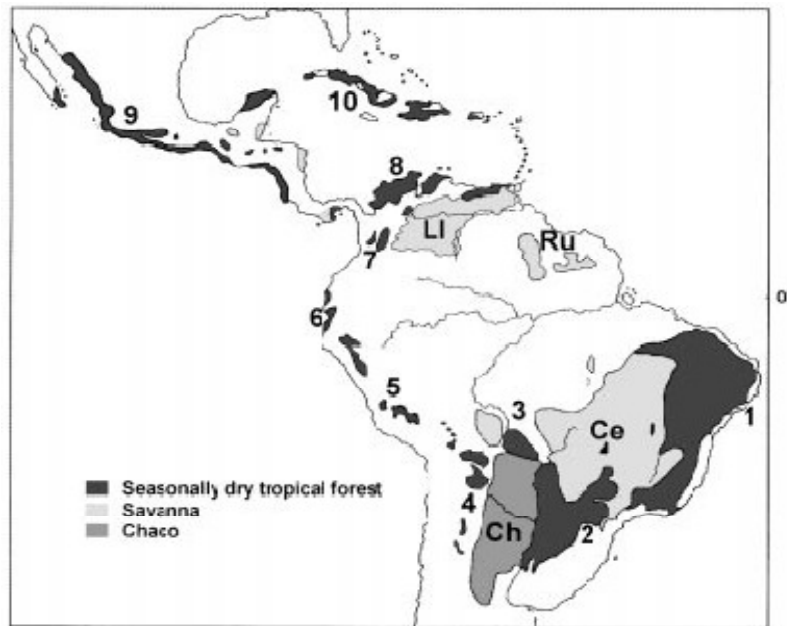


Figure 1.4 Mapa das Américas Central e do Sul indicando a distribuição das florestas tropicais sazonais secas (representadas em preto), que no Brasil fazem parte do Complexo Mata Atlântica onde ocorre a espécie *Caesalpinia echinata* Lam.

1: Caatinga. 2: Misiones Nucleus. 3, Região Boliviana Chiquitano. 4, Piedmont Nucleus. 5, Vales Interandinos Bolivianos e Peruanos. 6, Costa do Pacífico do Equador. 7, Vale Interandino Colombiano. 8, Costa caribenha da Colômbia e Venezuela. 9, América Central. 10, Antilhas. Tipos de savanas (em cinza claro): Ce, Cerrado. Ll, Llanos. Ru, Rupununi. Ch, Chaco (em cinza escuro). Fonte: Pennington *et al.* (2000).

Técnicas moleculares, juntamente com metodologias de modelagem, fornecem novas ferramentas promissoras para responder a questões ambientais mais complexas (Topping *et al.*, 2003). METAPOPOP é um programa que simula os processos demográficos e genéticos de uma metapopulação⁸, sob efeito de forças evolutivas como deriva gênica, mutação e fixação de genes; a ponto de conseguir um resultado fiel ao que teria ocorrido na natureza ao longo das gerações da planta em estudo (Le Corre *et al.*, 1997; Austerlitz *et al.*, 2000; Machon *et al.*, 2003).

⁸ Consiste num grupo de populações da mesma espécie separadas espacialmente e que podem interagir de algum modo.

A Figura 1.5 é um esquema simplificado do funcionamento do programa METAPOPOP. As frequências alélicas iniciais são definidas pelo usuário e, a partir delas, o programa forma os genótipos de cada indivíduo de forma aleatória. Em seguida, a cada geração são calculados os dados genéticos com base nas características demográficas e ecológicas da espécie, que foram inicialmente definidas pelo usuário.

Primeiramente, são calculados os tamanhos populacionais da próxima geração através de logística e do *input* de sementes vindas de outras populações. Depois são definidos aleatoriamente os pais que formarão os gametas e a união dos gametas produzindo os zigotos. Por último, a dispersão das sementes formadas será simulada conforme a matriz de dispersão definida pelo usuário do programa.

Simulações e modelagens há muito tempo são utilizadas para diferentes fins como a elaboração de estratégias de conservação, que exige um bom conhecimento da estrutura populacional e da diversidade genética; e também a estimativa do impacto de alterações ambientais naturais ou induzidas pelo Homem na condição genética das populações (Topping *et al.*, 2003). Estudos de viabilidade populacional usando dados genéticos e demográficos são raros, devido à necessidade de: (1) dados exaustivos sobre a demografia de uma espécie amostrada por longos períodos de tempo; (2) conhecimento da biologia reprodutiva da espécie; (3) estudo sobre a estrutura e variabilidade genética, além dos efeitos ambientais e demográficos sobre a sua biologia reprodutiva (Volis *et al.*, 2005). Por isso, as simulações são ferramentas muito importantes nos estudos populacionais, permitindo análises demográficas e genéticas num curto espaço de tempo e com baixo custo operacional.

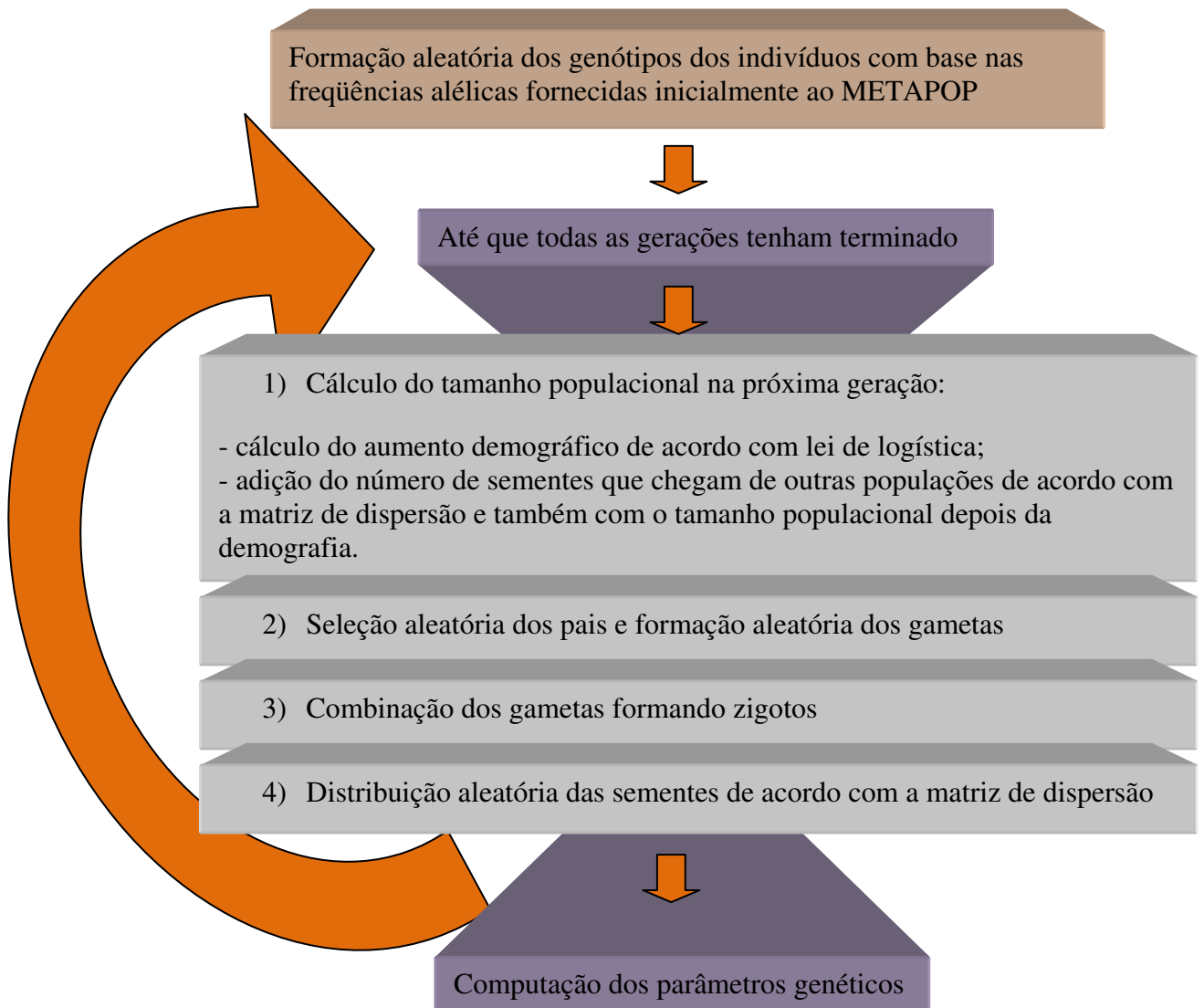


Figura 1.5 Esquema das etapas do programa METAPOP que calcula parâmetros genéticos populacionais a partir de dados demográficos e ecológicos como capacidade suporte, tamanho populacional, taxa de sobrevivência, produção de sementes e taxa de dispersão de pólen e de semente. Fonte: Le Corre *et al.* (1997).

1.2 Objetivos

Baseado em dados genéticos obtidos através de estudos anteriores com marcadores moleculares da espécie *Caesalpinia echinata* Lam. (Pau-Brasil), o objetivo deste capítulo foi compreender o papel de características ecológicas de uma espécie, como taxa de dispersão, sistema de cruzamento e ciclo de vida, na diferenciação genética nuclear e citoplasmática entre populações usando o programa METAPOP como ferramenta.

Os objetivos específicos foram:

- Simular os efeitos de diferentes sistemas de cruzamento desde alogamia até autogamia sobre a diferenciação genética entre as populações;
- Simular os efeitos de diferentes taxas de dispersão, representando a fragmentação do habitat, sobre a diferenciação genética entre as populações;
- Simular os efeitos da duração do ciclo de vida das plantas arbóreas sobre a diferenciação genética entre as populações;
- Correlacionar os resultados obtidos simulando os diferentes sistemas de cruzamento, taxas de dispersão e ciclos de vida com os resultados genéticos experimentais obtidos para populações naturais de Pau-Brasil;
- Simular a evolução da diferenciação genética encontrada atualmente entre populações naturais do Pau-Brasil nos próximos anos, assumindo um total isolamento reprodutivo entre suas populações sendo a dispersão dada como nula.

1.3 Materiais e Métodos

A evolução demográfica e genética de uma metapopulação foi simulada usando o programa METAPOPOP (Machon *et al.*, 1995; Le Corre *et al.*, 1997), que assume populações de plantas hermafroditas caracterizadas pelo seu genótipo em *loci*⁹ neutros ou seletivos. Este programa calcula a diferenciação genética entre demes¹⁰ de uma metapopulação assumindo os efeitos de forças evolutivas naturais, como mutação, deriva gênica e migração. Através do modelo demográfico dependente de densidade com gerações sobrepostas desenvolvido por Austerlitz *et al.* (2000), foram feitas simulações de uma metapopulação com 10 demes conectados de acordo com o modelo *stepping stone*¹¹ (Kimura and Weiss, 1964). Neste modelo, as populações ou demes estão dispostas numa linha reta ocorrendo troca de pólen e semente, apenas entre as duas demes mais próximas (Figura 1.6). As populações localizadas nas extremidades do modelo só dispersam para aquela população mais próxima. Este modelo foi usado por ser similar ao comportamento das populações naturais de Pau-Brasil localizadas ao longo da costa brasileira. A capacidade biótica de cada deme foi limitada em 300 plantas, de acordo com as pequenas populações do Pau-Brasil.

⁹ Local ou posição constante no cromossomo, onde localizam-se determinado gene ou marcador genético em todos os indivíduos de uma mesma espécie.

¹⁰ Uma população local panmítica, ou seja, que faz cruzamentos aleatórios entre seus indivíduos.

¹¹ Modelo teórico de estrutura populacional e fluxo gênico onde a espécie é subdividida em populações de tamanhos idênticos que trocam alelos apenas com as demes adjacentes.

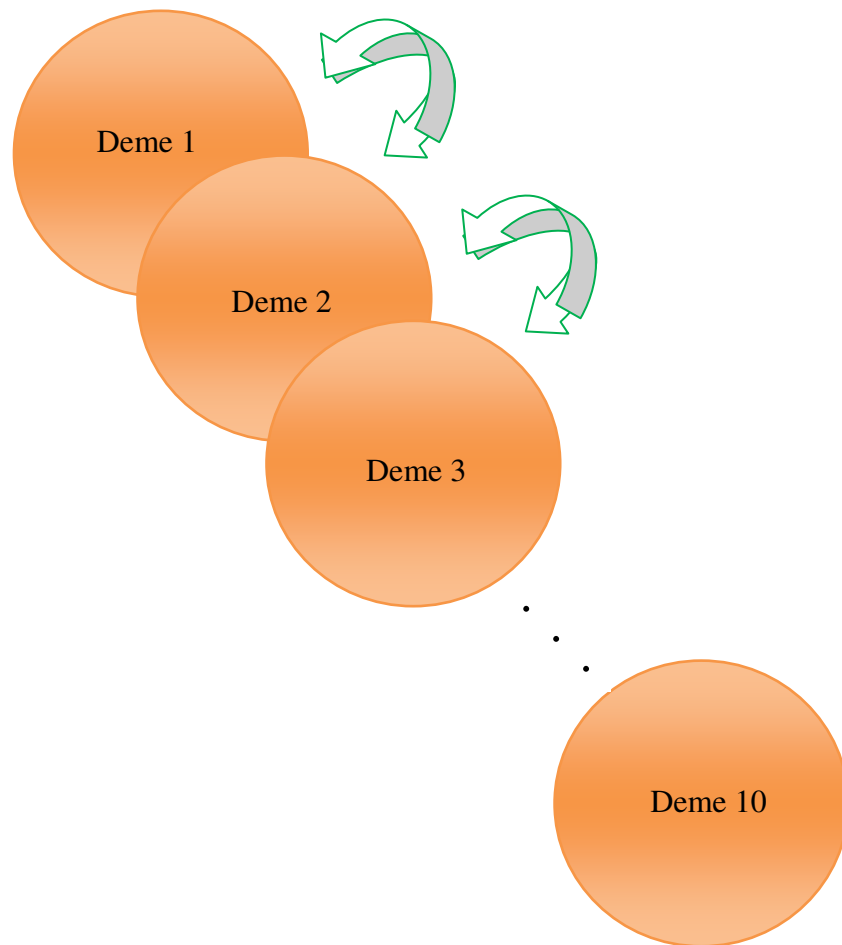


Figura 1.6 Esquema do modelo *stepping stone* utilizado pelo programa METAPOPOP nas simulações das populações naturais de *Caesalpinia echinata* Lam. Apenas as demes mais próximas podem trocar pólen ou semente entre elas, ou seja, a deme 1 por ser a primeira população do modelo troca apenas com a deme 2, esta por sua vez troca com as demes 1 e 3, que são vizinhas a ela e assim por diante até a décima deme.

1.3.1 Parâmetros demográficos e ciclo de vida

No modelo de Austerlitz *et al.* (2000), cada indivíduo pertence a uma determinada classe, que representa o período de um ano completo. De um ano para outro, cada indivíduo pode: (a) manter-se na mesma classe, (b) migrar para a classe seguinte, ou (c) morrer. Mas a probabilidade de se manter na mesma classe foi definida como zero, justamente para simular plantas perenes e, por isso, todos os indivíduos moviam-se para a próxima classe ou morriam. Impreterivelmente, todos os indivíduos da última classe morriam ao final das gerações, consideradas aqui como anos, porque atingiam a idade máxima de seu ciclo de vida. Portanto, as classes indicavam estritamente a idade de cada indivíduo e a idade máxima a ser atingida era equivalente ao total de classes assumido na simulação. Foram testados cinco ciclos de vida de 30, 60, 120, 300 e 500 classes cada um, simulando a idade máxima que estas plantas atingiriam: 30, 60, 120, 300 e 500 anos (Tabela 1.1). Já que existem indícios de que o Pau-Brasil vive até 300 anos de vida, esta duração de ciclo de vida foi definida como padrão nas outras simulações.

Tabela 1.1 Parâmetros demográficos usados nas simulações de populações de *C. echinata* com o programa METAPOP.

Classes	30, 60, 120, 300, 500
Fases juvenis/classe	7, 15, 15, 15, 30
Número de indivíduos/população	50 (10% juvenis e 90% adultos)
Fecundidade materna	3
Fecundidade paterna	100%

A definição da fase adulta reprodutiva e da fase juvenil pré-reprodutiva foi dependente da duração do ciclo de vida. Logicamente, quanto menor o ciclo de vida estabelecido para a simulação, mais curta a fase juvenil das plantas. Portanto, para as simulações de 30 anos, as sete primeiras classes foram definidas como juvenis, ou seja, a fertilidade masculina e feminina era equivalente a zero, enquanto que todas as outras classes eram compostas por plantas adultas que poderiam participar na reprodução via pólen ou semente. Para as simulações de plantas com 60, 120 e 300 anos, o número de classes juvenis foi 15 e para o ciclo de vida mais longo de 500 anos foi 30 (Tabela 1.1). Estes ajustes visavam que as populações não se extinguissem por produtividade insuficiente ou o inverso, ultrapassando a capacidade biótica de cada deme.

As populações eram compostas inicialmente por 50 indivíduos, distribuídos de forma aleatória e homogênea entre as classes de idade, sendo 10% do total de indivíduos alocados nas classes juvenis (pré-reprodutivas) e os outros 90% nas classes adultas (Tabela 1.1). O número total de indivíduos usado nas simulações foi independente da duração do ciclo de vida, mantendo-se sempre o mesmo valor de 50 plantas.

Para que as simulações fossem mais similares a uma população vegetal real, foram definidas diferentes probabilidades de sobrevivência e tamanho das plântulas com relação aos indivíduos. Para este, as plantas juvenis foram consideradas 10, 20 e 40% menores do que as adultas, dependendo da idade das mesmas até atingirem a fase adulta reprodutiva. Quanto à probabilidade de sobrevivência, ou seja, probabilidade de migrar para a classe seguinte a cada ano, foi definida de 50 a 70% para as plântulas e de 100% para adultos de qualquer classe, o que possibilitou as populações atingirem a sua capacidade suporte máxima sem risco de se extinguirem. Apesar das probabilidades de sobrevivência das classes adultas serem de 100%, estas dependiam da densidade populacional e, portanto diminuía com o crescimento da

população (Austerlitz *et al.*, 2000).

Os valores de fecundidade materna são absolutos no METAPOP. Eles correspondem ao número de descendentes, que sobrevivem e alcançam a primeira classe, produzidos por cada fêmea de forma independente da densidade. Foi assumido que todos os adultos tinham uma fertilidade materna de três (3), um valor consistente com as observações de campo (Tabela 1.1). METAPOP não assume limitação de pólen, e por isso a fertilidade masculina é relativa sendo definida como um (1) em todas as classes adultas, significando que, na prática, todos os adultos têm a mesma probabilidade de participar na reprodução via pólen (Tabela 1.1).

1.3.2 Parâmetros genéticos e sistema de cruzamento

Os cálculos de diferenciação genética se basearam em dois marcadores: nuclear e citoplasmático; a fim de reproduzir dados obtidos experimentalmente por Cardoso *et al.* (2005) e Lira *et al.* (2003) após a investigação de sete populações naturais de *Caesalpinia echinata* ao longo da costa brasileira. Os marcadores nucleares usados foram do tipo AFLP (Amplified Fragment Length Polymorphism), ou seja, marcadores dominantes que resultam em grande quantidade de *loci* altamente polimórficos. Eles são especialmente utilizados quando não há conhecimento prévio de seqüência de DNA e tem baixo custo inicial, o que facilita o estudo de espécies selvagens (Bensch and Åkesson, 2005). Por outro lado, os marcadores citoplasmáticos, geralmente presente nos cloroplastos das plantas, são bastante utilizados em estudos genéticos, pois o genoma desta organela não sofre recombinação, é herdado de forma uniparental e é haplóide (Provan *et al.*, 2001). Para ambos marcadores moleculares, nucleares e citoplasmáticos, os genótipos iniciais foram igualmente distribuídos

nas dez populações resultando em G_{ST}^{12} inicial próximo de zero, a fim de simular uma situação há 20.000 anos atrás quando as populações de Pau-Brasil não estariam diferenciadas.

Duzentos (200) marcadores nucleares foram simulados como marcadores AFLP com dois alelos por *locus*, devido à dominância deste tipo de marcador onde há apenas a análise de presença ou ausência de fragmentos. A sua taxa de mutação foi definida em 10^{-6} (Tabela 1.2; Mariette *et al.* 1999). Não foi assumida nenhuma ligação entre os *loci*, o que é compatível com o que ocorre na natureza, apesar de haver a possibilidade de no total de 200 *loci* de AFLP aleatórios, alguns estarem tão próximos no genoma a ponto de estarem ligados. Mas nosso resultado continua sendo válido, porque a ligação só aconteceria com uma pequena parte dos 200 *loci*, não afetando o resultado final que é baseado em 20 replicatas de cada simulação.

Foram definidos nas simulações dez *loci* de marcadores citoplasmáticos com 10 possíveis alelos em cada um, a fim de simular os marcadores encontrados no Pau-Brasil por Lira *et al.* (2003). A herança nestes *loci* era estritamente materna e a taxa de mutação foi definida em 10^{-4} , valor próximo a vários outros calculados na literatura (Tabela 1.2; Provan *et al.*, 1999; Navascués and Emerson, 2005).

Tabela 1.2 Parâmetros genéticos usados nas simulações de populações de *C. echinata* com o programa METAPOP.

Marcador Molecular	AFLP	Microssatélite de cloroplasto
Número de <i>loci</i>	200	10
Número de alelos/<i>locus</i>	2	10
Taxa de mutação	10^{-6}	10^{-4}

¹² Índice de diferenciação populacional baseado em polimorfismo genético equivalente ao F_{ST} .

O sistema de cruzamento foi testado com cinco diferentes taxas de autofecundação: 0, 25%, 50%, 75% e 100%, simulando desde uma espécie alógama até uma autógama passando por três sistemas mistos onde ocorre menos autofecundação do que fecundação cruzada, meio a meio, ou menos fecundação cruzada. O resultado mais próximo do experimental foi obtido com a simulação de uma espécie alógama (sem autofecundação), como esperado para Pau-Brasil (Neto *et al.*, 2005), e por isso este foi definido como sistema de cruzamento padrão para as outras simulações.

1.3.3 Dispersão do pólen e da semente

A dispersão da semente e do pólen assumiu o modelo *stepping stone* (Kimura and Weiss, 1964), no qual as populações são posicionadas em linha reta ocorrendo dispersão apenas entre indivíduos de populações vizinhas, mas não ocorrendo nenhuma troca com as outras populações (ver Figura 1.5). Este modelo é condizente com o mecanismo de dispersão de Pau-Brasil que é bastante restrito, limitado pela distância entre as populações.

A dispersão das plantas pode ser dividida em dois componentes, via pólen e via semente, e por este motivo foi utilizada uma razão entre fluxo de pólen e de semente como 10:1 (Lira *et al.* 2003). Seguindo sempre esta razão, foram testadas cinco diferentes taxas de dispersão, a fim de simular também a fragmentação da Mata Atlântica: $10^{-1}/10^{-2}$, $10^{-3}/10^{-4}$, $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ e zero (Anexo II). A anotação $10^{-1}/10^{-2}$ significa uma dispersão do pólen de 10^{-1} e da semente de 10^{-2} para as populações mais próximas, seguindo o modelo do *stepping stone*.

1.3.4 Simulações

Foram feitos quatro grupos de simulações usando o programa METAPOPOP e cada uma delas com 20 réplicas. Em todos os casos, foram computados dois índices de diferenciação genética (G_{ST}), um para cada marcador molecular, nuclear e citoplasmático, e o índice final utilizado nas análises foi calculado pela média das 20 réplicas de cada simulação. Estas simulações tinham como objetivo investigar (1) o efeito do sistema de cruzamento; (2) o impacto da intensidade de fragmentação (através da dispersão de pólen e semente); (3) o impacto do tempo de vida; e (4) o impacto do desmatamento recente; sobre a evolução da diferenciação genética nuclear e citoplasmática entre as populações, através dos índices de G_{ST} .

Nos primeiros três grupos de simulações, foi estudada a influência de cada parâmetro populacional (sistema de cruzamento, dispersão e tempo de vida) na diferenciação genética desde o último período glacial, assumindo que as populações não eram fragmentadas inicialmente e por isso não eram diferenciadas antes deste período. Então, o valor de G_{ST} inicial foi sempre próximo de zero, ou seja, todos os genótipos em todas as populações tinham aproximadamente as mesmas frequências alélicas.

Nestas três primeiras simulações, apenas um parâmetro foi analisado por vez usando como referência um cenário padrão, com ciclo de vida de 300 anos, taxa de autofecundação zero, dispersão de pólen à 10^{-3} e dispersão de semente à 10^{-4} . Estes dados padrões foram utilizados baseados em simulações preliminares e dados esperados para o Pau-Brasil. Para cada grupo de simulação foram feitas cinco simulações testando: taxas de autofecundação de 0%, 25%, 50%, 75% e 100%; taxas de dispersão do pólen e da semente de $10^{-1}/10^{-2}$, $10^{-3}/10^{-4}$, $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ e zero; e ciclos de vida de 30, 60, 120, 300 e 500 anos (Tabela 1.3). Estes

dados possibilitaram também inferir certas características ecológicas da biologia reprodutiva do Pau-Brasil que ainda não foram bem definidas ou estudadas.

Tabela 1.3 Parâmetros usados nas diferentes simulações sobre os efeitos de diferentes taxas de autofecundação, taxas de dispersão e ciclos de vida sobre a diferenciação genética entre populações. Valores sublinhados foram usados no cenário-padrão das simulações.

Taxas de autofecundação	<u>0</u> , 25%, 50%, 75%, 100%
Taxas de dispersão (pólen/semente)	$10^{-1}/10^{-2}$, <u>$10^{-3}/10^{-4}$</u> , $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ e zero
Ciclos de vida	30, 60, 120, <u>300</u> , 500 anos

A quarta e última simulação, analisou o impacto do desmatamento recente da Mata Atlântica, na evolução da diversidade genética do Pau-Brasil. Assumimos que o processo de forte desmatamento causou um isolamento completo das populações remanescentes desta espécie, por causa da destruição em massa da Mata Atlântica (Saatchi *et al.*, 2001). A fim de prever as conseqüências desta fragmentação, as simulações tinham índice de diferenciação inicial (G_{ST}) semelhantes àqueles obtidos experimentalmente com marcadores AFLP (Cardoso *et al.*, 2005) e microssatélites de cloroplasto (Lira *et al.*, 2003). Para isso, os genótipos finais obtidos depois de uma simulação de 20.000 anos usando os três parâmetros padrões (matriz de dispersão de $10^{-3}/10^{-4}$, alogamia e ciclo de vida de 300 anos), que forneceram os resultados mais próximos dos experimentais, foram usados como genótipos iniciais nestas simulações. E os parâmetros foram mantidos como padrão para taxa de autofecundação e ciclo de vida, mas alterados para taxas de dispersão, definidas agora como zero, antes de rodar o programa por mais 20.000 anos. Os valores de G_{ST} nuclear e G_{ST} citoplasmático de 20 replicatas foram registrados a cada 5.000 anos de simulação.

1.3.5 Análise dos resultados

Usando os dados de G_{ST} obtidos com as 20 réplicas de cada grupo de simulação, após 20.000 anos ou a cada 5.000 anos, no caso do quarto grupo de simulações, foram calculados a média e o desvio padrão de cada grupo de dados.

Posteriormente, testes estatísticos de regressão linear e análise de variância total e entre as médias com o teste de Tukey, usando 95% de intervalo de confiança, foram feitos através do programa R versão 2.6.0 (The R Foundation for Statistical Computing, Viena, Áustria, <http://www.Rproject.org>).

1.4 Resultados

O programa METAPOPOP foi utilizado para analisar o efeito de parâmetros ecológicos inerentes de cada espécie vegetal, como sistema de cruzamento, dispersão ou fragmentação do habitat e ciclo de vida, sobre a diferenciação genética nuclear e citoplasmática. Vários testes iniciais foram realizados até alcançarmos valores de G_{ST} mais próximos daqueles obtidos experimentalmente para as populações naturais de *Caesalpinia echinata*: G_{ST} nuclear = 0,62 (Cardoso *et al.*, 2005) e G_{ST} citoplasmático = 0,91 (Lira *et al.*, 2003). Os parâmetros demográficos, como tamanho equivalente entre juvenis e adultos e taxa de sobrevivência das classes juvenis foram ajustados para que os indivíduos pudessem alcançar a idade máxima mesmo com a baixa capacidade suporte ($k = 300$) das populações nas simulações.

Com base nos resultados destes primeiros testes foram definidos os parâmetros demográficos ideais para o Pau-Brasil, que fossem mais próximos da realidade das populações naturais desta espécie na Mata Atlântica. E posteriormente foram realizados quatro grupos de simulações, para cada situação: (1) sistema de cruzamento; (2) dispersão ou fragmentação; (3) ciclo de vida; e (4) fragmentação recente; abordadas separadamente a seguir.

1.4.1 Sistema de cruzamento

O primeiro grupo de simulações visou analisar o efeito de diferentes tipos de sistema de cruzamento sobre os níveis de diferenciação genética nuclear e citoplasmática. Foram testados cinco sistemas, desde alogamia até autogamia com mais três sistemas mistos, com taxas de autofecundação definidas em: 0, 25%, 50%, 75% e 100%. Os outros parâmetros

ecológicos foram mantidos como padrão com dispersão do pólen de 10^{-3} , dispersão da semente de 10^{-4} e ciclo de vida de 300 anos.

Os valores de G_{ST} nucleares foram: 0,63 para espécie alógama, 0,68 para 25%; 0,75 para 50%; 0,82 para 75% e 0,88 para 100% de autofecundação. A regressão linear, $r^2 = 0,8885$ ($P < 0,0001$) mostrou uma alta correlação entre o aumento da diferenciação genética nuclear e o aumento da taxa de autofecundação nas simulações (Figura 1.7). Os testes de variância e Tukey mostraram uma diferença significativa entre todos os valores de G_{ST} nucleares ($P < 0,0001$).

Como esperado, os valores médios do G_{ST} citoplasmático permaneceram inalterados nos diferentes tipos de sistema de cruzamento simulados pelo METAPOP (Figura 1.8). Os valores médios foram: 0,8681 para alogamia; 0,8681 para 25% de autofecundação; 0,8664 para 50%; 0,8803 para 75%; e 0,8632 para autogamia.

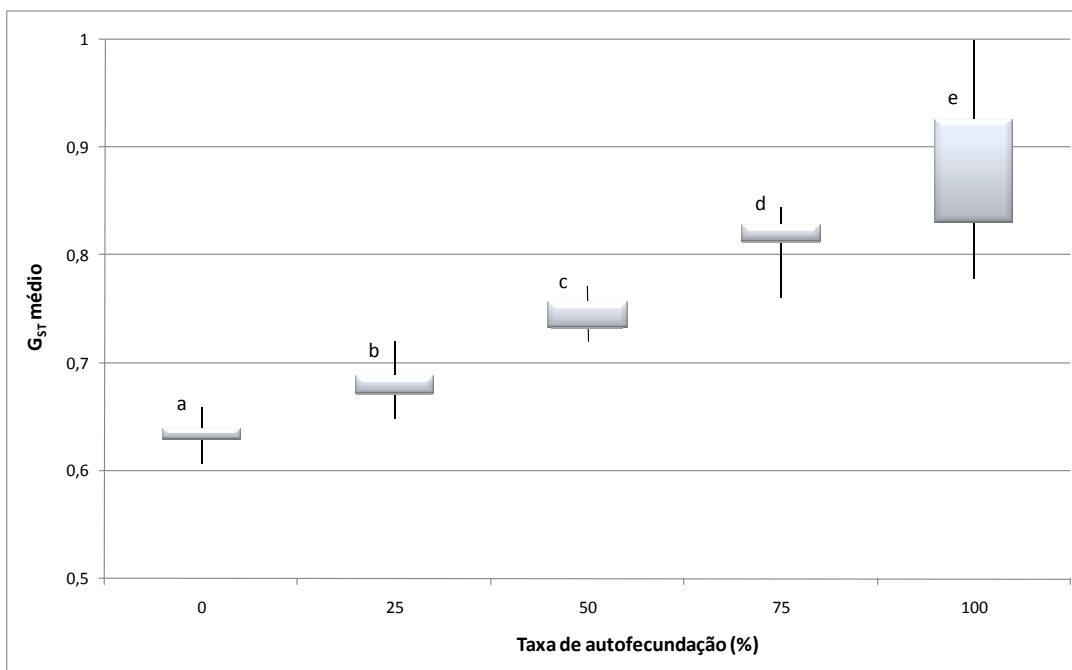


Figura 1.7 Gráfico com os valores médios de G_{ST} nuclear para as cinco taxas de autofecundação simuladas: 0, 25%, 50%, 75% e 100%. Barras de erro com valores mínimos e máximos de G_{ST} nas 20 réplicas de cada simulação. Letras diferentes indicam valores significativamente diferentes.

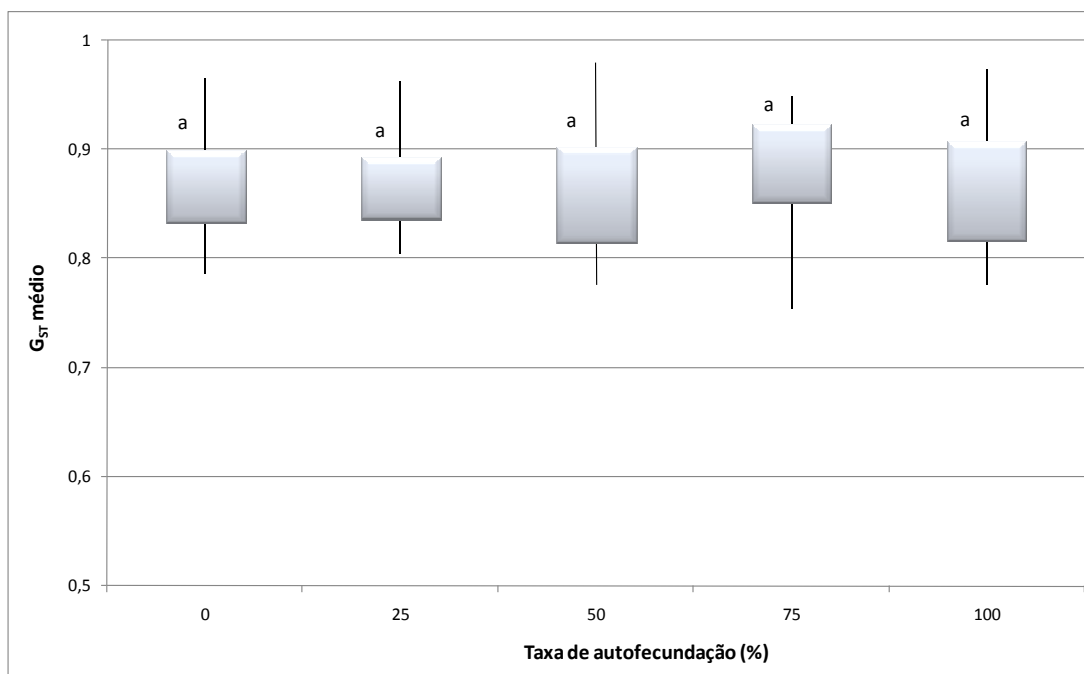


Figura 1.8 Gráfico com os valores médios de G_{ST} citoplasmático para as cinco taxas de autofecundação simuladas: 0, 25%, 50%, 75% e 100%. Barras de erro com valores mínimos e máximos de G_{ST} nas 20 réplicas de cada simulação. Mesma letra (a) indica que não houve diferenças significativas entre valores de G_{ST} .

1.4.2 Taxa de dispersão e fragmentação

Foram simuladas cinco taxas de dispersão de pólen e semente: $10^{-1}/10^{-2}$, $10^{-3}/10^{-4}$, $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ e zero, seguindo sempre a razão calculada para o Pau-Brasil entre pólen e semente de aproximadamente 10:1 (Lira *et al.*, 2003). Os valores de G_{ST} mostraram sofrer grande efeito causado por baixas taxas de dispersão, ou seja, maior fragmentação das populações. Os valores de G_{ST} nucleares e citoplasmáticos foram cada vez mais altos, quanto menor as taxas de dispersão do pólen e da semente simuladas.

Os valores médios de G_{ST} nuclear foram: 0,08 para dispersão $10^{-1}/10^{-2}$; 0,63 para $10^{-3}/10^{-4}$; e 0,79 para $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ e zero. Os valores de G_{ST} citoplasmático seguiram a mesma tendência, porém foram ainda mais altos, alcançando praticamente seu valor máximo de 1: 0,32 para $10^{-1}/10^{-2}$; 0,87 para $10^{-3}/10^{-4}$; e 0,92 para $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ e zero. A figura 1.9 mostra o gráfico dos valores de G_{ST} nucleares e citoplasmáticos para as diferentes taxas de dispersão. As regressões foram significativas com $r^2 = 0,9989$ ($P < 0,0001$) para os dados nucleares e $r^2 = 0,9522$ ($P < 0,0001$) para os dados citoplasmáticos.

O teste de Tukey mostrou que não havia diferença significativa apenas entre valores obtidos com as três menores taxas de dispersão ($10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ e zero), mas que as comparações entre os dados obtidos com dispersão $10^{-1}/10^{-2}$ e $10^{-3}/10^{-4}$ eram significativamente diferentes do restante. No caso dos marcadores nucleares, todas as diferenças foram altamente significativas ($P < 0,0001$). Por outro lado, com relação ao marcador citoplasmático, o valor de G_{ST} obtido com dispersão $10^{-1}/10^{-2}$ diferiu com $P < 0,0001$ dos outros valores, enquanto que o G_{ST} calculado para a taxa de dispersão $10^{-3}/10^{-4}$ diferiu dos valores de $10^{-5}/10^{-6}$ ($P = 0,011$), de $10^{-7}/10^{-8}$ ($P = 0,009$) e de zero ($P = 0,033$), com valores de P ainda dentro do erro de 5%.

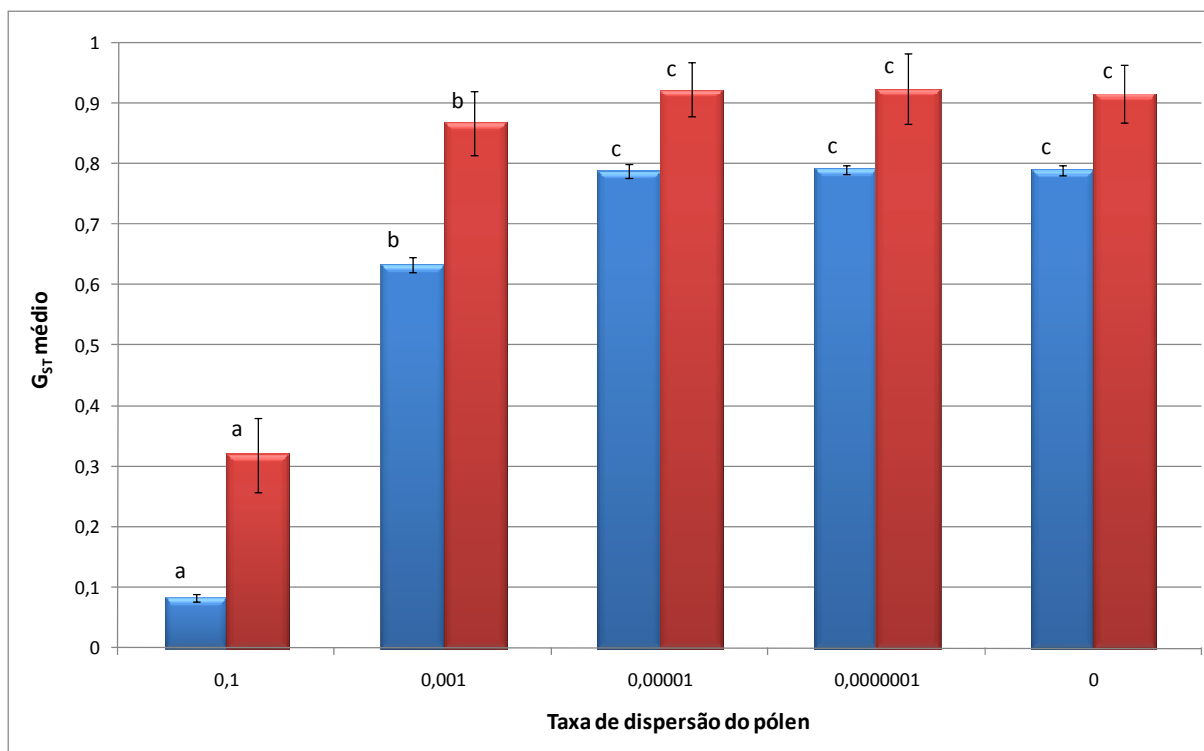


Figura 1.9 Gráfico com os valores médios de G_{ST} nuclear (azul) e citoplasmático (vermelho) para as taxas de dispersão testadas. Barras de erro com desvio padrão. Letras diferentes indicam diferenças significativas entre valores de G_{ST} .

1.4.3 Ciclo de vida

O terceiro grupo de simulações feitas no programa METAPOP simulou cinco ciclos de vida diferentes: 30, 60, 120, 300 e 500 anos. O G_{ST} nuclear médio foi 0,74 para 30 anos, 0,66 para 60 anos, 0,64 para 120 anos, 0,63 para 300 anos e 0,57 para 500 anos. E o G_{ST} citoplasmático médio foi 0,91 para 30 anos, 0,92 para 60 anos, 0,90 para 120 anos, 0,87 para 300 anos e 0,80 para 500 anos. Estes dados mostraram que quanto mais longo o ciclo de vida, menor a diferenciação genética principalmente nuclear, o que era esperado para espécies arbóreas.

Os valores de G_{ST} nuclear foram significativamente diferentes ($P < 0,0001$), exceto entre 120 e 300 anos ($P = 0,112$). E para os valores de G_{ST} citoplasmático, a diferença foi significativa apenas para os valores: 30 anos X 500 anos ($P < 0,0001$); 60 anos X 500 anos ($P < 0,0001$); 120 anos X 500 anos ($P < 0,0001$); 300 anos X 500 anos ($P = 0,001$); e 60 anos X 300 anos ($P = 0,026$).

A correlação entre longevidade e diferenciação genética inter-populacional foi muito forte para genes nucleares ($r^2 = 0,9582$, $P < 0,0001$), mas fraca para genes citoplasmáticos ($r^2 = 0,369$, $P < 0,0001$) (Figura 1.10).

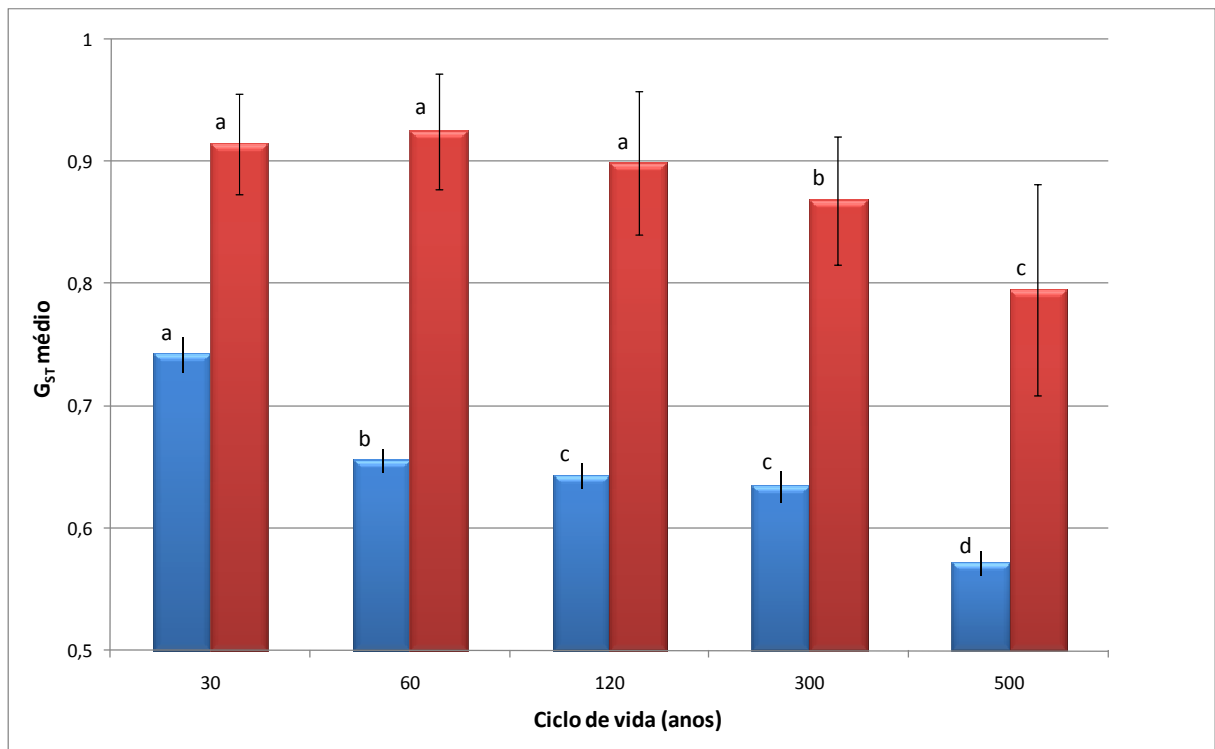


Figura 1.10 Gráfico com os valores médios de G_{ST} nuclear (azul) e citoplasmático (vermelho) para os diferentes ciclos de vida. Barras de erro com desvio padrão. Letras diferentes indicam diferenças significativas entre valores de G_{ST} .

1.4.4 Fragmentação recente

Este último conjunto de simulações tinha como objetivo prever o que acontecerá com a espécie *C. echinata* numa situação com grande fragmentação de seu habitat, na qual não ocorreria nenhum fluxo gênico entre as populações. Para isso, foram feitas simulações com valores iniciais de G_{ST} próximos aos encontrados experimentalmente com marcadores moleculares AFLP: 0,62 (Cardoso *et al.*, 2005) e microssatélites de cloroplasto: 0,91 (Lira *et al.*, 2003), e parâmetros padrões de sistema de cruzamento e ciclo de vida, porém com dispersões do pólen e da semente iguais a zero. Em última análise, estas simulações mostraram o que poderá acontecer com as populações naturais de *C. echinata* nos próximos 20.000 anos, um espaço de tempo curto em termos evolutivos, como resultado direto da fragmentação causada pelo desmatamento recente da Mata Atlântica.

Os valores médios de G_{ST} nuclear a cada 5.000 anos foram: 0,63 no tempo zero; 0,75 em 5.000 anos; 0,83 em 10.000 anos; 0,88 em 15.000 anos e 0,92 em 20.000 anos; todos significativos pelo teste de Tukey ($P < 0,0001$). Já os valores de G_{ST} citoplasmáticos variaram de 0,88 no tempo zero, para 0,93 em 5.000 anos; 0,95 em 10.000 anos e 0,96 em 15.000 e 20.000 anos; mantendo diferença significativa entre os valores de zero e 5.000 com relação ao restante ($P < 0,0001$), porém sem diferença significativa entre os três últimos tempos de 10.000, 15.000, 20.000. Figura 1.11 mostra um gráfico com os valores médios de G_{ST} nuclear e citoplasmático destas simulações.

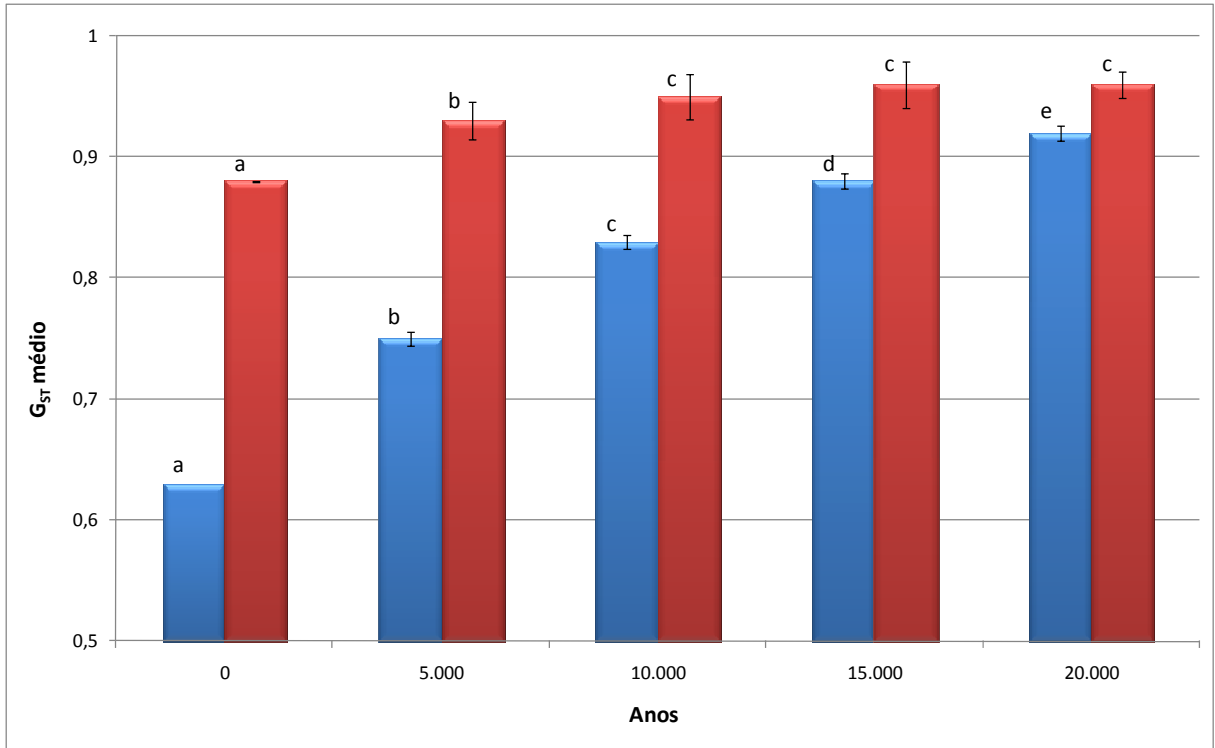


Figura 1.11 Gráfico com os valores médios de G_{ST} nuclear (azul) e citoplasmático (vermelho) durante a simulação de fragmentação recente, a cada 5.000 anos. Barras de erro com desvio padrão. Letras diferentes indicam diferenças significativas entre valores de G_{ST} .

1.5 Discussão

Foi usado um modelo demográfico e genético através do programa METAPOP para entender a influência de características ecológicas como sistema de cruzamento, dispersão ou fragmentação e ciclo de vida na diferenciação genética dentro de uma metapopulação. Nosso principal objetivo foi definir a contribuição de cada uma destas características em espécies arbóreas na diferenciação dos genes. Para isso, usamos como modelo a espécie *Caesalpinia echinata* (Pau-Brasil) para tentar também desvendar mais sobre a biologia reprodutiva ainda pouco estudada desta espécie. O Pau-Brasil ocorre no Complexo Mata Atlântica, porém em áreas mais secas que podem ter sofrido expansões e contrações durante os períodos de glaciações no Pleistoceno (Pennington *et al.* 2000). Com a ajuda de estudos genéticos realizados com populações naturais de Pau-Brasil no Brasil (Cardoso *et al.* 2005; Lira *et al.*, 2003), foi possível entender mais profundamente o comportamento desta espécie fazendo uso das simulações obtidas pelo METAPOP.

Apesar de pouco se saber sobre a biologia reprodutiva do Pau-Brasil, uma teoria foi postulada por Lira *et al.* (2003), na qual as populações naturais desta espécie estariam fragmentadas desde a última glaciação, em torno de 20.000 anos atrás, ou até antes disso. Esta hipótese explica a alta estruturação genética encontrada entre suas populações naturais: G_{ST} nuclear = 0,62 (Cardoso *et al.*, 2005) e G_{ST} citoplasmático = 0,91 (Lira *et al.*, 2003).

O programa METAPOP mostrou-se bastante eficiente no cálculo da diferenciação genética com marcadores nucleares e citoplasmáticos através do índice de G_{ST} . O mais importante é que este programa usa fatores evolutivos como mutação, deriva gênica e migração; para calcular estocasticamente a diferenciação dos genes com características programadas pelo usuário. Foram feitos quatro grupos de simulações com duração de 20.000

anos a fim de acessar o efeito de certas características ecológicas sobre a diferenciação dos genes desde a última glaciação até os dias atuais: (1) sistema de cruzamento, (2) taxa de dispersão ou fragmentação e (3) ciclo de vida; mas também dos dias atuais até 20.000 anos no futuro através do modelo (4) de fragmentação recente.

1.5.1 Sistema de cruzamento

O primeiro grupo de simulações analisou a diferenciação genética inter-populacional, através de marcadores nucleares e citoplasmáticos, comparando diferentes tipos de reprodução em espécies vegetais, permitindo definir a contribuição da autofecundação sobre a estrutura populacional causada pela diferenciação dos genes.

A diferenciação genética nuclear foi significativamente afetada pelo sistema de cruzamento, justificando 89% de toda a variação genética encontrada (Figura 1.7). Portanto, quanto maior a contribuição da autofecundação na biologia reprodutiva da espécie, maior a diferenciação entre as populações. Este resultado já era esperado porque o cruzamento entre parentes e a autofecundação tendem a homogeneizar os alelos dentro das populações, provocando uma diferenciação maior ainda entre elas, devido à falta de troca de informação genética entre as populações via fecundação cruzada.

Por outro lado, os marcadores citoplasmáticos não foram afetados pelo sistema reprodutivo (Figura 1.8). Isso se dá ao fato destes marcadores presentes nos cloroplastos serem herdados apenas maternalmente, nas angiospermas e, portanto, são passados de uma geração para a outra pela semente (Provan *et al.*, 2001), que geralmente possuem dispersão limitada, como no caso do Pau-Brasil.

Os resultados sugerem que o Pau-Brasil possui o seu sistema reprodutivo baseado em fecundações cruzadas, pois os valores de G_{ST} calculados pelas simulações que mais se aproximaram daqueles obtidos experimentalmente por Lira *et al.* (2003) (G_{ST} nuclear = 0,62) e Cardoso *et al.* (2005) (G_{ST} citoplasmático = 0,91) foram obtidos nas simulações com autofecundação igual a zero. Quanto maior a contribuição da autofecundação no sistema de cruzamento testado no METAPOP, menor foi a diferença relativa entre os valores de G_{ST} nucleares e citoplasmáticos, sugerindo ainda mais que o Pau-Brasil tenha preferência pela fecundação cruzada. Porém, acreditamos que, devido aos mecanismos de dispersão limitados do Pau-Brasil e em espécies correlatas juntamente com o fato de suas populações estarem distantes geograficamente umas das outras e fragmentadas, o Pau-Brasil seja uma espécie que faça muito endocruzamento, apesar de baixas taxas de autofecundação desta espécie (Neto *et al.* 2005).

