

RESEARCH

Open Access



Microgeographical structure in the major Neotropical malaria vector *Anopheles darlingi* using microsatellites and SNP markers

Melina Campos¹, Jan E. Conn^{2,3}, Diego Peres Alonso¹, Joseph M. Vinetz^{4,5}, Kevin J. Emerson⁶ and Paulo Eduardo Martins Ribolla^{1*}

Abstract

Background: In recent decades, throughout the Amazon Basin, landscape modification contributing to profound ecological change has proceeded at an unprecedented rate. Deforestation that accompanies human activities can significantly change aspects of anopheline biology, though this may be site-specific. Such local changes in anopheline biology could have a great impact on malaria transmission. The aim of this study was to investigate population genetics of the main malaria vector in Brazil, *Anopheles darlingi*, from a microgeographical perspective.

Methods: Microsatellites and ddRADseq-derived single nucleotide polymorphisms (SNPs) were used to assess levels of population genetic structuring among mosquito populations from two ecologically distinctive agricultural settlements (~60 km apart) and a population from a distant (~700 km) urban setting in the western Amazon region of Brazil.

Results: Significant microgeographical population differentiation was observed among *Anopheles darlingi* populations via both model- and non-model-based analysis only with the SNP dataset. Microsatellites detected moderate differentiation at the greatest distances, but were unable to differentiate populations from the two agricultural settlements. Both markers showed low polymorphism levels in the most human impacted sites.

Conclusions: At a microgeographical scale, signatures of genetic heterogeneity and population divergence were evident in *Anopheles darlingi*, possibly related to local environmental anthropic modification. This divergence was observed only when using high coverage SNP markers.

Keywords: *Anopheles darlingi*, Amazonian Brazil, Malaria, Microsatellite markers, SNPs, DdRADseq

Background

The prevalence of malaria in tropical and subtropical regions [1] is due mainly to environmental conditions that are suitable for the survival of the vector anopheline mosquitoes through the extrinsic incubation period of *Plasmodium* [2]. Among Neotropical countries, Brazil has the highest proportion of malaria cases, and nearly all transmission occurs in the Amazon region [1] where *Anopheles darlingi* is the primary vector. Four main factors of *An. darlingi*'s life history have contributed to its

pivotal role in *Plasmodium* transmission: susceptibility to human *Plasmodium* species; anthropophilic or opportunistic behavior [3–5]; rapid adaptability to local environmental modification [6, 7]; and the ability to blood feed successfully inside and outside houses [8, 9].

Deforestation and microclimate change that accompany human activity can significantly increase the human biting rate and other vector biology parameters in anopheline vectors across the globe [6, 7, 10, 11], though this may be site-specific [12–16]. Differences in environmental conditions have contributed to *An. darlingi* population structure spatially [6, 17, 18] and temporally [19, 20]. In Amazonian Brazil, rural settlements are subjected to geographical change by human interventions,

* Correspondence: pribolla@ibb.unesp.br

¹Biotechnology Institute (IBTEC) & Biosciences Institute at Botucatu (IBB), Sao Paulo State University (UNESP), Sao Paulo, Brazil

Full list of author information is available at the end of the article



for example, agriculture development, forest degradation and increases in house numbers. In general, the more recently occupied settlements, covered by a greater proportion of forest, have the greatest abundance of *An. darlingi* and the highest proportion of malaria cases compared with older settlements where there is increased deforestation and urbanization, and fewer malaria cases, a phenomenon described as frontier malaria [21, 22]. In the present study, we analyze *An. darlingi* populations from three endemic areas, ranging from a rural to an urban environment. Different proportions of anthropogenic (built) environment between urban and rural settings may lead to ecological segregation in breeding sites, resulting in divergence/speciation, as observed in *An. gambiae* (*s.l.*) in Cameroon [23].

Population genetic studies in the context of vector biology have used a variety of molecular markers, among them microsatellites and single nucleotide polymorphisms (SNPs). The former is a multiallelic marker that provides valuable polymorphism information, and it has been an important tool for numerous population genetics studies in *An. darlingi* and other vector species [24–26]. In *An. darlingi*, microsatellite markers have revealed moderate to high levels of genetic heterogeneity; subpopulations have been found at a macrogeographical scale (greater than 150 km apart) [17, 27, 28], and more surprisingly, seasonally related genetic subpopulations [20]. Single nucleotide polymorphisms have become popular for population genomics due to improvements in next generation sequencing and progressive cost reduction [29, 30]. Restriction-site Associated DNA sequencing (RADseq) and derivative approaches that generate SNP datasets have been used successfully to investigate genetic features in anophelines [31, 32]. For example, a recent study used SNPs to solve a long-standing controversy about the presence or absence of a species complex in *An. darlingi* by supporting the existence of three genetic clusters (putative species) within this vector in Brazil at a large scale [33]. This study may explain some previously incongruous findings [34, 35], but does little to clarify population structure of *An. darlingi* populations at a fine geographical scale in a heterogeneous landscape such as the Amazon region. Here, we inferred genetic divergence in *An. darlingi* populations at a local scale.

Methods

Mosquito collections

Mosquitoes were collected outdoors (peridomestic, within 10 m of each house) in two rural settlements, Granada and Remansinho in March 2012. Outdoor samples from the urban site of Cruzeiro do Sul were collected in March 2013 (Fig. 1). Collections were performed using human landing catch by the authors

(MC and PERM). All specimens were morphologically identified [36] as *An. darlingi* and stored at -20 °C (Table 1).

