



HAL
open science

Dramatic Changes in Malaria Population Genetic Complexity in Dielmo and Ndiop, Senegal, Revealed Using Genomic Surveillance

Amy K. Bei, Makhtar Niang, Awa B. Deme, Rachel F. Daniels, Fatoumata D. Sarr, Cheikh Sokhna, Cheikh Talla, Joseph Faye, Nafissatou Diagne, Souleymane Doucoure, et al.

► **To cite this version:**

Amy K. Bei, Makhtar Niang, Awa B. Deme, Rachel F. Daniels, Fatoumata D. Sarr, et al.. Dramatic Changes in Malaria Population Genetic Complexity in Dielmo and Ndiop, Senegal, Revealed Using Genomic Surveillance. *Journal of Infectious Diseases*, 2018, 217 (4), pp.622-627. 10.1093/infdis/jix580 . hal-01780650

HAL Id: hal-01780650

<https://hal.science/hal-01780650>

Submitted on 11 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dramatic Changes in Malaria Population Genetic Complexity in Dielmo and Ndiop, Senegal, Revealed Using Genomic Surveillance

Amy K. Bei,^{1,2,3,a} Makhtar Niang,^{4,a} Awa B. Deme,^{2,3} Rachel F. Daniels,^{1,5} Fatoumata D. Sarr,⁶ Cheikh Sokhna,⁷ Cheikh Talla,⁶ Joseph Faye,⁶ Nafissatou Diagne,⁷ Souleymane Doucoure,⁷ Souleymane Mboup,^{2,8} Dyann F. Wirth,^{1,5} Adama Tall,⁶ Daouda Ndiaye,³ Daniel L. Hartl,^{1,5,9} Sarah K. Volkman,^{1,5,10,a} and Aissatou Toure-Balde^{4,a}

¹Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; ²Laboratory of Bacteriology and Virology, Le Dantec Hospital, and ³Laboratory of Parasitology and Mycology, Faculty of Medicine and Pharmacy, Cheikh Anta Diop University, and ⁴Immunology Unit, Institut Pasteur de Dakar, Senegal; ⁵Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge; ⁶Epidemiology Unit, Institut Pasteur de Dakar, ⁷French National Research Institute for Sustainable Development, URMITE, and ⁸Institut de Recherche en Santé, de Surveillance Epidémiologique et de Formations, Dakar, Senegal; and ⁹Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, and ¹⁰School of Nursing and Health Sciences, Simmons College, Boston, Massachusetts

Dramatic changes in transmission intensity can impact *Plasmodium* population diversity. Using samples from 2 distant time-points in the Dielmo/Ndiop longitudinal cohorts from Senegal, we applied a molecular barcode tool to detect changes in parasite genotypes and complexity of infection that corresponded to changes in transmission intensity. We observed a striking statistically significant difference in genetic diversity between the 2 parasite populations. Furthermore, we identified a genotype in Dielmo and Ndiop previously observed in Thiès, potentially implicating imported malaria. This genetic surveillance study validates the molecular barcode as a tool to assess parasite population diversity changes and track parasite genotypes.

Keywords. malaria; Dielmo/Ndiop; Senegal; genomic surveillance; molecular barcode.

Traditional estimates of malaria transmission are often difficult and costly to perform, especially in areas of low transmission [1]. The ability to monitor and predict changes in transmission intensity in real time is critical to evaluating malaria parasite biology and malaria control efforts. Recently, a 24-single-nucleotide polymorphism (SNP) molecular barcode tool that correlates with whole-genome identity has assessed changes in malaria population

genomics and malaria incidence rates from cross-sectional studies in Thiès, Senegal [2], a region with very low malaria endemicity based upon an annual entomological inoculation rate (EIR) of 1–5.

Dielmo and Ndiop are 2 extensively described Senegalese villages 5 km apart where longitudinal follow-up studies on determinants of malaria infection have been conducted since 1990 and 1993, respectively [3–5]. Until the mid-2000s, Ndiop and Dielmo had contrasting malaria endemicity due to continual anopheline breeding year-round in Dielmo [4], thus sustaining intense and perennial malaria transmission (EIR of 258 infected bites/person/year during 1990–2006) [5], whereas transmission was moderate and seasonal (EIR = 80 infectious bites/person/year) in Ndiop [6]. Malaria transmission intensity has since declined substantially in both villages: EIR was 0.00 in Ndiop and 26.3 in Dielmo in 2014 (unpublished data).

