INSTITUTO NACIONAL DE PESQUISAS DA AMAZÔNIA
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA, CONSERVAÇÃO E BIOLOGIA EVOLUTIVA – PPG GCBEv

GENÉTICA POPULACIONAL DE *Aedes aegypti* (DIPTERA: CULICIDAE) DE DIFERENTES REGIÕES DO BRASIL, COM MARCADORES MICROSSATÉLITES E MITOCONDRIAL

AHANA MAITRA

Manaus, Amazonas
2019

AHANA MAITRA

GENÉTICA POPULACIONAL DE *Aedes aegypti* (DIPTERA: CULICIDAE) DE DIFERENTES REGIÕES DO BRASIL, COM MARCADORES MICROSSATÉLITES E MITOCONDRIAL

Orientadora: Dra.Vera Margarete Scarpassa (INPA/AM)

Tese apresentada ao Programa de Pós-Graduação em Genética, Conservação e Biologia Evolutiva do Instituto Nacional de Pesquisas da Amazônia como parte dos requisitos para obtenção do título de **Doutor em Genética, Conservação e Biologia Evolutiva**.

Manaus, Amazonas

2019

FICHA CATALOGRÁFICA

M232g Maitra, Ahana
Genética populacional de Aedes aegypti (Diptera: Culicidae) de diferentes regiões do Brasil, com marcadores microssatélites e mitocondrial / Ahana Maitra; orientadora Vera Margarete Scarpassa. -- Manaus:[s.l], 2019.
106 f.

Tese (Doutorado - Programa de Pós Graduação em Genética, Conservação e Biologia Evolutiva) -- Coordenação do Programa de Pós-Graduação, INPA, 2019.

1. Aedes aegypti. 2. Genética populacional . 3. Microssatélites. I. Scarpassa, Vera Margarete , orient. II. Título.

CDD: 597.80415

Sinopse:

Estudos de genética populacional foram realizados com amostras de *Aedes aegypti* de 15 localidades do Brasil. Neste estudo foram empregados dois marcadores moleculares: 12 locos de microssatélites (DNA nuclear) e a região do DNA *Barcode* do gene *COI* do DNA mitocondrial. Os resultados revelaram altos níveis de diferenciação genética e confirmaram a existência de dois grupos genéticos no Brasil. Além disso, três eventos independentes de introdução foram observados para este vetor no Brasil.

Palavras-chave: Dispersão passiva, Diferenciação genética, DNA *Barcode*, Microssatélites, Controle do vetor.

“Dedico esta tese com muito carinho aos meus pais, meu marido e a minha filha, Aadriya”

AGRADECIMENTOS

Inicialmente agradeço à minha família que sempre me incentivou nesse caminho, especialmente aos meus pais, Goutami e Sushanta, que sempre acreditaram em mim, que estiveram sempre ao meu lado e aconselhando-me nos momentos de dificuldades. Além de terem sido um suporte para que eu pudesse finalizar esse ciclo. Agradeço, em especial, à minha irmã Ushoshi que sempre demonstrou orgulho e admiração por mim. Essa tese é dedicada a vocês, como forma de retribuir o amor incondicional e dedicação a mim.

Ao meu marido Anibrata, quero agradecê-lo por me motivar nos momentos insolúveis, pelo amor, apoio e companheirismo. Por estar sempre ao meu lado e mesmo durante as madrugadas esteve ao meu lado, auxiliando-me na redação da tese. O tempo dedicado a mim foi essencial para me manter forte durante esses anos e sempre te amarei por isso.

À minha filha Aadriya pela paciência enorme e apoio total. Eu te amo.

Agradeço ao INPA (Instituto Nacional de Pesquisas da Amazônia), ao PPG-GCBEv (Programa de Pós-Graduação em Genética, Conservação e Biologia Evolutiva), ao Laboratório de Genética de Populações e Evolução de Vetores de Malária e Dengue (LGPEV), e ao Laboratório Temático de Biologia Molecular (LTBM) por fornecerem todo o apoio para realização dessa pesquisa. À Dona Elci, por ajudar a resolver as burocracias na Secretaria do Programa e por ser sempre prestativa.

À Dra. Vera Margarete Scarpassa, minha orientadora, agradeço por ter me aceitado orientar e pela confiança depositada em mim. Agradeço pela orientação recebida e amizade, aprendi muito em todos estes anos. Obrigada por tornar possível a realização desta etapa tão importante na minha vida.

Agradeço ao pessoal do LTBM - a Kyara, Paula e Gisele pelos constantes suportes prestados no Laboratório Temático de Biologia Molecular (LTMB) do INPA. Ao Fabio Costa, André, Gersonval Leandro, José Saraiva e à Dra. Vera Scarpassa pelo auxílio nas coletas das amostras em campo.

À Dra. Vera Margarete Scarpassa por proporcionar a maior parte das amostras (mosquitos e DNA) do banco de amostras de *Aedes aegypti* de armazenados em ultra freezer a -81ºC no Laboratório de Genética de Populações e Evolução de Mosquitos Vetores do INPA.

Aos colegas e amigos de laboratório Ronildo, José, Gersonval, Camila, Leonardo e Manuelina, obrigada por serem pessoas tão especiais. Essa caminhada foi com certeza mais fácil com vocês, muito obrigado pelos bons momentos. Pelo companheirismo e pela grande força que me deram nos momentos difíceis e decisivos do meu doutorado. Vocês me ajudaram muito nessa caminhada e sempre serei grata.

Agradeço aos meus amigos Dayse, Veracilda, Sandra, Galileu e Gilcelia, vocês são minha família em Manaus. Obrigada por me ajudar a superar a saudade e a distância da minha família. Obrigada pela amizade e pelo carinho nesses anos.

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela bolsa concedida e à FAPEAM (Fundação de Amparo à pesquisa do Estado do Amazonas) pelos projetos que financiaram essa tese.

Por fim, gostaria de agradecer a todos que, direta ou indiretamente, contribuíram para a realização e conclusão dessa pesquisa.

RESUMO

Aedes (*Stegomyia*) *aegypti* (Linnaeus, 1762) é o principal vetor da febre amarela urbana, quatro sorotipos do vírus da dengue (DENV1-4), chikungunya (CHKYV) e Zika (ZIKV) e é considerado um dos mais importantes mosquitos vetores. Com exceção da febre amarela, até a presente data, não há vacina disponível contra os CHKYV e ZIKV e, até mesmo a vacina contra a dengue ainda não provou ser eficaz contra os quatro sorotipos; portanto, a ferramenta mais importante para o controle dessas doenças é o combate do principal vetor, *Ae. aegypti*. No Brasil, *Ae. aegypti* está presente em todos os estados e apesar dos programas regulares de controle, sua densidade permanece alta e não tem sido possível impedir surtos de dengue em muitos centros urbanos. Neste estudo, a variabilidade genética e genética populacional foram avaliadas a partir de amostras de *Ae. aegypti* para entender sua estrutura genética e fluxo gênico entre 15 populações brasileiras desse vetor, com o emprego de 12 locos microssatélites e DNA mitocondrial (*COI* – região do DNA *Barcode*). Os marcadores microssatélites revelaram uma estrutura genética significante entre todas as localidades estudadas, evidenciando uma acentuada divergência genética entre Macapá e as demais localidades. A distância genética (valores pareados de F_{ST}) e a análise de variância molecular (AMOVA) foram estatisticamente significantes, independente das distâncias geográficas entre os locais analisados, indicando à presença de um processo dinâmico complexo que influencia os níveis de fluxo gênico dentro e entre diferentes regiões do Brasil. A análise Bayesiana realizada no programa *Structure* recuperou dois grandes grupos genéticos, assim como revelou a presença de uma sub-estruturação genética dentro de cada grupo. A análise de sequências de DNA *Barcode* corroborou a existência de duas linhagens genéticas principais de *Ae. aegypti* circulando no país. Apesar disso, a análise do *Bayesian Analysis of Population Structure* (*BAPS*) recuperou a presença de cinco grupos genéticos no Brasil, que foram quase semelhantes aos agrupamentos recuperados no *Structure* com dados microssatélites, exceto para os agrupamentos de Taubaté e Macapá. Essas diferenças na estrutura genética das populações podem afetar a competência vetorial de *Ae. aegypti* em transmitir os DENV, CHKYV, ZIKV e outras arboviroses e também sua resposta aos programas de controle com métodos genéticos direcionados para suprimir ou modificar geneticamente as populações de vetores com o objetivo de reduzir suas competências em transmitir patógenos.

ABSTRACT

Aedes (*Stegomyia*) *aegypti* (Linnaeus, 1762) is the main vector of urban yellow fever, four serotypes of dengue virus (DENV1-4), chikungunya (CHKYV) and Zika (ZIKV) virus, and is considered as one of the most important mosquito vectors. With the exception of yellow fever to date, there is no vaccine available against the CHKYV and ZIKV, and even the dengue vaccine has not yet proved effective against all the four serotypes of dengue virus; therefore, the most important tool for the control of these diseases is through combatting the main vector, *Ae. aegypti*. In Brazil, *Ae. aegypti* is present in all states and despite regular control programs, its density remains high and it has not been possible to prevent outbreaks of dengue in many urban centers. In this study, genetic variability and population genetics of *Ae. aegypti* were assessed to understand its genetic structure and gene flow among 15 Brazilian populations of this vector with 12 microsatellite loci and mitochondrial DNA (*COI* - DNA *Barcode* region). Microsatellite markers revealed a significant genetic structure among all the studied localities, evidencing a marked genetic divergence between Macapá and the other localities. Genetic distance (pairwise values of F_{ST}) and analysis of molecular variance (AMOVA) were statistically significant, regardless of the geographic distances between the analyzed sites, indicating the presence of a complex dynamic process that influences the levels of gene flow within and between different regions of Brazil. The Bayesian analysis performed in *Structure* retrieved two major genetic groups, as well as the presence of genetic sub-structuring within each group. The analysis of DNA *Barcode* sequences corroborated the existence of two major genetic groups of *Ae. aegypti* circulating in this country. However, the analysis of *Bayesian Analysis of Population Structure* (*BAPS*) revealed the presence of five genetic groups in Brazil, which were almost similar to those retrieved by *Structure* with microsatellite data, except the clustering of Taubaté and Macapá. These differences in genetic population structure may affect the vector competence of *Ae. aegypti* in transmitting DENV, CHKYV, ZIKV and other arboviruses, as well as its response to various control programs with genetic methods, which target at suppressing or genetically modifying the vector populations to reduce their competence in transmitting pathogens.

SUMÁRIO

Lista de Figuras

Introdução Geral

Capítulo I

Capítulo 2

Lista de Tabelas

Capítulo I

Capítulo 2

1. Introdução Geral

1.1. *Aedes aegypti*: origem, histórico e distribuição

Aedes (*Stegomyia*) *aegypti* (Linnaeus, 1762) (‘*aedes*’, do grego, odioso e ‘*aegypti*’, do latim, do Egito) pertence à classe Insecta, ordem Diptera, família Culicidae, subfamília Culicinae, gênero *Aedes* e subgênero *Stegomyia* (Consoli e Oliveira, 1994). De acordo com a literatura, este mosquito é de origem do Velho Mundo, provavelmente da região da Etiópia, África subsaariana e adaptou-se ao ambiente urbano, tornando-se altamente antropofílico. *Aedes aegypti* provavelmente foi introduzida na América, a partir da África, entre os séculos XV ao XIX, devido ao transporte marítimo entre esses continentes, especialmente navios europeus, que realizavam intenso comércio de escravos entre África ocidental e a América, durante os eventos de colonização (Eltis et al., 2010). Na Ásia, acredita-se que a introdução desse mosquito tenha ocorrido devido ao desenvolvimento da indústria e da navegação (Forattini, 2002; Smith, 1956). Atualmente o *Aedes aegypti* está distribuído nas regiões Tropical e Subtropical, entre os paralelos de 45º de latitude Norte e 40º de latitude Sul. Essa distribuição, quase cosmopolita, foi o resultado do aumento nas atividades humana e a facilidade dos meios de transportes por via aérea, marítima e terrestre (Rebêlo et al., 1999), o que provavelmente contribuiu para a dispersão passiva desse vetor, pois a capacidade de dispersão de voo ativo geralmente não excede 800 metros (Forattini, 2002).

Em 1957, Mattingly descreveu duas formas, posteriormente tratadas como subespécies, presentes no Complexo Aegypti, com base em padrões de coloração do corpo do adulto. Uma delas, a forma típica (*Ae. aegypti aegypti*), apresenta comportamento antropofílico e é de coloração pálida ou amarronzada, ou não enegrecida, e com somente escamas claras no tergito do primeiro segmento do abdômen. Esta forma apresenta elevada competência vetorial para o vírus da dengue (Failloux et al., 2002) e da febre amarela no meio urbano (Lourenço-de-Oliveira et al., 2002). Ao contrário, a forma *Ae. aegypti formosus* apresenta uma forte tonalidade enegrecida do tórax e do abdômen e ausência de escamas claras e vive em habitats florestais. Porém, vários estudos mostraram a presença dessas duas formas somente na África (Failloux et al., 2002; Wallis e Tabachnick, 1990) e, com base na literatura atual cita-se apenas o nome específico de *Ae. aegypti* em estudos feitos fora do continente africano (Forattini, 2002). Existe um consenso de que a partir da população silvestre, devido às pressões humanas decorrentes da destruição dos habitats naturais,

uma variedade genética desse mosquito teria sofrido um processo seletivo, adaptando-se às áreas alteradas e, posteriormente, teria encontrado nos aglomerados humanos o ambiente adequado à sua sobrevivência. Além disso, a adaptação aos criadouros artificiais teria sido um grande passo em direção ao comportamento sinantrópico (Christophers, 1960).

1.2. Aspectos Biológicos e Ecológicos do *Aedes aegypti*

Aedes aegypti é um mosquito pequeno com manchas brancas, tendo em média 0,5 cm de comprimento. As manchas brancas são intercaladas e encontradas na região posterior da cabeça, nos segmentos abdominais e nas pernas. No tórax, precisamente na região do escudo, apresenta escamas brancas dispostas em linhas laterais longitudinais, formando um desenho em forma de lira (Forattini, 2002) (Figura 1A). O macho é distinguido da fêmea, pela presença da genitália exposta, por apresentar as antenas mais plumosas e porte menor do que a fêmea (Figura 1B).

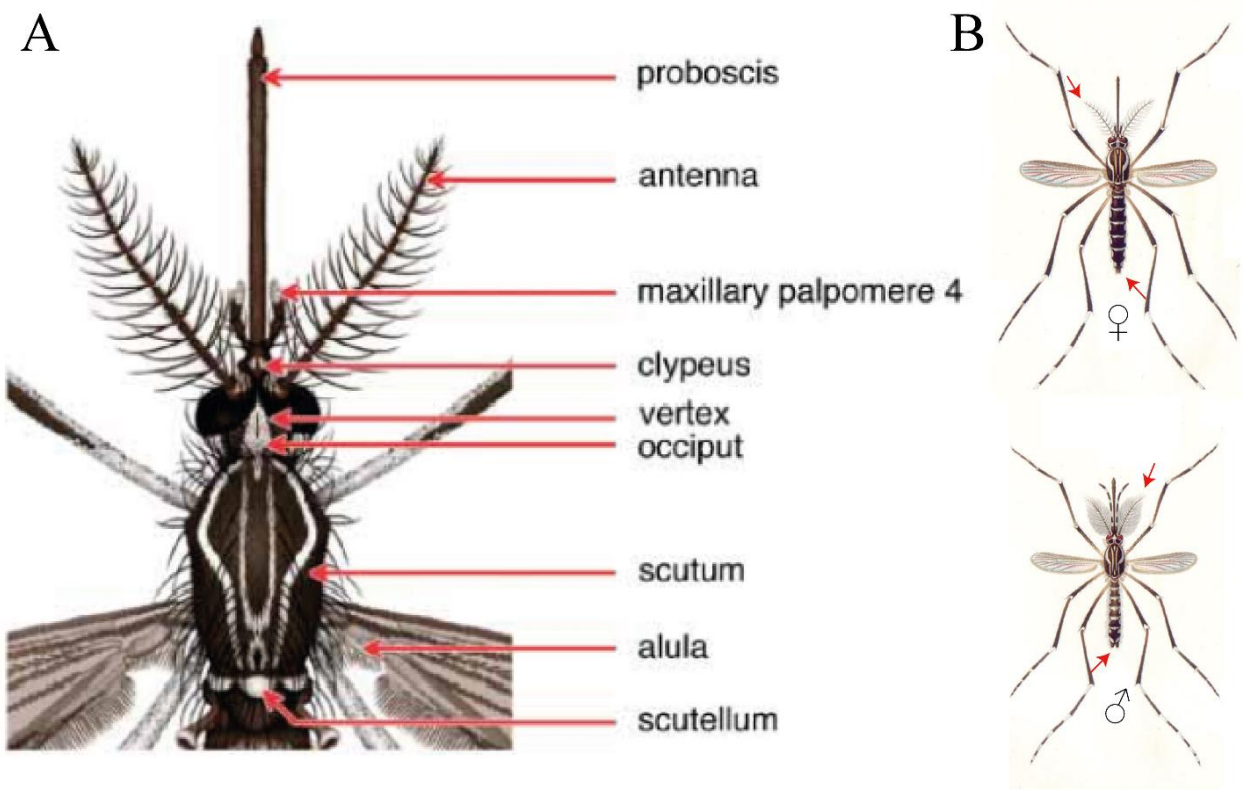

Figura 1. Características morfológicas do *Aedes aegypti* **A**. alguns caracteres de diagnose – tórax e cabeça (Fonte: Rueda, 2004), **B**. diferenças morfológicas e tamanho entre macho e fêmea – antenas e genitálias (Fonte: Goeldi, 1905).

Aedes aegypti apresenta desenvolvimento holometábolo, com duas fases distintas no seu ciclo de vida: uma aquática (ovos, larva e pupa) e a outra terrestre (adulto alado). Após o acasalamento, a fêmea necessita de sangue (fonte proteica) para o desenvolvimento dos ovos. O intervalo entre a alimentação sanguínea e a oviposição varia de dois a três dias. A fêmea grávida é atraída principalmente para recipientes escuros e com superfícies ásperas, sombreados, úmidos ou com água, nas quais depositam os ovos. As fêmeas costumam invadir caixas d'água e cisternas mal vedadas ou piscinas, aquários malcuidados, vasos com água no interior de residências e nos ornamentos de cemitérios. As fêmeas têm capacidade de colocar ovos de um mesmo ciclo gonadotrófico (amadurecimento dos ovos) em vários recipientes, processo conhecido como "saltos de oviposição" o que garante maior probabilidade de sobrevivência e a dispersão de sua prole (Rey e O'Connell, 2014).

Os ovos são depositados normalmente em áreas urbanas, geralmente aderidos às paredes dos recipientes, próximo ao espelho d'água e estes levam entre dois e três dias para eclodirem. Os ovos são menores que 1 mm de comprimento, inicialmente de coloração branca, mas após duas horas tornam-se quase negros devido ao processo de oxidação. Após a eclosão, surgem as larvas, as quais passam por quatro estádios, caracterizadas por grande mobilidade e elevado crescimento, exigindo, para tal, um bom suprimento alimentar e temperatura média entre 24ºC à 29ºC. O estágio de larva pode durar de 4 a 8 dias. O estágio de pupa não requer alimentação e seu desenvolvimento dura em média de 1 a 3 dias. Nesse estágio (no final) já é possível observar as características dos adultos como as asas, a proboscídea e as pernas. Após esta fase, o mosquito adulto emerge e este se alimenta de carboidratos, néctar e sucos vegetais (Consoli e Oliveira, 1994; Forattini, 2002). Assim que o ciclo gonadotrófico é completado, a fêmea está apta para a postura. O tempo médio de vida do mosquito adulto é de 30 dias e a cada oviposição, a fêmea coloca em média de 50 a 200 ovos (revisão em Gadelha e Toda, 1985). O desenho esquemático mostrando as fases de desenvolvimentos do *Ae. aegypti* está ilustrado na figura 2.

Aedes aegypti é um mosquito de hábito diurno com atividade no período vespertino. Os ovos são extremamente resistentes à dessecação, podendo permanecer viáveis por cerca de 18 meses até que a chegada de água propicie a eclosão dos ovos e o desenvolvimento larval (Consoli e Oliveira, 1994). A tendência do *Ae. aegypti* é permanecer onde nasceu, abrigado dentro das habitações. Quando a densidade de mosquitos é elevada, os espécimes se espalham para diversos pontos (dispersa) num raio de voo em torno de 100 metros (Consoli e Oliveira, 1994).

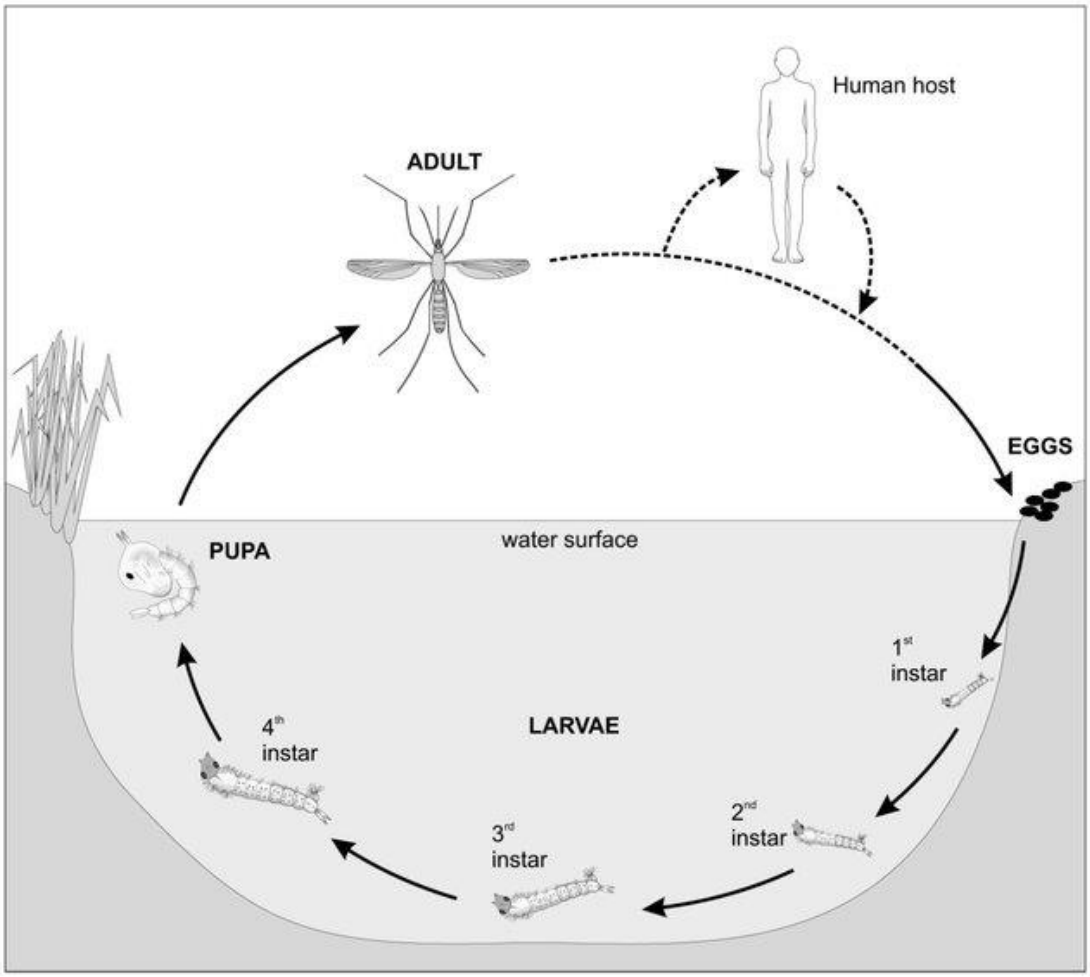

Figura 2 Ciclo de vida de *Aedes aegypti*, mostrando os estágios: ovo, os quatro estádios (*instar*) larvais (L1 a L4), pupa e a forma adulta alada (Fonte: Stanczyk, 2011).

1.3. Aspectos epidemiológicos do *Aedes aegypti*

Atualmente, o *Ae. aegypti* é considerada um dos principais vetores entre os Artrópodes. É o principal vetor dos quatro sorotipos do vírus de dengue (DENV-1, DENV-2, DENV-3 e DENV-4), do vírus da febre amarela (YFV) (ciclo urbano), do vírus da chikungunya (CHKYV) e do vírus da Zika (ZIKV). A capacidade vetorial deste mosquito pode se estender também à transmissão de outros arbovírus e a helmintos filarídeos, os quais podem atingir o homem e outros animais (Forattini, 2002).

Entretanto, até pouco tempo o principal problema epidemiológico estava relacionado à transmissão do DENV, com ênfase especial a forma mais severa da dengue, a febre hemorrágica do dengue (FHD). O vírus da dengue é o arbovírus humano mais frequente com aproximadamente 390 milhões de infecções a cada ano e em mais de 125 países, especialmente em regiões tropicais e subtropicais do mundo (Bhatt et al., 2013). Contudo, na última década, o CHKYV também emergiu como uma das principais causas de preocupação, causando epidemias na Ásia, Ilhas do Oceano Índico, Sul da Europa e nas Américas, incluindo o Brasil (Fernández-Salas et al., 2015; Madariaga et al., 2016).

Em 2007, um surto de Zika foi reportado a partir dos Estados Federados da Micronésia, marcando o primeiro registro do ZIKV em humanos. A partir desta o ZIKV se espalhou para outras ilhas do Pacífico. Desde então, não parou de avançar as barreiras geográficas chegando na Polinésia Francesa, Nova Caledônia, Ilhas Cook, Ilha de Páscoa (Chile) e, posteriormente, no Brasil e Colômbia. Posteriormente, durante um novo surto na Polinésia Francesa em 2013, a infecção pelo ZIKV foi associada a síndrome neurológica Guillain-Barré (Oehler et al., 2014). No Brasil, em 2015, o Ministério da Saúde confirmou a transmissão autóctone do ZIKV no estado de Rio Grande do Norte, na região Nordeste do país (Zanluca et al., 2015), seguido de uma expansão rápida para todas as outras regiões do país. Este foi o primeiro surto registrado nas Américas. Mais recentemente, uma epidemia generalizada do ZIKV ocorreu nas demais partes das Américas do Sul e Central e no Caribe e tem sido associada a anormalidades cerebrais fetais (Rubin et al., 2016).

O controle do *Ae. aegypti* é laborioso e difícil, devido ao fato de ser um mosquito muito versátil na escolha dos criadouros (lugares escolhidos pelas fêmeas para efetuar a postura de seus ovos). Até o momento, nenhuma vacina efetiva está disponível contra essas doenças, com exceção da febre amarela. Portanto, a ferramenta mais eficaz para a prevenção destas doenças é o controle do mosquito vetor (Donalísio e Glasser, 2002; Tauil, 2001).

1.4. Estudos genéticos do *Aedes aegypti* e os marcadores moleculares

O conhecimento da estrutura genética populacional e os padrões de fluxo gênico de *Ae. aegypti* é fundamental para formular e implementar estratégias mais efetivas de controle deste

vetor, consequentemente prevenindo a transmissão de arboviroses por este vetor. Avanços recentes, tanto na tecnologia da genotipagem molecular quanto nos desenvolvimentos teóricos têm gerado enormes avanços para a entomologia médica. O padrão de estrutura genética pode servir como um guia para desenhar estratégias de controle apropriadas e necessárias para interromper a transmissão do patógeno e controlar a doença. Com o advento da biologia molecular durante o século XX foi desenvolvido o conceito de marcador genético que trouxe enormes benefícios e conhecimento acerca da genética molecular. Nos últimos anos, diversos marcadores genéticos foram utilizados para examinar a estrutura genética das populações do *Ae. aegypti* (Ayres et al., 2003; Huber et al., 1999; Scarpassa et al., 2008; Tabachnick e Powell, 1978).

Durante o período de 1970-1980, isoenzimas foram utilizadas para determinar a estrutura genética deste do *Ae. aegypti*, tanto na escala microgeográfica (Tabachnick, 1982; Tabachnick e Powell, 1978) quanto na macrogeográfica (Powell et al., 1980; Tabachnick, 1991). A existência de vários grupos genéticos desse vetor em todo o continente americano foi obtida pela primeira vez com marcadores isoenzimáticos (Tabachnick e Powell, 1979).

Em 1999, Yan et al. estudaram a genética de populações de *Ae. aegypti* de Trinidad utilizando marcadores AFLP e RFLP. RAPD-PCR também foi usado para examinar a estrutura genética dessa espécie em Porto Rico, com amostras de 16 localidades diferentes (Apostol et al., 1996) e na Argentina com amostras de cinco cidades (Sousa et al., 2001). Com base em análises de genética de populações, utilizando RAPD, Wesson e Ocampo (2004) demonstraram que na Colômbia *Ae. aegypti* foi composta por populações geneticamente distintas com diferentes níveis de resistência a inseticidas e competência vetorial para os vírus da dengue. Subsequentemente, outros estudos também foram realizados utilizando vários marcadores moleculares, por exemplo, marcadores enzimáticos (Fraga et al., 2003; Paupy et al., 2004), o polimorfismo de fragmentos de restrição (RFLP) e marcadores do Polimorfismo de DNA Amplificado ao Acaso (RAPD) (Santos et al., 2003). Ainda no Brasil, um estudo foi realizado em cinco Estados, utilizando 47 locos de RAPD, revelando diferenciação genética e elevado polimorfismo nas populações do *Ae. aegypti* (Ayres et al., 2003). Paupy et al. (2004) relataram o uso de AFLP para determinar a estrutura genética do *Ae. aegypti* em Phnom Penh, Camboja. A quantidade de variação e os padrões de fluxo gênico detectados foram comparados com os obtidos para outros dois marcadores, isoenzimas e

microssatélites. Os autores concluíram que os marcadores de AFLP são úteis para estudos de genética de populações do *Ae. aegypti*.

DNA mitocondrial

Nas últimas décadas, um grande número de estudos envolvendo genética de populações de *Ae. aegypti* tem sido realizado em várias partes do mundo e nestes estudos os marcadores genes *COI*, *COII*, respectivamente, Citocromo C Oxidase Subunidade I e Subunidade II, gene *NADH*, subunidade 4 (*ND4*) ou subunidade 5 (*ND5*) do DNA mitocondrial foram empregados. O DNAmt é uma molécula haploide e de herança materna, portanto, não apresenta recombinação gênica. O DNAmt é constituído por uma molécula de DNA circular com aproximadamente 16-20 kb de tamanho, com conteúdo gênico conservado (apenas 37 genes); estrutura gênica simples (não possui DNA repetitivo, transposons, íntros ou pseudogenes), possui genes codificadores de duas subunidades ribossômicas (12S e 16S), 22 genes para RNAs transportadores, subunidades 6 e 8 de ATP *sintase* (ATP6 e ATP8), sete subunidades da NADH desidrogenase (*ND1-ND6* e *ND4L*), uma região rica em A + T (em vertebrados é chamada D-loop) não-codificadora e que parece conter o controle da replicação e transcrição do DNAmt, Citocromo b (*Cytb*) e três subunidades da enzima citocromo c oxidase (*COI*, *COII* e *COIII*) (Flynn e Nedbal, 1998; Moritz et al., 1987); porém, muito importante para o metabolismo celular (Lang et al., 1999). Em invertebrados, o DNAmt é muito parecido com o dos vertebrados, estando as diferenças principalmente relacionadas com o rearranjo dos genes na fita circular (Snustad e Simmons, 2008). O DNAmt é uma molécula de fácil amplificação; possui uma alta taxa de mutação quando comparado com os genes nucleares, o que resulta em altos graus de polimorfismo intraespecífico e divergência, o que é importante em estudos evolutivos (Avise, 1994).

Na Austrália, Beebe *et al*. (2005) utilizaram o gene *COI* em 46 amostras do *Ae. aegypti* do norte e de algumas ilhas próximas e encontraram 8 haplótipos, sugerindo uma discreta variação genética entre as populações analisadas. Mousson et al. (2005) analisaram os níveis de variabilidade dos genes mitocondriais *COI*, *Cytb* e *ND5*, com base em uma hipótese de múltiplas introduções do *Ae. aegypti* na América do Sul. Os autores sugeriram que as populações deste vetor na América do Sul podem ter sido estabelecidas a partir de diferentes linhagens fundadoras, representado por ondas de colonização sucessivas antes e depois dos programas de erradicação em massa. Um estudo de genética de populações conduzido no México utilizando RAPD e o gene

ND4 do DNAmt reportaram 25 haplótipos, onde demonstraram que nesta parte do mundo o *Ae. aegypti* varia geograficamente a distâncias maiores que 250 Km (Gorrochotegui-Escalante et al., 2002). Na América do Sul, Costa-da-Silva et al. (2005) utilizaram o gene *ND4* para estimar a variabilidade genética do *Ae. aegypti* de três cidades do Perú (Lima, Piura e Iquitos). Esses autores encontraram três haplótipos em 55 indivíduos sequenciados, indicando variação genética interpopulacional. Na Venezuela, Herrera et al. (2006) também utilizaram *ND4* do DNAmt para um estudo de genética de populações e detectaram sete haplótipos, o que foram distribuídos em dois clados. Diferenciação significativa foi detectada entre as coletas e estas foram geneticamente isoladas pela distância. Na última década, análises de sequências de *ND4* e *CO1* mostraram a coexistência de duas linhagens genéticas distintas do *Ae. aegypti* no Brasil (Bracco et al., 2007; Lima Júnior e Scarpassa, 2009; Scarpassa et al., 2008). Scarpassa et al. (2008) utilizou o gene *COI* para examinar o fluxo gênico entre 14 localidades do Brasil, identificou dois clados com 81% e 96% de suporte de *bootstrap*, respectivamente. Comparando com sequências depositadas no GenBank, estes clados foram associados com as regiões leste e oeste da África. Similarmente, comparando com as sequências do *GenBank* para o gene *COI* (Delatte et al., 2011), *Ae. aegypti* coletados de várias partes da África, indicou dois clados, um clado designado como" Leste Africano" e o outro como "Oeste Africano".