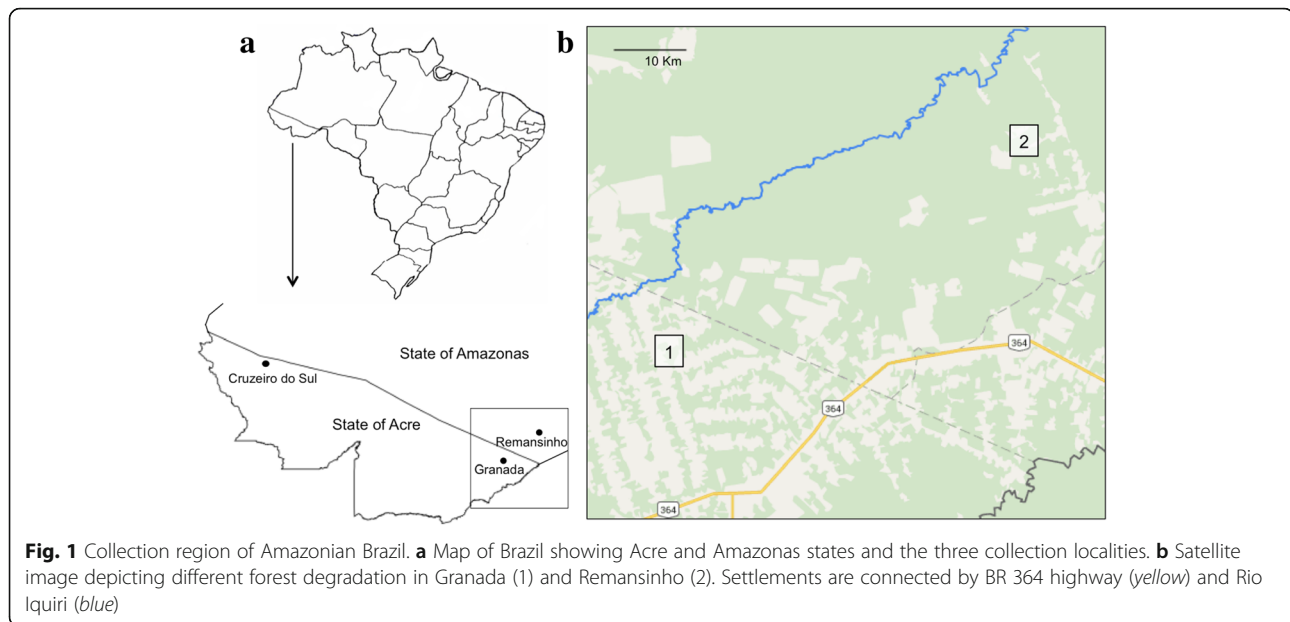
Microsatellite genotyping

DNA was prepared from each mosquito with 5% Chelex solution (BioRad, Hercules, USA). Nine microsatellite loci were genotyped for 175 *An. darlingi* specimens by PCR using fluorescently labeled reverse primers (FAM, NED, or HEX; Applied Biosystems, Foster City, USA) previously described [20, 27]. Amplified fragments were separated by capillary electrophoresis in an ABI 3700 Applied Biosystems and analyzed with GeneMarker software (SoftGenetics, State College, USA). The presence of null alleles was tested in MICRO-CHECKER [37]. Estimates of expected heterozygosity (H_E), allele richness (R_s), and private allele (P) were performed in FSTAT v 2.9.3.2 [38].

SNP genotyping

Double digest restriction associated DNA sequencing (ddRADseq)

DNA from 45 individual *An. darlingi* specimens (see Table 1 for sample sizes in three locations) was extracted using ReliaPrep™ Blood gDNA kit (Promega, Madison, USA) and its concentration was estimated using a Qubit fluorometer (Invitrogen, Carlsbad, USA). The sample size for ddRADseq analysis was based on previous study with *Anopheles darlingi* in Brazil [33]. Double restriction digestion of 200 ng of high quality genomic DNA with *EcoRI-MspI* restriction enzymes was performed in a 40 µl reaction volume and then purified with AMPure XP beads following the manufacturer's protocol. A pair of customized adapters (P1 and P2) were designed including Nextera® Index Primers (Illumina, San Diego USA) complement sequence, to perform the indexing with Nextera® DNA Sample Preparation Kit (Illumina) (Additional file 1: Table S1). The working stock dilution of hybridized adapters P1 (0.3 µM) and P2 (4.8 µM) was ligated to the digested DNA (T4 DNA Ligase, Promega). After another purification with AMPure XP beads, DNA was size selected on an agarose gel to 350–400 bp and purified again. PCR amplification for Nextera® indexing was carried out to generate Illumina sequencing libraries, according to these cycling conditions: an initial denaturation step at 72 °C for 3 min and at 95 °C for 30 s, followed by 16 cycles of 95 °C for 10 s, annealing at 55 °C for 30 s, elongation at 72 °C for 30 s, and a final extension cycle at 72 °C for 5 min, then each PCR product was purified one last time. Samples with distinct multiplexing indices were combined in equimolar ratios to compose a final library for sequencing. The library quantification was made with KAPA library quantification kit in a qPCR reaction. The samples were pooled, normalized and denatured, and finally loaded on the Illumina reagent cartridge. One library was paired-end sequenced in 150-cycles in a Miseq (Genetic Department Facility, Sao Paulo State University).



