

KARINA DOS SANTOS PADUAN

Desenvolvimento de marcadores SNPs (Single Nucleotide
Polymorphisms) para estudos populacionais do mosquito
Aedes aegypti no Brasil

Botucatu - SP

2008

KARINA DOS SANTOS PADUAN

Desenvolvimento de marcadores SNPs (Single Nucleotide
Polymorphisms) para estudos populacionais do mosquito
Aedes aegypti no Brasil

Tese apresentada ao Curso de Pós-graduação em Ciências
Biológicas, Área de Concentração Genética, do Instituto de
Biociências de Botucatu, Universidade Estadual Paulista –
UNESP, para obtenção do título de doutor.

Orientador: Prof. Dr. Paulo Eduardo Martins Ribolla

Botucatu - SP

2008

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉCNICA DE AQUISIÇÃO E TRATAMENTO
DA INFORMAÇÃO
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CAMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: SELMA MARIA DE JESUS

Paduan, Karina dos Santos.

Desenvolvimento de marcadores SNPs (Single Nucleotide Polymorphism) para estudos populacionais do mosquito *Aedes aegypti* no Brasil/ Karina dos Santos Paduan. – Botucatu : [s.n.], 2008.

Tese (doutorado) – Universidade Estadual Paulista, Instituto de Biociências de Botucatu 2008

Orientador: Paulo Eduardo Martins Ribolla

Assunto CAPES: 20200005

1. *Aedes aegypti* - Genética 2. Mosquitos

CDD 597.7

Palavras-chave: *Aedes aegypti*; Brasil; Dengue; Marcadores moleculares; *Single nucleotide polymorphisms*

*Aos meus pais com todo meu coração
por todo amor e incentivo irrestritos*

Dedico

Agradecimentos

A Deus, em primeiro lugar, pela minha existência e pela força e saúde que me permitiram concretizar este trabalho.

Ao professor Paulo Ribolla pela orientação, sugestões e apoio durante a realização deste doutorado e, sobretudo pela oportunidade de realizar um trabalho gratificante e enriquecedor à minha formação profissional.

Aos meus pais, irmão, cunhada e sobrinho que sempre me deram bons exemplos, e principalmente muito amor. Por todo apoio e incentivo em todos os momentos da minha vida, pela dedicação e amizade sempre presentes. Por mais que eu agradeça a cada realização, ou a cada dia, certamente nunca será suficiente.

Ao meu namorado Adriano, pelo apoio, compreensão e companheirismo durante as fases difíceis pelas quais passei e pelas grandes alegrias dos bons momentos.

A todos os colegas do Entomolab: Alberto, Aline, Bianca, Diego, Jayme, Leticinka, Leti's pela amizade, troca de experiências e convívio, que sem dúvida contribuíram muito para o meu desenvolvimento e conclusão deste trabalho.

A todos os amigos da salinha e do Departamento que tornaram mais alegres e divertidos os meus dias Betina, Bruna, Juliana, Karina, Tati, e tantos outros, pela amizade e agradável convivência.

A minha família 'botucatuense' Passarinho, Tia Sílvia, Lilika, Netão ... e Agregados, por me acolherem em suas vidas sempre com todo carinho, apoio e incentivo em todos os momentos.

A minha grande amiga Tânia Raso estimuladora dos meus projetos, por toda colaboração e amizade em todos estes anos.

Em especial a toda minha família... avós, tios, primos pelo apoio e carinho.

Aos professores do Departamento de Parasitologia da UNESP-Botucatu pelos ensinamentos e receptividade.

Aos técnicos do Departamento de Parasitologia pela amizade e apoio durante todos estes anos de estudos.

Meu carinho e gratidão a todos que de alguma forma participaram da minha vida nestes anos de doutoranda!!!

A CAPES pelo apoio à realização desta pesquisa com a concessão da bolsa de estudos.

*Se pude enxergar mais longe é porque
estava sobre os ombros de gigantes
'Isaac Newton'*

RESUMO

Aedes aegypti é o principal vetor dos vírus dengue e representa cada vez mais um problema de saúde pública. O uso de inseticidas de origem orgânica ou inorgânica é uma das metodologias mais adotadas como parte do controle de vetores, entretanto, processos de resistência têm sido detectados freqüentemente. O desenvolvimento de ferramentas moleculares tem facilitado o entendimento das relações entre o vetor, patógeno e o homem. Os marcadores genéticos mais utilizados em estudos de genética de populações geralmente representam um trabalho laboratorial intenso e, em alguns casos, são limitados pela quantidade de DNA genômico extraído por indivíduo. Recentemente, a tecnologia disponível, permitiu o desenvolvimento dos marcadores SNPs (*Single Nucleotide Polymorphisms*) que correspondem a diferenças de um único nucleotídeo na seqüência de DNA. Assim, o presente trabalho objetivou desenvolver marcadores SNPs localizados em genes nucleares, do mosquito *Ae. aegypti*, para aplicação na genotipagem de populações deste vetor no Brasil. Para tanto, foram investigadas seqüências em dezoito genes nucleares de dezesseis populações do *Ae. aegypti*, e nove destes genes foram selecionados para a presença de SNPs. Os oito marcadores mais polimórficos foram utilizados na caracterização de três populações do mosquito anteriormente analisadas utilizando-se as técnicas de RAPD (*Random Amplified Polymorphic DNA*) e seqüenciamento de genes mitocondriais. Os resultados obtidos com os SNPs mostraram um alto grau de polimorfismo intra e inter populacional, mais significativo quando comparado com os estudos realizados anteriormente, além da presença de duas diferentes linhagens genéticas entre as populações geográficas estudadas.

Palavras-chave: *Aedes aegypti*, Brasil, Dengue, Polimorfismos de único nucleotídeo, *multilocus*.

ABSTRACT

Aedes aegypti is the primary vector of dengue fever virus and it represents a major problem of Public Health. The use of insecticides of organic or inorganic origin is one of the more accepted methodologies as part of this control, however, resistance processes have been detected constantly. The development of molecular tools has made possible the understanding of the relationships among vector, pathogen and man. The genetic markers more used in studies of genetics populations usually represent an intense laboratorial work and in some cases they are limited for the amount of genomic DNA extracted by individual. Recently, the available technology, allowed the development of the markers SNPs (Single Nucleotide Polymorphism) that correspond to differences of a single nucleotide in the sequence of DNA. Thus, this work aims to develop of SNPs markers in nuclear genes of the *Ae. aegypti* for application in genotyping of populations of this vector in Brazil. For so much, were investigated sequences in eighteen nuclear genes of sixteen populations of *Ae. aegypti*, and nine of this genes were selected to presence of the SNPs. Eight more polymorphic markers were used in the characterization of three populations of the mosquito previously analyzed using the techniques of RAPD (Random Amplified Polymorphic DNA) and sequencing of mitochondrial genes. Results obtained with the SNPs showed a high degree of polymorphic intra and interpopulations, more significantly when compared with the prior studies, besides the presence of two different genetic lineages distributed among the studied geographical populations.

Key words: *Aedes aegypti*, Brazil, Dengue, Single Nucleotide Polymorphism, *multilocus*.

SUMÁRIO

	Página
RESUMO	vii
ABSTRACT	viii
LISTA DE FIGURAS	ix
LISTA DE TABELAS	xii
LISTA DE SÍMBOLOS E ABREVIATURAS	xiii
I. REVISÃO BIBLIOGRÁFICA	1
1.1 O <i>Aedes aegypti</i> (Linnaeus, 1762)	1
1.2 Importância Médica	4
1.3 Dengue	5
1.4 Expansão da Dengue	6
1.5 Dengue e o mosquito <i>Aedes aegypti</i>	9
1.6 Epidemiologia	9
1.7 Medidas de Controle e Vigilância Epidemiológica	10
1.8 Estrutura Genética nas Populações de Mosquito	11
1.9 Marcadores Moleculares	12
1.10 Marcadores Moleculares Dominantes	13
1.11 Marcadores Moleculares Codominantes	13
1.12 Marcadores Moleculares Haplóides	14
1.13 Marcadores Moleculares <i>Multilocus</i> Codominante	17
II. JUSTIFICATIVA	22
III. OBJETIVOS	24
IV. MATERIAL E MÉTODOS	26
4.1 Estudo de seleção dos SNPs	26
4.1.1 Amostragem	26
4.1.2 Preparação do DNA	28
4.1.3 Seleção dos oligonucleotídeos iniciadores	28
4.1.4 Condições de Amplificação	30
4.1.5 Purificação dos produtos de PCR	30
4.1.6 Reação de seqüenciamento	30
4.1.7 Precipitação das reações de seqüenciamento	31
4.1.8 Análise das seqüências	31
4.1.9 Análises dos SNPs	31

4.2 Estudo de Genotipagem por SNPs	32
4.2.1 Descrição dos genes selecionados	33
4.2.2 Caracterização dos SNPs por um método semi-automático	36
4.2.3 Seleção dos oligonucleotídeos flanqueadores	37
4.2.4 Reação de minisequenciamento	37
4.2.5 Purificação das reações de minisequenciamento	37
4.3 Análises da estrutura genética	38
4.3.1 Obtenção dos dados	38
4.3.2 Identidade de nucleotídeos e aminoácidos	38
4.3.3 Padrões de substituição	38
4.3.4 Desequilíbrio de ligação	38
4.3.5 Frequências alélicas	39
4.3.6 Equilíbrio de Hardy-Weinberg	39
4.3.7 Análise da estrutura genética populacional	40
4.3.8 Estruturação populacional	41
V. RESULTADOS	43
5.1 Análises dos oligonucleotídeos	43
5.2 Análises dos genes do estudo	45
5.3 Identidade de nucleotídeos/aminoácidos	47
5.4 Polimorfismos SNPs e genotipagem	53
5.5 Desequilíbrio de ligação	71
5.6 Frequências alélicas	75
5.7 Equilíbrio de Hardy-Weinberg	77
5.8 Estrutura genética populacional	80
5.10 F_{ST} pairwise	81
5.11 Estruturação populacional	82
VI. DISCUSSÃO	84
VII. CONCLUSÃO	91
VIII. REFERÊNCIAS BIBLIOGRÁFICAS	93
ANEXO 1 – PAPER ACEITO PARA PUBLICAÇÃO -	

LISTA DE FIGURAS

Figura 1	Mosquito <i>Aedes aegypti</i>	01
Figura 2	Países/áreas com risco de transmissão de dengue no ano de 2006.....	02
Figura 3	Criadouros naturais e artificiais do mosquito <i>Aedes aegypti</i>	03
Figura 4	Ciclo evolutivo do mosquito <i>Aedes aegypti</i>	04
Figura 5	Distribuição da dengue no mundo no ano de 2005.....	08
Figura 6	Organização do DNA mitocondrial com seus genes codificantes e a região controle (Control). lrRNA, subunidade ribossomal grande; srRNA, subunidade ribossomal pequena; ND1, ND2, ND3, ND4, ND4L, ND5 e ND6, subunidades da NADH desidrogenase; Cyt b, citocromo b; COI, COII e COIII, subunidades da citocromo c oxidase; ATP6 e ATP8, subunidades ATPase.....	15
Figura 7	Cromatograma com a representação de um SNP (Polimorfismo de Único Nucleotídeo) (C↔T) em seqüências do mosquito <i>Aedes aegypti</i>	18
Figura 8	Mapa do Brasil com as localidades de coleta das populações de <i>Aedes aegypti</i>	27
Figura 9	Localização cromossômica dos genes selecionados para o estudo.....	32
Figura 10	Visualização em gel de agarose 1% dos produtos de PCR do gene Na/K. Poços de 1 a 7: espécimes provenientes de Botucatu; PM: Marcador de Peso Molecular 250 pb.....	45

Figura 12	Alinhamento múltiplo da região amplificada do gene da <i>Ferritin heavy chain</i> realizado com o programa Clustal X. <i>Indel</i> presente nas seqüências está indicado em vermelho. <i>Indel</i> com 1 pb.....	50
Figura 13	Alinhamento múltiplo da região amplificada do gene da <i>Mucin-like protein</i> realizado com o programa Clustal X. <i>Indel</i> presente nas seqüências está indicado em vermelho. <i>Indel</i> com 24 pb.....	51
Figura 14	Alinhamento múltiplo da região amplificada do gene da <i>Phosphoglycerate kinase</i> realizado com o programa Clustal X. <i>Indel</i> presente nas seqüências está indicado em vermelho. <i>Indel</i> com 10pb.....	52
Figura 15	Regiões amplificadas de cada um dos nove genes nucleares. Íntros aparecem em letras minúsculas, oligonucleotídeos das reações de PCR estão sublinhados, oligonucleotídeos das reações de SNaPshot estão em negrito e itálico. Sítios SNPs estão sublinhados, sítios SNPs testados estão em amarelo, sítios SNPs utilizados na genotipagem estão em vermelho. SNPs não sinônimos estão indicados com a cor cinza.....	61
Figura 16	Histograma das freqüências alélicas dos oito SNPs estimados para os 76 indivíduos das populações de Maceió, Botucatu e Juazeiro do Norte. O eixo Y indica a freqüência alélica e o eixo X as populações.....	75
Figura 17	Resultado da análise de <i>clusters</i> . Cada um dos 76 indivíduos incluídos na análise é representado por uma linha vertical, dividida em duas cores. Estas cores representam a probabilidade do indivíduo pertencer a um dos <i>clusters</i>	82

LISTA DE TABELAS

Tabela I	Localização e coordenadas dos sítios de coleta do mosquito <i>Aedes aegypti</i> no Brasil.....	27
Tabela II	Seqüências do GenBank utilizadas na confecção dos oligonucleotídeos iniciadores, para cada gene nuclear do estudo.....	29
Tabela III	Oligonucleotídeos utilizados para amplificar genes nucleares contendo polimorfismos de único nucleotídeo.....	44
Tabela IV	Genes, número de seqüências, indivíduos e tamanho das seqüências utilizadas para validação dos oligonucleotídeos.....	46
Tabela V	Distribuição dos SNPs identificados nos 9 genes nucleares do mosquito <i>Aedes aegypti</i>	47
Tabela VI	Diversidade de nucleotídeos para os nove genes nucleares do mosquito <i>Aedes aegypti</i>	48
Tabela VII	Polimorfismos em nucleotídeos para os nove genes nucleares do mosquito <i>Aedes aegypti</i>	49
Tabela VIII	Setenta e um polimorfismos de único nucleotídeo identificados em nove genes de dezesseis populações naturais do mosquito <i>Aedes aegypti</i> no Brasil.....	53
Tabela IX	Polimorfismos de único nucleotídeo selecionados para o estudo de caracterização de três populações do mosquito <i>Aedes aegypti</i> no Brasil.....	56
Tabela X	Oligonucleotídeos utilizados para a reação de extensão alelo-específica dos polimorfismos de único nucleotídeo.....	57
Tabela XI	Planilha do Excel montada através da análise multi- <i>locus</i> de três populações	

	naturais do mosquito <i>Aedes aegypti</i> com sete SNPs de sete genes nucleares.....	59
Tabela XII	Histograma do número de <i>loci</i> ligados por <i>locus</i> para as populações de Maceió (A), Botucatu (B) e Juazeiro do Norte (C). (nível de significância = 0.0500).....	71
Tabela XIII	SNPs utilizados no estudo de caracterização das populações de Botucatu, Juazeiro do Norte e Maceió, do mosquito <i>Aedes aegypti</i> no Brasil.....	74
Tabela XIV	Frequência do alelo raro, heterozigosidades observada e esperada, estimativas de FIS (coeficiente de endogamia) e significância dos testes, para 8 SNPs de genes nucleares na população de <i>Aedes aegypti</i> de Botucatu.....	78
Tabela XV	Frequência do alelo raro, heterozigosidades observada e esperada, estimativas de FIS (coeficiente de endogamia) e significância dos testes, para 8 SNPs de genes nucleares na população de <i>Aedes aegypti</i> de Juazeiro do Norte.....	79
Tabela XVI	Frequência do alelo raro, heterozigosidades observada e esperada, estimativas de FIS (coeficiente de endogamia) e significância dos testes, para 8 SNPs de genes nucleares na população de <i>Aedes aegypti</i> de Maceió.....	80
Tabela XVII	Análise de variância molecular (AMOVA) da frequência dos marcadores SNPs entre populações de <i>Aedes aegypti</i> , sem agrupamento.....	81
Tabela XVIII	Matriz de diferenciação genética (F_{ST} pairwise) das populações de <i>Aedes aegypti</i> calculado a partir do método da estatística F	81

LISTA DE SÍMBOLOS E ABREVIATURAS

- BLAST - *Basic Local Alignment Search Tool*
- °C - graus Celsius
- COI - Citocromo c oxidase I;
- DENV (Dengue virus) - vírus Dengue;
- DNA - Ácido desoxirribonucléico;
- DNAm - DNA mitocondrial;
- dNTP - *deoxynucleotide triphosphates* - (dATP, dTTP, dCTP, dGTP);
- EDTA - ácido etilendiaminotetracético sal dissódico 2-hidrato;
- FA - Febre amarela;
- Kb - kilobases;
- KCl - Cloreto de potássio;
- L1 - Larva de primeiro estágio;
- L2 - Larva de segundo estágio;
- L3 - Larva de terceiro estágio;
- L4 - Larva de quarto estágio;
- M - Molar;
- mM - Milimolar (10^{-3} molar);
- µl - Microlitros (10^{-6} litros);
- MgCl₂ - Cloreto de magnésio;
- min - Minutos
- NaCl - Cloreto de sódio;
- NaOH - Hidróxido de sódio;
- ND4 - Nicotinamida adenina dinucleotídeo desidrogenase subunidade 4;
- ng - nanogramas (10^{-9} g)
- nm - nanômetros;
- pb - pares de bases;
- pmol - picomol;
- OPAS - Organização Panamericana de Saúde;
- PCR (Polymerase Chain Reaction) - Reação em Cadeia da DNA Polimerase;
- QTL (Quantitative Trait Loci) - Locos de Caracteres Quantitativos
- RAPD (Random Amplified Polymorphic DNA) - Polimorfismo de DNA Amplificado Randomicamente;

- RFLP (Restriction Fragment Length Polymorphism) - Polimorfismo de tamanho dos fragmentos de restrição;
- RNA - Ácido ribonucleico;
- RNAm - RNA mensageiro;
- RNAr - RNA ribossomal;
- RNAt - RNA transportador;
- rpm - Rotações por minuto;
- RT-PCR - (Reverse Transcriptase-Polymerase Chain Reaction) - PCR por Transcriptase Reversa;
- seg - Segundos;
- SNP - (Single Nucleotide Polymorphism) - Polimorfismo de único nucleotídeo;
- SSCP (Single-Strand Conformation Polymorphism) - Polimorfismo de conformação da fita simples;
- T.A. - Temperatura ambiente;
- U - unidades;
- UV - Ultra-violeta;
- xg - Força centrífuga relativa.

I. REVISÃO BIBLIOGRÁFICA

1.1 O *Aedes aegypti* (Linnaeus, 1762)

O *Aedes aegypti* é um díptero da família Culicidae, pertencente à subfamília Culicinae e ao subgênero *Stegomyia*, sendo considerado o principal vetor da febre amarela (FA) urbana e dos vírus dengue sorotipos 1-4 (DENV 1-4) (Gubler, 2002). A partir do estudo de padrões de variações morfológicas, biológicas e de distribuição geográfica de populações africanas do vetor, Mattingly (1957) registrou a natureza politípica do *Ae. aegypti*. Morfologicamente, as formas típicas adultas apresentam-se freqüentemente ornamentadas com manchas e desenhos de escamas claras no tórax e nas pernas, sendo que existem duas subespécies baseadas na morfologia, ecologia e distribuição geográfica. A *Ae. aegypti formosus* é de coloração escura e usualmente encontrada em áreas silvestres com distribuição ao sul do deserto do Saara e a *Ae. aegypti aegypti* apresenta coloração castanha e é encontrada em ambientes domiciliares com distribuição nas planícies costeiras da África, sendo essa última a forma encontrada em localidades geográficas fora do Continente Africano (Figura 1) (Tabachnick & Powell, 1979).



Figura 1. Mosquito *Aedes aegypti*. Fonte: www.fotografos.com.br

Evidências sugerem que este mosquito é originário do Continente Africano, especialmente no território onde hoje se situa a Etiópia. É nessa região faunística que se encontra o maior número de representantes do subgênero *Stegomyia* e onde ocorrem as espécies mais estreitamente relacionadas ao *Ae. aegypti* (Mattingly, 1957). Autores consideram que esta espécie tornou-se essencialmente doméstica em ambiente africano,

adaptando-se ao ambiente humano e procriando em reservatórios de água estocada nos povoados. A partir do Continente Africano, o mosquito se dispersou pelo mundo ao longo das rotas comerciais durante os séculos XVII, XVIII e XIX. A presença do vetor no continente americano data dos séculos XVII e XVIII, e foi associada ao intenso tráfego de navios negreiros procedentes da África Ocidental (Tabachnick, 1991). Próximo do ano de 1800, a espécie *Ae. aegypti* já tinha se estabelecido em muitas cidades tropicais, especialmente no Novo Mundo (Smith, 1956). Após seu estabelecimento distribuiu-se amplamente nas regiões tropicais e subtropicais, principalmente entre os paralelos de 35° N e 35° S (Slosek, 1986) (Figura 2).

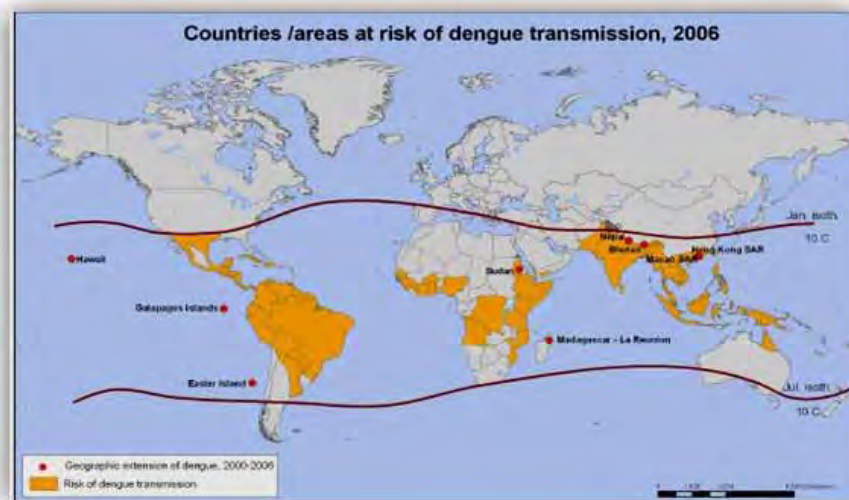


Figura 2. Países/áreas com risco de transmissão de dengue no ano de 2006. Fonte: Organização Mundial de Saúde

Acredita-se que o vetor tenha sido introduzido no Continente Americano em época precoce da colonização, graças ao tráfico comercial que se estabeleceu entre as Américas e as áreas de onde ele é natural (Forattini, 1965). Trata-se talvez da única espécie de mosquito a espalhar-se por todo o globo terrestre, com o concurso da atividade humana. No entanto, a distribuição deste vetor deve ser vista sob aspecto salpicado, antes do que contínuo e extenso, devido ser um mosquito de hábitos essencialmente domésticos e dotado de grande antropofilia, embora existam subespécies com alguma afinidade por ambientes silvestres (Mattingly, 1957).

De hábito diurno, este culicídeo é considerado hematófago voraz, sendo muito eclético quanto ao hospedeiro e sugando grande variedade de animais e o homem. O habitat desta

espécie está intimamente ligado as condições domiciliares ou peridomiciliares, decorrentes da ação antrópica (Figura 3). Sua preferência pelos depósitos artificiais como local de ovoposição faz com que a concentração populacional advinda com a urbanização, ao lado da larga utilização moderna de recipientes artificiais, seja fator determinante na sua crescente proliferação nos centros urbanos das regiões tropicais e subtropicais do planeta. Como exemplos de criadouros artificiais podemos citar: pneus, latas, caixas d'água, piscinas. Nesses criadouros, as fêmeas depositam seus ovos, isoladamente, em um substrato úmido, próximo à água ou em local inundável (Consoli & Lourenço-de-Oliveira, 1994).



Figura 3. Criadouros naturais e artificiais do mosquito *Aedes aegypti*. Fonte: www.novaimpressa.inf.br

O ciclo de desenvolvimento do *Ae. aegypti* compreende quatro estágios: ovo, larva (4 estádios: L1, L2, L3 e L4), pupa e adulto (Figura 4). O crescimento e o desenvolvimento das larvas variam com a temperatura da água, disponibilidade de alimento e densidade populacional no criadouro. Este ciclo geralmente se desenvolve em um curto espaço de tempo, podendo variar entre 10 e 15 dias, contribuindo para o aumento da densidade populacional desta espécie. O período de vida do mosquito adulto é de poucas semanas, podendo chegar, entretanto a 45 dias. Os ovos de *Ae. aegypti* têm alta capacidade de resistir à dessecação. Esta capacidade deve-se a um fenômeno denominado diapausa que é caracterizado pela suspensão de função e desenvolvimento desses ovos. Nos culicídeos este fenômeno ocorre quando a larva de primeiro estágio já está formada, mantendo-se viável na ausência de água por até 450 dias. Os ovos com desenvolvimento completo quando colocados em contato com a água, desencadeiam o processo de eclosão das larvas dando início ao ciclo de desenvolvimento (Honório, 1999).

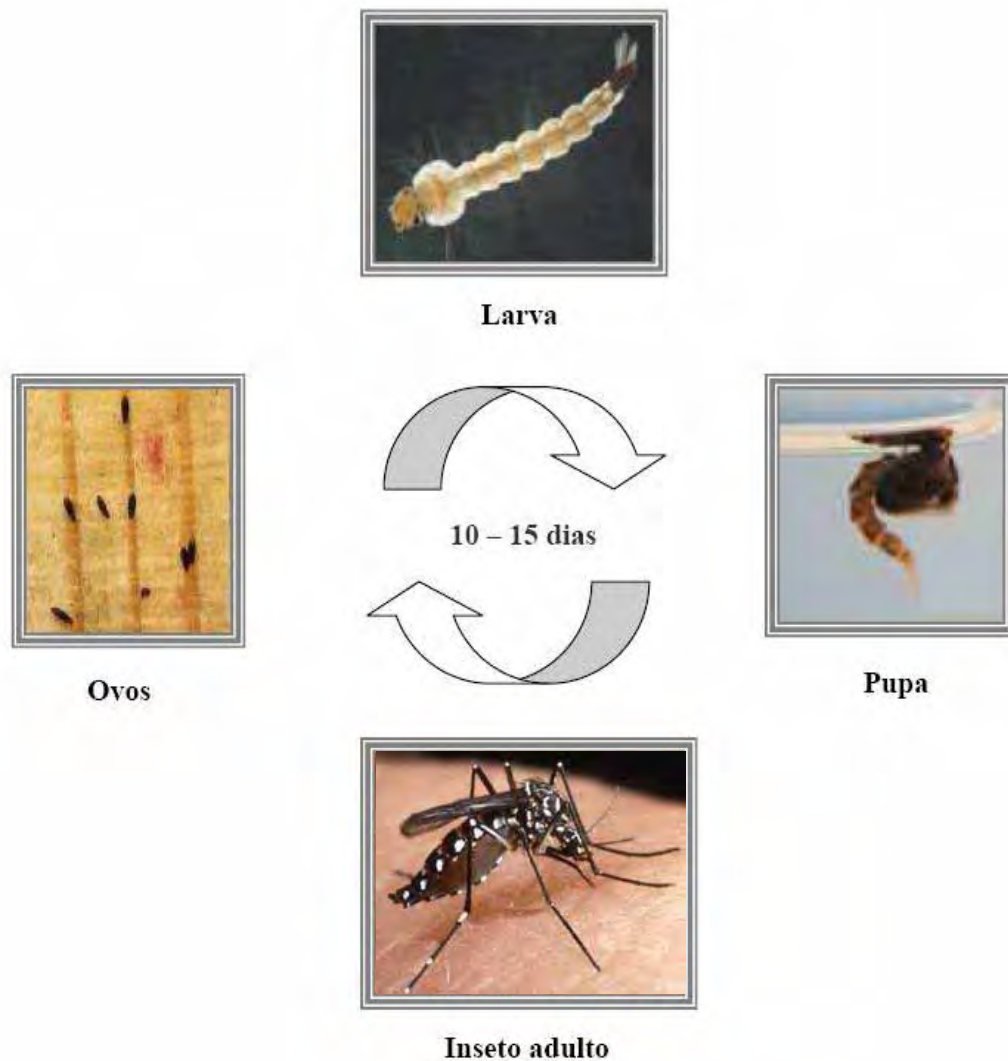


Figura 4. Ciclo de vida do mosquito *Aedes aegypti*.

1.2 Importância Médica

Segundo Mitchell (1995) uma espécie de mosquito só poderá ser considerada vetor de determinado arbovirus se: 1) ocorrer o isolamento do agente em espécimes coletados diretamente da natureza; 2) for demonstrado que o mosquito se infectou após uma alimentação utilizando o sangue de um hospedeiro na fase de viremia ou uma suspensão de vírus; 3) for demonstrado que o vírus pode ser transmitido para um hospedeiro através de uma picada ou se for encontrado vírus nas glândulas salivares do mosquito e 4) ocorrerem evidências em campo confirmando uma associação entre a espécie de mosquito estudada e a população de vertebrados na qual a infecção viral esteja ocorrendo.

Atualmente, dengue é a arbovirose mais comum que atinge o homem. A denominação "arbovírus" foi recomendada internacionalmente em 1963, para substituir a expressão

"arthropod-borne-virus", adotada em 1942, para designar um grupo de vírus capaz de se multiplicar em artrópodes e de ser transmitido através da picada destes aos hospedeiros vertebrados (Calisher, 1986). Os arbovírus são encontrados em quase todas as regiões do mundo. Atualmente, são conhecidos ao redor de quinhentos vírus, entre os quais, mais de duzentos são veiculados por mosquitos. Os vírus dengue sorotipos 1-4 são causadores de infecções clínicas e subclínicas que se manifestam sob a forma de quatro síndromes: encefalites, febres benignas de curta duração, febres hemorrágicas e poliartrite acompanhada de erupção cutânea. O ciclo de transmissão do vírus dengue inclui o homem infectado, o mosquito vetor do gênero *Aedes* e o homem susceptível. Os arbovírus multiplicam-se nos tecidos dos organismos dos artrópodes, que se infectam tornando-se vetores depois de sugarem sangue de hospedeiros, em período de viremia (Rigau-Pérez et al., 2001).

1.3 Dengue

Sabe-se hoje que o agente etiológico da dengue é representado por um arbovírus da família Flaviviridae, gênero *Flavivirus*, que inclui quatro tipos sorológicos: 1, 2, 3 e 4. O vírus dengue é composto por RNA simples fita, mede 40-50 nm de diâmetro e possui genoma de aproximadamente 11 kb. A infecção por qualquer sorotipo promove imunidade homóloga duradoura contra o mesmo sorotipo, mas não há proteção cruzada para os outros sorotipos. Pessoas habitantes de uma área endêmica podem ser infectadas por diferentes sorotipos da dengue durante a vida. O vírus tem como hospedeiro vertebrado o homem e outros primatas, mas somente o primeiro apresenta manifestação clínica da infecção e período de viremia de aproximadamente sete dias. Nos demais primatas a viremia é baixa e de curta duração (Gubler, 1998a).