Gonçalves da Silva et al. (2012) usaram sequências de *ND4* e uma estrutura de coalescência Bayesiana para testar um conjunto de hipóteses sobre o fluxo gênico entre populações de *Ae. aegypti* das Américas. Os autores avaliaram os padrões de fluxo gênico nas escalas continental e sub-regional (Bacia Amazônica) e sugeriram uma conectividade substancial em relação às populações de *Ae. aegypti* nas Américas. Como a dispersão ativa de longa distância não foi observado nesta espécie, esse trabalho apoiou a dispersão passiva como um determinante da estrutura genética das populações do *Ae. aegypti* nas Américas.

Recentemente, um estudo utilizou *ND4* para relacionar os dois clados do *Ae. aegypti*: um da África Ocidental e outro da África Oriental e demonstrou que as populações deste vetor fora da África consistem de misturas de mosquitos de ambos os clados (Moore et al., 2013).

Microssatélites

Ravel et al., (2001) afirmaram que os microssatélites podem ser adequados para identificar várias populações do *Ae. aegypti* em uma escala microgeográfica, assim como são úteis para

avaliar variação temporal (sazonal) dentro de populações de mosquitos. Os microssatélites são sequências de DNA repetidas de 1 a 6 pb (Tautz, 1993). As repetições de nucleotídeo dos microssatélites podem ser mono-, di-, tri-, tetra-, penta- ou hexa-nucleotídeos. Estes marcadores têm sido usados extensivamente em mapeamento genético e estudos forenses e populacionais. Eles têm altas taxas mutacionais, mas são conservados nas regiões franqueadoras. Uma premissa por trás do uso de microssatélites como marcadores genéticos é que seus alelos diferem apenas no número de unidades que eles contêm de uma única repetição (Guyer e Collins, 1993). Eles possuem herança codominante, são altamente polimórficos e densamente distribuídos nos genomas eucariotos e procariotos (Li et al., 2004). Eles são altamente variáveis e polimórficos e, portanto, podem potencialmente distinguir populações que podem ter divergido recentemente. Por isto, são ótimo marcadores para estudos intraespecíficos (Ferreira e Grattapaglia, 1998), diagnósticos de doenças, estudos de mapeamento genético, relações de parentesco, estrutura genética inter- e intra-populacional, genética forense e filogeografia.

Huber et al. (2002) enfatizaram a importância dos marcadores microssatélites para diferenciar populações do *Ae. aegypti* da cidade de Ho Chi Minh. Os autores descobriram diferenciação substancial entre amostras do *Ae. aegypti* das periferias da cidade, ao passo que as populações do centro mostraram menor diferenciação. Estes resultados implicam atividades humanas associadas à urbanização, como fatores que moldam a estrutura genética do *Ae. aegypti*. Marcadores microssatélites também foram utilizados para caracterizar diferentes populações do *Ae. aegypti* em Cote d' Ivoire (Ravel et al., 2002). Três locos dos microssatélites foram utilizados para analisar *Ae. aegypti* de quatro locais distintos e realizar análise genética temporal deste vetor. Huber et al. (2004) analisaram o polimorfismo em *Ae. aegypti* do Vietnã, Camboja e Tailândia utilizando marcadores microssatélites e encontraram baixos níveis de diferenciação genética dentro da população. Costa-Ribeiro et al. (2006) estudaram a variação genética do *Ae. aegypti* na cidade do Rio de Janeiro utilizando marcadores microssatélites e isoenzimas. Naquela cidade, os mosquitos foram coletados de cinco distritos e foram analisados utilizando-se cinco locos de microssatélites e seis sistemas isoenzimáticos para avaliar a quantidade de variação e os padrões de fluxo gênico em níveis locais, sendo a variação genética maior para os microssatélites.

Chambers et al. (2007) caracterizaram e isolaram sequências de microssatélites do *Ae. aegypti*. Este estudo permitiu ampliar o número de marcadores disponíveis para *Ae. aegypti* e, portanto, forneceram ferramentas adicionais para o estudo da variabilidade genética de populações deste mosquito. Slotman et al. (2007) avaliaram a variação genética de 17 marcadores microssatélites identificados no genoma de *Ae. aegypti*. Nove locos com pelo menos cinco alelos foram identificados em amostras coletadas na Tailândia. Endersby et al. (2009) avaliaram a estrutura genética de *Ae. aegypti* da Austrália e do Vietnã e compararam a diferenciação genética entre mosquitos dessas áreas e os de uma população na Tailândia. Seis locos de microssatélites e dois *EPIC* (*exon-primed-intron-crossing*) marcadores foram utilizados para avaliar o isolamento por distância em todas as populações e também dentro das amostras australianas. Hlaing et al. (2010) utilizaram 13 locos de microssatélites para estudar a estrutura genética espacial do *Ae. aegypti* em 36 localidades de Mianmar, Camboja e Tailândia, além de dois locais do Sri Lanka e Nigéria. Os resultados revelaram a presença de heterogeneidade genética nas principais cidades portuárias e similaridade genética entre os locais distantes, mas que estão interligados por estradas principais, sugerindo assim que as rotas de transporte conduzidas pelo homem resultaram na migração passiva por longas distâncias destes mosquitos.

Brown et al. (2011) analisaram 24 localidades ao redor do mundo de *Ae. aegypti* com o emprego de 12 locos microssatélites. As medidas de diversidade genética apoiaram a ideia de que as populações africanas representam a forma ancestral da espécie. Eles identificaram dois grupos genéticos distintos: um incluiu todas as populações domésticas fora da África e o outro incluiu populações domésticas e florestais na África. Outro estudo foi conduzido por Brown et al. (2014) que analisaram a filogeografia e variação genética espaço-temporal das populações de *Ae. aegypti* da Flórida, EUA, com 12 locos microssatélites para examinar a estrutura genética de 10 populações do *Ae. aegypti*. Os resultados mostraram similaridade genética entre as amostras do *Ae. aegypti* de Florida Keys e do sudeste da Flórida. Os mosquitos coletados em Florida Keys após o surto de dengue (2011) foram geneticamente semelhantes aos coletados antes do surto (2009). A homogeneidade genética das populações de *Ae. aegypti* nesta região sugere que uma migração passiva de mosquitos é muito comum no sul da Flórida. No Brasil, Monteiro et al. (2014) realizaram um estudo de âmbito nacional com base em 12 locos microssatélites. Os autores reafirmaram a existência de duas linhagens genéticas do vetor e também concluíram que os padrões genéticos nas populações atuais do *Ae. aegypti* no Brasil são mais consistente com a completa

erradicação da espécie no passado recente, seguida pela recolonização, em vez da possibilidade alternativa de expansão de bolsões residuais de refúgio. Pelo menos duas colonizações provavelmente aconteceram: um dos países do norte da América do Sul (por exemplo, Venezuela) que fundou o grupo do Norte do Brasil, e um do Caribe que fundou o grupo do Sudeste do Brasil. Um estudo mais recente de Kotsakiozi et al. (2017), também com 12 locos microssatélites, confirmaram a presença de dois grandes grupos genéticos do *Ae. aegypti* nas populações do Brasil, um conjunto do Nordeste e outro do Sul e também detectaram subdivisão genética dentro de cada um dos dois principais agrupamentos genéticos.

Conforme descrito acima, nos últimos anos o Brasil tem enfrentado frequentes surtos de dengue, Zika e chikungunya que resultaram em muitas mortes, sequelas neurológicas e nascimentos de bebês com microcefalia. Apesar desta situação, o controle destas doenças não tem sido possível devido à ausência de vacinas para o Chikungunya e Zika e de uma vacina eficaz para os sorotipos da dengue. Este quadro torna-se ainda mais preocupante, pois apesar dos programas regulares de controle do mosquito vetor, sua densidade permanece alta e tem sido difícil controlar este mosquito com base em métodos convencionais. Em decorrência desta situação, vários métodos alternativos de controle do vetor têm sido propostos até o momento, incluindo o controle genético.

O monitoramento eficiente das populações deste vetor pode auxiliar na prevenção da propagação das arboviroses transmitidas por este eficiente vetor. A caracterização genética é especialmente importante para se conhecer a dinâmica populacional e padrões de fluxo gênico que podem influenciar na capacidade do *Ae. aegypti* em transmitir as arboviroses. A Competência vetorial (CV) é caracterizada pela permissividade intrínseca de um vetor para infecção, replicação, e a transmissão de um vírus (Hardy 1988). A competência vetorial dos mosquitos é controlada geneticamente e varia conforme a região geográfica. A competência vetorial do *Ae. aegypti* para DENV pode variar de acordo com a região geográfica, assim como entre populações de diferentes países ou mesmo de diferentes estados do mesmo país. Variações na competência vetorial do *Ae. aegypti* podem ocorrer na mesma cidade ou em ilhas. Conhecendo a composição genética dos mosquitos recém introduzidos e suas procedências podem nos dar indicações do potencial da competência vetorial e também auxiliar no planejamento mais efetivo de controle. Portanto, estudos de diversidade

genética e estrutura populacional, incluindo padrões de fluxo gênico, em nível nacional, são fundamentais para conhecer e entender a dinâmica populacional do *Ae. aegypti* de uma determinada área ou região, o que é de suma importância para o desenho de uma vigilância entomológica mais efetiva e o desenvolvimento de novas medidas de controle desse vetor, especialmente o controle genético.

No presente estudo, analisamos a diversidade genética intra e interpopulacional, a estrutura populacional e os padrões de fluxo gênico de 15 populações de *Ae. aegypti* de diferentes regiões do Brasil, utilizando 12 loci de microssatélites e marcador mitocondrial (*CO1*-DNA *Barcode*) para avaliar os grupos genéticos existentes deste vetor circulando pelo país.

2. Objetivos

2.1. Objetivo Geral:

Analisar a variabilidade genética de amostras do *Ae. aegypti* coletadas em todas as regiões do Brasil, revelando questões como a estrutura genética, os padrões de dispersão e os possíveis eventos de colonização desse mosquito na Amazônia e em outras regiões no Brasil.

2.2. Objetivos Específicos:

1. Estimar o grau de variabilidade genética intra e interpopulacional de populações do *Ae. aegypti*;
2. Estimar o grau de divergência genética e os padrões de fluxo gênico das amostras do *Ae. aegypti*;
3. Propor hipóteses sobre as origens das populações do *Ae. aegypti* no Brasil e verificar os possíveis eventos de colonização desse mosquito;
4. Verificar a existência de linhagens ou grupos geneticamente estruturados para esta espécie.

3. Material e Métodos

A metodologia deste estudo encontra-se descrita em detalhes nos capítulos 1 e 2.

4. Resultados e Discussão

Os resultados e a discussão dos dados obtidos neste estudo estão apresentados em dois capítulos, nos formatos de artigos científicos intitulados abaixo:

Capítulo 1: Exploring deeper genetic structures: *Aedes aegypti* in Brazil

- Estimar o grau de diversidade genética intra e interpopulacional de populações do *Ae. aegypti* do Brasil;
- Estimar a divergência genética e os padrões de fluxo gênico das populações do *Ae. aegypti* do Brasil;

Capítulo 2: Capítulo 2: Genetic diversity and gene flow patterns of *Aedes aegypti* (Diptera: Culicidae) in Brazil

- Propor hipóteses sobre as origens das populações do Ae. aegypti do Brasil e verificar os possíveis eventos de colonização desse mosquito no país;
- Verificar a existência de linhagens ou grupos geneticamente estruturados para esta espécie.

4.1. Capítulo 1: Exploring deeper genetic structures: *Aedes aegypti* in Brazil

Artigo submetido para a revista *Acta Tropica*.

Ahana Maitra[a], Antônio Saulo Cunha-Machado[a], André de Souza Leandro[b], Fábio Medeiros da Costa[c], Vera Margarete Scarpassa[a,d,*]

[a] Programa de Pós-Graduação em Genética, Conservação e Biologia Evolutiva, Instituto Nacional Pesquisas da Amazônia, Manaus, CEP 69.067-375, Amazonas, Brazil.

[b] Centro de Zoonoses, Secretaria Municipal de Saúde e Saneamento, Prefeitura Municipal de Foz do Iguaçu, Paraná, Brazil.

[c] Oikos Consultoria e Projetos, Departamento de Meio Ambiente, Estrada de Santo Antônio, 3903 Apto 103 – Triângulo, Porto Velho, CEP 76.805 - 696, Rondônia, Brazil.

[d] Laboratório de Genética de Populações e Evolução de Vetores de Malária e Dengue, Instituto Nacional de Pesquisas da Amazônia, Manaus, CEP 69.067-375, Amazonas, Brazil.

* Corresponding author: Dr. Vera Margarete Scarpassa, Laboratório de Genética de Populações e Evolução de Vetores de Malária e Dengue, Coordenação de Biodiversidade. Instituto Nacional de Pesquisas da Amazônia, Manaus, CEP 69.067-375, Amazonas, Brazil. Tel.: +55 92 3643 3046. *E-mail address:* vera@inpa.gov.br

Abstract

Aedes aegypti, being the principal vector of dengue (DENV1 to 4), chikungunya and Zika viruses, is considered as one of the most important mosquito vectors. In Brazil, despite regular vector control programs, *Ae. aegypti* still persists with high urban density in all the states; consequently, the country has experienced successive outbreaks of these diseases. The aim of this study was to estimate the intra and inter population genetic diversity and genetic structure among 15 Brazilian populations of *Ae. aegypti* based on 12 microsatellite loci. A total of 510 specimens and 6,210 genotypes were analyzed: eight locations from northern (Itacoatiara, Manaus, Novo Airão, Boa Vista, Rio Branco, Porto Velho, Guajará-Mirim and Macapá), three from southeastern (Araçatuba, São José de Rio Preto and Taubaté), one from southern (Foz do Iguaçú), one from central west (Cuiabá) and two from northeastern (Campina Grande and Teresina) regions of Brazil. Genetic distance (pairwise values of F_{ST} and Nm) and the analysis of molecular variance (AMOVA) were statistically significant, independent of geographic distances among the sites analyzed, indicating that them are under a complex dynamic process that influence the levels of gene flow within and among different regions and localities of the country. Supporting these findings, only 31% of the overall samples were found to support the isolation by distance (IBD) model. Bayesian analysis in STRUCTURE revealed the existence of two major genetic groups, as well as there was genetic substructure within them, and these results were confirmed by the BAPS and DAPC analyses. These differences probably are the cumulative results of multiple introductions, fluctuations of population sizes due passive dispersal, climatic conditions, cycles of extinction and re-colonization due to the extensive use of insecticides in vector control activities, followed by founder effects, bottleneck events and genetic drift, throughout the country. These genetic differences may affect its vector competence to transmit dengue, chikungunya, Zika and other arboviruses, and also in their response to control programs; and may guide towards more effective integrated vector control strategies in Brazil.

Key words: Dengue vector, Population genetics, Brazil, Genetic differentiation, Gene flow, microsatellites.

4.1.1. Introduction

Aedes aegypti is the most important vector of human arboviruses, which transmits dengue (DENV), urban yellow fever (YFV), chikungunya (CHKYV) and Zika (ZIKV) viruses; consequently, it has been the most studied mosquito in recent times (Brown et al., 2014; Gloria-soria et al., 2016). Dengue virus is the most common human arbovirus, which causes approximately 390 million infections every year in more than 125 countries, especially in tropical and subtropical regions (Bhatt et al., 2013). Over the last decade, CHKYV has also emerged as a major cause of concern, causing epidemics in Asia, Indian Ocean islands, southern Europe and the Americas, including Brazil (Fernández-Salas et al., 2015; Madariaga et al., 2016). More recently, a widespread epidemic of ZIKV has occurred across South and Central Americas and the Caribbean, and has been linked to fetal brain abnormalities (Rubin et al., 2016). In 2015, Brazil experienced a massive ZIKV outbreak, which led to an epidemic resulting in several thousand cases of neonatal microcephaly (Kindhauser et al., 2016). Till date, 198,810 confirmed cases of ZIKV have been reported from Latin America and the Caribbean (PAHO and WHO, 2017). Brazil is, by far one the most affected country, not only by Zika, but also by dengue and chikungunya, as well as there have been instances of yellow fever outbreak between 2016 – 2017 (PAHO and WHO, 2017). In 2017, for example, 251,711 dengue and 185,854 chikungunya cases were recorded in Brazil. In addition to this, since 2015, the country has experienced massive ZIKV outbreaks (233,801 cases) (Epidemiol, 2018).

To date, there is no vaccine against CHKYV or ZIKV, and even the vaccine against dengue has not yet been proven to be effective against all four serotypes of DENV (WHO, 2016); therefore, the most important tool for the control of these diseases is to combat the main vector, *Ae. aegypti* (Gubler, 2002, 1998).

In Brazil, *Ae. aegypti* is present in all states (Neis Ribeiro et al., 2013) and despite regular control programs, its density remains high and it has not been possible to prevent dengue outbreaks in many urban centers (Scarpassa et al., 2008). Because of the diurnal feeding habit, extremely anthropophilic behavior, ability to lay eggs in multiple containers showing an incredible adaptive plasticity, desiccation-resistant eggs, fast life cycle, high passive dispersal and rapidly growing insecticide resistance of *Ae. aegypti* (Bass and Field, 2011; Linss et al., 2009), it has been very difficult to control this mosquito based on conventional methods, especially in Latin America

(Maciel-de-freitas et al., 2012). A lot of alternative vector control methods has been proposed so far, including genetic control, such as the use of transgenic mosquitoes, Release of Insects carrying Dominant Lethal gene (RIDL) (Laith and Walker, 2016) and *Wolbachia* infected *Ae. aegypti* lineages (Moreira et al., 2009).

Vector competence, as a component of vectoral capacity, is governed by intrinsic (genetic) factors that influence the ability of a vector to transmit a pathogen (Hardy et al., 1983; Woodring et al., 1996). The vector competence of *Ae. aegypti* for DENV has been seen to vary according to geographical region as well as between populations from different countries or even from different states of the same country (Bennett et al., 2002). Variations have also been observed in the vector competence of *Ae. aegypti* from the same city (Gonçalves et al., 2014) or from a small island (Vazeille-falcoz and Mousson, 1999). Differential transmission of Asian and African lineages of ZIKV by *Ae. aegypti* has been observed in South – Pacific islands (Calvez et al., 2018). Therefore, studies of genetic diversity and population structure, including the estimates of gene flow, are important to understand the population dynamics of *Ae. aegypti* in a given region, which is paramount for designing better effective entomological surveillance and control measures of this vector, especially the genetic control.

During 1970-1980, isozymes markers were used to determine the genetic structure of this vector on micro geographic (Tabachnick, 1982; Tabachnick and Powell, 1978) and macro geographic (Powell et al., 1980; Tabachnick, 1991) scales. The existence of several genetic groups of *Ae. aegypti* throughout the American continent was first suggested by a study with isozymes markers (Tabachnick and Powell, 1979). In Brazil, the earlier studies of genetic structure of *Ae. aegypti* populations involved isozymes (Dinardo-Miranda and Contel, 1996; Fraga et al., 2003) and RAPD (Ayres et al., 2003; Paduan et al., 2006) markers which demonstrated high polymorphism and genetic differentiation among different populations of this vector. In the last decade, analyses of mitochondrial DNA sequences (*ND4* and *CO1* genes) revealed the co-existence of two distinct genetic clusters of *Ae. aegypti* in Brazil (Bracco et al., 2007; Lima and Scarpassa, 2009; Scarpassa et al., 2008). Linss et al. (2014) also confirmed the existence of two distinct genetic groups of *Ae. aegypti* in Brazil, using *kdr* (knock-down-resistance) mutant alleles that confer pyrethroid resistance.

Monteiro et al. (2014) conducted a nationwide study based on microsatellites markers, which reaffirmed the existence of two distinct genetic groups of this vector in Brazil. These authors

also proposed that the *Ae. aegypti* populations from eastern, central and southern Brazil are genetically closer to the populations from Caribbean islands, whereas the populations from northern Brazil are genetically closer to the populations of Venezuela and other North American countries. A recent study with microsatellites (Kotsakiozi et al., 2017), revealed the time frame of re-invasion and re-colonization of *Ae. aegypti* in Brazil; the authors suggested that the populations of northern Brazil may have been introduced from neighboring countries, where the complete eradication of this vector had never been possible, whereas the *Ae. aegypti* populations of southern Brazil may have been recently migrated from the northern areas of Brazil itself.

Microsatellites or simple sequence repeats (SSR) markers are very useful for genetic diversity studies because of their abundance, high polymorphism, codominance and easy detection. Due to their biparental inheritance, they are efficient markers to reveal patterns of male- and female-mediated gene flow and understanding genetic relationships between populations of a species. Therefore, they provide useful information on the dispersion pattern of the vectors, which is critical for vector control efforts in the present time. These markers, due to their high mutation rates, can also detect differentiation even in weakly structured species. In the present study, we analyzed the intrapopulation genetic diversity, population structure and the gene flow patterns among 15 populations of *Ae. aegypti* from different regions of Brazil using 12 microsatellites loci in order to re-assess the existence of two previously established distinct genetic clusters of the vector. This knowledge can help in devising strategies to control the spread of the vector as well as of the diseases it transmits.

4.1.2. Material and Methods

Sample Collection

Aedes aegypti samples were collected from 15 sites of Brazil, covering all regions of country, as follows: eight from the states of Brazilian Amazon [Itacoatiara (IT), Manaus (MA), and Novo Airão (NA), state of Amazonas; Boa Vista (BV), state of Roraima; Rio Branco (RB), state of Acre; Porto Velho (PV) and Guajará-Mirim (GM), state of Rondonia and Macapá (MP), state of Amapá], one from the central west region of the country [Cuiabá (CU), state of Mato Grosso], two from the northeastern region [Campina Grande (CG), state of Paraiba; Teresina (TS), state of Piauí], three from the southeastern region [Araçatuba (AR), São José do Rio Preto (RP)

and Taubaté (TA), state of São Paulo] and one from the southern region [Foz do Iguaçú (IG), state of Paraná] (Figure 1). The details regarding specimen collection including the state, geographical coordinates, year of collection and sample size for each site are shown in Table 1.

Figure 1. Collection sites of *Aedes aegypti*. AR: Araçatuba, RP: São José do Rio Preto, TA: Taubaté, RB: Rio Branco, PV: Porto Velho, GM: Guajará-Mirim, IG: Foz do Iguaçú, CU: Cuiabá, CG: Campina Grande, TS: Teresina, BV: Boa Vista, IT: Itacoatiara, MA: Manaus, RB: Rio Branco, MP: Macapá, NA: Novo Airão.

Table 1: Localities sampled for *Ae. aegypti* from different states of Brazil

State	Locality	Abbreviation	Coordinates (Lat./Long.)	Sample Size	Year of collection
São Paulo	Araçatuba	AR	21°13'42.35"S 50°27'5.15"W	32	2009
São Paulo	São José de Rio Preto	RP	20°52'20.56"S 49°22'25.60"W	31	2005
São Paulo	Taubaté	TA	23° 2'26.05"S 45°33'23.08"W	32	2005
Paraná	Foz de Iguaçu	IG	25°32'24.34"S 54°35'1.29"W	36	2017
Mato Grosso	Cuiabá	CU	15°40'40.56"S 56° 5'42.24"W	36	2007
Paraíba	Campina Grande	CG	7°15'5.88"S 35°52'42.45"W	36	2017
Piauí	Teresina	TS	5° 7'14.19"S 42°48'18.58"W	35	2016
Roraima	Boa Vista	BV	2°46'56.04"N 60°41'2.67"W	36	2005
Amazonas	Itacoatiara	IT	3° 8'42.04"S 58°26'18.20"W	32	2011
Amazonas	Manaus	MA	3°11'34.36"S 60° 1'23.25"W	32	2011
Acre	Rio Branco	RB	10° 1'25.00"S 67°49'19.63"W	35	2006
Amapá	Macapá	MP	0° 1'18.79"N 51° 4'11.36"W	36	2017
Rondônia	Porto Velho	PV	8°45'39.68"S 63°54'1.59"W	34	2018
Rondônia	Guajará-mirim	GM	10°47'21.27"S 65°19'48.09"W	35	2018
Amazonas	Novo Airão	NA	2°38'12.57"S 60°56'38.95"W	32	2018

The mosquitoes were collected from fields as larva or adults. Larvae were collected from various artificial recipients near human dwellings and also by using oviposition traps (ovitraps) for 2-7 days, depending on the location. All the samples were collected from multiple breeding sites (15 to 20 breeding sites) per location to prevent sampling of related individuals. The breeding sites were widespread over different neighborhoods in each city, the geographic distances between breeding sites ranging from 20 meters to 5 km (Scarpassa et al., 2008). To further avoid the chances of sampling related individuals, only 2-3 specimens from each breeding site were used in the analyses. The specimens were morphologically identified using the taxonomic key of Forattini (2002), preserved in 95% ethanol and stored in freezer –20° C or stored dry in ultra-freezer -80° C, until DNA extraction.

DNA extraction, PCR and microsatellite genotyping

Genomic DNA was extracted individually from the larva or adult mosquitoes using phenol and chloroform method (Sambrook and Russell, 2001), with minor modifications. Twelve previously studied microsatellite loci (Brown et al., 2011; Chambers et al., 2007; Slotman et al., 2007) (Table S1) were genotyped for individual samples, as described by Brown et al. (2011). The *AC1, AC2, AC5, AG2, AG3, AG4, AG5, AG7, B07* and *CT2* loci consist of dinucleotide repeats, whereas *A1* and *B3* loci consist of tri-nucleotide repeats. For the PCR reactions, a total volume of 10 µL of reaction solution was prepared, containing 1 µL of 10-20 ng DNA template, 1 µL 10x buffer, 0.3 µL 50 mM $MgCl_2$, 2.1 µL 1 mM dNTPs, 0.4 µL 4 mM M13-tailed forward primer (Schuelke, 2000), 0.4 µL 4 mM fluorescent M13-labelled primer (FAM, HEX and TAMRA), 0.8 µL 4 mM reverse primer, 0.2 µL of 5U/µL Platinum Taq DNA polymerase (Invitrogen Inc., Carlsbad, CA, USA) and 3.8 µL sterile water to complete the final volume. The microsatellite loci were amplified in 96 well thermocycler Veriti^{TM} Thermal Cycler (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) according to thermocycle conditions described by Slotman et al. (2007). The PCR products were analyzed in an automated ABI 3130 xl Genetic Analyzer (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA), available at Instituto Nacional de Pesquisas da Amazônia, in Manaus, Brazil. The allele sizes were scored using GeneScan 500 ROX dye (Applied Biosystems) and genotyped in software GENEMAPPER version 4.0 (Applied Biosystems).

Microsatellite analyses

The dataset generated was initially checked with MICROCHECKER v. 2.23 (Oosterhout et al., 2004) for potential genotyping errors, as stuttering and large allele dropout, and also for the presence of null alleles. The same software was used to calculate the probability of occurrence and frequency of null alleles for each locus and each population. The measures of intrapopulation genetic diversity, such as the number of alleles per locus (*N*a), number of effective alleles per locus (*Ne*), private alleles and their frequencies were estimated using the GENALEX v. 6.41 (Peakall and Smouse, 2012). Allelic richness (*A*r) and inbreeding coefficient (*F*is) were calculated in FSTAT v. 2.9.3 (Goudet, 1995). The measures of observed (Ho) and expected (He) heterozygosities, linkage disequilibrium (*LD*) and the probability tests for Hardy-Weinberg Equilibrium (HWE) were estimated in ARLEQUIN v. 3.5 (Excoffier and Lischer, 2010).

The genetic structure, based on pairwise F_{ST} and *N*m values, and Analyses of Molecular Variance (AMOVA), were estimated in ARLEQUIN, v.3.5 (Excoffier and Lischer, 2010), with 10,000 permutations. AMOVA was analyzed on three different hierarchical levels of population structure, to partition the total molecular variance: (1) all samples (non-grouped) were analyzed as a unique group to test the overall genetic differences among samples; (2) two populations groups (Group 1: AR, RP, CU, CG, TA, IT, MA, NA, BV, RB, PV, GM and Group 2: IG, TS, MP) and (3) four population groups (Group 1: AR, RP, CU, CG ; Group 2: TA, IT, MA, NA, BV ; Group 3: RB, PV, GM and Group 4: IG, TS, MP) were tested to check the level of substructure within or between them. In this analysis, the two populations groups and four populations groups were obtained from Structure analysis. The sequential Bonferroni correction was applied for all cases of multiple comparisons (Holm, 1979).

The population structure was evaluated using the Bayesian clustering method calculated in STRUCTURE v. 2.3 (Pritchard et al., 2000). This method identifies clusters of genetically similar individuals from multilocus genotypes, without any prior knowledge about the sampling location. This model assumes K genetic clusters, each having a characteristic set of allele frequencies at each locus. The analysis was performed for ten independent runs for each K (from K=1 to K=15, the maximum number of populations used for this study). Consistent results were obtained across runs using a burn-in period of 100,000 permutations, followed by 1,000,000 Markov Chain Monte Carlo (MCMC) repeats. The optimal value of K was estimated by Evanno et al. (2005) method using online version of Structure Harvester v. 0.6.94 (Earl and vonHoldt, 2012). The program

CLUMPP v1.1.2 (Jakobsson and Rosenberg, 2007) was used to summarize the results from the 10 independent STRUCTURE runs and the results were plotted using DISTRUCT v.1.1 (Rosenberg, 2003). Depending on the number of clusters (value of K) obtained, we further partitioned our data set for a better understanding of the underlying genetic structure of each population.

A second analysis for test these clusters was accessed by Bayesian Analysis of Population Structure (BAPS) (Corander et al., 2008; Corander and Tang, 2007) to estimate the number of genetic groups along the area of study. In this analysis, 1 to 15 clusters were employed (the upper corresponding to the total number of sampled localities), and five independent runs were implemented. The most probable genetic cluster configuration was prepared by comparing the log-likelihood values of the best models.

In order to further explore the genetic structure, we also conducted Discriminant Analysis of Principal Components (DAPC) which submits genetic data to a principal component analysis (PCA) before conducting discriminant analysis (DA) on those principal components (Jombart et al., 2010). In contrast to Bayesian clustering methods, DAPC neither rely on a population genetic model (Hardy-Weinberg or gametic equilibrium, as does STRUCTURE) (Pritchard et al., 2000), nor this multivariate discriminant method is too computationally intense, and it is better at handling hierarchical structure or clinal variation within the populations (Jombart et al., 2010). The discriminants functions are based on linear combinations of alleles harboring the greatest variation between the clusters while minimizing variations between them (Jombart et al., 2010). This method differs from traditional PCA analysis as it minimizes within-group variability. We implemented DAPC in the ADEGENET package (Jombart, 2008) on R version 3.4.1 (R Core Team, 2017). The multivariate analysis defines the groups low levels of support for different numbers of potentially distinct genetic clusters in the absence of a priori population designation (Jombart et al., 2010). The optimal number of Principal Components (PCs) retained and used in analyses was optimized using a cross-validation method with the function *xvalDapc,* using varying numbers of PCs (and keeping the number of discriminant functions fixed, proposed by the R's package ADEGENET). The number of PCs retained can have a substantial impact on the results of the analysis. Indeed, retaining too many components with respect to the number of individuals can lead to over-fitting and instability (Jombart and Ahmed, 2011). To explore how the genetic variation was partitioned among sampling sites, DAPC was executed by the function *dapc* using the sampling sites as a prior. Then Bayesian Inference Criterion (BIC) was used to determine

optimal number of population clusters (K) using k values from 1 to 20 with functions *find. clusters* and *k-means,* to provide comparison with STRUCTURE and BAPS. The optimal number of clusters, K was chosen from the lowest value of BIC and subsequent scatterplots and bar plots were prepared. For a detailed cluster-assignment of specimens, we used the function *table(pop(x), grp$grp)* which display a table of the cluster vs sampling site, duly plotted on a comparative graph (function *table.value*).

