Trends in Parasitology

Contrasting Epidemiology and Genetic Variation of Duffy Negative *Plasmodium vivax* across Africa

Eugenia Lo[1*], Kareen Pestana[1], Daniel Kepple[1], Beka Raya[2], Lemu Golassa[3], Muzamil Mahdi Hamid[4], Delenasaw Yewhalaw[2], Gianluca Russo[5], Giacomo Paganotti[6]

[1] Biological Sciences, University of North Carolina at Charlotte, USA

[2] Tropical Infectious Disease Research Center, Jimma University, Jimma, Ethiopia

[3] Aklilu Lemma Institute of Pathobiology, Addis Ababa University, Addis Ababa, Ethiopia

[4] Department of Parasitology and Medical Entomology, Institute of Endemic Diseases, University of Khartoum, Khartoum, Sudan

[5] Public Health and Infectious diseases, Sapienza University, Rome, Italy

[6] Botswana-UPenn Partnership, University of Botswana Main Campus, Gaborone, Botswana

* Correspondence author: Eugenia Lo; email: eugenia.lo@uncc.edu

Abstract

Plasmodium vivax uses the Duffy antigen/chemokine receptor for entry into human erythrocytes. However, several recent studies have showed that *P. vivax* can infect Duffy-negative Africans and that they are no longer completely resistant to *P. vivax*. While the reports of *P. vivax* cases in different parts of Africa raise serious public health concerns of a potential spread across the continent, the epidemiology and biology of *P. vivax* malaria in Africa remain largely unclear. Here, we compare the prevalence, parasitemia, and microsatellite-based variation of *P. vivax* in Duffy negative individuals from West-Central, Southern, and East Africa, as well as with Duffy positive *P. vivax* infections collected in the same areas. The proportion of Duffy negativity ranges widely from 99% in West African to 30-40% in East African countries. Considerable differences were observed in Duffy-negative *P. vivax* prevalence among Cameroon, Botswana, and Ethiopia. *Plasmodium vivax* parasite density in Duffy-negative infections is significantly lower than in Duffy-positive infections, regardless of geographical or ethic group differences. Maximum likelihood analyses of the Duffy binding protein (DBP) sequences indicated that Duffy-negative *P. vivax* isolates were not monophyletic and did not share from a single source origin. Instead, they were found in multiple well-supported clades without clear geographical boundary. By contrast, analyses of five microsatellite loci showed clear genetic structure among the African isolates. The Duffy-negative *P. vivax* from Cameroon, Botswana, and Ethiopia each constituted distinct genetic clusters, different from the Duffy-positive *P. vivax*. *P. vivax* from Ethiopia displays the greatest diversity that constituted admixed clusters resembling both the Duffy-positive and Duffy-negative isolates from Botswana. *P. vivax* from Cameroon forms distinct lineages. These

findings help clarify the genetic origin and spreading pathways of *P*. *vivax* in Africa. This information will contribute to our limited knowledge of *P. vivax* epidemiology and biology in Africa.

Introduction

Vivax malaria was previously thought to be rare or absent in African populations who lack the Duffy blood group antigen expression [1,2]. A point mutation (c.1-67T>C; rs2814778) in the GATA-1 transcription factor binding site of the Duffy antigen/receptor for chemokines (*DARC*) gene promoter alters erythroid expression, eliminating Duffy antigen expression on the surface of the red blood cells [3-5]. However, recent studies reported several cases of *P. vivax* infection in Duffy-negative people in different parts of Africa [6-8], including countries where Duffy-negatives are predominant [9-13]. In addition, 29 African countries including six previously undocumented endemic countries (Benin, Comoros, Mozambique, Senegal, Zambia and Zimbabwe) have reported *P*. *vivax* clinical cases, infected vectors or asymptomatic parasitemia [13]. These reports clearly indicate that the endemic range of *P. vivax* has extended beyond East Africa and penetrated in areas of very high (>95%) Duffy-negativity, where *P. vivax* was previously thought to be absent [6,14]. Such a distribution raises important questions of how *P. vivax* infects Duffy-negative erythrocytes, whether *P. vivax* infections transmit in Duffy-negative populations and between Duffy-negative and Duffy-positive individuals, and what factors govern *P. vivax* transmission in Africa.

While *P. falciparum* is considered to be the deadliest malaria parasite with the most severe clinical outcomes, *P. vivax* is more widespread and often associated with high levels of morbidity. Despite the fact that several case reports from almost all countries across the African continent are emerging from various entomological and serological studies, community surveys, and clinical records [6,14,15], the epidemiology of *P. vivax* in Africa and the parasite invasion mechanism in Duffy-negative individuals remain poorly

understood. Compared to *P. falciparum*, *P. vivax* has a broader temperature tolerance, an earlier onset of gametocyte development, and can form dormant hypnozoites causing relapse [16], enabling *P. vivax* to spread through the diverse African climate and outcompete *P. falciparum* [17]. These biological factors certainly make *P. vivax* malaria difficult to control and eliminate, highlighting the concern of this 'new' *P. vivax* strains that infect Duffy-negative hosts to spread through much of Africa and result in substantial, negative public health and economic impacts [18].

Erythrocyte invasion by *Plasmodium* involves multiple interactions between parasite ligands and host receptors, some of which have overlapping and partially redundant roles [19-21]. In *P. falciparum*, there are several established invasion ligands such as EBA-175, EBA-181/JESEBL and EBA-140/BAEBL; and different strains use different subsets of ligands to catalyze invasion [22-25]. In *P. vivax*, only a single *P. vivax* ligand-receptor interaction has so far been studied in any detail, *P. vivax* Duffy Binding Protein (*PvDBP*1). Previous study has shown that mutations in *PvDBP*1 region II unique to *P. vivax* in Duffy-negative people in Ethiopia did not lead to binding of Duffy-negative erythrocytes [26]. Salvador (Sal) I *P. vivax* infects Squirrel monkeys without *PvDBP*1 binding to Squirrel monkey erythrocytes [26]. Further, EBP/DBP2 region II, a paralog of *PvDBP*1, was shown to bind to Duffy-positive and Duffy-null human erythrocytes at low frequency [26,27], despite being deleted in Sal-I *P. vivax*. These findings suggested that there are other Duffy-independent pathways that enable erythrocyte invasion and explain the widespread phenomenon of *P. vivax* infections in Africa.

