Author's personal copy

Molecular Phylogenetics and Evolution 57 (2010) 1341–1346

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Short Communication

Late Pleistocene environmental changes lead to unstable demography and population divergence of *Anopheles albimanus* in the northern Neotropics

Jose R. Loaiza [a,c,d,*], Marilyn E. Scott [b], Eldredge Bermingham [c], Oris I. Sanjur [c], Richard Wilkerson [e], Jose Rovira [f], Lina A. Gutiérrez [g], Margarita M. Correa [g], Mario J. Grijalva [h,i], Lotty Birnberg [h], Sara Bickersmith [j], Jan E. Conn [j,k]

[a] *Department of Natural Resource Sciences, McGill University, 21,111 Lakeshore Road, Ste. Anne de Bellevue, Que., Canada H9X 3V9*
[b] *Institute of Parasitology, McGill University, Canada*
[c] *Smithsonian Tropical Research Institute, Balboa Ancon, Unit 0948, Panama*
[d] *Programa Centroamericano de Maestría en Entomología, Vicerrectoría de Investigación y Postgrado, Universidad de Panamá, Panama*
[e] *Smithsonian Institution, Washington, DC, USA*
[f] *Instituto Conmemorativo Gorgas de Estudios de la Salud, Ciudad de Panamá, Panama*
[g] *Grupo de Microbiología Molecular, Escuela de Microbiología, Universidad de Antioquia, Medellín, Colombia*
[h] *Tropical Disease Institute, College of Osteopathic Medicine, Ohio University, USA*
[i] *Center for Infectious Disease Research, School of Biological Sciences, Pontifical Catholic University of Ecuador, Quito, Ecuador*
[j] *Wadsworth Center, Griffin Laboratory, New York State Department of Health, Slingerlands, NY, USA*
[k] *Department of Biomedical Sciences, School of Public Health, State University of New York-Albany, Albany, NY, USA*

ARTICLE INFO

Article history:
Received 23 April 2010
Revised 8 September 2010
Accepted 18 September 2010
Available online 1 October 2010

Keywords:
Anopheles albimanus
COI
White
ITS2
Pleistocene environmental changes
Geographic fragmentation
Population expansion

ABSTRACT

We investigated the historical demography of *Anopheles albimanus* using mosquitoes from five countries and three different DNA regions, the mitochondrial cytochrome oxidase subunit I gene (*COI*), the single copy nuclear *white* gene and the ribosomal internal transcribed spacer two (ITS2). All the molecular markers supported the taxonomic status of a single species of *An. albimanus*. Furthermore, agreement between the *COI* and the *white* genes suggested a scenario of Pleistocene geographic fragmentation (i.e., population contraction) and subsequent range expansion across southern Central America.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Several phylogeographic studies have depicted a complex biogeographical history across the Isthmus of Panama and northern Colombia where significant genetic differentiation is seen at short distances, and waves of colonization, extinctions and re-invasions appear to be the main forces shaping the distribution of genetic diversity. Most of these studies have placed diversification events in a time frame earlier than the complete formation of the Isthmus of Panama (3–3.5 mya) (Bermingham and Martin, 1998; Zeh et al., 2003; Weigt et al., 2005; Miller et al., 2008), and therefore, much less is known about the impact of more recent environmental changes across this region during the Pleistocene (0.01–1.8 mya). Pleistocene climatic oscillations have had a profound effect on the genetic structure of vectors of pathogens, such as *Anopheles* mosquitoes, worldwide. For instance, in Southeast Asia, *Anopheles dirus* and *Anopheles baimaii*, two important malaria vectors, have expanded demographically due to changes in forest structure and climatic conditions in the Pleistocene, 172,202 and 234,443 years ago (ya), respectively (O'Loughlin et al., 2008). Likewise, *Anopheles darlingi*, an important malaria vector in most of the Neotropics, has undergone Pleistocene population expansion in parts of South America likely due to changes in climatic conditions leading to forest fragmentation and refugia isolation, although in a much more recent time frame when compared with the Asian mosquitoes (25,312 ya) (Mirabello and Conn, 2006a,b). These and other studies

* Corresponding author at: Department of Natural Resource Sciences, McGill University, 21,111 Lakeshore Road, Ste. Anne de Bellevue, Que., Canada H9X 3V9.
E-mail addresses: jose.loaiza@mail.mcgill.ca (J.R. Loaiza), marilyn.scott@mcgill.ca (M.E. Scott), bermingham@si.edu (E. Bermingham), sanjuro@si.edu (O.I. Sanjur), wilkersonr@si.edu (R. Wilkerson), jrovira@gorgas.gob.pa (J. Rovira), liangutibui@gmail.com (L.A. Gutiérrez), mcorrea@quimbaya.udea.edu.co (M.M. Correa), grijalva@ohiou.edu (M.J. Grijalva), lottycita@yahoo.com (L. Birnberg), sab19@health.state.ny.us (S. Bickersmith), jconn@wadsworth.org (J.E. Conn).

1055-7903/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2010.09.016

Author's personal copy

have uncovered historical changes in demography that have ultimately shaped the geographical distribution and relative abundance of *Anopheles* species, and perhaps their capabilities to vector malarial parasites at a regional scale (Walton et al., 2000; Mirabello and Conn, 2008). Similar studies are scarce in southern Central America, despite the acknowledged geological and environmental complexity of this region during the Pleistocene (Crawford, 2003; Cortes-Ortíz et al., 2003; Nettel et al., 2008; Miller et al., 2008).