No homem a dengue manifesta-se como uma enfermidade infecciosa aguda caracterizada por um amplo espectro clínico que varia desde formas de infecção assintomática ou febre indiferenciada, até as formas de hemorrágicas e/ou choque. Os casos típicos de dengue podem ser agrupados em duas categorias principais: a) dengue clássico caracteriza-se por uma febre alta de início abrupto, cefaléia intensa, dor retroorbitária, dores articulares e musculares, prostração, podendo ocorrer alguns fenômenos hemorrágicos sem maiores conseqüências; b) dengue hemorrágica/ síndrome de choque da dengue é caracterizada por quadro de febre alta, inicialmente indiferenciável da forma clássica, que se segue - quando da normalização da temperatura, entre o terceiro e o quinto

dia da enfermidade - de fenômenos hemorrágicos (petéquia, púrpura, equimose, epistaxe, sangramento gengival, sufusões hemorrágicas, hematêmese, melena) e/ou insuficiência circulatória com ou sem choque hipovolêmico (OMS, 1987).

A origem exata da doença não está clara. Alguns indícios sugerem que ela tenha se originado na África e depois se espalhado pelo mundo com o comércio de escravos (Ehrenkranz et al., 1971; Smith, 1956). Entretanto, uma vez que são altamente adaptados ao hospedeiro invertebrado, tem sido proposto que os vírus podem ter se originado em ciclos silvestres envolvendo 'primatas inferiores' e mosquitos, habitantes da copa das árvores na Península Malaia (Smith, 1956; Rudnick, 1986; Halstead, 1992). As primeiras epidemias de uma doença compatível com o dengue, referidas em literatura, datam de 1779 em Jacarta no Cairo, Indonésia. No ano seguinte, surge a primeira descrição clínica desta enfermidade, feita por Benjamin Rush, durante uma epidemia na Filadélfia. No século XIX e nas primeiras décadas do século XX, foram registradas diversas epidemias atribuídas a dengue: Zanzibar (1823; 1870), Calcutá (1824; 1853; 1871; 1905), Antilhas (1827); Hong Kong (1901); Estados Unidos (1922), Austrália (1925-26; 1942), Grécia (1927-28); Japão (1942-45) (Henchal & Putnak, 1990).

1.4 Expansão da Dengue

Nas Américas, a dengue como entidade clínica é conhecida há pelo menos 200 anos. A primeira pandemia registrada na região, tomando como referência critérios clínicos e epidemiológicos, ocorreu nas Antilhas-Golfo do México-Atlântico, em 1827-28. Depois da década de 50 até os dias atuais observou-se intensificação da atividade do dengue no continente americano, verificando-se intervalos mais curtos entre os períodos pandêmicos, e o envolvimento de novas áreas geográficas anteriormente não atingidas como a América do Sul (Ehrenkranz, 1971; Pinheiros, 1989; OPAS, 1992).

Os primeiros casos de dengue no Brasil foram relatados baseados em critérios clínicos, em Niterói, Rio de Janeiro, em 1923. A partir daí, somente de novembro de 1981 a março de 1982, ocorreu uma epidemia na cidade de Boa Vista, Roraima, com circulação dos sorotipos 1 e 4 (Osanai, 1984), provavelmente uma expansão da onda epidêmica que atingiu vários países da América Central e do norte da América do Sul nos finais da década de 70. No ano de 1986, outra epidemia de dengue ocorreu no Brasil, na região metropolitana do Rio de Janeiro e municípios vizinhos (Schatzmayer et al., 1986; Nogueira et al., 1988). Em 1987, mais quatro Estados foram atingidos: Pernambuco, Bahia, Minas Gerais e São Paulo. Em

1990, a dengue atingiu o Estado de Mato Grosso do Sul e novamente com maior intensidade, o Estado de São Paulo; em 1991 foram registrados casos no Estado do Tocantins e, em 1992, em Mato Grosso. Pode-se considerar, tomando como referência as séries históricas de dados de notificação para o período 1980-1991, divulgados pelo Ministério da Saúde, que a dengue tornou-se endêmica e epidêmica na maioria dos Estados do Brasil (Ministério da Saúde, 1992).

No Estado de São Paulo, de acordo com os dados do Centro de Vigilância Epidemiológica da Secretaria de Saúde (CVE, 1987), os primeiros casos confirmados de dengue aconteceram em 1986, num total de 32, todos importados (28 do Rio de Janeiro, 2 do Ceará e 2 de Alagoas). No ano de 1987 verificou-se a primeira epidemia em território paulista, em Guararapes e em Araçatuba. Nos anos de 1988, 1989 e 1990 até novembro houve confirmação apenas de casos importados. A partir de novembro de 1990 e nos primeiros meses de 1991 ocorreu epidemia de importante magnitude partindo de Ribeirão Preto, localizada em região próxima ao Mato Grosso do Sul, onde estava ocorrendo grande epidemia de dengue que atingiu diversas cidades do interior paulista. Nestas situações relatadas foi identificada apenas a circulação do vírus sorotipo 1 (Pontes, 1991).

Nos dias atuais, em território brasileiro, os dados mostram que existem três sorotipos do vírus Dengue circulando: DENV-1, DENV-2 e DENV-3. Entretanto, desde 2005 pesquisadores relatavam a possibilidade da introdução do sorotipo 4 em razão do intenso transporte aéreo e marítimo entre o Brasil e outros países do continente americano, bem como países de outros continentes afetados pelo DENV-4 (Teixeira et al., 2005). Recentemente Figueiredo et al. (2008) reportaram a circulação do sorotipo DEN-4 na cidade de Manaus, que se confirmado, seria o primeiro registro deste sorotipo no Brasil em 25 anos. A co-circulação de mais de um sorotipo e o aumento da densidade populacional dos vetores responsáveis pela transmissão da doença são fatores que podem favorecer o risco de mais uma epidemia com um aumento no número de casos da forma hemorrágica dengue. Atualmente a dengue é endêmica em países tropicais do Pacífico Sul, Ásia, África, Américas, Caribe, Oceania e Austrália (Figuras 5-A/B). Tem-se observado novas regiões e diversos países antes não atingidos com atividade epidêmica, e circulação simultânea ou sucessiva de vários sorotipos virais, indicando crescente disseminação da dengue no mundo (PAHO, 2006).

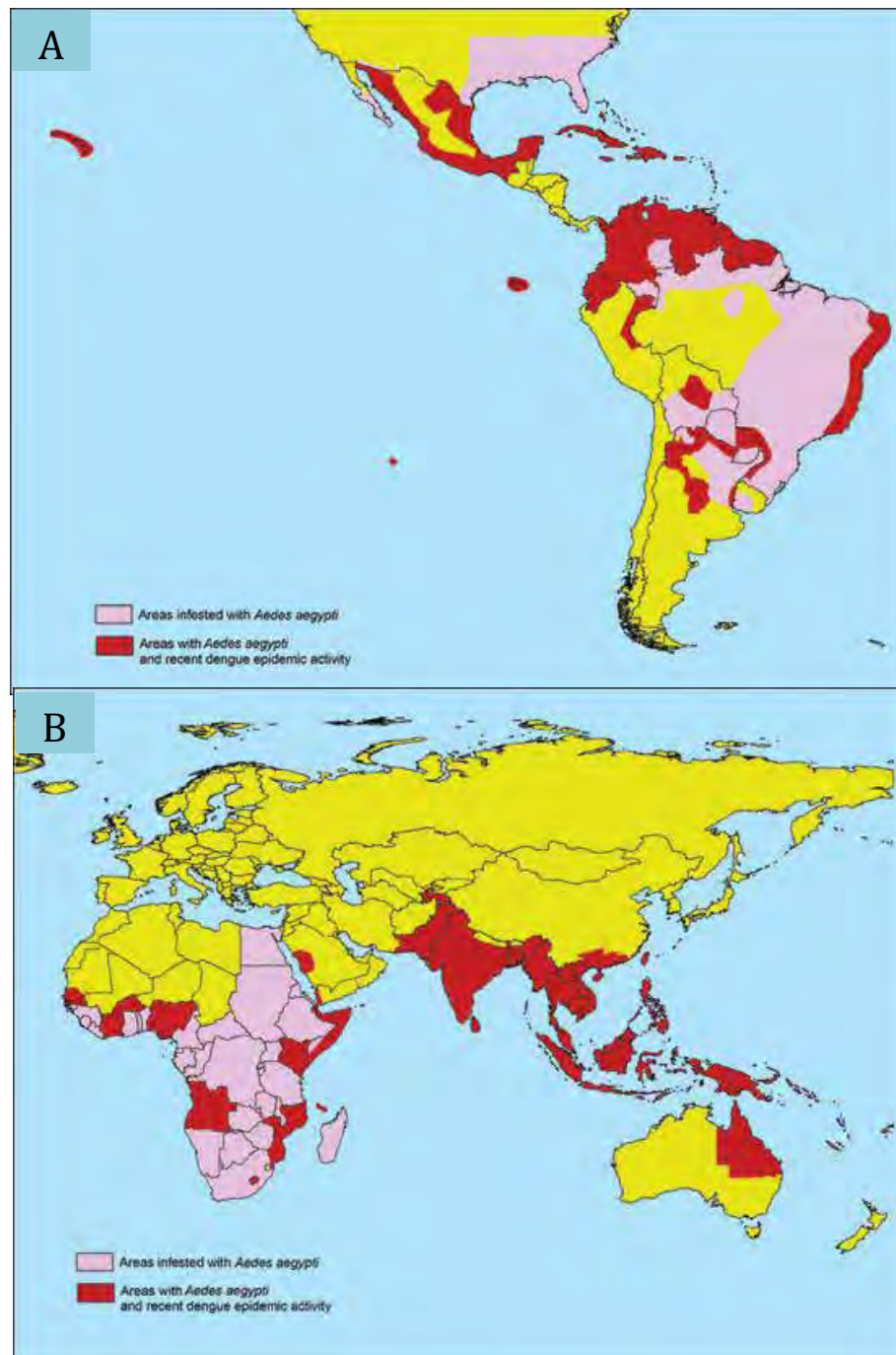


Figura 5 – A/B. Distribuição da dengue no mundo no ano de 2005. Fonte: www.cdc.gov

1.5 Dengue e o mosquito *Aedes aegypti*

Em meados da década de 40, até os anos 60, a dengue foi praticamente eliminada de 18 países incluindo o Brasil, em virtude do combate ao vetor em escala continental. O Brasil conseguiu a eliminação da transmissão em 1955, referendada oficialmente pela Organização Panamericana de Saúde (OPAS) em 1958. Devido à redução da vigilância e dos programas de controle, observou-se a reinfestação pelo *Ae. aegypti* em 1967, provavelmente originada dos

países vizinhos que não obtiveram êxito em sua erradicação (Franco, 1976). As primeiras reinfestações por este culicídeo no Brasil, ocorreram em Belém (Pará) em 1967, e em São Luiz (Maranhão) em 1968, tendo sido eliminados os respectivos focos em 1973. No ano de 1976 constatou-se a presença do vetor em Salvador (Bahia) e em 1977 no Rio de Janeiro (Rio de Janeiro) (Ministério da Saúde, 1991). A primeira epidemia de dengue foi registrada na cidade de Boa Vista, Estado de Roraima (na região norte da Amazônia Brasileira), com circulação dos sorotipos I e 4 (Osanai, 1984).

No Estado de São Paulo, os primeiros registros da presença do *Ae. aegypti* datam de 1980 e 1981 na zona portuária de Santos. Em 1985, a espécie foi observada na região oeste do Estado, provavelmente introduzida a partir do Mato Grosso e/ou Paraná (Glasser, 1997). A progressão geográfica da infestação pelo *Ae. aegypti* ocorreu na direção dos municípios situados de oeste para leste. Desde a re-introdução em 1967, não houve mais sucesso na eliminação deste culicídeo que atualmente está presente em condição domiciliada em 26 Estados e aproximadamente 4000 municípios (Penna, 2003).

1.6 Epidemiologia

A dengue que foi reconhecida há aproximadamente 200 anos tem apresentado caráter epidêmico e endêmico variado, com tendência ao agravamento nos últimos anos. Atualmente, é uma das mais importantes arboviroses transmitidas por insetos. Estima-se que 80 milhões de pessoas sejam infectadas por ano em 100 países de todos os continentes, exceto a Europa. Cerca de 550 mil casos de dengue hemorrágica são confirmados e pelo menos 12 mil pessoas morrem a cada ano em decorrência da doença (OMS, 2008). O impacto da dengue na população humana é notado, não apenas pelos sintomas, como pela ocorrência de óbitos, principalmente entre crianças. Existem prejuízos econômicos expressivos em decorrência de gastos com: tratamento, hospitalização, controle dos vetores, abstenção ao trabalho e perdas com turismo (WHO, 1997; Holmes et al., 1998).

De acordo com Gubler (1998b), os fatores responsáveis pelo ressurgimento e emergência da dengue epidêmica em escala global, nos últimos 50 anos, estão atribuídos a diversos fatores ainda não bem definidos. Os mais importantes estão relacionados com:

- ✚ As medidas de controle dos vetores de dengue, nos países endêmicos, são poucas ou inexistentes;

- ✚ O crescimento desordenado da população humana global causando mudanças demográficas;

✚ A expansão e alteração desordenadas do ambiente urbano, com infra-estrutura sanitária deficiente, propiciando aumento da densidade das populações do vetor;

✚ O aumento do intercâmbio comercial entre múltiplos países e conseqüente aumento do número de viagens aéreas, marítimas e fluviais, favorecendo a dispersão dos vetores e dos agentes infecciosos.

1.7 Medidas de controle e vigilância do *Aedes aegypti*

Até os dias atuais ainda não estão disponíveis um tratamento específico e uma vacina eficiente contra a dengue. Dessa maneira a única medida empregada depende do controle do vetor urbano, o *Ae. aegypti*. O uso de inseticidas orgânicos ou inorgânicos é o método mais adotado para o controle de insetos vetores de importância em Saúde Pública (Rose, 2001). Entretanto, têm sido detectados mecanismos de resistência para todas as classes de inseticidas, afetando, direta e profundamente, a re-emergência de doenças causadas por microorganismos transmitidos por mosquitos (Brogdon & McAllister, 1998). Atualmente, o desenvolvimento de ferramentas moleculares que possibilitem o entendimento das relações entre o vetor, o patógeno e o homem tem se mostrado cada vez mais necessário.

Gubler (1989) tem proposto diante das dificuldades encontradas, na luta contra a dengue, a utilização de um sistema de vigilância ativo com o objetivo de detectar precocemente a ocorrência de epidemias. Isto possibilitaria por em prática medidas de controle imediatas com o objetivo de reduzir a incidência e, desta forma, o risco de ocorrência de dengue hemorrágica. Esta vigilância ativa estaria centrada principalmente na organização de um sistema de vigilância epidemiológica, que englobaria práticas de vigilância virológica, clínica, sorológica e entomológica.

A vigilância epidemiológica envolve uma seqüência de ações que dependem da situação epidemiológica do município. Nas áreas com transmissão endêmica de dengue, o objetivo da vigilância é fornecer as informações necessárias sobre a situação epidemiológica. A finalidade desta ação é interromper a transmissão viral, através da diminuição da infestação e/ou eliminação do *Ae. aegypti*, visando impedir o aparecimento da dengue hemorrágica, já que a população está sensibilizada por um ou mais sorotipos do vírus. Esta vigilância deve ser realizada de forma contínua, principalmente nos períodos interepidêmicos, quando também existe a circulação do vírus (Fouque et al., 2004). Além disso, instituir a vigilância sorológica com o monitoramento e identificação dos diferentes sorotipos circulantes do vírus também é uma ação necessária. Análises deveriam ser

realizadas tanto em amostras de soro de pacientes, com suspeita clínica da doença, bem como em mosquitos coletados em campo ou em armadilhas de oviposição com o objetivo de elaborar um plano de prevenção para as epidemias de dengue. Ações como monitorar os sorotipos circulantes em mosquitos do gênero *Aedes* coletados em áreas infectadas pelo Dengue podem ser relevantes para determinar a magnitude e gravidade das epidemias (Teixeira et al., 2005).

A diversidade genética das populações de mosquito e das condições ambientais no Brasil, aliada aos aspectos climáticos muito diferentes em nosso país podem favorecer o surgimento da dengue e dificultar o seu controle. Assim, a integração de diversas áreas poderá contribuir como uma parte importante aos avanços da genética molecular no desenvolvimento de marcadores genéticos mais eficientes.

1.8 Estrutura Genética nas Populações de Mosquito

De modo geral, mosquitos vetores de agentes infecciosos manifestam alto grau de 'plasticidade' genética, com a capacidade de gerarem mutantes resistentes a diversas condições ambientais (Besansky et al., 1992). Uma das conseqüências desta elevada 'plasticidade' é a alta freqüência com que populações de mosquito adquirem resistência múltipla a inseticidas químicos ou biológicos utilizados no seu controle (Groetes & Tabachnik, 2000; Hemingway & Ranson, 2000; Ferre & Van Rie, 2002). Diferenças genéticas podem ser responsáveis por diferenças na capacidade e na competência vetorial ou na susceptibilidade do vetor aos agentes infecciosos ou aos inseticidas (Lourenço-de-Oliveira et al., 2002). Dessa maneira, a avaliação da variação genética de populações locais de um vetor pode ajudar a planejar medidas de controle adequadas às condições reais de cada área de interesse.

Desde o final da década de 1970, vários estudos têm demonstrado a variabilidade genética presente em populações do *Ae. aegypti* (Tabachnick & Powell, 1978; 1979), e atualmente existem vários métodos que possibilitam a análise desta variabilidade. Loxdale & Lushai (1998) publicaram uma extensa revisão sobre marcadores moleculares usualmente utilizados para examinar a variabilidade genética em populações de mosquito.

1.9 Marcadores Moleculares

Até o ano de 1966, o estudo de polimorfismos em populações era realizado com base nos fenótipos dos organismos. Entretanto, como os variantes fenotípicos são raros, as

populações eram frequentemente classificadas como homogêneas (Futuyma, 1992). Um dos primeiros ensaios de variação genética envolvendo o uso de um marcador molecular foi realizado com isoenzimas em populações de *Drosophila pseudobscuras*, e se observou que cerca de 30% dos *loci* de cinco populações eram polimórficos (Hubby & Lewontin, 1966). Entretanto, o poder de detecção da variabilidade existente diretamente ao nível do DNA só foi alcançado com o desenvolvimento das técnicas de biologia molecular na década de 80. O desenvolvimento do processo de amplificação em cadeia utilizando uma DNA polimerase (PCR) foi concebido por Kary Mullis e representou grande avanço na detecção de polimorfismos genéticos (Mullis & Faloona, 1987; Saiki et al., 1988). A partir daí, diversas técnicas de biologia molecular foram desenvolvidas, e estão hoje disponíveis para o estudo genético.

Marcador molecular é todo e qualquer fenótipo molecular oriundo de um gene expresso ou de uma sequência de DNA, correspondente a uma região codificante ou não (Ferreira & Grattapaglia, 1996). Um marcador para ser considerado bom e informativo deve ser: 1) variável; 2) abundante; 3) de comportamento previsível quanto à taxa de mutação, padrão de herança e localização cromossômica e, 4) de fácil análise. Os marcadores mais empregados em estudos de polimorfismos são: Isoenzimas, Polimorfismo de DNA Amplificado Randomicamente (RAPD), Polimorfismo de Tamanho dos Fragmentos de Restrição (RFLP), Microsatélites, Polimorfismo de Conformação de Fita Simples (SSCP), além do Sequenciamento do DNA mitocondrial (mtDNA) e de genes de RNA ribossomal (rRNA). Entretanto, a escolha do marcador depende do objetivo do estudo, do nível de observação desejada e da espécie estudada (Roderick, 1996). O estudo da variabilidade genética em populações naturais de *Ae. aegypti* através de marcadores moleculares tem sido utilizado com o objetivo de se estimar a variabilidade genética populacional e analisar os padrões de semelhança/diferença de fragmentos genômicos entre populações geográficas desta espécie. Alguns desses marcadores serão descritos a seguir.

1.10 Marcadores Moleculares Dominantes

Marcadores RAPD (Williams et al., 1990) têm sido muito empregados no estudo de populações naturais nos últimos anos. Eles são tecnicamente mais simples que outros tipos de marcadores (ex. RFLP), pois facilitam o estudo de um grande número de *loci* e fornecem uma amostragem aleatória do DNA. Entretanto, os marcadores RAPD têm limitações significativas quando comparados com marcadores codominantes e/ou haplóides. A

dominância limita as inferências feitas com RAPD, especialmente em estudos de organismos diplóides, e podem fornecer viés em parâmetros genéticos quando comparados com marcadores multialélicos ou codominantes (Lynch & Milligan, 1994). Marcadores RAPD possuem caráter dominante limitando o seu uso na estimativa de parâmetros genéticos, além de apresentar problemas de reprodutibilidade (Karp et al., 1996).

Pesquisas que utilizam esta técnica têm sido empregadas no estudo de populações de *Ae. aegypti* em diversas partes do mundo (Ballinger-Crabtree et al., 1992; Apostol et al., 1994; Gorrochotegui-Escalante et al., 2000; Paduan et al., 2006). Em recente estudo realizado em nosso laboratório, 11 populações naturais de *Ae. aegypti* foram analisadas utilizando-se marcadores RAPD. Os resultados indicaram que as populações brasileiras deste vetor estão estruturadas em 3 regiões geográficas (Norte, Nordeste e Sudeste), sugerindo a presença de subpopulações estruturadas com significativo nível de isolamento genético por distância. Entretanto, apesar dos resultados significativos, não foi possível detectar nenhuma diferenciação genética entre os espécimes dentro de uma mesma localidade (Paduan et al., 2006).

1.11 Marcadores Moleculares Codominantes

Microsatélites têm sido utilizados como marcadores genéticos para um grande número de artrópodes vetores de doenças humanas. Cento e cinquenta *loci* microsatélites foram caracterizados para o vetor da malária na África o *An. gambiae* s.l. (Zheng et al., 1993/1996) e tantos outros foram selecionados em estudos de QTL para susceptibilidade ao *Plasmodium* (Gorman et al., 1997; Zheng et al., 1997) e resistência à permetrina (Ranson et al., 2004). No entanto, em contraste com muitos artrópodes, microsatélites parecem estar sub-representados dentro de alguns membros da subfamília Culicinae. Um número limitado destes marcadores tem sido identificado em *Culex pipiens pipiens* e *Culex quinquefasciatus* (Fonseca et al., 1998; Keyghobadi et al., 2004). Dentro de *Ae. aegypti*, seis microsatélites foram identificados em seqüências depositadas nas bases de dados (Barbazan et al., 1999; Huber et al., 2001; Ravel et al., 2001).

Embora muitos pesquisadores tenham se empenhado em descobrir microsatélites, utilizando a técnica de bibliotecas enriquecidas, apenas um pequeno número de marcadores polimórficos foi selecionado para *Ae. aegypti* (Huber et al., 1999/2001; Fagerberger et al., 2001; Chambers et al., 2007). Ainda assim, este limitado número de microsatélites é aplicado a vários estudos com o objetivo de se conhecer melhor a estrutura populacional do

mosquito (Ravel et al., 2001/2002; Huber et al., 2002/2004). Costa-Ribeiro et al. (2006) testaram os marcadores microssatélites descritos por Huber et al. (2001) em populações de *Ae. aegypti* de cinco distritos no Rio de Janeiro. Apesar de ter obtido resultados promissores em relação aos marcadores isoenzimáticos um dos seis *loci* testados foi monomórfico e outro apresentou déficit de heterozigotos. Além do que, o grau de polimorfismo em populações brasileiras do mosquito *Ae. aegypti* foi muito menor quando comparado com as populações do mosquito estudadas na Ásia (Huber et al., 2002/2004; Paupy et al., 2004).

Na tentativa de se diferenciar populações locais, os mesmos espécimes de *Ae. aegypti* estudados por Paduan et al. (2006) foram submetidos a análises utilizando marcadores microssatélites de acordo com Huber et al. (2001) e Ravel et al. (2002). Entretanto, nenhum dos marcadores testados apresentou polimorfismo entre os indivíduos analisados (dados não publicados).

1.12 Marcadores Moleculares Haplóides

Muitas regiões do DNA do vetor *Ae. aegypti* foram clonadas, seqüenciadas e caracterizadas. Entre elas estão seqüências de genes do DNA mitocondrial que apresentam grande variabilidade e, portanto, são úteis para o reconhecimento de espécies crípticas, estudos filogenéticos ou de estrutura genética de populações (Walton et al., 2001; Alvarez & Hoy, 2002). O DNA mitocondrial (mtDNA) descoberto em 1960 tem organização relativamente simples, constituída por uma dupla fita circular contendo 37 genes, dos quais 2 codificam RNAs ribossômicos, 22 codificam RNAs transportadores e 13 codificam proteínas. Possui uma região não-codificante denominada Região Controle ou D-Loop que é responsável pelo controle da replicação e transcrição do genoma da mitocôndria. (Hoy, 2003) (Figura 6).

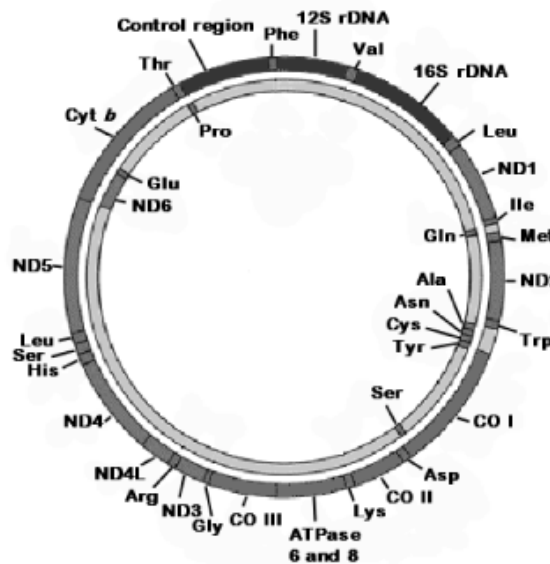


Figura 6. Organização do DNA mitocondrial com seus genes codificantes e região controle. ND1, ND2, ND3, ND4, ND4L, ND5 e ND6, subunidades da NADH desidrogenase; Cyt b, citocromo b; COI, COII e COIII, subunidades da citocromo c oxidase; ATP6 e ATP8, subunidades ATPase. Fonte: Pereira, 2000.

O genoma mitocondrial é herdado maternalmente e não sofre recombinação. A herança materna do DNA mitocondrial possibilita que esta molécula seja usada como um importante marcador molecular com o objetivo de traçar a história evolutiva de uma determinada espécie. É utilizado também para determinar a filogenia entre populações, independente de terem sofrido efeitos de adaptação e/ou seleção. Outra característica importante do DNA mitocondrial é a sua alta taxa de evolução. O DNA mitocondrial evolui 10 vezes mais rápido que o DNA nuclear, pois os mecanismos de reparo nesta molécula são pouco eficientes. As enzimas envolvidas neste processo são codificadas por genes nucleares e dependem de serem transportadas até as mitocôndrias, possibilitando com o acúmulo de mutações (Matioli, 2001). Por essa razão, as variações intraespecíficas são facilmente detectadas (Tabachnick & Black, 1996).

Os genes codificadores das subunidades da Nicotinamida adenina dinucleotídeo desidrogenase (NADH desidrogenase), citocromo c oxidase (CO) e RNAt (RNA transportador) são os genes que acumulam mais rapidamente as substituições de bases. Porém, os genes codificadores para as unidades ribossômicas e para o citocromo b estão entre os genes mais conservados do genoma mitocondrial. As regiões A+T ou D-loop são as mais variáveis, e podem acumular mutações do tipo substituições de bases ou inserção/deleção (Arias & Infante-Malachias, 2001).

A diversidade genética, as relações entre populações geográficas, a história das diferentes linhagens e os efeitos de fenômenos demográficos na composição genética das mesmas estão entre os aspectos que podem ser estudados usando marcadores mitocondriais. Diversas pesquisas têm sido realizadas em populações de diferentes partes do mundo utilizando esta classe de marcadores, todavia, os resultados têm permitido inferir principalmente sobre a presença de linhagens mitocondriais e filogeografia destas populações (Bosio et al., 2005; Gorrochotegui-Escalante et al. 2000/2002; Mousson et al., 2005). Utilizando a técnica de SSCP para analisar um fragmento do gene mitocondrial Nicotinamida adenina dinucleotídeo desidrogenase 4 (ND4) Gorrochotegui-Escalante et al. (2000) determinaram as taxas de fluxo gênico entre 10 populações de sete localidades ao longo da costa nordeste do México. Sete haplótipos foram detectados e as análises filogenéticas indicaram a presença de dois clados independentes bem suportados. Em 2002, na tentativa de estabelecer a estrutura de reprodução desta espécie Gorrochotegui – Escalante et al., compararam novamente populações de *Ae. aegypti* de diferentes regiões do México estudando os polimorfismos do gene ND4. Nesta ocasião foram identificados 25 haplótipos com isolamento genético significativo, indicando que o padrão de estruturação variava em cada região estudada. Diferente do estudo realizado em 2000, as análises filogenéticas indicaram uma única origem monofilética entre os haplótipos detectados.

Bosio et al. (2005) caracterizaram a estrutura genética de populações de *Ae. aegypti* na Tailândia utilizando também o gene ND4. Os autores concluíram que a estruturação desta espécie não variava por região, ao contrário do que aconteceu no México. Este resultado pode ter ocorrido por três fatores: primeiro, o transporte destes mosquitos através das atividades humanas pode ter aumentado a similaridade genética entre populações geograficamente distantes. Segundo, um processo estocástico como a deriva genética poderia ter influenciado no padrão genético dessas populações alterando a variabilidade das mesmas independente das distâncias geográficas. E terceiro, os efeitos dos esforços no controle destes mosquitos teriam levado as populações locais a sofrerem o efeito de “gargalo de garrafa” afetando profundamente a variabilidade genética.

Estudos realizados com populações de *Ae. aegypti* no Brasil têm revelado, principalmente, que existem duas linhagens mitocondriais que evoluem de forma independente e apresentam origem histórica distinta (Bracco et al., 2007; Paduan & Ribolla, 2008). Análises de populações brasileiras de *Ae. aegypti*, por sequenciamento de fragmentos dos genes COI e ND4, foram realizadas em nosso laboratório utilizando as mesmas

populações do estudo realizado por Paduan et al. (2006) e nas análises de microssatélites. O estudo consistiu da análise de 12 populações naturais coletadas em 6 Estados. Os resultados indicaram a presença de duas linhagens mitocondriais independentes e correlação positiva dos haplótipos com a localização geográfica das amostras em determinadas localidades. Contudo, apesar de ser detectada relação de alguns haplótipos com a localização geográfica houve pouca ou nenhuma diferenciação genética entre indivíduos dentro da mesma localidade não sendo possível inferir sobre a discriminação genética e geográfica intrapopulacional (Paduan & Ribolla, 2008).

Devido ao fato do DNAm ser de herança materna, estudar a diferenciação dessa molécula equivale a estudar a população de fêmeas. Isso, combinado com o fato do DNAm ser haplóide, faz com que o tamanho efetivo da população estudada seja quatro vezes menor do que quando se utiliza um marcador nuclear diplóide (Haavie et al., 2000).