The goal of this study was to apply the 24-SNP molecular barcode [7] to blinded samples from a longitudinal cohort spanning 25 years to evaluate the correlation between key genetic signals and malaria transmission and parasite population diversity in areas with dramatically different malaria endemicities. Specifically, the study aimed to characterize (1) the complexity of infection (COI); (2) the presence of any shared parasite clusters; and (3) whether the barcodes (genomic signature) of parasites from Dielmo/Ndiop match those previously identified throughout Senegal.

METHODS

Ethics Statement

The Ethics Committee of the Ministry of Health in Senegal approved this study. All samples were collected with informed consent per ethical requirements of the National Ethics Committee of Senegal, Pasteur Institutes of Dakar and Paris, and the Institut de Recherche pour le Développement.

Patient DNA Samples

All samples microscopically positive for *P. falciparum* infection were confirmed using quantitative polymerase chain reaction [8] following DNA extraction (QIAamp, Qiagen). Forty randomly selected blood samples from an archived collection (2001–2002 [n = 20] and 2014 [n = 20]) corresponding to periods of high transmission (Dielmo EIR = 353.8 [2001] and 409.9 [2002] [5]; Ndiop EIR = 171.9 [2001] and 16.9 [2002]) or extremely low transmission (Dielmo EIR = 26.3 and Ndiop EIR = 0.0 [2014, unpublished]), with sample sets blind-coded prior to analysis. DNA samples were preamplified as needed [7, 9].

Genotyping

24-SNP Molecular Barcoding

DNA was diluted, used directly, or preamplified as previously described [9]. Molecular barcoding was performed

Received 17 August 2017; editorial decision 31 October 2017; accepted 26 December 2017; published online January 9, 2018.

^aA. K. B., M. N., S. K. V., and A. T.-B. contributed equally to this work.

Correspondence: S. K. Volkman, ScD Harvard T.H. Chan School of Public Health, 665 Huntington Ave, Bldg 1, Rm 715, Boston, MA 02115 (svolkman@hsph.harvard.edu).

The Journal of Infectious Diseases® 2018;217:622–7

© The Author(s) 2018. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com. DOI: 10.1093/infdis/jix580

using high-resolution melting methodology [10] in 96-well plates on a Roche Light Cycler 96, with 8 assays per plate for 10 unknown samples, and 2 controls (either 3D7 and Dd2, or 3D7 and Tm90C6B, depending upon assay) in a final volume of 5 μ L. Occasionally assay confirmation was performed on a Light Scanner-32. Here, a single assay was run on 30 unknown samples and 2 controls (either 3D7 and Dd2, or 3D7 and Tm90C6B, depending upon assay) in a final volume of 10 μ L [11]. A novel SNP was identified in the IP20 sample for the A1 assay (Figure 1B) and confirmed by sequencing.

Barcode Phylogenetic Analysis

A neighbor-joining phylogenetic tree using the Jukes-Cantor method was created using Geneious (version 9.1.5, created by Biomatters).

Complexity of Infection Estimates

Complexity of infection was determined using the COIL algorithm [12], with monogenomic infections having <2 Ns and polygenomic infections having ≥ 2 Ns in the 24-SNP barcode.

Sample Size Estimates and Statistics

Sample size estimates and statistical significance was determined using Fisher exact test. The 2 blinded populations (IP1–IP20 and IP21–IP40) were evaluated first for bias, then for differences regarding (1) proportion of polygenomic infections and (2) repeated barcodes. To achieve 5% significance level for polygenomic vs monogenomic, 10 samples from the IP21–IP40 set and 7 samples from the IP1–IP20 set would be required, thus 20 samples per set exceeds the required number. To achieve 5% significance for unique monogenomic vs repeated monogenomic samples would require 5 samples from IP21–IP40 and 7 samples from IP1–IP20, such that 20 samples per set exceeds the required number. Thus, 20 samples per set is about 2 times the minimum required to detect statistically significant differences at the 5% level.

RESULTS

Patient Population

Demographic characteristics of the study population (Table 1) indicate the sex ratio (male/female) favored males (1.5 and 1.22 for the 2001–2002 and 2014 groups, respectively). The mean age was 6.37 (range, 0.7–16.1) years and 19.5 (range, 1.9–44.6) years for the 2001–2002 and 2014 groups, respectively. This shift in the age distribution is representative of the overall population, and is expected as malaria cases occur more frequently in older individuals as transmission declines [5]. Other statistically significant differences (highest in 2001–2002) included the mean parasite density, EIR, and malaria incidence.

