A Physical Map for an Asian Malaria Mosquito, *Anopheles stephensi*

Maria V. Sharakhova,[†] Ai Xia,[†] Zhijian Tu, Yogesh S. Shouche, Maria F. Unger, and Igor V. Sharakhov*

Department of Entomology, Virginia Tech, Blacksburg, Virginia; Department of Biochemistry, Virginia Tech, Blacksburg, Virginia;
Lab 3, National Center for Cell Science, Pune, India; Eck Institute for Global Health, University of Notre Dame, Notre Dame, Indiana

Abstract. Physical mapping is a useful approach for studying genome organization and evolution as well as for genome sequence assembly. The availability of polytene chromosomes in malaria mosquitoes provides a unique opportunity to develop high-resolution physical maps. We report a 0.6-Mb-resolution physical map consisting of 422 DNA markers hybridized to 379 chromosomal sites of the *Anopheles stephensi* polytene chromosomes. This makes *An. stephensi* second only to *Anopheles gambiae* in density of a physical map among malaria mosquitoes. Three hundred sixty-three (363) probes hybridized to single chromosomal sites, whereas 59 clones yielded multiple signals. This physical map provided a suitable basis for comparative genomics, which was used for determining inversion breakpoints, duplications, and origin of novel genes across species.

INTRODUCTION

Malaria, with its great biological and social complexity, is one of the most important global health problems. The arsenal of tools available for fighting malaria is very limited; vector control still remains to be the major approach for reducing or eliminating malaria. The genomic era opens an exciting opportunity to uncover the molecular mechanisms contributing to vectorial capacity and to identify targets potentially useful for vector management. Knowledge of the chromosomal location of genes and other DNA markers has important applications for taxonomy, systematics, ecological and population genetics, evolutionary genomics, and map-based cloning. The availability of well-polytenized chromosomes in anopheline mosquitoes provides a great opportunity to develop high-resolution physical maps. Polytene chromosomes are produced by multiple DNA replications without cell division and are characterized by large size and distinguishable banding patterns. They are present in various tissues of mosquitoes including salivary glands, gut, Malpighian tubules, and ovarian nurse cells.¹ Various molecular markers can be successfully used for physical mapping. The first *Anopheles gambiae* physical map included 46 clones microdissected from the different divisions of polytene chromosomes.² Later, microsatellite markers,³ randomly amplified polymorphic DNA (RAPD) markers,⁴ cosmids, and complementary DNA (cDNA) markers⁵ have been physically mapped to the *An. gambiae* chromosomes. The most detailed mosquito physical map has been developed for the *An. gambiae* genome project.⁶ *In situ* hybridization of 2,000 end-sequenced bacterial artificial chromosome (BAC) clones allowed researchers to assign 67 scaffolds of 227 megabase pairs (Mbp) in total length to the chromosome regions. In addition, 16 scaffolds, from 50 to 600 kb in length, were localized in pericentromeric regions using cDNA clones.⁷

In this study, we report a 0.6-Mb-resolution physical map for *Anopheles stephensi*, an important vector of *Plasmodium vivax* and *Plasmodium falciparum* in the Indo-Pakistan subcontinent and Middle-East. The map consists of 422 BAC clones and cDNA clones hybridized to 379 chromosomal sites. This makes *An. stephensi* only second to *An. gambiae*

in density of a physical map among malaria mosquitoes. The basic local alignment search tool (BLAST) of the mapped DNA markers against the *An. gambiae* genome identified species-specific and conserved sequences. The comparative analysis of chromosomal location of the markers with the *An. gambiae* genome found inversion breakpoints and possible gene duplications with no indication of interchromosomal transpositions.

MATERIAL AND METHODS

Mosquito strain and chromosome preparation. The Indian wild-type strain of *An. stephensi* was used for the physical mapping. The standard chromosome map for this strain was previously developed.⁸ Ovaries from half-gravid females prefixed in Carnoy's fixative solution were dissected in 50% propionic acid under a Leica MZ6 dissection microscope (Leica Microsystems GmbH, Wetzlar, Germany). A cover slide was placed on the follicles and pressed to squash the cells. The banding pattern of polytene chromosomes was examined using Olympus CX-41 phase-contrast microscope (×1,000) (Olympus America Inc., Melville, NY). Slides with good chromosomal preparations were dipped in liquid nitrogen, then cover slips were removed and slides were dehydrated in 50%, 70%, 95%, and 100% ethanol.