To date, the documentation of *P. vivax* infections across Africa is diverse, context-specific, and primarily driven by the specific objectives of isolated clinical or

epidemiological activities [8]. The varied diagnostic and methodological approaches used across studies have limited our ability to identify distinct epidemiological characteristics of *P*. *vivax* between regions. This situation is concerning because there is no comprehensive genetic and epidemiological data of *P. vivax* in Africa available to National Malaria Programs or World Health Organization to assess impacts and confer control strategies. There is yet limited data that illustrates the ability of *P. vivax* to spread to populations such as those in West Africa where the Duffy-negative allele is fixed. The lack of scientific insights and public awareness to this phenomenon can lead to an increased risk and spread of *P. vivax* infections in Africa. Therefore, this study utilized a standardized assay and provided comparisons of epidemiological attributes of *P. vivax* among West-Central, Southern, and East Africa, where *P. vivax* cases were previously reported. Specifically, we aimed to (1) determine the prevalence of Duffy negativity and *P. vivax* infections across different African countries; (2) compare *P. vivax* parasitemia between Duffy negative and Duffy positive infections collected from the same area; and (3) infer the genetic relationships among the African *P. vivax* as well as with the other global isolates. In addition, we provided an overview of some recent findings and major knowledge gaps in the evolutionary and epidemiological aspects of *P. vivax* in Africa. Comparing the epidemiological and genetic features of *P. vivax* from different parts of Africa will fill critical gap in understanding how widespread this phenomenon is impacting malaria control.

Materials and Methods

Ethics statement

Scientific and ethical clearance was given by the institutional scientific and ethical review boards of Jimma University (Ethiopia), The Cameroon Bioethics Initiative (Ref. CBI/249/ERCC/CAMBIN), The Human Research Ethnics Committee of the Ministry of Health (Botswana), University of Botswana, and University of North Carolina at Charlotte (USA). Written informed consent/assent for study participation was obtained from all consenting heads of households, parents/guardians (for minors under 18 years old), and each individual who was willing to participate in the study.

Study sites and sample collection

A total of 838 febrile patients collected from four study sites including Jimma and Bonga in Ethiopia, Dschang in Cameroon, and Tutume in Botswana were included in this study (**Figure 1**). Finger-prick blood samples were obtained from patients who visited the health centers/hospitals at each of the study sites. Thick and thin blood smears were prepared for each subject to screen for *P. vivax* by microscopy. Parasites were counted against 200 leukocytes and a slide was considered negative when no parasites were observed after counting over 100 microscopic fields. All slides were read in duplicate by two microscopists at the time of sample collection. The density of parasitemia was expressed as the number of asexual *P. vivax* per microliter of blood, assuming a leukocyte count of 8000 per microliter. Three to four blood spots, each equivalent to ~50µl of blood, were blotted on Whatman 3MM filter paper from each participating individual. Parasite DNA was extracted from dried blood spots by the Saponin/Chelex method [28] and genomic

DNA was eluted in a total volume of 200 µl TE buffer. Eluted DNA was used for PCR diagnosis and genotyping of malaria parasites.

Molecular screening of *P. vivax*

Parasite gene copy number was estimated using qPCR, specifically the SYBR Green detection method [29,30] using published primers (forward: 5′-GAATTTTCTCTTCGGAGTTTATTCTTAGATTGC-3′; reverse: 5′GCCGCAAGCTCCACGCCTGGTGGTGC-3′) specific to *P. vivax* that targeted the 18S rRNA genes. Amplification was conducted in a 20 µl reaction mixture containing 2 µl of genomic DNA, 10 µl SYBR Green qPCR Master Mix (Thermo Scientific), and 0.5 uM primer. The reactions were performed in QuantStudio Real-Time PCR Detection System (Thermo Fisher), with an initial denaturation at 95°C for 3 min, followed by 45 cycles at 94°C for 30 sec, 55°C for 30 sec, and 68°C for 1 min with a final 95°C for 10 sec. This was followed by a melting curve step of temperature ranging from 65°C to 95°C with 0.5°C increments to determine the melting temperature of each amplified product. Each assay included positive controls of *P. vivax* Pakchong (MRA-342G) and Nicaragua (MRA-340G) isolates, in addition to negative controls, including uninfected samples and water. A standard curve was produced from a ten-fold dilution series of the *P. vivax* control plasmid to determine the efficiency and detection limit of the qPCR. Melting curve analyses were performed for each amplified sample to confirm specific amplifications of the target sequence. The slope of the linear regression of threshold cycle number (*Ct*) versus log10 (gene copy number) was used to calculate amplification efficiency of each plate run based on internal standard controls. For the measure of reproducibility of the threshold cycle

number, the mean *Ct* value and standard error was calculated from three independent assays of each sample. A cut-off threshold of 0.02 fluorescence units that robustly represented the threshold cycle at the log-linear phase of the amplification and above the background noise was set to determine *Ct* value for each assay. Samples yielding *Ct* values higher than 40 (as indicated in the negative controls) were considered negative for *Plasmodium* species. The amount of parasite density in a sample was quantified by converting the *Ct* values into gene copy number (GCN) using the follow equation: $GCN_{sample} = 2^{E\times(40-Ctsample)}$; where GCN stands for gene copy number, *Ct* for the threshold cycle of the sample, and E for amplification efficiency. The differences in the log-transformed parasite GCN between samples among the study sites were assessed for significance at the level of 0.05 by one-tailed t-tests. Variations in GCN among samples were presented as boxplots showing the median and interquartile range values.

Duffy blood group genotyping

For all *P. vivax* positive samples, a 1,100-bp fragment of the human *Duffy blood group antigen* gene that encompasses the GATA1 transcription factor-binding site of the gene promoter was amplified using previously published primers [31]. PCR reaction contained 20µl DreamTaq PCR Mastermix, 1µl DNA template, and 0.5µl each primer. PCR conditions were: 94°C for 2-min, followed by 35 cycles of 94°C for 20s, 58°C for 30s, and 68°C for 60s, followed by a 4-min extension. PCR products were sequenced, and the chromatograms were visually inspected to determine the *Fy* genotypes.