Barcode Comparisons for COI and Genetic Diversity

We first tested whether the 2 barcode sample sets were balanced and found no significant difference ($P = .73$, Fisher exact test),

indicating no sampling bias between the 2 sample sets. This is consistent with neighbor-joining analysis (Figure 1C) that suggests admixing of parasite genotypes between the 2 study populations—Dielmo and Ndiop. We next estimated COI using the COIL algorithm for the 2 blinded sample sets. For samples IP1–IP20, COI was 1.35 (10/20 [50%] polygenomic samples), whereas samples IP21–IP40 had a COI of 1, as all samples contain single genotypes. These COI differences represent highly significant differences ($P = .0016$, Fisher exact test) in the number of monogenomic and polygenomic infections between 2001–2002 and 2014.

We also observed dramatic differences between these 2 sample populations when comparing genetic diversity and identity of specific parasite types. For samples IP1–IP20 there were no repeated barcodes (all were unique), whereas the IP21–IP40 set contained 3 repeated barcode clusters ($n = 2$, $n = 2$, and $n = 6$) that represented 66.6% (10/15) of the population (Figure 1A). While the number of barcode clusters in IP21–IP40 ($n = 3$) compared to IP1–IP20 ($n = 0$) was only borderline significant ($P = .069$, Fisher exact test), the difference in the number of repeated barcode samples for IP21–IP40 ($n = 10$) compared to IP1–IP20 ($n = 0$) was highly significant ($P = .0010$, Fisher exact test). These results support the hypothesis that samples IP1–IP20 originated from a period of relatively high malaria transmission based upon increased genomic diversity, whereas samples IP21–IP40 were collected during a period of relatively low malaria transmission given decreased genomic diversity. When samples were unblinded, the findings confirmed the hypothesis that the IP1–IP20 sample set was from a highly malaria-endemic 2001–2002 period characterized by high transmission intensity, whereas the IP21–IP40 set was from the lowly malaria endemic 2014 collection, where transmission intensity had dropped substantially.

Specific Barcode Identity Compared to Known Barcodes Implicates Possible Imported Malaria

In addition to genetic diversity, we also asked whether any of these genotypes matched a large database ($n = 1034$ monogenomic barcodes) of previously published [2] and unpublished barcodes (Deme and Daniels, unpublished data: $n = 215$ monogenomic barcodes) from other regions of Senegal. For samples from the high transmission period (2001–2002), none of the individual barcodes matched previously observed haplotypes. For samples from the low transmission period (2014), 1 of the 3 repeated barcode clusters ($n = 6$) corresponded to a parasite type (haplotype 3), observed in Thiès in both 2007 and 2010 [2]. The other 2 clusters (IP1, $n = 2$; IP2, $n = 2$) had not previously been observed, with one set (IP1) having a member each from Dielmo and Ndiop and the other set (IP2) having both cluster members from Ndiop.

We found no other matching barcodes in a large database ($n = 1249$) of published [2] and unpublished Senegal barcodes

Table 1. Baseline Demographic Characteristics of the Study Population

Characteristic	2001–2002 (n = 20)	2014 (n = 20)	OR (95% CI) P Value
Sex, No. (%)			
Male	12 (60)	11 (55)	1
Female	8 (40)	9 (45)	1.22 (.29–5.16)
Ratio (M/F)	1.5	1.22	
Age, y			
Mean	6.37	19.5	.00048
Range	(0.7–16.1)	(1.9–44.6)	
Parasite density/μL			
Mean	11 757.5	6283.2	.0038
Range	(150–29 450)	(25–45 000)	
Location, No. (%)			
Dielmo	10 (50)	11 (55)	
Ndiop	10 (50)	9 (45)	
mEIR			
Dielmo	29.6 ^a	2.19	<.0001
Ndiop	7.9 ^b	0.0 ^c	.06
IR, mean (95% CI)			
Dielmo	2.25 (2.13–2.37)	0.136 (0.099–0.173)	<.0001
Ndiop	1.98 (1.865–2.087)	0.104 (0.072–0.136)	<.0001

Characteristics of the sample populations from Dielmo and Ndiop are listed for the 2 time periods described in this study, 2001–2002, and 2014. Kruskal–Wallis and Fisher exact tests were used to compare distribution of quantitative and qualitative variables respectively between subgroups. For all statistical tests, $\alpha = .05$.

