Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ
Pós-Graduação em Biologia Celular e Molecular

RACHEL MAZZEI MOURA DE ANDRADE LINS

Genética de populações do gene *paralytic* e análise multilocus de genes que controlam o som de corte no complexo *Lutzomyia longipalpis* (Diptera: Psychodidae)

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Biologia Celular e Molecular

Orientador: Dr. Alexandre Afranio Peixoto

RIO DE JANEIRO
2012

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Pós-Graduação em Biologia Celular e Molecular

AUTOR: RACHEL MAZZEI MOURA DE ANDRADE LINS

GENÉTICA DE POPULAÇÕES DO GENE *paralytic* E ANÁLISE MULTILOCUS DE GENES QUE CONTROLAM O SOM DE CORTE NO COMPLEXO *Lutzomyia longipalpis* (DIPTERA: PSYCHODIDAE)

ORIENTADOR: Dr. Alexandre Afranio Peixoto

Aprovada em: 10 / 09 / 2012

EXAMINADORES:

Prof. Dr. Elisa Cupolillo - Presidente
Prof. Dr. Carlos Eduardo Guerra Schrago
Prof. Dr. Cristiano Valentim da Silva Lazoski
Prof. Dr. Alberto Martín Rivera Dávila - Suplente
Prof. Dr. Reginaldo Peçanha Brazil - Suplente

Rio de Janeiro, 10 de Setembro de 2012

Dedico esta tese à minha família,
ao meu marido Thiago e à minha filha Sarah

Agradeço ao programa de pós graduação em Biologia Celular e Molecular pela contribuição com minha formação como profissional, assim como aos seus funcionários pela ajuda com todas as questões burocráticas.

Agradeço aos órgãos de fomento científico CAPES e CNPq pelas bolsas de doutorado e doutorado sanduíche respectivamente.

Aos membros da banca pelo pronto comprometimento na formação desta banca de doutorado.

Ao meu orientador, Dr Alexandre Afranio Peixoto pela oportunidade de desenvolver minha tese em seu laboratório, contribuindo significativamente para minha formação científica.

Ao meu orientador do período sanduíche, Dr Gregory Lanzaro por todo apoio e ensinamento durante meu trabalho na Universidade da Califórnia (DAVIS). Com certeza este período foi fundamental não apenas para meu amadurecimento científico, mas também como pessoa, abrindo portas e trazendo oportunidades que eu jamais imaginaria ter.

Aos companheiros de laboratório pelos momentos agradáveis de conversas "científicas" ou não e pela troca de experiências: Tamara querida, amiga de todas as horas, Elisa (minha aluna querida), Luísa "guria", Rachel Canto, Samiruda, Gabriel, Rafa, Kiko, Paulo Amoretty, Karine, Israel, Ricardo, Gustavo, Rayane, Felipe Vigoder, Lourdes, Pablo, Saori, Saullo, e ao amigo Robson, peça fundamental no laboratório, sem o qual muitas teses e dissertações não andariam.

À minha família por seu apoio e compreensão durante todo esse período. À minha mãe pelo incentivo e troca de experiências em pesquisa, meu pai, minhas irmãs Mariana e Bianca, meus sobrinhos Iuri e Davi, e aos meus avós queridos, meus bichinhos Lupi e Pity.

Agradeço especialmente ao meu marido Thiago, pela paciência, incentivo e apoio incondicional em todos os momentos. Por fim, agradeço à minha filha Sarah, que desde a barriga já me inspirava a ser uma pessoa melhor e me motiva diariamente a batalhar por um lugar ao sol.

מאמין אני שאפשר

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

GENÉTICA DE POPULAÇÕES DO GENE *paralytic* E ANÁLISE MULTILOCUS DE GENES QUE CONTROLAM O SOM DE CORTE NO COMPLEXO *Lutzomyia longipalpis* (DIPTERA: PSYCHODIDAE)

RESUMO
TESE DE DOUTORADO

Rachel Mazzei Moura de Andrade Lins

Lutzomyia longipalpis é o principal vetor da leishmaniose visceral americana (LVA), e representa um complexo de espécies crípticas. Trabalhos anteriores sugeriram a existência de quatro espécies com distribuições geográficas distintas: (1) espécie A- *Lu. longipalpis* sensu stricto – Brasil (Exceto Roraima), (2) Laran – espécie B = *Lu. pseudolongipalpis* (Venezuela), (3) cis-Andina – espécie C (Colômbia, noroeste da Venezuela e incluíram Roraima no Brasil) e (4) trans-Andina – espécie D (América Central). Além disso, experimentos com cruzamentos, análise dos sons de cópula, feromônios, microsatelites e genes que controlam o som de corte, sugeriram que as populações brasileiras deste vetor representam diferentes espécies que podem ser divididas em dois grupos principais de acordo com o tipo de som (*Burst* ou Pulsado) que machos produzem durante a cópula. No entanto, nenhuma diferença diagnóstica foi observada entre esses dois grupos na maioria dos marcadores moleculares utilizados até o momento. Inicialmente analisamos a divergência molecular em um fragmento do gene *paralytic*, um *locus* envolvido no controle dos sons de corte em *Drosophila*, entre diversas populações de *Lu. longipalpis* do Brasil produzindo sons do tipo *Burst* ou Pulsado. Comparamos também *Lu. longipalpis* com uma espécie muito próxima, *Lutzomyia cruzi*, que produz sons do tipo *Burst*. Os resultados indicaram um número maior de diferenças fixas entre *Lu. cruzi* e as populações de *Lu. longipalpis* com som tipo Pulsado que com aquelas produzindo som do tipo *Burst*. Os dados confirmaram que as diferentes espécies do complexo *Lu. longipalpis* no Brasil podem ser divididas em dois grupos, um representando uma única especie e o segundo grupo mais heterogêneo que provavelmente representa diversas espécies incipientes. Também fizemos uma análise macrogeográfica da divergência e fluxo gênico no gene *paralytic* entre populações de *Lu. longipalpis* das Américas do Sul e Central. Nossos resultados corroboraram a classificação prévia das espécies B, C e D, e também confirmaram que este vetor representa um complexo de espécies no Brasil com dois grupos principais. Finalmente, fizemos uma análise multilocus usando ortólogos dos genes *period*, *paralytic*, *cacophony*, *ebony*, *fruitless* and *slowpoke*, que estão associados com a produção do som em *Drosophila*, para avaliar a divergência e fluxo gênico entre as espécies simpátricas do complexo *Lu. longipalpis* das localidades de Sobral, Jaíba e Estrela no Brasil. Altos níveis de divergência genética foram detectados com evidência de introgressão assimétrica na maioria dos *loci*. Além disso, nossos dados confirmaram que genes envolvidos na produção do som tendem a mostrar maiores níveis de divergência genética e menos introgressão entre as espécies do complexo *Lu. longipalpis* que outros marcadores.

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

POPULATION GENETICS OF THE *paralytic* GENE AND MULTILOCUS ANALYSIS OF NCOURTSHIP SONG GENES IN THE *Lutzomyia longipalpis* COMPLEX (DIPTERA: PSYCHODIDAE)

ABSTRACT

TESE DE DOUTORADO

Rachel Mazzei Moura de Andrade Lins

Lutzomyia longipalpis is the main vector of American Visceral Leishmaniasis, and represents a complex of cryptic sibling species. Previous studies suggested the existence of four main clades in Latin America representing different species with distinct geographical ranges: (1) Brazil (except Roraima State) – Species A = *Lu. longipalpis sensu stricto*, (2) Laran – Species B = *Lu. pseudolongipalpis* (Venezuela), (3) cis-Andean – Species C (Colombia and northwestern Venezuela, and Roraima State in Brazil) and (4) trans-Andean – Species D (Central America). In addition, cross-mating experiments, analysis of copulation songs, pheromones, microsatellites and lovesong genes, suggested that the Brazilian populations of this vector represent different species that can be divided into two main groups according to the type of song (Burst vs. Pulse) males produce during copulation. Nevertheless, no diagnostic differences have been observed between these two groups with most molecular markers used to date. We initially analyzed the molecular divergence in a fragment of the *paralytic* gene, a locus involved in the control of courtship songs in *Drosophila*, among a number of *Lu. longipalpis* populations from Brazil producing Burst and Pulse-type songs. We also compared *Lu. longipalpis* with a very closely related species, *Lutzomyia cruzi*, which produces Burst-type song. The results indicated a higher number of fixed differences between *Lu. cruzi* and the Pulse-type populations of *Lu. longipalpis* than with those producing Burst-type songs. The data confirmed that the different sibling species of the *Lu. longipalpis* complex in Brazil can be divided into two main groups, one representing a single species and a second, a more heterogeneous group, that probably represents a number of incipient species. We also performed a macrogeographic analysis of the divergence and gene flow in the *paralytic* gene among *Lu. longipalpis* populations from South and Central America. Our results corroborated the previous classification of Species B, C and D and also confirmed that this vector represents a species complex in Brazil with two main groups. Finally, we carried out a multilocus analysis using orthologues of the *Drosophila* genes *period*, *paralytic*, *cacophony*, *ebony*, *fruitless* and *slowpoke*, which are associated with courtship song production, to evaluate the divergence and gene flow between *Lu. longipalpis* sympatric siblings from the localities Sobral, Jaíba and Estrela in Brazil. High levels of genetic divergence were detected with evidence of asymmetric introgression in most loci. In addition, our data confirmed that genes involved in song production tend to show higher levels of genetic divergence and less introgression in the *Lu. longipalpis* complex than other markers.

SUMÁRIO

1. INTRODUÇÃO

1.1 – O gênero *Lutzomyia* e o complexo Longipalpis

O gênero *Lutzomyia* (França, 1924) (Diptera: Psychodidae: Phlebotominae) compreende um grande número de espécies de flebotomíneos do Novo Mundo e é considerado um dos mais importantes grupos de insetos hematófagos das Américas por possuir, em sua composição, diversos vetores associados à transmissão das leishmanioses visceral e cutânea.

De acordo com Young & Duncan (1994) este gênero está taxonomicamente dividido em diversos subgêneros, possuindo, ainda, várias espécies isoladas (Barretto 1962, Theodor 1965, Lewis et al. 1977, Martins et al. 1978, Young & Duncan 1994). Mais recentemente Galati (2003) em uma analise filogenética elevou vários subgêneros de *Lutzomyia* a gêneros.

Dentre as espécies associadas à transmissão da leishmaniose visceral americana (LVA) destaca-se *Lutzomyia longipalpis* Lutz & Neiva 1912. Este vetor possui uma ampla distribuição, sendo encontrado desde o México até a Argentina e Uruguai (Young & Duncan 1994; Salomón et al. 2011) havendo um grau considerável de isolamento entre as suas numerosas populações devido a sua baixa mobilidade e à existência de barreiras geográficas e climáticas (Lanzaro et al. 1993; Alexander et al. 1998). Esta espécie já foi encontrada naturalmente infectada em diversas regiões da América do Sul (e.g. Feliciangeli et al. 2003; Paiva et al. 2010; Acardi et al. 2010; Missawa et al. 2010). Em virtude de sua importância epidemiológica, numerosos estudos foram conduzidos com o objetivo de esclarecer seu real status taxonômico. Experimentos de cruzamentos, assim como análises de sinais acústicos, feromônios e marcadores moleculares, foram realizados e reforçam a ideia de que *Lu. longipalpis* representa um complexo de espécies no Brasil (revisto em Bauzer et al. 2007; Maingon et al. 2008).

Outro vetor da LVA no Brasil é *Lutzomyia cruzi* Mangabeira 1938 (Galati et al. 1997; Santos et al. 1998; de Pita-Pereira et al. 2008). Esta espécie possui uma distribuição restrita, se concentrando na região centro-oeste do Brasil, e na Bolívia (Brazil et al. 2003, 2010). Diversos estudos demonstraram que *Lu. longipalpis* e *Lu. cruzi* são bastante relacionadas (Brazil & Hamilton 2002; Spiegel et al. 2004; Watts et

al. 2005; Pinto et al. 2010; Vigoder et al. 2010a). Fêmeas de ambas as espécies são indistinguíveis e os machos apresentam diferenças sutis na genitália (Young & Duncan 1994). O fato de que algumas populações de *Lu. longipalpis s.l.* apresentam divergência maior do que entre algumas destas e *Lu. cruzi* indica que está última deva ser considerada um membro do Complexo Longipalpis (Vigoder et al. 2010a).

Provavelmente o primeiro trabalho sugerindo que *Lu. longipalpis* deveria ser considerada um complexo de espécies foi publicado por Mangabeira (1969) (post mortem), quando este descreveu diferenças no fenótipo de pintas nos tergitos abdominais entre populações de *Lu. longipalpis*. Machos coletados no Estado do Ceará (Nordeste) possuíam duas pintas, uma no terceiro e outra no quarto tergito abdominal (sendo este fenótipo chamado de 2S) enquanto machos coletados no Pará (Norte) possuíam apenas uma pinta no quarto tergito abdominal (fenótipo chamado de 1S) (Figura 1). Nesse trabalho também se ressaltou o fato das duas formas serem encontradas em condições ecológicas distintas e, assim, se especulou que elas poderiam representar diferentes espécies ou variedades.

Em uma revisão feita por Bauzer e colaboradores (Bauzer et al. 2007), onde os principais trabalhos envolvendo esta espécie foram analisados, todos os estudos comparando amostras da América Central com a América do Sul sugerem que *Lu. longipalpis* de fato seja um complexo de espécies (e.g. Lanzaro et al. 1993; Arrivillaga et al. 2003). Contudo, no caso das populações brasileiras, até o final do século passado a controvérsia com relação à existência ou não de mais de uma espécie ainda era grande, já que a interpretação dos resultados obtidos com a análise de cruzamentos e feromônios de um lado (e.g. Ward et al. 1983, 1988) e isoenzimas do outro (Mukhopadhyay et al. 1998; Mutebi et al. 1999; de Azevedo et al. 2000) eram contraditórios.

De acordo com os autores contrários a ideia da existência de um complexo no Brasil, os seus estudos não detectaram uma divergência genética suficiente para indicar a presença de uma ou mais espécies entre as populações brasileiras (Mukhopadhyay et al. 1998; Mutebi et al. 1999; de Azevedo et al. 2000). Contudo, como apresentado abaixo, as evidências obtidas com diferentes tipos de análise (cruzamentos, feromônios, sinais acústicos e marcadores moleculares), apoiam fortemente a existência de um complexo no Brasil (Ward et al. 1983, 1988; Hamilton et al. 1999a,b, 2004, 2005; Souza et al. 2002, 2004; Bauzer et al. 2002a, b, 2007; Maingon et al. 2003, 2008; Bottecchia et al. 2004; Watts et al. 2005; Souza et al. 2008, Lins et al. 2008; Araki et al.

2009; Vigoder et al. 2010a).

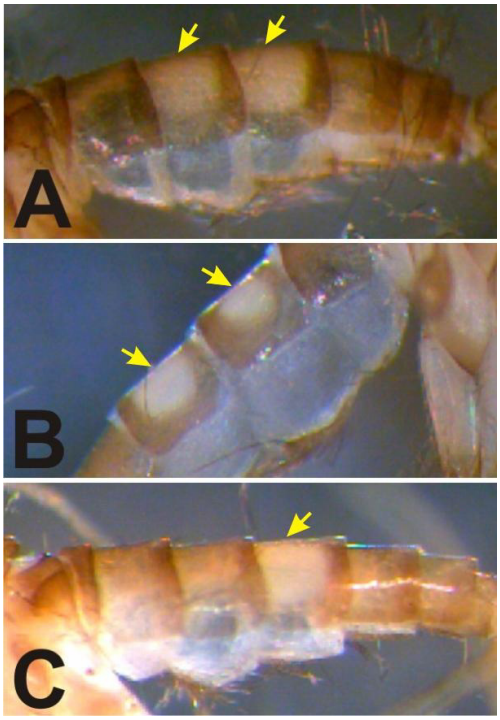


Figura 1 - Variação no número de pintas nos tergitos abdominais de machos de *Lu. longipalpis*: A – duas pintas (2S), B – intermediário e C – uma pinta (1S) (retirado de Araki 2009).

1.2 – Experimentos de cruzamentos.

Ward e colaboradores (1983) testaram, por análise de cruzamentos, populações brasileiras simpátricas e alopátricas que diferiam pelo número de pintas abdominais. Eles observaram que entre duas colônias oriundas de Sobral (Estado do Ceará) com machos possuindo uma (1) e duas (2) pintas havia baixa taxa de inseminação após a cópula, sugerindo que elas poderiam representar diferentes espécies crípticas em simpatria. Também se verificou uma baixa taxa de inseminação entre colônias contendo machos com duas pintas (2S) de Morada Nova (CE) e machos contendo uma pinta (1S) de Lapinha (Estado de Minas Gerais) e entre machos 1S da Ilha de Marajó e machos 1S de Lapinha. Esses resultados foram confirmados posteriormente, incluindo outras populações (Ward et al. 1988) e validaram a hipótese de que *Lu. longipalpis* representa um complexo de espécies no Brasil. Ward e colaboradores (1983, 1988) também verificaram a presença de uma forma intermediária entre o fenótipo 1S e 2S (uma

pequena pinta no terceiro tergito e uma pinta no quarto) (Figura 1) em altas frequências em algumas localidades, principalmente na região Nordeste, indicando um polimorfismo que também era verificado entre cruzamentos de algumas estirpes 1S e 2S (sugerindo ser este um caráter semidominante). Contudo, ainda que este fenótipo não seja espécie-específico, ele pode ser útil para identificar espécies simpátricas em algumas localidades onde os fenótipos intermediários sejam raros, como em Sobral (Ward et al. 1988) (ver abaixo).

Recentemente Souza e colaboradores (2008) realizaram novos experimentos de cruzamentos envolvendo populações simpátricas de Sobral (Sobral 1S e Sobral 2S) e alopátricas e Lapinha (MG), Jacobina (BA) e Natal (RN). Os resultados mostraram que após uma cópula todas as fêmeas que cruzaram com machos homoespecíficos e sobreviveram tempo suficiente para deixar ovos, produziram larvas viáveis. Contudo, os ovos de cruzamentos heteroespecíficos de populações alopátricas não eclodiram sugerindo falha na inseminação como observado por Ward e colaboradores (1983, 1988). Além disso, o isolamento entre as populações simpátricas de Sobral foi ainda maior já que nenhuma cópula foi observada em quase 180 ensaios de 20 minutos, sugerindo a evolução de reforço no isolamento reprodutivo nesta localidade (Souza et al. 2008).

Os resultados destes experimentos de cruzamentos não só suportam a hipótese da ocorrência de mais de uma espécie do complexo Longipalpis no Brasil como também indicam mecanismos de isolamento reprodutivo pré e pós-cópula entre estas populações. Isto levou a procura destes mecanismos, sendo os feromônios sexuais os primeiros a serem estudados neste complexo.

1.3 – Feromônios

A utilização de uma comunicação química como forma de transferência de informações entre e dentro de diferentes espécies já foi amplamente discutida e estudada para diversas classes de organismos, incluindo os insetos (e.g. Steiger et al. 2011). Além disso, a identificação, caracterização e síntese de feromônios promoveram uma profunda revolução no âmbito do controle de pragas e vetores uma vez que os pesquisadores puderam mimetizar substâncias naturalmente produzidas por esses animais para o seu

próprio controle, reforçando e estimulando investimentos em ecologia química (e.g. Reddy & Guerrero 2010).

Os estudos envolvendo a identificação e caracterização de feromônios sexuais produzidos pelos machos de *Lu. longipalpis* contribuíram bastante para a ideia de um complexo de espécies no Brasil. Ward e colaboradores (1988), estudando os feromônios de diferentes populações, encontraram uma correlação positiva entre estes e o isolamento reprodutivo observado em alguns cruzamentos. Análises posteriores mostraram que machos da localidade de Lapinha (MG), Sobral 1S (CE), Teresina (PI) e Barra de Guaratiba produzem o feromônio 9-metil-germacreno-B, machos de Jacobina (BA) produzem o 3-metil-α-himacaleno (+ α-himacaleno), machos de Natal (RN), Ilha de Marajó (PA), Sobral 2S (CE), Jaíba 2S (MG), Estrela de Alagoas 1S e 2S (AL) e Pancas (ES) produzem diterpenos do tipo cembreno-1 e em Jaíba 1S (MG) produzem o cembreno-2 (Ward et al. 1988; Hamilton et al. 1996, 1999a,b, 2004, 2005; Souza et al. 2002; Araki et al. 2009).

Além das populações do complexo no Brasil, *Lu. longipalpis* de outros países tiveram seus feromônios sexuais estudados. Em Posadas, Argentina e no Paraguai os machos produzem o feromônio 9-metil-germacreno-B, o mesmo produzido pelas populações brasileiras (Salomón et al. 2010; Brazil et al. 2009). Em Honduras, o feromônio predominante produzido por *Lu. longipalpis* é o 9-metil-germacreno-B. Já em Libéria, na Costa Rica, a situação é mais complexa, onde 3 tipos diferentes de feromônios foram encontrados: o 9-metil-germacreno-B, um tipo de homossesquiterpeno e um diterpeno, sugerindo que, nessa localidade, existem pelo menos 2, possivelmente 3, espécies do complexo ocorrendo em simpatria (Hamilton et al. 1996).

A grande maioria dos dados de feromônios se concentra em *Lu. longipalpis* e poucos estudos foram feitos com outras espécies de flebotomíneos. Dentre elas, duas espécies do complexo Longipalpis já descritas morfologicamente tiveram seus feromônios analisados: *Lu. cruzi* e *Lu. pseudolongipalpis* (uma espécie do complexo que ocorre na Venezuela), produzem, respectivamente, o 9-metil-germacreno-B (Brazil & Hamilton 2002), o mesmo encontrado em diversas populações de *L. longipalpis*, e 3-metil-α-himacaleno (Watts et al. 2005), semelhante ao encontrado na localidade de Jacobina.

Em sistemas de comunicação química de diversos taxa animais se especula, por parte de alguns autores, que ambos os sinais quanto as respostas a estes estão sob forte seleção estabilizadora (e.g. Brooks et al. 2005; Groot et al. 2006, 2010). Isso ocorre porque tanto o "sinalizador" quanto o "receptor" precisam estar finamente sintonizados para que o reconhecimento mútuo seja adequado, levando a população a convergir para a combinação sinal-resposta mais atrativa. Entretanto, alguns organismos apresentam uma variação intraespecífica significativa na sinalização envolvendo feromônios sexuais, como, por exemplo a mariposa *Heliothis subflexa* (Groot et al. 2010). Além disso, feromônios sexuais estão provavelmente sob seleção sexual e assim podem esta evoluindo rapidamente (Shirangi et al. 2009) e podem ter um papel importante em processos de especiação incipiente (e.g. Grillet et al 2012). Entre os membros que formam o complexo Longipalpis no Brasil, se observa uma grande variabilidade de feromônios sendo produzidos pelas diferentes populações estudadas.

Além dos feromônios sexuais, sinais acústicos produzidos pelos machos durante a cópula representam outro tipo comunicação entre os sexos que tem um papel potencialmente muito importante como mecanismo de isolamento reprodutivo no processo de especiação deste complexo.

1.4 – Análise do som de cópula produzido por machos de *L. longipalpis*

Outra forte evidência corroborando a existência de um complexo de espécies no Brasil foi obtida através da analise do som de cópula produzido por machos de *Lu. longipalpis* (Souza et al. 2002; 2004). Em *Drosophila* se sabe que o som produzido durante a corte aumenta a receptividade das fêmeas e foi descrito como um dos sinais que elas usam para reconhecer os machos da própria espécie (Kyriacou & Hall 1982, 1986; Ritchie et al. 1999). Deste modo, variações neste som de corte são importantes no isolamento reprodutivo entre espécies próximas de *Drosophila* (Kyriacou & Hall 1982, 1986; Ritchie et al. 1999). Em populações brasileiras de *Lu. longipalpis*, os dados de som e de feromônio estão parcialmente correlacionados (Souza et al. 2004; Bauzer et al. 2007), sendo dois sinais que podem ter um papel importante no isolamento reprodutivo das espécies do complexo .

Recentemente, Araki e colaboradores (2009) reuniram os dados de som de cópula e feromônios descritos até o presente momento para diferentes espécies de *Lu.*

longipalpis, que juntamente com os dados do gene *period* (ver abaixo) indicam a presença de dois grandes grupos no Brasil. Um grupo formado por populações distribuídas ao longo da costa do Espirito Santo até o Pará e formado pelas populações da Ilha de Marajó (PA), Sobral 2S (CE), Natal (RN) Estrela 2S (AL), Jaíba 2S (MG) e Pancas (ES), e outro, mais continental, formado pelas populações de Sobral 1S (CE), Teresina (PI), Estrela 1S (AL), Jacobina (BA), Jaíba 1S (MG) e Lapinha (MG). O primeiro se caracteriza por insetos que produzem essencialmente o mesmo tipo de som de cópula, o som classificado como do tipo *burst* e produzem o feromônio cembreno-1, representando provavelmente uma única espécie. Já no segundo grupo, observamos uma situação muito mais complexa, onde 5 variações do som de cópula do tipo pulsado (P1 a P5) são encontrados, assim como diferentes feromônios são produzidos (Figura 2) representando provavelmente 5 espécies incipientes do complexo.

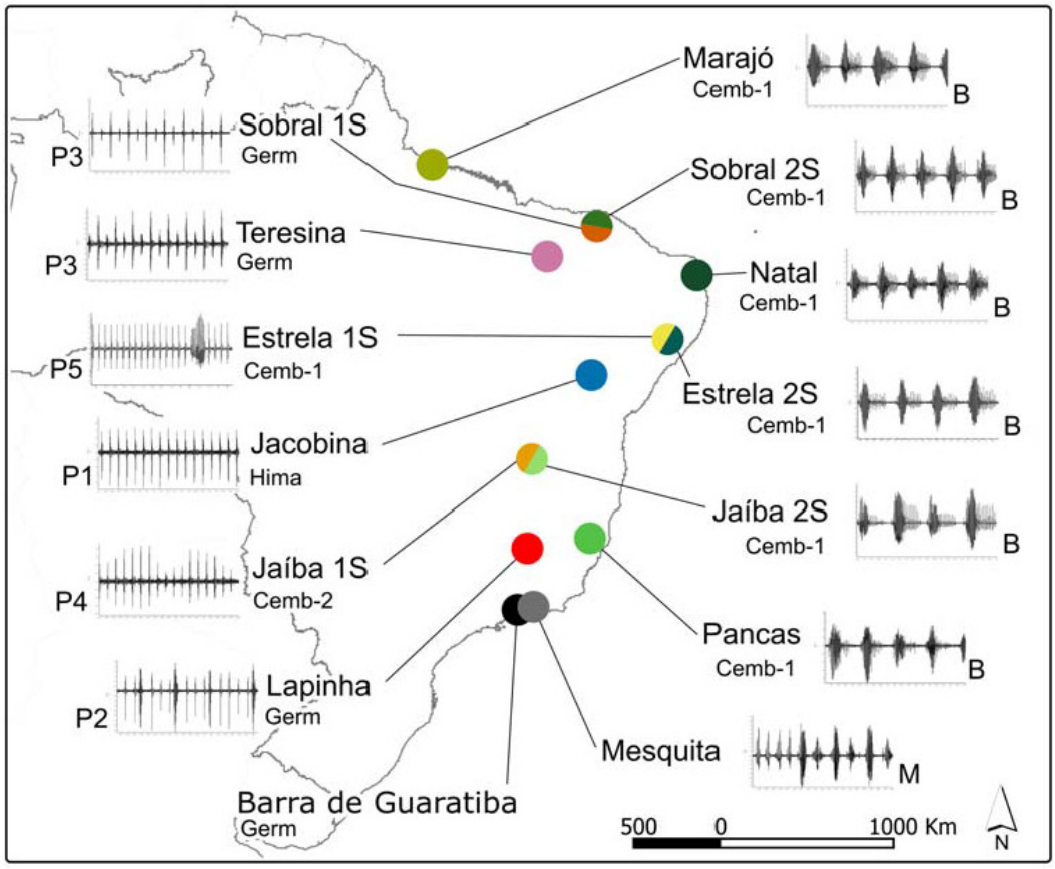


Figura 2 – Sumário das informações disponíveis para os diferentes sons de cópula e feromônios produzidos por populações de *Lu. longipalpis* no Brasil – Sons de cópula – *burst* (B) e pulsados (P1, P2, P3, P4 e P5); Feromônios – Cemb-1 e Cemb-2 – cembrenos, Germ – 9-metil-germacreno-B e Hima – himacaleno (Araki et al. 2009).

A população de Barra de Guaratiba (RJ) ainda não teve seu som analisado,

entretanto, as análises de feromônio demonstraram que machos dessa região produzem o 9-metil-germacreno-B, mas com quantidades significativas de cembreno. Em Mesquita, também no Rio de Janeiro, análises de som indicaram a existência de um novo tipo de som de cópula, chamado de misto (M), formado por características combinadas do som pulsado (P) e *burst* (B).

Outra espécie do complexo que teve seu som de cópula estudado foi *Lu. cruzi*. Em um recente estudo envolvendo esta espécie (Vigoder et al. 2010a) foi verificado que essa produz o som de cópula do tipo *burst*.

A utilização de marcadores genéticos que possam estar diretamente associados ao processo de especiação se torna particularmente importante no estudo de espécies crípticas. Dentre os marcadores mais interessantes neste aspecto, estão os genes envolvidos no controle do comportamento sexual, como, por exemplo, os genes associados ao controle do som. Por este motivo, fragmentos destes genes, inicialmente identificados no inseto modelo *Drosophila melanogaster*, foram isolados e utilizados no estudo do complexo Longipalpis.

1.5 – Genes envolvidos no controle do som e o complexo Longipalpis

Diversos genes de *Drosophila* que controlam o som de corte já foram identificados e clonados (Hall 1994) e dentre estes se destacam *period (per)*, *cacophony (cac)*, e *paralytic (para)*. Além do som de corte de *Drosophila*, os genes *per*, *cac* e *para* estão associados a diversas outras funções: *per* é um fator de transcrição (repressor) que está envolvido no controle dos ritmos circadianos (Konopka & Benzer 1971; Hardin 2011), o gene *cac* codifica a subunidade α-1 de um canal de cálcio dependente de voltagem (Smith et al. 1998; Peixoto & Hall 1998) envolvido em neurotransmissão (Hille 1992), e *para* codifica a subunidade α de um canal de sódio (Loughney et al. 1989; Littleton & Ganetzky 2000) e está associado a resistência a inseticidas da classe dos piretróides (Pittendrigh et al. 1997). Estes são, potencialmente, excelentes marcadores genéticos no estudo do processo de especiação não só de *Drosophila* como também de outros insetos que utilizam sinais acústicos durante a corte e cópula, como é o caso de flebotomíneos vetores (Ward et al. 1988; Souza et al. 2002, 2004; Vigoder et al. 2010a, b, 2011), e por este motivo eles foram isolados em *Lu. longipalpis* (Peixoto et al. 2001).

A primeira evidência molecular mostrando que *Lu. longipalpis* é de fato um complexo de espécies no Brasil foi obtida com a análise do gene *per* (Bauzer et al. 2002a,b) que mostrou uma clara diferenciação não só entre populações alopátricas de Jacobina (BA), Lapinha (MG) e Natal (RN) (Bauzer et al. 2002a) como também entre as populações simpátricas 1S e 2S de Sobral (CE) (Bauzer et al. 2002b). O gene *per* foi posteriormente estudado em diversas outras populações de *Lu. longipalpis* e os resultados obtidos com a análise da diferenciação genética neste marcador foram consistentes com aqueles obtidos com os estudos de feromônios e som de cópula (Araki et al. 2009).

Os resultados obtidos com *per* foram confirmados com a análise de fragmentos dos genes *cac* e *para.* Estudos prévios utilizando a região IVS6 de *cac* mostraram que este serviria como um ótimo marcador molecular em estudos de genética de populações e especiação em flebotomíneos, pois inclui um íntron com alta variabilidade e divergência entre espécies próximas (Lins et al. 2002). Isto veio a se confirmar com a análise de populações naturais de *Lu. longipalpis* (Bottecchia et al. 2004). Já o gene *para* foi o único marcador a apresentar diferenças fixas entre as populações simpátricas de Sobral (Lins et al. 2008), mostrando o potencial deste gene para análises futuras do complexo Longipalpis, como será apresentado adiante.

1.6 - Introgressão no complexo Longipalpis

Os resultados obtidos com genes potencialmente envolvidos com o controle do som de cópula, mostrando a ocorrência de diferentes espécies do complexo Longipalpis no Brasil, foram corroborados com análise de microsatélites (Maingon et al. 2003; Watts et al. 2005). No entanto, os dados obtidos com *cac, per* e microsatélites sugerem também a ocorrência de introgressão entre alguns os membros do complexo. Esta observação tem importantes consequências epidemiológicas já que torna possível não só a passagem de genes envolvidos na competência vetorial entre uma espécie e outra, como também a disseminação de genes controlando a resistência a inseticidas (Weill et al. 2000) que estão sendo ou poderão ser utilizados no controle de *Lu. longipalpis* (e.g. De Silans et al. 1998). Além disso, a ocorrência de introgressão poderia explicar porque genes mitocondriais, que são marcadores particularmente susceptíveis a fluxo gênico entre espécies próximas nem sempre apresentaram diferenças consistentes entre as

espécies do complexo Longipalpis (Soto et al. 2001; Hodgkinson et al. 2003; Arrivillaga et al. 2002, 2003; Rocha et al. 2011).

Apesar da contribuição de estudos envolvendo análises de um ou poucos genes, estes muitas vezes tem demonstrado serem insuficientes para o entendimento mais completo acerca da estruturação e padrões de fluxo gênico e divergência em complexos de espécies. No caso do complexo Longipalpis alguns marcadores evidenciam introgressão e outros uma divergência genética mais acentuada. Esses resultados são esperados em um processo recente de especiação onde as barreiras ao fluxo gênico não estão totalmente consolidadas e é possível se observar regiões no genoma onde a introgressão ainda está presente. Ou seja, é possível se constatar a ocorrência de introgressão diferencial, causando um mosaico de divergência genética ao longo do genoma conforme já foi demonstrado para outros flebotomíneos e insetos como *Drosophila melanogaster* e *Anopheles gambiae* (Machado et al. 2002; Besansky et al. 2003; Slotman et al. 2005; Stump et al. 2005; Tripet et al. 2005; Wang-Sattler et al. 2007; Mazzoni et al. 2008) e mais recentemente sugerido para *Lu. longipalpis* (Araki et al. 2009). Por este motivo, abordagens envolvendo análises multilocus se apresentam mais informativas, principalmente quando processos recentes de especiação estão em ocorrência.

Em um estudo recente, Araki e colaboradores (em preparação) utilizaram uma abordagem multilocus para estimar e comparar os níveis de fluxo gênico e divergência genética em 21 loci nucleares entre espécies simpátricas de *Lu. longipalpis* da localidade de Sobral (CE) e duas espécies alopátricas das localidades de Lapinha (MG) e Pancas (ES). Pancas e Sobral 2S, como já mencionado, representam membros da mesma espécie, cujos machos produzem o som do tipo *burst* e o feromônio cembreno-1. Lapinha e Sobral 1S, por outro lado, são duas espécies próximas que produzem o mesmo feromônio, o germacreno e o som pulsado P2 e P3 respectivamente. Os resultados desse trabalho sugerem fortemente a ocorrência de fluxo gênico entre as espécies analisadas, com diferentes níveis de introgressão entre os loci. Os resultados também indicam que esta introgressão parece ser assimétrica, com um fluxo gênico estimado maior na direção das espécies que produzem o som *burst* e feromônio cembreno-1 e de cinco a dez vezes maior entre as populações simpátricas. Esses dados sugerem que a ocorrência de introgressão durante o processo de separação parece ter desempenhado um papel crucial na estruturação do genoma das espécies que compõem

o complexo Longipalpis.

1.7 Justificativa

Apesar dos recentes avanços no estudo de *Lu. longipalpis* muito resta entender em relação ao processo de especiação neste complexo de espécies. Se compararmos a *An. gambiae*, por exemplo, relativamente poucos estudos foram realizados, ainda que *Lu. longipalpis* seja um importante vetor em termos de saúde pública. Por se tratar de um processo provavelmente recente, os resultados encontrados até o momento apresentam uma situação bastante interessante, onde algumas espécies do complexo podem ser mais facilmente identificadas, tendo sido, por isso, descritas morfologicamente, como *Lu. cruzi* e *Lu. pseudolongipalpis*, ao passo que a existência de outras, como é o caso das espécies crípticas ocorrendo no Brasil, foram demonstradas através de análises combinadas de dados genéticos e comportamentais.

Nesta tese, procuramos ampliar nossos conhecimentos do complexo Longipalpis estendendo a análise inicial do gene *para* realizada com amostras de Sobral (Lins et al 2008) a outras populações brasileiras e de outros países das Américas do Sul e Central (artigos 1 e 2). Realizamos também uma análise multilocus de pares de espécies simpátricas utilizando genes que controlam o som (artigo 3).

2. OBJETIVOS

2.1 - Objetivos gerais

Os objetivos desta tese foram os de: (1) realizar uma análise macrogeográfica do complexo Longipalpis envolvendo diferentes populações brasileiras, da América Central e do Sul desse vetor utilizando o gene *para* como marcador molecular e (2) avaliar, em *Lu. longipalpis*, o grau de divergência genética e fluxo gênico de novos marcadores associados ao controle do som já caracterizados em *D. melanogaster* através de uma análise multilocus envolvendo espécies simpátricas do complexo, comparando com outros dados gerados para essas espécies.

2.2 - Objetivos específicos

2.2.1 – Ampliar as análises envolvendo o gene *paralytic* à populações de *Lu. longipalpis* já estudadas com o gene *period*;

2.2.2 – Avaliar o grau de estruturação deste vetor representado pela divergência genética e a existência de fluxo gênico entre outras populações de *Lu. longipalpis* através de uma análise macrogeográfica envolvendo amostras de diferentes regiões brasileiras, da América Central e do Sul;

2.2.3 – Verificar a existência de introgressão e estimar a divergência genética entre espécies simpátricas de *Lu. longipalpis* das localidades de Sobral (CE), Jaíba (MG) e Estrela de Alagoas (AL), através de uma análise multilocus com marcadores associados ao controle do som.

3. APRESENTAÇÃO DO ARTIGO E MANUSCRITOS

Esta tese é composta por um artigo aceito para publicação (Capítulo 4) e dois artigos em preparação (Capítulo 5 e Capítulo 6).

Capítulo 4 – Lins RMMA, Souza NA, Brazil RP, Maingon RDC, Peixoto AA; **Fixed differences in the *paralytic* gene define two lineages within the *Lutzomyia longipalpis* species complex producing different types of lovesongs.** Artigo aceito para publicação na revista PLOS ONE: PONE-D-12-15761R2.

Página: 15

Capítulo 5 – Lins RMMA, Brazil RP, Souza NA, Ribolla PEM, Salomón OD, Lanzaro GC, Peixoto AA; **Macrogeographic analysis of the divergence in the *paralytic* gene among populations of the *Lutzomyia longipalpis* complex.** Manuscrito em preparação.

Página: 66

Capítulo 6 – Lins RMMA, Peixoto AA; **Multilocus analysis of sympatric sibling species of the *Lutzomyia longipalpis* complex using lovesong genes.** Manuscrito em preparação.

Página: 112

CAPÍTULO 4: "Diferenças fixas no gene *paralytic* definem duas linhagens que produzem diferentes sons de cópula no complexo de espécies *Lutzomia longipalpis*"

Lins RMMA, Souza NA, Brazil RP, Maingon RDC, Peixoto AA; **Fixed differences in the *paralytic* gene define two lineages within the *Lutzomyia longipalpis* species complex producing different types of lovesongs.** Artigo aceito para publicação: PONE-D-12-15761R2.

Este artigo foi aceito para publicação na revista PLOS ONE: PONE-D-12-15761R2 e é referente ao objetivo específico 2.2.1.

No presente artigo o gene *paralytic*, associado ao controle dom som, foi utilizado para estudar populações brasileiras de *Lu. longipalpis* e uma amostra de *Lu. cruzi*, analisadas anteriormente com o gene *period*, também associado ao controle do som. Nesse estudo se comprova a existência de dois grandes grupos de populações brasileiras produzindo diferentes sons de cópula e feromônios. Além da grande divergência genética encontrada entre esses dois grupos, foram encontradas diferenças fixas que definem, então, duas linhagens diferentes dentro do complexo.

Fixed differences in the *paralytic* gene define two lineages within the *Lutzomyia longipalpis* complex producing different types of courtship songs

Rachel MMA Lins[1], Nataly A Souza[2], Reginaldo P Brazil[3], Rhayza DC Maingon[4], Alexandre A Peixoto[1,5§]

[1] Laboratório de Biologia Molecular de Insetos, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil

[2] Laboratório de Transmissores de Leishmanioses, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil

[3] Laboratório de Bioquímica e Fisiologia de Insetos, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil

[4] Centre for Applied Entomology & Parasitology, Institute for Science & Technology in Medicine, School of Life Sciences, Keele University, Keele, United Kingdom

[5] Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular (INCT-EM), Rio de Janeiro, Brazil

§Corresponding author

Email addresses:

RMMAL: rlins@ioc.fiocruz.br

NAS: souzana@ioc.fiocruz.br

RPB: rpbrazil@ioc.fiocruz.br

RDCM: r.d.c.maingon@biol.keele.ac.uk

AAP: apeixoto@ioc.fiocruz.br

Abstract

The sand fly *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae), the most important vector of American visceral leishmaniasis, is widely distributed in Latin America. There is currently a consensus that it represents a species complex, however, the number and distribution of the different siblings is still uncertain. Previous analyses have indicated that Brazilian populations of this vector can be divided into two main groups according to the type of courtship song (Burst vs. Pulse) males produce during copulation. Nevertheless, no diagnostic differences have been observed between these two groups with most molecular markers used to date. We analyzed the molecular divergence in a fragment of the *paralytic* (*para*) gene, a locus involved in the control of courtship songs in *Drosophila*, among a number of *Lu. longipalpis* populations from Brazil producing Burst and Pulse-type songs. Our results revealed a very high level of divergence and fixed differences between populations producing the two types of songs. We also compared *Lu. longipalpis* with a very closely related species, *Lutzomyia cruzi*, which produces Burst-type songs. The results indicated a higher number of fixed differences between *Lu. cruzi* and the Pulse-type populations of *Lu. longipalpis* than with those producing Burst-type songs. The data confirmed our previous assumptions that the presence of different sibling species of the *Lu. longipalpis* complex in Brazil that can be divided into two main groups, one representing a single species and a second more heterogeneous group that probably represents a number of incipient species. We hypothesize that *para* might be one of the genes directly involved in the control of the

courtship song differences between these two groups or that it is linked to other loci associated with reproductive isolation of the Brazilian species.

Introduction

The study of species complexes provides an opportunity to investigate a number of unanswered questions about speciation [1]. Divergence between populations resulting in recent or incipient speciation can eventually lead to a number of molecular, behavioral and morphological changes, but very often these characters do not evolve at similar rates. This is particularly true in cases of cryptic speciation [2] where morphologically indistinguishable species can show striking behavioral differences, especially in aspects of courtship.

Acoustic communication is an important aspect of sexual behavior in a large number of insects [3], including disease vectors (e.g. [4-6]), and it has also a role in the reproductive isolation of many closely related species. In *Drosophila,* for example, courtship song is usually species-specific, being one of the signals females use to recognize males of their own species (e.g. [7-10]). *Drosophila* studies have also identified a number of genes controlling features of courtship songs (reviewed by [11-12]).

Acoustic signals can be also useful as one of the markers in an integrative analysis for species identification where classic morphologic differences fail to differentiate incipient sibling species [13]. One example in blood-sucking insects involves study of male copulation songs in the *Lutzomyia longipalpis* species complex [14-16], the main neotropical vector of *Leishmania infantum*, the etiological agent of American visceral leishmaniasis (AVL) [17]. As the main vector of an important parasitic disease, the existence of cryptic species in this insect may have important epidemiologic consequences [18-19] since divergence caused by genetic drift and/or natural selection may affect genes controlling aspects of the disease vector potential,

resulting in sibling species that are more efficient as vectors than others as has been shown in the *Anopheles gambiae* species complex [20-21].

Although *Lu. longipalpis* is a species complex [22-25], the number and distribution of the different sibling species is still uncertain. Previous studies using a combination of crossing experiments [22,26], analyses of acoustic signals [14-16], male sex pheromones [16,27-30] and molecular markers including orthologues of *Drosophila* courtship song genes *period* (*per*) and *cacophony* (*cac*) [16, 31-33], and microsatellites [34-35], have indicated that the Brazilian populations of this vector can be divided into two main groups according to the type of copulation song (Burst vs. Pulse) and pheromones that males produce [16]. Males of the first group of populations produce Burst-type song and the diterpene Cembrene-1 pheromone and probably represent a single species while the second group consists of populations producing different subtypes of Pulse-type song in combination with different pheromones that probably represent a number of incipient species [16]. However, *Lu. longipalpis* genetic structure in Brazil is rather complex with evidence of incomplete reproductive isolation and introgression [16,22,33] and no observed diagnostic differences between these two groups in most molecular markers used so far that would allow for a rapid identification of the different species.

The only potential exception so far is the *paralytic* (*para*) gene, a locus also involved in the control of courtship songs in *Drosophila* [36], characterized by fixed differences between a pair of sympatric sibling species of the *Lu. longipalpis* complex from Sobral (Ceára State, Brazil), that produce different copulation songs and male sex pheromones [37]. In the present study, we have extended the analysis of the *para* gene to a number of other *Lu. longipalpis* populations from Brazil. In addition, we have also analyzed the differentiation between *Lu. longipalpis* and *Lutzomyia cruzi*, a closely

related species [38] that also acts as a vector of *Le. infantum* in a region of Brazil [39]. Analyses of copulation songs, pheromones and molecular markers have indicated that *Lu. cruzi* is another species of the *Lu. longipalpis* complex [35,40-41].