1.13 Marcadores Moleculares *Multilocus* Codominante

Com relação ao conhecimento da estrutura genética de populações, os marcadores moleculares utilizados em pesquisas recentes, citadas anteriormente, não permitiram inferir sobre a discriminação genética em populações locais do mosquito *Ae. aegypti*. Na tentativa de se obter resultados que possibilitem estudar os processos comportamentais, evolucionários e históricos das populações, a inclusão de dados genealógicos de *loci* nucleares como, por exemplo, os SNPs (Polimorfismos de Único Nucleotídeo) tem se tornado claramente necessária (Zhang & Hewitt, 2003).

SNP é uma variação na sequência do DNA que ocorre quando um único nucleotídeo – A, T, C, ou G – no genoma difere entre membros de uma espécie, ou entre pares de cromossomos de um indivíduo. Por exemplo, duas sequências de um fragmento de DNA de diferentes indivíduos contêm uma diferença em um único nucleotídeo, GGGGCCG e GGGGTCG. Neste caso diz-se que existem dois alelos: C ou T (Figura 7). Esses marcadores são denominados bialélicos devido ao fato de que, geralmente, são observados apenas dois alelos (Brookes, 1999).

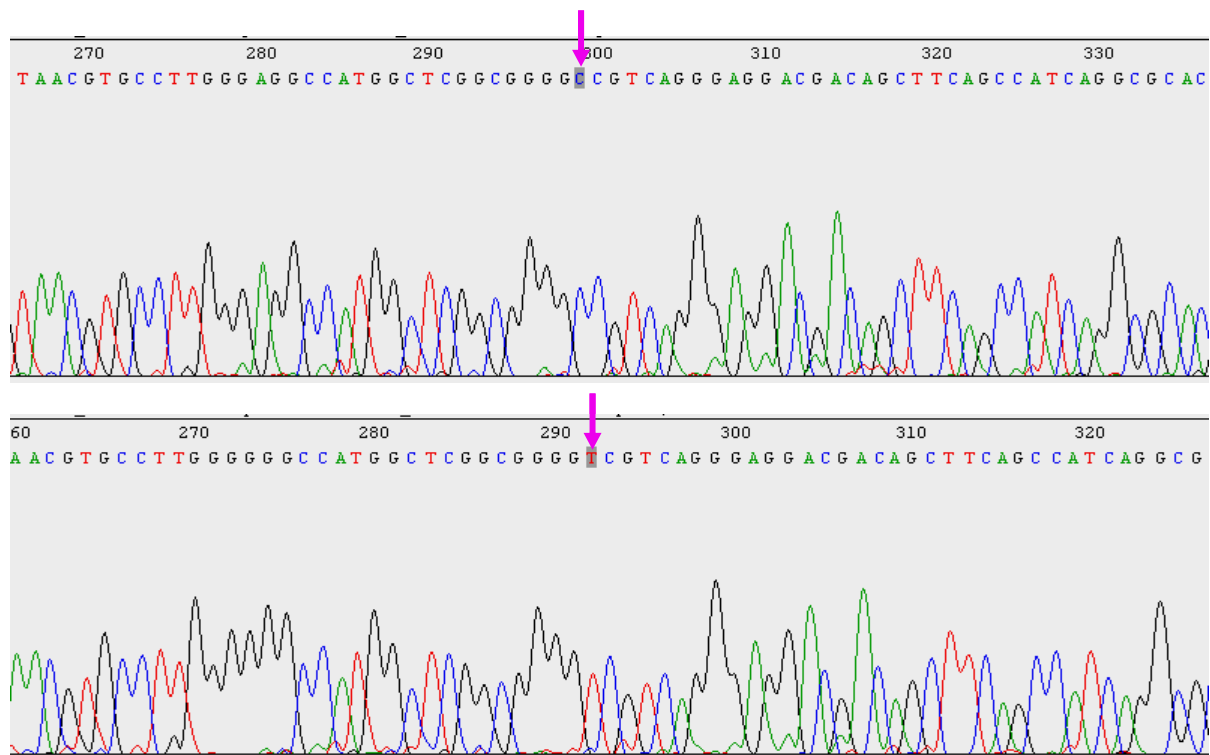


Figura 7. Cromatograma com a representação de um SNP (Polimorfismo de Único Nucleotídeo) (C↔T) em seqüências do mosquito *Aedes aegypti*.

SNPs são as formas mais freqüentes de variação genética em genomas de vertebrados e invertebrados; são codominantes e de fácil detecção quando comparados, por exemplo, com microssatélites (Taillon-Miller et al., 1998). São gerados por substituições de uma única base ou pequenos eventos de deleção/inserção, ocorrendo a uma taxa muito baixa de aproximadamente 10^{-7} alterações por sítio por geração, ou menos. SNPs podem estar presentes dentro de seqüências codificantes, regiões de genes não-codificantes, ou em regiões intergênicas. Como apenas uma pequena porcentagem dos genomas codifica proteínas, a maior parte deles localiza-se fora do DNA codificador. SNPs localizados dentro desta região têm particular importância porque é mais provável que as variações tenham um significado funcional, em nível do fenótipo. Entretanto, SNPs de seqüências codificantes não necessariamente acarretarão em mudanças na seqüência de aminoácidos de uma proteína que é produzida, devido à degeneração do código genético (Hammer, 1995).

Um SNP em que ambas as formas levam à mesma seqüência de polipeptídeos é denominado sinônimo (também denominado de mutação silenciosa) – e se uma seqüência diferente de polipeptídeos é produzida eles são denominados não-sinônimos. SNPs localizados em regiões não-codificantes do genoma ou em sítios sinônimos em regiões codificantes sem impacto direto no genoma fornecem ferramentas de grande utilidade para

estudos de evolução e genética de populações. Enquanto que, SNPs de sítios não sinônimos que podem alterar a estrutura e a função potencial das proteínas codificantes, representam marcadores alvo para se identificar variações genéticas associadas com características fenotípicas complexas (Hammer, 1995).

A descoberta de SNPs que afetam a atividade biológica tem se tornado cada vez mais importante e vem sendo facilitada pela disponibilidade de um vasto mapeamento do genoma. Os métodos de detecção de sítios únicos são, potencialmente, mais fáceis de automatizar e de aplicar na análise genética em larga escala. Deste modo, a informação gerada sobre os SNPs pode ser, posteriormente, utilizada em vários níveis nas análises do DNA como: marcadores no mapeamento genético; material de estudo na identificação de SNPs funcionais e seu respectivo fenótipo; ferramenta para estudos populacionais, por comparação do genoma de indivíduos de diferentes populações (Zhang & Hewitt, 2003).

Recentemente, a tecnologia disponível e a emergência de ferramentas de genotipagem de alta resolução têm acelerado o desenvolvimento de um grande número de programas de pesquisa, gerados a partir de projetos genoma, incluindo ferramentas de genotipagem *high-throughput* para identificar polimorfismos de único nucleotídeo (Syvanen, 2001). Isto os tem tornado marcadores genéticos de escolha para o mapeamento de alta-resolução, genética de populações e estudos de associação alélica (Berger et al. 2001; Black et al. 2001; Botstein & Risch 2003). Além disso, os estudos que empregam a detecção dos SNPs são utilizados para mapear e identificar genes específicos associados aos mosquitos de importância médica como *Ae. aegypti* (Morlais & Severson, 2003), *Anopheles gambiae* (Morlais et al., 2004) e *An. funestus* (Wondji et al., 2007).

O genoma do mosquito *An. gambiae*, por exemplo, foi seqüenciado utilizando três diferentes cepas desta espécie. Um total de 443.963 SNPs foram identificados ao longo de 278 Mb de comprimento, representando uma frequência de $1,6 \times 10^{-3}$, com uma média de 1 SNP a cada 125 pb codificantes. Altos níveis de polimorfismos foram identificados em genes relacionados com imunidade e receptores para o reconhecimento de patógenos, e esta quantidade de polimorfismos elevou as possibilidades de expansão dos estudos baseados nestes marcadores (Holt et al., 2002).

Consideramos que amplas análises de genomas utilizando marcadores nucleares poderiam facilitar a descoberta de fatores genéticos associados com características complexas relacionadas ao comportamento das populações naturais. Desta forma, estudando as variações genéticas do *Ae. aegypti* poderíamos correlacionar características

genéticas com fatores fenotípicos importantes que auxiliariam fornecendo dados para o desenvolvimento de estratégias para se reduzir a transmissão do vírus dengue. Por isso, neste trabalho nós descrevemos o desenho de oligonucleotídeos gene-específicos para amplificação de fragmentos de 9 genes nucleares com a presença de SNPs, além das condições de realização da genotipagem pela técnica de minisequenciamento em mosquitos *Ae. aegypti* individuais.

II. JUSTIFICATIVA

Dentre os agentes transmitidos pelo *Ae. aegypti*, a dengue é a que desempenha o papel de maior importância no âmbito da saúde pública. Cerca de dois bilhões de pessoas vivem em áreas onde ocorre a sua transmissão. Como não existem vacinas nem tratamento específico, a prevenção e o controle da doença dependem do controle do vetor. Inseticidas químicos têm sido largamente empregados para o controle do mosquito, entretanto, têm sido detectados processos de resistência, que afetam, direta e profundamente, as estratégias de controle e aumentam o risco de epidemias a cada ano em nosso país.

O conhecimento aprofundado sobre a estrutura genética e dinâmica populacional do *Ae. aegypti* é crucial, uma vez que populações geneticamente diferentes podem apresentar diferenças quanto a sua capacidade e competência vetoras. Assim, quando os organismos de interesse pertencem a populações de uma mesma espécie ou a espécies muito próximas, os marcadores nucleares serão informativos na medida em que sejam altamente variáveis. Por outra parte, a possibilidade de avaliar a variabilidade genética em vários cromossomos aumenta a probabilidade de recuperar informações relevantes sobre as dinâmicas de evolução do genoma. Finalmente, as inferências sobre estruturação populacional, fluxo gênico ou dinâmicas demográficas são mais robustas quando baseadas nas análises de *loci* independentes (não ligados), neutros (não submetidos a pressão de seleção), transmitidos por herança mendeliana e codominantes.

III. OBJETIVOS

GERAIS

1. Realizar estudo populacional em mosquitos da espécie *Aedes aegypti*;
2. Caracterizar a diversidade genética e a estrutura populacional de algumas populações naturais do mosquito *Ae. aegypti* coletadas no Brasil.

ESPECÍFICOS

1. Estimar o grau de variabilidade genética de populações do *Ae. aegypti*;
2. Demonstrar a utilização de marcadores SNPs na caracterização da estrutura genética de populações de *Ae. aegypti*;
3. Descrever a diversidade genética em genes nucleares do mosquito *Ae. aegypti*.

IV. MATERIAL E MÉTODOS

4.1 Estudo de seleção dos SNPs

Com o objetivo de investigar SNPs em genes nucleares, por comparação de seqüências, foram realizadas análises preliminares deste estudo, de mosquitos individuais, em 16 populações do *Ae. aegypti* no Brasil. A diversidade genética do *Ae. aegypti* foi estimada em regiões codificantes e não-codificantes de nove genes nucleares. Para assegurar o uso dos marcadores de único nucleotídeo em genética de populações, os SNPs selecionados foram, após triagem, utilizados para genotipagem de três populações do mosquito.

4.1.1 Amostragem

Amostras de populações do mosquito *Ae. aegypti* foram coletadas de janeiro a março de 2002, em cinco Estados brasileiros. Estas populações foram escolhidas porque já haviam sido comparadas em estudos anteriores, realizados em nosso laboratório utilizando marcadores RAPD, microssatélites e seqüenciamento de genes mitocondriais.

Para o estudo de triagem dos SNPs foram utilizados 15 indivíduos, escolhidos aleatoriamente, em 16 populações coletadas nas localidades de: Bauru (BAU-SP), Barretos (BRR-SP), Campinas (CAM-SP), Marília (MAR-SP), Presidente Prudente (PPR-SP), Ribeirão Preto (RPR-SP), Santos (SP), São José do Rio Preto (SJR-SP), Londrina LON-(PR), Maringá (MAG-PR), Juazeiro do Norte (JUA-CE), (gentilmente cedidas pela SUCEN); Botucatu (BOT-SP), Campo Grande (CGR-MS) (Prof. Dr. Newton G. Madeira); Maceió (MAC-AL), Porto Velho (PVE-RO), Vilhena (VIL-RO) (gentilmente cedidas pela Fundação Nacional de Saúde), além da linhagem de laboratório Rockfeller (RCK) (Figura 8; Tabela 1). Para o estudo de caracterização foram escolhidos indivíduos (machos e fêmeas) de três populações nas localidades de Botucatu, Juazeiro do Norte e Maceió.



Figura 8. Mapa do Brasil com as localidades de coleta das populações de *Aedes aegypti*.

Tabela I. Localização e coordenadas dos sítios de coleta do mosquito *Aedes aegypti* no Brasil.

Estado	Cidade	Latitude/Longitude
Alagoas	Maceió	-09°39'57"/35°44'07"
Ceará	Juazeiro do Norte	-07°12'47"/39°18'55"
Mato Grosso Sul	Campo Grande	-20°26'34"/54°38'47"
Paraná	Londrina	-23°18'37"/51°09'46"
Paraná	Maringá	-23°25'31"/51°56'19"
Rondônia	Vilhena	-12°44'26"/60°08'45"
Rondônia	Porto Velho	-08°45'43"/63°54'14"
São Paulo	Barretos	-20°33'26"/48°34'04"
São Paulo	Bauru	-22°18'53"/49°03'38"
São Paulo	Botucatu	-22°53'09"/48°26'42"
São Paulo	Campinas	-22°54'20"/47°03'39"
São Paulo	Marília	-22°12'50"/49°56'45"
São Paulo	Presidente Prudente	-22°07'32"/51°23'20"
São Paulo	Ribeirão Preto	-21°10'39"/47°48'37"
São Paulo	Santos	-23°57'39"/46°20'01"
São Paulo	São José do Rio Preto	-20°49'11"/49°22'46"

No insetário do Departamento de Parasitologia do IB da Unesp-Botucatu os ovos foram mantidos até o estágio adulto em condições adequadas (temperatura de $25 \pm 1^\circ\text{C}$, umidade relativa de $80 \pm 10\%$ e fotoperíodo de 12h). Foram utilizados para a pesquisa apenas mosquitos adultos, originados dos ovos coletados, obtidos em laboratório.

4.1.2 Preparação do DNA

O DNA genômico foi preparado de mosquitos individuais e inteiros, triturados com auxílio de um pistilo em tampão de extração utilizando-se resina Chelex100[®] *Molecular Biology Grade Resin* (Bio-Rad Laboratories), preparada a 5% conforme as recomendações do fabricante.

4.1.3 Seleção dos oligonucleotídeos iniciadores

Os genes analisados no estudo foram escolhidos por três motivos: 1) por estarem espalhados no genoma do *Ae. aegypti*, na tentativa de se evitar a ligação genética; 2) por se mostrarem envolvidos em processos importantes como: digestão, resistência a inseticidas, susceptibilidade a patógenos, resposta imune, etc; 3) para possibilitar futuros estudos de associação. Sequências específicas foram pesquisadas na base de dados do GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/index.html>), e utilizadas para desenho de oligonucleotídeos iniciadores para dezenove genes: AEG18 (*Transcription factor*), AelMUC1 (*Mucin-like protein*), APN (*Aminopeptidase N*), apolLp-II (*Apolipoprotein II*), CecA (*Cecropin A*), Chym (*Chymotrypsin*), DCE (*Dopachrome conversion enzyme*), DDC (*Dopa decarboxylase*), DefA (*Defensin A*), Ef-2 (*Elongation factor*), FerH (*Ferritin heavy chain*), LAP (*Lysosomal aspartic protease*), Mall (*Maltase-like I*), Na/K (*Sodium/potassium channel*), para (*DDT/pyrethroid resistance*), PGK (*Phosphoglycerate kinase*), CYP9J2 (*Cytochrome P450*), TSF (*Transferrin*) e TrypB (*Trypsin*) (Tabela II). Todas as sequências dos oligonucleotídeos utilizadas neste estudo foram confirmadas no Blast (*Basic local alignment search tool*) (<http://www.ncbi.nlm.nih.gov/BLAST>).

Tabela II. Seqüências do GenBank utilizadas na confecção dos oligonucleotídeos iniciadores, para cada gene nuclear do estudo.

Gene	Identidade	Número de Acesso no GenBank
AEG18	<i>Transcription factor</i>	AF326339 AY064076
MUCI	<i>Mucin-like protein</i>	AF387486 AY008350 AY064110 AY133345
APN	<i>Aminopeptidase N</i>	AF378117 AF390100 AY064078
apoLp-II	<i>Apolipophorin II</i>	AF038654
CecA	<i>Cecropin A</i>	AY064080
Chym	<i>Chymotrypsin</i>	AY038039 AY008349 AY064082 AF288384 AY064099 AY064098
DCE	<i>Dopachrome conversion enzyme</i>	AY064097 AY064096 AY064095
DDC	<i>Dopa decarboxylase</i>	AY064100 U27581
DefA	<i>Defensin A</i>	AF392802
Ef-2	<i>Elongation factor 2</i>	AF331798 AY040342.1 AY064103
FerH	<i>Ferritin heavy chain</i>	AF326341 AF126431 AY064105
LAP	<i>Lysosomal aspartic protease</i>	M95187
Mall	<i>Maltase-like I</i>	M30442 M30443
NaK	<i>Sodium/potassium channel</i>	AY663378 AY663385
para	<i>DDT/pyrethroid resistance</i>	AF534112 AY663377
PGK	<i>Phosphoglycerate kinase</i>	AY043171 AY064542
CYP9J2	<i>Cytocrome P450</i>	AY064093 AY064092
Tsf	<i>Transferrin</i>	AF387489 AY064537
TryB	<i>Trypsin</i>	M77814 X64362 M77814

Seqüências de cada gene foram alinhadas utilizando o programa ClustalX versão 1.8 (Thompson et al., 1997). Pares de oligonucleotídeos foram construídos para amplificarem produtos de 400-1200 pb (pares de bases) utilizando o programa GeneRunner versão 3.05 (Hasting Software, Inc., 1994). A melhor temperatura de reação para cada par de oligonucleotídeos foi determinada via gradiente de temperatura utilizando-se um termociclador *Eppendorf MasterCycler Gradient* (Eppendorf, Hamburg, Germany). Cada oligonucleotídeo testado foi selecionado conforme a qualidade obtida na amplificação e polimorfismos presentes nas populações naturais do mosquito *Ae. aegypti*.

Oligonucleotídeos que apresentaram amplificação inespecífica ou ineficiente para alguns indivíduos testados foram descartados.

4.1.4 Condições de Amplificação

Reações de amplificação foram realizadas com volume final de 25 μ l, contendo 3 μ l do DNA molde, 50 mM Tris-HCl, pH 8.0, 50 mM KCl, 2,5 mM $MgCl_2$, 2,5 mM de dNTPs (PCR nucleotide Mix, Invitrogen), 20 μ M de cada oligonucleotídeo específico e 1.5 U de Platinum® *Taq* DNA polimerase (Invitrogen). As reações foram realizadas em termociclador *Eppendorf MasterCycler Gradient* (Eppendorf, Hamburg, Germany). Os ciclos de amplificação foram os mesmos para cada par de oligonucleotídeos programados para: 1 ciclo inicial de 95°C por 5 min, seguidos por 30 ciclos de 95°C por 1 min, 56°C por 1 min, 72°C por 2 min, e um ciclo de extensão final de 72°C por 10 min. Dez microlitros do produto de PCR foram submetidos à eletroforese em gel de agarose 1%, corado com brometo de etídio 0,5 μ g/ml segundo Sambrook et al. (1989). Os fragmentos amplificados foram comparados com o DNA *Ladder* de 250 pb (Invitrogen), utilizado como padrão. O restante dos produtos de PCR foi utilizado para o seqüenciamento utilizando-se cada oligonucleotídeo iniciador específico *forward* e *reverse*.

4.1.5 Purificação dos produtos de PCR

Os fragmentos de PCR de mosquitos individuais foram purificados por reação enzimática adicionando-se 2 μ l da enzima ExoSAP IT (GE Healthcare) a 8 μ l do produto de PCR de acordo com as recomendações do fabricante, sendo a seguir submetidos ao seqüenciamento.

4.1.6 Reação de Seqüenciamento

As seqüências de DNA foram determinadas em seqüenciador automático ABI 377 (Applied Biosystems®) utilizando-se 4 μ l de 2,5 x Save Money (400 mM Tris-HCl pH 9,0, 10 mM $MgCl_2$), 4 μ l BigDye™ *Terminator Cycle Sequencing Ready Reaction Kit* versão 3.1 (Applied Biosystems, Foster City, CA), 3,2 pmol dos oligonucleotídeos *forward* e *reverse*, 4 μ l do produto de PCR purificado e H_2O q.s.p.. As reações foram realizadas em termociclador *Eppendorf MasterCycler Gradient* (Eppendorf, Hamburg, Germany) com os ciclos de temperatura programados para: 25 ciclos de 95°C por 10 seg, 50°C por 5 seg, 60°C por 4, com rampa de 1°C/seg, como recomendado pelo fabricante. Após a amplificação as

amostras foram mantidas a 4°C até a precipitação. Para cada amostra foram realizadas duas reações, sendo uma para o oligonucleotídeo *forward* e outra para o *reverse*.

4.1.7 Precipitação das reações de seqüenciamento

Às reações de seqüenciamento foram adicionados 80 µl de isopropanol 65%, incubando a T.A. por 20 min. Em seguida, as amostras foram centrifugadas a velocidade de 10.000xg por 25 min a T.A.. O isopropanol foi removido invertendo os tubos e, a seguir, adicionou-se 200 µl de etanol 60% e centrifugou-se a velocidade de 10.000xg por 5 min a T.A.. Todo o etanol foi removido com auxílio de micropipeta, pois qualquer etanol residual resultaria em manchas fluorescentes. As amostras foram secas em T.A. e o produto ressuspenso em 2 µl de tampão de amostra contendo Formamida Hi-Di™ (Applied Biosystems) + *Loading Buffer* (25 mM EDTA pH 8,0 contendo 50 mg/ml de Blue Dextran) (5:1). No momento da aplicação em seqüenciador automático ABI PRISM® 377 (Applied Biosystems, USA) as amostras foram aquecidas a 95°C por 5 min e rapidamente transferidas para o gelo.

4.1.8 Alinhamento das seqüências

Os eletroferogramas das seqüências *forward* e *reverse*, gerados automaticamente, foram inspecionados visualmente e heterozigotos, detectados. Seqüências de diversos indivíduos de diferentes populações foram alinhadas utilizando os programas MERGER (<http://bioweb.pasteur.fr/seganal/alignment/intro-uk.html>) e CLUSTALX versão 1.8 (Thompson et al., 1997), sendo identificadas utilizando-se o BLAST e comparadas com outras disponíveis no GenBank.

4.1.9 Análises dos SNPs (Polimorfismos de único nucleotídeo)

O alinhamento das seqüências foi realizado para possibilitar a identificação visual dos SNPs como transições ou transversões em regiões codificantes e não-codificantes. SNPs localizados dentro de regiões codificantes foram classificados como sinônimos e não-sinônimos e a posição de cada um deles no códon também foi determinada.

4.2 Estudo de genotipagem por SNPs

Após o estudo de seleção, foram utilizados SNPs de nove genes nucleares na caracterização das populações de Botucatu, Juazeiro do Norte e Maceió. A localização cromossômica de cada um destes genes está indicada no mapa físico do mosquito *Ae. aegypti* (Figura 9).

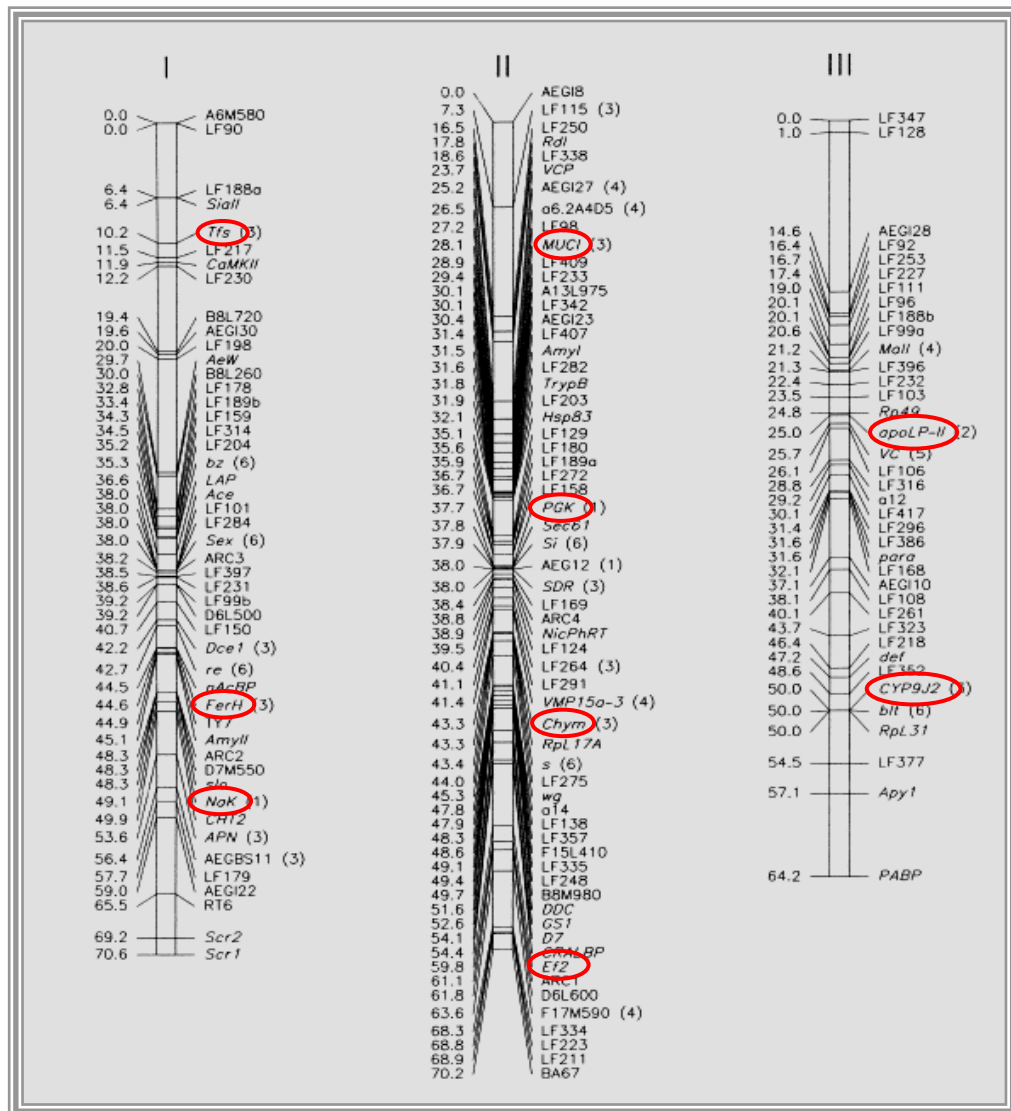


Figura 9. Localização cromossômica dos genes selecionados para o estudo de genotipagem por SNPs do mosquito *Aedes aegypti*. Fonte: Severson et al., 2002.

4.2.1 Descrição dos genes selecionados

Cromossomo 1

Tsf - Tranferrin. Soro transferrina (Tsf) de vertebrados é uma proteína utilizada para o transporte de ferro. Proteínas Tsf, bem como os cDNAs que codificam Tsfs, têm sido identificados para várias espécies de insetos (Nichol et al., 2002). Nestes, a função exata destas proteínas não está bem determinada, mas sabe-se que elas estão envolvidas no transporte de íons ferro, na oogênese, e na defesa imune inata contra parasitas e patógenos. Pesquisas sugerem que a Tsf é expressa relativamente tarde no ciclo de vida do mosquito, sendo produzida principalmente no corpo gorduroso e secretada dentro da hemolinfa. A molécula de transferrina é capaz de transportar íons ferro na forma férrica (Fe^{3+}) (Harizanova et al., 2005). Para este gene foram inicialmente selecionados 3 SNPs.

FerH - Ferritin heavy chain. A ferritina é a mais importante proteína de reserva do ferro e é encontrada especialmente nas células envolvidas na síntese de compostos férricos, no metabolismo e na reserva do ferro. A ferritina é o meio pelo qual os insetos hematófagos lidam com a alta concentração de íons ferro decorrente da alimentação sanguínea necessária à maturação dos ovos (Geiser et al., 2003). A ferritina citoplasmática dos vertebrados geralmente consiste de 24 subunidades representadas por uma cadeia pesada (21 kDa) e uma cadeia leve de polipeptídeos (19 kDa) (Chasteen & Harrison, 1999). Dados recentes revelaram que a ferritina secretada por insetos retém uma estrutura similar de 24 subunidades (Hamburger et al., 2005), mas a composição das subunidades difere. No caso do mosquito *Ae. aegypti* a ferritina é composta das subunidades 24, 26, 28 e 30 kDa (Dunkov et al., 1995). Para este gene foram inicialmente selecionados 3 SNPs.

NaK - Sodium/potassium channel. Sódio-potássio ATPase (ou NaK) é uma bomba iônica localizada na membrana plasmática e responsável pela manutenção das diferenças de potencial eletroquímico através da membrana. NaK é um heterodímero composto de subunidade-alfa catalítica ativa (contendo aproximadamente 1000 aminoácidos residuais) e uma subunidade beta glicosilada, de função desconhecida (Reeves & Yamanaka, 1993; Emery et al., 1998). Em insetos, embora a enzima seja abundante no sistema nervoso, a expressão em outros tecidos varia filogeneticamente e fisiologicamente. A expressão de Na^+/K^+ em insetos de modo geral é sugestiva de um papel primário em tecidos nervosos (Jungreis & Vaughan, 1977; Lebovitz et al., 1989; Baumann & Takeyasu, 1993). Em insetos

que se alimentam de sangue a Na/K será expressa em alta concentração dentro do intestino médio, onde age para remover Na⁺ e a água do lúmen intestinal (Emery et al., 1998). Durante os primeiros 20 minutos depois da alimentação sanguínea o mosquito *Ae. aegypti* excreta um volume de urina igual ao volume de hemolinfa antes da alimentação (Williams et al., 1983). Em *An. freeborni*, a diurese se inicia imediatamente após o início da alimentação (Nijhout & Carrow, 1978). Para este gene foi inicialmente selecionado 1 SNP.

Cromossomo 2

MUCI - Mucin-like protein. Mucinas são proteínas que contêm regiões altamente glicosiladas e longas seqüências de aminoácidos repetidos, ambas as estruturas altamente mutáveis. Tal variabilidade em estrutura também pode ser traduzida para diferenças em funções. No mosquito *Ae. aegypti* o gene que codifica para MUCI tem sido nomeado como mucina intestinal (AelMUC1), que é específica de intestino médio de fêmeas adultas e compartilha similaridades com proteínas da Matriz Peritrófica. A estrutura da AelMUC1 sugere que esta proteína é uma peritrofina de mosquitos fêmeas adultos de *Ae. aegypti*. A localização cromossômica e os polimorfismos alélicos aumentam a possibilidade de que a proteína de intestino médio esteja envolvida nas interações parasita vetor. Alguns estudos sugerem que ela poderia estar envolvida na susceptibilidade do mosquito ao *Plasmodium gallinaceum* e, ultimamente, na competência vetorial em transmiti-lo aos vertebrados (Morlais & Severson, 2001). Para este gene foram inicialmente selecionados 4 SNPs.