To assess the significance of correlation between geographic and genetic (F_{ST}) distance matrices among the 15 sampling sites, based on the results from the genetic structure analyses, Isolation by distance (IBD) was estimated using the Mantel test (Mantel, 1967) between a matrix of genetic and geographic distances by the regression of $F_{ST}/(1 - F_{ST})$ on the natural logarithm (ln) of straight-line geographical distance with the ADEGENET package of R, with the *mantel.randtest* function (999 permutations). As the correlation between genetic and geographic distances can occur under different biological scenarios, like continuous clines or distant patches, we visualized local densities of scatterplot of genetic and geographic distances using a two-dimensional kernel density estimation (function *kde2d*) in the MASS R package (Venables and Ripley, 2002).

Effective population size (*N*e) for each sample, based on the linkage disequilibrium (*LD*) model, was calculated in NeEstimator v. 1.3 (Peel et al., 2004). The dataset was also used to estimate demographic process, such as recent population bottleneck and/or population expansion, and the heterozygosity tests were used to analyze deviations from Mutation-Drift Equilibrium (MDE) for each sample across all loci. At selectively neutral loci, the expected heterozygosity calculated from allele frequencies data (He) assuming HWE, and from the number of alleles and sample sizes (Heq), assuming a population at MDE, are expected not to be significantly different. Thus, if a significant number of loci show He>Heq, this indicates that the population recently experienced a bottleneck. Conversely, a significant number of loci showing He<Heq may suggest population expansion. Therefore, the estimates of expected heterozygosity were calculated for each sample across all loci using three mutation models: infinite alleles model (IAM), stepwise mutation model (SMM) and the two-phase model (TPM). In this study, we used TPM with 95% single-step mutations and 5% multiple-step mutations as recommended for microsatellites (Piry et al., 1999). The statistical significance of deviation from MDE were assessed using Wilcoxon's s signed rank test, which has been recommended by Piry et al. (1999) for analyses with less than 20

(< 20) polymorphic microsatellite loci. While testing for bottlenecks, the null hypothesis assumed for the Wilcoxon's test is "no significant heterozygosity excess (on average across loci) is observed". Thus, the alternate hypothesis is the presence of significant heterozygosity excess (evidence of a recent bottleneck). This is a one-tailed test that requires at least four polymorphic loci to have significant ($P < 0.05$) test result (Piry et al., 1999). These analyses were performed using BOTTLENECK v. 1.2.02 (Piry et al., 1999).

4.1.3. Results

A total of 510 specimens of *Ae. aegypti* encompassing 15 Brazilian cities (Table 1 and Figure 1) were genotyped for 12 microsatellite loci, totaling 6,210 genotypes. All loci were found to be polymorphic (Table 2), with a total of 238 alleles, varying from 3 (locus *AC5* from Itacoatiara and Taubaté) to 14 (locus *AG2* from Boa Vista). Considering all the 15 samples, the highest number of alleles were observed for locus *AG2* (28) and the lowest number of alleles for locus *A1* (13). The highest mean allelic richness was observed in the samples from Cuiabá (CU) (8.690). Locus *AC1* was found to be in HW disequilibrium in most of the populations (12 out of 15), and also showed highest probability of null alleles in 6 out of 15 samples, followed by *B07* locus with probability of null alleles in 5 out of 15 samples. Out of 180 tests, 80 (44.44%) tests showed significant deviations for HWE, after the Bonferroni correction, most of them (56) suggesting heterozygote deficiency. The samples of Teresina (TS) and Macapá (MP) exhibited highest number of loci deviating from HWE, 11 and 8, respectively, likely indicating heterozygote deficiencies. The mean observed heterozygosity ranged from 0.526 [Teresina (TS)] to 0.800 [São José de Rio Preto (RP)], whereas mean expected heterozygosity ranged from 0.693 [Porto Velho (PV)] to 0.795 [Foz do Iguaçú (IG)]. The highest mean inbreeding coefficient (F_{IS}=0.308) was observed in the samples of Teresina (TS), followed by Foz do Iguaçú (IG) (F_{IS}= 0.150). The remaining samples showed slightly lower mean F_{IS} (0.004-0.085). Populations of Araçatuba (AR), São José de Rio Preto (RP), Itacoatiara (IT), Novo Airão (NA), Boa Vista (BV), Porto Velho (PV) and Guajará-Mirim (GM) showed mean negative values of F_{IS} (-0.001 to -0.095), indicating excess of heterozygotes.

Linkage disequilibrium (*LD*) analysis was carried out to confirm whether the deviations from HWE within samples were due to Wahlund effect, inbreeding, selection or the presence of

null alleles. A total of 149 out of 990 (15.05 %) locus-by-locus tests for *LD* were found to be significant ($P < 0.0005$), after the Bonferroni correction. The samples of Macapá (MP) and Campina Grande (CG) had the greatest number of significant pair-loci (both with 18), followed by Boa Vista (BV), with 15 (Table S2). No loci pair was consistently significant for *LD* across all 15 populations, indicating that close physical genetic linkage is unlikely. Therefore, the most of the deviations from HWE observed within samples may be attributed to presence of null alleles.

The number of private alleles by locus and by sample is shown in Table 3. A total of 51 private alleles were observed in 15 populations. The sample of Macapá (MP) had highest number of private alleles (13), but at low frequencies, followed by Teresina (TS) and Foz do Iguaçú (IG) (both with 7), whereas Itacoatiara did not show any private allele (0). The highest frequency of private alleles was observed for Novo Airão (NA) (for one allele at locus *AC5*, with 19.4%), followed by Teresina (TS) (one allele at locus *B07*, with 18.6%; one allele at locus *AG5*, with 15.7%). Considering all 12 loci analyzed, *AC2* locus had the highest number of private alleles (8) followed by locus *B07* (7).

Table 2: Intra-population genetic diversity of the 15 analyzed Brazilian populations of *Ae. aegypti*

Locus	**AR**	**RP**	**CU**	**CG**	**TA**	**IT**	**MA**	**NA**	**BV**	**RB**	**PV**	**GM**	**IG**	**TS**	**MP**
	2N = 64	**2N = 62**	**2N = 72**	**2N = 72**	**2N = 64**	**2N = 64**	**2N = 64**	**2N = 64**	**2N = 72**	**2N = 70**	**2N=68**	**2N=70**	**2N = 72**	**2N = 70**	**2N = 72**
AC1															
Na	7	9	8	6	6	8	8	7	9	7	4	8	12	8	9
Ne	3.638	4.631	4.880	4.596	3.354	5.607	6.545	5.273	7.364	4.90	2.960	4.522	9.143	5.523	5.411
Ar	6.828	8.791	7.867	5.985	5.873	7.927	7.995	6.997	8.932	6.980	3.995	7.977	11.728	7.927	8.682
r	0.148	-	0.092	0.142	-	-	-	0.123	-	-	0.097	-	0.140	0.219	-
Ho	0.468	0.741	0.628	0.527	0.548	0.800	0.766	0.586	0.871	0.750	0.500	0.7692	0.625	0.419	0.861
He	0.736	0.796	0.806	0.793	0.713	0.835	0.861	0.824	0.878	0.808	0.672	0.794	0.904	0.832	0.826
P-HWE	**0.0009**	0.0106	**0.0003**	**0.0000**	0.0075	0.3544	**0.0002**	**0.0000**	**0.0000**	**0.0000**	**0.0005**	**0.0006**	**0.0000**	**0.0000**	**0.0000**
Fis	0.367	0.070	0.223	0.338	0.234	0.043	0.112	0.293	0.009	0.073	0.260	0.032	0.313	0.500	-0.042
A1															
Na	6	6	7	6	4	6	5	5	8	8	7	6	8	7	5
Ne	3.670	3.469	4.735	3.34	2.864	3.215	3.442	4.038	3.687	4.237	3.945	4.056	4.270	3.171	1.663
Ar	5.847	5.981	6.967	5.951	3.924	5.843	4.924	4.985	7.693	7.808	6.856	5.849	7.699	6.693	4.815
r	-	-	-	-	-	-	-	-	-	-	-	-	-	0.092	-
Ho	0.781	0.838	0.757	0.777	0.875	0.843	0.875	0.741	0.833	0.818	0.911	0.800	0.833	0.529	0.416
He	0.739	0.723	0.801	0.710	0.661	0.699	0.720	0.764	0.739	0.775	0.757	0.764	0.776	0.694	0.404
P-HWE	0.0134	0.0752	0.0081	0.0623	**0.0002**	0.1822	**0.0000**	0.0591	0.020	0.0309	0.0126	0.0366	0.0485	**0.0003**	0.1587
Fis	-0.058	-0.162	0.055	-0.096	-0.330	-0.210	-0.218	0.030	-0.130	-0.056	-0.207	-0.047	-0.074	0.241	-0.031
AC2															
Na	5	6	6	4	8	4	5	5	7	6	5	4	7	7	6
Ne	3.483	2.994	3.33	3.143	5.625	2.381	3.718	2.656	3.939	3.228	2.736	2.151	3.711	3.402	2.173
Ar	4.905	5.483	5.885	3.998	7.715	3.919	4.999	4.904	6.955	5.429	4.982	3.969	6.828	7.000	5.695
r	-	-	-	-	-	-	-	-	-	-	-	-	-	0.133	-
Ho	0.937	0.903	0.625	0.787	0.800	0.687	0.742	0.812	0.818	0.600	0.764	0.676	0.794	0.478	0.647
He	0.724	0.676	0.710	0.692	0.836	0.589	0.743	0.633	0.757	0.700	0.644	0.543	0.741	0.721	0.547
P-HWE	0.0343	**0.0001**	**0.0000**	0.4065	**0.0012**	0.6144	**0.0005**	0.1644	**0.0016**	0.0434	0.2055	0.0926	0.0241	**0.0000**	0.4977
Fis	-0.301	-0.342	0.122	-0.141	0.044	-0.170	0.001	-0.289	-0.081	0.145	-0.191	-0.250	-0.072	0.342	-0.184

AC5															
Na	6	5	7	4	3	3	6	8	6	5	7	5	6	7	10
Ne	2.256	2.326	2.46	2.579	1.331	1.665	4.555	4.928	2.038	3.094	4.129	3.525	4.571	3.657	4.861
Ar	5.698	4.959	6.83	4.000	2.924	2.998	5.985	7.806	5.855	4.976	6.806	4.988	5.960	6.922	9.390
r	-	-	0.085	0.121	-	-	0.112	-	-	-	-	-	-	0.185	-
Ho	0.468	0.517	0.457	0.416	0.218	0.300	0.581	0.806	0.416	0.781	0.882	0.742	0.8000	0.406	0.657
He	0.565	0.580	0.602	0.621	0.252	0.406	0.793	0.810	0.516	0.687	0.769	0.726	0.79255	0.738	0.805
P-HWE	**0.0036**	0.1570	**0.0028**	**0.0000**	0.4786	**0.0022**	**0.0000**	**0.0000**	0.09500	0.4556	0.0740	**0.0034**	0.01914	**0.0000**	0.0580
Fis	0.173	0.110	0.243	0.332	0.135	0.265	0.271	0.005	.195	-0.139	-0.150	-0.023	-0.010	0.454	0.187
AG2															
Na	11	11	16	9	11	8	8	5	14	11	7	11	13	6	8
Ne	3.252	4.643	13.5	4.329	7.015	3.385	3.546	2.779	6.545	5.506	3.410	4.188	5.565	3.729	3.635
Ar	10.683	10.589	15.511	8.572	10.713	7.359	7.851	4.918	12.703	10.276	6.827	10.431	11.962	5.848	7.542
r	-	-	0.077	-	-	-	0.099	-	-	-	-	-	0.142	-	-
Ho	0.838	0.774	0.777	0.80	0.838	0.718	0.548	0.906	0.944	0.800	0.912	0.971	0.563	0.628	0.794
He	0.703	0.797	0.939	0.780	0.871	0.71500	0.729	0.650	0.859	0.830	0.717	0.772	0.833	0.742	0.735
P-HWE	0.83433	0.01436	**0.0004**	0.0470	0.0058	.5415	0.0087	**0.0001**	**0.0000**	**0.0006**	0.2439	0.1202	**0.0000**	0.2628	**0.0000**
Fis	-0.195	0.030	0.174	-0.026	0.038	-0.004	0.252	-0.402	-0.101	0.037	-0.276	-0.263	0.329	0.155	-0.081
AG3															
Na	9	10	8	5	8	9	7	9	7	7	10	8	8	9	12
Ne	5.505	7.308	4.624	2.973	4.633	6.169	5.341	5.389	3.863	5.267	5.898	5.806	6.533	4.263	7.806
Ar	8.822	9.937	7.94	4.885	7.955	8.946	6.897	8.789	6.731	6.895	9.563	7.812	7.966	8.827	11.593
r	-	-	-	-	-	0.116	-	-	-	-	-	-	-	0.098	0.126
Ho	0.718	0.838	0.818	0.657	0.687	0.625	0.800	0.687	0.944	0.823	0.882	0.828	0.785	0.592	0.636
He	0.831	0.877	0.795	0.673	0.796	0.851	0.826	0.827	0.751	0.822	0.842	0.839	0.862	0.779	0.885
P-HWE	0.0100	0.039	0.0181	0.6543	0.0081	**0.0001**	0.1802	0.06	**0.0003**	0.2698	0.0354	0.2895	**0.0005**	**0.0024**	**0.0000**
Fis	0.137	0.045	-0.029	0.024	0.139	0.269	0.033	0.171	-0.261	-0.002	-0.048	0.014	0.090	0.244	0.284
AG4															
Na	6	9	8	7	6	7	6	7	8	5	5	4	7	6	10
Ne	4.024	4.699	3.883	3.24	2.876	3.783	3.004	3.913	3.926	2.737	1.406	2.397	4.520	4.056	3.721
Ar	5.905	8.742	7.694	6.812	5.974	6.858	5.996	6.938	7.502	4.985	4.524	3.997	6.870	5.995	9.856
r	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.223
Ho	0.906	0.870	0.771	0.628	0.531	0.612	0.714	0.733	0.714	0.571	0.323	0.580	0.628	0.657	0.344
He	0.763	0.800	0.753	0.705	0.662	0.747	0.679	0.757	0.756	0.643	0.293	0.592	0.790	0.764	0.744
P-HWE	0.5040	0.0112	**0.0009**	0.0077	0.0233	0.0183	0.7649	0.0784	0.0071	0.0887	1.0000	0.3215	0.0251	**0.0019**	**0.0000**
Fis	-0.191	-0.090	-0.025	0.111	0.201	0.183	-0.053	0.032	0.056	0.114	-0.105	0.020	0.207	0.142	0.541

AG5															
Na	5	10	9	6	9	5	8	7	7	8	7	7	6	9	7
Ne	3.537	5.682	4.767	5.184	6.006	3.241	5.224	4.541	4.342	5.444	4.188	4.667	4.719	7.753	3.165
Ar	4.976	9.930	8.583	5.873	8.808	4.923	7.899	6.919	6.852	7.644	6.344	6.913	5.979	8.925	6.392
r	-	-	-	-	-	-	-	-	-	0.088	-	-	0.179	0.099	-
Ho	0.687	0.827	0.828	0.750	0.750	0.812	0.812	0.906	0.833	0.657	0.882	0.885	0.468	0.685	0.942
He	0.728	0.838	0.801	0.818	0.846	0.702	0.821	0.792	0.780	0.828	0.772	0.797	0.800	0.883	0.694
P-HWE	0.6142	0.0075	0.0292	0.0069	0.0134	0.5769	0.0232	0.0097	0.0375	**0.0000**	0.1667	0.0517	**0.0000**	**0.0000**	**0.0006**
Fis	0.057	0.013	-0.034	0.085	0.116	-0.160	0.011	-0.147	-0.069	0.209	-0.145	-0.113	0.418	0.227	-0.366
AG7															
Na	6	11	11	4	9	9	7	7	10	8	8	9	8	9	8
Ne	2.909	6.471	6.331	2.637	4.055	6.896	4.655	4.247	4.224	4.271	5.898	5.594	4.533	6.671	3.834
Ar	5.941	10.764	10.422	3.985	8.742	8.895	6.996	6.913	9.088	7.597	7.794	8.524	7.524	8.843	7.518
r	-	0.109	-	-	-	-	0.177	-	-	-	-	-	0.107	0.139	-
Ho	0.593	0.645	0.787	0.583	0.687	0.750	0.468	0.896	0.685	0.741	0.794	0.742	0.588	0.593	0.742
He	0.666	0.859	0.855	0.629	0.765	0.868	0.797	0.778	0.774	0.778	0.842	0.833	0.791	0.863	0.749
P-HWE	0.0674	**0.0004**	**0.0004**	0.2521	**0.0007**	**0.0020**	**0.0000**	0.5895	**0.0001**	**0.0002**	0.0374	**0.0016**	**0.0002**	**0.0000**	**0.0000**
Fis	0.111	0.252	0.080	0.074	0.103	0.138	0.416	-0.156	0.116	0.048	0.059	0.110	0.259	0.316	0.010
B3															
Na	7	7	5	6	5	6	6	7	4	5	6	6	8	6	9
Ne	3.574	3.971	3.196	5.126	4.063	3.977	5.005	4.154	3.161	2.589	3.247	3.141	4.614	3.111	5.104
Ar	6.941	6.979	4.971	5.996	4.924	5.904	5.997	6.895	3.999	4.924	5.797	5.797	7.805	5.688	8.840
r	-	-	-	0.082	-	-	-	-	-	-	-	-	-	0.134	-
Ho	0.625	0.741	0.805	0.657	0.781	0.750	0.806	0.843	0.555	0.625	0.970	0.882	0.714	0.454	0.764
He	0.731	0.760	0.696	0.816	0.765	0.760	0.813	0.771	0.693	0.6235	0.702	0.691	0.794	0.689	0.816
P-HWE	0.0059	0.0778	**0.0033**	0.0191	0.1040	0.6247	0.1011	**0.0000**	0.072	0.6855	**0.0000**	0.0165	0.0062	**0.0000**	0.0108
Fis	0.148	0.025	-0.159	0.198	-0.020	0.014	0.009	-0.096	0.201	-0.002	-0.390	-0.281	0.102	0.344	0.064
B07															
Na	8	12	14	10	11	12	10	8	11	9	6	11	11	7	12
Ne	4.697	7.567	8.281	6.715	5.642	7.136	6.759	3.141	5.459	6.331	2.769	7.020	6.397	3.505	9.075
Ar	7.846	11.759	13.023	9.617	10.516	11.600	9.90	5.887	10.632	8.605	5.887	10.427	9.816	6.889	11.757
r	-	-	0.128	-	-	-	-	0.219	-	0.131	-	-	0.148	0.183	-
Ho	0.906	0.967	0.638	0.812	0.906	0.875	0.812	0.312	0.939	0.600	0.558	0.800	0.571	0.400	0.969
He	0.799	0.882	0.891	0.864	0.835	0.873	0.865	0.692	0.829	0.854	0.648	0.869	0.855	0.725	0.903
P-HWE	0.0143	**0.0004**	**0.0000**	0.0330	**0.0028**	0.0752	0.2392	**0.0000**	**0.0000**	**0.0002**	**0.0002**	**0.0013**	**0.0000**	**0.0000**	0.9117
Fis	-0. 136	-0.099	0.286	0.061	-0.086	-0.002	0.062	0.553	-0.135	0.301	0.140	0.082	0.336	0.452	-0.075

CT2															
Na	4	7	9	6	5	6	8	6	6	6	7	8	8	8	10
Ne	2.107	3.636	3.821	1.676	2.532	3.089	5.139	2.913	2.821	3.513	4.211	4.132	3.801	5.560	5.864
Ar	3.919	6.893	8.585	5.686	4.918	5.436	7.907	5.917	5.948	5.885	6.905	7.814	7.694	7.962	9.643
r	-	-	0.103	-	-	-	-	-	-	-	-	-	-	0.195	-
Ho	0.687	0.933	0.558	0.444	0.806	0.812	0.903	0.700	0.757	0.750	0.705	0.857	0.694	0.464	0.916
He	0.533	0.737	0.749	0.408	0.615	0.687	0.818	0.667	0.655	0.726	0.773	0.768	0.747	0.835	0.841
P-HWE	0.0111	**0.0001**	**0.0000**	0.8010	**0.0000**	**0.0002**	**0.0025**	0.1390	**0.0000**	0.0635	0.2566	0.0048	0.0522	**0.0000**	**0.0000**
Fis	-0.294	-0.272	0.257	-0.088	-0.318	-0.186	-0.105	-0.049	-0.159	-0.033	0.089	-0.117	0.072	0.449	-0.091
Mean															
Na	6.667	8.58	9.000	6.083	7.083	6.917	7	6.750	8.083	7.083	6.583	7.25	8.50	7.417	8.833
Ne	3.554	4.783	5.317	3.798	4.166	4.212	4.744	3.998	4.281	4.260	3.733	4.266	5.198	4.534	4.693
Ar	6.526	8.401	8.690	5.947	6.916	6.717	6.946	6.489	7.741	6.834	6.357	7.0142	8.153	7.293	8.477
Ho	0.718	0.8	0.705	0.654	0.703	0.716	0.736	0.744	0.776	0.710	0.757	0.795	0.672	0.526	0.725
He	0.699	0.765	0.772	0.699	0.707	0.717	0.776	0.735	0.738	0.745	0.693	0.738	0.795	0.760	0.735
Fis	-0.030	-0.051	0.085	0.059	0.006	-0.001	0.050	-0.020	-0.044	0.043	-0.095	-0.085	0.150	0.308	0.004

N: number of mosquitoes analyzed; Na: number of alleles; Ne: number of effective alleles; *H*o: observed heterozygosity; *H*e: expected heterozygosity; Ar: allelic richness; *r*: estimated frequency of null alleles; *F*is: inbreeding coefficient. P-HWE: The values in bold indicate disequilibrium in the loci according to Hardy-Weinberg Equilibrium, after adjustment by the sequential Bonferroni correction ($P < 0.0042$). See Table 1 for locality abbreviations

Table 3: List of private alleles and their respective frequencies in parentheses

Pop/Locus	AC1	A1	AC2	AC5	AG2	AG3	AG4	AG5	AG7	B3	B07	CT2	Total
AR			163 (0.031)										1
RP	173 (0.032)				175 (0.032)		181 (0.032)					193 (0.10)	4
CU	207 (0.029)										203(0.028) 205(0.028)		3
CG					133 (0.029)						195 (0.016)		2
TA						194(0.125)							1
IT													0
MA												225 (0.032)	1
NA				129(0.194)			185(0.083)						2
BV			167(0.106) 185(0.045)	173 (0.056)								181(0.076) 219(0.045)	5
RB						158(0.088)							1
PV						156 (0.029)			173(0.029) 193(0.103)				3
GM	175 (0.038)												1
IG	211 (0.031)	168 (0.083)	165(0.059) 195(0.044)		143 (0.031)					183(0.043)	183 (0.043)		7
TS	179 (0.065)		199(0.087)	127 (0.0141)	171 (0.086)	194(0.125)		192 (0.157)			171(0.186)		7
MP	219 (0.069)		197(0.029) 215(0.029)	133(0.029) 137(0.029)		150(0.076) 192(0.030)	143(0.034) 179(0.069)			141(0.088) 144(0.118)	141(0.121) 143(0.076)		13
Total	6	1	8	5	4	6	4	1	2	3	7	4	51

See Table 1 for locality abbreviations.

Table 4 presents the estimates of genetic differentiation and gene flow among the samples. All pairwise F_{ST} values were observed to be highly significant (P = 0.0000 ± 0.0000), before and after the Bonferroni correction. The highest value (F_{ST} = 0.1845) was observed between the samples from the Macapá (MP) and Campina Grande (CG), consequently they had lowest level of gene flow (Nm = 2.2048). On the other hand, the lowest value of genetic distance (F_{ST} = 0.0317) was observed between Manaus (MA) and Itacoatiara (IT), both in state of Amazonas, consequently they showed the highest gene flow (Nm = 15.2987).

AMOVA analysis including all 15 samples (no grouping) revealed highly significant genetic differentiation among them (F_{ST} = 0.1072; P = 0.00000 ± 0.0000) (Table 5). For hierarchical level with two groups (Group 1: AR, RP, CU, CG, TA, IT, MA, NA, BV, RB, PV, GM; Group 2: IG, TS, MP), the analyses revealed no significant genetic structure between them (F_{CT} = 0.0219); but there was a highly significant structure among samples within groups (F_{SC} = 0.10068; P = 0.00000 ± 0.0000) attributed to the differences within group 1 and within group 2. For the hierarchical level with four population groups (Group 1: AR, RP, CU, CG; Group 2: TA, IT, MA, NA, BV; Group 3: RB, PV, GM; Group 4: IG, TS, MP), the percentage of variation for all levels were highly significant (F_{CT} = 0.03448; F_{SC} = 0.08188; P = 0.00000 ± 0.0000). In all three hierarchical levels though, the highest percentages of genetic variation were observed within samples (89.28%; 88.15%; 88.65%, respectively).

Table 4: Genetic differentiation (F_{ST}, below diagonal), gene flow (*Nm*, above diagonal) and approximate geographical distance in km (in parentheses)

Population	AR	RP	CU	CG	TA	IT	MA	NA	BV	RB	PV	GM	IG	TS	MP
AR		10.4168	7.9640	6.2825	4.6821	4.8986	4.6410	3.2018	2.8985	3.0870	2.7725	3.9135	3.2528	2.5912	2.5627
RP	0.0458 (145)		14.9064	7.3297	5.5461	6.0330	7.2320	4.0760	3.9241	5.2377	3.7245	6.1813	5.0469	4.2126	2.9756
CU	0.0591 (910)	0.0325 (925)		9.8967	6.3560	9.0046	8.6515	5.3348	6.3694	7.5578	4.2372	8.6137	5.6310	4.1759	3.8775
CG	0.0737 (2217)	0.0639 (2100)	0.0481 (2392)		5.8790	5.1921	5.3491	4.2199	3.1180	3.8552	3.3210	4.5942	4.1152	3.1181	2.2048
TA	0.0965 (550)	0.0827 (470)	0.0729 (1375)	0.0784 (2040)		8.6112	6.9198	4.1130	5.4872	2.9261	2.9478	3.6654	3.1733	3.2755	2.2661
IT	0.0926 (2205)	0.0765 (2200)	0.0526 (1423)	0.0878 (2552)	0.0549 (2607)		15.2987	7.0463	4.6058	4.2959	3.2185	6.4084	3.9033	3.5794	2.7006
MA	0.0973 (2352)	0.0647 (2305)	0.0546 (1453)	0.0855 (2720)	0.0674 (2695)	0.0317 (178)		10.8797	5.1986	8.0376	5.6493	11.2978	5.2780	5.6567	2.8892
NA	0.1351 (2285)	0.1093 (2370)	0.0857 (1539)	0.1059 (2819)	0.1084 (2796)	0.0663 (284)	0.0439 (116)		3.8155	4.5224	3.5831	6.3545	3.8602	3.5268	2.4464
BV	0.1471 (2870)	0.1130 (2880)	0.0728 (2055)	0.1382 (2960)	0.0835 (3285)	0.0979 (706)	0.0877 (665)	0.1159 (605)		3.6950	2.7586	3.5779	3.0235	3.4636	2.3489
RB	0.1394 (2257)	0.0871 (2315)	0.0621 (1387)	0.1148 (3500)	0.1460 (2770)	0.1043 (1290)	0.0586 (1142)	0.0996 (1118)	0.1192 (1632)		6.1251	9.2992	4.3533	3.6907	2.9791
PV	0.1528 (1995)	0.1184 (2055)	0.1056 (1135)	0.1309 (3088)	0.1450 (2511)	0.1345 (869)	0.0813 (750)	0.1225 (755)	0.1534 (1328)	0.0755 (454)		11.0947	3.4162	3.3412	2.2116
GM	0.1133 (1960)	0.0748 (2043)	0.0549 (1138)	0.0982 (3261)	0.1200 (2502)	0.0724 (1134)	0.0424 (1027)	0.0730 (1021)	0.1226 (1575)	0.0510 (288)	0.0431 (274)		5.0000	4.3035	3.1400
IG	0.1332 (600)	0.0901 (715)	0.0816 (1110)	0.1083 (2895)	0.1361 (935)	0.1136 (2504)	0.0865 (2580)	0.1147 (2622)	0.1419 (3205)	0.1030 (2260)	0.1277 (2103)	0.0909 (1994)		3.4521	2.7665
TS	0.1618 (1975)	0.1061 (1912)	0.1069 (1850)	0.1382 (802)	0.1324 (2035)	0.1226 (1755)	0.0812 (1930)	0.1242 (2037)	0.1262 (2165)	0.1193 (2825)	0.1302 (2370)	0.1041 (2562)	0.1265 (2601)		2.4958
MP	0.1633 (2870)	0.1439 (2328)	0.1142 (1790)	0.1849 (1870)	0.1808 (2625)	0.1562 (885)	0.1475 (1030)	0.1697 (1144)	0.1755 (1112)	0.1437 (2170)	0.1844 (1721)	0.1374 (1975)	0.1531 (2872)	0.1669 (1085)	

All *Fst* values were highly significant. $P = 0.0000 \pm 0.0000$. See Table 1 for locality abbreviations.

Table 5: Hierarchical analysis (AMOVA) of the genetic variation in the *Ae. aegypti* samples

Groups of Samples	Source of variation	Degrees of freedom	Percentage Variation (%)	Fixation index
No Grouping (All) AR, RP, CU, CG, TA, IT, MA, NA, BV, RB, PV, GM, IG, TS, MP	Among populations	14	10.72	$Fst = 0.1072^{***}$
	Within populations	1005	89.28	
Two Groups (1) AR, RP, CU, CG, TA, IT, MA, NA, BV, RB, PV, GM (2) IG, TS, MP	Among groups	1	1.96	$Fct = 0.01958$
	Among populations within groups	13	9.89	$Fsc = 0.10068^{***}$
	Within populations	1005	88.15	$Fst = 0.11846^{***}$
Four Groups (1) AR, RP, CU, CG (2) TA, IT, MA, NA, BV (3) RB, PV, GM (4) IG, TS, MP	Among groups	3	3.45	$Fct = 0.03448^{***}$
	Among populations within groups	11	7.91	$Fsc = 0.08188^{***}$
	Within populations	1005	88.65	$Fst = 0.11354^{***}$

Significance test 10,000 permutations, F_{ST} fixation index within samples, F_{CT} fixation index between regions, *Fsc* fixation index among samples within regions. *** $P = 0.00000 \pm 0.00000$. See Table 1 for locality abbreviations.

As per the genetic structure analyzed by the Bayesian approach in STRUCTURE (Figure 2A), Evanno's method identified K=2 as the most probable number of genetic clusters. Two distinct genetic clusters can be seen and all populations showed more or less mixed ancestry. The first cluster grouped all samples from southeast [Araçatuba (AR), São José de Rio Preto (RP), Taubaté (TA)] and central west [Cuiabá (CU)], seven populations from north [Boa Vista (BV), Itacoatiara (IT), Manaus (MA), Novo Airão (NA), Rio Branco (RB), Porto Velho (PV), Guajará-Mirim (GM)] and one from northeast [Campina Grande (CG)]. The second cluster grouped one population each from north [Macapá (MP)], northeast [Teresina (TS)] and south [Foz de Iguaçú (IG)] of Brazil. The geographic distribution of the clusters is presented in Figure 2C.

We also separately analyzed the clusters (Cluster 1 and Cluster 2) generated from the first STRUCTURE analysis. These analyses revealed the presence of a genetic sub-structure within them (Figure 2B). By the method of Evanno et al. (2005), Cluster 1 was further subdivided into three genetic sub-clusters (K=3) and Cluster 2 was subdivided in two smaller genetic sub-clusters (K=2). The geographic distribution of the sub-clusters is shown in Figure 2D.