Methods

We analyzed samples of *Lu. longipalpis* from eight different Brazilian localities: Lapinha, Minas Gerais State; Jaíba, Minas Gerais State; Jacobina, Bahia State; Pancas, Espírito Santo State; Estrela de Alagoas, Alagoas State; Natal, Rio Grande do Norte State; Marajó Island (Salvaterra), Pará State; and Teresina, Piauí State (Figure 1). A permit for sand fly collection in Brazil was obtained from the Brazilian Ministry of Environment (SISBIO #26066-1). Sand flies were captured using CDC light-traps near human habitation with permission from local homeowners. In addition, the collections were usually supported by the local vector surveillance authorities from local State Health Departments. Male *Lu. longipalpis* are characterized by polymorphism in the number of abdominal spots [22]: although this phenotype cannot be used to identify different allopatric species of the complex, it can be useful in some cases of sympatry, as previous work in Sobral (reviewed in [19,25]) and, more recently, in Estrela de Alagoas and Jaíba [16] has shown. In these three localities, the sympatric one spot (1S) and two spot (2S) males produce different copulation songs (Pulse-type and Burst-type, respectively) and represent different species [16]. Therefore, these samples were analyzed separately. Males from Natal that are highly polymorphic for the number of spots including very high numbers of intermediate forms which are rare in Sobral, Estrela de Alagoas and Jaíba, and Pancas (1S) produced Burst-type song, while males of Lapinha (1S), Jacobina (2S) and Teresina (1S) that represents a majority of this locality produce different subtypes of Pulse-type song [16]. We also analyzed a sample of *Lu.*

cruzi from Corumbá, Mato Grosso do Sul State and two males of *Lutzomyia pseudolongipalpis* from Curarigua, Venezuela, used as an outgroup in the genealogical analysis.

Genomic DNA was isolated according to [42] and the PCR Master Mix (Promega) was used to perform PCR according to [37]. PCR products were purified using the Wizard SV Gel and PCR Clean-up System (Promega) or GFX PCR DNA and Gel Band Purification Kit (GE Healthcare). Purified fragments were cloned using the pMOS*Blue* Blunt Ended Cloning Kit (GE Healthcare) or TOPO TA Cloning Kit (Invitrogen). Plasmid DNA was isolated using Flexiprep Kit (GE Healthcare) or using 96 well microplates and the alkaline lysis method [43] followed by filtration in Millipore Multiscreen filter plates. DNA sequencing was carried out with an ABI 3730 sequencer using the Big Bye 3.1 Kit (Applied Biosystems).

Lu. longipalpis para gene fragments from all populations were initially processed using BioEdit Sequence Alignment Editor [44] before population genetics analyses, which also included previously published sequences from Sobral [37]. A minimum of eight sequences per individual were aligned to obtain two consensus sequences corresponding to the two alleles, A and B, or one consensus sequence where flies were treated as homozygotes and the sequences were duplicated. The estimated probability of misclassifying a heterozygous fly as a homozygous with this procedure was less than 1%.

Both polymorphism and population structure analyses were carried out using DnaSP v5 [45] and Proseq 2.91 [46]. A Minimum Evolution tree based on p distances was estimated using MEGA5 [47]. All sequences were submitted to GenBank (accession numbers JQ359112-JQ359437). Analysis of molecular variance (AMOVA) was carried out with Arlequin 3.11 [48]. A non-recombinant block of the initial

fragment was obtained using the IMgc program [49] to construct the haplotype network with TCS v1.21 software [50].

Results

We analyzed a total 298 allele sequences from 149 males of a fragment of the *para* gene of *Lu. longipalpis* [37] of approximately 385 bp, including a variable sized intron of ~220 bp. Analyses included previously published and new sequences from the two Sobral sympatric sibling species [37] and new sequences from samples of the eight Brazilian localities analyzed here (Figure 1). Sympatric one spot (1S) and two spot (2S) males found in Estrela de Alagoas and Jaíba were analyzed separately since these males produce different copulation songs (Pulse and Burst, respectively) and represent different species, as previously observed in Sobral [16]. We also analyzed 24 allele sequences of *Lu. cruzi* males from Corumbá, State of Mato Grosso do Sul, a closely related sibling of the *Lu. longipalpis* complex producing Burst-type song [41], and used 4 sequences obtained from two males of *Lu. pseudolongipalpis*, a more distantly related sibling species [35,51], as an outgroup in the genealogical analysis (see below). Figure S1 shows the alignment of the whole fragment. Most of the variation was found within the intron, that included a number of indels. However, some rare non-synonymous substitutions were also observed.

Table 1 shows a summary of the polymorphisms for each population analyzed, excluding the regions with gaps. Populations of *Lu. longipalpis* were grouped according to the type of copulation song they produce: Pulse or Burst. The results showed that Lapinha was the least polymorphic among the Pulse song populations while Jacobina had the highest values of π and θ. Among the Burst song populations, Marajó and Jaiba 2S were the least and most polymorphic samples, respectively. Tajima´s D and Fu and

Li's D* and F* tests of neutrality [52-53] were performed for each population. Although one value was significant at a 5% level, all values were non-significant after Bonferroni correction.

Molecular differentiation analysis was performed for all pairwise comparisons involving the *Lu. longipalpis* populations, except for the small sample of Marajó. Again, the populations were grouped according to the song type they produce, Pulse or Burst. Table 2 shows the fixation indexes (Fst) as well as the number of fixed differences (Sf) in each comparison. The lowest pairwise Fst values were obtained between populations producing the same song type, while very high values of differentiation were observed in the comparisons involving populations producing either Burst or Pulse-type copulation songs. Indeed, fixed differences were found in those latter comparisons, except for the Estrela 1S sample. However, when sequences of a single fly (sequences Est1S8A and Est1S8B) were excluded from the analysis, this Pulse song population also showed fixed differences when compared to all other Burst song populations (numbers within brackets). Previous analysis by Araki et al. [16] suggested that the spot phenotype in this locality might not be as reliable for identifying the two sympatric sibling species as in Sobral [15,22,30,32-34,37]. However, *para* gene Fst values clearly confirm the presence of two sympatric species in Estrela, i.e. Estrela 1S and 2S (Table 2 and below).

Smaller sequence differences in *para* were observed between Burst-type populations than between Pulse-type populations. Indeed, the mean pairwise Fst value among Burst-type populations was 0.063 ± 0.067 compared with 0.147 ± 0.104 among Pulse-type populations. In contrast, the mean pairwise differentiation between populations with the two main song types was much higher (0.790 ± 0.044). These results were corroborated by AMOVA performed to examine the partition of *para*

sequence variation within *Lu. longipalpis* (Table 3). Most of the molecular variation (64.95%) was observed between the two main song types (Burst x Pulse), reflecting a clear separation between these groups. In addition, the results revealed a small part of this variation (7.0 %) distributed among populations within groups.

The same 383 bp *para* gene fragment studied in *Lu. longipalpis* was also amplified in *Lu. cruzi* from Corumbá, State of Mato Grosso do Sul. As shown in Table 1, *Lu. cruzi* showed levels of polymorphism in *para* that were similar to the lowest values observed among the *Lu. longipalpis* samples. Higher differentiation and fixed nucleotide differences between *Lu. cruzi* and all *Lu. longipalpis* populations with high Fst values (ranging from 0.7139 and 0.9271) were also observed (Table 2). However, a number of Fst values were smaller than comparisons between Burst-type and Pulse-type populations of *Lu. longipalpis*. Furthermore, *Lu. cruzi* that produces Burst-type songs showed two fixed differences compared with Burst-type populations and four to six differences compared with Pulse-type populations, whereas comparisons between the two song types of *Lu. longipalpis* displayed two to four fixed differences.

A Minimum Evolution tree including all *Lu. longipalpis* and *Lu. cruzi* sequences and those from the more distant sibling *Lu. pseudolongipalpis* (Figure 2) showed clear separation between the two main groups producing different copulation songs. In the tree, the two sequences (E1S8A and E1S8B) belonging to one Estrela de Alagoas 1S fly that were excluded from the Fst analysis (Table 2) clustered with the sequences corresponding to the Burst-type populations indicating that this individual probably represents a case where the spot phenotype did not match the correct song type in this locality. Sequences of *Lu. cruzi* that also produce Burst-type songs (light green circles) were grouped together with the *Lu. longipalpis* Burst-type sequences. As expected, *Lu. pseudolongipalpis* (open circles) were isolated from all other populations.

Finally, a haplotype network (Figure 3) was constructed based on a 249 bp non-recombinant fragment generated from the original segment of the *para* gene to avoid ambiguities due to recombinant events. A total of 40 haplotypes with 37 segregating sites were identified (Table S1) and a single network was generated using a 95% connection limit, except for *Lu. pseudolongipalpis*, which did not group in the same network.

The two main haplotypes generated were H13 and H28. Haplotype 13 corresponds to Burst-type populations and was composed of sequences of Sobral 2S, Estrela 2S, Natal and Pancas. Haplotype 28 was the most frequent haplotype of Pulse song populations from Sobral 1S, Jaíba 1S, Lapinha and Teresina. There was clear separation between the two groups producing different song types. These groups were connected by a single mutation between H11 and H4. H11 represents sequences of Sobral 2S, Pancas, Estrela 2S and Marajó. Interestingly, most of the sequences corresponding to H4 are from Estrela 1S, whose males produced the same type of pheromone, Cembrene-1 [16] found in populations with the H11 haplotype. In addition, nearly all *Lu. cruzi* haplotypes appeared as a separate cluster more closely related to the Burst-type populations of *Lu. longipalpis.*

Discussion

Understanding the structure of sibling species complexes is a difficult task for evolutionary biologists and this is particularly true in the case of cryptic species [2]. The lack of diagnostic morphological changes coupled with incomplete reproductive isolation and introgression, a common phenomenon among very closely related siblings [54-55], makes the identification and delimitation of the different species a difficult assignment.

Combined analyses using molecular markers, particularly the *per* gene [16] and microsatellites [34-35], and behavioral traits (songs and pheromones) strongly suggest that Brazilian *Lu. longipalpis* populations can be divided into two main groups according to the type of song (Burst vs. Pulse) males produce during copulation [16]. Fixed *para* gene differences between these two main lineages further support this notion. Indeed, the haplotype networks obtained with *per* [16] and *para* (Fig 3) showed a clear separation between the two population groups. In addition, although no fixed differences between the two lineages were observed in *per*, the pairwise divergence between *Lu. longipalpis* populations measured by Fst values in these two genes were highly correlated (Mantel test, $r= 0.819$, $p < 0.01$). Furthermore, both genes show a higher level of divergence among Pulse-type than among Burst-type song populations, consistent with the idea that the latter populations that produce the same song-type and the same pheromone (Cembrene-1) belong to a single species [16]. However, data from both genes indicate that the relationship among populations producing the different subtypes of Pulse-type song is more complex and heterogeneous. For example, males from Jacobina produce the P1 song and a combination of alpha-himachalene and 3-methyl-alpha-himachelene sex pheromones; Lapinha males produce the P2 song and 9-methyl-germacrene-B, (9MGB), sex pheromone; and Sobral 1S and Teresina produce the same P3 song associated with 9MGB sex pheromone [16]. Jaíba 1S males produce the P4 song and Cembrene-2 sex pheromone whereas in Estrela, 1S males produce the P5 song and the Cembrene-1 sex pheromone. Thus, combined molecular and behavioral data strongly suggest that these populations belong to five different incipient sibling species [16]. Indeed, for at least one pair of Pulse-type song populations (Jacobina and Lapinha) crossing experiments [26] and cytogenetic analysis [56] support this hypothesis.

Comparative *para* and *per* data ([41], this study) also suggest that *Lu. cruzi* is another member of the *Lu. longipalpis* complex. However, *per* analysis indicated higher genetic differentiation between *Lu. cruzi* and Burst-type song populations where the present results with *para* showed a higher Fst value between the former and Pulse-type populations. *Lu. cruzi* males produce a variation of the Burst-type song with shorter inter-burst intervals [41] and the 9MGB sex pheromone [40] also found in many Pulse-type populations of *Lu. longipalpis* [16]. Considering that *Lu. cruzi* males produce Burst-type song, it is tempting to speculate that *para* might be an important genetic determinant of song type (Burst vs. Pulse) between the two groups of *Lu. longipalpis* populations. Alternatively, *para* and *per* might be linked, with different levels of linkage disequilibrium and/or ancestral polymorphisms, to other loci associated with the reproductive isolation between the Brazilian sibling species.

The *D. melanogaster* courtship song genes are involved in a number of different molecular functions (reviewed in [12]). The three song genes used so far to study the *Lu. longipalpis* complex, *para, cac* and *per* encode, respectively, a voltage-gated sodium channel, a voltage-gated calcium channel, and a transcriptional repressor primarily involved in the circadian clock. It is possible that future RNA interference experiments (e.g. [57]) will help to confirm the potential role of these and other song genes in controlling copulation song differences among *Lu. longipalpis* sibling species. In addition, playback experiments (e.g. [7-9,58]) should also be carried out to directly infer whether copulation songs are involved in mate choice and reproductive isolation.

Finally, our *para* data also confirm existence of three localities (Sobral, Jaiba and Estrela) where pairs of species carrying different spot phenotypes and producing either Burst-type or Pulse-type songs occur in sympatry [16]. The existence of fixed differences in *para*, allowing easy genotyping of females of the different species, will be

particularly useful in these three localities to investigate whether the Burst-type and Pulse-type song females show any differences in other aspects of behavior when they occur sympatrically. The study of such phenotypic differences among closely related or incipient vector species is necessary because of the evolutionary and epidemiological implications of traits such as host or habitat preferences that have potential roles in ecological speciation [59] and/or in vector capacity [20].

Acknowledgements

We are grateful to Dora Feliciangeli (BIOMED, Maracay, Venezuela) and Jan Fish (Keele University) for *Lu. pseudolongipalpis* sand flies and *para* clones, respectively. We would also like to thank Robson Costa da Silva for technical assistance at FIOCRUZ, Brazil, Gabriel E. M. Ferreira for helping with the figures, and two anonymous reviewers and the editor William Etges for helpful comments on the manuscript.

References

1. Marie Curie SPECIATION Network, Butlin R, Debelle A, Kerth C, Snook RR, et al. (2012) What do we need to know about speciation? Trends Ecol Evol 27: 27-39.
2. Bickford D, Lohman DJ, Sodhi NS, Ng PK, Meier R, et al. (2007) Cryptic species as a window on diversity and conservation. Trends Ecol Evol 22: 148-155.
3. Ewing AW (1989) Arthropod Bioacoustics. Ithaca, New York: Cornell

University Press.

4. Cator LJ, Arthur BJ, Harrington LC, Hoy RR (2009) Harmonic convergence in the love songs of the dengue vector mosquito. Science 323: 1077-1079.
5. Warren B, Gibson G, Russell IJ (2009) Sex recognition through midflight mating duets in *Culex* mosquitoes is mediated by acoustic distortion. Curr Biol 19: 485-491.
6. Pennetier C, Warren B, Dabire KR, Russell IJ, Gibson G (2010) "Singing on the wing" as a mechanism for species recognition in the malarial mosquito *Anopheles gambiae*. Curr Biol 20: 131-136.
7. Ritchie MG, Halsey EJ, Gleason JM (1999) *Drosophila* song as a species-specific mating signal and the behavioural importance of Kyriacou & Hall cycles in *D. melanogaster* song. Anim Behav 58: 649-657.
8. Saarikettu M, Liimatainen JO, Hoikkala A (2005) The role of male courtship song in species recognition in *Drosophila montana*. Behav Genet 35: 257-263.
9. Klappert K, Mazzi D, Hoikkala A, Ritchie MG (2007) Male courtship song and female preference variation between phylogeographically distinct populations of *Drosophila montana*. Evolution 61: 1481-1488.
10. Wen SY, Yamada H, Li YF, Kimura MT, Oguma Y, Sawamura K, Toda MJ (2011) Copulatory courtship behavior and sine song as a mate recognition cue in *Drosophila lini* and its sibling species. Zoolog Sci 28: 469-475.
11. Peixoto AA (2002) Evolutionary behavioral genetics in *Drosophila*. Adv Genet 47: 117-150.
12. Gleason JM (2005) Mutations and natural genetic variation in the courtship song of *Drosophila*. Behav Genet 35: 265-277.
13. Dayrat B (2005) Towards integrative taxonomy. Biol J Linnean Soc 85: 407-

415.

14. de Souza NA, Ward RD, Hamilton JG, Kyriacou CP, Peixoto AA (2002) Copulation songs in three siblings of *Lutzomyia longipalpis* (Diptera: Psychodidae). Trans R Soc Trop Med Hyg 96: 102-103.
15. Souza NA, Vigoder FM, Araki AS, Ward RD, Kyriacou CP, et al. (2004) Analysis of the copulatory courtship songs of *Lutzomyia longipalpis* in six populations from Brazil. J Med Entomol 41: 906-913.
16. Araki AS, Vigoder FM, Bauzer LG, Ferreira GE, Souza NA, et al. (2009) Molecular and behavioral differentiation among Brazilian populations of *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae). PLoS Negl Trop Dis 3:e365.
17. Lainson R, Rangel EF (2005) *Lutzomyia longipalpis* and the eco-epidemiology of American visceral leishmaniasis, with particular reference to Brazil: a review. Mem Inst Oswaldo Cruz 100: 811-827.
18. Lanzaro GC, Warburg A (1995) Genetic variability in phlebotomine sandflies: possible implications for leishmaniasis epidemiology. Parasitol Today 11: 151-154.
19. Maingon RD, Ward RD, Hamilton JG, Bauzer LG, Peixoto AA (2008) The *Lutzomyia longipalpis* species complex: does population sub-structure matter to Leishmania transmission? Trends Parasitol 24: 12-17.
20. Lehmann T, Diabate A (2008) The molecular forms of *Anopheles gambiae*: a phenotypic perspective. Infect Genet Evol 8: 737-746.
21. Cohuet A, Harris C, Robert V, Fontenille D (2010) Evolutionary forces on *Anopheles*: what makes a malaria vector? Trends Parasitol 26: 130-136.
22. Ward RD, Phillips A, Burnet B, Marcondes CB (1988) The *Lutzomyia*

longipalpis complex: reproduction and distribution. In: Service MW, editor. Biosystematics of Haematophagous Insects. Oxford: Oxford University Press. pp. 258-269.

23. Lanzaro GC, Ostrovska K, Herrero MV, Lawyer PG, Warburg A (1993) *Lutzomyia longipalpis* is a species complex: genetic divergence and interspecific hybrid sterility among three populations. Am J Trop Med Hyg 48: 839-847.
24. Arrivillaga J, Mutebi JP, Pinango H, Norris D, Alexandre B, et al. (2003) The taxonomic status of genetically divergent populations of *Lutzomyia longipalpis* (Diptera: Psychodidae) based on the distribution of mitochondrial and isozyme variation. J Med Entomol 40: 615-627.
25. Bauzer LG, Souza NA, Maingon RDC, Peixoto AA (2007) *Lutzomyia longipalpis* in Brazil: a complex or a single species? A mini-review. Mem Inst Oswaldo Cruz 102: 1-12.
26. Souza NA, Andrade-Coelho CA, Vigoder FM, Ward RD, Peixoto AA (2008) Reproductive isolation between sympatric and allopatric Brazilian populations of *Lutzomyia longipalpis s.l.* (Diptera: Psychodidae). Mem Inst Oswaldo Cruz 103: 216-219.
27. Hamilton JG, Hooper AM, Mori K, Pickett JA, Sano A (1999a) 3-Methyl-α-himachalene confirmed, and the relative stereochemistry defined, by synthesis as the sex pheromone of the sandfly *Lutzomyia longipalpis* from Jacobina, Brazil. Chem Commun 355-356.
28. Hamilton JG, Ibbotson HC, Hooper AM, Mori K, Picket JA, et al. (1999b) 9-Methylgermacrene-B confirmed by synthesis as the sex pheromone of the sandfly *Lutzomyia longipalpis* from Lapinha, Brazil, and the absolute stereochemistry defined as 9S. Chem Commun 2335-2336.

29. Hamilton JG, Brazil RP, Maingon R (2004) A fourth chemotype of *Lutzomyia longipalpis* (Diptera: Psychodidae) from Jaibas, Minas Gerais State, Brazil. J Med Entomol 41: 1021-1026.

30. Hamilton JG, Maingon RD, Alexander B, Ward RD, Brazil RP (2005) Analysis of the sex pheromone extract of individual male *Lutzomyia longipalpis* sandflies from six regions in Brazil. Med Vet Entomol 19: 480–488.

31. Bauzer LG, Souza NA, Ward RD, Kyriacou CP, Peixoto AA (2002a) The *period* gene and genetic differentiation between three Brazilian populations of *Lutzomyia longipalpis*. Insect Mol Biol 11: 315-323.

32. Bauzer LG, Gesto JS, Souza NA, Ward RD, Hamilton JG, et al. (2002b) Molecular divergence in the *period* gene between two putative sympatric species of the *Lutzomyia longipalpis* complex. Mol Biol Evol 19: 1624-1627.

33. Bottecchia M, Oliveira SG, Bauzer LG, Souza NA, Ward RD, et al. (2004) Genetic divergence in the *cacophony* IVS6 intron among five Brazilian populations of *Lutzomyia longipalpis*. J Mol Evol 58: 754-761.

34. Maingon RD, Ward RD, Hamilton JG, Noyes HA, Souza N, et al. (2003) Genetic identification of two sibling species of *Lutzomyia longipalpis* (Diptera: Psychodidae) that produce distinct male sex pheromones in Sobral, Ceará State, Brazil. Mol Ecol 12: 1879-1894.

35. Watts PC, Hamilton JG, Ward RD, Noyes HA, Souza NA, et al. (2005) Male sex pheromones and the phylogeographic structure of the *Lutzomyia longipalpis* species complex (Diptera: Psychodidae) from Brazil and Venezuela. Am J Trop Med Hyg 73: 734-743.

36. Peixoto AA, Hall JC (1998) Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of *Drosophila.* Genetics 148: 827-838.

37. Lins RM, Souza NA, Peixoto AA (2008) Genetic divergence between two sympatric species of the *Lutzomyia longipalpis* complex in the paralytic gene, a locus associated with insecticide resistance and lovesong production. Mem Inst Oswaldo Cruz. 103: 736-740.

38. Pinto IS, Filho JD, Santos CB, Falqueto A, Leite YL (2010) Phylogenetic relationships among species of *Lutzomyia*, subgenus *Lutzomyia* (Diptera:Psychodidae). J Med Entomol 47: 16-21.

39. de Pita-Pereira D, Cardoso MA, Alves CR, Brazil RP, Britto C (2008) Detection of natural infection in *Lutzomyia cruzi* and *Lutzomyia forattinii* (Diptera: Psychodidae: Phlebotominae) by *Leishmania infantum chagasi* in an endemic area of visceral leishmaniasis in Brazil using a PCR multiplex assay. Acta Trop. 107: 66-69.

40. Brazil RP, Hamilton JG (2002) Isolation and identification of 9-methylgermacrene-B as the putative sex pheromone of *Lutzomyia cruzi* (Mangabeira, 1938) (Diptera: Psychodidae). Mem Inst Oswaldo Cruz 97: 435-436.

41. Vigoder FM, Araki AS, Bauzer LG, Souza NA, Brazil RP, Peixoto AA (2010) Lovesongs and period gene polymorphisms indicate *Lutzomyia cruzi* (Mangabeira, 1938) as a sibling species of the *Lutzomyia longipalpis* (Lutz and Neiva, 1912) complex. Infect Genet Evol 10: 734-739.

42. Jowett T (1998) Preparation of nucleic acids. In: Roberts DB, editor. *Drosophila*: A practical approach. Oxford: IRL Press. pp. 347-371.

43. Sambrook D, Russell J (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor: CSHL Press. 2100 p.

44. Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and

analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41: 95-98.

45. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.
46. Filatov DA (2002) ProSeq: A software for preparation and evolutionary analysis of DNA sequence data sets. Mol Ecol Notes 2: 621–624.
47. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.
48. Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.0: An integrated software package for population genetics data analysis. Evol Bioinformatics Online 1: 47-50.
49. Woerner A, Cox M, Hammer M (2007) Recombination-Filtered Genomic Datasets by Information Maximization. Bioinformatics 23: 1851-1853.
50. Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9: 1657-1659.
51. Arrivillaga JC, Feliciangeli MD (2001) *Lutzomyia pseudolongipalpis*: the first new species within the longipalpis (Diptera: Psychodidae: Phlebotominae) complex from La Rinconada, Curarigua, Lara State, Venezuela. J Med Entomol 38: 783-790.
52. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.
53. Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. Genetics 133: 693-709.
54. Pinho C, Hey J (2010) Divergence with Gene Flow: Models and Data. Annu Rev

Ecol Evol Syst 41:215-230.

55. Hausdorf B (2011) Progress toward a general species concept. Evolution 65:923-931.

56. Yin H, Mutebi JP, Marriott S, Lanzaro GC (1999) Metaphase karyotypes and G-banding in sandflies of the *Lutzomyia longipalpis* complex. Med Vet Entomol 13:72-77.

57. Sant'Anna MR, Alexander B, Bates PA, Dillon RJ (2008) Gene silencing in phlebotomine sand flies: Xanthine dehydrogenase knock down by dsRNA microinjections. Insect Biochem Mol Biol 38:652-660.

58. Ritchie MG, Townhill RM, Hoikkala A (1998) Female preference for fly song: playback experiments confirm the targets of sexual selection. Anim Behav 56: 713-717.

59. Schluter D (2009) Evidence for ecological speciation and its alternative. Science 323: 737-741.

Figure Legends

Figure 1 – Map of Brazil with the approximate location of the studied samples.

Figure 2 – Minimum Evolution tree of sequences from Brazilian populations of *Lu. longipalpis* producing Burst-type (dark green circles) and Pulse-type songs (red circles), *Lu. cruzi* (light green circles) and the more distant sibling species *Lu. pseudolongipalpis* (open circles) used as outgroup. The sequences E1S8A and E1S8B are the only red circles that cluster with the Burst-type sequences. Bootstrap values based on 1000 replications (values below 50% are not shown).

Figure 3 - Haplotype network of Brazilian populations of *Lu. longipalpis* and *Lu. cruzi*. Each population is represented by a different color and each node represents a unique haplotype.

Table 1 – Polymorphism summaries of the *para* gene fragment from populations of *Lu. longipalpis* and *Lu. cruzi*.

Population	Song-type	n	S	π	θ	D_T	D*	F*
Sobral 1S	P	32	12 (12)	0.0033 (0.0032)	0.0080 (0.0079)	-2.0334*	-1.2001	-1.7165
Lapinha	P	28	2 (2)	0.0016 (0.0016)	0.0014 (0.0014)	0.3094	-0.7144	-0.4930
Jacobina	P	22	9 (9)	0.0051 (0.0050)	0.0067 (0.0066)	-0.8129	-1.2837	1.3311
Teresina	P	24	9 (9)	0.0037 (0.0036)	0.0065 (0.0064)	-1.4200	-2.4491	-2.4955
Jaíba 1S	P	24	6 (6)	0.0023 (0.0023)	0.0043 (0.0043)	-1.3944	-0.9729	-1.2699
Estrela 1S	P	22	5 (5)	0.0019 (0.0019)	0.0037 (0.0037)	-1.4525	-0.4601	-0.8577
Sobral 2S	B	28	8 (9)	0.0040 (0.0057)	0.0056 (0.0073)	-0.8846	0.0821	-0.2423
Estrela 2S	B	32	6 (7)	0.0038 (0.0056)	0.0041 (0.0058)	-0.1481	1.2092	0.9335
Jaíba 2S	B	24	11 (11)	0.0047 (0.0047)	0.0080 (0.0080)	-1.3831	-1.3688	-1.5994
Natal	B	24	11 (12)	0.0047 (0.0049)	0.0080 (0.0087)	-1.4081	-1.8366	-1.9906
Pancas	B	32	10 (11)	0.0047 (0.0051)	0.0067 (0.0074)	-0.9295	-1.3358	-1.4167
Marajó	B	6	1 (3)	0.0016 (0.0048)	0.0012 (0.0044)	-1.7188	1.0525	1.1577
Lu. cruzi	B	24	5 (5)	0.0019 (0.0018)	0.0037 (0.0036)	-1.4315	-2.1728	-2.2703

B. Burst type song; **P.** Pulse type song; **n**. number of sequences; **S**. number of segregating sites; **π**. nucleotide diversity; **θ**. neutral parameter based on the segregating sites; $\mathbf{D_T}$. Tajima test of neutrality. **D*** and **F***. Fu and Li's tests of neutrality. Numbers in parentheses represent the analysis of nucleotide diversity considering the regions with gaps. * $p<0.05$

Table 2 – Pairwise differentiation between Pulse-type and Burst-type populations of *Lu. longipalpis* and *Lu. cruzi*.

		[Pulse-type populations					] [Burst-type populations					]	
		S1S	**Lap**	**Jac**	**Ter**	**J1S**	**E1S**	**S2S**	**E2S**	**J2S**	**Natal**	**Pancas**	***Lu. cruzi***
Pulse-type populations	**S1S**		0.2077***	0.1261**	0.0315^{ns}	0^{ns}	0.6171*** (0.4914****)	0.7819****	0.8257****	0.8103****	0.8009****	0.8003****	0.8695****
	Lap	0		0.2915***	0.2633*	0.2007*	0.8079*** (0.6491***)	0.8285****	0.8704****	0.8504****	0.8455****	0.8443****	0.9083****
	Jac	**0**	**0**		0.1697**	0.1706**	0.3251** (0.2343**)	0.6899****	0.7414***	0.7335***	0.7146****	0.7152****	0.8002***
	Ter	0	0	0		0.0050^{ns}	0.6002*** (0.4938***)	0.7701****	0.8122***	0.7987***	0.7886***	0.7883****	0.8558****
	J1S	0	0	0	0		0.7031*** (0.5603***)	0.8026****	0.8458***	0.8282****	0.8209****	0.8200****	0.8871***
	E1S	0 (0)	1 (1)	0 (0)	0 (0)	1 (1)		0.8356*** (0.7312***)	0.8913** (0.7922**)	0.8627** (0.7741***)	0.8579** (0.7581***)	0.8561*** (0.7577****)	0.9271** (0.8504**)
Burst-type populations	**S2S**	3	4	2	3	4	3 (0)		0.0635^{ns}	0.2272**	0.0380^{ns}	0.0914*	0.7139****
	E2S	3	4	2	3	4	3 (0)	0		0.0744^{ns}	0^{ns}	0^{ns}	0.7761***
	J2S	3	4	2	3	4	3 (0)	0	0		0.0737^{ns}	0.0633*	0.7534****
	Natal	3	4	2	3	4	3 (0)	0	0	0		0^{ns}	0.7387****
	Pancas	3	4	2	3	4	3 (0)	0	0	0	0		0.7364****
	Lu. cruzi	5	6	4	5	6	5 (2)	2	2	2	2	2	

Upper right matrix – pairwise differentiation (Fst) and significance (P values were obtained with 10,000 random permutations). Lower left matrix – fixed differences between samples. S1S – Sobral 1S, Lap – Lapinha, Jac – Jacobina, Ter – Teresina, J1S – Jaíba 1S, S2S – Sobral 2S, E2S – Estrela 2S, J2S – Jaíba 2S. S1S – Sobral 1S. Lap – Lapinha. Jac – Jacobina. Ter – Teresina. J1S – Jaíba 1S. S2S – Sobral 2S. E2S – Estrela 2S. J2S – Jaíba 2S. Significance of pairwise Fst values was estimated with 10,000 random permutations. Values between brackets included the single E1S fly which probably represents a case were the spot phenotype does not match the song type in this population.

ns - non-significant; * p<0.05; **p<0.01; *** p<0.001; **** p<0.0001

Table 3 – AMOVA statistics

Source of variation	Percentage of variation
Among groups	64.95
Among populations within groups	7.00
Within populations	28.06
Fsc (haplotypes/populations within groups)	0.1996***
Fst (haplotypes/populations/groups)	0.7194***
Fct (populations/groups)	0.6495*

Copulation song groups: Burst-type: Sobral 2S, Estrela 2S, Jaíba 2S, Natal and Pancas; **Pulse- type:** Sobral 1S, Jaíba 1S, Estrela 1S, Lapinha, Jacobina and Teresina. Significance corresponding to the fixation indexes was obtained through 10,000 permutations. *$p<0.01$; *** $p<0.0001$

Supporting Information

Figure S1. Alignment of the *paralytic* gene whole fragment. Intron sequence is highlighted in grey and non-recombinant block used to construct the haplotype network is highlighted in yellow. Dots indicate the same nucleotide and dashes indicate gaps.

Table S1. Distribution of the 40 haplotypes found among *Lu. longipalpis* and *Lu. cruzi* samples, segregating sites within a 251 bp non-recombinant fragment and number of sequences represented in each sample.

Fig 1

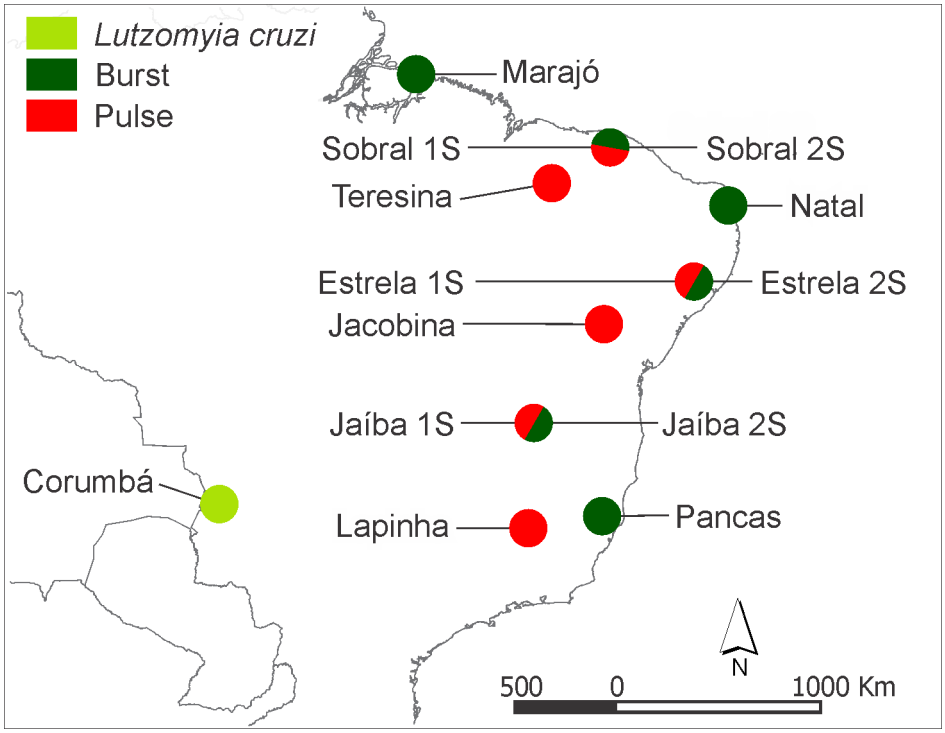

Lutzomyia cruzi
Burst
Pulse
Marajó
Sobral 1S
Sobral 2S
Teresina
Natal
Estrela 1S
Estrela 2S
Jacobina
Jaíba 1S
Jaíba 2S
Corumbá
Lapinha
Pancas
N
500
0
1000 Km

Fig 2

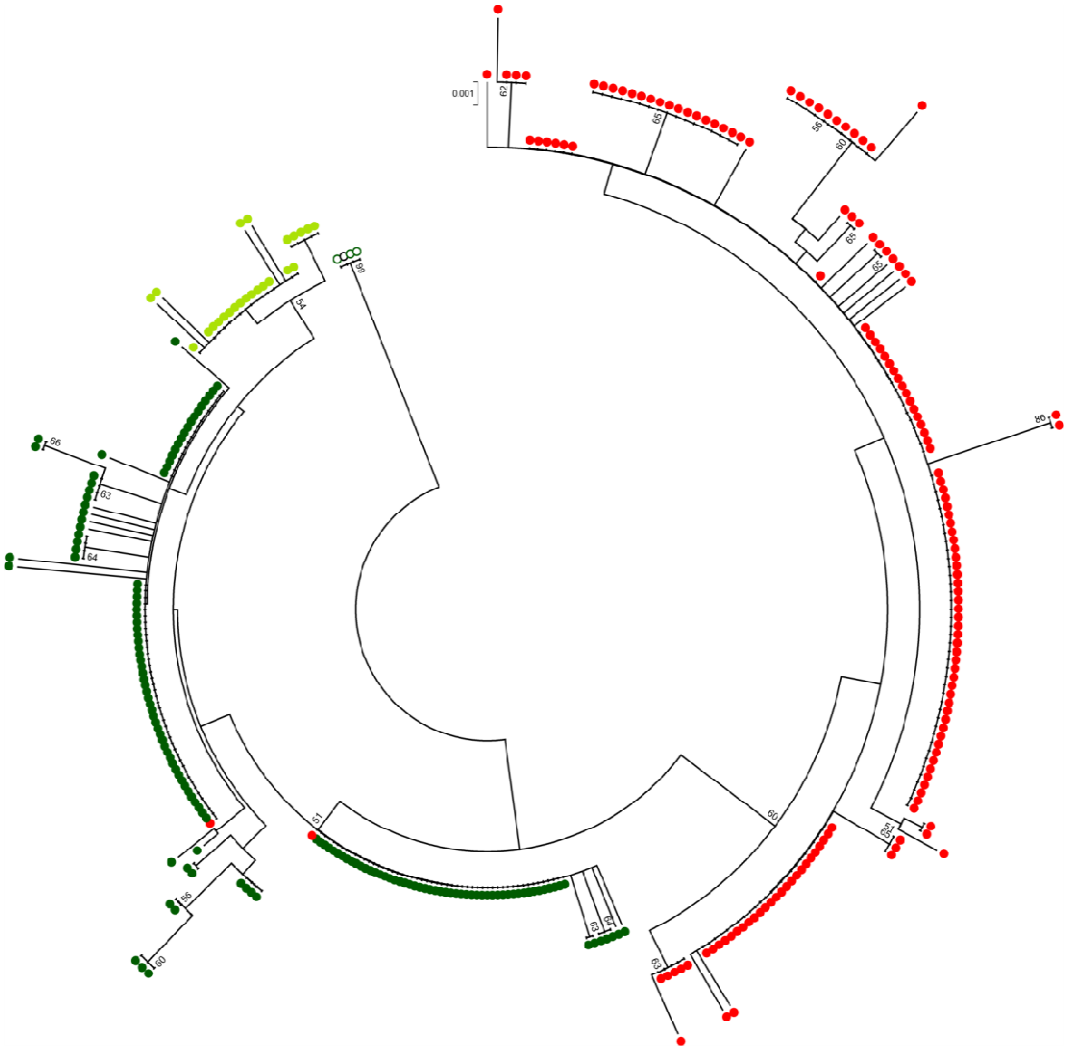


Fig 3

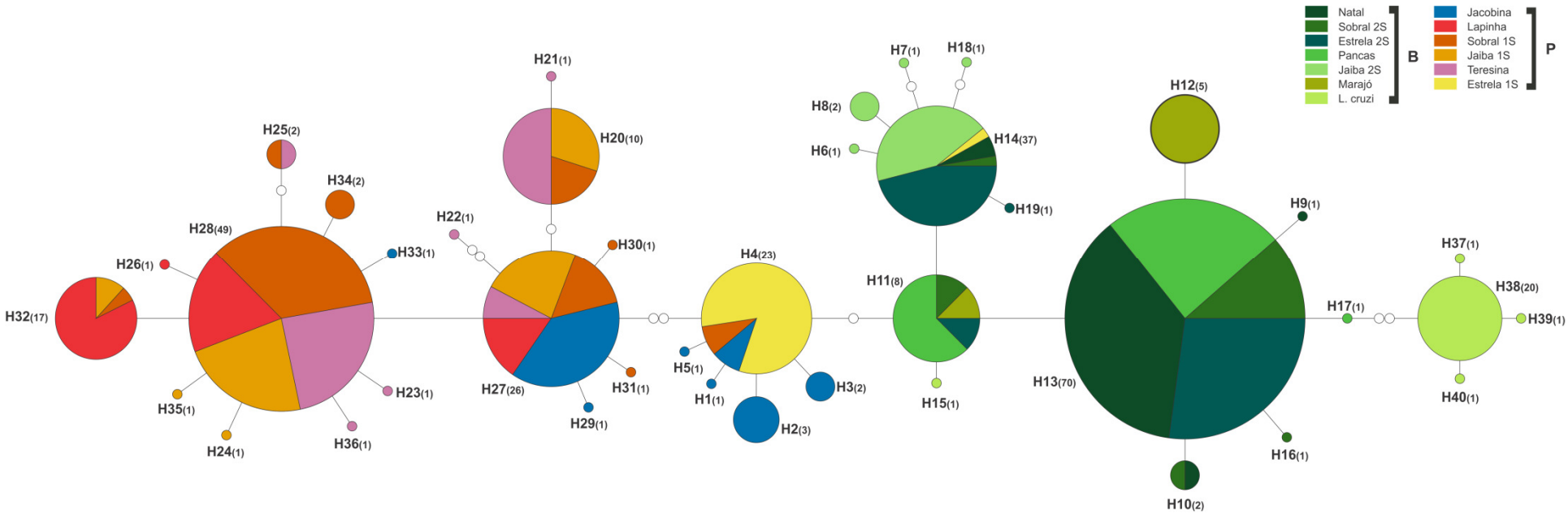
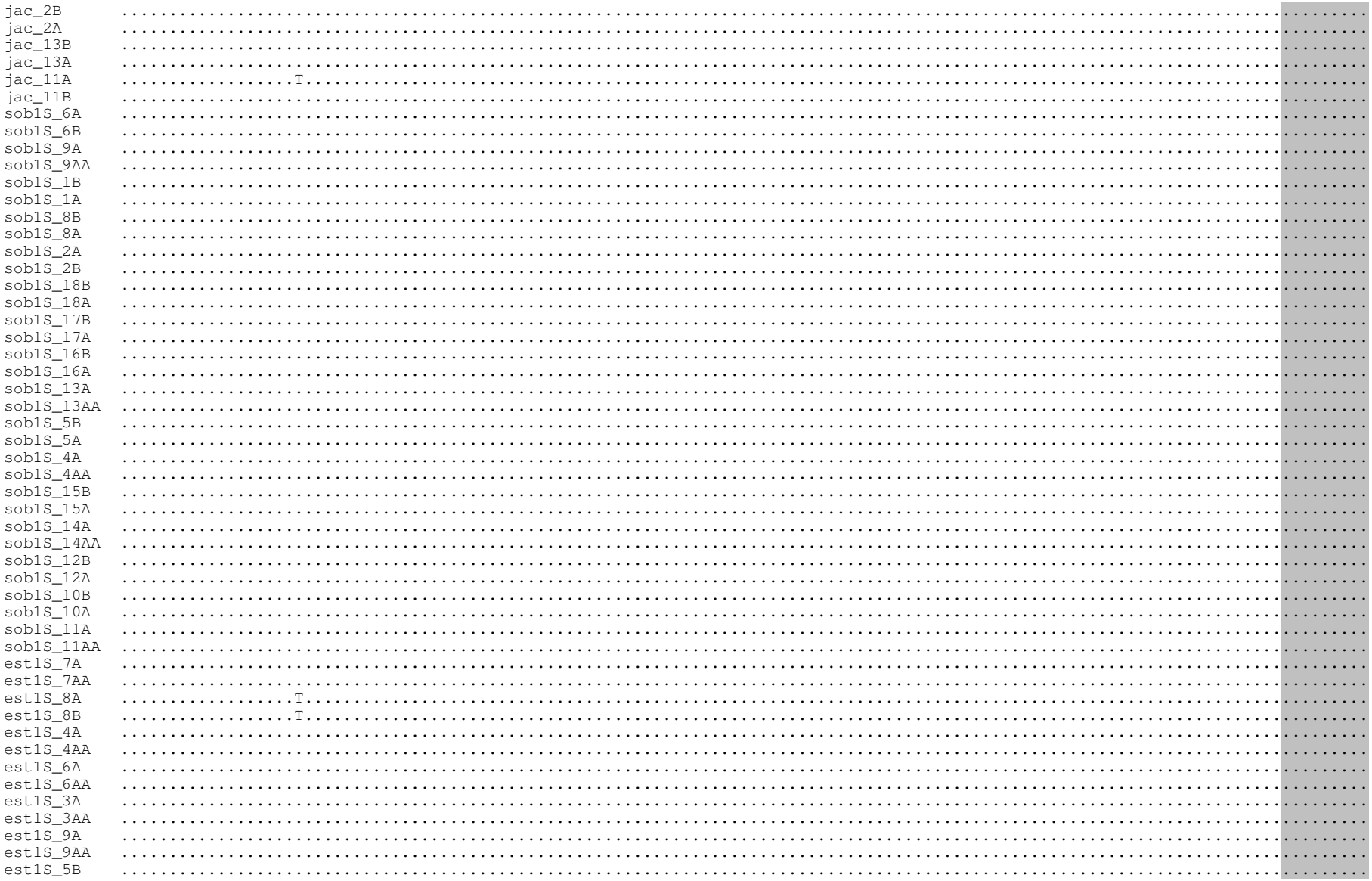

Natal
Sobral 2S
Estrela 2S
Pancas
Jaiba 2S
Marajó
L. cruzi
B
Jacobina
Lapinha
Sobral 1S
Jaiba 1S
Teresina
Estrela 1S
P
H1(1)
H2(3)
H3(2)
H4(23)
H5(1)
H6(1)
H7(1)
H8(2)
H9(1)
H10(2)
H11(8)
H12(5)
H13(70)
H14(37)
H15(1)
H16(1)
H17(1)
H18(1)
H19(1)
H20(10)
H21(1)
H22(1)
H23(1)
H24(1)
H25(2)
H26(1)
H27(26)
H28(49)
H29(1)
H30(1)
H31(1)
H32(17)
H33(1)
H34(2)
H35(1)
H36(1)
H37(1)
H38(20)
H39(1)
H40(1)

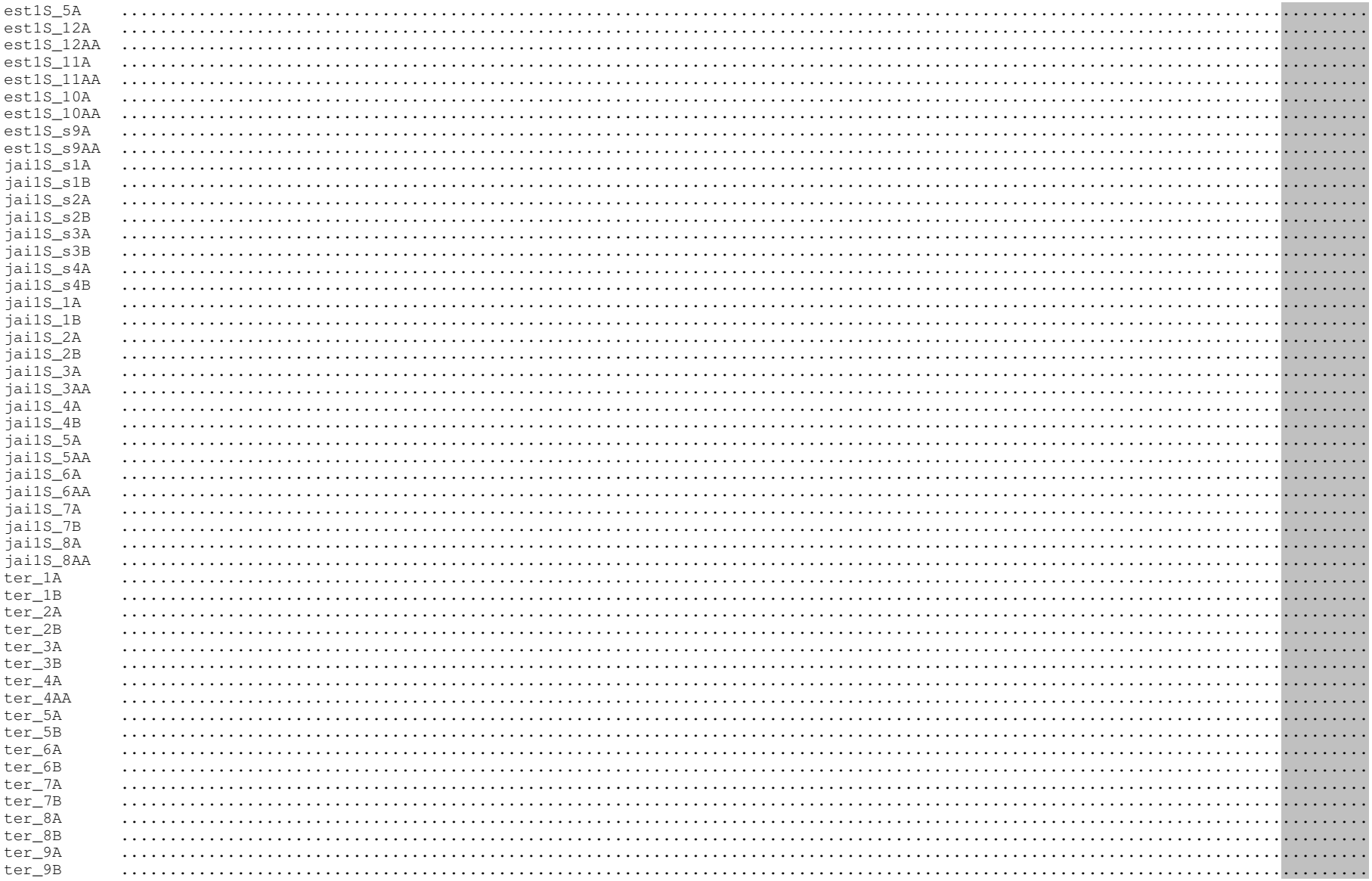
Table S1: Distribution of the 40 haplotypes found among *Lu. longipalpis* and *Lu. cruzi* samples, segregating sites within a 251-bp non-recombinant fragment and number of sequences represented in each sample.