PGK - Phosphoglycerate kinase. Phosphoglycerato quinase é uma enzima monomérica, chave da cadeia glicolítica, que catalisa a conversão de glucose (ou frutose) em piruvato. É utilizada em sete passos da glicólise, sendo necessária para a geração de ATP no estágio terminal da cadeia glicolítica. Uma molécula simples de ATP é gerada durante esta reação. Estudos com espécimes mutantes de *Drosophila* que tiveram diminuição na produção de PGK apresentaram comportamentos anormais, sugerindo que a redução dos níveis de ATP tem efeito significativo na locomoção das larvas, consistentes com funções sinápticas secundárias anormais provocadas pela diminuição de PGK (Wang et al., 2004). Para este gene foram inicialmente selecionados 3 SNPs.

Chym - Chymiotrypsin. Chymiotripsinas são conhecidas pelo papel desempenhado na digestão dos insetos (Terra & Ferreira, 1994). Como a *early* tripsina (tripsina inicial) (Noriega

et al., 1996), o gene da chimiotripsina é primeiro transcrito no início da vida adulta; o nível de RNA mensageiro (RNAm) atinge o máximo antes da alimentação e o produto de proteínas não é traduzido até a ingestão de alimento. Diferente da *early* tripsina, o RNAm não diminui após o repasto e a proteína chimiotripsina, a expressão de proteínas e a atividade enzimática são produzidas continuamente durante a digestão do alimento. O fato de que a expressão de proteínas e a atividade enzimática aumentem em extratos brutos depois da ingestão de proteínas, sugere um envolvimento na digestão destas (Jiang et al., 1997). Para este gene foi inicialmente selecionado 1 SNP.

Ef-2- Elongation Factor 2. A biosíntese de proteínas ocorre em três fases: início, elongação e terminação, cada uma das quais envolve um grupo distinto de fatores de proteínas. O Fator de Elongação-2 (Ef-2) é uma grande proteína com mais de 800 aminoácidos que estabelece complexas interações com os ribossomos a fim de catalisar as mudanças conformacionais necessárias para a elongação da cadeia de polipeptídeos. Ele se liga aos ribossomos e catalisa a translocação do tRNA do aminoácido para o sítio do peptídeo, participando da síntese de proteínas (Grinblat et al., 1989). Tem uma sequência de proteínas altamente conservada na qual mudanças evolucionárias fornecem sinais filogenéticos. Por isso tem sido utilizada como genes candidatos a estudos de análises filogenéticas principalmente em artrópodes (Friedlander et al., 1992; 1994). O gene codificante para Ef-2 não exibe uma expressão diferencial para o mosquito *Ae. aegypti*. Para este gene foram inicialmente selecionados 2 SNPs.

Cromossomo 3

apoLP-II – Apolipophorin II. O transporte extracelular de lipídios não solúveis em água, via sistema circulatório dos animais, constitui uma função vital que utiliza lipoproteínas complexas altamente especializadas. Em insetos, uma única lipoproteína, lipoforina, é utilizada para transportar uma variedade de biomoléculas lipofílicas entre tecidos. Componentes de apolipoproteínas em genes servem para estabilizar os lipídios e modular o metabolismo de partículas de lipoproteínas. Estudos demonstram que duas apolipoproteínas de lipoforinas, não permutáveis, originam-se de uma proteína precursora comum, apolipoforina II/I (apoLp-II/I) (Chapman, 1986; Weers et al., 1993; Babin et al., 1999). Apolipoforina-II é, portanto, um tipo de lipoproteína que nos insetos é denominada lipoforina, sendo utilizada para efetuar o transporte de uma variedade de moléculas

lipofílicas na hemolinfa. Este veículo de transporte multifuncional é sintetizado no corpo gorduroso que, combina muitas das funções exercidas pelo fígado e tecido adiposo de vertebrados. Para este gene foram inicialmente selecionados 3 SNPs.

CYP9J2 – Citocrome P450. Constituem uma superfamília multigênica de enzimas onipresentes na natureza com uma grande diversidade de funções. As citocromos P450 têm sido encontradas em todos os organismos vivos examinados, incluindo mamíferos, peixes, artrópodes, fungos, plantas e bactérias (Agosin, 1985). Em insetos, mais de 300 P450s já foram identificadas e estão distribuídas através de 27 famílias *CYP* em mais de 70 famílias de genes P450 identificadas. Na maioria dos casos, elas estão envolvidas tanto no metabolismo endógeno como no metabolismo de xenobióticos (Nelson et al., 1996; Feyereisen, 1999). Em insetos suas funções incluem: o metabolismo de esteróides, embora sejam particularmente conhecidas pela resistência a inseticidas e pela tolerância a toxinas de plantas (Gould, 1984). Em mosquitos da espécie *Ae. aegypti* este gene é especificamente expresso em intestino médio de fêmeas adultas. Para este gene foram inicialmente selecionados 2 SNPs.

4.2.2 Caracterização dos SNPs por um método semi-automático

O método de genotipagem deve fornecer estimativas acuradas da frequência de alelos, além de ter tempo e custo baixos para sua realização. Atualmente existem diferentes métodos disponíveis para genotipagem por SNPs em amostras individuais (Syvanen, 2001). O método de escolha para este trabalho se baseia na extensão alelo-específica ou minisequenciamento de um oligonucleotídeo adjacente ao sítio SNP denominado SNaPshot™ (Applied Biosystems). A utilização do kit SNaPshot™ permite a detecção de mutações, pelo método dideoxi marcado com fluorescência e técnica de extensão de uma única base que tem se mostrado altamente eficiente (Le Hellard et al., 2002).

4.2.3 Seleção dos oligonucleotídeos flanqueadores

No presente estudo cada oligonucleotídeo flanqueador foi construído adjacente aos sítios mutacionais, com até 19 pb, utilizando o programa GeneRunner versão 3.05 (Hasting Software, Inc., 1994) de acordo com as recomendações da técnica de minisequenciamento (Applied Biosystems). Para facilitar a discriminação dos SNPs de um mesmo gene alguns oligonucleotídeos foram construídos com diferentes tamanhos, pela inserção de cauda poli C na extremidade 5'. A melhor temperatura de reação para cada oligonucleotídeo foi

determinada utilizando o termociclador *Eppendorf MasterCycler Gradient* (Eppendorf, Hamburg, Germany).

4.2.4 Reação de minisequenciamento

Produtos de PCR foram purificados por reação enzimática com ExoSAP IT (GE Healthcare) de acordo com as recomendações do fabricante. Reações de minisequenciamento foram montadas com volume final de 10 µl, contendo 2 µl do produto de PCR purificado, 1 µl de 2,5x save money (400 mM Tris-HCl pH 9,0, 10 mM MgCl₂), 1 µl de Snapshot™ Multiplex mix, 0,4 pmol de cada oligonucleotídeo e água q.s.p.. As reações foram realizadas em termociclador *Eppendorf MasterCycler Gradient* (Eppendorf, Hamburg, Germany) programado para: 40 ciclos de 95°C por 10 seg, 58°C por 5 seg, 60°C por 30 seg, com rampa de 1°C/seg.

4.2.5 Purificação das reações de minisequenciamento

Reações de minisequenciamento foram purificadas com 1 U de *Shrimp Alkaline Phosphatase* (GE Helthcare) conforme as recomendações do fabricante. A seguir, as reações foram diluídas 1:1 com Formamida Hi-Di™ (Applied Biosystems) + *Loading buffer* (25 mM EDTA pH 8,0 contendo 50 mg/ml de Blue Dextran). No momento da aplicação em seqüenciador automático ABI PRISM® 377 (Applied Biosystems) as amostras foram aquecidas a 95°C por 3 min e rapidamente transferidas para o gelo.

4.3 Análises da Estrutura Genética Populacional

4.3.1 Obtenção dos dados

Após a corrida eletroforética, os SNPs foram identificados com base no tamanho do produto amplificado e na fluorescência. Planilhas com dados genotípicos diplóides foram obtidas por contagem direta e montadas no formato Excel, sendo convertidas com auxílio do programa Convert (Glaubitz, 2004), que auxilia na utilização de programas como Arlequin 3.0 (Excoffier et al., 2005) e Structure 2.0 (Pritchard et al., 2000).

4.3.2 Identidade de nucleotídeos e aminoácidos

A identidade de nucleotídeos e aminoácidos foi obtida pela contagem direta dos sítios (nucleotídeos ou aminoácidos) conservados das seqüências de DNA do mosquito *Ae.*

aegypti, quando comparadas par a par com outras da mesma espécie depositadas no GenBank.

4.3.3 Padrões de substituição

As substituições sinônimas e não-sinônimas foram determinadas por contagem direta também por comparação com seqüências disponíveis no GenBank. O programa DNAsp 4.1 (Rozas et al., 2003) foi utilizado para determinar o número de sítios segregantes, o número de haplótipos, a diversidade de haplótipos (H_d), o número médio de substituições nucleotídicas por sítio (π); e o número médio de diferenças nucleotídicas (κ), das seqüências obtidas para cada gene.

4.3.4 Desequilíbrio de ligação

O programa Arlequin 3.0 permitiu o cálculo do desequilíbrio de ligação (LD) entre todos os pares de alelos. O LD foi calculado para determinar quais alelos dos vários *loci* SNPs analisados segregavam independentemente uns dos outros. Esta análise é importante porque mesmo que dois *loci* ligados possam estar nas proporções genotípicas de Hardy-Weinberg, alelos em cada um deles podem estar associados de forma não aleatória com alelos de outro *locus*, ou seja, podem estar sendo transmitidos juntos com uma freqüência diferente da esperada sob segregação independente. Alelos que estão extremamente ligados entre si, localizados em um mesmo cromossomo, tendem a não ser separados durante a meiose, aumentando a chance da ocorrência de gametas com freqüências diferentes daquelas esperadas sob segregação independente. Tal fenômeno acontece devido à redução na freqüência de recombinação na região entre os *loci*. Assim, quando dois marcadores genéticos são transmitidos juntos com uma freqüência diferente da esperada sob segregação independente (não-aleatória), diz-se que os marcadores estão em desequilíbrio de ligação (Liu, 1998).

4.3.5 Freqüências alélicas

As freqüências alélicas foram obtidas por contagem direta através do pacote de programas Arlequin 3.0. Este cálculo pode ser obtido somando-se o número de indivíduos heterozigotos para um determinado alelo mais duas vezes o número de indivíduos homozigotos, e dividindo-se o resultado pelo dobro do número total de indivíduos. De acordo com a fórmula:

$$F(p) = (p + 2 pq)/2N$$

onde,

- $F(p)$: frequência do alelo p ;
- p : número de indivíduos homozigotos do alelo p ;
- pq : número de indivíduos heterozigotos;
- N : número de indivíduos na amostra.

A avaliação da frequência de um alelo particular em uma população, chamada frequência gênica ou alélica, é considerada fundamental nos estudos evolutivos, pois a mudança genética de uma população pode ser avaliada pela mudança nas suas frequências gênicas (Nei, 1978). O conhecimento da frequência de heterozigotos apresenta importância na medida em que cada heterozigoto carrega diferentes alelos, os quais demonstram a existência de variação genética na população (Weir, 1996).

4.3.6 Equilíbrio de Hardy-Weinberg

A diversidade gênica é a medida que indica a heterozigosidade esperada para um dado *locus* sob equilíbrio de Hardy-Weinberg e, consiste na proporção esperada de indivíduos heterozigotos para as frequências alélicas observadas. A verificação do equilíbrio de Hardy-Weinberg (HW) foi feita pela comparação da frequência observada dos genótipos com a frequência esperada em equilíbrio, calculada utilizando o programa Arlequin 3.0, que aplica o método de Guo & Thompson (1992). O F_{IS} (coeficiente de endogamia) é utilizado para estimar a redução mínima em heterozigosidade de um indivíduo devido ao cruzamento aleatório, dentro da população, e foi calculado para testar as proporções do equilíbrio de HW entre os genótipos obtidos, de acordo com a fórmula:

$$F_{IS} = 1 - (\text{heterozigosidade observada} / \text{heterozigosidade esperada})$$

onde,

- Se existe um excesso de heterozigotos, então $F_{IS} < 0$;
- Se existe um excesso de homozigotos, então $F_{IS} > 0$;
- O teste do qui-quadrado foi utilizado para testar a hipótese nula, onde $F_{IS} = 0$.

O princípio de HW afirma que as frequências gênicas e genotípicas de um único *locus* estarão fixadas e em equilíbrio numa população numerosa, que não está em processo evolutivo franco, e após uma geração de acasalamento randômico.

4.3.7 Análise da estrutura genética populacional

No presente estudo, a análise da estrutura populacional foi calculada empregando a análise de variância molecular (AMOVA), conforme descrito por Excoffier et al., (1992) utilizando o programa Arlequin 3.0, que permite verificar o grau de significância genética inter e intrapopulacional. Este método estatístico se baseia na geração de estimativas da variância genética em diferentes níveis hierárquicos (ou seja, de como a diversidade genética encontra-se distribuída dentro das populações e entre as populações), além de valores de estatística F , que indicam o grau de subdivisão geográfica.

Os parâmetros estatísticos foram calculados de acordo com Weir & Cockerham (1984). O F_{ST} (índice de fixação) é utilizado para estimar a extensão da diferenciação genética entre as subpopulações (agrupamentos). O índice F_{ST} entre pares de populações (*pairwise* F_{ST}) também foi calculado, e pode ser utilizado como uma estimativa da distância genética entre elas. Todos os valores foram verificados através de testes não paramétricos de permutação (Excoffier et al., 1992). O valor de P igual a 0,05 foi adotado como limiar de significância para todos os testes estatísticos realizados.

4.3.8 Estruturação populacional

O programa Structure 2.0 (Pritchard et al., 2000) foi utilizado na análise de estruturação de populações com base em dados de genotipagem *multilocus*. A utilização deste teste possibilita a construção de *clusters* genéticos e estima a fração do genótipo de cada indivíduo que pertence a cada *cluster* na ausência de qualquer informação prévia sobre a estrutura da população. A análise se baseia na combinação de todos os genótipos individuais *multilocus* e calcula a probabilidade de um número pré-determinado de *clusters* (K), assumindo que misturas (híbridos) podem existir (ex. não se assume o isolamento genético). A seguir, obtém-se o valor mais provável de K pela análise das probabilidades e distribuição dos genótipos. Os resultados são convertidos para um gráfico com a probabilidade do ancestral de cada *cluster* para cada indivíduo.

As aplicações deste método incluem inferências sobre a presença de populações distintas, designação de indivíduos às populações, estudo de zonas de híbridos, identificação

de migrantes e indivíduos não-puros, e cálculos de frequências alélicas de população em situações onde muitos indivíduos são migrantes ou híbridos. O poder deste método de associação estruturada baseia-se na análise efetiva da estrutura da população, agrupamento correto e inclusão dos indivíduos em grupos apropriados, utilizando um pequeno número de marcadores moleculares não-ligados, além da identificação de migrantes ou indivíduos com mistura genética. Em consequência, os indivíduos são atribuídos (probabilisticamente) a uma população, ou a várias se seus genótipos indicam que são mestiços (Pritchard & Wen, 2003). Este programa foi inicialmente utilizado com sucesso em estudos envolvendo populações humanas (Zöllner et al., 2004; Rosenberg et al., 2002), testes de associação (Falush et al., 2003; Wall & Pritchard, 2003; Pritchard & Rosenberg, 1999), e tem sido também aplicado em estudos de estrutura populacional de mosquitos do gênero *Anopheles* (Boyd & Foley 2007; Moreno et al. 2007) e para a espécie *Culex pipiens* (Fonseca et al., 2004).

V. RESULTADOS

5.1 Análises dos oligonucleotídeos

Após os testes dos dezanove oligonucleotídeos construídos no início do projeto, nove (47,6%) foram selecionados por amplificarem produtos de PCR específicos e apresentarem SNPs nas seqüências. De todos os oligonucleotídeos construídos, dez (52,7%) foram descartados por não apresentarem o padrão necessário para seleção. Desta forma, os oligonucleotídeos dos genes AEG18, APN, CecA, DCE, DDC, DefA, LAP, Mall, para e TrypB foram testados mas não utilizados neste estudo. A tabela III contém os oligonucleotídeos das regiões amplificadas para os genes: apolLP-II (AAEL009955), Chym (AAEL003060), Ef-2 (AAEL004500), FerH (AAEL007385), NaK (AAEL012062), MUCI (AAEL004798), CYP9J2 (AAEL006805), PGK (AAEL004988) e Tsf (AAEL015639).

Tabela III. Oligonucleotídeos utilizados para amplificar genes nucleares contendo polimorfismos de único nucleotídeo.

Gene, VectorBase gene ID n.	Oligonucleotídeo	Seqüências dos oligonucleotídeos
Lipoprotein N-terminal domain AAEL009955	ApollpF	5'-GTAAGTCTATTTCGATGC-3'
	ApollpR	5'-TCAGAACAGCGACGATAG-3'
Serine-type endopeptidase AAEL5576966	ChymF	5'-TCCCTACCAAAATCTCGCTGC-3'
	ChymR	5'-GAGCACTGTCGTTGTCCCA-3'
Elongation factor AAEL004500	Ef2F	5'-GATTGAGTTTCCGGTGTG-3'
	Ef2R	5'-TCGTCA TCCTTGGTCTTG-3'
Ferritin subunit 1 AAEL007385	FerHF	5'-AAGTATGCCGCCCTATTTTGC-3'
	FerHR	5'-GGAGCGCTCTTCTTCATCT-3'
Mucin-like protein AAEL004798	MUCF	5'-GCAAAATGTCCCGATATCTTC-3'
	MUCR	5'-ACGGTTTCACGGTCCTTAG-3'
Sodium/potassium channel AAEL012062	NaKF	5'-CCCCGAATGGGTCAAGTTC-3'
	NaKR	5'-GTCAACCGATGACCCAGATC-3'
Cytocrome P450 AAEL006805	P450F	5'-CGAGGTTGGCACTGTTGG-3'
	P450R	5'-CGGTAACATCGACTTTTGC-3'
Phosphoglycerate kinase AAEL58082	PGKF	5'-CAAGGTGAACGAAATGATC-3'
	PGKR	5'-CGTGTGATACCTGAGATT-3'
Transferrin AAEL015639	TsF	5'-GTAGTGGCAGTGCTTCTG-3'
	TsR	5'-CTGCCCTTGCCTTTCATG-3'

5.2 Análises dos genes do estudo

Na figura 10 está apresentada a foto de um gel de agarose com os fragmentos de DNA amplificados do gene Tsf pelos oligonucleotídeos selecionados.

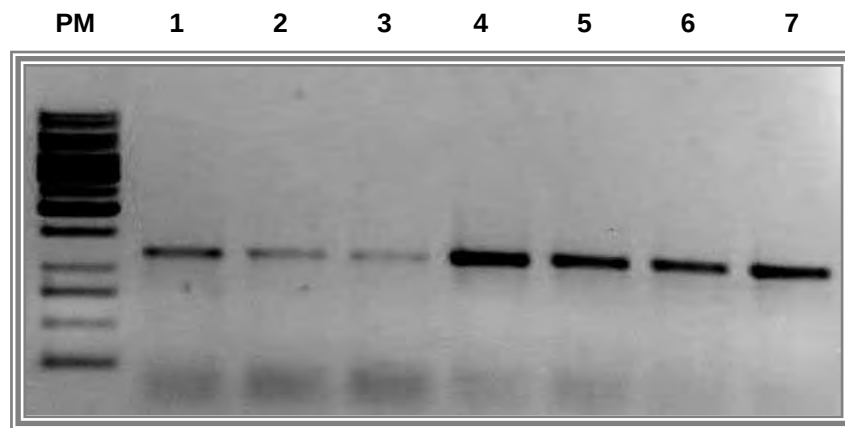


Figura 10. Visualização em gel de agarose a 1% dos produtos de PCR do gene TSF. *Lanes* de 1 a 7: espécimes provenientes de Botucatu; PM: Marcador de Tamanho Molecular 250 pb.

Após os ciclos de amplificação, fragmentos de nove genes foram seqüenciados para diversos espécimes das dezesseis populações naturais do mosquito *Ae. aegypti* (Tabela IV).

Tabela IV. Genes, número de seqüências, indivíduos e tamanho das seqüências utilizadas para validação dos oligonucleotídeos.

Gene	N	Indivíduos	L	C
apolp-II	11	BOT7, CGR7, MAC7, MAR7, MAG4, SAN4, SJR4, PPR4, JUA7, RPR4, RCK6	350	332
Chym	10	BOT2, BOT3, CGR1, CGR2, CGR3, CGR9, MAC4, MAC5, MAC6, MAC7	525	502
Ef-2	13	BOT7, CGR7, MAC7, MAG4, PVE7, PPR1, RPR4, SAN4, SJR1, SJR4, RCK6, RCK7	390	381
FerH	8	CGR1, CGR2, CGR3, CGR9, MAC4, MAC5, MAC6, MAC7	530	510
MUCI	6	BAU7, CGR7, LON10, PPR4, RPR4, RCK7	500	445
Na/k	14	BAU5, BAU6, BRR5, CAM4, CAM5, MAR1, PPR1, PVE5, PVE6, SAN5, SJR1, PPR1, RPR7, VIL4	497	387
PGK	12	BOT7, CGR7, MAC7, MAG4, PVE7, PPR4, RPR1, RPR4, RCK6, RCK7, SJR1, SJR4	470	460
CYP9J2	6	MAC3, MAC4, MAC6, MAC7, MAC8, SAN10	360	335
TSF	6	BAU4, BAR4, CAM4, PVE4, PVE6, SAN5	1199	536

N = número de indivíduos analisados; L = tamanho do fragmento de PCR (pb); C = comprimento da seqüência alinhada (pb).

SNPs foram detectados em todos os *loci* examinados para os nove genes selecionados. A comparação das seqüências permitiu a identificação de 71 SNPs no total de 3.888 pb examinados. O maior número de SNPs foi encontrado no fragmento do gene PGK, com 45 SNPs/Kb, e o menor no fragmento do gene NaK com apenas 2 SNPs/Kb (Tabela V).

Tabela V. SNPs identificados em regiões codificantes e não-codificantes dos 9 genes nucleares do mosquito *Ae. aegypti*.

Gene	N	S _T	S _A	S _S	SNPs/kb
apoLp-II	11	8	3	1	24
Chym	10	2	1	1	16
Ef-2	13	2	2	1	05
FerH	8	14	3	2	27
MUCI	6	12	4	1	18
NaK	14	1	1	1	02
CYP9J2	6	3	2	1	09
PGK	12	21	4	2	45
TSF	6	8	2	1	24
Total	86	71	22	11	18

N = número de seqüências; S_T = SNPs totais; S_A = SNPs analisados; S_S = SNPs selecionados.

5.3 Identidade de nucleotídeos/aminoácidos

Foram identificados 59 SNPs em regiões codificantes e 12 SNPs em regiões não-codificantes, além de 35 polimorfismos por inserção/deleção (indels). A média da ocorrência dos SNPs nas seqüências analisadas foi de um a cada 55 pb, ou de 18 SNPs/kb. A identidade dos nucleotídeos e aminoácidos foi testada visualmente para cada seqüência obtida no estudo. Como o número de seqüências analisadas por gene foi pequeno, em alguns casos, apenas 2 haplótipos foram detectados entre as populações naturais (ex. NaK, FerH). Entretanto, apesar disso, 14 SNPs foram identificados para o gene FerH. A diversidade média do número de nucleotídeos por gene foi de $9,9 \times 10^{-3}$,

variando de $5,3 \times 10^{-3}$ para o gene NaK a 75×10^{-3} para o gene FerH. O número médio de substituições nucleotídicas por sítio (π) foi maior para o gene FerH, que também apresentou o maior número médio de diferenças nucleotídicas (κ), apesar do baixo número de haplótipos detectados (Tabela VI).

Tabela VI. Diversidade de nucleotídeos para os nove genes nucleares do mosquito *Aedes aegypti*.

Gene	S	NH	Hd	π	K
apoLp-II	8	7	0,872	0,0063	2,103
Chym	3	4	0,711	0,0002	1,022
Ef-2	2	4	0,758	0,0026	1,022
FerH	14	2	0,536	0,0147	7,500
MUCI	12	3	0,600	0,0087	4,000
NaK	1	2	0,538	0,0013	0,538
PGK	21	11	0,985	0,0098	4,394
CYP9J2	3	4	0,900	0,0068	1,400
Tsf	8	6	1,000	0,0078	4,200

S = número de sítios segregantes; NH = número de haplótipos; Hd = diversidade de haplótipos; π = número médio de substituições nucleotídicas por sítio; κ = número médio de diferenças nucleotídicas.

A maior frequência dos SNPs detectados em regiões codificantes ocorreu na terceira posição do códon (68,6%), ao invés da 1st (primeira) (17,1%) ou 2nd (segunda) (14,3%) posições e, na maior parte das vezes, como transição (59,3%). Como transversão, os SNPs ocorreram a uma taxa menor do que as transições (40,7%), também com maior frequência na terceira posição do códon (62,5%) e em taxas menores na 1st (primeira) (12,5%) ou 2nd (segunda) (25%) posições. Mutações sinônimas ou silenciosas foram mais comuns (61%) do que mutações não-sinônimas (39%) (Tabela VII).

Tabela VII. Polimorfismos em nucleotídeos para os nove genes nucleares do mosquito *Aedes aegypti*.

		Regiões Codificantes											
		Sítios Polimórficos											
Gene	Indels	Transições				Transversões				S	Ns	Total	
		1st	2nd	3rd	Σ	1st	2nd	3rd	Σ				
apoLp-II		0	1	3	4	0	1	3	4	5	3	8	
Chym		1	0	1	2	0	0	0	0	1	1	2	
Ef-2		0	0	2	2	0	0	0	0	2	0	2	
FerH	1 pb	0	1	2	3	0	0	1	1	3	1	4	
MUCI	24 pb	1	2	5	8	1	1	2	4	6	6	12	
NaK		0	0	0	0	0	0	0	0	0	0	0	
PGK	10 pb	2	1	8	11	2	2	6	10	11	10	21	
CYP9J2		0	0	2	2	0	0	0	0	2	0	2	
Tsf		2	0	1	3	0	2	3	5	6	2	8	
Total	35	6	5	24	35	3	6	15	24	36	23	59	

Indels = polimorfismos por inserção/deleção; Σ = total; S = substituições sinônimas; Ns = substituições não-sinônimas.

Indels foram detectados em três (33,3%) dos nove genes estudados, sendo eles: FerH, Muc e PGK. Para o fragmento amplificado do gene FerH o *indel* está localizado em um íntron (Figura 12). Entretanto, para os fragmentos dos genes MUC (Figura 13) e PGK (Figura 14) foram amplificados apenas éxons, assim, os *indels* ocorreram em regiões codificantes.

```

CGR1 CCGCCTATTTTGGCCAGGAAAAGATCAACCTGCCCGGGTTCGAAAAGTTTTCTTCCACG
CGR2 CCGCCTATTTTGGCCAGGAAAAGATCAACCTGCCCGGGTTCGAAAAGTTTTCTTCCACG
MAC7 CCGCCTATTTTGGCCAGGAAAAGATCAACCTGCCCGGGTTCGAAAAGTTTTCTTCCACG
*****

CGR1 CCGCCGCTGAAGAGCGGGAGCACGGCATCAAGCTGATCGAGTACGCTCTTATGCGTGGA
CGR2 CCGCCGCTGAAGAGCGGGAGCACGGCATCAAGCTGATCGAGTACGCTCTTATGCGTGGA
MAC7 CCGCCGCTGAAGAGCGGGAGCACGGCATCAAGCTGATCGAGTACGCTCTTATGCGTGGA
*****

CGR1 AGGCTCCCGTCGACAAGCACTTCAAGCTGAAGTATGTAAGTCGGTCACATCGGTGAAGAG
CGR2 AGGCTCCCGTCGACAAGCACTTCAAGCTGAAGTATGTAAGTCGGTCACATCGGTGAAGAG
MAC7 AGGCTCCCGTCGACAAGCACTTCAAGCTGAAGTACGTAAGTCGGTCACATCGGTGAAGAA
*****

CGR1 A-CCTCCAGATCCACTTTGTGCTAAACTTTTCGACATCTTTTACTGTATGCTTTCAGGACC
CGR2 A-CCTCCAGATCCACTTTGTGCTAAACTTTTCGACATCTTTTACTGTATGCTTTCAGGACC
MAC7 ATCCTCCTGAATCAGTTTGTGCTAAACTTTTCGACATATCTTACTGTATGCTTTCAGGACC
*****

CGR1 ACGAGGTGCCAACCGTCACCACTGGTGAATCCGCTCTGGAAACCGCCCTACAGAAGGAGG
CGR2 ACGAGGTGCCAACCGTCACCACTGGTGAATCCGCTCTGGAAACCGCCCTACAGAAGGAGG
MAC7 ACGAGGTGCCAACCGTCACCACTGGTGAATCCGCTCTGGAAACCGCCCTACAGAAGGAGG
*****

CGR1 TAGAAGTGACGAAGAGCATCCGCGGAGTCATCAAGGCCTGCGAAGATGGCAGCAATGACT
CGR2 TAGAAGTGACGAAGAGCATCCGCGGAGTCATCAAGGCCTGCGAAGATGGCAGCAATGACT
MAC7 TAGAAGTGACGAGGAGCATCCGCGGAGTCATCAAGGCCTGCGAAGATGGCAGCAATGACT
*****

CGR1 TCCACGTAAGCTTCAAATCATCTGCACCGAAGGTCTTCCCCATTGTCATGCATTTTGTCT
CGR2 TCCACGTAAGCTTCAAATCATCTGCACCGAAGGTCTTCCCCATTGTCATGCATTTTGTCT
MAC7 TCCACGTAAGATTCAAATCATCTGCACCGAAGGTCTTCCCCATTGTCATGCATTTTGTCT
*****

CGR1 TCTTCCTCTTTTCTTTCAAATCAACAGTTGGCCGACTACTTGACCGGCGAATATCTGGAC
CGR2 TCTTCCTCTTTTCTTTCAAATCAACAGTTGGCCGACTACTTGACCGGCGAATATCTGGAC
MAC7 TCTTCTTCTTCTCTCTCAAATCAACAGTTGGCCGACTACTTGACCGGCGAATACCTGGAC
*****

CGR1 GAACAGCACAAAGGGACAGCGCGAGCTGGCC
CGR2 GAACAGCACAAAGGGACAGCGCGAGCTGGCC
MAC7 GAACAGCACAAAGGGACAGCGCGAGCTGGCC
*****