Anopheles albimanus is a primary malaria vector throughout the northern Neotropics, yet its population history with respect to Pleistocene environmental changes has not been examined in depth. Furthermore, no accurate information exists on the geographical origin of *An. albimanus*, and its likely initial colonization path throughout the Americas. Recent research on *An. albimanus* uncovered considerable geographic structuring across the Isthmus of Panama and northern Colombia (Gutiérrez et al., 2009; Loaiza et al., 2010). Several divergent groups of mtDNA *COI* haplotypes were found co-occurring across this region, and were hypothesized to be the result of late Pleistocene geographic fragmentation and multiple re-introductions via demographic expansion. The population expansion, in central-eastern Panama and the Caribbean coast of Colombia, was believed to be due to climatic oscillation around 22,000 ya, and thus *An. albimanus* is not at mutation-drift equilibrium regionally (Gutiérrez et al., 2009; Loaiza et al., 2010). The co-occurrence of several maternal lineages of *An. albimanus* across Panama and northern Colombia may have important implications for malaria control, especially if they are differentially involved in malaria transmission. Nevertheless, three hypotheses remain to be tested: (1) Is the population divergence in recent studies of *An. albimanus* due to cryptic speciation?; (2) Is this divergence associated with Pleistocene environmental changes?; and (3) Is the population expansion of *An. albimanus* supported by a single copy nuclear gene?. To answer these questions, we combined the data from Gutiérrez et al. (2009) and Loaiza et al. (2010), with 123 additional mtDNA *COI* sequences from Nicaragua, eastern Panama and Ecuador, expanding the sampling scheme and augmenting the statistical power of the analyses. In addition, we sequenced a subset of individuals for the single copy nuclear *white* gene; because the *white* locus mutates more slowly than the *COI* gene, we used it to infer the initial colonization path of *An. albimanus* across southern Central America. Finally, we looked for fixed substitutional changes and length differences in the ribosomal DNA ITS2 marker as this may indicate the existence of cryptic species or restricted gene flow in *An. albimanus* (Collins and Paskewitz, 1996).

2. Materials and methods

2.1. Sample processing and laboratory procedures

Information on mosquito sampling, collection sites, laboratory procedures and GenBank accession numbers for the *COI* gene can be found in Gutiérrez et al. (2009) and Loaiza et al. (2010). Additional samples were obtained either as adults or in larval collections. The latter were reared to adulthood and processed as in Loaiza et al. (2010). A total of 612 sequences of the *COI*, all shortened to 776 bp, were analyzed in this study (new *COI* sequence submission GenBank accession Nos. HM030881–HM030907). We sequenced the *white* gene for 175 individuals (HM042289–HM042297) and the ITS2 for 173 (HM042298–HM042301) from a subset of randomly selected samples (Table 1). Information on PCR-amplification conditions, and sequencing reactions for the ITS2 region and the *white* gene can be found in Linton et al. (2002) and Mirabello and Conn (2008), respectively.

2.2. Neutral expectation and genetic diversity

The program DNASP v4.50.02 (Rozas et al., 2003) was used to calculate Tajima's *D* to determine whether or not the *COI* and the *white* gene sequences conformed to neutral expectations. Because the *white* gene is nuclear we allowed for recombination while testing for significant deviation from neutrality, thus accounting for overestimation of directional selection. We computed basic sequence statistics for all the markers in ARLEQUIN v3.11 (Excoffier et al., 2005) (Table 1).

2.3. Population divergence and historical demography

2.3.1. COI

We employed the spatial analysis of molecular variance (SAMOVA v1.0) (Dupanloup et al., 2002) to define aggregates of collection sites that are geographically homogeneous, but genetically differentiated from other similar aggregates. We ran SAMOVA from $K = 2$ to 30 and implemented 10,000 independent simulated annealing steps each starting from 100 random sets of initial conditions. The Mantel analysis was used to test for the isolation by

Table 1
Summary of diversity measures for *Anopheles albimanus*.

Gene marker	*N*	No. of *COI* haplotypes	*K*	$H_d \pm SD$	$\pi \pm SD$
COI	612	191 (131)	11.23	0.91 ± 0.008	0.014 ± 0.0003
NCRWP	135	55 (30)	6.76	0.95 ± 0.011	0.008 ± 0.0003
CEPCO	309	93 (65)	3.41	0.79 ± 0.024	0.004 ± 0.0004
PCOLE	168	53 (36)	7.42	0.94 ± 0.009	0.009 ± 0.0005
		White alleles			
White	175	I – XVIII	0.66	0.50 ± 0.001	0.0008 ± 0.0005
NCRWP	56	**I**, **II**, **III**, **IV**, **V**, XIII, XVIII	0.57	0.39 ± 0.001	0.0007 ± 0.0001
CEPCO	61	**I**, **II**, **III,** X, XI, XII, **XIII**, XIV, XV, XVI, XVII,	0.12	0.12 ± 0.003	0.0001 ± 0.0002
PCOLE	58	**I**, **II**, **III**, **IV**, VI, VII, VIII, IX	1.06	0.73 ± 0.004	0.0014 ± 0.0002
		ITS2 variants			
ITS2	173	a, b	0.52	0.54 ± 0.00	0.0011 ± 0.00
NCRWP	56	**a** – *2*; 17	0.42	0.33 ± 0.00	0.0014 ± 0.00
CEPCO	61	**a** – *2*; 21	0.11	0.11 ± 0.00	0.0003 ± 0.00
PCOLE	56	**a** – *2*, b – 3; 0	0.03	0.04 ± 0.00	0.0001 ± 0.00