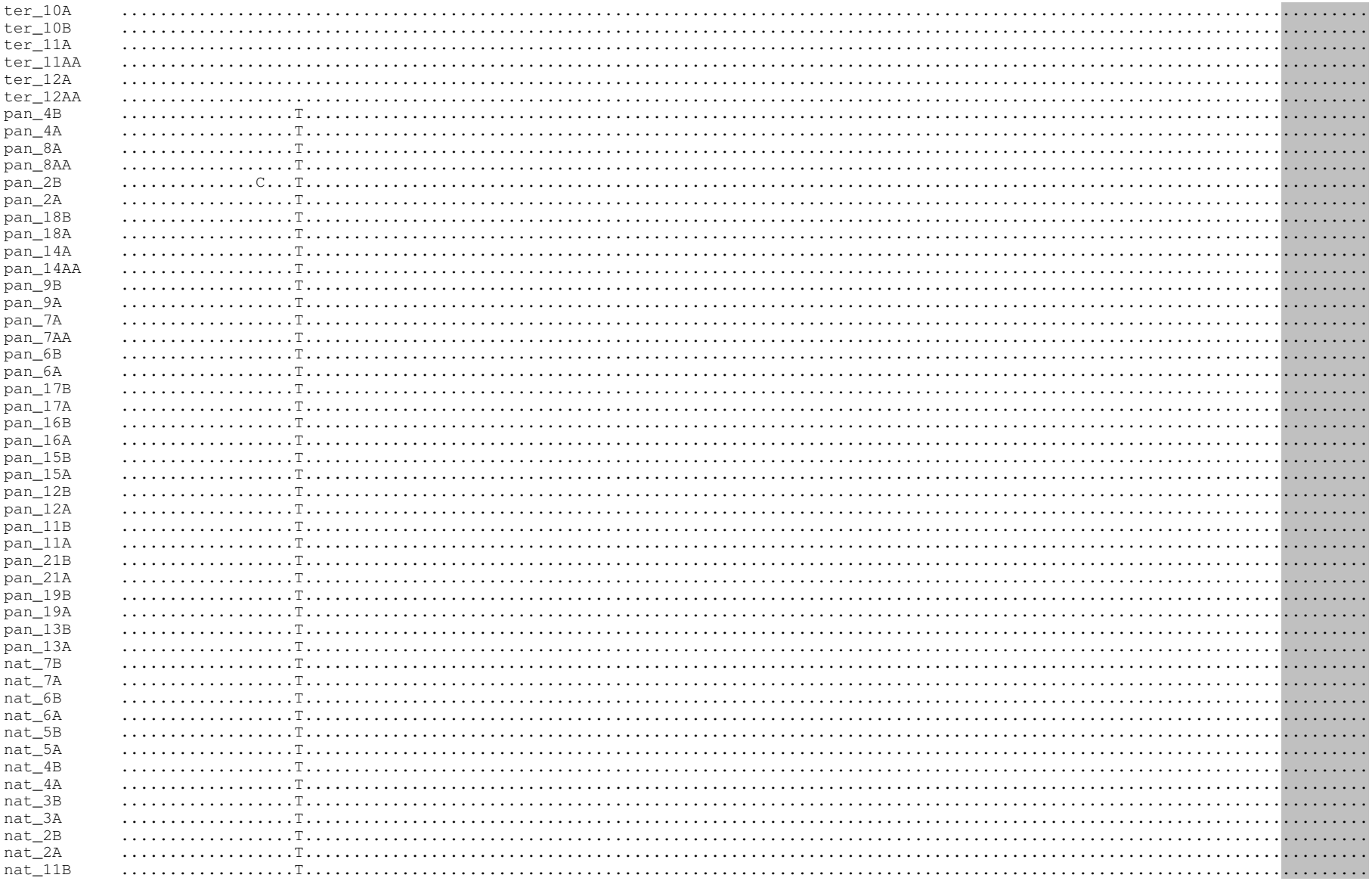
	Number of sequences represented in each sample													
H	**S1S**	**S2S**	**J1S**	**J2S**	**E1S**	**E2S**	**Lap**	**Jac**	**Nat**	**Pan**	**Ter**	**Mar**	***Lu. cruzi***	**Total**
H1								1						1
H2								3						3
H3								2						2
H4	2				19			2						23
H5								1						1
H6				1										1
H7				1										1
H8				2										2
H9									1					1
H10		2							1					2
H11		1				1				5		1		8
H12												5		5
H13		19				8			17	26				70
H14		1		16	1	17			2					37
H15													1	1
H16		1												1
H17										1				1
H18				1										1
H19						1								1
H20	2		3								5			10
H21											1			1
H22											1			1
H23											1			1
H24			1											1
H25	1										1			2
H26							1							1
H27	4		6				4	10			2			26
H28	17		11				9				12			49
H29								1						1
H30	1													1
H31	1													1
H32	1		2				14							17
H33								1						1
H34	2													2
H35			1											1
H36											1			1
H37													1	1
H38													20	20
H39													1	1
H40													1	1

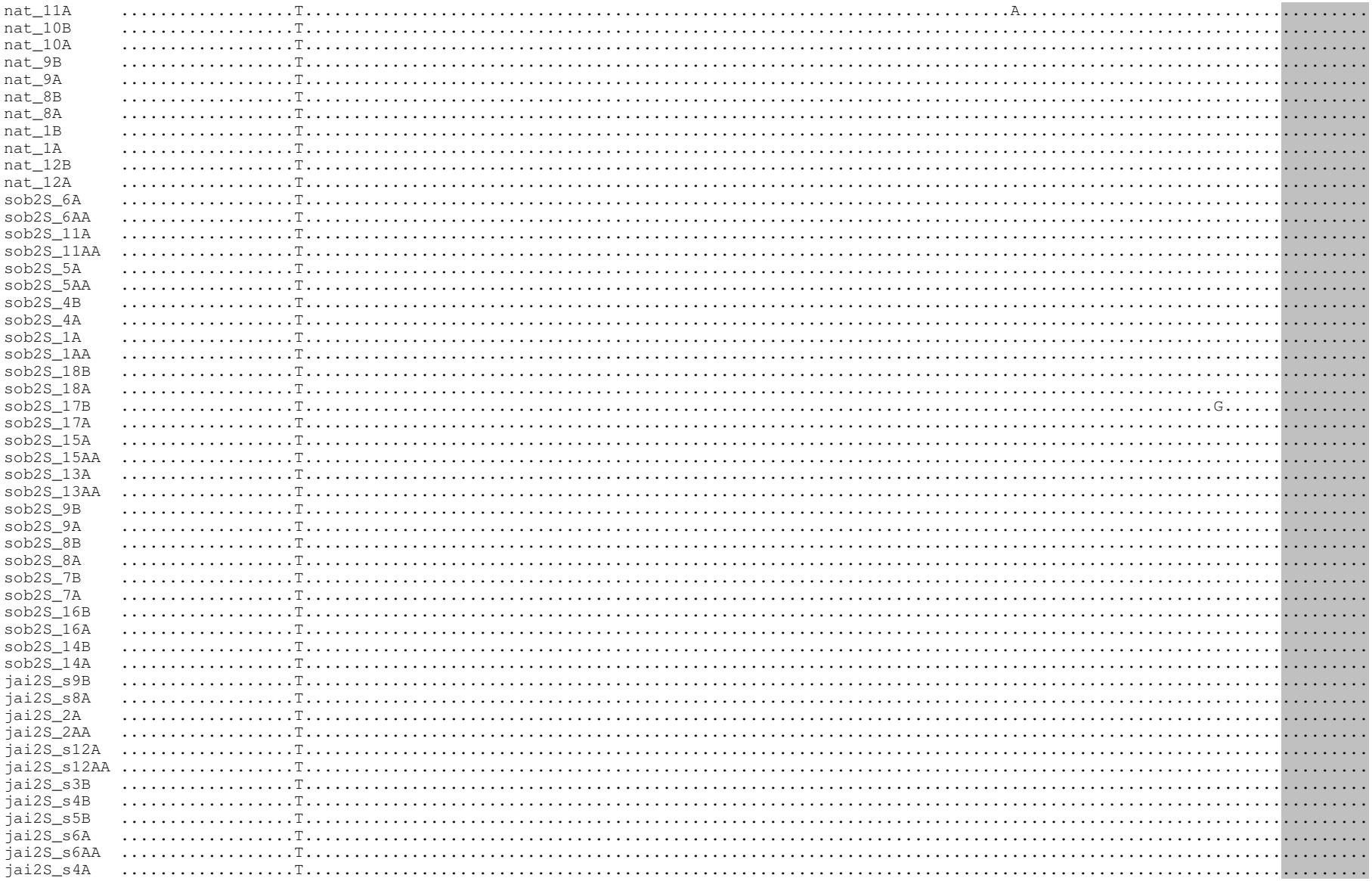
H: Haplotype, S1S: Sobral 1S, S2S: Sobral 2S, J1S: Jaíba 1S, J2S: Jaíba 2S, E1S: Estrela 1S, E2S: Estrela 2S, Lap: Lapinha, Jac: Jacobina, Nat: Natal, Pan: Pancas, Ter: Teresina, Mar: Marajó.

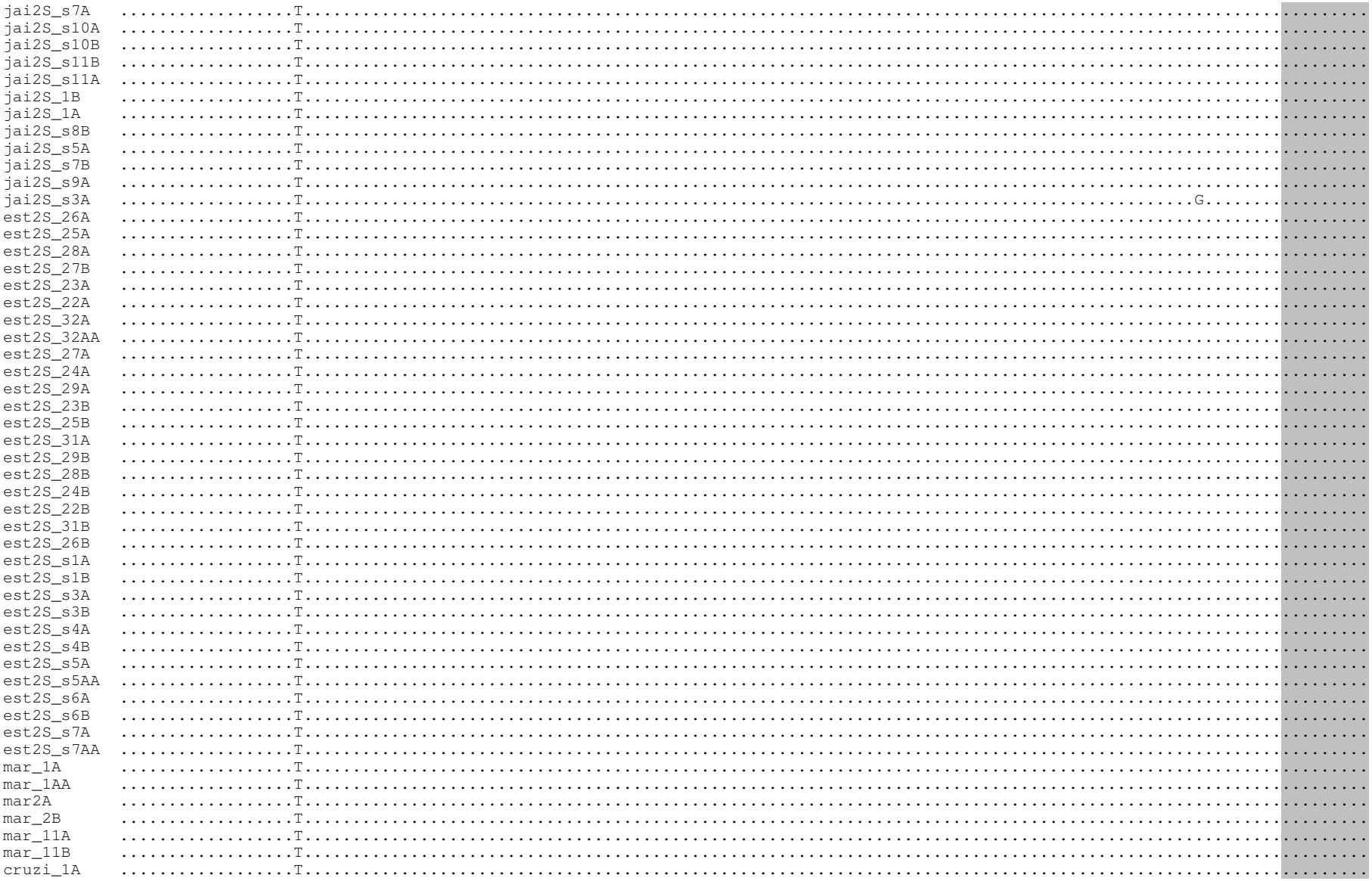
Figure S1: Alignment of the *paralytic* gene whole fragment. Intron sequence is highlighted in grey and non-recombinant block used to construct the haplotype network is highlighted in yellow. Dots indicate the same nucleotide and dashes indicate gaps.

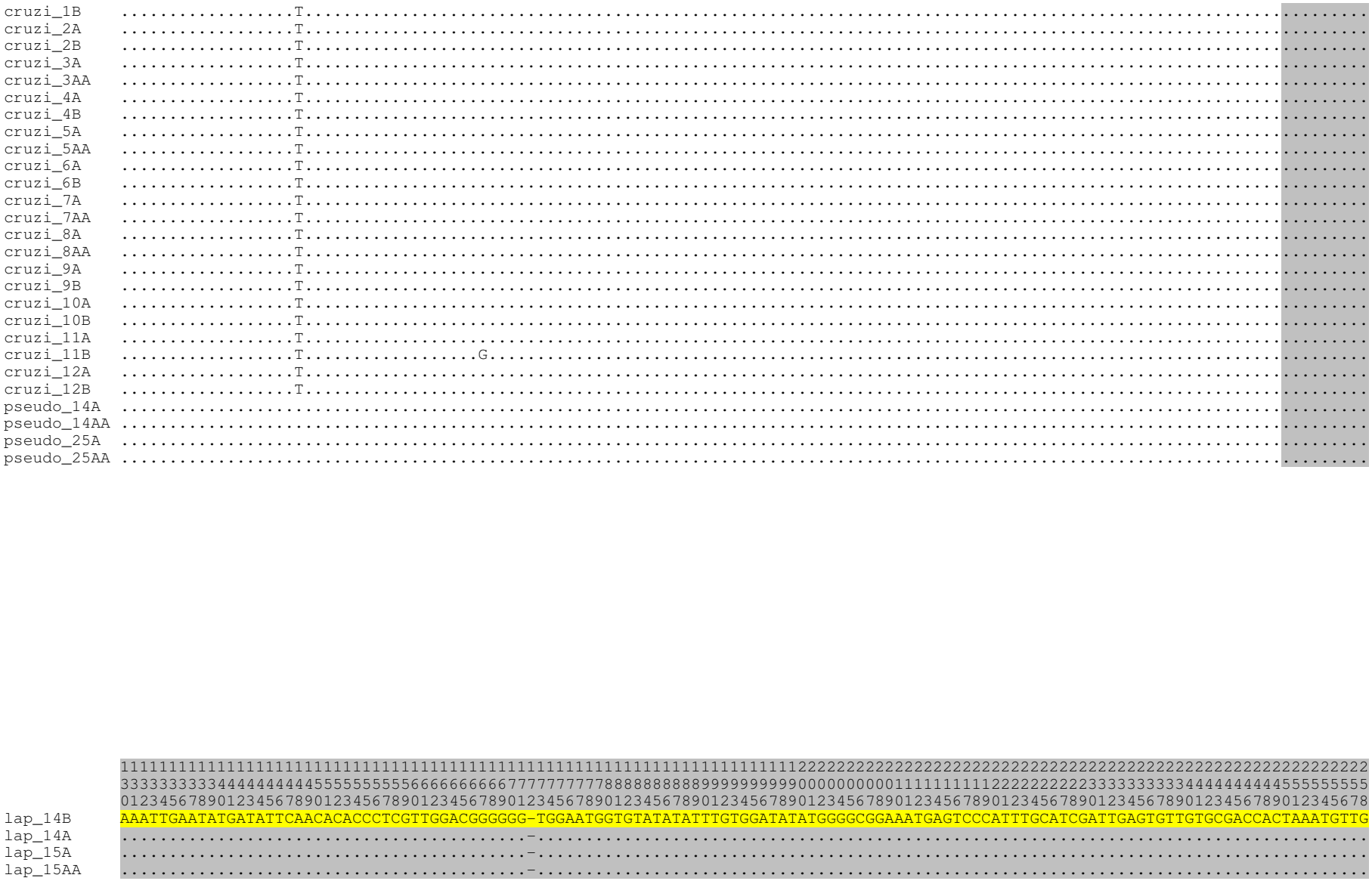

jac_2B
jac_2A
jac_13B
jac_13A
jac_11A
jac_11B
sob1S_6A
sob1S_6B
sob1S_9A
sob1S_9AA
sob1S_1B
sob1S_1A
sob1S_8B
sob1S_8A
sob1S_2A
sob1S_2B
sob1S_18B
sob1S_18A
sob1S_17B
sob1S_17A
sob1S_16B
sob1S_16A
sob1S_13A
sob1S_13AA
sob1S_5B
sob1S_5A
sob1S_4A
sob1S_4AA
sob1S_15B
sob1S_15A
sob1S_14A
sob1S_14AA
sob1S_12B
sob1S_12A
sob1S_10B
sob1S_10A
sob1S_11A
sob1S_11AA
est1S_7A
est1S_7AA
est1S_8A
est1S_8B
est1S_4A
est1S_4AA
est1S_6A
est1S_6AA
est1S_3A
est1S_3AA
est1S_9A
est1S_9AA
est1S_5B

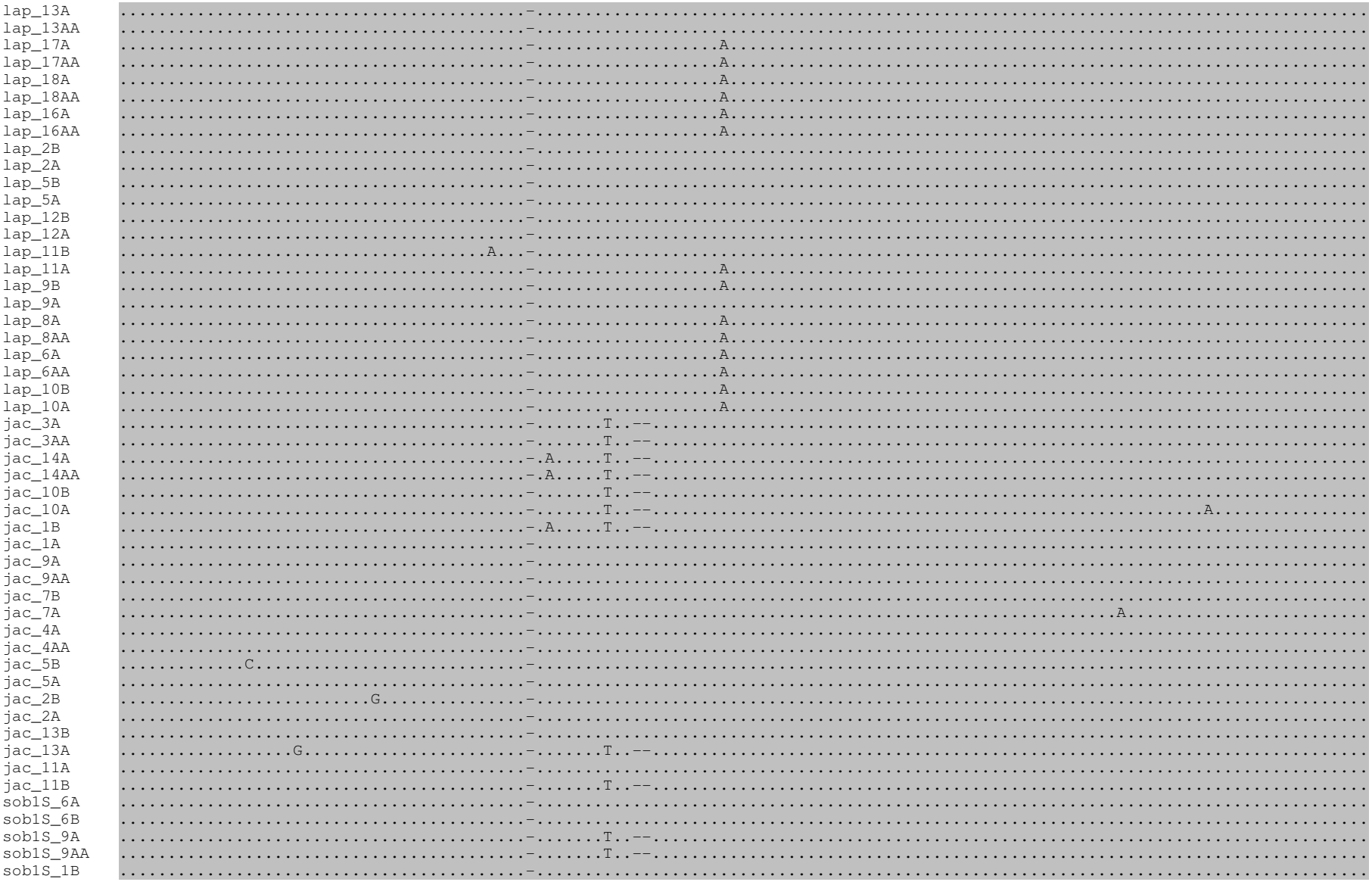

est1S_5A
est1S_12A
est1S_12AA
est1S_11A
est1S_11AA
est1S_10A
est1S_10AA
est1S_s9A
est1S_s9AA
jai1S_s1A
jai1S_s1B
jai1S_s2A
jai1S_s2B
jai1S_s3A
jai1S_s3B
jai1S_s4A
jai1S_s4B
jai1S_1A
jai1S_1B
jai1S_2A
jai1S_2B
jai1S_3A
jai1S_3AA
jai1S_4A
jai1S_4B
jai1S_5A
jai1S_5AA
jai1S_6A
jai1S_6AA
jai1S_7A
jai1S_7B
jai1S_8A
jai1S_8AA
ter_1A
ter_1B
ter_2A
ter_2B
ter_3A
ter_3B
ter_4A
ter_4AA
ter_5A
ter_5B
ter_6A
ter_6B
ter_7A
ter_7B
ter_8A
ter_8B
ter_9A
ter_9B

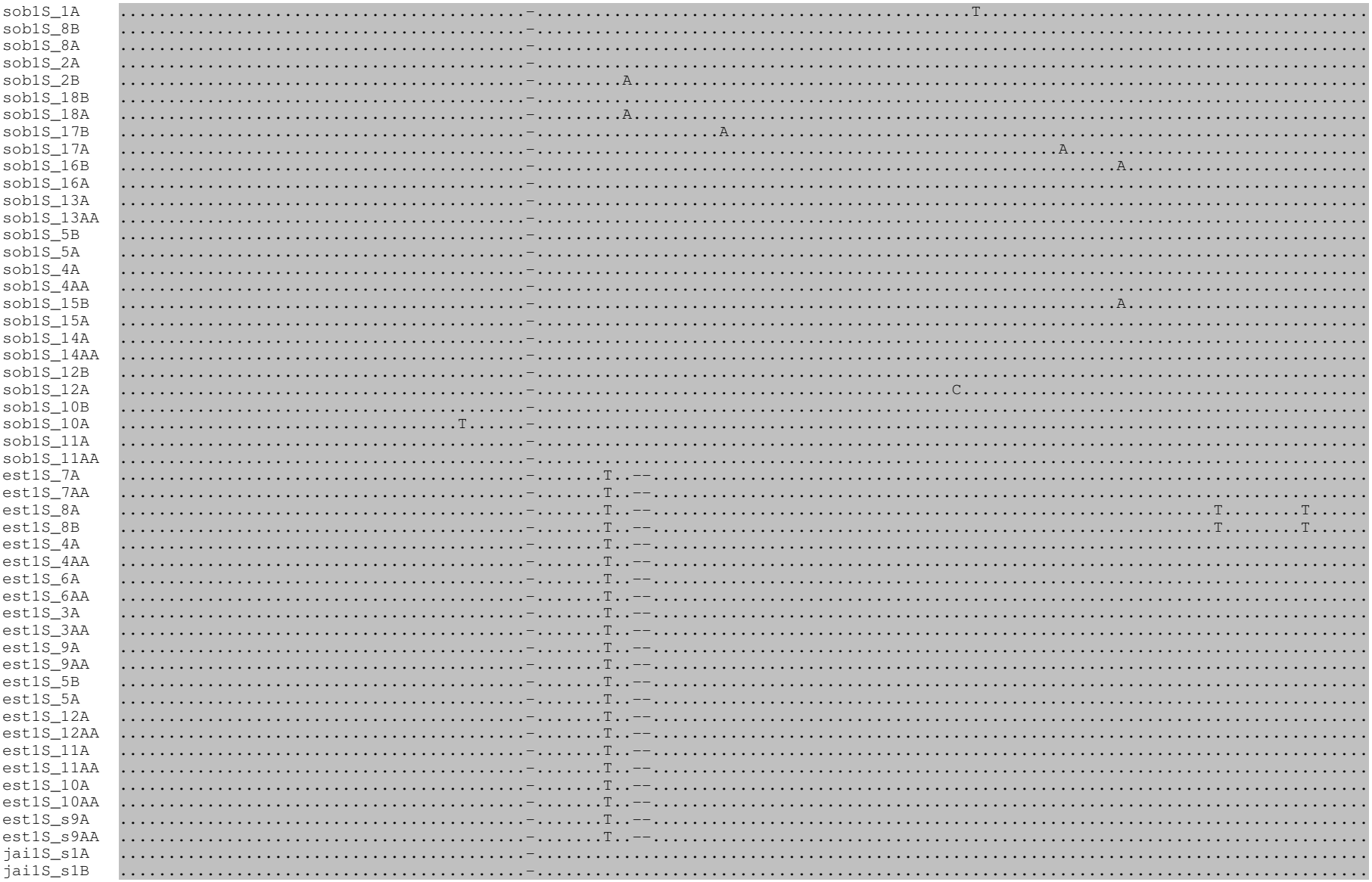

ter_10A
ter_10B
ter_11A
ter_11AA
ter_12A
ter_12AA
pan_4B
pan_4A
pan_8A
pan_8AA
pan_2B
pan_2A
pan_18B
pan_18A
pan_14A
pan_14AA
pan_9B
pan_9A
pan_7A
pan_7AA
pan_6B
pan_6A
pan_17B
pan_17A
pan_16B
pan_16A
pan_15B
pan_15A
pan_12B
pan_12A
pan_11B
pan_11A
pan_21B
pan_21A
pan_19B
pan_19A
pan_13B
pan_13A
nat_7B
nat_7A
nat_6B
nat_6A
nat_5B
nat_5A
nat_4B
nat_4A
nat_3B
nat_3A
nat_2B
nat_2A
nat_11B

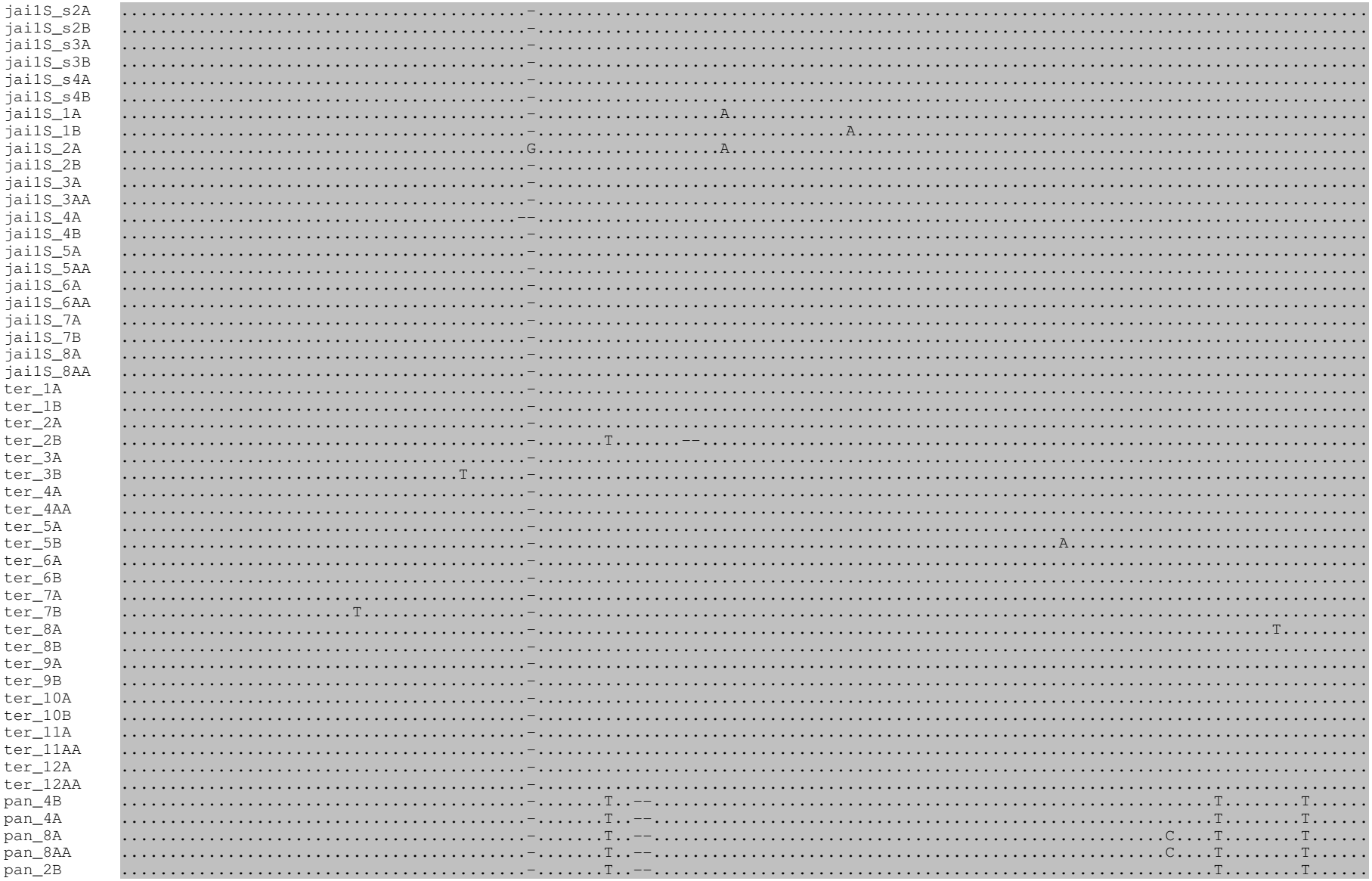

nat_11A
nat_10B
nat_10A
nat_9B
nat_9A
nat_8B
nat_8A
nat_1B
nat_1A
nat_12B
nat_12A
sob2S_6A
sob2S_6AA
sob2S_11A
sob2S_11AA
sob2S_5A
sob2S_5AA
sob2S_4B
sob2S_4A
sob2S_1A
sob2S_1AA
sob2S_18B
sob2S_18A
sob2S_17B
sob2S_17A
sob2S_15A
sob2S_15AA
sob2S_13A
sob2S_13AA
sob2S_9B
sob2S_9A
sob2S_8B
sob2S_8A
sob2S_7B
sob2S_7A
sob2S_16B
sob2S_16A
sob2S_14B
sob2S_14A
jai2S_s9B
jai2S_s8A
jai2S_2A
jai2S_2AA
jai2S_s12A
jai2S_s12AA
jai2S_s3B
jai2S_s4B
jai2S_s5B
jai2S_s6A
jai2S_s6AA
jai2S_s4A

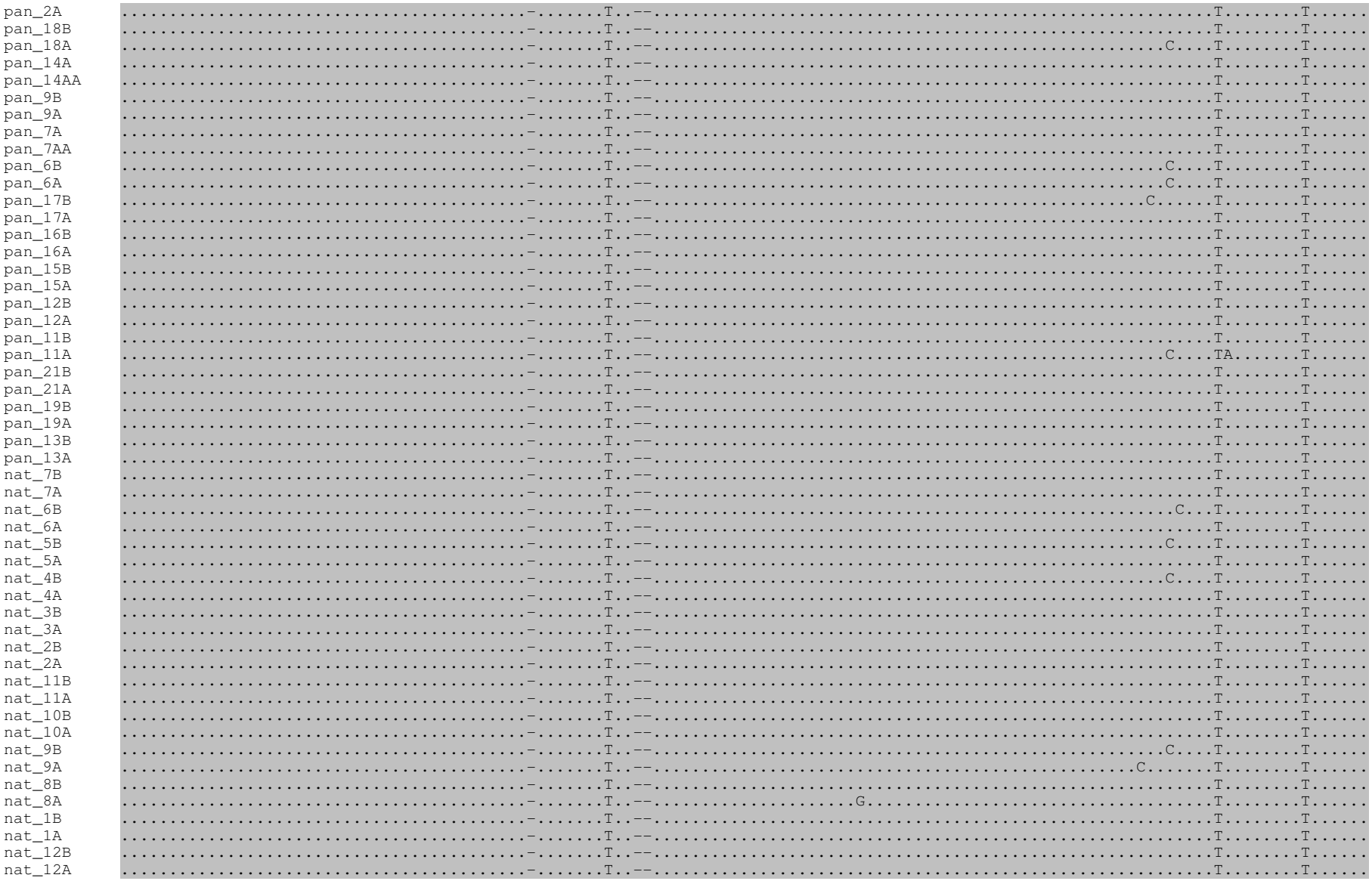

jai2S_s7A
jai2S_s10A
jai2S_s10B
jai2S_s11B
jai2S_s11A
jai2S_1B
jai2S_1A
jai2S_s8B
jai2S_s5A
jai2S_s7B
jai2S_s9A
jai2S_s3A
est2S_26A
est2S_25A
est2S_28A
est2S_27B
est2S_23A
est2S_22A
est2S_32A
est2S_32AA
est2S_27A
est2S_24A
est2S_29A
est2S_23B
est2S_25B
est2S_31A
est2S_29B
est2S_28B
est2S_24B
est2S_22B
est2S_31B
est2S_26B
est2S_s1A
est2S_s1B
est2S_s3A
est2S_s3B
est2S_s4A
est2S_s4B
est2S_s5A
est2S_s5AA
est2S_s6A
est2S_s6B
est2S_s7A
est2S_s7AA
mar_1A
mar_1AA
mar2A
mar_2B
mar_11A
mar_11B
cruzi_1A

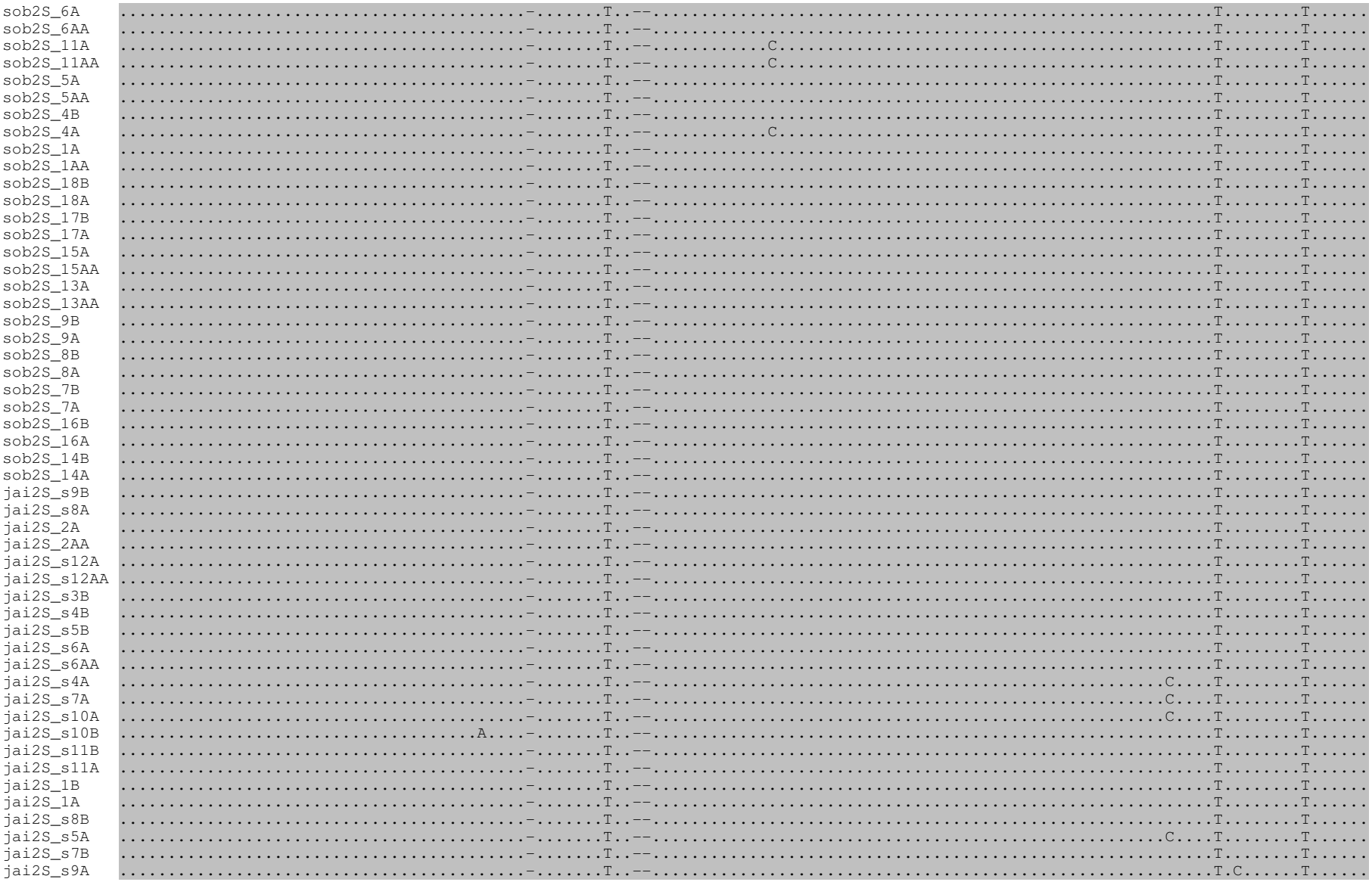

cruzi_1B
cruzi_2A
cruzi_2B
cruzi_3A
cruzi_3AA
cruzi_4A
cruzi_4B
cruzi_5A
cruzi_5AA
cruzi_6A
cruzi_6B
cruzi_7A
cruzi_7AA
cruzi_8A
cruzi_8AA
cruzi_9A
cruzi_9B
cruzi_10A
cruzi_10B
cruzi_11A
cruzi_11B
cruzi_12A
cruzi_12B
pseudo_14A
pseudo_14AA
pseudo_25A
pseudo_25AA
lap_14B
AAATTGAATATGATATTCAACACACCCTCGTTGGACGGGGGG-TGGAATGGTGTATATATTTGTGGATATATGGGGCGGAAATGAGTCCCATTTGCATCGATTGAGTGTTGTGCGACCACTAAATGTTG
lap_14A
lap_15A
lap_15AA