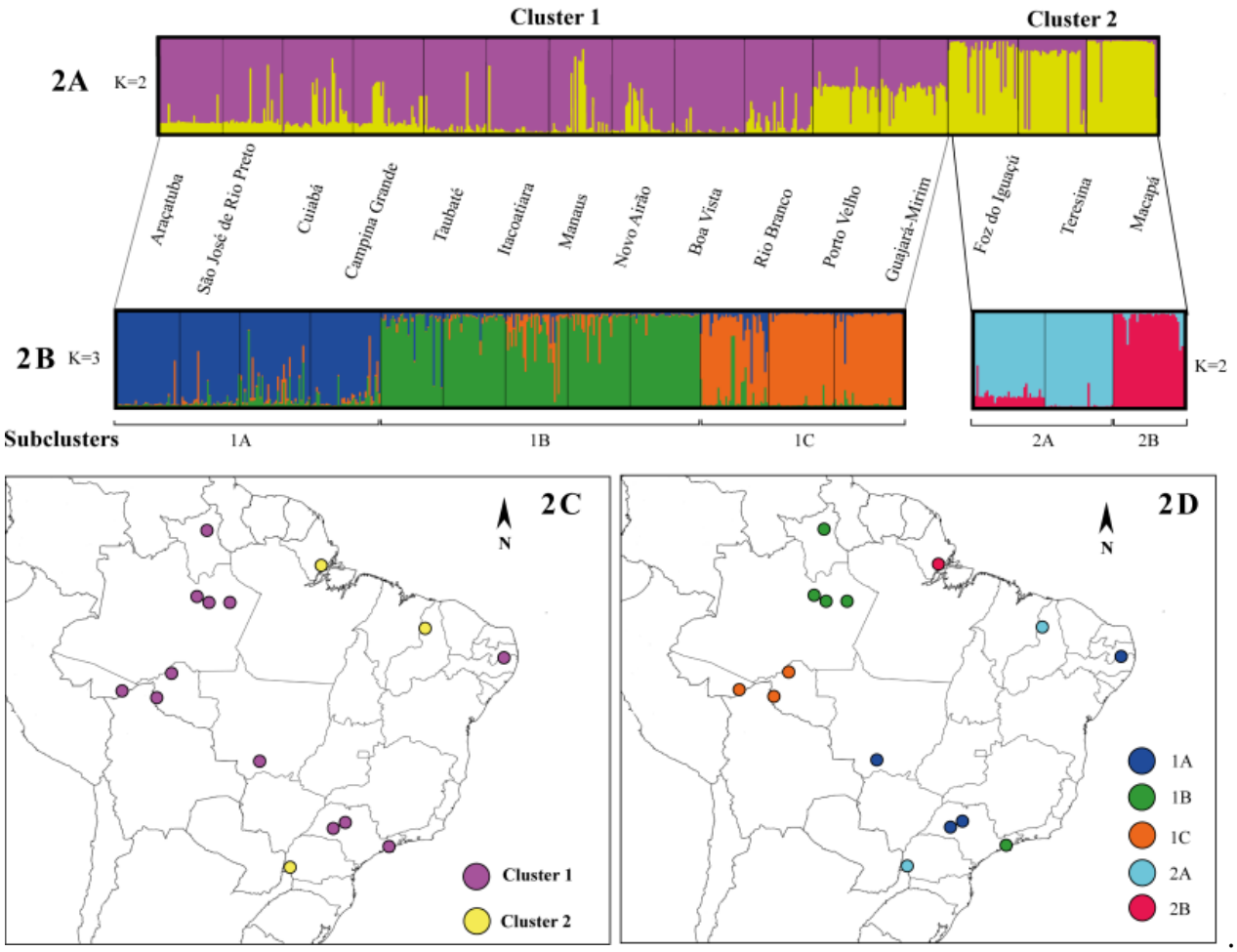

Figure 2. STRUCTURE bar plot for all *Aedes aegypti* populations studied. Individuals are represented by vertical bars along the plot. The height of each color represents the probability of assignment to a specific cluster. The black lines within the plots indicate population limits. A) Subdivision of all the individuals into K = 2 clusters. B) Sub-cluster 1, K = 5; Sub-cluster 2, K = 2. C) Geographical distribution of the 2 clusters. D) Geographical representation of the 5 sub- clusters.

Thirteen genetic clusters (C1 to C13) were identified by the second Bayesian approach implemented in BAPS [Log (marginal likelihood) of optimal partition = -23513.2655; probabilities for number of clusters = 13 (0.999)]. Out of these, only two clusters (C1 and C5) clustered two samples: Araçatuba (AR) and São José de Rio Preto (RP), both from state of São Paulo, were grouped in cluster C1; Itacoatiara (IT) and Manaus (MA), both from the state of Amazonas, were grouped in cluster C5. Each of the remaining samples was assigned in a different cluster (Figure S1).

The number of PCs retained for DAPC analyses were calculated using a cross validation method implemented in 'xvalDapc' function from R adegenet package. In this study, 100 PCs were retained with median and confidence interval for random chance de 97.5 % (0.0880) (Figure S2). Based on the values obtained from 100 PCs, the first DAPC was implemented to observe the genetic variation among the sampling sites. Also, a bar plot of eigenvalues for the discriminant analysis was used to select 14 discriminant functions to be retained (Figure 3). This analysis yielded results similar to those retrieved in STRUCTURE (Figure 2B); the same colors of sub-clusters have been used to facilitate the comparison. The scatterplot of DAPC based on the sampling sites (*a priori*), revealed that specimens from Macapá (MP) were markedly isolated from specimens of other sampling sites (Figure 3). The remaining samples seem to appear in three following groups: 1) TA, BV, IT, MA and NA; 2) RB, PV and GM plus TS; 3) AR, RP, CU and CG. The scatterplot ellipse represented by Foz de Iguaçú (IG) was positioned between groups 2 and 3.

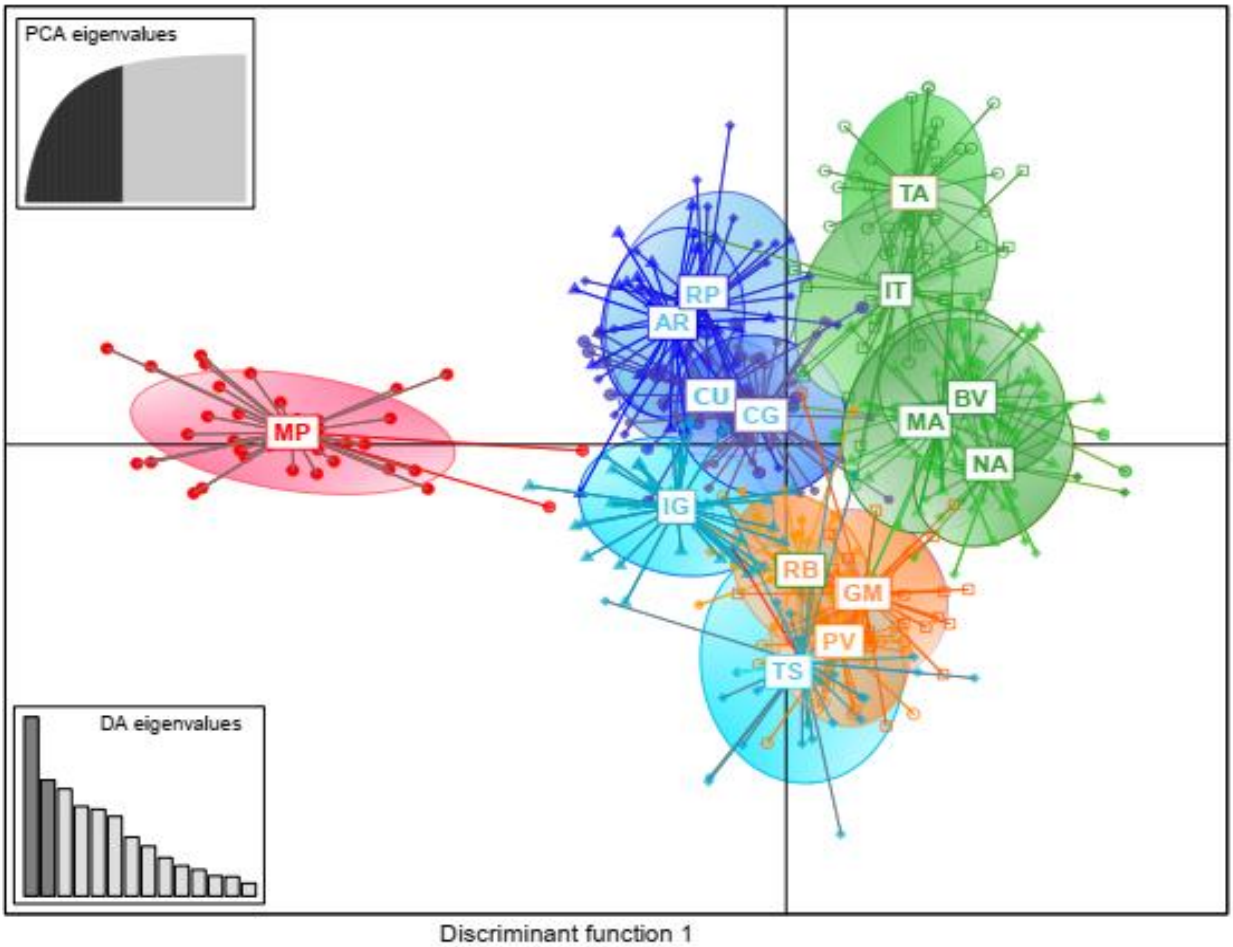

Figure 3. Discriminant analysis of principal components (DAPC) scatterplot for the 15 sampling sites of *Aedes aegypti* and populations as priors. The optimal number of principal components (PCs = 100) was retained as determined by DAPC cross-validation and 14 discriminant functions. DA and PCA eigenvalues of the analysis are displayed in insets. In this plot, populations were selected a priori based on regional location and the first two principle components served as the axes. The graphs represent the individuals as dots and the groups as inertia ellipses. The colors correspond to the five sub-clusters identified by STRUCTURE analysis (dark blue 1A; green 1B; orange 1C; light blue 2A and red 2B).

Based on the Bayesian Inference Criterion (BIC), 13 genetic clusters were identified (BIC=783.442; K=13) (Figure S3), which were similar to the analysis of BAPS. Though some of the scatterplot ellipses were overlapped with each other, at least four clearly defined groupings of ellipses can be observed (Figure S4): cluster 13 (represented by 33 specimens of Macapá), cluster 5 (represented by 14 specimens of Teresina), cluster 11 (represented by 14 specimens of Teresina) and remaining overlapped ellipses representing one group. Most of the samples were attributed their respective clusters with > 85% of membership probability as can be observed by comparing the composition of each genetic cluster (Figure 4A) as well as in the bar plot (Figure 4B). Moreover, it was observed that the composition of some genetic clusters can be compared with the sub-clusters retrieved by STUCTURE (Figure 2B). For example, the clusters C9 and C10 (Figure 4A) correspond to the sub-cluster 1A (Figure 2B), represented by samples from AR, RP, CU and CG; the clusters C3 and C4 correspond to the sub-cluster 1B (Figure 2B), represented by a major quantity of samples from TA, IT, MA and NA; the clusters C6, C7 and C8 correspond to the sub-cluster 1C (Figure 2B), represented by a major quantity of samples from RB, PV and GM; the cluster C2 correspond to the sub-clusters 2A (Figure 2B), represented by a major quantity of samples from IG and, the cluster 13 correspond to the sub-cluster 2B (Figure 2B) , represented by a major quantity of samples from MP. Only the specimens from BV (cluster C12) and TS (clusters C5 and C11), could not be compared with the sub-clusters analyzed by STUCTURE (Figure 2B).

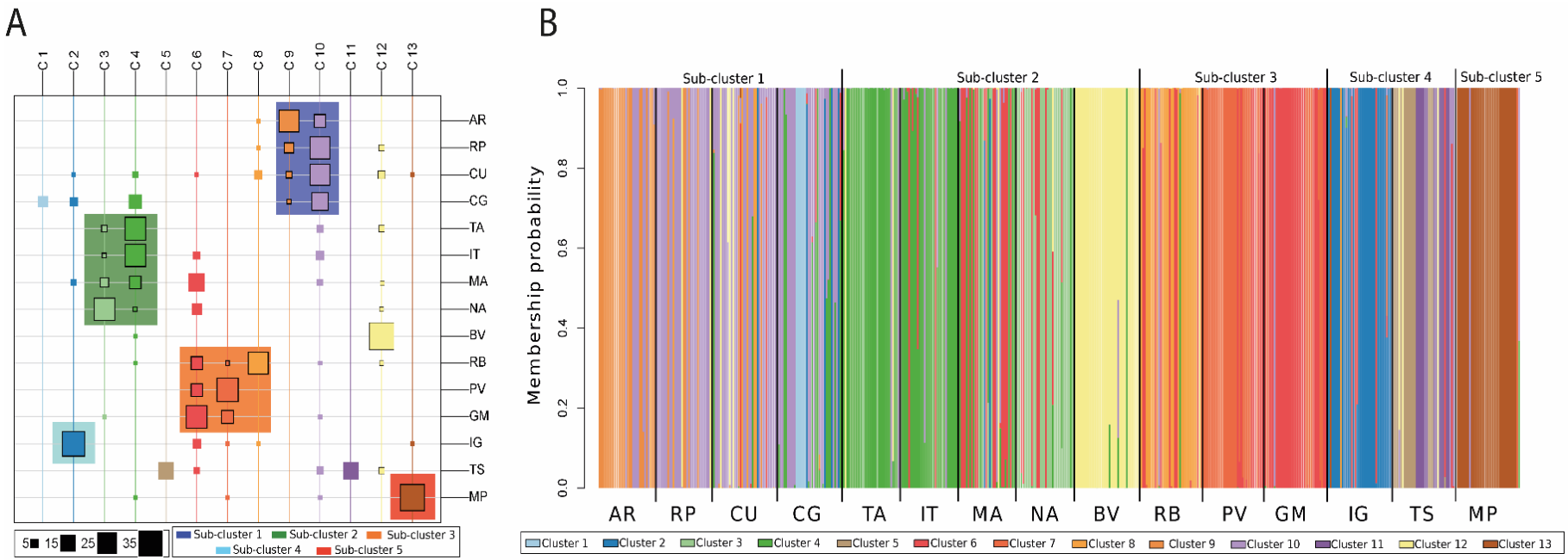

Figure 4. Population structure of Brazilian populations using DAPC. A) The composition of each of the 13 inferred genetic clusters; larger black box indicates more individuals. The scale of the quantity of specimens for each cluster is presented by the squares from 5 to 35 at the bottom of the figure. For example, inferred cluster 1(C1) contains only few individuals from Campina Grande and inferred cluster 9 (C9) contains a mixture of the individuals from Araçatuba, São José de Rio Preto, Cuiabá and Campina Grande. The colors of sub-clusters are the same as sub-clusters of Structure as in Fig. 2B. B) Membership probabilities to 13 inferred genetic clusters for *Aedes aegypti* individuals. See Table 1 for population codes.

The Mantel test analysis for the *Ae. aegypti* revealed a weak, but statistically significant correlation between genetic and geographic distances, (r = 0.31, P = 0.004) among the 15 sampling sites covering a range of ~145 to 3500 km. This result indicates that ~ 31% genetic differentiation observed can be explained by IBD model (Figure 5A). The two-dimensional kernel density estimation indicated a patched pattern of genetic differentiation among sampling sites, due to the two clouds observed, a medium density (between 500 and 1500 km) and the other with a high density (between 1800 and 2800 km) (Figure 5B).

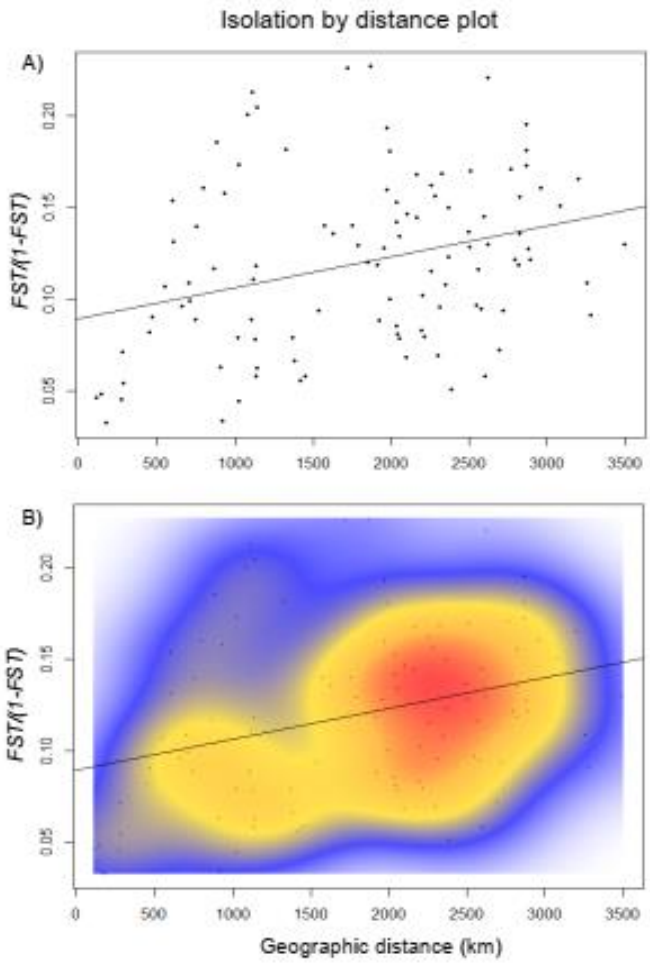

Figure 5. Isolation by distance (IBD) analysis among 15 samples of *Aedes aegypti* (r = 0.031, p = 0.004). A) Pairwise F_{ST}/ (1 – F_{ST}) distances plotted against geographic distances. B) Local density of points plotted using a two-dimensional kernel density estimation. (Line correlation trend is shown; colors represent the relative density of points: blue low density, yellow medium density, red high density).

Table 6 portrays the effective population size (*Ne*) estimates, based on the *LD* model. The lowest *Ne* value was observed in the samples of Campina Grande (CG) (11.8) and the highest value was verified in the samples of Itacoatiara (IT) (∞). The overall average *Ne* value was 75.1 with a 95% CI from 67.8 to 83.2.

The heterozygosity tests performed under IAM, SMM and TPM models are also shown in Table 6. Using IAM, heterozygote excess was detected for all the populations tested as well as all of them were statistically significant ($P < 0.05$). Under SMM model, two [Araçatuba (AR), and Manaus (MA)] of the 15 populations showed significant heterozygosity excess. For the TPM (95%), six populations [Araçatuba (AR), Campina Grande (CG), Itacoatiara (IT), Manaus (MA), Porto Velho (PV) and Guajará-Mirim (GM)] showed significant heterozygote excess. The populations of Araçatuba (AR) and Manaus (MA) had significant heterozygosity excess for all three mutation models tested (IAM, SMM and TPM), indicating that these populations might have experienced a recent bottleneck.

Table 6: Effective population size (*N*e) based on the linkage disequilibrium (LD) model and heterozygosity tests of *Aedes aegypti* samples, based on three mutational models (IAM, SMM and TPM)

Model		AR	RP	CU	CG	TA	IT	MA	NA	BV	RB	PV	GM	IG	TS	MP	Total
LD		26.4	51.3	82.2	11.8	43.5	∞	33.4	26.7	38.2	75.7	37.0	29.4	111.5	23.2	15.5	75.1
95%CI		18.7-39.9	35.2-86.2	49.7-194.0	9.2-15.2	28.8-76.7	121.0-∞	24.3-49.5	19.2-39.8	27.0-59.2	44.4-193.9	25.3-60.4	21.7-42.2	58.2-571.1	17.9-31.3	12.6-19.3	67.8-83.2
IAM	$H_e < H_{eq}$	0	0	1	0	1	1	1	1	1	1	0	1	0	0	1	
	$H_e > H_{eq}$	12	12	11	12	11	11	11	11	11	11	12	11	12	12	11	
	P ($H_e > H_{eq}$)	**0.0001**	**0.0001**	**0.0012**	**0.0001**	**0.0002**	**0.0017**	**0.0002**	**0.0009**	**0.0009**	**0.0002**	**0.0001**	**0.0006**	**0.0001**	**0.0001**	**0.0004**	
SMM	$H_e < H_{eq}$	4	4	7	3	5	5	3	4	6	8	4	3	5	7	9	
	$H_e > H_{eq}$	8	8	5	9	7	7	9	8	6	4	8	9	7	5	3	
	P ($H_e > H_{eq}$)	**0.0320**	0.2349	0.9451	0.0647	0.3667	0.2120	**0.0017**	0.3955	0.6614	0.8098	0.1167	0.2593	0.1902	0.8982	0.9988	
TPM (95%)	$H_e < H_{eq}$	3	4	7	2	4	2	2	3	5	5	3	3	4	6	9	
	$H_e > H_{eq}$	9	8	5	10	8	10	10	9	7	7	9	9	8	6	3	
	P ($H_e > H_{eq}$)	**0.0031**	0.1018	0.8303	**0.0171**	0.1506	**0.0320**	**0.0009**	0.1018	0.3386	0.2349	**0.0386**	**0.0461**	0.0647	0.3667	0.9983	

CI: Confidence interval. IAM: Infinite Alleles Model; SMM: Stepwise Mutation Model; TPM: Two Phase Model; H_e: expected heterozygosity; H_{eq}: expected heterozygosity under mutation and drift equilibrium; $H_e > H_{eq}$: number of loci showing heterozygote excess, P ($H_e > H_{eq}$): P-value of Wilcoxon tests (one tail for heterozygote excess. Values highlighted in bold are those indicative of a bottleneck ($P < 0.05$). See Table 1 for locality abbreviations.

4.1.4. Discussion

Genetic clusters and gene flow patterns

The results of this study, based on the pair-wise genetic distances and AMOVA, indicated a low to moderate genetic differentiation among the 15 Brazilian populations of *Ae. aegypti*, where ~ 31% of this difference can be explained by IBD model. The geographically closer populations showed lowest distance values [between Araçatuba (AR) and São José de Rio Preto (RP), Porto Velho (PV) and Guajará-Mirim (GM), and between Manaus (MA) and Itacoatiara (IT)]. Considering that Manaus was the first city in the state of Amazonas to be infested with *Ae. aegypti* in 1996, it is possible that the population of Itacoatiara (IT) was founded by individuals from Manaus (MA). The lack of private alleles and the largest *Ne* in the samples of Itacoatiara suggest either an extensive gene flow with Manaus (*Nm*=15.2987) or that the events of colonization are not old enough to generate private alleles and to achieve a proper genetic structure.

On the contrary, the populations of Manaus (MA), Itacoatiara (IT), Novo Airão (NA) and Taubaté (TA), which are geographically distant from each other, also showed genetic homogeneity and were included in the same genetic cluster and sub-cluster. Altogether, these findings suggest genetic connectivity among the populations of the localities, with or without geographical proximity. Considering that *Ae. aegypti* has a low flight range varying from 50 to 800 m (Harrington et al., 2005; Honório et al., 2003), therefore, this genetic connectivity between localities distant might be attributed to its dispersal across long distances via passive migration and it may be correlated with the human migrations and commercial traffic by roads, highways and fluvial networks (Bosio et al., 2005; Gonçalves da Silva et al., 2012; Kotsakiozi et al., 2017).

The samples of Cuiabá (CU) were observed to be genetically closer to Araçatuba (AR), São José do Rio Preto (RP) and Campina Grande (CG), which can be attributed to the central position of Cuiabá, serving as connection route between southeastern and northeastern regions of Brazil. This hypothesis also may help to explain the highest number of alleles and allelic richness obtained for this population. Interestingly, the populations of Teresina (TS) and Campina Grande (CG), which are geographically closer to each other, were found to be genetically distant, which was noted as well between the populations of Taubaté (TA) and Araçatuba (AR)/São José de Rio Preto (RP), also observed by Scarpassa et al. (2008). The city of Taubaté (TA) has a huge network of highways with an intense flow of people and trade activities connecting it to other major cities

of Brazil. Thus, the population of *Ae. aegypti* from Taubaté may have been colonized by individuals introduced from another region of Brazil. Alternatively, as Manaus was infested by *Ae. aegypti* in 1996, after the infestation of southeastern Brazil, it is possible that the mosquitos which have colonized Manaus came from southeastern Brazil.

The clustering of Rio Branco (RB), Porto Velho (PV) and Guajará-Mirim (GM) had low values of F_{ST} (0.04 - 0.07) among them, may be because of their geographical proximity and well-connected network among these cities; Guajará-Mirim (GM), being located at the border between Brazil and Bolivia, may serve as an entry point for new genetic groups of *Ae. aegypti*.

The population of Foz do Iguaçú (IG) was found to be genetically different from other populations analyzed in this study, which may be due to its location at the border of Brazil, Argentina and Paraguay and also, as it is a huge tourist attraction, thereby, bringing in more flux of human as well as trade. Gloria-Soria et al. (2016) had suggested the existence three genetic groups in South America: a northern group including Colombian, Venezuelan and northern Brazil populations; a southern Brazil group; and an Argentina group. They had also observed that Argentina was the only region outside Africa with the evidence of recent ancestry of mosquitoes from Africa, which can extend to populations of southern Bolivia and Paraguay (Llinas and Gardenal, 2011; Paupy et al., 2012; Rondan Dueñas et al., 2009).

The population of Teresina (TS), on the contrary, seems to be genetically isolated, may be because of its geographical location, or, due to its hot and dry climate resulting to a differentiated genetic structure as a consequence of low gene flow with the other populations analyzed here. Similar genetic isolation has also been observed by Gloria-Soria et al. (2016b) for the population of *Ae. aegypti* in Goudiri (Africa). Also, local adaptations and micro geographic factors like, urbanization and subsequent increase in favorable breeding sites can affect the microevolution and genetic structuring of this vector (Louise et al., 2015; Wilke et al., 2017).

However, all the fourteen samples analyzed in this study were found to be genetically closer to each other when compared to the population of Macapá (MP). The largest number of private alleles and the highest values of genetic distance (F_{ST}) with all other samples of this study, for the sample of Macapá (MP) could explain its highest divergence, possibly due to the colonization of individuals introduced from a different geographic region. Kotsakiozi et al. (2017) reported similar results with microsatellites loci; Rio Branco and Sao José de Rio Preto were

grouped in different sub-clusters. These authors also found highest differences, based on the cluster analyses, between Macapá/Belém/Santarém and remaining samples from Brazil, whereas Monteiro et al. (2014), also using microsatellites loci, reported genetic differences between Marabá/Tucuruí and other samples from Brazil. Altogether, it is possible to infer that the samples of *Ae. aegypti* from Belém, Santarém, Marabá and Tucuruí (state of Pará) and Macapá (state of Amapá) consist of a separate genetic group from the remaining populations of Brazil. Lima Júnior and Scarpassa, (2009), using the *ND4* gene, demonstrated that Belém had the highest number of private haplotypes, where most of them did not share any haplotype with those of the previous studies. Belém is situated on the bank of a large river (Amazonas river) and near Atlantic Ocean, whereas state of Amapá (Macapá) shares borders with Suriname and French Guiana, and also it is situated on the bank of Amazonas river, favoring the introduction of new genetic groups of this vector.

The analyses of genetic structure of this study indicated the existence of two major genetic clusters (Fig. 2), supporting the previous studies of Brazilian *Ae. Aegypti* (Bracco et al., 2007; Monteiro et al., 2014; Paduan et al., 2008, 2006; Scarpassa et al., 2008) and confirming the existence of two distinct groups of *Ae. aegypti* circulating in the country. However, the results from this study also highlight the existence of population sub-structuring within these two established genetic structures. The presence of distinct genetic clusters and sub-clusters of *Ae. aegypti* in Brazil has also been discussed in earlier studies of Monteiro et al. (2014), Gloria-soria et al. (2016) and Kotsakiozi et al. (2017). These genetic differences probably result from cumulative events of multiple introductions, urbanization, passive dispersal associated with intensive human migration and trade networks, distinct climatic conditions and extensive use of insecticides for the vector control activities throughout the country, followed by cycles of extinction and re-colonization and genetic drift, as has been proposed by Scarpassa et al. (2008). Another explanation is the probable existence of a relict population which might have escaped the eradication efforts, as many regions due to their remoteness could not have achieved a total eradication. Moreover, eradication was never fully achieved in northern South America, Caribbean and the United States (Gubler, 1998). These cumulative factors influence the levels of gene flow and genetic differentiation among and within different regions and localities of this country and, consequently, the genetic structure of this vector.

Demographic changes among the populations

The relatively small *N*e values observed in this study are in accordance with four others samples of this vector from Manaus, as studied by Mendonça et al. (2014) and the two Brazilian samples from Jacobina and Cachoeiro, as studied by Saarman et al. (2017); however, the *N*e values were relatively low when compared with *N*e values observed in anopheline malaria vectors (Scarpassa and Conn, 2007), which are seven-fold to infinity higher. This suggests that these mosquitoes form localized breeding units even in large cities where the regional census size may be larger (Saarman et al., 2017).

The genetic bottleneck analysis indicated that none of the samples analyzed in this study appear to be expanding. On the contrary, a significant signal of bottleneck effect was observed in the samples of Araçatuba (AR) and Manaus (MA), under the three mutation models tested, and in the sample of Campina Grande (CG), Itacoatiara (IT), Porto Velho (PV) and Guajará-Mirim (GM) under the two models tested (IAM and TPM). These tests indicate that they experienced a recent bottleneck event, which are mostly in agreement with the effective population sizes observed in these samples (Table 6), Itacoatiara (IT) being an exception though. In spite of showing indications of recent bottleneck events, the *N*e was observed to be ∞; which may be attributed to the stochastic sampling errors of this population. Every population of this study have exhibited different demographic dynamics, like, low *N*e with signs of bottleneck [Campina Grande (CG), and Aracatuba (AR)], high *N*e with no bottleneck [Foz do Iguacu (IG), São José de Rio Preto (RP), Taubaté (TA), Cuiabá (CU), Rio Branco (RB)], low *N*e with no signs of bottleneck [Macapá (MP) and Teresina (TS)] and high *N*e with signs of bottleneck [Itacoatiara(IT)]; re-enforcing the hypothesis of microevolution and micro–geographic population structuring due to the involvement of non-neutral characters and occasionally affecting the susceptibility of mosquitoes to insecticides and chemical repellents (Louise et al., 2015), leading to possible events of bottleneck in a population. We can't deny the possibility of independent events of multiple introductions either, which is usually portrayed by severe bottlenecks that may reduce the genetic variability of founding populations to levels incompatible with an expansion phase and thereby, may help to attain sufficient additive genetic variability for efficient expansion and adaptation to novel environments (Kamgang et al., 2011).

4.1.5. Conclusion

Considering all analyses of this study, we can summarize that the population of *Ae. aegypti* in Brazil have undergone huge changes in its genetic structure since its eradication in the 1950s, its re-introduction in the country and till the recent country-wide outbreaks of dengue and zika. These changes can be attributed to the combined effect of many anthropogenic factors including the overuse of insecticides as vector control measures resulting in repeated extinction, re-colonization events, events of multiple introduction from both inside and outside Brazil, immigration of mosquitoes due to increased urbanization and, improved commercial roadways, highways, airways and river networks. We conclude that despite the existence of two major genetic clusters observed for *Ae. aegypti* in Brazil, also exists significant genetic structuring within them. These differences in the population genetic structure may affect its vector competence to transmit DENV, CHKYV, ZIKV and other arboviruses, and also in their response to control programs using genetic methods targeted to suppress or genetically modify the vector populations and subsequently decrease their efficiency to transmit pathogens (McGraw and O'Neill, 2013). This study enforces the need for more intensive and fine - scale studies throughout the country, which can help in planning and formulation of better vector control strategies and surveillance protocols, specifically for Brazil, and consequently reduce and prevent massive outbreaks of associated arboviral diseases.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This study was funded by MCTIC/INPA (Process number: PRJ 12.311) and by grants from Amazonas Research Foundation (FAPEAM, process number**:** 3111/2012) and from PRÓ-EQUIPAMENTO-CAPES to Vera Margarete Scarpassa. We thank Renata da Silva Leandro, Marllon Rinaldo de Lima Andrade, Steffany Larissa da Silva Galdino for the field collections from Campina Grande; Dr. Raimundo Nonato Picanço Souto and Dr. Fred Júlio Costa Monteiro for the field collcetions from Macapá; and MSc. Gersonval L.S. Monte for the field collections from Teresina. We also thank MSc. Jose F. Saraiva, MSc. Sandra M. H. Rangel and Dr. Ronildo B. Alencar, for helping with the analyses, figures and maps.

Additional Information

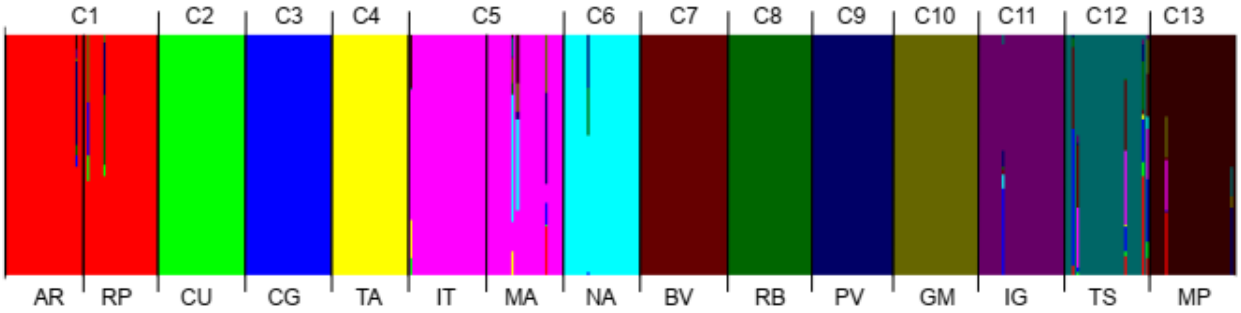

Figure S1. Bayesian genetic cluster analysis of BAPS with K = 13 clusters. The height of each color represents the probability of assignment to a specific cluster. The black lines within the plots indicate population limits. AR: Araçatuba, RP: São José do Rio Preto, TA: Taubaté, RB: Rio Branco, PV: Porto Velho, GM: Guajará-Mirim, IG: Foz do Iguaçú, CU: Cuiabá, CG: Campina Grande, TS: Teresina, BV: Boa Vista, IT: Itacoatiara, MA: Manaus, RB: Rio Branco, MP: Macapá, NA: Novo Airão.