N = number of individuals of *An. albimanus* sequenced per gene marker and population deme as define by SAMOVA (see Fig. 1). The number in parentheses indicates the number of *COI* singletons. Bold letters are shared *white* alleles or ITS2 variants and plain letters are unique in that population deme. (a) and (b) are insertions in positions 281–283 (ATG) and 291–292 (AG) of the ITS2; these are followed by their frequencies shown by the number in italics. Underlined numbers are the number of individuals with multiple gene copies of the ITS2. *K* is the average number of nucleotide differences among sequences. H_d, haplotype diversity ± standard deviation; π, nucleotide diversity ± standard deviation.

Author's personal copy

distance (IBD) pattern, using a pairwise matrix of linearized genetic distances, estimated by Φ_{ST}, and the natural log-transformed geographic distance. Because *An. albimanus* has a coastal distribution we used the shortest geographic distances among populations along the shore instead of straight-line distances. The significance of the Mantel test was determined by $n = 10{,}000$ permutations using the IBD web service v3.15 (http://ibdws.sdsu.edu; Jensen et al., 2005). Genetic structure within population demes as defined by SAMOVA was further explored using the Mantel test as explained above. The program MDIV was used to distinguish between (1) a scenario of past isolation with no subsequent migration from (2) a scenario with no isolation, but contemporary gene flow or secondary contact among populations using MDIV@BioHPC (http://cbsuapps.tc.cornell.edu/mdiv.aspx; Nielsen and Wakeley, 2001). We ran MDIV under the HKY model, each simulation 6×10^6 times and assumed a 10% burn-in period with priors for Max $T = 5$–10, and Max $M = 5$–12. Estimates of θ, M and T were taken from the highest posterior probability. The net divergence D_A (Nei, 1987), the between population distance minus the within population distance, was calculated using the Tamura and Nei model in MEGA v4.0 (Tamura et al., 2007). To assess the long-term stability of *An. albimanus* we used the mismatch distribution and the raggedness (r) statistic (Rogers and Harpending, 1992), both calculated in ARLEQUIN v3.11 (Excoffier et al., 2005). Neutrality tests of Fu's F_S (1997) and R_2 (Ramos-Onsins and Rozas, 2002) were obtained in DNASP v4.50.02 (Rozas et al., 2003), and 10,000 coalescence simulations assessed significance. Dates of population expansion were estimated with the formula $T = \tau/2u$ (Rogers and Harpending, 1992) using 10 generations per year (Walton et al., 2000) and the mutation rate of *Drosophila* estimated at 1.2×10^{-8} substitutions per site per year (Powell et al., 1986). We examined changes in the effective population size through time implementing the Bayesian skyline plot (BSKP). The software BEAST v.1.4.2. (Drummond and Rambaut, 2007) was used to generate the BSKP with the SRD06 model of sequence evolution as suggested when analyzing mitochondrial protein coding genes. The analysis was implemented under a relaxed clock with rate for each branch drawn from a log normal distribution. The Markov Chain Monte Carlo algorithm was iterated for 10×10^7generations with a burn-in of 2×10^4 generations. We assessed the genealogical relationship among *COI* haplotypes using the statistical parsimony (SP), and the median-joining (MJ) network methods. These analyses were implemented in the programs TCS v1.12 (Clement et al., 2000) and NETWORK v4.2.0.1 (http://www.fluxus-engineering; Bandelt et al., 1999), respectively, and *Anopheles triannulatus*, in the same subgenus *Nyssorhynchus*, was used as the outgroup in the MJ network to provide directionality.

2.3.2. White

For the *white* gene data set we computed the SP network (Clement et al., 2000), the neutrality tests Fu's F_S (1997) and R_2 (Ramos-Onsins and Rozas, 2002), the mismatch distribution (Rogers and Harpending, 1992) and the BSKP (Drummond and Rambaut, 2007). We used the *Drosophila* substitution rate of 0.004–0.008 per site per million years (Langley and Aquadro, 1987; Miyashita and Langley, 1988), 10 generations per year and 759 bp of the *white* gene to calculate time of population growth or decline in the BSKP.

3. Results

3.1. Neutral expectation and genetic diversity

In total, 191 *COI* haplotypes and eighteen *white* alleles were recovered in this study (Table 1 and Fig. 1). Both markers conformed to neutral expectations, as estimates of Tajima's D were not significantly different from zero (Tajima's D, *COI* = −1.38, $P > 0.05$ and *white* $D = -0.59$, $P > 0.05$). Two short insertions were found in the ITS2 region of *An. albimanus*. These insertions were located in positions 281–283 (ATG) and 291–292 (AG), respectively and were found in 9 (5.2%) of 173 individuals. Furthermore, one base pair ambiguity in position 234 (i.e., multiple copies of the ITS2) separated 38 (21.9%) individuals from Nicaragua (4), Costa Rica (13) and Panama (21) from the remaining 135, all of which had a single peak for that nucleotide and were distributed over the entire geographic area (Table 1).