Phylogenetic analyses

We amplified and obtained *PvDBP* sequences of 4 Duffy-positive and 4 Duffy-negative *P. vivax* samples from Botswana, 107 Duffy-positive and 9 Duffy-negative *P. vivax* samples from Ethiopia, and 53 Duffy-positive and 16 Duffy-negative *P. vivax* samples from Sudan. We were unable to obtain unambiguous, high-quality DNA sequences from the *P. vivax* Cameroon samples, and thus these samples were not used for phylogenetic analyses. Sequences obtained in the present study were aligned with 36 previously published *P. vivax* isolates from Africa (Uganda: *n*=31, Madagascar: *n*=4, and Mauritania: *n*=1) as well as from South America (*n*=81), East and Southeast Asia (*n*=64), Oceania Pacific (*n*=31), and Central Asia (*n*=77; **Supplementary Table 1**). Duffy status of these published sequences are unknown. The DBP sequence of Sal-1 (NC_009911.1) and EBP sequence of *P. cynomolgi* (Y11396.1) were used as outgroups. Alignment was performed with MUSCLE based on default settings [32] and manually edited in BioEdit v7.2.5 [33]. Phylogenetic trees were reconstructed using the maximum likelihood method implemented in RAxML v8.0 [34] with 500 bootstrap replicates to assess clade support on two datasets: (1) *PvDBP* sequences of only the African isolates and (2) *PvDBP* sequences of all global isolates.

Results and Discussion

Contrasting proportion of Duffy-negatives and *P. vivax* prevalence

Duffy genotyping for a collection of 176-390 febrile patients shows diverse proportions of Duffy-negatives in Cameroon, Botswana, Ethiopia, and Sudan (**Figure 1**). In Cameroon,

over 90% of the populations is at risk of malaria and 41% of the individuals have at least one episode of malaria each year [35]. *Plasmodium vivax* has been recently reported across multiple sites in southwestern Cameroon [11,36,37]. In Dschang, the clinical incidence of *P. vivax* ranged from 5.6% during the dry season to 36.2% during the rainy season (**Table 1)**. Based on Duffy genotyping, over 99% of the febrile patients are Duffy-negative (**Table 1; Figure 1**). During a collection of 500 clinical cases over a 6-month period in 2017, 273 (54.6%) were diagnosed as malaria infections by microscopy and 181 (36.2%) were confirmed as *P. vivax* by PCR. Almost all of the *P. vivax* cases were submicroscopic and detected in Duffy-negative individuals. The ability of *P. vivax* to invade Duffy-negative erythrocytes and the reports of clinical cases may imply a large population of asymptomatic *P. vivax* carriers in the community. According to a previous active case surveillance in Bolifamba, Cameroon, the prevalence of *P. vivax* is 4.4% (12/269 individuals), where Duffy-negatives accounted for 50% (6/12) of the *P. vivax* infections [36].

In Botswana, the proportion of Duffy-negative was 78.4% (236/301) among asymptomatic individuals and 83.5% (147/176) among febrile patients (**Figure 1**). Vivax malaria was first reported in asymptomatic children in a survey during the 2012-2013 transmission season [10] (**Table 1**). The average rate of asymptomatic *P. vivax* cases was 4.7%, but with large variation among districts. Tutume and Kweneng East (**Table 1**), which accounted for most of the *P. vivax* cases in Botswana, had 16.9% (54/320) and 13.6% (93/686) infection rates [10]. In Kweneng East, our analyses of 301 febrile patients indicated that 3% (9/301) of the asymptomatic schoolchildren were detected as *P. vivax*-positive by qPCR assay. Among them, eight were Duffy-negative (C/C) and one was

Duffy-positive (T/C) (**Table 1; Supplementary Table 2**). In Tutume, 6.8% (12/176) of the febrile patients were detected with *P. vivax* and 10 of them were Duffy-negative.

Vivax malaria is a significant problem in Ethiopia [30,38,39]. According to our passive surveillance in Jimma, among a collection of 358 febrile patients over a 7-month period in 2017, 36% (129/358) were Duffy-negatives (**Figure 1**) and 37.4% (134/358) were detected with *P. vivax* (**Table 1**). Among the confirmed *P. vivax* infections, 11.9% (16/134) were confirmed as being from Duffy-negatives. Likewise, in Bonga, 30.3% (125/413) of the febrile patients were detected with *P. vivax* and 3.2% (4/125) were from Duffy-negatives (**Table 1**). For these 20 Duffy-negatives *P. vivax* infections, microscopy, nested and quantitative PCRs indicated that 16 were single infections and four were mixed with *P. falciparum*. According to an earlier surveillance in Asendabo, 35% (137/390) of the general population is Duffy-negative [30]. The asymptomatic prevalence of *P. vivax* is 5.9% (23/390) and Duffy-negatives accounted for 8.7% (2/23) of the *P. vivax* infections (**Table 1**). These two asymptomatic cases were single infections. The finding of a Duffy-negative rate of 35% in the general population in Ethiopia is consistent with the high ethnic diversity and complex admixture history in East Africa [40].

In Sudan, most cases of malaria are caused by *P. falciparum* but in recent years there has been an increase in detection of *P. vivax* [77, 86]. In a passive surveillance study of febrile patients conducted over a 6-month period between 2017 and 2018, 992 samples were found to be microscopic-positive for malaria. Of those 992, 190 were confirmed to be *P. vivax* positive by PCR and qPCR assays [77] (**Table 1**). Further testing revealed that a subset of 4 were mixed infections with *P. vivax* and *P. falciparum* [77]. In Khartoum, the prevalence of *P. vivax* was XX% (101/XX) and in Duffy-negatives

accounted for 8.9% (9/101) of infections [77]. During the same collection period in New Halfa and River Nile, the *P. vivax* rate of infection was XX% (40/XX) and XX% (49/XX), respectively. While Duffy-negatives only accounted for a small percentage of infections in Khartoum, they represented 17.5% (7/40) of *P. vivax* infections in New Halfa and 36.7% (18/49) and River Nile [77]. The varying rates of Duffy-negative individuals among the different study sites may be explained by various groups and admixture patterns. Many population groups in Sudan are dominated by Nilotic and Euroasian admixtures with minimal West African affinity. One such exception is the Afro-Asiatic speaking Hausa population in Northeastern Sudan which have recently migrated from West Africa within the past 300 years. Although evidence has shown the Hausa have remain largely genetically isolated from neighboring populations [87].