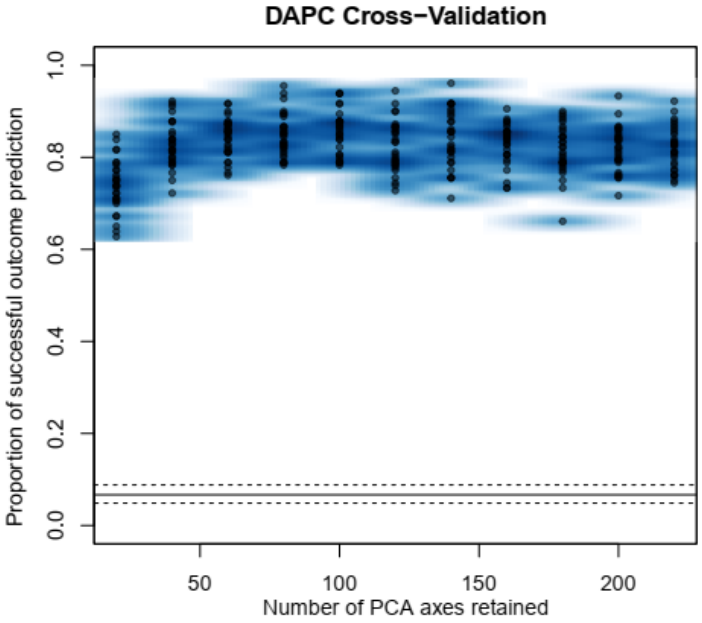

Figure S2. Cross-validation procedure to choose the optimal number of Principal Components for the DAPC analysis.

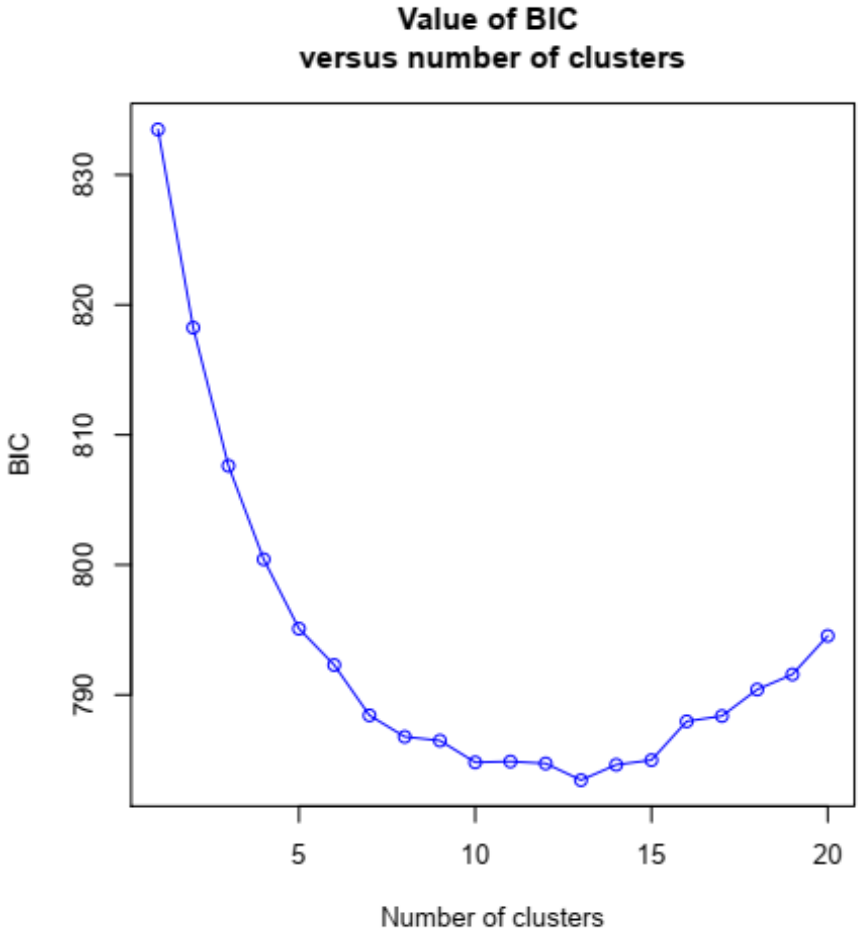

Figure S3. Selection of the optimal number of clusters in the DAPC using the lowest Bayesian Information Criterion (BIC).

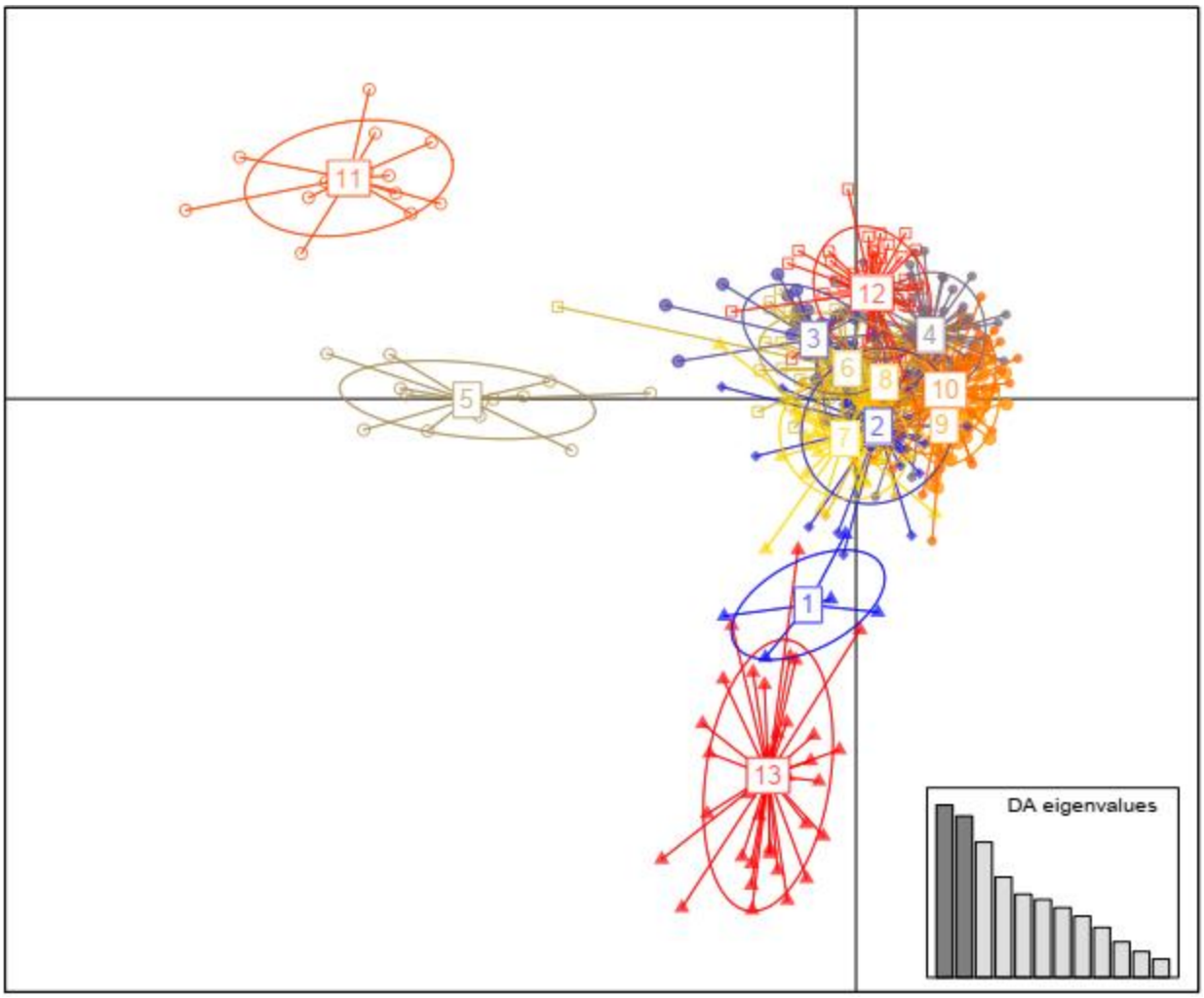

Figure S4. DAPC scatterplots representing the inferred genetic clusters by using the first two principle components as axes. The graphs represent the individuals as dots and the groups as inertia ellipses. DA eigenvalues of the analysis are displayed in insets.

Table S1 Sequences of microsatellite loci used in the study.

Locus	Sequences 5'-3'	Annealing temperature (°C)	Fluorescent Primer	References
A1	F: GACGTAAACCGAGTGGGAGA R: GCATTTAACCGCGCTAGAAC	55	M13- FAM	Brown et al. (2011)
AC1	F: TCCGGTGGGTTAAGGATAGA R: ACTTCACGCTCCAGCAATCT	55	M13- TAM	Slotman et al. (2007)
AC2	F: AATACAACGCGATCGACTCC R: AACGATTAGCTGCTCCGAAA	55	M13- TAM	Slotman et al. (2007)
AC5	F: TGGATTGTTCTTAACAAACACGAT R: CGATCTCACTACGGGTTTCG	60	M13- TAM	Slotman et al. (2007)
AG2	F: TCCCCTTTCAAACCTAATGG R: TTTGCCCTCGTATGCTCTCT	52	M13- FAM	Slotman et al. (2007)
AG3	F: CGCCAAAACTGAAAACTGAA R: AAGGGCGGTGATGACTTTCT	60	M13- HEX	Slotman et al. (2007)
AG4	F: AAACCTGCGCAACAATCAT R: AAGGACTCCGTATAATCGCAAC	60	M13- TAM	Slotman et al. (2007)
AG5	F: TGATCTTGAGAAGGCATCCA R: CGTTATCCTTTCATCACTTGTTTG	52	M13-FAM	Slotman et al. (2007)
AG7	F: CGTGCGAGTGAATGAGAGAC R: CATCCTCTCATCAGCTTCTAATAA	52	M13-HEX	Slotman et al. (2007)
B3	F: GCAAGTTGCAAAGTGCTCAA R: ACCCACCGTTTGCTTTGTAG	55	M13-HEX	Brown et al. (2011)
B07	F: CAAACAACGAACTGCTCACG R: TCGCAATTTCAACAGGTAGG	60	M13- FAM	Chambers et al. (2007)
CT2	F: CGCAGTAGGCGATATTCGTT R: ACCACCACCAACACCATTCT	55	M13- HEX	Slotman et al. (2007)

F: forward; R: reverse; All forward primers were designed with a short M13 tail at the start (TCCCAGTCACGACGT)

Table S2. Significant loci-pairs of Linkage Disequilibrium (LD) test.

Population	Pair-loci
Araçatuba	1) AC1-AC5
	2) AG7-B3
	3) A1-B3
	4) AC5-CT2
	5) AG4-CT2
	6) B07-CT2
Boa Vista	1) AC1-AC2
	2) A1-AC5
	3) AC1-AG2
	4) A1-AG3
	5) AG2-AG3
	6) AC1- AG5
	7) AG4 - AG5
	8) AC1 - AG7
	9) AC2- AG7
	10) B3- AG7
	11) AG4- B3
	12) AG5- B3
	13) AC1- B07
	14) AG2- B07
	15) B07- CT2
Campina Grande	1) AC1-A1
	2) A1- AC2
	3) AC1- AC5
	4) A1- AC5
	5) AC2- AC5
	6) AC5 - AG2
	7) AC5 - AG4
	8) A1 - AG5
	9) AC5 - AG5
	10) AG2 - AG5
	11) AG4 - AG5
	12) A1 - B3
	13) AC5 - B3
	14) AG4 - B3
	15) AG5 - B3
	16) A1 - B07
	17) AC5 - B07
	18) A1 - CT2
Cuiabá	1) AC1 - AC2
	2) AC1 - AC5
	3) AC5 - AG4
	4) AG3 - AG4
	5) A1 - AG7
	6) AG4 - AG7
	7) A1 - B3
	8) AC1 - B07
	9) B07 - CT2

Guajará-Mirim	1) AC1- AG2 2) AG2- AG3 3) AC2- AG5 4) AC5- AG5 5) AG5 - B07 6) AG7 - B07 7) B3 - B07 8) AC2 - CT2 9) AG7 - CT2 10) B3 - CT2 11) B07 - CT2
Foz do Iguaçú	1) AG2 - AG5 2) AG2 - B07 3) B3 - B07
Macapá	1) AC1 - AG2 2) A1 - AG2 3) A1 - AG3 4) AC1 - AG5 5) AC5 - AG5 6) AG2 - AG5 7) AG3 - AG5 8) AC1 - AG7 9) AC2 - AG7 10) AG2 - AG7 11) AG5 - AG7 12) AC1 - B3 13) AG5 - B3 14) AG7 - B3 15) B3 - B07 16) AG2 - CT2 17) AG5 - CT2 18) AG7 - CT2
Manaus	1) AC1 - A1 2) AC1 - AC2 3) AC1 - AG2 4) A1 - AG3 5) A1 - AG3 6) AC2 - AG3 7) AC2 - AG4 8) AC2 - AG5 9) AC2 - AG7 10) A1 - B07 11) A1 - CT2 12) AC2 - CT2 13) AG5 - CT2 14) AG7 - CT2
Novo Airão	1) AC1 - AC2 2) AC5 - AG2 3) AC2 - AG4 4) AC1 - AG5 5) AC2 - AG5 6) AC5 - AG5 7) AG2 - AG5 8) AC2 - B3

	9) AC1 - CT2
Porto Velho	1) AG2 - AG3 2) AC2 - AG5 3) AG5 - AG7 4) AG3 - B3 5) A1 CT2 6) AC5 - CT2
Rio Branco	1) A1 - AG3 2) AG2 - AG5 3) AC1 - AG7 4) AG2 - AG7 5) AG5 - AG7 6) AG2 - CT2 7) AG7 - CT2
São José do Rio Preto	1) A1 - AC5 2) AC2 - AG2 3) A1 - AG4 4) AG2 - AG4 5) AG2 - AG7 6) AG4 - AG7 7) A1 - B3 8) AG3 - B3 9) AC1 - B07 10) A1 - CT2 11) AC2 - CT2 12) AG4 - CT2 13) B07 - CT2
Teresina	1) A1 - AC5 2) AC5 - AG5 3) AG5 - B3 4) A1 - B07 5) AG5 - B07 6) AC1 - CT2 7) AG5 - CT2 8) B3 - CT2 9) B07 - CT2
Taubaté	1) AC1 - AC2 2) AC2 - AC5 3) A1 - AG2 4) AC2 - AG4 5) A1 - B3 6) AC2 - B3 7) AC2 - CT2 8) AC5 - CT2 9) AG2 - CT2

4.2. Capítulo 2: Genetic differentiation and gene flow patterns of *Aedes aegypti* (Diptera: Culicidae) in Brazil

Artigo a ser submetido para a revista *Parasites and Vectors*.

Abstract

Aedes aegypti, being the principal vector of dengue (DENV1 to 4), chikungunya (CHKYV) and Zika (ZIKV) viruses, is considered as one of the most important mosquito vectors. Between the 16th and 18th centuries, during the period of colonization and slave trade, *Aedes aegypti* (Diptera: Culicidae), the mosquito native to Sub-Saharan Africa, was considered to be introduced to Brazil. The species was supposedly eradicated between 1950 – 1960 from numerous American countries including Brazil, followed by events of multiple introductions of this vector and subsequent incidences of dengue, chikungunya and Zika, throughout the country. In this study, we assessed the genetic variability and population genetics of *Ae. aegypti* samples from 14 Brazilian cities, to explore their genetic structure and probable geographic origin, using mitochondrial DNA marker (COI- *Barcode*). A total of 161 mtDNA sequences were analyzed: eight locations from northern (Itacoatiara, Manaus, Novo Airão, Boa Vista, Rio Branco, Porto Velho, Guajará-Mirim and Macapá), three from southeastern (Araçatuba, São José de Rio Preto and Taubaté), one from central west (Cuiabá) and two from northeastern (Campina Grande and Teresina) regions of Brazil. Our analysis with mtDNA sequences retrieved 20 Brazilian haplotypes of *Ae. aegypti* and also confirmed the existence of two major genetic groups circulating in Brazil. We also observed 3 possible entry points of introduction of this vector in Brazil. Moreover, the analysis of the Bayesian Analysis of Population Structure (BAPS) recovered the presence of five genetic groups in Brazil. AMOVA revealed significant genetic variation for almost all the hierarchical levels, indicating high level of genetic differentiation; which may result in different levels of vectoral competence for arboviruses. Therefore, these results are deemed to be necessary to identify the difference in the population genetic structure of this mosquito which can potentially impact the epidemiology of dengue, Chikungunya and Zika in Brazil.

Key words: Dengue vector, Population genetics, Brazil, Genetic structure, Mitochondrial DNA.

4.2.1. Introduction

Aedes aegypti, considered the main vector of DENV in the Americas, is also the principal vector of Zika (ZIKV), chikungunya (CHKYV) and urban yellow fever (YFV). This mosquito species originated in sub-Saharan Africa, where domesticated and adapted to the urban environment, becoming anthropophilic, and their larvae are usually found in artificial breeding grounds, predominating in urban environments (Barreto and Teixeira, 2008). Human migrations around the world have caused huge changes in its landscape and favored the dispersion of *Ae. aegypti*.

Aedes aegypti was introduced in Brazil during the colonial period, probably at the time of the slave trade (Consoli and Oliveira, 1994; Figueiredo, 2003). However, before the arrival of dengue virus, it is believed that this vector already existed for many years in the American continent (Franco, 1976). The eradication campaign of *Ae. aegypti* in Brazil began with Oswaldo Cruz in 1904 as a fight against the yellow fever (Figueiredo, 2003). In the following years, until 1920, with the help of Rockefeller Foundation, Brazil was able to keep away yellow fever epidemics for several years (Franco, 1976). Between 1950s and 1960s, *Ae. aegypti* was considered to be eradicated from Brazil and from several other countries of the American continent (MS 2004). However, eradication has not been very successful in Suriname, Guyana and Venezuela, likely causing the reintroduction of this vector in Brazil in 1967, in the State of Pará (Consoli and Oliveira, 1994). Subsequently, this mosquito was found in Bahia (1976), in Rio de Janeiro (1977) and in the State of Roraima at the beginning of the 1980 (Schatzmayr, 2000) Since, 1998, *Ae. aegypti* is present in all Brazilian states (Scarpassa et al., 2008).

In recent years, a large number of population genetics studies have been performed with *Ae. aegypti*, which it has been the most studied vector since the later ten years. The use of isozyme markers has been important in the genetic characterization of the two subspecies: *Ae. aegypti formosus* and *Ae. aegypti aegypti* (Powell et al., 1980; Tabachnick and Powell, 1978; Wallis and Tabachnick, 1990). Later, several studies have been carried out using DNA sequences to assess the genetic variability of these mosquitoes, for example, by using Restricted fragmented length Polymorphism (RFLP) (Yan et al., 1999), Amplified Fragments Length Polymorphism (AFLP) (Merrill et al., 2005), and using Random amplification of DNA (RAPD) (Apostol et al., 1996; De

Sousa et al., 2001; Gorrochotegui-Escalante et al., 2002; Paduan et al., 2006; Santos et al., 2003; Wesson and Ocampo, 2004). In Brazil, the genetic diversity of populations of *Ae. aegypti* of five Brazilian states was accessed with RAPD markers by Ayres et al. (2003), that revealed high polymorphism and genetic differentiation among populations of this vector in Brazil.

The Control region (also known as the A + T region or D-loop region) of mitochondrial DNA (mtDNA) has been used to analyze in six populations of *Ae. aegypti* in Argentina revealing significant genetic differentiation and absence of any correlation between geographic and genetic distances (Rondan-Dueñas et al., 2009). In Australia, Beebe et al. (2005) used the *COI* gene of mtDNA to analyze the genetic structure of local *Ae. aegypti* and of some nearby islands, and found eight haplotypes with a discrete genetic variation among the analyzed samples. Among the mitochondrial genes used in population genetics studies *of Ae. aegypti*, the NADH dehydrogenase, subunit 4 (ND4) mtDNA has also been seen to be an excellent marker for detect genetic differentiation. Gorrochotegui-Escalante et al. (2002) conducted a genetic study in Mexico using both RAPD and ND4 markers and observed 25 haplotypes. In Thailand, a study of nineteen populations using the same marker (ND4) generated seven haplotypes and a lower genetic diversity, compared to the populations of Mexico (Bosio et al., 2005). In South America, Costa-da-Silva et al. (2005) also used ND4 to estimate the genetic variability of *Ae. aegypti* from three cities in Peru (Lima, Piura and Iquitos), three haplotypes were found. Mousson et al (2005) investigated the relationships between *A. aegypti* and *A. albopictus* mosquitoes based on three mitochondrial-DNA genes (cytochrome b, cytochrome oxidase I and NADH dehydrogenase subunit 5) and observed that most South American populations of *A. aegypti* were genetically similar to South-East Asian populations (Thailand and Vietnam), except for one sample from Boa Vista (northern Amazonia), which was closely related to samples from Africa (Guinea and Ivory Coast). They suggested that African populations of *A. aegypti* introduced during the slave trade have persisted in Boa Vista, resisting eradication campaigns. Paduan et al. (2006), with the Random Amplified Polymorphic DNA (RAPD) markers, found high polymorphism and high levels of genetic differentiation in populations of *Ae. aegypti* from six Brazilian states, the population being divided into two main groups, one comprising the populations from Northeast and the other, comprising the populations from North, Central-West and Southeast regions of Brazil. Bracco et al. (2007) analyzed the genetic relatedness and phylogeographic structure of *Aedes aegypti* from Brazil, Peru, Venezuela, Guatemala, US, Africa and Asia using ND4 gene.

The authors described 20 distinct haplotypes divided in two clades and hypothesized the occurrence of two subspecies of *Ae. aegypti* in the Americas. Lima-Júnior and Scarpassa (2009), used ND4 for a population study of this vector in the Brazilian Amazon, found high number of haplotypes (13), high average values of genetic diversity and a relative gene flow among the ten sampled localities in their study. Scarpassa et al. (2008), using the *COI* gene of mtDNA, also analyzed the population structure and dispersion patterns of *Ae. aegypti* in four Brazilians regions. The authors found ten haplotypes distributed in two major genetic groups, suggesting that the populations of *Ae. aegypti* from these regions have originated from east and west African genetic clusters. Paupy et al (2012), using *COI* and *ND4*, also found two major genetic clades of *Ae. aegypti* in Brazil. The authors observed a significant genetic structure among geographic populations suggesting a recent re-expansion of *Ae. aegypti* in Bolivia and the existence of two genetic groups, one dominant group recovered throughout Bolivia, and the second (related to West African *Ae. aegypti* specimens) restricted to the rural localities in South Bolivia. Sousa et al. (2017) investigated the genetic differentiation of this vector in the Brazilian state of Maranhão, based on the mitochondrial *ND4* gene and observed also the existence of two distinct clades. All the above studies demonstrated that most populations of *Ae. aegypti* are genetically structured in different regions of the world and confirmed the presence of two distinct genetic groups in Brazil.

In the present study, we investigated the genetic variability and population structure of *Ae. aegypti* from 14 Brazilian cities, with sequences of the *CO1*gene (Barcode region), a more conserved region than the other mtDNA markers, used for the population studies of *Ae. aegypti* in Brazil, to evaluate the existing genetic groups/ lineages of this vector circulating throughout the country.

4.2.2. Materials and Methods

Sample collection

Aedes aegypti samples were collected from 14 Brazilian cities, covering the major regions of the country, as follows: eight from the states of Brazilian Amazon [Itacoatiara (IT), Manaus (MA), and Novo Airão (NA), state of Amazonas; Boa Vista (BV), state of Roraima; Rio Branco (RB), state of Acre; Porto Velho (PV) and Guajará-Mirim (GM), state of Rondônia and Macapá (MP), state of Amapá], one from the central west region of the country [Cuiabá (CU), state of Mato Grosso], two from the northeastern region [Campina Grande (CG), state of Paraiba; Teresina

(TS), state of Piauí], and three from the southeastern region [Araçatuba (AR), São José do Rio Preto (RP) and Taubaté (TA), state of São Paulo] (Figure 1). The sample of Foz do Iguaçú, however, was not included in this study because no individual amplified for the Barcode region, despite several attempts. The details regarding specimen collection, geographical coordinates, year of collection and sample size for each site are shown in Table 1.

The mosquitoes were collected from fields as larva or adults. Larvae were collected from various artificial recipients near human dwellings and also by using oviposition traps (ovitraps) for 2 to 7 days, depending on the location. All the samples were collected from multiple breeding sites (15 – 20 breeding sites) per location to prevent sampling of related individuals. The breeding sites were widespread over different neighborhoods in each city, the geographic distances between sites ranging from 20m to 5 km (Scarpassa et al., 2008). To further avoid the chances of sampling related individuals, only 1 specimen from each breeding site were used in the analyses. The specimens were morphologically identified using the taxonomic key of Forattini (2002), preserved in 95% ethanol and stored in freezer –20º C or stored dry in freezer -80º C, until DNA extraction.

Figure 1. Collection sites of *Aedes aegypti*. AR: Araçatuba, RP: São José do Rio Preto, TA: Taubaté, RB: Rio Branco, PV: Porto Velho, GM: Guajará-Mirim, CU: Cuiabá, CG: Campina Grande, TS: Teresina, BV: Boa Vista, IT: Itacoatiara, MA: Manaus, RB: Rio Branco, MP: Macapá, NA: Novo Airão.

DNA extraction, PCR and sequencing

Genomic DNA was extracted individually from the larva or adult mosquitoes using phenol and chloroform method (Sambrook and Russell, 2001), with minor modifications. The COI barcode region was amplified for 161 individuals using the primers LCO 1490 and HCO 2198 (Folmer et al., 1994) and according to the thermocyclic condition described by Folmer et al. (1994). All PCR reactions included negative controls. The PCR products stained with GelRed ™, visualized in 1% agarose gels under UV light to verify their expected size and then purified with PEG precipitation (20% polyethylene glycol 8000/2.5 M NaCl). Sequencing reactions were carried out for both reverse and forward DNA strands with Big Dye Terminator Kit ®and analyzed in an automated sequencer ABI 3130 xl Genetic Analyzer (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA), available at Instituto Nacional de Pesquisas da Amazônia, in Manaus, Brazil.

Data analyses

All the DNA sequences were automatically aligned with the help of ClustalW tool (Larkin et al., 2007) and manually edited with BIOEDIT v. 7.2.5 (Hall, 1999) with the help of Chromas Lite®. The consensus sequences generated a fragment size of 646 base pairs (bp) which was translated into amino acids to verify the presence of stop codons, pseudo genes and Numts. All the sequences were confirmed using Basic Local Alignment Search Tool (BLAST) platform available at the National Centre for Biotechnology Information (NCBI) website (http://blast.ncbi.nlm.gov/Blast.cgi).

The intra-population and overall genetic diversity indices such as, number of haplotypes (*NH*), number of transitions/transversions (*T*s/*T*v), number of segregating sites (*S*), number of private sites (*NPS*), average number of nucleotides differences (*K*), haplotype diversity ($H \pm$ SD) and nucleotide diversity ($\pi \pm$ SD), with their respective standard deviations were estimated in DnaSP v. 5.10 (Librado and Rozas, 2009) and Arlequin v. 3.1 (Excoffier et al., 2007). The neutrality tests of Tajima's *D*, Fu and Li's *D* and *F* and Fu's *Fs* were estimated in DnaSP v. 5.10 (Librado and Rozas, 2009) and Arlequin v. 3.1 (Excoffier et al., 2007). These neutrality tests were used to test population equilibrium and are expected to yield high negative values for expanding populations.

The genealogical relationships between haplotypes were evaluated by constructing a minimum haplotype network, using the median-joining method implemented in the Network v. 4.6 software, available at http://www.fluxus-engineering.com (Bandelt et al., 1999). Two sequences of *Aedes albopictus* were used as outgroups.

The Bayesian population mixture analysis was implemented in BAPS v. 5.3.1 (Bayesian Analysis of Population Structure) (Corander et al., 2008) to identify the population genetic grouping. First a mixture analysis was carried out and the results were used in the admixture analysis, the number of groups (k) was set from 2 to 10.

Genetic differentiation (pairwise F_{ST} values), gene flow (*Nm* values), hierarchical analysis (AMOVA) and Mantel test were estimated using Arlequin v. 3.1 (Excoffier et al., 2007). AMOVA analyses were performed among all samples (non-grouped populations) as well as the dataset was partitioned according to the geographic regions and subsequent AMOVA was performed among all the groups at various hierarchical levels. Isolation-by-distance (IBD) was estimated by Mantel's test, using the correlation between genetic and geographic distances by the regression of F_{ST} /(1-F_{ST}) on the natural logarithm (ln) of straight-line geographic distance, and the significance level was tested using 10,000 permutations. Straight-line geographic distances between the sites were obtained using Google Earth and GPS. The Bonferroni correction was applied to multiple comparisons (Holm, 1979)

The dataset was tested for saturations levels with DAMBE (Xia and Xie, 2001) This analysis detects the presence of any saturation between transition and transversion rates in relation to genetic distances. The phylogenetic relationships were inferred using the neighbor-joining (NJ) in MEGAv.6 (Tamura et al., 2013) and maximum likelihood (ML) in Garli v.0.95 (Zwickl, 2006), and Bayesian inference (BI) in MrBayes v.3.2.5 (Ronquist et al., 2012). The nucleotide substitution model Kimura-2 Parameters (K-2P) (Kimura 1980) were used in the NJ analysis with 2,000 replicates, whereas the ML and BI analyses were performed using the nucleotide substitution model GTR+G (Hasegawa et al., 1985) previously selected with the Akaike Information Criterion (AIC) in the jModelTest (Darriba et al., 2012). We also aligned sequences of this study with the sequences of *Ae. aegypti* deposited in the GenBank (Paupy et al., 2012), with accession numbers JQ926685 (Vietnam), JQ926688 (Cambodia), JQ926691 (Thailand), JQ926704 (Tanzania), JQ926702 (Cameroon), JQ926701 (Venezuela), JQ926698 (Mexico), JQ926696 (Martinique), JQ926693 (Cote D'Ivoire), JQ926684 (USA),

JQ926683 (Bolivia) and JQ926682 (Bolivia); and BI analysis was performed using the nucleotide substitution model HKY + I + G (Hasegawa et al., 1985) previously selected with the Akaike Information Criterion (AIC) in the jModelTest (Darriba et al., 2012). This model assumes variable base frequencies, symmetrical substitution matrix over the sites following a gamma distribution rate variation among sites. In the phylogenetic relationship using the ML, the branch supports (bootstrap) were assessed with 2,000 replicates. In the BI analyses, two simultaneous independent runs of the Markov Chain Monte Carlo (MCMC) were performed for 5 million generations, while sampling every 1,000 generations with a burn-in of 25%. Posterior probabilities (BPP) were used to assess nodal support. The intra and inter-samples genetic distances of the fourteen localities and among the two groups were calculated using Mega v.6.0 (Tamura et al., 2013), based on the K-2P evolutionary model.

The bGMYC (a Bayesian implementation of the general mixed yule-coalescent model for species delimitation) analysis was carried out to test the preliminary hypothesis of species/ lineage delimitation (Reid and Carstens, 2012). The bGMYC package was used in the R (R Core Team, 2017) program. The analysis was generated based on the *CO1* barcode, using multiple ultra-metric trees calculated by the Bayesian MCMC analysis of 100 post-burn seeds in Beast 2 (Bouckaert et al., 2014) with the previously selected GTR + G model. Nodal support levels (later traits) were included in the branches of the tree, and the bars at the right of the dendrogram represent clusters of species/hypothesized lineages, based on the Bayesian posterior probability of co-specificity. To delimit species, it is necessary to specify a probability threshold above which the individuals will be considered heterospecific; we adopt a threshold of $P = 0.95$ (posterior probability).

4.2.3. Results

Our dataset consisted 161 sequences of a 646 bp fragment of the DNA barcode (*CO1*) of *Ae. aegypti* from 14 sampling sites of Brazil (Figure 1; Table 1). The amino acid translation revealed no stop codons ensuring absence of pseudogenes or nuclear mitochondrial DNA sequences (Numts). Analysis of the transition and transversion rates in relation to the genetic distances (K-2P) did not reveal saturation (Figure S1), suggesting that this dataset is informative for the phylogenetic analyses. All sequences had 19 variable (polymorphic) sites, 15 of which were parsimoniously informative. Out of 116 base substitutions identified, 112 (96.55%) were

transitions and only four (3.45%) were transversions. The mean composition of the nucleotides was A = 27.8%, C = 17.3%, G = 16% and T = 38.9%, with high content of A+T = 66.7%, especially at third codon position (87.3%).