3.2. Population divergence and historical demography

In agreement with previous studies the *COI* showed considerable population structure in *An. albimanus*. SAMOVA defined three population demes with the highest among group variation at 64.75% ($F_{CT} = 0.6475$, $P > 0.0001$). These population demes were clearly segregated geographically indicating either the existence of specific barriers to gene flow or separate demographic origins. The Mantel analyses to test for IBD among all populations, and within each population deme were all statistically insignificant (all populations, $R^2 = 0.0152$, $P = 0.813$; NCRWP, $R^2 = 0.0134$, $P = 0.931$; CEPCO, $R^2 = 0.0612$, $P = 0.429$; PCOLE, $R^2 = 0.0074$, $P = 0.701$), suggesting no association between geographic and genetic distances overall, and shallow genetic structure within population demes. These results have to be interpreted cautiously though, as population demes may be better defined by current migration rather than by historical gene flow.

MDIV analysis supported secondary contact as a better model to explain the *COI* geographical pattern versus historical isolation without subsequent gene flow among population demes (Fig. S1 Supplementary data). The net nucleotide substitution per site between NCRWP and CEPCO ($D_A = 0.005 \pm 0.002$, 95% CI) placed the time of divergence between these population demes around 250,000 (285,000 ± 215,000) years ago in the late Pleistocene. Furthermore, divergence times between NCRWP vs. PCOLE ($D_A = 0.017 \pm 0.004$, 95% CI) and CEPCO vs. PCOLE ($D_A = 0.016 \pm 0.003$, 95% CI) were around 850,000 (975,000 ± 725,000) and 827,000 (952,000 ± 702,000) years ago, respectively, in the middle Pleistocene. Fu's F_S (1997) neutrality test was negative and highly significant in all of the population demes (NCRWP $Fs = -24.444$, $P < 0.0001$; CEPCO $Fs = -25.719$, $P < 0.0001$; PCOLE $Fs = -24.633$, $P < 0.0001$), therefore rejecting the mutation-drift equilibrium assumption in *An. albimanus* and strongly favoring the interpretation of population expansion. Similarly, the R_2 test (Ramos-Onsins and Rozas, 2002) was low, positive and significant (NCRWP $R_2 = 0.0628$, $P = 0.0053$; CEPCO $R_2 = 0.0224$, $P = 0.0012$; PCOLE $R_2 = 0.0634$, $P = 0.0032$), thus reinforcing a scenario of population expansion. It is noteworthy that the mismatch distribution for the entire *COI* data statistically fit the model of sudden expansion, but showed three defined peaks reflecting the existence of three population demes, each of them expanding, albeit at different times (Fig. S3 Supplementary data).

The SP and the MJ networks further supported the population divergence defined by SAMOVA, and provided deeper resolution of the geographic structure of *An. albimanus* uncovering four different *COI* haplogroups (Table 1, Fig. 1, and Fig. S4 Supplementary data). These four maternal lineages were separated by 6–18 mutational steps and segregated among the three geographical regions defined by SAMOVA. Haplogroups (B) and (C) co-occurred and were more prevalent in NCRWP; (A) was prevalent in CEPCO, and (D) was largely predominant in PCOLE (Fig. 1). The net nucleotide substitution per site among these haplogroups ($D_A = 0.008$–0.026) placed the time of divergence between them around (400,000–1300,000) years ago, in the late and middle Pleistocene. We re-calculated the mismatch and neutrality tests for each