1.5.2 Taxa de dispersão e fragmentação

O segundo grupo de simulações teve como objetivo analisar a dispersão via pólen e semente o que reflete diretamente na fragmentação do habitat de uma espécie. As taxas de dispersão utilizadas no METAPOP foram: $10^{-1}/10^{-2}$, $10^{-3}/10^{-4}$, $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ e zero. Foi utilizada a razão de 10:1 entre as taxas de dispersão do pólen e da semente, calculado por Lira *et al.* (2003) para o Pau-Brasil. Uma forte diferenciação genética nuclear e citoplasmática foi observada, relacionada à diminuição da dispersão, ou seja, aumento da fragmentação. Isto é resultado direto da diminuição do fluxo gênico que contrabalança os efeitos negativos da deriva gênica e da erosão genética de populações pequenas (Despres *et al.*, 2002).

A taxa de dispersão mais alta de $10^{-1}/10^{-2}$ levou a uma diferenciação genética menor

em ambos marcadores (G_{ST} nuclear = 0,08 e G_{ST} citoplasmático = 0,32), muito abaixo dos resultados experimentais do Pau-Brasil, e por isso fora da realidade desta espécie. Com a diminuição das taxas de dispersão, a diferenciação genética aumentou, seguindo uma correlação de aproximadamente 100% para os marcadores nucleares e 95% para os citoplasmáticos (Figura 1.9). Os marcadores citoplasmáticos sofreram maior diferenciação do que os marcadores nucleares, alcançando valores próximos ao valor máximo (1) na simulação com taxas de dispersão de $10^{-5}/10^{-6}$. Isto é resultado de uma forte deriva sobre os marcadores citoplasmáticos, cujo tamanho populacional efetivo é duas vezes menor quando comparado com marcadores nucleares em espécies hermafroditas, porque eles são transportados apenas pelas sementes e têm origem materna (Lira *et al.* 2003).

Resultados de G_{ST} nuclear e citoplasmático calculados experimentalmente para o Pau-Brasil foram similares aos resultados com taxas de dispersão moderada de $10^{-3}/10^{-4}$ para pólen e semente respectivamente. Os mecanismos de dispersão do Pau-Brasil são restritos: sementes são dispersas através da explosão do fruto (Carvalho, 1994) e pólenes são dispersos por abelhas dos gêneros *Xylocopa* e *Centris*, mas também pela espécie *Apis mellifera* (Borges *et al.*, 2008). No entanto, alguns estudos mostram que árvores em populações de baixa densidade demográfica, como o Pau-Brasil, recebem pólen de distâncias maiores do que plantas mais densamente povoadas promovendo a fecundação cruzada mesmo quando os polinizadores são pequenos insetos (Ward *et al.*, 2005; Gonzales *et al.*, 2006). Observações em campo e outros experimentos são necessários para resolver este ponto da biologia reprodutiva do Pau-Brasil.

1.5.3 Ciclo de vida

O efeito da longevidade na diferenciação dos marcadores nucleares e citoplasmáticos foi avaliado no terceiro grupo de simulações. Esta característica ecológica das plantas é importante na manutenção da diversidade genética, especialmente em ambientes fragmentados (Hamrick, 2004). O ciclo de vida mais longo provocou uma diminuição da diferenciação genética em ambos marcadores, especialmente para os nucleares que justificou 96% da variação genética encontrada (Figura 1.10). Por outro lado, os marcadores citoplasmáticos foram pouco afetados pelo ciclo de vida e por isso foram fracamente correlacionados com a diferenciação genética.

Valores de G_{ST} obtidos nas simulações de 300 anos foram equivalentes àqueles obtidos experimentalmente, indicando que o ciclo de vida do Pau-Brasil deve ser longo como pressuposto. A diferença entre os valores de G_{ST} para ambos os marcadores nas simulações de 300 e 500 anos foi significativa, mostrando que indivíduos de Pau-Brasil não viveriam até 500 anos. Mas os valores calculados para 120 e 300 anos de vida não foram significativamente diferentes, portanto sendo possível que esta espécie tenha o ciclo de vida com duração entre 120 e 300 anos.

1.5.4 Fragmentação recente

O quarto grupo de simulações tinha por objetivo calcular a diferenciação genética entre as populações naturais de Pau-Brasil dos dias atuais até 20.000 anos no futuro, para simular o efeito da fragmentação e destruição recentes de seu habitat. Atualmente, a Mata Atlântica é altamente fragmentada principalmente devido à ação antropogênica (Pennington *et*

al., 2004). Uma simulação usando os parâmetros-padrão, que são exatamente aqueles cujos resultados foram equivalentes aos dados experimentais de *C. echinata*, foi feita para que fossem utilizados os dados demográficos e genéticos gerados após 20.000 anos (que seria equivalente aos dados desta espécie na natureza) em uma nova simulação. Isto é, estaríamos utilizando dados similares aos atuais de valores demográficos e genéticos como iniciais para simulações que avaliam o efeito da fragmentação recente à qual as populações de *C. echinata* foram submetidas em seu habitat, que resultou no completo isolamento de suas populações.

Os resultados mostraram que a diferenciação de ambos marcadores nucleares e citoplasmáticos aumentou significativamente após 20.000 anos. Os marcadores citoplasmáticos alcançaram seu valor máximo de diferenciação após 10.000 anos, na metade da simulação, e assim se mantiveram até o final. Claramente, podemos extrapolar esta futura diferenciação genética obtida com as simulações para a situação do Pau-Brasil na natureza e também de muitas outras espécies do mesmo ambiente ou em situação parecida. Estas populações, com baixo fluxo gênico, podem estar submetidas ao processo de deriva ecológica e especiação (Lavin *et al.*, 2004). Torna-se importante, neste caso, quando é pouco provável que haja diminuição da diferenciação genética, tentar ao menos freá-la, a fim de manter o curso natural da evolução, mesmo que isso resulte na formação de novas espécies ou subespécies.

1.5.5 Questões conservacionistas

Populações remanescentes de uma espécie sem fluxo gênico tornam-se diferenciadas geneticamente e, conseqüentemente, diminuem suas chances de se adaptarem, tornando-se, portanto, mais vulneráveis às alterações ambientais e possivelmente podem ser ameaçadas de

extinção em seu próprio habitat (Yao *et al.*, 2007). Para contornar este tipo de situação, algumas técnicas de manejo e conservação podem ser empregadas em populações naturais.

Um método usado para aumentar a diversidade genética de populações é a transferência de indivíduos de uma população para outra, mas isso necessita de conhecimento dos processos de adaptação local para evitar uma possível depressão causada pela fecundação cruzada entre indivíduos muito diferenciados geneticamente (Ramp *et al.*, 2006; Raabova *et al.*, 2007). O fato das populações naturais de Pau-Brasil estarem altamente diferenciadas nas três regiões de ocorrência no Brasil (Cardoso *et al.*, 1998; Lira *et al.*, 2003; Cardoso *et al.*, 2005), junto com os resultados de fragmentação recente obtidos pelo METAPOPOP, torna esta técnica pouco favorável para a conservação do Pau-Brasil. A distância genética das populações naturais desta espécie deve ser levada em consideração para o sucesso de um reflorestamento.

A proteção e conservação *in situ* são ferramentas bastante promissoras para o Pau-Brasil e especialmente para espécies raras e ameaçadas de extinção de uma forma geral. Uma rede de áreas protegidas e o manejo sustentável de florestas são essenciais para a manutenção da variabilidade genética nas espécies florestais (Rajora and Mosseler, 2001). Porém, apenas estudos moleculares podem ajudar na determinação de áreas prioritárias de conservação, através da busca por haplótipos em baixa frequência, que deveriam mais urgentemente ser conservados (Castiñeiras *et al.*, 2007). Seguindo este princípio, no caso do Pau-Brasil seria necessário conservar no mínimo uma população de cada uma das três regiões que possuem haplótipos exclusivos investigados pelos marcadores microssatélites de cloroplasto (Lira *et al.*, 2003).

Capítulo 2

Análise genética e epigenética de espécies vegetais de manguezal

Manuscrito submetido ao periódico Ecology Letters (Anexo III):

DNA methylation of *Laguncularia racemosa* populations affected by different salinity regimes

2.1 Introdução

Manguezais são formações vegetais que ocorrem nas costas tropicais e subtropicais, e que diariamente são submetidos a variações de salinidade da água e oscilações do nível do mar (Saenger, 2003; Barth *et al.*, 2006). As espécies de manguezal têm ampla tolerância ambiental, ou seja, adaptam-se ao crescimento em diferentes ambientes. Por conseguinte, é natural que florestas de mangue em diferentes zonas geográficas possuam características estruturais distintas (Schaeffer-Novelli *et al.*, 1990). Em geral, manguezais atingem seu máximo desenvolvimento quando submetidos a grandes marés, mas com boa influência das águas fluviais, precipitação, nutrientes e sedimentos, alcançando mais de 40 metros de altura (Schaeffer-Novelli *et al.*, 1990). Por outro lado, manguezais também povoam regiões onde existem fatores limitantes, como secas periódicas, furacões, alta salinidade, e acúmulo de poluentes como metais pesados.

As espécies de mangue pertencem a famílias taxonomicamente distintas, porém desenvolveram especializações em comum tais como viviparidade¹³, exclusão de sal e raízes aéreas, para sobreviver às condições ambientais extremas dos manguezais (Tomlinson, 1986; Shi *et al.*, 2005). Embora o crescimento de espécies halófitas¹⁴ seja influenciado positivamente por salinidade de nível baixo a moderado (Ball, 2002), altos níveis de salinidade podem afetar negativamente processos metabólicos nestas espécies, tais como síntese protéica e fotossíntese, retardando o crescimento e o desenvolvimento vegetal (Parida

¹³ Fenômeno durante o qual o embrião dentro da semente se desenvolve ainda na planta-mãe, sendo chamado de propágulos no caso das espécies de mangue.

¹⁴ Plantas adaptadas quanto à morfologia e fisiologia para crescer em solos ricos em sal.

and Das, 2004, Parida *et al.*, 2004). Nestas condições ambientais de alto estresse salino, plantas formam florestas de estrutura reduzida, atingindo 1,5 a 3 metros de altura na forma de arbustos (Schaeffer-Novelli *et al.*, 1990).

Existem três principais gêneros de espécies bem adaptadas ao ecossistema manguezal: *Laguncularia* ou mangue branco, *Avicennia* ou mangue preto, e *Rhizophora* ou mangue vermelho. No Brasil, são encontradas no total seis espécies: *Laguncularia racemosa*, *Avicennia schaueriana*, *A. germinans*, *Rhizophora mangle*, *R. racemosa* e *R. harrisonii*. No estado do Rio, existem as espécies *L. racemosa*, *A. schaueriana*, *A. germinans* (apenas nos manguezais mais ao norte do Estado) e *R. mangle*. *L. racemosa* (Família: Combretaceae) é uma espécie cosmopolita, freqüente em manguezais do mundo todo, enquanto que *A. schaueriana* (Família: Avicenniaceae) é uma espécie endêmica do Brasil. O estudo genético destas duas espécies pode elucidar como ocorre o fluxo gênico entre as populações de manguezais de forma mais efetiva e quais os efeitos da seleção natural sobre a sobrevivência destas espécies em ambientes tão variáveis e limitantes.

O manguezal de Guaratiba, no município do Rio de Janeiro (RJ), encontrado na Baía de Sepetiba, possui uma formação de mangue chamada apicum¹⁵, regiões presentes mais para dentro do continente que sofrem inundações. Porém, por serem planícies, as águas da maré ficam retidas no apicum sofrendo evaporação devido ao forte calor das regiões tropicais, e conseqüentemente há acúmulo de sal no solo. Geralmente plantas que habitam as regiões de mangue próximas aos apicuns são menos desenvolvidas assumindo um porte arbustivo

¹⁵ Planícies hipersalinas que são áreas de inundação de maré com vegetação arbustiva adaptada para solos salinos.

(Saenger, 2003), provavelmente como consequência do forte estresse salino ao qual estão constantemente submetidas, que provocaria pressões seletivas divergentes para plantas localizadas próximas do apicum ou próximas do rio. A figura 2.1 mostra um esquema da fisionomia do manguezal de Guaratiba, desde a beira do rio Piracão até o apicum, com as três espécies de mangue presentes nesta área. Duas populações possuem indivíduos bastante diferentes morfologicamente e são separadas fisicamente pelo “Mar de *Rhizophora*”, uma região que se estende por 100 a 200 metros onde são encontrados somente indivíduos da espécie *R. mangle*.

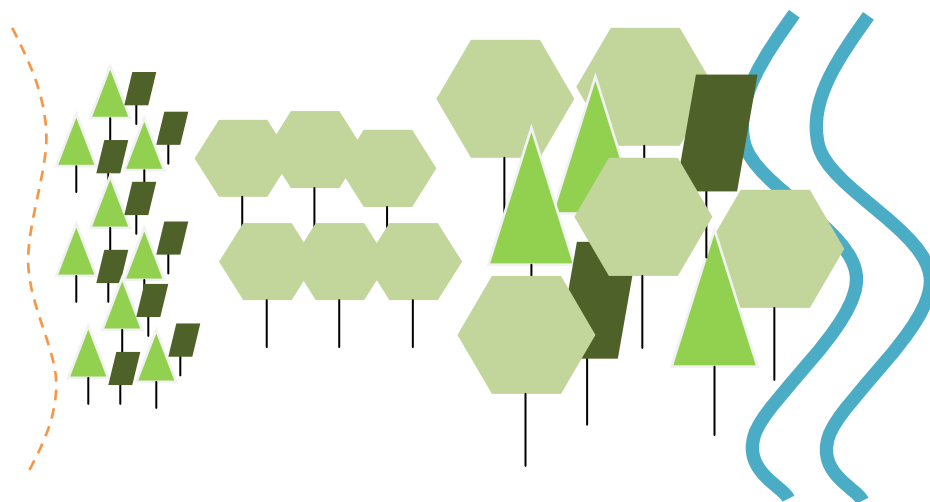


Figura 2.1 Esquema da fisionomia do manguezal de Guaratiba (RJ), desde o limite do apicum (tracejado à esquerda) até a beira do rio Piracão (em azul à direita). Próximas ao rio estão as plantas chamadas ribeirinhas com maior desenvolvimento vegetal, *L. racemosa* (retângulo verde escuro), *A. schaueriana* (triângulo verde claro) e *R. mangle* (hexágono), e próximas ao apicum estão as plantas menos desenvolvidas, somente *L. racemosa* e *A. schaueriana* são encontradas ali. Entre elas está representado o “Mar de *Rhizophora*”, separando as duas populações de *L. racemosa* e *A. schaueriana*.

No manguezal de Guaratiba, tanto *L. racemosa* quanto *A. schaueriana* localizadas na população próxima ao apicum, atingem no máximo 3 metros de altura e possuem uma série de galhos ramificados ao longo do tronco com caule pouco desenvolvido (Figura 2.2a). Por outro lado, indivíduos localizados na população próxima do rio, chamadas de ribeirinhas, possuem em torno de 10 metros de altura com galhos no dossel (Figura 2.2b). Divergências estruturais como estas de Guaratiba podem estar diretamente associadas à variação genética ou epigenética da população, como resultado da pressão seletiva de cada um dos ambientes.

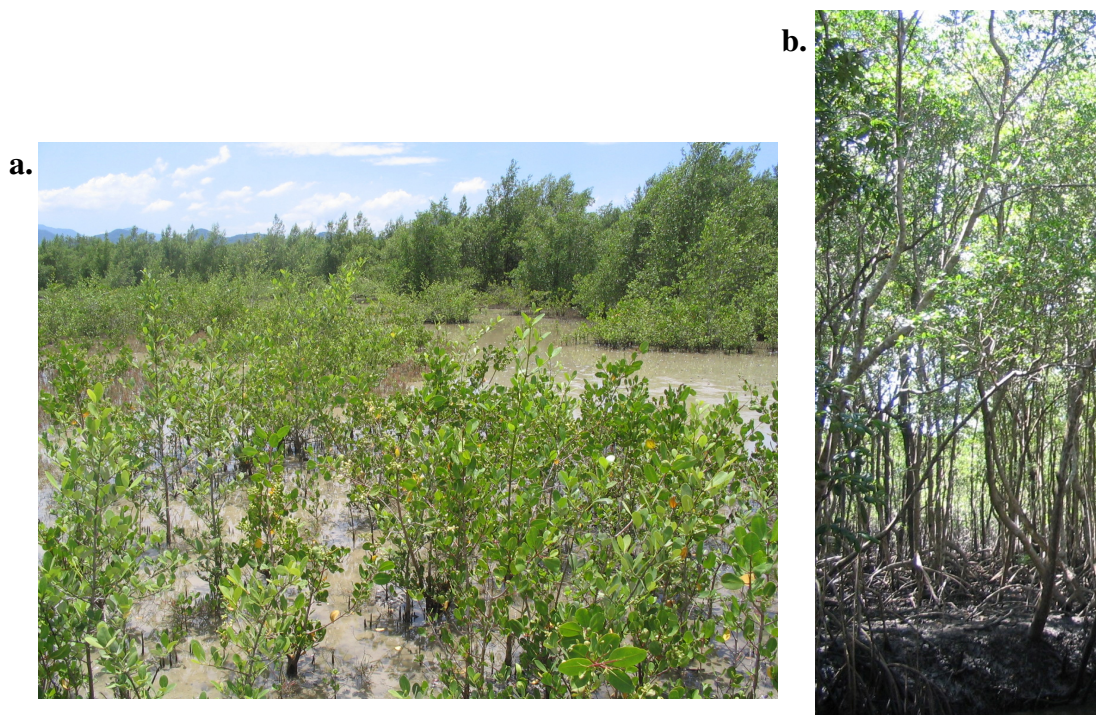


Figura 2.2 Fotos mostrando o contraste morfológico de plantas das duas populações de Guaratiba: **(a)** plantas de *L. racemosa* na região próxima ao apicum; e **(b)** indivíduos ribeirinhos de todas as espécies.

Marcadores moleculares AFLP (*Amplified Fragment Length Polymorphism*) são muito eficientes nos estudos de espécies selvagens cujo DNA ainda não foi seqüenciado, pois permitem a análise de grande quantidade de polimorfismo com o uso de pouca quantidade de DNA genômico, baixo custo inicial e resultados obtidos mais rapidamente (Ouborg *et al.*, 1999; Bensch and Åkesson 2005). Dentre os principais usos desta técnica estão: cálculo da diversidade genética, avaliação da estrutura populacional, busca por *loci* que afetem diretamente os fenótipos e formação de mapas de ligação genéticos (Bensch and Åkesson 2005).

Baseado nos marcadores AFLP, outro método que possibilita a análise da variação epigenética, chamado MSAP (*Methyl-Sensitive Amplified Polymorphism*), foi desenvolvido (Xiong *et al.* 1999). Esta técnica fornece dados genéticos e principalmente epigenéticos de forma muito parecida do AFLP. Através do uso de enzimas isosquisômeras, que reconhecem o mesmo sítio de restrição, mas possuem diferentes sensibilidades à metilação de citosinas do DNA (Xiong *et al.*, 1999; Azhikina and Sverdlov, 2005), torna-se possível identificar polimorfismo em regiões afetadas pela metilação do DNA (Cervera *et al.*, 2002). A maioria dos estudos até a presente data, que avaliaram alterações de metilação nas citosinas, estavam voltados para a herdabilidade dos padrões de metilação e níveis de polimorfismo em ecótipos ou estirpes de espécies cultivadas e/ou espécies poliplóides (Zhao *et al.*, 2007; Keyte *et al.*, 2006; Salmon *et al.*, 2005; Takata *et al.*, 2005; Hao *et al.*, 2004; Jaligot *et al.*, 2004). Porém estudos populacionais epigenéticos utilizando-se destes marcadores ainda não são freqüentes na literatura científica, apesar do grande potencial desta técnica (Bossdorf *et al.*, 2008).

A metilação das citosinas do DNA é uma das principais e melhor estudadas alterações epigenéticas (Schones and Zhao 2008). Ela é responsável por uma grande variedade de processos biológicos, incluindo a manutenção da estabilidade do genoma, através do controle da proliferação de elementos transponíveis; e a regulação da expressão gênica e da estrutura da cromatina (Finnegan *et al.*, 2000; Lee and Chen, 2001; Hao *et al.*, 2004, Lukens and Zhan 2007). Isto ocorre através da metilação de dinucleotídeos CG e, em menor frequência, de sítios não-CG como seqüências GNC e CNN (onde N é um A, T ou C) em plantas (Huettel *et al.* 2006; Zhang *et al.* 2006).

Alguns dos estudos realizados sobre a metilação de citosinas relacionando-a com o desenvolvimento de plantas mostraram a sua importância na formação das sementes (Huh *et al.* 2008), no desenvolvimento floral (Finnegan *et al.*, 2000; Habu *et al.*, 2001) e também na resposta ao estresse biótico e abiótico (Madlung and Comai, 2004). Outro fato muito interessante sobre a metilação do DNA em uma planta é que seus padrões podem ser herdados, sendo passados de uma geração para a outra (Bender, 2000; Gehring and Henikoff, 2007). Isso permite que a seleção natural atue sobre características moduladas por fenômenos epigenéticos sem que haja modificações nas seqüências de DNA (Rapp and Wendel, 2005). A figura 2.3 exemplifica casos onde características controladas por fenômenos epigenéticos geram novos fenótipos que são filtrados pela seleção natural.

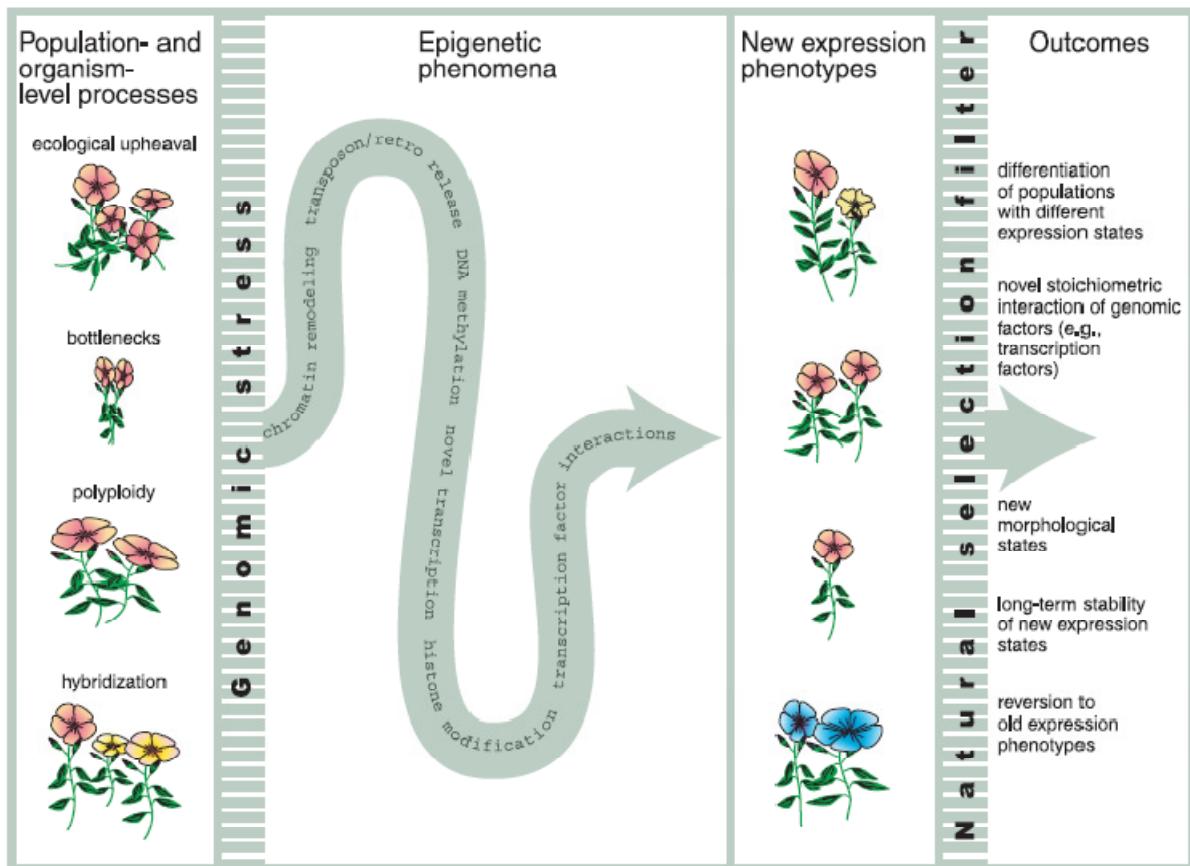


Figura 2.3 Esquema sobre a interação entre fenômenos epigenéticos e seleção natural. Uma série de processos a nível populacional (à esquerda) causa estresse genômico, levando à indução de fenômenos epigenéticos (grande seta ao centro). Estes vários fenômenos operam num contexto ecológico e evolutivo para produzir novos fenótipos (ao centro à direita), desde moleculares a morfológicos. Estes novos fenótipos são filtrados pela seleção natural, ou seja, aqueles que sobreviverem podem sofrer um processo evolutivo como retenção ou perda de estados epigenéticos fixados anteriormente. Fonte: Rapp and Wendell (2005).

As alterações epigenéticas possibilitam o desenvolvimento de novos fenótipos tanto morfológicos quanto moleculares, sem que ocorra nenhuma alteração de seqüência de nucleotídeos no genoma (Jablonka and Lamb, 1998). Ainda assim são responsáveis por uma parte das diferenças fenotípicas herdadas entre os indivíduos (Rapp and Wendel, 2005). Os fenômenos epigenéticos, conhecidos por controlarem a expressão de genes são metilação de citosinas, modificações covalentes de histonas de acetilação ou metilação e os processos mediados por RNA de interferência (Zhang 2008). Porém as suas implicações na evolução de populações naturais de plantas e seu desenvolvimento são ainda inexplorados (Kalisz and Purugganan, 2004).

Muitos avanços ocorreram com o objetivo de compreender melhor a seleção natural e seu efeito sobre as interações genótipo/ambiente, correlacionando a quantidade e estruturação da variação genética, variação ambiental e fenótipos (Le Rouzic and Carlborg, 2008). Com isso, novos conceitos de evolução, plasticidade ambiental e epigenética estão sendo estudados a fim de se formular uma nova Teoria Evolutiva mais completa e abrangente (Pigliucci, 2007). A Teoria Evolutiva Moderna foi postulada através da combinação da teoria da evolução de espécies por meio de seleção natural de Charles Darwin, a genética como base da herança biológica de Gregor Mendel e a genética populacional mais recente (detalhes em Pigliucci, 2007). Porém ainda estão fora desta teoria os novos conceitos sobre caracteres adquiridos ou assimilados resultantes de alterações epigenéticas, que apesar de soarem como idéias Lamarckistas atualmente existem indícios da sua existência e importância na adaptação do indivíduo ao ambiente. A epigenética tornou-se, portanto, alvo de interesse em estudos evolutivos, ligados a herança, adaptação e seleção natural.

A variabilidade total de um indivíduo é decorrente da parte epigenética do genoma e também da variação nos componentes genéticos, ou o próprio DNA, que possui potencial para formar novos fenótipos. Parte desta variabilidade do DNA está escondida porque não causa alterações fenotípicas ou seus efeitos só são visíveis sobre uma condição genética ou ambiental específica (Le Rouzic and Carlborg 2008), como resultado inclusive de fenômenos epigenéticos. Por isso, a relação entre a variabilidade genética e epigenética é bastante íntima, o que torna interessante o estudo comparativo entre elas numa população.

2.2 Objetivos

O objetivo geral deste capítulo foi avaliar a diversidade e estrutura genética e epigenética de duas populações de espécies de mangue, *L. racemosa* e *A. schaueriana*, em Guaratiba (Rio de Janeiro - RJ) através do uso de dois marcadores dominantes, o AFLP e o MSAP.

Os objetivos específicos foram:

- Avaliar a diversidade e a estrutura genética de indivíduos de *L. racemosa* das duas populações com marcadores AFLP;
- Avaliar a diversidade e a estrutura genética de indivíduos de *A. schaueriana* provenientes das duas populações com marcadores AFLP;
- Correlacionar os resultados obtidos com ambas as espécies usando marcadores AFLP;
- Avaliar a estrutura genética e epigenética entre as duas populações de *L. racemosa* com marcadores MSAP.

2.3 Materiais e Métodos

2.3.1 Material Vegetal

O alvo do estudo foram indivíduos de *L. racemosa* e *A. schaueriana* que ocorrem no manguezal de Guaratiba, no município do Rio de Janeiro (RJ – 23°00'S-23°05'S e 43°30'O-43°35'W – Figura 2.4). Este manguezal é caracterizado pela existência do apicum ou planície hipersalina, onde plantas próximas a esta área possuem um desenvolvimento arbustivo enquanto plantas ribeirinhas têm porte arbóreo. Foram realizadas diferentes coletas de folhas jovens e inteiras de plantas adultas aleatoriamente dispostas nas duas populações. Para as análises genéticas de *A. schaueriana* foram coletadas dez plantas ribeirinhas e doze do apicum, enquanto que de *L. racemosa* foram doze ribeirinhas e nove do apicum. Para as análises de metilação do DNA, pelo MSAP, foram amostradas 17 árvores adultas de *L. racemosa* de cada população. As amostras foram imediatamente acondicionadas em sílica gel e conservadas à temperatura ambiente até a extração do DNA.

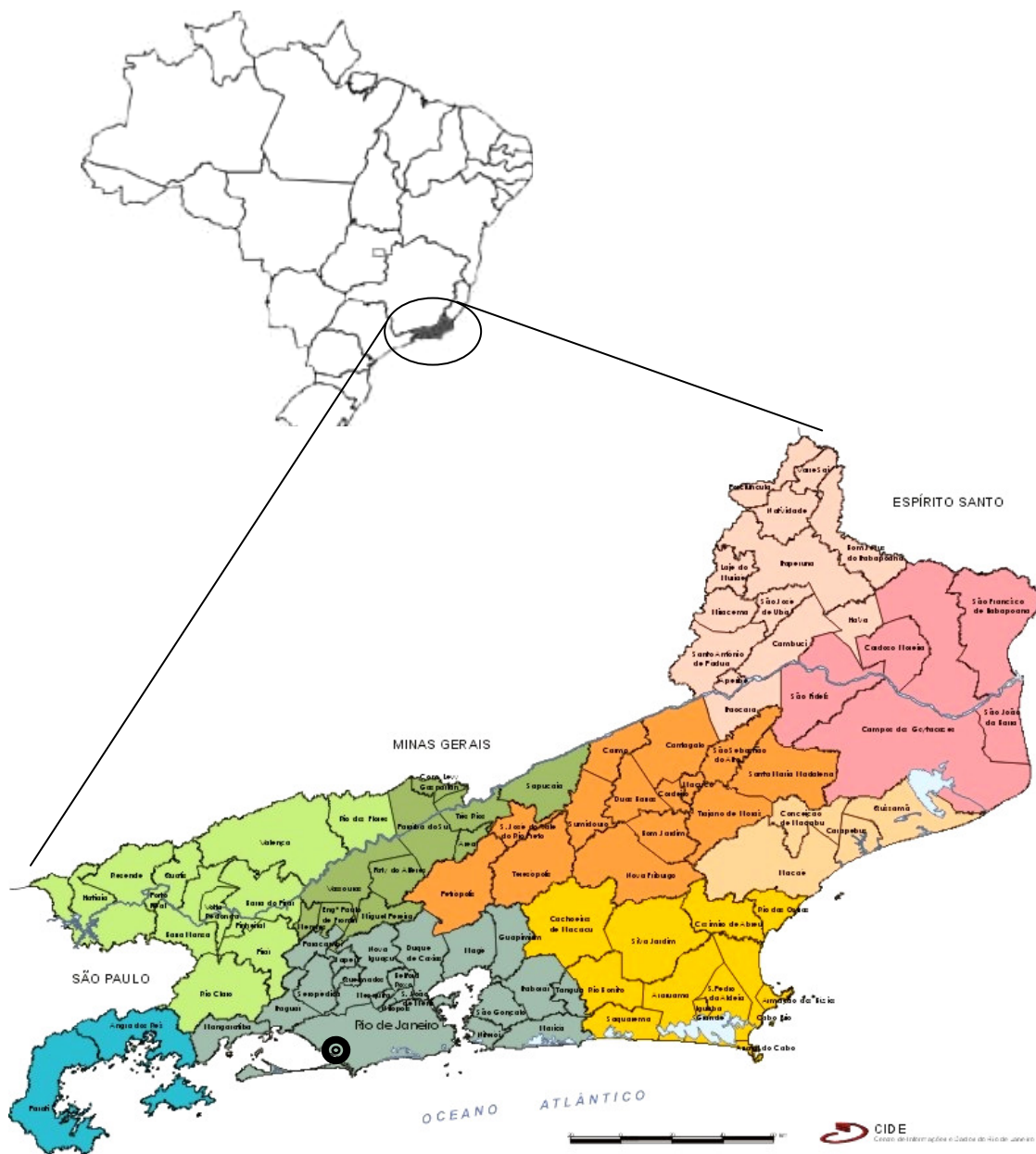


Figura 2.4 Mapa do Brasil com o Estado do Rio de Janeiro em destaque. Na Baía de Sepetiba, fica o manguezal de Guaratiba, marcado com um círculo preto, dentro do município do Rio de Janeiro.

2.3.2 Extração do DNA

Um total de 50mg de folhas secas em sílica-gel de cada indivíduo de *L. racemosa* e *A. schaueriana* coletados foram maceradas no mixer Retsch MM 301 por 1 minuto a 30 Hz. Foram adicionados 700µl de tampão MATAB (Tris-HCl 0,1M pH 8,0, NaCl 1,25M, EDTA 0,02M, MATAB 4%) pré-aquecido com β-mercaptoetanol (0,1%), para as amostras de *L. racemosa* foram ainda adicionados 10mg de PVP 40.000. As amostras foram incubadas a 65°C por 1 hora e posteriormente submetidas a duas extrações com 600µl de clorofórmio: álcool isoamílico (24:1) seguidos de agitação manual por 10 minutos e centrifugação a 1530 rcf por 10 minutos. Na seqüência, a primeira precipitação foi feita com 0,8 volumes de isopropanol gelado, seguido de centrifugação 12.000 rcf por 15 minutos. Após solubilização do precipitado em água estéril, a segunda precipitação foi feita com 0,1 volumes de NaCl 5M e 2 volumes de etanol absoluto gelado. Estas precipitações foram incubadas por 30 minutos no freezer (-20°C) e lavadas com etanol 70%. O DNA foi dissolvido em 100µl de água estéril. A qualidade e a quantidade de DNA foram verificadas em gel de agarose 1%.

2.3.3 AFLP

2.3.3.1 Protocolo de amplificação

O AFLP foi feito como descrito por Vos *et al.* (1995) usando o kit AFLP™ System Analysis Kit (Life Technologies, Inc.) com modificações descritas por Cardoso *et al.* (2005). Esta técnica se baseia no corte enzimático do DNA genômico a fim de obter marcadores nucleares neutros e bem distribuídos ao longo de todo o genoma sem conhecimento prévio de qualquer seqüência de nucleotídeos (Figura 2.5).

Utilizamos 300ng de DNA genômico nas digestões de volume final de 15µl com as enzimas *EcoRI* e *MseI* usando 10U de cada enzima. As digestões ocorreram a 37°C por 3 horas e depois as enzimas foram desnaturadas por 15 minutos a 70°C. Após a confirmação da digestão total de cada amostra através de géis de agarose 1%, foram ligados adaptadores de cada enzima usando 6µl da solução de adaptadores, 6µl do DNA digerido e mais 0,3µl da enzima T4 DNA ligase. Esta reação foi incubada a 20°C por 2 horas.

Reações de pré-amplificações foram feitas usando os *primers*¹⁶ complementares aos adaptadores de *EcoRI* e *MseI*, cada um com uma base seletiva. Usamos o DNA ligado diluído de 5 a 10 vezes, dependendo da intensidade da digestão verificada no gel, com o objetivo de manter as amostras com aproximadamente a mesma concentração de DNA. O programa de PCR da pré-amplificação consistiu em 29 ciclos de 30 segundos a 94°C, 60 segundos a 56°C e 60 segundos a 72°C, com extensão final de 2 minutos. Os produtos do PCR foram checados em gel de agarose 1% e diluídos proporcionalmente de 10 a 50 vezes, para que todas as amostras ficassem com concentrações de DNA similares.

¹⁶ Oligonucleotídeos de DNA usados como iniciadores numa reação de PCR.

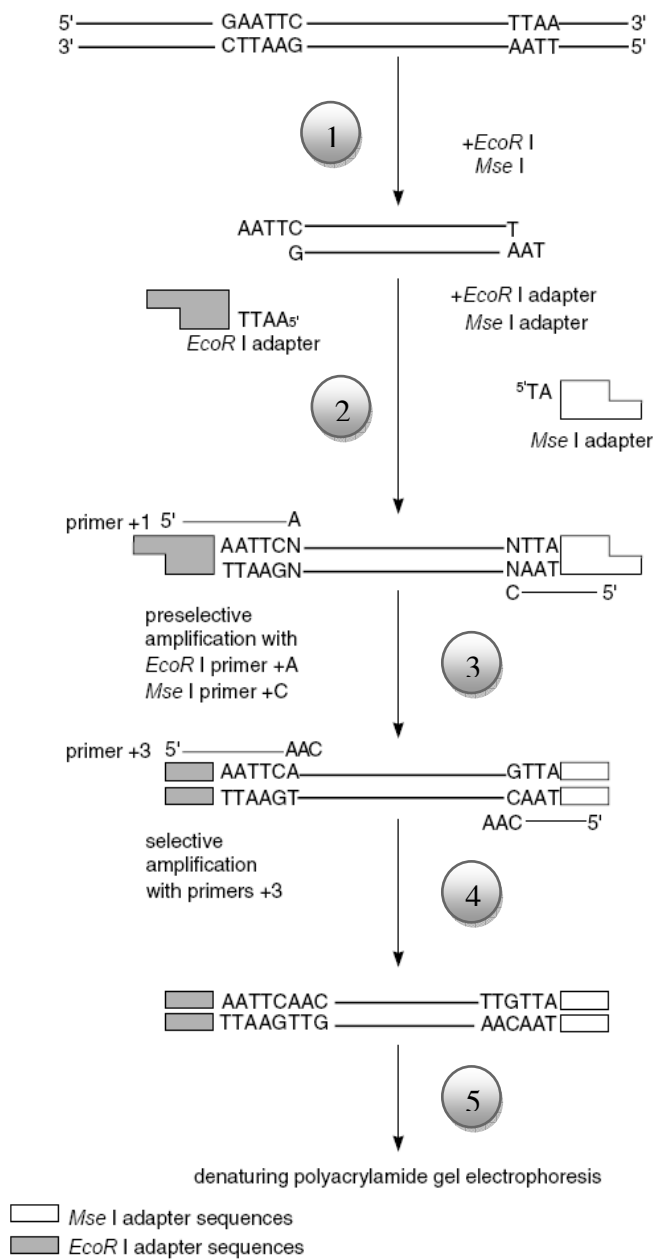


Figura 2.5 Esquema do protocolo de AFLP, mostrando as principais etapas desta técnica. (1) Digestão do DNA genômico; (2) Ligação dos adaptadores; (3) Pré-amplificação; (4) Amplificação final; e (5) Separação por eletroforese. Fonte: Invitrogen Life Technologies, Inc.

A reação de amplificação final foi feita em 20µl de volume final. O *primer* complementar ao adaptador *EcoRI* foi marcado radioativamente com 1µCi de $\gamma^{32}\text{P-ATP}$ e 0,2U de enzima T4 Polinucleotídeo Kinase por reação. Estas marcações ficaram a 37°C por 1 hora, depois a 70°C por 10 minutos e depois mais 10 minutos no gelo. Para a reação de PCR usamos 5µl do DNA pré-amplificado e diluído, *primers* complementares aos adaptadores *EcoRI* e *MseI* com três bases seletivas num volume final de 20µl. O programa desta amplificação é do tipo *touchdown*¹⁷, ou seja, ocorre diminuição da temperatura de anelamento a cada ciclo até atingir a temperatura desejada, que é mantida nos ciclos seguintes. Os primeiros 12 ciclos tiveram etapas de 30 segundos a 94°C, 30 segundos de anelamento, começando em 65°C e diminuindo 0,7°C por ciclo e 1 minuto a 72°C. O programa continuou por mais 25 ciclos com temperatura de anelamento de 56°C. As combinações de *primers* utilizadas foram *EcoRI-ACC/MseI-CAT*, *EcoRI-ACG/MseI-CAG*, *EcoRI-AAG/MseI-CAT*.

Os resultados foram analisados após separação por eletroforese a 60W por 2 horas, em gel de poliacrilamida 6% desnaturante com 7M de uréia e TBE 1X. O gel foi exposto a um filme tipo Kodak XK1 de um a dois dias. A revelação do filme foi feita usando soluções de revelador e fixador da Kodak.

¹⁷ Programa de PCR que reduz o anelamento de *primers* não-específicos pela diminuição gradual de temperatura de anelamento entre os ciclos.

2.3.3.2 Análise dos dados

Os dados foram analisados pela presença (1) ou ausência (0) de bandas nos géis. Apenas fragmentos claramente marcados radioativamente foram usados nas análises estatísticas. Os valores de diversidade dentro das populações foram calculados pelos índices de diversidade de Nei (Nei, 1973) e de Shannon usando o programa POPGENE (versão 1.21; Yeh *et al.*, 1997). Além disso, com o mesmo programa obtivemos os índices de similaridade e distância genética calculados segundo Nei (1978) e o índice de diferenciação G_{ST} . Estas análises assumem o equilíbrio de Hardy-Weinberg, o que não é aconselhável quando se usa marcadores dominantes, como o AFLP, porque pressupõe que não existe endocruzamento na população calculando assim as frequências alélicas (Foll *et al.*, 2008).

A fim de contornar esta situação e aumentar a robustez dos resultados foram realizados testes estatísticos do tipo Bayesiano e de multivariável. A análise Bayesiana usa dados de marcadores dominantes para calcular as frequências alélicas e os valores de endocruzamento, assumindo o desvio do equilíbrio de Hardy-Weinberg (Holsinger *et al.*, 2002; Foll *et al.*, 2008). O índice de diferenciação bayesiano, chamado θ_{ST} , foi calculado pelo programa Hickory versão 1.1 que fez a análise através da cadeia MCMC (*Monte-Carlo Markov Chain*) após gerar 250.000 cadeias e descartar as 50.000 primeiras (Holsinger *et al.*, 2002). O modelo usado foi *full model* que incorpora incertezas tanto sobre os valores de F_{ST} quanto de F_{IS} ¹⁸, o coeficiente de endogamia.

¹⁸ Índice que representa a taxa de endogamia nas populações de uma dada espécie.

O outro índice de diferenciação, chamado de β_{ST} , foi calculado pela análise de multivariável através do programa ADE-4 (Thioulouse *et al.*, 1996). Esta análise ignora o equilíbrio de Hardy-Weinberg, simplesmente porque se baseia na presença ou ausência de bandas ao invés das frequências alélicas.

As divergências significativas entre os índices de diversidade genética de Shannon e de Nei foram avaliadas por testes de *Kruskal-Wallis rank sum test*¹⁹, usando o programa R versão 2.6.0 (The R Foundation for Statistical Computing, Viena, Áustria, <http://www.Rproject.org>). Os valores de P foram corrigidos pelo método de Bonferroni através da multiplicação do valor de P calculado pelo teste de *Kruskal-Wallis rank sum test* pelo número total de genes (Rice, 1989).

2.3.4 MSAP

2.3.4.1 Protocolo de amplificação

A análise dos níveis de metilação no DNA foi realizada usando a técnica do MSAP, como recomendado por Xiong *et al.* (1999). Este protocolo é uma modificação do AFLP (Vos *et al.*, 1995), onde as digestões por enzimas de corte freqüente são realizadas em paralelo com dois isosquisômeros *MspI* e *HpaII*. Estas enzimas reconhecem e clivam a mesma seqüência de nucleotídeos (5'-CCGG-3'), mas mostram diferentes sensibilidades à metilação no DNA.

¹⁹ Teste estatístico de significância que compara cada elemento ao invés de usar a média.

Enquanto *HpaII* é sensível à metilação de qualquer citosina em ambas as fitas de DNA, *MspI* cliva citosinas internas metiladas (5'-C^mCGG-3'), mas não as externas. Por conseguinte, a comparação entre os perfis de marcadores MSAP gerados pelos dois isosquisômeros permite avaliar o estado de metilação das citosinas internas nas ilhas CpG dos sítios de restrição, permitindo analisar a estrutura populacional genética e epigenética de *L. racemosa*. No entanto, *MspI* é sensível à metilação da citosina externa (5'-5^mCCGG-3', conhecido como metilação assimétrica de CCG ou hemimetilação), e ambos *HpaII* e *MspI* são sensíveis à metilação completa de ambas as citosinas, levando a subestimar o nível de metilação global do genoma (Figura 2.6).

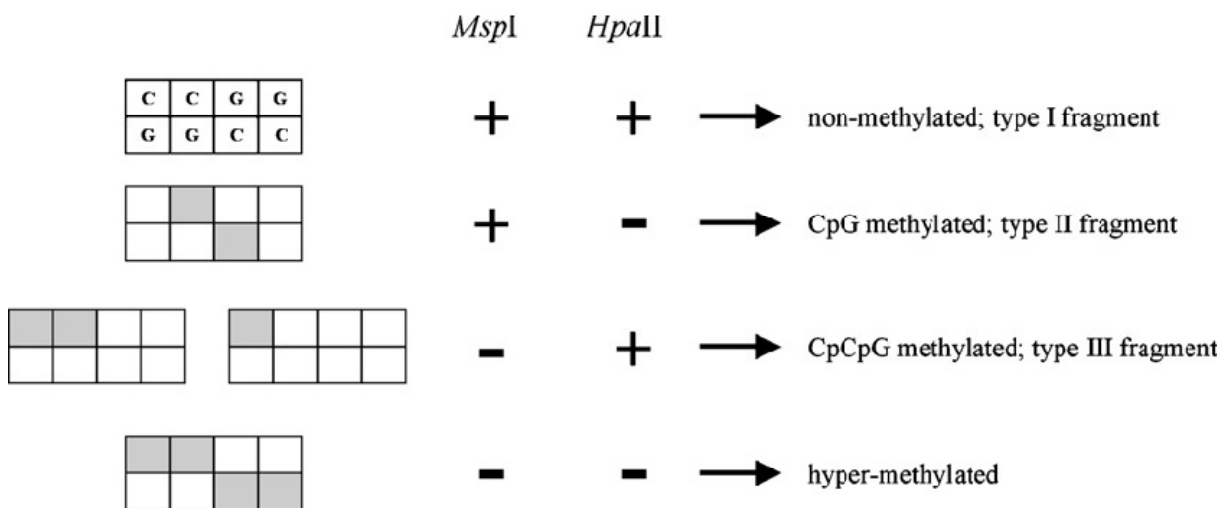


Figura 2.6 Tipos de fragmentos obtidos pela técnica do MSAP, devido à diferença de sensibilidade dos isosquisômeros *MspI* e *HpaII*: (+) clivagem da enzima, e (-) não há clivagem. Áreas em cinza no esquema representam citosinas metiladas. Fonte: Salmon *et al.* (2008).

O adaptador da enzima *EcoRI* (5'-CTCGTAGACTGCGTACC-3' e 5'-AATTGGTACGCAGTCTAC-3') e a sequência básica dos *primers* complementares a ele (5'-GACTGCGTACCAATTC-3') foram os mesmos que os utilizados no protocolo original do AFLP (Vos *et al.*, 1995). Os adaptadores das enzimas *MspI* e *HpaII* (5'-GATCATGAGTCCTGCT-3' e 5'-CGAGCAGGACTCATGA-3') e a sequência básica dos *primers* (5'-ATCATGAGTCCTGCTCGG-3') foram descritos por Xiong *et al.* (1999).

As amostras foram submetidas a digestão com a enzima *EcoRI* usando 1µg de DNA genômico e 10U de enzima (Promega^R) em tampão H 1X (Promega^R) num volume final de 200µl. Os DNAs digeridos foram precipitados utilizando 0,1 volume de acetato de sódio 3M e 2,5 volumes de etanol 100% e depois lavado em etanol 70%. Metade do DNA digerido com *EcoRI* foi usado em cada uma das reações seguintes com 5U de *MspI* ou *HpaII* (Promega^R) em tampão multicore ou tampão B 1X (Promega^R), respectivamente, num volume final de 50µl. As incubações das digestões foram todas realizadas a 37°C durante 6 horas e as enzimas foram posteriormente desnaturadas a 65°C durante 20 minutos. Assim, cada amostra tinha dois perfis *MspI*-MSAP e *HpaII*-MSAP, gerados a partir das digestões com *EcoRI/MspI* e *EcoRI/HpaII*, respectivamente, para serem amplificados e analisados. A ligação dos adaptadores foi realizada com 20µl do DNA digerido, tampão para T4 DNA ligase 1X (Promega^R), 1U de T4 DNA ligase (Promega^R), 5pmol de cada adaptador *EcoRI* e 50pmol de cada adaptador *MspI/HpaII* numa reação de 30µl a 20°C por 3 horas.

As pré-amplificações foram realizadas em reações de 20µl usando 2µl do DNA digerido e ligado, tampão de PCR 1X, 0,4mM de dNTPs, 30ng do *primer* básico de *EcoRI* e *MspI/HpaII* e 2U de Taq polimerase (Ludwig). A reação ocorreu em 25 ciclos de 94°C por 1

minuto, 56°C por 1 minuto e 72°C por 2 minutos, com uma extensão final de 10 minutos. Os produtos da pré-amplificação foram diluídos 20 vezes e 5µl destas amostras foram utilizadas nas amplificações seletivas finais. Estas reações de 20µl, por sua vez, continham 30ng de cada *primer* seletivo *EcoRI* e *MspI/HpaII*, 0,2mM dNTPs, tampão de PCR 1X e 2U de Taq polimerase (Ludwig). O programa do tipo *touchdown* começou a 94°C por 30 segundos, 65°C por 30 segundos e 72°C por 1 minuto com decréscimo de 0,7°C na temperatura de anelamento por ciclo durante 12 ciclos, seguidos de 24 ciclos de 94°C por 30 segundos, 56°C por 1 minuto e 72°C por 2 minutos, com extensão final de 5 minutos. Quatro combinações de *primers* seletivos foram utilizados: *EcoRI*+AG com *MspI/HpaII*+TCAA, *EcoRI*+AC com *MspI/HpaII*+TCAA, *EcoRI*+AAC com *MspI/HpaII*+TCAA e *EcoRI*+AG com *MspI/HpaII*+AAT.

Os produtos amplificados foram separados por eletroforese por 2 horas e 30 minutos a 60W em gel de poliacrilamida 4% com 7,5M de uréia. Os géis foram corados com uma solução de nitrato de prata 0,1% acrescida de formaldeído 0,5% por 30 minutos após a fixação do gel em solução de ácido acético 10% por 20 minutos. Para a revelação utilizou-se uma solução de carbonato de sódio 6% com formaldeído 0,5% e 2µg de tiosulfato de sódio por 3 minutos e em seguida a revelação foi interrompida pela adição de ácido acético 10%. Os géis corados com prata foram documentados através de imagens para posterior análise.

2.3.4.2 Análise dos dados

As 34 amostras foram analisadas segundo a presença (1) ou ausência (0) de fragmentos amplificados pela técnica do MSAP a partir de ambas as clivagens *EcoRI/MspI* e *EcoRI/HpaII*, respectivamente chamados de *MspI*-MSAP e *HpaII*-MSAP. A presença de fragmentos em ambos os perfis *MspI*-MSAP e *HpaII*-MSAP indicavam sítios não-metilados, enquanto que sítios metilados foram caracterizados pela presença de fragmentos *MspI*-MSAP, mas ausência dos mesmos em *HpaII*-MSAP (Figura 2.6; Peraza-Echevarria *et al.*, 2001). A presença de fragmentos no perfil *HpaII*-MSAP que estavam ausentes em *MspI*-MSAP representaram hemimetilações da citosina externa (Salmon *et al.*, 2005) e foram excluídos das análises por serem um estado raro e geralmente não herdado entre as plantas (Figura 2.6; Salmon *et al.*, 2008).

Tanto metilados quanto não-metilados, os *loci* foram considerados polimórficos quando, pelo menos, uma amostra divergiu das outras com relação à presença ou ausência de fragmentos nos géis. Os *loci* polimórficos de *MspI*-MSAP foram analisados como o componente genético de *L. racemosa*, enquanto que os *loci* polimórficos de *HpaII*-MSAP representaram o componente epigenético.

Marcadores dominantes, tais como o MSAP, não permitem a definição de heterozigotos diretamente pelos resultados moleculares. Portanto a fim de estimar as estruturas populacionais genéticas e epigenéticas de *L. racemosa*, três abordagens estatísticas foram feitas.

Em primeiro lugar, perfis individuais foram investigados pela análise de multivariável, uma abordagem baseada na presença ou ausência de bandas que não assume o equilíbrio de Hardy-Weinberg, pois não calcula as frequências alélicas. A análise de componente principal (PCA - *Principal Component Analysis*) foi feita entre os perfis individuais sobre uma matriz de covariância seguida da análise feita entre os grupos, chamada EigenAnalysis (BPCA, ver Parisod and Christin 2008 para detalhes), computada com os dados de *MspI*-MSAP e *HpaII*-MSAP utilizando o programa ADE-4 (Thioulouse *et al.*, 1996). BPCA calcula um índice de diferenciação genética chamado β_{ST} entre as populações para ambos os dados genéticos e epigenéticos. O significado estatístico deste índice foi avaliado pelo teste de randomização Romesburg com 9999 permutações.

A diversidade genética foi avaliada pelo índice de diversidade de Shannon calculado com base na frequência de cada *locus* dos 17 indivíduos para cada perfil: *MspI*-MSAP e *HpaII*-MSAP. Tal como recomendado por Bussel (1999), $\log_2(0)$ foi substituído por 0 para bandas ausentes fixas ($\text{freq} = 0$). Com base no índice de diversidade Shannon, o índice de diferenciação G_{ST} foi calculado assumindo que as populações estão em equilíbrio de Hardy-Weinberg (Bussel, 1999).