Fluorescence *in situ* hybridization. In this study, 181 *An. stephensi*, *An. gambiae*, and *Anopheles funestus* cDNA and BAC clones were hybridized to polytene chromosomes of *An. stephensi* (Table S1). The 241 previously mapped markers^{8–10} were also added to the physical map of *An. stephensi* (Table S2). Recombinant cDNA and BAC clones were isolated using PhasePrep BAC DNA Kit (Sigma-Aldrich Corp., St. Louis, MO). The isolated DNA was labeled with Cy5-AP3-dUTP and Cy3-AP3-dUTP (GE Healthcare UK Ltd., Buckinghamshire, England) using a modified Nick translation labeling protocol or labeled with Biotin-16-dUTP by a modified Nick Translation Mix protocol (Roche Applied Science, Indianapolis, IN). The DNA probes were hybridized to the chromosomes at 39°C overnight in hybridization solution (Invitrogen Corp., Carlsbad, CA). The chromosomes were then washed in 0.2XSSC (saline-sodium citrate: 0.03 M sodium chloride, 0.003 M sodium citrate) counterstained with YOYO-1 and mounted in DABCO. Fluorescent signals were detected and recorded using a Zeiss LSM 510 Laser Scanning Microscope (Carl Zeiss MicroImaging, Inc., Thornwood, NY). The exact localization of each signal was determined within a

*Address correspondence to Igor V. Sharakhov, Department of Entomology, 203 Fralin Life Science Institute, West Campus Drive, MC 0346, Virginia Tech, Blacksburg, VA 24061. E-mail: igor@vt.edu

[†]These authors contributed equally to this work.

subdivision, and no variation in signal localization was detected among all the nuclei examined for a given clone. Localization of a signal was done using a standard cytogenetic map for *An. stephensi*.⁸ The BLASTN and BLASTX algorithms were used to identify homologous sequences in the *An. gambiae* genome, which is available at VectorBase.¹¹

RESULTS

The polytene chromosome complement of *Anopheles* consists of five arms: X, 2R, 2L, 3R, and 3L. The X chromosome is the shortest and the 2R chromosome the longest. We present a 0.6-Mb-resolution physical map for *An. stephensi* that consists of 173 BAC clones and 249 cDNA clones (Figure 1, Figure S1). In this study, 181 BAC and cDNA clones were mapped to the polytene chromosomes from ovarian nurse cells of *An. stephensi* using fluorescence *in situ* hybridization (Figure 2, Table S1). Together with 241 previously published markers,^{8–10} these clones occupy 379 chromosomal sites on five chromosome arms of *An. stephensi* (Table S2). The genome size of *An. stephensi* is ~240 Mbp.¹² Therefore, the resolution of the current map is about 633 kb on the average. The Cy3 and Cy5 labeled probes hybridized to both euchromatic and heterochromatic regions and to various places with respect to polymorphic inversions (Figure 2). Of 422 probes, 363 had unique locations on chromosome arms. Forty-nine probes of a total 59 probes with multiple locations on *An. stephensi* chromosomes had one or more hits in the *An. gambiae* genome (Table S3) and 10 multiply located clones were specific to *An. stephensi*. The multiple locations may result from inversions, transpositions, or gene/segmental duplications. The only transpositions that we can safely detect with our mapping procedure are those that took place between unique locations in different chromosomal arms. In our study, out of 422 probes, no interchromosomal transposition event has been identified between *An. gambiae* and *An. stephensi*. Eight *An. gambiae* BAC clones, 10E06, 109B13, 127F13, 141A14, 146D17, 155I2, 25D14, and 31H07, produced two distinct signals on the chromosomal arms of *An. stephensi* but unique BLAST hits in the *An. gambiae* genome (Table S3). The data suggest that these BAC clones captured breakpoints of fixed inversions between the two species. Our previous study found that the 141A14 and 146D17 BAC clones are located in the breakpoints of the inversions *2La* and *2Ro*, respectively.¹⁰ These inversions are fixed in *Anopheles merus*, a member of the *An. gambiae* complex. Thus, physical mapping of an outgroup species, *An. stephensi*, supported the ancestral state of these arrangements in the *An. gambiae* complex. Twenty-two of 249 cDNA clones (8.8%) were localized to multiple chromosome sites in *An. stephensi* but were localized to unique locations in the *An. gambiae* genome. In contrast, only three cDNA clones (1.2%) with unique locations in the *An. stephensi* had multiple hits in the *An. gambiae* genome. These data indicate that the rate of possible gene duplication events was higher in the *An. stephensi* lineage. The BLAST analysis against the databases of all the available genomes (<http://blast.ncbi.nlm.nih.gov/>) identified 24 *An. stephensi* cDNAs with no hits (Table S4). Thus, almost 10% of the mapped *An. stephensi* cDNA clones were species-specific. The cDNAs were obtained from unique transcripts identified in the midgut tissue of the adult *An. stephensi* female.¹³ The study of chromosomal localization of the 24 cDNA clones