Stacks v1.31 [39] pipeline was used to identify SNP loci within and between individuals. Briefly, all sequence reads were quality filtered using the default parameters of stacks component *process_radtags*. Then, each individual's sequence reads were aligned to the *An. darlingi* reference genome [40] using Bowtie2 with default parameters [41], and stacks component *ref_map.pl* was used to generate the genotype data (see Additional file 2: Text S1 for parameters used). Stacks was used to generate genotypes from a single SNP position (parameter - *write_single_snp* from stacks component *populations*) for each RAD locus, which passed through a minimum allele sequence depth of 5, as used by Emerson and collaborators [33], that was called in at least 50% of individuals, considering only one population. The last parameter certified no population bias in the SNPs selection.

Statistical and structural analyses

A Bayesian clustering analysis with STRUCTURE [42] was performed assuming the admixture model and assuming correlated allele frequencies among populations. We conducted 20–40 independent runs for each *K* value (ranging from 1–4) using a 100,000 'burn-in' period and 1,000,000 generations. The optimal value of *K* was

inferred using the Evanno method [43] implemented in structureHarvester [44]. Locus-specific and pairwise F_{ST} estimates of genetic diversity, as well as Hardy-Weinberg (HW) equilibrium tests and linkage disequilibrium (LD) between pairs of microsatellite loci were computed using ARLEQUIN 3.5 [45]. The nominal significance level was $\alpha = 0.05$; when multiple tests were performed, the sequential Bonferroni procedure was applied. In addition, as microsatellites are known to give precise, but often downwardly biased estimates of genetic differentiation [46], we include estimates of corrected Hendrick G_{ST} (G''_{ST}) [47], using GenoDive package [48], that standardize the differentiation estimate relative to the maximum differentiation possible for the level of homozygosity observed. Adegnet package [49] in R software [50] was used to perform principal components analysis (PCA) and discriminant analysis of principal components (DAPC).

Results

Genetic diversity and structure of microsatellite data in *An. darlingi*

One hundred and seventy-five specimens from rural settlements Granada and Remansinho, and urban Cruzeiro do Sul in western Amazonian Brazil were genotyped

Table 1 Location, number of specimens, and genetic marker used for microsatellite and ddRADseq analyses (for additional details see Additional file 3: Table S6)

Code	Location	Region	State	Latitude	Longitude	Microsatellites	ddRADseq
GRA	Granada	Acrelandia	Acre	-9.752	-67.071	59	15
REM	Remansinho	Labrea	Amazon	-9.497	-66.582	60	16
CZS	Cruzeiro do Sul	Cruzeiro do Sul	Acre	-7.625	-72.673	56	14

using nine microsatellite markers that were polymorphic in all groups analyzed. Estimates of H_E , F_{is} and allelic richness (R_s) per locality and sampling period are presented in Additional file 1: Table S2. The number of alleles per locus present within a population ranged from 4 to 43. Significant departures from Hardy-Weinberg equilibrium were detected at loci ADC29 and ADC138 in all samples (Additional file 1: Table S2) and these markers were excluded in population structure analysis. The highest values of allelic richness (R_s) and number of private alleles (P) were observed in Remansinho (Additional files 1: Tables S2, S3). Grouping the rural settlements of Remansinho and Granada resulted in 46 private alleles (Additional file 1: Table S3). Estimates of F_{ST} were significant only between Cruzeiro do Sul and each of the two rural settlements Granada and Remansinho (Table 2). In the locus-by-locus analysis, F_{ST} ranged from 0.019 ($P < 0.0001$) to 0.133 ($P < 0.0001$) (Additional file 1: Table S4). STRUCTURE analysis of microsatellite alleles revealed two genetic clusters consisting of Cruzeiro do Sul and Granada + Remansinho (Fig. 2c); nonetheless PCA did not clearly separate the three locations (Fig. 2a). DAPC showed evidence of four genetic clusters, with all clusters represented in all three geographical locations (Fig. 3a, b).

Genetic diversity and structure of ddRADseq data in *An. darlingi*

From 54,616,244 ddRAD tag sequences (NCBI SRA BioProject PRJNA298241), around 46 million sequences passed several levels of quality filtering in the process_radtags program (Stacks v. 1.31 [39], details in Additional file 3: Table S6), and 33.9% (± 2.06 SD) of this set of reads was aligned to the *An. darlingi* genome [40]. An average of 17,401 ($\pm 6,248$ SD) ddRAD loci were genotyped per sample. After filtering, 2185 SNPs were found in at least 50% of all 45 individuals. Pairwise F_{ST} values were significant between Cruzeiro do Sul and both rural settlements, as well as between settlements (Table 2). Remansinho had the highest number of private alleles and polymorphic sites (Additional file 1: Table S5).

STRUCTURE analysis of SNP variation revealed three genetic clusters ($K = 3$), which were assigned to each collection point (Fig. 2d). PCA separated the three

Table 2 Pairwise F_{ST} values from microsatellite and ddRADseq data of *An. darlingi* populations. Lower left values are from microsatellites while upper right values are from ddRADseq data

	Granada	Remansinho	Cruzeiro do Sul
Granada		0.072*	0.181*
Remansinho	0.002		0.118*
Cruzeiro do Sul	0.043*	0.042*	

* $P < 0.001$

populations based on SNP variation (Fig. 2b). DAPC partitioned the genetic variation into three genetic clusters, where each contains a unique collection point (Fig. 3).