Table 1. Collection sites and haplotype frequency observed for the COI gene (Barcode region) in 161 *Aedes aegypti* samples from Brazil

Localities, States	**Abbreviation**	**Coordinates (Lat./Long.)**	***N***	**Haplotype frequency**
Rio Branco, Acre	RB	10° 1'25.00"S 67°49'19.63"W	15	***H1*** (13), *H13* (2)
Porto Velho, Rondônia	PV	8°45'39.68"S 63°54'1.59"W	9	***H1*** (9)
Guajará Mirim, Rondônia	GM	10°47'21.27"S 65°19'48.09"W	11	***H1*** (10), H2 (1)
Boa Vista, Roraima	BV	2°46'56.04"N 60°41'2.67"W	11	***H10*** (9), *H11* (1), H12 (1)
Itacoatiara, Amazonas	IT	3° 8'42.04"S 58°26'18.20"W	10	***H1***(6), *H10* (2), *H11* (2)
Manaus, Amazonas	MA	3°11'34.36"S 60° 1'23.25"W	17	***H1*** (8), *H10* (7), H18 (1), H19 (1)
Novo Airão, Amazonas	NA	2°38'12.57"S 60°56'38.95"W	10	***H10*** (9), *H11* (1)
Taubaté, São Paulo	TA	23° 2'26.05"S 45°33'23.08"W	14	***H4*** (13), H20 (1)
São José de Rio Preto, São Paulo	RP	20°52'20.56"S 49°22'25.60"W	13	***H1*** (5), *H3* (1), *H4* (7)
Araçatuba, São Paulo	AR	21°13'42.35"S 50°27'5.15"W	12	*H1* (4), ***H4*** (8)
Cuiabá, Mato Grosso	CU	15°40'40.56"S 56° 5'42.24"W	11	*H1* (2), ***H4*** (7), *H13* (1), H14 (1)

Campina Grande, Paraíba	CG	7°15'5.88"S 35°52'42.45"W	13	***H1*** (5), *H3* (2), *H4* (3), H5 (1), H6 (1), H7 (1)
Teresina, Piauí	TS	5° 7'14.19"S 42°48'18.58"W	5	*H1* (1), H15 (1), H16 (1), **H17** (2)
Macapá, Amapá	MP	0° 1'18.79"N 51° 4'11.36"W	10	***H1*** (7), *H3* (1), H8 (1), H9 (1)
Total			161	

In parentheses, number of individuals observed for each haplotype. The italics haplotypes are shared among samples. The bold haplotypes indicate the most frequent haplotype for the respective localities. *N* number of specimens sequenced.

Of the 20 haplotypes observed, four (20%) were shared among all 14 sample sites and 16 (80%) were singletons and exclusive for the specific sample site (Table 1). Campina Grande (CG) had the largest number of haplotypes (6), followed by Teresina (TS), Manaus (MA), Cuiabá (CB) and Macapá (MP), all with four haplotypes. The lowest number of haplotypes (1) was observed in Porto Velho (PV). The haplotype H1 was the most frequent and it was shared by all populations, except Boa Vista (BV), Novo Airão (NA) and Taubaté (TA). H4, the second most frequent haplotype, was shared among the individuals from Taubaté (TA), São José de Rio Preto (RP), Araçatuba (AR), Cuiabá and Campina Grande (CG). H10 was shared among the specimens from BV, IT, MA and NA (northern region). H11 was shared among individuals from BV and IT, NA (both from the State of Amazonas); but it was absent in Manaus (MA). Interestingly, the individuals from MA did not share haplotypes with NA, both situated in the state of Amazonas and geographically close to each other. Also, the samples RB, BV and TA did not share any haplotypes among them, and PV did not share haplotypes with NA, BV and TA (Table 1; Figure 2).

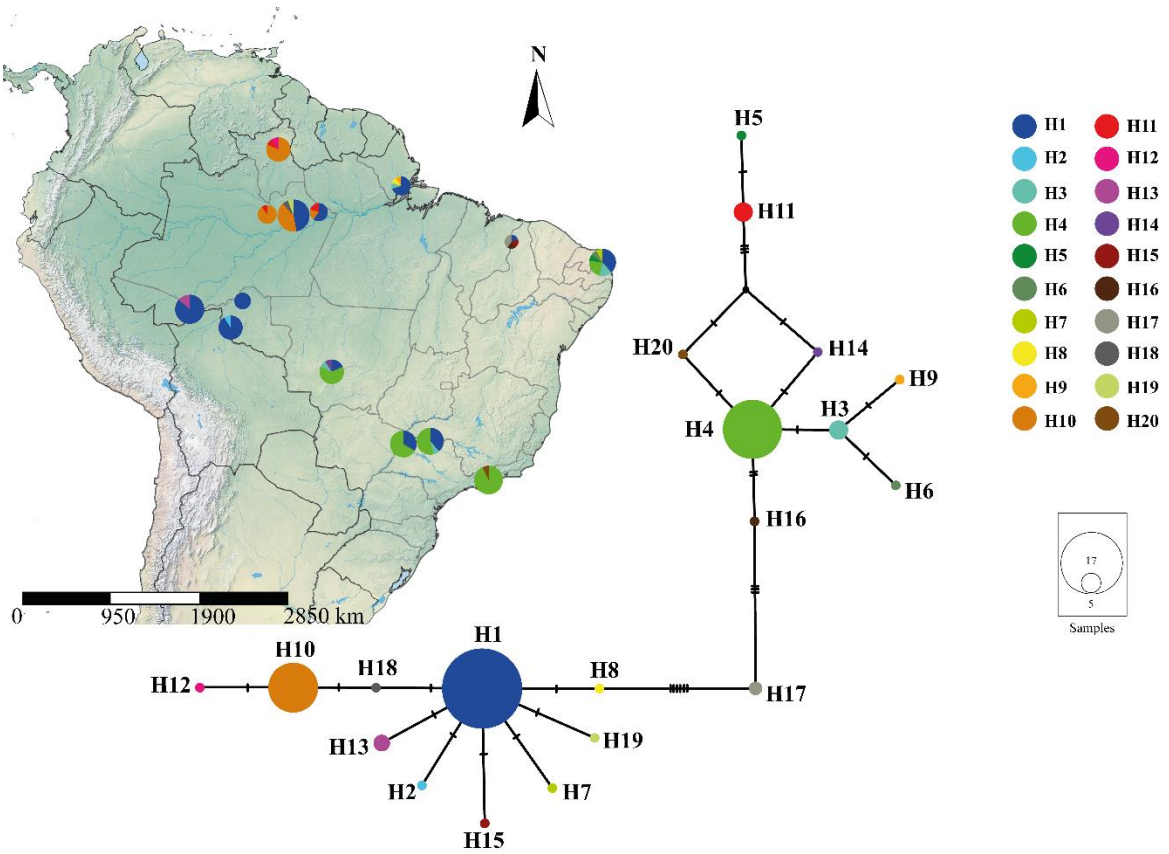

Figure 2. Haplotype network generated with 95% confidence and their geographic distribution. Each haplotype is represented by a different color. The circle size indicated the frequency of the individuals observed in each haplotype.

The genetic population structure analysis using BAPS indicated that the optimal number of clusters was 5 (K= 5, logML= -279.5051). The BAPS bar plot shows a considerable level of admixture among the clusters (Figure 3) and divided the fourteen samples into 5 haplotypes clusters: Cluster 1 (H4, H3, H9, H6, H20 and H14), Cluster 2 (H 16 and H17), Cluster 3 (H1, H2, H7, H8, H13, H15 and H19), Cluster 4 (H10, H12 and H18) and Cluster 5 (H5 and H11). The haplotype network (Figure S3) exhibited two main genetic groups, with certain levels of subdivision within them. The samples from northern (BV, IT, MA and NA) and southwestern Amazon (RB, PV and GM) were clearly clustered in separate groups. The southeastern (AR, RP and TA), central west (CU) and northeastern (CG and TS) samples seem to be genetically closer to each other than the northern clusters; but, the northeastern (CG and TS) samples had significant level of isolation which corresponds to the presence of exclusive haplotypes (H15, H16, H17

observed in TS and H5, H6, H7 observed in CG). In spite of being situated in north of Brazil, the samples of MP seem to be genetically more isolated from the remaining samples of the northern region (BV, IT, MA and NA).

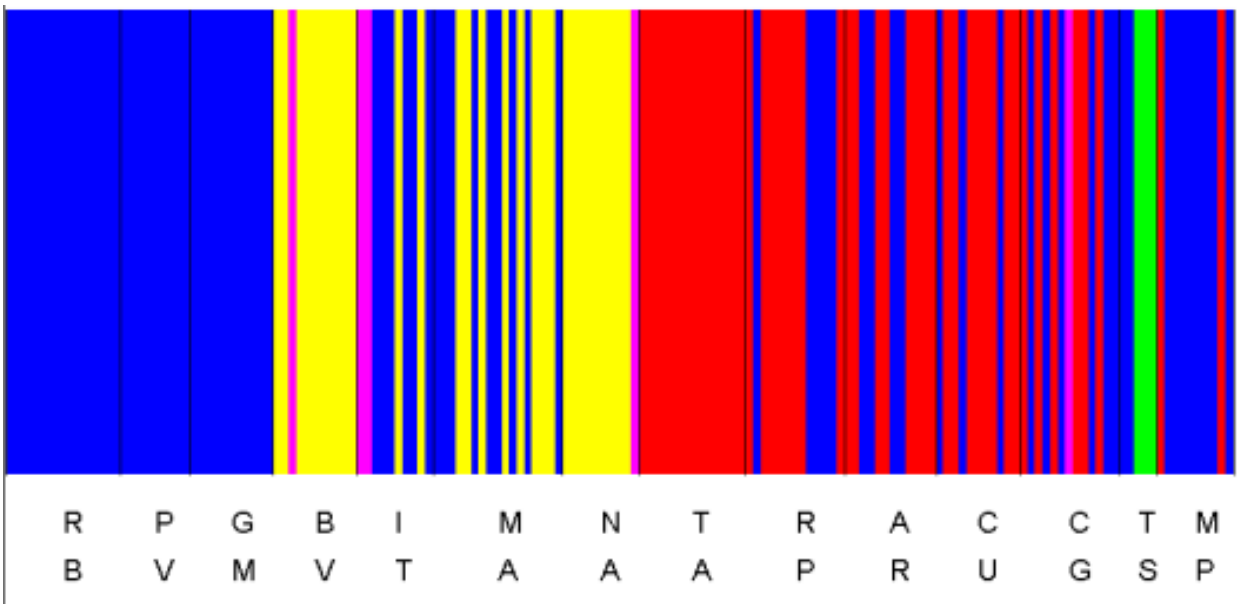

Figure 3. Bayesian genetic cluster analysis of BAPS with K = 5 clusters. The height of each color represents the probability of assignment to a specific cluster. The black lines within the plots indicate population limits. AR: Araçatuba, RP: São José do Rio Preto, TA: Taubaté, RB: Rio Branco, PV: Porto Velho, GM: Guajará-Mirim, CU: Cuiabá, CG: Campina Grande, TS: Teresina, BV: Boa Vista, IT: Itacoatiara, MA: Manaus, RB: Rio Branco, MP: Macapá, NA: Novo Airão.

The BI (Figure S2) and NJ (Figure S3) trees generated almost identical topologies and, as observed in the haplotypes network, retrieved two main genetic groups. The ML tree also generated two genetic groups (Groups 1 and 2); however, the H17 (TS) was included in a separate branch (basal). In the BI tree, the apical clade (BPP: 0.95) clustered most of the samples from the Amazonian states (RB, PV, GM, BV, MA, NA, IT) and the basal clade clustered most of the specimens from southeastern and northeastern regions of Brazil, except the well supported group of some of the samples from BV, IT and NA (H5 and H11). The BI tree with the sequences this study and those of GenBank (Fig. 4) also retrieved two main genetic groups. The DNA sequences from East (Cameroon and Ivory Coast) and West (Tanzania) Africa were observed to be in separate

groups. The sequences from Bolivia, French Guiana, USA, Asia (India, Cambodia, Thailand and Vietnam), Australia and East Africa (Cameroon and Ivory Coast) were grouped together (Group1) with the samples from the Amazonian states (RB, PV, GM, BV, MA, NA, IT) of Brazil, whereas the sequences from Martinique (Caribbean), Mexico and Tanzania (West Africa) were grouped together (Group 2) with most specimens from the southeastern and northeastern regions of Brazil. H17 (represented by 2 individuals from TS) diverged from the rest of the second group (BPP 0.68), as was observed in Figure S2.

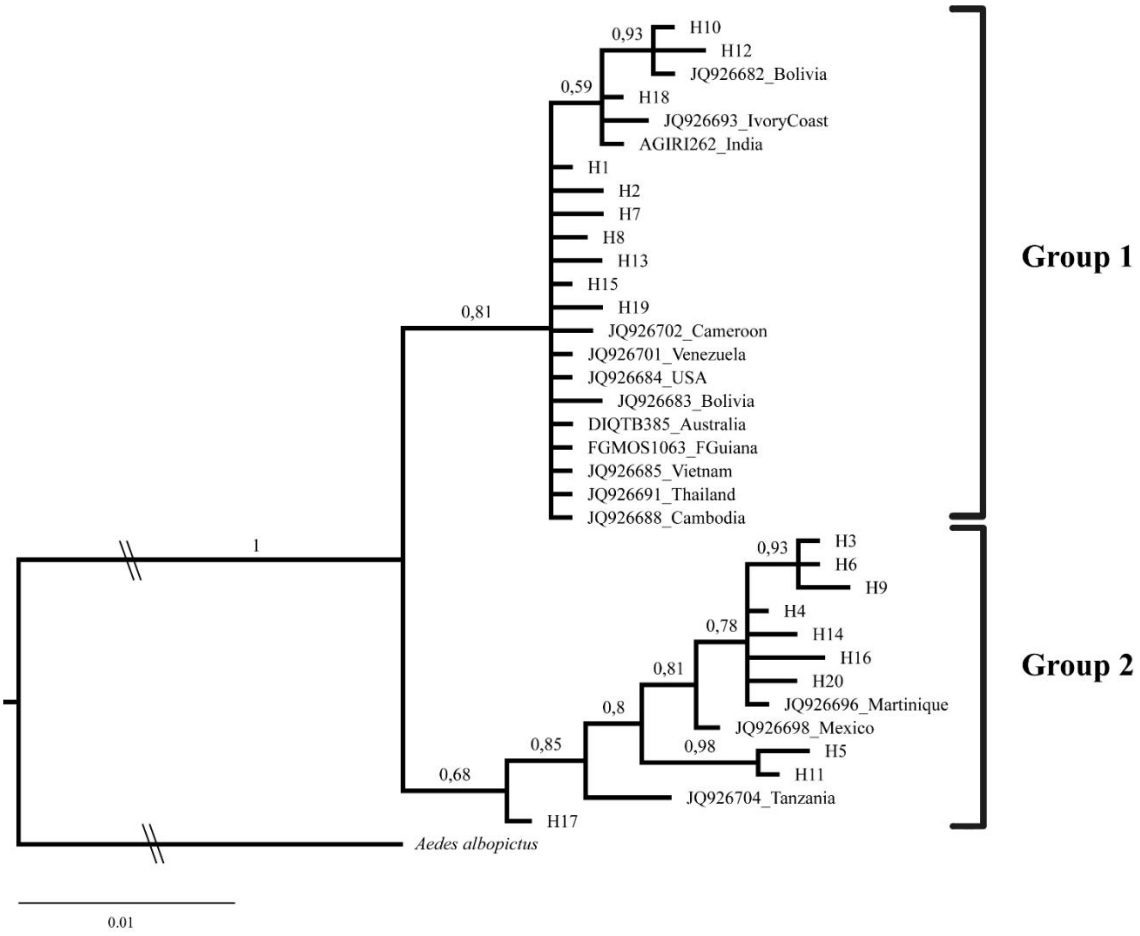

Figure 4. Bayesian Inference (BI) tree inferred for the 20 haplotypes from Brazil and other sequences accessed from GenBank, with HKY + I +G evolutionary model. Values above each branch represent the BPP. The two groups are identified as Group 1 and Group 2. *Aedes albopictus* was used as the outgroup.

The bGMYC results are summarized in Figure S5. The colored matrix compares individual specimens (mtDNA sequences), with colors corresponding to the posterior probability and indicating that they are conspecific. Only *Aedes albopictus* (outgroup) was delimited as heterospecific with $P > 0.95$. The other haplotypes were separated into four clades, but, with $P < 0.95$ and; hence, could not be delimited as heterospecific.

The average intrapopulation genetic distances (Table S1) was 0.51%, ranging from 0 (PV) to 1.11% in the samples of CG. The interpopulation distance varied from 0 (between populations of GM, PV and RB) to 1.9% (between populations of TA and PV, GM, RB) (Table S1). The intergroup genetic distance was 1.8% and the intragroup distance was observed to be 0.34% and 0.63% for Groups 1 and 2, respectively (Table S2).

Haplotype and nucleotide diversities were used as measures of genetic diversity in *Ae. aegypti* (Table 2). Total haplotype diversity (H) was 0.729, ranging from 0.000 in PV to 0.900 in TS. The highest nucleotide diversity was observed in the samples of CG (0.01072), followed by the samples of RP (0.00957). Maximum number of haplotypes (6) were also observed in the samples of CG. Table 2 also shows the neutrality tests for the 14 samples. The neutrality test for the samples of BV (Tajima's *D* negative and significant) and NA (negative and significant values of Tajima's *D* and Fu & Li's *D*) suggested deviations from the neutral model, probably indicating recent population expansion, positive selection or selective sweep. The positive and significant values of Fu & Li's *D* (IT, RP, CU, MP) and Tajima's *D* (RP) suggest a possible balancing selection, population sub-division or recent bottleneck in these samples and; therefore, the molecular polymorphism cannot be explained by the neutral model.

Table 2. Intra-population genetic diversity and neutrality tests estimated for the COI gene in *Aedes aegypti* samples from Brazil

Samples	***N***	***Ts/Tv***	***NH***	***S***	***NPS***	***K***	***H ± SD***	***π ± SD***	***Tajima's D***	***Fu and Li's D***	***Fu and Li's F***	***Fu's Fs***
Rio Branco	15	1/0	2	1	0	0.2476	0.248±0.131	0.0004±0.0002	-0.39883	0.70104	0.4775	0.133
Porto Velho	9	0/0	1	0	0	0.0000	NC	NC	NC	NC	NC	NC
Guajará-Mirim	11	0/1	2	1	0	0.1818	0.182±0.144	0.0003±0.0002	-1.1285	-1.2895	-1.3992	-0.410
Boa Vista	11	11/0	3	11	0	2.1454	0.345± 0.172	0.0033± 0.0024	-1.8448*	-2.0747	-2.2809	2.342
Itacoatiara	10	13/0	3	13	0	4.9778	0.622±0.138	0.0077±0.0029	0.3787	1.4995**	1.3754	4.859
Manaus	17	2/1	4	3	0	1.1618	0.640±0.073	0.0018±0.0002	0.8757	-0.0627	0.2156	0.152
Novo Airão	10	11/0	2	11	0	2.2000	0.200±0.154	0.0034±0.0026	-1.9443*	-2.2759**	-2.4683	4.292
Taubaté	14	1/0	2	1	0	0.1428	0.143±0.119	0.0002±0.0002	-1.1552	-1.3975	-1.5139	-0.595
São Jose de Rio Preto	13	12/0	3	12	0	6.1795	0.603±0.088	0.0096±0.0015	2.4416**	1.4649*	1.9657	7.010
Araçatuba	12	12/0	2	12	0	5.8182	0.485± 0.106	0.0090± 0.002	1.9499	1.4693*	1.8116	9.242
Cuiabá	11	13/0	4	13	0	5.5636	0.600±0.154	0.0086±0.0025	1.1096	1.4899**	1.5769	3.712
Campina Grande	13	15/1	6	16	0	6.9231	0.821±0.082	0.0107±0.0011	1.4399	0.4308	0.7978	2.279
Teresina	5	10/1	4	11	0	5.8000	0.900±0.161	0.0089±0.0022	0.7088	0.7088	0.7495	0.726
Macapá	10	11/0	4	11	0	3.9111	0.533±0.180	0.0061±0.0026	0.0262	1.4721*	1.2537	2.256
Total	161	8/0.286	20	19	0	5.8812	0.729±0.025	0.0091±0.0005	2.0394	-0.3147	0.7161	0.245

N sample size. *T*s/*T*v transitions/transversions. *NH* Number of haplotypes. *S* Number of segregating sites. *NPS* Number of private sites. *K* Average number of nucleotides differences. *H* ± SD and *π* ± SD Haplotype and nucleotide diversities, respectively, with respective standard deviations (SD), *NC* not calculated. * $P < 0.05$; ** $P < 0.02$

AMOVA revealed significant genetic variation for almost all the hierarchical levels (Table 3), indicating high level of genetic structuring. Significant genetic difference (Φ_{ST} = 0.5010) was observed among the 14 samples (no grouping). Within the northern region, though 70.25% of variation was observed "within populations", a significant genetic differentiation was also observed among the populations. When the populations of northern and southeastern regions of the country were compared, highest significant genetic variation (62.1%) among the two geographical groups and lowest genetic variation within population (29.23%) was observed. When we divided all the samples into four geographic regions (North Amazon, Southwest Amazon, Northeastern Brazil and Central & Southeastern Brazil), the percentage of genetic variation "among groups" and "within populations" were significant and almost equal (47.06% and 45.56%, respectively). These results were corroborated with the pair-wise F_{ST} (P = 0.05) and *Nm* values (Table 4). They showed similar significant genetic differentiation among groups of populations (as seen in AMOVA), with pairwise values ranging between -0.05 and 0.99 and varied levels of gene flow (*Nm* ranging from infinity to 0.003). The Mantel test results indicated a weak but, significant correlation between genetic and geographic distances (r = 0.442, P = 0.0003), suggesting that 44.2% of genetic differentiation can be explained by the IBD model.

Table 3. Hierarchical analysis (AMOVA) of the genetic variation in the *Ae. aegypti* samples

Groups of Samples	Source of variation	Degrees of freedom	Percentage Variation (%)	Fixation index
No Grouping (All) AR, RP, CU, CG, TA, IT, MA, NA, BV, RB, PV, GM, TS, MP	Among populations	13	50.10	$\Phi st = 0.5010^{***}$
	Within populations	114	49.90	
Within the northern region (IT, MA, NA, BV, RB, PV, GM, MP)	Among populations	7	21.75	$\Phi st = 0.2975^{***}$
	Within populations	85	70.25	
Between northern (IT, MA, NA, BV, RB, PV, GM, MP) and southeastern (AR, RP, TA) regions	Among groups	1	62.10	$\Phi ct = 0.6210^{**}$
	Among populations within groups	9	8.68	$\Phi sc = 0.2290^{***}$
	Within populations	160	29.23	$\Phi st = 0.7077^{***}$
Between north (BV, IT, MA, NA, MP) and south (RB, PV, GM) of Amazon	Among groups	1	27.07	$\Phi ct = 0.2707$
	Among populations within groups	6	14.56	$\Phi sc = 0.1996^{**}$
	Within populations	105	58.37	$\Phi st = 0.4163^{***}$

Between northern (IT, MA, NA, BV, RB, PV, GM, MP) and north-eastern (CG, TS) regions	Among groups	1	32.94	Φct =0.3294*
	Among populations within groups	8	11.90	Φsc =0.1775**
	Within populations	119	55.16	Φst =0.4484***
Between northern/north-eastern and central/southeastern regions	Among groups	1	51.66	Φct =0.5166**
	Among populations within groups	12	10.74	Φsc =0.2222***
	Within populations	197	37.59	Φst =0.6241***
Four Geographic Groups (1) North Amazon (MP, BV, IT, NA, MA) (2) South Amazon (RB, PV, GM) (3) North eastern Brazil (TS, CG) (4) Central and Southeastern Brazil (CU, AR, RP, TA)	Among groups	3	47.06	Φct = 0.4706**
	Among populations within groups	10	7.37	Φsc = 0.1393***
	Within populations	353	45.56	Φst = 0.5444***

See Table 1 for locality abbreviations. Significance test 10,000 permutations, Φ_{ST} fixation index within samples, Φ_{CT} fixation index between regions, Φsc fixation index among samples within regions. *** P = 0.00000 ± 0.00000. ** P < 0.01. * P < 0.05.

Table 4. Effective number of migrants (*Nm*) and genetic distances (*F*st values), above and below the diagonal, respectively, among the samples of *Aedes aegypti* from Brazil

	RB	**PV**	**GM**	**BV**	**IT**	**MA**	**NA**	**TA**	**RP**	**AR**	**CU**	**CG**	**TS**	**MP**
RB		20.8929	12.6655	0.2592	1.9093	1.0040	0.2508	0.0082	0.3505	0.2621	0.2000	0.5957	0.2963	2.6460
PV	0.02337		Inf.	0.2987	2.7347	1.1783	0.2931	0.0037	0.4514	0.3365	0.2547	0.7852	0.4063	4.6753
GM	0.03798	0.01957		0.2883	2.2971	1.1409	0.2831	0.0067	0.4137	0.3092	0.2344	0.7094	0.3680	3.7141
BV	**0.65860***	**0.62603***	**0.63429***		3.4425	1.2220	Inf.	0.0488	0.5456	0.4357	0.3546	0.8903	0.4516	0.9435
IT	0.38882	0.15458	0.17876	0.12682		5.1279	3.8090	0.1300	1.3900	1.0220	0.7860	3.3153	2.9210	Inf.
MA	**0.67304***	0.29792	0.30471	**0.29036***	0.08884		1.2427	0.0309	0.4093	0.3132	0.2429	0.6947	0.3726	2.2630
NA	**0.66596***	**0.63046***	**0.63850***	-0.10131	0.11604	**0.28691***		0.0482	0.5815	0.4629	0.3738	0.9575	0.4845	1.0022
TA	**0.98385***	**0.99268***	**0.98686***	**0.91111***	**0.79361***	**0.94182***	**0.91211***		0.9385	1.2184	1.7173	0.5795	0.1509	0.1066
RP	**0.58791***	0.52555	**0.54725***	**0.47821***	0.26455	**0.54987***	**0.46231***	0.34758		Inf.	Inf.	Inf.	6.3680	1.5074
AR	**0.65611***	0.59777	**0.61791***	**0.53438***	0.32851	**0.61484***	**0.51927***	0.29097	0.07826		Inf.	Inf.	3.6804	1.0397
CU	**0.71425***	**0.66252***	**0.68081***	**0.58505***	0.38882	**0.67304***	**0.57223***	0.22550	0.05144	0.07943		12.9625	2.2359	0.7556
CG	**0.45632***	0.38903	0.41344	0.35965	0.13105	**0.41852***	0.34306	**0.46319***	0.04253	0.00645	0.03714		22.1936	4.0211
TS	**0.62793***	0.55170	0.57607	0.52545	0.14616	**0.57302***	0.50790	**0.76821***	0.07280	0.11961	0.18276	0.02203		2.9317
MP	0.15893	0.09661	0.11865	**0.34638***	-0.01010	0.18096	0.33285	**0.82427***	0.24908	0.32474	0.39823	0.11059	0.14570	

Bold values indicate significant genetic distances. *$P < 0.003$, after Bonferroni correction; Inf. = infinity. See Table 1 for locality abbreviations.

4.2.4. Discussion

The present study revealed a significant genetic differentiation and moderately high genetic variability among the samples of *Ae. aegypti* throughout Brazil. The samples from Campina Grande (CG) showed the highest *K*, *H* and π values (indicating the presence of genetically distant haplotypes in sympatry); whereas PV, followed by RB and GM, showed the lowest values, indicating reduced gene flow with other localities or a substantial recent bottleneck effect due to intense vector control by insecticides (Scarpassa et al., 2008). However, samples from Teresina (TS), appeared to be genetically isolated from the other samples of Brazil; including the samples from CG, which is geographically closer to TS. This may indicate genetic isolation of local vector population, probably attributed to under sampling or local mutational forces, including environmental factors (long and dry summer periods) and specific vector control measures of that locality (over-usage of insecticides leading to extinction and recolonization). The samples from Teresina were genetically studied for the first time, so they could not be compared with the results from earlier studies. However, with microsatellite markers (Maitra et al. unpublished data), we have observed similar divergence of TS with other samples from Brazil.

Tajima's *D*, and Fu and Li's *D* and *F* neutrality tests were mostly positive and significant for the samples of RP, AR, CU and MP, indicating a genetic sub-division supported by the existence of 2 haplotype groups within them. The samples of BV and NA, on the other hand, were negative and significant for these neutrality tests, suggesting a recent population expansion. This may be due to colonization, as BV may be treated as a possible entry point due to its proximity to Venezuela, Guianas and Suriname, as suggested by Osanai et al. (1983), or rapid urbanization, as NA is gaining popularity as a tourist city in Amazonas, increasing the genetic flow to other cities, in recent times. Although, Fu's *Fs* test did not generate any significant values, hence, none of the samples indicate a significant demographic expansion. Similar results have been observed by Scarpassa et al. (2008) (with *CO1* gene) as well as Lima Junior and Scarpassa (2009) (with *ND4* gene) for most of the Brazilian samples of *Ae. aegypti*.

The present study also revealed the existence of two distinct mtDNA groups of *Ae. aegypti* from 14 localities of Brazil, with 1.8% genetic distance among them, confirming the results of

similar previous works in Brazil (Bracco et al., 2007; Lima Junior and Scarpassa, 2009; Paduan et al., 2008; Scarpassa et al., 2008). However, five genetic clusters were identified by the Bayesian analysis and supported by the pairwise values of F_{ST} and AMOVA results, which showed significant genetic differentiation in most comparisons.

In this study, the haplotypes (20) were divided in two genetic groups. Although H1 (the most frequent haplotype observed) had the highest frequencies in the samples of RB, PV and GM (almost 100%), it was totally absent from the samples of BV, NA and TA; suggesting an absence of gene flow between BV and RB (see Table 1). Lima Junior and Scarpassa (2009) also observed very low level of gene flow between BV and RB, with ND4 marker.

H4 (the second most frequent haplotype), on the contrary, manifested the highest frequencies in the samples of TA (100%), RP, AR and CU (southeast and west central regions of Brazil), whereas it was totally absent from the samples of RB, PV, GM, BV, IT, MA, NA and MP (northern and south-western Amazon of Brazil), suggesting an absence of gene flow between them. This has also been supported by AMOVA with a significant level of variation among these groups. Similar results have been observed with microsatellite markers (Maitra et al. unpublished data), where the southwestern Amazon (RB, PV and GM) were observed to be genetically differentiated from northern Amazon (BV, IT, NA, MA and MP) and also from the southeastern Brazil (RP, AR and TA).

Although, with microsatellite markers (unpublished data), we have observed a genetic connection between TA and MA (Southeastern Brazil and Brazilian Amazon); which may be recent and can be supported by the fact that *Ae. aegypti* was reported first time in 1996 from Manaus (MA), much later than it was reported from the southeastern regions of Brazil. The connection between the Brazilian Amazon and southeastern Brazil has been implied by Vasconcelos et al. (1999), especially by the support of the history of re-invasion of *Ae. aegypti* (after the declared eradication in 1955), which is believed to have begun in southeastern Brazil, and have reached the central-west and northeastern regions by the mid-1980s, finally reaching all Brazilian states by 1998 (Figueiredo, 2003). Scarpassa et al. (2008) also observed a genetic connection between the samples of one of the neighborhoods of MA (Praça 14 de Janeiro) and TA.

The samples of MA used in this study were from a different neighborhood (Compensa), which did not share any haplotypes with the samples of TA, as was verified by Scarpassa et al. (2008).

However, H10 (the third most frequent haplotype observed) was exclusively observed in all the samples from northern Amazon (BV, IT, MA and NA), except MP. This haplotype sharing represents a profound and complex structure of genetic differentiation and gene flow among the samples from different geographical regions of the country. The samples from MP, in spite of being situated in the northern region of the country, portrayed genetic homogeneity with the samples of south-eastern and north-eastern regions of Brazil. Though, according to our analysis with microsatellites (unpublished data), MP was observed to be a different genetic group altogether (isolated from the samples of northern Brazil as well as samples from south-eastern Brazil). Kotsakiozi et al. (2017), Gloria–Soria et al. (2016) and Monteiro et al. (2014), reported similar results with microsatellites loci. These studies found highest differences between Marabá/Tucuruí/ Castanhal/Macapá/Santarém and remaining samples from Brazil. Altogether, it is possible to infer that the samples of *Ae. aegypti* from Belém, Santarém, Marabá and Tucuruí (state of Pará) and Macapá (state of Amapá) consist of a well differentiated genetic group from the remaining populations of Brazil. Lima Júnior and Scarpassa (2009), using the *ND4* gene, demonstrated that sample of Belém had the highest number of private haplotypes, where most them did not share any haplotype with those of the previous studies.