Author's personal copy

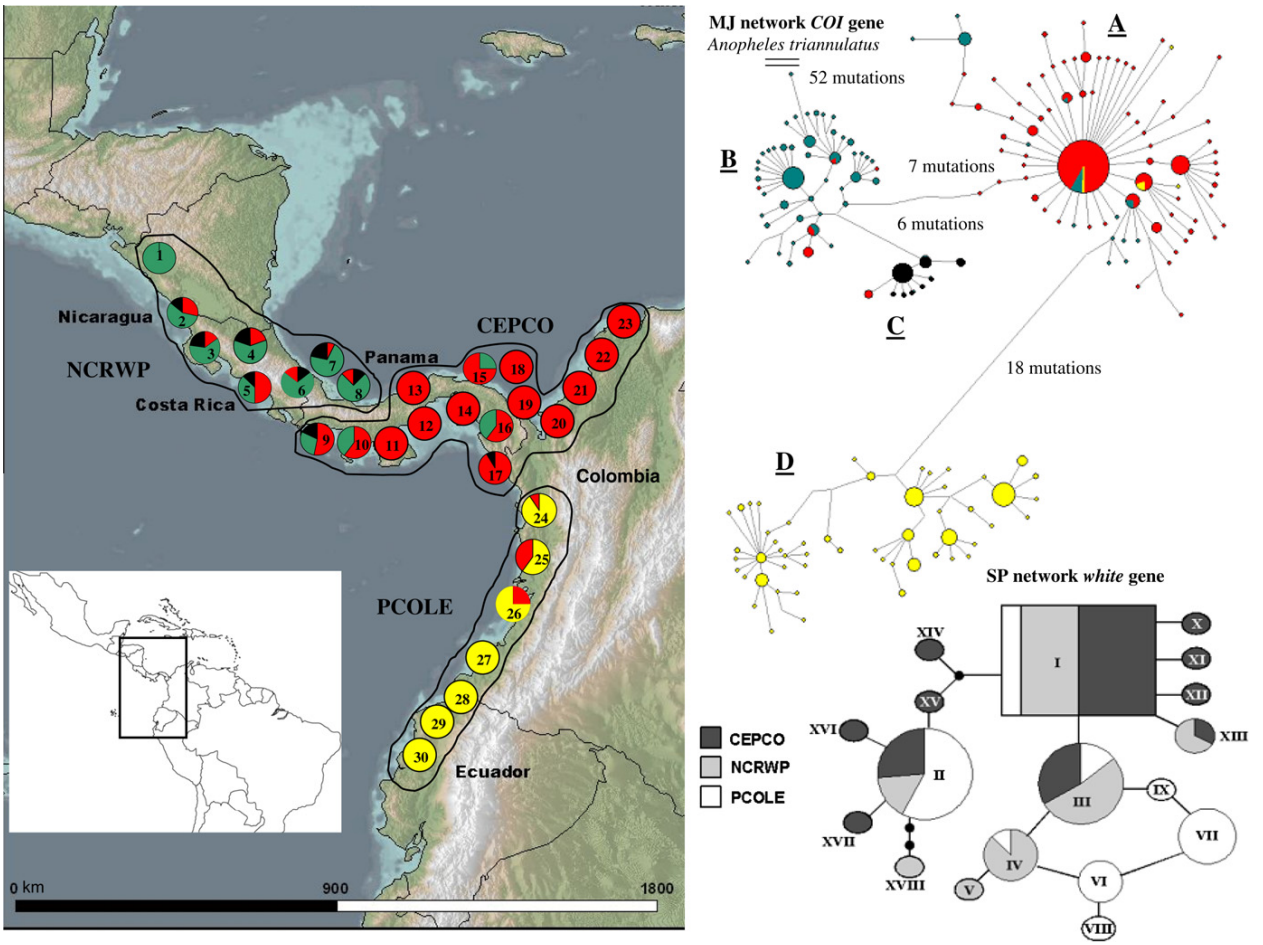

Fig. 1. The map shows the geographic distribution of four *COI* haplogroups (A, red), (B, green), (C, black) and (D, yellow) from the MJ haplotype network of *Anopheles albimanus*, located in the upper right corner. The circled numbers correspond to 30 localities, positioned on the map according to the longitude and latitude of each site. Three population demes defined by SAMOVA are outlined: Nicaragua, Costa Rica and the Atlantic coast of western Panama (NCRWP = localities 1–8); the Pacific coast of western Panama, central-eastern Panama and the Caribbean coast of Colombia (CEPCO = localities 9–23); and the Pacific coast of Colombia and Ecuador (PCOLE = localities 24–30). The statistical parsimony network of eighteen *white* gene alleles (I–XVIII) and their geographical distributions are shown in the lower right corner. The inset map depicts the geographic position of the study area. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

mitochondrial haplogroup separately, because the coalescence signal of the *COI* gene (i.e., the time since expansion) may be confounded by subsequent gene flow and more recent secondary contact among population demes. Both the Fu's F_S and the R_2 tests rejected the mutation-drift equilibrium assumption in all the *COI* haplogroups, likely due to population expansion, and this was corroborated by the non-significant raggedness indexes and unimodal mismatch distributions (Fig. S5 Supplementary data). The expansion times for A ($\tau = 3.658$), B ($\tau = 5.132$), C ($\tau = 2.929$) and D ($\tau = 4.765$) were 19,641 (95% CI 9324–39,641), 27,555 (95% CI 14,177–47,533), 15,727 (95% CI 6324–35,641) and 25,585 (95% CI 10,077–46,533), respectively, all in the late Pleistocene. The BSKPs further supported significant population growth and showed an increase in effective population size in each of these four *COI* haplogroups, all within 150,000–550,000 years before the present (Fig. S2 Supplementary data).

In contrast to the *COI* gene, genetic diversity was lower for the *white* gene (Table 1). Eighteen alleles were joined by 1–3 nucleotide differences in a SP network (Fig. 1). These alleles were very closely related and not segregated geographically as were the four *COI* haplogroups. Instead, one ancestral allele (I) was present in all the sample locations, but more frequently in localities across NCRWP, and CEPCO, and almost fixed for those in Costa Rica and Panama. Three other alleles, of intermediate-frequency (II, III, and IV), were shared among regions whereas the rest were less frequent and more restricted geographically (Table 1 and Fig. 1). It is noteworthy that seven out of ten *white* gene singletons were recovered exclusively from CEPCO, and this may suggest a recent pattern of regional population growth (Fig. 1). Furthermore, Fu's F_S (1997) neutrality test was negative and significant ($Fs = -7.65$, $P < 0.0001$), and the R_2 test (Ramos-Onsins and Rozas, 2002) was low, positive and highly significant ($R_2 = 0.0112$, $P = 0.0013$). The expansion time for the *white* alleles based on $\tau = 2.55$ from the mismatch distribution (Fig. S5 Supplementary data), was between 20,586 and 41,996 years ago, in the late Pleistocene. Nevertheless, the BSKP for the nuclear marker depicted a recent decrease in the effective population size of *An. albimanus* around 50,000 years before the present, although it was not statistically significant (Fig. S2 Supplementary data). Given the use of non-anopheline mutation rates in the present study, our estimates of the divergence and expansion times for the *COI* and the *white* genes have to be interpreted with caution as they may reflect only approximate values.