Por último, as análises estatísticas Bayesianas foram utilizadas por estimar o desvio do equilíbrio de Hardy-Weinberg mesmo quando são usados marcadores dominantes (Holsinger *et al.*, 2002; Parisod *et al.*, 2005). Na verdade, o cálculo Bayesiano da estrutura genética incorpora a incerteza sobre os valores de F_{ST} e F_{IS} através do modelo *full model*. A cadeia MCMC foi gerada 250.000 vezes, a fim de garantir a convergência da cadeia em sua distribuição estacionária, sendo rejeitadas as primeiras 50.000 simulações. Através do

programa Hickory versão 1.1 (Holsinger *et al.*, 2002), o índice de diferenciação θ_{ST} foi estimado com suas variâncias e um intervalo de confiança de 95%.

A fim de contrastar as estruturas genética e epigenética obtidas através dos dados de *MspI*-MSAP e *HpaII*-MSAP foi feita uma análise de Co-Inércia utilizando o programa ADE-4 (Thioulouse *et al.*, 1996). A significância desta associação foi testada pela randomização Romesburg com 9999 permutações.

As divergências significativas entre os índices de diversidade de Shannon baseada nos perfis de *MspI*-MSAP e *HpaII*-MSAP em ambas as populações foram avaliadas por testes de *Kruskal-Wallis rank sum test* usando o programa R versão 2.6.0 (The R Foundation for Statistical Computing, Viena, Áustria, <http://www.Rproject.org>). Os valores de P foram corrigidos pelo método de Bonferroni (Rice, 1989).

2.4 Resultados

2.4.1 Marcadores AFLP

Três pares de *primers* foram utilizados para obter marcadores AFLP a fim de analisar a diversidade genética de duas espécies de mangue *Laguncularia racemosa* e *Avicennia schaueriana*. A primeira espécie tem maior distribuição geográfica, ocorrendo no continente Americano e Africano. A segunda ocorre especialmente no Brasil e em alguns pontos da América Central. No Estado do Rio de Janeiro, ambas ocorrem no manguezal de Guaratiba, na Baía de Sepetiba. Curiosamente neste local, ambas as espécies possuem diferenças morfológicas bastante evidentes entre indivíduos que ocorrem na população ribeirinha ou próxima do apicum mesmo estando estas populações separadas por algumas centenas de metros apenas.

2.4.1.1 Diversidade genética

Para a espécie *L. racemosa* foram encontrados 262 *loci* de AFLP, dos quais 96,2% (252 *loci*) eram polimórficos para a população ribeirinha e 88,2% (231 *loci*) para a população próxima ao apicum (Tabela 2.1). Os índices de diversidade de Shannon e de Nei mostraram maior diversidade genética nas plantas do apicum do que as ribeirinhas, porém com valores de desvio padrão bastante altos. Os valores de Shannon foram 0,36 para as plantas ribeirinhas e

0,42 para aquelas do apicum. Enquanto que os valores de Nei foram de 0,22 para as ribeirinhas e de 0,27 para as plantas do apicum (Tabela 2.1). Tanto os valores de Nei quanto de Shannon variaram de foram significativa entre as duas populações segundo o teste de *Kruskal-Wallis* ($P < 0,0001$).

Tabela 2.1 Dados obtidos com marcadores AFLP para duas populações da espécie *L. racemosa*: número e porcentagem de *loci* polimórficos e os índices de diversidade genética de Shannon e de Nei. Letras diferentes indicam valores significativamente diferentes.

População	<i>Loci</i> polimórficos	Índice de Shannon	Índice de Nei
Ribeirinha	252 (96,2%)	0,36 a	0,22 a
Apicum	231 (88,2%)	0,42 b	0,27 b

Para a espécie *A. schaueriana* foram encontrados 212 *loci* de AFLP, sendo que as plantas ribeirinhas tiveram mais polimorfismo (93,4% - 198 *loci*) do que as do apicum (46,2% - 98 *loci*) (Tabela 2.2). Seguindo um padrão inverso ao obtido com a espécie *L. racemosa*, os índices de diversidade genética de Shannon e de Nei tiveram valores maiores para indivíduos ribeirinhos, Shannon = 0,49 e Nei = 0,33, do que para as plantas do apicum, Shannon = 0,23 e Nei = 0,15. Os testes de *Kruskal-Wallis* mostraram que os valores de diversidade genética tanto de Nei quanto de Shannon de ambas as populações foram significativamente diferentes ($P < 0,0001$).

Tabela 2.2 Dados obtidos com marcadores AFLP para duas populações da espécie *A. schaueriana*: número e porcentagem de *loci* polimórficos e os índices de diversidade genética de Shannon e de Nei. Letras diferentes indicam valores significativamente diferentes.

População	<i>Loci</i> polimórficos	Índice de Shannon	Índice de Nei
Ribeirinha	198 (93,4%)	0,49 a	0,33 a
Apicum	98 (46,2%)	0,23 b	0,15 b

2.4.1.2 Estrutura populacional

Os valores do índice de diferenciação genética G_{ST} foram bem diferentes nas duas espécies, sendo igual a 0,032 para *L. racemosa* e 0,191 para *A. schaueriana*, indicando uma diferenciação de 3,2% entre as populações de *L. racemosa* e de 19,1% entre as populações de *A. schaueriana* (Tabela 2.3) Os dados de identidade e distância genética de Nei também mostraram que as populações ribeirinhas e do apicum de *L. racemosa* são mais semelhantes entre si do que as populações de *A. schaueriana*. Os índices de identidade entre as populações de *L. racemosa* foi de 0,996 e de distância foi de 0,004; enquanto que para a espécie *A. schaueriana*, suas populações tiveram identidade genética de 0,871 e distância de 0,138 (Tabela 2.4).

Tabela 2.3 Índices de diferenciação genética entre as populações ribeirinha e próxima do apicum calculados pelas análises de frequência alélica (G_{ST}), Bayesiana (θ_{ST}) e de multivariável (β_{ST}) para as duas espécies de mangue: *L. racemosa* e *A. schaueriana*.

Espécie	G_{ST}	θ_{ST}	β_{ST}
<i>L. racemosa</i>	0,032	0,004	0,054 ^{NS}
<i>A. schaueriana</i>	0,191	0,086	0,098*

* P = 0,0132

NS – Não significativo

Tabela 2.4 Índices de identidade e distância genéticas calculados segundo Nei (1978) entre as populações das duas espécies: *L. racemosa* e *A. schaueriana*.

Espécie	Identidade genética	Distância genética
<i>L. racemosa</i>	0,996	0,871
<i>A. schaueriana</i>	0,004	0,138

O índice de diferenciação genética, calculado pelo método Bayesiano, entre as populações de *A. schaueriana* foi maior do que entre as populações de *L. racemosa*: $\theta_{ST} = 0,086$ (SD = 0,012, limite menor de 0,063 e maior de 0,111 no intervalo de confiança de 95%), e $\theta_{ST} = 0,004$ (SD = 0,003, limite menor de 0,0002 e maior de 0,0113 no intervalo de confiança de 95%), respectivamente (Tabela 2.3).

A terceira análise estatística, de multivariável, mostrou uma diferenciação não significativa entre as populações de *L. racemosa* ($\beta_{ST} = 0,054$, não significativo) e bastante baixa para a espécie *A. schaueriana* ($\beta_{ST} = 0,098$, $P = 0,0132$) (Tabela 2.3).

2.4.2 Marcadores MSAP

As análises dos marcadores MSAP foram feitas com as duas populações: ribeirinha e próxima do apicum; da espécie *Laguncularia racemosa* de Guaratiba (RJ). Esta espécie apresentou baixo índice de diferenciação entre as populações com marcadores AFLP e, portanto, as alterações morfológicas que ocorrem neste manguezal não poderiam ser justificadas apenas pela variação genética. O uso dos marcadores MSAP permitiu o estudo de variações genéticas e epigenéticas, moduladas pela metilação de citosinas no DNA, entre as populações de *L. racemosa*, controle este muito importante na formação de novos fenótipos.

2.4.2.1 Níveis de polimorfismo

As quatro combinações de *primer* produziram 97 *loci* de MSAP nas 34 amostras. A população ribeirinha teve 41 *loci* metilados, ou seja, bandas presentes em *MspI*-MSAP e ausentes em *HpaII*-MSAP em pelo menos uma amostra, enquanto que a população do apicum teve apenas 25 *loci* metilados, correspondendo a 42,3% e 25,8% de todos os *loci* analisados (Figura 2.7). Os outros *loci* foram denominados como não-metilados, quando tinham presença de banda em ambos os perfis da mesma amostra ou o inverso na ausência das bandas. *Loci* foram considerados polimórficos quando uma ou mais amostras tinham um padrão de banda diferente das outras amostras da população. O nível de polimorfismo foi maior na população ribeirinha do que do apicum (Figura 2.8). Na população ribeirinha, 85,4% dos *loci* metilados e 5,4% dos *loci* não-metilados eram polimórficos. Entretanto, no apicum, indivíduos tiveram apenas 64% dos *loci* metilados e 1,4% dos *loci* não-metilados com algum polimorfismo.

Em conclusão, a população ribeirinha teve mais *loci* polimórficos do que a população do apicum com relação a ambos *loci* metilados e não-metilados, embora este último tenha sido em menor proporção. Isso indica uma maior variabilidade genética entre plantas ribeirinhas, mas principalmente da variabilidade epigenética. Além desta diferença com relação ao polimorfismo, houve também diferenças entre o número de *loci* metilados que foi maior nas ribeirinhas com 41 contra 25 nas plantas do apicum.

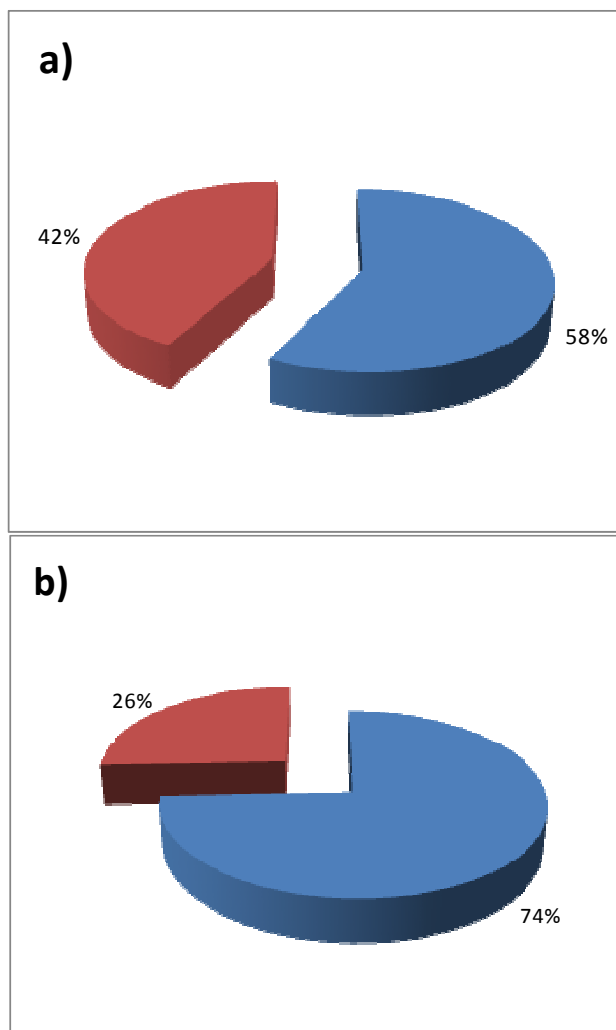


Figura 2.7 Gráficos ilustrando o nível de metilação do DNA dentre os 97 *loci* analisados para ambas as populações: (a) ribeirinha e (b) próxima do apicum. *Loci* metilados (vermelho) e não-metilados (azul). Valores das contribuições de cada um no total de *loci* analisados expressos em porcentagem.

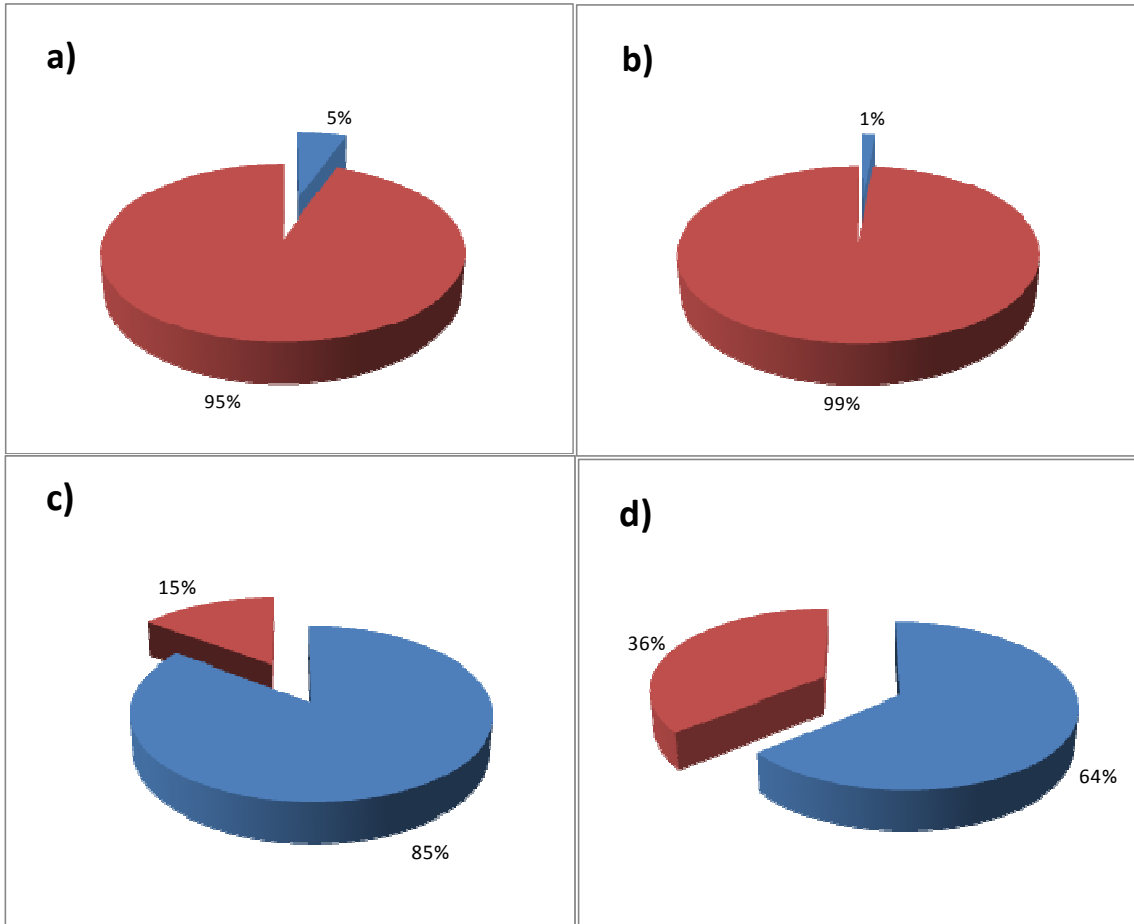


Figura 2.8 Gráficos mostrando o nível de polimorfismo por população para os *loci* metilados e não-metilados. Porcentagem de *loci* monomórficos (vermelho) e *loci* polimórficos (azul) são exibidos nos gráficos. (a) *Loci* não-metilados nas ribeirinhas; (b) *Loci* não-metilados no apicum; (c) *Loci* metilados nas ribeirinhas; e (d) *Loci* metilados no apicum.

2.4.2.2 Estrutura genética populacional

Três métodos foram utilizados para o cálculo do índice de diferenciação entre as populações ribeirinha e do apicum (Tabela 2.5). A fim de evitar as hipóteses do equilíbrio de Hardy-Weinberg como tamanho de população infinito e cruzamento aleatório, a análise de multivariável foi utilizada calculando o índice de diferenciação denominado β_{ST} . Os dados de *MspI*-MSAP e *HpaII*-MSAP, correspondentes às porções genética e epigenética do genoma de *L. racemosa*, respectivamente, resultou em $\beta_{ST} = 0,148$ e $0,212$ ($P < 0,001$), indicando uma maior diferenciação populacional no nível epigenético (Tabela 2.5).

Tabela 2.5 Índice de diversidade de Shannon calculado para cada população e índices de diferenciação das análises de multivariável (β_{ST}), de frequência alélica (G_{ST}) e Bayesiana (θ_{ST}) calculados para os componentes genético (*MspI*-MSAP) e epigenético (*HpaII*-MSAP) do genoma de *L. racemosa*.

Perfil	População	β_{ST}	Shannon	G_{ST}	θ_{ST}
<i>MspI</i>-MSAP	Ribeirinha		0,048		
	Apicum	0,148*	0,035	0,111	0,119
<i>HpaII</i>-MSAP	Ribeirinha		0,101*		
	Apicum	0,212*	0,043*	0,254	0,160

* Diferenças significativas

A diversidade genética dentro das populações foi avaliada pelo índice de diversidade de Shannon, um método baseado na frequência de alelos que pode ser diretamente estimado através do uso de marcadores dominantes. Após a correção dos valores de P pelo método Bonferroni, os índices de Shannon não foram significativamente diferentes sobre o nível genético (*MspI*-MSAP) entre as populações, mas sim no nível epigenético (*HpaII*-MSAP). Cabe ressaltar também que entre os dados genéticos e epigenéticos da mesma população houve diferença significativa nos índices de diversidade de Shannon. Os dados de *MspI*-MSAP resultaram em 0,048 e 0,035 para as populações ribeirinha e do apicum, respectivamente ($P = 29.1776$; não significativo) e dados de *HpaII*-MSAP resultaram em 0,101 e 0,043 para ribeirinhas e do apicum, respectivamente ($P < 0,025$; valor significativo para $\alpha = 5\%$) (Tabela 2.5). As diferenças entre os índices de Shannon de *MspI*-MSAP e *HpaII*-MSAP entre amostras da mesma população foram significativas ($P < 0,001$ e $P = 0.0025$, respectivamente para ribeirinhas e apicum). O índice de diferenciação G_{ST} , calculado com base na frequência alélica, mostrou que as populações foram mais diferenciadas no nível epigenético ($G_{ST} = 0,254$ para *HpaII*-MSAP) do que no nível genético ($G_{ST} = 0,111$ para *MspI*-MSAP) (Tabela 2.5).

O terceiro índice de diferenciação, θ_{ST} , foi calculado pela abordagem Bayesiana e mostrou diferenciação genética de 0,119 (*MspI*-MSAP; SD = 0,034, 95% intervalo de confiança: 0,055 limite inferior e 0,186 limite superior) e diferenciação epigenética de 0,160 (*HpaII*-MSAP; SD = 0,026, limite inferior de 0,110, limite superior de 0,210) (Tabela 2.5).

Portanto, as semelhanças entre os diferentes índices de diferenciação calculados neste trabalho mostraram maior diferenciação entre as populações ribeirinha e do apicum de *L. racemosa* no nível epigenético (Figura 2.9).

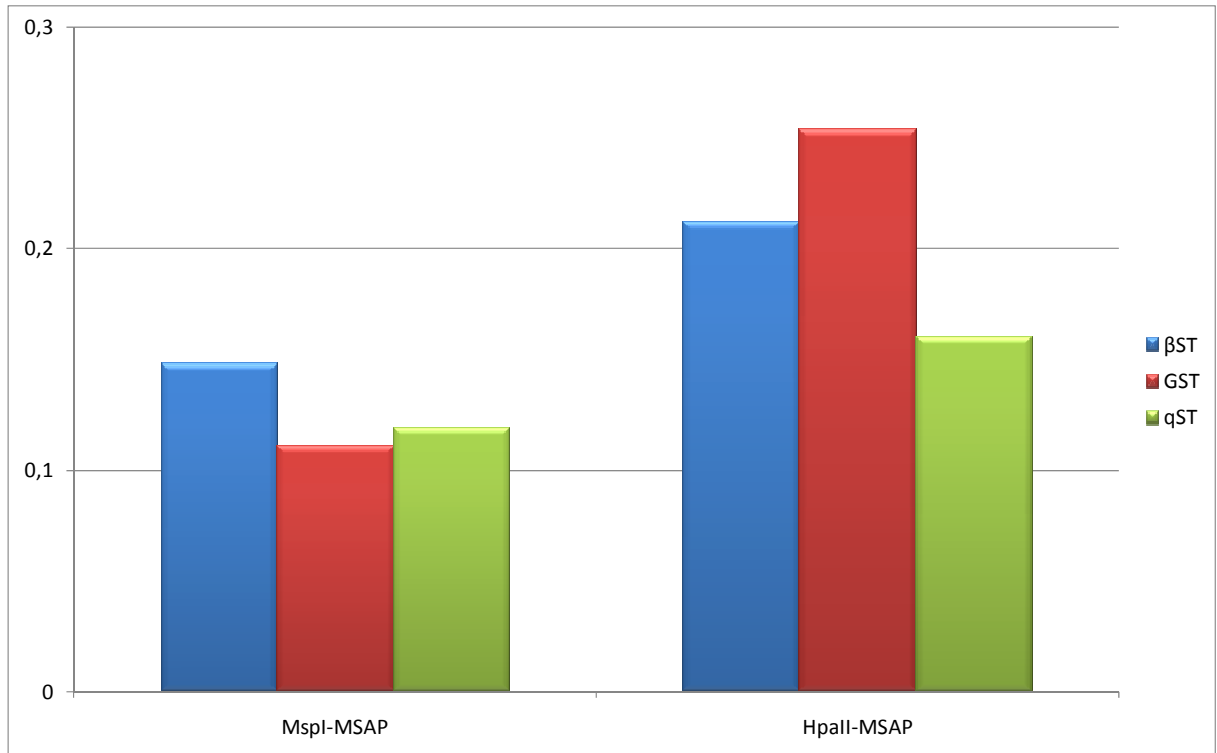


Figura 2.9 Gráfico com todos os índices de diferenciação calculados pelos três métodos estatísticos: análise de multivariância (β_{ST}), de frequência alélica (G_{ST}) e Bayesiana (θ_{ST}), para ambos os componentes genético (*MspI*-MSAP) e epigenético (*HpaII*-MSAP).

2.4.2.3 Análise de componente principal (PCA)

Os dados do MSAP foram organizados espacialmente pela análise de PCA com base na matriz de covariância. As diversidades genéticas e epigenéticas estão resumidas sobre os dois componentes principais, representando 40,2% e 35,7% da inércia total, respectivamente (Figuras 2.10). Figura 2.10 (a) representa a estrutura genética do genoma de *L. racemosa*, com base em dados de *MspI*-MSAP, onde a proximidade genética de indivíduos de cada uma das populações ribeirinha e do apicum (representados como 1 e 2, respectivamente) pode ser observada, apesar de ainda ter alguma mistura entre as populações. Por outro lado, a estrutura epigenética obtida com dados de *HpaII*-MSAP (Figura 2.10b) diferencia mais expressivamente indivíduos de cada população, formando dois grupos separados.

Os dois primeiros eixos da análise de Co-Inércia explicam 74,4% do total da variação genética. Na figura 2.10 (c), os círculos representam a posição de cada perfil genético no gráfico, baseado em dados de *MspI*-MSAP, de ambas as populações (representadas como 1 e 2 para ribeirinha e do apicum, respectivamente), enquanto que a cabeça das setas indicam a posição do perfil epigenético (*HpaII*-MSAP). O comprimento das setas representa a distância entre os perfis genéticos e epigenéticos. Curiosamente, existem três indivíduos ribeirinhos e três do apicum que têm perfis genéticos muito semelhantes entre eles, compartilhando o mesmo círculo na figura 2.10 (c), porém as suas setas apontam para direções praticamente opostas, indicando perfis epigenéticos distintos referentes a cada população.

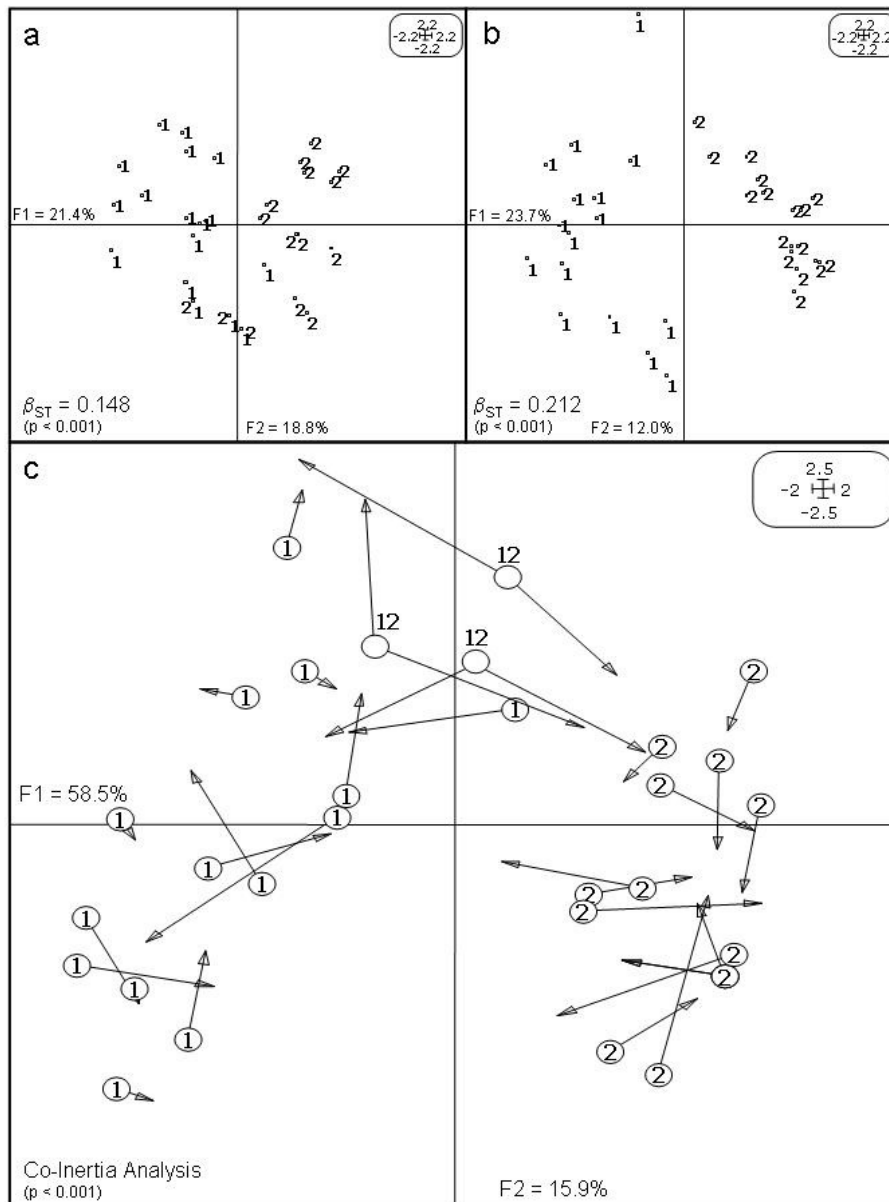


Figura 2.10 Análise de componente principal (PCA) entre indivíduos e entre grupos: (a) PCA baseado no perfil genético (*MspI*-MSAP) em ambas as populações (1) ribeirinhas e (2) apicum; (b) PCA baseado no perfil epigenético (*HpaII*-MSAP) em ambas as populações (1) ribeirinhas e (2) apicum; (c) Análise de Co-Inércia onde círculos correspondem ao perfil genético e cabeça das setas ao perfil epigenético. F1 e F2 mostram os valores de contribuição dos componentes na variância total da análise. β_{ST} foi calculado com base na análise entre-grupo para ambos os perfis genético e epigenético e foram testados significativamente.

2.5 Discussão

Utilizando os marcadores AFLP e MSAP, analisamos os níveis de diferenciação genética em duas espécies de mangue, *Laguncularia racemosa* e *Avicennia schaueriana*, e genética *versus* epigenética apenas para *L. racemosa*. Foram analisadas amostras das populações naturais de Guaratiba (Rio de Janeiro-RJ), que possuem diferenças morfológicas bastante evidentes entre elas. Uma população, chamada ribeirinha, é caracterizada pela sua estrutura arbórea, enquanto a outra população, localizada próxima ao apicum, tem plantas arbustivas (mais detalhes na introdução). Devido à diferença nos sítios de clivagem das enzimas de corte freqüente usadas pelos dois marcadores AFLP e MSAP, não serão feitas análises comparativas entre seus resultados de diversidade e de diferenciação genética entre as populações de *L. racemosa*.

2.5.1 Marcadores AFLP

Os marcadores AFLP mostraram alto grau de polimorfismo nas populações ribeirinhas das duas espécies (96,2% e 93,4% para *L. racemosa* e *A. schaueriana*), porém níveis mais baixos nas populações do apicum (88,2% e 46,2%, para cada espécie respectivamente) (Tabelas 2.1 e 2.2). Os índices de diversidade genética de Shannon e de Nei calculados para cada amostra tiveram valores significativamente diferentes entre as populações da mesma espécie ($P < 0,0001$). Curiosamente, indivíduos da população ribeirinha de *L. racemosa*

mostraram uma diversidade genética mais baixa do que daqueles da população do apicum, que possuem um desenvolvimento anormal de seus indivíduos (Tabela 2.1). Por outro lado, a população ribeirinha de *A. schaueriana* possui valores de diversidade genética maiores do que a população do apicum, que pode estar sofrendo ação de pressões naturais seletivas que estariam diminuindo a sua diversidade genética provavelmente como resultado de sua proximidade à planície hipersalina do apicum (Tabela 2.2).

Os índices de diferenciação genética obtidos pelos diferentes testes estatísticos foram moderados a baixos entre as populações de *A. schaueriana* e bastante baixos entre as populações de *L. racemosa* (Tabela 2.3). Estes resultados sugerem que as duas populações de *L. racemosa* sejam muito semelhantes geneticamente, enquanto que as populações de *A. schaueriana* possuem algum grau de diferenciação genética entre elas. Estes resultados foram corroborados pelos índices de identidade e distância genética de Nei, que mostraram uma alta semelhança entre as populações de *L. racemosa*, mas entre as populações de *A. schaueriana* os valores de identidade e distância genética foram menos expressivos, mesmo assim obtivemos alta semelhança entre as populações desta espécie como era esperado devido à grande proximidade delas (Tabela 2.4).

Os resultados obtidos com marcadores AFLP podem ser explicados pela ocorrência de cada espécie. *L. racemosa*, sendo uma espécie mais comumente encontrada nos manguezais do Oceano Atlântico, e possui uma melhor distribuição genética entre suas populações e por isso, as populações de Guaratiba que são muito próximas, poderiam ser consideradas como uma única população. Por outro lado, *A. schaueriana* é uma espécie de ocorrência mais restrita, o que limitaria em termos geográficos o fluxo gênico e a manutenção da variabilidade

genética de suas populações, e como conseqüência uma pequena estruturação genética pôde ser evidenciada entre as populações de Guaratiba que são bastante próximas. Geralmente, indivíduos de *A. schaueriana* são mais afetados pelo ambiente, tendo maior sensibilidade às limitações dos manguezais, além disso, outro limitante pode ser o tipo de propágulo²⁰ gerado por esta espécie, provavelmente disperso de forma menos eficiente quando comparado ao da espécie *L. racemosa*. A localização dos indivíduos com relação à proximidade do rio é uma característica importante na dispersão dos propágulos. As plantas ribeirinhas de *A. schaueriana* podem estar sendo favorecidas na troca com outras populações resultando assim num maior índice de diversidade genética em relação à população do apicum.

Em conclusão, os marcadores AFLP podem explicar parte da variação morfológica encontrada entre as populações ribeirinha e do apicum da espécie *A. schaueriana*, mas o mesmo não acontece para as populações de *L. racemosa*. Esta mostrou uma alta semelhança genética e índices de variabilidade genética maiores na população próxima ao apicum, o que não era esperado devido ao grande estresse ao qual as plantas neste ambiente são submetidas constantemente.

²⁰ Nome dado às sementes que se desprendem da planta-mãe já germinadas, que é resultado da viviparidade nas espécies de mangue.

2.5.2 Marcadores MSAP

Os marcadores MSAP foram utilizados a fim de identificar a diversidade e estrutura genética *versus* epigenética entre as populações de *L. racemosa*, não diferenciadas pelo AFLP. Os fenômenos epigenéticos podem aumentar a plasticidade fenotípica das plantas em situações de flutuações ambientais ou de estresse (Rapp and Wendel, 2005; Salmon *et al.*, 2005), o que poderia estar acontecendo com esta espécie nas populações de Guaratiba, que são submetidas a diferentes condições ambientais (Figura 2.11). Marcadores MSAP permitem a análise do nível de metilação do DNA, que podem atuar diretamente sobre a expressão gênica, gerando variações fenotípicas induzidas por perturbações ambientais e genéticas e que podem ser transmitidas de uma geração para a outra (Pavet *et al.*, 2006; Gehring and Henikoff 2007; Lukens and Zhan 2007).

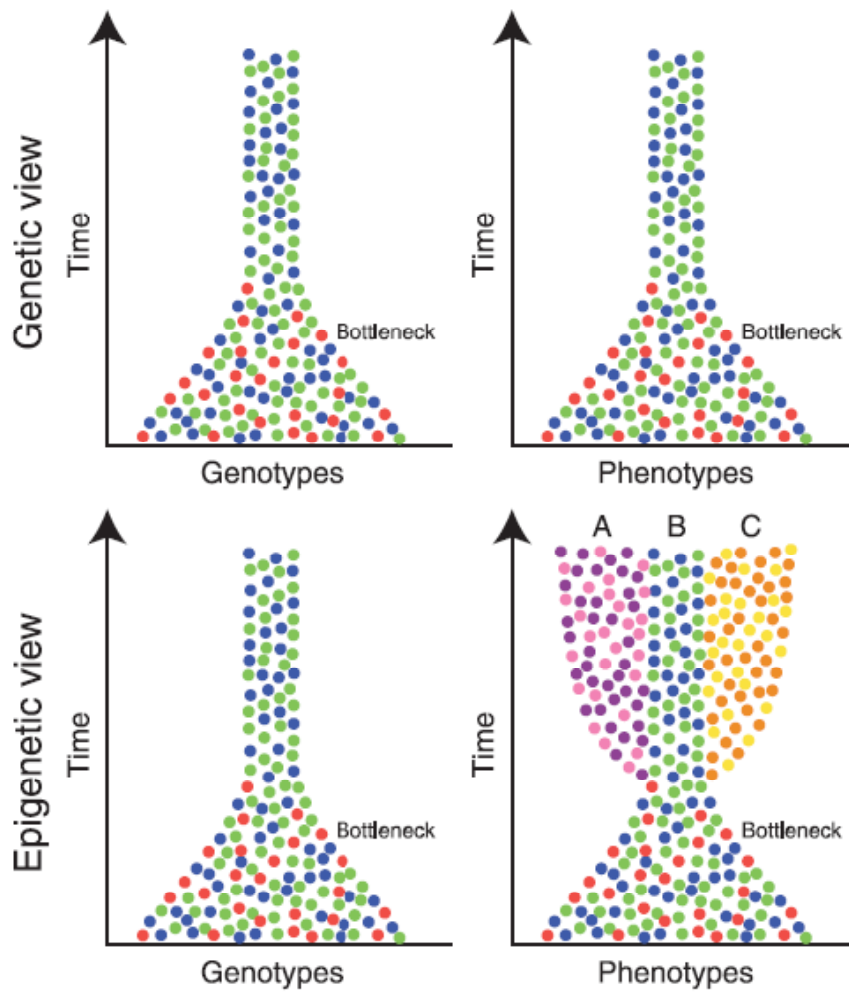


Figura 2.11 Visão genética e epigenética dos efeitos da deriva, mais especificamente, de gargalos. Nos gráficos do alto, a visão genética de perda de alelos pela deriva e/ou seleção natural, corresponde diretamente à perda de fenótipos também (círculos vermelhos se perderam, aumentando a frequência dos círculos azuis e verdes). Mas a visão epigenética, nos gráficos de baixo, mostra que mesmo após a perda de alelos na população (círculos vermelhos), o processo seletivo pode ser melhorado com o surgimento de novos fenótipos (A e C) controlados por alterações epigenéticas, além dos fenótipos correspondentes à visão genética do processo (B).

A proporção de *loci* metilados encontrados na população ribeirinha de *L. racemosa* foi 42% (Figura 2.7a), equivalente àquela encontrada para ecótipos de *Brassica oleracea*, 30-41% (Salmon *et al.*, 2008), algodão *Gossypium hirsutum*, 32% (Keyte *et al.*, 2006) e *Arabidopsis thaliana*, 35-43% (Cervera *et al.*, 2002). Por outro lado, a população do apicum teve menor proporção de *loci* metilados, 26% do total analisado (Figura 2.7b), indicando uma possível hipometilação no genoma destas plantas. O processo de demetilação do DNA ativo ocorre em plantas de forma intrinsecamente associada com o estresse biótico e abiótico, afetando funções básicas do genoma (Pavet *et al.*, 2006). E, como resultado deste processo, a expressão de muitos genes de desenvolvimento em plantas, como *flowering locus wa* (FWA), *superman* (SUP), *agamous* (AG) e *Linaria vulgaris* homologue of *cycloidea* (Lcyc), pode ser afetada (Habu *et al.*, 2001). Em *Nicotiana tabaccum*, a expressão de uma enzima que tem atividade de demetilação foi aumentada em folhas tratadas com alumínio, frio e sal, como resultado de um controle epigenético baseado na metilação de citosinas (Choi *et al.*, 2007).

Três métodos estatísticos foram utilizados para estudar a diversidade e diferenciação genética *versus* epigenética entre as duas populações de *L. racemosa* de Guaratiba. A análise de multivariável calculou o índice β_{ST} com base na presença ou ausência de banda que é mais indicado para marcadores dominantes como o MSAP. Este índice de diferenciação foi significativamente diferente quando comparados os perfis genético (*MspI*-MSAP) e epigenético (*HpaII*-MSAP) (Tabela 2.5). Este resultado é corroborado pelos índices de G_{ST} , calculado com base nas freqüências dos alelos, e θ_{ST} , uma abordagem Bayesiana, mostrando uma forte estruturação populacional sobre o nível epigenético (Tabela 2.5). Isto pode ser

evidência de que as populações que ocorrem em ambientes tão distintos enfrentem diferentes processos evolutivos, sobre os quais a metilação do DNA é provavelmente afetada.

Os índices de Shannon, calculados com base na frequência dos alelos de ambos os marcadores *MspI*-MSAP e *HpaII*-MSAP, foi maior na população ribeirinha do que do apicum, mas estes valores foram significativos apenas no nível epigenético (Tabela 2.5). Os índices de Shannon dos perfis genético e epigenético de cada população também foram significativamente diferentes, indicando mais uma vez que estes dois componentes do genoma sofrem processos evolutivos independentes. Estas diferenças sugerem ainda a existência de microambientes na região ribeirinha, mais afetada pelo regime de marés; enquanto indivíduos da população do apicum estão em um ambiente mais homogêneo apesar dos altos níveis de salinidade do solo.

No PCA baseado em análise de multivariável, há claros indícios de uma maior diferenciação entre as populações ribeirinha e do apicum no nível epigenético, e, mais importante ainda, indivíduos do apicum estão mais estreitamente relacionados sob o ponto de vista epigenético do que genético (Figura 2.10a e 2.10b). Este padrão também pôde ser observado na análise de Co-Inércia, um método mais robusto que associou diretamente os dados genéticos e epigenéticos sobre a mesma escala espacial. Alguns indivíduos das populações ribeirinhas e do apicum tinham o perfil genético semelhante, representado por círculos, mas perfil epigenético distinto, representado pelas cabeças das setas (Figura 2.10c). O comprimento da seta é compatível com a distância entre os perfis genéticos e epigenéticos e as direções da flecha são bons indicadores da relação entre a estrutura genética e epigenética nas populações. Ao estudar os comprimentos e direções das setas de indivíduos da população

próxima do apicum (figura 2.10c), parece que elas estão convergindo para um ponto comum. Isso sugere que existe um perfil melhor adaptado às condições ambientais de estresse deste ambiente, resultado de modificações genômicas importantes, principalmente no que diz respeito aos fenômenos epigenéticos. Estas mudanças podem resultar numa variação fenotípica não correspondente à variação genética, mas que ainda assim é plausível a sofrer seleção (Lukens and Zhan 2007). Portanto, os resultados deste marcador sugerem que os indivíduos do apicum são fenotipicamente diferentes dos indivíduos ribeirinhos como consequência de alterações epigenéticas de longo prazo passadas ao longo das gerações, que podem ter sofrido o efeito da seleção natural diretamente.

Capítulo 3

Uma revisão sobre microssatélites nucleares e o isolamento destes marcadores em duas espécies de mangue

Uso de microssatélites em estudos de conservação

Floresta e Ambiente, 2006, volume 13, número 2, p. 25 – 36 (Anexo IV)

3.1 Introdução

3.1.1 Marcadores Moleculares

As populações naturais possuem em geral altos níveis de variabilidade genética entre indivíduos da mesma população que é introduzida continuamente por mutação, migração ou fluxo gênico (Morand *et al.*, 2002). Por outro lado, as espécies que têm preferência pela autofecundação, que limita a migração e o fluxo gênico, podem ter a variabilidade genética em suas populações reduzida devido ao cruzamento entre co-sangüíneos. Vários estudos já comprovaram que o fluxo gênico em florestas tropicais pode alcançar longas distâncias, sendo importante na manutenção da variabilidade genética (White *et al.*, 2002; Chase *et al.*, 1996). Esta, por sua vez, é fundamental para a manutenção e sobrevivência da espécie e de suas populações às ameaças causadas principalmente pela destruição do seu habitat e por coletas predatórias (Heuertz *et al.*, 2001).

Dentre os diferentes níveis hierárquicos de biodiversidade, uma importância especial tem se focado na diversidade genética, que é a base desta hierarquia. Ela geralmente é observada ao nível de espécie, dentro de uma mesma população ou entre diferentes populações. A diversidade genética tem papel importante na conservação de espécies ameaçadas de extinção e na manutenção de suas populações ao longo do tempo, pois espécies com baixa variação genética têm geralmente uma redução no *fitness*, ou seja, na habilidade de sobreviver a mudanças ambientais durante o seu processo evolutivo (Frankham, 1995).

Tradicionalmente, a diversidade dentro de uma espécie era medida apenas pelas diferenças morfológicas, mas atualmente os métodos de detecção, baseados em dados moleculares, vêm sendo utilizados e considerados como vantajosos sobre os métodos antigos devido ao maior número de caracteres estudados (Ferreira and Grattapaglia, 1998; Karp *et al.*, 1996).

Dentre os seis marcadores moleculares tradicionais, três não utilizam o PCR, isoenzimas, RFLP (*Restriction Fragment Length Polymorphism*) e minissatélites, e por isso possuem limitações quanto ao número de indivíduos analisados por estudo, principalmente por causa da grande quantidade de material genético necessário e do excessivo trabalho laboratorial com resultados mais demorados e custosos (Ferreira and Grattapaglia, 1998).

Por outro lado, outros três marcadores são baseados em PCR, RAPD (*Random Amplified Polymorphic DNA*), AFLP (*Amplified Fragment Length Polymorphism*) e microssatélites ou SSR (*Simple Sequence Repeat*), e permitem a obtenção de resultados mais rapidamente, além de possibilitar estudos populacionais com grande número de indivíduos e também outros tipos de estudos como de genética humana, forense, de biologia evolutiva e de biologia do desenvolvimento, devido à sensibilidade, rapidez e versatilidade do PCR (Ferreira and Grattapaglia, 1998; White *et al.*, 1989; Williams *et al.*, 1990).

Estes marcadores podem ser divididos em dois tipos: os dominantes, como RAPD e AFLP, que não permitem a análise de heterozigotos na população, apenas a presença ou ausência de marcadores ao longo do genoma; e os co-dominantes, como microssatélites, que permitem a análise de todos os alelos de cada *loci* de um indivíduo, e assim obtendo dados de frequência de alelos na população (Ouborg *et al.*, 1999).

Marcadores dominantes como RAPD e AFLP têm a vantagem de não necessitarem de conhecimento prévio de seqüência de DNA e por isso podem ser utilizados com as mais diferentes espécies. Além disso, grandes quantidades de marcadores são obtidas com baixo custo de desenvolvimento da técnica. Eles já foram usados em estudos com diferentes espécies de plantas como *Eucalyptus* (Kirst *et al.*, 2005), *Caesalpinia echinata* (Cardoso *et al.*, 1998; Cardoso *et al.*, 2005), *Euterpe edulis* (Cardoso *et al.*, 2000), *Veronica alpina* (Albach *et al.*, 2006), *Viola sp.* (Eckstein *et al.*, 2006), *Cedrus atlantica* (Renau-Morata *et al.*, 2005), *Croton alabamensis* (Van EE *et al.*, 2006) com o objetivo de avaliar a variabilidade genética e a estruturação populacional.

Os microssatélites ou SSR são repetições em tandem²¹ de 1 a 6 nucleotídeos, encontrados em todos os procariotos e eucariotos estudados até o momento (Zane *et al.*, 2002). Os microssatélites foram identificados no genoma de eucariotos no final da década de 80 (Tautz, 1989), e logo os pesquisadores encontraram altos níveis de polimorfismo neste marcador, comprovando que se tratava de uma nova ferramenta molecular bastante eficiente, com pelo menos o dobro de informação obtida, e duas vezes mais rapidamente do que outros marcadores (Weber, 1990). A co-dominância destes marcadores é uma importante característica que possibilita as análises do sistema reprodutivo de plantas e animais e suas conseqüências como porcentagem de heterozigotos na população (Morand *et al.*, 2002; Collevatti *et al.*, 2001; Heuertz *et al.*, 2001; White *et al.*, 1999). Atualmente, sabemos que existem microssatélites tanto no genoma nuclear das plantas (Condit and Hubbell, 1991; Wang *et al.*, 1994) como no genoma de cloroplastos (Powel *et al.*, 1995).

²¹ Seqüências repetidas lado a lado no genoma.

O fato de ter microssatélites em cloroplasto possibilitou a utilização desta organela em estudos populacionais, que eram pouco produtivos devido à baixa taxa de mutação em seu genoma (Provan *et al.*, 2001). Os microssatélites encontrados nos cloroplastos, também chamados de marcadores citoplasmáticos, são ótimas ferramentas para estudos de fluxo gênico e padrões de cruzamento, pois estas organelas são herdadas de forma uniparental: maternalmente em angiospermas e paternalmente em gimnospermas (Provan *et al.*, 2001; Provan, 2000; Vendramin *et al.*, 2000). Por isso, diversos estudos de estrutura genética foram feitos usando estes marcadores citoplasmáticos tanto em gimnospermas: *Picea abies* K (Vendramin *et al.*, 2000), *Abies sp* (Clark *et al.*, 2000; Vendramin *et al.*, 1999); quanto em angiospermas: *Silene paradoxa* L (Mengoni *et al.*, 2001), *Fraxinus sp* (Morand-Prieur *et al.*, 2002) e *Caesalpinia echinata* (Lira *et al.*, 2003).

Os marcadores microssatélites tornaram-se extremamente difundidos, apesar do alto custo inicial para o desenvolvimento da técnica, sendo atualmente utilizados em um grande número de estudos para responder a perguntas específicas nas áreas da biologia evolutiva e da biologia da conservação (Balloux and Lugon-Moulin, 2002). Isto se dá devido a três características presentes nos microssatélites: (1) a neutralidade, pois representam o genoma como um todo; (2) a abundância com ampla distribuição pelo genoma; e (3) o alto polimorfismo encontrado mesmo em populações que passaram por eventos de gargalo ou em populações que possuem baixo polimorfismo detectado por outros marcadores (Maudet *et al.*, 2002).

A maior limitação ao uso de marcadores microssatélites é o grande trabalho necessário para o desenvolvimento e isolamento dos *loci* que contêm estes marcadores, principalmente

para plantas, cujo número de repetições de dinucleotídeos é 10 vezes menor do que nos primatas (Zane *et al.*, 2002). Para contornar esta situação foram desenvolvidos diferentes métodos de enriquecimento de bibliotecas genômicas (White and Powell, 1997; Edwards *et al.*, 1996; Ostrander *et al.*, 1992). Zane *et al.* (2002) indicam diferentes métodos de isolamento de microssatélites, porém todos bastante custosos, o que gera um empecilho para se trabalhar com este tipo de marcador molecular. O método tradicional para o desenvolvimento de marcadores microssatélites envolve a criação de bibliotecas genômicas de inserto pequeno, a posterior hibridização com oligonucleotídeos repetidos em tandem e o seqüenciamento de clones candidatos, tornando assim o processo lento, trabalhoso e caro (Thiel *et al.*, 2003). Após seu isolamento, ainda é preciso caracterizar os *primers*, ou seja, definir quais *primers* são eficientes na amplificação dos marcadores e quais marcadores são polimórficos para a espécie. Esta etapa pode tornar-se tão problemática quanto o isolamento, em termos de dinheiro e tempo gasto (Squirrell *et al.*, 2003).

Com o estabelecimento de bancos de dados de seqüências expressas do genoma (EST – *Expressed Sequence Tags*), uma nova modalidade de microssatélites foi descoberta, os microssatélites gênicos ou SSR-EST. EST são porções seqüenciadas de cDNA (DNA complementar), que são cópias do mRNA (RNA mensageiro). Por isso, eles representam parte da porção transcrita do genoma em certas condições e são bastante conservados entre espécies próximas (Poncet *et al.*, 2006). O Instituto TIGR (*The Institute for Genomic Research* – www.tigr.org) possui bancos de ESTs de 232 espécies de plantas, dentre elas: 10 coníferas, 150 dicotiledôneas e 41 monocotiledôneas. A grande maioria destas espécies é de importância

econômica e geralmente são espécies cultivadas o que limita o uso destes bancos de dados para estudos populacionais em espécies selvagens.

O desenvolvimento dos SRS-ESTs é relativamente fácil e de baixo custo porque são utilizadas seqüências disponibilizadas na internet e programas de computador distribuídos gratuitamente para identificar os marcadores (Thiel *et al.*, 2003; Varshney *et al.*, 2005). As principais vantagens desta técnica são: 1) menor taxa de mutação em regiões codificantes, o que diminui a ocorrência de alelos nulos²² e ainda permite uma maior transferabilidade destes marcadores entre espécies; 2) menor homoplasia²³, resultado também da taxa de mutação mais baixa; 3) alguns trabalhos obtiveram fragmentos mais limpos com bandas melhor definidas do que aquelas obtidas com os microsatélites nucleares aleatórios (Woodhead *et al.* 2005).

Alguns programas (Tabela 3.1) usados no isolamento de microsatélites a partir de seqüências de EST foram criados na linguagem Pearl, como o MISA (MIcroSATellite) que é um módulo de busca com várias características úteis para o controle de qualidade das seqüências de EST e para o desenvolvimento de *primers* específicos para microsatélites (Varshney *et al.*, 2005). Em plantas, os microsatélites isolados a partir de seqüências de EST estão sendo bastante utilizados em espécies cultivadas como trigo, cevada, café, laranja, amêndoa e girassol (Gao *et al.*, 2004; Thiel *et al.*, 2003; Poncet *et al.*, 2006; Chen *et al.*, 2006; Xie *et al.*, 2006; Pashley *et al.*, 2006).

²² Alelos que não amplificam durante o PCR por problema de anelamento do *primer*.

²³ Similaridade genética entre duas espécies que não é resultante da herança de um ancestral comum.

Tabela 3.1 Ferramentas para busca de microssatélites em bancos de dados de EST. Fonte: Varshney *et al.* (2005).

Programa	Referências
MicroSAteLLite (MISA)	http://pgrc.ipk-gatersleben.de/misa/ ; Thiel <i>et al.</i> (2003)
SSRFinder	Gao <i>et al.</i> (2003)
BuildSSR	Rungis <i>et al.</i> (2004)
SSR Identification Tool (SSRIT)	Kantety <i>et al.</i> (2002)
Tandem Repeat Finder (TRF)	Benson (1999)
Tandem Repeat Occurrence Locator (TROLL)	Castelo <i>et al.</i> (2002)
CUGIssr	http://www.genome.clemson.edu/projects/ssr/
Sputnik	http://abajian.net/sputnik/index.html
Modified Sputnik	Morgante <i>et al.</i> (2002)
Modified Sputnik II	http://wheat.pw.usda.gov/ITMI/EST-SSR/LaRota/
SSRSEARCH	ftp://ftp.gramene.org/pub/gramene/software/scripts/ssr.pl

Uma das grandes vantagens destes marcadores provenientes de regiões expressas do genoma é a transferabilidade entre espécies relacionadas em contraste aos microssatélites de regiões anônimas, refletindo a natureza conservadora das regiões codificantes quando comparada às regiões não-codificantes (Varshney *et al.*, 2005b). A posição dos microssatélites gênicos com relação à região codificante resulta em diferentes níveis de polimorfismo, onde as regiões 3'UTR (*Untranslated Regions*) são mais polimórficas entre cultivares, as regiões 5'UTR são mais polimórficas entre cultivares e espécies, e a sequência codificante do gene é mais polimórfica entre espécies e gêneros (Varshney *et al.*, 2005).

3.1.2 Aplicações ecológicas dos microssatélites

Segundo Liu and Cordes (2004), os marcadores moleculares AFLP e microssatélites são os principais marcadores da atualidade, pois estão em crescimento exponencial no número de publicações usando estas ferramentas em diferentes tipos de estudo, tanto de estrutura populacional quanto de genômica funcional. Os marcadores neutros são especialmente usados em estudos ecológicos com o objetivo de responder a perguntas importantes sobre a biologia da espécie e suas populações (Tabela 3.2).

Devido à alta variabilidade dos microssatélites e à possibilidade de se obter marcadores mendelianos multi-alélicos, eles são especialmente úteis para inferir eventos demográficos recentes, inclusive detecção de impactos induzidos pelas comunidades humanas nas populações (Pearse and Crandall, 2004). Os microssatélites emergiram como uma das opções mais populares para estudos genéticos com questões ecológicas, em parte porque eles possuem potencial para determinar estimativas de migração, têm poder para distinguir relativamente altas taxas de migração para panmixia, e podem estimar a relação de parentesco entre indivíduos (Selkoe and Toonen, 2006). Como resultado da localização destes marcadores em regiões gênicas, estudos de diversidade funcional em plantas fenotipicamente bem caracterizadas poderão ser realizados com a ajuda dos SSR-EST presentes em genes de função conhecida, correlacionando assim a frequência alélica com variação fenotípica, ou ainda, com a variação biológica (Eujayl *et al.*, 2002).

Tabela 3.2 Algumas questões ecológicas que podem ser respondidas com marcadores genéticos neutros, separados por tipo de dados necessário. Fonte: Selkoe and Toonen (2006).

Necessita de dados de frequência alélica multi-*locus**

De qual população estes indivíduos se originaram?

Quantas populações existem?

Necessita de dados de microssatélites ou alto polimorfismo de sequência**

Esta população se expandiu ou se contraiu nos últimos anos?