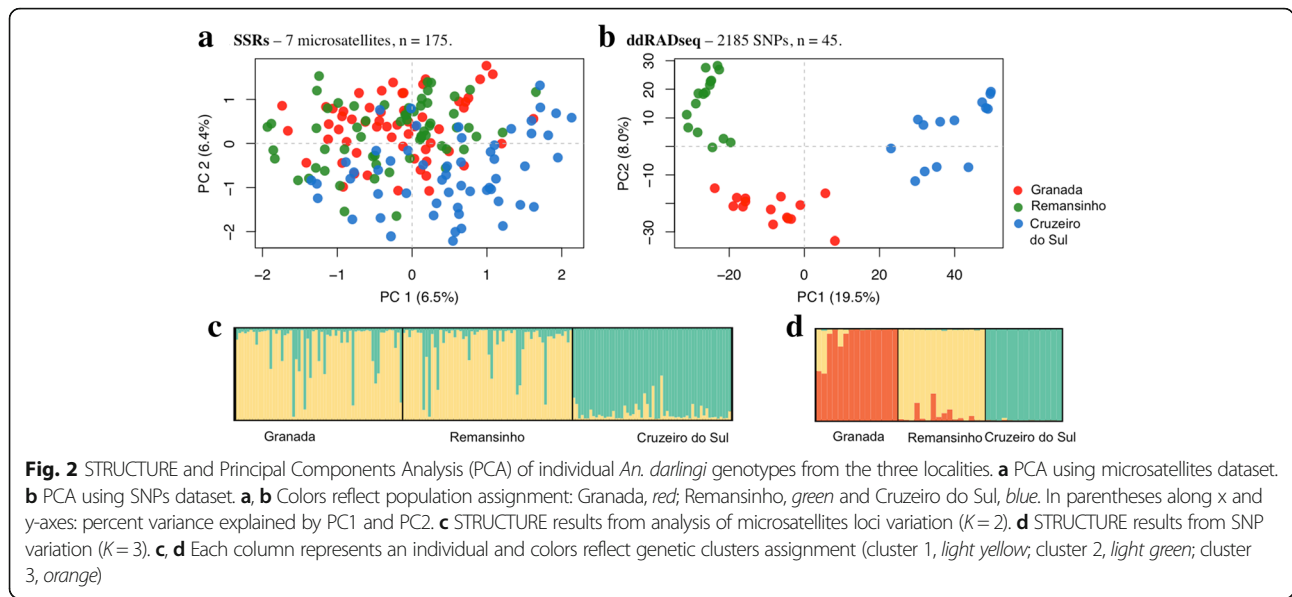
Overall diversity was also calculated using G''_{ST} index. The results also showed a higher level of diversity with SNPs $G''_{ST} = 0.138$, 0.121–0.155 (2.5–97.5% CI) than with microsatellites $G''_{ST} = 0.119$, 0.065–0.219 (2.5–97.5% CI), but the values were not so different as in F_{ST} estimates.

Discussion

The observed population genetic divergence among collection localities was higher with SNPs than with microsatellites markers in both model-based analysis using Bayesian Analysis (BA) with STRUCTURE, F_{ST} estimation, and non-model-based analysis by DAPC. BA assumes a model-clustering method based on allele frequencies at each locus, and probabilistically each individual is assigned to a number of genetically distinct clusters (K) [42]. In the present study, BA revealed two clusters by microsatellites; one essentially characterized the specimens from urban Cruzeiro do Sul and the other categorized specimens from both rural settlements (Fig. 2c). However, the optimal number of clusters based on BA analysis of SNPs was three, and each cluster defined only one location (Fig. 2d). It is also worth noting that the admixture between clusters was lower in the SNP analysis, highlighting the discrimination among genetic clusters. Both marker types showed higher F_{ST} estimates between Cruzeiro do Sul and the two rural settlements than between the two rural settlements, however the estimates based on SNPs were more than 4-fold higher than those based on microsatellites between Cruzeiro do Sul and rural settlements, and 35-fold higher between Granada and Remansinho (Table 2).

High-throughput methods using next generation sequencing that analyze a subsample of the genome, such as ddRADseq, have two major advantages compared to microsatellites, the need of smaller sample sizes and also no need of prior knowledge of the genomic sequence [30]. In the present study, the number of SNPs generated and used was much higher than the number of microsatellites, which could contribute to increased statistical power in the analysis. Nevertheless, other studies have shown the efficiency of SNP genotyping even when a small number of SNPs are used [51–53]. SNP analyses have corroborated microsatellite-based findings, and have presented superior accuracy, robustness and recovered finer population structure when compared to microsatellite analysis [53, 54].