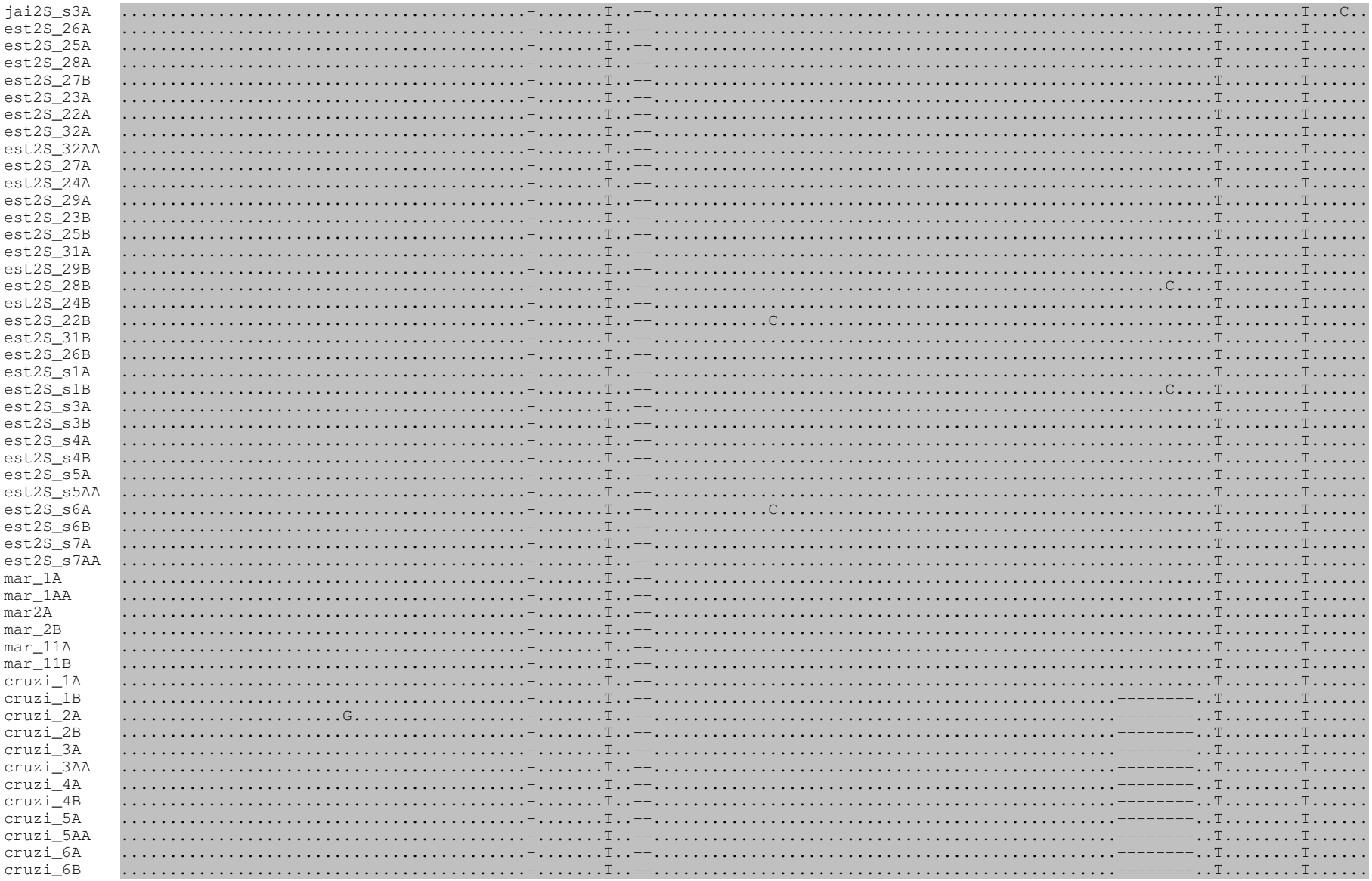

lap_13A
lap_13AA
lap_17A
lap_17AA
lap_18A
lap_18AA
lap_16A
lap_16AA
lap_2B
lap_2A
lap_5B
lap_5A
lap_12B
lap_12A
lap_11B
lap_11A
lap_9B
lap_9A
lap_8A
lap_8AA
lap_6A
lap_6AA
lap_10B
lap_10A
jac_3A
jac_3AA
jac_14A
jac_14AA
jac_10B
jac_10A
jac_1B
jac_1A
jac_9A
jac_9AA
jac_7B
jac_7A
jac_4A
jac_4AA
jac_5B
jac_5A
jac_2B
jac_2A
jac_13B
jac_13A
jac_11A
jac_11B
sob1S_6A
sob1S_6B
sob1S_9A
sob1S_9AA
sob1S_1B

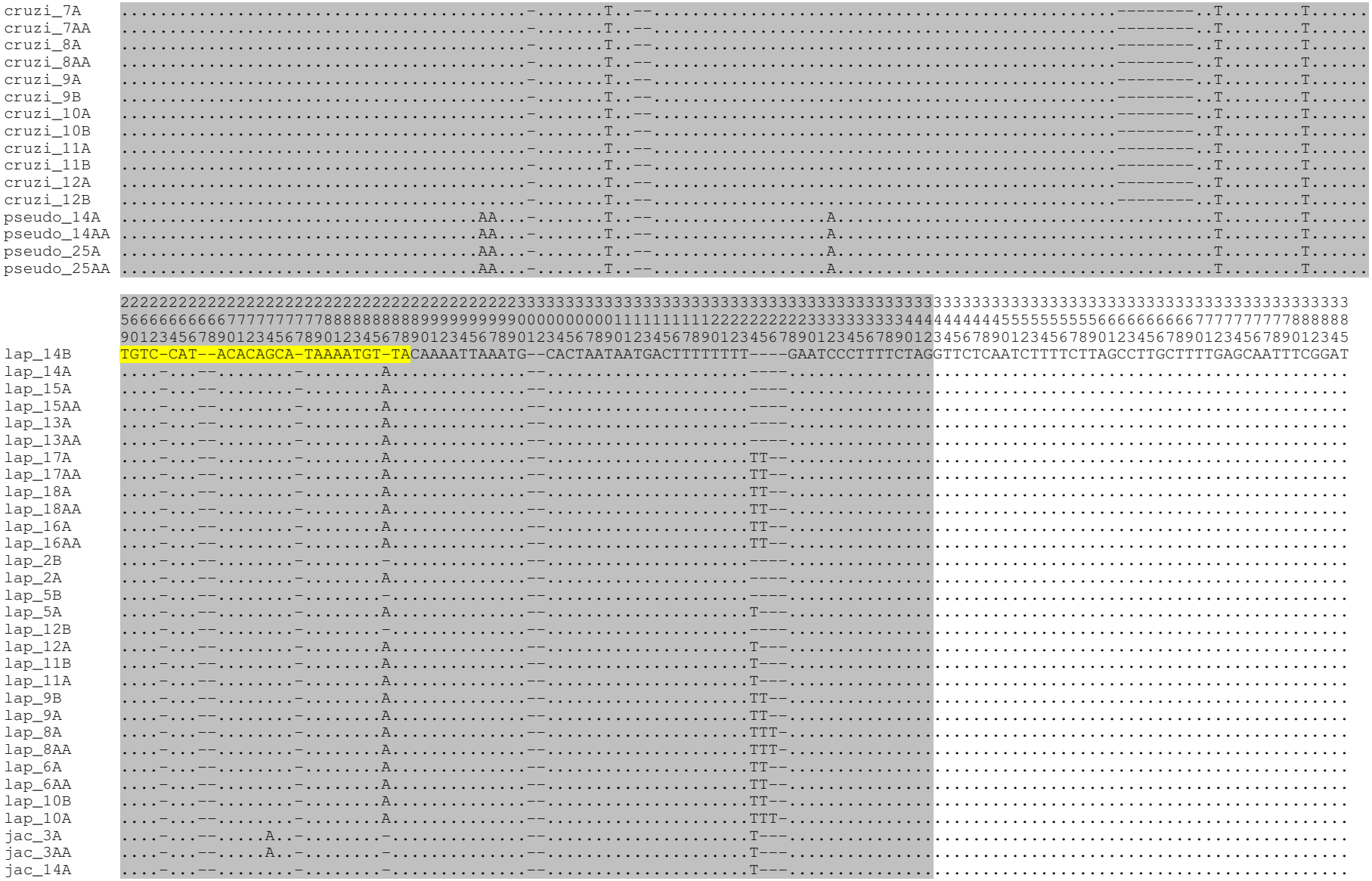

sob1S_1A
sob1S_8B
sob1S_8A
sob1S_2A
sob1S_2B
sob1S_18B
sob1S_18A
sob1S_17B
sob1S_17A
sob1S_16B
sob1S_16A
sob1S_13A
sob1S_13AA
sob1S_5B
sob1S_5A
sob1S_4A
sob1S_4AA
sob1S_15B
sob1S_15A
sob1S_14A
sob1S_14AA
sob1S_12B
sob1S_12A
sob1S_10B
sob1S_10A
sob1S_11A
sob1S_11AA
est1S_7A
est1S_7AA
est1S_8A
est1S_8B
est1S_4A
est1S_4AA
est1S_6A
est1S_6AA
est1S_3A
est1S_3AA
est1S_9A
est1S_9AA
est1S_5B
est1S_5A
est1S_12A
est1S_12AA
est1S_11A
est1S_11AA
est1S_10A
est1S_10AA
est1S_s9A
est1S_s9AA
jai1S_s1A
jai1S_s1B

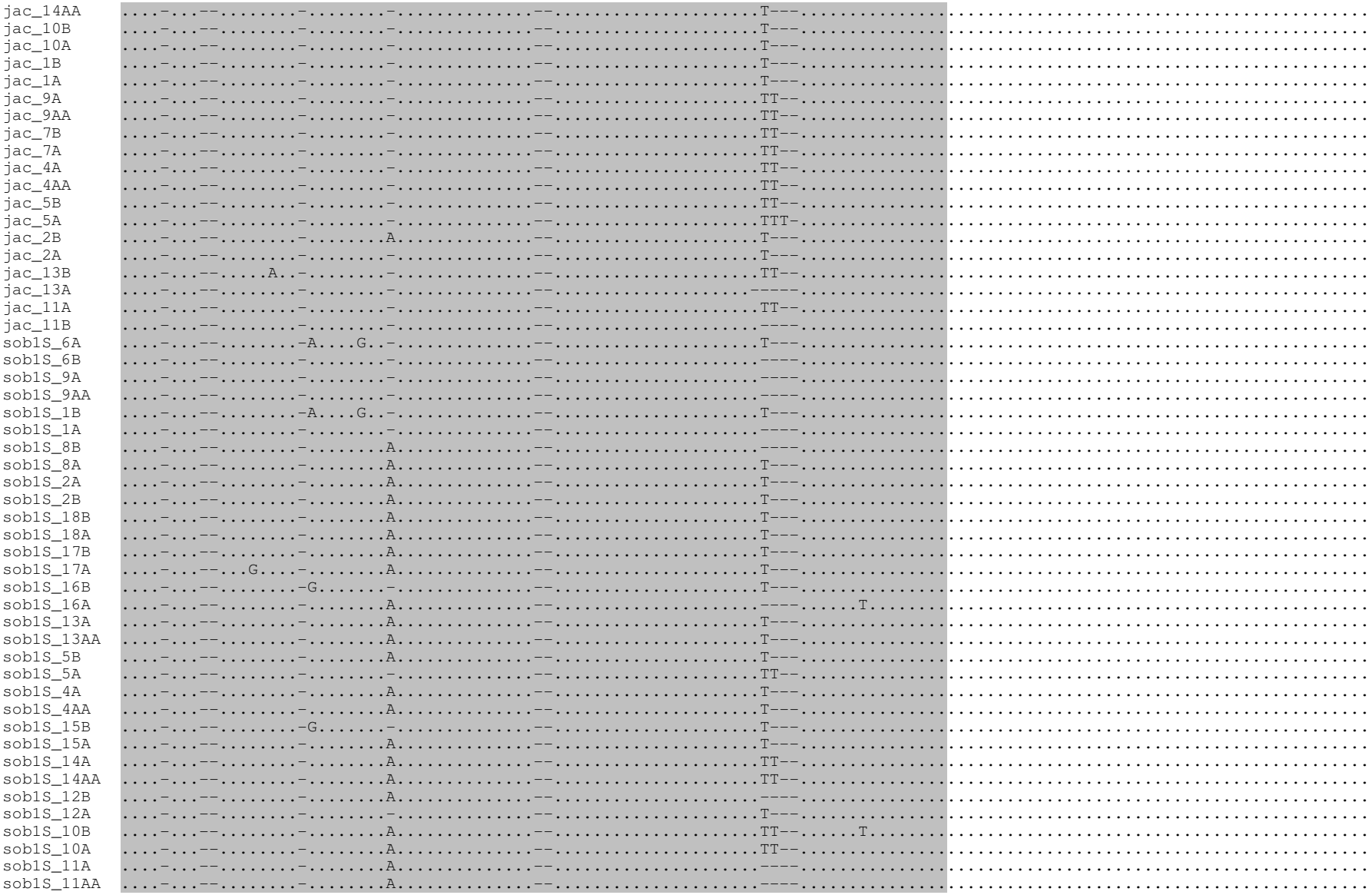

jai1S_s2A
jai1S_s2B
jai1S_s3A
jai1S_s3B
jai1S_s4A
jai1S_s4B
jai1S_1A
jai1S_1B
jai1S_2A
jai1S_2B
jai1S_3A
jai1S_3AA
jai1S_4A
jai1S_4B
jai1S_5A
jai1S_5AA
jai1S_6A
jai1S_6AA
jai1S_7A
jai1S_7B
jai1S_8A
jai1S_8AA
ter_1A
ter_1B
ter_2A
ter_2B
ter_3A
ter_3B
ter_4A
ter_4AA
ter_5A
ter_5B
ter_6A
ter_6B
ter_7A
ter_7B
ter_8A
ter_8B
ter_9A
ter_9B
ter_10A
ter_10B
ter_11A
ter_11AA
ter_12A
ter_12AA
pan_4B
pan_4A
pan_8A
pan_8AA
pan_2B

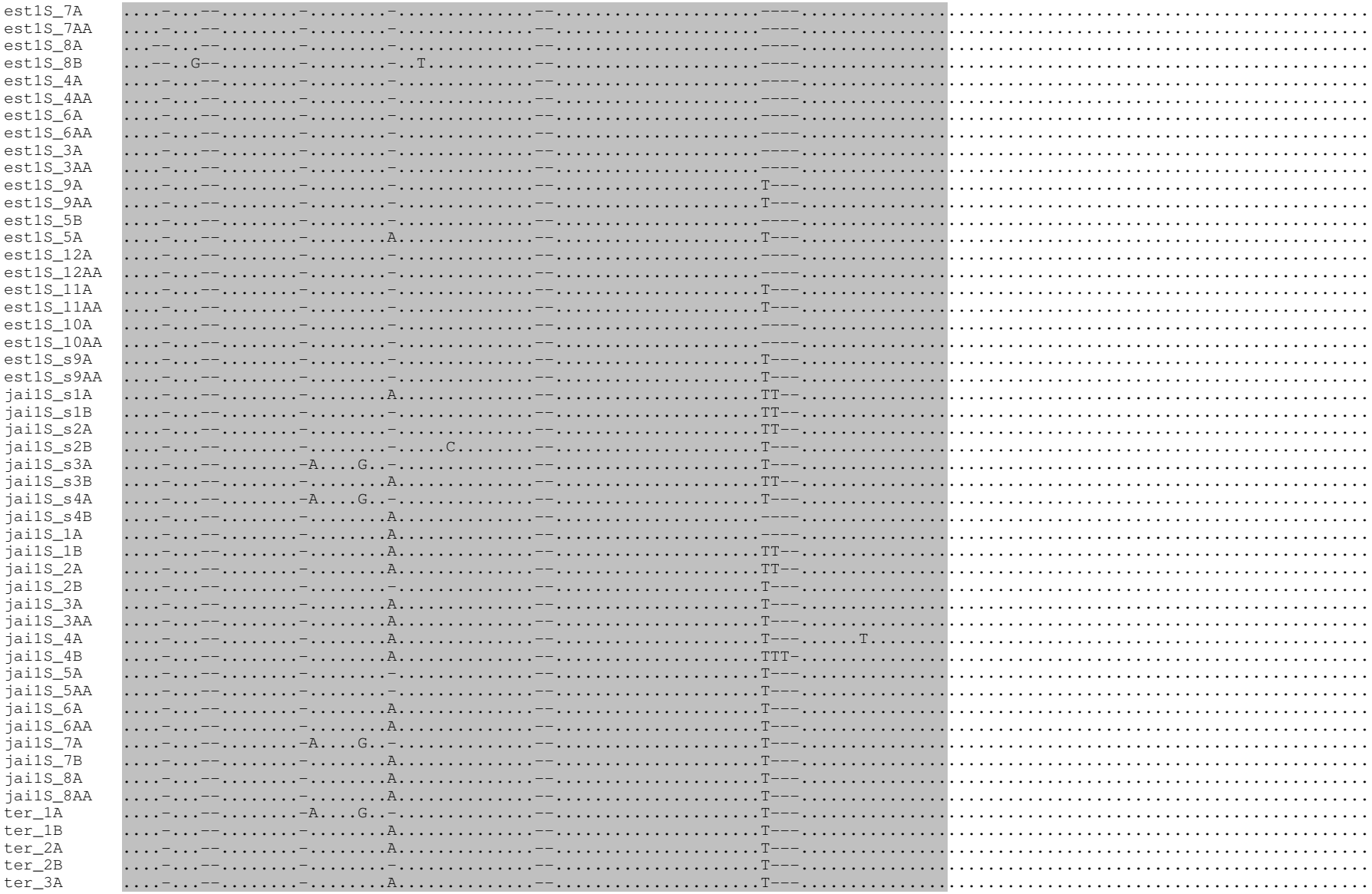

pan_2A
pan_18B
pan_18A
pan_14A
pan_14AA
pan_9B
pan_9A
pan_7A
pan_7AA
pan_6B
pan_6A
pan_17B
pan_17A
pan_16B
pan_16A
pan_15B
pan_15A
pan_12B
pan_12A
pan_11B
pan_11A
pan_21B
pan_21A
pan_19B
pan_19A
pan_13B
pan_13A
nat_7B
nat_7A
nat_6B
nat_6A
nat_5B
nat_5A
nat_4B
nat_4A
nat_3B
nat_3A
nat_2B
nat_2A
nat_11B
nat_11A
nat_10B
nat_10A
nat_9B
nat_9A
nat_8B
nat_8A
nat_1B
nat_1A
nat_12B
nat_12A

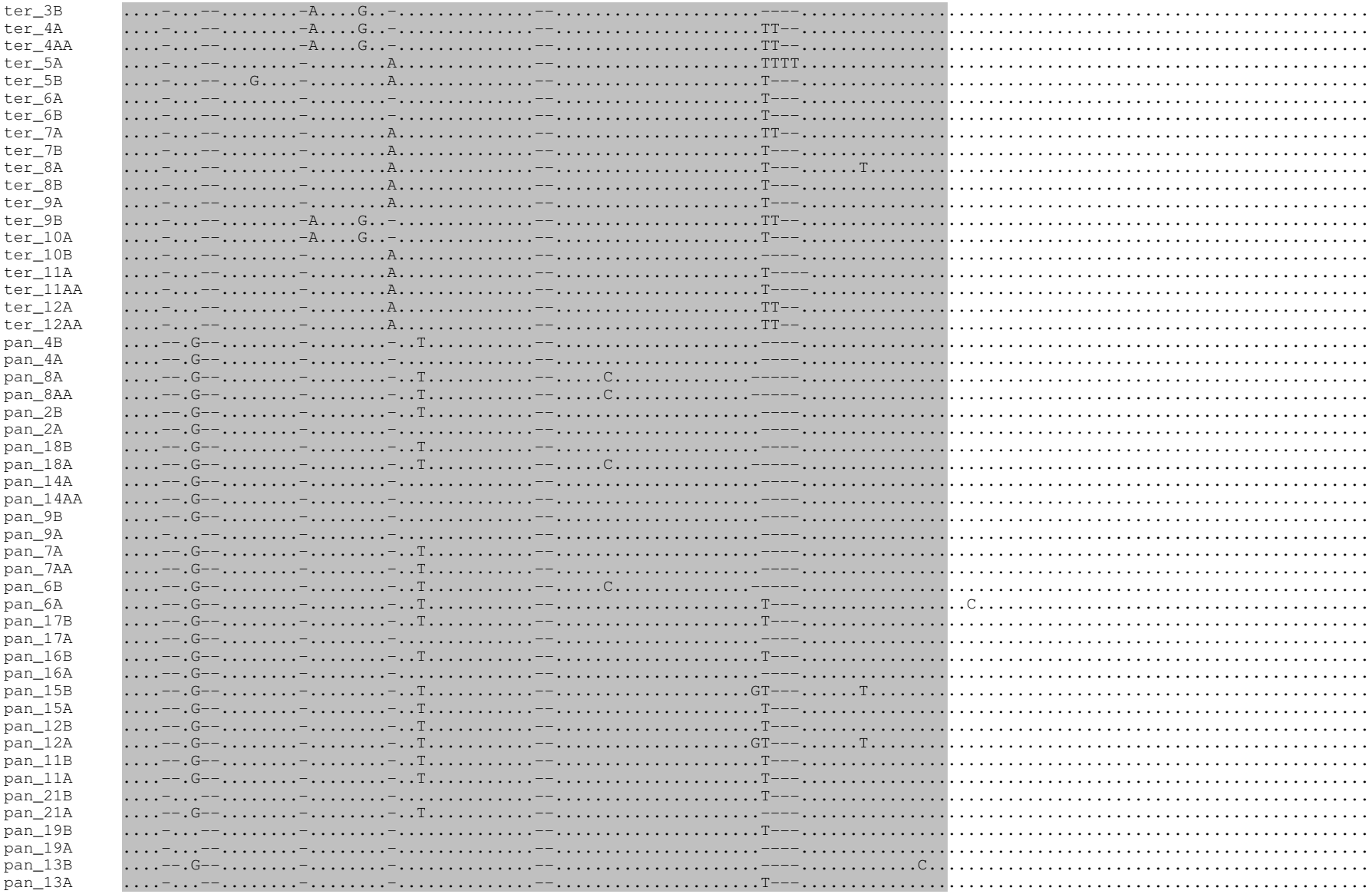

sob2S_6A
sob2S_6AA
sob2S_11A
sob2S_11AA
sob2S_5A
sob2S_5AA
sob2S_4B
sob2S_4A
sob2S_1A
sob2S_1AA
sob2S_18B
sob2S_18A
sob2S_17B
sob2S_17A
sob2S_15A
sob2S_15AA
sob2S_13A
sob2S_13AA
sob2S_9B
sob2S_9A
sob2S_8B
sob2S_8A
sob2S_7B
sob2S_7A
sob2S_16B
sob2S_16A
sob2S_14B
sob2S_14A
jai2S_s9B
jai2S_s8A
jai2S_2A
jai2S_2AA
jai2S_s12A
jai2S_s12AA
jai2S_s3B
jai2S_s4B
jai2S_s5B
jai2S_s6A
jai2S_s6AA
jai2S_s4A
jai2S_s7A
jai2S_s10A
jai2S_s10B
jai2S_s11B
jai2S_s11A
jai2S_1B
jai2S_1A
jai2S_s8B
jai2S_s5A
jai2S_s7B
jai2S_s9A

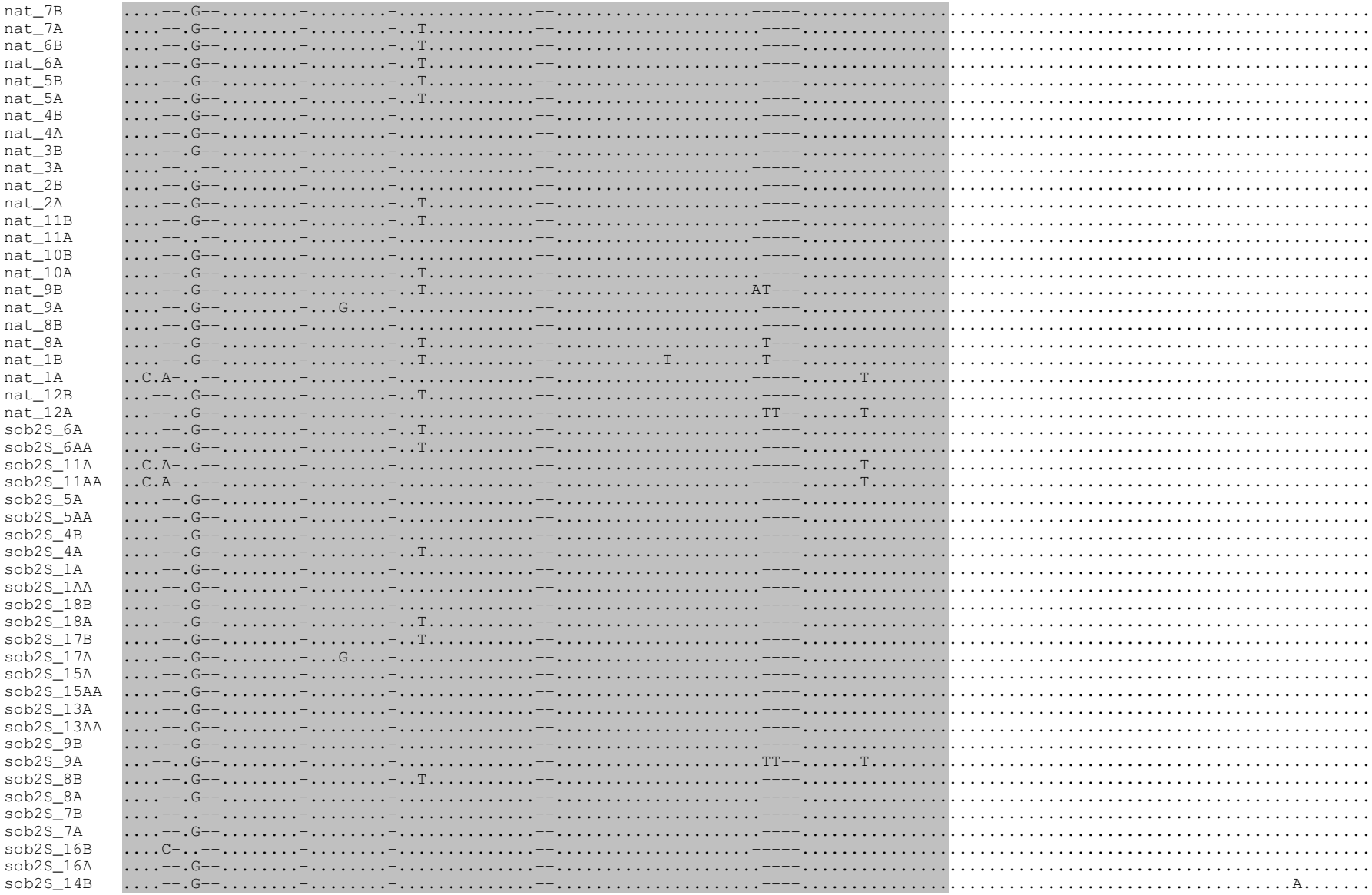

jai2S_s3A
est2S_26A
est2S_25A
est2S_28A
est2S_27B
est2S_23A
est2S_22A
est2S_32A
est2S_32AA
est2S_27A
est2S_24A
est2S_29A
est2S_23B
est2S_25B
est2S_31A
est2S_29B
est2S_28B
est2S_24B
est2S_22B
est2S_31B
est2S_26B
est2S_s1A
est2S_s1B
est2S_s3A
est2S_s3B
est2S_s4A
est2S_s4B
est2S_s5A
est2S_s5AA
est2S_s6A
est2S_s6B
est2S_s7A
est2S_s7AA
mar_1A
mar_1AA
mar2A
mar_2B
mar_11A
mar_11B
cruzi_1A
cruzi_1B
cruzi_2A
cruzi_2B
cruzi_3A
cruzi_3AA
cruzi_4A
cruzi_4B
cruzi_5A
cruzi_5AA
cruzi_6A
cruzi_6B

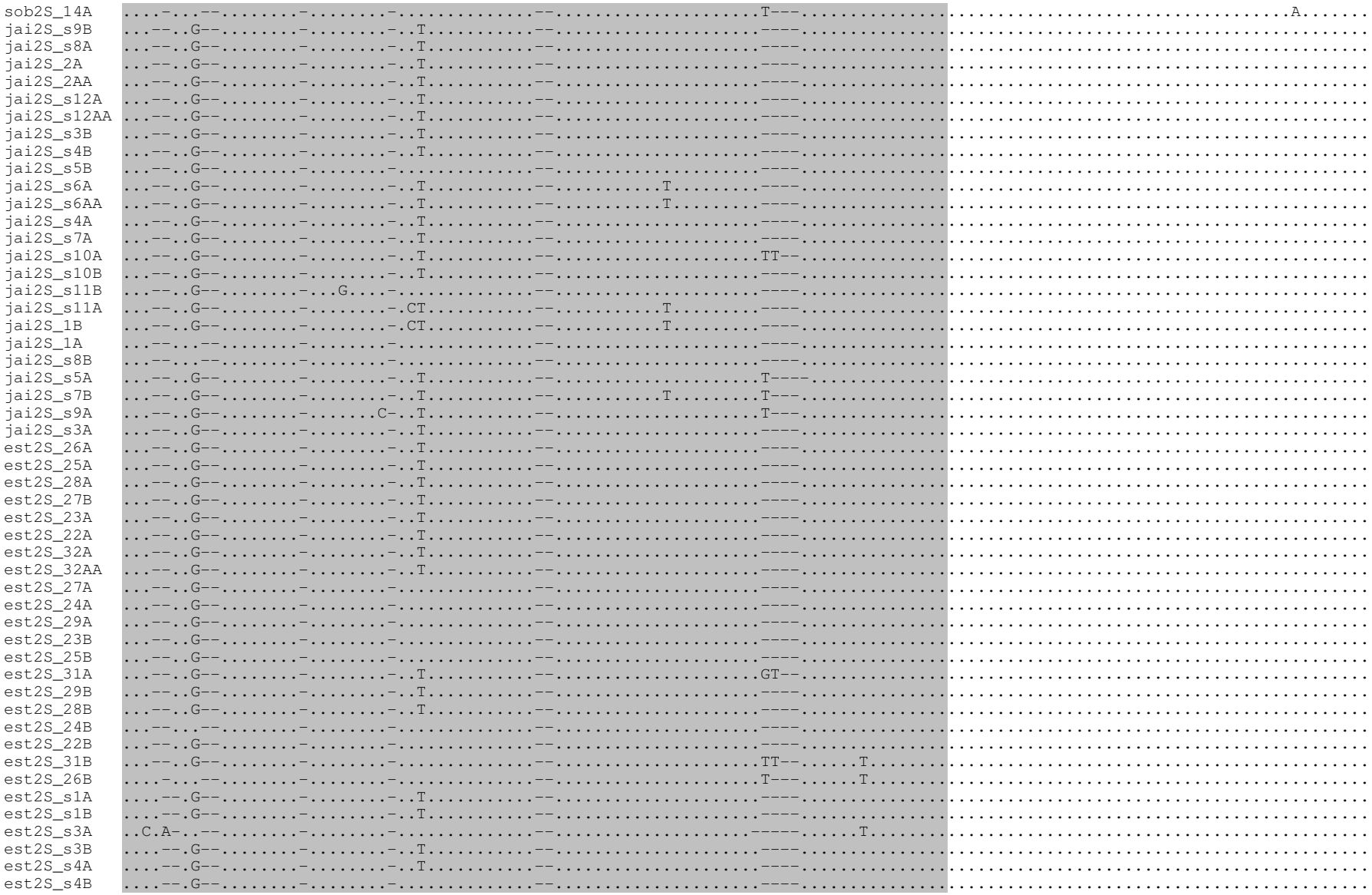


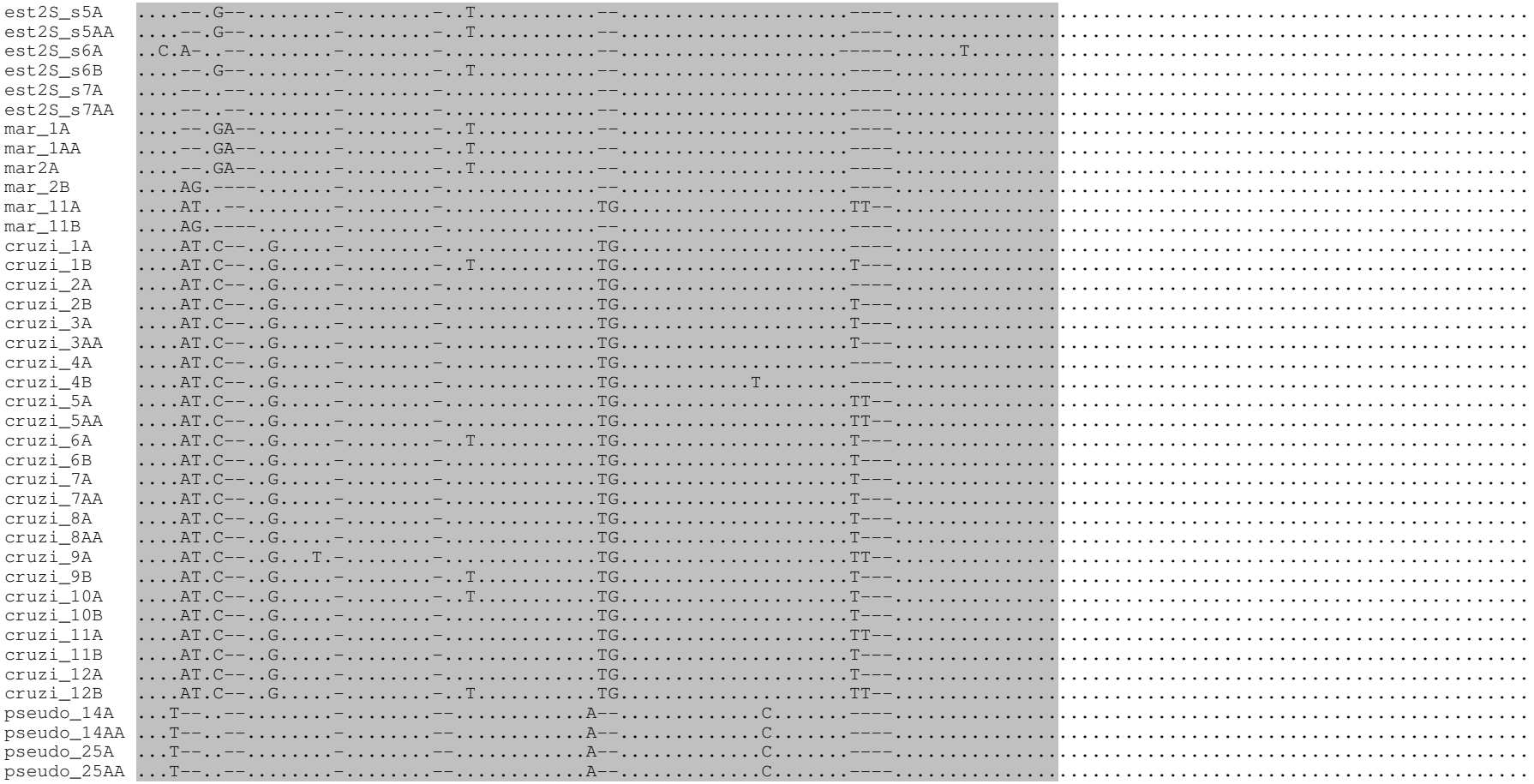

jac_14AA
jac_10B
jac_10A
jac_1B
jac_1A
jac_9A
jac_9AA
jac_7B
jac_7A
jac_4A
jac_4AA
jac_5B
jac_5A
jac_2B
jac_2A
jac_13B
jac_13A
jac_11A
jac_11B
sob1S_6A
sob1S_6B
sob1S_9A
sob1S_9AA
sob1S_1B
sob1S_1A
sob1S_8B
sob1S_8A
sob1S_2A
sob1S_2B
sob1S_18B
sob1S_18A
sob1S_17B
sob1S_17A
sob1S_16B
sob1S_16A
sob1S_13A
sob1S_13AA
sob1S_5B
sob1S_5A
sob1S_4A
sob1S_4AA
sob1S_15B
sob1S_15A
sob1S_14A
sob1S_14AA
sob1S_12B
sob1S_12A
sob1S_10B
sob1S_10A
sob1S_11A
sob1S_11AA


est1S_7A
est1S_7AA
est1S_8A
est1S_8B
est1S_4A
est1S_4AA
est1S_6A
est1S_6AA
est1S_3A
est1S_3AA
est1S_9A
est1S_9AA
est1S_5B
est1S_5A
est1S_12A
est1S_12AA
est1S_11A
est1S_11AA
est1S_10A
est1S_10AA
est1S_s9A
est1S_s9AA
jai1S_s1A
jai1S_s1B
jai1S_s2A
jai1S_s2B
jai1S_s3A
jai1S_s3B
jai1S_s4A
jai1S_s4B
jai1S_1A
jai1S_1B
jai1S_2A
jai1S_2B
jai1S_3A
jai1S_3AA
jai1S_4A
jai1S_4B
jai1S_5A
jai1S_5AA
jai1S_6A
jai1S_6AA
jai1S_7A
jai1S_7B
jai1S_8A
jai1S_8AA
ter_1A
ter_1B
ter_2A
ter_2B
ter_3A


ter_3B
ter_4A
ter_4AA
ter_5A
ter_5B
ter_6A
ter_6B
ter_7A
ter_7B
ter_8A
ter_8B
ter_9A
ter_9B
ter_10A
ter_10B
ter_11A
ter_11AA
ter_12A
ter_12AA
pan_4B
pan_4A
pan_8A
pan_8AA
pan_2B
pan_2A
pan_18B
pan_18A
pan_14A
pan_14AA
pan_9B
pan_9A
pan_7A
pan_7AA
pan_6B
pan_6A
pan_17B
pan_17A
pan_16B
pan_16A
pan_15B
pan_15A
pan_12B
pan_12A
pan_11B
pan_11A
pan_21B
pan_21A
pan_19B
pan_19A
pan_13B
pan_13A


nat_7B
nat_7A
nat_6B
nat_6A
nat_5B
nat_5A
nat_4B
nat_4A
nat_3B
nat_3A
nat_2B
nat_2A
nat_11B
nat_11A
nat_10B
nat_10A
nat_9B
nat_9A
nat_8B
nat_8A
nat_1B
nat_1A
nat_12B
nat_12A
sob2S_6A
sob2S_6AA
sob2S_11A
sob2S_11AA
sob2S_5A
sob2S_5AA
sob2S_4B
sob2S_4A
sob2S_1A
sob2S_1AA
sob2S_18B
sob2S_18A
sob2S_17B
sob2S_17A
sob2S_15A
sob2S_15AA
sob2S_13A
sob2S_13AA
sob2S_9B
sob2S_9A
sob2S_8B
sob2S_8A
sob2S_7B
sob2S_7A
sob2S_16B
sob2S_16A
sob2S_14B


sob2S_14A
jai2S_s9B
jai2S_s8A
jai2S_2A
jai2S_2AA
jai2S_s12A
jai2S_s12AA
jai2S_s3B
jai2S_s4B
jai2S_s5B
jai2S_s6A
jai2S_s6AA
jai2S_s4A
jai2S_s7A
jai2S_s10A
jai2S_s10B
jai2S_s11B
jai2S_s11A
jai2S_1B
jai2S_1A
jai2S_s8B
jai2S_s5A
jai2S_s7B
jai2S_s9A
jai2S_s3A
est2S_26A
est2S_25A
est2S_28A
est2S_27B
est2S_23A
est2S_22A
est2S_32A
est2S_32AA
est2S_27A
est2S_24A
est2S_29A
est2S_23B
est2S_25B
est2S_31A
est2S_29B
est2S_28B
est2S_24B
est2S_22B
est2S_31B
est2S_26B
est2S_s1A
est2S_s1B
est2S_s3A
est2S_s3B
est2S_s4A
est2S_s4B


est2S_s5A
est2S_s5AA
est2S_s6A
est2S_s6B
est2S_s7A
est2S_s7AA
mar_1A
mar_1AA
mar2A
mar_2B
mar_11A
mar_11B
cruzi_1A
cruzi_1B
cruzi_2A
cruzi_2B
cruzi_3A
cruzi_3AA
cruzi_4A
cruzi_4B
cruzi_5A
cruzi_5AA
cruzi_6A
cruzi_6B
cruzi_7A
cruzi_7AA
cruzi_8A
cruzi_8AA
cruzi_9A
cruzi_9B
cruzi_10A
cruzi_10B
cruzi_11A
cruzi_11B
cruzi_12A
cruzi_12B
pseudo_14A
pseudo_14AA
pseudo_25A
pseudo_25AA

CAPÍTULO 5: "Análise macrogeográfica da divergência no gene *paralytic* entre populações do complexo de espécies *Lutzomyia longipalpis*"

Lins RMMA, Brazil RP, Souza NA, Ribolla PEM, Salomón OD, Lanzaro GC, Peixoto AA; **Macrogeographic analysis of the divergence and gene flow in the *paralytic* gene among *Lutzomyia longipalpis* populations.**

Este manuscrito está em elaboração e corresponde aos objetivos específicos 2.2.1 e 2.2.2.

O presente artigo expande as análises realizadas com o gene *paralytic* à outras populações de *Lu. longipalpis* da América Central e do Sul com o objetivo de avaliar o grau de divergência genética entre elas. Uma amostra de *Lu. cruzi*, uma espécie do complexo Longipalpis foi incluída nesse trabalho. Os dados obtidos reforçam a classificação proposta anteriormente para algumas espécies distribuídas na América Central e do Sul. Além disso, os dados confirmam a existência de um complexo de espécies no Brasil e reforçam a existência de dois grupos distintos de populações brasileiras de *Lu. longipalpis* cujos machos produzem sons do tipo *Burst* e do tipo Pulsado, com alta diferenciação genética e diferenças fixas entre elas. Os dados gerados corroboram também a hipótese de que *Lu. cruzi* representa mais uma espécie do complexo Longipalpis.

Macrogeographic analysis of the divergence in the *paralytic* gene among populations of the *Lutzomyia longipalpis* species complex.

Rachel MMA Lins[1], Reginaldo P Brazil[2], Nataly A Souza[3], Paulo EM Ribolla[4], Oscar D Salomón[5], Gregory C Lanzaro[6], Alexandre A Peixoto[1§]

[1] Laboratório de Biologia Molecular de Insetos, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, RJ, Brazil

[2] Laboratório de Bioquímica e Fisiologia de Insetos, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, RJ, Brazil

[3] Laboratório de Transmissores de Leishmanioses, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, RJ, Brazil

[4] Departamento de Parasitologia, Instituto de Biociências, UNESP, Botucatu, SP, Brasil

[5] Centro Nacional de Diagnóstico e Investigación en Endemo-epidemias, ANLIS, Ministerio de Salud, Buenos Aires, Argentina.

[6] Vector Genetics Laboratory, Department of Pathology, Microbiology and Immunology, School of Veterinary Medicine, University of California – Davis, Davis, CA, USA

§Corresponding author

Email addresses:

RMMAL: rlins@ioc.fiocruz.br

RPB: rpbrazil@ioc.fiocruz.br

NAS: souzana@ioc.fiocruz.br

PEMR: pribolla@ibb.unesp.br

ODS: odanielsalomon@gmail.com

GCL: gclanzaro@ucdavis.edu

AAP: apeixoto@ioc.fiocruz.br

Abstract

American Visceral Leishmaniasis (AVL) is caused by *Leishmania infantum* and is mainly transmitted by *Lu. longipalpis,* which represents a complex sibling species. Previous studies suggested the existence of four main clades in Latin America representing different species with distinct geographical ranges: (1) Brazil (except Roraima State) – Species A = *Lu. longipalpis sensu stricto*, (2) Laran – Species B = *Lu. pseudolongipalpis* (Venezuela), (3) cis-Andean – Species C (Colombia and northwestern Venezuela, and Roraima State in Brazil) and (4) trans-Andean – Species D (Central America). In addition, cross-mating experiments, analysis of copulation song, pheromones, microsatellites and *period*, *cacophony* and *paralytic* genes, suggested that the Brazilian clade did not represent in fact a single species. Here, we performed a macrogeographic analysis of the divergence and gene flow in the *paralytic* gene among *Lu. longipalpis* populations from South and Central America. Our results corroborated the previous classification of Species B, Species C and Species D and also confirmed that this vector represents a species complex in Brazil with two main groups.

Introduction

Visceral leishmaniasis is a very important neglected tropical disease prevalent in many regions of the World (WHO, 2010) and *Lutzomyia longipalpis* (Diptera: Psychodidae) is its main vector in Latin America (Lainson and Rangel 2005). The geographical distribution of this sand fly extends from central Mexico to northern Argentina and Paraguay (Young and Duncan, 1994) occurring in a variety of habitats and in close association with man. The existence of geographic and climatic barriers together with the vector low mobility resulted in a considerably degree of isolation among its numerous populations. The question whether *Lu. longipalpis* would in fact comprise a species complex, has led to a number of studies aiming at determining the real taxonomic status of this vector (reviewed in Bauzer et al 2007).