As populações são diferentes em tamanho no passado e no presente?

Necessita de identificação genotípica multi-*locus****

Qual a relação genética entre os indivíduos?

Quais indivíduos migraram?

Quais indivíduos são clones?

Pode ser feito com diferentes marcadores moleculares**

Qual a distância de dispersão média dos descendentes (ou gametas)?

Quais são as relações fonte-sumidoro entre as populações?

Como condições do meio-ambiente causam impacto na estrutura populacional e migração?

Qual a dinâmica de extinção e colonização da metapopulação?

A estrutura populacional ou sua conectividade mudou nos últimos anos?

*Estas análises podem necessitar de >10 microssatélites – o número está inversamente correlacionado com o grau de diferenciação genética entre as populações. Espécies com baixa taxa de migração e/ou populações pequenas necessitam de menos *loci*.

**Usando >1 *locus* diminuirá substancialmente erro amostral inter*locus*.

***Geralmente necessita de microssatélites, mas também é possível com AFLP e RAPD

Dentre as principais aplicações dos microssatélites estão mapeamento genético, identificação individual por DNA e estudos de parentesco, filogenia, genética de populações e genética da conservação, epidemiologia e patologia molecular, mapeamento de QTLs (*Quantitative Trait Loci*), e seleção assistida por marcador (Chistiakov *et al.*, 2006). Mas estes marcadores vêm sendo muito utilizados para responder várias perguntas relacionadas à genética de populações, como análises de fluxo gênico, paternidade e estruturação

populacional, que resultam em dados sobre a distribuição da variabilidade genética entre e dentro de populações naturais, que são essenciais para a adoção de medidas de conservação tanto *ex situ* quanto *in situ* (Oliveira *et al.*, 2006). E estes resultados obtidos através dos marcadores moleculares, por outro lado, permitem um melhor entendimento de outros processos populacionais como dispersão e migração, história demográfica de populações como gargalo e efeito fundador, detecção de híbridos entre espécies, fluxo gênico, deriva gênica, filogeografia entre outros (DeYoung and Honeycutt, 2005).

3.1.3 Publicações atuais sobre conservação

O número de publicações utilizando ferramentas moleculares em estudos de conservação é cada vez maior, de 373 artigos publicados em 1990 houve um salto para 1.466 artigos em 2007 (Figura 3.1). A evolução do número de artigos publicados a cada dois anos, nos últimos 17 anos, mostra alguns aumentos mais expressivos em 1992 com 75% a mais de publicações do que 1990; 1994 com 27,5% a mais que 1992; e 2004 com 21% a mais que 2002. Isso mostra que a descoberta de novas técnicas pode causar um aumento não-linear no número de publicações anuais. A constante evolução das técnicas moleculares e de investigação do DNA, aumentando a eficiência e diminuindo os custos, impulsionou as publicações neste século a partir de 2001, tendo uma média de 1258 artigos desde 2002 até 2007, que antes era de 781 artigos de 1990 até 2000.

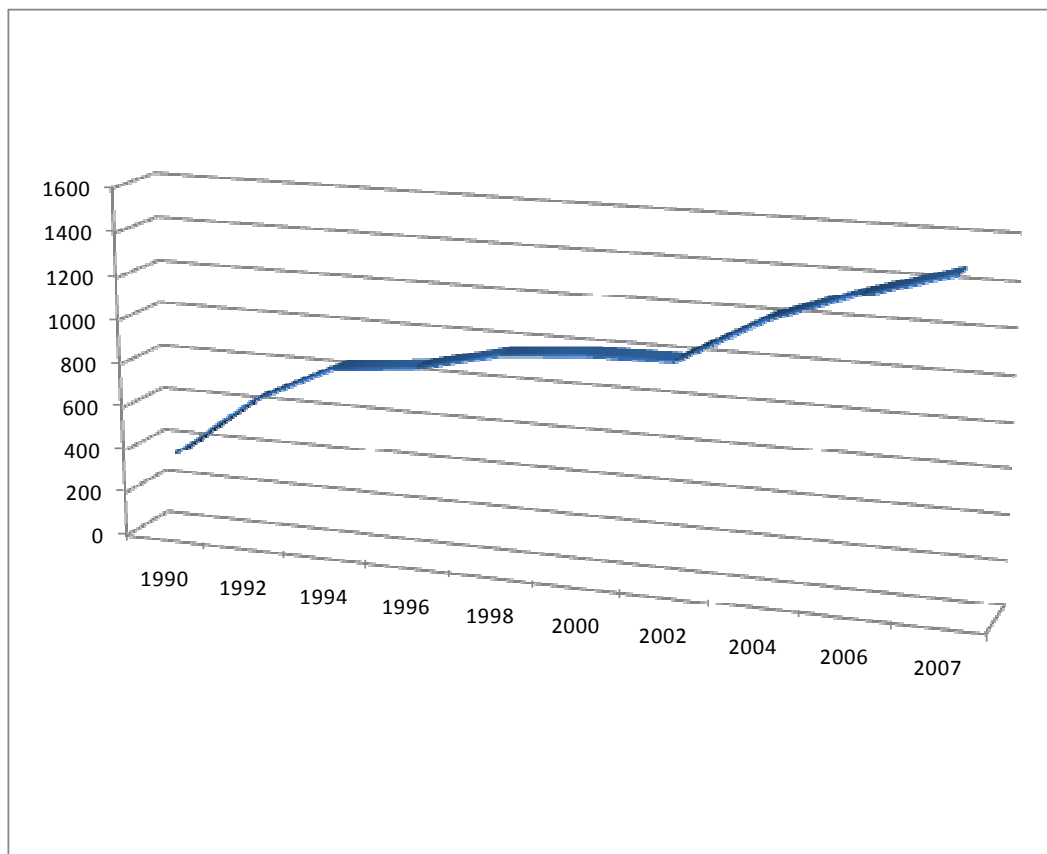


Figura 3.1 Gráfico do número de publicações por ano, a cada dois anos de 1990 a 2007, pesquisados usando as palavras-chave conservação (*conservation*) e DNA na ferramenta de busca do ISI Web of Science.

De todas estas publicações com DNA, as mais significativas com relação à conservação utilizam os marcadores moleculares microssatélites como ferramenta. Alguns marcadores tradicionais como RAPD e AFLP, baseados em PCR, são também bastante representativos nas publicações, enquanto que os outros marcadores que não usam PCR, como RFLP, isoenzimas e minissatélites, têm menor contribuição para este tipo de estudo (Figura 3.2).

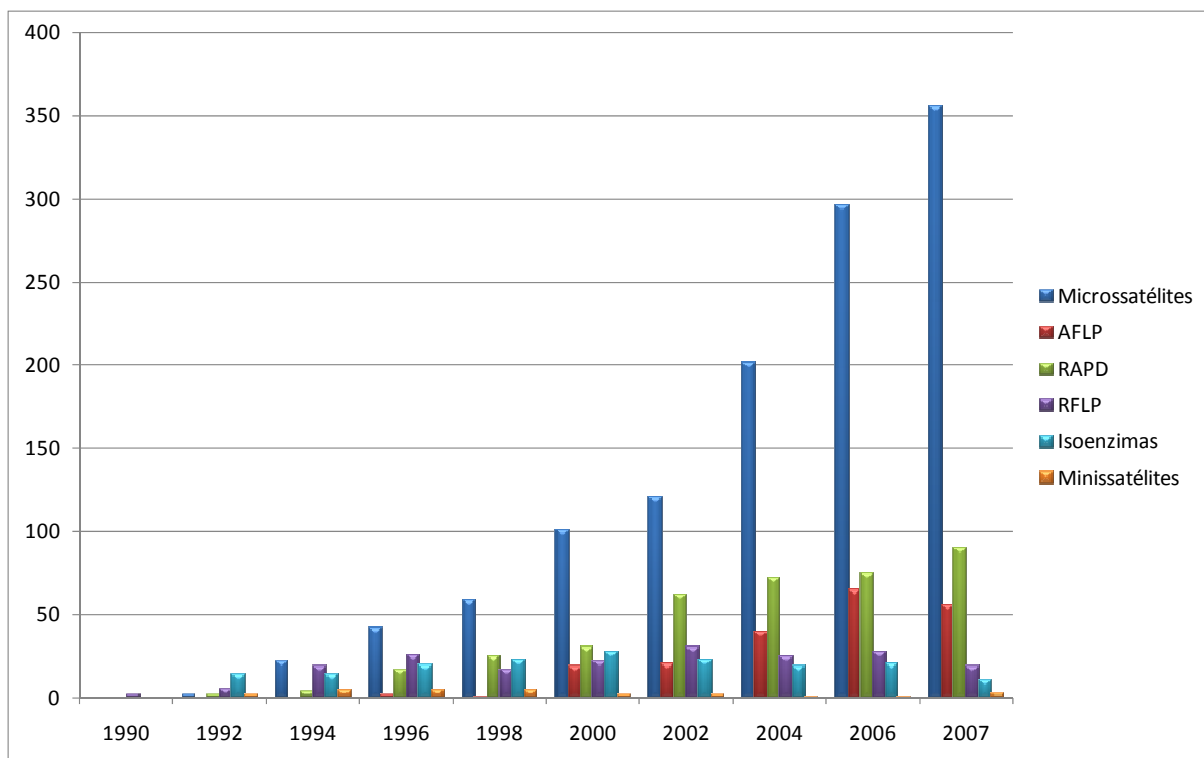


Figura 3.2 Gráfico do número de artigos desde 1990 a 2007, a cada dois anos, onde cada barra é referente a um marcador molecular diferente pesquisado usando a ferramenta de busca do ISI Web of Science por palavras-chaves da técnica e conservação (*conservation*).

Observa-se um grande aumento no número de artigos com microssatélites ao longo dos anos, sendo esta a mais importante técnica utilizada atualmente em estudos de conservação. Por outro lado, os marcadores moleculares, que não dependem do PCR, estão sendo cada vez menos utilizados neste tipo de estudo, principalmente por causa das dificuldades com relação à grande quantidade de DNA necessária para análise e ao trabalho intenso.

Conseqüentemente, os microssatélites são os marcadores moleculares com maior contribuição na totalidade dos estudos de conservação, tendo mais de 1400 artigos publicados nos últimos 18 anos, número três vezes maior do que do RAPD, o segundo marcador mais

utilizado (Figura 3.3). Outro dado interessante desta pesquisa é que os microssatélites, além de possuírem a maior contribuição nas publicações científicas dentre todos os marcadores, ele e o RAPD são os únicos marcadores em ascensão nos últimos anos, sendo estes valores muito mais expressivos para os microssatélites (Tabela 3.3).

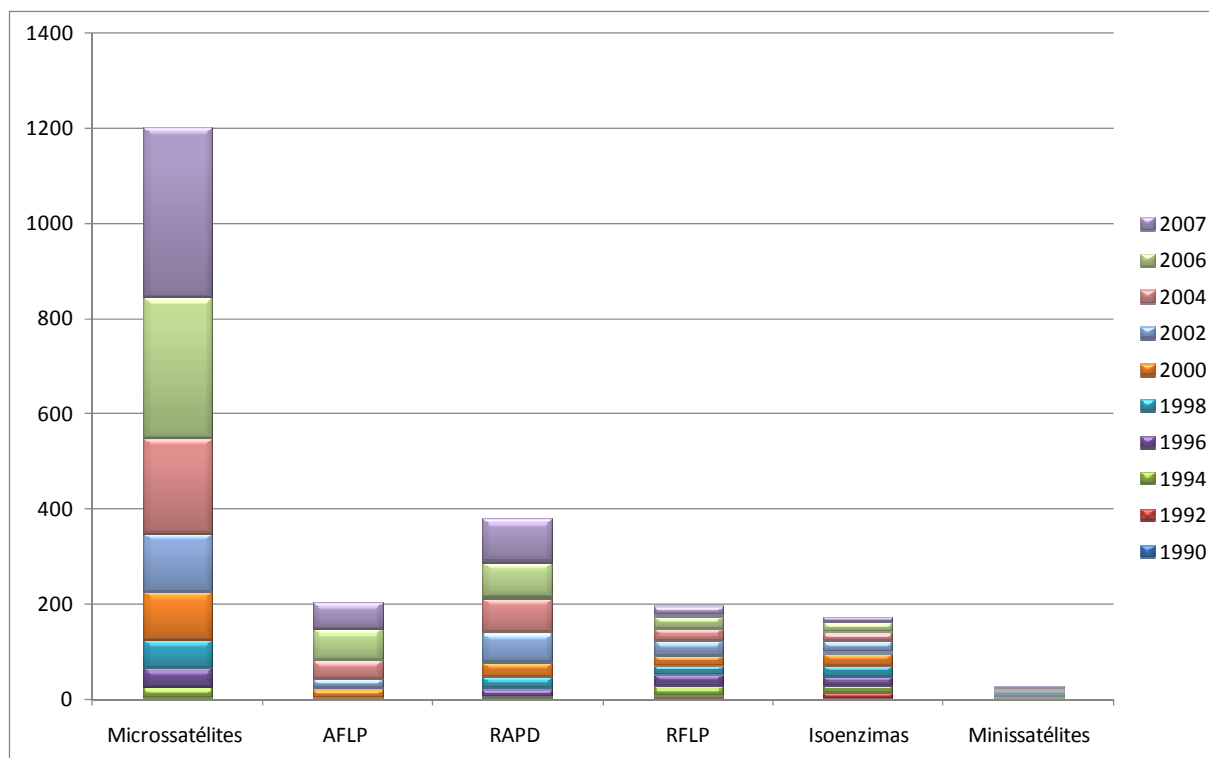


Figura 3.3 Gráfico do número total de artigos desde 1990 a 2007, a cada dois anos, por marcador molecular. As cores são correspondentes ao ano de publicação dos artigos como indicado na legenda. Busca feita pelo ISI Web of Science usando as palavras-chaves da técnica e conservação (*conservation*).

Tabela 3.3 Tabela do número de artigos publicados na área da conservação em cada ano pesquisado desde 1990 a 2007, para cada marcador molecular individualmente.

	Microssatélites	AFLP	RAPD	RFLP	Isoenzimas	Minissatélites
1990				2		
1992	2		2	6	14	2
1994	22		4	19	14	5
1996	42	2	17	26	20	5
1998	59	1	25	17	23	5
2000	101	19	31	22	28	2
2002	121	21	62	31	23	2
2004	202	40	72	25	19	1
2006	296	65	75	28	21	1
2007	356	56	90	19	11	3
Total	1201	204	378	195	173	27

A popularidade dos microssatélites em estudos genéticos conservacionistas pode ser evidenciado pela criação da revista *Molecular Ecology Notes*, dedicada quase que exclusivamente a publicações de isolamento e caracterização de *primers* e dados de frequência alélica obtidos para uma grande variedade de espécies (Liu and Cordes, 2004). Esta revista foi criada em 2001, devido à grande quantidade de publicações do tipo “*Primer Notes*” na revista *Molecular Ecology*, tendo como proposta ser um veículo de rápida disseminação dos avanços tecnológicos em ecologia molecular, como programas de computador, inovações metodológicas e desenvolvimento de marcadores moleculares (detalhes em <http://www.blackwellpublishing.com>). Esta revista, chamada em 2008 de *Molecular Ecology Resources*, detém boa parte de todas as publicações usando microssatélites em estudos de conservação (Figura 3.4) e seu fator de impacto é 1,22.

O periódico *Molecular Ecology*, que aceita publicações que utilizam técnicas moleculares para responder a questões ecológicas, de evolução, comportamento ou conservação, tem fator de impacto 4,825 sendo a 10^a no ranking de citações em ecologia e biologia evolutiva. Apesar de seu bom fator de impacto e desempenho no ranking das citações, sua contribuição com trabalhos de conservação usando microssatélites como ferramenta diminuiu neste ano (Figura 3.4). Por outro lado, a revista *Conservation Genetics* que tem fator de impacto 1,429 sendo a 7^a no ranking de citações em conservação da biodiversidade, teve um aumento em sua contribuição no número de publicações com microssatélites representando 27% de todas as publicações deste assunto em 2008, valor bem mais alto do que 2006 (11%) e 2007 (16%) (Figura 3.4). Mas outras revistas também podem ser citadas como importantes na divulgação de trabalhos moleculares que visam o maior conhecimento da biologia da espécie e as possíveis estratégias conservacionistas a serem adotadas, como *Conservation Biology*, *Biological Conservation*, *Heredity* e *Genetics*.

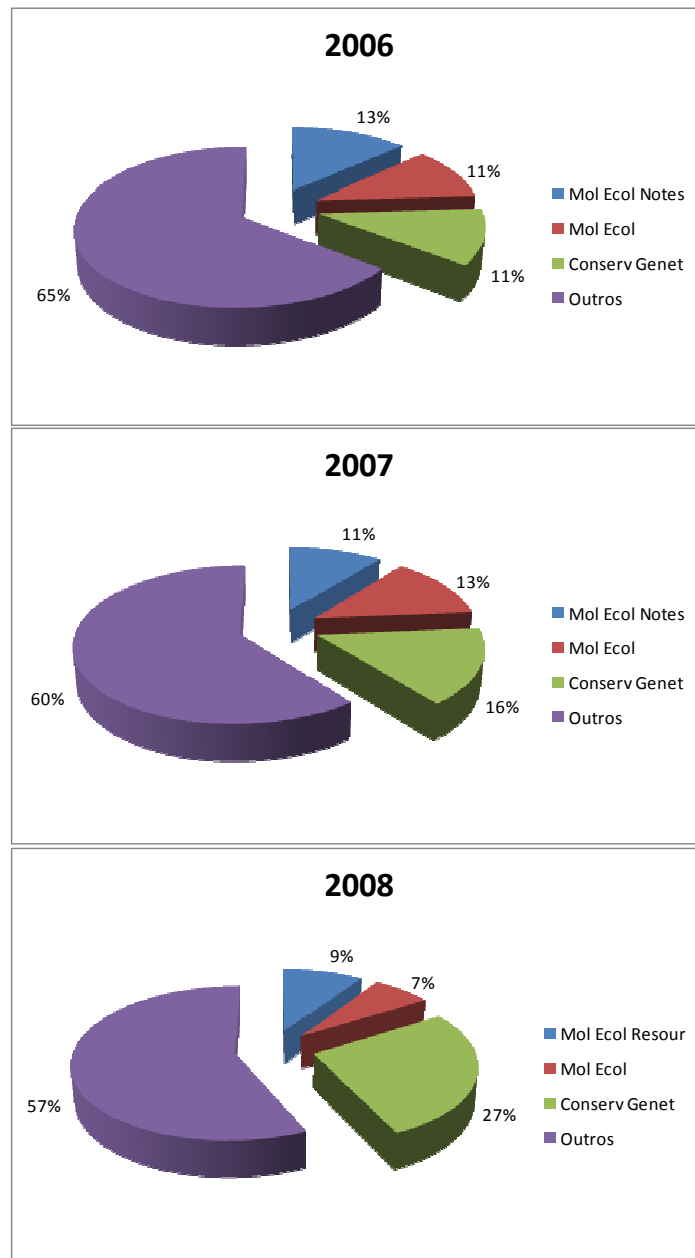


Figura 3.4 Gráficos da contribuição relativa de cada periódico: *Molecular Ecology Notes* (*Mol Ecol Notes*) ou *Molecular Ecology Resources* (*Mol Ecol Resour*), *Molecular Ecology* (*Mol Ecol*) e *Conservation Genetics* (*Conserv Genet*) com publicações sobre conservação usando marcadores microssatélites em 2006, 2007 e primeiro semestre de 2008.

3.1.4 Espécies de mangue

O manguezal é um ecossistema bastante dinâmico e altamente resiliente, exposta aos processos de transição entre os ambientes marinho, estuarino e lagunar, com alternância de inundações derivadas da atuação das marés em regime mixohalino (Saenger, 2003). Eles ocorrem nas costas tropicais e subtropicais, e diariamente são submetidos a variações de salinidade da água e oscilações do nível do mar (Saenger, 2003; Barth 2006). Suas florestas caracterizam-se por espécies vegetais lenhosas típicas, que possuem adaptações quanto à sustentação em substrato não consolidado, ação das marés, aeração do sistema radicular e tolerância à salinidade e ao estresse hídrico (Tomlinson, 1986).

No Estado do Rio de Janeiro, o ecossistema manguezal apresenta grandes formações na região Norte Fluminense (foz do rio Paraíba do Sul), na Baía de Guanabara, na Baía de Sepetiba (Guaratiba) e na Baía da Ilha Grande (municípios de Angra dos Reis e Paraty), região Sul Fluminense. Com relação às espécies, ocorrem *Avicennia schaueriana*, *A. germinans*, *Laguncularia racemosa*, e *Rhizophora mangle*, sendo *A. germinans* restrita ao norte do Estado.

Já existem marcadores microssatélites caracterizados para algumas espécies de mangue, como *A. marina* (Maguire *et al.* 2000; Geng *et al.*, 2007), *A. alba* (Teixeira *et al.* 2003), *A. germinans* (Nettel *et al.*, 2005; Céron-Souza *et al.* 2006), *R. stylosa* (Islam *et al.* 2004), *R. mangle* (Rosero-Galindo *et al.*, 2002) e mais recentemente de *L. racemosa* também (Nettel *et al.*, 2008). Apesar disso, não existem trabalhos sobre a estrutura populacional ou o fluxo gênico de nenhuma destas espécies nos manguezais do Brasil.

3.2 Objetivos

O principal objetivo deste capítulo foi isolar marcadores microssatélites de duas espécies de mangue visando a conservação deste ecossistema.

Os objetivos específicos são:

- Isolar marcadores do tipo SSR-EST a partir do banco de EST de *Avicennia marina*;
- Testar os marcadores SSR-EST isolados nas espécies brasileiras de *Avicennia*: *A. schaueriana* e *A. germinans*;
- Isolar marcadores SSR nucleares aleatórios de *A. schaueriana* e *L. racemosa* através da criação de bibliotecas de DNA enriquecidas com microssatélites;
- Caracterizar os marcadores SSR aleatórios passíveis de serem utilizados em estudos populacionais.

3.3 Materiais e Métodos

3.3.1 SSR-EST

3.3.1.1 Isolamento dos marcadores SSR-EST

O isolamento de marcadores do tipo SSR-EST tem sido feito como uma forma de contornar as desvantagens do isolamento baseado na criação de bibliotecas genômicas enriquecidas com microssatélites. A existência de bancos de EST disponíveis para a comunidade científica, principalmente de plantas cultivadas, facilita o isolamento de marcadores microssatélites de regiões gênicas. Estes marcadores, além de serem mais conservados entre espécies próximas por estarem presentes em regiões transcritas, aparentemente eles são também mais facilmente caracterizados em laboratório, devido à maior clareza dos fragmentos amplificados e menor taxa de mutação (Woodhead *et al.* 2005).

O instituto TIGR (*The Institute for Genomic Research* - <http://www.tigr.org/>) possui uma grande coleção de bancos de dados sobre genômica de plantas. Dentre os bancos de DNA das espécies cultivadas já seqüenciadas, também podemos encontrar bancos de EST (na seção *TIGR Plant Transcript Assemblies*) de diversas espécies vegetais. A espécie de mangue *Avicennia marina* possui algumas seqüências de EST depositadas no banco de dados do instituto TIGR.

Seqüências de *A. marina* foram utilizadas na busca por marcadores microssatélites através do programa MISA (<http://pgrc.ipk-gatersleben.de/misa/>; Thiel *et al.*, 2003). Este

programa permite definir o tipo de microssatélite que será buscado na biblioteca de EST, com o número mínimo de repetições de bases. Neste trabalho, buscamos repetições de mononucleotídeos com no mínimo 10 repetições; de dinucleotídeos com no mínimo 6 repetições; e para os outros microssatélites tri-, tetra-, penta- e hexanucleotídeos com no mínimo 5 repetições; possibilitando inclusive combinações de diferentes motivos.

Após a identificação das seqüências com microssatélites, elas foram analisadas usando a ferramenta BlastX, que compara a seqüência de nucleotídeos traduzida em proteína com o banco de dados de proteínas, pelo site do NCBI (*National Center for Biotechnology Information* – <http://www.ncbi.nlm.nih.gov>). E assim foi possível identificar as seqüências de maior homologia com proteínas (*e-value*²⁴ maior que $1e^{-5}$) e suas regiões 5'UTR, 3'UTR e região codificante do gene, a fim de confirmar a localização dos microssatélites em regiões transcritas de um gene.

A partir destes resultados, foram desenhados *primers* para todas estas seqüências através do programa *Primer 3*, versão 0.4.0 (Rozen and Skaletsky 2000), com dados-padrão do programa, exceto pelo tamanho médio do *primer* definido em 24 nucleotídeos e da T_m , definida em 65°C (Tabela 3.4).

²⁴ Um parâmetro que descreve o número de seqüências homólogas que seriam esperadas de ocorrer ao acaso durante a busca em um banco de dados de seqüências.

Tabela 3.4 Sequências dos quinze pares de *primers* desenhados para amplificar marcadores microssatélites do tipo SSR-EST nas espécies de *Avicennia*.

<i>Primer</i>	Forward	Reverse
SSRAM01	5'AATTTGCTCGCAGAAGTGTGAGTG3'	5'GGTGAAGTAAACTCCGGAATCCT3'
SSRAM02	5' AGGTCTCCCCAAGAAATA3'	5' TTCGAGCTCCCATATGATCC3'
SSRAM03	5'TAAATGAGCCCCATTCAAGGAAAA3'	5'AATTTAAAACCGCCCTGGTTCAAAG3'
SSRAM04	5'GGTGGCCGTAAGATCAATGAGAAC3'	5'TACACACACGCACATAGCCAAGAA3'
SSRAM05	5'GCTCGTTCGAGAAGATCAAGGAGAA3'	5'CTTATTGCAACCCAGTAGCAACC3'
SSRAM06	5'AAGCCCAGCCCTTTTTACTCTTTG3'	5'TCCATAGTGACGCGTTCCTTACAA3'
SSRAM07	5'TCGCTTATCCGTTCTTGACTACCC3'	5'TCCATCACTATCCTCTCCCTCACC3'
SSRAM08	5'TACATTTGCCACCTCCTTCAGTTG3'	5'AAGCTCAAGAGCAGCAGCAAAGAC3'
SSRAM09	5'GCGTGGGGCAGTTATACAAGCAT3'	5'TTGTCCAAGGTCATCATCTTTTGC3'
SSRAM10	5'CAGTGGACTGCTTTGGTGCATAAC3'	5'CCATTTCTGCAAGAGAAGTGGAGA3'
SSRAM11	5'GCACGCTTTCAATGCACTACCATA3'	5'CTTCCCCTTCTTCGCTGACTTTTT3'
SSRAM12	5'AATCCCACAGACCACCAACTGAAC3'	5'CAACCAAATCTTGCGCTCTTTTT3'
SSRAM13	5'GCAGCTGCAATGGAGATAAATTCG3'	5'TTCTTGTCCAAGGGTCAAAGGTGT3'
SSRAM14	5'TCAAGTGATCGTCGTGGTGGTTAT3'	5'AACAAGGCCATATGTTTCTTCACC3'
SSRAM15	5'GTCGAATTCTACCTCCGCCGTAAG3'	5'TCGATGATCAAGAAACCCTTAAGCA3'

3.3.1.2 Caracterização dos marcadores SSR-EST

Os *primers* desenhados para os microssatélites SSR-EST foram testados em ambas as espécies *A. schaueriana* e *A. germinans*, que ocorrem no Brasil. Foram feitos vários PCRs de teste para encontrar a concentração de DNA ideal, além da temperatura de anelamento ótima para cada par de *primer* e a quantidade de ciclos do programa.

As reações de PCR ideais usavam 12ng de DNA genômico, 0,5µM de cada *primer*, 200µM de dNTP, tampão de Taq 1X, 1,5mM de MgCl₂ e 0,8U de Taq Polimerase num volume total de 10µl. Foram usados três programas de PCR, dependendo do marcador analisado. O primeiro consistia em desnaturação inicial de 1 minuto, 38 ciclos de 94°C por 15 segundos, 63°C por 45 segundos e 72°C por 45 segundos, seguidos de extensão final de 5

minutos e foi usado com os *primers* SSR-AM01, SSR-AM02, SSSR-AM07, SSR-AM11, SSR-AM12 e SSR-AM13. O segundo programa era um *touchdown* com desnaturação inicial de 1 minuto, 7 ciclos com diminuição de 0,7°C por ciclo com os seguintes parâmetros 94°C por 15 segundos, 63,6°C por 45 segundos e 72°C por 45 segundos, seguidos de 35 ciclos com temperatura de anelamento fixa em 58°C e extensão final de 5 minutos, que funcionou com os *primers* SSR-AM09 e SSR-AM10. O terceiro programa é igual ao segundo exceto pelos tempos, onde são usados 15 segundos de desnaturação, 15 segundos de anelamento e 30 segundos de extensão, usado para amplificar os marcadores SSR-AM03. Os produtos de PCR foram visualizados em gel de agarose 1,5% (TBE 0,5X) corado com brometo de etídio.

3.3.2 SSR nucleares aleatórios

O isolamento de marcadores do tipo microssatélites nucleares aleatórios das espécies de mangue *Laguncularia racemosa* e *Avicennia schaueriana* foi feito através da varredura de uma biblioteca de DNA enriquecida com microssatélites (Noel *et al.*, 2005; Le Cadre *et al.*, 2005). As primeiras etapas da criação do banco de DNA enriquecido foram: (1) clivagem do DNA; (2) ligação de adaptadores ao DNA digerido; (3) enriquecimento dos fragmentos com microssatélites; e (4) clonagem e transformação da biblioteca em bactérias.

Uma vez criada a biblioteca de fragmentos de DNA enriquecidos com microssatélites, foi necessário fazer uma busca pelos clones positivos que seriam seqüenciados. Os clones que possuíam plasmídeos com microssatélites foram então avaliados quanto à possibilidade de desenhar *primers* específicos para a região de interesse.

Esta parte do trabalho foi desenvolvida em colaboração com o laboratório “Conservation des Espèces, Restauration et Suivi des Populations” do Museu Nacional de História Natural de Paris, na França, pelo programa de doutorado sanduíche da CAPES sob a co-orientação da Dra Nathalie Machon.

3.3.2.1 Criação da biblioteca de DNA enriquecida com microssatélites

O primeiro passo na construção da biblioteca de DNA foi a clivagem de 1µg de DNA genômico total com 5U das enzimas *AluI*, *HaeIII* e *RsaI* separadamente numa reação de 50µl, à 37°C por 4 horas. Após confirmação das digestões do DNA, foram selecionados fragmentos entre 300 e 1000 pares de base através de gel de agarose 1%. Com o DNA eluído do gel, mas também utilizando DNA digerido total, foram feitas ligações dos adaptadores com 10µl do DNA digerido, 5ng de cada adaptador MLU1A (5' CTCTTGCTTACGCGTGGACTA 3') e MLU1B (5' TAGTCCACGCGTAAGCAAGAGCACA 3'), 2U de T4 DNA ligase numa reação de 50µl e incubamos por 2 horas à 37°C. Depois a enzima foi inativada à 65°C por 10 minutos. Em seguida, foram feitas seis reações de PCR para todos os DNAs ligados a fim de obter quantidade de DNA suficiente para um enriquecimento (10µg de produto de PCR). Usou-se então 1µl do DNA ligado, 1,8mM de MgCl₂, tampão QBio Taq 1X, BSA na concentração final de 200µg/µl, 0,1mM de dNTP, 250ng de adaptador MLU1A (usado aqui como *primer*) e 1U de Taq QBio, numa reação de 50µl. O programa teve uma primeira desnaturação por 1 minuto, seguido de 23 ciclos de 94°C por 1 minuto, 55°C por 1 minuto e 72°C por 2 minutos, com extensão final de 10 minutos. As amplificações das ligações foram

checadas em gel de agarose 2%. A fim de aumentar a diversidade dos fragmentos digeridos em cada enriquecimento, foram misturados os produtos de três PCRs com DNA ligado e digerido com duas enzimas diferentes, após purificação dos mesmos.

Para enriquecer os fragmentos de DNA digeridos e ligados, foram utilizadas duas misturas de oligonucleotídeos²⁵ de repetição: Pool A com os trinucleotídeos (GTT)₁₀, (GTC)₈ e (GAT)₉; e Pool C com os dinucleotídeos (TC)₈ e (TG)₈. A captura dos fragmentos com microssatélites foi feita usando micropartículas imantadas, que graças aos anticorpos (estreptavidina) presentes em sua superfície, podem se associar a moléculas marcadas com biotina, neste caso os oligonucleotídeos de repetição. Para tal, usamos o sistema “Streptavidin Magnesphere^R Paramagnetic Particles” da Promega^R. Primeiro as partículas magnéticas foram preparadas retirando a solução de estocagem, após juntar as partículas com a ajuda de um ímã na parede do tubo. Adicionou-se 300µl de SSC 0,5X, misturando gentilmente e novamente as partículas foram separadas na parede do tubo com o ímã. A solução de SSC foi então retirada do tubo e nova solução foi adicionada. Este processo se repetiu três vezes até a lavagem completa das partículas imantadas. Ao final, as partículas foram ressuspensas em 100µl de SSC 0,5X.

As amostras foram preparadas adicionando-se 400µl de água aos 100µl de produto de PCR purificado, desnaturando à 94°C por 15 minutos e juntando cada oligonucleotídeo marcados com biotina na concentração final de 0,3µM e também SSC na concentração final de 0,5X. Misturar gentilmente e incubar por 20 minutos à temperatura ambiente, misturando novamente a cada 2 minutos. As amostras foram adicionadas aos tubos contendo as partículas

²⁵ Moléculas curtas, geralmente de 6 a 100 nucleotídeos, de DNA fita simples.

imantadas e misturados com a pipeta. Incubar por 10 minutos à temperatura ambiente girando o tubo gentilmente para evitar o acúmulo das partículas no fundo. Usar o ímã para juntar as partículas magnéticas, que estavam neste ponto ligadas covalentemente aos oligonucleotídeos marcados com biotina que por sua vez estão ligados aos fragmentos de DNA complementar às seqüências dos oligonucleotídeos, e em seguida remover a solução com a pipeta. Fazer quatro lavagens com 300µl de SSC 0,1X, incubando por 10 minutos cada lavagem e removendo a solução com a ajuda do ímã. Por último, adicionar 100µl de água, misturar, incubar por 1 minuto à temperatura ambiente e transferir o líquido, que é o enriquecimento, para um tubo de 1,5ml novo. Repetir a eluição mais uma vez com 100µl de água e depois com 50µl. Juntar todo o DNA enriquecido num único tubo e guardar na geladeira a 4°C.

Foi feito então um novo PCR para confirmar a captura dos fragmentos com microssatélites usando 10µl do DNA enriquecido e as mesmas condições do PCR da ligação, mas com volume final de 40µl. O programa teve uma desnaturação inicial de 1 minuto seguida de 21 ciclos a 94°C por 1 minuto, 60°C por 1 minuto e 72°C por 2 minutos, com extensão final de 1 hora. Os PCRs foram checados em gel de agarose 1,5%. Foram feitos cinco PCRs para cada enriquecimento, sendo todos do mesmo DNA purificados juntos. As clonagens foram feitas usando o kit “TOPO TA Cloning” da Invitrogen^R usando 4µl de produtos de PCR, 1µl de solução salina e 1µl do vetor TOPO e incubando por 30 minutos à temperatura ambiente. As células quimicamente competentes “One Shot Mach 1” da Invitrogen^R foram usadas na transformação com 2µl de reação de clonagem. Após incubar no gelo por 30 minutos, a 42°C por 30 segundos e gelo novamente, foram adicionados 250ml de meio SOC à temperatura ambiente e os tubos colocados em agitação a 37°C por 1 hora antes

de espalhar sobre placas de Petri com meio LB mais 0,1mg/ml de ampicilina, 15µg/ml de Xgal e 10µg/ml de IPTG. Incubar as placas durante a noite a 37°C. Foram feitas três placas para cada transformação usando 50µl, 100µl e 150µl das células transformadas.

3.3.2.2 Busca de clones positivos

Após o crescimento das bactérias transformadas com os fragmentos de DNA enriquecidos com microssatélites, foram selecionadas todas as colônias brancas, ou seja, todas as colônias cujas bactérias possuem plasmídeos clonados no seu interior. Estas colônias brancas foram repicadas em duas placas de petri com meio LB e 0,1mg/ml de ampicilina, que possuíam um papel quadriculado com marcação de 1 a 145, onde cada colônia foi posicionada de forma idêntica em duas placas iguais. As placas foram incubadas a 37°C durante a noite. Uma das placas em replicata foi usada na transferência das colônias para membranas de Nylon Hybond N+ (Amersham^R). As membranas foram deixadas por 5-10 minutos em contato com as colônias e colocadas na solução de lise e desnaturação (NaOH 0,5M e NaCl 1,5M) por 10 minutos, depois na solução de neutralização (Tris-HCl 0,5M pH 7,5 e NaCl 1,5M) por 10 minutos e por último na solução de lavagem (SSC 2X) por 5 minutos. Depois as colônias foram fixadas nas membranas através de incubação a 80°C por 1 hora e estas foram guardadas à temperatura ambiente.

Estas membranas foram usadas em reações de hibridização com a mistura dos oligonucleotídeos em repetição marcados radioativamente com $\gamma^{32}\text{P}$ -ATP e purificados através de coluna de poliacrilamida. Primeiro as membranas foram incubadas em agitação a

45°C por 45 minutos com a solução R1 (SSC 2X e SDS 0,1%) mais 0,1mg/ml de proteinase K. Após lavar duas vezes com solução R1 nova a 45°C por 10 minutos, as membranas foram transferidas para os tubos de hibridização com a solução de hibridização (SSC 6X e SDS 0,1%) mais 0,2mg/ml de esperma de salmão pré-fervido (90°C por 5 minutos). E foram incubadas por 2 horas a 45°C (pool C de dinucleotídeos) e a 65°C (pool A de trinucleotídeos). A solução foi descartada e nova solução com esperma de salmão a 0,01mg/ml pré-fervido e os oligonucleotídeos marcados radioativamente foi colocada no lugar da antiga. As membranas foram incubadas durante a noite nas mesmas temperaturas da incubação anterior, dependendo do pool de oligonucleotídeos utilizado. No dia seguinte, foram feitas três lavagens com SSC 6X a 40-45°C por 15 minutos cada. As membranas foram colocadas em filme plástico e reveladas no *PhosphorImage* (FujiFilm^R).

3.3.2.3 Seqüenciamento da biblioteca de DNA

Para fazer o seqüenciamento dos fragmentos de DNA clonados, as colônias da placa de Petri em duplicata, que não foi usada para as membranas, foram coletadas com um palito e colocadas em 50µl de água milliQ, incubadas a 98°C por 10 minutos e centrifugadas a 12.000 rcf por 5 minutos. Em seguida, foi feito um PCR com tampão de Taq 1X, 1,5mM de MgCl₂, 0,25µM dos *primers* M13 (sense) e M13 (anti-sense), 0,1mM de dNTP, 1,5U de Taq polimerase e 1,5µl do clone desnaturado em 30µl de volume final. O programa utilizado foi 96°C por 5 minutos, 5 ciclos de 96°C por 1 minuto, 52°C por 1 minuto e 72°C por 3 minutos, 20 ciclo de 96°C por 1 minuto, 52°C por 1 minuto e 72°C por 2 minutos, com extensão final

de 10 minutos. Os produtos de PCR foram checados em gel de agarose 1,5% corado com brometo de etídio e as amostras com produto de amplificação maior de 500pb foram selecionadas para um segundo PCR feito com a mistura dos oligonucleotídeos de repetição. Neste PCR usamos tampão de Taq 1X, 1,5mM de MgCl₂, 0,15μM de *primer* M13 (sense ou anti-sense), 0,15μM de cada oligonucleotídeo do Pool A ou C, correspondentes ao enriquecimento da amostra, 0,1mM de dNTP, 1,5U de Taq polimerase e 1μl do produto do PCR anterior. Foi usado o mesmo programa de PCR da amplificação M13-M13. Os PCRs foram analisados em gel de agarose 1,2% com TBE 0,5X e corado com brometo de etídio. As amostras que conseguiram ser amplificadas pelo segundo PCR (M13-oligo) foram seqüenciadas pela empresa Genoscreen.

3.3.2.4 Análise das seqüências e desenho dos *primers*

As análises das seqüências dos clones foram feitas através da busca das seqüências dos adaptadores MLU1A e MLU1B e em seguida nucleotídeos com repetição em tandem como os microssatélites. Todos os fragmentos de DNA de ambas as espécies *L. racemosa* e *A. schaueriana* que possuíam microssatélites foram analisados pelo programa *Primer 3*, versão 0.4.0 (Rozen and Skaletsky 2000), para buscar *primers* que amplifiquem a região dos microssatélites. Foram definidos o tamanho médio dos *primers* (24 nucleotídeos) e a T_m (65°C), além do conteúdo GC (aceitável entre 40 e 60%) para obter *primers* com boa amplificação e reprodutibilidade. Os *primers* sense foram desenhados com uma cauda de

M13, que será utilizada para a marcação com fluorescência nas genotipagens (Schuelke, 2000).

3.3.2.5 Caracterização dos *primers* SSR nucleares

De todos os *primers* desenhados foram testados 12 pares, metade proveniente do banco de DNA de cada uma das espécies: LA01 e LA14 (Pool A); LC01, LC02, LC44 e LC55 (Pool C) para *L. racemosa* (Tabela 3.5) e AC56, AC82, AC101, AC108, AC128 e AC177 (Pool C) para *A. schaueriana* (Tabela 3.6). Todos os *primers* foram testados nas condições de PCR com 12ng de DNA genômico, tampão da Taq 1X, 1,5mM de MgCl₂, 0,5μM de cada *primer*, 200μM de dNTP e 0,8U de Taq polimerase num volume final de 10μl. Os programas de PCR tiveram temperaturas de anelamento diferentes conforme o *primer* usado, mas seguiu sempre os mesmos parâmetros básicos com desnaturação inicial de 2 minutos, seguido de 32 ciclos de 94°C por 15 segundos, anelamento por 15 segundos, 72°C por 10 segundos com extensão final de 30 minutos. Os produtos de PCR foram visualizados em gel de agarose 2% (TAE 0,5X) corado com brometo de etídio.

Tabela 3.5 Sequências dos *primers* desenhados para amplificar marcadores SSR específicos para *L. racemosa* e as temperaturas de anelamento (T) ideais. A cauda dos *primers* sense é a parte sublinhada das sequências.

Primer	Sequências	T (°C)
LA01-	5' <u>TG</u> TAAAACGACGGCCAGTAAATTGGGAAAGACAAAACCTCC3'	56
LA01-	5'TGAGGAGGAGGAAGAGTACGATG3'	
LA14-	5' <u>TG</u> TAAAACGACGGCCAGTTGAGAGAAGGGATTGGTTGTATTTG3'	58
LA14-	5'TTTCACGGCAAGAATTGGTAAG3'	
LC01-	5' <u>TG</u> TAAAACGACGGCCAGTGCCTAACACCACAAGAAAGGAG3'	60
LC01-	5'TGCCACCAAATTCACAGAAAG3'	
LC02-	5' <u>TG</u> TAAAACGACGGCCAGTAACAGATATTCAAGTACTAAACAACG3'	56
LC02-	5'CACATACACACCACCACACG3'	
LC44-	5' <u>TG</u> TAAAACGACGGCCAGTGGTAGATTGCAATGTCCTAGTTTC3'	60
LC44-	5'CCTAAGGGAAGCCAAGTCTG3'	
LC55-	5' <u>TG</u> TAAAACGACGGCCAGTGAGGGAGAGTGAAGTGGGTTTCC3'	60
LC55-	5'CTTCAACCAGCCCATATCCATC3'	

Tabela 3.6 Sequências dos *primers* desenhados para amplificar marcadores SSR específicos para *A. schaueriana*, com suas respectivas temperaturas de anelamento (T) ideais ou testadas nos *primers* que não funcionaram. A cauda dos *primers* sense é a parte sublinhada das sequências.

Primer	Sequências	T (°C)
AC56-F	5' <u>TG</u> TAAAACGACGGCCAGTGAGGAGGAGCACGACAAATTC3'	60
AC56-R	5'TCCAGAAAGAAAGGAACACACTC3'	
AC82-F	5' <u>TG</u> TAAAACGACGGCCAGTCAATGCACAAATTTCTAGTGAACG3'	56
AC82-R	5'GGCAACTTAGAGTTGGATTGG3'	
AC101-F	5' <u>TG</u> TAAAACGACGGCCAGTGTTTCGTTTCATGAGTAAACTCGAC3'	56
AC101-R	5'CATTTTCTTTGGACAAACATGG3'	
AC108-F	5' <u>TG</u> TAAAACGACGGCCAGTCCTAACCCGACTACTGTGTGGTG3'	60
AC108-R	5'CCAAATCATTTCGAGGCAACTTAG3'	
AC128-F	5' <u>TG</u> TAAAACGACGGCCAGTCACATTTGCACAGACACAAAC3'	60
AC128-R	5'TCCCATTGTCTAACCCAACC3'	
AC177-F	5' <u>TG</u> TAAAACGACGGCCAGTCGTTCTCCATCCATCCTTG3'	60
AC177-R	5'GCGGCTTAGAGTTGGGTTG3'	

3.4 Resultados

3.4.1 Isolamento dos marcadores microssatélites do tipo SSR-EST

Avicennia é um gênero que ocorre exclusivamente em ecossistemas de manguezais. No Brasil, existem duas espécies pertencentes a este gênero: *A. schaueriana*, que ocorre em todo o Brasil; e *A. germinans*, que ocorre mais abundantemente nas regiões norte e nordeste. A espécie *A. marina*, porém, possui um banco de dados de EST disponível pelo instituto TIGR (*The Institute for Genomic Research* - <http://www.tigr.org/>). Com o objetivo de isolar marcadores microssatélites do tipo SSR-EST para serem usados nas espécies brasileiras de *Avicennia*, 1.539 seqüências de EST de *A. marina* foram utilizadas na busca destes microssatélites.

De um total de 688.537 pares de bases de nucleotídeos, foram encontrados 69 microssatélites presente em 62 seqüências, dos quais 16 eram mononucleotídeos (23,2%), 31 eram dinucleotídeos (44,9%) e 22 trinucleotídeos (31,9%) (Tabela 3.7).

Tabela 3.7 Ocorrência de microssatélites nas seqüências de EST de *A. marina* organizadas por motivo e número de repetições.

SSR	Número de repetições											Total
	5	6	7	8	9	10	11	12	13	14	>15	
A							2		1	2	1	6
C								1				1
T						5	1	1	1	1		9
AC											1	1
AT		2	1		3						1	7
CA		1		2	1							4
CT		2	1		1				3			7
GA									1			1
GT		1		1								2
TA		2		1								3
TC		2	1	1						1		5
TG				1								1
AAG	1											1
AGA	1	1	1									3
ATA	1											1
ATC					1							1
CAG	1		1									2
CCA		1										1
CCG	1											1
CCT	4											4
GAT	1											1
GCG		1										1
GGC			2	1								3
TCT		1		1								2
TTA					1							1

As 62 seqüências de EST com microssatélites foram avaliadas pela ferramenta Blastx, que compara as seqüências de nucleotídeos traduzidas em proteínas com um banco de proteínas. Foi possível identificar 22 seqüências com microssatélites que possuíam alta homologia (*e-value* maior que e^{-5}) com proteínas. Através do programa *Primer 3*, foram definidos 15 pares de *primers* de regiões flangeadoras dos microssatélites, dos quais seis estavam presentes na região 3'UTR do gene, outros seis na região 5'UTR e três dentro do gene.

Dos 15 pares de *primers* desenvolvidos, nove amplificaram bandas únicas para as duas espécies de *Avicennia* do Brasil: *A. schaueriana* e *A. germinans*, apesar de terem sido isolados de seqüências provenientes da espécie *A. marina*, mostrando a alta transferabilidade dos marcadores SSR-EST entre espécies relacionadas (Tabela 3.8). As ampliações dos pares de *primers* SSR-AM01 e SSR-AM02 funcionaram com todas as amostras testadas de *A. schaueriana* e *A. germinans*, mas não com *L. racemosa* (Figuras 3.5 e 3.6, respectivamente), como era esperado porque esta espécie pertence a outra família de plantas. Isso ocorreu também com os *primers* SSR-AM03, SSR-AM07, SSR-AM09, SSR-AM10, SSR-AM11, SSR-AM13. Porém o *primer* SSR-AM12, fugindo à regra, amplificou bandas de tamanhos parecidos aos da *Avicennia* em *L. racemosa* também (Figura 3.7).

Tabela 3.8 Dados dos quinze pares de *primers* desenvolvidos para SSR-EST com a localização, o motivo, tamanho esperado e amplificação boa (+) ou ruim (-) nas espécies *A. schaueriana* (AS), *A. germinans* (AG) e *L. racemosa* (LR).

<i>Primer</i>	Localização	Motivo	Tamanho	AS	AG	LR
SSRAM01	3'UTR	(GAT)5	221pb	+	+	-
SSRAM02	3'UTR	(AT)7	173pb	+	+	-
SSRAM03	3'UTR	(T)4CC(T)13	183pb	+	+	-
SSRAM04	3'UTR	(C)12	187pb	-	-	-
SSRAM05	3'UTR	(AAG)5	196pb	-	-	-
SSRAM06	3'UTR	(AT)6	171pb	-	-	-
SSRAM07	5'UTR	(GCG)6	221pb	+	+	-
SSRAM08	5'UTR	(TC)8	151pb	-	-	-
SSRAM09	5'UTR	(T)11	238pb	+	+	-
SSRAM10	5'UTR	(AT)6	173pb	+	+	-
SSRAM11	5'UTR	(TC)7	203pb	+	+	-
SSRAM12	5'UTR	(GT)6	194pb	+	+	+
SSRAM13	Gene	(CA)6	150pb	+	+	-
SSRAM14	Gene	(GGC)7	222pb	-	-	-
SSRAM15	Gene	(CA)9	175pb	-	-	-

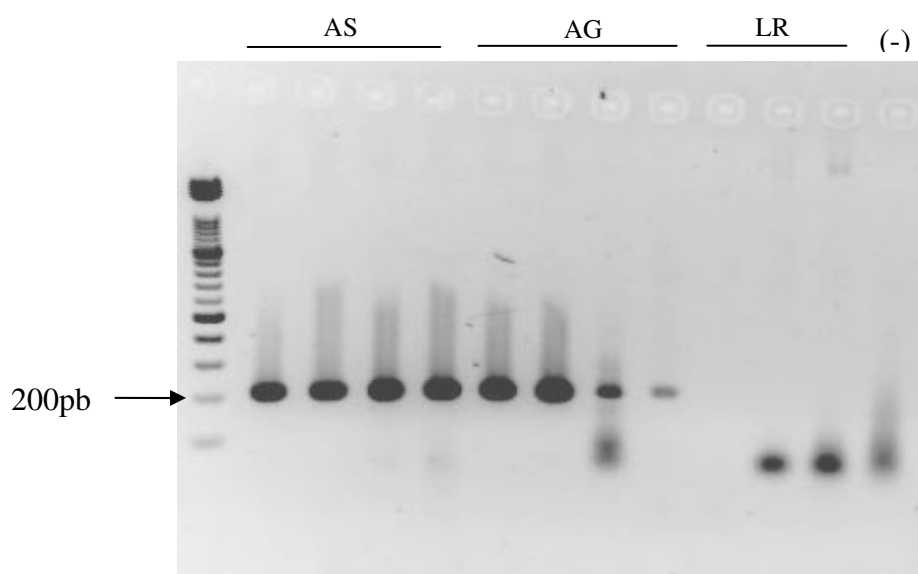


Figura 3.5 Gel de agarose 1,5% com produtos de amplificação do *primer* SSR-AM01 para quatro amostras da *A. schaueriana* (AS), quatro amostras de *A. germinans* (AG), três de *L. racemosa* (LR) e um controle negativo (-). Marcador de peso molecular: XIV.

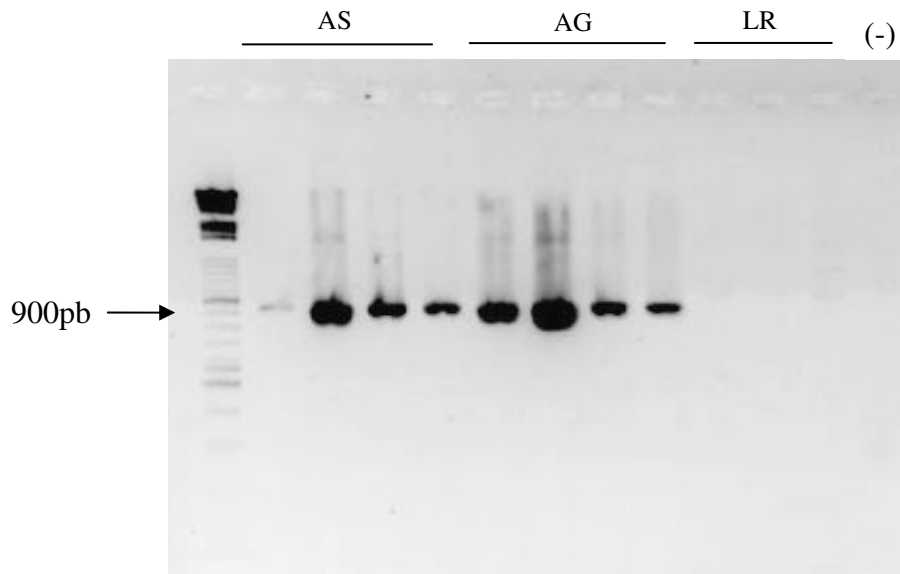


Figura 3.6 Gel de agarose 1,5% com produtos de amplificação do *primer* SSR-AM02 para quatro amostras da *A. schaueriana* (AS), quatro amostras de *A. germinans* (AG), três de *L. racemosa* (LR) e um controle negativo (-). Marcador de peso molecular: XIV.