Author's personal copy

4. Discussion

The present study provides further evidence that late Pleistocene environmental changes had a profound effect on the demography and regional geographic structure of *An. albimanus* in the northern Neotropics. The *COI* findings appear to be the result of historical fragmentation leading to four *COI* haplogroups that were geographically isolated during the middle Pleistocene, but reconnected across southern Central America via more recent secondary contact and a common late Pleistocene expansion event. Although the mismatch distributions for NCRWP, CEPCO and PCOLE were not statistically different from a model of sudden population expansion, they were all visually bimodal, perhaps reflecting the admixing of *COI* haplogroups among different regions. In addition, MDIV estimated higher secondary contact between NCRWP and CEPCO, and this is supported by less divergence between haplogroups (A), (B) and (C) (6–13 mutations) than between (A, B, C) and (D) (18–25 mutations). Clearly, this pattern reflects the fact that haplogroup D is the most geographically restricted because none of its haplotypes were encountered outside PCOLE (Fig. 1).

Alternatively, the *COI* divergence may be due to directional selection in the mitochondrial genome, for instance, an advantageous mutation sweeping to fixation across southern Central America. However, selection is generally a locus specific force, whereas demographic changes affect the entire genome. The latter is more consistent with the evidence of population expansion in both the *COI* and *white* genes. Moreover, all the *COI* haplogroups showed an excess of low frequency haplotypes indicating a common pattern of recent population growth regardless of their geographical distributions. This would make a selective sweep unlikely because the same advantageous mutation would have to appear independently in each of NCRWP, CEPCO and PCOLE.

On the other hand, two short insertions, found in the ITS2 of *An. albimanus*, may support the presence of unidentified cryptic taxa, but very few individuals harbored these indels, and only one of these insertions was found exclusively in PCOLE (Table 1). Besides these indels, no fixed substitutional changes or length differences were encountered in the ITS2 sequences of 52, 59, and 52 individuals from NCRWP, CEPCO and PCOLE, respectively. Although based on our ITS2 analysis, we cannot reject substantial historical gene flow among populations of *An. albimanus*; we cannot entirely reject the existence of multiple taxa (i.e., cryptic speciation) either. For instance, the levels of divergence between our sympatrically occurring mtDNA lineages are far in excess of those observed between sister taxa in other *Anopheles* species (Foley et al., 2006). The restricted occurrence of intragenomic variability in the ITS2 of samples from NCRWP and CEPCO, and an exclusive insertion in three samples from PCOLE, may suggest past geographical fragmentation as proposed for the *COI* data, though the alternative hypothesis of recent, perhaps slight limits to gene flow seems equally likely (Bower et al., 2008).

Anopheles albimanus is believed to have originated in the Caribbean islands and then colonized the American continent, and a founder effect may have been a main factor shaping its continental population structure (Molina-Cruz et al., 2004). The colonization routes suggested by these authors hypothesize that *An. albimanus* invaded South America from northern Central America. In the present analysis, *Anopheles triannulatus* joined through 52 mutations to haplogroup B in the *COI* MJ network (Fig. 1) indicating that haplogroups (C) and (A) originated from (B), and then (D) originated from (A). This pattern is consistent with the scenario suggested by Molina-Cruz et al. (2004). The co-occurrence of two distinct *COI* haplogroups (B and C) in NCRWP supports this view and indicates that successful initial colonization of South America was achieved by two separate introductions of *An. albimanus* from Nicaragua into Colombia and Ecuador. Nevertheless, genetic diversity as measured by both genes does not decrease progressively from north to south, as would be expected under a founder effect and sequential bottleneck in the colonizing front. Instead genetic diversity is lower in CEPCO than in Nicaragua (NCRWP) and Ecuador (PCOLE) (Table 1).

Coalescent theory predicts that the most frequent and widely distributed gene variants are ancestral, because they have had more time to disperse (Uthicke and Benzie, 2003). In our data, both of these variants, the *COI* haplotype (A1) and the *white* allele (I), were prevalent across Panama, implying a possible Panamanian geographical origin for *An. albimanus* (Fig. 1 and Fig. S4 Supplementary data). This view is unlikely, though, given the recent geological origin of the Isthmus of Panama, dated approximately 3–3.5 mya (Bermingham and Martin, 1998). Due to the non-equilibrium frequencies of variants in both genes and our incomplete sampling of the geographic range of *An. albimanus*, we do not have enough evidence to draw firm conclusions about its geographical origin and initial colonization path across southern Central America.