```

Figura 12. Alinhamento múltiplo da região amplificada do gene da *Ferritin heavy chain* realizado com o programa Clustal X. *Indel* presente nas seqüências está indicado em vermelho. *Indel* com 1 pb.

```

LON10  TTCGACTCCAACCACTTGGTCTTCCTGCCCCACGAGGACTGCACAAAGTTCTATTTGTGC
CGR7    TTGGACTCCAACCACTTGGTCTTCCTGCCCCACGAGGACTGCACAAAGTTCTATTTGTGC
RBP4    TTCGACTCCAACCACTTGGTCTTCCTGCCCCACGAGGACTGCACAAAGTTCTATTTGTGC
      ** *****

LON10  GGTCTAATGGCCCGGTTGAGAAACAGTGTCCATCGGGACTGCACTGGAACCTCGCAAGCC
CGR7    GGTCTAATGGCCCGGTTGAGAAACAGTGTCCATCGGGACTGCACTGGAACCTCGCAAGCC
RBP4    GGTCTAATGGCCCGGTTGAGAAACAGTGTCCATCGGGACTGCACTGGAACCTCGCAAGCC
      *****

LON10  AGCGTTTGTGATTGGCCAGAGCTGGCCGGATGTAGTGGTGGAAGTTCTGTGCCCCCAACT
CGR7    AGCGTTTGTGATTGGCCAGAGCTGGCCGGATGTAGTGGTGGAAGTTCTGTGCCCCCAACT
RBP4    AGCGTTTGTGATTGGCCAGAGCTGGCCGGATGTAGTGGTGGAAGTACTGTGCCTCCAACCT
      *****

LON10  GTCACAGTAACTCCGGAGCCTGTAAGC-----ACCACAACG
CGR7    GTCACAGTAACTCCGGAGCCTGTAAGC-----ACCACAACG
RBP4    GTCACAGTAACTCCGGAACCTGTAACCTCTACTACGGCATCTCCTGCAGTCACCACAACG
      *****

LON10  GCACCAGCTGCGACCACCAAGTGCTCCTCCCTCTTCAACCGTGGCACCAACCAACAAGTGT
CGR7    GCACCAGCTGCGACCACCAAGTGCTCCTCCCTCTTCAACCGTGGCACCAACCAACAAGTGT
RBP4    GCACCAGCTGCGACCACCAAGTGCTCCTCCCTCTTCAACTGTGGCACCAACCAACAAGTGT
      *****

LON10  CCTGAATTCTTCAACCCCGATCACGTATCTTTTCATACCCCATGCCGATTGTTCAAAGTTC
CGR7    CCTGAATTCTTCAACCCCGATCACGTATCTTTTCATACCCCATGCCGATTGTTCAAAGTTC
RBP4    CCTGAATTCTTCAACCCCGATCACGTATCTTTTCATGCCCATGCCGATTGTTCAAAGTTC
      *****

LON10  TACGTATGCACCCAGGAAGGTCCCGTTGAAAAGAGCTGTCCATCTGGACTTCACTGGAAT
CGR7    TACGTATGCACCCAGGAAGGTCCCGTTGAAAAGAGCTGTCCATCTGGACTTCACTGGAAT
RBP4    TACGTATGCACCCAGGAAGGTCCCGTTGAAAAGAGCTGTCCATCTGGACTTCACTGGAAT
      *****

LON10  CAACAGGGTAGCATCTGCGATTGGCCGGAGGTAGCTGGATGTGTGGCCAGTGCATCGAT
CGR7    CAACAGGGTAGCATCTGCGATTGGCCGGAGGTAGCTGGATGTGTGGCCAGTGCATCGAT
RBP4    CAACAGGGTAGCATCTGCGATTGGCCGGCGGTAGCTGGATGTGTGGCCAGTGCATCGAT
      *****

```

Figura 13. Alinhamento múltiplo da região amplificada do gene da *Mucin-like protein* realizado com o programa Clustal X. *Indel* presente nas seqüências está indicado em vermelho. *Indel* com 24 pb.

```

BTU7 AATGGCATTTCACGCTCCTTAAGGTATTGAACAATATGGAAATTGGAGGATCATTGTTTCGA
CGR7 AATGGCATTTCACGTTCTTAAGGTATTGAACAATATGGAAATTGGAGGATCATTGTTTCGA
RCK6 AATGGCATTTCACGTTCTTAAGGTATTGACCTATATGGAAATTGGAGGATCATTTTTTCGA
*****
BTU7 C-----GAGGAGGGTTCAAAGATCGTTCAGAAGCTGGTAGACAAAGCCAAGGCCA
CGR7 C-----GAGGAGGGTTCAAAGATCGTTCAGAAGCTGGTAGACAAAGCCAAGGCCA
RCK6 CCAAGGGGGACGAGGAGGGTTCAAAGATCGTTCAGAAGCTGGTAGACAAAGCCAAGGCCA
*
BTU7 ACAACGTGCAACTGCATTTACCAGTGGATTTCGTAAGCTGGTGATAAATTTGCCGAAGATG
CGR7 ACAACGTGCAACTGCATTTACCAGTGGATTTCGTAAGCTGGTGATAAATTTGCCGAAGATG
RCK6 ACAACGTGCAACTGCATTTACCGTGGATTTCGTAAGCTGGTGATAAATTTGCCGAAGATG
*****
BTU7 CTGCCGTTGGAGCTGCTACCGTCGAGAGCGGAATCCCGCTGGACACATGGGTCTCGACG
CGR7 CTGCCGTTGGAGCTGCTACCGTCGAGAGCGGAATCCCGCTGGACACATGGGTCTCGACG
RCK6 CTGCCGTTGGAGCTGCTACCGTCGAGAGCGGAATCCCGCTGGACACATGGGTCTCGACG
*****
BTU7 TTGGACCGAAGTCTCGTGAAGCATTCGCTGCTCCGATTGCCCGCTCCAAGATCATTGTTT
CGR7 TTGGACCCAAGTCTCGTGAAGCATTCGCTGCTCCGATTGCCCGCTCCAAGATCATTGTTT
RCK6 TTGGACCCAAGTCTCGTGAAGCATTCGCTGCTCCGATTGCCCGGTCCAAGATCATTGTTT
*****
BTU7 GGAACGGTCCACCCGGCGTTTTTTGAGTTCCTCAATTTTCGCCAACGGAACGAAAGCTCTTA
CGR7 GGAACGGTCCACCCGGCGTTTTTTGAGTTCCTCAATTTTCGCCAACGGAACGAAAGCTCTTA
RCK6 GGAACGGTCCACCCGGCGTTTTTTGAGTTCCTCAATTTTCGCCAACGGAACGAAAGCTCTTA
*****
BTU7 TGGATGCCGTTGTCTGCTGCAACAAAGGCAGGCACTGTATCTATTATTGGCGGTGGAGATA
CGR7 TGGATGCCGTTGTCTGCTGCAACAAAGGCAGGCACTGTATCTATTATTGGCGGTGGAGATA
RCK6 TGGATGCCGTTGTCTGCTGCAACAAAGGCAGGCACTGTATCTATAATTGGGGGTGGAGATA
*****
BTU7 CCGCTTCGTGTTGCGCCAAGTGGGGAACAGAATCTC
CGR7 CCGCTTCGTGTTGCGCCAAGTGGGGAACAGAATCTC
RCK6 CCGCTCCTTGGTGCGCCAAGTGGGGAACAGAATCTC
*****

```

Figura 14. Alinhamento múltiplo da região amplificada do gene da *Phosphoglycerate kinase* realizado com o programa Clustal X. *Indel* presente nas seqüências está indicado em vermelho. *Indel* com 10pb.

5.4 Polimorfismos SNPs e Genotipagem

A análise das seqüências nucleares para os nove genes selecionados evidenciou a presença de SNPs tanto inter quanto intrapopulacional nos mosquitos analisados. A Tabela VIII contém todos os SNPs identificados para cada um dos fragmentos de genes nucleares estudados para o mosquito *Ae. aegypti*.

Tabela VIII. Setenta e um polimorfismos de único nucleotídeo identificados em nove genes de dezesseis populações naturais do mosquito *Aedes aegypti* no Brasil.

Locus / Chromosome	SNP position: identity	SNP location
ApoLp-II	031: T/C	Éxon
	068: A/C	Éxon
	101: A/G	Éxon
	125: C/G	Éxon
	161: A/G	Éxon
	248: C/T	Éxon
	284: C/A	Éxon
	325: G/C	Éxon
Chym	013: C/T	Éxon
	218: A/G	Éxon
CYP9J2	134: C/T	Éxon
	164: A/G	Éxon
	268: T/A	3'UTR
Ef2	054: C/T	Éxon
	096: C/T	Éxon
FerH	162: C/T	Éxon
	187: A/G	Íntron
	194: T/A	Íntron
	197: A/T	Íntron
	198: T/C	Íntron
	201: G/C	Íntron

	223: A/C	Íntron
	225: C/T	Íntron
	319: G/A	Éxon
	377: A/C	Éxon
	432: T/C	Íntron
	437: C/T	Íntron
	441: C/T	Íntron
	480: C/T	Éxon
NaK	272: A/T	Íntron
PGK	041: A/C	Éxon
	044: T/C	Éxon
	052: G/C	Éxon
	054: T/C	Éxon
	060: A/C	Éxon
	062: A/T	Éxon
	064: T/C	Éxon
	085: G/T	Éxon
	103: A/G	Éxon
	163: A/G	Éxon
	181: T/C	Éxon
	238: C/T	Éxon
	268: C/G	Éxon
	304: C/G	Éxon
	406: G/A	Éxon
	424: T/A	Éxon
	430: C/G	Éxon
	446: T/C	Éxon
	448: G/T	Éxon
	451: T/G	Éxon
	471: G/A	Éxon
MUC1	020: C/G	Éxon

Tsf	159: C/T	Éxon
	183: T/A	Éxon
	191: C/T	Éxon
	215: G/A	Éxon
	223: G/C	Éxon
	251: C/T	Éxon
	272: C/T	Éxon
	320: A/T	Éxon
	321: T/A	Éxon
	329: A/G	Éxon
	442: A/C	Éxon
	612: T/A	Éxon
	666: A/T	Éxon
	696: C/G	Éxon
	813: T/C	Éxon
	917: A/T	Éxon
	966: T/C	Éxon
	1045: C/T	Éxon
	1083: C/G	Éxon

A Tabela IX contém a descrição dos 22 sítios SNPs previamente selecionados para o estudo de caracterização das populações de Botucatu, Juazeiro do Norte e Maceió do mosquito *Ae. aegypti*.

Tabela IX. Polimorfismos de único nucleotídeo selecionados para o estudo de caracterização de três populações do mosquito *Aedes aegypti* no Brasil.

Gene	Nº acesso ao GenBank*	Ns	Posição	Localização		
apoLp-II	EU427570 (332 bp)	1	031: T/C	Éxon		
		2	068: A/C	Éxon		
		3	170: G/A	Éxon		
Ef2	EU427571 (381 bp)	1	054: C/T	Éxon		
		2	096: C/T	Éxon		
Chym	EU427572 (445 bp)	1	442: A/G	Éxon		
FerH		1	162: C/T	Éxon		
		2	377: A/C	Íntron		
		3	441: C/T	Íntron		
		CYP9J2	EU427574 (335 bp)	1	164: A/G	Éxon
				2	286: T/A	3'UTR
MUC	EU427572 (445 bp)	1	159: C/T	Éxon		
		2	215: G/A	Éxon		
		3	272: C/T	Éxon		
		4	329: A/G	Éxon		
NaK	EU427573 (387 bp)	1	342: A/T	Íntron		
PGK	EU427575 (460 bp)	1	044: T/C	Éxon		
		2	083: A/G	Éxon		
		3	161: T/C	Éxon		
		4	268: C/G	Éxon		
Tsf	EU427576 (536 bp)	1	813: G/A	Éxon		
		2	966: T/C	Éxon		

Ns = número do SNP testado para cada gene. Alelos mais comuns estão na primeira posição. * Número correspondente às seqüências obtidas no estudo e submetidas ao GenBank.

Baseado em cada um dos pontos de mutação descritos na tabela IX oligonucleotídeos adjacentes aos sítios SNPs foram desenhados e, a seguir, testados para genotipagem das 3 populações naturais do mosquito *Ae. aegypti*.

Tabela X. Oligonucleotídeos utilizados para a reação de extensão alelo-específica dos polimorfismos de único nucleotídeo.

Gene	Oligonucleotídeos	Seqüências dos Oligonucleotídeos	Mutação
apoLp-II	ApoLp-1	TTTCGATGCCTTGTTCCGAG	Y (T ou C)
	ApoLp-2	(CCC)GTTGAAGCTTTGGCCAATCT	M (A ou C)
	ApoLp-3	(CCCCC)CCAGCATTGAAAGCCGCCAA	R (C ou T)
Chym	Chym-1	CCTTGATAGTCCTCGTGGA	R
Ef2	Ef2-1	CCGGCTTGATACGTTCGGC	Y
	Ef2-2	(CCC)GCTGTTCATGAACAAGATGGA	Y
FerH	FerH-1	CAAGCACTTCAAGCTGAACTA	Y
	FerH-2	(CCC)AGCAATGACTTCCACGTAAG	M (A ou C)
	FerH-3	(CCCCC)TCGGCCAACTGTTGATTTGA	Y
MUC	MUC-1	GCGTTTGTGATTGGCCAGAG	Y
	MUC-2	(CCC)CTGTCACAGTAACTCCGGA	R
	MUC-3	(CCCCC) GTGCTCCTCCCTCTTCAAC	Y
	MUC-4	(CCCCCCCC)GCCACACATCCAGCTACC	M
NaK	NaK-1	TCGCCAAAAGCCCTCCAAC	W (A ou T)
CYP9J2	P450-1	GAACTGAAGGCTAGACCGAA	R
	P450-2	(CCC)GCAAACAAAGTTATCCGCGA	W
PGK	PGK-1	GAGGTGGAATGGCATTACG	Y
	PGK-2	(CCC)ACCAGCTTCTGAACGATCTT	R
	PGK-3	(CCCCC)CCAGTTACGAAATCCAC	R
	PGK-4	(CCCCCCCC)GCGAATGCTTCACGAGACCTT	S (G ou C)
Tsf	Tsf-1	TACGTCTACCTATGTGAAGA	Y
	Tsf-2	(CCC)CCACATCTTGAGGGCCTTGTC	Y

Após a padronização das reações de minisequenciamento SNPs localizados nos genes ApoLp-II, Ef-2, MUC, NaK, CYP9J2, PGK, Tsf, foram utilizados para a genotipagem das populações de Maceió, Botucatu e Juazeiro do Norte. Os SNPs selecionados para os genes Chymotripsina e Ferritina não foram utilizados porque o DNA, que já havia sido

utilizado em pesquisas anteriores, foi insuficiente para realizar as reações de minisequenciamento.

A tabela XI contém os dados obtidos após a genotipagem das três populações do mosquito *Ae. aegypti* com dezoito SNPs.

Tabela XI. Planilha montada através da análise *multilocus* de três populações do mosquito *Aedes aegypti* com dezoito SNPs de sete genes nucleares.

	GENE	Nak	TSF-1	TSF-2	P450-1	P450-2	Apol-1	Apol-2	Apol-3		Ei2-1	Ei2-2	PGK-1	PGK-2	PGK-3	PGK-4	MUC1	MUC2	MUC3	MUC4
MACEIO		A/T	G/A	T	G/A	A/T	C/T	C/A	G/A				C/T	G/A	C/T	G/C	G/A	G/A	G/A	C/T
	1 AL1	A	G	T	A	A	T	A	G/A	AL1	G/A	C/T	T	G	C	C	A	A	A	C
	2 AL2	A	G	T	A	A	T	A	G	AL2	G	C/T	T	G	C	C	G/A	G/A	G/A	C/T
	3 AL3	A	G	T	G/A	A/T	C/T	A	G	AL3	G	C/T	T	G	C	C	G/A	G/A	G/A	C/T
	4 AL4	A/T	G	T	A	A	C/T	C/A	G/A	AL4	G/A	C/T	T	G	C	C	G/A	G/A	G/A	C/T
	5 AL5	A	G	T	G/A	A/T	C/T	C/A	G/A	AL5	G/A	C/T	T	G	C	C	G/A	G/A	G/A	C/T
	6 AL6	A	G	T	A	A	C/T	C/A	G/A	AL6	A	C	T	G/A	T	G	G/A	G/A	G/A	C/T
	7 AL7	A/T	G	T	G/A	A/T	C/T	A	G	AL7	G	C/T	T	G/A	T	G	G	G	G	T
	8 AL8	A	G	T	G	T	T	C/A	G/A	AL8	G	T	T	G	C/T	C/G	A	A	A	C
	9 AL9	A	G	T	G	T	C/T	C/A	G/A	AL9	G/A	C/T	T	G	C	C	G/A	G/A	G/A	C/T
	10 AL10	A	G	T	G	T	T	A	G	AL10	G	C/T	T	G/A	T	G	G/A	G/A	G/A	C
	11 AL11	A	G/A	T	A	A	C/T	A	G	AL11	G	T	T	G	T	G	A	A	A	C
	12 AL12	A	G	C/T	A	A	C/T	C/A	G/A	AL12	G	C/T	T	G	C	C	G/A	G/A	G/A	C/T
	13 AL13	A	G	T	A	A	T	C/A	G/A	AL13	G/A	C/T	T	G	C	C	A	A	A	C
	14 AL14	A	G	T	G	T	C/T	A	G	AL14	G/A	C/T	T	G	C	C	A	A	A	C
	15 AL15	A	G	T	G/A	A/T	C	A	G	AL15	G	T	T	G	C	C	G/A	G/A	G/A	C/T
	16 AL16	A	G/A	C/T	A	A	C/T	A	G	AL16	G/A	C/T	T	G/A	C/T	G	A	A	A	C
	17 AL17	A	G/A	C/T	A	A	T	C/A	G/A	AL17	G/A	C/T	T	G	C	C/G	A	A	A	C
	18 AL18	A	G/A	C/T	G	T	T	C/A	G/A	AL18	G/A	C/T	T	G	C	C	A	A	A	C
	19 AL19	A	G/A	C/T	G/A	A/T	C/T	A	G	AL19	G	T	T	G	C	C	G/A	G/A	G/A	C/T
	20 AL20	A	G/A	T	G/A	A/T	C/T	C/A	G/A	AL20	G	T	T	G	C	C	G/A	G/A	G/A	C/T
	21 AL21	A	G	T	A	A	T	C/A	G/A	AL21	G/A	C/T	T	G	C	C	A	A	A	C
	22 AL22	A	G/A	C/T	A	A	C/T	A	G	AL22	G/A	C/T	T	G/A	C	G	G/A	G/A	G/A	C/T
	23 AL23	A	G/A	C/T	A	A	C	A	G	AL23	G	T	T	G	C	C	G/A	G/A	G/A	C/T
	24 AL24	A	G/A	C/T	A	A	C/T	A	G	AL24	G/A	C/T	T	G	C	C	A	G/A	G/A	C/T
BOTUCATU																				
	1 BTU1	T	G/A	C	G	T	T	A	G	BTU1	A	C	C	A	T	C	A	A	A	C
	2 BTU2	A/T	G/A	C	G	T	C/T	A	G	BTU2	A	C	C	A	T	C	G/A	G/A	G/A	C/T
	3 BTU3	A/T	A	C	G	T	T	C/A	G/A	BTU3	G/A	C/T	T	A	T	G	A	A	A	C
	4 BTU4	A/T	G	C	G	T	T	C/A	G/A	BTU4	G	T	T	G	T	G	A	A	A	C
	5 BTU5	A/T	G/A	C	G	T	C/T	A	G	BTU5	A	C	C	A	T	G	G/A	G/A	G/A	C/T
	6 BTU6	A/T	A	C	G	T	T	C/A	G/A	BTU6	A	T	C	A	T	G	A	A	A	C
	7 BTU7	A/T	A	C	G	T	T	C/A	G/A	BTU7	G/A	C/T	C	A	T	G	A	A	A	C
	8 BTU8	T	A	C	G/A	A/T	T	C/A	G/A	BTU8	G/A	C/T	C	A	T	G	A	A	A	C
	9 BTU9	A/T	A	C	G/A	A/T	T	A	G	BTU9	G/A	C/T	C/T	A	T	G	G/A	G/A	G/A	C/T
	10 BTU10	T	A	C	G/A	A/T	T	C/A	G/A	BTU10	G	C	T	A	T	G	G/A	G/A	G/A	C/T
	11 BTU11	A/T	A	C	A	A/T	T	C/A	G	BTU11	G/A	C/T	C/T	A	T	G	A	A	A	C
	12 BTU12	A/T	G	C	G/A	A/T	T	A	G	BTU12	A	C	C/T	A	T	G	G/A	G/A	G/A	C/T
	13 BTU13	A/T	A	C	G	T	T	C/A	G/A	BTU13	G/A	C/T	T	A	T	G	G/A	G/A	G/A	C/T
	14 BTU14	A	G	C	G	T	T	C/A	G/A	BTU14	A	C	C	A	T	G	G/A	G/A	G/A	C/T
	15 BTU15	A/T	G	C	G	T	T	C/A	G	BTU15	G/A	C/T	C	A	T	G	A	A	A	C
	16 BTU16	A	G/A	C	G/A	A/T	T	A	G	BTU16	A	C	C	A	T	G	A	A	A	C
	17 BTU17	A/T	G/A	C	G	T	T	C/A	G/A	BTU17	G/A	C/T	T	A	T	G	G/A	G/A	G/A	C/T
	18 BTU18	A/T	G	C	G	T	T	C/A	G/A	BTU18	G	T	C	A	T	G	G	G	G	T
	19 BTU19	A	A	C	G	T	T	C/A	G/A	BTU19	A	C	C/T	A	T	G	G	G	G	T
	20 BTU20	A/T	A	C	G/A	A/T	T	C/A	G/A	BTU20	A	C	C	A	T	G	G/A	G/A	G/A	C/T
	21 BTU21	A	G/A	C	G	T	T	A	G	BTU21	G/A	C/T	C	A	T	G	G/A	G/A	G/A	C/T
	22 BTU22	A/T	A	T	G	T	T	C/A	G/A	BTU22	G	T	C/T	A	T	G	G	G	G	T
	23 BTU23	A	A	C	G	T	T	A	G	BTU23	A	C	C/T	A	T	G	G/A	G/A	G/A	C/T
	24 BTU24	A	G/A	T	G	T	T	A	G/A	BTU24	A	C	T	A	T	G	G/A	G/A	G/A	C/T

Tabela XI. Planilha montada através da análise *multilocus* de três populações naturais do mosquito *Aedes aegypti* com dezoito SNPs em sete genes nucleares. (continuação)

	GENE	Nak	TSF-1	TSF-2	P450-1	P450-2	Apol-1	Apol-2	Apol-3		Ef2-1	Ef2-2	PGK-1	PGK-2	PGK-3	PGK-4	MUC1	MUC2	MUC3	MUC4
JUAZEIRO	1 JUA1	A/T	G/A	C/T	A/T	G/A	C/T	C/A	G/A		G/A	C/T	C/T	G/A	C/T	G/C	G/A	G/A	G/A	C/T
	2 JUA2	A	G	T	A	A	T	C/A	G/A	JUA1	G/A	C/T	C	A	C/T	C/G	A	A	A	C
	3 JUA3	A	G/A	C/T	G	T	C/T	A	G/A	JUA2	G/A	C/T	T	A	T	G	G/A	G/A	G/A	C/T
	4 JUA4	A	A	C	G	T	C/T	C/A	G/A	JUA3	G	T	T	A	T	G	A	A	A	C
	5 JUA5	A	G	T	G	T	C	C	A	JUA4	G/A	C/T	T	G	C	C	A	A	A	C
	6 JUA6	A	G/A	T	A	A	T	C	A	JUA5	G	T	C/T	G/A	C/T	C/G	A	A	A	C
	7 JUA7	A	G/A	T	G/A	A/T	A	T	A	JUA6	G/A	C/T	C/T	G/A	C/T	C/G	A	A	A	C
	8 JUA8	A	A	C	G	T	C/T	C/A	G/A	JUA7	G/A	C/T	C/T	A	T	G	A	A	A	C
	9 JUA9	A	G/A	C/T	G	T	C/T	C/A	G/A	JUA8	G/A	C/T	C/T	G/A	C/T	C/G	A	A	A	C
	10 JUA10	A	A	C	G	T	C	C	A	JUA9	G/A	C/T	T	A	T	G	G	G	G	T
	11 JUA11	A	G/A	C/T	G	A/T	T	C	A	JUA10	G	T	C/T	G/A	T	G	A	A	A	C
	12 JUA12	A	G	T	G/A	A/T	T	C	A	JUA11	G	T	T	A	C	C	G	G	G	T
	13 JUA13	A	G/A	T	G	T	T	C	A	JUA12	G/A	C/T	C/T	G/A	C/T	C/G	G/A	G/A	G/A	C/T
	14 JUA14	A	G/A	C/T	G/A	A/T	T	C	A	JUA13	G/A	C/T	T	A	T	G	G/A	G/A	G/A	C/T
	15 JUA1M	A	A	T	A	A/T	C/T	C/A	G/A	JUA14	A	C	C/T	G/A	C/T	C/G	A	A	A	C
	16 JUA2M	A	G	T	G	T	C/T	C/A	G/A	JUA1M	A	C	T	A	C	C	A	A	A	C
	17 JUA3M	A	A	C/T	G	T	C/T	C/A	G/A	JUA2M	G	T	C/T	A	T	G	A	A	A	C
	18 JUA4M	A	G	T	G/A	A/T	T	A	G	JUA3M	G	T	C/T	G/A	C/T	C/G	A	A	A	C
	19 JUA5M	A	G	C/T	G/A	A/T	T	A	G	JUA4M	G/A	C/T	C/T	G/A	C/T	C/G	A	A	A	C
	20 JUA6M	A	G	G	G	T	T	A	G	JUA5M	G	T	T	G	C	C	A	A	A	C
	21 JUA7M	A	G/A	C/T	G	T	T	C	A	JUA6M	G	T	T	A	G	C	A	A	A	C
	22 JUA8M	A	G/A	C	G	T	T	C	A	JUA7M	G	T	T	A	T	G	G/A	G/A	G/A	C/T
	23 JUA9M	A	G	T	A	A	T	C	A	JUA8M	G/A	C/T	T	A	T	G	A	A	A	C
	24 JUA10M	A	G/A	C/T	A	A/T	T	C/A	A	JUA9M	G/A	C/T	C/T	G/A	C/T	C/G	G/A	G/A	G/A	C/T
	25 JUA11M	A	A	C	G	T	T	C/A	G/A	JUA10M	G	T	C/T	G/A	C/T	C/G	G/A	G/A	G/A	C/T
	26 JUA12M	A	G/A	C/T	A	A/T	T	C	A	JUA11M	G/A	C/T	C/T	A	T	G	A	A	A	C
	27 JUA13M	A	G	T	A	A	T	C	A	JUA12M	G	T	T	A	C/T	C/G	A	A	A	C
	28 JUA14M	A	G/A	C/T	G/A	A/T	T	C	A	JUA13M	G/A	C/T	T	G/A	C/T	C/G	G/A	G/A	G/A	C/T

Após a conclusão dos resultados pelas reações de minisequenciamento foi possível observar que todos os genótipos obtidos utilizando a técnica de SNaPshot™ foram idênticos aos obtidos anteriormente pelo seqüenciamento. Na figura 15 estão apresentadas todas as regiões amplificadas (sítios dos oligonucleotídeos sublinhados), íntrons, indicação de todos os SNPs, e modificações de aminoácidos com os respectivos SNPs identificados durante o estudo.

Lipoprotein N-terminal domain (AAEL009955) – frame 3

60

GTAAGGTCTA**TTTCGATGCCTTGTTCGAGY**TGGAACCGGTGCCTCC**GTTGAAGCTTTGG**
 K V Y F D A L F R V G T G A S V E A L

120

CCAATCTMTACAAGAACAAGGAGGTATCTGACGCAAAGGAR**C**CAGAAGCTTCTCTTCGTTT
 A N L Y K N K E V S D A K E Q K L L F V

180

CTCT**S**AACCTGGTCACTTCTATGACTAAA**CCAGCATTGAAAGCCGCCAAR**CTGCTGCTCG
 S L N L V T S M T K P A L K A A K L L L

240

ACGGAAACCCATCTCGTGAAGCTTACCTCAGCGTCGGAAGCTTAGTCAACAAGTACTGCC
 D G N P S R E A Y L S V G S L V N K Y C

300

AGAAGTT**Y**GGATGTGAATCAGCCGATGTTAAAGAAATCTCTGAM**M**AAGTTCTCTGCAAAGC
 Q K F G C E S A D V K E I S **DE** K F S A K

360

TTGGCAAGTGCCAACCGACCACCC**S**CCAAGAAGAGGATACTATCGTCGCTGTTCTGA
 L G K C Q P T T **RP** Q E E D T I V A V L

Figura 15. Regiões amplificadas de cada um dos nove genes nucleares. Íntros aparecem em letras minúsculas, oligonucleotídeos das reações de PCR estão sublinhados, oligonucleotídeos das reações de SNaPshot estão em negrito e itálico. SNPs identificados estão sublinhados, SNPs testados estão em amarelo, SNPs utilizados na genotipagem estão em vermelho. SNPs não-sinônimos estão indicados com a cor cinza.