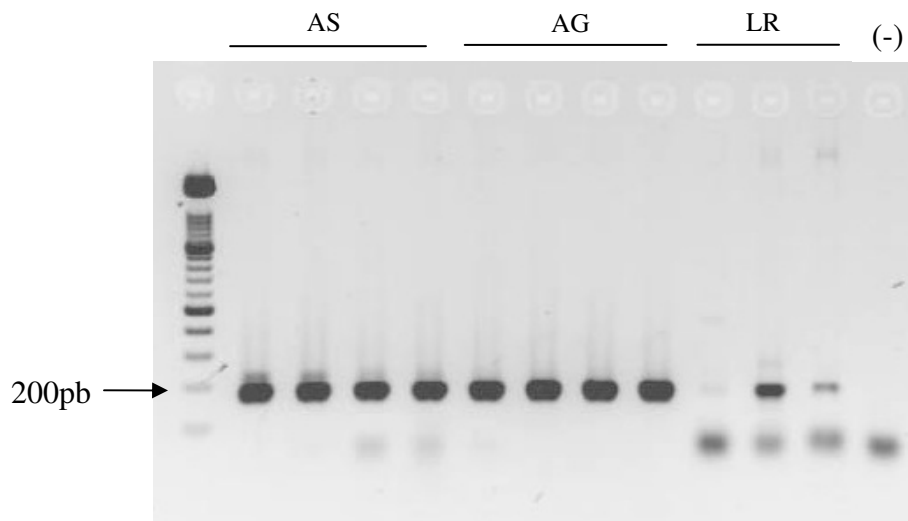


Figura 3.7 Gel de agarose 1,5% com produtos de amplificação do *primer* SSR-AM12 para quatro amostras da *A. schaueriana* (AS), quatro amostras de *A. germinans* (AG), três de *L. racemosa* (LR) e um controle negativo (-). Marcador de peso molecular: XIV.

A agarose não possui resolução suficiente a ponto de ver diferenças muito pequenas, ou seja, de alguns pares de base apenas, por isso não é suficiente para a identificação de polimorfismo entre os indivíduos amostrados ou suas populações. Futuramente faremos a genotipagem de cada *primer* nas populações de *A. schaueriana* e *A. germinans* para definir quais *loci* de marcadores microssatélites isolados possuem polimorfismo nestas espécies para serem usados em estudos populacionais.

3.4.2 Isolamento dos marcadores microssatélites nucleares aleatórios

O isolamento de marcadores de microssatélites nucleares aleatórios foi feito através da criação de um banco de DNA enriquecido com microssatélites. A fim de isolar marcadores tipo SSR aleatórios específicos para espécies de mangue como *L. racemosa* e *A. schaueriana* foram criadas quatro bibliotecas. O primeiro passo foi a digestão do DNA genômico com três enzimas de restrição, *AluI*, *HaeIII* e *RsaI*. O tamanho ideal dos fragmentos para clonagem a fim de obter marcadores SSR é de 300 a 1000 pares de base. Porém, alguns testes com a eluição dos fragmentos a partir de gel de agarose não deram certo porque os produtos de PCR ficaram muito concentrados numa pequena faixa de tamanho, o que não é favorável para obter fragmentos variados na biblioteca de DNA, e por isso algumas amostras não-eluídas do gel também foram utilizadas nas etapas seguintes (Figura 3.8).

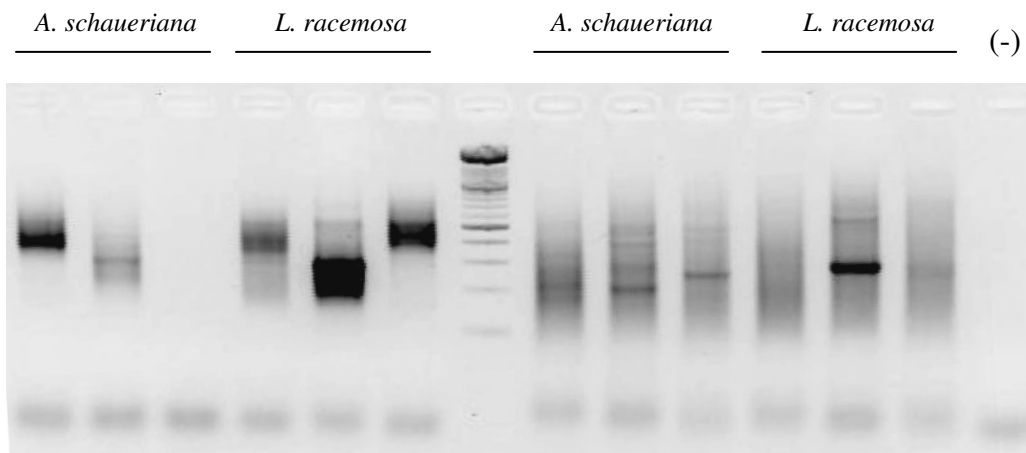


Figura 3.8 Gel de agarose 2% com os produtos da amplificação da ligação das amostras de *L. racemosa* e *A. schaueriana* eluídas do gel (à esquerda do marcador) e sem eluição (à direita do marcador). Digestões na seguinte ordem: *AluI*, *HaeIII* e *RsaI*; para cada amostra. Marcador de peso molecular: XIV. Controle negativo com água (-).

Para a etapa seguinte, foram misturados produtos de PCR do DNA digerido e ligado de duas enzimas diferentes a fim de aumentar a variedade de fragmentos selecionados. Juntamos as amostras de *A. schaueriana* digerida com *AluI* eluídas do gel com aquelas digeridas com *RsaI* sem eluição e chamamos de A1. As amostras de *A. schaueriana* sem eluição digeridas com *AluI* e *HaeIII* foram juntadas e chamadas de A2. As amostras de *L. racemosa* chamadas de L1 foram digeridas com *AluI* e *RsaI* ambas eluídas do gel. E a L2 foram as amostras de *L. racemosa* sem eluição digeridas com *AluI* e *RsaI*. As digestões com *HaeIII* de *L. racemosa* tiveram bandas menores que as outras e por isso decidimos não utilizá-las nas análises seguintes.

Cada uma destas amostras A1, A2, L1 e L2 foram enriquecidas com uma mistura de oligonucleotídeos repetitivos como microssatélites através do uso de partículas magnéticas com anticorpo estreptavidina e oligonucleotídeos biotinilados. A afinidade covalente entre estreptavidina e biotina faz com que os oligonucleotídeos sejam capturados pelas partículas magnéticas e conseqüentemente que as seqüências complementares aos oligonucleotídeos também sejam capturadas. O PCR do enriquecimento mostrou fragmentos com tamanhos entre 200 e 700 pares de base para os enriquecimentos L1 e A1 com o pool A e entre 200 e 900 pares de base para os enriquecimentos de L2 e A2 com o pool C (Figura 3.9).

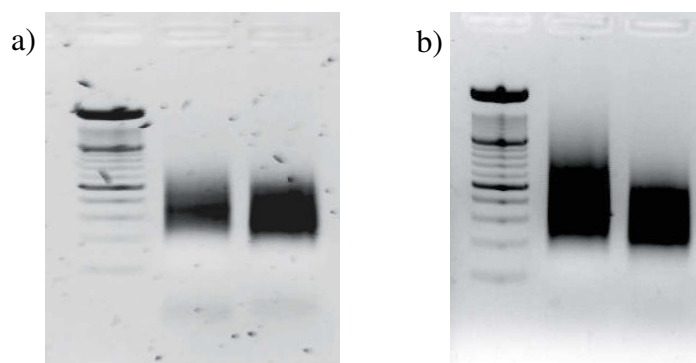


Figura 3.9 Gel de agarose 1,5% com os produtos de PCR do enriquecimento com: a) pool A de trinucleotídeos das amostras L1 e A1, respectivamente; e b) pool C de dinucleotídeos das amostras L2 e A2, respectivamente. Marcador de peso molecular: XIV.

Após as transformações, todas as colônias brancas foram repicadas para duas placas idênticas com espaço para 145 clones cada uma. Uma destas placas foi usada na transferência para as membranas de hibridização com os oligonucleotídeos. E a outra placa ficou guardada

com as réplicas dos clones para serem usadas no seqüenciamento dos plasmídeos. Montamos quatro bibliotecas de DNA: oito placas com clones da amostra L1-PoolA, quatro placas da amostra L2-PoolC, seis placas da amostra A1-PoolA e mais oito placas da amostra A2-PoolC, totalizando 26 placas com aproximadamente 145 clones cada. As hibridizações das membranas com os oligonucleotídeos marcados radioativamente para identificar os clones positivos foram repetidas duas vezes, mas não tiveram um bom resultado. Por isso, a seleção dos clones positivos foi baseada em duas ampliações via PCR: a primeira feita com os *primers* M13 sense e anti-sense para checar o tamanho do inserto nos clones; e a segunda feita com o produto do primeiro PCR, usando um *primer* M13 sense ou anti-sense e os mesmos oligonucleotídeos do enriquecimento a fim de amplificar apenas clones com microsatélites dentro da seqüência inserida no plasmídeo. Só foram usados no segundo PCR clones com insertos maiores do que 450 pares de base. Para o primeiro PCR (M13-M13), foram selecionados aleatoriamente 95 clones do Pool A e 129 clones do Pool C da biblioteca da espécie *L. racemosa* (Figura 3.10), e mais 107 clones do Pool A e 211 clones do Pool C da biblioteca de *A. schaueriana* (Figura 3.11). Dentre os clones com produtos de PCR maiores que 450 pares de base, foram feitas 129 ampliações via PCR com *primers* M13 e oligonucleotídeos (Figura 3.12). Conseguimos seqüenciar 100 clones, dos quais 48 possuíam verdadeiros microsatélites em suas seqüências, ou seja, 47,5% dos clones seqüenciados. As porcentagens de clones positivos das duas espécies foram bastante parecidas de 43,6% e 50,8% para *A. schaueriana* e *L. racemosa* respectivamente. Através do programa *Primer 3* foram desenhados 51 pares de *primers*, pois alguns clones possuíam mais de um

microsatélite na mesma seqüência. Destes *primers*, doze pares foram testados com as amostras de *A. schaueriana*, *A. germinans* e *L. racemosa*.

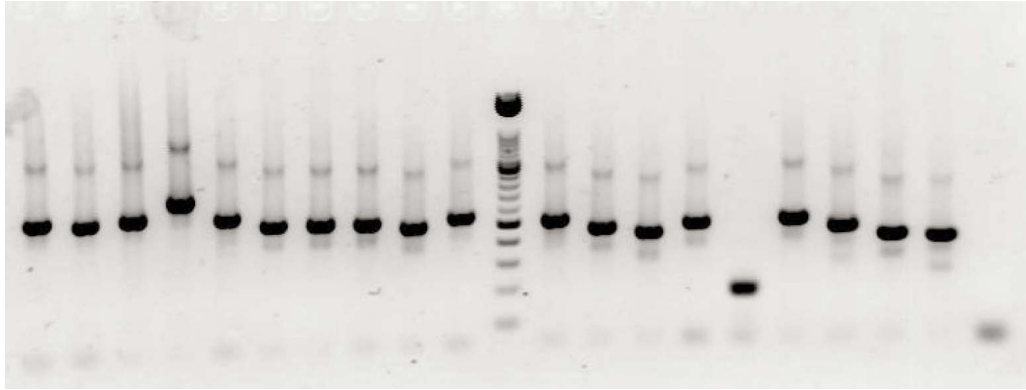


Figura 3.10 Gel de agarose 1,2% com os produtos de PCR (M13-M13) dos clones de *L. racemosa* selecionados aleatoriamente da biblioteca genômica enriquecida com microsatélites. Marcador de peso molecular: XIV (bandas mais fortes de 500 e 1000 pares de base).

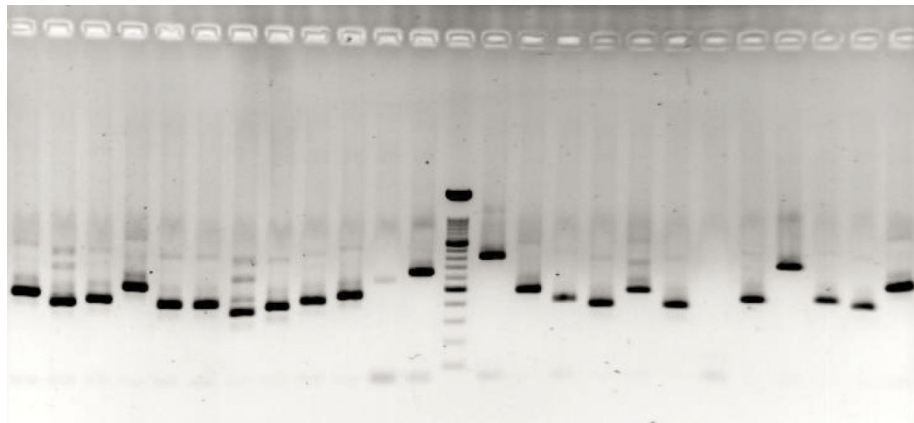


Figura 3.11 Gel de agarose 1,2% com os produtos de PCR (M13-M13) dos clones de *A. schaueriana* selecionados aleatoriamente da biblioteca genômica enriquecida com microsatélites. Marcador de peso molecular: XIV (bandas mais fortes de 500 e 1000 pares de base).

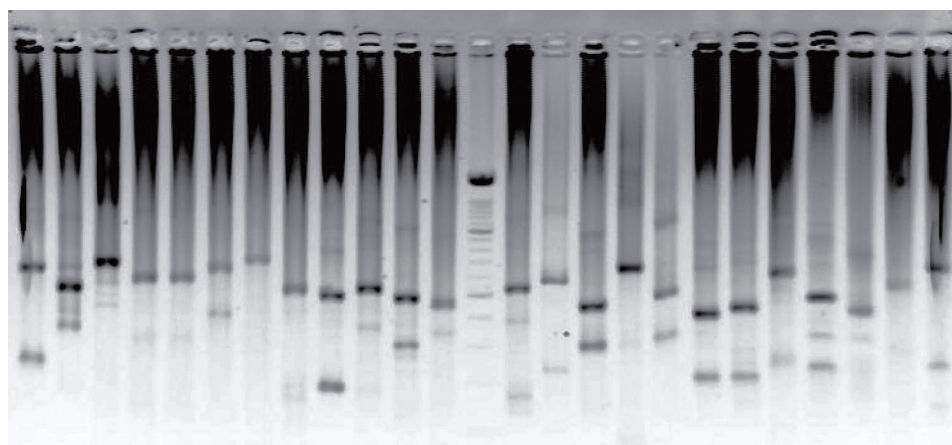


Figura 3.12 Gel de agarose 1,2% com os produtos de PCR (M13-oligos) dos clones de *A. schaueriana* e *L. racemosa* selecionados baseado no primeiro PCR (M13-M13). Marcador de peso molecular: XIV.

Os primeiros resultados foram mais satisfatórios para a espécie *L. racemosa* do que *A. schaueriana* (Tabela 3.9). Todos os seis pares de *primers* de *L. racemosa* (LA01, LA14, LC01, LC02, LC44 e LC55) obtiveram produtos de amplificação de tamanhos próximos do esperado (Figura 3.13). Com relação aos *primers* desenhados para *A. schaueriana*, três deles (AC56, AC128 e AC177) amplificaram produtos de tamanho esperado e estão prontos para a genotipagem (Figura 3.14). O *primer* AC56 também teve resultado positivo com amostras da espécie *A. germinans* (Figura 3.15). Com relação aos outros três *primers*, um deles (AC82) não gerou produto de amplificação, o outro (AC108) amplificou duas bandas fortes no gel com tamanhos de 200 e 300 pares de base, e o último *primer* (AC101) também não teve boa amplificação.

Tabela 3.9 Dados dos doze pares de *primers* desenvolvidos para SSR com o motivo, tamanho esperado e amplificação boa (+) ou ruim (-) nas espécies *A. schaueriana* (AS), *A. germinans* (AG) e *L. racemosa* (LR). NT – Não testado.

<i>Primer</i>	Motivo	Tamanho esperado	AS	AG	LR
LA01	(CAT) ₈	205pb	NT	NT	+
LA14	(GTT) ₆	272pb	NT	NT	+
LC01	(CT) ₁₂ (GT) ₆	296pb	NT	NT	+
LC02	(AT) ₄ (AC) ₂ (AT) ₄	107pb	NT	NT	+
LC44	(AC) ₅ (ATGC) ₂ (AC) ₇	180pb	NT	NT	+
LC55	(GT) ₇	295pb	NT	NT	+
AC56	(GT) ₈	191pb	+	+	NT
AC82	(GT) ₉	276pb	-	-	NT
AC101	(GT) ₉	183pb	-	-	NT
AC108	(GT) ₁₁	207pb	-	-	NT
AC128	(AC) ₁₁	264pb	+	-	NT
AC177	(GT) ₇	236pb	+	-	NT

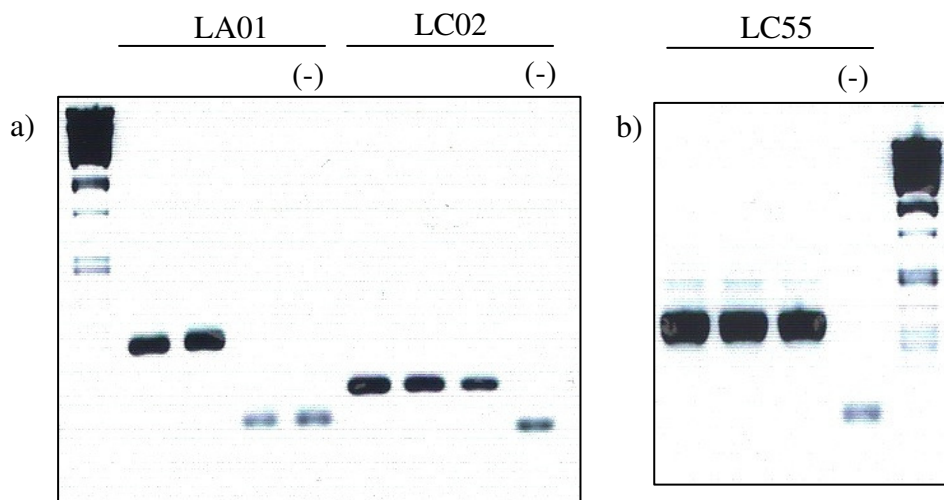


Figura 3.13 Géis de agarose 2% com os produtos da amplificação com os *primers* LA01 (a), LC02 (a) e LC55 (b) de três amostras de *L. racemosa* e controle negativo com água (-). Marcadores de peso molecular: DNA λ digerido com PstI.

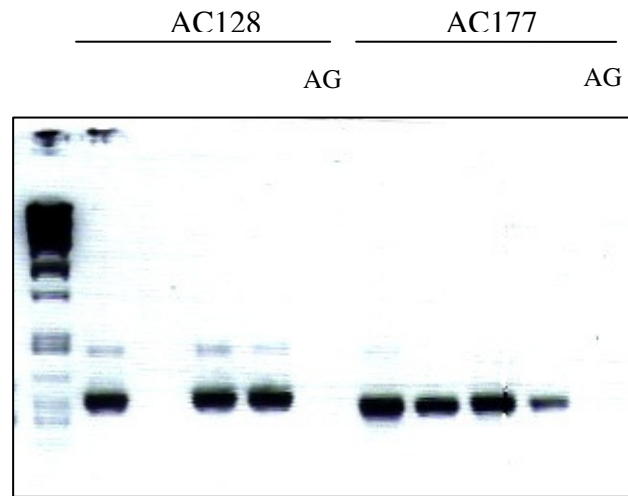


Figura 3.14 Gel de agarose 2% com os produtos da amplificação dos *primers* AC128 e AC177 com quatro amostras de *A. schaueriana* e uma amostra de *A. germinans* (AG), que não teve amplificação. Marcador de peso molecular: DNA λ digerido com PstI.



Figura 3.15 Gel de agarose 2% com produtos da amplificação do *primer* AC56 de duas amostras de *A. schaueriana* (AS) e uma de *A. germinans* (AG). Marcador de peso molecular: DNA λ digerido com PstI.

3.5 Discussão

Apesar do evidente crescimento do uso de microssatélites em estudos populacionais conservacionistas, um dos limitantes desta técnica é justamente o isolamento de marcadores específicos a partir de bibliotecas de DNA enriquecidas com microssatélites (Squirrel *et al.*, 2003). Com o surgimento de bancos de dados de EST, que são seqüências transcritas do genoma, e com o constante aumento no número de espécies com seqüências depositadas nestes bancos, é possível obter microssatélites do tipo SSR-EST usando programas desenvolvidos com o objetivo de facilitar o isolamento destes marcadores (Varshney *et al.*, 2005).

Atualmente, a maioria das publicações utilizando microssatélites do tipo SSR-EST visa o mapeamento genético e o estudo da diversidade funcional de espécies cultivadas, sendo os microssatélites aleatórios os mais utilizados em estudos de *fingerprinting*²⁶ e identificação parental (Varshney *et al.*, 2005). No entanto, atualmente os SSR-EST representam uma ferramenta adicional para geneticistas populacionais ou ecologistas moleculares, devido a sua alta transferabilidade entre espécies, e porque alguns trabalhos mostraram que estes marcadores são muito mais estáveis e fáceis de amplificar do que os microssatélites nucleares aleatórios (Woodhead *et al.*, 2005; Ellis *et al.*, 2006).

Um limitante dos microssatélites gênicos seria justamente a possível seleção que este marcador sofreria, mas Woodhead *et al.* (2005) provaram que a diferenciação genética não estava sendo afetada pela seleção, pois os índices de diferenciação genética, F_{ST} , calculados

²⁶ Método que compara similaridades e diferenças genéticas entre indivíduos baseado no polimorfismo das seqüências de DNA únicas para cada indivíduo.

com resultados dos marcadores SSR-EST, SSR e AFLP foram similares. Aparentemente, portanto, os microssatélites derivados de EST parecem ser geralmente neutros podendo ser usados em estudos de efeito demográfico sobre o nível de variação genética, visando a conservação de espécies (Ellis *et al.*, 2006).

Na busca por marcadores microssatélites do tipo SSR-EST usando seqüências da espécie *A. marina*, encontramos microssatélites em 4% do total de seqüências de EST analisadas e desenhamos 15 pares de *primers* para regiões com alta homologia a proteínas. O rendimento foi bastante baixo quando comparado ao das bibliotecas de DNA enriquecidas, porém tivemos as grandes vantagens de não ter tido nenhum custo laboratorial durante seu isolamento e ainda a alta transferabilidade entre espécies correlatas, que possibilitou a amplificação destes *primers* em duas espécies de *Avicennia* encontradas no Brasil: *A. schaueriana* e *A. germinans*.

A partir das bibliotecas de DNA enriquecidas com microssatélites, a porcentagem de clones positivos foi de 47,5% no total para as duas espécies com relação aos clones seqüenciados, porém o trabalho laboratorial foi bastante dispendioso de tempo e dinheiro. Além disso, algumas das etapas não ocorreram da forma esperada por motivos técnicos como a etapa de seleção do tamanho dos fragmentos obtidos com a digestão e de hibridização na busca por clones positivos durante este trabalho, sendo esta mais uma desvantagem desta técnica. Por outro lado, a abordagem usada para definir os clones positivos através dos PCRs M13-M13 e M13-oligo foi bastante eficiente, pois a porcentagem de clones seqüenciados que possuíam microssatélites para análise genética de *L. racemosa* foi 50,8%, sendo maior do que a encontrada por Nettel *et al.* (2008) com a mesma espécie (apenas 28%). Para outras espécies

de mangue, o rendimento deste tipo de isolamento de microssatélites foi bem maior com 84% do total, onde dez de 26 pares de *primers* foram polimórficos em *Rhizophora mangle* (Rosero-Galindo *et al.*, 2002); e de 63,5% a 100% para *A. germinans*, seis pares de *primers* do total de 13 tiveram alto polimorfismo (Ceron-Souza *et al.*, 2006).

Apesar das grandes discussões sobre o uso de microssatélites de regiões gênicas (SSR-EST) ou microssatélites nucleares aleatórios em estudos populacionais, estas duas vertentes possuem vantagens e desvantagens importantes a serem levadas em consideração quando é necessário desenvolver este tipo de marcador molecular. Os microssatélites gênicos são desenvolvidos mais facilmente sem custo inicial, porém partindo do pressuposto da existência de um banco de EST para a espécie que se deseja pesquisar ou alguma espécie correlata, o que limita muito o trabalho com espécies selvagens. Por outro lado, os microssatélites nucleares aleatórios podem ser desenvolvidos para qualquer espécie com a qual se deseja trabalhar, porém os custos iniciais de isolamento e caracterização dos marcadores são bastante altos além de ser necessário ter bom conhecimento sobre as técnicas em laboratório. Outra desvantagem desta técnica é que geralmente não ocorre transferabilidade entre espécies, mesmo aquelas bastante próximas taxonomicamente, o que torna necessário isolar microssatélites para cada espécie em estudo.

Como perspectivas, serão feitas as genotipagens dos marcadores SSR-EST em *Avicennia* e SSR nuclear aleatório em *A. schaueriana*, *A. germinans* e *L. racemosa*, a fim de fazer um estudo do fluxo gênico e estruturação genética populacional entre manguezais do Brasil. E assim comparar a biologia reprodutiva de uma espécie com distribuição restrita como *A. schaueriana* com outra espécie de maior distribuição como *L. racemosa*, mas que

ocorrem no mesmo habitat. Além de comparar também a neutralidade dos dois tipos de microssatélites isolados, o que é importante para que novos estudos populacionais, usando técnicas mais simples de isolamento de microssatélites, possam ser feitos.

Conclusão geral

Os resultados obtidos com o programa METAPOPOP, apresentados no primeiro capítulo da tese, provaram ser possível através de ferramentas computacionais simular a evolução de uma espécie tendo em foco a diferenciação genética entre suas populações. Usando a espécie *Caesalpinia echinata* Lam. (Pau-Brasil) como modelo, foi possível elucidar um pouco mais sobre sua biologia reprodutiva, além de identificar a contribuição de algumas características ecológicas, como ciclo de vida, taxa de autofecundação e taxa de dispersão, sobre a diferenciação genética entre populações.

Os resultados obtidos demonstraram que os marcadores citoplasmáticos foram fortemente afetados pelas diferentes taxas de dispersão que simularam a fragmentação do habitat, sendo mais diferenciados com o aumento da fragmentação. O ciclo de vida afetou de forma menos evidente sendo menor a diferenciação quanto maior a longevidade da planta. Por último, a taxa de autofecundação não causou nenhum efeito sobre a diferenciação de marcadores citoplasmáticos. Estes marcadores são passados de forma uniparental entre as plantas o que os torna mais susceptíveis às diferentes taxas de dispersão do que às outras características estudadas. Por outro lado, os marcadores nucleares foram largamente afetados por todas as três características, tendo 99,9% de sua variação explicada pela taxa de dispersão, 95,8% explicada pela longevidade e 88,9% explicada pelo sistema de cruzamento nas análises de regressão.

Com todos estes resultados foi possível inferir aspectos da biologia reprodutiva do Pau-Brasil em função dos estudos genéticos feitos anteriormente usando marcadores AFLP (Cardoso *et al.*, 2005) e microssatélites de cloroplasto (Lira *et al.*, 2003) com sete populações naturais desta espécie ao longo da costa brasileira. A comparação dos índices de G_{ST} obtidos

experimentalmente com aqueles obtidos nas simulações sugere que esta espécie seja preferencialmente alógama com taxas de dispersão moderadas, dados pouco esperados devido aos limitados mecanismos de dispersão de pólen e semente desta espécie. No caso do Pau-Brasil, suas populações estão espalhadas, porém formando pequenos grupos ou agregados, o que poderia estar contribuindo para o sistema de cruzamento alógamo e moderadas taxas de dispersão como encontrados nas simulações. Quanto à longevidade, não foi possível sugerir de forma conclusiva sobre a duração do ciclo de vida do Pau-Brasil. A simulação mais próxima do G_{ST} experimental foi aquela de uma espécie que vive até 300 anos. Porém não houve diferença significativa entre os resultados de 120 e 300 anos de vida, o que gera dúvidas quanto a esse valores.

Por último, foram feitas simulações que objetivaram fazer uma projeção da evolução das populações naturais de Pau-Brasil nos próximos 20.000 anos, levando em consideração que estas populações estariam hoje completamente isoladas devido à alta fragmentação de seu habitat, ou seja, nenhum fluxo gênico ocorreria entre elas. Estes resultados foram bastante surpreendentes, pois mostraram que em 10.000 anos os marcadores citoplasmáticos estariam totalmente diferenciados alcançando $G_{ST} = 0,95$. Ao mesmo tempo, os marcadores nucleares sofreriam uma diferenciação igualmente dramática saltando de 0,63 para 0,92 após 20.000 anos sem fluxo gênico. Estes resultados mostram claramente a importância da manutenção do cruzamento entre indivíduos de populações diferentes contrabalanceando os efeitos da fragmentação do habitat, a fim de manter a integridade genética da espécie.

No segundo capítulo foram usadas duas metodologias para analisar duas populações de manguezais no município do Rio de Janeiro. Estas populações apresentam uma característica especial dos mangues que é a presença do apicum, uma região com altos níveis de salinidade no solo. As plantas que ocorrem próximas ao apicum são bastante diferentes

morfologicamente daquelas que ocorrem próximas ao rio e por isso a estrutura genética de *A. schaueriana* e *L. racemosa*, com marcadores AFLP, assim como a estrutura epigenética *versus* genética desta última foram investigadas usando também marcadores MSAP.

Os marcadores AFLP mostraram que a espécie *A. schaueriana* possui uma estrutura populacional de 8,6% a 19% entre suas populações, dependendo do método estatístico abordado, enquanto que as populações de *L. racemosa* mostraram pouca estrutura genética que variou de 0 a 3,2%. A população próxima ao apicum está mais susceptível a processos seletivos naturais devido ao estresse ambiental causado pela alta salinidade, o que diminuiria a diversidade genética e aumentaria a estrutura e diferenciação genética entre suas populações. Estes resultados esperados ocorreram com as populações de *A. schaueriana*, porém o inverso ocorreu com as populações de *L. racemosa* que mostrou índices de diversidade genética maiores na população do apicum e nenhuma estrutura genética significativa entre as duas populações que justificasse a variação morfológica encontrada entre elas.

Isso nos levou a utilizar marcadores moleculares MSAP aprofundando nosso estudo sobre *L. racemosa* nas populações de Guaratiba. Estes marcadores possibilitaram uma análise comparativa dos padrões de metilação do DNA – uma alteração epigenética que pode ser responsável pela expressão ou silenciamento de genes, além do silenciamento de elementos transponíveis e rearranjo da cromatina – através do uso de duas enzimas de restrição que clivam o mesmo sítio, porém possuem diferentes sensibilidades à metilação da citosina. Os resultados obtidos com a enzima *MspI*, sensível apenas à metilação na citosina externa de uma das fitas de DNA, ou seja, à hemimetilação, que é um evento raro na natureza, foram considerados como estrutura genética. Por outro lado, os resultados obtidos pela enzima *HpaII*, sensível à metilação das citosinas internas de ambas as fitas, foram considerados como

estrutura epigenética. A comparação das duas estruturas permitiu uma análise mais robusta da espécie *L. racemosa*, englobando as alterações genéticas e epigenéticas do genoma.

O resultado do MSAP foi surpreendente, pois mostrou que existe uma diferenciação epigenética (*HpaII*-MSAP) de 16% a 25,4%, sendo estes valores maiores do que a diferenciação genética (*MspI*-MSAP) de 11% a 15% entre as populações de *L. racemosa*. Ainda mais interessante foi o resultado da análise de componente principal (PCA) que mostrou de forma espacial a estrutura genética e epigenética mas também a união entre as duas pela análise de Co-Inércia. De forma evidente, a variação epigenética se divide em dois grupos distintos formados por indivíduos de cada população, enquanto que a variação genética não consegue atingir tal definição. A análise de Co-Inércia mostrou ainda que indivíduos das duas populações geneticamente similares possuíam padrão epigenético bastante distinto dependente da população de ocorrência destes indivíduos, provavelmente afetado pelo próprio ambiente onde vivem.

Este trabalho demonstrou, por conseguinte, que a variação epigenética pode estar sofrendo pressões evolutivas e seletivas divergentes dependentes do ambiente. Este é um importante passo para um entendimento maior sobre a seleção de genes expressos ou silenciados durante o desenvolvimento vegetal, mostrando que o controle epigenético via metilação de citosinas pode ser afetado por alterações ambientais sendo passados de geração para geração. Isso pode inclusive explicar os maiores índices de diversidade genética na população do apicum de *L. racemosa* obtidos com os marcadores AFLP, que seriam resultado de uma baixa pressão seletiva, pois esta estaria especialmente ativa sobre o componente epigenético do genoma desta espécie. Este é provavelmente o principal responsável pela sobrevivência desta espécie num ambiente com alto estresse abiótico, por afetar diretamente a expressão de genes envolvidos no desenvolvimento vegetal.

O terceiro capítulo abordou o isolamento de marcadores microssatélites nucleares aleatórios e de regiões gênicas (SSR-EST) através da criação de banco de DNA enriquecidos com microssatélites para ambas as espécies de mangue, *L. racemosa* e *A. schaueriana*, e do uso do banco de EST de *A. marina*, respectivamente. Uma revisão da literatura mostrou que os marcadores RAPD, AFLP e microssatélites são os mais utilizados em estudos genéticos que visam a conservação de espécies, porém os microssatélites vêm sendo usados de forma crescente porque são marcadores codominantes e, por conseguinte, possuem uma contribuição muito maior em artigos científicos ao longo dos anos recentes.

O isolamento dos SSR-EST de *A. marina* foi feita de forma rápida usando um programa chamado MISA e os testes com os *primers* isolados foram bastante positivos, sendo possível amplificar fragmentos das duas espécies de *Avicennia* ocorrentes no Brasil: *A. schaueriana* e *A. germinans*. Por outro lado, a criação das bibliotecas de DNA mostrou-se demorada e custosa, além de problemas técnicos que impediram certas etapas do protocolo que tiveram que ser suprimidas ou substituídas por outros métodos. Por fim, foram obtidos muitos clones com microssatélites, que foram seqüenciados. Dos *primers* testados para *L. racemosa*, todos amplificaram bandas do tamanho esperado. Já para *A. schaueriana*, metade deles amplificaram bandas do tamanho esperado e apenas um destes amplificou banda do mesmo tamanho em amostras de *A. germinans*, evidenciando a baixa transferabilidade entre espécies deste tipo de marcador quando contrastado com os marcadores SSR-EST. Futuramente, serão feitas as genotipagens com todos os marcadores isolados por ambos os métodos a fim de determinar aqueles que possuem polimorfismo e que serão utilizados em análises genéticas populacionais de mangue. Outro objetivo futuro é comparar os resultados obtidos com os dois tipos de microssatélites para identificar aqueles que são realmente neutros.

Esta tese, portanto, mostrou três abordagens diferentes com relação ao estudo genético populacional. A primeira usou métodos computacionais para simular populações sofrendo efeitos evolutivos naturais como deriva gênica e mutação a fim de compreender o papel de características ecológicas na diferenciação genética entre as populações. A segunda abordagem usou dois marcadores moleculares parecidos, o AFLP e o MSAP, para investigar as espécies de mangue, *L. racemosa* e *A. schaueriana*, a fim de identificar a contribuição genética e epigenética na diferenciação morfológica de indivíduos de duas populações. Por fim, foi feito o isolamento de marcadores microssatélites nucleares através de dois métodos, a criação de bibliotecas genômicas e o uso de bancos de EST, a fim de relacionar as vantagens e desvantagens de cada metodologia e seus resultados interespecíficos.

Referências bibliográficas

- AKKAYA, M.S.; BHAGWAT, A.A.; CREGAN, P.B. Length polymorphisms of simple sequence repeat DNA in Soybean. **Genetics**, v.132, n.4, p.1131-1139, 1992.
- ALBACH, D.C.; SCHÖNSWETTER, P.; TRIBSCH, A. Comparative phylogeography of the *Veronica alpina* complex in Europe and North America. **Molecular Ecology**, v.15, n.11, p.3269-3286, 2006.
- ALONGI, D.M. Mangrove forests: Resilience, protection from tsunamis, and responses to global climate change. **Estuarine, Coastal and Shelf Science**, v.76, p.1-13, 2008.
- AUSTERLITZ, F.; MARIETTE, S.; MACHON, N.; GOUYON, P.-H.; GODELLE, B. Effects of colonization processes on genetic diversity: differences between annual plants and tree species. **Genetics**, v.154, p.1309-1321, 2000.
- AZHIKINA, T.L. & SVERDLOV, E.D. Study of Tissue-Specific CpG Methylation of DNA in Extended Genomic *Loci*. **Biochemistry (Moscow)**, v.70, n.5, p.596-603, 2005.
- BALL, M.C. Interactive effects of salinity and irradiance on growth: implications for mangrove forest structure along salinity gradients. **Trees**, v.16, p.126-139, 2002.
- BALLOUX, F.; LUGON-MOULIN, N. The estimation of population differentiation with microsatellite markers. **Molecular Ecology**, v.11, n.2, p.155-165, 2002.
- BARTH, O.M.; SÃO-THIAGO, L.E.U.; BARROS, M.A. Paleoenvironment interpretation of a 1760 years B.P. old sediment in a mangrove area of the Bay of Guanabara, using pollen analysis. **Anais da Academia Brasileira de Ciências (Annals of the Brazilian Academy of Sciences)**, v.78, p.227-229, 2006.
- BENDER, J. Plant epigenetics. **Current Biology**, v.12, p.R412-4, 2000.
- BENSCH, S. & ÅKESSON, M. Ten years of AFLP in ecology and evolution: why so few animals? **Molecular Ecology**, v.14, p.2899-2914, 2005.
- BENSON, G. Tandem repeats finder: a program to analyze DNA sequences. **Nucleic Acids Research**, v.27, n.2, p.573-580, 1999 *apud* VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, v.23, n.1, p.48-55, 2005.
- BORGES, L.A.; SOBRINHO, M.S.; LOPES, A.V. Phenology, pollination, and breeding system of the threatened tree *Caesalpinia echinata* Lam. (Fabaceae), and a review of studies on the reproductive biology in the genus. **Flora** (in press), 2008.
- BOSSDORF, O.; RICHARDS, C.L.; PIGLIUCCI, M. Epigenetics for ecologists. **Ecology Letters**, v.11, p.106-115, 2008.

- BOSSUYT, B. Genetic rescue in an isolated metapopulation of a naturally fragmented plant species, *Parnassia palustris*. **Conservation Biology**, v.21, p.832-841, 2007.
- BUSSELL, J.D. The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). **Molecular Ecology**, v.8, p.775-789, 1999.
- BYARS, S.G.; PAPST, W.; HOFFMANN, A.A. Local adaptation and cogradient selection in the alpine plant, *Poa hiemata*, along a narrow altitudinal gradient. **Evolution**, v.61, n.12, p.2925-2941, 2007.
- CARDOSO, M.A.; PROVAN, J.; POWELL, W.; FERREIRA, P.C.G.; DE OLIVEIRA, D.E. High genetic differentiation among remnant populations of the endangered *Caesalpinia echinata* Lam. (Leguminosae Caesalpinioideae). **Molecular Ecology**, v.7, n.5, p.601-608, 1998.
- CARDOSO, S.R.S.; ELOY, N.B.; PROVAN, J.; CARDOSO, M.A.; FERREIRA, P.C.G. Genetic differentiation of *Euterpe edulis* Mart. populations estimated by AFLP analysis. **Molecular Ecology**, v.9, n.11, p.1753-1760, 2000.
- CARDOSO, S.R.S.; PROVAN, J.; LIRA, C.F.; PEREIRA, L.O.R.; FERREIRA, P.C.G.; CARDOSO, M.A. High levels of genetic sub-structuring as a result of population fragmentation in the tropical tree species *Caesalpinia echinata* Lam. **Biodiversity and Conservation**, v.14, p.1047-1057, 2005.
- CARVALHO, P.E.R. **Espécies florestais brasileiras. Recomendações silviculturais, potencialidades e uso da madeira**. EMBRAPA-CNPQ/SP1, Brasília, 1994.
- CASTELO, A.T.; MARTINS, W.; GAO, G.R. TROLL – Tandem Repeat Occurrence Locator. **Bioinformatics**, v.18, n.4, p.634-636, 2002 *apud* VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, v.23, n.1, p.48-55, 2005.
- CASTIÑEIRAS, L.; GUZMÁN, F.A.; DUQUE, M.C.; SHAGARODSKY, T.; CRISTÓBAL, R.; DE VICENTE, M.C. AFLPs and morphological diversity of *Phaseolus lunatus* L. in Cuban home gardens: approaches to recovering the lost *ex situ* collection. **Biodiversity and Conservation**, v.16, p.2847-2865, 2007.
- CERÓN-SOUZA, I.; RIVERA-OCASIO, E.; FUNK, S. M.; MCMILLAN, W.O. Development of six microsatellite *loci* for black mangrove (*Avicennia germinans*). **Molecular Ecology Notes**, v.6, p.692-694, 2006.
- CERVERA, M.T.; RUIZ-GARCÍA, L.; MARTÍNEZ-ZAPATER, J.M. Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. **Molecular Genetics and Genomics**, v.268, n.4, p.543-552, 2002.
- CHASE, M.R.; MOLLER, C.; KESSELI, R.; BAWA, K.S. Distant gene flow in tropical trees. **Nature**, v.383, n.6599, p.398-399, 1996.

- CHEN, C.; ZHOU, P.; CHOI, Y.A.; HUANG, S.; GMITTER, F.G. Mining and characterizing microsatellites from citrus ESTs. **Theoretical and Applied Genetics**, v.112, n.7, p.1248-1257, 2006.
- CHISTIAKOV, D.A.; HELLEMANS, B.; VOLCKAERT, F.A.M. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. **Aquaculture**, v.255, n.1-4, p.1-29, 2006.
- CHOI, C.-S.; SANO, H. Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. **Molecular Genetics and Genomics**, v.277, p.589–600, 2007.
- CLARK, C.M.; WENTWORTH, T.R.; O`MALLEY, D.M. Genetic discontinuity revealed by chloroplast microsatellites in eastern north american *Abies* (Pinaceae). **American Journal of Botany**, v.87, n.6, p.774-782, 2000.
- COLLEVATTI, R.G.; GRATTAPAGLIA, D.; HAY, J.D. High resolution microsatellite based analysis of the mating system allows the detection of significant biparental inbreeding in *Caryocar brasiliense*, an endangered tropical tree species. **Heredity**, v.86, n.1, p.60-67, 2001.
- COLLEVATTI, R.G.; GRATTAPAGLIA, D.; HAY, J.D. Evidences for multiple maternal lineages of *Caryocar brasiliense* populations in the Brazilian Cerrado based on the analysis of chloroplast DNA sequences and microsatellite haplotype variation. **Molecular Ecology**, v.12, p.105-115, 2003.
- CONDIT, R.; HUBBELL, S.P. Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. **Genome**, v.34, n.1, p.66-71, 1991.
- COUVET, D. Deleterious effects of restricted gene flow in fragmented populations. **Conservation Biology**, v.16, p.369-376, 2002.
- DESPRES, L.; LORIOT, S.; GAUDEUL, M. Geographic pattern of genetic variation in the European globeflower *Trollius europaeus* L. (Ranunculaceae) inferred from amplified fragment length polymorphism markers. **Molecular Ecology**, v.11, p.2337-2347, 2002.
- DEYOUNG, R.W.; HONEYCUTT, R.L. The molecular toolbox: genetic techniques in wildlife ecology and management. **Journal of Wildlife Management**, v.69, n.4, p.1362-1384, 2005.
- DUTECH, C.; JOLY, H.I.; JARNE, P. Gene flow, historical population dynamics and genetic diversity within French Guianan populations of a rainforest tree species, *Vouacapoua americana*. **Heredity**, v.92, p.69-77, 2004.
- DUTECH, C.; MAGGIA, L.; TARDY, C.; JOLY, H.I.; JARNE, P. Tracking a genetic signal of extinction-recolonization events in a neotropical tree species: *Vouacapoua americana* Aublet in French Guiana. **Evolution**, v.57, p.2753-2764, 2003.

- ECKSTEIN, R.L.; O'NEILL, R.A.; DANIHELKA, J.; OTTE, A.; KÖHLER, W. Genetic structure among and within peripheral and central populations of three endangered floodplain violets. **Molecular Ecology**, v.15, n.9, p.2367-2379, 2006.
- EDWARDS, K.J.; BARKER, J.H.A.; DALY, A.; JONES, C.; KARP, A. Microsatellite libraries enriched for several microsatellite sequences in plants. **Biotechniques**, v.20, n.5, p.758-760, 1996.
- ELLIS, J.R.; PASHLEY, C.H.; BURKE, J.M.; MCCAULEY, D.E. High genetic diversity in a rare and endangered sunflower as compared to a common congener. **Molecular Ecology**, v.15, n.9, p.2345-2355, 2006.
- ESCUADERO, A.; IRIONDO, J.M.; TORRES, M.E. Spatial analysis of genetic diversity as a tool for plant conservation. **Biological Conservation**, v.113, p.351-365, 2003.
- EUJAYL, I.; SORRELLS, M.E.; BAUM, M.; WOLTERS, P.; POWELL, W. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. **Theoretical and Applied Genetics**, v.104, n.2-3, p.399-407, 2002.
- FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao Uso de Marcadores Moleculares em Análise Genética**. 3^a ed. Brasília: EMBRAPA-CENARGEN, 1998.
- FINNEGAN, E.J.; PEACOCK, W.J.; DENNIS, E.S. DNA methylation, a key regulator of plant development and other processes. **Current Opinion in Genetics & Development**, v.10, p.217-223, 2000.
- FOLL, M.; BEAUMONT, M.A. ; GAGGIOTTI, O. An approximate bayesian computation approach to overcome biases that arise when using amplified fragment length polymorphism markers to study population structure. **Genetics**, v.179, p:927-939, 2008.
- FRANKHAM, R. Conservation genetics. **Annual Review of Genetics**, v.29, p.305-327, 1995.
- GAO, L.; TANG, J.; LI, H.; JIA, J. Analysis of microsatellites in major crops assessed by computational and experimental approaches. **Molecular Breeding**, v.12, n.3, p.245-261, 2003 *apud* VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, v.23, n.1, p.48-55, 2005.
- GAO, L.F.; JING, R.L.; HUO, N.X.; LI, Y.; LI, X.P.; ZHOU, R.H.; CHANG, X.P.; TANG, J.F.; MA, Z.Y.; JIA, J.Z. One hundred and one new microsatellite *loci* derived from ESTs (EST-SSRs) in bread wheat. **Theoretical and Applied Genetics**, v.108, n.7, p.1392-1400, 2004.
- GEHRING, M.; & HENIKOFF, S. DNA methylation dynamics in plant genomes. **Biochimica et Biophysica Acta**, v.1769, p.276-286, 2007.

- GONZALES, E.; HAMRICK, J.L.; SMOUSE, P.E.; DYER, R.J. Pollen-mediated gene dispersal within continuous and fragmented populations of a forest understorey species, *Trillium cuneatum*. **Molecular Ecology**, v.15, p.2047-2058, 2006.
- HABU, Y.; KAKUTANI, T.; PASZKOWSKI, J. Epigenetic developmental mechanisms in plants: molecules and targets of plant epigenetic regulation. **Current Opinion in Genetics & Development**, v.11, p.215–220, 2001.
- HAMRICK, J.L. Response of forest trees to global environmental changes. **Forest Ecology and Management**, v.197, p.323-335, 2004.
- HAMRICK, J.L.; GODT, M.J.W. Allozyme diversity in plant species, in: Brown, A.H.D., Clegg, M.T., Kahler, A.L., Weir, B.S. (Eds.), **Plant Population Genetics, Breeding and Genetic Resources**. Sinauer Associates Inc., Sunderland, MA, pp. 43-63, 1990.
- HAMRICK, J.L.; GODT, M.J.W.; SHERMAN-BROYLES, S.L. Factors influencing levels of genetic diversity in woody plant species. **New Forests**, v.6, p.95-124, 1992.
- HAO, Y.J.; WEN, X.P.; DENG, X.X. Genetic and epigenetic evaluations of citrus calluses recovered from slow-growth culture. **Journal of Plant Physiology**, v.161, n.4, p.479-484, 2004.
- HEUERTZ, M.; HAUSMAN, J.F.; TSVETKOV, I.; FRASCARIA-LACOSTE, N.; VEKEMANS, X. Assessment of genetic structure within and among Bulgarian populations of the common ash (*Fraxinus excelsior* L.). **Molecular Ecology**, v.10, n.7, p.1615-1623, 2001.
- HOLSINGER, K.E.; LEWIS, P.O.; DEY, D.K. A Bayesian approach to inferring population structure from dominant markers. **Molecular Ecology**, v.11, p.1157–1164, 2002.
- HONNAY, O.; JACQUEMYN, H. Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. **Conservation Biology**, v.21, p.823-831, 2007.
- HUETTEL, B.; KANNO, T.; DAXINGER, L.; AUFSATZ, W.; MATZKE, A.J.; MATZKE, M. Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. **The EMBO Journal**, v.25, p.2826-2836, 2006.
- HUH, J.H.; BAUER, M.J.; HSIEH, T.F.; FISCHER, R.L. Cellular programming of plant gene imprinting. **Cell**, v.132, p.735-745, 2008.
- JABLONKA, E. & LAMB, M.J. Epigenetic inheritance in evolution. **Journal of Evolutionary Biology**, v.11, p.159-183, 1998.
- JALIGOT, E.; BEULÉ, T.; BAURENS, F.C.; BILLOTTE, N.; RIVAL, A. Search for methylation-sensitive amplification polymorphisms associated with the mantled variant phenotype in oil palm (*Elaeis guineensis* Jacq). **Genome**, v.47, n.1, p.224-8, 2004.

- KALISZ, S. & PURUGGANAN, M.D. Epialleles via DNA methylation: consequences for plant evolution. **TRENDS in Ecology and Evolution**, v.19, p.309-314, 2004.
- KANTETY, R.V.; LA ROTA, M.; MATTHEWS, D.E.; SORRELLS, M.E. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. **Plant Molecular Biology**, v.48, n.5-6, p.501-510, 2002 *apud* VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, v.23, n.1, p.48-55, 2005.
- KARP, A.; SEBERG, O.; BUIATTI, M. Molecular techniques in the assessment of botanical diversity. **Annals of Botany**, v.78, n.2, p.143-149, 1996.
- KEYTE, A.L.; PERCIFIELD, R.; LIU, B.; WENDEL, J.F. Intraspecific DNA Methylation Polymorphism in Cotton (*Gossypium hirsutum* L.). **Journal of Heredity**, v.97, n.5, p.444-450, 2006.
- KIMURA, M.; WEISS, G.H. The stepping stone model of population structure and the decrease of genetic correlation with distance. **Genetics**, v.49, p.561-576, 1964.
- KIRST, M.; CORDEIRO, C.M.; REZENDE, G.D.S.P.; GRATTAPAGLIA, D. Power of microsatellite markers for fingerprinting and parentage analysis in *Eucalyptus grandis* breeding populations. *Journal of Heredity*, V.96, n.2, p.161-166, 2005.
- LAVIN, M.; SCHRIRE, B.P.; LEWIS, G.; PENNINGTON, R.T.; DELGADO-SALINAS, A.; THULIN, M.; HUGHES, C.E.; MATOS, A.B.; WOJCIECHOWSKI, M.F. Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. **Philosophical Transactions of the Royal Society of London. Series B; Biological sciences**, v.359, p.1509-1522, 2004.
- LE CORRE, V.; MACHON, N.; PETIT, R.J.; KREMER, A. Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. **Genetical Research**, v.69, p.117-125, 1997.
- LE ROUZIC, A. & CARLBORG, Ö. Evolutionary potential of hidden genetic variation. **Trends in Ecology & Evolution**, v.23, p.33-37, 2008.
- LEE, H.-S. & CHEN, Z.J. Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. **Proceedings of the National Academy of Sciences USA**, v.98, p.6753-6758, 2001.
- LIRA, C.F. **Análise genética de populações naturais de Pau-Brasil utilizando marcadores AFLP e microssatélites de cloroplasto**. Rio de Janeiro, 84f. Dissertação (mestrado em Química Biológica). Instituto de Ciências Biomédicas, Universidade Federal do Rio de Janeiro, 2003.
- LIRA, C.F.; CARDOSO, S.R.S.; FERREIRA, P.C.G.; CARDOSO, M.A.; PROVAN, J. Long-term population isolation in the endangered tropical tree species *Caesalpinia*

- echinata* Lam. revealed by chloroplast microsatellite. **Molecular Ecology**, v.12, n.12, p.3219-3225, 2003.
- LIU, Z.J.; & CORDES, J.F. DNA marker technologies and their applications in aquaculture genetics. **Aquaculture**, v.238, n.1-4, p.1-37, 2004.
- LOVELESS, M.D. Isozyme variation in tropical trees: patterns of genetic organization. **New Forests**, v.6, p.67-94, 1992.
- LUKENS, L.N. & ZHAN, S. The plant genome's methylation status and response to stress: implications for plant improvement. **Current Opinion in Plant Biology**, v.10, p.317-322, 2007.
- MACHON, N.; BARADAT, D.; GODELLE, B. **METAPOPOP a program simulating evolutionary processes acting on metapopulations**. Paris, France: Laboratoire ES, Paris-Sud, 1995.
- MACHON, N.; BARDIN, P.; MAZER, S.J.; MORET, J.; GODELLE, B.; AUSTERLITZ, F. Relationship between genetic structure and seed and pollen dispersal in the endangered orchid *Spiranthes spiralis*. **New Phytologist**, v.157, p.677-687, 2003.
- MADLUNG, A. & COMAI, L. The Effect of Stress on Genome Regulation and Structure. **Annals of Botany**, v.94, p.481-495, 2004.
- MÄKINEN, H.S.; CANO, J.M.; MERILÄ, J. Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. **Molecular Ecology**, v.17, n.15, p.3565-3582, 2008.
- MARIETTE, S.; LE CORRE, V.; KREMER, A. **Sampling within the genome for measuring within-population diversity: Trade-offs between markers** *apud* GILLET, E.M. (Eds.). **Sampling Strategies for Marker Analysis**. Compendium of the Research Project Development, optimization and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity, 1999.
- MAUDET, C.; MILLER, C.; BASSANO, B.; BREITENMOSER-WÜRSTEN, C.; GAUTHIER, D.; OBEXER-RUFF, G.; MICHALLET, J.; TABERLET, P.; LUIKART, G. Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [*Capra ibex (ibex)*]. **Molecular Ecology**, v.11, n.3, p.421-436, 2002.
- MENGONI, A.; BARABESI, C.; GONNELLI, C.; GALARDI, F.; GABBRIELLI, R.; BAZZICALUPO, M. Genetic diversity of heavy metal-tolerant populations in *Silene paradoxa* L. (Caryophyllaceae): a chloroplast microsatellite analysis. **Molecular Ecology**, v.10, n.8, p.1909-1916, 2001.