Lu. longipalpis was first suggested as not being a single species by Mangabeira (1969), when he identified differences in the number of abdominal spots among males from different Brazilian populations. Later, cross-mating experiments conducted with sympatric and allopatric populations from Brazil demonstrated a reduced degree of insemination in some crosses, suggesting the occurrence of different sibling species (Ward et al 1983). These results were later confirmed for other populations and validated the hypothesis that *Lu. Longipalpis* could represent a species complex in Brazil (Ward et al 1988).

Further investigations carried out with populations from Central and South America demonstrated a clear divergence between *Lu. longipalpis* from different areas (reviewed in Bauzer et al 2007). For example, colonies from Costa Rica, Colombia and Brazil had their genetic variability assessed from 27 enzyme-coding loci. The results demonstrated a high genetic distance among these populations and cross-mating experiments resulted in sterile male progeny, suggesting that these colonies represent

three species (Lanzaro et al 1993). These populations also show divergence in salivary maxadilan gene expression and metaphase karyotypes (Yin et al 1999, 2000).

Analyses conducted with mitochondrial genes, including the cytochrome c oxidase I (COI) gene, isozyme loci and populations from South and Central America supported the existence of four species with distinct geographical ranges: (1) Brazil (except Roraima State) – Species A = *Lu. longipalpis sensu stricto*, (2) Laran – Species B = *Lu. pseudolongipalpis* (Venezuela), (3) cis-Andean – Species C (which may be further subdivided into two groups – Colombia and northwestern Venezuela, and included Roraima State in Brazil) and (4) trans-Andean – Species D (Central America) (Arrivillaga et al 2002, 2003).

Although the mitochondrial and isozyme data seem to suggest the occurrence of a single species in most of Brazil, many other studies support the existence of a complex among the Brazilian populations (reviewed in Bauzer et al 2007; Maingon et al 2008). Crossing experiments (Ward et al 1988; Souza et al 2008), analyses of acoustic signals (Souza et al 2002, 2004; Araki et al 2009), male sex pheromones (Hamilton et al 1999a,b, 2004, 2005; Araki et al 2009), analysis of orthologues of *Drosophila* courtship song genes *period* (*per*), *cacophony* (*cac*), *paralytic* (*para*) (Bauzer et al 2002a,b; Bottecchia et al 2004; Araki et al 2009; Lins et al 2008, 2012) and microsatellites (Maingon et al 2004, Watts et al 2005), have indicated that the Brazilian populations of this vector can be divided in two main groups according to the type of copulation song and pheromone males produce (Araki et al 2009; Lins et al 2012).

One such group, probably representing a single species, is distributed from Espírito Santo State to Pará State and its males produce Burst-type song and the Cembrene-1 pheromone. The second group, more continental in its distribution, exhibits more complex song repertoires with the males producing five variations of the Pulse-type song (P1 to P5) in combination with four different types of pheromones

(Germacrene, Himachalene, and Cembrenes 1 and 2). These populations probably represent different incipient species (Araki et al 2009).

Lutzomyia cruzi is another species of the *Lu. longipalpis* complex from Brazil (Watts et al 2005; Brazil & Hamilton 2002; Vigoder et al 2010). This species was also implicated as a vector of *Le. infantum* in a region of the country (de Pita-Pereira et al 2008). Its males produce the Germacrene pheromone and Burst-type song similar, albeit with a shorter inter-burst interval, to that produced by many *Lu. longipalpis* populations (Vigoder et al 2010).

Currently there is a consensus that *Lu. longipalpis* represent a species complex (Ward et al 1998; Lanzaro et al 1993; Arrivillaga et al 2003; Bauzer et al 2007; Maingon et al 2008, Araki et al 2009; Lins et al 2012). However, the number and the geographic range of the different siblings is still uncertain. Although high levels of genetic divergence between sympatric and allopatric populations have been detected, reproductive isolation is not complete and introgression has played a role in shaping this vector genome. In a recent study, Araki et al (in preparation) analyzed 21 nuclear loci using a multilocus approach with Brazilian sympatric and allopatric populations of *Lu. longipalpis* and detected the occurrence of asymmetric gene flow with variable levels of introgression among the different loci.

Although much progress in the understanding of *Lu. longipalpis* species complex in Brazil was made, no diagnostic differences were observed between these two groups in most molecular markers studied. The only exception so far is the *paralytic* (*para*) gene, a locus involved in the control of courtship songs in *Drosophila* (Peixoto et al 1998), which showed fixed differences between Brazilian allopatric and sympatric populations producing different copulation songs (Lins et al 2008, 2012). In the present study we have extended the analysis of the *para* gene to many other populations of the *Lu. longipalpis* complex from Central and South America.

Methods

We analyzed forty-one sandfly samples from South and Central America including those previous published in Lins et al (2008, 2012): from Brazil - Sobral 1S and 2S, Sobral, Ceará State; Natal, Rio Grande do Norte State; Passira 2S, and Itamaracá 2S, Pernambuco State; Estrela de Alagoas 1S and 2S, Alagoas State; Jacobina 1S and Jequié 2S, Bahia State; Nova Porteirinha 2S, Ipanema 1S, Jaíba 1S and 2S, Lassance 1S, Lapinha 1S, Minas Gerais State; Afonso Cláudio 1S and Pancas 1S, Espírito Santo State; Barra de Guaratiba 1S and Rio Bonito 1S, Rio de Janeiro State; Araçatuba, São Paulo State; Campo Grande 2S, Mato Grosso do Sul State; Barcarena 1S, Camará 2S, Marajó 1S and Cametá 1S, Pará State; Palmas 1S, 2S and intermediate (int), Tocantins State; Teresina 1S, Piauí State; Pirenópolis 1S, Goiania State; Roraima, Roraima State; Palo Gordo and Durania, Colombia; Trujillo, Venezuela; Brasilito, Costa Rica; San Juan Bautista and Isla del Tigre, Honduras; Missiones, Argentina; Paraguay 2S; *Lu. cruzi* from Mato Groso do Sul State, Brazil; *Lu. pseudolongipalpis* from Curarigua, Venezuela. The 1S and 2S refer to the male spot phepotype, one or two abdominal spots, in the fourth or third and foruth tergite, respectively (Ward et al 1988, Araki et al 2009, Lins et al 2012). In the intermediate phenotype (Int) the spot in the third tergite has about half or less the size of that on the fourth tergite. Most males from the Brazilian samples had their spot pattern identified. Although this phenotype cannot be used to identify different species, in some cases, as reported for Sobral (reviewed in Maingon et al 2008; Bauzer et al 2007), Jaíba and Estrela de Alagoas (Araki et al 2009; Lins et al 2012), the sympatric 1S and 2S males produce different copulation songs (Pulse-type and Burst-type, respectively) and represent different species.

Genomic DNA was isolated according to Jowett et al 1998 and the PCR Mater Mix or Go Taq Green Master Mix (Promega) was used to perform the PCR according to Lins et al (2008). PCR products were purified using the Wizard SV Gel and PCR Clean-

up System (Promega), GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) or QIAquick PCR purification Kit (QIAGEN). Purified fragments were cloned using the pMOS*Blue* Blunt Ended Cloning Kit (GE Healthcare) or TOPO TA Cloning Kit (Invitrogen). Plasmid DNA was isolated using Flexiprep Kit (GE Healthcare) or using 96 wells microplates and the alkaline lysis method (Sambrook & Russell, 2001) followed by filtration in Millipore Multiscreen filter plates. DNA sequencing was carried out in an ABI Prism 3730 sequencer using the Big Bye 3.0 or 3.1 Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and was done by the PDTIS sequencing facility at Fundação Oswaldo Cruz.

Before the population genetic analyses, BioEdit Sequence Alignment Editor (Hall, 1999) was used to process *Lu. longipalpis para* gene fragments from all populations, which also included the previously published sequences from Sobral, Natal, Pancas, Jaíba, Estrela de Alagoas, Lapinha, Jacobina, Teresina, Marajó and *Lu. cruzi* (Lins et al, 2008, 2012). A minimum of six sequences per individual were aligned to obtain two consensus sequences corresponding to the two alleles, named as A and B, or one consensus sequence in which case the flies were treated as homozygotes and the sequences were duplicated. The estimated probability of misclassifying a heterozygous fly as a homozygous with this procedure is less than 5%. Sequences were submitted to the GenBank (accession numbers XXXX-XXXX).

The polymorphism and population structure analyses were carried out using DnaSP v5 (Librado & Rozas, 2009), Proseq 2.91 (Filatov, 2002) and Arlequin 3.11 (Excoffier et al 2005) programs. The Neighbor Joining tree based on Fst values was estimated using MEGA5 (Tamura et al 2011). The analysis of molecular variance (AMOVA) was carried out with Arlequin 3.11 (Excoffier et al 2005). A non-recombinant block of the initial fragment was obtained using the IMgc program

(Woerner et al 2007) to construct the haplotype network with TCS v1.21 software (Clement et al 2000).

Results

In this study, we have analyzed 832 consensus sequences in total, corresponding to both alleles of each individual. This *para* fragment of 394bp includes an intron of variable size of approximately 230bp. Sequences corresponding to populations previously published (Lins et al 2008, 2012) were included in the analyses. Figure 1 presents the approximate geographic position of *Lu. longipalpis*, *Lu. cruzi* and *Lu. pseudolongipalpis* populations analyzed. For most Brazilian populations it also contains information on the type of copulation song produced by the males, i.e., Burst, Pulse or Mixed type song (Araki et al 2009; Vigoder et al, in preparation). Additional file S1 shows an alignment of the polymorphic sites for the whole fragment haplotypes (WFHap), which differ from the haplotypes based on the smaller non-recombinant fragment (see below). Sites in the intron region are highlighted in grey. Most part of the variation is found within the intron portion with a number of indels events. A few non-synonymous substitutions were also detected in both exon fragments. The distribution and number of haplotypes represented in each sample are shown in additional file S2. The most represented haplotypes are WFHap_3 and WFHap_5, found in the Pulse type populations and Burst type populations respectively. WFHap_5 is also represented in the populations of Araçatuba (no recorded song), Argentina, Paraguay and *Lu. cruzi*, indicating that they probably represent siblings of the same species. Interestingly, all sequences of Rio Bonito correspond to WFHap_1, which is also represented in Barra de Guaratiba, Roraima, Trujillo (Venezuela), Palo Gordo and Durania (Colombia) and the

samples from Central America (Isla del Tigre, San Juan Bautista – Honduras; Brasilito – Costa Rica).-

Table 1 presents a summary of the polymorphisms found in the *Lu. longipalpis*, *Lu. cruzi* and *Lu. pseudolongipalpis para* fragment. Overall, low levels of genetic polymorphisms were encountered in the *Lu. longipalpis* Brazilian populations *para* gene fragment; Palmas 2S ($\pi = 0.00694$, $\theta = 0.00941$) population was the most polymorphic whereas Campo Grande was the least one ($\pi = 0.00059$, $\theta = 0.00155$). Among other *Lu. longipalpis* populations from Central and South America, Palo Gordo (South America, Colombia) showed the lowest level of genetic polymorphism ($\pi = 0.00323$, $\theta = 0.00294$) and San Juan Bautista (Central America, Honduras) was the most polymorphic one ($\pi = 0.00748$, $\theta = 0.00669$). *Lu. cruzi* and *Lu. pseudolongipalpis* presented low levels of genetic polymorphism ($\pi = 0.00185$, $\theta = 0.00366$; $\pi = 0.00154$, $\theta = 0.00229$ respectively), similar to those found in some Brazilian *Lu. longipalpis* populations. Tajima's D and Fu & Li's D tests of neutrality (Tajima 1989; Fu & Li 1993) indicated no deviation of the neutral model of evolution for most of the populations analyzed. For Sobral 1S (S1S), Palmas intermediate (PALi) and Pirenópolis (PIR) the tests were not significant after Bonferroni's correction.

Among Brazilian samples of *Lu. longipalpis*, the genetic differentiation expressed by Fst values (Table 2) indicated high levels of divergence between populations that produce different copulation songs, ranging from 0.7589 (ACLA – Burst X E1S – Pulse) to 0.9838 (MAR – Burst X LAP - Pulse). Indeed, mean pairwise Fst values revealed that the highest level of genetic divergence as well as fixed differences occur between populations producing different patterns of copulation songs (Burst x Pulse = 0.8613) when compared to those that produce the same song (Pulse x Pulse = 0.08973; Burst x Burst = 0.0506). The only exceptions are the populations of Estrela 1S where the spot phenotype is no completely reliable separating the sympatric

siblings (see Araki et al 2009; Lins et al 2012) and Itamaracá, which produces the Burst type song but show no fixed differences in the comparisons with Pulse type populations, despite the high degree of genetic differentiation. Interestingly enough, a single sequence found in an heterozygous fly carrying also a typical "Burst-type haplotype, is responsible for the lack of fixed differences between Itamaraca and Pulse-type song populations, suggesting that this fly might represent an hybrid of introgression event.

Only one Brazilian population producing a rare Mixed type of copulation song (RBON) was included in this analysis. This rare song was only previously found in a small sample from another locality (Mesquita) also in Rio de Janeiro state (Araki et al 2009). When compared to other *Lu. longipalpis* populations from Brazil that produce either Pulse or Burst-type songs, RBON revealed very high levels of genetic differentiation in all comparisons ranging from 0.3924 (RBON x BGUA) to 1 (RBON x MAR). Nevertheless, no fixed differences were found when RBON was compared to BGUA and J1S. On the other hand, RBON presented 1 fixed difference when compared to other Pulse type populations and 4 fixed differences in most comparisons with Burst type populations. Interestingly, no fixed difference was detected when RBON, BGUA and J1S were compared to *Lu. longipalpis* populations from Colombia (PG and DUR), Venezuela (TRU) and Central America – Honduras (IDT, BRA and SJB). Argentina and Paraguay populations have shown low values of Fst and no fixed differences when compared to Brazilian Burst type populations. Significance Fst p values are shown in additional file S3.

Consistent with the findings of Arrivillaga et al (2002, 2003), populations defined as Species C - cis-Andean (Colombia: PG and DUR; and northwestern Venezuela: TRU) have shown low levels of genetic differentiation and no fixed difference. Also, as suggested by the authors mentioned above, our data reinforces the subdivision of Species C into two groups as no divergence was found between the

Colombian populations of *Lu. longipalpis* and a moderate genetic differentiation was detected when they were compared to TRU, from Venezuela. A Brazilian sample apparently related to species C is Roraima, also confirming Arrivillaga et al (2003). This population presented lower Fst values compared with TRU, DUR and PG as when compared with other *Lu. longipalpis* populations. Accordingly, populations classified as Species D – trans-Andean (IDT, BRA, SJB) (Arrivillaga et al 2002, 2003) presented low levels of divergence and no fixed difference.

Lu. cruzi, a close related species to Brazilian *Lu. longipalpis* populations that produces Burst type song, has shown low levels of divergence not only against Burst type populations previously published (Lins et al 2012) but also against the new ones producing this song type included in the present study. Although Lins et al (2012) have found 2 fixed differences between *Lu. cruzi* and Burst type populations, no fixed differences were encountered in all comparisons with *Lu. cruzi* and these populations, as a result of new indel regions including these fixed differenced previously detected.

As expected, comparisons between *Lu. pseudolongipalpis* and other *Lu. longipalpis* and *Lu. cruzi* populations have shown very high genetic differentiation and variable number of fixed differences.

Figure 2 shows a Neighbor-Joining tree constructed with the Fst values of all comparisons of *Lu. longipalpis*, *Lu. cruzi*, *Lu. pseudolongipalpis* populations. Brazilian *Lu. longipalpis* populations were grouped according to the copulation song they produce (Burst or Pulse). Rio Bonito population, which produces the Mixed-type copulation song, was grouped with Barra de Guaratiba (no recorded song to date).

Analysis of molecular variance (AMOVA) was performed to examine the partition of *para* sequence variation within *Lu. longipalpis* populations studied (Table 3). Most part of the variation (80.44%) was observed between the proposed groups: 1 - Pulse-type populations (LAP, JAC, S1S, J1S, TER, JEQ, LASS, PAL1S, PALi, PIR,

E1S), 2 – Burst-type *Lu. longipalpis* populations (S2S, PAN, NAT, J2S, E2S, BAR, PASS, CGRA, NPOR, ACLA, CMR, CMT, IPA, PAL2S, ITA, MAR), 3 – RBON and BGUA, 4 – IDT, BRA, SJB and 5 – RR, PG, DUR, TRU. This result reinforces not only the species division proposed by Arrivillaga et al (2002, 2003) but also the existence of the two main song groups (Burst x Pulse) in Brazil. Furthermore, the smallest part of the molecular variation was detected among populations within groups (1.41%) whereas 18.15% of this variation was within populations.

A haplotype network (Figure 3) was constructed based on a 254bp non-recombinant *para* fragment. In total were analyzed 748 sequences corresponding to 57 haplotypes. A single network was generated with the 95% connection limit. *Lu. pseudolongipalpis* did not connect to this network even with the reduction of the connection limit to 90% (the minimum allowed by the program). Haplotype classification here has no relation with the previous whole fragment haplotype numbering. The main haplotype generated in this network was H1, corresponding to Brazilian Pulse-type song *Lu. longipalpis* populations, Roraima and the other samples from the northern South (Colombia and Venezuela) and Central (Honduras) America. The main haplotype of Burst-type song populations was H6. It is possible do observe a clear separation between the two groups producing different copulation songs. These two groups are connected by three mutations between H8 and H11. Interestingly, H8 contains a single sequence of Estrela 1S (E1S8A), which produces the Pulse-type song. The other allele of the same fly (E1S8B) is found in the H7 haplotype. This probably represents a case where the spot phenotype does not match the song phenotype (Araki et al 2009; Lins et al 2002). The sequences from Araçatuba (Brazil, no recorded song to date), Argentina and Paraguay clustered with the Burst-type song populations. *Lu. cruzi* (which also produces the Burst-type song) sequences, classified as haplotypes H30,

H32, H36 and H37, were connected by a single mutation to H23 and placed with the *Lu. longipalpis* Burst-type song group of populations.

Discussion

Neglected tropical diseases have been receiving increasingly attention as they blight the lives of a billion people worldwide and threaten the health of millions more (WHO, 2010). Understanding the biology and the genetic structure of a particular vector is an important component for explaining and predicting the epidemiology of a number of diseases. In this context, vectors comprising species complexes have driven the attention of epidemiologists and health organizations because of the complexities they bring in understanding the spreading of these tropical diseases. Knowledge on the population structure of the *Lu. longipalpis* species complex may ultimately shed light on anomalies in the epidemiology of visceral leishmaniasis in the New World (Arrivillaga et al 2003).

Although a number of studies support the existence of the *Lu. longipalpis* complex, the number of species and their geographical range still remains unknown. Arrivillaga et al (2002, 2003) proposed the existence of four different species: Species A (comprising most populations of Brazil except Roraima State) - *Lu. longipalpis sensu stricto*; Species B (Laran) - *Lu. pseudolongipalpis*, Species C (cis-Andean) and Species D (trans-Andean). Although no fixed difference was encountered, the genetic divergence provided by the molecular analysis of the *para* gene strongly supports Arrivillaga's classification of Species C and D. Comparisons between *Lu. pseudolongipalpis* and all other *Lu. longipalpis* populations revealed very high levels of genetic differentiation and fixed differences, supporting the previous classification as Species B (Arrivillaga et al 2003).

However, as mentioned before a number of studies indicate that "Species A" does not represent a single species in Brazil (Ward et al 1998; Bauzer et al 2007; Araki et al 2009; Lins et al 2012). Combined molecular analyses of *per* and *para* genes (Bauzer et al 2002a,b; Araki et al 2009; Lins et al 2008, 2012a), microsatellites (Maingon et al 2003; Watts et al 2005), copulation songs (Souza et al 2002; Araki et al 2009) and pheromones (Hamilton et al 1999a,b, 2004, 2005; Araki et al 2009) have provided sufficient evidence to support a complex of Brazilian sibling species which can be divided in two main groups according to the song males produce (Burst or Pulse). Together with Lins et al (2012) findings, our data strongly supports this idea, as *para* revealed high levels of genetic divergence and fixed differences between the *Lu. longipalpis* populations producing different songs.

Samples from Araçatuba, Argentina and Paraguay had no recorded copulatory song recorded to date, nonetheless, these populations share the same WFhap_5 additional files S1 and S2) as the other Burst type *Lu. longipalpis* populations and were grouped with these populations in both the NJ Fst tree (Figure 2) and the haplotype network (Figure 3). In addition, low levels of genetic differentiation as well as the absence of fixed differences suggested that they might represent siblings of the same species. On the other hand, WFhap_3 (additional files S1 and S2) was shared by all Pulse type populations previously published: LAP, JAC, S1S, TER, J1S, E1S (Lins et al 2012) and other included in this study: JEQ, LASS, PAL1S, PALiS, PIR. Together with the occurrence of fixed differences, the existence of specific haplotypes for each song group raises the question whether *paralytic* gene would be one of the genes modulating the song type in *Lu. longipalpis* populations, further allowing diagnostic assays to identify females from different phenotype groups and also males which have not had their song recorded.

All sequences of Rio Bonito population, which produces a Mixed pattern copulation song, correspond to WFhap_1 (additional files S1 and S2). It is suggestive that this population might represent a different species of the *Lu. longipalpis* complex. Interestingly, the population of Barra de Guaratiba (Rio de Janeiro State) also shares this haplotype and was grouped with RBON in both the NJ Fst tree and the haplotype network. Furthermore, no fixed difference was found between these populations. WFHap_1 is also shared by *Lu. longipalpis* populations of Roraima (Brazil),Trujillo (Venezuela), Durania (Colombia), Palo Gordo (Colombia), Brasilito (Costa Rica), Isla del Tigre and San Juan Bautista (Venezuela). These populations were grouped in a single branch in the NJ Fst tree.

A number of studies have provided for the existence of sympatric siblings in the localities of Sobral, Jaíba and Estrela (e.g, Araki et al 2009, Lins et al 2012, for two of the most recent papers). In these localities, males can be differentiated by the presence of 1 or 2 abdominal spots with rare or relative small number of males with the intermediate phenotype. They produce different copulation songs (Burst, 2S males and Pulse, 1S males, in all three localities, albeit with different subtype of Pulse song) and pheromones: Germacrene (S1S) and Cembrene 1 (S2S) in Sobral, Cembrene 2 (J1S) and Cembrene 1 (J2S) in the Jaíba. A particular case occurs in Estrela though, whereas the same Cembrene 1 pheromone is produced by both 1S and 2S males. Our data confirms the existence of a new pair of sympatric siblings in the region of Palmas (State of Tocantins). Males 1S and 2S from Palmas produce different copulation songs (Vigoder et al, in preparation): Burst in PAL2S and Pulse in PAL1S. Moreover, high genetic divergence (82.77%) and fixed differences (3) were found between them. Comparisons between PALi and PAL2S have also shown high levels of genetic differentiation (81.75%) and fixed differences (3), similar to those obtained for PAL1S and PAL2S, suggesting that (most) intermediate flies in this locality belong to the PAL1S species.

Previous analyses with *per* and *para* genes indicated that *Lu. cruzi* is another member of the *Lu. longipalpis* complex (Araki et al 2009; Lins et al 2012a). *Lu. cruzi* males produce a variation of the Burst-type song with shorter inter-burst interval (Vigoder et al 2010) and the Germacrene sex pheromone (Brazil et al 2002), also found in many Pulse-type populations of *Lu. longipalpis* (Araki et al 2009). Lins et al (2012) demonstrated that this species is more closely related to the Brazilian *Lu. longipalpis* that produce the Burst type song showing a lower Fst value and smaller number of fixed differences between them than those found between *Lu. cruzi* and Pulse-type populations. Here we have encountered even lower levels of genetic differentiation and no fixed difference between *Lu. cruzi* and *Lu. longipalpis* Burst type populations, a result of new indels including these previously detected fixed differences. In addition, *Lu. cruzi* also shares WFHap_5, the most frequent in Burst type *Lu. longipalpis* populations.

Finally, the complexity found in complexes of sibling species represents a difficult task for evolutionary biologists, as species delimitation is conflicting specially when reproductive isolation is still incomplete and introgression occurs. Combined molecular and behavioral (song and pheromones) analyses have provided the most consistent results about the structure of *Lu. longipalpis* species complex in Brazil. Data provided by the molecular analysis of *para* gene in this macrogeographic perspective which included samples from other South and Central American localities have provided important clues about the distribution of the different sibling species of the complex, conciliating previous published data obtained with different types molecular markers (Arrivilaga et al 2003; Watts at al 2005; Araki et al 2009; Lins et al 2012).

Acknowledgements

We would like to thank Dr Gabriel E. M. Ferreira for collecting some of the *Lu. longipalpis* sand flies samples, Dr Yoosook Lee and Catelyn Neiman (University of California, Davis) for the laboratory assistance at UC Davis and Robson Costa da Silva for technical assistance at FIOCRUZ, Brazil.

References

Araki AS, Vigoder FM, Bauzer LG, Ferreira GE, Souza NA, et al. (2009) Molecular and behavioral differentiation among Brazilian populations of *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae). PLoS Negl Trop Dis 3:e365.

Arrivillaga J, Mutebi JP, Pinango H, Norris D, Alexandre B, et al. (2003) The taxonomic status of genetically divergent populations of *Lutzomyia longipalpis* (Diptera: Psychodidae) based on the distribution of mitochondrial and isozyme variation. J Med Entomol 40: 615-627.

Arrivillaga JC, Norris DE, Feliciangeli MD, Lanzaro GC (2002) Phylogeography of the neotropical sand fly *Lutzomyia longipalpis* inferred from mitochondrial DNA sequences. Infect Genet Evol Dec;2(2):83-95.

Bauzer LG, Gesto JS, Souza NA, Ward RD, Hamilton JG, et al. (2002b) Molecular divergence in the *period* gene between two putative sympatric species of the *Lutzomyia longipalpis* complex. Mol Biol Evol 19: 1624-1627.

Bauzer LG, Souza NA, Maingon RDC, Peixoto AA (2007) *Lutzomyia longipalpis* in Brazil: a complex or a single species? A mini-review. Mem Inst Oswaldo Cruz 102: 1-12.

Bauzer LG, Souza NA, Ward RD, Kyriacou CP, Peixoto AA (2002a) The *period* gene and genetic differentiation between three Brazilian populations of *Lutzomyia longipalpis*. Insect Mol Biol 11: 315-323.

Bottecchia M, Oliveira SG, Bauzer LG, Souza NA, Ward RD, Garner KJ, Kyriacou CP, Peixoto AA (2004) Genetic divergence in the cacophony IVS6 intron among five

Brazilian populations of *Lutzomyia longipalpis*. J Mol Evol Jun;58(6):754-61.

Brazil RP, Hamilton JG (2002). Isolation and identification of 9-methylgermacrene-B as the putative sex pheromone of *Lutzomyia cruzi* (Mangabeira, 1938) (Diptera: Psychodidae). Mem Inst Oswaldo Cruz 97: 435-436.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9: 1657-1659.

de Pita-Pereira D, Cardoso MA, Alves CR, Brazil RP, Britto C (2008) Detection of natural infection in *Lutzomyia cruzi* and *Lutzomyia forattinii* (Diptera: Psychodidae: Phlebotominae) by Leishmania infantum chagasi in an endemic area of visceral leishmaniasis in Brazil using a PCR multiplex assay. Acta Trop. 107: 66-69.

de Souza NA, Ward RD, Hamilton JG, Kyriacou CP, Peixoto AA (2002) Copulation songs in three siblings of *Lutzomyia longipalpis* (Diptera: Psychodidae). Trans R Soc Trop Med Hyg 96: 102-103.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.0: An integrated software package for population genetics data analysis. Evol Bioinformatics Online 1: 47-50.

Filatov DA (2002) ProSeq: A software for preparation and evolutionary analysis of DNA sequence data sets. Molecular Ecology Notes 2: 621–624.

Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. Genetics 133: 693-709.

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95-98.

Hamilton JG, Brazil RP, Maingon R (2004) A fourth chemotype of *Lutzomyia longipalpis* (Diptera: Psychodidae) from Jaibas, Minas Gerais State, Brazil. J Med Entomol 41: 1021-1026.

Hamilton JG, Hooper AM, Mori K, Pickett JA, Sano A (1999a) 3-Methyl-α-himachalene confirmed, and the relative stereochemistry defined, by synthesis as the sex pheromone of the sandfly *Lutzomyia longipalpis* from Jacobina, Brazil. Chem Commun 355-356.

Hamilton JG, Ibbotson HC, Hooper AM, Mori K, Picket JA, et al. (1999b) 9-Methylgermacrene-B confirmed by synthesis as the sex pheromone of the sandfly *Lutzomyia longipalpis* from Lapinha, Brazil, and the absolute stereochemistry defined as 9S. Chem Commun 2335-2336.

Hamilton JG, Maingon RD, Alexander B, Ward RD, Brazil RP (2005) Analysis of the sex pheromone extract of individual male *Lutzomyia longipalpis* sandflies from six regions in Brazil. Med Vet Entomol 19: 480–488.

Lanzaro GC, Ostrovska K, Herrero MV, Lawyer PG, Warburg A (1993) *Lutzomyia longipalpis* is a species complex: genetic divergence and interspecific hybrid sterility among three populations. Am J Trop Med Hyg 48: 839-847.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Lins RM, Souza NA, Peixoto AA (2008) Genetic divergence between two sympatric species of the *Lutzomyia longipalpis* complex in the paralytic gene, a locus associated with insecticide resistance and lovesong production. Mem Inst Oswaldo Cruz. 103: 736-740.

Lins RMMA, Souza NA, Brazil RP, Maingon RDC, Peixoto AA (2012) Fixed differences in the *paralytic* gene define two lineages within the *Lutzomyia longipalpis* complex producing different types of courtship songs. PLOS one

Maingon RD, Ward RD, Hamilton JG, Bauzer LG, Peixoto AA (2008) The *Lutzomyia longipalpis* species complex: does population sub-structure matter to Leishmania transmission? Trends Parasitol 24: 12-17.

Mangabeira O (1969) Sobre a sistemática e biologia de Phlebotomus do Ceará. Rev Brasil Malariol Doenças Trop 21: 3-26.

Peixoto AA, Hall JC (1998) Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of *Drosophila.* Genetics 148: 827-838.

Sambrook D, Russell J (2001) Molecular cloning: a laboratory manual. Cold Spring Harbour: CSHL Press. 2100 p.

Souza NA, Andrade-Coelho CA, Vigoder FM, Ward RD, Peixoto AA (2008) Reproductive isolation between sympatric and allopatric Brazilian populations of *Lutzomyia longipalpis s.l.* (Diptera: Psychodidae). Mem Inst Oswaldo Cruz 103: 216-219.

Souza NA, Vigoder FM, Araki AS, Ward RD, Kyriacou CP, et al. (2004) Analysis of the copulatory courtship songs of *Lutzomyia longipalpis* in six populations from Brazil. J Med Entomol 41: 906-913.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.

Vigoder FM, Araki AS, Bauzer LG, Souza NA, Brazil RP, Peixoto AA (2010) Lovesongs and period gene polymorphisms indicate *Lutzomyia cruzi* (Mangabeira, 1938) as a sibling species of the *Lutzomyia longipalpis* (Lutz and Neiva, 1912) complex. Infect Genet Evol 10: 734-739.

Ward RD, Phillips A, Burnet B, Marcondes CB (1988) The *Lutzomyia longipalpis* complex: reproduction and distribution. In: Service MW, editor. Biosystematics of Haematophagous Insects. Oxford: Oxford University Press. pp. 258-269.

Ward RD, Ribeiro AL, Ready PD, Murtagh A (1983) Reproductive isolation between

different forms of *Lutzomyia longipalpis* (Lutz & Neiva) (Diptera: Psy- chodidae), the vector of *Leishmania donovani chagasi* Cunha & Chagas and its significance to kala- azar distribution in South America. Mem Inst Oswaldo Cruz 78: 269-280.

Watts PC, Hamilton JG, Ward RD, Noyes HA, Souza NA, et al. (2005) Male sex pheromones and the phylogeographic structure of the *Lutzomyia longipalpis* species complex (Diptera: Psychodidae) from Brazil and Venezuela. Am J Trop Med Hyg 73: 734-743.

WHO (2010) First WHO report on neglected tropical diseases: working to overcome the global impact of neglected tropical diseases. http://www.who.int/neglected_diseases/2010report/en/

Woerner A, Cox M, Hammer M (2007) Recombination-Filtered Genomic Datasets by Information Maximization. Bioinformatics 23: 1851-1853.

Yin H, Mutebi JP, Marriott S, Lanzaro GC (1999) Metaphase karyotypes and G-banding in sandflies of the *Lutzomyia longipalpis* complex. Med Vet Entomol. 13: 72-77.

Yin H, Norris DE, Lanzaro GC (2000) Sibling species in the *Lutzomyia longipalpis* complex differ in levels of mRNA expression for the salivary peptide, maxadilan. Insect Mol Biol Jun;9(3):309-14.

Young DG, Duncan M (1994) Guide to the identification and geographic distribution of *Lutzomyia* sand flies in Mexico, the West Indies, Central and South America (Diptera: Psychodidae). Mem Am Entomol Inst 54: 1-881.

Figure legends

Figure 1 – Approximate geographic location of samples from South and Central America studied.

Figure 2 – Neighbor Joining tree of sequences from Brazilian *Lu. longipalpis* producing the Burst-type song (green circles), Pulse-type song (orange circles), Mixed-type song (magenta circle) and no recorded song (black circles). South (not Brazilian) and Central America samples are represented in squares: Arrivillaga's Species C are represented in yellow squares, Species D in blue squares, and Species B in red square. The populations of Argentina and Paraguay are represented in open squares. *Lu. cruzi*, which also produces the burst type song is represented in green triangle.

Figure 3 - Haplotype network of populations of *Lu. longipalpis* and *Lu. cruzi*. Each population is represented by a different color and each node represents a unique haplotype.

Supporting information

S1 – Whole Fragment Haplotype polymorphic sites alignment (WFHap). The region corresponding to the intron sequence is highlighted in grey. Dots indicate the same nucleotide.

S2 - Distribution of the 57 haplotypes found among *Lu. longipalpis*, *Lu. cruzi* and *Lu. pseudolongipalpis* samples and the number of sequences represented in each population.

S3 - Fst significance (p values) obtained with 1023 random permutations.

Table 1: Polymorphism summaries of the *para* gene fragment from populations of *Lu. longipalpis*, *Lu. cruzi* and *Lu. pseudolongipalpis*.

Population	n	S	π	θ	D_T	D_{FL}
LAP	28	2	0.00158	0.00138	0.30938	-0.71444
JAC	22	9	0.00505	0.00667	-0.81288	-1.28372
S1S	32	12	0.00327	0.00803	-2.03335	-1.20010
J1S	24	6	0.00233	0.00432	-1.39438	-0.97286
TER	24	9	0.00369	0.00651	-1.42001	-2.44907
LASS	22	4	0.00180	0.00297	-1.10208	-0.81047
PAL1S	24	6	0.00256	0.00437	-1.25116	-0.23775
PALi	14	4	0.00154	0.00338	-1.79759	-2.27380
PIR	24	10	0.00324	0.00720	-1.83317	1.61487
E1S	20	5	0.00208	0.00382	-1.39152	-0.41302
S2S	28	8	0.00399	0.00559	-0.88246	0.08210
PAN	32	10	0.00473	0.00673	-0.92953	-1.33584
NAT	24	11	0.00468	0.00800	-1.40808	-1.83657
J2S	24	11	0.00472	0.00798	-1.38309	-1.36881
E2S	20	5	0.00297	0.00381	-0.67163	-0.41302
BAR	22	4	0.00218	0.00297	-0.74755	-0.81047
PASS	12	6	0.00346	0.00540	-1.37073	-1.10843
CGRA	18	2	0.00059	0.00155	-1.50776	-1.98899
NPOR	22	13	0.00514	0.00988	-1.69741	-2.09264
ACLA	24	10	0.00541	0.00728	-0.85812	-1.61487
CMR	24	3	0.00187	0.00219	-0.37127	-1.35733
CMT	8	5	0.00484	0.00523	-0.33547	0,12651
IPA	8	6	0.00502	0.00625	-0.92039	-0.72189
PAL2S	24	13	0.00694	0.00941	-0.90925	-0.96989
ITA	22	9	0.00416	0.00669	-1.26541	-1.28372
MAR	6	3	0.00342	0.00355	-0.18544	-0.37481
ARG	24	10	0.00514	0.00718	-0.94840	-2.11862
ARA	24	7	0.00356	0.00512	-0.95068	-0.68707
BGUA	12	1	0.00134	0.00092	1.06589	0.75202
RR	18	7	0.00492	0.00546	-0.33106	0.70089
PAR	8	4	0.00468	0.00412	0.58623	0.56807
TRU	16	5	0.00395	0.00404	-0.06961	-0.29672
PG	22	4	0.00323	0.00294	0.27174	0.14251
DUR	14	4	0.00289	0.00337	-0.47373	-0.55476
IDT	22	6	0.00529	0.00457	0.48733	0.52088
BRA	24	6	0.00651	0.00448	1.37450	1.23249
SJB	24	9	0.00748	0.00669	0.38342	-0.26490
CRU	24	5	0.00185	0.00366	-1.43150	-2.17281
PSEUDO	20	3	0.00154	0.00229	-0.87550	1.00649

n. number of sequences; **S**. number of segregating sites; **π**. nucleotide diversity; **θ**. neutral parameter based on the segregating sites; D_T. Tajima test of neutrality. D_{FL} Fu and Li's test of neutrality. **LAP**: lapinha, **JAC**: Jacobina, **S1S**: Sobral 1S, **J1S**: Jaíba 1S, **TER**: Teresina, **LASS**: Lassance, **PAL 1S**: Palmas 1S, **PALi**: Palmas intermediate, **PIR**: Pirenópolis, **E1S**: Estrela 1S, **S2S**: Sobral 2S, **PAN**: Pancas, **NAT**: Natal, **J2S**: Jaíba 2S, **E2S**: Estrela 2S, **BAR**: Barcarena, **PASS**: Passira, **CGRA**: Campo Grande, **NPOR**: Nova Porteirinha, **ACLA**: Afonso Cláudio, **CMR**: Camará, **CMT**: Cametá, **IPA**: Ipanema, **PAL 2S**: Palmas 2S, **ITA**: Itamaracá, **MAR**: Marajó, **ARG**: Argentina, **ARA**: Araçatuba, **BGUA**: Barra de Guaratiba, **RR**: Roraima, **PAR**: Paraguay, **TRU**: Trugillo, **PG**: Palo Gordo, **DUR**: Durania, **IDT**: Isla del Tigre, **BRA**: Brasilito, **SJB**: San Juan Bautista, **CRU**: Lu. cruzi, **PSEUDO**: Lu. pseudolongipalpis.

Table 2: Pairwise differentiation and fixed differences between populations of *Lu. Longipalpis*, *Lu. Cruzi*, and *Lu. pseudolongipalpis*.