Serine-type endopeptidase (AAEL5576966) – frame 2

60
TCCCTACCAAAT**Y**TCGCTGCAGTGGAACTTCAACGATGGACAAACGGAGACCATGCACTT
P Y Q I S L Q W N F N D G Q T E T M H F

120
CTGCGGAGCTTCGGTGTGTAACGAAAACCTTCGTCCTGACGGCTGCTCACTGCAAGACCGC
C G A S V L N E N F V L T A A H C K T A

180
ATACTCCAATACCGGGTTCATCGAAGTGGTTGCCGCTGAACATGATGTGGCCGTTGCGGA
Y S N T G F I E V V A A E H D V A V A E

240
AGGATCCGAACAGCGTCGTTTGGTTGCGGAATTCATC**R****TCCACGAGGACTATCAAGG**GGG
G S E Q R R L V A E F I **IV** H E D Y Q G G

300
AGTCAGTCCCGATGACATTGCCGTCATTTCGTGTGGACAAACCCTTCGAATTGAACGATAA
V S P D D I A V I R V D K P F E L N D K

360
GGTGAAGGCCGTTAAGCTGCCCAAGCAGTTGGAACAATTCGATGGCGATGTTACTCTGAG
V K A V K L P K Q L E Q F D G D V T L S

420
TGGATGGGGATCCGTATCGACGACGGTGTTCGCGGACTATCCTGACAAACTGAGGgtaag
G W G S V S T T V F P D Y P D K L R -----

480
ttcggtagttgatttgattggagacgagttactgacgaaccgctaactttgcaGAAAGTC
----- V T

540
GTGCTTCCGCTGGTAGACTACGAACAATGTAACGCCCTGTGGGACAACGACAGTGCTC
D E P L T L Q K V V L P L V D Y E Q C

Figura 15. (continuação)

Elongation factor (AAEL004500) – frame 1

60
GATTGCGTTTCCGGTGTGTGCGTACAGACCGAAACCGTGCTGCGTCAGGCCAT**Y**GCCGAA
D C V S G V C V Q T E T V L R Q A I A E

120
CGTATCAAGCCGGTGCTGTT**CA**TGAACAAGATGGAYCGTGCTCTGCTGGAGCTGCAGCTG
R I K P V L F M N K M D R A L L E L Q L

180
GATGCTGAAGATCTGTACCAGACCTTCCAGCGTATCGTTGAAAACGTCAACGTCATCATT
D A E D L Y Q T F Q R I V E N V N V I I

240
GCCACCTACAACGACGATGGCGGCCCAATGGGAGAAGTCCGTGTCGATCCATCCAAGGGT
A T Y N D D G G P M G E V R V D P S K G

300
TCCGTCGGTTTCCGGTCTGCACGGATGGGCCTTCACTCTCAAGCAGTTCGCTGAG
S V G F G S G L H G W A F T L K Q F A E

360
ATGTACGCCGCTATGTTCAAGATCGATGTCGTCAAGCTCATGAACCGTCTGTGGGGAGAG
M Y A A M F K I D V V K L M N R L W G E

410
AACTTCTTCAACCCCAAGACCAAGAAGTGGGCCAAGACCAAGGATGACGAC
N F F N P K T K K W A K T K D D D

Figura 15. (continuação)

Ferritin subunit 1 (AAEL007385) – frame 1

60
AAGTATGCCGCCTATTTTGCCCAGGAAAAGATCAACCTGCCCGGGTTCGAAAAGTTTTC
K Y A A Y F A Q E K I N L P G F E K F F

120
TTCCACGCCGCGCTGAAGAGCGGGAGCACGGCATCAAGCTGATCGAGTACGCTCTTATG
F H A A A E E R E H G I K L I E Y A L M

180
CGTGGAAGGCTCCCGTCGA**CAAGCACTTCAAGCTGAACTA****Y**gtaagtcggtcacatcgg
R G K A P V D K H F K L N Y -----

240
tgaaga**r**acctcc**w**ga**wy**ca**s**tttgtgctaaactttcgacat**mt****y**ttactgtatgctttc

300
agGACCACGAGGTGCCAACCGTCACCACTGGTGAATCCGCTCTGGAAACCGCCCTACAGA
-- D H E V P T V T T G E S A L E T A L Q

360
AGGAGGTAGAAGTGACGAR**R**GAGCATCCGCGGAGTCATCAAGGCCTGCGAAGATGGC**AGCA**
K E V E V T K S I R G V I K A C E D G S

420
ATGACTTCCACGTAAG**M**ttcaaatcatctgcaccgaaggtcttccccattgtcatgcatt
N D F H V **SR** -----

480
ttgtcttcttct**y**tctt**y**tct**y**t**caaatcaacagTTGGCCGA**CTACTTGACCGGCGAATA**Y**
----- L A D Y L T G E Y

540
CTGGACGAACAGCACAAGGGACAGCGGAGCTGGCCGAGAAGATCGCCACGCTTAAAA**AG**
L D E Q H K G Q R E L A E K I A T L K K

550
ATGAAGAAGAGCGCTC
M K K S V

Figura 15. (continuação)

Mucin-like protein (AAEL004798) – frame 3

60
GCAAATGTCCCGATATCTT**S**GACTCCAACCACTTGGTCTTCCTGCCCCACGAGGACTGCA
K C P D I **FL** D S N H L V F L P H E D C

120
CAAAGTTCTATTTGTGCGGTCATAATGGCCCGGTTGAGAAACAGTGTCCATCGGGACTGC
T K F Y L C G H N G P V E K Q C P S G L

180
ACTGGAACCTCGCAAGCCA**GCGTTTGTGATTGGCCAGAGY**TGGCCGGATGTAGTGGTGGAA
H W N S Q A S V C D W P E L A G C S G G

240
GT**W**CTGTGCC**Y**CCAA**CTGTCACAGTAACTCCGGAR**CCTGTAAS**S**CACCACAACGGCACCAG
ST S V P P T V T V T P E P V **ST** T T T A P

300
CTGCGACCAC**Y**A**GTGCTCCTCCCTCTTCAACY**GTGGCACCAACCAACAAGTGTCTCTGAAT
A A T T S A P P S S T V A P T N K C P E

360
TCTTCAACCCCGATCACGT**WW**CTTTCAT**R**CCCCATGCCGATTGTTCAAAGTTCTACGTAT
F F N P D H V **ST** F **IM** P H A D C S K F Y V

420
GCACCCAGGAAGGTCCCGTTGAAAAGAGCTGTCCATCTGGACTTCACTGGAATCAACAGG
C T Q E G P V E K S C P S G L H W N Q Q

480
GTAGCATCTGCGATTGGCCGG**MGGTAGCTGGATGTGTGGCC**AGTGCATCGATTCCACCTA
G S I C D W P **EA** V A G C V A S A S I P P

490
AGGACCGTGAAACCGT
K D R E T

Figura 15. (continuação)

Sodium/potassium channel (AAEL012062) – frame 1

60
CCCGAATGGGTCAAGTTCTGTAAGAATCTCTTCGGTGGCTTCGCTCTGCTGCTGTGGATC
P E W V K F C K N L F G G F A L L L W I

120
GGTGCTATCCTGTGTTTCATTGCCTACTCGATCCTGGCCAGTACCGTCGAGGAACCGGCC
G A I L C F I A Y S I L A S T V E E P A

180
GACGATAACCTGTACCTCGGCATCGTGCTGACCGCCGTCGTGATAGTTACCGGTATTTTC
D D N L Y L G I V L T A V V I V T G I F

240
TCGTATTATCAGGTAAGGaatttggttacattcgacgatacggttaagttttgtcgacttt
S T T G V A -----

300
tcaaccttatgtcgactgtttgaaatcttatgtcgacttttgtaacctcacgacaaaatc

360
atgcctgcaagctaccggag**tcgccaaaagccctccaac**ttaacatgatcgttgcaa

420
ttgaaaactatcattttcttccctctgcccttttcataggaatcgaAAAGTTCGAAGATTA
----- K V R R L

480
TGGAATCGTTCAAGAACATGGTGCCCCAGTTTGCGACCGTACTGCGTGAGGGCGAGAAGC
W N R S R M V P Q F A T V L R E G E K L

510
TGACCCTGCGCGCCGAAGATCTGGTCATCGGTGAC
T L R A E D L V I G D

Figura 15. (continuação)

Citocrome P450 (AAEL006805) – frame 3

60

CGAGGTTGGCACTGTTGGAGGTCAAAATGGCTCTGTACCACCTGGTCAAAGATTTCAACC

R L A L L E V K M A L Y H L V K D F N L

120

TGCAACCGTCGGAGAAAACGCAGATCCCGCTGCGGCTGTCCAAGAGTGCGTTCACCATGC

Q P S E K T Q I P L R L S K S A F T M Q

180

AAGCGGAAAATGGYGTGGCTG**GAACTGAAGGCTAGACCGAAR**GCttagaccaaatgat

A E N G V W L E L K A R P K A -----

240

ggaatgggtttatctactggctctgatggctataccgcaaagttatcttatcaattcttc

300

atccgtg**gcaaacaaagttatccgcgaw**acgtaagtcacattgactacaatcatggtatg

350

tttaataaacacaaatatgtggaagtcatccaaaacgcaaagtcgatggtaccg

Figura 15. (continuação)

Phosphoglycerate kinase (AAEL58082) – frame 2

60
CAAGGTGAACGAAATGATCATTG**GAGGTGGAATGGCATTTC****CGY**TCCTTAAS**GY**ATTGAM
K V N E M I I G G G M A F **TP** **FL** L **KN** **VA** L **NT**

120
C**WY**ATGGAAATTGGAGGATCATT**K**TTTCGACGAGGAGGGTTC**R**AAGATCGTTCAGAAGCT
NY M E I G G S **LF** F D E E G S K I V Q K L

180
GGTAGACAAAGCCAAGGCCAACAACGTGCAACTGCATTTACC**RG**TGGATTTCGTA**ACTGG**
V D K A K A N N V Q L H L P V D F V T G

240
YGATAAATTTGCCGAAGATGCTGCCGTTGGAGCTGCTACCGTCGAGAGCGGAATTCC**Y**GC
D K F A E D A A V G A A T V E S G I P A

300
TGGACACATGGGTCTCGACGTTGGACC**S**AAGTCTCGTGAAGCATT**CGC**TGCTCCGATTGC
G H M G L D V G P K S R E A F A A P I A

360
CCG**S**TCCAAGATCATTGTTTGAACGGTCCACCCGGCGTTTTTGAGTTCCCCAATTT**CGC**
R S K I I V W N G P P G V F E F P N F A

420
CAACGGAACGAAAGCTCTTATGGATGCCGTTGTCGCTGCAACAAAR**RG**CAGGCACTGTATC
N G T K A L M D A V V A A T K A G T V S

480
TAT**W**ATTGG**S**GGTGGAGATACCGCT**YCK**T**GK**TGCGCCAAGTGGGGAACAA**R**ATCTCAGGT
I I G G G D T A **SP** **CW** C A K W G T **KE** S Q V

490
ATCACACG
S H

Figura 15. (continuação)

Transferrin (AAEL015639) – frame 1

60
GTAGTGGAAGTGCTTCTGCTGATCGGTCATAGTCATGGCCAACAGAAGGAAACGTTCAAA
V V E V L L L I G H S H G Q Q K E T F K

120
TTGTGCGTTCCGCATCAGATCATGGACGCGTGTCAGGATCTGATGGCCAAACCGGATGCG
L C V P H Q I M D A C Q D L M A K P D A

180
GCCATCCAGGTTTCAGTGCATTGCCGGCCGCGATCGTATGGAGTGTCTGGAGAAGGTCAAG
A I Q V Q C I A G R D R M E C L E K V K

240
GCACGGGAAGCGGACTTTGTGGCCGTCGATCCGGAAGACATGTACGTGGCCTATCACATG
A R E A D F V A V D P E D M Y V A Y H M

300
GCCAATCAGGACTTTAGCGTGTTACCGAGTTCCGCACGCTGGAGGAACCGAAGGCCGAG
A N Q D F S V F T E F R T L E E P K A E

360
TTCCGTTACGAGGGCATCATTCTGGTCCGCAAGAGCGACAACTTCCGCAGTTTGGCTGAT
F R Y E G I I L V R K S D N F R S L A D

420
CTCAGGGGTAAGAAGTCCTGCCACACGGGATACGGAAGAAATGTTGGCTATAAAATCCCG
L R G K K S C H T G Y G R N V G Y K I P

480
ATCACGAAACTGAAGTCGGCGGGAGTTTTCAAACCTGGCCACTGATTCGGAACCTGTCTCCG
I T K L K S A G V F K L A T D S E L S P

540
CTCGAGGAAGAACTGAAGGGGCTCTCCGACCTGTTTCGGAAGTGCGTGCTTGGTTGGAAAG
L E E E L K G L S D L F G S A C L V G K

600
TACTCTCCGAACGATGAAGTGAACCGATTGCTCAAGAAACGTTACTCGAACTTGTGTGCT
Y S P N D E V N R L L K K R Y S N L C A

660
CTGTGCGAACGWCCGGAAGTATGTGACTACCCGGACAAGTATTCCGGATATGATGGAGCG
L C E R P E V C D Y P D K Y S G Y D G A

720
ATCCGWTGTTTGGTAGAGAACAAACGGCGATGTGGCSTTTACTAAGGTCATCTACGTGAAC
I R C L V E N N G D V A F T K V I Y V N

780
AAGTACTTCGGTCTTCCGGTAGGAGGCGCCCCAGCCAAACCAGCCCTGAATCCAAATGCG
K Y F G L P V G G A P A K P A L N P N A

840
CGTACCGAAGAC**TACGTCTACCTATGTGAAGAY**GGCTCAACCCGCCCGATCACCGGACCG
R T E D Y V Y L C E D G S T R P I T G P

900
GCCTGCTCGTGGGCTCAACGACCTTGGCAGGGATACATGGGCAACGGCGACATCAACAGC
A C S W A Q R P W Q G Y M G N G D I N S

960
CGTTTCCAGCGTCTGCWGCAGCGCCTTCAGCAGTTCTACCAGGACGCCAAGAACTCGGCC
R F Q R L QL Q R L Q Q F Y Q D A K N S A

1020
GACACY**GACAAGGCCCTCAAGATGTGG**GTCGATCGCAAGAATGTTCTGGTTGATTGCGAA
D T D K A L K M W V D R K N V L V D C E

1080
GTTCCGGTACAACCTGGAGATCATYTGAAACGCGCCCAGTACAAGGACGTGATCGAGCGC
V P V Q P G D H L N R A Q Y K D V I E R

1140
GASGGTCCGTTCCAGAACAAGATCAAACCTCTGCGTAACTTCCCTGATCGAGCTGAACAAG
DE G P F Q N K I K L C V T S L I E L N K

1200
TGCGAAGTCATGCAGAAGGCTGCCTACTCTCGTGATGTCCGTCCGGCCTTTGAGTGTGTC
C E V M Q K A A Y S R D V R P A F E C V

1210
ATGAAAGGCAAGGGCAG
M K G K G

Figura 15. (continuação)

5.5 Desequilíbrio de Ligação

Os dados obtidos com os 18 SNPs genotipados foram analisados com o programa Arlequin e demonstraram um significativo Desequilíbrio de Ligação (LD) entre SNPs de um mesmo gene (Tabela XII), mas nenhum LD foi detectado entre SNPs de genes diferentes para as três populações analisadas.

Tabela XII. Histograma do número de *loci* ligados por *locus* para as populações de Maceió (A), Botucatu (B) e Juazeiro do Norte (C). (nível de significância = 0.0500)

A

Locus	#	N	T1	T2	C1	C2	A1	A2	A3	E1	E2	P1	P2	P3	P4	M1	M2	M3	M4
N		*	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
T1		-	*	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T2		-	+	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C1		-	-	-	*	+	-	-	-	-	-	-	-	-	-	-	-	-	-
C2		-	-	-	+	*	-	-	-	-	-	-	-	-	-	-	-	-	-
A1		-	-	-	-	-	*	+	+	-	-	-	-	-	-	+	+	+	+
A2		-	-	-	-	-	+	*	+	-	-	-	-	-	-	-	-	-	-
A3		-	-	-	-	-	+	+	*	-	-	-	-	-	-	-	-	-	-
E1		-	-	-	-	-	-	-	-	*	+	-	+	-	-	-	-	-	-
E2		-	-	-	-	-	-	-	-	+	*	-	+	-	-	-	-	-	-
P1		-	-	-	-	-	-	-	-	-	*	-	-	-	-	-	-	-	-
P2		-	-	-	-	-	-	-	-	+	+	-	*	+	+	-	-	-	-
P3		-	-	-	-	-	-	-	-	-	-	-	+	*	+	-	-	-	-
P4		-	-	-	-	-	-	-	-	-	-	-	+	+	*	-	-	-	-
M1		+	-	-	-	-	+	-	-	-	-	-	-	-	-	*	+	+	+
M2		+	-	-	-	-	+	-	-	-	-	-	-	-	-	+	*	+	+
M3		-	-	-	-	-	+	-	-	-	-	-	-	-	-	+	+	*	+
M4		+	-	-	-	-	+	-	-	-	-	-	-	-	-	+	+	+	*

Locus: N– NaK; T1- Tsf(1); T2- Tsf(2); C1- CYP9J2(1); C2- CYP9J2(2); A1- apolLpII-(1); A2– apolLpII-(2); A3– apoLpII-(3); E1- Ef2(1); E2- Ef2(2); P1- PGK(1); P2- PGK(2); P3- PGK(3); P4- PGK(4); M1- MUCI(1); M2- MUCI(2); M3- MUCI(3); M4- MUCI(4).

Tabela XII. Histograma do número de *loci* ligados por *locus* para as populações de Maceió (A), Botucatu (B) e Juazeiro do Norte (C). (nível de significância = 0.0500) (continuação)

B

Locus	#	N	T1	T2	C1	C2	A1	A2	A3	E1	E2	P1	P2	P3	P4	M1	M2	M3	M4
-N		*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T1		-	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T2		-	-	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C1		-	-	-	*	+	-	-	-	-	-	-	-	-	-	-	-	-	-
C2		-	-	-	+	*	-	-	-	-	-	-	-	-	-	-	-	-	-
A1		-	-	-	-	-	*	-	-	-	-	-	-	-	-	-	-	-	-
A2		-	-	-	-	-	-	*	+	+	+	-	-	-	-	-	-	-	-
A3		-	-	-	-	-	-	+	*	-	-	-	-	-	-	-	-	-	-
E1		-	-	-	-	-	-	+	-	*	+	+	-	-	-	-	-	-	-
E2		-	-	-	-	-	-	+	-	+	*	-	-	-	-	-	-	-	-
P1		-	-	-	-	-	-	-	-	+	-	*	-	-	-	-	-	-	-
P2		-	-	-	-	-	-	-	-	-	-	-	*	-	-	-	-	-	-
P3		-	-	-	-	-	-	-	-	-	-	-	-	*	-	-	-	-	-
P4		-	-	-	-	-	-	-	-	-	-	-	-	-	*	-	-	-	-
M1		-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	+	+	+
M2		-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	*	+	+
M3		-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	*	+
M4		-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	*

Locus: N- NaK; T1- TsF(1); T2- TsF(2); C1- CYP9J2(1); C2- CYP9J2(2); A1- apolLpII-(1); A2- apolLpII-(2); A3- apoLpII-(3); E1- Ef2(1); E2- Ef2(2); P1- PGK(1); P2- PGK(2); P3- PGK(3); P4- PGK(4); M1- MUCI(1); M2- MUCI(2); M3- MUCI(3); M4- MUCI(4).

Tabela XII. Histograma do número de *loci* ligados por *locus* para as populações de Maceió (A), Botucatu (B) e Juazeiro do Norte (C). (nível de significância = 0.0500) (continuação)

C

Locus	#	N	T1	T2	C1	C2	A1	A2	A3	E1	E2	P1	P2	P3	P4	M1	M2	M3	M4
-N		*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T1		-	*	+	-	+	-	-	-	-	-	-	+	+	+	-	-	-	-
T2		-	+	*	+	+	-	-	-	-	-	-	-	+	+	-	-	-	-
C1		-	-	+	*	+	-	-	-	-	-	-	-	+	+	-	-	-	-
C2		-	+	+	+	*	+	-	-	-	-	-	-	+	+	-	-	-	-
A1		-	-	-	-	+	*	+	+	-	-	-	-	-	-	-	-	-	-
A2		-	-	-	-	-	+	*	+	-	-	-	-	-	-	-	-	-	-
A3		-	-	-	-	-	+	+	*	-	-	-	-	-	-	-	-	-	-
E1		-	-	-	-	-	-	-	-	*	+	-	-	-	-	-	-	-	-
E2		-	-	-	-	-	-	-	-	+	*	-	-	-	-	-	-	-	-
P1		-	-	-	-	-	-	-	-	-	-	*	-	-	-	-	-	-	-
P2		-	+	-	-	-	-	-	-	-	-	-	*	+	+	-	+	+	-
P3		-	+	+	+	+	-	-	-	-	-	-	+	*	+	-	-	-	-
P4		-	+	+	+	+	-	-	-	-	-	-	+	+	*	-	-	-	-
M1		-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	+	+	+
M2		-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	*	+	+
M3		-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	*	+
M4		-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	*

Locus: N- NaK; T1- Tsf(1); T2- Tsf(2); C1- CYP9J2(1); C2- CYP9J2(2); A1- apolLpII-(1); A2- apolLpII-(2); A3- apoLpII-(3); E1- Ef2(1); E2- Ef2(2); P1- PGK(1); P2- PGK(2); P3- PGK(3); P4- PGK(4); M1- MUCI(1); M2- MUCI(2); M3- MUCI(3); M4- MUCI(4).

Quando SNPs identificados no mesmo gene apresentaram o mesmo padrão de ligação, apenas o mais polimórfico foi escolhido para as análises. Oito SNPs foram utilizados na caracterização das populações de Botucatu, Juazeiro do Norte e Maceió e compõe as análises posteriores: apolLp-II(2), Ef2(1), MUCI(1), NaK(1), CYP9J2(1), PGK(1), PGK(2), Tsf(2) (Tabela XIII). Os SNPs Chym(1), FerH(1) e FerH(2) foram selecionados, mas não utilizados nas análises das populações porque o DNA utilizado no estudo foi insuficiente para realizar todas as reações.

Tabela XIII. SNPs utilizados no estudo de caracterização das populações de Botucatu, Juazeiro do Norte e Maceió, do mosquito *Aedes aegypti* no Brasil.

Gene	Ns	Posição	S/V	S/N	Localização
apoLp-II	2	068: A/C	V	S	Éxon
Chym	1	442: A/G	S	N	Éxon
Ef2	1	054: C/T	S	S	Éxon
FerH	1	162: C/T	S	S	Éxon
	2	377: A/C	V	-	Éxon
CYP9J2	1	164: A/G	S	S	Éxon
MUC	1	159: C/T	S	S	Éxon
NaK	1	342: A/T	V	-	Íntron
PGK	1	044: T/C	S	N	Éxon
	2	083: A/G	S	S	Éxon
Tsf	2	966: T/C	S	S	Éxon

Ns = número do SNP selecionado. S/V = (S) transição / (V) transversão; S/N = (S) sinônima / (N) não-sinônima.

5.6 Frequências alélicas

Após o cálculo das frequências alélicas e dos histogramas observamos que as frequências para os alelos de cada loco variaram entre as três populações estudadas (Figura 16).

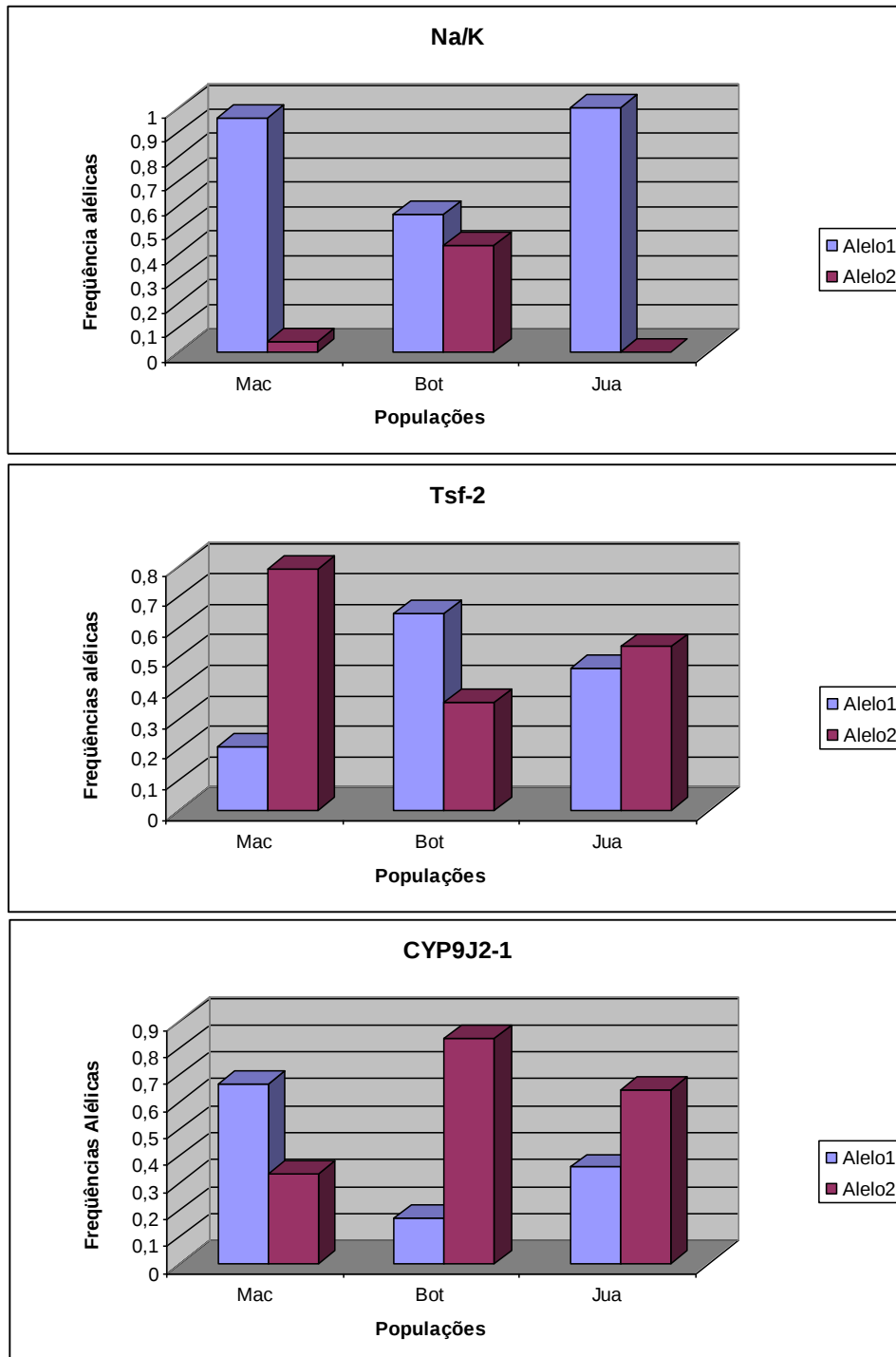


Figura 16. Histogramas das freqüências alélicas dos oito SNPs estimados para os 76 indivíduos das populações de Maceió, Botucatu e Juazeiro do Norte. O eixo Y indica a freqüência alélica e o eixo X as populações.

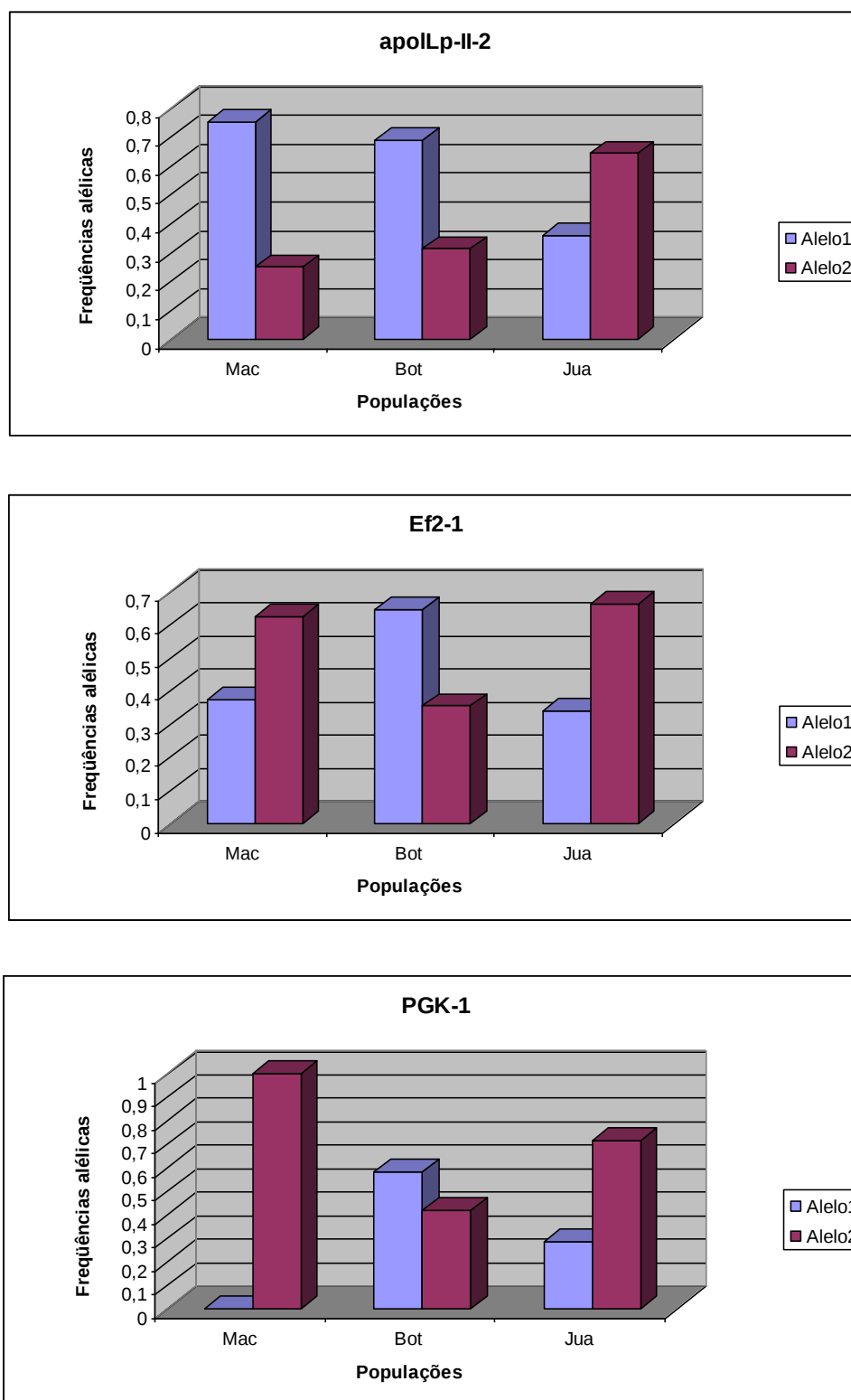


Figura 16. (continuação)

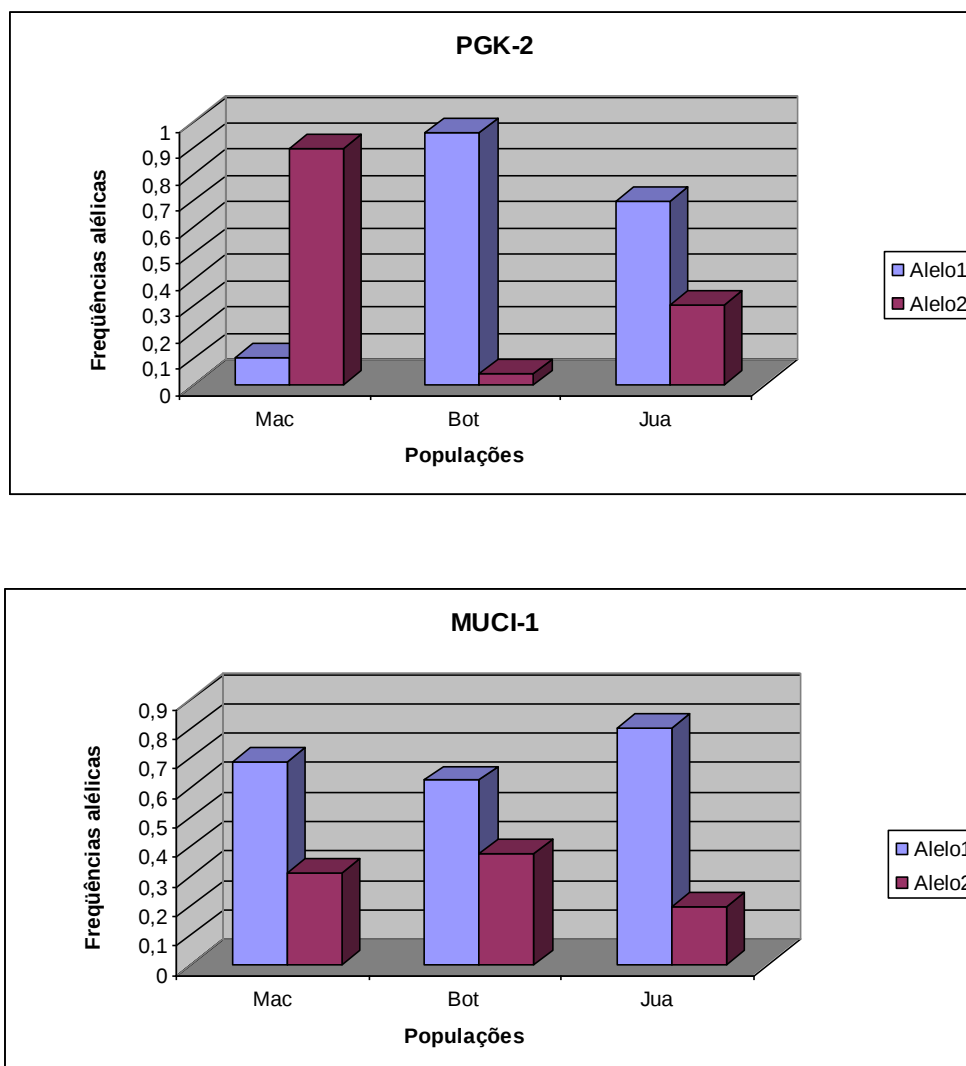


Figura 16. (continuação)

5.7 Equilíbrio de Hardy-Weinberg

A maioria dos SNPs genotipados nas três populações não mostrou nenhum desvio significativo do equilíbrio de Hardy-Weinberg. A população de Botucatu apresentou 5 *loci* com desvio entre a heterozigosidade observada (H_O) e a heterozigosidade esperada (H_E), mas o desvio do número de heterozigotos foi significativo apenas para o *locus* PGK-2, mostrando que nesta população este marcador não se apresenta em equilíbrio de HW (Tabela XIV). A população de Juazeiro do Norte apresentou apenas o *locus* CYP9J2 com desvio do número de heterozigotos e teste estatístico significativo, demonstrando que o marcador não está em equilíbrio de HW nesta população (Tabela XV). Para a população de Maceió, o cálculo para as heterozigosidades observadas (H_O) e esperadas (H_E) mostrou

desvio apenas para o *locus* CYP9J2, mas o teste estatístico não foi significativo sugerindo que nesta população todos os marcadores testados estão em equilíbrio de HW (Tabela XVI).