- MORAND, M.E.; BRACHET, S.; ROSSIGNOL, P.; DUFOUR, J.; FRASCARIA-LACOSTE, N. A generalized heterozygote deficiency assessed with microsatellites in French common ash populations. **Molecular Ecology**, v.11, n.3, p.377-385, 2002.
- MORAND-PRIEUR, M.E.; VEDEL, F.; RAQUIN, C.; BRACHET, S.; SIHACHAKR, D.; FRASCARIA-LACOSTE, N. Maternal inheritance of a chloroplast microsatellite marker in controlled hybrids between *Fraxinus excelsior* and *Fraxinus angustifolia*. **Molecular Ecology**, v.11, n.3, p.613-617, 2002.
- MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially present with non-repetitive DNA in plant genomes. **Nature Genetics**, v.30, p.194-200, 2002 *apud* VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, v.23, n.1, p.48-55, 2005.
- MULLIS, K.B.; FALOONA, F.A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. **Methods in Enzymology**, v.155, Part F, p.335-350, 1987.
- MYERS, N.; MITTERMEIER, R.A.; MITTERMEIER, C.G.; DA FONSECA, G.A.B.; KENT, J. Biodiversity hotspots for conservation priorities. **Nature**, v.403, p.853-858, 2000.
- NAVASCUÉS, M. & EMERSON, B.C. Chloroplast microsatellites: measures of genetic diversity and the effect of homoplasmy. **Molecular Ecology**, v.14, p.1333-1341, 2005.
- NEI, M. Analysis of gene diversity in subdivided populations. **Proceedings of the National Academy of Sciences USA**, v.70, p.3321-3323, 1973.
- NEI, M. Estimation of average heterozygosity and genetic distances from a small number of individual. **Genetics**, v.89, p.583-590, 1978.
- NETO, J.D.G.; SEBBENN, A.M.; KAGEYAMA, P.Y. Sistema de reprodução em *Caesalpinia echinata* Lam. implantada em arboreto experimental. **Revista Brasileira de Botânica**, v.28, p.409-418, 2005.
- NETTEL, A.; DODD, R.S.; CID-BECERRA, J.A.; DE LA ROSA-VELEZ, J. Development of microsatellite markers for the white mangrove (*Laguncularia racemosa* C.F. Gaertn., Combretaceae). **Conservation Genetics**, v.9, n.4, p.1037-1038, 2008.
- NOVICK, R.R.; DICK, C.W.; LEMES, M.R.; NAVARRO, C.; CACCONE, A.; BIRMINGHAM, E. Genetic structure of Mesoamerican populations of big-leaf mahogany (*Swietenia macrophylla*) inferred from microsatellite analysis. **Molecular Ecology**, v.12, p.2885-2893, 2003.
- OLIVEIRA, E.J.; PÁDUA, J.G.; ZUCCHI, M.I.; VENCOVSKY, R.; VIEIRA, M.L.C. Origin, evolution and genome distribution of microsatellites. **Genetics and Molecular Biology**, v.29, n.2, p.294-307, 2006.

- OSTRANDER, E.A.; JONG, P.M.; RINE, J.; DUYK, G. Construction of small-insert genomics DNA libraries highly enriched for microsatellite repeat sequences. **Proceedings of the National Academy of Science USA**, v.89, n.8, p.3419-3423, 1992.
- OUBORG, N.J.; PIQUOT, Y.; VAN GROENENDAEL, J.M. Population genetics, molecular markers and the study of dispersal in plants. **Journal of Ecology**, v.87, p.551-568, 1999.
- PARIDA, A.K. & DAS, A.B. Effects of NaCl stress on nitrogen and phosphorous metabolism in a true mangrove *Bruguiera parviflora* grown under hydroponic culture. **Journal of Plant Physiology**, v.161, p.921-928, 2004.
- PARIDA, A.K.; DAS, A.B.; MITTRA, B. Effects of salt on growth, ion accumulation, photosynthesis and leaf anatomy of the mangrove, *Bruguiera parviflora*. **Trees**, v.18, p.167-174, 2004.
- PARISOD, C. & CHRISTIN, P.-A. Genome-wide association to fine-scale ecological heterogeneity within a continuous population of *Biscutella laevigata* (Brassicaceae). **New Phytologist**, v.178, p.436-447, 2008.
- PARISOD, C.; TRIPPI, C.; GALLAND, N. Genetic variability and founder effect in the pitcher plant *Sarracenia purpurea* (Sarraceniaceae) in populations introduced into Switzerland: from inbreeding to invasion. **Annals of Botany**, v.95, p.277-286, 2005
- PASHLEY, C.H.; ELLIS, J.R.; MCCAULEY, D.E.; BURKE, J.M. EST databases as a source for molecular markers: Lessons from *Helianthus*. **Journal of Heredity**, v.97, n.4, p.381-388, 2006.
- PAVET, V.; QUINTERO, C.; CECCHINI, N.M.; ROSA, A.L.; ALVAREZ, M.E. *Arabidopsis* displays centromeric DNA hypomethylation and cytological alterations of heterochromatin upon attack by *Pseudomonas syringae*. **Molecular Plant-Microbe Interactions**, v.19, n.6, p.577-587, 2006.
- PEARSE, D.E.; CRANDALL, K.A. Beyond F_{ST} : Analysis of population genetic data for conservation. **Conservation Genetics**, v.5, n.5, p.585-602, 2004.
- PENNINGTON, R.T.; LAVIN, M.; PRADO, D.E.; PENDRY, C.A.; PELL, S.K.; BUTTERWORTH, C.A. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. **Philosophical Transactions of the Royal Society of London. Series B; Biological sciences** 359, 515-537, 2004.
- PENNINGTON, R.T.; PRADO, D.E.; PENDRY, C.A. Neotropical seasonally dry forests and Quaternary vegetation changes. **Journal of Biogeography**, v.27, p.261-273, 2000.

- PEREZA-ECHEVARRIA, S.; HERRERA-VALENCIA, V.A.; KAY, A. Detection of DNA methylation changes in micropropagated banana plants using methylation-sensitive amplification polymorphism (MSAP). **Plant Science**, v.161, n.2, p.359-367, 2001.
- PIGLIUCCI, M. Do we need an extended evolutionary synthesis? **Evolution**, v.61, n.12, p.2743–2749, 2007.
- PONCET, V.; RONDEAU, M.; TRANCHANT, C.; CAYREL, A.; HAMON, S.; DE KOCHKO, A.; HAMON, P. SSR mining in coffee tree EST databases: potential use of EST-SSRs as markers for the *Coffea* genus. **Molecular Genetics and Genomics**, v.276, n.5, p.436-449, 2006.
- POWELL, W.; MORGANTE, M.; ANDRE, C.; MCNICOL, J.W.; MACHRA, Y.; DOYLE, J.J.; TINGEY, S.V.; RAFALSKI, J.A. Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. **Current Biology**, v.5, n.9, p.1023-1029, 1995.
- PROVAN, J. Novel chloroplast microsatellites reveal cytoplasmic variation in *Arabidopsis thaliana*. **Molecular Ecology**, v.9, n.12, p.2183-2185, 2000.
- PROVAN, J.; POWELL, W.; HOLLINGSWORTH, P.M. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. **Trends in Ecology and Evolution**, v.16, n.3, p.142-147, 2001.
- PROVAN, J.; SORANZO, N.; WILSON, N.J.; GOLDSTEIN, D.B.; POWELL, W. A low mutation rate for chloroplast microsatellites. **Genetics**, v.153, p.943-947, 1999.
- RAABOVÁ, J.; MÜNZZBERGOVÁ, Z.; FISCHER, M. Ecological rather than geographic or genetic distance affects local adaptation of the rare perennial herb, *Aster amellus*. **Biological Conservation**, v.139, p.348-357, 2007.
- RAJORA, O.P.; & MOSSELER, A. Challenges and opportunities for conservation of forest genetic resources. **Euphytica**, v.118, p.197-212, 2001.
- RAMP, J.M.; COLLINGE, S.K.; RANKER, T.A. Restoration genetics of the vernal pool endemic *Lasthenia conjugens* (Asteraceae). **Conservation Genetics**, v.7, p.631-649, 2006.
- RAPP, R.A. & WENDEL, J.F. Epigenetics and plant evolution. **New Phytologist**, v.168, p.81–91, 2005.
- RENAU-MORATA, B.; NEBAUER, S.G.; SALES, E.; ALLAINGUILLAUME, J.; CALIGARI, P.; SEGURA, J. Genetic diversity and structure of natural and managed populations of *Cedrus atlantica* (Pinaceae) assessed using random amplified polymorphism DNA. **American Journal of Botany**, v.92, n.5, p.875-884, 2005.
- RICE, W. R. Analyzing tables of statistical tests. **Evolution**, v.43, p.223-225, 1989.

- ROSERO-GALINDO, C.; GAITAN-SOLIS, E.; CÁRDENAS-HENAO, H.; TOHME, J.; TORO-PEREA, N. Polymorphic microsatellites in a mangrove species, *Rhizophora mangle* L. (Rhizophoraceae). **Molecular Ecology Notes**, v.2, p.281–283, 2002.
- ROZEN, S. & SKALETSKY, H.J. **Primer3 on the WWW for general users and for biologist programmers**, v.132, Part 3, p.365-386, 1999. *apud* KRAWETZ, S.; MISENER, S. (eds) **Bioinformatics Methods and Protocols: Methods in Molecular Biology**. Humana Press, Totowa, NJ, 500pp, 1999.
- RUNGIS, D.; BÉRUBÉ, Y.; ZHANG, J.; RALPH, S.; RITLAND, C.E.; ELLIS, B.E.; DOUGLAS, C.; BOHLMANN, J.; RITLAND, K. Robust simple sequence repeat markers for spruce (*Picea* spp.) from expressed sequence tags. **Theoretical and Applied Genetics**, v.109, n.6, p.1283-1294, 2004 *apud* VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, v.23, n.1, p.48-55, 2005.
- SAATCHI, S.; AGOSTI, D.; ALGER, K.; DELABIE, J.; MUSINSKY, J. Examining fragmentation and loss of primary forest in the southern Bahian Atlantic forest of Brazil with radar imagery. **Conservation Biology**, v.15, p.867-875, 2001.
- SAENGER, P. **Mangrove Ecology, Silviculture and Conservation**. Kluwer Academic Publishers, The Netherlands, 2003.
- SALMON, A.; AINOUCHE, M.L.; WENDEL, J.F. Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). **Molecular Ecology**, v.14, p.1163–1175, 2005.
- SALMON, A.; CLOTAULT, J.; JENCZEWSKI, E.; CHABLE, V.; MANZANARES-DAULEUX, M.J. *Brassica oleracea* displays a high level of DNA methylation polymorphism. **Plant Science**, v.174, p.61–70, 2008.
- SCHAEFFER-NOVELLI, Y.; CINTRÓN-MOLERO, G.; ADAIME, R.R.; CAMARGO, T.M. Variability of mangrove ecosystems along the Brazilian coast. **Estuaries and Coasts**, v.13, p.204-218, 1990.
- SCHONES, D.E. & ZHAO, K. Genome-wide approaches to studying chromatin modifications. **Nature Review Genetics**, v.9, p.179-191, 2008.
- SELKOE, K.A.; & TOONEN, R.J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. **Ecology Letters**, v.9, n.5, p.615-629, 2006.
- SHEMSKE, D.W.; HUSBAND, B.C.; RUCKELSHAUS, M.H.; GOODWILLIE, C.; PARKER, I.M.; BISHOP, J.G. Evaluating Approaches to the Conservation of Rare and Endangered Plants. **Ecology**, v.75, n.3, p.584-606, 1994.
- SHI, S.; HUANGA, Y.; ZENGA, K.; TANA, F.; HEA, H.; HUANGA, J.; FU, Y. Molecular phylogenetic analysis of mangroves: independent evolutionary origins of vivipary and salt secretion. **Molecular Phylogenetics and Evolution**, v.34, p.159–166, 2005.

- SOS Mata Atlântica, INPE. Atlas dos remanescentes florestais da Mata Atlântica, período 1995–2000. **Fundação SOS Mata Atlântica & INPE**, São Paulo, 2002.
- SQUIRRELL, J.; HOLLINGSWORTH, P.M.; WOODHEAD, M.; RUSSELL, J.; LOWE, A.J.; GIBBY, M.; POWELL, W. How much effort is required to isolate nuclear microsatellites from plants? **Molecular Ecology**, v.12, n.6, p.1339-1348, 2003.
- TAKATA, M.; KISHIMA, Y.; SANO, Y. DNA methylation polymorphisms in rice and wild rice strains: detection of epigenetic markers. **Breeding Science**, v.55, p.57-63, 2005.
- TAUTZ, D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. **Nucleic Acids Research**, v.17, n.16, p.6463-6471, 1989.
- THIEL, T.; MICHALEK, W.; VARSHNEY, R.K.; GRANER, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). **Theoretical and Applied Genetics**, v.106, n.3, p.411-422, 2003.
- THIOLOUSE, J.; CHESSEL, D.; DOLÉDEC, S.; OLIVIER, J.M. ADE-4: a multivariate analysis and graphical display software. **Statistics and Computing**, v.7, p.75-83, 1996.
- TOMLINSON, P.B. **The Botany of Mangroves**. Cambridge University Press, 413pp, 1986.
- TOPPING, C.; OSTERGAARD, S.; PERTOLDI, C.; BACH, L.A. Modelling the loss of genetic diversity in vole populations in a spatially and temporally varying environment. **Annales zoologici Fennici**, v.40, p.255-267, 2003.
- VAN EE, B.W.; JELINSKI, N.; BERRY, P.E.; HIPPI, A.L. Phylogeny and biogeography of *Croton alabamensis* (Euphorbiaceae), a rare shrub from Texas and Alabama, using DNA sequence and AFLP data. **Molecular Ecology**, v.15, n.9, p.2735-2751, 2006.
- VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, v.23, n.1, p.48-55, 2005.
- VARSHNEY, R.K.; SIGMUND, R.; BORNER, A.; KORZUN, V.; STEIN, N.; SORRELLS, M.E.; LANGRIDGE, P.; GRANER, A. Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. **Plant Science**, v.168, n.1, p.195-202, 2005b.
- VENDRAMIN, G.G.; ANZIDEI, M.; MADAGHIELE, A.; SPERISEN, C.; BUCCI, G. Chloroplast microsatellite analysis reveals the presence of population subdivision in Norway spruce (*Picea abies* K.). **Genome**, v.43, n.1, p.68-78, 2000.
- VENDRAMIN, G.G.; DEGEN, B.; PETIT, R.J.; ANZIDEI, M.; MADAGHIELE, A.; ZIEGENHAGEN, B. High level of variation at *Abies alba* chloroplast microsatellite loci in Europe. **Molecular Ecology**, v.8, n.7, p.1117-1126, 1999.

- VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRIJTERS, A.; POT, J.; PELEMAN, J.; KUIPER, M.; *et al.* AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, v.23, n.21, p.4407-4414, 1995.
- WANG, Z.; WEBER, J.L.; ZHONG, G.; TANKSLEY, S.D. Survey of plant short tandem DNA repeats. **Theoretical and Applied Genetics**, v.88, n.1, p.1-6, 1994.
- WARD, M.; DICK, C.W.; GRIBEL, R.; LOWE, A.J. To self, or not to self... A review of outcrossing and pollen-mediated gene flow in neotropical trees. **Heredity**, v.95, p.246-254, 2005.
- WEBER, J.L. Human DNA polymorphisms and methods of analysis. **Current Opinion in Biotechnology**, v.1, n.2, p.166-171, 1990.
- WEISING, K. & GARDNER, R.C. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. **Genome**, v.42, n.1, p.9-19, 1999.
- WHITE, G. & POWELL, W. Isolation and characterization of microsatellite *loci* in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species. **Molecular Ecology**, v.6, n.9, p.851-860, 1997.
- WHITE, G.M.; BOSHIER, D.H.; POWELL, W. Genetic variation within a fragmented population of *Swietenia humilis* Zucc. **Molecular Ecology**, v.8, n.11, p.1899-1909, 1999.
- WHITE, G.M.; BOSHIER, D.H.; POWELL, W. Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. **Proceedings of the National Academy of Sciences USA**, v.99, n.4, p.2038-2042, 2002.
- WHITE, T.J.; ARNHEIM, N.; ERLICH, H.A. The polymerase chain reaction. **Trends in Genetics**, v.5, n.6, p.185-189, 1989.
- WILLIAMS, J.G.; KUBELIK, A.R.; LIVAK, K.J.; RAFALSKI, J.A.; TINGEY, S.V. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research**, v.18, n.22, p.6531-6535, 1990.
- WOODHEAD, M.; RUSSELL, J.; SQUIRRELL, J.; HOLLINGSWORTH, P.M.; MACKENZIE, K.; GIBBY, M.; POWELL, W. Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. **Molecular Ecology**, v.14, n.6, p.1681-1695, 2005.
- XIE, H.; SUI, Y.; CHANG, F.Q.; XU, Y.; MA, R.C. SSR allelic variation in almond (*Prunus dulcis* Mill.). **Theoretical and Applied Genetics**, v.112, n.2, p.366-372, 2006.

- XIONG, L.Z.; XU, C.G.; SAGHAI-MAROOF, M.A.; ZHANG, Q. Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. **Molecular Genetics and Genomics**, v.261, n.3, p.439-446, 1999.
- YAO, X.; YE, Q.; KANG, M.; HUANG, H. Microsatellite analysis reveals interpopulation differentiation and gene flow in the endangered tree *Changiostyrax dolichocarpa* (Styracaceae) with fragmented distribution in central China. **New Phytologist**, v.176, p.472-480, 2007.
- YEH, F.C.; YANG, R.C.; BOYLE, T. **POPGENE, Version 1.21: Software Microsoft Windowbased Freeware for Population Genetic Analysis**. University of Alberta, Edmonton, Alberta, Canada, 1997.
- YOUNG, A.; BOYLE, T.; BROWN, T. The population genetic consequences of habitat fragmentation for plants. **Trends in Ecology and Evolution**, v.11, p.413-418, 1996.
- ZANE, L.; BARGELLONI, L.; PATARNELLO, T. Strategies for microsatellite isolation: a review. **Molecular Ecology**, v.11, n.1, p.1-16, 2002.
- ZHANG, X. The Epigenetic Landscape of Plants. **Science**, v.320, p.489-492, 2008.
- ZHANG, X.; YAZAKI, J.; SUNDARESAN, A.; COKUS, S.; CHAN, S.W.; CHEN, H.; HENDERSON, I.R.; SHINN, P.; PELLEGRINI, M.; JACOBSEN, S.E.; ECKER, J.R. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. **Cell**, v.126, p.1189-1201, 2006.
- ZHAO, X.; CHAI, Y.; LIU, B. Epigenetic inheritance and variation of DNA methylation level and pattern in maize intra-specific hybrids. **Plant Science**, v.172, p.930-938, 2007.

Glossário

Alelo nulo – Alelo que não amplifica durante o PCR por problema de anelamento do primer.

Apicum – Planícies hipersalinas que são áreas de inundação de maré com vegetação arbustiva adaptada para solos salinos.

Deme – Uma população local panmítica, ou seja, que faz cruzamentos aleatórios entre seus indivíduos.

Deriva gênica – Flutuações estocásticas na frequência alélica dentro de uma população, mais evidente em populações pequenas.

Epistático - Fenômeno em que vários pares de genes interagem entre si para influenciar uma única característica fenotípica.

EST – Um segmento de cDNA com sequência idêntica à parte da região transcrita de um gene.

e-value – Um parâmetro que descreve o número de sequências homólogas que seriam esperadas de ocorrer ao acaso durante a busca em um banco de dados de sequências.

Fingerprinting – Método que compara similaridades e diferenças genéticas entre indivíduos baseado no polimorfismo das sequências de DNA únicas para cada indivíduo.

Fitness – A medida da habilidade do indivíduo de uma espécie em sobreviver e se reproduzir.

F_{IS} – Índice que representa a taxa de endogamia nas populações de uma dada espécie.

F_{ST} – Índice de diferenciação populacional baseado em polimorfismo genético.

G_{ST} – Índice de diferenciação populacional baseado em polimorfismo genético equivalente ao F_{ST} .

Halófitas – Plantas adaptadas quanto a morfologia e fisiologia para crescer em solos ricos em sal.

Homoplasia – Similaridade genética entre duas espécies que não é resultante da herança de um ancestral comum.

Hotspot – Áreas consideradas prioritárias para conservação, que possuem altos índices de biodiversidade.

Kruskal-Wallis rank sum test - Teste estatístico de significância que compara cada elemento ao invés de usar a média.

Locus (ou *loci*) – Local ou posição constante no cromossomo, onde localizam-se determinado gene ou marcador genético em todos os indivíduos de uma mesma espécie.

Metapopulação - Consiste num grupo de populações da mesma espécie separadas espacialmente e que podem interagir de algum modo.

Oligonucleotídeo ou oligo – Molécula curta, geralmente de 6 a 100 nucleotídeos, de DNA fita simples.

Pleiotrópico - Quando uma variação devida a um só gene não se traduz num, mas em vários efeitos fenotípicos, diz-se que o gene tem um efeito *pleiotrópico*.

Primer – Oligonucleotídeo de DNA usado como iniciador numa reação de PCR.

Propágulo – Nome dado às sementes que se desprendem da planta-mãe já germinadas, que é resultado da viviparidade nas espécies de mangue.

Repetições em tandem – Sequências repetidas lado a lado no genoma.

Stepping Stone – Modelo teórico de estrutura populacional e fluxo gênico onde a espécie é subdividida em populações de tamanhos idênticos que trocam alelos apenas com as demes adjacentes.

Touchdown – Programa de PCR que reduz o anelamento de *primers* não-específicos pela diminuição gradual de temperatura de anelamento entre os ciclos.

Viviparidade – Fenômeno durante o qual o embrião dentro da semente se desenvolve ainda na planta-mãe, sendo chamado de propágulos no caso das espécies de mangue.

Anexo I

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

Manuscript submitted to Biological Conservation:

“How life-cycle, mating system and fragmentation processes can influence nuclear and cytoplasmic genetic differentiation? A modeling study based on *Caesalpinia echinata* Lam. (Leguminosae)”

Author’s information:

1) Catarina DA FONSECA LIRA-MEDEIROS^{a27}

^aLaboratoire de Conservation des Espèces, Restauration et Suivi des Populations – UMR 5173 CNRS – Muséum National d’Histoire Naturelle – 61, rue Buffon – 75005 Paris, France. catilira@bioqmed.ufrj.br/catarina_lira@yahoo.com.br

2) Frédéric AUSTERLITZ^b

^bCNRS, Laboratoire Ecologie Systématique et Evolution, UMR 8079, Orsay, F-91405; Univ Paris-Sud, Orsay, F-91405; AgroParisTech, Paris, F-75231. Frederic.Austerlitz@u-psud.fr

3) Paulo CAVALCANTI GOMES FERREIRA^{cd}

^cLaboratório de Biologia Molecular de Plantas – Instituto de Pesquisas do Jardim Botânico do Rio de Janeiro – Rua Pacheco Leão, 915 – 22460-030 Rio de Janeiro, Brasil

²⁷ Present Address: Laboratório de Biologia Molecular de Plantas – Instituto de Pesquisas do Jardim Botânico do Rio de Janeiro – Rua Pacheco Leão, 915 – 22460-030 Rio de Janeiro, Brasil. catilira@bioqmed.ufrj.br

22 ^dInstituto de Bioquímica Médica – Universidade Federal do Rio de Janeiro – Avenida
23 Bauhinia, 400 – CCS, Bloco D(SS) sala 05 – Ilha do Fundão – 21491-590 Rio de
24 Janeiro, Brasil. paulof@bioqmed.ufrj.br

25 **4) Monica AIRES CARDOSO^c**

26 ^cLaboratório de Biologia Molecular de Plantas – Instituto de Pesquisas do Jardim
27 Botânico do Rio de Janeiro – Rua Pacheco Leão, 915 – 22460-030 Rio de Janeiro,
28 Brasil. mcardoso@jbrj.gov.br

29 **5) Nathalie MACHON^a**

30 ^aLaboratoire de Conservation des Espèces, Restauration et Suivi des Populations –
31 UMR 5173 CNRS – Muséum National d’Histoire Naturelle – 61, rue Buffon – 75005
32 Paris, France. machon@mnhn.fr

33

34 ***Corresponding author:***

35 Catarina DA FONSECA LIRA-MEDEIROS

36 *catilira@bioqmed.ufrj.br/catarina_lira@yahoo.com.br*

37 Address: Laboratório de Biologia Molecular de Plantas – Instituto de Pesquisas do
38 Jardim Botânico do Rio de Janeiro – Rua Pacheco Leão, 915 – 22460-030 Rio de
39 Janeiro – RJ – Brasil.

40 Telephone: 55 21 3204 2092

41 Fax: 55 21 3204 2071

42

43 **Title:** “How can plant life-cycle, mating system and fragmentation processes influence
44 nuclear and cytoplasmic genetic differentiation? A modeling study based on
45 *Caesalpinia echinata* Lam. (Leguminosae)”

46

47 **Abstract**

48 Using a simulation study we investigated the effects of mating system, life cycle and
49 population fragmentation during the last 20 000 years, since the last glaciation, on the
50 genetic differentiation of nuclear and cytoplasmic loci of plants. Our model was based
51 on *Caesalpinia echinata*, an endemic tree species of the Brazilian Atlantic Rainforest.
52 Ecological features such as population distribution and fertility success of *C. echinata*,
53 as well as their level of genetic differentiation for both nuclear and cytoplasmic markers
54 were used to create the simulation model. We found that mating system and
55 fragmentation (evaluated by different dispersal rates) have a strong impact on nuclear
56 genetic differentiation between populations. On the other hand, they had no or limited
57 impact on cytoplasmic differentiation. Moreover, longer life cycle counteracted genetic
58 differentiation for both nuclear and cytoplasmic markers. We obtained the closest
59 results to the empirical data with simulation parameters corresponding to a long-living
60 and allogamous species with moderately levels of pollen and seed flow. This provides
61 good insights about *C. echinata*'s reproductive biology, which is still unclear to some
62 extent. Finally, we investigated the impact of complete isolation of *C. echinata*
63 populations, with no gene flow between populations, induced by the currently
64 accelerated fragmentation of the Atlantic Rainforest habitat. The nuclear and
65 cytoplasmic genetic differentiation between populations increased rapidly under this

66 scenario. In order to avoid an extreme genetic differentiation between *C. echinata*'s
67 natural populations, we discuss potential management strategies.

68

69 **Keywords:** genetic differentiation, reproductive system, metapopulation simulation,
70 Atlantic Rainforest, forest management

71

72 **1. Introduction**

73

74 Habitat fragmentation may lead for many species to the formation of small and
75 isolated populations that are likely to undergo a decrease of genetic diversity within
76 populations and an increase of genetic differentiation among populations (Young et al.,
77 1996). The level of genetic diversity is, however, critically important for the species to
78 maintain their evolutionary potential, in order to cope with the changing environment,
79 and a loss of genetic diversity is often associated with reduced fitness and adaptation
80 capacity (Couvét, 2002; Yao et al., 2007). The assessment of the genetic and
81 reproductive status within populations of fragmented species is crucial to formulate
82 conservation plans for their genetic resources (Rajora and Mosseler, 2001). Genetic
83 diversity in fragmented forest species will depend on several features, in particular their
84 genetic diversity before fragmentation, the species geographical range, the level of gene
85 flow, as well as their breeding system and longevity (Hamrick and Godt, 1990;
86 Hamrick, 2004; Bossuyt, 2007; Honnay and Jacquemyn, 2007; Yao et al., 2007).

87 The rare species *Caesalpinia echinata* Lam. (Family: Leguminosae) is an endemic
88 species of the Brazilian Atlantic Rainforest, a very fragmented ecosystem that,
89 nevertheless, contains one of the highest level of biological diversity in the world
90 (Myers et al., 2000). *Caesalpinia echinata*, also known as brazil-wood, is an
91 economically and culturally important tree species. In the past 500 years it has been
92 extremely exploited for the production of a pigment dye, being substituted by artificial
93 dyes in the end of the 19th century (Carvalho, 1994). Furthermore, there are less than 7
94 percent of the natural Atlantic Rainforest original area remaining in Brazil, as
95 consequence of the fragmentation of its populations and destruction of its habitat (SOS

96 Mata Atlântica INPE, 2002). And these are the main threats for the maintenance of *C.*
97 *echinata*'s natural populations. To formulate and establish efficient conservation
98 strategies for this species it is necessary further investigation on its reproductive
99 biology, which remains still unclear.

100 Previous genetic studies on seven *C. echinata*'s natural populations distributed
101 along the Brazilian coast revealed divergent levels of genetic diversity, but high levels
102 of genetic differentiation for two dominant markers ($H_e = 0.636$ and 0.125 ; $F_{ST} = 0.58$
103 and 0.62 ; with RAPD and AFLP respectively; Cardoso et al., 1998; Cardoso et al.,
104 2005). They also showed a strong isolation by distance (IBD) signal, characterized by a
105 significant correlation between geographic distances and genetic differentiation
106 (Cardoso et al., 2005). The IBD signal was even stronger for chloroplast microsatellite
107 markers on the same populations, with very low genetic diversity ($H_e = 0.085$) and very
108 high genetic differentiation ($F_{ST} = 0.91$) (Lira et al., 2003). For these cytoplasmic
109 markers, seven of the eight observed haplotypes were population-specific and only two
110 populations were not fixed for a given haplotype (Lira et al., 2003).

111 This strong IBD between natural populations of *C. echinata* indicates a high genetic
112 structure, potentially resulting from an ancient fragmentation, possibly before last
113 glaciations event. The nuclear differentiation of *C. echinata* is much higher than what
114 has been observed for many tropical forest tree species, for which allozyme-based
115 studies showed an average F_{ST} value of ~ 0.1 (Hamrick et al, 1992; Loveless et al, 1992).
116 A similarly low level of differentiation has also been observed on forest tree species
117 with nuclear microsatellites (Novick et al, 2003; Dutech et al, 2004), even if these
118 values might be underestimates because of the high allelic diversity of microsatellites
119 (Hedrick 1999, 2005). This hypothesis that natural ancient fragmentation would have

120 provoked a high genetic differentiation today is also supported by the fact that *C.*
121 *echinata* habitat is defined as seasonally dry tropical forest on the Atlantic Rainforest
122 Complex (Pennington et al., 2004). These dry areas were probably widespread during
123 glacial events, but probably contracted during the inter-glacial periods to a few sites
124 where similar conditions to the drier paleoclimatic periods prevailed (Cardoso et al.,
125 1998).

126 Molecular techniques, together with modeling methodologies, provide promising
127 new tools for addressing complex environmental questions (Topping et al., 2003). The
128 objective of our study was to understand the effect of life cycle, mating system and
129 fragmentation process on genetic differentiation among *C. echinata* populations. With
130 this goal, we have used previous *C. echinata* genetic and ecological data (Cardoso et al.,
131 1998; Lira et al., 2003; Cardoso et al., 2005) and the METAPOP software that simulates
132 evolutionary processes both demographic and genetic acting on metapopulations (Le
133 Corre et al., 1997; Austerlitz et al., 2000; Machon et al., 2003). More specifically, we
134 used the simulated data to infer which values for some reproductive traits of *C. echinata*
135 such as selfing rate and life span were the most compatible with the empirical data. We
136 simulated the evolution of *C. echinata* populations since the end of the last glacial
137 period, assuming that after Last Glacial Maximum (LGM) *C. echinata* populations were
138 restricted to dry forest fragments of the Atlantic Rainforest. The tropical deglaciation
139 was calculated between 23 000 and 20 000 calendar years before present, preceding the
140 Northern Hemisphere warming by several thousands of years (Seltzer et al., 2002).
141 Finally, we simulated the evolution of the same populations from present to the next 20
142 000 years, assuming a complete isolation between populations, in order to assess the
143 impact of the ongoing anthropogenic fragmentation of *C. echinata* populations.

144

145 **2. Materials and Methods**

146

147 The demographic and genetic processes were simulated using the software
148 METAPOPOP (Machon et al., 1995; Le Corre et al., 1997), which assumes populations of
149 hermaphroditic plants characterized by their genotypes at neutral loci and at loci
150 submitted to natural selection, but we focused here on neutral loci. The simulated
151 populations exchange migrants via seeds and pollen dispersal, according to migration
152 matrices. To model demography, we used the density dependent demographic model
153 with overlapping generations developed by Austerlitz et al. (2000) (see below). We
154 assumed a metapopulation consisting of ten demes following a one-dimension stepping
155 stone model (Kimura and Weiss, 1964). The maximum biotic capacity of each deme
156 was set to 300 plant individuals, in accordance with the small natural populations of *C.*
157 *echinata* studied previously (T.S. Pereira, personal communication).

158 ***Demographic parameters***

159 In the model of Austerlitz et al. (2000), each individual belongs to a given size
160 class. The time step of the model corresponds to one year. From one year to the next,
161 each individual can either die, stay in the same class or move to the upper class. We set
162 here the probabilities to stay in the same class (including the last one) to zero. Thus, all
163 individuals either moved to the upper class or died, and all individuals of the last class
164 died the year after they reached this class. Therefore, the size classes were here strictly
165 equivalent to age classes and the maximum possible age for an individual was the total
166 number of classes assumed. The simulations started with 50 individuals in each
167 population with 10% being juveniles, i.e. their male and female fertilities were assumed

168 to be zero, randomly distributed among juvenile-age classes, and 90% of adults that
169 could participate to the reproduction both through seeds and pollen, also randomly
170 distributed along adult age classes. These populations would then grow until they
171 reached their equilibrium size of 300 individuals after few years.

172 The juveniles were considered smaller than adults, and they had a surviving
173 probability of only 50% to 70% every year in absence of density dependence. This
174 probability was set to 100% for the adult classes. However, because of the density-
175 dependence model (for details see Austerlitz et al., 2000), these survival probabilities
176 decreased as population sizes increased, so the populations would not grow above their
177 carrying capacity of 300 individuals. The values set for female fecundity are absolute in
178 Metapop. They correspond to the number of offspring that reach an age of one year
179 produced by each female in absence of density dependence. Here, we assumed that all
180 adults had a female fertility of 3.0, a value roughly consistent with natural populations
181 (T.S. Pereira, personal communication). Metapop assumes no pollen limitation,
182 therefore male fertilities are relative. Here, they were set to 1.0 for all adult classes,
183 meaning that all adult individuals had the same probability to participate in reproduction
184 via pollen.

185 *Genetic parameters*

186 METAPOP allows us to simulate two types of molecular markers, namely
187 biparental nuclear markers and maternally-inherited cytoplasmic markers, in order to
188 calculate the genetic differentiation between populations. For the nuclear loci, we
189 simulated 200 AFLP markers with two alleles per locus, simulating presence or absence
190 of band, assuming a mutation rate of 10^{-6} (Mariette et al., 1999). For simplicity, all
191 AFLP loci were assumed to be independent and in linkage equilibrium, although this

192 assumption may not be fully realistic because among 200 random AFLP loci some
193 could be not independent and/or in linkage disequilibrium. However, this assumption
194 cannot undermine our results, since we compute only G_{ST} mean values across loci (see
195 below), and so they are unaffected by a possible correlation between loci. We also
196 simulated ten haploid chloroplast microsatellite loci with 10 possible alleles for each
197 locus. We assumed a strictly maternal inheritance for these loci and a mutation rate of
198 10^{-4} (Provan et al., 1999; Navascués and Emerson, 2005). Initial allelic frequencies of
199 all loci were equally and randomly distributed within the populations, leading to G_{ST}
200 values of almost zero on the first generation, which means no genetic differentiation
201 between populations at the beginning of simulations. We thus assumed that brazil-wood
202 populations were not genetically differentiated during last glaciation (20 000 years ago),
203 when they possibly suffered expansions on dry patches of Atlantic Forest Complex.

204 *Simulations performed*

205 We performed four sets of simulations using METAPOP. In all cases we computed
206 the nuclear and cytoplasmic genetic differentiation among populations (G_{ST}). These
207 simulation sets aimed to investigate (1) the effect of breeding system; (2) the impact of
208 fragmentation intensity; (3) the impact of life span; and (4) the impact of recent
209 deforestation on the future evolution of genetic differentiation. For all simulations we
210 assumed initial G_{ST} values close to zero, i.e. allelic frequencies were randomly and
211 equally distributed between loci and populations, except for the fourth group of
212 simulations where genetic differentiation levels were similar to that empirically
213 calculated for brazil-wood natural populations. Simulations were repeated 20 times for
214 each scenario and nuclear and cytoplasmic G_{ST} values after 20 000 years of simulation
215 (first three scenarios) or every 5 000 years (fourth scenario) were recorded.

216 **1) Effect of mating system on the genetic differentiation among populations**

217 The breeding system was tested for five different levels of selfing rate (s),
218 corresponding to pure allogamy ($s = 0\%$), mixed mating systems ($s = 25\%$, 50% , or
219 75%) or pure autogamy ($s = 100\%$). Meanwhile, the other three ecological parameters
220 were set as default: life span of 300 years, pollen migration rate of 10^{-3} and seed
221 migration rate of 10^{-4} .

222 **2) Impact of fragmentation intensity on the genetic differentiation among**
223 **populations**

224 Seed and pollen dispersal was assumed to follow a linear stepping stone model
225 (Kimura and Weiss, 1964), to reflect the disposition of the natural populations of *C.*
226 *echinata*, which are located along the Brazilian Atlantic coast. Thus, only nearest
227 populations exchanged seeds and pollen in our simulations. The ratio of pollen flow to
228 seed flow was set to 10:1, consistent with the results obtained by Lira et al. (2003).

229 In the simulations, we assumed five different sets of migration rates $10^{-1}/10^{-2}$, $10^{-3}/10^{-4}$,
230 $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ and zero (for instance, the notation $10^{-1}/10^{-2}$ means a pollen
231 migration rate of 10^{-1} and a seed migration rate of 10^{-2} to the nearest populations on left
232 and right). Breeding system was set to 'no selfing' (allogamous) as default parameter
233 like suggested by previous studies (Giudice-Neto et al., 2005) and life span remained as
234 300 years.

235 **3) Impact of life cycle on the genetic differentiation among populations**

236 Accordingly to our age class model suggested by Austerlitz et al. (2000), we set the
237 total number of classes to 30, 60, 120, 300 or 500, simulating a corresponding life span
238 of 30, 60, 120, 300 and 500 years. For each life cycle, we assumed a specific number of
239 juvenile classes in order to be more realistic. In the 30-age class model, the first seven

240 classes were assumed to be juveniles, while the other classes were assumed to be adults.
241 In the 60-, 120-, 300-age classes models, the number of juvenile classes was set to 15,
242 while in the 500-age class model, it was set to 30. Such settings were assumed in order
243 to guarantee that the populations would not go extinct because of insufficient seed
244 production. The other ecological parameters were set as allogamous species, pollen
245 migration rate of 10^{-3} and seed migration rate of 10^{-4} .

246 ***4) Impact of recent deforestation, predictions for the future***

247 We studied the potential impact of recent deforestation on the future evolution of
248 genetic diversity in *C. echinata*. We considered that this drastic deforestation process
249 would cause a complete isolation of the remnant populations of this species, because of
250 the massive destruction of the Atlantic Rainforest (Saatchi et al., 2001). To predict the
251 consequences of such fragmentation, simulations were run with initial parameters as
252 close as possible from the experimental results obtained with RAPD (Cardoso et al.,
253 1998), AFLP (Cardoso et al., 2005) and chloroplast microsatellite (Lira et al., 2003)
254 markers. For this purpose, we used the allelic frequencies and demographic results
255 obtained after a 20 000-year simulation using the default parameters that were more
256 reliable to *C. echinata* species ($10^{-3}/10^{-4}$ dispersal rates, allogamy and 300-years life
257 span) as initial genetic and demographic parameters. This scenario was run again on
258 METAPOP for 20 000 years, using default selfing rate (0%) and life span (300 years),
259 but dispersal rates set as zero. We performed 20 replicates and for each of them we
260 recorded nuclear and cytoplasmic G_{ST} values every 5 000 years.

261 ***Statistical analyses***

262 The software R (R Development Core Team, 2004) was used to calculate
263 significance differences between G_{ST} values by the Tukey multiple comparisons of

264 means with 95% of confidence interval. Linear correlations were calculated using
265 factorial variables and Multiple R-Squares and P values were registered.

266

267 3. Results

268

269 1) *Effect of mating system on the genetic differentiation among populations.*

270 The mean nuclear G_{ST} observed at the end of the simulations increased with the
271 level of selfing rate (Figure 1a): 0.63 (for complete allogamy; SD = 0.01), 0.68 (SD =
272 0.02), 0.75 (SD = 0.02), 0.82 (SD = 0.02) and 0.88 (for complete autogamy; SD = 0.06).
273 All nuclear G_{ST} values were significantly different among selfing levels tested ($P <$
274 0.0001) and highly correlated ($R^2 = 0.89$; $P < 0.0001$) (Figure 1a). As expected, the
275 cytoplasmic G_{ST} did not depend on selfing rate: 0.87 (for allogamy; SD = 0.05), 0.87
276 (SD = 0.04), 0.87 (SD = 0.05), 0.88 (SD = 0.06) and 0.86 (for autogamy; SD = 0.06); all
277 values were not significantly divergent or correlated (Figure 1b). This was expected
278 since selfing only reduces the effective pollen migration rate and cytoplasmic genes do
279 not migrate through pollen.

280 2) *Impact of fragmentation intensity on the genetic differentiation among* 281 *populations.*

282 The mean G_{ST} values increased strongly as the dispersal rates decreased: 0.08
283 (SD = 0.006), 0.63 (SD = 0.01), 0.79 (SD = 0.01), 0.79 (SD = 0.007) and 0.79 (SD =
284 0.008) for nuclear markers and 0.32 (SD = 0.06), 0.87 (SD = 0.05), 0.92 (SD = 0.04),
285 0.92 (SD = 0.06) and 0.92 (SD = 0.05) for cytoplasmic markers (Figure 2). Both nuclear
286 and cytoplasmic G_{ST} values were not significantly divergent between the three
287 simulations with stronger fragmentation (dispersal rates of $10^{-5}/10^{-6}$, $10^{-7}/10^{-8}$ and zero).

288 Genetic differentiation and dispersal rate were highly correlated: $R^2 = 0.9989$ ($P <$
289 0.0001) for nuclear markers; $R^2 = 0.95$ ($P < 0.0001$) for cytoplasmic markers, indicating
290 a strong impact of fragmentation on genetic population differentiation.

291 **3) *Impact of life cycle on the genetic differentiation among populations.***

292 Differentiation was quite high among population for both nuclear and
293 cytoplasmic markers, whatever life span was assumed. As expected, the life cycle
294 counteracted the genetic differentiation and mean G_{ST} values decreased on higher life
295 spans: 0.74 (SD = 0.01), 0.66 (SD = 0.01), 0.64 (SD = 0.01), 0.63 (SD = 0.01) and 0.57
296 (SD = 0.01), for nuclear markers; and 0.91 (SD = 0.04), 0.92 (SD = 0.05), 0.9 (SD =
297 0.06), 0.87 (SD = 0.05) and 0.80 (SD = 0.09) for cytoplasmic markers (Figure 3). Tukey
298 tests showed no significant difference between 120 years and 300 years life span on
299 nuclear G_{ST} values ($P = 0.11$). However, cytoplasmic G_{ST} values were significantly
300 different between 500 years life span and all the others simulated (30, 60, 120 and 300
301 years; $P < 0.001$), and also between 60 and 300 years ($P < 0.05$). The correlation
302 between longevity and population genetic differentiation was very strong for nuclear
303 markers ($R^2 = 0.96$; $P < 0.0001$) but much weaker for cytoplasmic markers
304 ($R^2 = 0.3945$; $P < 0.0001$).

305 **4) *Impact of recent deforestation, predictions for the future.***

306 We simulated the future evolution of diversity during the next 20 000 years,
307 assuming that seed and pollen dispersal had completely stopped. The mean G_{ST} values
308 were: 0.63 (SD = 0.0005), 0.75 (SD = 0.005), 0.83 (SD = 0.006), 0.88 (SD = 0.006) and
309 0.92 (SD = 0.006) every 5 000 years (Figure 4). And the mean cytoplasmic G_{ST} values
310 also increased: 0.88 (SD = 0.0008) at time zero, 0.93 (SD = 0.02) after 5 000 years, 0.95
311 (SD = 0.02) after 10 000 years, 0.96 (SD = 0.02) after 15 000 years and 0.96 (SD =

312 0.01) after 20 000 years (Figure 4). Correlation was higher on nuclear markers ($R^2 =$
313 0.9974; $P < 0.0001$) than cytoplasmic markers ($R^2 = 0.7991$; $P < 0.0001$), which were
314 not significantly divergent between 10 000, 15 000 and 20 000 years predictions.

315

316 **4. Discussion**

317

318 We assessed the effect of breeding system, fragmentation intensity, and the impact
319 of life span on the nuclear and cytoplasmic genetic differentiation using a demo-genetic
320 model based on the genetic differentiation of *Caesalpinia echinata*'s natural
321 populations. This allowed us to study the potential impact of the recent deforestation on
322 this species. Our aim was to test the hypothesis that *C. echinata* populations have been
323 fragmented since the last glaciation, almost 20 000 years ago, when occurred the last
324 expansion of seasonally dry tropical forests followed by population isolation when
325 humid Atlantic forest expanded again (Pennington et al.,2004). This hypothesis could
326 explain the high genetic structure found with molecular markers: nuclear $F_{ST} = 0.58$ and
327 0.62, respectively for RAPD and AFLP (Cardoso et al., 1998; Cardoso et al., 2005) and
328 cytoplasmic $F_{ST} = 0.91$ (Lira et al., 2003).

329 We used ten demes with maximum biotic capacity of 300 plants. While this may
330 seem too few plants, this number was chosen because we intended to simulate the same
331 populations already analyzed by molecular markers on previous studies. Furthermore,
332 the data on brazil-wood abundance on Atlantic Rainforest is controversial. Although
333 brazil-wood was over-exploited during almost 500 years, there may be an unreal
334 perception – not properly documented – that *C. echinata* plants were easily identified on
335 Atlantic Rainforest, and that they occurred in dense populations. In addition, because of

336 better quality and higher amounts of the pigment dye, the vast majority of individuals
337 collected at that time were probably about 20 years old, suggesting that many
338 generations could have been explored during 500 years (Bueno et al., 2002). These
339 observations indicate that *C. echinata* populations could have been small-sized even if
340 great amounts of wood were extracted from Atlantic Rainforest.

341 Population size is very important on the maintenance of demographic and
342 genetic authenticity of a species because smaller populations are more affected by
343 stochastic events that cause impoverishment of plant populations of either common or
344 rare species (Broadhurst & Young, 2006; Honnay & Jacquemyn, 2007; Broadhurst et
345 al., 2008). This impoverishment can result of poor seed and pollen dispersal or less
346 seedling recruitment (Jacquemyn et al., 2007; Kolb, 2008). Although species with good
347 gene flow can counteract the effects small population size (Honnay et al., 2006), for *C.*
348 *echinata* we believe that simulations with larger populations would not directly affect
349 the genetic differentiation caused by the ecological traits addressed here. When
350 simulations with larger populations were carried out, the results showed periods longer
351 than 100 000 years would be needed to reach similar genetic differentiation parameters
352 achieved (data not shown). Although, a period larger than 20 000 years of fragmentation
353 for *C. echinata* populations is possible, still it is hard to know for sure.

354 **1) *Effect of mating system on the genetic differentiation among populations.***

355 On the first set of simulations analyzing different breeding-system species, we
356 observed that genetic differentiation increased with higher selfing rate, for nuclear genes
357 but not for cytoplasmic ones (Figure 1). Also, while the nuclear and cytoplasmic G_{ST}
358 values were very high, they showed substantial differences on allogamous species
359 (nuclear G_{ST} was 0.63, SD = 0.01; cytoplasmic G_{ST} was 0.86, SD = 0.05), what was not

360 found on autogamous species (nuclear $G_{ST} = 0.88$ and cytoplasmic $G_{ST} = 0.86$; SD =
361 0.06). *Caesalpinia echinata*'s natural populations have high levels of differentiation, but
362 the level of differentiation of nuclear and cytoplasmic markers is unlike, indicating that
363 this species is preferably an allogamous species, with very low level of selfing as
364 suggested by a study on experimental field plants of Giudice-Neto et al. (2005). While
365 there could be methods more reliable than simulations to study the brazil-wood mating
366 system, our result is still relevant to formulate conservation strategies for *C. echinata*
367 natural populations until further investigation is done.

368 **2) *Impact of fragmentation intensity on the genetic differentiation among***
369 ***populations.***

370 Another important feature investigated was the impact of pollen and seed
371 dispersal rates that may have counteracted the effects of post-glacial fragmentation,
372 genetic drift and population genetic erosion (Despres et al., 2002). A strong
373 differentiation on nuclear and cytoplasmic loci was observed. Gene flow counteracted
374 fragmentation in the simulations with the highest rate $10^{-1}/10^{-2}$ of pollen/seed dispersal,
375 leading to a very low genetic differentiation of both nuclear (mean $G_{ST} = 0.08$, SD =
376 0.006) and cytoplasmic loci (mean $G_{ST} = 0.32$, SD = 0.06), much below the empirical
377 results. As dispersal rates diminished and thus fragmentation effects increased, G_{ST}
378 values increased for both nuclear and cytoplasmic loci. This increase was much stronger
379 for cytoplasmic loci, as a consequence of a stronger drift caused by reduction of half of
380 the effective size than nuclear genes in a hermaphroditic species, because they are only
381 maternally transmitted.

382 Empirical *C. echinata* data on nuclear and cytoplasmic G_{ST} are closer to the
383 results with moderate rates of $10^{-3}/10^{-4}$ for pollen/seed dispersal. Dispersal mechanisms

384 are likely to be restricted because *C. echinata* seeds are dispersed by fruit explosion an
385 autochorous mechanism (Carvalho, 1994) and pollens are dispersed by medium- to
386 large-sized bees (*Centris sp* and *Xylocopa sp*) but also by the small honey bee *Apis*
387 *mellifera*, that might promote self-fertilization due to its visiting behavior (Borges et al.,
388 2008). This level of gene flow seems to be much lower than for many other tropical
389 species, who have reached a much lower G_{ST} value for nuclear genes (mean G_{ST} of 0.1;
390 Loveless et al., 1992). The level of cytoplasmic G_{ST} in *C. echinata* is nevertheless rather
391 similar to the ones observed in other tropical species (Collevatti et al., 2003; Dutech et
392 al., 2003), which indicates that it differs mainly for pollen flow, which thus seems to be
393 rather low in *C. echinata*, as compared to other species. Further field observation and
394 experiments (e.g. paternity analyses) are needed to confirm this pattern of *Caesalpinia*
395 *echinata*'s life history.

396 **3) *Impact of life cycle on the genetic differentiation among populations.***

397 Tree longevity is important for the persistence of genetic diversity, especially in
398 fragmented habitats (Hamrick, 2004). Our study indicated that longer life span
399 decreased genetic differentiation for both nuclear and cytoplasmic markers as fewer
400 generations are passed through time (Figure 3). Mean G_{ST} values decreased with life
401 span, for both nuclear and cytoplasmic markers. The values that were nearly consistent
402 with experimental data were obtained with simulations for 300-year life span (nuclear
403 $G_{ST} = 0.63$, cytoplasmic $G_{ST} = 0.87$), a value that is realistic for *C. echinata* (H.C. Lima,
404 personal communication), although we cannot discard a life span of 120 years that was
405 not significantly divergent from 300-year plants.

406 **4) *Impact of recent deforestation, predictions for the future.***

407 The fourth set of simulations intended to predict the future of *C. echinata*'s
408 natural populations in the Atlantic Rainforest. While this species is naturally isolated on
409 seasonally dry forests (Pennington et al., 2004), its level of fragmentation has strongly
410 increased because of deforestation due to anthropogenic action (Myers et al., 2000). Our
411 simulation predictions indicate that both nuclear and cytoplasmic genetic differentiation
412 would increase if seed and pollen dispersal were completely stopped as a result of this
413 recent fragmentation. Cytoplasmic G_{ST} reached almost the maximum value after 10 000
414 years of simulation without gene flow.

415 ***Conclusion***

416 Simulations helped us to infer some important ecological issues of *C. echinata*
417 plant communities in Brazil, such as allogamous breeding system, moderate gene flow
418 and long life span. Although much ecological and reproductive investigation is still
419 needed to clarify some points, *C. echinata*'s natural populations have already reached a
420 level of nuclear and cytoplasmic genetic differentiation so high that could hardly be
421 diminished. Nevertheless, it becomes crucial to slowdown the actual genetic
422 differentiation maintaining as much genetic diversity as possible. One method to
423 improve genetic diversity among populations is the transfer of individuals between
424 populations but it requires the knowledge of local adaptation process in order to avoid
425 outbreeding depression (Ramp et al., 2006; Raabová et al., 2007). This methodology
426 should be addressed with caution on *C. echinata*'s natural populations since trees from
427 different populations are genetically very divergent from each other, except populations
428 from very near locations (Cardoso et al., 1998; Cardoso et al., 2005; Lira et al. 2003).
429 The METAPOP predictions indicate a tendency to higher genetic isolation with time

430 leading to ecological drift and maybe even speciation in longer time ranges (Lavin et al.,
431 2004). It might also be necessary to do some ex situ conservation measures, such as
432 seed banking and vegetative propagation of key genotypes, as a complement to in situ
433 conservation of the species (Rajora and Mosseler, 2001; Castiñeiras et al., 2007; Yao et
434 al., 2007).

435

436 **5. Acknowledgements**

437 The simulations were performed at the Museum National d’Histoire Naturelle in Paris
438 using the linux-cluster facility. This cluster was partially financed by the Groupement
439 De Recherche CNRS “Génomique des Populations”. We thank the Brazilian Education
440 Council (CAPES) for the fellowship to CF Lira-Medeiros conduct the simulation
441 analyses at Museum d’Histoire Naturelle in Paris. Also the Brazilian research Council
442 (CNPq) for the research fellowship granted to PCG Ferreira. We are grateful to Monika
443 Zavodna from Museum National d’Histoire Naturelle in Paris that provided valuable
444 comments to previous versions of the manuscript, and Dr. Luiz Landau from
445 COPPE/UFRJ that provided facilities to run some of the simulations. The development
446 of the software Metapop is partially funded by the EVOLTREE network.