Duffy-negativity among ethnic groups and historical human movement in Africa

Historical human movement and human genetics are highly relevant to *P. vivax* distribution in Africa. Recent genome-wide studies of African populations have refined earlier models of the continent's history and its impact on genetic diversity. The Bantu expansion and population admixture are two main historical events that shape the present distribution and genetic make-up of ethnic groups across Africa. The Bantu and Khoisan are two major ethnic groups in West-Central and Southern Africa. The Bantu heartland was in the region between southern Nigeria and Cameroon [41]. Several hypotheses suggest that the Bantu migrated towards East Africa where other ethnic groups such as the Cushitic and Nilotic dominated, potentially around 2,000 years ago [42] (**Figure 2A**). A component of Bantu ancestry (in particular ancestry from Cameroon) has been found

in the Southern African Khoisan, indicating that the West-Central African populations may have also reached south of the continent within the last 750 years and mixed with the indigenous Khoisan [42,43] (**Figure 2A**). It is apparent that the Khoesan groups of Southern Africa were originally and mostly Duffy-positive ancestors [44]. The Southern African populations, particularly the Bantu-related population of Botswana, carry a variable Khoesan ancestry due to admixture events between Bantu and pre-existing hunter-gatherer populations [45-48]. The Duffy-null allele from Bantu of West-Central Africa, which is thought to have arisen at least 30,000 years ago, could have spread throughout the southern part of the continent through gene flow between populations [49,50]. Our data showing a rate of 82.9% Duffy negativity in the Botswanan population is consistent with the Bantu expansion and admixture theories. While the direction of the Bantu expansion is still in debate, there is evidence showing that the Ethiopian and Sudanese population, with an admixture of several Eurasian ancestries and some Nilotic and Semitic-Cushitic components, migrated south after the Bantu expansion 2-5 thousand years [43,51-53] (**Figure 2B**). These migrations could have spread *P. vivax* infections from West-Central to other parts of Africa. Our data showing a rate of 35% Duffy negativity in Southwestern Ethiopia and East Sudan is consistent with the complex admixture history.

Low parasitemia in symptomatic Duffy-negative *P. vivax* infections and implications on invasion mechanism

Plasmodium vivax parasite density in Duffy-negative infected individuals is significantly lower than the Duffy-positive infected individuals, regardless of geographical or ethic

group differences (**Figure 3**). This difference might suggest a lower invasion capability of *P. vivax* in Duffy-negative than Duffy-positive individuals. The Duffy-negative *P. vivax* samples in Ethiopia and Sudan showed a greater range of parasitemia variation than those in Cameroon and Botswana. This could be due to the difference in sample size (**Figure 3; Supplementary Table 2**). In Botswana and Ethiopia, Duffy-positive and Duffy-negative individuals live side-by-side. In very few cases the asexual parasites were detected by microscopy in Duffy-negative individuals. The public health burden, economic impact, and severity associated with vivax malaria has been underestimated for long until recent studies showing the association of *P. vivax* infections with severe malaria and death [54,55]. The clinical spectrum of *P. vivax* malaria ranges from asymptomatic parasitemia and uncomplicated febrile illness to severe and fatal malaria. Other severe clinical manifestations include multiorgan dysfunction with anemia and thrombocytopenia [56]. Another major public health concern of *P. vivax* is its association with spontaneous abortions, premature and low birth weight in pregnant women. These clinical features have mostly been described for Duffy-positive populations and may vary among Duffy-negative individuals in Africa. In this study, the Duffy-negative individuals who were infected with *P. vivax* were mostly submicrocopic and exhibited fever at the time of sample collection. It is unclear if the spectrum of clinical symptoms is different from those in the Duffy-posiitve patients.

Erythrocyte invasion involves multiple interactions between parasite ligands and host receptors [19-21]. Mutations in *PvDBP*1 region II unique to *P. vivax* in Duffy-negative people in Ethiopia did not lead to binding of Duffy-negative erythrocytes and Sal-I *P. vivax* infects Squirrel monkeys without *PvDBP*1 binding to Squirrel monkey erythrocytes [26],

suggesting that there are other Duffy-independent pathways that enable erythrocyte invasion. For example, EBP/DBP2 region II has shown to bind to Duffy-positive and Duffy-null human erythrocytes at low frequency [26,27], despite being deleted in Sal I *P. vivax*. CD71 (Transferrin Receptor 1, TfR1) has been shown to bind readily with the reticulocyte binding proteins based on *in-vitro* experiments [57,58]. Given reticulocytes constitute only a small fraction of all red blood cells, invasion via this RBP-TfR1 pathway may result in only a small number of infected erythrocytes and explain the considerably low parasitemia observed in Duffy-negative *P. vivax* infections (**Figure 2**). Further, recent transcriptomic study has also indicated that genes belonging to tryptophan-rich antigen and merozoite surface protein families were highly expressed in the *Saimiri*-infected *P. vivax* [59], of which erythrocytes did not bind to DBP1 from the Belem isolate of *P. vivax* [60]. There is growing evidence that members of the tryptophan-rich antigen gene family are involved in erythrocyte invasion [61,62]. Other proteins such as Reticulocyte Binding Surface Antigen (*RBSA*) [63], anchored micronemal antigen (*GAMA*) [64], and Rhoptry neck protein (*RON*) [65] have also been suggested to play a role in red cell invasion, especially in low-density *P. vivax* infections [8]. Future studies are merited to clarify the expression and role of these proteins in Duffy-negative erythrocyte invasion and their respective host receptors.