Tabela XIV. Frequência do alelo raro, heterozigosidades observada e esperada, estimativas de F_{IS} (coeficiente de endogamia) e significância dos testes, para 8 SNPs de genes nucleares na população de *Aedes aegypti* de Botucatu.

Botucatu						
Locus	N	R	H_O	H_E	F_{IS}	P
apoLp-II	24	0.312	0.625	0.467	-0.438	0.06
Ef2	24	0.339	0.375	0.493	+0.201	0.39
NaK	24	0.437	0.625	0.520	-0.250	0.40
MUCI	24	0.375	0.500	0.504	-0.045	1.00
CYP9J2	24	0.166	0.250	0.318	+0.121	0.50
PGK-1	24	0.583	0.333	0.537	+0.333	0.20
PGK-2	24	0.041	0.000	0.122	+1	0.02
Tsf	24	0.645	0.291	0.467	+0.381	0.08

N = tamanho da amostra; R = frequência do alelo raro; H_E e H_O = heterozigosidades esperada e observada para genes nucleares; F_{IS} = teste exato para as proporções de Hardy-Weinberg; P = significância dos testes.

Tabela XV. Frequência do alelo raro, heterozigosidades observada e esperada, estimativas de F_{IS} (coeficiente de endogamia) e significância dos testes, para 8 SNPs de genes nucleares na população de *Aedes aegypti* de Juazeiro do Norte.

Juazeiro do Norte						
Locus	N	R	H_O	H_E	F_{IS}	P
apoLp-II	28	0.642	0.357	0.467	+0.239	0.23
Ef2	28	0.645	0.535	0.456	-0.177	0.42
MUCI	28	0.232	0.196	0.350	+0.225	0.25
NaK	28	0.000	<i>Locus monomórfico</i>			
CYP9J2	28	0.357	0.214	0.502	+0.546	0.01
PGK-1	28	0.285	0.500	0.450	-0.208	0.37
PGK-2	28	0.303	0.392	0.465	+0.089	0.67
Tsf	28	0.464	0.428	0.525	+0.156	0.46

N = tamanho da amostra; R = frequência do alelo raro; H_E e H_O = heterozigosidades esperada e observada para genes nucleares; F_{IS} = teste exato para as proporções de Hardy-Weinberg; P = significância dos testes.

Tabela XVI. Frequência do alelo raro, heterozigosidades observada e esperada, estimativas de F_{IS} (coeficiente de endogamia) e significância dos testes, para 8 SNPs de genes nucleares na população de *Aedes aegypti* de Maceió.

Maceió						
Locus	N	R	H_O	H_E	F_{IS}	P
apoLp-II	24	0.250	0.500	0.382	-0.314	0.27
Ef2	24	0.375	0.666	0.478	-0.405	0.071
MUCI	24	0.333	0.312	0.438	-0.241	0.35
NaK	24	0.041	0.083	0.081	-0.022	1.00
CYP9J2	24	0.666	0.250	0.453	+0.455	0.06
PGK-1	24	0.000	Locus monomórfico			
PGK-2	24	0.895	0.208	0.190	-0.095	1.00
Tsf	24	0.208	0.416	0.336	-0.243	0.53

N = tamanho da amostra; R = frequência do alelo raro; H_E e H_O = heterozigosidades esperada e observada para genes nucleares; F_{IS} = teste exato para as proporções de Hardy-Weinberg; P = significância dos testes.

5.8 Estrutura genética populacional

Através da análise utilizando o teste AMOVA, com base nos dados dos 8 *loci* polimórficos, as estimativas da estatística F de Wright revelaram que as populações estão altamente estruturadas. Verificou-se que do total de diversidade genética, 23,65% está entre populações, e apenas 3,71% dentro das populações, sendo o valor de F_{ST} igual a 0,2364. Grande parte da variabilidade genética (72,01%) foi identificada dentro dos indivíduos, provavelmente por este ser um marcador diplóide de análise *multilocus* (Tabela XVII).

Tabela XVII. Análise de variância molecular (AMOVA) da frequência dos marcadores SNPs entre populações de *Aedes aegypti*, sem agrupamento.

Fonte de variação	GL	SQ	Porcentagem total da variação	P	Estatística <i>F</i>
Entre populações	2	111.073	23.65%	(< 0,01)	$F_{ST} = 0.2364$
Dentro de populações	73	254.631	3.71%	(> 0,05)	
Dentro dos indivíduos	76	240.500	72.64%	(< 0,01)	
Total	151	606.204			

GL = grau de liberdade; SQ = soma dos quadrados.

5.10 F_{ST} pairwise

A diferenciação genética entre pares de populações foi investigada pelas estimativas F_{ST} . Altos níveis de diferenciação genética foram obtidos nas três populações definidas de acordo com o modelo de agrupamento por distância genética. A média dos F_{ST} pairwise variou de + 0.15888 a + 0.37738, e foi altamente significativa ($P < 10^{-5}$). A população de Juazeiro compartilhou do mesmo ramo com a população de Botucatu, ou seja, se mostrou geneticamente mais próxima desta (+ 0.15888) e menos da população de Maceió (+ 0.20602). As populações de Botucatu e Maceió apareceram em ramos separados, e se mostraram geneticamente mais distantes (+ 0.37738), e todos os valores de diferenciação genética entre pares de populações foram altamente significativos ($P < 10^{-5}$) (Tabela XVIII).

Tabela XVIII. Matriz de diferenciação genética (F_{ST} pairwise) das populações de *Aedes aegypti* calculado a partir do método da estatística *F*.

Populações	Maceió	Botucatu	Juazeiro
Maceió	-		
Botucatu	0.37738 ^a	-	
Juazeiro	0.20602 ^a	0.15888 ^a	-

^a = $P < 10^{-5}$

5.11 Estruturação populacional

A estruturação genética das populações de mosquitos *Ae. aegypti* com base na informação individual dos genótipos para cada *locus*, foi possível por meio da utilização do programa Structure 2.0. O agrupamento foi realizado para os oito SNPs genotipados. Os resultados da análise de clusters revelaram a presença duas linhagens genéticas ($K = 2$) entre as populações do mosquito *Ae. aegypti*. Quando este número foi fixado, uma estrutura notável apareceu com 99,99% dos indivíduos sendo corretamente determinados para a linhagem a que pertencem (probabilidade de inclusão no grupo $> 99,99$). Uma das linhagens é representada pela população de Botucatu e a outra pela de Maceió. Os mosquitos de Juazeiro do Norte pareceram como um grupo de híbridos entre os dois anteriormente citados (Figura 17).

Figura 17. Resultado da análise de *clusters*. Cada um dos 76 indivíduos incluídos na análise é representado por uma linha vertical, dividida em duas cores. Estas cores representam a probabilidade do indivíduo pertencer a um dos *clusters*.



VI. DISCUSSÃO

A diferenciação genética espacial de populações de *Ae. aegypti* tem sido amplamente estudada em todo o mundo com diferentes classes de marcadores moleculares. Entretanto, devido às limitações apresentadas por cada uma das técnicas aplicadas, a utilização destas ferramentas em cada estudo tem revelado dados diversificados e pouco aplicáveis ao estudo genético das populações do *Ae. aegypti*, em nível local e mundial. Há ainda pouca relação dos dados obtidos com os fenômenos como migração, seleção e resistência aos inseticidas. Estudar e compreender os padrões de estrutura genética de populações locais deste culicídeo pode fornecer informações importantes sobre a dinâmica do *Ae. aegypti* nas cidades, tão importante para compreendermos aqueles fenômenos. Desta forma, a busca de novos marcadores moleculares têm se tornado cada vez mais necessária. Marcadores SNPs localizados em genes nucleares com disposição aleatória no genoma possibilitam a análise de diferenças associadas com herança biparental dos marcadores analisados.

A utilização de marcadores SNPs em estudos de genética de populações para o *Ae. aegypti* visa fornecer uma ferramenta para analisar a variabilidade genética ao longo do genoma. Este estudo é o primeiro que permite comparar diretamente a variação em *loci* SNPs com a variação observada, utilizando marcadores mitocondriais (Paduan & Ribolla, 2008) e RAPD (Paduan et al., 2006) em populações naturais de *Ae. aegypti* no Brasil. Esta pesquisa foi iniciada com a proposta de desenvolver marcadores moleculares nucleares, polimórficos, de fácil aplicação e genotipagem, com a capacidade de avaliar a estrutura genética de populações naturais deste culicídeo no Brasil. Para tanto, foram investigadas regiões codificantes e não-codificantes de nove genes nucleares distribuídos ao longo do genoma. As populações naturais analisadas apresentaram elevados índices de diversidade nucleotídica, confirmando a alta diversidade observada anteriormente em genomas de insetos (Aquadro et al., 2001; Holt et al., 2002; Morlais & Severson, 2003).

Estudos anteriores reportaram uma diversidade de 12SNPs/kb em populações mantidas em laboratório (Morlais & Severson, 2003) e de 30 a 124 SNPs/kb em populações naturais da Venezuela (Urdaneta-Marquez et al., 2008). Em nosso estudo foram identificados de 2 a 45 SNPs/kb nos 9 genes nucleares pesquisados. Esta alta diversidade de nucleotídeos detectada em estudos com *Ae. aegypti* vêm demonstrando que os SNPs poderiam ser extremamente úteis como marcadores genéticos, principalmente em estudos com tecnologias *high-throughput*, estudos de ligação e análises de características complexas

(Morlais et al., 2003). A diversidade intrapopulacional baseada nos marcadores SNPs investigados em nosso estudo foi maior quando comparada com marcadores RAPD e mitocondriais aplicados a estudos realizados com as mesmas populações (Paduan & Ribolla, 08; Paduan et al., 2006). Urdaneta-Marquez et al. (2008) também identificaram uma alta densidade de SNPs em populações naturais do *Ae. aegypti* na Venezuela, entretanto, os marcadores SNPs selecionados por eles apresentaram uma variância menor quando comparados com os haplótipos mitocondriais identificados.

A análise dos nossos dados mostrou uma distribuição heterogênea da presença de SNPs entre os diferentes genes analisados. Dois exemplos foram observados nos genes NaK e PGK que apresentaram um e 22 SNPs, respectivamente. Embora não diretamente comparáveis, o genoma do mosquito *An. gambiae*, principal vetor de malária na África, tem sido descrito com uma distribuição altamente variável para a presença de SNPs (Holt et al., 2002; Morlais et al., 2004). Holt et al. (2002) durante o seqüenciamento completo do genoma deste vetor descreveram que alguns genes apresentavam poucos SNPs e outros mais de 8 SNPs/kb. Em 2004, Morlais et al. (2004) utilizaram três cepas de laboratório para investigar a diversidade de nucleotídeos em 35 genes de regiões codificantes do genoma do *An. gambiae*. Os genes do estudo foram selecionados de acordo com um possível papel na interação patógeno-vetor ou no comportamento da espécie, e a média da diversidade de nucleotídeos nos indivíduos selecionados mostrou a presença de 1 SNP a cada 125 pb. Três *loci* SNPs selecionados em três diferentes genes foram utilizados para genotipagem de populações naturais do mosquito *An. gambiae* de três localidades com focos de malária. O padrão de polimorfismo das populações naturais foi contrastante ao padrão obtido com as cepas de laboratório, ou seja, as cepas analisadas apresentaram pouco polimorfismo e as populações naturais, SNPs altamente polimórficos. Os resultados obtidos nesta pesquisa demonstraram que estudos baseados em SNPs podem ser valiosos na identificação de variações nas seqüências de DNA visando a associação à traços fenotípicos ligados à características importantes como a competência vetora e a resistência aos inseticidas químicos.

Em nossa pesquisa os genes que apresentaram maiores taxas de substituição de nucleotídeos foram FerH, MUCI e PGK. Genes envolvidos em processos ligados à alimentação sanguínea (FerH e MUCI) e geração de ATP (PGK). Especificamente as mucinas têm se mostrado importantes moléculas na interação parasita-vetor (Morlais & Severson,

2001). De forma semelhante, altas taxas de polimorfismos por SNPs foram observadas em genes do mosquito *An. gambiae* com possível função receptora. Estes dados podem sugerir que genes associados com interações patógeno-hospedeiro podem ser bem especializados (Morlais et al., 2004).

Para todos os nove genes investigados foi observado que as substituições por transição foram mais comuns que as substituições por transversão, e que as substituições sinônimas foram mais frequentes que as substituições não-sinônimas. Além dos altos índices de SNPs detectados, consequência da substituição de nucleotídeos, três (33,3%) dos nove genes analisados neste estudo continham *indels*. Os dados obtidos neste e em outros estudos sugerem que este tipo de polimorfismo também pode ser frequente no genoma do *Ae. aegypti*. Morlais & Severson (2001) já haviam relatado sobre a presença de *indels* no genoma deste vetor; e Bhatramakki et al. (2002) relataram que este tipo de polimorfismo pode ser facilmente explorado como marcadores baseados em PCR para o mapeamento genético. Em nossa pesquisa foram identificados *indels* altamente significantes nos genes MUCI e PGK, respectivamente com 24 e 10pb. Para o gene MUCI, estudos anteriores haviam descrito este e outros *indels* dentro de seqüências codificantes. Modificações estas que resultaram em diferentes isoformas da proteína que poderiam levar a propriedades funcionais/estruturais distintas que influenciam a susceptibilidade ao *Plasmodium gallinaceum* (Morlais & Severson, 2001). Entretanto além do *indel* detectado e descrito para o gene da mucina, detectamos também a presença de um *indel*, e de substituições não-sinônimas significativas principalmente no gene PGK. Schmid & Tautz (1997), já relataram que genes que exibem um grande número de substituições não-sinônimas podem ser de evolução rápida, possivelmente porque evoluem em uma fase neutra ou por seleção positiva, com polimorfismos nas seqüências, dada a vantagem seletiva sobre uma variedade de condições. Entretanto, para verificar todas as mutações não-sinônimas e a presença de um *indel* com 10 pb em seqüências do gene PGK, além de outros, são necessários estudos mais detalhados. É necessário verificar se existe uma evolução molecular adaptativa que age em cada gene do mosquito e investigar a diversidade de nucleotídeos em espécies relacionadas como o *Ae. albopictus*, um eficiente vetor do Dengue no Continente Asiático.

Dados anteriores (Morlais e Severson, 03) e recentes (Urdaneta-Marquez et al., 2008; Paduan & Ribolla, aceito para publicação) obtidos com a identificação de SNPs ao longo do genoma do *Ae. aegypti* confirmaram o alto grau de polimorfismo que pode ser detectado

utilizando este tipo de marcador. Em nosso trabalho, as mesmas populações naturais do mosquito que haviam sido submetidas anteriormente às análises com marcadores RAPD, microssatélites e análise de genes mitocondriais apresentaram altos índices de polimorfismos em todos os sítios SNPs selecionados. Os oito marcadores SNPs estudados forneceram resultados promissores para a análise genética das populações do mosquito em escala inter e intrapopulacional. Indivíduos da população de Botucatu (BOT) que haviam sido submetidos a análise com marcadores RAPD, microssatélites e sequenciamento de genes mitocondriais, sem apresentar qualquer diferenciação entre indivíduos de uma mesma população, mostraram altas taxas de polimorfismo por SNPs nos indivíduos analisados. O mesmo foi observado nas populações de Maceió (MAC) e Juazeiro do Norte (JUA), que quando analisadas pelo sequenciamento de genes mitocondriais apresentaram quase que em sua totalidade, apenas um haplótipo. Entretanto, quando estas populações foram analisadas com os marcadores SNPs apresentaram índices elevados de polimorfismos intrapopulacionais.

A caracterização de indivíduos das populações de BOT, MAC e JUA pela análise de RAPD, demonstrou uma pequena diferenciação genética das populações do nordeste em relação a populações de outras regiões do país, com pouca ou nenhuma diferenciação de indivíduos dentro de cada população. A genotipagem dos indivíduos das populações de MAC (10 indivíduos), BOT (10 indivíduos) e JUA (10 indivíduos) pela análise do mtDNA evidenciou a presença de duas linhagens históricas principais introduzidas em nosso país. Na linhagem mais numerosa, denominada (1) estavam presentes todos os indivíduos das três populações analisadas, sendo que BOT e MAC, compartilhavam um único haplótipo. Apenas a população de JUA, apresentou cinco indivíduos que tinham três haplótipos diferentes da linhagem 1. Após a análise dos indivíduos de MAC (24 indivíduos), BOT (24 indivíduos) e JUA (28 indivíduos) pelos marcadores SNPs ficou evidente a presença de duas linhagens principais, sendo uma representada pela população de BOT, uma representada pela população de MAC, e espécimes de JUA se apresentaram como uma população intermediária a BOT e MAC. A utilização dos marcadores SNPs permitiu diferenciar principalmente as populações de BOT e JUA que com relação à análise de mtDNA apresentavam o mesmo haplótipo e, analisadas por RAPD e microssatélites se mostraram indiferenciáveis em nível local.

Esta diferença observada entre a análise do mtDNA e dos marcadores SNPs reflete o que podemos esperar das análises populacionais utilizando-se marcadores dominantes, não

mendelianos, do tipo mitocondrial. A similaridade encontrada entre as populações de BOT e MAC representa apenas o reflexo de uma história evolutiva passada, mas não a real distribuição destes indivíduos atualmente. Isto se deve ao fato de que o DNA mitocondrial não sofre hibridação, mas apenas traça o histórico materno dos indivíduos. Análises por marcadores SNPs dos indivíduos que contém outros haplótipos de mtDNA poderá mostrar a real situação destas diferentes populações no Brasil e, mesmo esclarecer se estes dois grupos distintos de indivíduos com haplótipos diferentes estão ou não em isolamento reprodutivo.

Enquanto a análise de genomas possibilita conhecer novos aspectos da biologia dos vetores, a variação genética oferece a possibilidade de se estudar a estrutura genética das populações em vários níveis de distribuição geográfica. Os numerosos polimorfismos encontrados nos genes do mosquito *Ae. aegypti* neste estudo poderão ser utilizados para iniciar um mapeamento das populações naturais deste vetor no Brasil e para possibilitar a realização de estudos de associação. Adicionalmente, para se estudar a estrutura populacional do *Ae. aegypti* dentro dos municípios, que é geralmente influenciada por desvios genéticos e migração, análises temporais podem fornecer informação sobre as forças responsáveis pelas alterações genéticas em uma pequena escala de tempo (ex. poucas gerações).

Análises refinadas ao longo do genoma utilizando SNPs poderão ainda auxiliar na descoberta de fatores genéticos associados com características complexas. O que se sugere é que a procura de genes responsáveis por estas características seja feita em primeiro plano com um estudo de ligação, de varredura genômica não muito sensível com a finalidade de se identificar as regiões cromossômicas potencialmente candidatas para alojar o gene condicionador das características. Após essa identificação deve se utilizar uma análise mais fina nas regiões promissoras usando densos mapas de polimorfismos de um único nucleotídeo (SNP). Estudos assim têm sido realizados com o intuito de se desvendar o mecanismo de resistência dos mosquitos á inseticidas como, por exemplo, para o mosquito *An. arabiensis* (Matambo et al., 2007) investigado quanto à associação da resistência, à diferentes classes de inseticidas, com mutações no gene *kdr*. Assim, a genotipagem de um grande número de SNPs poderia facilitar o estudo da estrutura genética de populações do *Ae. aegypti*, tão necessário para prever o espalhamento de genes de interesse, uma vez que o desenvolvimento de novas medidas de controle da dengue necessita da identificação de

genes candidatos que afetem fenótipos como: antropofilia, fertilidade, ou susceptibilidade aos parasitas.

VII. CONCLUSÃO

Neste estudo,

✚ Os SNPs apresentaram-se altamente polimórficos quando comparados com outras classes de marcadores moleculares utilizados anteriormente nas três populações do mosquito *Ae. aegypti*;

✚ Os marcadores de único nucleotídeo se mostraram eficientes para estudos de genética populacional nas três populações naturais do *Ae. aegypti* estudadas;

✚ As três populações de *Ae. aegypti* analisadas neste estudo apresentaram SNPs altamente polimórficos;

✚ Os dados obtidos com os SNPs possibilitaram a obtenção de resultados com evidências de níveis elevados de estruturação populacional nas três populações de *Ae. aegypti* estudadas.

VIII. REFERÊNCIAS BIBLIOGRÁFICAS

- AGOSIN, M. Role of microsomal oxidations in insecticide degradation. In: KERKUT, G.A.; GILBERT, L.I. (Eds). **Comprehensive Insect Physiology, Biochemistry and Pharmacology**. New York: Pergamon, 1985. p.647–712.
- ALVAREZ, J.M.; HOY, M.A. Evaluation of the ribosomal ITS2 DNA sequences in separating closely related populations of the parasitoid *Ageniaspis* (Hymenoptera: Encyrtidae). **Ann. Entomol. Soc. Am.**, v.95, n.2, p.250-256, mar. 2002.
- APOSTOL, B.L. et al. Use of randomly amplified polymorphic DNA amplified by polymerase chain reaction markers to estimate the number of *Aedes aegypti* families at oviposition sites in San Juan, Puerto Rico. **Am. J. Trop. Med. Hyg.**, v.51, n.1, p.89-97, jul. 1994.
- AQUADRO, C.F.; BAUER-DUMONT, V.; REED, F.A. Genome-wide variation in the human and fruitfly: a comparison. **Curr. Opin. Genet. Dev.**, v.11, n.6, p.627–634, dez. 2001.
- ARIAS, M.A.; INFANTE-MALACHIAS, M.E. RFLP: O emprego de enzimas de restrição para detecção de polimorfismos de DNA. In: MARTIOLI, S.R. (Ed). **Biologia Molecular e Evolução**. São Paulo: Holos editora, 2001. p.143-152.
- BABIN, P.J. et al. Apolipophorin II/I, apolipoprotein B, vitellogenin, and microsomal triglyceride transfer protein genes are derived from a common ancestor. **J. Mol. Evol.**, v.49, n.1, p.150-160, jul. 1999.
- BALLINGER-CRABTREE, M.E.; BLACK, W.C.; MILLER, B.R. Use of genetic polymorphisms detected by the random-amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) for differentiation and identification of *Aedes aegypti* subspecies and populations. **Am. J. Trop. Med. Hyg.**, v.47, n.6, p.893-901, dec. 1992.
- BARBAZAN, P. et al. Characterization of three microsatellite loci for *Aedes aegypti* (Diptera: Culicidae) and their use for population genetic study. **Southeast Asian J. Trop. Med. Public Health.**, v.30, n.3, p.482–483, sep. 1999.
- BAUMANN, O.; TAKEYASU, K. Polarized distribution of Na,K-ATPase in honeybee photoreceptors is maintained by interaction with glial cells. **J. Cell Science**, v.105, n.2, p.287–301, jun. 1993.
- BERGER, J.; SUZUKI, T.; SENTI, K.A. Genetic mapping with SNP markers in *Drosophila*. **Nat. Genet.**, v.29, n.4, p.475–481, nov. 2001.
- BESANSKY, N.J.; FINNERTY, V.; COLLINS, F.H. Molecular perspectives on the genetics of mosquitoes. **Adv. Genet.**, v.30, p.123-184, 1992.

- BHATTRAMAKKI, D.; DOLAN, M.; HANAFEY, M. Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers. **Plant. Mol. Biol.**, v.48, n.5-6, p.539-547, mar. 2002.
- BLACK, W.C. et al. Population genomics: genome-wide sampling of insect populations. **Ann. Rev. Entomol.**, v.46, p.441-469, jan. 2001.
- BOSIO, C.F. et al. Genetic structure of *Aedes aegypti* populations in Thailand using mitochondrial DNA. **Am. J. Trop. Med. Hyg.**, v.72, n.4, p.434-442, apr. 2005.
- BOTSTEIN, D.; RISCH, N. Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. **Nat. Genet. Suppl.**, v.33, p.228-237, mar. 2003.
- BOYD, A.M.; FOLEY, D.H. Distribution of sibling species of the *Anopheles annulipes* complex (Diptera: Culicidae) in the Townsville region of Australia. **Australian J. Entomol.**, v.46, n.2, p.133-139, may. 2007.
- BRACCO, J.E. et al. Genetic variability of *Aedes aegypti* in the Americas using a mitochondrial gene: evidence of multiple introductions. **Mem. Inst. Oswaldo Cruz**, v.102, n.5, p.573-580, aug. 2007.
- BROGDON, W.G.; MCALLISTER, J.C. Insecticide resistance and vector control. **Emerg. Infect. Dis.**, v.4, n.4, p.605-613, 1998.
- BROOKES, A.J. The essence of SNPs. **Gene**, v.234, n.2, p.177-186, jul.1999.
- CALISHER, C.H. Arthropod-Borne and Rodent-Borne Viral Diseases: WHO Technical Report Series No. 719. **Am. J. Trop. Med. Hyg.**, v.35, n.1, p.216-217, jan. 1986.
- CHAMBERS, E.W. et al. Microsatellite isolation and linkage group identification in the yellow fever mosquito *Aedes aegypti*. **J. Hered.**, v.98, n.3, p.202-210, may-jun. 2007.
- CHAPMAN, M.J. Comparative analysis of mammalian plasma lipoproteins. In: SEGREST, J.P.; ALBERTS, J.J. (Eds). **Methods in Enzymology**. Boston: Academic Press, 1986. p. 70-143.
- CHASTEEN, N.D.; HARRISON, P.M. Mineralization in ferritin: an efficient means of iron storage. **J. Struct. Biol.**, v.126, n.3, p.182-194, jun. 1999.
- CONSOLI, R.A.G.B.; LOURENÇO-DE-OLIVEIRA, R. **Principais Mosquitos de Importância Sanitária no Brasil**. 1.ed. Rio de Janeiro: Editora Fiocruz, 1994. 225p.
- COSTA-RIBEIRO, M.C.V.; LOURENÇO-DE-OLIVEIRA, R.; FAILLOUX, A.B. Higher genetic variation estimated by microsatellites compared to isoenzyme markers in *Aedes*

- aegypti* from Rio de Janeiro. **Mem. Inst. Oswaldo Cruz**, v.101, n.8, p.917-921, dec. 2006.
- DUNKOV, B.C. et al. Isolation and characterization of mosquito ferritin and cloning of a cDNA that encodes one subunit. **Arch. Insect Biochem. Physiol.**, v.29, n.3, p.293-307, feb. 1995.
- EHRENKRANZ, J.N. Pandemic dengue in Caribbean countries and the Southern United States Past: present and potential problems. **New Engl. J. Med.**, v.285, n.26, p.1460-1469, 1971.
- EMERY, A.M. et al. Insect Na⁺/K⁺ -ATPase. **J. Insect Physiol.**, v.44, n.3-4, p.197-210, mar. 1998.
- EXCOFFIER, L.; LAVAL, G.; SCHNEIDER S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. **Evol. Bioinf. Online**, v.1, p.47-50, 2005.
- EXCOFFIER, L.; SMOUSE, P.E.; QUATTRO, J.M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics**, v.131, n.2, p.479-491, jun. 1992.
- FAGERBERG, A.J.; FULTON, R.E.; BLACK, W.C. Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. **Insect Mol. Biol.**, v.10, n.3, p.225-236, jun. 2001.
- FALUSH, D.; STEPHENS, M.; PRITCHARD, J.K. Inference of population structure using multi-locus genotype data: linked loci and correlated allele frequencies. **Genetics**, v.164, n.4, p.1567-1587, aug. 2003.
- FERRE, J.; VAN RIE, J. Biochemistry and genetics of insect resistance to *Bacillus thuringiensis*. **Ann. Rev. Entomol.**, v.47, p.501-533, jan. 2002.
- FERREIRA, M.E.; GRATTAPLAGIA, D. **Introdução ao uso de marcadores moleculares na análise genética**. Brasília: EMBRAPA, CENARGEM, 1996, 220p.
- FEYEREISEN, R. Insect P450 enzymes. **Ann. Rev. Entomol.**, v.44, p.507-533, jan. 1999.
- FIGUEIREDO, R.M.P. et al. Dengue virus type 4, Manaus, Brazil. **Emerg. Infect. Diseases**, v.14, n.4, p.667-669, 2008.
- FONSECA, D.M. et al. Emerging vectors in the *Culex pipiens* complex. **Science**, v.3, p.1535-1538, mar. 2004.

- FONSECA, D.M.; ATKINSON, C.T.; FLEISCHER, R.C. Microsatellite primers for *Culex pipiens quinquefasciatus*, the vector of avian malaria in Hawaii. **Mol. Ecol.**, v.7, n.11, p.1617-1619, nov. 1998.
- FORATTINI, O.P. **Entomologia médica**. 2ed. São Paulo, Editora da Universidade de São Paulo, 1965. 506p.
- FOUQUE, F.; GARINCI, R.; GABORIT, P. Epidemiological and entomological surveillance of the co-circulation of DEN-1, DEN-2 and DEN-4 viruses in French Guiana. **Trop. Med. Int. Health**, v.9, n.1, p.41-46, jan. 2004.
- FRANCO, O. **História da febre amarela no Brasil**. Rio de Janeiro: Departamento de Endemias Rurais, 1976. 208p.
- FRIEDLANDER, T.P.; REGIER, J.C.; MITTER, C. Nuclear gene sequences for higher level phylogenetic analysis: 14 promising candidates. **Syst. Biol.**, v.41, n.4, p.483-490, dec. 1992.
- FRIEDLANDER, T.P.; REGIER, J.C.; MITTER, C. Phylogenetic information content of five nuclear gene sequences in animals: Initial assessment of character sets from concordance and diversity studies. **Syst. Biol.**, v.43, n.4, p.511-525, dec. 1994.
- FUTUYMA, D.J. **Biologia Evolutiva**. 2ed. Ribeirão Preto: Sociedade Brasileira de Genética, 1992, 613p.
- GEISER, D.L. *Aedes aegypti* ferritin: A cytotoxic protector against iron and oxidative challenge? **FEBS J.**, v.270, n.18, p.3667-3674, sep. 2003.
- GLASSER, C.M. **Estudo da infestação do Estado de São Paulo por *Aedes aegypti* e *Aedes albopictus***. 1997. 130f. Dissertação (Mestrado) - Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo.
- GLAUBITZ, J.C. A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. **Mol. Ecol. Notes**, v.4, n.2, p.309-310, jun. 2004.
- GORMAN, J.M. et al. Mapping quantitative trait locus involved in melanotic encapsulation of foreign bodies in malaria vector *Anopheles gambiae*. **Genetics**, v.146, n.3, p.965-971, jul. 1997.
- GORROCHOTEGUI-ESCALANTE, N. et al. Breeding structure of *Aedes aegypti* populations in Mexico varies by region. **Am. J. Trop. Med. Hyg.**, v.66, n.2, p.213-222, feb. 2002.