447

448 **6. References**

- 449 Austerlitz, F., Mariette, S., Machon, N., Gouyon, P.-H., Godelle, B., 2000. Effects of
450 colonization processes on genetic diversity: differences between annual plants and
451 tree species. *Genetics* 154, 1309-1321.
- 452 Borges, L.A., Sobrinho, M.S., Lopes, A.V., 2008. Phenology, pollination, and breeding
453 system of the threatened tree *Caesalpinia echinata* Lam. (Fabaceae), and a review
454 of studies on the reproductive biology in the genus. *Flora* (in press).
- 455 Bossuyt, B., 2007. Genetic rescue in an isolated metapopulation of a naturally
456 fragmented plant species, *Parnassia palustris*. *Conservation Biology* 21, 832-841.
- 457 Broadhurst, L.M., Young, A.G., 2006. Reproductive constraints for the long-term
458 persistence of fragmented *Acacia dealbata* (Mimosaceae) populations in southeast
459 Australia. *Biological Conservation* 133, 512-526.
- 460 Broadhurst, L.M., Young, A.G., Forrester, R., 2008. Genetic and demographic
461 responses of fragmented *Acacia dealbata* (Mimosaceae) populations in
462 southeastern Australia. *Biological Conservation*,
463 doi:10.1016/j.biocon.2008.08.021
- 464 Bueno, E., Roquero, A., Fernandes, F.L., Lima, H.C., Montaigne, J., Guedes, M.J.,
465 Manzano, N., 2002. *Paubrasil*. Axis Mundi, São Paulo (in portuguese).
- 466 Cardoso, M.A., Provan, J., Powell, W., Ferreira, P.C.G., de Oliveira, D.E., 1998. High
467 genetic differentiation among remnant populations of the endangered *Caesalpinia*
468 *echinata* Lam. (Leguminosae Caesalpinioideae). *Molecular Ecology* 7, 601-608.
- 469 Cardoso, S.R.S., Provan, J., Lira, C.F., Pereira, L.O.R., Ferreira, P.C.G., Cardoso, M.A.,
470 2005. High levels of genetic sub-structuring as a result of population

471 fragmentation in the tropical tree species *Caesalpinia echinata* Lam. Biodiversity
472 and Conservation 14, 1047-1057.

473 Carvalho, P.E.R., 1994. Espécies florestais brasileiras. Recomendações silviculturais,
474 potencialidades e uso da madeira. EMBRAPA-CNPQ/SP1, Brasília.

475 Castiñeiras, L., Guzmán, F.A., Duque, M.C., Shagardsky, T., Cristóbal, R., de Vicente,
476 M.C., 2007. AFLPs and morphological diversity of *Phaseolus lunatus* L. in Cuban
477 home gardens: approaches to recovering the lost ex situ collection. Biodiversity
478 and Conservation 16, 2847-2865.

479 Collevatti, R.G., Grattapaglia, D., Hay, J.D., 2003. Evidences for multiple maternal
480 lineages of *Caryocar brasiliense* populations in the Brazilian Cerrado based on
481 the analysis of chloroplast DNA sequences and microsatellite haplotype variation.
482 Molecular Ecology 12, 105-115.

483 Couvet, D., 2002. Deleterious effects of restricted gene flow in fragmented populations.
484 Conservation Biology 16, 369-376.

485 Despres, L., Lorient, S., Gaudeul, M., 2002. Geographic pattern of genetic variation in
486 the European globeflower *Trollius europaeus* L. (Ranunculaceae) inferred from
487 amplified fragment length polymorphism markers. Molecular Ecology 11, 2337-
488 2347.

489 Dutech, C., Joly, H. I., Jarne, P., 2004. Gene flow, historical population dynamics and
490 genetic diversity within French Guianan populations of a rainforest tree species,
491 *Vouacapoua americana*. Heredity 92, 69-77.

492 Dutech, C., Maggia, L., Tardy, C., Joly, H. I., Jarne, P., 2003. Tracking a genetic signal
493 of extinction-recolonization events in a neotropical tree species: *Vouacapoua*
494 *americana* Aublet in French Guiana. Evolution 57, 2753-2764.

495 Giudice-Neto, J.D.G., Sebbenn, A.M., Kageyama, P.Y., 2005 (in portuguese). Sistema
496 de reprodução em *Caesalpinia echinata* Lam. implantada em arboreto
497 experimental. Revista Brasileira de Botânica 28, 409-418.

498 Hamrick, J.L., 2004. Response of forest trees to global environmental changes. Forest
499 Ecology and Management 197, 323-335.

500 Hamrick, J.L., Godt, M.J.W., 1990. Allozyme diversity in plant species, in: Brown,
501 A.H.D., Clegg, M.T., Kahler, A.L., Weir, B.S. (Eds.), Plant Population Genetics,
502 Breeding and Genetic Resources. Sinauer Associates Inc., Sunderland, MA, pp.
503 43-63.

504 Hamrick, J. L., Godt, M. J. W., Sherman-Broyles, S. L., 1992. Factors influencing
505 levels of genetic diversity in woody plant species. New Forests 6, 95-124.

506 Hedrick, P.W., 1999. Perspective: Highly variable loci and their interpretation in
507 evolution and conservation. Evolution 53, 313-318.

508 Hedrick, P.W., 2005. A standardized genetic differentiation measure. Evolution 59,
509 1633-1638.

510 Honnay, O., Coart, E., Butaye, J., Adriaens, D., Van Glabeke, S., Roldán-Ruiz, I., 2006.
511 Low impact of present and historical landscape configuration on the genetics of
512 fragmented *Anthyllis vulneraria* populations. Biological Conservation 127, 411-
513 419.

514 Honnay, O., Jacquemyn, H., 2007. Susceptibility of common and rare plant species to
515 the genetic consequences of habitat fragmentation. Conservation Biology 21, 823-
516 831.

517 Jacquemyn, H., Vandepitte, K., Brys, R., Honnay, O., Roldán-Ruiz, I., 2007. Fitness
518 variation and genetic diversity in small, remnant populations of the food deceptive
519 orchid *Orchis purpurea*. *Biological Conservation* 139, 203-210.

520 Kimura, M., Weiss, G.H., 1964. The stepping stone model of population structure and
521 the decrease of genetic correlation with distance. *Genetics* 49, 561-576.

522 Kolb, A., 2008. Habitat fragmentation reduces plant fitness by disturbing pollination
523 and modifying response to herbivory. *Biological Conservation* 141, 2540-2549.

524 Lavin, M., Schrire, B.P., Lewis, G., Pennington, R.T., Delgado-Salinas, A., Thulin, M.,
525 Hughes, C.E., Matos, A.B., Wojciechowski, M.F., 2004. Metacommunity process
526 rather than continental tectonic history better explains geographically structured
527 phylogenies in legumes. *Philosophical Transactions of the Royal Society of*
528 *London. Series B; Biological sciences* 359, 1509-1522.

529 Le Corre, V., Machon, N., Petit, R.J., Kremer, A., 1997. Colonization with long-
530 distance seed dispersal and genetic structure of maternally inherited genes in
531 forest trees: a simulation study. *Genetical Research* 69, 117-125.

532 Lira, C.F., Cardoso, S.R.S., Ferreira, P.C.G., Cardoso, M.A., Provan, J., 2003. Long-
533 term population isolation in the endangered tropical tree species *Caesalpinia*
534 *echinata* Lam. revealed by chloroplast microsatellites. *Molecular Ecology* 12,
535 3219-3225.

536 Loveless, M. D., 1992. Isozyme variation in tropical trees: patterns of genetic
537 organization. *New Forests* 6, 67-94.

538 Machon, N., Baradat, D., Godelle, B., 1995. METAPOPOP a program simulating
539 evolutionary processes acting on metapopulations. Paris, France: Laboratoire ES,
540 Paris-Sud.

541 Machon, N., Bardin, P., Mazer, S.J., Moret, J., Godelle, B., Austerlitz, F., 2003.
542 Relationship between genetic structure and seed and pollen dispersal in the
543 endangered orchid *Spiranthes spiralis*. *New Phytologist* 157, 677-687.

544 Mariette, S., Le Corre, V., Kremer, A., 1999. Sampling within the genome for
545 measuring within-population diversity: Trade-offs between markers, in: Gillet,
546 E.M. (Eds.), 1999. *Sampling Strategies for Marker Analysis. Compendium of the*
547 *Research Project Development, optimization and validation of molecular tools for*
548 *assessment of biodiversity in forest trees in the European Union DGXII*
549 *Biotechnology FW IV Research Programme Molecular Tools for Biodiversity.*

550 Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000.
551 Biodiversity hotspots for conservation priorities. *Nature* 403, 853-858.

552 Navascués, M., Emerson, B.C., 2005. Chloroplast microsatellites: measures of genetic
553 diversity and the effect of homoplasmy. *Molecular Ecology* 14, 1333-1341.

554 Novick, R. R., Dick, C. W., Lemes, M. R., Navarro, C., Caccone, A. & Bermingham,
555 E., 2003. Genetic structure of Mesoamerican populations of Big-leaf mahogany
556 (*Swietenia macrophylla*) inferred from microsatellite analysis. *Molecular Ecology*
557 12, 2885-2893.

558 Pennington, R.T., Lavin, M., Prado, D.E., Pendry, C.A., Pell, S.K., Butterworth, C.A.,
559 2004. Historical climate change and speciation: neotropical seasonally dry forest
560 plants show patterns of both Tertiary and Quaternary diversification.
561 *Philosophical Transactions of the Royal Society of London. Series B; Biological*
562 *sciences* 359, 515-537.

563 Provan, J., Soranzo, N., Wilson, N.J., Goldstein, D.B., Powell, W., 1999. A low
564 mutation rate for chloroplast microsatellites. *Genetics* 153, 943-947.

565 R Development Core Team, 2004. R: A language and environment for statistical
566 computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-
567 900051-00-3, <http://www.Rproject.org>.

568 Raabová, J., Münzbergová, Z., Fischer, M., 2007. Ecological rather than geographic or
569 genetic distance affects local adaptation of the rare perennial herb, *Aster amellus*.
570 Biological Conservation 139, 348-357.

571 Rajora, O.P., Mosseler, A., 2001. Challenges and opportunities for conservation of
572 forest genetic resources. Euphytica 118, 197-212.

573 Ramp, J.M., Collinge, S.K., Ranker, T.A., 2006. Restoration genetics of the vernal pool
574 endemic *Lasthenia conjugens* (Asteraceae). Conservation Genetics 7, 631-649.

575 Saatchi, S., Agosti, D., Alger, K., Delabie, J., Musinsky, J., 2001. Examining
576 fragmentation and loss of primary forest in the southern Bahian Atlantic forest of
577 Brazil with radar imagery. Conservation Biology 15, 867-875.

578 Seltzer, G.O., Rodbell, D.T., Baker, P.A., Fritz, S.C., Tapia, P.M., Rowe, H.D., Dunbar,
579 R.B., 2002. Early warming of tropical South America at the last Glacial-
580 Interglacial transition. Science 296, 1685-1686.

581 SOS Mata Atlântica, INPE, 2002. Atlas dos remanescentes florestais da Mata Atlântica,
582 período 1995–2000. Fundação SOS Mata Atlântica & INPE. São Paulo.

583 Topping, C., Ostergaard, S., Pertoldi, C., Bach, L.A., 2003. Modeling the loss of genetic
584 diversity in vole populations in a spatially and temporally varying environment.
585 Annales zoologici Fennici 40, 255-267.

586 Yao, X., Ye, Q., Kang, M., Huang, H., 2007. Microsatellite analysis reveals
587 interpopulation differentiation and gene flow in the endangered tree

588 *Changiostyrax dolichocarpa* (Styracaceae) with fragmented distribution in central
589 China. *New Phytologist* 176, 472-480.
590 Young, A., Boyle, T., Brown, T., 1996. The population genetic consequences of habitat
591 fragmentation for plants. *Trends in Ecology and Evolution* 11, 413-418.

7. Captions:

Figure 1: Simulation of the effect of mating system with five different selfing rates on genetic differentiation among populations based on *C. echinata* modeling. G_{ST} values were recorded after 20 000 years of simulation and 20 replicates were used to calculate means and deviations plotted on graphics.

a) Mean nuclear G_{ST} values for five selfing rates tested by the simulations. b) Mean cytoplasmic G_{ST} values for five selfing rates.

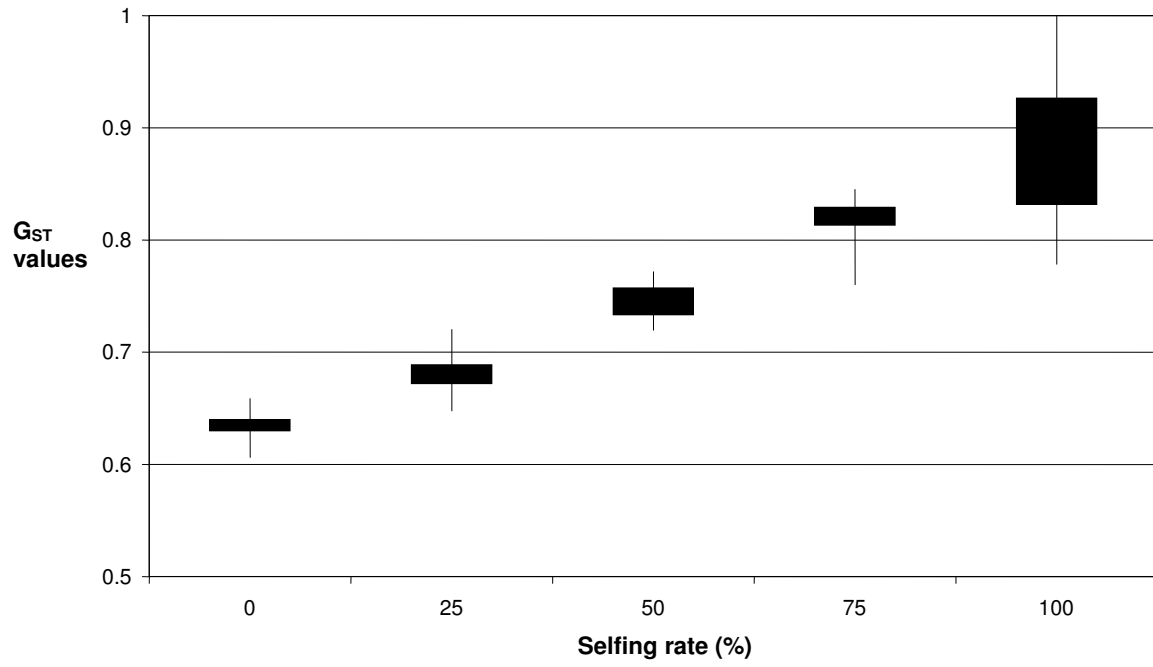
Figure 2: Simulations based on *C. echinata* model of the impact of fragmentation represented by five different pollen and seed dispersal rates on inter-population genetic differentiation. Nuclear (black columns) and cytoplasmic (white columns) mean G_{ST} values for the different gene flows tested of 20 replicates after 20 000 years of simulation. Errors bars represent standard deviations.

Figure 3: Simulation indicating the influence of different life spans on inter-population genetic differentiation based on *C. echinata* modeling. Black columns representing mean nuclear G_{ST} values and white columns represent mean cytoplasmic G_{ST} values calculated using 20 replicates after 20 000 years of simulation. Errors bars represent standard deviations.

Figure 4: Simulations obtained with *C. echinata* metapopulation model from present to 20 000 years ahead for future predictions of its natural populations. Mean G_{ST} values were calculated every 5 000 years of simulation for nuclear genes (black columns) and cytoplasmic genes (white columns) using 20 replicates. Errors bars represent standard deviations.

Figure 1:

a)



b)

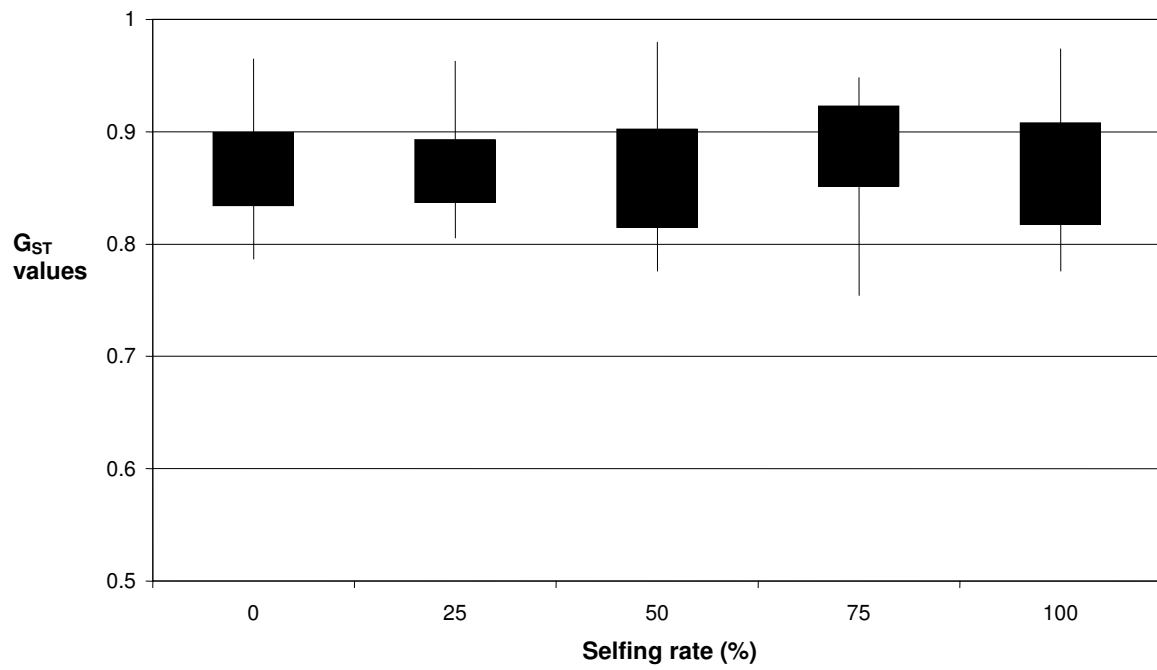


Figure 2:

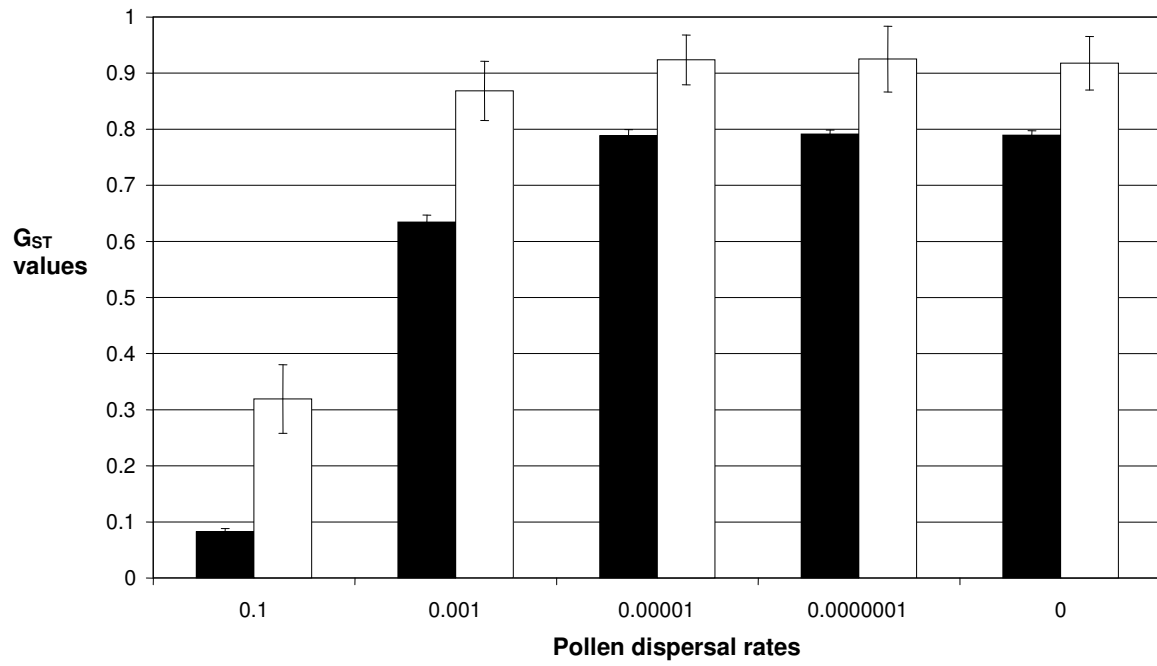


Figure 3:

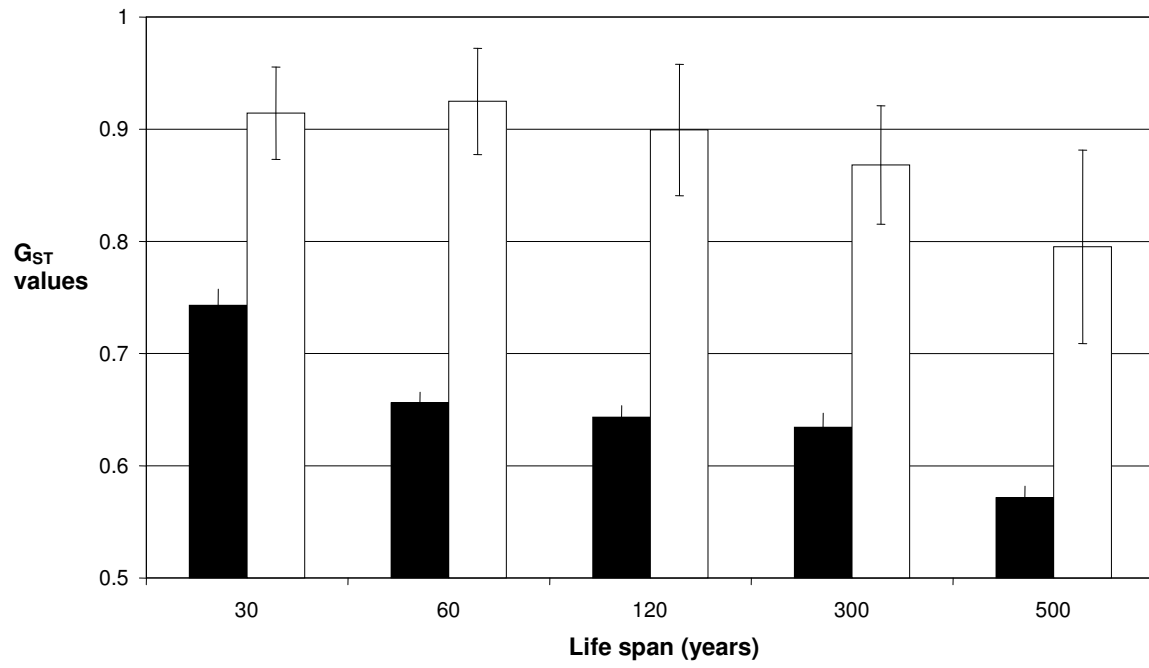
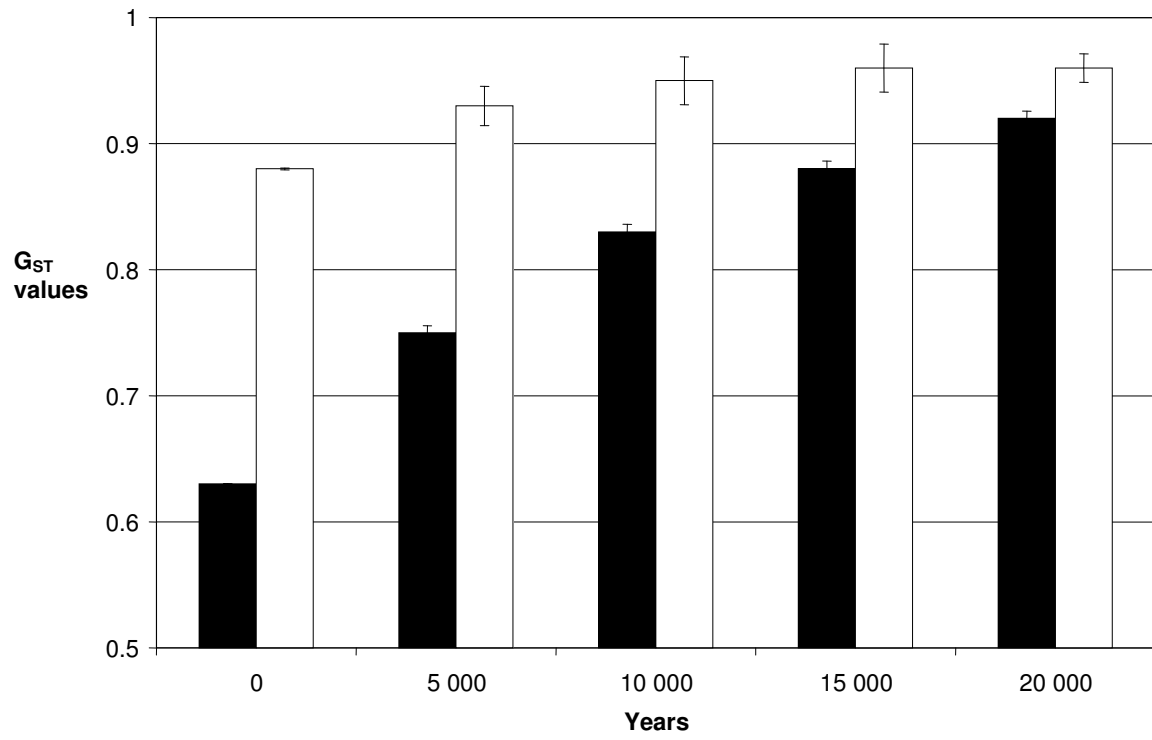


Figure 4:



Anexo II

Exemplo de matrizes de dispersão do pólen e da semente usadas no METAPOPOP para simulações com dispersão de $10^{-3}/10^{-4}$. Cada coluna corresponde a uma deme e as linhas mostram as possíveis combinações de dispersão entre as demes. Seguindo o modelo “Stepping Stone”, ocorre troca apenas entre as populações mais próximas. E dentro da população principal não há limite de troca de pólen e sementes, sendo o valor de dispersão definido como 1.

a) Matriz de dispersão do pólen de 10^{-3} .

Deme	1	2	3	4	5	6	7	8	9	10
	1	0,001	0	0	0	0	0	0	0	0
	0,001	1	0,001	0	0	0	0	0	0	0
	0	0,001	1	0,001	0	0	0	0	0	0
	0	0	0,001	1	0,001	0	0	0	0	0
	0	0	0	0,001	1	0,001	0	0	0	0
	0	0	0	0	0,001	1	0,001	0	0	0
	0	0	0	0	0	0,001	1	0,001	0	0
	0	0	0	0	0	0	0,001	1	0,001	0
	0	0	0	0	0	0	0	0,001	1	0,001
	0	0	0	0	0	0	0	0	0,001	1

b) Matriz de dispersão da semente de 10^{-4} .

Deme	1	2	3	4	5	6	7	8	9	10
	1	0,0001	0	0	0	0	0	0	0	0
	0,0001	1	0,0001	0	0	0	0	0	0	0
	0	0,0001	1	0,0001	0	0	0	0	0	0
	0	0	0,0001	1	0,0001	0	0	0	0	0
	0	0	0	0,0001	1	0,0001	0	0	0	0
	0	0	0	0	0,0001	1	0,0001	0	0	0
	0	0	0	0	0	0,0001	1	0,0001	0	0
	0	0	0	0	0	0	0,0001	1	0,0001	0
	0	0	0	0	0	0	0	0,0001	1	0,0001
	0	0	0	0	0	0	0	0	0,0001	1

Anexo III

DNA methylation of *Laguncularia racemosa* populations affected by different salinity regimes

Catarina da Fonseca Lira-Medeiros^{a,b*}, Christian Parisod^{c,d}, Ricardo Avancini Fernandes^a,
Camila Souza da Mata^a, Mônica Aires Cardoso^a, Paulo Cavalcanti Gomes Ferreira^{a,b}

^aDipeq – Instituto de Pesquisas do Jardim Botânico do Rio de Janeiro – Rua Pacheco Leão, 915 – 22460-030 Rio de Janeiro, Brasil.

^bInstituto de Bioquímica Médica – Universidade Federal do Rio de Janeiro – Avenida
Bauhinia, 400 – CCS, Bloco D(SS) sala 05 – Ilha do Fundão – 21491-590 Rio de Janeiro,
Brasil.

^cLaboratoire de Biologie Cellulaire – Institut J.-P. Bourgin - INRA centre de Versailles –
78026 Versailles, France.

^dPresent address: National Center for Biosystematics, University of Oslo, Norway.

E-mail Addresses:

catilira@bioqmed.ufrj.br

christian.parisod@nhm.uio.no

ricardo.avancini@yahoo.com.br

csouza@jbrj.gov.br

mcardoso@jbrj.gov.br

paulof@bioqmed.ufrj.br

Running title: Epigenetics of a natural salt stress species

Keywords: White mangrove, epigenetics, salinity stress, plant development, cytosine methylation, AFLP, MSAP

Type of article: *Letters*

Abstract: # words

Manuscript: # words

References: #

Figures and tables: #

* **Corresponding author:** Catarina DA FONSECA LIRA-MEDEIROS

Instituto de Bioquímica Médica – Universidade Federal do Rio de Janeiro – Avenida
Bauhinia, 400 – CCS, Bloco D(SS) sala 05 – Ilha do Fundão – 21491-590 Rio de Janeiro,
Brasil. Phone: 55 21 2562 6751; Fax: 55 21 2270 8647, Email: catilira@bioqmed.ufrj.br

Abstract

Mangroves are typical plant formations submitted to variations of salinity and nutrient input. *Laguncularia racemosa* occurs among contrasted habitats, on riverside (RS) or near salt marshes (SM), where plants present remarkable morphological differences. In plants, epigenetic modifications, which might be under natural selection, may occur in response to stress. To investigate the genetic and epigenetic contribution to the phenotypic differences observed among two natural populations of *L. racemosa*, we used AFLP and MSAP. SM population was hypomethylated and less polymorphic than RS population. Three statistical treatments were used to analyze genetic and epigenetic differentiation between these populations: frequency-based (G_{ST}), multivariate (β_{ST}) and Bayesian analysis (F_{ST}). All methods resulted on higher epigenetic than genetic divergence. Co-Inertia analysis explored jointly the genetic and epigenetic structure, showing a tendency of a common epigenetic profile on SM population, which might be a result from the stressful habitat selection occurring independently of their genetic profile.

Keywords: White mangrove, epigenetics, salinity stress, plant development, cytosine methylation, phenotype, natural selection

Introduction

Mangroves are ecosystems occurring on tropical and subtropical coastlines, which are submitted to daily variations of water salinity and sea level oscillations (Saenger 2003; Barth *et al.* 2006). Mangrove species tolerates a wide range of environmental conditions and can thus present divergent structural and morphological characteristics in different geographic zones (Schaeffer-Novelli *et al.* 1990). Most species living in these areas have developed specialized characteristics such as vivipary, salt exclusion and aerial roots to survive in extreme environmental conditions (Tomlinson 1986; Shi *et al.* 2005). In general, in regions with suitable topography allows large tidal ranges and abundant inputs of fresh water, rainfall, nutrients and sediments mangroves have better development, reaching their maximum development: over 40 meters high; (Schaeffer-Novelli *et al.* 1990). On the other hand, mangroves adaptations allow them to colonize even habitats with limiting factors, like periodic drought, hurricanes, and hyper-saline soils also called salt marshes. Although adapted to saline environments, in such extreme salt stress conditions, mangroves have abnormal development, reaching only 1.5 to 3 meters high; with shrub-like plant physiognomy (Schaeffer-Novelli *et al.* 1990). The growth and development reduction on salt-marsh mangroves can be explained by cellular and biochemical processes modifications caused by high salinity, such as protein synthesis and photosynthesis (Parida & Das 2004, Parida *et al.* 2004). These modifications can happen even on well-adapted halophyte species that are known to be positively influenced by low to moderate salinity levels (Ball 2002).

Laguncularia racemosa (L.) Gaertn. f. (Family: Combretaceae), also known as white mangrove, is a west-mangrove species, broadly found in Atlantic Ocean mangroves (Tomlinson, 1986). This species occurs abundantly in Brazil, either close to river basins or near salt marshes, suggesting that individuals from each site might be under divergent selective pressures caused by different salinity levels. *L. racemosa* individuals located on both

mangrove formations happen to be morphologically different concerning height and branch ramifications as postulated by Schaeffer-Novelli *et al.* (1990). Riverside (RS) plants are 10 meters high, having tree-like structure, with branches on the canopy. On the other hand, plants near salt marsh (SM) are shrub-like, reaching only 2 meters high with several basal stems. Epigenetic changes can alter phenotypes, both morphological and molecular traits, without changing nucleotide sequence on coding or promoter regions of a gene (Jablonka & Lamb 1998). Moreover, they are responsible for some heritable differences among individuals (Rapp & Wendel 2005). There are evidences that the regulation of gene expression and chromatin structure, as well as the maintenance of genome stability against the proliferation of transposable elements, are controlled by epigenetic changes such as DNA methylation that can be maintained among generations in plants (Bender 2000; Gehring & Henikoff 2007). There is also strong evidence that methylation of CpG dinucleotides is crucial for several plant developmental processes (Lee & Chen 2001; Hao *et al.* 2004; Huettel *et al.* 2006; Zhang *et al.* 2006). Therefore, natural selection could act on many traits modulated by epigenetics, without DNA sequence changes (Rapp & Wendel 2005). Although epigenetics mechanisms controlling DNA methylation on coding and promoter regions of genes and their effects on phenotypes have been investigated at some extent (Zhang 2008), their consequences on the evolution of natural plant populations is currently underexplored (Kalisz & Purugganan 2004; Lukens & Zhan 2007; Bossdorf *et al.* 2008).

Based on Amplified Fragment Length Polymorphism (AFLP; Vos *et al.* 1995), the Methyl-Sensitive Amplified Polymorphism (MSAP) technique provides informative epigenetic data (Xiong *et al.* 1999). The combination of AFLP and MSAP systems using two isoschizomer frequent-cutter enzymes (*MspI* and *HpaII*), which recognize the same restriction site but have different DNA methylation sensitivities, allows the identification of methylation polymorphisms (Cervera *et al.* 2002; Schones & Zhao 2008, Schellenbaum *et al.* 2008). Most

studies evaluating changes on cytosine methylation addressed inheritance of methylation patterns and polymorphisms among ecotypes or strains of crop and/or polyploid species (Hao *et al.* 2004; Jaligot *et al.* 2004; Salmon *et al.* 2005; Takata *et al.* 2005; Keyte *et al.* 2006; Zhao *et al.* 2007), but the potential to analyze epigenetic population structure and its correlation with natural phenotypic variation using MSAP methodology is still unexplored.

RS and SM populations thus provide an interesting framework to assess the genetic as well as epigenetic structure of natural populations, using AFLP and MSAP methodologies respectively, aiming to understand their contribution on phenotypic differences between individuals from these natural populations. This study is one of the first assessing the epigenetic diversity and structure of natural populations and brings fresh evidence that DNA methylation alterations are crucial for population sustenance in stressful habitats.

Materials and Methods

Plant Material

Two populations of *Laguncularia racemosa*, located in Sepetiba Bay's mangrove ecosystem – Rio de Janeiro Brazil (23°00'S–23°05'S and 43°30'W–43°35'W), were analyzed. The first population is located on the river basin (called riverside or RS) and individuals have a tree-like appearance. The other population occurs near a salt marsh (called SM) and individuals are abnormally developed, with a shrub-like appearance. Individuals from both populations are one or two hundreds of meters apart, separated by a *Rhizophora mangle* forest, typical on mangroves. The area analyzed had near 16 hectares limited by the salt marsh at one side and Piracão river at the other side. Young and undamaged leaves from 17 randomly chosen adult trees were sampled from each population and immediately conditioned on silica gel for posterior genetic analysis.

DNA extraction

DNA extraction was carried based on Cardoso *et al.* (1998) with modifications, including a scale down process using 50mg of dried leaves. PVP 40 000 was used to improve DNA quality. DNA was resuspended in 100µl of sterile water. Quantification of DNA was estimated on 1% agarose gels.

AFLP and MSAP analysis

AFLP and MSAP methodologies were based on Vos *et al.* (1995) and Xiong *et al.* (1999), respectively. Samples were submitted to *EcoRI* digestion using 1µg of genomic DNA and 10U of enzyme (Promega^R) with 1X buffer H in a final volume of 200µl. Digested DNA was precipitated with 0.1vol of 3M sodium acetate and 2.5vol of 100% ethanol and afterwards washed on 70% ethanol. Half of *EcoRI*-digested DNA was used on each isoschizomers enzymes digestion with 5U of *MspI* in 1X Multicore buffer (Promega^R) or 5U of *HpaII* (Promega^R) in 1X Buffer B (Promega^R) in a final volume of 50µl. Incubations were all performed at 37°C for 6 hours and enzymes were afterwards denatured at 65°C for 20 minutes. Adapter ligation was performed with 20µl of digested DNA, 1X T4 DNA ligase buffer (Promega^R), 1U T4 DNA ligase enzyme (Promega^R), 5pmol of each *EcoRI* adapter (Vos *et al.* 1995) and 50pmol of each *MspI/HpaII* adapters (Xiong *et al.* 1999) in a 30µl reaction for 3 hours at 20°C.

Pre-amplification was conducted on a 20µl reaction using 2µl of ligated and digested DNA, 1X PCR buffer, 0.4mM dNTPs, 30ng of *EcoRI* and *MspI/HpaII* basic primers (Table 1) and 2U Taq polymerase (Ludwig Biotecnologia Ltda). The reactions were carried out on 25 cycles of 94°C 1 minute, 56°C 1 minute and 72°C 2 minutes with a 10-minutes final extension. Pre-amplification products were diluted 20 times and 5µl was used on the selective amplifications. These 20µl reactions contained 30ng of each selective primer *EcoRI* and *MspI/HpaII* (Table 1), 0.2mM dNTPs, 1X PCR buffer and 2U Taq polymerase (Ludwig Biotecnologia Ltda). The

touchdown program performed was: 94°C 30 seconds, 65°C 30 seconds and 72°C 1 minute decreasing 0.7°C of the annealing temperature per cycle during 12 cycles and then 24 cycles of 94°C 30 seconds, 56°C 1 minute and 72°C 2 minutes with final extension of 5 minutes length. The final amplification products were separated by electrophoresis for 2:30 hours at 60W on a 4% denaturing polyacrylamide gel with 7.5M urea. Gels were stained with a 0.1% silver nitrate solution plus 0.5% formaldehyde for 30 minutes after gel fixation on 10% acetic acid solution for 20 minutes. Staining development used a 6% sodium carbonate solution with 0.5% formaldehyde and 2µg of sodium thiosulphate for 3 minutes and reaction was stopped with 10% acetic acid solution. Silver-stained gels were photo-documented for further scoring.

Data analysis

The 34 samples were scored for presence or absence of AFLP and MSAP fragments.

Polymorphic bands from *EcoRI/MspI* were used to assess the genetic structure of *L. racemosa* populations, while polymorphic *EcoRI/HpaII* profiles informed about the epigenetic structure.

Both digests analyzed on the same sample showed the level of methylation on *L. racemosa*

genome. Presence of fragments on both *EcoRI/MspI* and *EcoRI/HpaII* digests indicated a

non-methylated site, while methylated sites of both strand internal cytosines were

characterized by the presence of *EcoRI/MspI* fragments absent on *EcoRI/HpaII* profiles

(Peraza-Echeverria *et al.* 2001, Salmon *et al.*, 2008). Loci where the opposite profile was

found, fragment present on *EcoRI/HpaII* but absent on *EcoRI/MspI* digestion, were excluded

from analysis, since they represent hemimethylation of the external cytosine that are rare and

generally not inheritable between generations (Salmon *et al.*, 2008). Methylated loci fixed

within population, i.e. all samples from one population had the same fragment present on

EcoRI/MspI profile but absent on *EcoRI/HpaII* profile, as well as polymorphic methylated

loci were counted. Both methylated and non-methylated loci were considered polymorphic

when at least one sample had a different pattern of presence or absence of fragments from the rest of the same population.

Dealing with dominant markers, heterozygosis is not obtained directly. In order to estimate population genetic structure, deviation from Hardy-Weinberg equilibrium thus has to be either: (i) assumed as null, (ii) bypassed, or (iii) assessed by other means (Bonin *et al.* 2007; Parisod & Christin 2008). Here, these three possibilities have been explored and compared to provide robust estimates of the genetic and epigenetic structure of *L. racemosa* populations using AFLP and MSAP respectively.

First, genetic diversity (H_{pop}) was assessed by Shannon diversity index calculated based on the frequency of each band out of the 17 individuals for AFLP and MSAP systems. As recommended by Bussell (1999), $\log_2(0)$ was replaced by 0 for fixed absent bands.

Significant differences of Shannon index among populations for both *EcoRI/MspI* and *EcoRI/HpaII* profiles have been assessed by Kruskal-Wallis rank sum test, which calculates the significant differences between data from the same loci. Significance of the tests was adjusted with Bonferroni correction (Rice 1989). Assuming that populations are on Hardy-Weinberg equilibrium, genetic and epigenetic structure was computed for each AFLP and MSAP loci as $G_{ST} = (H_{tot} - H_{pop}) / H_{tot}$ (Bussell 1999).

Second, individual profiles were also investigated by multivariate analyses because it is a band-based approach that does not assume Hardy-Weinberg equilibrium. Principal Component Analysis (PCA) on inter-profile covariance matrix followed by Between-group Eigen Analysis (BPCA, see details in Parisod & Christin 2008) was computed on AFLP and MSAP data using ADE-4 (Thioulouse *et al.* 1996). BPCA (i.e. PCA among groups based on the PCA among individuals) divides the variance into within- and between-group components and, given that it is a Euclidian approach, can be considered as analogous to *F*-statistics (called β_{ST}). The statistical significance was assessed by the Romesburg randomization test

(9999 permutations). In addition, multivariate analyses allow for the joint analysis of genetic and epigenetic structure through statistical procedures maximizing and testing the common variance of different datasets. Here, the symmetrical Co-Inertia Analysis was used to investigate the association between AFLP and MSAP profiles by projecting the PCA scores of individuals into a new subspace maximizing their covariance. Unlike the related Canonical Correspondence Analysis, Co-Inertia analysis does not rely on linear regressions and thus can be safely used for any number of variables to be related (Dolédec & Chessel 1994).

Significance of this association has been tested in ADE-4 by a procedure in which rows of AFLP and MSAP tables were randomly permuted 9999 times.

Third, Bayesian statistical procedures have been used to investigate population genetic structure with dominant markers by estimating the deviation from Hardy-Weinberg equilibrium (Holsinger *et al.* 2002). Indeed, Bayesian estimates of genetic structure incorporate the uncertainty about F_{ST} and F_{IS} by Monte-Carlo Markov Chain (MCMC). After MCMC was run for 250,000 times to ensure convergence of the chain to its stationary distribution and discarding the first 50,000 simulations, estimates every five steps were retained to provide a full description of the genetic structure. Here, preliminary simulations (not shown) suggested that the full model (F_{ST} and $F_{IS} \neq 0$) should be used because it fitted the data well and incorporated the lowest complexity (DIC = 117 and 247, for AFLP and MSAP data, respectively; see Holsinger *et al.* 2002 for details). Although the f-free model and $f = 0$ model were relatively close, the full model was assumed as the most suitable because *L. racemosa* is self-fertile (Landry and Rathcke 2007). Thus, the most likely F_{ST} and F_{IS} were estimated with their variance and reliable 95% credible intervals, using Hickory v1.1 software (Holsinger *et al.* 2002).

Results

Methylation patterns and polymorphisms

Four selective primer-pair combinations (Table 1) provided 97 reliable loci among the 34 samples. Riverside (RS) population had 39 methylated loci, i.e. band present on *EcoRI/MspI* digests and absent on *EcoRI/HpaII* digests on at least one sample, while salt marsh (SM) population had only 20 methylated loci, accounting for 40.2% and 20.6% of all analyzed loci (Table 2). Level of polymorphism was higher on RS than SM population. On RS, 84.6% of methylated loci were polymorphic but six loci were fixed on their methylated state for all samples from this population (Table 2). Almost half of SM methylated loci (55%) were polymorphic leaving to nine loci fixed on this population (Table 2). AFLP markers showed 15 polymorphic loci (15.5%) on RS and only 10 (10.3%) on SM population. By the other hand, MSAP markers showed 37 polymorphic loci (38.1%) on RS and 13 (13.4%) on SM.

Genetic estimates (AFLP)

Three methods were used to calculate the genetic differentiation among RS and SM populations of *L. racemosa*. AFLP data had within-population Shannon diversity index = 0.048 and 0.035 for RS and SM populations, respectively (Table 3). Kruskal-Wallis rank sum test showed no significant difference between these indexes ($p = 29.2$). Population genetic structure was calculated by three different statistical methods; the first based on Hardy-Weinberg equilibrium assumptions, the second on multivariate analysis and the third on Bayesian analysis: $G_{ST} = 0.111$, $\beta_{ST} = 0.148$ ($p < 0.001$) and $F_{ST} = 0.119$ (SD = 0.034; 95% credible interval: 0.055 - 0.186), respectively (Table 3).

Epigenetic estimates (MSAP)

On MSAP data, the Shannon index of diversity was 0.101 and 0.043 for RS and SM populations, respectively, and was significantly different after Bonferroni correction ($p <$

0.025) (Table 3). These values were also significantly higher than AFLP indexes calculated for RS ($p < 0.001$) and SM ($p < 0.01$) populations. The three methods used to calculate epigenetic population structure had more expressive values than the genetic ones obtained with AFLP data, although only significantly divergent on multivariate analysis results. The epigenetic indexes was: $G_{ST} = 0.254$, $\beta_{ST} = 0.212$ ($p < 0.001$) and $F_{ST} = 0.160$ (SD = 0.026; 95% credible interval: 0.110 - 0.210) (Table 3).

Genetic versus epigenetic population structure

The PCA analysis on covariance matrix of AFLP data summarized 40.2% of the total inertia on the two first principal components (Figure 1a), and 35.7% of the MSAP data (Figure 1b). In both genetic (AFLP) and epigenetic (MSAP) PCA subspaces, individuals from RS and SM populations are clustered on two distinct groups along the first axis. This differentiation is more evident at the epigenetic level than at the genetic level, where similar MSAP profiles corresponds to samples collected on the same population.

The two first axis of the Co-Inertia analysis explained 74.4% of the genetic co-variation between AFLP and MSAP datasets and this association was significantly different than expected at random ($p < 0.001$). In the Co-Inertia subspace (Figure 1c), three individuals from RS population and three from SM population showed similar genetic (AFLP) profiles, but rather divergent epigenetic (MSAP) profiles. These epigenetic profiles points toward opposite directions clearly occupying positions near the other individuals from the same population.

Discussion

Individuals that have distinct physiographic patterns growing in two nearby populations of *Laguncularia racemosa* under different salinity conditions were investigated by AFLP and MSAP analysis. Environmental fluctuations, like that occurring in mangrove ecosystems,

could increase the functional plasticity due to genetic or epigenetic modifications (Rapp and Wendel 2005; Salmon *et al.* 2005). Epigenetic variation can generate novel and heritable phenotypic variation in response to environmental or genetic perturbations (Pavet *et al.* 2006; Gehring & Henikoff 2007; Lukens & Zhan 2007). There is strong evidence that methylation of CG dinucleotides and, to a lesser extent, non-CG sites like CNG and CNN sequences (where N is a A, T or C) are crucial for several plant developmental processes (Lee & Chen 2001; Hao *et al.* 2004; Huettel *et al.* 2006; Zhang *et al.* 2006), such as gene imprinting during seed development (Huh *et al.* 2008), vernalization (Finnegan *et al.* 2000; Habu *et al.* 2001) and the response to biotic and abiotic stresses (Madlung & Comai 2004).

The proportion of methylated loci investigated on *L. racemosa* RS population (40.2%) was similar than of *Brassica oleracea* ecotypes, 30-41% (Salmon *et al.* 2008), *Gossypium hirsutum*, 32% (Keyte *et al.* 2006) and *Arabidopsis thaliana*, 35-43% (Cervera *et al.* 2002). On the other hand, SM individuals had rather fewer methylated loci accounting to only 20.6% of total loci investigated, indicating a possible hypomethylation on plants located near the salt marsh. The process of active DNA demethylation occurring on plants is intrinsically associated with biotic or abiotic stresses and affecting also basic genomic functions (Pavet *et al.* 2006). And, in result, the expression of many plant development genes, such as *FLOWERING LOCUS WA (FWA)*, *SUPERMAN (SUP)*, *AGAMOUS (AG)* and *Linaria vulgaris* homologue of *cycloidea (Lcyc)* can be affected (Habu *et al.* 2001). In *Nicotiana tabaccum*, the expression of an enzyme that has demethylation activity was increased on aluminium, salt and cold treated-leaves as a result of epigenetic control based on cytosine methylation (Choi *et al.* 2007).

While the Shannon genetic diversity indexes based on AFLP data were similar for both RS and SM populations with no significant difference, the epigenetic diversity indexes based on MSAP data were significantly higher on RS population (Table 3). More importantly, genetic

versus epigenetic diversity indexes of both populations were significantly different, indicating a truly divergent structure obtained by genetic and epigenetic components of *L. racemosa* genome. This is an important indication that both components have independent evolutionary histories being differently affected by natural selection caused probably by environmental fluctuations and salt stress. While SM plants occurs on a not favorable saline habitat, RS population is located on a more variable habitat, typical of mangrove forests, daily submitted to salt water and nutrient input limitations. These different conditions could have lead SM population to a smaller MSAP-Shannon index as a consequence of an increased number of fixed methylated loci resulted by selection whereas RS maintained high MSAP-Shannon index due to more variable habitat.

Three statistical methods used to investigate the population genetic and epigenetic structure of *L. racemosa* showed congruent estimates (Table 3). All approaches indicated that genetic differentiation among populations was less structured than epigenetic one. The stronger divergence between AFLP and MSAP data was observed with G_{ST} index, which assumes that populations are on Hardy-Weinberg equilibrium. Although this could be true, other statistical analyses that ignore or calculate the deviation from Hardy-Weinberg equilibrium were also assessed using the multivariate and the Bayesian approach respectively, confirming the higher differentiation of the epigenetic component between populations. Since these results are not likely to be explained by stochastic processes such as genetic drift, they suggest that epigenetic modifications arise in *L. racemosa* in response to habitat variation.

The PCA projections of AFLP and MSAP profiles (Figure 1a and 1b) as well as the joint analysis of them by Co-Inertia (Figure 1c) support the higher differentiation of the epigenetic component between RS and SM populations. AFLP data show that plants from different populations present slightly differentiated genotypes. However, MSAP variation was highly contrasted, with convergent epigenetic profiles in the SM population. It was observed that SM

individuals had a propensity to converge into a common point represented by the arrow's directions on figure 1c, what may justify the lower MSAP-Shannon index for SM population. Interestingly, the Co-Inertia analysis had three individuals from each population with very similar genetic profile what could represent plants that recently dispersed from one population to the other. Nevertheless, the six individuals showed highly contrasted epigenetic profile, but correlated to the other samples from the same population. This corroborates the hypothesis that environmental conditions are playing a key role on epigenetic modifications of *L. racemosa* individuals. Even individuals with analogous genotypes collected on each population, and probably closely related, have divergent epigenetic profiles that are more related to the populations which they naturally belonged. This could represent further evidence that some methylation profiles are better adapted to the stressful environmental conditions of salt marsh areas. Moreover, the epigenetic modifications have probably strong influence on phenotypes of these natural populations since divergent epigenetic profiles were found on plants genetically similar but that occurred on different populations.

The direct or indirect relation between epigenetic changes and phenotypic variation may respond to selection, even when unlinked to genetic variation (Lukens & Zhan 2007). Epigenetic changes are heritable and reversible, but unlikely genetic variation they may represent convenient genomic alteration to explore variable environment and new adaptive landscapes. The phenotypic variation between natural populations of *L. racemosa* is probably governed by epigenetic factors without significant genetic changes, which deserves further investigations to uncover the genes controlled by these cytosine methylation changes.

References

- Azhikina TL & Sverdlov ED (2005) Study of Tissue-Specific CpG Methylation of DNA in Extended Genomic Loci. *Biochemistry (Moscow)* 70(5): 596-603.
- Ball MC (2002) Interactive effects of salinity and irradiance on growth: implications for mangrove forest structure along salinity gradients. *Trees* 16:126–139.
- Barth OM, São-Tiago LEU, Barros MA (2006) Paleoenvironment interpretation of a 1760 years B.P. old sediment in a mangrove area of the Bay of Guanabara, using pollen analysis. *Anais da Academia Brasileira de Ciências (Annals of the Brazilian Academy of Sciences)* 78: 227-229.
- Bender J (2000) Plant epigenetics. *Current Biology* 12:R412-4.
- Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology* 16, 3737–3758.
- Bossdorf, O., Richards, C.L., Pigliucci, M. (2008). Epigenetics for ecologists. *Ecology Letters*, 11, 106-115.
- Bussell JD (1999) The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). *Molecular Ecology* 8: 775-789.
- Cardoso MA, Provan J, Powell W, Ferreira PCG, de Oliveira DE (1998) High genetic differentiation among remnant populations of the endangered *Caesalpinia echinata* Lam. (Leguminosae Caesalpinioideae). *Molecular Ecology* 7(5): 601-608.
- Cervera MT, Ruiz-García L, Martínez-Zapater JM (2002) Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Molecular Genetics and Genomics* 268(4): 543-552.
- Charlesworth D (2006) Evolution of plant breeding systems. *Current Biology* 16: R726-R735.

- Choi C-S, Sano H (2007) Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. *Mol Genet Genomics* 277:589–600.
- Dolédec S, Chessel D (1994) Co-inertia analysis: an alternative method for studying species-environment relationships. *Freshwater Biology*, 31, 277-294.
- Finnegan EJ, Peacock WJ, Dennis ES (2000) DNA methylation, a key regulator of plant development and other processes. *Current Opinion in Genetics & Development* 10:217–223
- Gehring M, Henikoff S (2007) DNA methylation dynamics in plant genomes. *Biochimica et Biophysica Acta* 1769: 276–286.
- Habu Y, Kakutani T, Paszkowski J (2001) Epigenetic developmental mechanisms in plants: molecules and targets of plant epigenetic regulation. *Current Opinion in Genetics & Development* 11:215–220
- Hao YJ, Wen XP, Deng XX (2004) Genetic and epigenetic evaluations of citrus calluses recovered from slow-growth culture. *Journal of Plant Physiology* 161(4): 479-484.
- Holsinger KE & Wallace LE (2004) Bayesian approaches for the analysis of population genetic structure: an example from *Platanthera leucophaea* (Orchidaceae). *Molecular Ecology* 13: 887–894.
- Holsinger KE, Lewis PO, Dey DK (2002) A Bayesian approach to inferring population structure from dominant markers. *Molecular Ecology* 11: 1157–1164.
- Huettel B, Kanno T, Daxinger L, Aufsatz W, Matzke AJ, Matzke M (2006) Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. *The EMBO Journal* 25:2826-2836.
- Huh JH, Bauer MJ, Hsieh TF, Fischer RL (2008) Cellular programming of plant gene imprinting. *Cell* 132: 735-745.