Plasmodium vivax isolates could be different in their expression of various invasion ligands, including some that mediate the recognition and invasion to reticulocytes, providing a potential mechanism for variations in reticulocyte preference [66,67]. Moreover, successful schizont development was shown to be associated with increased younger reticulocytes in the Indian *P. vivax* isolates [66]. The low prevalence of schizonts

in peripheral blood has led to the hypothesis that *P. vivax* could be sequestering in reticulocyte-rich zones such as the bone marrow [68]. It is possible that with older reticulocytes in the peripheral blood, there may be a greater conversion rate to sexual stages, limiting the number of schizonts and thus a lower detectable parasitemia. Further investigations on the reticulocyte and parasite burden in bone marrow of Duffy-negative *P. vivax*-infected individuals, as well as the asexual and sexual development of these parasites would be highly informative.

Genetic relationships and origin hypotheses of *P. vivax* in Duffy-negative Africans

Maximum likelihood analyses of the African *P. vivax* isolates based on the Duffy binding protein (DBP) gene indicated that *P. vivax* isolates from Duffy-negative individuals were not monophyletic but found in multiple well-supported clades (clades II-VI in **Figure 4A**). These clades did not show clear geographical boundary. For instance, the two Duffy-negative *P. vivax* from Ethiopia (clade VI; bootstrap 91%) were closely related to Duffy-positive *P. vivax* from the same area, as well as to Duffy-positive and Duffy-negative *P. vivax* from neighboring Sudan. The present data may imply that Duffy-negative and Duffy-positive individuals shared similar *P. vivax* strains possibly by the same ancestral origin or through local transmission. Similar patterns were observed when other geographical isolates were included (**Figure 4B)**. The Duffy-negative *P. vivax* were clustered together with the Duffy-positive ones without genetic distinction. Such a clustering pattern may also suggest that evolution of *PvDBP* region II is more driven by functional selection rather than by geographical isolation. Duffy-negative *P. vivax* samples generally had a lower *PvDBP* gene diversity, likely due to smaller host reservoirs and/or limited samples.

Previous studies indicated that *P*. *vivax* in Southeast Asia and South America evolved in a clade of parasites that infect African monkeys [69,70]. *Plasmodium vivax* in African apes might present a substantial parasite reservoir from which Duffy-positive and Duffy-negative human infections arose from. However, the spread of Duffy-negative mutation across Africa could have eliminated the initial substantial *P. vivax* diversity in humans, and resulted in remarkable differences in genetic diversity between the populations of *P. vivax* that infect apes and humans [69,70]. There are two hypotheses concerning the origin of *P. vivax* in Duffy-negative Africans. The first hypothesis posits that the ancestral *P. vivax* infected all African primates including apes and Duffy-positive humans [71] (**Figure 5A**). One of these ancestral lineages evolved to a Duffy-independent pathway and subsequently spread to different parts of Africa via human migration [43] (**Figure 2**). The geographical overlap between apes and humans, e. g, in Cameroon and the Democratic Republic of Congo suggest a West-Central African origin of *P. vivax* in Duffy-negatives [72,73]. The second hypothesis posits that the ancestral *P. vivax* infected only non-human primates in Africa until some of the lineages crossed the species barrier and gave rise to the parasite population currently infecting Duffy-positive humans [69]. It is possible that Duffy-negative *P. vivax* observed today in West-Central, Southern, and East Africa represent separate lineages that were derived multiple times independently from Duffy-positive individuals (**Figure 5B**). Previous phylogenies based on a number of nuclear genes and partial mitochondrial genomes revealed incongruent genetic relationships [64,65]. It is particularly problematic for recently diverged lineages due to incomplete lineage sorting or lack of phylogenetic signal [74-76]. Moreover, no African *P. vivax* isolates from Duffy-positive and Duffy-negative individuals were included. This

study sheds light on the genetic relationships of the Duffy-negative and Duffy-positive African *P. vivax* together with the other geographical isolates based on *PvDBP.* However, this gene region may bias towards selection or with limited resolution. Extensive phylogenetic analyses using whole genome sequences of Duffy-negative *P. vivax* from West-Central, Southern, and East Africa, together with the existing data of the *P. vivax*-like isolates in African apes are needed to adjudicate the origin hypotheses of Duffy-negative *P. vivax*.

Uncertainty in the transmissibility of Duffy-negative *P. vivax*

In the parasite life cycle, a portion of the asexual stage schizonts converts into sexual gametocytes and contributes to transmission. In *P. vivax*, gametocytes appear simultaneously with the asexual schizonts [77] while in *P. falciparum*, gametocytes usually appear 10-14 days after infection. Such a difference in development not only increases the length of time the individual remains infectious, but also increases the likelihood of transmission before the infected individual seeks treatment [78] . The documentation of *P*. *vivax* infections in regions where the predominant population is Duffy-negative [8-13] suggested that there may not be a strong selection against Duffy-negative erythrocytes by the parasites. *P. vivax* in Duffy-negative hosts may be just as fit and share similar life cycle as in Duffy-positive hosts. The proportion of infections carrying gametocytes and gametocyte density in the general population are proxies for human-to-mosquito transmission potential [79-81]. In populations which is predominantly Duffy-negatives, infections could be transmitted back and forth among these individuals without a strong selective differential. However, in populations where Duffy-positive and Duffy-

negative individuals live side by side [30,31,82], it is unclear if there is a host preference and how often infections transmit between Duffy-positive and Duffy-negative individuals.

In Southeast Asia and the Peruvian Amazon, where *P. vivax* transmission is intense and stable, frequent recombination and human movements have led to high genetic diversity and lower differentiation among *P. vivax* populations [83-87]. By contrast, in Ethiopia, clear genetic structure was observed between the northern and southern *P. vivax* populations, likely due to the apparent differences in seasonality and landscape [88]. Future study should provide a rigid comparison and analysis of asymptomatic reservoirs and transmission mechanisms in different Duffy-negative populations across Africa. In addition, it is well known that *P. vivax* ability to relapse from dormant liver-stage hypnozoites, from weeks to years after clearance of the primary blood-stage infection is major obstacle to its control and elimination [89,90]. However, there is yet no information on the frequency and clinical impacts of relapse in Duffy-negative *P. vivax* infections. The increased risk of *P. vivax* infection and the growing clinical burden across regions as well as among Duffy phenotypes certainly highlight the public health concern of vivax malaria in Africa.