demonstrated that eight clones (33%) are X chromosome specific and nine clones (38%) have multiple locations on autosomal arms of *An. stephensi* (Table S4). Because of the limited information of gene annotations, the function of these 24 *An. stephensi*-specific clones remains unclear.

DISCUSSION

In species with polytene chromosomes, like those in the genera *Anopheles* and *Drosophila*, *in situ* hybridization has facilitated the comparative genome mapping. This approach can detect the macro-rearrangement, such as inversions and translocations, as well as transpositions between chromosomal arms. Our previous study showed that paracentric inversions and whole-arm translocations are the major types of chromosome rearrangements in *Anopheles*.⁹ The mean length of conserved syntenic blocks between *An. gambiae* and *An. stephensi* ranges from 0.6 Mb on the X chromosome to 3.8 Mb on the 3R arm. Thus, the large-scale genomic information cannot be easily transferred from the better studied *An. gambiae* to the less studied *An. stephensi*.

In this study, no interchromosomal transposition event has been identified between *An. gambiae* and *An. stephensi*. This agrees well with rare detection of transpositions in *Anopheles* and *Drosophila* using comparative mapping. Only two possible transposition events have been detected out of 157 clones hybridized to the *An. funestus* chromosomes.¹⁴ Similarly, a low rate of gene transposition has been found in genus *Drosophila*; two transposition events have been detected out of 328 clones, which resulted in 4.9×10^{-5} transpositions/gene/million years.¹⁵ One mechanism of gene transposition is retroposition that uses reverse transcription of RNA and insertion of the resulting cDNA into a different position.¹⁶ Another mechanism of transposition is transposon-mediated excision and insertion of genomic segments.¹⁷ A study on *Drosophila* suggested that ectopic exchange between highly similar small nuclear RNA (snRNA) sequences could be responsible for the transposition of gene *larval serum protein 1β*.¹⁸

Our study showed that the rate of possible gene duplication events was higher in the *An. stephensi* lineage than in the *An. gambiae* lineage. Thirty-six of 157 cDNA clones (22.9%) were localized to multiple chromosome sites in *An. funestus* and to unique locations in the *An. gambiae* genome. Only one clone (0.64%) had a unique location in the *An. funestus* while having multiple hits in the *An. gambiae* genome.¹⁴ Thus, both studies showed the low rate of gene duplication in the *An. gambiae* lineage. Although the gene duplication rate was higher in *An. funestus* than *An. stephensi*, the generation rate of species-specific genes had the opposite pattern. In our study, 24 of 246 (9.8%) genes were specific to *An. stephensi*. In a similar study, only four of 157 (2.5%) genes were specific to *An. funestus*.¹⁴ In a comparative study of transcriptomes, *An. funestus* and *An. gambiae* shared 94% of all transcripts derived from *An. funestus*,¹⁹ which shows the high degree of transcriptome similarity between the two African mosquitoes. In contrast, only 1,387 out of 3,946 unique transcripts of *An. stephensi* (35.1%) were mapped on the *An. gambiae* genome¹³ indicating the higher degree of transcriptome divergence between the Asian and African malaria vectors. Thus, the physical map presented here is useful for studying the transcriptome and will greatly complement a future

FIGURE 1. A physical map showing the locations of 345 markers on five chromosome arms of *Anopheles stephensi*. (') indicates multiple located clones with signals of equal intensity. (*) indicates primary hybridization sites of multiply located clones. Positions of breakpoints for polymorphic inversion are shown with lower case letters.