- Jablonka E & Lamb MJ (1998) Epigenetic inheritance in evolution. *Journal of Evolutionary Biology* 11: 159-183.
- Jaligot E, Beulé T, Baurens FC, Billotte N, Rival A (2004) Search for methylation-sensitive amplification polymorphisms associated with the mantled variant phenotype in oil palm (*Elaeis guineensis* Jacq). *Genome* 47(1):224-8
- Kalisz S and Purugganan MD (2004) Epialleles via DNA methylation: consequences for plant evolution. *TRENDS in Ecology and Evolution* 19: 309-314
- Keyte AL, Percifield R, Liu B, AND Wendel JF (2006) Intraspecific DNA Methylation Polymorphism in Cotton (*Gossypium hirsutum* L.). *Journal of Heredity* 2006:97(5):444–450
- Landry CL and Rathcke BJ 2007 Do inbreeding depression and relative male fitness explain the maintenance of androdioecy in white mangrove, *Laguncularia racemosa* (Combretaceae)? *New Phytologist* 176: 891–901
- Lee H-S and Chen Z J 2001 Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. *PNAS* 98: 6753-6758.
- Lukens LN and Zhan S (2007) The plant genome's methylation status and response to stress: implications for plant improvement. *Current Opinion in Plant Biology* 2007, 10:317–322
- Madlung A and Comai L (2004) The Effect of Stress on Genome Regulation and Structure. *Annals of Botany* 94: 481–495, 2004
- Parida AK, Das AB, Mitra B (2004) Effects of salt on growth, ion accumulation, photosynthesis and leaf anatomy of the mangrove, *Bruguiera parviflora*. *Trees* (2004) 18:167–174

- Parida AK and Das AB, 2004, Effects of NaCl stress on nitrogen and phosphorous metabolism in a true mangrove *Bruguiera parviflora* grown under hydroponic culture. *Journal of Plant Physiology* 161 (2004) 921–928
- Parisod C & Christin P-A 2008, Genome-wide association to fine-scale ecological heterogeneity within a continuous population of *Biscutella laevigata* (Brassicaceae). *New Phytologist* (2008) 178: 436–447
- Pavet V, Quintero C, Cecchini NM, Rosa AL, Alvarez ME (2006) *Arabidopsis* displays centromeric DNA hypomethylation and cytological alterations of heterochromatin upon attack by *Pseudomonas syringae*. *Molecular Plant-Microbe Interactions* 19(6): 577-587.
- Peraza-Echeverria S, Herrera-Valencia VA, Kay A 2001 Detection of DNA methylation changes in micropropagated banana plants using methylation-sensitive amplification polymorphism (MSAP). *Plant Sci.* 2001 Jul;161(2):359-367.
- Pigliucci M, 2007 Do we need an extended evolutionary synthesis? *Evolution* 61-12: 2743–2749
- Rapp RA and Wendel JF, 2005 Epigenetics and plant evolution. *New Phytologist* 168: 81–91
- Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution* 43:223-225.
- Saenger, P., 2003. *Mangrove Ecology, Silviculture and Conservation*. Kluwer Academic Publishers, The Netherlands.
- Salmon A, Ainouche ML, Wendel JF 2005 Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). *Molecular Ecology* (2005)14, 1163–1175
- Salmon A, Clotault J, Jenczewski E, Chable V, Manzanares-Dauleux MJ 2008 *Brassica oleracea* displays a high level of DNA methylation polymorphism. *Plant Science* 174 (2008) 61–70

- Schones DE & Zhao K 2008 Genome-wide approaches to studying chromatin modifications. *Nature Review Genetics* 9: 179-191
- Schaeffer-Novelli Y, Cintrón-Molero G, Adaime RR and Camargo TM 1990 Variability of mangrove ecosystems along the Brazilian coast. *Estuaries and Coasts* 13: 204-218.
- Schellenbaum P., Mohler V., Wenzel G., Walter B. 2008. Variation in DNA methylation patterns of grapevine somaclones (*Vitis vinifera* L.). *BMC Plant Biology* 8: 78-87.
- Shi S., Huang Y., Zeng K., Tana F., He H., Huang J., Fu Y. 2005. Molecular phylogenetic analysis of mangroves: independent evolutionary origins of vivipary and salt secretion. *Molecular Phylogenetics and Evolution* 34: 159–166.
- Takata M, Kishima Y, Sano Y 2005 DNA methylation polymorphisms in rice and wild rice strains: detection of epigenetic markers. *Breeding Science* 55: 57-63.
- Thioulouse, J., D. Chessel, S. Dolédec, and J. M. Olivier. 1996. ADE-4: a multivariate analysis and graphical display software. *Statistics and Computing* 7:75-83.
- Tomlinson PB (1986) *The Botany of Mangroves*. Cambridge University Press, 413pp.
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23(21): 4407-4414.
- Xiong LZ, Xu CG, Saghai-Maroo MA, Zhang Q (1999) Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. *Molecular Genetics and Genomics* 261(3): 439-446.
- Zhang X 2008 *The Epigenetic Landscape of Plants*. *Science* 320:489-492
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR. 2006 Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 126: 1189-1201

Zhao X, Chai Y, Liu B 2007 Epigenetic inheritance and variation of DNA methylation level and pattern in maize intra-specific hybrids. *Plant Science* 172: 930–938

Figure Legends

Figure 1: Multivariate analysis of the genetic and epigenetic components of riverside (1) and salt marsh (2) populations of *Laguncularia racemosa*: a) Principal Component Analysis (PCA) on covariance matrix for genetic profiles (AFLP); b) PCA on covariance matrix for epigenetic profiles (MSAP); c) Co-Inertia Analysis maximizing the covariance of PCA shown in (a) and (b), and testing for the significance of this association with 9999 permutations. Circles correspond to the projection of genetic (AFLP) profiles and arrow heads indicates the projection of epigenetic (MSAP) profiles. F1 and F2 values show the contribution of two principal components summarizing the total variance of each dataset. β_{ST} was calculated with Between EigenAnalysis (BPCA) for both genetic and epigenetic profiles and tested with 9999 permutations.

Table 1: AFLP and MSAP primer sequences.

Enzyme	Pre-amplification	Selective amplification
EcoRI	5'-GACTGCGTACCAATTC-3' (Vos <i>et al.</i> 2005)	E+AG
		E+AC
		E+AAC
MspI/HpaII	5'-ATCATGAGTCCTGCTCGG-3' (Xiong <i>et al.</i> 1999)	HM+TCAA
		HM+AAT

Table 2: Polymorphic and monomorphic methylated loci obtained comparing AFLP and MSAP profiles, polymorphic AFLP loci and polymorphic MSAP loci observed within 97 loci on samples of two *L. racemosa* natural populations (RS and SM). Loci with presence of AFLP fragment but absence of MSAP fragment were considered methylated. Loci with fragment difference on more than one sample were considered polymorphic.

Population	Methylated loci			Polymorphic	Polymorphic
	Polymorphic	Fixed	Total	AFLP	MSAP
RS	33	6	39 (40.2%)	15	37
SM	11	9	20 (20.6%)	10	13

Table 3: Within- and among-population diversity estimates for *Laguncularia racemosa* using genetic (AFLP) and epigenetic (MSAP) data. Values with different uppercase letters are significantly different.

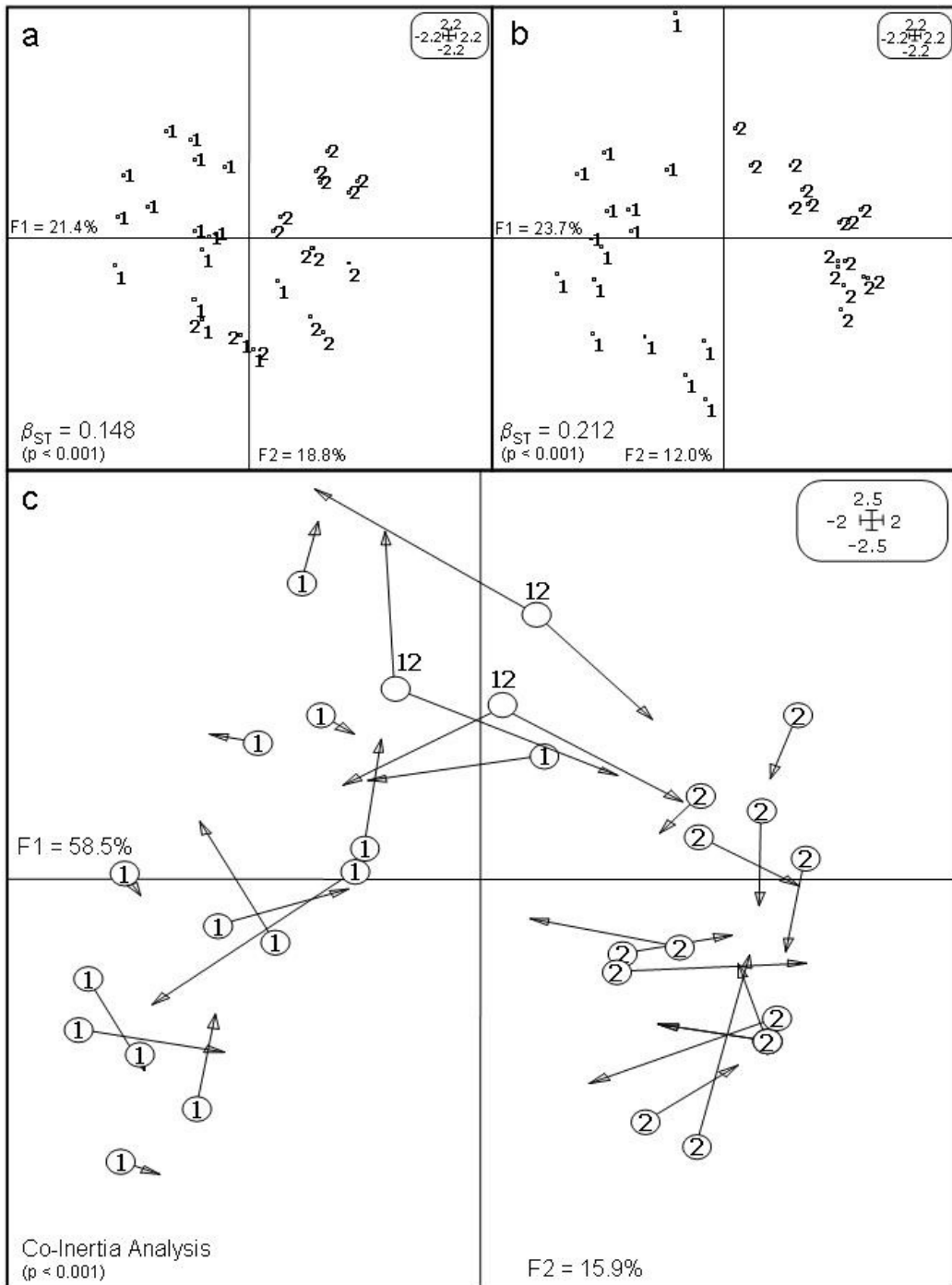
G_{ST} – frequency-based estimate from Shannon diversity index.

β_{ST} – band-based estimate from BPCA (multivariate analysis).

F_{ST} – F-estimate from Bayesian analysis.

Data	Population	Shannon			
		diversity	G_{ST}	β_{ST}	F_{ST}
AFLP	RS	0.048 ^a	0.111	0.148	0.119
	SM	0.035 ^a			
MSAP	RS	0.101 ^b	0.254	0.212	0.160
	SM	0.043 ^a			

Figure 1:



Anexo IV

Uso de microssatélites em estudos de biologia da conservação

Catarina da Fonseca Lira de Medeiros*^{1,2}, Monica Aires Cardoso¹,
Paulo Cavalcanti Gomes Ferreira²

*Laboratório de Biologia Molecular de Plantas, Instituto de Pesquisa Jardim Botânico do Rio de Janeiro,
Rio de Janeiro, Brasil¹*

Instituto de Bioquímica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil²

catilira@bioqmed.ufrj.br*

Recebido em 20 de outubro de 2006

Resumo

Os níveis de variabilidade genética populacionais podem ser determinados por marcadores moleculares, como RAPD e AFLP usados em estudos de fingerprinting, ou microssatélites, que são bastante utilizados em estudos de estrutura genética populacional, respondendo a questões ecológicas essenciais para a adoção de medidas conservacionistas. Atualmente são conhecidos microssatélites de regiões codificantes do genoma (SSR-EST), que têm como vantagem a transferabilidade entre espécies próximas em contraste aos microssatélites de regiões anônimas. Esta revisão visa divulgar esta nova modalidade de microssatélites, os SSR-ESTs, que estão sendo cada vez mais utilizados como ferramenta molecular eficiente para estudos populacionais com enfoque em conservação.

Palavras-chaves: SSR, marcadores moleculares, genética de populações

Use of microsatellite markers in conservation biology studies

Abstract

The levels of genetic population variability can be determined by molecular markers, as RAPD and AFLP in fingerprinting studies, or microsatellites, that are used in studies of population genetic structure, answering ecological questions that are essential for the adoption of conservation measures. Nowadays coding region microsatellites were discovered (SSR-EST), and they have the advantage of transferability between related species in contrast to the microsatellites of anonymous regions. This review aims the spreading of this new modality of microsatellites, SSR-ESTs, which are being more and more used as an efficient molecular tool for population studies with approach on conservation.

Key-words: SSR, molecular markers, population genetics

Introdução aos microssatélites

Dentre os diferentes níveis hierárquicos de biodiversidade, uma importância especial tem sido dado à

diversidade genética, que é a base desta hierarquia. Ela geralmente é observada ao nível de espécie, dentro de uma mesma população ou entre diferentes populações. A diversidade genética tem papel

importante na conservação de espécies ameaçadas e na manutenção de suas populações ao longo do tempo, pois espécies com baixa variação genética têm geralmente uma redução na habilidade de sobreviver a mudanças ambientais durante o seu processo evolutivo (Frankham, 1995)

As populações naturais possuem em geral altos níveis de variabilidade genética intrapopulacional, que é introduzida continuamente por mutação, migração ou fluxo gênico (Morand *et al.*, 2002), com exceção das espécies que têm reprodução preferencial por endogamia, o que contribui bastante para a diminuição da variabilidade genética em suas populações. Vários estudos já comprovaram que o fluxo gênico em florestas tropicais pode alcançar longas distâncias, sendo importante na manutenção da variabilidade genética (White *et al.*, 2002; Chase *et al.*, 1996), que é fundamental para a manutenção e sobrevivência da espécie e de suas populações, às ameaças causadas principalmente pelo homem (Heuertz *et al.*, 2001).

Tradicionalmente, a diversidade dentro de uma espécie era medida apenas pelas diferenças morfológicas, mas atualmente os métodos de detecção, baseados em dados moleculares, vêm sendo utilizados e considerados como vantajosos sobre os métodos antigos devido ao maior número de caracteres estudados (Ferreira & Grattapaglia, 1998; Karp *et al.*, 1996).

Uma das técnicas mais importantes desenvolvida para estudos de marcadores moleculares é o PCR (Polymerase Chain Reaction – Reação de Polimerase em Cadeia) que consiste numa reação em cadeia por ação da enzima DNA polimerase que copia fragmentos de DNA, desde que exposta a uma fita de DNA-molde, desoxirribonucleotídeos e oligonucleotídeos usados como iniciadores pela enzima polimerásica. O PCR foi inventado em 1983 por Kary Mullis, com base no seu conhecimento sobre seqüenciamento e oligonucleotídeos e, em 1987, junto com o matemático Fred Faloona foi possível provar que o PCR era verossímil e bastante eficiente, marcando o início da era genômica (Mullis & Faloona, 1987).

Dentre os seis principais marcadores moleculares, três não se utilizam do PCR – isoenzimas, RFLP

(Restricted Fragment Length Polymorphism – Polimorfismo por Tamanho de Fragmento de Restrição) e minissatélites – e por isso possuem limitações quanto ao número de indivíduos analisados por estudo, principalmente por causa da grande quantidade de material genético necessário e do excessivo trabalho laboratorial com resultados demorados e custosos (Ferreira & Grattapaglia, 1998).

Por outro lado, os outros três são baseados em PCR – RAPD (Random Amplified DNA – DNA Amplificado ao Acaso), AFLP (Amplified Fragment Length Polymorphism – Polimorfismo por Tamanho de Fragmento Amplificado) e microssatélites – e permitem a obtenção do resultado mais rapidamente, além de possibilitar estudos populacionais com grande número de indivíduos e também outros tipos de estudo como de genética humana, forenses, de biologia evolutiva e de biologia do desenvolvimento, devido à sensibilidade, velocidade e versatilidade do PCR (Ferreira & Grattapaglia, 1998; White *et al.*, 1989; Williams *et al.*, 1990).

Marcadores dominantes como RAPD e AFLP são usados para fazer análises de *fingerprinting* em diferentes organismos, inclusive plantas como *Eucalyptus* (Kirst *et al.*, 2005), *Caesalpinia echinata* (Cardoso *et al.*, 1998; Cardoso *et al.*, 2005), *Euterpe edulis* (Cardoso *et al.*, 2000), *Veronica alpina* (Albach *et al.*, 2006), *Viola sp.* (Eckstein *et al.*, 2006), *Cedrus atlantica* (Renau-Morata *et al.*, 2005), *Croton alabamensis* (Van EE *et al.*, 2006), com o objetivo de avaliar a variação genética e a estruturação populacional. Estas ferramentas são bastante eficientes em estudos com espécies selvagens, pois não necessitam de conhecimento prévio de seqüência de DNA e permitem a análise de uma grande quantidade de indivíduos. Além disso, são técnicas com baixos custos de desenvolvimento e operacional e que necessitam de pouca quantidade de DNA. Porém, uma desvantagem dessas técnicas é de não possuírem definição suficiente para estudos de parentesco ou mapeamento genético, por não possibilitarem a visualização dos alelos por locus, ou seja, a determinação de heterozigotos e homozigotos.

Os microssatélites ou SSR (Simple Sequence

Repeats – Repetições Simples de Sequência) são repetições em tandem de 1 a 6 nucleotídeos, encontrados em todos os procariotos e eucariotos estudados até o momento (Zane *et al.*, 2002). Os microssatélites foram encontrados no genoma de eucariotos no final da década de 80 (Tautz, 1989) e logo os pesquisadores perceberam os altos níveis de polimorfismo deste marcador, comprovando que se tratava de uma nova ferramenta molecular bastante eficiente, com pelo menos o dobro de informação obtida duas vezes mais rapidamente do que outros marcadores (Weber, 1990).

Um dos trabalhos mais significativos da época, provou a hipervariabilidade destes marcadores em *Drosophila melanogaster* e em baleias e a sua hereditariedade no homem (Tautz, 1989). Já na década de 90, trabalhos encontraram estes marcadores moleculares no genoma de plantas ocorrendo com frequência similar à encontrada no genoma de vertebrados (Akkaya *et al.*, 1992). Atualmente, sabemos que existem microssatélites tanto no genoma nuclear das plantas (Condit & Hubbell, 1991; Wang *et al.*, 1994) como no genoma de cloroplastos (Powel *et al.*, 1995).

A taxa de mutação do genoma de cloroplasto é baixa, o que era uma barreira para sua utilização em estudos populacionais até serem descobertos microssatélites de repetições mononucleotídicas ao longo deste genoma (Provan *et al.*, 2001). Acredita-se que os marcadores microssatélites de cloroplasto sejam uma ótima ferramenta em estudos de fluxo gênico e padrões de cruzamento (Provan *et al.*, 2001; Provan, 2000; Vendramin *et al.*, 2000). Além de estudos de avaliação da estrutura genética de populações vegetais e modo de herança do cloroplasto (Weising & Gardner, 1999) em gimnospermas *Picea abies* K (Vendramin *et al.*, 2000), *Abies sp* (Clark *et al.*, 2000; Vendramin *et al.*, 1999) e em angiospermas *Silene paradoxa* L (Mengoni *et al.*, 2001), *Fraxinus sp* (Morand-Prieur *et al.*, 2002) e *Caesalpinia echinata* (Lira *et al.*, 2003).

Os marcadores microssatélites tornaram-se extremamente difundidos na biologia, apesar do alto custo de desenvolvimento, sendo atualmente utilizados em

um grande número de estudos para investigar a estrutura genética de populações, respondendo a perguntas específicas nas áreas da biologia evolutiva e da biologia da conservação (Balloux & Lugon-Moulin, 2002). Isto se dá devido a três características presentes nos microssatélites: (1) a neutralidade, pois representam o genoma inteiro; (2) a abundância com ampla distribuição pelo genoma; e (3) o alto polimorfismo encontrado mesmo em populações que passaram por eventos como gargalos ou em populações que possuem baixo polimorfismo detectado por outros marcadores (Maudet *et al.*, 2002).

Microssatélites são marcadores codominantes, ou seja, podem ser observados ambos os alelos presentes no mesmo *locus*, possibilitando a sua aplicação em estudos genéticos e ecológicos através da detecção da estrutura genética de populações isoladas e fragmentadas, possibilitando análises do sistema reprodutivo de plantas e animais e suas conseqüências como porcentagem de heterozigosidade nas populações (Morand *et al.*, 2002; Collevatti *et al.*, 2001; Heuertz *et al.*, 2001; White *et al.*, 1999). Através dos oligonucleotídeos complementares às seqüências flanqueadoras das repetições dos microssatélites, são obtidos produtos de tamanhos variados correspondendo às expansões e/ou contrações da região repetitiva (Provan *et al.*, 2001).

A maior limitação ao uso de marcadores microssatélites é o grande trabalho necessário para o desenvolvimento e isolamento dos *loci* que contêm estes marcadores (Zane *et al.*, 2002), principalmente para plantas, cujo número de repetições de dinucleotídeos é 10 vezes menor do que nos primatas. Para contornar esta situação foram desenvolvidos diferentes métodos de enriquecimento de bibliotecas genômicas (White & Powell, 1997; Edwards *et al.*, 1996; Ostrander *et al.*, 1992). Zane *et al.* (2002) indicam diferentes tipos de isolamento de microssatélites, porém todos são bastante custosos, o que gera um empecilho para se trabalhar com estes marcadores moleculares. Após seu isolamento, faz-se necessário tornar os iniciadores desenvolvidos para a amplificação via PCR dos *loci* microssatélites funcionais através da otimização das condições ideais do próprio PCR. Este segundo passo pode tornar-se tão problemático quanto o primeiro

Tabela 1. Ferramentas para busca de microssatélites em bancos de dados de ESTs. Adaptado de Varshney *et al.* (2005).

Table 1. Tools for microsatellite research in EST’s databases. Adapted from Varshney *et al.* (2005).

Programa	Ref.
MicroSATellite (MISA)	http://pgrc.ipk-gatersleben.de/misa/ ; Thiel <i>et al.</i> (2003)
SSRFinder	Gao <i>et al.</i> (2003)
BuildSSR	Rungis <i>et al.</i> (2004)
SSR Identification Tool (SSRIT)	Kantety <i>et al.</i> (2002)
Tandem Repeat Finder (TRF)	Benson (1999)
Tandem Repeat Occurrence Locator (TROLL)	Castelo <i>et al.</i> (2002)
CUGIssr	http://www.genome.clemson.edu/projects/ssr/
Sputnik	http://abajian.net/sputnik/index.html
Modified Sputnik	Morgante <i>et al.</i> (2002)
Modified Sputnik II	http://wheat.pw.usda.gov/ITMI/EST-SSR/LaRota/
SSRSEARCH	ftp://ftp.gramene.org/pub/gramene/software/scripts/ssr.pl

passo de isolamento, em termos de dinheiro e tempo gasto (Squirrell *et al.*, 2003).

Com o estabelecimento de bancos de dados de seqüências expressas do genoma (EST – Expressed Sequence Tags – Partes de Seqüências Expressas), uma nova modalidade de microssatélites foi descoberta, os microssatélites gênicos ou EST-SSR. O desenvolvimento destes marcadores é relativamente fácil e tem baixo custo porque são utilizadas seqüências disponibilizadas na internet e programas de computador distribuídos gratuitamente para isolar o microssatélite gênico e desenhar os oligonucleotídeos iniciadores do PCR (Varshney *et al.*, 2005). As principais vantagens desta técnica segundo Woodhead *et al.* (2005) são: 1) regiões codificantes possuem menor taxa de mutação o que diminuiria a ocorrência de alelos nulos e ainda permitiria uma maior transferabilidade destes marcadores entre espécies; 2) o baixo nível de mutação resultaria também em menor homoplasia; 3) alguns trabalhos com espécies cultivadas obtiveram fragmentos mais limpos com bandas melhor definidas do que aquelas obtidas com os microssatélites genômicos.

Microssatélites em regiões não-codificantes e codificantes

O método tradicional para o desenvolvimento de marcadores microssatélites envolve a criação de

bibliotecas genômicas de inserto pequeno, a posterior hibridização com oligonucleotídeos repetidos em tandem e o seqüenciamento de clones candidatos, tornando assim o processo lento, trabalhoso e caro (Thiel *et al.*, 2003).

EST (Expressed Sequence Tags – Partes de Seqüências Expressas) são porções seqüenciadas de cDNA (DNA complementar), que são cópias do mRNA (RNA mensageiro). Por isso, eles representam parte da porção transcrita do genoma em certas condições e são bastante conservados entre espécies próximas (Poncet *et al.*, 2006). O aumento no número de banco de dados de ESTs disponíveis permitiu a busca de microssatélites derivados de ESTs (SSR-EST), diminuindo assim o tempo e dinheiro necessários para o seu isolamento (Thiel *et al.*, 2003).

O Instituto TIGR (The Institute for Genomic Research – www.tigr.org) possui um banco de dados de ESTs com 232 espécies de plantas, dentre elas estão 10 coníferas, 150 dicotiledôneas e 41 monocotiledôneas. A grande maioria destas espécies é de importância econômica, e por isso é difícil encontrar seqüências de espécies selvagens que não são cultivadas ou não têm interesse econômico cujo estudo visa apenas a sua conservação.

Para a identificação dos microssatélites derivados de ESTs, alguns programas foram criados na linguagem Pearl como o MISA (MicroSATellite) que é um

Tabela 2. Resumo de algumas questões ecológicas que podem ser respondidas com marcadores genéticos neutros, separados por tipo de dados necessário. Adaptado de Selkoe & Toonen (2006).
Table 2. Summary of some ecological questions that can be answered with neutral genetic markers, separated for the type of data necessary. Adapted from Selkoe & Toonen (2006).

Necessita de dados de frequência alélica multi-locus*

- De qual população estes indivíduos se originaram?
- Quantas populações existem?

Necessita de dados de microssatélites ou alto polimorfismo de sequência**

- Esta população se expandiu ou se contraiu nos últimos anos?
- As populações são diferentes em tamanho no passado e no presente?

Necessita de identificação genotípica multi-locus***

- Qual a relação genética entre os indivíduos?
- Quais indivíduos migraram?
- Quais indivíduos são clones?

Pode ser feito com diferentes marcadores moleculares**

- Qual a distância de dispersão média dos descendentes (ou gametas)?
 - Quais são as relações fonte-sumidoro entre as populações?
 - Como condições do meio-ambiente causam impacto na estrutura populacional e migração?
 - Qual a dinâmica de extinção e recolonização da metapopulação?
 - A estrutura populacional ou sua conectividade mudou nos últimos anos?
-

*Estas análises podem necessitar de >10 microssatélites – o número está inversamente correlacionado com o grau de diferenciação genética entre as populações. Espécies com baixa taxa de migração e/ou populações pequenas necessitam de menos *loci*.

**Usando >1 *locus* diminuirá substancialmente erro amostral interlocus.

***Geralmente necessita de microssatélites, mas também é possível com AFLP e RAPD

módulo de busca com várias características úteis para o controle de qualidade das seqüências de ESTs e para o desenvolvimento de iniciadores que amplificarão os *loci* contendo microssatélites (Varshney *et al.*, 2005). Na Tabela 1 estão algumas destas ferramentas de busca.

A avaliação de germoplasmas usando microssatélites derivados de ESTs pode aumentar o poder dos marcadores moleculares permitindo examinar a diversidade funcional em plantas fenotipicamente bem caracterizadas, pois expansões e contrações das repetições destes SSR-ESTs de genes de função conhecida podem ser relacionadas com variação fenotípica ou, melhor ainda, com função biológica (Eujayl *et al.*, 2002).

Com relação às plantas, os microssatélites isolados a partir de seqüências ESTs estão sendo bastante utilizados em espécies cultivadas como trigo, cevada,

café, laranja, amêndoa e girassol (Gao *et al.*, 2004; Thiel *et al.*, 2003; Poncet *et al.*, 2006; Chen *et al.*, 2006; Xie *et al.*, 2006; Pashley *et al.*, 2006).

Uma das grandes vantagens destes marcadores provenientes de regiões expressas do genoma é a transferabilidade entre espécies relacionadas em contraste aos microssatélites de regiões anônimas, refletindo a natureza conservadora das regiões codificantes quando comparada às regiões não-codificantes (Varshney *et al.*, 2005b). A posição dos microssatélites gênicos com relação à região codificante resulta em diferentes níveis de polimorfismo, onde as regiões 3'UTR (Untranslated Regions – Regiões Não-Traduzidas) são mais polimórficas no nível de cultivares, as regiões 5'UTR são mais polimórficas entre cultivares e espécies, e a seqüência codificante do gene é mais polimórfica entre espécies e gêneros (Varshney *et al.*, 2005).

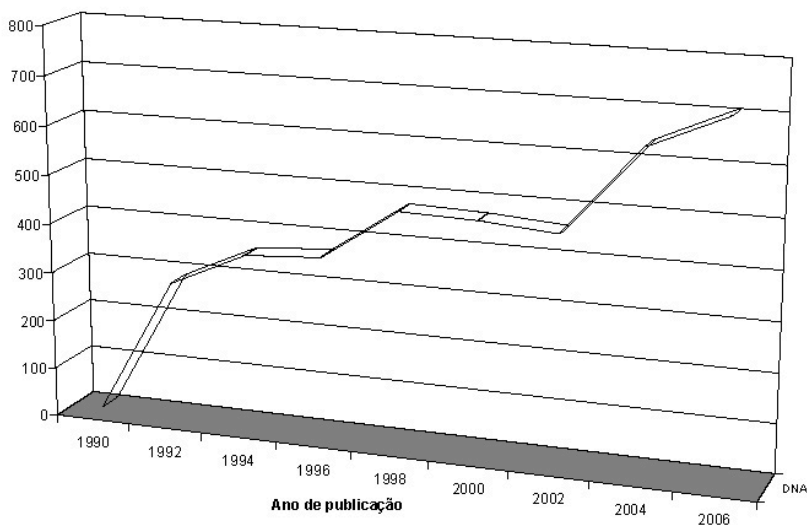


Figura 1. Número de artigos referente à pesquisa realizada pela ferramenta Web of Science do site ISI Web of Knowledge [v3.0] (<http://isiknowledge.com>) utilizando como palavras-chaves conservação (conservation) e DNA.

Figure 1: The number of articles resulting from the research made by the tool search Web of Science of the site ISI Web of Knowledge [v3.0] (<http://isiknowledge.com>) using the keywords conservation and DNA.

Aplicações dos microssatélites para conservação

Segundo Liu & Cordes (2004), os marcadores moleculares AFLP e microssatélites estão ainda em fase exponencial de crescimento com relação ao número de trabalhos publicados utilizando estas ferramentas e por isso eles parecem ser os principais atores na revolução genômica.

Devido à alta variabilidade dos microssatélites e à possibilidade de se obter marcadores mendelianos multi-alélicos, eles são especialmente úteis para inferir eventos demográficos recentes, inclusive detecção de impactos induzidos pelas comunidades humanas nas populações (Pearse & Crandall, 2004). Os microssatélites emergiram como uma das opções mais populares para estudos genéticos com questões ecológicas, em parte porque eles possuem potencial para determinar estimativas de migração, têm poder para distinguir relativamente altas taxas de migração para panmixia, e podem estimar a relação de parentesco entre os indivíduos (Selkoe & Toonen, 2006).

Os principais marcadores moleculares citados na introdução, que são considerados marcadores neutros, podem ser utilizados em diferentes tipos de

estudos ecológicos com o objetivo de responder a perguntas importantes sobre a biologia da espécie e suas populações. Na Tabela 2, temos algumas destas questões ecológicas e o tipo de dados moleculares que são necessários para respondê-las.

Dentre as principais aplicações dos microssatélites estão mapeamento genético, identificação individual por DNA e estudos de parentesco, filogenia, genética de populações e genética da conservação, epidemiologia e patologia molecular, mapeamento de QTLs (Quantitative Trait Loci – Loci de Traços Quantitativos), e seleção assistida por marcador (Chistiakov *et al.*, 2006).

Os microssatélites vêm sendo muito utilizados para responder várias perguntas relacionadas à genética de populações, como análises de fluxo gênico, paternidade e estruturação populacional, que resultam em dados sobre a distribuição da variabilidade genética entre e dentro de populações naturais, que são essenciais para a adoção de medidas de conservação tanto *ex situ* quanto *in situ* (Oliveira *et al.*, 2006).

A obtenção de dados sobre a variabilidade genética permite um melhor entendimento de processos

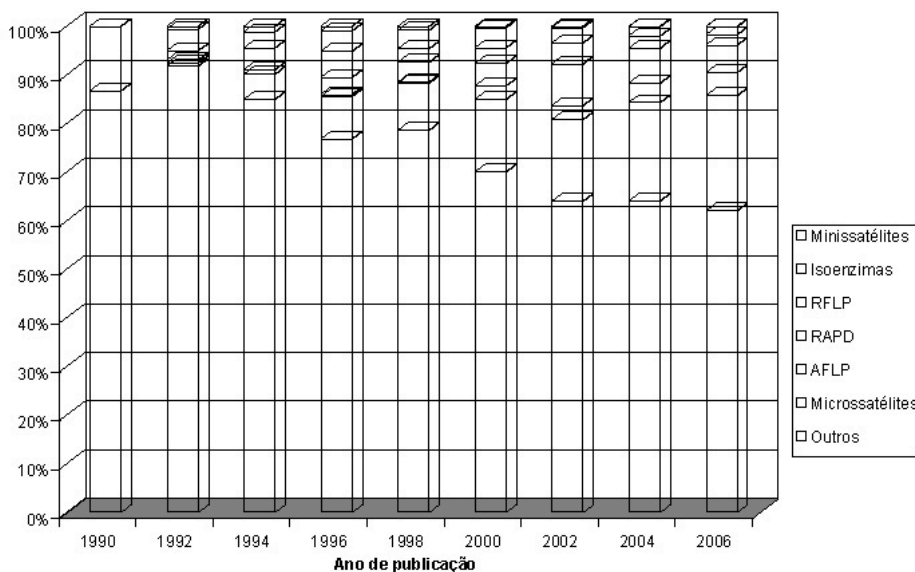


Figura 2. Número de artigos referente à pesquisa realizada pela ferramenta Web of Science do site ISI Web of Knowledge [v3.0] (<http://isiknowledge.com>) utilizando como palavras-chaves DNA, microsatélites (microsatellite), AFLP, RAPD, RFLP, isoenzimas (isozyme), minissatélites (minisatellite), juntamente com a palavra conservação (conservation).

Figure 2. The number of articles resulting from the research made by the tool search Web of Science of the site ISI Web of Knowledge [v3.0] (<http://isiknowledge.com>) using the keywords DNA, microsatellite, AFLP, RAPD, RFLP, isozyme, minisatellite, all with the word conservation.

populacionais como dispersão e migração, história demográfica de populações como gargalos e efeito fundador, detecção de híbridos entre espécies, fluxo gênico e deriva gênica, filogeografia, ecologia de paisagem como distribuição, movimentos e dispersão dos animais, e ecotoxicologia (deYoung & Honeycutt, 2005).

Publicações atuais sobre conservação

O número de publicações utilizando ferramentas moleculares em estudos de conservação é cada vez maior, de apenas 13 artigos em 1990 aumentou para 710 artigos em 2006. Na figura 1, o gráfico com a evolução do número de artigos publicados bianualmente nos últimos 16 anos mostra três saltos significativos quanto ao número de publicações utilizando o DNA como ferramenta: de 1990 para 1992, com a diferença de 276 artigos; de 1996 para 1998, com diferença de 106 artigos; e de 2002 para 2004, com diferença de

178 artigos. Uma possível explicação para este rápido incremento na quantidade de artigos publicados por ano, é a implementação e o desenvolvimento de novas técnicas que, além de aumentarem a eficiência dos procedimentos, tornaram-se capazes de reduzir seus custos.

De todas estas publicações com DNA, as mais significativas com relação à conservação utilizam microsatélites como ferramenta. E os três principais marcadores moleculares baseados em PCR, RAPD, AFLP e os próprios microsatélites, são os mais representativos, enquanto que os outros marcadores, RFLP, isoenzimas e minissatélites, têm menor difusão para este tipo de estudo.

A Figura 2 mostra a tendência de uma expansão na utilização dos marcadores microsatélites em estudos de conservação, seguidos de forma menos expressiva pelos marcadores AFLP e RAPD. Além disso, pode ser ainda observado o uso cada vez menor dos marcadores RFLP, isoenzimas e minissatélites, devido às

suas duas maiores desvantagens: grande quantidade de material necessário e trabalho laboratorial intenso e custoso.

Conclusão

Os microssatélites são ferramentas bastante robustas e importantes para estudos genéticos de diversidade visando a conservação das espécies (Selkoe & Toonen, 2006). Um dos limitantes de uso desta ferramenta é justamente o isolamento dos marcadores microssatélites a partir de DNA genômico em bibliotecas enriquecidas com repetições de di ou trinucleotídeos (Squirrel *et al.*, 2003).

Com o surgimento de bancos de dados de ESTs, que são seqüências codificantes do genoma, e com o constante aumento no número de espécies estudadas e seqüências depositadas, foi possível obter microssatélites presentes nestas regiões do genoma através de uma busca relativamente simples dentro dos bancos de dados. Muitos programas já foram desenvolvidos com o objetivo de facilitar o isolamento destes marcadores (Varshney *et al.*, 2005).

Atualmente a maioria das publicações utilizando-se desta ferramenta visa o mapeamento genético e o estudo da diversidade funcional de espécies cultivadas ou com importância econômica, sendo os microssatélites de regiões anônimas os mais importantes em estudos de *fingerprinting* e identificação parietal (Varshney *et al.*, 2005). No entanto, os SSR-ESTs representam uma ferramenta adicional para geneticistas populacionais de plantas ou ecologistas moleculares, pois trabalhos examinando a sua transferibilidade entre espécies e os resultados com populações naturais indicam que estes marcadores são muito mais estáveis e fáceis de amplificar do que os microssatélites de regiões anônimas do genoma (Woodhead *et al.*, 2005; Ellis *et al.*, 2006).

Um limitante dos microssatélites gênicos seria justamente a possível seleção que este marcador sofreria, mas Woodhead *et al.* (2005) perceberam que a diferenciação genética não estava sendo afetada pela seleção, pois os valores de F_{st} – definido por Pearse & Crandall (2004) como o parâmetro padrão usado

para descrever o nível de diferenciação entre subpopulações – baseados em SSR-EST, SSR genômico e AFLP foram parecidos. Aparentemente, portanto, os microssatélites derivados de ESTs parecem ser geralmente neutros podendo ser usados em estudos de efeito demográfico sobre o nível de variação genética, muito comuns em estudos de conservação de espécies (Ellis *et al.*, 2006).

Bibliografia

ALBACH DC, SCHÖNSWETTER, P.; TRIBSCH, A. Comparative phylogeography of the *Veronica alpina* complex in Europe and North America. **Molecular Ecology**, V.15, n.11, p.3269-3286, 2006.

AKKAYA, M.S.; BHAGWAT, A.A.; CREGAN P.B. Length polymorphisms of simple sequence repeat DNA in Soybean. **Genetics**, V.132, n.4, p.1131-1139, 1992.

BALLOUX, F. ; LUGON-MOULIN, N. The estimation of population differentiation with microsatellite markers. **Molecular Ecology**, V.11, n.2, p.155-165, 2002.

BENSON, G. Tandem repeats finder: a program to analyze DNA sequences. **Nucleic Acids Research**, V.27, n.2, p.573-580, 1999 **apud** VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, V.23, n.1, p.48-55, 2005.

CARDOSO, M.A.; PROVAN, J.; POWELL, W.; FERREIRA, P.C.G.; DE OLIVEIRA, D.E. High genetic differentiation among remnant populations of the endangered *Caesalpinia echinata* Lam. (Leguminosae Caesalpinioideae). **Molecular Ecology**, V.7, n.5, p.601-608, 1998.

CARDOSO, S.R.S.; ELOY, N.B.; PROVAN, J.; CARDOSO, M.A.; FERREIRA, P.C.G. Genetic differentiation of *Euterpe edulis* Mart. populations estimated by AFLP analysis. **Molecular Ecology**,

V.9, n.11, p.1753-1760, 2000.

CARDOSO, S.R.S.; PROVAN, J.; LIRA, C.F.; PEREIRA, L.O.R.; FERREIRA, P.C.G.; CARDOSO, M.A. High levels of genetic structuring as a result of population fragmentation in the tropical tree species *Caesalpinia echinata* Lam. **Biodiversity and Conservation**, V.14, n.5, p.1047-1057, 2005.

CASTELO, A.T. ; MARTINS, W. ; GAO, G.R. TROLL – Tandem Repeat Occurrence Locator. **Bioinformatics**, V.18, n.4, p.634-636, 2002 **apud** VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, V.23, n.1, p.48-55, 2005.

CHASE, M.R.; MOLLER, C.; KESSELI, R.; BAWA, K.S. Distant gene flow in tropical trees. **Nature**, V.383, n.6599, p.398-399, 1996.

CHEN, C.; ZHOU, P.; CHOI, Y.A.; HUANG, S.; GMITTER, F.G. Mining and characterizing microsatellites from citrus ESTs. **Theoretical and Applied Genetics**, V.112, n.7, p.1248-1257, 2006.

CHISTIAKOV, D.A.; HELLEMANS, B.; VOLCKAERT, F.A.M. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. **Aquaculture**, V.255, n.1-4, p.1-29, 2006.

CLARK, C.M.; WENTWORTH, T.R.; O'MALLEY, D.M. Genetic discontinuity revealed by chloroplast microsatellites in eastern north american *Abies* (Pinaceae). **American Journal of Botany**, V.87, n.6, p.774-782, 2000.

COLLEVATTI, R.G.; GRATTAPAGLIA, D.; HAY, J.D. High resolution microsatellite based analysis of the mating system allows the detection of significant biparental inbreeding in *Caryocar brasiliense*, an endangered tropical tree species. **Heredity**, V.86, n.1, p.60-67, 2001.

CONDIT, R.; HUBBELL, S.P. Abundance and DNA

sequence of two-base repeat regions in tropical tree genomes. **Genome**, V.34, n.1, p.66-71, 1991.

DEYOUNG, R.W.; HONEYCUTT, R.L. The molecular toolbox: genetic techniques in wildlife ecology and management. **Journal of Wildlife Management**, V.69, n.4, p.1362-1384, 2005.

ECKSTEIN, R.L.; O'NEILL, R.A.; DANIHELKA, J.; OTTE, A.; KÖHLER, W. Genetic structure among and within peripheral and central populations of three endangered floodplain violets. **Molecular Ecology**, V.15, n.9, p.2367-2379, 2006.

EDWARDS, K.J.; BARKER, J.H.A.; DALY, A.; JONES, C.; KARP, A. Microsatellite libraries enriched for several microsatellite sequences in plants. **Biotechniques**, V.20, n.5, p.758-760, 1996.

ELLIS, J.R.; PASHLEY, C.H.; BURKE, J.M.; MCCAULEY, D.E. High genetic diversity in a rare and endangered sunflower as compared to a common congener. **Molecular Ecology**, V.15, n.9, p.2345-2355, 2006.

EUIJAYL, I.; SORRELLS, M.E.; BAUM, M.; WOLTERS, P.; POWELL, W. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. **Theoretical and Applied Genetics** V.104, n.2-3, p.399-407, 2002.

FRANKHAM, R. Conservation genetics. **Annual Review of Genetics**, V.29, p.305-327, 1995.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao Uso de Marcadores Moleculares em Análise Genética**. 3ª ed. Brasília: EMBRAPA-CENARGEN, 1998.

GAO, L.; TANG, J.; LI, H.; JIA, J. Analysis of microsatellites in major crops assessed by computational and experimental approaches. **Molecular Breeding**, V.12, n.3, p.245-261, 2003 **apud** VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, V.23, n.1, p.48-55,

2005.

GAO, L.F.; JING, R.L.; HUO, N.X.; LI, Y.; LI, X.P.; ZHOU, R.H.; CHANG, X.P.; TANG, J.F.; MA, Z.Y.; JIA, J.Z. One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. **Theoretical and Applied Genetics**, V.108, n.7, p.1392-1400, 2004.

HEUERTZ, M.; HAUSMAN, J.F.; TSVETKOV, I.; FRASCARIA-LACOSTE, N.; VEKEMANS, X. Assessment of genetic structure within and among Bulgarian populations of the common ash (*Fraxinus excelsior* L.). **Molecular Ecology**, V.10, n.7, p.1615-1623, 2001.

KANTETY, R.V.; LA ROTA, M.; MATTHEWS, D.E.; SORRELLS, M.E. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. **Plant Molecular Biology**, V.48, n.5-6, p.501-510, 2002 **apud** VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, V.23, n.1, p.48-55, 2005.

KARP, A.; SEBERG, O.; BUIATTI, M. Molecular techniques in the assessment of botanical diversity. **Annals of Botany**, V.78, n.2, p.143-149, 1996.

KIRST, M.; CORDEIRO, C.M.; REZENDE, G.D.S.P.; GRATTAPAGLIA, D. Power of microsatellite markers for fingerprinting and parentage analysis in *Eucalyptus grandis* breeding populations. **Journal of Heredity**, V.96, n.2, p.161-166, 2005.

LIRA, C.F.; CARDOSO, S.R.S.; FERREIRA, P.C.G.; CARDOSO, M.A.; PROVAN, J. Long-term population isolation in the endangered tropical tree species *Caesalpinia echinata* Lam. revealed by chloroplast microsatellite. **Molecular Ecology**, V.12, n.12, p.3219-3225, 2003.

LIU, Z.J.; CORDES, J.F. DNA marker technologies and their applications in aquaculture genetics.

Aquaculture, V.238, n.1-4, p.1-37, 2004.

MAUDET, C.; MILLER, C.; BASSANO, B.; BREITENMOSER-WÜRSTEN, C.; GAUTHIER, D.; OBEXER-RUFF, G.; MICHALLET, J.; TABERLET, P.; LUIKART, G. Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [*Capra ibex (ibex)*]. **Molecular Ecology**, V.11, n.3, p.421-436, 2002.

MENGGONI, A.; BARABESI, C.; GONNELLI, C.; GALARDI, F.; GABBRIELLI, R.; BAZZICALUPO, M. Genetic diversity of heavy metal-tolerant populations in *Silene paradoxa* L. (Caryophyllaceae): a chloroplast microsatellite analysis. **Molecular Ecology**, V.10, n.8, p.1909-1916, 2001.

MORAND, M.E.; BRACHET, S.; ROSSIGNOL, P.; DUFOUR, J.; FRASCARIA-LACOSTE, N. A generalized heterozygote deficiency assessed with microsatellites in French common ash populations. **Molecular Ecology**, V.11, n.3, p.377-385, 2002.

MORAND-PRIEUR, M.E.; VEDEL, F.; RAQUIN, C.; BRACHET, S.; SIHACHAKR, D.; FRASCARIA-LACOSTE, N. Maternal inheritance of a chloroplast microsatellite marker in controlled hybrids between *Fraxinus excelsior* and *Fraxinus angustifolia*. **Molecular Ecology**, V.11, n.3, p.613-617, 2002.

MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially present with non-repetitive DNA in plant genomes. **Nature Genetics**, V.30, p.194-200, 2002 **apud** VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, V.23, n.1, p.48-55, 2005.

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

- OLIVEIRA, E.J.; PÁDUA, J.G.; ZUCCHI, M.I.; VENCOVSKY, R.; VIEIRA, M.L.C. Origin, evolution and genome distribution of microsatellites. **Genetics and Molecular Biology**, V.29, n.2, p.294-307, 2006.
- OSTRANDER, E.A.; JONG, P.M.; RINE, J.; DUYK, G. Construction of small-insert genomics DNA libraries highly enriched for microsatellite repeat sequences. **Proceedings of the National Academy of Science of the United States of America**, V.89, n.8, p.3419-3423, 1992.
- PASHLEY, C.H.; ELLIS, J.R.; MCCAULEY, D.E.; BURKE, J.M. EST databases as a source for molecular markers: Lessons from *Helianthus*. **Journal of Heredity**, V.97, n.4, p.381-388, 2006.
- PEARSE, D.E.; CRANDALL, K.A. Beyond Fst: Analysis of population genetic data for conservation. **Conservation Genetics**, V.5, n.5, p.585-602, 2004.
- PONCET, V.; RONDEAU, M.; TRANCHANT, C.; CAYREL, A.; HAMON, S.; DE KOCHKO, A.; HAMON, P. SSR mining in coffee tree EST databases: potential use of EST-SSRs as markers for the *Coffea* genus. **Molecular Genetics and Genomics**, V.276, n.5, p.436-449, 2006.
- POWELL, W.; MORGANTE, M.; ANDRE, C.; MCNICOL, J.W.; MACHRA, Y.; DOYLE, J.J.; TINGEY, S.V.; RAFALSKI, J.A. Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. **Current Biology**, V.5, n.9, p.1023-1029, 1995.
- PROVAN, J. Novel chloroplast microsatellites reveal cytoplasmic variation in *Arabidopsis thaliana*. **Molecular Ecology**, V.9, n.12, p.2183-2185, 2000.
- PROVAN, J.; POWELL, W.; HOLLINGSWORTH, P.M. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. **Trends in Ecology and Evolution**, V.16, n.3, p.142-147, 2001.
- RENAU-MORATA, B.; NEBAUER, S.G.; SALES, E.; ALLAINGUILLAUME, J.; CALIGARI, P.; SEGURA, J. Genetic diversity and structure of natural and managed populations of *Cedrus atlantica* (Pinaceae) assessed using random amplified polymorphism DNA. **American Journal of Botany**, V.92, n.5, p.875-884, 2005.
- RUNGIS, D.; BÉRUBÉ, Y.; ZHANG, J.; RALPH, S.; RITLAND, C.E.; ELLIS, B.E.; DOUGLAS, C.; BOHLMANN, J.; RITLAND, K. Robust simple sequence repeat markers for spruce (*Picea* spp.) from expressed sequence tags. **Theoretical and Applied Genetics**, V.109, n.6, p.1283-1294, 2004.
- VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, V.23, n.1, p.48-55, 2005.
- SELKOE, K.A.; TOONEN, R.J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. **Ecology Letters**, V.9, n.5, p.615-629, 2006.
- SHEMSKE, D.W.; HUSBAND, B.C.; RUCKELSHAUS, M.H.; GOODWILLIE, C.; PARKER, I.M.; BISHOP, J.G. EVALUATING APPROACHES TO THE CONSERVATION OF RARE AND ENDANGERED PLANTS. **ECOLOGY**, V.75, n.3, p.584-606, 1994.
- SQUIRRELL, J.; HOLLINGSWORTH, P.M.; WOODHEAD, M.; RUSSELL, J.; LOWE, A.J.; GIBBY, M.; POWELL, W. How much effort is required to isolate nuclear microsatellites from plants? **Molecular Ecology**, V.12, n.6, p.1339-1348, 2003.
- TAUTZ, D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. **Nucleic Acids Research**, V.17, n.16, p.6463-6471, 1989.
- THIEL, T.; MICHALEK, W.; VARSHNEY, R.K.; GRANER, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.).

Theoretical and Applied Genetics, V.106, n.3, p.411-422, 2003.

Van EE, B.W.; JELINSKI, N.; BERRY, P.E.; HIPPI, A.L. Phylogeny and biogeography of *Croton alabamensis* (Euphorbiaceae), a rare shrub from Texas and Alabama, using DNA sequence and AFLP data. **Molecular Ecology**, V.15, n.9, p.2735-2751, 2006.

VARSHNEY, R.K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, V.23, n.1, p.48-55, 2005.

VARSHNEY, R.K.; SIGMUND, R.; BORNER, A.; KORZUN, V.; STEIN, N.; SORRELLS, M.E.; LANGRIDGE, P.; GRANER, A. Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Science*, V.168, n.1, p.195-202, 2005b.

VENDRAMIN, G.G. ; DEGEN, B. ; PETIT, R.J. ; ANZIDEI, M. ; MADAGHIELE, A. ; ZIEGENHAGEN, B. High level of variation at *Abies alba* chloroplast microsatellite loci in Europe. **Molecular Ecology**, V.8, n.7, p.1117-1126, 1999.

VENDRAMIN, G.G.; ANZIDEI, M.; MADAGHIELE, A.; SPERISEN, C.; BUCCI, G. Chloroplast microsatellite analysis reveals the presence of population subdivision in Norway spruce (*Picea abies* K.). **Genome**, V.43, n.1, p.68-78, 2000.

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

WEBER, J.L. Human DNA polymorphisms and methods of analysis. **Current Opinion in Biotechnology**, V.1, n.2, p.166-171, 1990.

WEISING, K.; GARDNER, R.C. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. **Genome**, V.42, n.1,

p.9-19, 1999.

WHITE, G.; POWELL, W. Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species. **Molecular Ecology**, V.6, n.9, p.851-860, 1997.

WHITE, T.J.; ARNHEIM, N.; ERLICH, H.A. The polymerase chain reaction. **Trends in Genetics**, V.5, n.6, p.185-189, 1989.

WHITE, G.M.; BOSCHER, D.H.; POWELL, W. Genetic variation within a fragmented population of *Swietenia humilis* Zucc. **Molecular Ecology**, V.8, n.11, p.1899-1909, 1999.

WHITE, G.M.; BOSCHER, D.H.; POWELL, W. Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. **Proceedings of the National Academy of Sciences of the United States of America**, V.99, n.4, p.2038-2042, 2002.

WILLIAMS, J.G.; KUBELIK, A.R.; LIVAK, K.J.; RAFALSKI, J.A.; TINGEY, S.V. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research**, V.18, n.22, p.6531-6535, 1990.

WOODHEAD, M.; RUSSELL, J.; SQUIRRELL, J.; HOLLINGSWORTH, P.M.; MACKENZIE, K.; GIBBY, M.; POWELL, W. Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. **Molecular Ecology**, V.14, n.6, p.1681-1695, 2005.

XIE, H. ; SUI, Y. ; CHANG, F.Q. ; XU, Y. ; MA, R.C. SSR allelic variation in almond (*Prunus dulcis* Mill.). **Theoretical and Applied Genetics**, V.112, n.2, p.366-372, 2006.

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

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)