Abbreviations: CI, confidence interval; IR, annual incidence rate (number of episodes per person per year); mEIR, mean entomological inoculation rate (number of infected bites per person per year, averaged over the year); ND, not determined; OR, odds ratio.

^aFrom previously published data [5].

^bFrom previously published data [6].

^cUnpublished data.

from Thiès, Dakar, Kaolack, Kaffrine, Kedougou, Kolda, Saint Louis (Figure 1D). Travel histories for the 6 individuals with the haplotype 3 genotype, previously identified in Thiès (Figure 1D), indicated both extensive local travel as well as travel to the Thiès region. Thus, genetic evidence of shared parasite genotypes between Dielmo/Ndiop and Thiès was confirmed by travel history.

DISCUSSION

Changes in transmission can dramatically influence the genetic diversity of circulating *P. falciparum* parasites. To test this hypothesis, we characterized the genetic complexity at 2 time points from a well-characterized longitudinal cohort to determine if genomic approaches can predict transmission intensity [2]. We characterized COI, polygenomic vs monogenomic infections, and evidence of shared genetic clusters, and asked whether the barcodes of these Dielmo/Ndiop parasites match parasite types previously identified throughout Senegal.

We observed a striking and statistically significant difference in the genetic signatures between the 2 blinded sample populations, with one (2001–2002) population having a high percentage of polygenomic infections (50%), no shared clusters, and no previously observed barcodes, whereas the other (2014) population contained only monogenomic infections and 3 shared clusters representing two-thirds (10/15 [66.6%]) of the population diversity. Such changes in genetic signatures were true for

both Dielmo and Ndiop if considered separately. Monogenomic proportions went from 40% (Ndiop) and 60% (Dielmo) in 2001–2002 to 100% monogenomic for both populations in 2014. Conversely, barcode clonality went from 0 to 66% in both Dielmo and Ndiop across this same time, with both signals mirroring decline in EIR as a measure of transmission. These trends in genotypic diversity and complexity mirror expectations for a population with high outcrossing (2001–2002) or increased selfing (2014) [13, 14]. These striking changes represent statistically significant differences despite relatively small sample sizes, which is one of the advantages of genetic data. Indeed, for the genetic signals we were testing, we only needed approximately ≤ 10 samples to achieve significance (5% level with a Fisher exact test).

Parasite genotypes are shared between Dielmo and Ndiop based upon neighbor-joining tree analysis (Figure 1C) that revealed no strict clustering of barcodes by locality. Nevertheless, we saw distinct patterns across time that corresponded to dramatic changes in transmission levels. Among these patterns, we detected a cluster of parasites (haplotype 3) that was detected in Thiès in 2007 and 2010 [2], suggesting possible migration of parasites either by infected humans or infected mosquitoes. These genetic findings are supported by travel history, as a subset of individuals who harbored these shared barcode infections (ND14 and DI12) had traveled to the Thiès region. Thus, genetic signatures may help in tracking parasite infections and networks of transmission.

Our findings are generalizable and consistent with observations both in other parts of Senegal [2] as well in other geographic areas. For example, a recent study used barcode identity to suggest that the source of an outbreak of *P. falciparum* malaria in Panama was possibly from neighboring Colombia [10]. The finding of shared barcodes between Thiès and Dielmo/Ndiop, in addition to shared barcodes between Dielmo and Ndiop emphasize the need to evaluate domestic importation of parasite types, which may be interrogated with other surveillance methods [15].

This study validates and confirms the use of genetic surveillance to monitor malaria transmission changes, but additional studies are needed to more comprehensively investigate parasite population dynamics with changing malaria epidemiology resulting from malaria-reducing interventions. Additional investigation will extend this study to the larger longitudinal cohort to determine precise inflection points of transmission intensity across time and their correlation with parasite population genomic complexity. Such sampling will better inform the minimal essential data required to detect changes in R_0 through epidemiological modeling [2]. Such work will expand the barcode database for Dielmo and Ndiop and inform the underlying genetic diversity of the region across time. Integration of genetic data with patient clinical and behavioral data (such as travel) can guide definition of “local” circulating vs imported malaria genotypes for parasite tracking, informing transmission networks and identifying “sink” and “source” populations toward successful malaria elimination [2, 15].