FIGURE 2. Fluorescent *in situ* hybridization of *Anopheles stephensi* cDNA and BAC clones labeled with Cy3 (red) and Cy5 (blue) to polytene chromosomes from ovarian nurse cells. (A) Hybridization of *An. stephensi*-specific cDNA to the heterochromatin of the X chromosome. (B) Mapping of the *An. stephensi* cDNA to a region outside of the common polymorphic inversion 2Rb. (C) Hybridization of two *An. stephensi* BAC clones with euchromatic regions on the 3R chromosomal arm. Arrows indicate signals of hybridization.

genome sequencing project for *An. stephensi* by facilitating the genome assembly.

Received June 28, 2010. Accepted for publication August 30, 2010.

Note: Supplemental tables and figure appear at www.ajtmh.com.

Acknowledgments: We thank Nora J. Besansky, Frank H. Collins, Abraham Eappen, Marcelo Jacobs-Lorena, and the Malaria Research and Reference Reagent Resource Center (MR4) for providing DNA clones for physical mapping. We thank Melissa Wade for editing the text.

Financial support: This work was supported by National Institutes of Health grant 1R21AI081023-01 and startup funds from Virginia Tech to Igor V. Sharakhov.

Authors' addresses: Maria V. Sharakhova, Ai Xia, and Igor V. Sharakhov, Department of Entomology, Virginia Tech, Blacksburg, VA, E-mails: msharakh@vt.edu, xia@vt.edu, and igor@vt.edu. Zhijian Tu, Department of Biochemistry, Virginia Tech, Blacksburg, VA, E-mail: jaketu@vt.edu. Yogesh S. Shouche, Lab 3, National Center for Cell Science, Pune - 411007, India, E-mail: yogesh@nccs.res.in. Maria F. Unger, Eck Institute for Global Health, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, E-mail: munger1@nd.edu.

REFERENCES

- Zhimulev IF, 1996. *Morphology and Structure of Polytene Chromosomes*. San Diego, CA: Academic Press, 1–490.
- Zheng L, Saunders RD, Fortini D, della Torre A, Coluzzi M, Glover DM, Kafatos FC, 1991. Low-resolution genome map of the malaria mosquito *Anopheles gambiae*. *Proc Natl Acad Sci USA* 88: 11187–11191.
- Zheng L, Benedict MQ, Cornel AJ, Collins FH, Kafatos FC, 1996. An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics* 143: 941–952.
- Dimopoulos G, Zheng L, Kumar V, della Torre A, Kafatos FC, Louis C, 1996. Integrated genetic map of *Anopheles gambiae*: use of RAPD polymorphisms for genetic, cytogenetic and STS landmarks. *Genetics* 143: 953–960.
- della Torre A, Favia G, Mariotti G, Coluzzi M, Mathiopoulos KD, 1996. Physical map of the malaria vector *Anopheles gambiae*. *Genetics* 143: 1307–1311.
- Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburger P, Atkinson PW, Baden H, de Berardinis V, Baldwin D, Benes V, Biedler J, Blass C, Bolanos R, Boscos D, Barnstead M, Cai S, Center A, Chaturverdi K, Christophides GK, Chrystal MA, Clamp M, Cravchik A, Curwen V, Dana A, Delcher A, Dew I, Evans CA, Flanagan M, Grundschober-Freimoser A, Friedli L, Gu Z, Guan P, Guigo R, Hillenmeyer ME, Hladun SL, Hogan JR, Hong YS, Hoover J, Jaillon O, Ke Z, Kodira C, Kokoza E, Koutsos A, Letunic I, Levitsky A, Liang Y, Lin JJ, Lobo NF, Lopez JR, Malek JA, McIntosh TC, Meister S, Miller J, Mobarry C, Mongin E, Murphy SD, O'Brochta DA, Pfannkoch C, Qi R, Regier MA, Remington K, Shao H, Sharakhova MV, Sitter CD, Shetty J, Smith TJ, Strong R, Sun J, Thomasova D, Ton LQ, Topalis P, Tu Z, Unger MF, Walenz B, Wang A, Wang J, Wang M, Wang X, Woodford KJ, Wortman JR, Wu M, Yao A, Zdobnov EM, Zhang H, Zhao Q, Zhao S, Zhu SC, Zhimulev I, Coluzzi M, della Torre A, Roth CW, Louis C, Kalush F, Mural RJ, Myers EW, Adams MD, Bork P, Brey PT, Venter JC, Weissenbach J, Kafatos FC, Collins FH, Hoffman SL, 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298: 129–149.
- Sharakhova M, Hammond MP, Lobo NF, Krzywinski J, Unger MF, Hillenmeyer ME, Bruggner RV, Birney E, Collins FH, 2007. Update of the *Anopheles gambiae* PEST genome assembly. *Genome Biol* 8: R5.
- Sharakhova MV, Xia A, McAlister SI, Sharakhov IV, 2006. A standard cytogenetic photomap for the mosquito *Anopheles stephensi* (Diptera: Culicidae): application for physical mapping. *J Med Entomol* 43: 861–866.
- Xia A, Sharakhova M, Leman S, Tu Z, Bailey J, Smith C, Sharakhov IV, 2010. Genome landscape and evolutionary plasticity of chromosomes in malaria mosquitoes. *PLoS ONE* 5: e10592.
- Xia A, Sharakhova MV, Sharakhov IV, 2008. Reconstructing ancestral autosomal arrangements in the *Anopheles gambiae* complex. *J Comput Biol* 15: 965–980.
- Lawson D, Arensburger P, Atkinson P, Besansky NJ, Bruggner RV, Butler R, Campbell KS, Christophides GK, Christley S, Dialynas

