Infection, Genetics and Evolution 16 (2013) 87–92

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

No evidence for positive selection at two potential targets for malaria transmission-blocking vaccines in *Anopheles gambiae s.s*

Jacob E. Crawford [a,*], Susan M. Rottschaefer [a], Boubacar Coulibaly [b], Madjou Sacko [b], Oumou Niaré [b], Michelle M. Riehle [c], Sékou F. Traore [b], Kenneth D. Vernick [c,d], Brian P. Lazzaro [a]

[a] Department of Entomology, Cornell University, Ithaca, NY, USA
[b] Malaria Research and Training Center, University of Bamako, Bamako, Mali
[c] Department of Microbiology, University of Minnesota, St. Paul, MN, USA
[d] Unit for Genetics and Genomics of Insect Vectors, Institut Pasteur, Paris, France

ARTICLE INFO

Article history:
Received 29 October 2012
Received in revised form 4 January 2013
Accepted 5 January 2013
Available online 26 January 2013

Keywords:
Malaria
Anopheles gambiae
Transmission-blocking vaccine
Plasmodium falciparum
Natural selection
Population genetics

ABSTRACT

Human malaria causes nearly a million deaths in sub-Saharan Africa each year. The evolution of drug-resistance in the parasite and insecticide resistance in the mosquito vector has complicated control measures and made the need for new control strategies more urgent. *Anopheles gambiae s.s.* is one of the primary vectors of human malaria in Africa, and parasite-transmission-blocking vaccines targeting *Anopheles* proteins have been proposed as a possible strategy to control the spread of the disease. However, the success of these hypothetical technologies would depend on the successful ability to broadly target mosquito populations that may be genetically heterogeneous. Understanding the evolutionary pressures shaping genetic variation among candidate target molecules offers a first step towards evaluating the prospects of successfully deploying such technologies. We studied the population genetics of genes encoding two candidate target proteins, the salivary gland protein saglin and the basal lamina structural protein laminin, in wild populations of the M and S molecular forms of *A. gambiae* in Mali. Through analysis of intraspecific genetic variation and interspecific comparisons, we found no evidence of positive natural selection at the genes encoding these proteins. On the contrary, we found evidence for particularly strong purifying selection at the laminin gene. These results provide insight into the patterns of genetic diversity of saglin and laminin, and we discuss these findings in relation to the potential development of these molecules as vaccine targets.

1. Introduction

Mosquito-stage, transmission blocking vaccines have been proposed as an alternative novel technology for blocking transmission of human malaria parasites (Carter and Chen, 1976; Gwadz, 1976; Barreau et al., 1995; Brennan et al., 2000). The logic of this technology is that mosquito proteins essential for parasite development could be targeted, for example by human derived antibodies, and blocked such that transmission is halted within the mosquito. In the case of mosquito midgut receptors exploited by the ookinete stage of the parasite, anti-receptor anti-bodies delivered in the bloodmeal could directly compete with and inhibit parasite development, while anti-salivary gland receptor anti-bodies will have to be delivered to the mosquito through bloodmeals subsequent to the initial exposure (Brennan et al., 2000; Lavazec and Bourgouin, 2008). At the heart of this hypothetical technology is the assumption that the vaccine target (i.e. mosquito proteins) can be easily identified and blocked by antibodies, but genetic variation segregating in natural populations could result in a heterogeneous molecule population that is difficult to target. Such genetic variation could be neutral with respect to natural selection or it could adaptively evolving, including through selection to resist parasite establishment. Substantial traction has been made in identifying potential target molecules within the mosquito, but most of that work has relied on genetically inbred lab strains of both mosquito and the parasite (Brennan et al., 2000; Arrighi et al., 2005; Saul, 2007; Dinglasan and Jacobs-Lorena, 2008), providing little information on potential genetic heterogeneity and its consequences in nature. Understanding the evolutionary pressures shaping genetic variation among candidate target molecules offers a first step towards evaluating the prospects of successfully deploying such technologies.

The suggestions that pathogen-related selection pressures are likely to drive host evolution date back 100 years (Biffen, 1905; Haldane, 1949; Lederberg, 1999). As one of the most permissive and common vectors of the human malaria parasite *Plasmodium falciparum*, *Anopheles gambiae sensu stricto* is a good candidate for experiencing such selection. Several non-immune mosquito proteins directly interact with the developing malaria parasite and

* Corresponding author. Present address: Department of Integrative Biology, University of California, Berkeley, CA, USA
E-mail address: j.crawford@berkeley.edu (J.E. Crawford).

1567-1348/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2013.01.006

are in some cases required for successful parasite tissue invasion and development. One of the first *Anopheles* proteins shown to interact directly with both ookinete and oocyst stages of *Plasmodium* parasites is laminin, a component of basal laminae in the mosquito, including the one surrounding the midgut (Adini and Warburg, 1999; Vlachou et al., 2001; Arrighi and Hurd, 2002; Dessens et al., 2003). Subsequent studies of laminin suggest that this protein may act as a trigger for the transition from ookinete to oocyst development, or even as a protective coating that masks the parasite from immune detection (Arrighi et al., 2005; Warburg et al., 2007). Further evidence for the intimate and perhaps protective role of laminin was provided by the observations that laminin becomes localized within the oocysts and sporozoites, and it is incorporated into the oocyst capsule (Nacer et al., 2008).