	Pulse-type populations								N	Pulse-type		Burst-type populations									
	LAP	JAC	S1S	J1S	TER	JEQ	LASS	PAL1S	PALi	PIR	E1S	S2S	PAN	NAT	J2S	E2S	BAR	PASS	CGRA	NPOR	ACLA
LAP	0.0000	0	0	0	0	0	0	0	0	0	0	3	3	3	3	3	3	3	3	3	3
JAC	0.3316	0.0000	0	0	0	0	0	0	0	0	0	2	2	2	2	2	2	2	2	2	2
S1S	0.0423	0.2150	0.0000	0	0	0	0	0	0	0	0	3	3	3	3	3	3	3	3	3	3
J1S	0.0579	0.2043	0.0000	0.0000	0	0	0	0	0	0	0	3	3	3	3	3	3	3	3	3	3
TER	0.1590	0.1911	0.0134	0.0000	0.0000	0	0	0	0	0	0	3	3	3	3	3	3	3	3	3	3
JEQ	0.6560	0.0146	0.0450	0.0861	0.0000	0.0000	0	0	0	0	0	3	3	3	3	3	3	3	3	3	3
LASS	0.0024	0.2772	0.0091	0.0214	0.1091	0.3759	0.0000	0	0	0	0	3	3	3	3	3	3	3	3	3	3
PAL1S	0.0522	0.2301	0.0000	0.0000	0.0029	0.1293	0.0132	0.0000	0	0	0	3	3	3	3	3	3	3	3	3	3
PALi	0.0167	0.2046	0.0000	0.0000	0.0189	0.2289	0.0000	0.0000	0.0000	0	0	3	3	3	3	3	3	3	3	3	3
PIR	0.0865	0.2198	0.0000	0.0068	0.0307	0.0000	0.0252	0.0000	0.0000	0.0000	0	3	3	3	3	3	3	3	3	3	3
E1S	0.0676	0.1827	0.0698	0.0654	0.1213	0.1141	0.0505	0.0692	0.0338	0.0840	0.0000	0	0	0	0	0	0	0	0	0	0
S2S	0.9118	0.7700	0.8625	0.8639	0.8494	0.8527	0.8901	0.8721	0.8749	0.8546	0.8131	0.0000	0	0	0	0	0	0	0	0	0
PAN	0.9060	0.7745	0.8619	0.8624	0.8494	0.8493	0.8859	0.8700	0.8711	0.8544	0.8146	0.0263	0.0000	0	0	0	0	0	0	0	0
NAT	0.9161	0.7626	0.8624	0.8639	0.8477	0.8496	0.8929	0.8729	0.8763	0.8538	0.8116	0.0000	0.0076	0.0000	0	0	0	0	0	0	0
J2S	0.9089	0.7570	0.8573	0.8575	0.8418	0.8355	0.8854	0.8663	0.8671	0.8480	0.8047	0.0660	0.0547	0.0330	0.0000	0	0	0	0	0	0
E2S	0.9470	0.7842	0.8842	0.8899	0.8710	0.8992	0.9246	0.9001	0.9128	0.8782	0.8412	0.0000	0.0069	0.0000	0.0218	0.0000	0	0	0	0	0
BAR	0.9702	0.8196	0.9073	0.9168	0.8979	0.9500	0.9516	0.9265	0.9461	0.9050	0.8758	0.0467	0.0619	0.0387	0.1020	0.0588	0.0000	0	0	0	0
PASS	0.9408	0.7455	0.8686	0.8717	0.8488	0.8490	0.9125	0.8837	0.8928	0.8579	0.8131	0.0486	0.0283	0.0143	0.0000	0.0006	0.1278	0.0000	0	0	0
CGRA	0.9652	0.7984	0.8972	0.9058	0.8853	0.9340	0.9439	0.9166	0.9358	0.8930	0.8595	0.0003	0.0090	0.0000	0.0358	0.0000	0.0368	0.0290	0.0000	0	0
NPOR	0.8942	0.7360	0.8429	0.8407	0.8248	0.8012	0.8682	0.8496	0.8458	0.8311	0.7833	0.0473	0.0380	0.0207	0.0000	0.0000	0.0827	0.0000	0.0149	0.0000	0
ACLA	0.8713	0.7136	0.8238	0.8189	0.8047	0.7677	0.8441	0.8277	0.8194	0.8103	0.7589	0.0176	0.0129	0.0011	0.0282	0.0000	0.0919	0.0031	0.0215	0.0105	0.0000
CMR	0.9627	0.8199	0.9044	0.9125	0.8950	0.9373	0.9445	0.9216	0.9377	0.9016	0.8721	0.1563	0.1611	0.1520	0.1918	0.2015	0.0786	0.2283	0.2232	0.1657	0.1638
CMT	0.9337	0.7138	0.8543	0.8546	0.8295	0.7851	0.8999	0.8683	0.8725	0.8393	0.7854	0.0268	0.0486	0.0121	0.0366	0.0324	0.0788	0.0234	0.0725	0.0188	0.0000
IPA	0.9566	0.7408	0.8727	0.8779	0.8517	0.8650	0.9263	0.8917	0.9079	0.8615	0.8131	0.0000	0.0203	0.0000	0.0489	0.0165	0.0700	0.0472	0.0176	0.0256	0.0059
PAL2S	0.8697	0.7188	0.8255	0.8197	0.8056	0.7605	0.8430	0.8277	0.8175	0.8114	0.7638	0.1141	0.1095	0.0847	0.0983	0.0863	0.1668	0.0188	0.1078	0.0729	0.0624
ITA	0.9086	0.7431	0.8500	0.8511	0.8335	0.8311	0.8827	0.8609	0.8635	0.8402	0.7919	0.0284	0.0352	0.0068	0.0332	0.0122	0.0377	0.0275	0.0000	0.0293	0.0466
MAR	0.9838	0.7598	0.8914	0.9016	0.8724	0.9586	0.9564	0.9165	0.9479	0.8834	0.8411	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0083	0.0000	0.0000	0.0000
ARG	0.8292	0.6700	0.7883	0.7778	0.7651	0.6936	0.7995	0.7861	0.7673	0.7706	0.7119	0.1079	0.1385	0.0948	0.1330	0.1187	0.1718	0.0996	0.1298	0.1150	0.0869
RBON	0.9714	0.6529	0.8000	0.8203	0.7826	0.9675	0.9284	0.8595	0.9253	0.8106	0.8080	0.9317	0.9246	0.9349	0.9280	0.9626	0.9834	0.9576	0.9795	0.9146	0.8916
ARA	0.8118	0.6573	0.7744	0.7617	0.7503	0.6657	0.7811	0.7697	0.7468	0.7550	0.6929	0.0731	0.0980	0.0630	0.1006	0.0769	0.1193	0.0678	0.0828	0.0850	0.0610
BGUA	0.8889	0.4997	0.6884	0.6892	0.6489	0.7231	0.8204	0.7399	0.7833	0.6848	0.6733	0.8913	0.8864	0.8908	0.8820	0.9225	0.9535	0.9016	0.9435	0.8619	0.8361
RR	0.6908	0.4387	0.5848	0.5609	0.5428	0.3683	0.6295	0.5945	0.5695	0.5627	0.5562	0.8223	0.8244	0.8146	0.8095	0.8296	0.8605	0.7935	0.8412	0.7912	0.7705
PAR	0.8950	0.6631	0.8211	0.8140	0.7898	0.6607	0.8556	0.8276	0.8141	0.8003	0.7389	0.2166	0.2475	0.1964	0.2283	0.2586	0.3747	0.1867	0.3009	0.1900	0.1382
TRU	0.7003	0.4345	0.5834	0.5598	0.5444	0.3774	0.6333	0.5963	0.5742	0.5592	0.5485	0.8275	0.8293	0.8208	0.8153	0.8383	0.8705	0.8024	0.8514	0.7959	0.7726

PG	0.7395	0.4571	0.6057	0.5891	0.5697	0.4744	0.6782	0.6314	0.6298	0.5909	0.5983	0.8578	0.8567	0.8540	0.8485	0.8732	0.9003	0.8492	0.8865	0.8323	0.8107
DUR	0.8056	0.4830	0.6502	0.6398	0.6096	0.5093	0.7377	0.6809	0.6870	0.6293	0.6321	0.8642	0.8621	0.8604	0.8533	0.8848	0.9164	0.8560	0.9017	0.8338	0.8091
IDT	0.7132	0.4436	0.5921	0.5786	0.5535	0.4747	0.6576	0.6113	0.6022	0.5818	0.5679	0.8467	0.8467	0.8424	0.8364	0.8620	0.8891	0.8349	0.8748	0.8197	0.8002
BRA	0.6038	0.3606	0.4931	0.4690	0.4541	0.3268	0.5474	0.5065	0.4850	0.4841	0.4746	0.8233	0.8251	0.8173	0.8121	0.8336	0.8610	0.8034	0.8445	0.7955	0.7773
SJB	0.5935	0.3525	0.4836	0.4587	0.4444	0.2997	0.5367	0.4966	0.4745	0.4746	0.4631	0.8205	0.8227	0.8145	0.8095	0.8308	0.8585	0.8004	0.8419	0.7928	0.7743
CRU	0.8176	0.6718	0.7818	0.7695	0.7584	0.6763	0.7878	0.7769	0.7544	0.7625	0.7021	0.0798	0.0931	0.0711	0.0998	0.0776	0.1057	0.0677	0.0764	0.0861	0.0762
PSEUDO	0.9729	0.8476	0.9174	0.9253	0.9081	0.9517	0.9563	0.9336	0.9506	0.9146	0.8995	0.8987	0.8946	0.9014	0.8945	0.9323	0.9576	0.9210	0.9507	0.8781	0.8561

	Burst-type populations						**N**	**M**	**N**	**N**	**N**	**N**	**N**	**N**	**N**	**N**	**N**	**N**	**B**	**N**
	CMR	**CMT**	**IPA**	**PAL2S**	**ITA**	**MAR**	**ARG**	**RBON**	**ARA**	**BGUA**	**RR**	**PAR**	**TRU**	**PG**	**DUR**	**IDT**	**BRA**	**SJB**	**CRU**	**PSEUDO**
LAP	3	3	3	3	0	3	3	1	3	1	1	3	1	1	1	1	1	1	3	5
JAC	2	2	2	2	0	2	2	1	2	1	1	2	1	1	1	1	1	0	2	6
S1S	3	3	3	3	0	3	3	1	3	1	1	3	1	1	1	1	1	1	3	6
J1S	3	3	3	3	0	3	3	0	3	0	0	3	0	0	0	0	0	0	3	6
TER	3	3	3	3	0	3	3	1	3	1	1	3	1	1	1	1	1	1	3	6
JEQ	3	3	3	3	0	3	3	1	3	1	1	3	1	1	1	1	1	1	3	6
LASS	3	3	3	3	0	3	3	1	3	1	1	3	1	1	1	1	1	1	3	6
PAL1S	3	3	3	3	0	3	3	1	3	1	1	3	1	1	1	1	1	1	3	5
PALi	3	3	3	3	0	3	3	1	3	1	1	3	1	1	1	1	1	1	3	6
PIR	3	3	3	3	0	3	3	1	3	1	1	3	1	1	1	1	1	1	3	6
E1S	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1	1	1	1	0	4
S2S	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
PAN	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
NAT	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
J2S	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	4
E2S	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
BAR	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
PASS	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
CGRA	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
NPOR	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	4
ACLA	0	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
CMR	0.0000	0	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
CMT	0.0969	0.0000	0	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
IPA	0.2038	0.0204	0.0000	0	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
PAL2S	0.2224	0.0646	0.0943	0.0000	0	0	0	4	0	4	3	0	3	3	3	3	3	3	0	5
ITA	0.1481	0.0489	0.0000	0.1089	0.0000	0	0	1	0	1	1	0	1	1	1	1	1	1	0	4
MAR	0.2160	0.0265	0.0000	0.0605	0.0000	0.0000	0	4	0	4	3	0	3	3	3	3	3	3	0	5
ARG	0.2188	0.0704	0.0594	0.1466	0.1207	0.0648	0.0000	4	0	4	3	0	3	3	3	3	3	3	0	5

	Burst-type populations						N	M	N	N	N	N	N	N	N	N	N	B	N
	CMR	CMT	IPA	PAL2S	ITA	MAR	ARG	RBON	ARA	BGUA	RR	PAR	TRU	PG	DUR	IDT	BRA	SJB	CRU	PSEUDO
RBON	0.9763	0.9523	0.9749	0.8866	0.9312	1.0000	0.8533	0.0000	4	0	0	4	0	0	0	0	0	0	4	7
ARA	0.1677	0.0374	0.0219	0.1242	0.0842	0.0164	0.0000	0.8376	0.0000	4	3	0	3	3	3	3	3	3	0	5
BGUA	0.9461	0.8817	0.9172	0.8304	0.8822	0.9529	0.7884	0.3924	0.7701	0.0000	0	4	0	0	0	0	0	0	4	7
RR	0.8607	0.7624	0.7885	0.7707	0.8009	0.8012	0.7350	0.5279	0.7217	0.4088	0.0000	3	0	0	0	0	0	0	3	7
PAR	0.4042	0.1215	0.1464	0.1887	0.2248	0.1900	0.0000	0.9134	0.0000	0.8221	0.7170	0.0000	3	3	3	3	3	3	0	5
TRU	0.8698	0.7690	0.7956	0.7746	0.8068	0.8126	0.7359	0.5519	0.7208	0.4246	0.1594	0.7231	0.0000	0	0	0	0	0	3	7
PG	0.8981	0.8283	0.8528	0.8096	0.8435	0.8699	0.7750	0.3479	0.7615	0.2683	0.1878	0.7869	0.1520	0.0000	0	0	0	0	3	7
DUR	0.9124	0.8291	0.8610	0.8063	0.8496	0.8850	0.7666	0.5825	0.7500	0.4352	0.1453	0.7762	0.1177	0.0000	0.0000	0	0	0	3	7
IDT	0.8874	0.8140	0.8369	0.7990	0.8302	0.8539	0.7631	0.6651	0.7504	0.5078	0.3199	0.7724	0.3806	0.4020	0.4259	0.0000	0	0	3	7
BRA	0.8609	0.7801	0.8017	0.7780	0.8032	0.8155	0.7428	0.4978	0.7313	0.3383	0.2559	0.7407	0.3006	0.2744	0.3143	0.0175	0.0000	0	3	6
SJB	0.8586	0.7766	0.7983	0.7752	0.8001	0.8122	0.7395	0.4946	0.7279	0.3351	0.2480	0.7372	0.2867	0.2660	0.3040	0.0127	0.0000	0.0000	3	7
CRU	0.1552	0.0452	0.0234	0.1311	0.0800	0.0027	0.0455	0.8449	0.0000	0.7793	0.7338	0.0673	0.7323	0.7720	0.7609	0.7597	0.7415	0.7382	0.0000	5
PSEUDO	0.9505	0.9086	0.9338	0.8532	0.8982	0.9638	0.8144	0.9842	0.7959	0.9560	0.8811	0.8681	0.8896	0.9120	0.9269	0.9062	0.8780	0.8785	0.8002	0.0000

Upper right matrix – fixed differences between samples; Lower left matrix – pairwise differentiation (Fst). M: Mixed song; N: no recorded song; B: Burst-type song.

Table 3: AMOVA statistics

Source of variation	Percentage of variation
Among groups	80.53
Among populations within groups	1.51
Within populations	17.96
Fsc (haplotypes/populations within groups)	0.07772***
Fst (haplotypes/populations/groups)	0.82045***
Fct (populations/groups)	0.80532***

Groups considered: Group 1: Rio Bonito, Barra de Guaratiba; **Group 2:** Lapinha, Jacobina, Sobral 1S, Jaíba 1S, Teresina, Jequié, Lassance, Palmas 1S, Pirenópolis, Estrela 1S; **Group 3:** Sobral 2S, Pancas, Natal, Jaíba 2S, Estrela 2S, Barcarena, Passira, Campo Grande, Nova Porteirinha, Afonso Cláudio, Camará, Cametá, Ipanema, Palmas 2S, Itamaracá, Marajó; **Group 4:** Isla del Tigre, Brasilito, San Juan Bautista; **Group 5:** Roraima, Trujillo, Palo Gordo, Durania. Significance corresponding to the fixation indexes were obtained through 10.100 permutations *** $p<0.0001$.

Fig 1

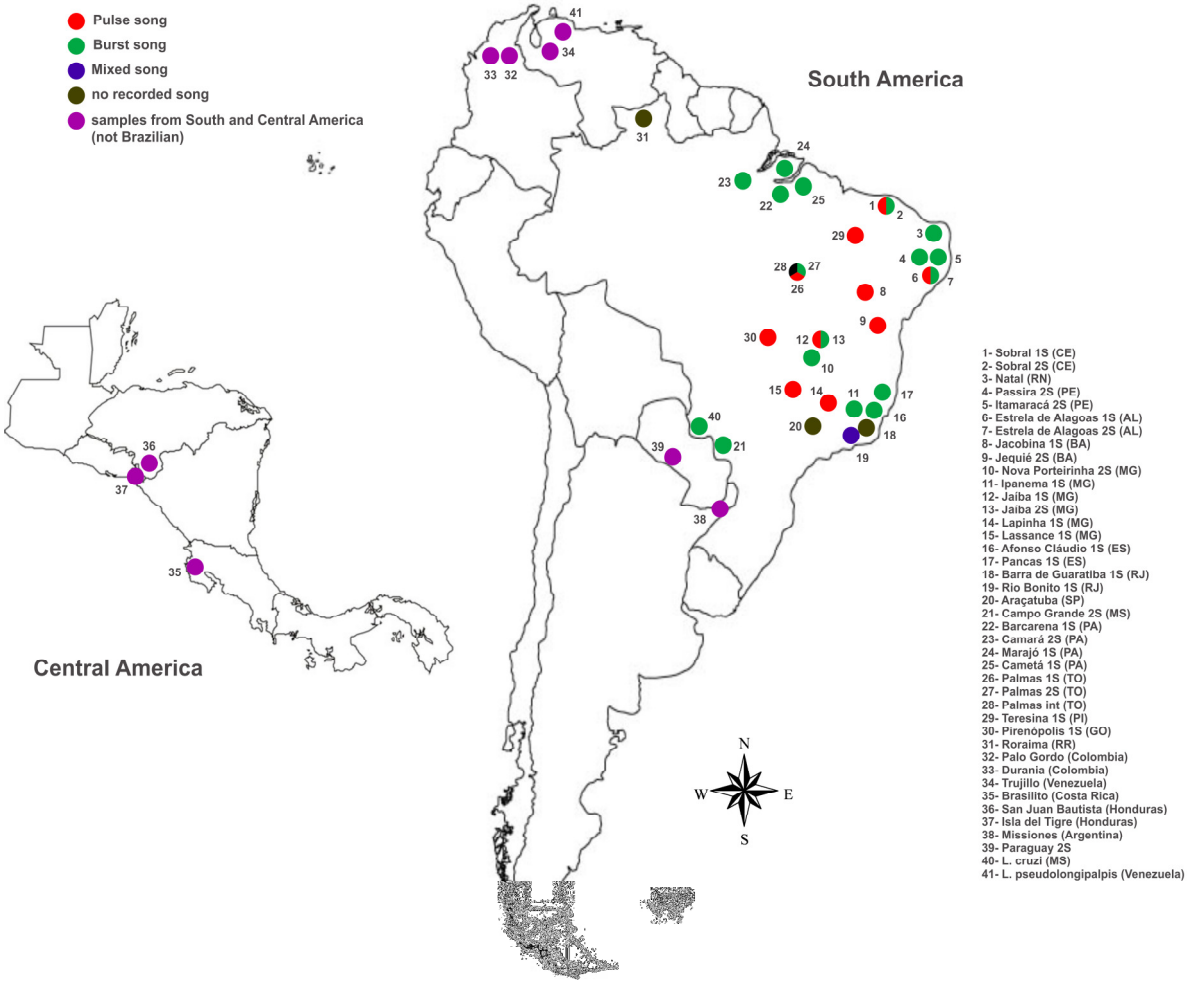

Pulse song
Burst song
Mixed song
no recorded song
samples from South and Central America (not Brazilian)
South America
Central America
N
W
E
S
1- Sobral 1S (CE)
2- Sobral 2S (CE)
3- Natal (RN)
4- Passira 2S (PE)
5- Itamaracá 2S (PE)
6- Estrela de Alagoas 1S (AL)
7- Estrela de Alagoas 2S (AL)
8- Jacobina 1S (BA)
9- Jequié 2S (BA)
10- Nova Porteirinha 2S (MG)
11- Ipanema 1S (MG)
12- Jaíba 1S (MG)
13- Jaíba 2S (MG)
14- Lapinha 1S (MG)
15- Lassance 1S (MG)
16- Afonso Cláudio 1S (ES)
17- Pancas 1S (ES)
18- Barra de Guaratiba 1S (RJ)
19- Rio Bonito 1S (RJ)
20- Araçatuba (SP)
21- Campo Grande 2S (MS)
22- Barcarena 1S (PA)
23- Camará 2S (PA)
24- Marajó 1S (PA)
25- Cametá 1S (PA)
26- Palmas 1S (TO)
27- Palmas 2S (TO)
28- Palmas int (TO)
29- Teresina 1S (PI)
30- Pirenópolis 1S (GO)
31- Roraima (RR)
32- Palo Gordo (Colombia)
33- Durania (Colombia)
34- Trujillo (Venezuela)
35- Brasilito (Costa Rica)
36- San Juan Bautista (Honduras)
37- Isla del Tigre (Honduras)
38- Misiones (Argentina)
39- Paraguay 2S
40- L. cruzi (MS)
41- L. pseudolongipalpis (Venezuela)

Fig 2

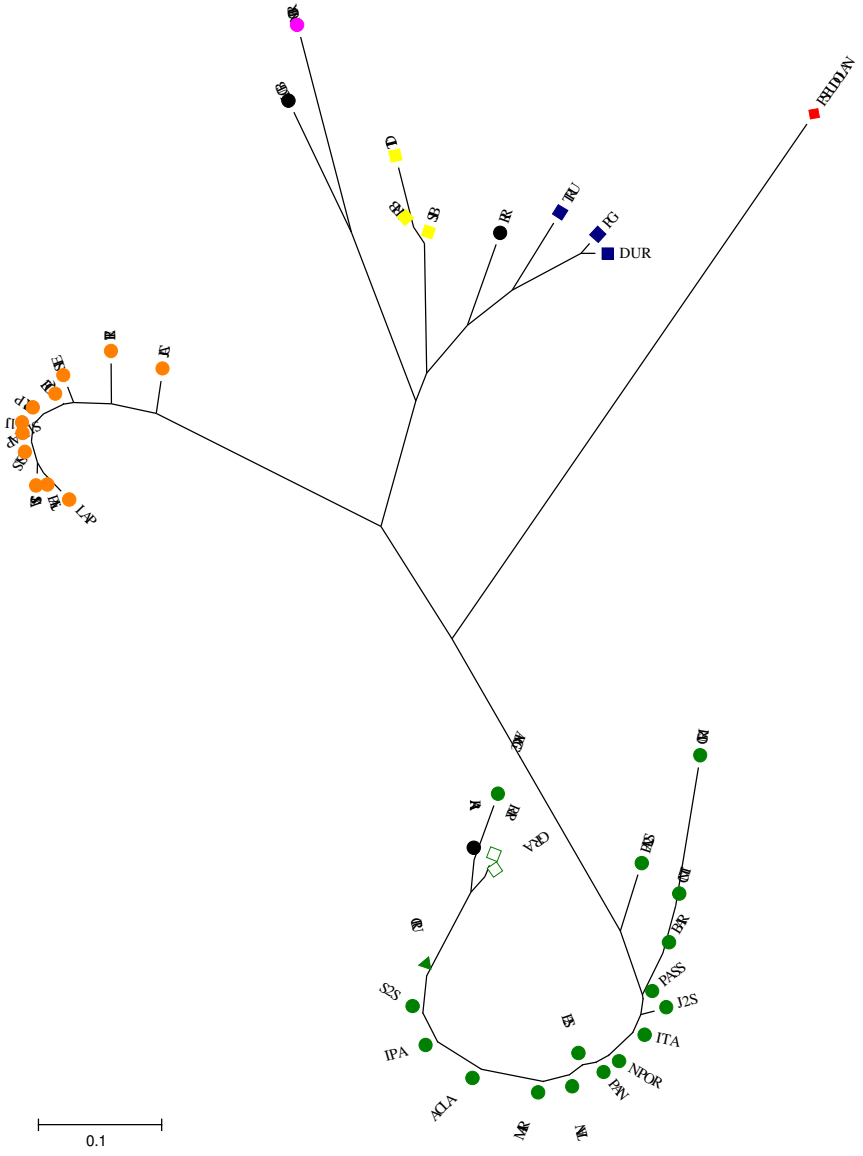
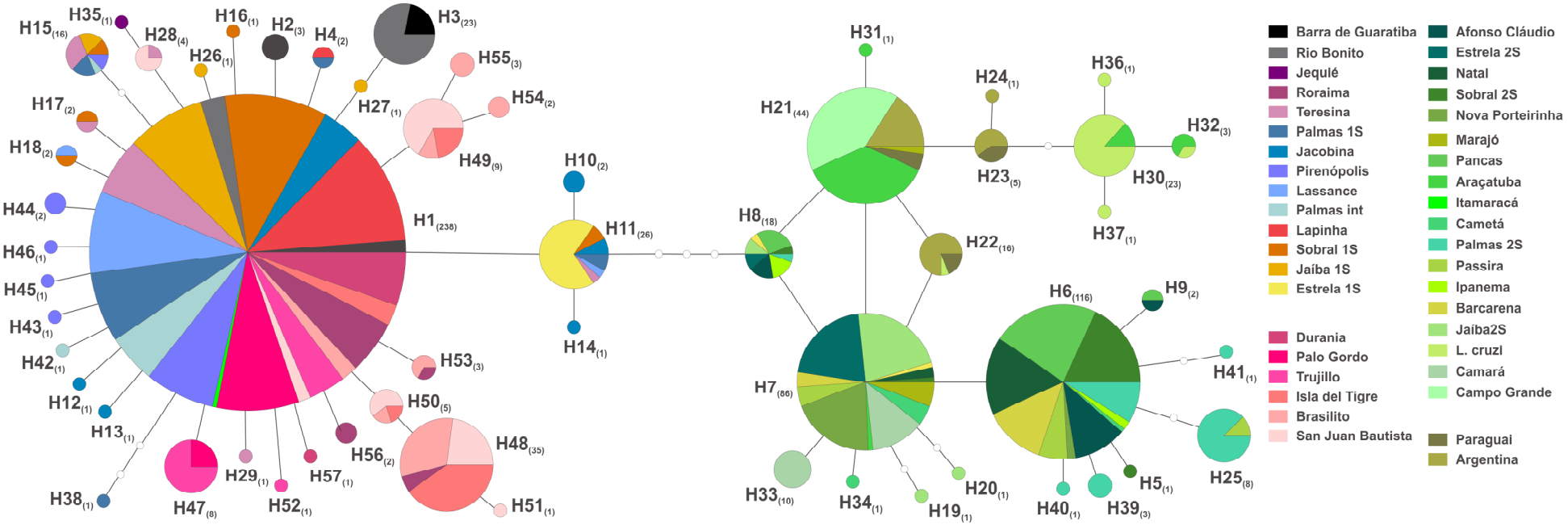
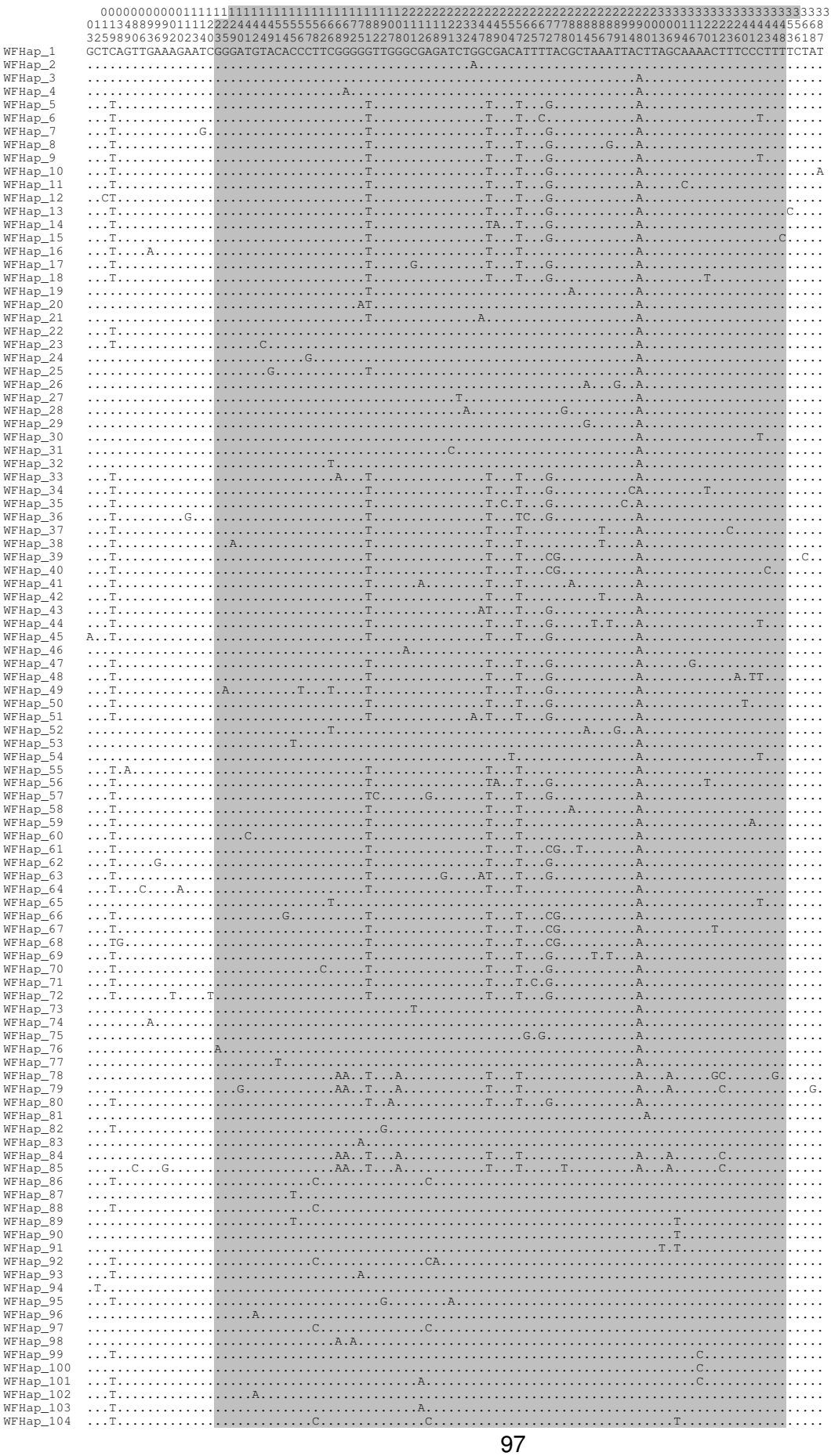
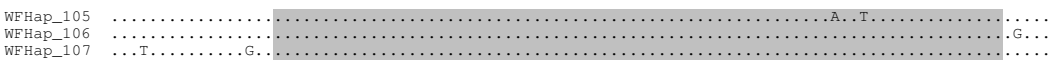


Fig 3


H35(1)
H15(16)
H28(4)
H16(1)
H2(3)
H4(2)
H3(23)
H26(1)
H55(3)
H27(1)
H54(2)
H17(2)
H49(9)
H18(2)
H10(2)
H44(2)
H1(238)
H11(26)
H46(1)
H14(1)
H45(1)
H43(1)
H53(3)
H42(1)
H50(5)
H12(1)
H48(35)
H13(1)
H56(2)
H29(1)
H57(1)
H38(1)
H47(8)
H52(1)
H51(1)
H31(1)
H24(1)
H36(1)
H21(44)
H32(3)
H23(5)
H30(23)
H8(18)
H37(1)
H22(16)
H6(116)
H9(2)
H41(1)
H7(86)
H25(8)
H5(1)
H33(10)
H34(1)
H20(1)
H40(1)
H39(3)
H19(1)
Barra de Guaratiba
Rio Bonito
Jequié
Roraima
Teresina
Palmas 1S
Jacobina
Pirenópolis
Lassance
Palmas int
Lapinha
Sobral 1S
Jaíba 1S
Estrela 1S
Durania
Palo Gordo
Trujillo
Isla del Tigre
Brasilito
San Juan Bautista
Afonso Cláudio
Estrela 2S
Natal
Sobral 2S
Nova Porteirinha
Marajó
Pancas
Araçatuba
Itamaracá
Cametá
Palmas 2S
Passira
Ipanema
Barcarena
Jaiba2S
L. cruzi
Camará
Campo Grande
Paraguai
Argentina

S1: Alignment of polymorphic sites of whole fragment haplotypes (WFHap)




WFHap_105 .....A..T.....
WFHap_106 .....G...
WFHap_107 ...T.....G.....

S2: Number of sequences represented in each sample (WFHAp)

	Pulse-type populations										Burst-type populations										
Haplotype	LAP	JAC	S1S	J1S	TER	JEQ	LASS	PAL1S	PIR	E1S	S2S	PAN	NAT	J2S	E2S	BAR	PASS	CGRA	NPOR	ACLA	CMR
WFHap_1				1																	
WFHap_2																					
WFHap_3	27	7	22	18	15	1	20	18	15	18											
WFHap_4	1							1													
WFHap_5										2	21	22	18	15	18	17	9	17	14	16	12
WFHap_6											2		1								
WFHap_7											1										
WFHap_8											1		1	1	2	1					1
WFHap_9											1	2	1					1		4	
WFHap_10											2										
WFHap_11												4								1	
WFHap_12												1									
WFHap_13												1									
WFHap_14												1									
WFHap_15												1									
WFHap_16													1								
WFHap_17													1								
WFHap_18													1	3					2		
WFHap_19		3																			
WFHap_20		3																			
WFHap_21		1																			
WFHap_22		5																			
WFHap_23		1																			
WFHap_24		1																			
WFHap_25		1																			
WFHap_26			2	3	5			3	2												
WFHap_27			1																		
WFHap_28			1		1																
WFHap_29			2																		
WFHap_30			2	1			1	2	3												
WFHap_31			1																		
WFHap_32			1																		
WFHap_33														1					1		
WFHap_34														2							
WFHap_35														1							
WFHap_36														1							
WFHap_37																					
WFHap_38																					
WFHap_39																					
WFHap_40																					
WFHap_41																					
WFHap_42																					

WFHap_43															4						11
	Pulse-type populations										Burst-type populations										
Haplotype	LAP	JAC	S1S	J1S	TER	JEQ	LASS	PAL1S	PIR	E1S	S2S	PAN	NAT	J2S	E2S	BAR	PASS	CGRA	NPOR	ACLA	CMR
WFHap_44																	1				
WFHap_45																	2				
WFHap_46				1																	
WFHap_47																			1		
WFHap_48																			1		
WFHap_49																			1		
WFHap_50																			1		
WFHap_51																			1		
WFHap_52					1																
WFHap_53					1	1															
WFHap_54					1																
WFHap_55																				1	
WFHap_56																				1	
WFHap_57																				1	
WFHap_58																					
WFHap_59																					
WFHap_60																					
WFHap_61																					
WFHap_62																					
WFHap_63																					
WFHap_64																					
WFHap_65							1														
WFHap_66																					
WFHap_67																					
WFHap_68																					
WFHap_69																					
WFHap_70																					
WFHap_71																					
WFHap_72																					
WFHap_73																					
WFHap_74									1												
WFHap_75									1												
WFHap_76									1												
WFHap_77									1												
WFHap_78																					
WFHap_79																					
WFHap_80																					
WFHap_81																					
WFHap_82																					
WFHap_83																					
WFHap_84																					
WFHap_85																					

WFHap_86																					
WFHap_87																					
	Pulse-type populations										Burst-type populations										
Haplotype	LAP	JAC	S1S	J1S	TER	JEQ	LASS	PAL1S	PIR	E1S	S2S	PAN	NAT	J2S	E2S	BAR	PASS	CGRA	NPOR	ACLA	CMR
WFHap_88																					
WFHap_89																					
WFHap_90																					
WFHap_91																					
WFHap_92																					
WFHap_93																					
WFHap_94																					
WFHap_95																					
WFHap_96																					
WFHap_97																					
WFHap_98																					
WFHap_99																					
WFHap_100																					
WFHap_101																					
WFHap_102																					
WFHap_103																					
WFHap_104																					
WFHap_105																					
WFHap_106																					
WFHap_107																					

	Burst-type populations						N	M	N	N	N	N	N	N	N	N	N	N	B	N	
Haplotype	CMT	IPA	PAL2S	ITA	MAR	ARG	RBON	PALi	ARA	BGUA	RR	PAR	TRU	PG	DUR	IDT	BRA	SJB	CRU	PSEUDO	Total
WFHap_1							24			8	2		3	11	5	5	7	7			73
WFHap_2										4											4
WFHap_3				1				12													174
WFHap_4																					2
WFHap_5	3	7	9	18	6	17			18			5							20		264
WFHap_6																					3
WFHap_7																					1
WFHap_8	2																				9
WFHap_9			1																		10
WFHap_10																					2
WFHap_11																					5
WFHap_12																					1
WFHap_13																					1
WFHap_14																					1
WFHap_15																					1
WFHap_16																					1

Haplotype	CMT	IPA	PAL2S	ITA	MAR	ARG	RBON	PALi	ARA	BGUA	RR	PAR	TRU	PG	DUR	IDT	BRA	SJB	CRU	PSEUDO	Total
	Burst-type populations					N	M	N	N	N	N	N	N	N	N	N	N	N	B	N	
WFHap_17																					1
WFHap_18			1	2																	9
WFHap_19																					3
WFHap_20																					3
WFHap_21																					1
WFHap_22																					5
WFHap_23																					1
WFHap_24																					1
WFHap_25																					1
WFHap_26								1													16
WFHap_27																					1
WFHap_28																					2
WFHap_29																					2
WFHap_30																					9
WFHap_31																					1
WFHap_32																					1
WFHap_33																					2
WFHap_34																					2
WFHap_35																					1
WFHap_36																					1
WFHap_37						1															1
WFHap_38						1															1
WFHap_39						1															1
WFHap_40						1															1
WFHap_41						1															1
WFHap_42						2						2									4
WFHap_43	1																				16
WFHap_44			6																		7
WFHap_45																					2
WFHap_46																					1
WFHap_47																					1
WFHap_48																					1
WFHap_49																					1
WFHap_50			1																		2
WFHap_51																					1
WFHap_52																					1
WFHap_53																					2
WFHap_54																					1
WFHap_55																					1
WFHap_56																					1
WFHap_57																					1
WFHap_58									1			1									2

WFHap_59									2											2
WFHap_60									1											1
WFHap_61									2									1		2
WFHap_62	1																			1
WFHap_63	1																			1
WFHap_64		1																		1
WFHap_65																				1

	Burst-type populations					**N**	**M**	**N**	**N**	**N**	**N**	**N**	**N**	**N**	**N**	**N**	**N**	**B**	**N**	
Haplotype	**CMT**	**IPA**	**PAL2S**	**ITA**	**MAR**	**ARG**	**RBON**	**PALi**	**ARA**	**BGUA**	**RR**	**PAR**	**TRU**	**PG**	**DUR**	**IDT**	**BRA**	**SJB**	**CRU**	**PSEUDO**	**Total**
WFHap_66																			1		0
WFHap_67																			1		0
WFHap_68																			1		0
WFHap_69			1																		1
WFHap_70			3																		3
WFHap_71			1																		1
WFHap_72			1																		1
WFHap_73								1													1
WFHap_74																					1
WFHap_75																					1
WFHap_76																					1
WFHap_77																					1
WFHap_78																					0
WFHap_79																					0
WFHap_80				1																	1
WFHap_81														1							1
WFHap_82											4		5	8	7						24
WFHap_83														2							2
WFHap_84																				18	0
WFHap_85																				2	0
WFHap_86											2					13	10	8			33
WFHap_87																		1			1
WFHap_88																1	1	3			5
WFHap_89																		2			2
WFHap_90																1		1			2
WFHap_91																		1			1
WFHap_92																		1			1
WFHap_93													6								6
WFHap_94													1								1
WFHap_95													1								1
WFHap_96																	2				2
WFHap_97																	1				1
WFHap_98																	3				3
WFHap_99											5										5
WFHap_100											2										2

WFHap_101	1				1
WFHap_102	1				1
WFHap_103	1				1
WFHap_104			1		1
WFHap_105			1		1
WFHap_106		1			1
WFHap_107		1		1	

M: Mixed song; N: no recorded song; B: Burst-type son

S3	RBON	BGUA	LAP	JAC	S1S	J1S	TER	JEQ	LASS	PALL1S	PALI
RBON	x										
BGUA	0.00879+-0.0025	X									
LAP	0.00000+-0.0000	0.00000+-0.0000	x								
JAC	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	x							
S1S	0.00000+-0.0000	0.00000+-0.0000	0.02832+-0.0051	0.00000+-0.0000	x						
J1S	0.00000+-0.0000	0.00000+-0.0000	0.03516+-0.0061	0.00000+-0.0000	0.83691+-0.0094	x					
TER	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.21680+-0.0141	0.40234+-0.0145	x				
JEQ	0.00098+-0.0010	0.00977+-0.0029	0.14941+-0.0083	0.51660+-0.0136	0.24707+-0.0099	0.25879+-0.0097	0.29199+-0.0140	x			
LASS	0.00000+-0.0000	0.00000+-0.0000	0.45020+-0.0168	0.00000+-0.0000	0.27344+-0.0139	0.21484+-0.0134	0.01367+-0.0031	0.16211+-0.0111	x		
PALL1S	0.00000+-0.0000	0.00000+-0.0000	0.08496+-0.0095	0.00000+-0.0000	0.88965+-0.0112	0.91797+-0.0084	0.34766+-0.0139	0.18848+-0.0135	0.28418+-0.0176	x	
PALI	0.00000+-0.0000	0.00000+-0.0000	0.27148+-0.0120	0.00098+-0.0010	0.88672+-0.0079	0.87598+-0.0106	0.24512+-0.0147	0.33496+-0.0133	0.58301+-0.0142	0.84766+-0.0129	x
PIR	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0022	0.00000+-0.0000	0.43262+-0.0137	0.30371+-0.0152	0.13086+-0.0063	0.43066+-0.0148	0.13965+-0.0135	0.61719+-0.0161	0.38281+-0.0157
E1S	0.00000+-0.0000	0.00000+-0.0000	0.16602+-0.0120	0.00000+-0.0000	0.00391+-0.0019	0.00586+-0.0022	0.00195+-0.0014	0.25977+-0.0149	0.10742+-0.0102	0.01465+-0.0039	0.21191+-0.0126
S2S	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00391+-0.0019	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
PAN	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
NAT	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00195+-0.0014	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
J2S	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
E2S	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
ARG	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00195+-0.0014	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
BAR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0026	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
PASS	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00684+-0.0027	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
CGRA	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00684+-0.0023	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
NPOR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00391+-0.0019	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
ACLA	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
ARA	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00977+-0.0036	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
CMR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
CMT	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.01562+-0.0039	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
IPA	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.02148+-0.0047	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
CRU	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.01758+-0.0039	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
PAL2S	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00098+-0.0010	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
PAR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.01953+-0.0041	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
ITA	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0026	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
MAR	0.00000+-0.0000	0.00195+-0.0014	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.02539+-0.0061	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
IDT	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
BRA	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.01074+-0.0033	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
SJB	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.01953+-0.0041	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
RR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00977+-0.0029	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000

TRU	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0022	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
PG	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0022	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
DUR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0022	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
PSEUDO	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00391+-0.0019	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000

	PIR	**E1S**	**S2S**	**PAN**	**NAT**	**J2S**	**E2S**	**ARG**	**BAR**	**PASS**	**CGRA**
RBON											
BGUA											
LAP											
JAC											
S1S											
J1S											
TER											
JEQ											
LASS											
PALL1S											
PALI											
PIR	x										
E1S	0.00195+-0.0014	X									
S2S	0.00000+-0.0000	0.00000+-0.0000	x								
PAN	0.00000+-0.0000	0.00000+-0.0000	0.08887+-0.0077	x							
NAT	0.00000+-0.0000	0.00000+-0.0000	0.89844+-0.0094	0.30176+-0.0133	x						
J2S	0.00000+-0.0000	0.00000+-0.0000	0.00684+-0.0023	0.01758+-0.0046	0.05859+-0.0066	x					
E2S	0.00000+-0.0000	0.00000+-0.0000	0.43555+-0.0165	0.32812+-0.0150	0.66113+-0.0122	0.17480+-0.0098	x				
ARG	0.00000+-0.0000	0.00000+-0.0000	0.02148+-0.0040	0.00098+-0.0010	0.02051+-0.0041	0.00977+-0.0033	0.01367+-0.0037	x			
BAR	0.00000+-0.0000	0.00000+-0.0000	0.04004+-0.0071	0.01367+-0.0046	0.03125+-0.0065	0.00098+-0.0010	0.04004+-0.0047	0.01074+-0.0033	x		
PASS	0.00000+-0.0000	0.00000+-0.0000	0.09375+-0.0085	0.14844+-0.0091	0.27441+-0.0122	0.41895+-0.0151	0.32324+-0.0119	0.06738+-0.0075	0.01270+-0.0034	x	
CGRA	0.00000+-0.0000	0.00000+-0.0000	0.43457+-0.0140	0.28809+-0.0116	0.65723+-0.0127	0.11035+-0.0114	0.99902+-0.0002	0.02734+-0.0055	0.09668+-0.0111	0.29883+-0.0159	x
NPOR	0.00000+-0.0000	0.00000+-0.0000	0.01660+-0.0042	0.04004+-0.0055	0.11816+-0.0093	0.86914+-0.0138	0.39844+-0.0174	0.00977+-0.0033	0.00098+-0.0010	0.76758+-0.0137	0.19922+-0.0127
ACLA	0.00000+-0.0000	0.00000+-0.0000	0.17871+-0.0102	0.22656+-0.0130	0.33496+-0.0191	0.10352+-0.0100	0.37695+-0.0164	0.01465+-0.0039	0.01758+-0.0039	0.33496+-0.0147	0.16211+-0.0128
ARA	0.00000+-0.0000	0.00000+-0.0000	0.02441+-0.0045	0.00391+-0.0023	0.03516+-0.0053	0.00488+-0.0020	0.02930+-0.0055	0.55566+-0.0148	0.01367+-0.0034	0.07715+-0.0093	0.06055+-0.0074
CMR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.07227+-0.0071	0.00000+-0.0000	0.00000+-0.0000
CMT	0.00000+-0.0000	0.00000+-0.0000	0.13867+-0.0107	0.11914+-0.0101	0.26367+-0.0130	0.16113+-0.0106	0.24414+-0.0158	0.13086+-0.0123	0.07812+-0.0083	0.25781+-0.0099	0.11328+-0.0090
IPA	0.00000+-0.0000	0.00000+-0.0000	0.61426+-0.0158	0.24512+-0.0122	0.69727+-0.0124	0.14844+-0.0093	0.33301+-0.0169	0.20215+-0.0153	0.17676+-0.0136	0.26855+-0.0124	0.51562+-0.0149
CRU	0.00000+-0.0000	0.00000+-0.0000	0.04004+-0.0068	0.02441+-0.0055	0.10254+-0.0080	0.00684+-0.0027	0.10449+-0.0102	0.08398+-0.0093	0.06445+-0.0066	0.10547+-0.0092	0.12598+-0.0113
PAL2S	0.00000+-0.0000	0.00000+-0.0000	0.00195+-0.0014	0.00098+-0.0010	0.00684+-0.0027	0.00195+-0.0014	0.02051+-0.0045	0.00098+-0.0010	0.00000+-0.0000	0.20215+-0.0137	0.01465+-0.0034
PAR	0.00000+-0.0000	0.00098+-0.0010	0.01367+-0.0034	0.00586+-0.0026	0.02148+-0.0045	0.01074+-0.0036	0.03320+-0.0055	0.78711+-0.0140	0.00293+-0.0016	0.06250+-0.0079	0.02344+-0.0044
ITA	0.00000+-0.0000	0.00000+-0.0000	0.04883+-0.0071	0.03613+-0.0068	0.29297+-0.0167	0.07031+-0.0080	0.31836+-0.0134	0.00977+-0.0029	0.02246+-0.0042	0.17285+-0.0120	0.51855+-0.0138
MAR	0.00000+-0.0000	0.00000+-0.0000	0.78418+-0.0141	0.67090+-0.0135	0.99902+-0.0002	0.43457+-0.0165	0.99902+-0.0002	0.27637+-0.0107	0.64355+-0.0185	0.69434+-0.0141	0.99902+-0.0002

IDT	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
BRA	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
SJB	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
RR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
TRU	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
PG	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
DUR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000
PSEUDO	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000

	NPOR	**ACLA**	**ARA**	**CMR**	**CMT**	**IPA**	**CRU**	**PAL2S**	**PAR**	**ITA**	**MAR**
RBON											
BGUA											
LAP											
JAC											
S1S											
J1S											
TER											
JEQ											
LASS											
PALL1S											
PALI											
PIR											
E1S											
S2S											
PAN											
NAT											
J2S											
E2S											
ARG											
BAR											
PASS											
CGRA											
NPOR	x										
ACLA	0.23438+-0.0128	X									
ARA	0.01367+-0.0031	0.03516+-0.0058	x								
CMR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	x							
CMT	0.24707+-0.0118	0.45703+-0.0154	0.19727+-0.0132	0.06836+-0.0070	x						
IPA	0.21387+-0.0104	0.40723+-0.0171	0.28711+-0.0074	0.01172+-0.0033	0.30176+-0.0116	x					
CRU	0.05078+-0.0066	0.04102+-0.0070	0.42090+-0.0149	0.00000+-0.0000	0.08008+-0.0076	0.55371+-0.0185	x				

PAL2S	0.00977+-0.0033	0.02344+-0.0052	0.00000+-0.0000	0.00000+-0.0000	0.10547+-0.0095	0.04785+-0.0056	0.00000+-0.0000	x				
PAR	0.01367+-0.0034	0.03418+-0.0061	0.46680+-0.0178	0.00195+-0.0014	0.16406+-0.0092	0.20703+-0.0130	0.18945+-0.0132	0.00586+-0.0022	x			
ITA	0.08008+-0.0076	0.07227+-0.0080	0.01562+-0.0042	0.00000+-0.0000	0.09180+-0.0076	0.42480+-0.0175	0.08594+-0.0086	0.00195+-0.0014	0.00977+-0.0029	x		
MAR	0.54590+-0.0155	0.53418+-0.0143	0.35059+-0.0142	0.05859+-0.0079	0.42871+-0.0168	0.99902+-0.0002	0.54980+-0.0157	0.19727+-0.0102	0.23145+-0.0148	0.99902+-0.0002	x	
IDT	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	
BRA	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	
SJB	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	
RR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	
TRU	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	
PG	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	
DUR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	
PSEUDO	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	

	IDT	BRA	SJB	RR	TRU	PG	DUR	PSEUDO
RBON								
BGUA								
LAP								
JAC								
S1S								
J1S								
TER								
JEQ								
LASS								
PALL1S								
PALI								
PIR								
E1S								
S2S								
PAN								
NAT								
J2S								
E2S								
ARG								
BAR								
PASS								
CGRA								
NPOR								
ACLA								
ARA								
CMR								
CMT								
IPA								
CRU								
PAL2S								
PAR								
ITA								
MAR								
IDT	x							
BRA	0.21777+-0.0173	x						
SJB	0.25293+-0.0151	0.56250+-0.0179	x					
RR	0.00098+-0.0010	0.00000+-0.0000	0.00000+-0.0000	x				
TRU	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00098+-0.0010	x			
PG	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0026	x		
DUR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00977+-0.0036	0.03223+-0.0063	0.36621+-0.0109	x	
PSEUDO	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	x

	IDT	BRA	SJB	RR	TRU	PG	DUR	PSEUDO
RBON								
BGUA								
LAP								
JAC								
S1S								
J1S								
TER								
JEQ								
LASS								
PALL1S								
PALI								
PIR								
E1S								
S2S								
PAN								
NAT								
J2S								
E2S								
ARG								
BAR								
PASS								
CGRA								
NPOR								
ACLA								
ARA								
CMR								
CMT								
IPA								
CRU								
PAL2S								
PAR								
ITA								
MAR								
IDT	x							
BRA	0.21777+-0.0173	x						
SJB	0.25293+-0.0151	0.56250+-0.0179	x					
RR	0.00098+-0.0010	0.00000+-0.0000	0.00000+-0.0000	x				
TRU	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00098+-0.0010	x			
PG	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0026	x		
DUR	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00977+-0.0036	0.03223+-0.0063	0.36621+-0.0109	x	
PSEUDO	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	x

CAPÍTULO 6: "Análise multilocus de espécies simpátricas do complexo *Lutzomyia longipalpis* usando genes que controlam o som de corte"

Lins RMMA, Peixoto AA; **Multilocus analysis of sympatric sibling species of the *Lutzomyia longipalpis* complex using lovesong genes.**

Este manuscrito corresponde ao objetivo específico 2.2.3.