- GORROCHOTEGUI-ESCALANTE, N. et al. Genetic isolation by distance among *Aedes aegypti* populations along the northeastern coast of Mexico. **Am. J. Trop. Med. Hyg.**, v.62, n.2, p.200-209, feb. 2000.
- GOULD, F. Mixed function oxidases and herbivore polyphagy: the devil's advocate position. **Ecol. Entomol.**, v.9, p.29-34, 1984.
- GRINBLATT, Y.; BROWNT, N.H.; KAFATOST, C.F. Isolation and characterization of the *Drosophila* translational elongation factor 2 gene. **Nucleic Acids Res.**, v.17, n.18, p.7303-7314, sep. 1989.
- GROETES, F.R.; TABACHNICK, B.E. Roles of selection intensity, major genes, and minor genes in evolution of insecticide resistance. **J. Econ. Entomol.**, v.93, n.6, p.1580-1587, dec. 2000.
- GUBLER, D.J. Dengue and Dengue Hemorrhagic Fever. **Clin. Microb. Rev.**, v.11, n.3, p.480-495, jul. 1998a.
- GUBLER, D.J. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. **Trends Microbiol.**, v.10, n.2, p.100-103, feb. 2002.
- GUBLER, D.J. Resurgent vector-borne diseases as a global health problem. **Emerg. Infect. Dis.**, v.4, n.3, p.442-450, jul. 1998b.
- GUBLER, D.J. Vigilancia activa del dengue y de la fiebre hemorrágica del dengue. **Bol. Of. Sanit. Panam.**, v.107, n.1, p.22-30, 1989.
- GUO, S.W.; THOMPSON, E.A. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. **Biometrics**, v.48, n.2, p.361-372, jun. 1992.
- HAAVIE, J.; SAETRE, G.P.; MOUNT, T. Discrepancies in population differentiation at microsatellites, mitochondrial DNA and plumage color in the pied flycatcher – inferring evolutionary processes. **Mol. Ecol.**, v.9, n.8, p.1137-1148, aug. 2000.
- HALSTEAD, S.B. The XXth century dengue pandemic: need for surveillance and research. **World Health Stat. Q.**, v.45, p.292-298, n.2/3, 1992.
- HAMBURGER, A.E. et al. Crystal structure of a secreted insect ferritin reveals a symmetrical arrangement of heavy and light chains. **J. Mol. Biol.**, v.349, n.3, p.558-569, jun. 2005.
- HAMMER, M.F. A recent common ancestry for human Y chromosomes. **Nature**, v.378, n.6555, p.376-378, nov. 1995.

- HARIZANOVA, N. et al. *Aedes aegypti* transferrin. Gene structure, expression pattern, and regulation. **Insect Mol. Biol.**, v.14, n.1, p.79-88, jan. 2005.
- HEMINGWAY, J.; RANSON, H. Insecticide resistance in insect vectors of human disease. **Ann. Rev. Entomol.**, v.45, p.371-391, jan. 2000.
- HENCHAL, E.A.; PUTNAK, J.R. The dengue viruses. **Clin. Microbiol. Rev.**, v.3, n.4, p.376-396, oct. 1990.
- HOLMES, E.C.; BARTLEY, L.M.; GARNET, G.P. The emergence of dengue past, present and future. In: KRAUSE, R.M. (Ed). **Emerging Infectors**. London: Academic Press. 1998. p.301-325.
- HOLT, R.A. et al. The genome sequence of the malaria mosquito *Anopheles gambiae*. **Science**, v.298, n.5591, p.129-149, oct. 2002.
- HONÓRIO, N.A. **Estudos de aspectos da biologia do *Aedes aegypti* (Linnaeus, 1762) e *Aedes albopictus* (Skuse, 1984), em área endêmica de dengue, no estado do Rio de Janeiro**. 1999. Dissertação (Mestrado em Biologia Parasitária) – Instituto Oswaldo Cruz, Rio de Janeiro, 1999.
- HOY, M.A. **Insect molecular genetics: an introduction to principles and applications**. Academic Press, New York. 2nd Edition. 600p. 2003.
- HUBBY, J.L.; LEWONTIN, R.C. A molecular approach to the study of genic heterozygosity in natural populations. ii. amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. **Genetics**, v.54, n.2, p.595-609, aug. 1966.
- HUBER, K. et al. Human transportation influences *Aedes aegypti* gene flow in Southeast Asia. **Acta Trop.**, v.90, n.1, p.23-29, mar. 2004.
- HUBER, K. et al. Isolation and variability of polymorphic microsatellite loci in *Aedes aegypti*, the vector of dengue viruses. **Mol. Ecol. Notes**, v.1, n.4, p.219-222, dec. 2001.
- HUBER, K. et al. Short report: microsatellite sequences as markers for population genetic studies of the mosquito *Aedes aegypti*, the vector of dengue viruses. **Am. J. Trop. Med. Hyg.**, v.61, n.6, p.1001-1003, dec. 1999.
- HUBER, K. et al. Temporal genetic variation in *Aedes aegypti* populations in Ho Chi Minh City (Vietnam). **Heredity**, v.89, n.1, p.7-14, jul. 2002.
- JIANG, Q. et al. cDNA cloning and patterns of expression of an adult, female-specific chymotrypsin from *Aedes aegypti* midgut. **Insect Biochem. Molec. Biol.**, v.27, n.4, p.283-289, apr. 1997.

- JUNGREIS, A.M.; VAUGHAN, G.L. Insensitivity of lepidopteran tissue to ouabain: absence of ouabain binding and Na⁺/K⁺-ATPases in larval and adult midgut. **J. Insect Physiol.**, v.23, n.4, p.503-509, 1977.
- KARP, A.; SEBERG, O.; BUIATTI, M. Molecular techniques in the assessment of botanical diversity. **Ann. Bot.**, v.78, n.2, p.143-149, aug. 1996.
- KEYGHOBADI N, et al. Microsatellite loci from the northern house mosquito (*Culex pipiens*), a principal vector of West Nile virus in North America. **Mol. Ecol. Notes.**, v.4, n.1, p.20–22, mar. 2004.
- LE HELLARD, S. et al. SNP genotyping on pooled DNAs: comparison of genotyping technologies and a semi automated method for data storage and analysis. **Nucleic Acids. Res.**, v.30, n.15, p.e74, aug. 2002.
- LEBOVITZ, R.M.; TAKEYASU, K.; FAMBROUGH, D.M. Molecular characterization of the (Na⁺ + K⁺)-ATPase α -subunit in *Drosophila melanogaster*. **EMBO J.**, v.8, n.1, p.193-202, jan. 1989.
- LIU, B.H. **Statistical genomics**. New York: CRC, 1998. 610p.
- LOURENÇO-DE-OLIVEIRA, R. et al. *Aedes aegypti* in Brazil: genetically differentiated populations with high susceptibility to dengue and yellow fever viruses. **Trans. R. Soc. Trop. Med. Hyg.**, v.98, n.1, p.43-54, jan. 2002.
- LOXDALE, H.D.; LUSHAI, G. Molecular markers in entomology. **Bull. Entomol. Res.**, v.88, p.577-600, 1998.
- LYNCH, M.; MILLIGAN, B.G. Analysis of population genetic structure with RAPD markers. **Mol. Ecol.**, v.3, n.2, p.91-99, apr. 1994.
- MATAMBO, T.S. et al. Insecticide resistance in the malarial mosquito *Anopheles arabiensis* and association with the *kdr* mutation. **Med. Vet. Entomol.**, v.21, n.1, p.97-102, mar. 2007.
- MATIOLI, S. R. Glossário. In: Sergio Russo Matioli. (Org.). *Biologia molecular e evolução*. 1 ed. Ribeirão Preto: Holos, editora, 2001, v. 1, p. 193-198.
- MATTINGLY, P.E. General aspects of the *Aedes aegypti* problem. In *Taxonomy and Bionomics*. **Ann. Trop. Med. Parasitol.**, v.51, p.392-408, 1957.
- MINISTÉRIO DA SAÚDE. Relatório final de Reunião Técnica do Programa de Combate à Febre Amarela e Dengue. Brasília, 1991.

- MINISTÉRIO DA SAÚDE. Séries históricas de agravos e doenças transmissíveis. **Inf. Epidemiol. SUS**, v.1, p.25, 1992.
- MITCHELL, C.J. Geographic spread of *Aedes albopictus* and potential for involvement in arbovirus cycles in Mediterranean Basin. **J. Vector Ecol.**, v.20, n.1, p.44-58, jun. 1995.
- MORENO, M. Genetic population structure of *Anopheles gambiae* in Equatorial Guinea. **Malaria J.**, v.6, p.137-147, oct. 2007.
- MORIYAMA, E.N.; POWELL, J.R. Intraspecific nuclear DNA variation in *Drosophila*. **Mol. Biol. Evol.**, v.13, n.1, p.261-277, jan. 1996.
- MORLAIS, I. et al. A targeted approach to the identification of candidate genes determining susceptibility to *Plasmodium gallinaceum* in *Aedes aegypti*. **Mol. Genet. Genomics**, v.269, n.6, p.753-764, aug. 2003.
- MORLAIS, I. et al. Intraspecific nucleotide variation in *Anopheles gambiae*: new insights into the biology of malaria vectors. **Am. J. Trop. Med. Hyg.**, v.71, n.1, p.795-802, dec. 2004.
- MORLAIS, I.; SEVERSON, D.W. Identification of a polymorphic mucin-like gene expressed in the midgut of the mosquito, *Aedes aegypti*, using an integrated bulked segregant and differential display analysis. **Genetics**, v.158, p.1125-1136, jul. 2001.
- MORLAIS, I.; SEVERSON, D.W. Intraespecific DNA variation in nuclear genes of the mosquito *Aedes aegypti*. **Insect. Mol. Biol.**, v.12, n.6, p.631-639, dec. 2003.
- MOUSSON, L. et al. Phylogeography of *Aedes* (*Stegomyia*) *aegypti* (L.) and *Aedes* (*Stegomyia*) *albopictus* (Skuse) (Diptera: Culicidae) based on mitochondrial DNA variations. **Genet. Res.**, v.86, n.1, p.1-11, aug. 2005.
- MULLIS, K.B.; FALOONA, F.A. Specific Synthesis of DNA in vitro via a Polymerase-catalyzed Chain Reaction. **Methods Enzymol.**, v.155, p.335-350, 1987.
- NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. **Genetics**, v.89, n.3, p.583-590, jul. 1978.
- NELSON, D.R. et al. P450 Superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature. **Pharmacogenetics**, v.6, n.1, p.1-42, feb. 1996.
- NICHOL, H.; LAW, J.H.; WINZERLING, J.J. Iron metabolism in insects. **Annu. Rev. Entomol.**, v.47, n.1, p.535-559, jan. 2002.
- NIJHOUT, H.F.; CARROW, G.M. Diuresis after a bloodmeal in female *Anopheles freeborni*. **J. Insect Physiol.**, v.24, n.4, p.293-298, 1978.

- NOGUEIRA, R.M.R. et al. Virological study of a dengue type I epidemic at Rio de Janeiro. **Mem. Inst. Oswaldo Cruz.**, v.83, p.219-25, 1988.
- NORIEGA, F.G. et al. Early tripsin, a female-specific midgut protease in *Aedes aegypti*: Isolation, amino-terminal sequence determination, and cloning and sequencing of the gene. **Insect Biochem. Mol. Biol.**, v.26, n.2, p.119-126, feb. 1996.
- OMS - ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Dengue hemorrágico: diagnóstico, tratamento e controle.** Genebra, 1987.
- OMS - Organização Mundial de Saúde. **Número de casos de dengue.** Disponível em: <http://www.who.int/mediacentre/factsheets/fs117/en/> (Acessado em: Junho de 2008).
- OPAS - Organización Panamericana de La Salud. El dengue y la fiebre hemorrágica de dengue em las Américas: uma visão geral Del problema. **Bol. Epidemiol.**, v.13, n.1, p.9-10, 1992.
- OSANAI, C.H. **A epidemia de dengue em Boa Vista, Território Federal de Roraima, 1981-1982.** 1984. Dissertação (Mestrado) - Escola Nacional de Saúde Pública, Rio de Janeiro.
- OSANAI, C.H. **A epidemia de dengue em Boa Vista, Território Federal de Roraima, 1981-1982.** Rio de Janeiro, 1984. [Dissertação de Mestrado-Escola Nacional de Saúde Pública].
- PADUAN, K.S.; ARAÚJO-JÚNIOR, J.P.; RIBOLLA, P.E.M. Genetic and variability in geographical populations of *Aedes aegypti* in Brazil elucidated by molecular markers. **Genet. Mol. Biol.**, v.29, n.2, p.391-395, 2006.
- PADUAN, K.S.; RIBOLLA, P.E.M. Characterisation of 8 single nucleotide polymorphism markers in *Aedes aegypti*. **Mol Ecol. Notes**, aceito para publicação.
- PADUAN, K.S.; RIBOLLA, P.E.M. Mitochondrial DNA polymorphism and heteroplasmy in populations of *Aedes aegypti* in Brazil. **J. Med. Entomol.**, v.45, n.1, p.59-67, jan. 2008.
- PAHO - Pan American Health Organization. **Number of reported cases of dengue & dengue hemorrhagic fever (DHF), Region of the Americas (by country and subregion).** 2006.
- PAUPY, C. et al. Influence of breeding sites features on genetic differentiation of *Aedes aegypti* populations analyzed on a local scale in Phnom Penh municipality of Cambodia. **Am. J. Trop. Med. Hyg.**, v.71, n.1, p.73-81, jul. 2004.
- PENNA, M.L.F. Um desafio para a saúde pública brasileira: o controle do dengue. **Cad. Saúde Publ.**, v.19, n.1, p.305-309, 2003.

- PEREIRA, S.L. Mitochondrial genome organization and vertebrate phylogenetics. **Genet. Mol. Biol.**, v.23, n.4, p.745-752, dec. 2000.
- PINHEIROS, P.P. Dengue in the Americas: 1980-1987. **Epidemiol. Bull.**, v.10, n.1, p.1-8, 1989.
- PONTES R.J.S. Epidemia de Dengue em Ribeirão Preto, São Paulo, Brasil: Nota Prévia. **Rev. Saúde Pública**, v.25, n.4, p.315-317, oct. 1991.
- PRITCHARD, J.K.; ROSENBERG, N.A. Use of unlinked genetic markers to detect population stratification in association studies. **Am. J. Hum. Genet.**, v.65, n.1, p.220–228, jul. 1999.
- PRITCHARD, J.K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v.155, n.2, p.945-959, jun. 2000. (program structure, version 2.1 available from: http://pritch.bsd.uchicago.edu/software/structure2_1.html).
- PRITCHARD, J.K.; WEN, W. **Documentation for STRUCTURE Software**. Version 2 (<http://pritch.bsd.Uchicago/edu>). abr. 2007.
- RANSON, H. et al. Genetic mapping of genes conferring permethrin resistance in the malaria vector *Anopheles gambiae*. **Insect Mol. Biol.**, v.13, n.4, p.379-386, aug. 2004.
- RAVEL, S. et al. A preliminary study of the population genetics of *Aedes aegypti* (Diptera: Culicidae) from Mexico using microsatellite and AFLP markers. **Acta Trop.**, v.78, n.3, p.241-250, mar. 2001.
- RAVEL, S. et al. Microsatellite markers for population genetic studies in *Aedes aegypti* (Diptera: Culicidae) from Côte d'Ivoire: evidence for a microgeographic genetic differentiation of mosquitoes from Bouaké. **Acta Trop.**, v.82, n.1, p.39-49, apr. 2002.
- REEVES, S.A.; YAMANAKA, M.K. Cloning and sequence analysis of the α subunit of the cat flea sodium pump. **Insect Biochem. Mol. Biol.**, v.23, n.7, p.809-814, oct. 1993.
- RIGAU-PÉREZ, J.G.; VORNDAM, A.V.; CLARK, G.G. The dengue and dengue hemorrhagic fever epidemic in Puerto Rico. **Am. J. Trop. Med. Hyg.**, v.64, n.1, p.67-74, jan. 2001.
- RODERICK, G.K. Geographic structure of insect population: gene flow, phylogeography and their uses. **Ann. Rev. Entomol.**, v.41, p.325-352, jan. 1996.
- ROSE, R.I. Pesticides and public health: integrated methods of mosquito management. **Emerg. Infect. Diseases.**, v.7, n.1, p.17-23, jan-feb. 2001.
- ROSENBERG, N.A. et al. Genetic structure of human populations. **Science**, v.298, n.5602, p.2381-2385, dec. 2002.

- ROZAS, J. et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods. **Bioinformatics**, v.19, n.18, p.2496-2497, dec. 2003.
- RUDNICK, A. Dengue fever epidemiology in Malaysia 1901–1980. In: RUDNICK, A.; LIM, T.W. (Eds). **Dengue Fever Studies in Malaysia, Bulletin**. Kuala Lumpur: Institute of Medical Research. 1986. p.9-38.
- SAIKI, R.K. et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. **Science**, v.239, n.4839, p.487-491, jan. 1988.
- SAMBROOK, J.; FRITSH, E.F.; MANIATS, T. **Molecular Cloning: a laboratory manual**. 2nd ed., Cold Spring Harbor Laboratory Press, NY, USA. 1989.
- SCHATZMAYR, H.G. et al. An outbreak of dengue virus at Rio de Janeiro. **Mem. Inst. Oswaldo Cruz**, v.81, p.245-6, 1986.
- SCHMID, K.J.; TAUTZ, D. A screen for fast evolving genes from *Drosophila*. **Proc. Natl. Acad. Sci. USA**, v.94, n.12, p.9746-9750, jun. 1997.
- SEVERSON, D.W. et al. Linkage map organization of expressed sequence tags and sequence tagged sites in the mosquito, *Aedes aegypti*. **Insect Mol. Biol.**, v.11, n.4, p.371-378, aug. 2002.
- SLOSEK, J. *Aedes aegypti* mosquitoes in the Americas: a review of their interactions with the human population. **Soc. Sci. Med.**, v.23, n.3, p.249-257, 1986.
- SMITH, C.E.G. The history of dengue in tropical Asia and its probable relationship to the mosquito *Aedes aegypti*. **Am. J. Trop. Med. Hyg.**, v.5, n.2, p.243-251, mar. 1956.
- SUCEN - SECRETARIA DE ESTADO DA SAÚDE DE SÃO PAULO. Centro de Vigilância Epidemiológica. Manual de vigilância epidemiológica: dengue. São Paulo. 1987.
- SYVANEN, A.C. Accessing genetic variation: genotyping single nucleotide polymorphisms. **Nat. Rev. Genet.**, v.2, n.12, p.930-942, dec. 2001.
- TABACHNICK, W.J. Evolutionary genetics and arthropod-borne disease. The yellow fever mosquito. **Am. Entomologist.**, v.37, n.1, p.14-24, 1991.
- TABACHNICK, W.J.; BLACK, W.C. Population genetics in vector biology. In: MARQUARDT, W.C.; BEATY B.J. (Eds). **The Biology of Disease Vectors**. Niwot: University Press of Colorado, 1996. p.417-437.
- TABACHNICK, W.J.; POWELL, J.R. A world-wide survey of genetic variation in the yellow fever mosquito, *Aedes aegypti*. **Genet. Res.**, v.34, n.3, p.215-229, dec. 1979.

- TABACHNICK, W.J.; POWELL, J.R. Genetic structure of the East African domestic populations of *Aedes aegypti*. **Nature**, v.272, p.535-537, apr. 1978.
- TAILLON-MILLER, P. et al. Overlapping genomic sequences: a treasure trove of single nucleotide polymorphisms. **Genome Res.**, v.8, n.7, p.748-754, jul. 1998.
- TEIXEIRA, M.G. et al. Dengue and dengue hemorrhagic fever epidemics in Brazil: what research is needed based on trends, surveillance, and control experiences? **Cad. Saude Publica**, v.21, n.5, p.1307-1315, sep. 2005.
- TERRA, W.R.; FERREIRA, C. Insect digestive enzymes: properties, compartmentalization and function. **Comp. Biochem. Physiol.**, v.109, n.1, p.1-62, sep. 1994.
- THOMPSON, J.D. et al. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Res.**, v.25, n.24, p.4876-4882, dec. 1997.
- WALL, K.D.; PRITCHARD, J.K. Haplotype blocks and linkage disequilibrium in the human genome. **Nat. Rev. Genet.**, v.4, n.8, p.587-597, aug. 2003.
- WALTON, C. et al. Genetic population structure and introgression in *Anopheles dirus* mosquitoes in South-east Asia. **Mol. Ecol.**, v.10, n.3, p.569-580, mar. 2001.
- WANG, P. et al. A Drosophila temperature-sensitive seizure mutant in phosphoglycerate kinase disrupts ATP generation and alters synaptic function. **J. Neurosci.**, v.12, n.24, p.4518-29, may 2004.
- WEERS, P.M.M. et al. Biosynthesis of locust lipophorin: apolipophorin I and II originate from a common precursor. **J. Biol. Chem.**, v.268, n.6, p.4300-4303, feb. 1993.
- WEIR, B.S. **Genetic Data Analysis II – Methods for Discrete Population Genetic Data**. Sinauer Associates, Sunderland. 1996. 337p.
- WEIR, B.S.; COCKERHAM C.C. Estimating *F*-statistics for the analysis of population structure. **Evolution**, v.38, n.6, p.1358-1370, 1984.
- WILLIAMS, J.C.; HAGEDORN, H.H.; BEYENBACH, K.W. Dynamic changes in flow-rate and composition of urine during the post-bloodmeal diuresis in *Aedes-aegypti* (L). **J. Comparative Physiol.**, v.153, n.2, p.257-265, jun. 1983.
- WILLIAMS, J.G.K. et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic. Acids. Res.**, v.18, n.22, p.6531-6535, nov. 1990.

- WONDJI, C.S.; HEMINGWAY, J.; RANSON, H. Identification and analysis of Single Nucleotide Polymorphisms (SNPs) in the mosquito *Anopheles funestus*, malaria vector. **BMC Genomics**, v.8, p.5-18, jan. 2007.
- WORLD HEALTH ORGANIZATION. **Dengue hemorrhagic fever: diagnosis, treatment, prevention and control**. 2nd ed. Geneva: WHO; 1997. p.84.
- ZHANG, D.X.; HEWITT, G.M. Molecular DNA analyses in genetic studies of populations: practice, problems and prospects. **Mol. Ecol.**, v.12, n.3, p.563-584, jun. 2003.
- ZHENG, L. et al. A detailed genetic map for the X chromosome of the malaria vector, *Anopheles gambiae*. **Science**, v.30, n.261, p.605–608, jul. 1993.
- ZHENG, L. et al. An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. **Genetics**, v.143, n.2, p.941-952, jun. 1996.
- ZHENG, L. et al. Quantitative trait loci for refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B. **Science**, v.276, n.5311, p.425-428, apr. 1997.
- ZÖLLNER, S. et al. Evidence for extensive transmission distortion in the human genome. **Am. J. Hum. Genet.**, v.74, n.1, p.62-72, jan. 2004.

ARTIGO ACEITO PARA PUBLICAÇÃO
MOLECULAR ECOLOGY NOTES - *IN PRESS*

Single Nucleotide Polymorphism Markers in *Aedes aegypti*
Populations in Brazil

Paduan KS and Ribolla PEM

ABSTRACT

DNA sequences were obtained from GenBank, and primer sets were designed using GeneRunner software (Hasting Software, Inc. 1994). A total of ≤ 15 individual mosquitoes chosen at random from 16 cities in different state of Brazil were analysed for the screening of SNPs. Genomic DNA was prepared from single mosquitoes using 300 μ L 5% Chelex (Biorad®) in water, homogenised, and incubated for 30 min at 80°C. Each PCR reaction was performed in a final volume of 25 μ l containing approximately 100 ng genomic DNA, 1 x buffer, 2 mM MgCl₂, 2.5 mM dNTP (GE Healthcare), 20 μ M of each primer and 1.5 U Platinum Taq DNA polymerase (Invitrogen). The PCR program involved 1 cycle of 95°C for 1 min followed by 35 cycles of 95°C for 1 min, 56°C for 1 min, 72°C for 1 min, and a final extension step of 72°C for 10 min.

Nuclear loci were amplified by PCR, checked on a 1% agarose gel, sequenced using a Big Dye^M Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, Foster City, CA, USA) and analysed in an automated ABI Prism 377 DNA sequencer (Applied Biosystems). Consensus sequences were obtained and aligned using the MERGER (<http://bioweb.pasteur.fr/seqanal/alignment/intro-uk.html>) and CLUSTAL W programs (Thompson *et al.* 1994). Sequence comparisons identified one single nucleotide polymorphism (SNP) every 52 coding base pairs, and 36 indels. The loci, primers, the positions of all SNP sites, and novel sequences deposited in GenBank are listed in Table 1.

The SNP genotyping assay was conducted using an Applied Biosystems SNaPshotTM kit. A total of 76 individuals collected from three natural populations in the cities of Botucatu, Juazeiro do Norte and Maceió were characterised with the eight most polymorphic SNPs. PCR products were incubated with ExoSapIT (GE Healthcare) according to the manufacturer's recommendations. Reactions were carried out using

SNaPshot™ protocol (Applied Biosystems). After cycling, the unincorporated fluorescent ddNTPs were removed by shrimp alkaline phosphatase (GE Healthcare), also according to the manufacturer's recommendations. 11 µl formamide was added to each SNaPshot reaction and loaded on an ABI Prism 377 automated DNA sequencer (Applied Biosystems). The relative proportion of each allele was measured by the height and fluorescence of the corresponding peaks. Genepop (Raymond & Rousset 1995) was used to calculate expected and observed heterozygosities (HE and HO), and to test for genotyping linkage disequilibrium (LD) and departures from the Hardy-Weinberg equilibrium (HWE). Significance levels for all tests were corrected for multiple simultaneous tests (Rice 1989).

The intraspecific diversity of *Ae. aegypti* was assessed in coding and non-coding regions of seven nuclear genes in different mosquito populations in Brazil. All genotypes obtained with the SNP assays were identical to those obtained by sample sequencing. Only two non-significant ($\alpha = 0.05$) departures from HWE were detected, in only 2 loci in different populations. Significant LD was observed for SNPs within loci (i.e. between SNPs in a gene), except for the PGK gene, but LD was not detected among SNPs from different sequences in different populations (Table 2). Despite the relatively recent introduction of *Ae. aegypti* into Brazil, we were able to find highly polymorphic markers that will be useful to characterise different mosquito populations, and will assist in future mosquito surveys and epidemiological studies.

REFERENCES

ACKNOWLEDGMENTS

This study was funded by a grant from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP Proc. No. 01/06647-9). KSP and PEMR are supported by

scholarships from CAPES and CNPq, respectively. We thank Maria de Lourdes G. Maçoris (from SUCEN) and Prof. Dr. Newton G. Madeira (from UNESP) for providing mosquito samples.

Table 1. Fifty-five single nucleotide polymorphisms in *Aedes aegypti* native populations.

GenBank Accession	Locus / Chromosome	Oligonucleotide sequences (5'-3')	SNP position: identity	SNP location
EU427573 (387 bp)	NaK / I	F: CCCGAAATGGGTCAAGTTC R: GTCACCGATGACCAGATC	272: A/T	Intron
EU427576 (536 bp)	Tsf / I	F: GTAGTGGCAGTGCTTCTG R: CTGCCCTTGCCTTTCATG	048: G/C 086: G/A 165: G/A 214: T/C 318: A/G 435: C/G 465: A/T 519: T/A	Exon Exon Exon Exon Exon Exon Exon Exon
EU427574 (335 bp)	CYP9J2 / III	F: TGCATCAAAACCCACCTTC R: CGGTAAACATCGACTTTTGC	048: C/T 078: A/G	Exon Exon
EU427570 (332 bp)	apoLp-2 / III	F: GTAAGGTCTATTTCGATGC R: TCAGAACAGCGACGATAG	182: T/A 022: T/C 059: A/C 092: A/G 116: C/G 161: A/G	3' UTR Exon Exon Exon Exon Exon

EU427571 (381 bp)	Ef2 / II	F: GATTGAGTTTCCGGTGTG R: TCGTCATCCTTGGTCTTG	239: C/T	Exon	
			275: C/A	Exon	
			316: G/C	Exon	
EU427575 (460 bp)	PGK / II	F: CAAGGTGAACGAAATGATC R: CGTGTGATACCTGAGATTTC	042: C/T	Exon	
			084: C/T	Exon	
			011: A/C	Exon	
			014: C/T	Exon	
				022: G/C	Exon
				024: T/C	Exon
				030: A/C	Exon
				032: A/T	Exon
				034: T/C	Exon
				055: G/T	Exon
				083: A/G	Exon
				143: A/G	Exon
				161: T/C	Exon
				218: C/T	Exon
				248: C/G	Exon
				284: C/G	Exon
				386: G/A	Exon
				404: T/A	Exon

EU427572 (445 bp)	AeIMUC1 / II	F: GCAAATGTCCCGATATCTTC R: ACGGTTTTCACGGTCCTTAG	410: C/G	Exon
			426: T/C	Exon
			428: G/T	Exon
			431: T/G	Exon
			450: G/A	Exon
			003: C/G	Exon
			142: C/T	Exon
			166: T/A	Exon
			174: C/T	Exon
			198: G/A	Exon
			206: G/C	Exon
			258: C/T	Exon
			279: C/T	Exon
			327: A/T	Exon
			328: T/A	Exon
			336: A/G	Exon
			449: A/C	Exon

F: forward primer, R: reverse primer. Single nucleotide polymorphisms used in analyses are in bold. The most common allele is listed first.

Table 2. Characterisation of 8 single nucleotide polymorphisms using high-throughput genotyping. Collections were taken from Botucatu (state of São Paulo), Juazeiro do Norte (Ceará) and Maceió (Alagoas). N is sample size, R is rare allele frequency, H_E and H_O are expected and observed heterozygosities for nuclear genes.

Locus	Botucatu					Juazeiro do Norte					Maceió				
	N	R	H_O	H_E	P	N	R	H_O	H_E	P	N	R	H_O	H_E	P
NaK	24	0.437	0.625	0.520	0.40	28	0.000	Monomorphic locus			24	0.041	0.083	0.081	1.00
Tsf	24	0.645	0.291	0.467	0.08	28	0.464	0.428	0.525	0.46	24	0.208	0.416	0.336	0.53
CYP9J2	24	0.166	0.250	0.318	0.50	28	0.357	0.214	0.502	0.01	24	0.666	0.250	0.453	0.06
apoLp-II	24	0.312	0.625	0.467	0.06	28	0.642	0.357	0.467	0.23	24	0.250	0.500	0.382	0.27
Ef2	24	0.339	0.375	0.493	0.39	28	0.645	0.535	0.456	0.42	24	0.375	0.666	0.478	0.071
PGK-1	24	0.583	0.333	0.537	0.20	28	0.285	0.500	0.450	0.37	24	0.000	Monomorphic locus		
PGK-2	24	0.041	0.000	0.122	0.02	28	0.303	0.392	0.465	0.67	24	0.895	0.208	0.190	1.00
MUCI	24	0.375	0.500	0.504	1.00	28	0.232	0.196	0.350	0.25	24	0.333	0.312	0.438	0.35

P = significance of tests