- E, Hammond M, Hill CA, Konopinski N, Lobo NF, MacCallum RM, Madey G, Megy K, Meyer J, Redmond S, Severson DW, Stinson EO, Topalis P, Birney E, Gelbart WM, Kafatos FC, Louis C, Collins FH, 2009. VectorBase: a data resource for invertebrate vector genomics. *Nucleic Acids Res* 37: D583–D587.
12. Rai KS, Black WC 4th, 1999. Mosquito genomes: structure, organization, and evolution. *Adv Genet* 41: 1–33.
13. Patil DP, Atanur S, Dhotre DP, Anantharam D, Mahajan VS, Walujkar SA, Chandode RK, Kulkarni GJ, Ghate PS, Srivastav A, Dayananda KM, Gupta N, Bhagwat B, Joshi RR, Mourya DT, Patole MS, Shouche YS, 2009. Generation, annotation, and analysis of ESTs from midgut tissue of adult female *Anopheles stephensi* mosquitoes. *BMC Genomics* 10: 386.
14. Sharakhov IV, Serazin AC, Grushko OG, Dana A, Lobo N, Hillenmeyer ME, Westerman R, Romero-Severson J, Costantini C, Sagnon N, Collins FH, Besansky NJ, 2002. Inversions and gene order shuffling in *Anopheles gambiae* and *A. funestus*. *Science* 298: 182–185.
15. Gonzalez J, Ranz JM, Ruiz A, 2002. Chromosomal elements evolve at different rates in the *Drosophila* genome. *Genetics* 161: 1137–1154.
16. Xing J, Wang H, Belancio VP, Cordaux R, Deininger PL, Batzer MA, 2006. Emergence of primate genes by retrotransposon-mediated sequence transduction. *Proc Natl Acad Sci U S A* 103: 17608–17613.
17. Long M, Betran E, Thornton K, Wang W, 2003. The origin of new genes: glimpses from the young and old. *Nat Rev Genet* 4: 865–875.
18. Gonzalez J, Casals F, Ruiz A, 2004. Duplicative and conservative transpositions of larval serum protein 1 genes in the genus *Drosophila*. *Genetics* 168: 253–264.
19. Serazin AC, Dana AN, Hillenmeyer ME, Lobo NF, Coulibaly MB, Willard MB, Harker BW, Sharakhov IV, Collins FH, Ribeiro JM, Besansky NJ, 2009. Comparative analysis of the global transcriptome of *Anopheles funestus* from Mali, West Africa. *PLoS ONE* 4: e7976.