The observed geographic structure seems to be the result of population contraction of an ancestral and perhaps stable population of *An. albimanus* across the Isthmus of Panama (i.e., the four *COI* haplogroups diverged and expanded within the same time frame). This view is supported by lower diversity in the *COI* and *white* genes across Panama (CEPCO), which may indicate that *An. albimanus* went through a population bottleneck followed by a subsequent demographic expansion toward Nicaragua and Ecuador (Table 1). Previous work using four microsatellite loci and the mtDNA *ND5* gene also indicated significantly reduced diversity for one locality in central Panama, thus supporting our hypothesis (Molina-Cruz et al., 2004). Our results from two independent molecular markers may represent two temporal pictures of the demography of *An. albimanus* in the study area, with the *white* gene showing a bottleneck and the *COI* depicting a subsequent demographic expansion, and this may be caused by different mutation rates, and effective population sizes in these genes. The bottleneck hypothesis may be supported by a drop of 35% precipitation and changes in vegetation and faunal structure in central-eastern Panama in the late Pleistocene (Piperno and Jones, 2003; Gonzáles et al., 2006). Additional signals of Pleistocene geographic fragmentation across Panama have been demonstrated in howler monkeys, black mangroves and dirt frogs, and hypothesized to be caused by forest fragmentation, sea level changes, and climatic refugia, respectively (Crawford, 2003; Cortes-Ortíz et al., 2003; Nettel et al., 2008).

5. Conclusions

Anopheles albimanus appears to be a single, albeit polymorphic, species that is not at mutation-drift equilibrium due to past geographic fragmentation and regional fluctuation in its effective population size. There is a strong geographic component in its genetic structure with three population demes and an admixture zone across eastern Panama and northern Colombia. The *COI* gene suggests a common pattern of historical isolation, subsequent haplotype mixing, and population expansion of four mtDNA *COI* lineages in the late Pleistocene. Data from the *white* gene do not reflect the high genetic diversity of the *COI*, but are consistent with the scenario of late Pleistocene population expansion, thus supporting demographic phenomena as the cause of structure rather than a selective sweep. Finally, lower genetic diversity by both the *COI* and *white* genes across Panama suggests that *An. albimanus* populations contracted, then subsequently expanded toward Nicaragua and Ecuador.

Author's personal copy

Acknowledgments

Financial support was obtained partially from the Secretariat for Science, Technology and Innovation of Panama (SENACYT) through a research grant (COL08-066) awarded to Jose R. Loaiza. Additional financial support was provided by the Smithsonian Tropical Research Institute (STRI), National Institutes of Health (NIH) grant (AI) R0154139-02 to JEC, Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología (COLCIENCIAS) grant 1115-05-16879 to MCO, Comité para el Desarrollo de la Investigación (CODI) Universidad de Antioquia grant 8700-039 and E-01233, to MCO, and Natural Sciences and Engineering Research Council of Canada (NSERC) to MES. Research at the Institute of Parasitology is supported by a regroupement stratégique from Fonds Québécois de la recherche sur la nature et les technologies (FQRNT). We would like to acknowledge Dr. Efrain Beltran from the National Malaria Eradication Service (SNEM), in Ecuador, and Luis Guillermo Chaverri Sánchez from Instituto Nacional de Biodiversidad de Costa Rica (INBio) for technical assistance. We are grateful to Matthew Miller and Laura B. Geyer from STRI and to Martin Donnelly from the Liverpool School of Tropical Medicine for helpful discussions and comments on earlier versions of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2010.09.016.

References

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Bermingham, E., Martin, A.P., 1998. Comparative mtDNA phylogeography of Neotropical fresh water fishes: testing shared history to infer the evolutionary landscape of lower Central America. Mol. Ecol. 7, 499–517.

Bower, J.E., Dowton, M., Cooper, R.D., Beebe, N.W., 2008. Intraspecific concerted evolution of the rDNA ITS1 in *Anopheles farauti* sensu stricto (Diptera: Culicidae) reveals recent patterns of population structure. J. Mol. Biol. 67, 397–411.

Collins, F.H., Paskewitz, S.M., 1996. A review of the use of ribosomal DNA (rDNA) to differentiate among cryptic *Anopheles* species. Insect. Mol. Biol. 5 (1), 1–9.

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

Cortes-Ortíz, L., Bermingham, E., Rico, C., Rodríguez-Luna, E., Sampaio, I., Ruiz-Garcíae, M., 2003. Molecular systematics and biogeography of the Neotropical monkey genus *Alouatta*. Mol. Phyl. Evol. 26, 64–81.

Crawford, A.J., 2003. Huge populations and old species of Costa Rican and Panamanian dirt frogs inferred from mitochondrial and nuclear gene sequences. Mol. Ecol. 12, 2525–2540.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Dupanloup, I., Schneider, S., Excoffier, L., 2002. A simulated annealing approach to define the genetic structure of populations. Mol. Ecol. 11, 2571–2581.

Excoffier, L., Laval, G., Schmeider, S., 2005. Arlequin ver 3.0: an integrated software package for population genetic data analysis. EBO 1, 47–50.

Foley, D.H., Wilkerson, R.C., Cooper, R.D., Volovsek, M.E., Bryan, J.H., 2006. A molecular phylogeny of *Anopheles annulipes* (Diptera: Culicidae) sensu lato: the most species-rich anopheline complex. Mol. Phylogenet. Evol. 43, 283–297.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.

Gonzáles, C., Urrego, L.E., Martínez, J.I., 2006. Late quaternary vegetation and climate change in the Panama basin: Palynological evidence from marine cores ODP 677 and TR 163-38. PALAEO 234, 62–80.