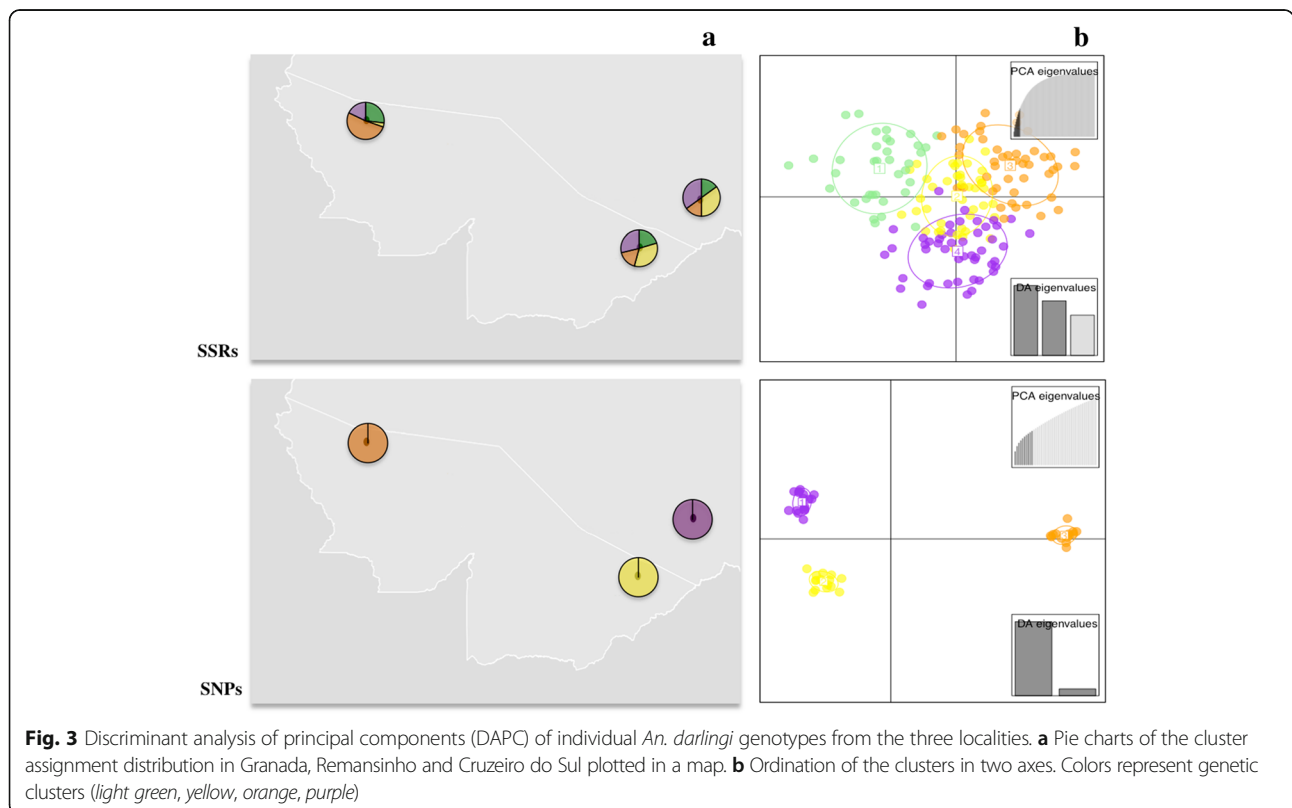
PCA of the SNP data separates individuals originating in Cruzeiro do Sul from those in Granada and Remansinho (separated by ~ 700 km) along the first



principal component that explains 19.5% of the variation (Fig. 2b). Nonetheless, at a finer scale (~60 km apart), individuals from Granada and those from Remansinho were also separated along principal component 2, which accounted for 8% of the total variation. This was validated by DAPC analysis, which found three distinct clusters, uniquely identifying individuals to their appropriate

geographical population. No clear separation of the populations was reflected in the PCA for the microsatellite data (Fig. 2a), and DAPC revealed four distinct genetic clusters, equally partitioned among the three geographical locations.

Rural settlements are in constant flux due to human interventions such as agricultural development, forest degradation, and increased and often mobile human



populations [22, 55]. Such anthropogenic environmental changes along with seasonal climate variation, such as temperature, rainfall and humidity, influence the survival, density and distribution of mosquitoes [16, 20]. The highest risk of malaria transmission in such settlements, common in Amazonian Brazil and Peru [7, 22, 56], is typically in the newest settlements where migrants have no previous exposure to *Plasmodium* and little effective shelter from bites of infected *An. darlingi*. Proximity of residences to potential mosquito breeding-sites is also associated with the likelihood of becoming infected with *Plasmodium* [57–59]. The two rural settlements of the present study, Granada and Remansinho, have experienced anthropogenic landscape modification to different degrees because of their relative ages; Granada was initiated in 1982 and Remansinho 25 years later [55, 60]. Regardless of the genetic marker used (microsatellites or SNPs) the two rural settlements samples presented higher genetic diversity than the sample from the urban area. Even between the rural settlements, genetic diversity was highest in mosquitoes from the newer settlement, Remansinho, which has a greater proportion of intact forest compared to the older settlement of Granada [60]. Our findings in *An. darlingi* support the hypothesis that deforestation may be associated with a loss of genetic diversity [61, 62]. Deforestation enhanced survivorship, reproductive fitness and increased population growth potential of *An. gambiae* in the western Kenyan highlands [63, 64]. A similar scenario may be occurring in *An. darlingi* in settlements that are at different temporal points in the frontier malaria model. Once roads have been built for settlements, deforestation to clear space for housing and crop planting is a priority.

Small-area interventions may be an effective approach for malaria control and elimination in the neotropics and globally, once transmission pockets have been identified and characterized [65–67]. Each locality has peculiar environmental characteristics and thus, it might have different anopheline population genetic backgrounds, which may lead to differences in vector capacity and competitiveness. For an intervention to be successful, it is essential to be able to precisely identify genetic differences between vector populations and subpopulations at a microgeographical scale.

Conclusion

In this study, we provide evidence that the detection of microgeographical population structure at a fine scale is only robust when we apply high-resolution molecular typing techniques, since conventional approaches based on microsatellite markers may underestimate overall genetic distances in closely related vector populations. In

our view the application of ddRADtag sequencing for genetic analysis of mosquito populations represents a suitable molecular tool to further elucidate vector population dynamics in malaria endemic areas.