These findings suggest 3 possible events of introduction of *Ae. aegypti* in Brazil from following probable entry points: northwestern region (PV, RB, GM), extreme northern region (BV, which is near Pacaraima, which is situated at the border of Brazil and Venezuela) and southeastern region (Taubaté, which is near to the port of Santos, Rio de Janeiro and Espirito Santo). Our results also confirm the circulation of two distinct African genetic groups (East and West African) as observed by Scarpassa et al. (2008). Most of the specimens from northern regions of Brazil were observed to be well connected with Venezuela, Bolivia, French Guiana, Asia, Australia, USA and East African groups of *Ae. aegypti*. On the other hand, the samples from south-eastern and north-eastern Brazil seem to be more connected with the Caribbean, Mexican and West African groups of the vector. Gonçalves da Silva et al. (2012) also observed similar gene flow networks between southeastern Brazil and Mexico-North America, along with connections between Mexico/North

America and Venezuela, Venezuela and the Brazilian Amazon, and the Brazilian Amazon/Peru and southeastern Brazil. There is good evidence in support of these genetic connections; for instance, land connection between Venezuela and Boa Vista (northern Brazilian Amazon) was suggested by the presence of dengue serotypes in Boa Vista, previously only recorded in Venezuela (Codeço et al., 2009; Lourenço-de-Oliveira et al., 2004). Mousson et al. (2005) also observed that the specimens of *Ae. aegypti* from Boa Vista were genetically related to those from Guinea and Ivory Coast (West Africa). Failloux et al. (2002) observed high level of genetic similarity between the populations of *Ae. aegypti* from French Guiana and Southeast Asia, suggesting an alternative route of introduction of this vector in the Americas, via Asia. Colonization from West Africa could have occurred with the intensive shipping trade between West Africa and ports of South and Central America during the late 15th Century to early 19th Century (Powell and Tabachnick, 2013; Tabachnick, 1991) also migration from East Africa could have been possible via historic trading routes to India (Powell and Tabachnick, 2013). Moreover, an alternative or additional migration among these continents could have occurred via a direct route for spices and silver trade, which linked Manila (Philippines) to Acapulco (Mexico), Panama and Lima (Peru) between 1565 and 1815 (Giráldez, 2015).

Summarizing the results obtained in this study, we have observed a complex gene flow pattern among the populations of *Ae. aegypti* throughout Brazil, and it will be an over simplification of observed results, if we try to hypothesize a single pattern of introduction, colonization or gene flow network among these populations. Actually, our results indicate a strong influence of genetic factors (including mutation, migration, extinction and re -colonization), environmental factors (local climatic conditions of the specific population, such as long dry season, long rainy season or long winter), human traffic and goods dispersal (poor/well connectivity with other big urban centers) and local factors (specific to that locality, such as, the type and magnitude of implemented vector control, sanitary hygiene, population density) on the population dynamics of *Ae. aegypti*. Every single population of this vector seem to have undergone different mutational changes and hence, more intensive and fine-scale population studies should be encouraged to assess the population structure of this vector, because every population may attribute to different levels of vectoral competence for arboviruses, and identifying the difference in the population

genetic structure this mosquito is essential to formulate more effective and integrated control measures.

Acknowledgements

This study was funded by MCTIC/INPA (Process number: PRJ 12.311) and by grants from Amazonas Research Foundation (FAPEAM, process number**:** 3111/2012) and from PRÓ-EQUIPAMENTO-CAPES to Vera Margarete Scarpassa. We thank Renata da Silva Leandro, Marllon Rinaldo de Lima Andrade, Steffany Larissa da Silva Galdino for the field collections from Campina Grande; Dr. Raimundo Nonato Picanço Souto and Dr. Fred Júlio Costa Monteiro for the field collections from Macapá; and MSc. Gersonval L.S. Monte for the field collections from Teresina. We also thank MSc. Antonio Saulo Cunha-Machado and Dr. Ronildo B. Alencar, for helping with the analyses and maps.

Additional Information

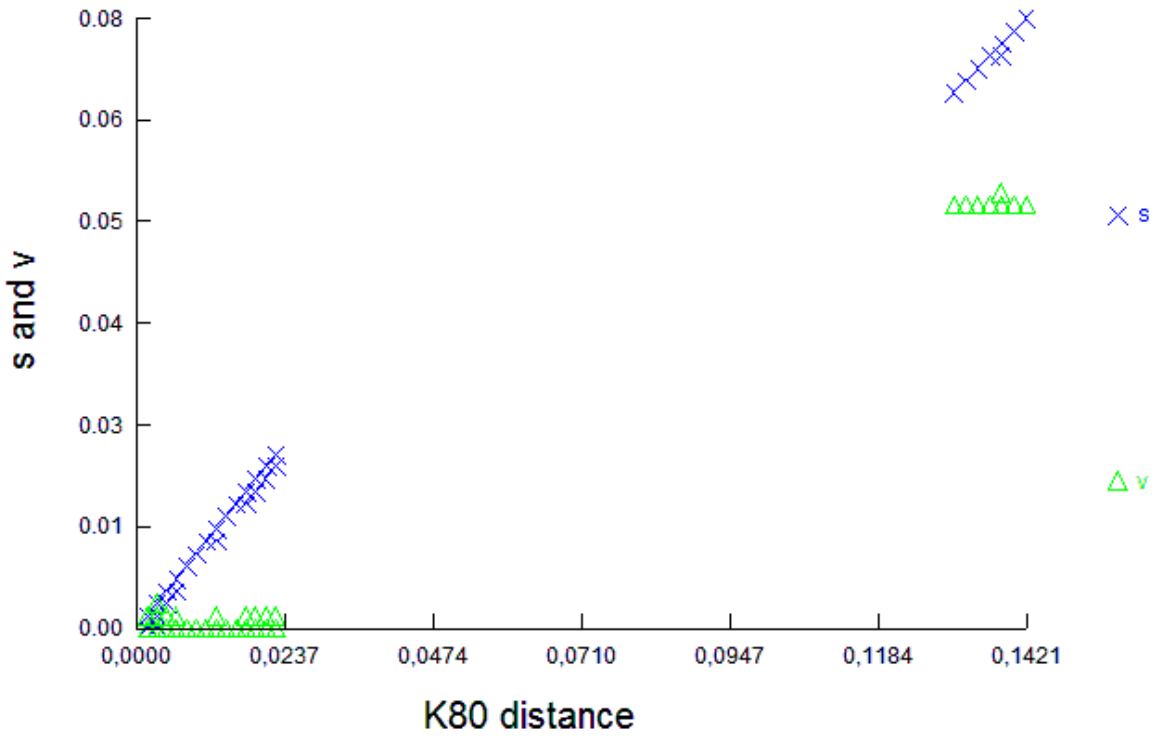

Fig S1. The graph generated by DAMBE. No saturation has been observed between transition and transversion rates in relation to genetic distances.

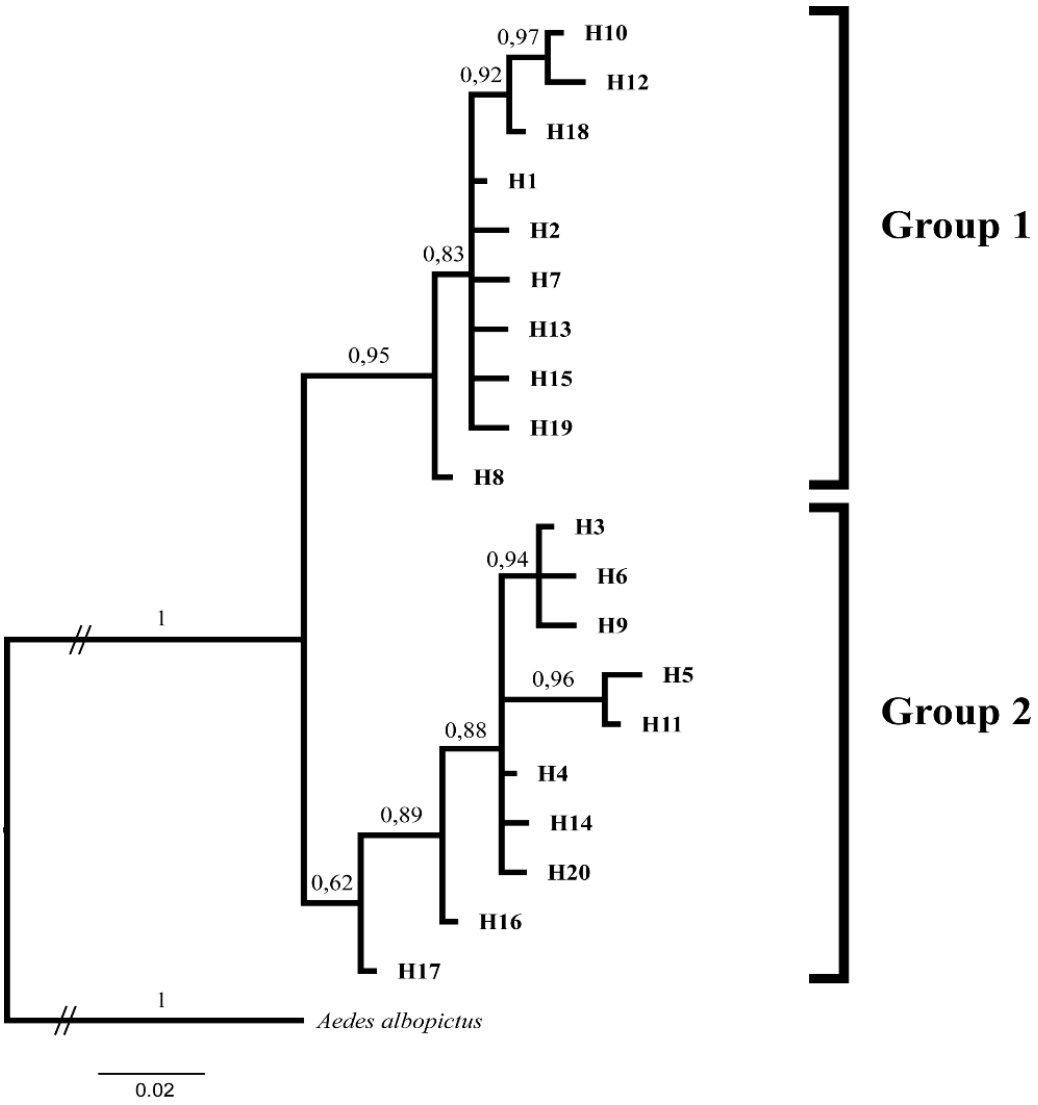

Figure S2. Bayesian Inference (BI) tree inferred for the 20 haplotypes from Brazil, with GTR + G model. Values above each branch represent the BPP. The two groups are identified as Group 1 and Group 2. *Aedes albopictus* is used as the outgroup.

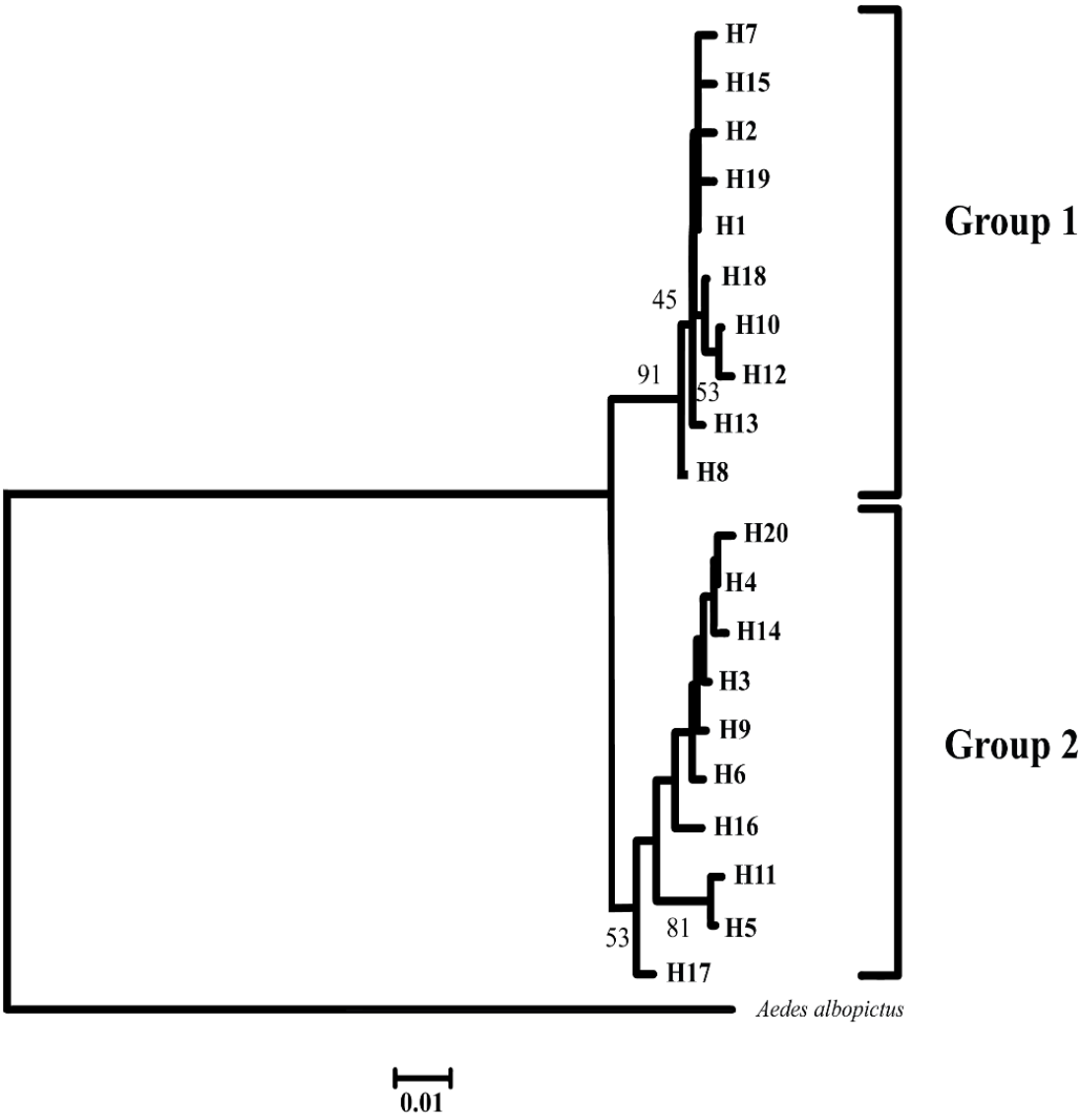

Figure S3. Neighbor Joining (NJ) tree inferred for the 20 haplotypes from Brazil. The values on the branches represent the bootstrap support. The two groups are identified as Group 1 and Group 2. *Aedes albopictus* is used as the outgroup.

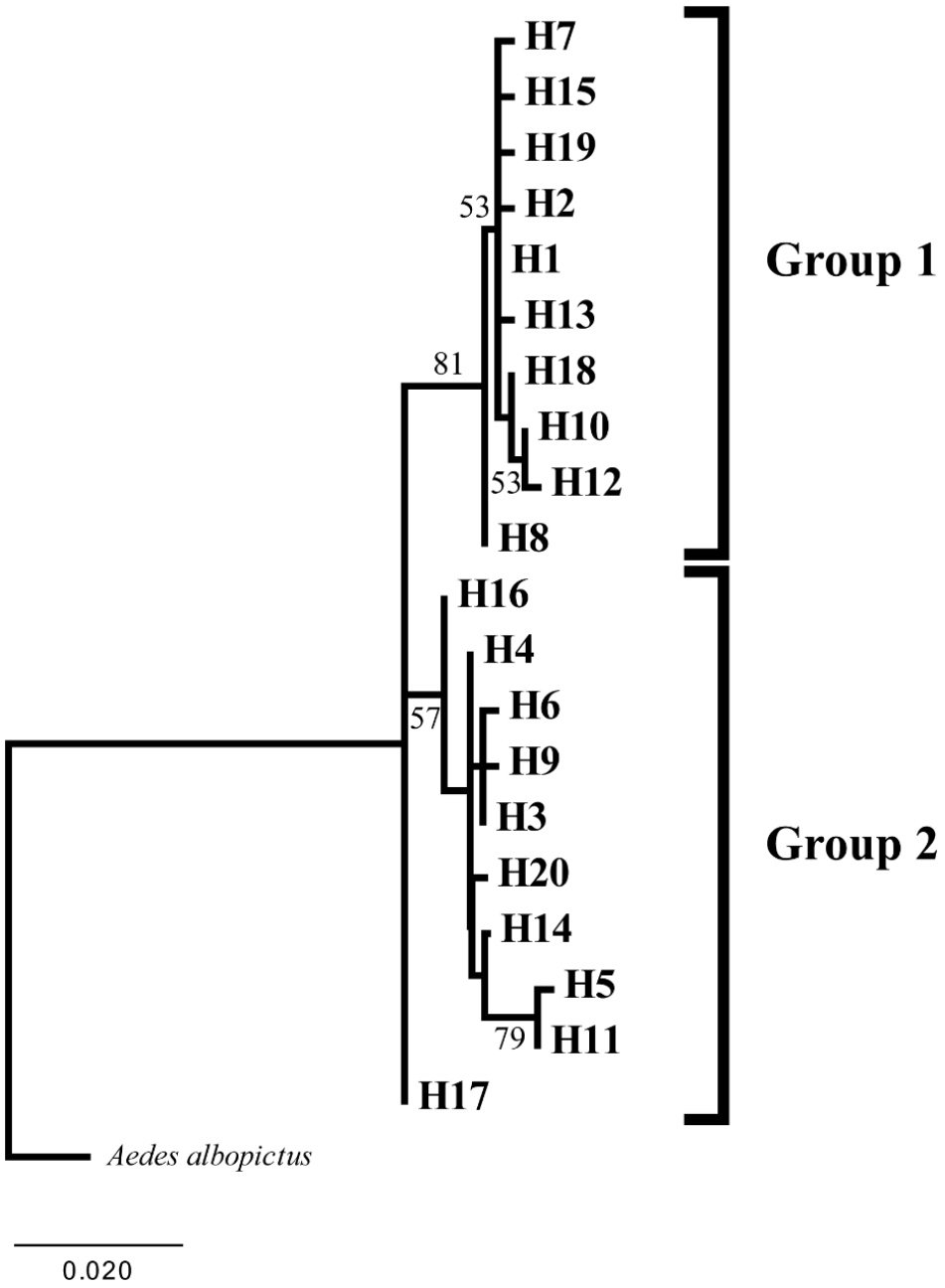

Fig S4. Maximum Likelihood (ML) tree inferred for the 20 haplotypes from Brazil, with GTR + G evolutionary model. Values above each branch represent the bootstrap support. The two groups are identified as Group 1 and Group 2. *Aedes albopictus* is used as the outgroup.

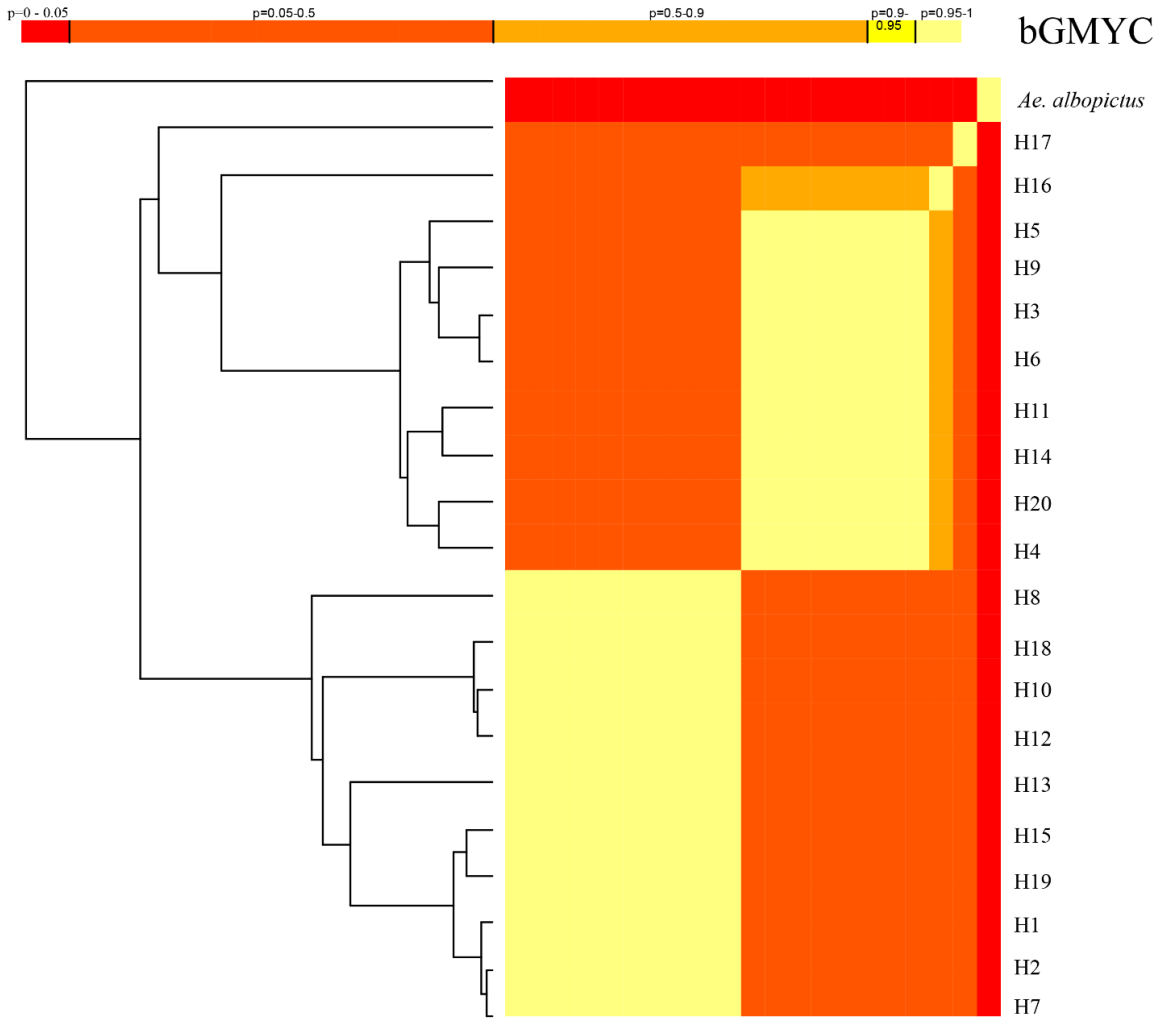

Fig S5. Species delimitation using bGMYC. The colored table is a sequence-by-sequence matrix. Cells are colored by the posterior probability that the corresponding sequences are conspecific, and this allows for the visualization of uncertainty in species limits. Only *Aedes albopictus* (outgroup) was separated as heterospecific (BPP > 95%).

Table S1 Intra and interpopulation genetic distances (K2P) obtained with the COI dataset.

	RB	**PV**	**GM**	**BV**	**IT**	**MA**	**NA**	**TA**	**RP**	**AR**	**CU**	**CG**	**TS**	**MP**
RB	**0.00038**													
PV	0.000	**0.00000**												
GM	0.000	0.000	**0.00028**											
BV	0.005	0.005	0.005	**0.00337**										
IT	0.005	0.005	0.005	0.006	**0.00785**									
MA	0.002	0.001	0.002	0.004	0.005	**0.00180**								
NA	0.005	0.005	0.005	0.003	0.006	0.004	**0.00346**							
TA	0.019	0.019	0.019	0.018	0.017	0.019	0.018	**0.00022**						
RP	0.012	0.012	0.012	0.013	0.012	0.012	0.013	0.008	**0.00974**					
AR	0.013	0.013	0.013	0.014	0.013	0.013	0.013	0.006	0.009	**0.00918**				
CU	0.014	0.014	0.014	0.015	0.014	0.014	0.015	0.006	0.009	0.008	**0.00878**			
CG	0.010	0.010	0.010	0.012	0.011	0.011	0.011	0.010	0.010	0.010	0.010	**0.01114**		
TS	0.008	0.008	0.008	0.011	0.010	0.009	0.011	0.012	0.010	0.010	0.011	0.010	**0.00909**	
MP	0.004	0.004	0.004	0.007	0.007	0.005	0.007	0.016	0.011	0.012	0.013	0.010	0.009	**0.00646**

K2P: Kimura 2 Parameters. Values in bold represent intra-population distances. See Table 1 for locality abbreviations.

Table S2 Genetic distances (K2P) between and within the two genetic groups obtained.

	Group 1	Group 2
Group 1	0.003382	**0.018**
Group 2	0.006346	

K2P: Kimura 2 Parameters. Value in bold represent the distance between the two groups.

5. Discussão Geral

Este é o primeiro estudo que estima a variação molecular e estrutura genética de populações de *Ae. aegypti* com o emprego de dois marcadores moleculares simultaneamente, microssatélites (herança bi parental) e DNAmt (herança materna) no Brasil. Os resultados revelaram um elevado e significativo nível de estrutura genética em escala macrogeográfica (do país) e confirmaram a existência de dois grupos genéticos para esta espécie no Brasil.

No capítulo 1, com base nas análises realizadas com 12 locos dos microssatélites em 15 populações de *Ae. aegypti* do Brasil, todos os locos mostraram-se úteis para estimar a variação genética. Os níveis de diferenciação encontrados mostraram uma alta estruturação entre as 15 populações. Também, foi indicada a existência de dois grandes grupos genéticos, mas com a existência de subestruturação populacional dentro de cada um deles, sendo a diferenciação genética mais significativa entre a população de Macapá e as demais populações.

No capítulo 2, com base nas análises realizadas com o gene *COI* (*Barcode*) de DNAmt nas 14 populações de *Ae. aegypti* do Brasil, foi encontrado 20 haplótipos. Os haplótipos H1, H4 e H10 foram os mais frequentes e foram os que mais contribuíram para a diferenciação genética entre as populações estudadas. Observou-se que há uma maior semelhança genética entre os espécimes das regiões do norte do Brasil e as linhagens de *Ae. aegypti* da Bolívia, Guiana Francesa, Ásia, Austrália, Venezuela, EUA e leste da África. Por outro lado, as amostras do sudeste e nordeste do Brasil mostraram maior semelhança genética com as linhagens caribenhas, mexicanas e oeste-africanas do vetor.

Considerando os resultados dos dois marcadores (nuclear e mitocondrial) conjuntamente utilizados nesse estudo, os resultados indicaram que a amostra de Teresina foi a mais geneticamente isoladas de todas as outras localidades destes estudos. Isolamento genético foi observado também entre a população de Macapá e as amostras do norte do Amazônia (Boa Vista, Manaus, Itacoatiara e Novo Airão). Os grupos genéticos observados com os dois marcadores também foram congruentes, exceto para os agrupamentos de Taubaté e Macapá. Com marcadores microssatélites, observou-se uma conexão genética entre as amostras de Taubaté (sudeste) e as amostras do norte da Amazônia (Boa Vista, Manaus, Itacoatiara e Novo Airão). Por outro lado, na análise com DNAmt não houve compartilhamento de haplótipos entre Taubaté e as amostras do

norte da Amazônia (Boa Vista, Manaus, Itacoatiara e Novo Airão). A amostra de Macapá foi a mais divergente das demais com o emprego de marcadores microssatélites. No entanto, na análise com o marcador mitocondrial, observou-se o contrário esta amostra compartilhou haplótipos com as amostras da região sudeste do Brasil (Taubaté, São José de Rio Preto e Araçatuba).

Em conclusão, com o emprego dos dois marcadores, observou-se uma estrutura genética complexa para *Ae. aegypti* do Brasil. A estrutura genética estimada com marcadores microssatélites provavelmente reflete um cenário recente, possibilitando inferir padrões de fluxo gênico contemporâneo, enquanto que os dados obtidos com a região do DNA Barcode do gene *COI*, refletem um cenário antigo, possibilitando inferir padrões históricos. Portanto, estes marcadores, por apresentarem modos distintos de herança e taxas de mutações distintas, proporcionam informações sobre a história demográfica e estrutura destas populações em períodos distintos de tempo, assim permitindo o conhecimento mais amplo sobre a trajetória evolutiva da espécie. Ainda, resultados distintos entre estes marcadores podem refletir taxas de migração diferentes entre machos e fêmeas, mas o que parece não ser o caso do *Ae. aegypti*, devido a sua dinâmica populacional peculiar.

A conectividade genética observada entre as amostras analisadas a partir dos locos microssatélites, pode sugerir que as duas linhagens genéticas observadas com o DNAmt estão livremente trocando genes, provavelmente, em consequência da intensa urbanização das cidades (construções, abastecimento de água e coletas de lixo regulares, densidade da população humana [fonte sanguínea], entre outros), efeitos de dispersão passiva e eventos múltiplos de introdução deste vetor dentro do país. Além disso, cada localidade mostrou uma dinâmica populacional diferente sugerindo a presença de fatores locais (ambientais, climáticos, culturais) atuando, o que pode exigir que a implementação de programas direcionados ao controle genético seja diferenciada nestas localidades.

No geral, as populações de *Aedes aegypti* deste estudo exibiram baixo tamanho efetivo populacional (*Ne)* que pode estar associado aos frequentes efeitos *bottleneck* e fundador. A frequência destes efeitos, no entanto, depende do envolvimento de fatores ambientais e climáticos locais, eventos de introduções e das medidas de controle do vetor entre as localidades.

No entanto, padrões distintos de dinâmica demográfica foram observados nestas populações, tais como baixo *Ne* com sinais de gargalo (Campina Grande e Araçatuba), alto *Ne* sem sinais de gargalo (Foz do Iguaçu, São José do Rio Preto, Taubaté, Cuiabá, Rio Branco), baixo *Ne*

sem sinais de gargalo (Macapá e Teresina) e alto *Ne* com sinais de gargalo (Itacoatiara). das amostras. Nas populações com *Ne* baixo, a deriva genética pode atuar mais intensamente na perda de diversidade genética e em menor tempo e, portanto, o tamanho do *Ne* da população alvo deve ser considerado em programas de modificação genética para este vetor.

6. Conclusão

Os resultados apresentados neste estudo demonstram importantes encontrados no que referem a diversidade genética, a estrutura genética e os padrões de dispersão de amostras do *Ae. aegypti*, assim como os possíveis eventos de colonização desse mosquito na Amazônia e em outras regiões no Brasil, como seguem:

- Alto e significativo nível de diferenciação genética foi observado entre as populações de Ae. aegypti em nosso estudo, confirmando estudos prévios;

- A existência de dois maiores grupos genéticos desta espécie foi confirmada no Brasil, mas com a existência de subestruturação populacional dentro de cada grupo genético.

- A maior divergência genética foi observada entre as amostras de Ae. aegypti de Macapá e as demais localidades do Brasil.

- Três prováveis eventos independentes de introdução deste vetor em Brasil foram observados.

7. Referências Bibliográficas

Apostol, B.L., Black, W.C., Reiter, P., Miller, B.R., 1996. Population genetics with RAPD-PCR markers: the breeding structure of *Aedes aegypti* in Puerto Rico. Heredity (Edinb). 76, 325–334. https://doi.org/10.1038/hdy.1996.50

Avise, J.C., 1994. Molecular Markers, Natural History and Evolution.

Ayres, C.F.J., Melo-santos, M.A.V., Solé-Cava, A.M., Furtado, A.F., 2003. Genetic Differentiation of *Aedes aegypti* (Diptera: Culicidae), the Major Dengue Vector in Brazil. J. Med. Entomol. 40, 430–435. https://doi.org/10.1603/0022-2585-40.4.430

Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Barreto, M.L., Teixeira, M. da G.L.C., 2008. Dengue no Brasil: situação epidemiológica e contribuições para uma agenda de pesquisa. Artig. Publicados em Periódicos Nac. 53–72.

Bass, C., Field, L.M., 2011. Gene amplification and insecticide resistance. Pest Manag. Sci. 67, 886–890. https://doi.org/10.1002/ps.2189

Beebe, N.W., Whelan, P.I., van den Hurk, A., Ritchie, S., Cooper, R.D., 2005. Genetic diversity of the dengue vector *Aedes aegypti* in Australia and implications for future surveillance and mainland incursion monitoring. Commun. Dseases Intell. Q. Rep. 29, 299–304.

Bennett, K.E., Olson, K.. E., Munoz, M.D.L., Ildefonso, F.-S., Farfan-ale, J.A., Higgs, S., Iv, W.C.B., Beaty, B.J., 2002. Variation in vector competence for dengue 2 virus among 24 collections of *Aedes aegypti* from mexico and the united states. Am. Soc. of Tropical Med. Hyg. 67, 85–92.

Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., Myers, M.F., George, D.B., Jaenisch, T., William Wint, G.R., Simmons, C.P., Scott, T.W., Farrar, J.J., Hay, S.I., 2013. The global distribution and burden of dengue. Nature 496, 504–507. https://doi.org/10.1038/nature12060

Bosio, C.F., Harrington, L.C., Jones, J.W., Norris, D.E., Scott, T.W., 2005. Genetic structure of *Aedes aegypti* populations in thailand using mitochondrial Dna. Am. J. Trop. Med. Hyg. 72, 434–442.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput. Biol. 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bracco, J.E., Capurro, M.L., Lourenço-de-Oliveira, R., Sallum, M.A.M., 2007. Genetic variability of *Aedes aegypti* in the Americas using a mitochondrial gene: evidence of multiple introductions. Mem Inst Oswaldo Cruz. 102, 573–580. https://doi.org/10.1590/S0074-02762007005000062

Brown, J.E., Evans, B.R., Zheng, W., Obas, V., Barrera-Martinez, L., Egizi, A., Zhao, H., Caccone, A., Powell, J.R., 2014. Human impacts have shaped historical and recent evolution in *Aedes aegypti*, the dengue and yellow fever mosquito. Evolution (N.Y). 68, 514–525. https://doi.org/10.1111/evo.12281

Brown, J.E., Mcbride, C.S., Johnson, P., Ritchie, S., Paupy, C., Bossin, H., Lutomiah, J., Fernandez-Salas, I., Ponlawat, A., Cornel, A.J., Black IV, W.C., Gorrochotegui-Escalante, N., Urdaneta-Marquez, L., Sylla, M., Slotman, M., Murray, K.O., Walker, C., Powell, J.R., 2011. Worldwide patterns of genetic differentiation imply multiple "domestications" of *Aedes aegypti*, a major vector of human diseases. Proc. R. Soc. B Biol. Sci. 278, 2446–2454. https://doi.org/10.1098/rspb.2010.2469

Calvez, E., Connor, O.O., Pol, M., Rousset, D., Faye, O., Richard, V., 2018. Differential transmission of Asian and African Zika virus lineages by *Aedes aegypti* from New Caledonia. Emerg. Microbes Infect. https://doi.org/10.1038/s41426-018-0166-2

Chambers, E.W., Meece, J.K., McGowan, J.A., Lovin, D.D., Hemme, R.R., Chadee, D.D., McAbee, K., Brown, S.E., Knudson, D.L., Severson, D.W., 2007. Microsatellite Isolation and Linkage Group Identification in the Yellow Fever Mosquito *Aedes aegypti*. J. Hered. 98, 202–210. https://doi.org/10.1093/jhered/esm015

Christophers, S.R., 1960. *Aedes aegypti* (L.) The yellow fever mosquito, Cambridge University Press.