Gutiérrez, L.A., Naranjo, N.J., Cienfuegos, A.V., Muskus, C.E., Luckhart, S., Conn, J.E., Correa, M.M., 2009. Population structure analyses and demographic history of the malaria vector *Anopheles albimanus* Wiedemann (Diptera: Culicidae) from the Caribbean and the Pacific regions of Colombia. Mal. J. 8, 259.

Jensen, J., Bohonak, A.J., Kelley, S.T., 2005. Isolation by distance, web service. BMC Genetics 6, 13.

Langley, C.H., Aquadro, C.F., 1987. Restriction-map variation in natural populations of *Drosophila melanogaster*: *white*-locus region. Mol. Biol. Evol. 151, 651–663.

Linton, Y.M., Samanidou-Voyadjoglou, A., Harbach, R.E., 2002. Ribosomal ITS2 sequence data for *Anopheles maculipennis* and *An. messeae* in northern Greece, with a critical assessment of previously published sequences. Insect. Mol. Biol. 11 (4), 379–383.

Loaiza, J.R., Scott, M.E., Bermingham, E., Rovira, R., Conn, J.E., 2010. Evidence for Pleistocene population divergence and expansion of *Anopheles albimanus* in Southern Central America. Am. J. Trop. Med. Hyg. 82 (1), 156–164.

Miller, M.J., Bermingham, E., Klicka, J., Escalante, P., Raposo do Amaral, F.S., Weir, J.T., Winker, K., 2008. Out of Amazonia again and again: episodic crossing of the Andes promotes diversification in a lowland forest flycatcher. Proc. R. Soc. B. 275, 1133–1142.

Mirabello, L., Conn, J.E., 2006a. Molecular population genetics of the malaria vector *Anopheles darlingi* in Central and South America. Heredity 96, 311–321.

Mirabello, L., Conn, J.E., 2006b. Correction to Heredity (2006) 96, 311–321: Molecular population genetics of the malaria vector *Anopheles darlingi* in Central and South America. Heredity 97, 438.

Mirabello, L., Conn, J.E., 2008. Population analysis using the nuclear *white* gene detects Pliocene/Pleistocene lineage divergence within *Anopheles nuneztovari* in South America. Med. Vet. Entomol. 22, 109–119.

Miyashita, N., Langley, C.H., 1988. Molecular and phenotypic variation of the *white* locus region in *Drosophila melanogaster*. Genetics 120, 199–212.

Molina-Cruz, A., De Merida, A.M., Mills, K., Rodríguez, F., Schoua, C., Yurrita, M.M., Molina, E., Palmieri, M., William, C.B., 2004. Gene flow among *Anopheles albimanus* populations in Central America, South America, and the Caribbean assessed by microsatellites and mitochondrial DNA. Am. J. Trop. Med. Hyg. 71, 350–359.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Nettel, A., Dodd, R.S., Afzal-Rafii, Z., Tovilla-Hernandez, C., 2008. Genetic diversity enhanced by ancient introgression and secondary contact in East Pacific black mangroves. Mol. Ecol. 17, 2680–2690.

Nielsen, R., Wakeley, J., 2001. Distinguishing migration from isolation: A Markov chain Monte Carlo approach. Genetics 158, 885–896.

O'Loughlin, S.M., Okabayashi, T., Honda, M., Kitazoe, Y., Kishino, H., Somboon, P., Sochantha, S., Nambanya, P.K., Saikia, V.D.E.V., Walton, C., 2008. Complex population history of two *Anopheles dirus* mosquito species in South Asia suggests the influence of Pleistocene climate change rather than human-mediated effects. J. Evol. Biol. 21, 1555–1569.

Piperno, D.R., Jones, J.G., 2003. Paleoecological and archaeological implications of the Late Pleistocene Early Holocene record of vegetation and climate from the Pacific coastal plain of Panama. Quaternary Res. 59, 79–87.

Powell, J.R., Caccone, A., Amato, G.D., Yoon, C., 1986. Rate of nucleotide substitution in *Drosophila* mitochondrial DNA and nuclear DNA are similar. Proc. Natl. Acad. Sci. USA 83, 9090–9093.

Ramos-Onsins, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19, 2092–2100.

Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 62, 122–127.

Rozas, J., Sanchez-Del Rio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescence and other methods. Bioinformatics 19, 2496–2497.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Walton, C., Handley, J.M., Tun-Lin, W., 2000. Population structure and population history of *Anopheles dirus* mosquitoes in Southeast Asia. Mol. Biol. Evol. 17, 962–974.

Uthicke, S., Benzie, J.A.H., 2003. Gene flow and population history in high dispersal marine invertebrates: mitochondrial DNA analysis of *Holothuria nobilis* (Echinodermata: Holoturoidea) populations from the Indo-Pacific. Mol. Ecol. 12, 2635–2648.

Weigt, L.A., Crawford, A.J., Stanley Rand, A., Ryan, M., 2005. Biogeography of the tungara frog, *Physalaemus pustulosus*: a molecular perspective. Mol. Ecol. 14, 3857–3876.

Zeh, J.A., Zeh, D.W., Bonilla, M.M., 2003. Phylogeography of the harlequin beetle-riding pseudoscorpion and the rise of the Isthmus of Panama. Mol. Ecol. 12, 2759–2769.