Additional files

Additional file 1: Table S1. Double digest RADseq primer and adapters sequences for *An. darlingi*. **Table S2.** Estimates of R_s , H_e and F_{IS} of *An. darlingi* microsatellite loci in three Brazilian populations. **Table S3.** Estimates of private alleles in *An. darlingi* using microsatellite loci. **Table S4.** Locus-by-locus analysis of *An. darlingi* microsatellite loci. **Table S5.** Summary of ddRADseq dataset containing all positions (variant and fixed) from the three *An. darlingi* populations. (DOCX 27 kb)

Additional file 2: Text S1. Bash script with commands used to run the Stacks pipeline and STRUCTURE analysis. (TXT 4 kb)

Additional file 3: Table S6. Per-individual *An. darlingi* detail of the number of sequence reads and unique stacks genotyped. (XLSX 18 kb)

Abbreviations

BA: Bayesian analysis; DAPC: Discriminant analysis of principal components; ddRADseq: double digest restriction associated DNA sequencing; LD: Linkage disequilibrium; PCA: Principal components analysis; SNP: Single nucleotide polymorphism

Acknowledgments

We are very grateful to Dr. Marinete Póvoa and her entomology team (Evandro Chagas Institute, Belém, Pará) who provided the specimens of *An. darlingi* from Cruzeiro do Sul.

Funding

MC was supported by FAPESP. JMV received funding from International Centers for Excellence in Malaria Research (grant U19AI089681). KJE received funding St. Mary's College of Maryland. PEMR has a CNPq fellowship.

Availability of data and materials

The datasets generated during and analyzed during the current study are available in the Sequence Read Archive (SRA), BioProject PRJNA298241 (<http://www.ncbi.nlm.nih.gov/bioproject/PRJNA298241/>)

Authors' contributions

PEMR and MC designed the field and laboratory work; PEMR, MC and DPA performed the laboratory research; PEMR, MC, DPA and KJE analyzed data. All authors actively contributed to the interpretation of the findings and development of the final manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Author details

¹Biotechnology Institute (IBTEC) & Biosciences Institute at Botucatu (IBB), Sao Paulo State University (UNESP), Sao Paulo, Brazil. ²Department of Biomedical Sciences, School of Public Health, University at Albany (State University of New York), Albany, NY, USA. ³New York State Department of Health, Wadsworth Center, Albany, NY, USA. ⁴Division of Infectious Diseases, Department of Medicine, University of California, La Jolla, San Diego, CA, USA. ⁵Instituto de Medicina Tropical "Alexander von Humboldt," and Departamento de Ciencias Celulares y Moleculares, Laboratorio de Investigación y Desarrollo, Universidad Peruana Cayetano Heredia, Lima, Peru. ⁶Biology Department, St. Mary's College of Maryland, St. Mary's City, MD, USA.