Conclusions

With the growing number of *P. vivax* cases reported in Duffy-negative individuals as well as across the continent, vivax malaria is no longer a rare but widespread phenomenon in Africa. This paper is the first to characterize and compare the epidemiological and genetic features of Duffy-negative *P. vivax* from different parts of Africa. It is unclear that *P. vivax* can exhibit both clinical symptoms and present as asymptomatic in the Duffy-negative

populations. The generally low parasitemia observed in the Duffy-negative infections may suggest a less efficient but continuously evolving invasion mechanism that allows a greater negative public health impact in coming years. While there is yet no direct evidence on transmission capability of Duffy-negative *P. vivax*, the genetic relatedness based on *PvDBP* sequences suggested similar strains shared between Duffy-negative and Duffy-positive populations. Further in-depth investigations are needed to unveil the invasion and transmission mechanisms of these infections. These data would help predict the scale of disease spread and improve existing malaria control measures, beyond *P. falciparum* in Africa. On the public health front-end, there should be more resources and training allocated to diagnosis and treatment of vivax malaria, given its unique ability in causing relapse and other longer-term health impacts. Duffy-negative Africans are not resistant to *P. vivax* infection and the public health significance of vivax malaria in Africa should no longer be neglected.

Acknowledgements

We thank the students and staffs from Jimma University, University of Khartoum, Evangelical University of Cameroon, University of Botswana, and University of North Carolina at Charlotte for their technical assistance in sample collection and processing; the communities and hospitals for their support and willingness to participate in this research. This research was funded by National Institutes of Health (R15 AI138002).

References

1. Price RN, Tjitra E, Guerra CA, Yeung S, White NJ, Anstey NM. Vivax malaria: neglected and not benign. *Am J Trop Med Hyg* 2007;77:79-87.

2. Howes RRE, Patil APAAP, Piel FBF, Nyangiri OAOOA, Kabaria CWCCW, Gething PW, et al. The global distribution of the Duffy blood group. Nat Commun. 2011.

3. Miller LH, Mason SJ, Clyde DF, McGinniss MH. The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. *N Engl J Med* 1976;295:302-304.

4. Tournamille C, Colin Y, Cartron JP, Le Van Kim C. Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nat Genet* 1995;10:224-228.

5. Hoher G, Fiegenbaum M, Almeida S. Molecular basis of the Duffy blood group system. Blood Transfus 2018; 16: 93-100.

6. Howes RE, Reiner Jr. RC, Battle KE, Longbottom J, Mapping B, Ordanovich D, Tatem AJ, Draeley C, Gething PW, Zimmerman PA, Smith DL, Hay SI. *Plasmodium vivax* transmission in Africa. *PLoS Negl Trop Dis* 2015;9:e0004222.

7. Zimmerman PA. *Plasmodium vivax* Infection in Duffy-Negative People in Africa. *Am J Trop Med Hyg* 2017;97:636-638.

8. Gunalan K, Niangaly A, Thera MA, Doumbo OK, Miller LH. *Plasmodium vivax* Infections of Duffy-negative erythrocytes: Historically undetected or a recent adaptation? *Trends Parasitol* 2018;34:420-429.

9. Mendes C, Dias F, Figueiredo J, Mora VG, Cano J, de Sousa B, do Rosario VE, Benito A, Berzosa P, Arez AP. Duffy negative antigen is no longer a barrier to *Plasmodium vivax*--molecular evidences from the African West Coast (Angola and Equatorial Guinea). *PLoS Negl Trop Dis* 2011;5:e1192.

10. Motshoge T, Ababio GK, Aleksenko L, Read J, Peloewetse E, Loeto M, Mosweunyane T, Moakofhi K, Ntebele DS, Chihanga S, Motlaleng M, Chinorumba A, Vurayai M, Pernica JM, Paganotti GM, Quaye IK. Molecular evidence of high rates of asymptomatic *P. vivax* infection and very low *P. falciparum* malaria in Botswana. *BMC Infect Dis* 2016;16:520.

11. Russo G, Faggiono G, Paganoti GM, Djeunang Dongho GB, Pomponi A, De Santis R, Tebano G, Mibida M, Sanou Sobze M, Vullo V, Rezza G, Lista FR. Molecular evidence of *Plasmodium vivax* infection in Duffy negative symptomatic individuals from Dschang, West Cameroon. *Malar J* 2017;16:74.

12. Niangaly A, Gunalan K, Ouattara A, Coulibaly D, MS´a J, Adams M, Travassos MA, Ferrero J, Laurens MB, Kone AK, Thera MA, Plowe CV, Miller LH, Doumbo OK. *Plasmodium vivax* infections over 3 years in Duffy blood group negative Malians in Bandiagara, Mali. *Am J Trop Med Hyg* 2017;97:744-752.

13. Brazeau NF, Whitesell AN, Doctor SM, Keeler C, Mwandagalirwa MK, Tshefu AK, Likwela JL, Juliano JJ, Meshnick SR. *Plasmodium vivax* Infections in Duffy-negative individuals in the Democratic Republic of the Congo. *Am J Trop Med Hyg* 2018;99:1128-1133.

14. Twohig KA, Pfeffer DA, Baird JK, Price RN, Zimmerman PA, Hay SI, Gething PW, Battle KE, Howes RE. Growing evidence of *Plasmodium vivax* across malaria endemic Africa. *PLoS Negl Trop Dis* 2019;13: e0007140.

15. Guerra CA, Howes RE, Patil AP, Gething PW, Van Boeckel TP, Temperley WH, et al. (2010) The International Limits and Population at Risk of *Plasmodium vivax* Transmission in 2009. PLoS Negl Trop Dis 4(8): e774.

16. Livingstone FB. The Duffy blood groups, vivax malaria, and malaria selection in human populations: a review. Hum Biol. 1984;56:413–25.

17. Battle KE, Lucas TCD, Nguyen M, Howes RE, Nandi AK, Twohig KA, et al. Mapping the global endemicity and clinical burden of Plasmodium vivax, 2000–17: a spatial and temporal modelling study. Lancet [Internet]. Lancet Publishing Group; 2019.