A second host protein, saglin, plays a crucial role in parasite localization and invasion of the salivary glands through a receptor-ligand interaction with the *Plasmodium* TRAP protein (Brennan et al., 2000; Korochkina et al., 2006; Okulate et al., 2007; Ghosh et al., 2009). *P. falciparum* invasion of the salivary gland is inhibited when saglin is blocked with either antibody interference or receptor saturation by SM1, a short peptide whose physical conformation resembles TRAP (Brennan et al., 2000; Ghosh et al., 2001, 2009). Moreover, point mutations in TRAP completely abrogate gland invasion (Matuschewski et al., 2002). Population genetic analysis of *Plasmodium falciparum* and *Plasmodium vivax* suggests adaptive maintenance of variation, especially in the A-domain that binds to saglin (Weedall et al., 2007; Barry et al., 2009).

Based on the evidence that laminin and saglin mediate *Plasmodium* infection in *Anopheles*, we hypothesized that these proteins may be under pathogen-related selection pressure. It should be noted that, while we hypothesize that *Plasmodium* may be a selectively driving force, it is in practice difficult to identify the proximal agent of selection. Nonetheless, the potential effects of selection on potential vaccine targets has important implications regardless of the agent of selection. To elucidate the selective history of the *saglin* and *laminin* genes, we sequenced alleles sampled from wild populations of the two incipient species of *A. gambiae*, the M and S molecular forms from Mali. We analyzed patterns of intraspecific polymorphism and divergence at these loci but found no significant evidence for adaptive evolution at these loci in either population.

2. Materials and methods

2.1. Mosquito samples

Anopheles gambiae individuals were collected inside dwellings from the villages of Bancoumana and N'gabakoro Droit outside the Malian capital city, Bamako (12°39′N 8°0′W), and an additional collection was drawn from Toumani-Oulena, Mali (10°83′N 7°81′W). In Mali and much of West Africa, *A. gambiae* is largely composed of two populations that are defined by fixed differences in rDNA on the X chromosome (Lehmann and Diabate, 2008). The M/S molecular form of each individual mosquito was determined using the PCR diagnostic developed by Favia et al. (2001). Of the mosquitoes sampled from Bancoumana, four were M form and 11 were S form. All mosquitoes sampled from N'gabakoro Droit were M form ($n = 10$), and all Toumani-Oulena individuals were S form ($n = 7$). *Anopheles merus* DNA from mosquitoes of the OPHANSI colony was obtained from MR4.

2.2. DNA extraction, PCR and sequencing

DNA was extracted from the mosquitoes using DNeasy kits (Qiagen) under slight modifications to the manufacturers' suggested protocols. PCR primers were designed based on the published *A. gambiae* genome sequence (Holt et al., 2002). Each gene was amplified from genomic DNA using iProof high fidelity DNA Polymerase (BioRad). PCR products were run out on a 1% agarose gel and the product fragments were excised and purified using the PureLink gel extraction kit (Invitrogen). Adenosine tails were added to the purified products by incubating for 20 min at 72 °C with PCR buffer, dATP and Taq polymerase. Products were then cloned using the TOPO XL cloning kit (Invitrogen). Colonies to be sequenced were grown overnight at 37 °C in liquid Luria–Bertani broth supplemented with 20 mg/ml kanamycin, and the plasmids were isolated using the Qiaprep spin miniprep kit (Qiagen). The products were then sequenced directly from the plasmids using the BigDye Terminator Cycle Sequencing Kit v3.1(ABI). The sequences were assembled using Sequencher (Gene Codes Corp.) and CodonCode Aligner (CodonCode Corp.). Only one of the two alleles at each gene was sequenced from any given mosquito in the study. All sequences have been deposited into Genbank under accession numbers KC438305–KC438359.

To control for sequencing error, all singleton polymorphisms were verified by re-amplification and direct sequencing of heterozygous PCR products. The entire gene was amplified from genomic DNA using iProof high fidelity DNA Polymerase (BioRad) and this full-length amplicon was then used as template in a secondary PCR using internally nested primers to robustly amplify the gene region containing the singleton to be validated. Unincorporated primers and dNTPs were inactivated from these secondary amplification products by incubation with ExoI and SAP (both manufactured by USB), and amplification products were then sequenced using the BigDye Terminator Cycle Sequencing Kit v3.1(ABI). To further avoid errors stemming from homopolymers, we deleted all homopolymer sequences from the alignment.

2.3. Loci analyzed

We analyzed the *SAG* locus (AGAP000610), which is X-linked, and *LANB2* (AGAP007629), which is found on the distal tip of chromosome 2L. Both loci were surveyed in the Bancoumana and N'gabakoro Droit populations, but only *SAG* sequences were generated from the Toumani-Oulena population. The original predicted exon structure of *SAG* was reannotated in a subsequent genome release as reflected in Ensemble, removing a predicted intron leaving a single coding sequence (Brennan et al., 2000). However, to be certain, we obtained cDNA from live *A. gambiae* females and directly sequenced *SAG* transcripts to identify prospective intron boundaries. We used the same PCR, cloning, and sequencing conditions and reagents described above, using first-strand cDNA from whole mosquitoes as template in the PCR reaction. After aligning the cDNA sequence to the Agam PEST reference sequence, we identified a single 175 basepair (bp) intron beginning at position 1007 of the coding sequence, consistent with the original Ensemble annotation prior to the re-annotation based on short peptide mapping (Brennan et al., 2000). According to our sequencing results, the final saglin protein is predicted to be 374 amino acids in length, and for the analyses here, we assumed that the putative 175 bp intron is non-coding and analyzed the sequence accordingly. For *LANB2*, we analyzed the genomic region according to the exon structure annotated in Vectorbase (www.vectorbase.org).

2.4. Population genetic analysis

Measures of nucleotide diversity estimated from the average number of differences between haplotypes (π) and the number of polymorphic sites (θ_W) were calculated on synonymous and nonsynonymous sites alone as well as on all sites combined using DnaSP version 5 (Librado and Rozas, 2009). Three neutrality tests that emphasize different features of the data including Tajima's *D*