Codeço, C.T., Honório, N.A., Ríos-Velásquez, C.M., Santos, M. da C.A. dos, Mattos, I.V. de, Luz, S.B., Reis, I.C., Cunha, G.B. da, Rosa-Freitas, M.G., Tsouris, P., Castro, M.G. de, Hayd, R.L.N., Luitgards-Moura, J.F., 2009. Seasonal dynamics of *Aedes aegypti* (Diptera: Culicidae) in the northernmost state of Brazil: a likely port-of-entry for dengue virus 4. Mem.

Inst. Oswaldo Cruz. 104, 614–620. https://doi.org/10.1590/S0074-02762009000400014

Consoli, R.A.G.B., Oliveira, R.L. De, 1994. Principais mosquitos de importância sanitária no Brasil.

Corander, J., Marttinen, P., Sirén, J., Tang, J., 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics. 9, 539. https://doi.org/10.1186/1471-2105-9-539

Corander, J., Tang, J., 2007. Bayesian analysis of population structure based on linked molecular information. Math. Biosci. 205, 19–31. https://doi.org/10.1016/j.mbs.2006.09.015

Costa-Ribeiro, M.C. V., Lourenco-de-Oliveira, R., Failloux, A.-B., 2006. Geographic and temporal genetic patterns of *Aedes aegypti* populations in Rio de Janeiro, Brazil. Trop. Med. Int. Heal. 11, 1276–1285. https://doi.org/10.1111/j.1365-3156.2006.01667.x

Da Costa-da-Silva, A.L., Capurro, M.L., Bracco, J.E., 2005. Genetic lineages in the yellow fever mosquito *Aedes* (Stegomyia) *aegypti* (Diptera: Culicidae) from Peru. Mem. Inst. Oswaldo Cruz. 100, 539–544. https://doi.org/10.1590/S0074-02762005000600007

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772–772. https://doi.org/10.1038/nmeth.2109

De Sousa, G.B., Blanco, A., Gardenal, C.N., 2001. Genetic Relationships Among *Aedes aegypti* (Diptera: Culicidae) Populations from Argentina Using Random Amplified Polymorphic DNA Polymerase Chain Reaction Markers. J. Med. Entomol. 38, 371–375. https://doi.org/10.1603/0022-2585-38.3.371

Delatte, H., Bagny, L., Brengue, C., Bouetard, A., Paupy, C., Fontenille, D., 2011. The invaders: Phylogeography of dengue and chikungunya viruses Aedes vectors, on the South West islands of the Indian Ocean. Infect. Genet. Evol. 11, 1769–1781. https://doi.org/10.1016/J.MEEGID.2011.07.016

Dinardo-Miranda, L.L., Contel, E.P.B., 1996. Enzymatic Variability in Natural Populations of *Aedes aegypti* (Diptera: Culicidae) from Brazil. J. Med. Entomol. 33, 726–733. https://doi.org/10.1093/jmedent/33.5.726

Donalísio, M.R., Glasser, C.M., 2002. Vigilância entomológica e controle de vetores do dengue. Rev. Bras. Epidemiol. 5, 259–279. https://doi.org/10.1590/S1415-790X2002000300005

Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eltis, D., Richardson, D., Davis, D.B., Blight, D.W., 2010. Atlas of the transatlantic slave trade.

Endersby, N.M., Rapley, L.P., Ryan, P.A., Nam, V.S., Yen, N.T., Kittiyapong, P., 2009. Genetic Structure of *Aedes aegypti* in Australia and Vietnam Revealed by Microsatellite and Exon Primed Intron Crossing Markers Suggests Feasibility of Local Control Options. J. Med. 46, 1074–1083.

Epidemiol, B., 2018. Boletim Epidemiológico 49, 1–14.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Laval, G., Schneider, S., 2007. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. Online 1, 47–50.

Failloux, A.-B., Vazeille, M., Rodhain, F., 2002a. Geographic Genetic Variation in Populations of the Dengue Virus Vector *Aedes aegypti*. J. Mol. Evol. 55, 653–663. https://doi.org/10.1007/s00239-002-2360-y

Failloux, A.-B., Vazeille, M., Rodhain, F., 2002b. Geographic Genetic Variation in Populations of the Dengue Virus Vector *Aedes aegypti*. J. Mol. Evol. 55, 653–663. https://doi.org/10.1007/s00239-002-2360-y

Fernández-Salas, I., Danis-Lozano, R., Casas-Martínez, M., Ulloa, A., Bond, J.G., Marina, C.F., Lopez-Ordóñez, T., Elizondo-Quiroga, A., Torres-Monzón, J.A., Díaz-González, E.E., 2015. Historical inability to control *Aedes aegypti* as a main contributor of fast dispersal of chikungunya outbreaks in Latin America. Antiviral Res. 124, 30–42. https://doi.org/10.1016/j.antiviral.2015.10.015

Ferreira, M.E., Grattapaglia, D., 1998. Introduccion al uso de marcadores moleculares en el analisis genetico. Brasilia: EMBRAPA-SPI/EMBRAPA-CENARGEN.

Figueiredo, L.T.., 2003. Dengue in Brazil: past, present and future perspective. Dengue Bolletin 27, 25–33.

Flynn, J.J., Nedbal, M.A., 1998. Phylogeny of the Carnivora (Mammalia): Congruence vs

Incompatibility among Multiple Data Sets. Mol. Phylogenet. Evol. 9, 414–426. https://doi.org/10.1006/MPEV.1998.0504

Folmer, 0, Black, M., Hoeh, W., Lutz, R., Vrijenhoek+, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299. https://doi.org/10.1371/journal.pone.0013102

Forattini, O.P., 2002. Culicidologia médica: Identificacao, biologia e epidemiologia. Editora da Universidade de São Paulo., São Paulo.

Fraga, C., Maria, J., Maia, J.D.F., 2003. Enzymatic variability in *Aedes aegypti* (Diptera : Culicidae) populations from Manaus AM, Brazil. Genetics and Molecular Biology. 187, 181–187.

Franco, O., 1976. Historia da Febre Amarela no Brasil. Ministério da Saúde, Rio Janeiro.

Gadelha, D.P., Toda, A.T., 1985. Biologia e comportamento do *Aedes aegypti*. Rev. Bras. Malariol. e Doencas Trop. 37, 29–36.

Giráldez, A., 2015. The Age of Trade: The Manila Galleons and the Dawn of the Global Economy, Journal of Jesuit Studies. https://doi.org/10.1163/22141332-00304009-16

Gloria-Soria, A., Ayala, D., Bheecarry, A., Calderon-arguedas, O., Chadee, D.D., Chiappero, M., Coetzee, M., Elahee, K.B.I.N., Fernandez-salas, I., Kamal, H.A., Kamgang, B., Khater, E.I.M., Kramer, L.D., Micieli, M.V., Paupy, C., Ponlawat, A., Sanchez-casas, R.M., Seixas, A.L.O., Sousa, C.A., Tropicales, E., Rica, U.D.C., Jos, S., Rica, C., 2016. Global genetic diversity of *Aedes aegypti*. Molecular Ecology. 25, 5377–5395. https://doi.org/10.1111/mec.13866

Gloria-Soria, A., Kellner, D.A., Brown, J.E., Kamgang, B., Powell, J.R., Biology, E., 2016b. Temporal genetic stability of *Stegomyia aegypti* (=*Aedes aegypti*) populations. Med Vet Entomol. 30, 235–240. https://doi.org/10.1111/mve.12153.

Goeldi, E.A., 1905. Os mosquitos no Para´. Mem Mus Goeldi Hist Nat Ethnog 4.

Gonçalves, C.M., Melo, F.F., Bezerra, J.M.T., Chaves, B.A., Silva, B.M., Silva, L.D., Pessanha, J.E.M., Arias, J.R., Secundino, N.F.C., Norris, D.E., Pimenta, P.F.P., 2014. Distinct variation in vector competence among nine field populations of *Aedes aegypti* from a Brazilian dengue-endemic risk city. Parasites & Vectors. 7, 320.

Gonçalves da Silva, A., Cunha, I.C.L., Santos, W.S., Luz, S.L.B., Ribolla, P.E.M., Abad-Franch,

F., 2012. Gene flow networks among american *Aedes aegypti* populations. Evol. Appl. 5, 664–676. https://doi.org/10.1111/j.1752-4571.2012.00244.x

Gorrochotegui-Escalante, N., Lozano-Fuentes, S., Black, W.C., Fernandez-Salas, L., Garcia-Rejon, J., Gomez-Machorro, C., Beaty, B.J., Farfan-Ale, J.A., De Lourdes Munoz, M., 2002. Breeding structure of *Aedes aegypti* populations in Mexico varies by region. Am. J. Trop. Med. Hyg. 66, 213–222. https://doi.org/10.4269/ajtmh.2002.66.213

Goudet, J., 1995. FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. J. Hered. 86, 485–486. https://doi.org/10.1093/oxfordjournals.jhered.a111627

Gubler, D.J., 2002. The global emergence/resurgence of arboviral diseases as public health problems. Arch. Med. Res. 33, 330–42.

Gubler, D.J., 1998. Dengue and dengue hemorrhagic fever. Clin. Microbiol. Rev. 11, 480–96.

Guyer, M.S., Collins, F.S., 1993. The Human Genome Project and the Future of Medicine. Arch. Pediatr. Adolesc. Med. 147, 1145. https://doi.org/10.1001/archpedi.1993.02160350019003

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98. https://doi.org/10.1108/S1876-0562(2012)0000006007

Hardy, J.L., Houk, E.J., Kramer, L.D., Reeves, W.C., 1983. Intrinsic Factors Affecting Vector Competence of Mosquitoes for Arboviruses. Annu. Rev. Entomol. 28, 229–262. https://doi.org/10.1146/annurev.en.28.010183.001305

Harrington, L.C., Scott, T.W., Lerdthusnee, K., Coleman, R.C., Costero, A., Clark, G.G., Jones, J.J., Kitthawee, S., Kittayapong, P., Sithiprasasna, R., Edman, J.D., 2005. Dispersal of the dengue vector *Aedes aegypti* within and between rural communities. Am. J. Trop. Med. Hyg. 72, 209–20.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174. https://doi.org/10.1007/BF02101694

Hlaing, T., Willoughby, T.L., Somboon, P., Walton, C., 2009. Mitochondrial pseudogenes in the nuclear genome of *Aedes aegypti* mosquitoes: implications for past and future population genetic studies. BMC Genet. 10. https://doi.org/10.1186/1471-2156-10-11

Holm, S., 1979. Board of the Foundation of the Scandinavian Journal of Statistics A Simple

Sequentially Rejective Multiple Test Procedure A Simple Sequentially Rejective Multiple Test Procedure. Source Scand. J. Stat. 6, 65–70. https://doi.org/10.2307/4615733

Honório, N.A., Silva, C., Leite, P.J., Monteiro, J., Lounibos, L.P., Lourenço-de-oliveira, R., 2003. Dispersal of *Aedes aegypti* and *Aedes albopictus* (Diptera : Culicidae) in an Urban Endemic Dengue Area in the State of Rio de Janeiro , Brazil. Mem Inst Oswaldo Cruz. 98, 191–198.

Huber, K., Le Loan, L., Huu Hoang, T., Khanh Tien, T., Rodhain, F., Failloux, A.-B., 2002. Temporal genetic variation in *Aedes aegypti* populations in Ho Chi Minh City (Vietnam). Heredity (Edinb). 89, 7–14. https://doi.org/10.1038/sj.hdy.6800086

Huber, K., Loan, L. Le, Chantha, N., Failloux, A.-B., 2004. Human transportation influences *Aedes aegypti* gene flow in Southeast Asia. Acta Trop. 90, 23–29. https://doi.org/10.1016/J.ACTATROPICA.2003.09.012

Huber, K., Mousson, L., Rodhain, F., Failloux, A.B., 1999. 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. 61, 1001–1003. https://doi.org/10.4269/ajtmh.1999.61.1001

Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components : a new method for the analysis of genetically structured populations.

Kamgang, B., Brengues, C., Fontenille, D., Njiokou, F., Simard, F., Paupy, C., 2011. Genetic Structure of the Tiger Mosquito , *Aedes albopictus* , in Cameroon (Central Africa). PLoS ONE 6(5): e20257. https://doi.org/10.1371/journal.pone.0020257.

Kindhauser, M.K., Allen, T., Frank, V., Santhana, R.S., Dye, C., 2016. Zika: the origin and spread of a mosquito-borne virus. Bull. World Health Organ. 94, 675–686C. https://doi.org/10.2471/BLT.16.171082

Kotsakiozi, P., Gloria-Soria, A., Caccone, A., Evans, B., Schama, R., Martins, A.J., Powell, J.R., 2017. Tracking the return of *Aedes aegypti* to Brazil , the major vector of the dengue , chikungunya and Zika viruses. PLoS Negl Trop Dis 11(7): e0005653. https://doi.org/10.1371/journal.pntd.0005653.

Laith, Y., Walker, T., 2016. Zika virus outbreak in the Americas : The need for novel mosquito control methods Zika virus outbreak in the Americas : the need for novel. Lancet Glob. Heal. 8–10. https://doi.org/10.1016/S2214-109X(16)00048-6

Lang, B.F., Gray, M.W., Burger, G., 1999. Mitochondrial Genome Evolution and the Origin of Eukaryotes. Annu. Rev. Genet. 33, 351–397. https://doi.org/10.1146/annurev.genet.33.1.351

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Li, Y.-C., Korol, A.B., Fahima, T., Nevo, E., 2004. Microsatellites Within Genes: Structure, Function, and Evolution. Mol. Biol. Evol. 21, 991–1007. https://doi.org/10.1093/molbev/msh073

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Lima, R.S., Scarpassa, V.M., 2009. Evidence of two lineages of the dengue vector *Aedes aegypti* in the Brazilian Amazon, based on mitochondrial DNA ND4 gene sequences. Genet. Mol. Biol. 32, 414–422. https://doi.org/10.1590/S1415-47572009005000036

Linss, J.G.B., Brito, L.P., Garcia, G.A., Araki, A.S., Bruno, R.V., Bento, J., Lima, P., Valle, D., Martins, A.J., 2014. Distribution and dissemination of the Val1016Ile and Phe1534Cys Kdr mutations in *Aedes aegypti* Brazilian natural populations. Parasites & Vectors. 7: 25. https://doi.org/10.1186/1756-3305-7-25

Linss, J.G.B., Martins, A.J., Valle, D., Peixoto, A.A., de Andrade Lins, R.M.M., 2009. Voltage-Gated Sodium Channel Polymorphism and Metabolic Resistance in Pyrethroid-Resistant *Aedes aegypti* from Brazil. Am. J. Trop. Med. Hyg. 81, 108–115. https://doi.org/10.4269/ajtmh.2009.81.108

Llinas, G.A., Gardenal, C.N., 2011. Phylogeography of *Aedes aegypti* in Argentina : Long-Distance Colonization and Rapid Restoration of Fragmented Relicts After a Continental Control Campaign. Vector-Borne Zoonotic Dis. 00. https://doi.org/10.1089/vbz.2011.0696

Louise, C., Vidal, P.O., Suesdek, L., 2015. Microevolution of *Aedes aegypti*. PLoS One 10, 1–16. https://doi.org/10.1371/journal.pone.0137851

Lourenço-de-Oliveira, R., Honório, N.A., Castro, M.G., Schatzmayr, H.G., Miagostovich, M.P., Alves, J.C., Silva, W.C., Leite, P.J., Nogueira, R.M., 2002. Dengue virus type 3 isolation from *Aedes aegypti* in the municipality of Nova Iguaçu, State of Rio de Janeiro. Mem. Inst. Oswaldo Cruz 97, 799–800. https://doi.org/10.1590/s0074-02762002000600009

Lourenço-de-Oliveira, R., Vazeille, M., de Filippis, A.M.., Failloux, A.., 2004. *Aedes aegypti* in Brazil: genetically differentiated populations with high susceptibility to dengue and yellow fever viruses. Trans. R. Soc. Trop. Med. Hyg. 98, 43–54. https://doi.org/10.1016/S0035-9203(03)00006-3

Maciel-de-freitas, R., Aguiar, R., Bruno, R. V, Guimarães, M.C., Lourenço-de-oliveira, R., Sorgine, M.H.F., Struchiner, C.J., Valle, D., Neill, S.L.O., Moreira, L.A., 2012. Why do we need alternative tools to control mosquito-borne diseases in Latin America. Mem Inst Oswaldo Cruz. 107(6), 828–829.

Madariaga, M., Ticona, E., Resurrecion, C., 2016. Chikungunya : bending over the Americas and the rest of the world. Braz J Infect Dis. 20(1), 91–98.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Research. 27, 209-220.

McGraw, E.A., O'Neill, S.L.O., 2013. Beyond insecticides : new thinking on an ancient problem. Nat Rev Microbio. 11(3), 181-93. https://doi.org/10.1038/nrmicro2968

Mendonça, B.A.A., de Sousa, A.C.B., de Souza, A.P., Scarpassa, V.M., 2014. Temporal genetic structure of major dengue vector *Aedes aegypti* from Manaus, Amazonas, Brazil. Acta Trop. 134, 80–88. https://doi.org/10.1016/j.actatropica.2014.02.014

Merrill, S.A., Ramberg, F.B., Hagedorn, H.H., 2005. Phylogeography and population strucure of *Aedes aegypti* in Arizona. Trop. Med. 72, 304–310.

Monteiro, F.A., Shama, R., Martins, A.J., Gloria-Soria, A., Brown, J.E., Powell, J.R., 2014. Genetic Diversity of Brazilian *Aedes aegypti*: Patterns following an Eradication Program.

PLoS Negl. Trop. Dis. 8, 1–10. https://doi.org/10.1371/journal.pntd.0003167

Moore, M., Sylla, M., Goss, L., Burugu, M.W., Sang, R., Kamau, L.W., Kenya, E.U., Bosio, C., Munoz, M. de L., Sharakova, M., Black, W.C., 2013. Dual African Origins of Global *Aedes aegypti* s.l. Populations Revealed by Mitochondrial DNA. PLoS Negl. Trop. Dis. 7, e2175. https://doi.org/10.1371/journal.pntd.0002175

Moreira, L.A., Iturbe-Ormaetxe, I., Jeffery, J.A., Lu, G., Pyke, A.T., Hedges, L.M., Rocha, B.C., Hall-Mendelin, S., Day, A., Riegler, M., Hugo, L.E., Johnson, K.N., Kay, B.H., McGraw, E.A., van den Hurk, A.F., Ryan, P.A., O'Neill, S.L., 2009. A Wolbachia Symbiont in *Aedes aegypti* Limits Infection with Dengue, Chikungunya, and Plasmodium. Cell 139, 1268–1278. https://doi.org/10.1016/J.CELL.2009.11.042

Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics. Annu. Rev. Ecol. Syst. 18, 269–292. https://doi.org/10.1146/annurev.es.18.110187.001413

Mousson, L., Dauga, C., Garrigues, T., Schaffner, F., Vazeille, M., Failloux, A.B., 2005. Phylogeography of Aedes (Stegomyia) aegypti (L.) and Aedes (Stegomyia) albopictus (Skuse) (Diptera: Culicidae) based on mitochondrial DNA variations. Genet. Res. 86, 1–11. https://doi.org/10.1017/S0016672305007627

Neis Ribeiro, A., Gassen Balsan, L., de Moura, G., 2013. Análise das políticas públicas de combate à dengue. Grupo EUMED.NET de la Universidad de Málaga.

Oehler, E., Watrin, L., Larre, P., Leparc-Goffart, I., Lastère, S., Valour, F., Baudouin, L., Mallet, H.P., Musso, D., Ghawche, F., 2014. Zika virus infection complicated by Guillain-Barré syndrome – case report, French Polynesia, December 2013. Eurosurveillance 19, 20720. https://doi.org/10.2807/1560-7917.ES2014.19.9.20720

Oosterhout, V., Hutchinson, W., Wills, D.P.M., SHIPLEY, P., 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Resour. 4, 535–538.

Osanai, C.H., Travassos Da Rosa, A.P.A., Tang, A.T., Amaral, R.S., Passos, A.D.C., Tauil, P.L., 1983. Outbreak of dengue in Boa Vista, Roraima. Preliminary report. Rev. Inst. Med. Trop. Sao Paulo 25, 53–54.

Paduan, K.D.S., Eduardo, P., Ribolla, M., 2008. Mitochondrial DNA Polymorphism and

Heteroplasmy in Populations of *Aedes aegypti* in Brazil. J Med Entomol. 45(1), 59–67.

Paduan, K.D.S., Jo Jr., J.P.A., M Ribolla, P.E., 2006. Genetic variability in geographical populations of *Aedes aegypti* (Diptera, Culicidae) in Brazil elucidated by molecular markers. Genet. Mol. Biol. 29, 391–395.

PAHO, World Health Organization (WHO), 2017. Zika situation report - Brazil 1–9. https://doi.org/10.1126/science.aaf5036

Paupy, C., Le, G., Brengues, C., Guerra, M., Revollo, J., Barja, Z., Hervé, J., Fontenille, D., 2012. Infection , Genetics and Evolution Genetic structure and phylogeography of *Aedes aegypti* , the dengue and yellow-fever mosquito vector in Bolivia. Infect. Genet. Evol. 12, 1260–1269. https://doi.org/10.1016/j.meegid.2012.04.012

Paupy, C., Orsoni, A., Mousson, L., Huber, K., 2004. Comparisons of Amplified Fragment Length Polymorphism (AFLP), Microsatellite, and Isoenzyme Markers: Population Genetics of *Aedes aegypti* (Diptera: Culicidae) from Phnom Penh (Cambodia). J. Med. Entomol. 41, 664–671. https://doi.org/10.1603/0022-2585-41.4.664

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—An Update. Bioinformatics 28, 2537–2539.

Peel, D., Ovenden, J., Peel, S., 2004. NeEstimator: Software for Estimating Effective Population Size. Queensl. Gov. Dep. Prim. Ind. Fish. St Lucia, Queensl.

Piry, S., Luikart, G., Cornuet, J.-M., 1999. Bottleneck : A Computer Program for Detecting Recent Reductions in the Effective Population Size Using Allele Frequency Data. J. Hered. 90, 502–503. https://doi.org/10.1093/jhered/90.4.502

Powell, J.R., Tabachnick, W.J., 2013. History of domestication and spread of *Aedes aegypti*-a review. Mem. Inst. Oswaldo Cruz 108, 11–17. https://doi.org/10.1590/0074-0276130395

Powell, J.R., Tabachnick, W.J., Arnold, J., 1980. Genetics and the origin of a vector population: *Aedes aegypti*, a case study. Science 208, 1385–7.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics 204, 391–393. https://doi.org/10.1534/genetics.116.195164

R Core Team, 2017. R: A Language and Environment for Statistical Computing.

Ravel, S., Hervé, J.-P., Diarrassouba, S., Kone, A., Cuny, G., 2002. 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. 82, 39–49. https://doi.org/10.1016/S0001-706X(02)00028-1

Ravel, S., Monteny, N., Olmos, D.V., Verdugo, J.E., Cuny, G., 2001. A preliminary study of the population genetics of *Aedes aegypti* (Diptera: Culicidae) from Mexico using microsatellite and AFLP markers. Acta Trop. 78, 241–250. https://doi.org/10.1016/S0001-706X(01)00083-3

Rebêlo, J.M.M., Costa, J.M.L., Silva, F.S., Pereira, Y.N.O., Silva, J.M. da, 1999. Distribuição de *Aedes aegypti* e do dengue no Estado do Maranhão, Brasil. Cad. Saude Publica 15, 477–486. https://doi.org/10.1590/S0102-311X1999000300004

Reid, N.M., Carstens, B.C., 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evol. Biol. 12, 196. https://doi.org/10.1186/1471-2148-12-196

Rey, J.R., O'Connell, S.M., 2014. Oviposition by *Aedes aegypti* and Aedes albopictus : Influence of congeners and of oviposition site characteristics. J. Vector Ecol. 39, 190–196. https://doi.org/10.1111/j.1948-7134.2014.12086.x

Rondan Dueñas, J., Albrieu Llinás, G., Panzetta-Dutari, GM, Gardenal, C., 2009. Two different routes of colonization of *Aedes aegypti* in Argentina from neighboring countries. J Med Entom 46, 1344–1354.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Rosenberg, N.A., 2003. distruct: a program for the graphical display of population structure. Mol. Ecol. Notes 4, 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Rubin, E.J., Ph, D., Greene, M.F., Baden, L.R., 2016. Zika Virus and Microcephaly. N Engl J Med. 374, 984-985. https://doi.org/10.1056/NEJMe1601379

Rueda, L.M., 2004. Pictorial Keys for the Identification of Mosquitoes (Diptera: Culicidae) Associated With Dengue Virus Transmission.Zootaxa. 589, 1-60.

Saarman, N.P., Gloria-Soria, A., Anderson, E.C., Evans, B.R., Pless, E., Cosme, L. V., Gonzalez-

Acosta, C., Kamgang, B., Wesson, D.M., Powell, J.R., 2017. Effective population sizes of a major vector of human diseases, *Aedes aegypti*. Evol. Appl. 10, 1031–1039. https://doi.org/10.1111/eva.12508

Sambrook, J.F., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, 3rd ed., V. ed. Cold Spring Harbor Laboratory Press.

Santos, V.M. dos, Macoris, M. de L. da G., Andrighetti, M.T.M., Avila, P.E., Kirchgatter, K., 2003. Analysis of genetic relatedness between populations of *Aedes aegypti* from different geographic regions of São Paulo state, Brazil. Rev. Inst. Med. Trop. Sao Paulo 45, 99–101. https://doi.org/10.1590/S0036-46652003000200009

Scarpassa, V.M., Cardoza, T.B., Junior, R.P.C., 2008. Population Genetics and Phylogeography of *Aedes aegypti* (Diptera : Culicidae) from Brazil. Am J Trop Med Hyg. 78, 895–903.

Scarpassa, V.M., Conn, J.E., 2007. Population genetic structure of the major malaria vector *Anopheles darlingi* (Diptera : Culicidae) from the Brazilian Amazon , using microsatellite markers. Mem Inst Oswaldo Cruz. 102, 319–327.

Schatzmayr, H.G., 2000. Dengue situation in Brazil by year 2000. Mem. Inst. Oswaldo Cruz 95, 179–181. https://doi.org/10.1590/S0074-02762000000700030

Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nat. Biotechnol. 18, 233–234. https://doi.org/10.1038/72708

Slotman, M.A., Kelly, N.B., Harrington, L.C., Kitthawee, S., Jones, J.W., Scott, T.W., Caccone, A., Powell, J.R., 2007. Polymorphic microsatellite markers for studies of *Aedes aegypti* (Diptera : Culicidae), the vector of dengue and yellow fever. Molecular Ecology Notes. 7, 168–171. https://doi.org/10.1111/j.1471-8286.2006.01533.x

Smith, C.E.G., 1956. The history of dengue in tropical Asia and its probable relationship to the mosquito *Aedes aegypti*. J. Trop. Med. Hyg. 59, 243–51.

Snustad, D.P., Simmons, M.J., 2008. Fundamentos de genética. 4. ed. Rio de Janeiro: Guanabara Koogan. Genética de organelas e cloroplastos.

Sousa, A.A. de, Fraga, E., Sampaio, I., Schneider, H., Barros, M.C., 2017. Genetic differentiation in populations of *Aedes aegypti* (Diptera, Culicidae) dengue vector from the Brazilian state of Maranhão. Rev. Bras. Entomol. 61, 51–59. https://doi.org/10.1016/j.rbe.2016.10.003

Stanczyk, N.M., 2011. An investigation of DEET-insensitivity in *Aedes aegypti*. PhD thesis,

University of Nottingham.

Tabachnick, W.J., 1991. Evolutionary Genetics and Arthropod-borne Disease: The Yellow Fever Mosquito. Am. Entomol. 37, 14–26. https://doi.org/10.1093/ae/37.1.14

Tabachnick, W.J., 1982. Geographic and temporal patterns of genetic variation of *Aedes aegypti* in New Orleans. Am. J. Trop. Med. Hyg. 31, 849–853. https://doi.org/10.4269/ajtmh.1982.31.849

Tabachnick, W.J., Powell, J.R., 1979. A world-wide survey of genetic variation in the yellow fever mosquito, *Aedes aegypti*. Genet. Res. 34, 215–29.

Tabachnick, W.J., Powell, J.R., 1978. Genetic structure of the East African domestic populations of *Aedes aegypti*. Nature 272, 535–537. https://doi.org/10.1038/272535a0

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30, 2725–2729. https://doi.org/10.1093/molbev/mst197

Tauil, P.L., 2001. Urbanização e ecologia do dengue. Cad. Saude Publica 17, S99–S102. https://doi.org/10.1590/S0102-311X2001000700018

Tautz, D., 1993. Notes on the definition and nomenclature of tandemly repetitive DNA sequences, in: DNA Fingerprinting: State of the Science. Birkhäuser Basel, Basel, pp. 21–28. https://doi.org/10.1007/978-3-0348-8583-6_2

Vasconcelos, P.F.., Rosa, A.P.A.., Pinheiro, F.P., Rodrigues, S.G., Rosa, E.S.., Cruz, A.C.., Rosa, J.F.S.., 1999. *Aedes aegypti*, dengue and re-urbanization of yellos fever in Brazil and other South American Countries: past and present situation and future perspectives. Dengue Bull. 23, 1–8.

Vazeille-Falcoz, M., Mousson, L., 1999. Variation in oral susceptibility to dengue type 2 virus of populations of *Aedes aegypti* from the Islands of Tahiti and Moorea , French Polynesia. Am J Trop Med Hyg. 60, 292–299.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York.

Wallis, G.P., Tabachnick, W.J., 1990. Genetic analysis of rock hole and domestic *Aedes aegypti* on the Caribbean island of Anguilla. J. Am. Mosq. Control Assoc. 6, 625–30.

Wesson, D.M., Ocampo, C.B., 2004. Population dynamics of *Aedes aegypti* from a Dengue

Hyperendemic Urban Setting In Colombia. Am. J. Trop. Med. Hyg. 71, 506–513. https://doi.org/10.4269/ajtmh.2004.71.506

WHO, 2016. Dengue vaccine: WHO position paper – July 2016 Introduction. Wkly. Epidemiol. Rec. 30, 349–364.

Wilke, A.B.B., Wilk-da-Silva, R., Marrelli, M.T., 2017. Microgeographic population structuring of *Aedes aegypti* (Diptera: Culicidae). PLoS One 12, e0185150. https://doi.org/10.1371/journal.pone.0185150

Woodring, J.L., Higgs, S., Beaty, B.J., 1996. Natural cycles of vector-borne pathogens. Biol. Dis. vectors 51–72.

Xia, X., Xie, Z., 2001. DAMBE: Software Package for Data Analysis in Molecular Biology and Evolution. J. Hered. 92, 371–373. https://doi.org/10.1093/jhered/92.4.371

Yan, G., Romero-Severson, J., Walton, M., Chadee, D.D., Severson, D.W., 1999. Population genetics of the yellow fever mosquito in trinidad: Comparisons of amplified fragment length polymorphism (AFLP) and restriction fragment length polymorphism (RFLP) markers. Mol. Ecol. 8, 951–963. https://doi.org/10.1046/j.1365-294X.1999.00647.x

Zanluca, C., Melo, V.C.A. de, Mosimann, A.L.P., Santos, G.I.V. dos, Santos, C.N.D. dos, Luz, K., Zanluca, C., Melo, V.C.A. de, Mosimann, A.L.P., Santos, G.I.V. dos, Santos, C.N.D. dos, Luz, K., 2015. First report of autochthonous transmission of Zika virus in Brazil. Mem. Inst. Oswaldo Cruz 110, 569–572. https://doi.org/10.1590/0074-02760150192

Zwickl, D.J., 2006. Garli—Genetic Algorithm for Rapid Likelihood Inference. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.