18. Anstey NM, Russell B, Yeo TW, Price RN. The pathophysiology of vivax malaria. Trends Parasitol. Elsevier; 2009. p. 220–7.

19. Zeeshan M, Tyagi RK, Tuagi K, Alam MS, Sharma YD. Host-parasite interaction: selective Pv-fam-a family proteins of *Plasmodium vivax* bind to a restricted number of human erythrocyte receptors. *J Infect Dis* 2015;211:1111-20.

20. Cowman A, Tonkin CJ, Tham WH, Duraisingh MT. The molecular basis of erythrocyte invasion by malaria parasites. *Cell Host Microbe* 2017;22:232-245.

21. Kumar H, Tolia NH. Getting in: The structural biology of malaria invasion. *PLoS Pathog* 2019; 15: e1007943.

22. Dolan SA, MIller LH, Wellems TE. Evidence for a switching mechanism in the invasion of erythrocytes by *Plasmodium falciparum* parasites. *J Clin Invest* 1990;86:618-624.

23. Stubbs J, Simpson KM, Triglia T, et al. Molecular mechanism for switching of *P. falciparum* invasion pathways into human erythrocytes. *Science* 2005;309:1384-1387.

24. Persson KE, McCallum F, Reiling L, et al. Variation in use of erythrocyte invasion pathways by Plasmodium *falciparum* mediates evasion of human inhibitory antibodies. *J Clin Invest* 2008;118:342-51.

25. Bowyer PW, Stewart LB, Aspeling-Jones H, et al. Variation in *Plasmodium falciparum* erythrocyte invasion phenotypes and merozoite ligand gene expression across different populations in areas of malaria endemicity. *Infect Immun* 2015;83:2575-82.

26. Gunalana K, Lo E, Hostetler JB, Yewhalaw D, Mu J, Neafsey DE, Yan G, Miller LH. Role of *Plasmodium vivax* Duffy-binding protein 1 in invasion of Duffy-null Africans. *Proc Natl Acad Sci USA* 2016;113:6271.

27. Hester J, Chan ER, Menard D, Mercereau-Puijaion O, Barnwell J, Zimmerman PA, Serre D. De novo assembly of a field isolate genome reveals novel *Plasmodium vivax* erythrocyte invasion genes. *PLoS Negl Trop Dis* 2013;7:e2569.

28. Bereczky S, Martensson A, Gil JP, Farnert A. Short report: Rapid DNA extraction from archive blood spots on filter paper for genotyping of Plasmodium falciparum. Am J Trop Med Hyg. 2005;72:249–51.

29. Perandin F, Manca N, Calderaro A, Piccolo G, Galati L, Medici MC, Arcangeletti MC, Snounou G, Dettori G, Chezzi C. Development of a real-time PCR assay for detection of *Plasmodium falciparum*, *Plasmodium vivax*, and *Plasmodium ovale* for routine clinical diagnosis. *J Clin Microbiol* 2004;42:1214-1219.

30. Lo EYY, Delenasaw Y, Zhong D, Zemene E, Degefa T, Tushune K, Ha M, Lee MC, James AA, Yan G. Molecular epidemiology of *Plasmodium vivax* and *Plasmodium falciparum* malaria among Duffy-positive and Duffy-negative populations in Ethiopia. *Malar J* 2015;14:84-94.

31. Menard D, Barnadas C, Bouchier C, Henry-Halldin C, Gray LR, Ratsimbasoa A, Thonier V, Carod JF, Domarie O, Colin Y, Bertrand O, Picot J, King CL, Grimberg BT, Mercereau-Puijalon O, Zimmerman PA. *Plasmodium vivax* clinical malaria is commonly observed in Duffy-negative Malagasy people. *Proc Natl Acad Sci USA* 2010;107:5967-5971.

32. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 2004;5:113.

33. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp* 1999; 41:95-98.

34. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312-1313.

35. Ngassa Mbenda HG, Awasthi G, Singh PK, Gouado I, Das A. Does malaria epidemiology project Cameroon as 'Africa in miniature'? *J Biosci* 2014;39:727-38.

36. Fru-Cho J, Bumah VV , Safeukui I, Nkuo-Akenji T, Titanji VPK, Haldar K. Molecular typing reveals substantial *Plasmodium vivax* infection in asymptomatic adults in a rural area of Cameroon. *Malar J* 2014;13:170.

37. Ngassa Mbenda HG, Das A. Molecular Evidence of *Plasmodium vivax* mono and mixed malaria parasite infections in Duffy-Negative native Cameroonians. *PLoS ONE* 2014;9:e103262.

38. Woldearegai TG, Kremsner PG, Kun JF, Mordmuller B. *Plasmodium vivax* malaria in Duffy-negative individuals from Ethiopia. *Trans R Soc Trop Med Hyg* 2013;107:328-31.

39. Zhou G, Yewhalaw D, Lo EYY, Zhong D, Wang X, Degefa T, Zemene E, Lee MC, Kenede E, Tushune K, Yan G. Analysis of asymptomatic and clinical malaria in urban and suburban settings of southwestern Ethiopia in the context of sustaining malaria control and approaching elimination. *Malar J* 2016;15:250.

40. Shriner D, Tekola-Ayele F, Adeyemo A, Rotimi CN. Genetic ancestry of Hadza and Sandawe peoples reveals ancient population structure in Africa. *Genome Biol Evol* 2018;10:875-882.

41. Pakendorf B, de Filippo C, Bostoen K. Molecular perspectives on the Bantu expansion: A synthesis. *Language dynamics and change*. 2011;1:50-88.

42. Grollemunda R, Branforda S, Bostoenb K, Meadea A, Vendittia C, Page M. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proc Natl Acad Sci USA* 2015;112:13296-301.

43. Busby GBJ, Band G, Le QS, Jallow M, Bougama E, Mangano VD, Amenga-Etego LN, Enimil A, Apinjoh T, Ndila CM, Manjurano A, Nyirongo V, Doumba O, Rockett KA, Kwiatkowski DP, Spencer CCA, Malaria Genomic Epidemiology Network. Admixture into and within sub-Saharan Africa. *eLife* 2016;5:e15266.