No presente artigo seis genes associados ao controle do som são utilizados em uma análise multilocus com populações de *Lu. longipalpis* das regiões de Sobral (CE), Jaíba (MG) e Estrela (AL), onde trabalhos anteriores indicam a existência espécies do complexo Longipalpis vivendo em simpatria. Os resultados gerados confirmam a existência de espécies simpátricas nessas regiões e demonstram a existência de um alto grau de diferenciação genética entre elas com a ocorrência de introgressão em alguns loci. Essa introgressão também se demonstrou ser diferencial e assimétrica em algumas comparações. Além disso, uma comparação com dados gerados por um ensaio multilocus anterior, sugere que genes associados ao controle do som apresentam, em média, maior valor de divergência genética e menor introgressão que outros marcadores.

Short Communication

Multilocus analysis of sympatric sibling species of the *Lutzomyia longipalpis* complex in Brazil using lovesong genes

Rachel MMA Lins[1], Alexandre A Peixoto[1,2§]

[1] Laboratório de Biologia Molecular de Insetos, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil

[2] Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular (INCT-EM), Rio de Janeiro, Brazil

§Corresponding author

Email addresses:

RMMAL: rlins@ioc.fiocruz.br

AAP: apeixoto@ioc.fiocruz.br

Abstract

Lutzomyia longipalpis is the main vector of American Visceral Leishmaniais (AVL) and represents a species complex in Brazil. Combined molecular and behavioral (copulatory courtship songs and pheromones) analyses have provided the best evidence supporting the existence of sympatric siblings. Molecular analysis of *period* and *paralytic* genes demonstrated high levels of genetic differentiation between the siblings from Sobral, Jaíba and Estrela, where males present different patterns of abdominal spots, copulatory courtship songs and pheromones. Recent multilocus data suggested that genes involved in the production of courtship songs might exhibit higher levels of divergence and lower introgression between the Sobral siblings when compared to other markers. Here we performed a multilocus analysis using orthologues of the *Drosophila* genes *period*, *paralytic*, *cacophony*, *ebony*, *fruitless* and *slowpoke*, which are associated with courtship song production to evaluate the divergence and gene flow between *Lu. longipalpis* sympatric siblings from Sobral, Jaíba and Estrela. High levels of genetic divergence were detected with evidence of asymmetric introgression in most loci. In addition, our data confirmed that genes involved in song production tend to show higher levels of genetic divergence when compared to other markers.

Background

The population genetics and taxonomic status of the American visceral leishmaniasis vector *Lutzomyia longipalpis* (Lainson & Rangel 2005) has been widely studied since early investigations suggested that it might comprise a species complex (Mangabeira 1969, Ward et al 1983, 1988; Lanzaro et al 1993). Currently there is a consensus that this vector represents a species complex (Ward et al 1988; Lanzaro et al 1993; Bauzer 2002a,b; Maingon et al 2003;

Arrivillaga et al 2003; Watts et al 2005; Bauzer et al 2007; Maingon et al 2008; Lins et al 2008, 2012; Araki et al 2009) with variable levels of divergence and gene flow among populations. Combined molecular and behavioral analyses have provided the best evidence supporting the existence of different sibling species occurring within Brazil (Hamilton et al 1999a, 1999b, 2004, 2005; Bauzer et al 2002a, 2002b; de Souza et al 2002, 2004, 2008; Maingon et al 2003; Bottecchia et al 2004; Watts et al 2005; Araki et al 2009, Lins et al 2012). Two recent studies (Araki et al, 2009; Lins et al 2012) demonstrated the existence of two distinctive groups clearly separated in Brazil, with the males producing different copulatory courtship songs and pheromones. Males of one group of populations, which likely represents a single species, produce Burst-type song and the Cembrene-1 pheromone, while in the other, more complex, males produce five variations of the Pulse-type song (P1 to P5) in combination with four different pheromone types (Germacrene, Hymachalene, and Cembrenes 1 and 2), probably representing five incipient species.

Recently, a multilocus analysis (Araki et al, in preparation) was performed to estimate and compare the levels of gene flow and genetic divergence between Brazilian populations in 21 nuclear loci comparing two sympatric species from Sobral, Ceará State, and two allopatric siblings from Lapinha (Minas Gerais State) and Pancas (Espírito Santo State). The results suggest the occurrence of gene flow between the species studied with different levels of introgression particularly in the sympatric pair depending on the locus analyzed, also indicating that this introgression seems to be asymmetric and has been an important factor in the evolution of this complex.

In addition, the data also suggest that three genes associated with song production used in the study, *paralytic (para), cacophony (cac)* and *period (per),* which encode, respectively, a voltage-gated sodium channel, a voltage-gated calcium channel, and a transcriptional repressor primarily involved in the circadian clock, have on average higher genetic divergence

and lower level of introgression between the Sobral sympatric sibling species than the other loci. Indeed, work carried out with *per* and *para* genes in a number of Brazilian *Lu. longipalpis* populations revealed very high levels of genetic differentiation between siblings producing different copulation songs (Bauzer et al 2002a,b; Araki et al 2009; Lins et al 2008, 2012).

The present study aims to estimate the levels of divergence and gene flow between sympatric siblings of *L. longipalpis* from the Brazilian localities of Sobral, Jaíba (Minas Gerais State) and Estrela de Alagoas (Alagoas State) applying a multilocus approach with the isolation with migration model using six nuclear loci associated with courtship song production: *para, cac and per*, already mentioned above, and the genes *ebony* (*e*), *slowpoke* (*slo*) and *fruitless* (*fru*), which encode respectively a beta-alanyl-dopamine-synthase, a calcium activated potassium channel and a transcription factor mainly involved in sex-determination.

Methods

The *per*, *para, cac* primers were the same used in previously published studies (Bauzer et al 2002a, 2002b; Bottecchia et al 2004; Lins et al 2008; 2012). Primers used to amplified fragments orthologs to three other *Drosophila* genes associated with courtship song production (*e*, *slo* and *fru*) were designed based on search for EST corresponding to these genes the VectorBase databank (http://www.vectorbase.org/). These primers sequences are the following: *e*, 5lle1- GGA GAA TCA ACG ATG GAA GC, 3lle1- AGG ACA ACA TTC TGC CCT G; *fru* - 5llfru1- TGC CAC GGT GGT GAA ATT G, 3llfru2- ATC TTT GGC CAA CGT TCA CCT CCT; *slow* - 5llslo1 TAT GGA ATG CTG TGC ATC GG, 3llslo1 TTA TCA TCT TTG TTG GAG CCA C.

PCR reactions were carried out with DNA extracted individually from single male sandflies according to Jowett et al (1998). Fragments obtained were purified using either QIAquick PCR purification kit (QIAGEN) or GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) according to the manufacturers' specifications and cloned using the Zero Blunt TOPO PCR cloning kit for sequencing (Invitrogen) or, alternatively, the CloneJET PCR Cloning Kit (Fermentas). Plasmidial DNA was isolated using the 96 wells microplates and the alkaline lysis method (Sambrook & Russel, 2001) followed by filtration in Millipore Multiscreen filter plates. DNA sequencing was carried out in an ABI Prism 3730 sequencer using the Big Bye 3.0 or 3.1 Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) at the PDTIS sequencing facility at Fundação Oswaldo Cruz. Sequences were submitted to GenBank (Accession numbers XXXX).

The homology of the initial fragments obtained was confirmed with Blastx software (http://blast.ncbi.nlm.nih.gov/Blast.cgi). At least eight clones were sequenced per fly per gene with the objective to determine, when possible, both alleles for each individual. Sequence formatting, editing and alignments were done with BioEdit Sequence Alignment Editor 7.1.3 (Hall 1999).

Molecular polymorphisms and genetic differentiation parameters were calculated using DnaSP v5 and ProSeq softwares (Librado & Rozas 2009; Filatov 2002). Neutrality tests were carried out with the DnaSP v5 program.

The largest non-recombining (NR) blocks as well as the removed sequences archives were generated with the IMgc software (Woerner et al 2007) since IM analyses require datasets with no evidence of recombination. The IM software (Hey & Nielsen 2004) applies the isolation with migration model and is based on Markov chain Monte Carlo (MCMC) simulations of gene genealogies. Six demographic parameters were estimated: effective population sizes for the ancestral (θ_A) and the two descendent populations (θ_1 and θ_2),

divergence time (t) and the migration in both directions (m_1 and m_2). The Infinite Sites model (IS) was chosen because the species studied here represent recent cases of population splitting and the loci analyzed are nuclear genes. Initial runs were conducted to get the ideal parameters and genealogies and to verify how strongly parameter values were auto correlated over the course of the run (to ensure good mixing, certifying the independency of the values) and to achieve, when possible, both stationary distribution and convergence. Runs have been repeated three times with different seed numbers with at least 20,000,000 steps after the burn-in period (100,000 steps) to obtain better estimates of the posterior probability distributions.

Results and Discussion

An isolation with migration model implemented in the IM software was used to estimate gene flow and other parameters using a specific class of markers associated with the control of courtship song. This model is applied to genetic data drawn from a pair of closely related species and its major assumptions involve selectively neutrality, no recombination within loci, free recombination between loci and that mutation has followed the model applied to the data (Hey & Nielsen 2004).

Initially non-recombining blocks were obtained for each locus and their respective length and position are listed in Table 1, which also presents the list of putative recombinant sequences removed and polymorphism parameters for each locus/population studied. An average of 24 consensus sequences were generated per gene per population. Tajima's D (Tajima, 1989) and Fu & Li's (Fu & Li, 1993) tests were performed for each locus and no deviation from the neutral model of evolution was detected and after Bonferroni's correction (Table 2). Comparisons between the pairs of sympatric siblings revealed a high level of genetic differentiation in all case except for *cac* in Sobral (S1S x S2S, Fst – 0.0778) and *fru* in Estrela (E1S x E2S, Fst – 0.0178) (Table 3). Fixed differences were also observed.

Figure 1 shows the marginal posterior probability distributions for the six demographic parameters using IM and the parameter estimates are listed in Table 4. The distributions for time of divergence and effective population size estimates (θ) for the ancestral populations did not present good results not allowing reliable estimates. Nevertheless, Sobral 1S apparently had an expansion of almost twice the size of the ancestral population after separation whereas Sobral 2S suffered a contraction of nearly half size of the same population. The same situation occurred with Jaíba 1S and Jaíba 2S. On the other hand, for the pair Estrela 1S and Estrela 2S we encountered the opposite situation.

For all combined loci analysis (Table 4, Figure 1), migration rate estimates indicate migration in both directions for Sobral 1S and Sobral 2S with Sobral 2S receiving more migrants than Sobral 1S. Estrela 1S and Estrela 2S also displayed migration in both directions, being approximately 20 times higher from Estrela 2S to Estrela 1S. Migration in only one direction (from J1S to J2S) was detected in the pair from Jaíba.

Our results confirm the existence of sympatric species in the three localities in accordance with previous studies performed with this vector (Bauzer et al 202b; Lins et al 2008; Araki et al 2009; Lins et al 2012). In addition, we have found a high degree of genetic divergence between the sympatric siblings analyzed with the occurrence of introgression in most loci. Nevertheless, compared to other loci used a previous multilocus analysis (Araki et al, in preparation) our results demonstrated a higher mean Fst value for genes associated with song production in Sobral (data not shown).

Different levels of gene flow were encountered depending on the marker and pair of populations considered. Asymmetric introgression, already described for other species complexes such as *An. gambiae* (Marsden et al 2011), was also detected between the *L. longipalpis* sympatric species. One possible explanation is the difference in the relative abundance of these species, where the direction of introgression is predicted to occur from the

most abundant to the least one. In fact, the effective population size estimates suggest that Sobral 1S, Jaiba 1S and Estrela 2S have larger population sizes than their sympatric siblings. Differences in the species abundance may also be leading to hybrids backcrossing with the most frequent species (Borge et al 2005). Asymmetric behavioral isolation (Egger et al 2010) may be playing a role in the dynamics of the Longipalpis complex reproductive isolation.

The results presented here shows that although courtship song genes present higher divergence than other loci they can also show signs of gene flow suggesting that introgression is still playing a role in the evolution of the *Lu. longipalpis* complex. Further investigations including allopatric populations and other markers involved in other aspects of the biology of these insects will be useful to better determine the population structure of this vector in Brazil. This will also help addressing questions such as the extent to which and whether genes associated with vectorial capacity cross the species boundaries, and what are the possible epidemiological consequences of such process.

Acknowledgements

We are grateful to Reginaldo Brazil (IOC-FIOCRUZ, Rio de Janeiro, Brazil) and Nataly Souza (IOC-FIOCRUZ, Rio de Janeiro, Brazil) for the *Lu. longipalpis* sandfly samples. We would also like to thank Robson Costa da Silva for technical assistance at FIOCRUZ, Brazil and the PDTIS (FIOCRUZ, Rio de Janeiro, Brazil) sequencing service.

References

Araki AS, Vigoder FM, Bauzer LG, Ferreira GE, Souza NA, et al. (2009) Molecular and behavioral differentiation among Brazilian populations of *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae). PLoS Negl Trop Dis 3:e365.

Arrivillaga J, Mutebi JP, Pinango H, Norris D, Alexandre B, et al. (2003) The taxonomic

status of genetically divergent populations of *Lutzomyia longipalpis* (Diptera: Psychodidae) based on the distribution of mitochondrial and isozyme variation. J Med Entomol 40: 615-627.

Bauzer LG, Souza NA, Ward RD, Kyriacou CP, Peixoto AA (2002a) The *period* gene and genetic differentiation between three Brazilian populations of *Lutzomyia longipalpis*. Insect Mol Biol 11: 315-323.

Bauzer LG, Gesto JS, Souza NA, Ward RD, Hamilton JG, et al. (2002b) Molecular divergence in the *period* gene between two putative sympatric species of the *Lutzomyia longipalpis* complex. Mol Biol Evol 19: 1624-1627.

Bauzer LG, Souza NA, Maingon RDC, Peixoto AA (2007) *Lutzomyia longipalpis* in Brazil: a complex or a single species? A mini-review. Mem Inst Oswaldo Cruz 102: 1-12.

Borge T, Lindroos K, Na´dvorn´ık P, Syva¨nen AC, Sætre GP (2005) Amount of introgression in flycatcher hybrid zones reflects regional differences in pre and post-zygotic barriers to gene exchange. Journal of Evolutionary Biology, 18, 1416– 1424.

Bottecchia M, Oliveira SG, Bauzer LG, Souza NA, Ward RD, et al. (2004) Genetic divergence in the *cacophony* IVS6 intron among five Brazilian populations of *Lutzomyia longipalpis*. J Mol Evol 58: 754-761.

de Souza NA, Ward RD, Hamilton JG, Kyriacou CP, Peixoto AA (2002) Copulation songs in three siblings of *Lutzomyia longipalpis* (Diptera: Psychodidae). Trans R Soc Trop Med Hyg 96: 102-103.

Egger B, Mattersdorfer K, Sefc KM (2010) Variable discrimination and asymmetric preferences in laboratory tests of reproductive isolation between cichlid colour morphs. Journal of Evolutionary Biology, 23, 433–439.

Filatov DA (2002) ProSeq: A software for preparation and evolutionary analysis of DNA sequence data sets. Molecular Ecology Notes 2: 621–624.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics 133:693-709.

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95-98.

Hamilton JG, Hooper AM, Mori K, Pickett JA, Sano A (1999a) 3-Methyl-α-himachalene confirmed, and the relative stereochemistry defined, by synthesis as the sex pheromone of the sandfly *Lutzomyia longipalpis* from Jacobina, Brazil. Chem Commun 355-356.

Hamilton JG, Ibbotson HC, Hooper AM, Mori K, Picket JA, et al. (1999b) 9-Methylgermacrene-B confirmed by synthesis as the sex pheromone of the sandfly *Lutzomyia longipalpis* from Lapinha, Brazil, and the absolute stereochemistry defined as 9S. Chem Commun 2335-2336.

Hamilton JG, Brazil RP, Maingon R (2004) A fourth chemotype of *Lutzomyia longipalpis* (Diptera: Psychodidae) from Jaibas, Minas Gerais State, Brazil. J Med Entomol 41: 1021-1026.

Hamilton JG, Maingon RD, Alexander B, Ward RD, Brazil RP (2005) Analysis of the sex pheromone extract of individual male *Lutzomyia longipalpis* sandflies from six regions in Brazil. Med Vet Entomol 19: 480–488.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of Drosophila pseudoobscura and D. persimilis. Genetics 167:747-760.

Jowett T (1998) Preparation of nucleic acids. In: Roberts DB, editor. *Drosophila*: A practical approach. Oxford: IRL Press. pp. 347-371.

Lainson R, Rangel EF (2005) *Lutzomyia longipalpis* and the eco-epidemiology of American visceral leishmaniasis, with particular reference to Brazil: a review. Mem Inst Oswaldo Cruz 100: 811-827.

Lanzaro GC, Ostrovska K, Herrero MV, Lawyer PG, Warburg A (1993) *Lutzomyia*

longipalpis is a species complex: genetic divergence and interspecific hybrid sterility among three populations. Am J Trop Med Hyg 48: 839-847.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Lins RM, Souza NA, Peixoto AA (2008). Genetic divergence between two sympatric species of the *Lutzomyia longipalpis* complex in the *paralytic* gene, a locus associated with insecticide resistance and lovesong production. Mem Inst Oswaldo Cruz 103:736-40.

Lins RMMA, Souza NA, Brazil RP, Maingon RDC, Peixoto AA (2012). Fixed differences in the *paralytic* gene define two lineages within the *Lutzomyia longipalpis* complex producing different types of courtship songs. PLOS one, in press.

Maingon RD, Ward RD, Hamilton JG, Noyes HA, Souza N, et al. (2003) Genetic identification of two sibling species of *Lutzomyia longipalpis* (Diptera: Psychodidae) that produce distinct male sex pheromones in Sobral, Ceará State, Brazil. Mol Ecol 12: 1879-1894.

Maingon RD, Ward RD, Hamilton JG, Bauzer LG, Peixoto AA (2008). The *Lutzomyia longipalpis* species complex: does population sub-structure matter to *Leishmania* transmission? Trends Parasitol 24:12-17.

Mangabeira O. Sobre a sistemática e biologia de Phlebotomus do Ceará (1969). Rev Brasil Malariol Doenças Trop 21:3-26.

Marsden CD, Lee Y, Nieman CC, Sanford MR, Dinis J, Martins C, Rodrigues A, Cornel AJ, Lanzaro GC (2011). Asymmetric introgression between the M and S forms of the malaria vector, *Anopheles gambiae*, maintains divergence despite extensive hybridization. Mol Ecol 20(23):4983-94.

Sambrook D, Russell J (2001) Molecular cloning: a laboratory manual. Cold Spring Harbour: CSHL Press. 2100 p.

Souza NA, Vigoder FM, Araki AS, Ward RD, Kyriacou CP, et al. (2004) Analysis of the copulatory courtship songs of *Lutzomyia longipalpis* in six populations from Brazil. J Med Entomol 41: 906-913.

Souza NA, Andrade-Coelho CA, Vigoder FM, Ward RD, Peixoto AA (2008) Reproductive isolation between sympatric and allopatric Brazilian populations of *Lutzomyia longipalpis s.l.* (Diptera: Psychodidae). Mem Inst Oswaldo Cruz 103: 216-219.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Ward RD, Ribeiro AL, Ready PD, Murtagh A (1983). Reproductive isolation between different forms of *Lutzomyia longipalpis* (Lutz & Neiva) (Diptera: Psychodidae), the vector of *Leishmania donovani chagasi* Cunha & Chagas and its significance to kala- azar distribution in South America. Mem Inst Oswaldo Cruz 78:269-80.

Ward RD, Phillips A, Burnet B, Marcondes CB (1988) The *Lutzomyia longipalpis* complex: reproduction and distribution. In: Service MW, editor. Biosystematics of Haematophagous Insects. Oxford: Oxford University Press. pp. 258-269.

Watts PC, Hamilton JG, Ward RD, Noyes HA, Souza NA, et al. (2005) Male sex pheromones and the phylogeographic structure of the *Lutzomyia longipalpis* species complex (Diptera: Psychodidae) from Brazil and Venezuela. Am J Trop Med Hyg 73: 734-743.

Woerner A, Cox M, Hammer M (2007) Recombination-Filtered Genomic Datasets by Information Maximization. Bioinformatics 23: 1851-1853.

Figure Legend

Figure 1: Demographic parameters estimates: Marginal posterior probability for the six demographic parameters estimated using IM: **theta –** effective population size, **divergence time** between **A:** Sobral 1S and Sobral 2S, **B:** Jaíba 1S and Jaíba 2S, **C:** Estrela 1S and Estrela 2S and **migration rates** in both directions for the three pairs of sympatric populations. All three IM simulations with different seed numbers are plotted.

Table 1: Sequences features and polymorphism parameters

Locus	Population	N	RM	NR blocks	Length (bp)	Removed sequences	S	π	θ
period	S1S	24 (17)	8	93-175	266 (83)	1S11br781, 1S17br850, 1S18br852, 1S25br784, 1S26br857, 1S19br1000, 1S29br1002	27 (7)	0.02415 (0.02268)	0.02718 (0.02495)
	S2S	23 (18)	7			2S13br916, 2S22br1012, 2S25br1014, 2S23br1025, 2S07br1021	22 (8)	0.01727 (0.01811)	0.02241 (0.02802)
	J1S	17 (7)	7			Jai1S08B, Jai1S08A, Jai1S07, Jai1S03B, Jai1S03A, Jai1S02B, Jai1S02A, Jai1S01B, Jai1S14B, Jai1S14A	20 (5)	0.02214 (0.02180)	0.02224 (0.02459)
	J2S	15 (10)	4			Jai2S14B, Jai2S09A, Jai2S12, Jai2S10A, Jai2S06	17 (5)	0.01733 (0.02008)	0.01966 (0.02129)
	E1S	22 (19)	1			Est1S08, Est1S07, Est1Ss09A	33 (6)	0.02415 (0.02295)	0.03429 (0.02146)
	E2S	30 (27)	5			Est2S26B, Est2S30A, Est2S30B	30 (10)	0.01875 (0.01800)	0.02858 (0.03164)
paralytic	S1S	32 (32)	0	1-285	385 (285)		12 (10)	0.00327 (0.00310)	0.00803 (0.00896)
	S2S	28 (22)	2			S2S11A, S2S11A2, S2S4A, S2S17A, S2S7B, S2S16B	8 (2)	0.00399 (0.00066)	0.00559 (0.00199)
	J1S	24 (24)	0				6 (4)	0.00233 (0.00251)	0.00432 (0.00384)
	J2S	24 (22)	0			J2Ss8B, J2S1A	11 (7)	0.00472 (0.00309)	0.00798 (0.00693)
	E1S	22 (21)	0			E1S8A	5 (4)	0.00190 (0.00138)	0.00371 (0.00401)
	E2S	32 (26)	2			E2S24B, E2Ss1B, E2Ss3A, E2Ss6A, E2Ss7A, E2Ss7A2	6 (3)	0.00384 (0.00084)	0.00405 (0.00285)

cacophony	S1S	23 (17)	2	23-86	116 (64)	1S15br875, 1S16br744, 1S21br1147, 1S30br1153, 1S18lab54, 1S24br1149	15 (2)	0.02859 (0.00490)	0.04726 (0.01232)
	S2S	22 (18)	1			2S11br948, 2S08br528, 2S13br582, 2S23br711	8 (3)	0.03185 (0.02440)	0.02522 (0.01938)
	J1S	20 (17)	3			J1S5B, J1Ss3A, J1Ss3B	16 (4)	0.03624 (0.02268)	0.04698 (0.02005)
	J2S	24 (23)	0			J2Ss8B	7 (5)	0.02570 (0.03271)	0.02205 (0.02336)
	E1S	24 (21)	1			E1S2A, E1S2A2, E1S8A	11 (6)	0.03852 (0.02540)	0.03386 (0.02926)
	E2S	22 (16)	1			E2S26B, E2S27A, E2S28A, E2S28A2, E2S31A, E2S31A2	10 (5)	0.03066 (0.02972)	0.03190 (0.02843)
ebony	S1S	24 (18)	5	1-170	566 (170)	S1S1A, S1S1B, S1S3B, S1S4A, S1S10A, S1S11A	37 (6)	0.01713 (0.00874)	0.01852 (0.01070)
	S2S	18 (18)	4				28 (7)	0.01269 (0.00569)	0.01536 (0.01322)
	J1S	24 (14)	10			J1Ss3A, J1S2B, J1S4A, J1S4B, J1S5A, J1S6B, J1S8A, J1S8B, J1S11A, J1S11A2	51 (8)	0.01897 (0.01282)	0.02562 (0.01613)
	J2S	22 (22)	7				17 (4)	0.00821 (0.00222)	0.00870 (0.00669)
	E1S	22 (20)	7			E1S3A, E1S5B	41 (7)	0.02121 (0.01201)	0.02110 (0.01265)
	E2S	22 (22)	3				26 (5)	0.01328 (0.00571)	0.01328 (0.00841)
fruitless	S1S	24 (22)	3	151-634	634 (484)	S1S16A, S1S20A	20 (10)	0.00641 (0.00387)	0.00927 (0.00629)
	S2S	18 (18)	2				12 (5)	0.00475 (0.00213)	0.00595 (0.00340)
	J1S	22 (14)	2			J1S3B, J1S4A, J1S8B, J1S10A, J1S10A2, J1S11A, J1S11B, J1Ss4B	22 (7)	0.00907 (0.00451)	0.01041 (0.00505)
	J2S	20 (12)	2			J2S1A, J2S2A, J2S2A2, J2Ss4A, J2Ss4B,	16 (1)	0.00827 (0.00038)	0.00772 (0.00076)

						J2Ss7A, J2Ss8B, J2Ss10B			
	E1S	24 (24)	2				14 (8)	0.00352 (0.00248)	0.00657 (0.00506)
	E2S	18 (18)	2				14 (5)	0.00497 (0.00270)	0.00695 (0.00341)
slowpoke	S1S	14 (13)	17	625-805	1073 (181)	S1S10A	141 (13)	0.06651 (0.06549)	0.06426 (0.05661)
	S2S	22 (21)	15			S2S11A	142 (14)	0.05070 (0.04668)	0.05797 (0.06379)
	J1S	18 (15)	23			J1S1A, J1S11A, J1Ss4A	147 (11)	0.06906 (0.07490)	0.06322 (0.05734)
	J2S	20 (11)	15			J2A, J2A2, J2Ss4A, J2Ss4B, J2Ss6A, J2Ss6B, J2Ss9A, J2Ss9B, J2Ss12A	149 (7)	0.06014 (0.04111)	0.05882 (0.03464)
	E1S	18 (16)	10			E1S5B, E1S10A	113 (9)	0.05572 (0.04986)	0.05078 (0.04676)
	E2S	14 (11)	17			E2S25A, E2S27A, E2S27B	154 (9)	0.07019 (0.03792)	0.06821 (0.04390)

n: number of sequences; **RM:** minimum number of recombinant events; **NR:** non-recombining blocks; **S:** number of segregating sites; **π:** nucleotide diversity based on the average number of differences between two sequences; **θ:** nucleotide diversity based on **S**.

Table 2: Neutrality tests for each locus

Population		Locus					
		period	***paralytic***	***cacophony***	***ebony***	***fruitless***	***slowpoke***
D_T	**S1S**	-0.5375	-2.03335	-1.40971	-0.55346	-1.24878	-0.28960
	S2S	-0.8520	-0.88246	0.40955	-0.69521	-0.74571	-0.92921
	J1S	-0.2060	-1.39438	-0.84965	-1.00784	-0.62606	-0.23468
	J2S	-0.4781	-1.38309	0.51732	-0.20455	0.26509	-0.31816
	E1S	-1.3018	-1.45254	0.14732	-0.24794	-1.62760	-0.03361
	E2S	-1.3977	-0.14812	-0.43710	-0.14242	-1.06799	-0.40084
D_{FL}	**S1S**	-0.2517	-1.20010	-1.70657	-0.43840	-0.86133	0.05908
	S2S	-0.9789	0.08210	0.31193	-0.82802	0.27430	-1.65187
	J1S	-0.8120	-0.97286	-0.40467	-1.21041	0.39513	0.18046
	J2S	-0.3821	-1.36881	0.62918	-0.67296	0.24140	1.35994
	E1S	-1.3399	-0.46069	1.02773	0.22496	-2.34918	1.30243
	E2S	-1.3871	1.20922	-0.83604	0.16181	-1.28210	0.57823

D_T: Tajimas's D test; **D_{FL}:** Fu & Li's test of neutrality; **S1S:** Sobral 1S; **S2S:** Sobral 2S; **J1S:** Jaíba 1S; **J2S:** Jaíba 2S; **E1S:** Estrela de Alagoas 1S; **E2S:** Estrela de Alagoas 2S.

Table 3: Genetic differentiation per locus between each pair of *Lu. longipalpis* sympatric species

Locus	Comparison (a x b)	Fst	Dxy	Da	Ss	Sf	Sa	Sb
period	S1S x S2S	0.3952**** (0.5132****)	0.0346 (0.0429)	0.0137 (0.0220)	6 (0)	0 (0)	22 (7)	16 (8)
	J1S x J2S	0.3749**** (0.5074*)	0.0315 (0.0436)	0.0118 (0.0221)	4 (0)	1 (1)	16 (5)	13 (5)
	E1S x E2S	0.4280**** (0.6043 ****)	0.0379 (0.0533)	0.0162 (0.0322)	21 (5)	0 (0)	14 (2)	11 (5)
paralytic	S1S x S2S	0.7740**** (0.8821****)	0.0163 (0.0198)	0.0126 (0.0174)	1 (0)	3 (3)	12 (12)	7 (2)
	J1S x J2S	0.8114**** (0.8665****)	0.0189 (0.0212)	0.0153 (0.0184)	0 (0)	4 (5)	6 (4)	11 (7)
	E1S x E2S	0.7589**** (0.9206****)	0.0120 (0.0140)	0.0091 (0.0129)	2 (1)	0 (0)	3 (3)	4 (2)
cacophony	S1S x S2S	0,0778* (0.0054^{ns})	0.0281 (0.1006)	0.0022 (0.0005)	4 (9)	0 (0)	8 (6)	0 (3)
	J1S x J2S	0.2894** (0.3146***)	0.0339 (0.0522)	0.0098 (0.0164)	4 (3)	0 (0)	5 (2)	0 (3)
	E1S x E2S	0.3781**** (0.5967****)	0.0500 (0.0937)	0.0189 (0.0559)	6 (6)	0 (0)	2 (3)	1 (1)
ebony	S1S x S2S	0.3934**** (0.0867**)	0.0230 (0.0125)	0.0090 (0.0011)	7 (1)	0 (0)	32 (7)	19 (16)
	J1S x J2S	0.2896**** (0.1300**)	0.0187 (0.0128)	0.0054 (0.0017)	7 (1)	1 (0)	40 16)	10 (3)
	E1S x E2S	0.3331**** (0.1388****)	0.0255 (0.0131)	0.0085 (0.0018)	10 (0)	1 (0)	30 (15)	17 (6)
fruitless	S1S x S2S	0.6482**** (0.1963*)	0.0142 (0.0250)	0.0092 (0.0049)	4 (15)	0 (0)	16 (8)	5 (11)
	J1S x J2S	0.5300**** (0.2697*)	0.0164 (0.0195)	0.0087 (0.0053)	3 (13)	1 (1)	18 (7)	11 (1)
	E1S x E2S	0.0178^{ns} (0.0098^{ns})	0.0037 (0.0203)	0.0001 (0.0002)	5 (23)	0 (0)	7 (6)	6 (3)
slowpoke	S1S x S2S	0.2393*** (0.1326*)	0.0562 (0.2411)	0.0134 (0.0320)	57 (51)	0 (0)	35 (4)	33 (19)
	J1S x J2S	0.1760**** (0.0314^{ns})	0.0567 (0.2191)	0.0100 (0.0069)	62 (52)	0 (0)	29 (15)	37 (3)
	E1S x E2S	0.1385** (0^{ns})	0.0546 (0.1521)	0.0076 (0)	47 (28)	0 (0)	34 (15)	47 (26)

S1S: Sobral 1S; **S2S:** Sobral 2S; **J1S:** Jaíba 1S; **J2S:** Jaíba 2S; **E1S:** Estrela de Alagoas 1S; **E2S:** Estrela de Alagoas 2S; **Fst:** pairwise differentiation; **Dxy:** average number of nucleotide substitutions per site; **Da:** number of net nucleotide substitutions per site**; Ss:** number of shared polymorphisms; **Sf:** number of fixed differences; **Sa:** number of sites exclusive to population a; **Sb:** number of sites exclusive to population b; P values were obtained with 10.000 random permutations (ns - non-significant; * p<0.05; **p<0.01; *** p<0.001; **** p<0.0001); numbers in parenthesis represent the estimates obtained using the non-recombining block (NR).

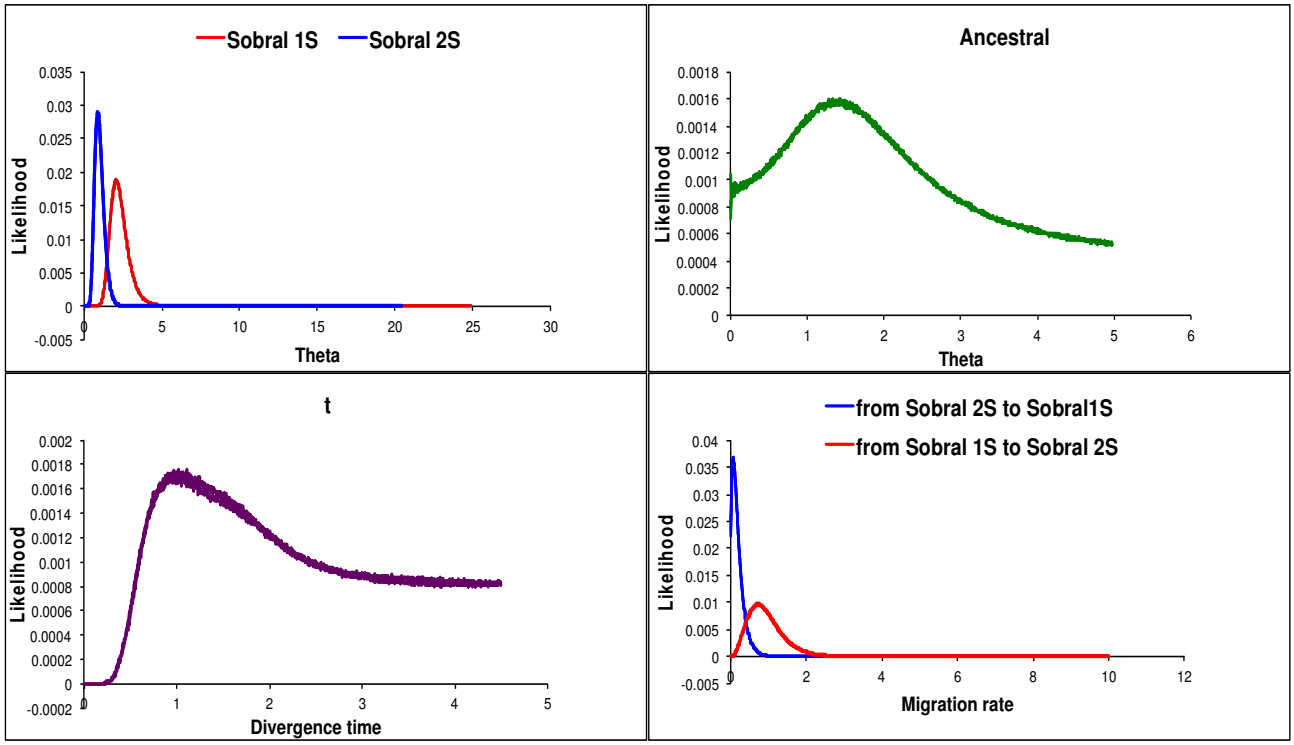
Table 4: IM parameters estimates for all loci per pair of *L. longipalpis* sympatric species

		θ_1			θ_2			θ_A			t			m_1			m_2		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
	Minbin	0.6846	0.6348	0.6348	0.1943	0.1943	0.1739	0.0025	0.0025	0.0025	0.1553	0.1373	0.1462	0.005	0.005	0.005	0.005	0.005	0.005
	Maxbin	18.0854	16.517	15.82	4.2241	4.4082	4.0809	4.9763	4.9763	4.9763	4.4977	4.4977	4.4977	3.145	2.875	2.975	9.935	8.825	7.425
	HiPt	2.0786	2.0288	2.0537	0.8694	0.8694	0.8898	1.3268	1.282	1.4961	1.0193	0.9698	0.9517	0.065	0.065	0.065	0.705	0.715	0.735
	HiSmth	2.0537	2.0537	2.0288	0.8694	0.8694	0.8694	1.4165	1.3866	1.3517	0.9607	0.9878	0.9473	0.055	0.065	0.055	0.725	0.705	0.725
S1S / S2S	**Mean**	2.228	2.2031	2.2031	0.9512	0.9512	0.9512	1.8845	1.8845	1.8944	2.0002	1.9868	2.0408	0.145	0.155	0.155	0.865	0.855	0.865
	95Lo	1.3816	1.3567	1.3567	0.5216	0.5216	0.5012	0.1369	0.1319	0.1319	0.5737	0.5737	0.5827	0.005	0.015	0.015	0.255	0.245	0.255
	95Hi	3.8212	3.8212	3.8212	1.7081	1.7285	1.7081	4.7473	4.7423	4.7473	4.3628	4.3628	4.3673	0.625	0.635	0.625	2.105	2.085	2.115
	HPD90Lo	1.3318	1.3069	1.3069	0.4807	0.4807	0.4807	**0.0025**	**0.0025**	0.0025	**0.5467**	**0.5467**	**0.5603**	0.005	0.005	0.005	0.205	0.205	0.205
	HPD90Hi	3.2237	3.2237	3.1989	1.483	1.483	1.483	**4.0900**	**4.0851**	4.095	**4.4977**	**4.4887**	**4.1963**	0.405	0.405	0.405	1.625	1.605	1.625
	Minbin	0.5015	0.5015	0.5248	0.1814	0.1987	0.1814	0.0029	0.0029	0.0029	0.1625	0.1675	0.1475	0.005	0.005	0.005	0.005	0.005	0.005
	Maxbin	10.7183	10.7183	11.5347	4.3619	4.621	5.0011	5.8286	5.8286	5.8286	4.9975	4.9975	4.9975	4.085	3.355	3.965	6.255	5.705	5.785
	HiPt	1.9244	1.9011	1.8777	0.9069	0.8897	0.8897	0.9068	0.8893	0.9884	1.3825	1.4825	1.3725	0.005	0.005	0.005	0.425	0.425	0.425
	HiSmth	1.9011	1.8777	1.9011	0.8724	0.8724	0.8724	0.901	0.8835	0.866	1.3775	1.3675	1.4125	0.005	0.005	0.005	0.435	0.425	0.435
J1S / J2S	**Mean**	2.0177	2.0177	2.0177	0.9588	0.9588	0.9588	1.6707	1.6766	1.6649	1.9675	1.9575	1.9525	0.125	0.125	0.125	0.555	0.555	0.555
	95Lo	1.2013	1.2013	1.2246	0.5096	0.5096	0.5096	0.0904	0.0904	0.0904	0.6375	0.6375	0.6375	0.005	0.005	0.005	0.125	0.115	0.125
	95Hi	3.3473	3.3473	3.3473	1.7534	1.7534	1.7361	5.4204	5.4204	5.4145	4.7975	4.7975	4.7975	0.755	0.775	0.765	1.455	1.445	1.455
	HPD90Lo	1.178	1.178	1.178	0.4751	0.4751	0.4751	**0.0029**	0.0029	**0.0029**	**0.5725**	**0.5675**	**0.5725**	0.005	0.005	0.005	0.085	0.085	0.085
	HPD90Hi	2.9041	2.9041	2.9041	1.4943	1.4943	1.4943	**4.3532**	4.3474	**4.3415**	**4.3675**	**4.3575**	**4.3275**	0.455	0.465	0.455	1.105	1.095	1.095
	Minbin	0.1627	0.1844	0.1193	0.3761	0.3761	0.3954	0.0027	0.0027	0.0027	0.1013	0.1238	0.1283	0.015	0.045	0.005	0.005	0.005	0.005
	Maxbin	3.8291	4.4583	3.8075	10.3671	8.0911	8.9012	5.421	5.421	5.421	4.4977	4.4977	4.4977	9.995	9.995	9.995	4.985	5.735	6.005
	HiPt	0.7919	0.8353	0.8136	1.4176	1.4176	1.4176	1.0766	0.9898	0.9898	3.2647	1.9732	1.8022	1.525	1.525	1.535	0.075	0.075	0.075
	HiSmth	0.8136	0.8136	0.8136	1.4176	1.3983	1.3983	1.0712	1.0549	1.0658	3.3142	2.2253	1.7393	1.515	1.545	1.545	0.065	0.065	0.065
E1S / E2S	**Mean**	0.8786	0.8786	0.8786	1.4948	1.4948	1.4948	2.1288	2.1234	2.1288	2.6168	2.5673	2.5852	1.785	1.785	1.785	0.185	0.185	0.185
	95Lo	0.4664	0.4664	0.4664	0.9162	0.9162	0.9162	0.122	0.122	0.1166	0.6052	0.6412	0.6142	0.745	0.745	0.755	0.015	0.015	0.015
	95Hi	1.5512	1.5512	1.5512	2.4399	2.4399	2.4592	5.2203	5.2203	5.2203	4.4032	4.4032	4.4032	3.895	3.895	3.885	0.865	0.875	0.865
	HPD90Lo	0.4447	0.4447	0.4447	0.8969	0.8969	0.8969	0.0027	**0.0027**	**0.0027**	1.0328	1.0328	1.0328	0.695	0.695	0.705	0.005	0.005	0.005
	HPD90Hi	1.3559	1.3559	1.3559	2.1506	2.1506	2.1506	4.6291	**4.6346**	**4.6400**	4.4977	4.4977	4.4977	3.085	3.085	3.085	0.525	0.535	0.525

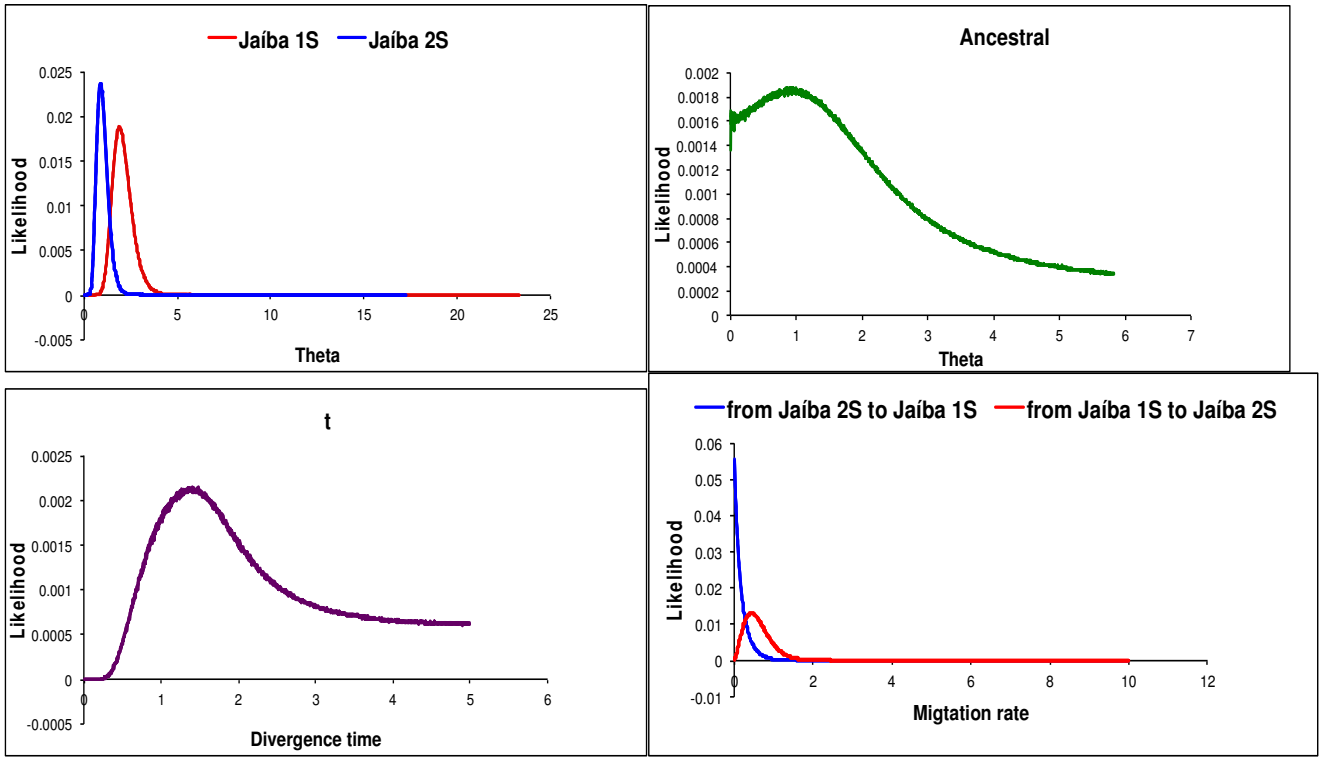
Notes: **S1S:** Sobral 1S; **S2S:** Sobral 2S; **J1S:** Jaíba 1S; **J2S:** Jaíba 2S; **E1S:** Estrela de Alagoas 1S; **E2S:** Estrela de Alagoas 2S;
a, b, c: three independent runs with different seed numbers;
θ: the upper bound of the population mutation parameter for population 1 (θ_1:S1S, J1S, E1S), 2 (θ_2:S2S, J2S, E2S) and ancestral (θ_A);
t: maximum time of population splitting;
m_1: migration rate estimate from 2 to 1; **m2:** migration rate estimate from 1 to 2
Minbin: the midpoint value of the lowest bin; **Maxbin:** the midpoint value of the highest bin
HiPt: the value of the bin with the highest count; **HiSmth:** the value of the bin with the highest count, after the counts have been smoothed by taking the average of 9 points centered on each bin
95Lo: the estimated point to witch 2.5% of the total area lies to the left; **95Hi:** the estimated point to witch 2.5% of the total area lies to the right
HPD90Lo: the lower bound of the estimated 90% highest posterior density (HPD) interval; **HPD90Hi:** the upper bound of the 90% HPD interval (numbers in bold: HPD values not reliable)

Figure 1: Demographic parameters estimates

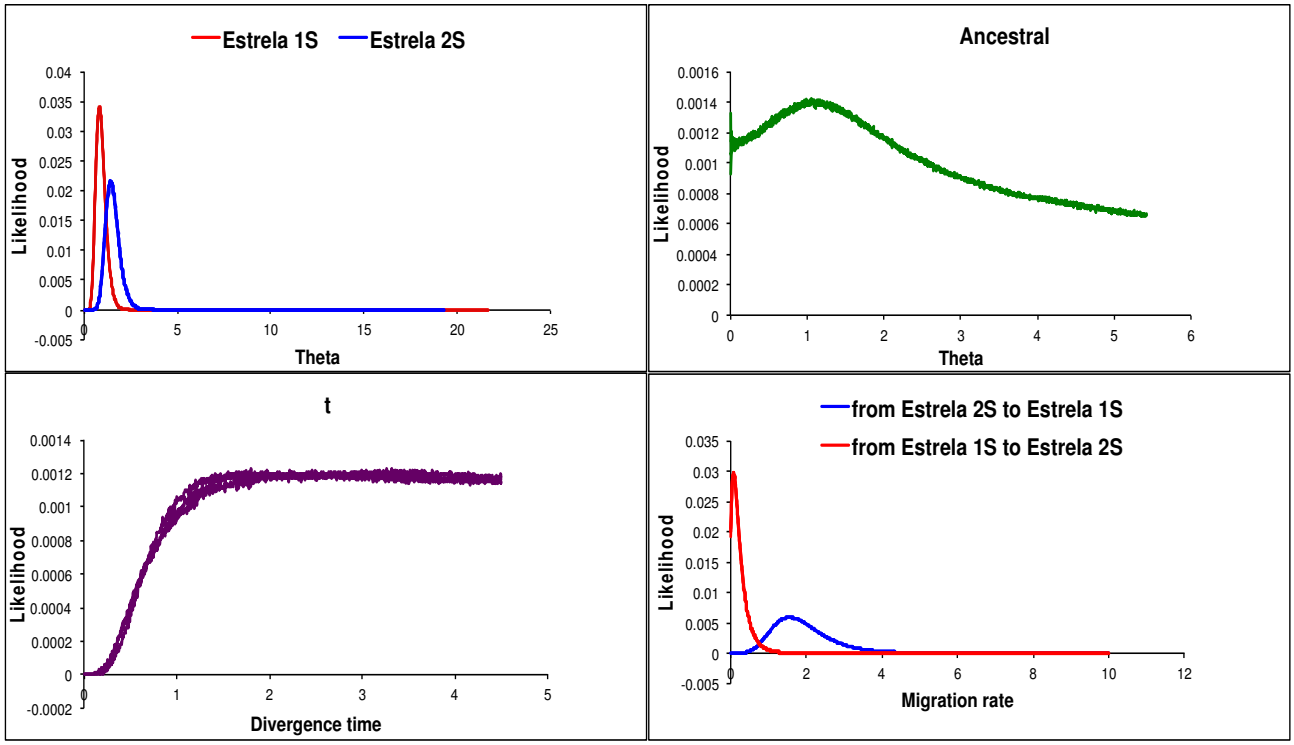
A



B



C



Marginal posterior probability for the six demographic parameters estimated using IM: **theta –** effective population size, **divergence time** between **A:** Sobral 1S and Sobral 2S, **B:** Jaíba 1S and Jaíba 2S, **C:** Estrela 1S and Estrela 2S and **migration rates** in both directions for the three pairs of sympatric populations. All three IM simulations with different seed numbers are plotted.