This study illustrates the utility of *Plasmodium* molecular genetic surveillance to assess parasite population complexity as transmission changes, to detect changes in parasite population structure that signal success or failure of interventions and impact policy decisions. Furthermore, this study illustrates the value of employing genomic surveillance to monitor the movement of specific parasite genotypes between regions that, with larger sample sizes and database of region-specific parasite genomes, could be used to track imported malaria strains from other regions and guide effective elimination activities.

Notes

Acknowledgments. The longitudinal cohorts in Dielmo and Ndiop were funded by The Pasteur Institutes of Dakar and Paris, the Institut de Recherche pour le Développement (IRD), and the Senegalese Ministry of Health. We acknowledge the villagers of Dielmo and Ndiop for their continuous support and participation in the cohort. We thank the creators of Geneious (Biomatters) for their generous donation of a 1-year license to scientists in malaria-endemic regions, including author M. N., attending the Harvard-Broad Barcode Workshop.

Financial support. The longitudinal cohorts in Dielmo and Ndiop were funded by the Pasteur Institutes of Dakar and Paris, IRD, and the Senegalese Ministry of Health. This work was supported by a grant from the Bill & Melinda Gates Foundation

(to D. F. W.). A. K. B. is supported by an International Research Scientist Development Award (award number 1K01TW010496). M. N.’s attendance at the Harvard-Broad Barcode Workshop was supported by the Bill & Melinda Gates Foundation.

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Tusting LS, Bousema T, Smith DL, Drakeley C. Measuring changes in *Plasmodium falciparum* transmission: precision, accuracy and costs of metrics. *Adv Parasitol* **2014**; 84:151–208.
2. Daniels RF, Schaffner SF, Wenger EA, et al. Modeling malaria genomics reveals transmission decline and rebound in Senegal. *Proc Natl Acad Sci U S A* **2015**; 112:7067–72.
3. Rogier C, Ly AB, Tall A, Cissé B, Trape JF. *Plasmodium falciparum* clinical malaria in Dielmo, a holoendemic area in Senegal: no influence of acquired immunity on initial symptomatology and severity of malaria attacks. *Am J Trop Med Hyg* **1999**; 60:410–20.
4. Trape JF, Rogier C, Konate L, et al. The Dielmo project: a longitudinal study of natural malaria infection and the mechanisms of protective immunity in a community living in a holoendemic area of Senegal. *Am J Trop Med Hyg* **1994**; 51:123–37.
5. Trape JF, Tall A, Sokhna C, et al. The rise and fall of malaria in a West African rural community, Dielmo, Senegal, from 1990 to 2012: a 22 year longitudinal study. *Lancet Infect Dis* **2014**, 14:476–488.
6. Rogier C. Natural history of *Plasmodium falciparum* malaria and determining factors of the acquisition of anti-malaria immunity in two endemic areas, Dielmo and Ndiop (Senegal). *Bull Mem Acad R Med Belg* **2000**; 155:218–26.
7. Daniels R, Volkman SK, Milner DA, et al. A general SNP-based molecular barcode for *Plasmodium falciparum* identification and tracking. *Malar J* **2008**; 7:223.
8. Canier L, Khim N, Kim S, et al. An innovative tool for moving malaria PCR detection of parasite reservoir into the field. *Malar J* **2013**; 12:405.
9. Mharakurwa S, Daniels R, Scott A, Wirth DF, Thuma P, Volkman SK. Pre-amplification methods for tracking low-grade *Plasmodium falciparum* populations during scaled-up interventions in southern Zambia. *Malar J* **2014**; 13:89.
10. Obaldia N 3rd, Baro NK, Calzada JE, et al. Clonal outbreak of *Plasmodium falciparum* infection in eastern Panama. *J Infect Dis* **2015**; 211:1087–96.
11. Daniels R, Ndiaye D, Wall M, et al. Rapid, field-deployable method for genotyping and discovery of single-nucleotide polymorphisms associated with drug resistance in *Plasmodium falciparum*. *Antimicrob Agents Chemother* **2012**; 56:2976–86.

12. Galinsky K, Valim C, Salmier A, et al. COIL: a methodology for evaluating malarial complexity of infection using likelihood from single nucleotide polymorphism data. *Malar J* **2015**; 14:4.
13. Conway DJ, Roper C, Oduola AM, et al. High recombination rate in natural populations of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **1999**; 96:4506–11.
14. Anderson TJ, Haubold B, Williams JT, et al. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Mol Biol Evol* **2000**; 17:1467–82.
15. Wesolowski A, Eagle N, Tatem AJ, et al. Quantifying the impact of human mobility on malaria. *Science* **2012**; 338:267–70.