Received: 24 September 2016 Accepted: 31 January 2017

Published online: 13 February 2017

References

- WHO. World Malaria Report 2015. In: World Malaria Report. Switzerland: World Health Organization; 2015. p. 280.
- Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI. The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature*. 2005;434(7030):214–7.
- Moreno M, Saavedra MP, Bickersmith SA, Lainhart W, Tong C, Alava F, et al. Implications for changes in *Anopheles darlingi* biting behaviour in three communities in the peri-Iquitos region of Amazonian Peru. *Malar J*. 2015;14:290.
- Sinka ME, Bangs MJ, Manguin S, Rubio-Palis Y, Chareonviriyaphap T, Coetzee M, et al. A global map of dominant malaria vectors. *Parasit Vectors*. 2012;5:69.
- Zimmerman RH, Galardo AK, Lounibos LP, Arruda M, Wirtz R. Bloodmeal hosts of *Anopheles* species (Diptera: Culicidae) in a malaria-endemic area of the Brazilian Amazon. *J Med Entomol*. 2006;43(5):947–56.
- Lainhart W, Bickersmith SA, Nadler KJ, Moreno M, Saavedra MP, Chu VM, et al. Evidence for temporal population replacement and the signature of ecological adaptation in a major Neotropical malaria vector in Amazonian Peru. *Malar J*. 2015;14:375.
- Vittor AY, Gilman RH, Tielsch J, Glass G, Shields T, Lozano WS, et al. The effect of deforestation on the human-biting rate of *Anopheles darlingi*, the primary vector of *Falciparum* malaria in the Peruvian Amazon. *Am J Trop Med Hyg*. 2006;74(1):3–11.
- Hiwat H, Bretas G. Ecology of *Anopheles darlingi* Root with respect to vector importance: a review. *Parasit Vectors*. 2011;4:177.
- Jimenez IP, Jimenez JE, Brochero H. Preliminary biological studies on larvae and adult *Anopheles* mosquitoes (Diptera: Culicidae) in Miraflores, a malaria endemic locality in Guaviare department, Amazonian Colombia. *J Med Entomol*. 2014;51(5):1002–9.
- Afrane YA, Githeko AK, Yan G. The ecology of *Anopheles* mosquitoes under climate change: case studies from the effects of deforestation in East African highlands. *Ann N Y Acad Sci*. 2012;1249:204–10.
- Saxena R, Nagpal BN, Singh VP, Srivastava A, Dev V, Sharma MC, et al. Impact of deforestation on known malaria vectors in Sonitpur district of Assam, India. *J Vector Borne Dis*. 2014;51(3):211–5.
- Parham PE, Hughes DA. Climate influences on the cost-effectiveness of vector-based interventions against malaria in elimination scenarios. *Phil Trans R Soc B*. 2015;370:20130557.
- Rottschaefer SM, Riehle MM, Coulibaly B, Sacko M, Niare O, Morlais I, et al. Exceptional diversity, maintenance of polymorphism, and recent directional selection on the APL1 malaria resistance genes of *Anopheles gambiae*. *PLoS Biol*. 2011;9(3):e1000600.
- Vittor AY, Pan W, Gilman RH, Tielsch J, Glass G, Shields T, et al. Linking deforestation to malaria in the Amazon: characterization of the breeding habitat of the principal malaria vector, *Anopheles darlingi*. *Am J Trop Med Hyg*. 2009;81(1):5–12.
- Walsh JF, Molyneux DH, Birley MH. Deforestation: effects on vector-borne disease. *Parasitology*. 1993;106(Suppl):S55–75.
- Yasuoka J, Levins R. Impact of deforestation and agricultural development on anopheline ecology and malaria epidemiology. *Am J Trop Med Hyg*. 2007;76(3):450–60.
- Malafrente RS, Marrelli MT, Marinotti O. Analysis of ITS2 DNA sequences from Brazilian *Anopheles darlingi* (Diptera: Culicidae). *J Med Entomol*. 1999;36(5):631–4.
- Santos LM, Gama RA, Eiras AE, Fonseca CG. Genetic differences based on AFLP markers in the mosquito species *Anopheles darlingi* collected in versus near houses in the region of Porto Velho, RO, Brazil. *Genet Mol Res*. 2010;9(4):2254–62.
- Angella AF, Gil LH, Silva LH, Ribolla PE. Population structure of the malaria vector *Anopheles darlingi* in Rondonia, Brazilian Amazon, based on mitochondrial DNA. *Mem Inst Oswaldo Cruz*. 2007;102(8):953–8.
- Angella AF, Salgueiro P, Gil LH, Vicente JL, Pinto J, Ribolla PE. Seasonal genetic partitioning in the Neotropical malaria vector, *Anopheles darlingi*. *Malar J*. 2014;13:203.
- da Silva-Nunes M, Codeco CT, Malafrente RS, da Silva NS, Juncansen C, Muniz PT, Ferreira MU. Malaria on the Amazonian frontier: transmission dynamics, risk factors, spatial distribution, and prospects for control. *Am J Trop Med Hyg*. 2008;79(4):624–35.
- de Castro MC, Monte-Mor RL, Sawyer DO, Singer BH. Malaria risk on the Amazon frontier. *Proc Natl Acad Sci U S A*. 2006;103(7):2452–7.
- Kamdem C, Tene Fossog B, Simard F, Etoua J, Ndo C, Kengne P, et al. Anthropogenic habitat disturbance and ecological divergence between incipient species of the malaria mosquito *Anopheles gambiae*. *PLoS One*. 2012;7(6):e39453.
- Lanzaro GC, Toure YT, Carnahan J, Zheng L, Dolo G, Traore S, et al. Complexities in the genetic structure of *Anopheles gambiae* populations in west Africa as revealed by microsatellite DNA analysis. *Proc Natl Acad Sci U S A*. 1998;95(24):14260–5.
- Onyabe DY, Conn JE. Genetic differentiation of the malaria vector *Anopheles gambiae* across Nigeria suggests that selection limits gene flow. *Heredity*. 2001;87(Pt 6):647–58.
- Slotman MA, Tripet F, Cornel AJ, Meneses CR, Lee Y, Reimer LJ, et al. Evidence for subdivision within the M molecular form of *Anopheles gambiae*. *Mol Ecol*. 2007;16(3):639–49.
- Conn JE, Vineis JH, Bollback JP, Onyabe DY, Wilkerson RC, Povoaa MM. Population structure of the malaria vector *Anopheles darlingi* in a malaria-endemic region of eastern Amazonian Brazil. *Am J Trop Med Hyg*. 2006;74(5):798–806.
- Scarpassa VM, Conn JE. Population genetic structure of the major malaria vector *Anopheles darlingi* (Diptera: Culicidae) from the Brazilian Amazon, using microsatellite markers. *Mem Inst Oswaldo Cruz*. 2007;102(3):319–27.
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, et al. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*. 2008;3(10):e3376.
- Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*. 2012;7(5):e37135.
- O'Loughlin SM, Magesa S, Mbogo C, Moshaf F, Midega J, Lomas S, Burt A. Genomic analyses of three malaria vectors reveals extensive shared polymorphism but contrasting population histories. *Mol Biol Evol*. 2014;31(4):889–902.
- Fouet C, Kamdem C, Gamez S, White BJ. Extensive genetic diversity among populations of the malaria mosquito *Anopheles moucheti* revealed by population genomics. *Infect Genet Evo*. 2017;48:27–33.
- Emerson KJ, Conn JE, Berge ES, Randel MA, Sallum MA. Brazilian *Anopheles darlingi* Root (Diptera: Culicidae) Clusters by Major Biogeographical Region. *PLoS One*. 2015;10(7):e0130773.
- Manguin S, Wilkerson RC, Conn JE, Rubio-Palis Y, Danoff-Burg JA, Roberts DR. Population structure of the primary malaria vector in South America, *Anopheles darlingi*, using isozyme, random amplified polymorphic DNA, internal transcribed spacer 2, and morphologic markers. *Am J Trop Med Hyg*. 1999;60(3):364–76.
- Pedro PM, Sallum MAM. Spatial expansion and population structure of the neotropical malaria vector, *Anopheles darlingi* (Diptera: Culicidae). *Biol J Linn Soc*. 2009;97(4):854–66.
- Consoli RL-d-O, R. Principais mosquitos de importância sanitária no Brasil. Rio de Janeiro: Fiocruz; 1994.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes*. 2004;4(3):535–8.
- Goudet J. FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered*. 1995;86(6):2.
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. *Mol Ecol*. 2013;22(11):3124–40.
- Marinotti O, Cerqueira GC, de Almeida LG, Ferro MI, Loreto EL, Zaha A, et al. The genome of *Anopheles darlingi*, the main neotropical malaria vector. *Nucleic Acids Res*. 2013;41(15):7387–400.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945–59.
- Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14(8):2611–20.
- Earl DA, Vonholdt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour*. 2012;4(2):359–61.
- Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 2010;10(3):564–7.
- Balloux F, Lugon-Moulin N. The estimation of population differentiation with microsatellite markers. *Mol Ecol*. 2002;11(2):155–65.