7. DISCUSSÃO

Nesta tese o expandimos a análise do gene *paralytic* (*para*) a populações já analisadas anteriormente com o gene *period* (*per*) e também a amostras da América Central e do Sul. Além disso, espécies simpátricas do complexo *Lu. longipalpis* das regiões de Sobral (CE), Jaíba (MG) e Estrela de Alagoas (AL) foram utilizadas em uma análise multilocus com genes envolvidos no controle do som, por este motivo a discussão que se segue foi dividida em dois tópicos principais.

7.1 – Análise da divergência e fluxo gênico no gene *para* entre populações de *Lu. longipalpis*

Um entendimento mais amplo de espécies vetoras de doenças torna possível um planejamento adequado de estratégias de controle mais eficientes, minimizando grandes consequências de impacto epidemiológico. Nesse sentido, insetos vetores de parasitas de doenças negligenciadas tem recebido uma atenção mais cuidadosa devido a ameaça que representam à vida de milhões de pessoas (WHO, 2010). Alguns desses vetores compõem complexos de espécies crípticas, trazendo grandes desafios a biólogos evolucionistas no que diz respeito à identificação e delimitação de sua distribuição geográfica devido à falta de uma característica morfológica diagnóstica associada a um isolamento reprodutivo incompleto e à ocorrência de introgressão (Bickford et al. 2007; Pinho & Hey 2010; Hausdorf 2011). *Lu. longipalpis*, vetora da Leishmaniose Visceral Americana (LVA) representa um complexo de espécies crípticas com algumas espécies distribuídas na América Central e do Sul. As evidências mais consistentes que sugerem a existência desse complexo foram obtidas com análises envolvendo os genes *per* e *para*, associados ao controle do som em *Drosophila*, análises de feromônios, som de cópula e microssatélites e suportam a existência de dois grandes grupos de populações no Brasil produzindo diferentes sons de cópula (Burst x Pulsado) e feromônios (cembreno1 x 9-metil-germacreno, himacaleno, cembrenos 1 e 2) (Araki et al. 2009).

Os resultados obtidos com a análise do gene *para* demonstram uma grande diferenciação genética e diferenças fixas entre as populações brasileiras de *Lu. longipalpis* que produzem sons de cópula diferentes, estando de acordo com a divisão em dois grupos principais proposta por Araki e colaboradores (2009). Da mesma forma que encontrado para o gene *per* (Bauzer et al. 2002b; Araki et al. 2009), o gene *para* evidenciou a existência de

espécies do complexo vivendo em simpatia nas regiões de Sobral (CE), Jaíba (MG) e Estrela (AL) com um alto grau de divergência genética, incluindo algumas diferenças fixas entre elas (Lins et al. 2012). Além disso, a análise de populações da região de Palmas (TO) revelou que, nesta região, também existe mais uma espécie simpátrica do complexo Longipalpis cujos machos produzem distintos sons de cópula: Palmas 1S produz o som do tipo Pulsado e Palmas 2S produz o som do tipo *Burst*. Nesta localidade os indivíduos com fenótipo intermediário (PALi) estão agrupados com os machos que possuem o fenótipo de uma pinta (PAL1S) sugerindo que provavelmente representam a mesma espécie (Lins et al. em preparação, pag 65).

Dados gerados em uma análise mais abrangente do gene *para* envolvendo amostras de *Lu. longipalpis* da Colômbia, Venezuela, Brasil e América Central reforçaram a classificação das espécies B (*Lu. pseudolongipalpis*), C – cis-andina (incluindo a população brasileira de Roraima) e D – trans-andina proposta por Arrivillaga e colaboradores (2002, 2003). Esses dados sugerem ainda uma subdivisão da espécie C em dois subgrupos distintos: Colômbia e nordeste da Venezuela. Ainda que não tenham sido encontradas diferenças fixas nessas comparações, a análise do gene *para* revelou uma diferenciação moderada (medida pelo índice Fst) entre essas populações: PG x TRU = 15,20% e DUR x TRU = 11,77%. O agrupamento da população de Roraima com a espécie C reflete possivelmente a sua proximidade geográfica com as populações da Colômbia e nordeste da Venezuela.

A grande diferenciação genética encontrada e a ocorrência diferenças fixas no gene *para* entre as populações estudadas indicam que, no Brasil, *Lu. longipalpis* não forma um único clado ou espécie A (*Lu. longipalpis sensu stricto*) como proposto por Arrivilaga e colaboradores (2002, 2003), formando dois grupos principais (já mencionado anteriormente) de acordo com o som de cópula produzido pelos machos dessas populações. Altos valores de diferenciação genética foram encontrados entre esses dois grupos (média de Fst *Burst* x Pulsado = 0.8613) em relação a comparações utilizando populações que produzem o mesmo som (média de Fst Pulsado x Pulsado = 0.08973 e *Burst* x *Burst* = 0.0506) (artigo 2: Lins et al., em preparação). Esse resultado foi reforçado pela existência de haplótipos distintos encontrados em altas frequências em cada um desses grupos: WFHap_3 para as populações que produzem o som Pulsado e WFHap_5 para as populações que produzem o som do tipo *Burst*. Esses haplótipos poderiam talvez servir como uma ferramenta diagnóstica para separar fêmeas dessas regiões assim como serem sugestivos sobre o fenótipo de som de populações que não tiveram seu som registrado, como é o caso das populações de Araçatuba: compartilha o haplótipo WFHap_5 e foi agrupada com as populações que produzem o som tipo *Burst* e

com as amostras da Argentina e Paraguai. Contudo, esta suposição deve ser tratada com cautela pois análise preliminar do som produzido por machos do Paraguai sugere a ocorrência de um novo som do tipo pulsado (Vigoder, dados não publicados).

Outros dois casos que precisam ser tratados com cuidado são as populações de Estrela 1S e Itamaracá. Estrela 1S parece demonstrar um fenótipo de pintas não completamente confiável na separação das espécies simpátricas nesta localidade uma vez que um dos seus indivíduos possui haplótipos compartilhados com as populações que produzem o som do tipo *Burst*. Itamaracá, apesar de produzir o som do tipo *Burst*, não apresentou nenhuma diferença fixa quando comparada a populações que produzem o som Pulsado apesar da grande diferenciação genética encontrada. Isto se deve provavelmente a uma sequência de um único inseto heterozigoto, sugerindo que este seja um híbrido ou o resultado de um evento de introgressão.

A população de Rio Bonito, que produz um som raro encontrado também na população de Mequita (RJ) (Araki et al. 2009) teve todos os seus haplótipos classificados como WHap_1 (artigo 2: Lins et al., em preparação). Interessantemente, as populações brasileiras de Barra de Guaratiba (RJ, sem som registrado) e Jaíba 1S (MG, som Pulsado) assim como as populações da Colômbia (Palo Gordo, Durania), nordeste da Venezuela (Trujillo), Costa Rica (Brasilito) e Honduras (Isla Del Tigre, San Juan Bautista) compartilham esse mesmo haplótipo e não possuem diferenças fixas entre elas, sugerindo que, talvez, essas populações tenham um ancestral comum, compartilhando polimorfismos ancestrais.

Análises utilizando o gene *per* e *Lu. cruzi* (Vigoder et al. 2010) indicaram que esta representa mais uma espécie do complexo Longipalpis uma vez que alguns valores de diferenciação genética encontrados na comparação entre ela e outras populações de *Lu. longipalpis* foram menores que aqueles encontrados quando somente populações de *Lu. longipalpis* foram consideradas. Esta análise sugeriu ainda que *Lu. cruzi* estaria mais próxima das populações de *Lu. longipalpis* que produzem o som do tipo Pulsado. Destaca-se que tanto *Lu. cruzi* quanto algumas populações de *Lu. longipalpis* que possuem o som pulsado produzem o 9-metil-germacreno como feromônio sexual. Entretanto, os resultados gerados através da análise do gene *para* e essas espécies não só confirmam o status de *Lu. cruzi* como mais uma espécie do complexo Longipalpis como demonstram que ela é uma espécie mais relacionada àquelas populações de *Lu. longipalpis* que produzem o som do tipo *Burst*, onde foram verificadas menor diferenciação genética e diferenças fixas entre elas (Lins et al. 2012). De fato, análises de som de cópula revelaram que *Lu. cruzi* produz um padrão de som do tipo *Burst* muito semelhante àquele produzido por populações brasileiras de *Lu. longipalpis*

(Vigoder et al. 2010). Na análise macrogeográfica, onde uma avaliação mais abrangente foi realizada, os resultados com o gene *para* mostraram uma menor diferenciação genética e ausência de diferenças fixas entre *Lu. cruzi* e populações brasileiras de *Lu. longipalpis*. Contudo, isto se deve a novas regiões de indels encontradas no novo alinhamento gerado que continham as diferenças fixas anteriormente detectadas.

O entendimento da origem, separação e espalhamento das diferentes espécies que formam o complexo Longipalpis representa um grande desafio uma vez que processos recentes de especiação envolvem espécies não completamente isoladas reprodutivamente, tornando difícil o estabelecimento concreto de sua delimitação e distribuição geográfica. Arrivillaga e colaboradores (2002) propuseram uma interpretação biogeográfica inicial, a partir dos dados obtidos da filogenia molecular do gene citocromo oxidase I (COI) e as quatro espécies descritas anteriormente (espécies A, B, C e D). De acordo com esse autores, seus dados sugerem que o processo de especiação no complexo *Lu. longipalpis* começou no Plioceno, a partir de um *pool* gênico sub-Andino-Amazonico resultante da orogenia dos Andes (formação da Cordilheira dos Andes oriental). Os quatro clados mencionados provavelmente divergiram como um resultado de eventos de vicariância que ocorreram durante os estágios finais do Plioceno e Pleistoceno. Considerando as populações brasileiras não representam uma única espécie (espécia A), análises mais abrangentes ainda restam a serem feitas. Para que uma análise biogeográfica mais detalhada seja realizada com as espécies brasileiras, é necessário que outras amostras de regiões intermediárias e de estados brasileiros ainda não estudados sejam incluídos.

7.2 – Análise multilocus de espécies simpátricas do complexo *Lu. longipalpis* utilizando genes associados ao controle do som

Processos recentes de especiação oferecem uma grande oportunidade e desafio para que estudos evolutivos envolvendo divergência genética, com ou sem a presença de fluxo gênico, sejam conduzidos de forma a se tentar elucidar o papel das diferentes forças evolutivas no processo de diferenciação e formação de espécies.

Alguns modelos explicando os fatores necessários para dirigir esse processo em casos de alopatria com contato secundário, simpatria, parapatria e todos aqueles casos onde há fluxo gênico intermitente, incluem a presença de forte seleção divergente e acasalamento seletivo (Maynard Smith 1966; Gavrilets 2006; Pinho & Hey 2010). Durante os estágios iniciais de um processo de divisão de uma população em duas, em um modelo de especiação alopátrica com

contato secundário quando o isolamento reprodutivo é ainda incompleto, os genomas vão, gradativamente, se tornando homogêneos, exceto naquelas regiões associadas ou envolvidas na diferenciação ecológica e isolamento reprodutivo, onde o fluxo gênico é restrito devido à atuação de forças seletivas. Posteriormente, quando a divergência e o isolamento reprodutivo se tornam mais intensos, é esperado que o nível de divergência ao longo do genoma aumente (Wu 2001; Pinho & Hey 2010).

No caso das espécies incipientes de *An. gambiae s.s.* dois modelos foram propostos para explicar os dados obtidos em estudos deste importante vetor da malária. No modelo especiação-com-fluxo-gênico (e.g. Turner et al. 2005), o fluxo gênico causado por hibridização resulta em baixa diferenciação ao longo do genoma, exceto em regiões gênicas envolvidas em dirigir a divergência onde a seleção atua prevenindo a introgressão, resultando em um mosaico de regiões de alta e baixa divergência. Em um outro modelo, proposto por White e colaboradores (2010), o processo de especiação entre as espécies incipientes de *An. gambiae* se desenvolve sem um papel importante para o fluxo gênico. Nesse modelo, o baixo valor adaptativo dos híbridos da F1 resulta em reduzido fluxo gênico efetivo da hibridização, com um consequente espalhamento da divergência ao longo do genoma.

Quando presente durante o processo de especiação, a introgressão pode não ocorrer de forma bidirecional na mesma frequência, sendo chamado este processo de introgressão assimétrica. Neste contexto, uma das espécies funciona como "doadora" de material genético e, a outra, como "receptora". Essa assimetria pode ser causada por diversos fatores, tais como: diferenças na abundância relativa das espécies em questão, o que resultaria no retro-cruzamento dos híbridos com maior frequência com a espécie mais comum (e.g. Borge et al. 2005); expansão do alcance de uma das espécies, nesse caso, a direção da introgressão ocorre da espécie local para a que está invadindo o território em questão (e.g. Excoffier et al. 2009); e ainda, isolamento comportamental assimétrico (Egger et al. 2010), fitness diferencial em cruzamentos *in situ*, seleção para introgressão de genes vantajosos de uma espécie em relação à outra, entre outros (Morgan et al. 2010). Evidência da ocorrência de introgressão assimétrica já foi reportada em alguns complexos de insetos vetores como, por exemplo, *An. gambiae* (Oliveira et al. 2008; Marsden et al. 2011), *An. albitarsis* (Krzywinski et al. 2011), *Cu. pipiens* (Gomes et al. 2009).

No caso do complexo Longipalpis, uma recente análise multilocus (Araki et al., em preparação) envolvendo 21 diferentes loci e espécies brasileiras simpátricas (Sobral 1S e Sobral 2S - CE) e alopátricas (Lapinha – MG e Pancas – ES) foi realizada com o objetivo de verificar a divergência e fluxo gênico existente entre elas. Os resultados desse trabalho

sugerem fortemente a ocorrência de introgressão entre as populações analisadas e, em algumas comparações, essa introgressão ocorre de forma assimétrica e diferencial, com um fluxo gênico maior na direção das espécies que produzem o som do tipo *Burst* e de cinco a dez vezes maior entre as populações simpátricas.

Os resultados obtidos nesta tese com a análise multilocus dos genes associados ao controle do som, confirmam a existência de espécies simpátricas do complexo Longipalpis nas regiões de Sobral (CE), Jaíba (MG) e Estrela (AL), havendo uma grande divergência entre elas com a ocorrência de introgressão em quase todos os loci. Da mesma forma que Araki e colaboradores (2009), nesta análise também se constatou que esse fluxo gênico ocorre de forma diferencial e assimétrica dependendo do locus analisado, onde Sobral 2S parece receber mais migrantes que Sobral 1S. O mesmo ocorre entre Jaiba 2S e 1S. No caso de Estrela, a introgressão parece ocorre numa intensidade maior de Estrela 2S para Estrela 1S. Esta diferença provavelmente está relacionada à abundância relativa das espécies analisadas, de fato tanto Sobral 1S, Jaiba 1S e Estrela 2S são aproximadamente duas vezes mais abundantes que as outras. Dentre os trabalhos realizados utilizando uma abordagem molecular e *Lu. longipalpis*, as análises multilocus revelaram os resultados mais consistentes, demonstrando que, de fato, existe uma grande divergência genética entre membros do complexo, principalmente quando as comparações são realizadas com populações que produzem sons de cópula diferentes. Entretanto, como esperado para espécies em processo recente de separação, a introgressão ainda é detectável em diversos marcadores, e parece ter contribuído grandemente para a estrutura atual do genoma dessas espécies.

7.3 – Considerações finais

Diversas abordagens têm sido utilizadas para que uma visão mais ampla acerca da estruturação do complexo Longipalpis seja obtida. Ainda que alguns trabalhos sugiram a existência de algumas espécies distribuídas na América Central e do Sul, a classificação de uma única espécie brasileira se mostra inconsistente com as diversas análises moleculares e comportamentais envolvendo este vetor. Um melhor entendimento e subsequente classificação das espécies brasileiras se faz necessário para que estratégias de controle epidemiológico mais eficientes sejam empregadas. Além disso, conhecer a intensidade na qual as diferentes espécies trocam genes pode permitir uma previsão da frequência de espalhamento de genes associados à resistência a inseticidas e competência vetorial. Também se fazem necessários estudos mais amplos correlacionando a distribuição da doença com a

ocorrência das diferentes espécies do complexo assim como um conhecimento mais detalhado a cerca de aspectos da capacidade vetorial de cada membro do complexo, como por exemplo o nível de infecção natural por *Leishmania infantum*. Estes dados possivelmente contribuirão para que as diferenças regionais na transmissão da leishmaniose visceral no Novo Mundo sejam esclarecidas. Por fim, sob o ponto de vista evolutivo, se faz necessário que outras localidades sejam estudadas no território brasileiro, para que estudos biogeográficos sejam realizados de forma consistente, permitindo que uma avaliação robusta sobre a possível origem e espalhamento das espécies do complexo Longipalpis no Brasil seja realizada.

8. REFERÊNCIAS GERAIS

Acardi SA, Liotta DJ, Santini MS, Romagosa CM, Salomón OD. Detection of *Leishmania infantum* in naturally infected *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae) and *Canis familiaris* in Misiones, Argentina: the first report of a PCR-RFLP and sequencing-based confirmation assay. Mem Inst Oswaldo Cruz 2010 Sep;105(6):796-9.

Alexander B, Mutebi JP, Hearne D, Lanzaro GC, Ward RD, Hamilton JGC. Current Status of the *Lutzomyia longipalpis* Species Complex. Mem Inst Oswaldo Cruz 1998;93:31-3.

Araki AS, Vigoder FM, Bauzer LG, Ferreira GE, Souza NA, Araújo IB, et al. Molecular and behavioral differentiation among Brazilian populations of *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae). PLoS Negl Trop Dis 2009;3:e365.

Arrivillaga JC, Norris DE, Feliciangeli MD, Lanzaro GC. Phylogeography of the neotropical sand fly *Lutzomyia longipalpis* inferred from mitochondrial DNA sequences. Infect Genet Evol 2002 Dec;2(2):83-95.

Arrivillaga J, Mutebi JP, Pinango H, Norris D, Alexandre B, Feliciangeli MD, et al. The taxonomic status of genetically divergent populations of *Lutzomyia longipalpis* (Diptera: Psychodidae) based on the distribution of mitochondrial and isozyme variation. J Med Entomol 2003;40:615-27.

Barretto MP. Novos subgêneros de *Lutzomyia* França, 1924 (Diptera, Psychodidae, subfamília Phlebotominae). Rev Inst Med Trop 1962;4:91-110.

Bauzer LG, Souza NA, Ward RD, Kyriacou CP, Peixoto AA. The *period* gene and genetic differentiation between three Brazilian populations of *Lutzomyia longipalpis*. Insect Mol Biol 2002(a);11:315-23.

Bauzer LG, Gesto JS, Souza NA, Ward RD, Hamilton JG, Kyriacou CP, et al. Molecular divergence in the *period* gene between two putative sympatric species of the *Lutzomyia longipalpis* complex. Mol Biol Evol 2002(b);19:1624-7.

Bauzer LG, Souza NA, Maingon RDC, Peixoto AA. *Lutzomyia longipalpis* in Brazil: a complex or a single species? A mini-review. Mem Inst Oswaldo Cruz 2007;102:1–12.

Besansky NJ, Krzywinski J, Lehmann T, Simard F, Kern M, Mukabayire O, et al. Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. Proc Natl Acad Sci USA 2003;100:10818-23.

Bickford D, Lohman DJ, Sodhi NS, Ng PK, Meier R, Winker K, et al. Cryptic species as a window on diversity and conservation. Trends Ecol Evol 2007;22:148–55.

Borge T, Lindroos K, Na´dvorn´ık P, Syva¨nen AC, Sætre GP. Amount of introgression in flycatcher hybrid zones reflects regional differences in pre and post-zygotic barriers to gene exchange. Journal of Evolutionary Biology 2005;18:1416–24.

Bottecchia M, Oliveira SG, Bauzer LG, Souza NA, Ward RD, Garner KJ, et al. Genetic divergence in the *cacophony* IVS6 intron among five Brazilian populations of *Lutzomyia longipalpis*. J Mol Evol 2004;58:754-61.

Brazil RP, Hamilton JG. Isolation and identification of 9-methylgermacrene- B as the putative sex pheromone of *Lutzomyia cruzi* (Mangabeira, 1938) (Diptera: Psychodidae). Mem Inst Oswaldo Cruz 2002;97:435–6.

Brazil RP, Gomes Brazil B. Bionomia. In Rangel EF, Lainson R, editors. Flebotomineos do Brasil. Rio de Janeiro: Fiocruz; 2003. p. 257-74.

Brazil RP, Caballero NN, Hamilton JG. Identification of the sex pheromone of *Lutzomyia longipalpis* (Lutz & Neiva, 1912) (Diptera: Psychodidae) from Asunción, Paraguay. Parasit Vectors 2009 Nov;2(1):51.

Brazil RP, Passos WL, Brazil BG, Temeljkovitch M, Andrade-Filho JD. Diptera, Psychodidae, *Phlebotominae Rondani*, 1840: Range extension and new records from lowland Bolivia. Check List 2010;6(4):587-8.

Brooks R, Hunt J, Blows MW, Smith MJ, Bussiere LF, Jennions MD. Experimental evidence for multivariate stabilizing sexual selection. Evolution 2005;59: 871–80.

de Azevedo AC, Monteiro FA, Cabello PH, Souza NA, Rosa-Freitas MG, Rangel EF. Studies on populations of *Lutzomyia longipalpis* (Lutz & Neiva, 1912) (Diptera: Psychodidae: Phlebotominae) in Brazil. Mem Inst Oswaldo Cruz 2000 May-Jun;95(3):305-22.

de Pita-Pereira D, Cardoso MA, Alves CR, Brazil RP, Britto C. Detection of natural infection in *Lutzomyia cruzi* and *Lutzomyia forattinii* (Diptera: Psychodidae: Phlebotominae) by *Leishmania infantum chagasi* in an endemic area of visceral leishmaniasis in Brazil using a PCR multiplex assay. Acta Trop 2008 Jul;107(1):66-9.

de Souza NA, Ward RD, Hamilton JG, Kyriacou CP, Peixoto AA. Copulation songs in three siblings of *Lutzomyia longipalpis* (Diptera: Psychodidae). Trans R Soc Trop Med Hyg 2002;96: 102-3.

De Silans LN, Dedet JP, Arias JR. Field monitoring of cypermethrin residual effect on the mortality rates of the Phlebotomine sand fly *Lutzomyia longipalpis* in the state of Paraíba, Brazil. Mem Inst Oswaldo Cruz 1998 May-Jun;93(3):339-44.

Egger B, Mattersdorfer K, Sefc KM. Variable discrimination and asymmetric preferences in laboratory tests of reproductive isolation between cichlid colour morphs. Journal of Evolutionary Biology 2010;23:433–9.

Excoffier L, Foll M, Petit RJ. Genetic consequences of range expansions. Annual Review of Ecology Evolution and Systematics 2009;40:481–501.

Feliciangeli MD, Mazzarri MB, Blas SS, Zerpa O. Control trial of *Lutzomyia longipalpis* s.l. in the Island of Margarita, Venezuela. Trop Med Int Health 2003 Dec;8(12):1131-6. Erratum in: Trop Med Int Health 2004 Feb;9(2):314.

Galati EAB, Nunes VL, Rego Jr FA, Oshiro ET, Rodrigues M. Estudo de flebotomíneos (Diptera, Psychodidae) em foco de leishmaniose visceral no Estado de Mato Grosso do Sul, Brasil. Rev Saúde Pública 1997;31(4):378–90.

Galati EAB. Morfologia e taxonomia 2.1: classificação de Phlebotominae. In Rangel EF, Lainson R, editors. Flebotomineos do Brasil. Rio de Janeiro: Fiocruz; 2003. p. 23-51.

Gavrilets S. The Maynard Smith model of sympatric speciation. Journal of Theoretical Biology 2006;239:172–82.

Gomes B, Sousa C, Novo MT, Freitas FB, Alves R, Côrte-Real AR, et al. Asymmetric introgression between sympatric molestus and pipiens forms of *Culex pipiens* (Diptera: Culicidae) in the Comporta region, Portugal. BMC Evolutionary Biology 2009;9:262.

Grillet M, Everaerts C, Houot B, Ritchie MG, Cobb M, Ferveur JF. Incipient speciation in *Drosophila melanogaster* involves chemical signals. Sci Rep 2012;2:224.

Groot AT, Horovitz JL, Hamilton J, Santangelo RG, Schal C, Gould F. Experimental evidence for interspecific directional selection on moth pheromone communication. Proc Natl Acad Sci USA 2006 Apr;103(15):5858-63.

Groot AT, Classen A, Staudacher A, Schal C, Heckel DG. Phenotypic plasticity in sexual communication signal of a noctuid moth. J Evol Biol 2010 Dec;23(12):2731-8.

Hall JC. The mating of a fly. Science 1994 Jun;264(5166):1702-14.

Hamilton JG, Ward RD, Dougherty MJ, Maignon R, Ponce C, Ponce E, et al. Comparison of the sex-pheromone components of *Lutzomyia longipalpis* (Diptera:Psychodidae) from areas of visceral and atypical cutaneous leishmaniasis in Honduras and Cost Rica. Ann Trop Med Parasitol 1996 Oct;90(5):533-41.

Hamilton JG, Hooper AM, Mori K, Pickett JA, Sano A. 3-Methyl-ahimachalene confirmed, and the relative stereochemistry defined, by synthesis as the sex pheromone of the sandfly *Lutzomyia longipalpis* from Jacobina, Brazil. Chem Commun 1999(a);355–6.

Hamilton JG, Ibbotson HC, Hooper AM, Mori K, Picket JA, Sano S. 9- Methylgermacrene-B confirmed by synthesis as the sex pheromone of the sandfly *Lutzomyia longipalpis* from Lapinha, Brazil, and the absolute stereochemistry defined as 9S. Chem Commun 1999(b);2335–6.

Hamilton JG, Brazil RP, Maingon R. A fourth chemotype of *Lutzomyia longipalpis* (Diptera: Psychodidae) from Jaibas, Minas Gerais State, Brazil. J Med Entomol 2004 Nov;41(6):1021-6.

Hamilton JG, Maingon RD, Alexander B, Ward RD, Brazil RP. Analysis of the sex pheromone extract of individual male *Lutzomyia longipalpis* sandflies from six regions in Brazil. Med Vet Entomol 2005 Dec;19(4):480-8.

Hardin PE. Molecular genetic analysis of circadian timekeeping in *Drosophila*. Adv Genet 2011;74:141-73.

Hausdorf B. Progress toward a general species concept. Evolution 2011;65:923-31.

Hille B. G protein-coupled mechanisms and nervous signaling. Neuron 1992 Aug;9(2):187-95.

Hodgkinson VH, Birungi J, Quintana M, Dietze R, Munstermann LE. Mitochondrial cytochrome b variation in populations of the visceral leishmaniasis vector *Lutzomyia longipalpis* across eastern Brazil. Am J Trop Med Hyg 2003;69:386-392.

Konopka RJ, Benzer S. Clock mutants of *Drosophila melanogaster*. Proc Natl Acad Sci USA 1971;68:2112–6.

Krzywinski J, Li C, Morris M, Conn JE, Lima JB, Povoa MM, et al. Analysis of the evolutionary forces shaping mitochondrial genomes of a Neotropical malaria vector complex. Mol Phylogenet Evol 2011 Mar;58(3):469-77.

Kyriacou CP, Hall JC. The function of courtship song rhythms in *Drosophila*. Anim Behav 1982;30:784–801.

Kyriacou CP, Hall JC. Interspecific genetic control of courtship song production and reception in *Drosophila*. Science 1986 Apr;232(4749):494-7.

Lanzaro GC, Ostrovska K, Herrero MV, Lawyer PG, Warburg A. *Lutzomyia longipalpis* is a Species Complex: Genetic Divergence and Interespecific Hybrid Sterility among Three Populations. Am J Trop Med Hyg 1993;48(6):839-47.

Lewis DJ, Young DG, Fairchild GB, Minter DM. Proposals for a stable classification of the phlebotomine sandflies (Diptera: Psychodidae). Syst Entomol 1977;2:319-32.

Lins RM, Oliveira SG, Souza NA, de Queiroz RG, Justiniano SC, Ward RD, et al. Molecular evolution of the cacophony IVS6 region in sandflies. Insect Mol Biol 2002 Apr;11(2):117-22.

Lins RM, Souza NA, Peixoto AA. Genetic divergence between two sympatric species of the *Lutzomyia longipalpis* complex in the *paralytic* gene, a locus associated with insecticide resistance and lovesong production. Mem Inst Oswaldo Cruz 2008;103:736-40.

Lins RMMA, Souza NA, Brazil RP, Maingon RDC, Peixoto AA. Fixed differences in the *paralytic* gene define two lineages within the *Lutzomyia longipalpis* complex producing different types of courtship songs. PLOS one 2012, in press.

Littleton JT, Ganetzky B. Ion channels and synaptic organization: analysis of the *Drosophila* genome. Neuron 2000 Apr;26(1):35-43.

Loughney K, Kreber R, Ganetzky B. Molecular analysis of the *para* locus, a sodium channel gene in *Drosophila*. Cell 1989 Sep;58(6):1143-54.

Machado CA, Kliman RM, Markert JA, Hey J. Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. Mol Biol Evol 2002;19:472-88.

Maingon RD, Ward RD, Hamilton JG, Bauzer LG, Peixoto AA. The *Lutzomyia longipalpis* species complex: does population sub-structure matter to *Leishmania* transmission? Trends Parasitol 2008;24:12-17.

Mangabeira O. Sobre a sistemática e biologia de Phlebotomus do Ceará. Rev Brasil Malariol Doenças Trop 1969;21:3-26.

Marsden CD, Lee Y, Nieman CC, Sanford MR, Dinis J, Martins C, et al. Asymmetric introgression between the M and S forms of the malaria vector, *Anopheles gambiae*, maintains divergence despite extensive hybridization. Mol Ecol 2011 Dec;20(23):4983-94.

Martins AV, Williams P, Falcão AL. American Sand Flies (Diptera: Psychodidae: Phlebotominae). Rio de Janeiro: Acad Bras Ciênc; 1978.

Mazzoni CJ, Araki AS, Ferreira GE, Azevedo RV, Barbujani G, Peixoto AA. Multilocus analysis of introgression between two sand fly vectors of leishmaniasis. BMC Evol Biol 2008;8:141.

Maynard Smith J. Sympatric speciation. American Naturalist 1966;100:637–50.

Missawa NA, Michalsky EM, Fortes-Dias CL, Santos Dias E. *Lutzomyia longipalpis* naturally infected by *Leishmania (L.) chagasi* in Várzea Grande, Mato Grosso State, Brazil, an area of intense transmission of visceral leishmaniasis. Cad Saude Publica 2010 Dec;26(12):2414-9.

Morgan K, Linton Y-M, Somboon P Somboon P, Saikia P, Dev V, et al. Inter-specific gene flow dynamics during the Pleistocene-dated speciation of forest-dependent mosquitoes in Southeast Asia. Molecular Ecology 2010;19:2269–85.

Mukhopadhyay J, Ghosh K, Rangel EF, Munstermann LE. Genetic variability in biochemical characters of Brazilian field populations of the Leishmania vector, Lutzomyia longipalpis (Diptera: Psychodidae). Am J Trop Med Hyg 1998;59:893-901.

Mutebi JP, Alexander B, Sherlock I, Wellington J, Souza AA, Shaw J, et al. Breeding structure of the sand fly Lutzomyia longipalpis (Lutz & Neiva) in Brazil. Am J Trop Med Hyg 1999;61:149-57.

Oliveira E, Salgueiro P, Palsson K, Vicente JL, Arez AP, Jaenson TG, et al. High levels of hybridization between molecular forms of *Anopheles gambiae* from Guinea Bissau. Journal of Medical Entomology 2008;45:1057–63.

Paiva BR, Oliveira AG, Dorval ME, Galati EA, Malafronte RS. Species-specific identification of Leishmania in naturally infected sand flies captured in Mato Grosso do Sul State, Brazil. Acta Trop 2010 Jul-Aug;115(1-2):126-30.

Peixoto AA, Hall JC. Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of *Drosophila*. Genetics 1998 Feb;148(2):827-38.

Peixoto AA, Gomes CA, de Amoretty PR, Lins RM, Meireles-Filho AC, de Souza NA, Kyriacou CP. New molecular markers for phlebotomine sand flies. Int J Parasitol 2001 May;31(5-6):635-9.

Pinho C, Hey J. Divergence with Gene Flow: Models and Data. Annual Review of Ecology, Evolution, and Systematics 2010;41:215-30.

Pinto IS, Andrade Filho JD, Santos CB, Falqueto A, Leite YLR. Phylogenetic relationships among species of *Lutzomyia*, subgenus *Lutzomyia* (Diptera: Psychodidae). J Med Entomol 2010;47:16–21.

Pittendrigh B, Reenan R, ffrench-Constant RH, Ganetzky B. Point mutations in the *Drosophila* sodium channel gene *para* associated with resistance to DDT and pyrethroid insecticides. Mol Gen Genet 1997 Nov;256(6):602-10.

Reddy GV, Guerrero A. New pheromones and insect control strategies. Vitam Horm 2010;83:493-519.

Ritchie MG, Halsey EJ, Gleason JM. *Drosophila* song as a species-specific mating signal and the behavioural importance of Kyriacou & Hall cycles in *D. melanogaster* song. Anim. Behav 1999 58:649–57.

Rocha L de S, Falqueto A, Dos Santos CB, Grimaldi GJ, Cupolillo E. Possible implication of the genetic composition of the *Lutzomyia longipalpis* (Diptera: Psychodidae) populations in the epidemiology of the visceral leishmaniasis. J Med Entomol 2011 Sep;48(5):1016-22.

Rona LD, Carvalho-Pinto CJ, Mazzoni CJ, Peixoto AA. Estimation of divergence time between two sibling species of the *Anopheles* (Kerteszia) *cruzii* complex using a multilocus approach. BMC Evol Biol 2010 Mar;10:91.

Salomón OD, Araki AS, Hamilton JG, Acardi SA, Peixoto AA. Sex pheromone and *period* gene characterization of *Lutzomyia longipalpis* sensu lato (Lutz & Neiva) (Diptera: Psychodidae) from Posadas, Argentina. Mem Inst Oswaldo Cruz 2010 Nov;105(7):928-30.

Salomón OD, Basmajdian Y, Fernández MS, Santini MS. *Lutzomyia longipalpis* in Uruguay: the first report and the potential of visceral leishmaniasis transmission. Mem Inst Oswaldo Cruz 2011 May;106(3):381-2.

Santos SO, Arias JR, Ribeiro AA, Hoffmann MP, Freitas RA, Malacco MAF. Incrimination of *Lutzomyia cruzi* as a vector of American Visceral Leishmaniasis. Med Vet Entomol 1988;12: 315–7.

Shirangi TR, Dufour HD, Williams TM, Carroll SB. Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. PLoS Biol. 2009 Aug;7(8):e1000168.

Slotman MA, Della Torre A, Calzetta M, Powell JR. Differential introgression of chromsomal regions between *Anopheles gambiae* and *An. arabiensis*. Am J Trop Med Hyg 2005 Aug;73(2):326-35.

Smith LA, Peixoto AA, Kramer EM, Villella A, Hall JC. Courtship and visual defects of *cacophony* mutants reveal functional complexity of a calcium-channel alpha1 subunit in *Drosophila*. Genetics 1998 Jul;149(3):1407-26.

Soto SI, Lehmann T, Rowton ED, Velez BI, Porter CH. Speciation and population structure in the morphospecies *Lutzomyia longipalpis* (Lutz & Neiva) as derived from the mitochondrial ND4 gene. Mol Phylogenet Evol 2001;18:84-93.

Souza NA, Vigoder FM, Araki AS, Ward RD, Kyriacou CP, Peixoto AA. Analysis of the copulatory courtship songs of *Lutzomyia longipalpis* in six populations from Brazil. J Med Entomol 2004;41:906-13.

Souza NA, Andrade-Coelho CA, Vigoder FM, Ward RD, Peixoto AA. Reproductive isolation between sympatric and allopatric Brazilian populations of *Lutzomyia longipalpis s.l.* (Diptera: Psychodidae). Mem Inst Oswaldo Cruz 2008;103:216-9.

Spiegel CN, Brazil RP, Soares MJ. Ultrastructural cytochemistry of the sex pheromone glands of *Lutzomyia cruzi* male sand flies (Diptera: Psychodidae: Phlebotominae). Arthropod Struct Dev 2004 Oct;33:399–404.

Steiger S, Schmitt T, Schaefer HM. The origin and dynamic evolution of chemical information transfer. Proc Biol Sci 2011 Apr;278(1708):970-9.

Stump AD, Shoener JA, Costantini C, Sagnon N, Besansky NJ. Sex-linked differentiation between incipient species of *Anopheles gambiae*. Genetics 2005 Mar;169(3):1509-19.

Theodor O. On the classification of American Phlebotominae. J Med Entomol 1965;2:171-197.

Tripet F, Dolo G, Lanzaro GC. Multilevel analyses of genetic differentiation in *Anopheles gambiae* s.s. reveal patterns of gene flow important for malaria-fighting mosquito projects. Genetics 2005 Jan;169(1):313-24.

Turner T, Hahn M, Nuzhdin S. Genomic islands of speciation in *Anopheles gambiae*. PLoS Biology 2005;3:e285.

Vigoder FM, Araki AS, Bauzer LG, Souza NA, Brazil RP, Peixoto AA. Lovesongs and *period* gene polymorphisms indicate *Lutzomyia cruzi* (Mangabeira, 1938) as a sibling species of the *Lutzomyia longipalpis* (Lutz and Neiva, 1912) complex. Infect Genet Evol 2010(a) Aug;10(6):734-9.

Vigoder FM, Souza NA, Peixoto AA. Copulatory courtship song in *Lutzomyia migonei* (Diptera: psychodidae). Mem Inst Oswaldo Cruz 2010(b) Dec;105(8):1065-7.

Vigoder FM, Souza NA, Peixoto AA. Acoustic signals in the sand fly *Lutzomyia (Nyssomyia) intermedia* (Diptera: Psychodidae). Parasit Vectors 2011 May;4:76.

Wang-Sattler R, Blandin S, Ning Y, Blass C, Dolo G, Toure YT, et al. Mosaic genome architecture of the *Anopheles gambiae* species complex. PLoSOne 2007; 2:e1249.

Ward RD, Ribeiro AL, Ready PD, Murtagh A. Reproductive isolation between different forms of *Lutzomyia longipalpis* (Lutz & Neiva) (Diptera: Psychodidae), the vector of *Leishmania donovani chagasi* Cunha & Chagas and its significance to kala- azar distribution in South America. Mem Inst Oswaldo Cruz 1983;78:269-80.

Ward RD, Phillips A, Burnet B, Marcondes CB. The *Lutzomyia longipalpis* complex: reproduction and distribution. In: Service MW, editor. Biosystematics of Haematophagous Insects. Oxford: Oxford University Press; 1988. p. 258-69.

Watts PC, Hamilton JG, Ward RD, Noyes HA, Souza NA, Kemp SJ, et al. Male sex pheromones and the phylogeographic structure of the *Lutzomyia longipalpis* species complex (Diptera: Psychodidae) from Brazil and Venezuela. Am J Trop Med Hyg 2005;73:734–43.

Weill M, Chandre F, Brengues C, Manguin S, Akogbeto M, Pasteur N, et al. The kdr mutation occurs in the Mopti form of *Anopheles gambiae* s.s. through introgression. Insect Mol Biol 2000 Oct;9(5):451-5.

White BJ, Cheng C, Simard F, Costantini C, Besansky NJ. Genetic association of physically unlinked islands of genomic divergence in incipient species of *Anopheles gambiae*. Molecular Ecology 2010;19:925–39.

World Heath Organization [homepage na internet]. First WHO report on neglected tropical diseases: working to overcome the global impact of neglected tropical diseases [acesso em julho de 2012]. Disponível em: http://www.who.int/neglected_diseases/2010report/en/

Wu C-I. The genic view of the process of speciation. Journal of Evolutionary Biology 2001;14: 851–65.

Young DG, Duncan MA. Guide to the Identification and Geographic Distribution of *Lutzomyia* Sand in Mexico, the West Indies, Central and South America (Diptera: Psychodidae). Gainesville: Mem Am Entomol Inst Associated Publishers; 1994.

Brazil RP, Gomes Brazil B. Bionomia. In Rangel EF, Lainson R, editors. Flebotomineos do Brasil. Rio de Janeiro: Fiocruz; 2003. p. 257-74.