47. Hedrick PW. A standardized genetic differentiation measure. *Evolution*. 2005; 59(8):1633–8.
48. Meirmans PG, Tienderen PHV. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes*. 2004;4:3.
49. Jombart T, Ahmed I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. 2011;27(21):3070–1.
50. R_Core_Team: R. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
51. Coates BS, Sumerford DV, Miller NJ, Kim KS, Sappington TW, Siegfried BD, Lewis LC. Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. *J Hered*. 2009;100(5):556–64.
52. Ryyanen HJ, Tonteri A, Vasemagi A, Primmer CR. A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). *J Hered*. 2007;98(7):692–704.
53. Telfer EJ, Stovold GT, Li Y, Silva-Junior OB, Grattapaglia DG, Dungey HS. Parentage reconstruction in *Eucalyptus nitens* using SNPs and microsatellite markers: a comparative analysis of marker data power and robustness. *PLoS One*. 2015;10(7):e0130601.
54. Jeffries DL, Copp GH, Lawson Handley L, Olsen KH, Sayer CD, Hanfling B. Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Mol Ecol*. 2016;25(13):2997–3018.
55. Moutinho PR, Gil LH, Cruz RB, Ribolla PE. Population dynamics, structure and behavior of *Anopheles darlingi* in a rural settlement in the Amazon rainforest of Acre, Brazil. *Malar J*. 2011;10:174.
56. de Castro MC, Sawyer DO, Singer BH. Spatial patterns of malaria in the Amazon: implications for surveillance and targeted interventions. *Health Place*. 2007;13(2):368–80.
57. Clark TD, Greenhouse B, Njama-Meya D, Nzarubara B, Maiteki-Sebuguzi C, Staedke SG, et al. Factors determining the heterogeneity of malaria incidence in children in Kampala, Uganda. *J Infect Dis*. 2008;198(3):393–400.
58. de Barros FS, Honorio NA, Arruda ME. Temporal and spatial distribution of malaria within an agricultural settlement of the Brazilian Amazon. *J Vector Ecol*. 2011;36(1):159–69.
59. Staedke SG, Nottingham EW, Cox J, Kanya MR, Rosenthal PJ, Dorsey G. Short report: proximity to mosquito breeding sites as a risk factor for clinical malaria episodes in an urban cohort of Ugandan children. *Am J Trop Med Hyg*. 2003;69(3):244–6.
60. Barbosa S, Gozde AB, Lima NF, Batista CL, Bastos Mda S, Nicolete VC, et al. Epidemiology of disappearing *Plasmodium vivax* malaria: a case study in rural Amazonia. *PLoS Negl Trop Dis*. 2014;8(8):e3109.
61. Gardner TA, Barlow J, Chazdon R, Ewers RM, Harvey CA, Peres CA, Sodhi NS. Prospects for tropical forest biodiversity in a human-modified world. *Ecol Lett*. 2009;12(6):561–82.
62. Vieira IC, Toledo PM, Silva JM, Higuchi H. Deforestation and threats to the biodiversity of Amazonia. *Braz J Biol*. 2008;68(4 Suppl):949–56.
63. Afrane YA, Lawson BW, Githeko AK, Yan G. Effects of microclimatic changes caused by land use and land cover on duration of gonotrophic cycles of *Anopheles gambiae* (Diptera: Culicidae) in western Kenya highlands. *J Med Entomol*. 2005;42(6):974–80.
64. Afrane YA, Zhou G, Lawson BW, Githeko AK, Yan G. Effects of microclimatic changes caused by deforestation on the survivorship and reproductive fitness of *Anopheles gambiae* in western Kenya highlands. *Am J Trop Med Hyg*. 2006;74(5):772–8.
65. Brooker S, Clarke S, Njagi JK, Polack S, Mugo B, Estambale B, et al. Spatial clustering of malaria and associated risk factors during an epidemic in a highland area of western Kenya. *Trop Med Int Health*. 2004;9(7):757–66.
66. Coulibaly D, Rebaudet S, Travassos M, Tolo Y, Laurens M, Kone AK, et al. Spatio-temporal analysis of malaria within a transmission season in Bandiagara, Mali. *Malar J*. 2013;12:82.
67. Zacarias OP, Andersson M. Spatial and temporal patterns of malaria incidence in Mozambique. *Malar J*. 2011;10:189.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