44. Petersen DC, Libiger O, Tindall EA, Hardie R-A, Hannick LI, Glashoff RH, et al. (2013) Complex Patterns of Genomic Admixture within Southern Africa. PLoS Genet 9(3): e1003309.

45. Schuster SC, Miller W, Ratan A, Tomsho LP, et al. Complete Khoisan and Bantu genomes from southern Africa. Nature 2010;**463**,: 943–947.

46. Tau T, Wally A Fanie TP, Ngono CL, Mpoloka SW, Davison S, D'Amato ME. 2017. Genetic variation and population structure of Botswana population. Scientific Reports 7:6768.

47. Bajić V, Barbieri C, Hübner A, Güldemann T, Naumann C, Gerlach L, Berthold F, Nakagawa H, Mpoloka SW, Roewer L, Purps J, Stoneking M, Pakendorf B. Genetic structure and sex-biased gene flow in the history of southern African populations. American Journal of Physical Anthropology 2018;167:656-671.

48. Tawe L, Motshoge T, Ramatlho P, Mutukwa N, Muthoga CW, Dongho GBD, Martinelli A, Peloewetse E, Russo G, Quaye IK, Paganotti GM. Human cytochrome P450 2B6 genetic variability in Botswana: a case of haplotype diversity and convergent phenotypes. Scientific Reports 2018;8:4912.

49. Hamblin MT, Di Rienzo A. 2000. Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus. American Journal of Human Genetics 66:1669–1679.

50. Hamblin MT, Thompson EE, Di Rienzo A. 2002. Complex signatures of natural selection at the Duffy blood group locus. American Journal of Human Genetics 70:369–383.

51. Pagani L, Kivisild T, Tarekegn A, Ekong R, Plaster C, Gallego Romero I, Ayub Q, Mehdi SQ, Thomas MG, Luiselli D, Bekele E, Bradman N, Balding DJ, Tyler-Smith C. Ethiopian genetic diversity reveals linguistic stratification and complex influences on the Ethiopian gene pool. *Am J Hum Genet* 2012;91:83-96.

52. Pickrell JK, Patterson N, Loh PR, Lipson M, Berger B, Stoneking M, Pakendorf B, Reich D. Ancient west Eurasian ancestry in southern and eastern Africa. *Proc Natl Acad Sci USA* 2014;111:2632-7.

53. Hollfelder N, Schlebusch CM, Günther T, Babiker H, Hassan HY, Jakobsson M (2017) Northeast African genomic variation shaped by the continuity of indigenous groups and Eurasian migrations. PLoS Genet 13(8): e1006976.

54. Anstey NM, Russell B, Yeo TW, Price RN. The pathophysiology of vivax malaria. Trends Parasitol. Elsevier; 2009. p. 220–7.

55. Naing C, Whittaker MA, Nyunt Wai V, Mak JW (2014) Is Plasmodium vivax Malaria a Severe Malaria?: A Systematic Review and Meta-Analysis. PLoS Negl Trop Dis 8(8): e3071.

56. Villamil-Gómez WE, Eyes-Escalante M, Franco-Paredes C. Severe and Complicated Malaria due to Plasmodium vivax Provisional chapter Severe and Complicated Malaria due to Plasmodium vivax. 2016; http://dx.doi.org/10.5772/64974

57. Ovchynnikova E, Aglialoro F, Bentlage AEH, et al. DARC extracellular domain remodeling in maturating reticulocytes explains *Plasmodium vivax* tropism. *Blood*. 2017;130(12):1441-1444.

58. Chan LJ, Dietrich MH, Nguitragool W, Tham WH. Plasmodium vivax Reticulocyte Binding Proteins for invasion into reticulocytes. Cell Microbiol 2020;22:e13110.

59. Gunalan K, Sá JM, Barros RRM, Anzick SL, Caleon RL, Mershon JP, et al. Transcriptome profiling of *Plasmodium vivax* in Samira monkeys identifies potential ligands for invasion. Proc. Natl. Acad. Sci. U.S.A. 2019;116: 7053–61.

60. Wertheimer SP, Barnwell JW (1989) Plasmodium vivax interaction with the human Duffy blood group glycoprotein: Identification of a parasite receptor-like protein. Exp Parasitol **69**:340–350.

61. Wang B, Lu F, Cheng Y, Chen J-H, Jeon H-Y, Ha K-S, et al. Immunoprofiling of the tryptophan-rich antigen family in *Plasmodium vivax*. Infect. Immun. 2015;83: 3083–95.

62. Tyagi K, Hossain ME, Thakur V, Aggarwal P, Malhotra P, Mohmmed A, et al. *Plasmodium vivax* tryptophan rich antigen PvTRAg36.6 interacts with PvETRAMP and PvTRAg56.6 interacts with PvMSP7 during erythrocytic stages of the parasite. Plos One. 2016;11: e0151065.

63. Moreno-Pérez DA, Baquero LA, Chitiva-Ardila DM, Patarroyo MA. Characterising PvRBSA: an exclusive protein from *Plasmodium* species infecting reticulocytes. Parasites Vectors. 2017;10: 243.

64. Baquero LA, Moreno-Pérez DA, Garzón-Ospina D, Forero-Rodríguez J, Ortiz-Suárez HD, Patarroyo MA. PvGAMA reticulocyte binding activity: predicting conserved functional regions by natural selection analysis. Parasites Vectors. 2017;10: 251.

65. Arévalo-Pinzón G, Bermúdez M, Curtidor H, Patarroyo MA. The *Plasmodium vivax* rhoptry neck protein 5 is expressed in the apical pole of *Plasmodium vivax* VCG-1 strain schizonts and binds to human reticulocytes. Malar. J. 2015;14: 106.

66. Lim, C., Pereira, L., Saliba, K. S., Mascarenhas, A., Maki, J. N., Chery, L., Gomes, E., Rathod P.K., Duraisingh, M. T. (2016). Reticulocyte preference and stage development of *Plasmodium vivax* isolates. The Journal of Infectious Diseases, 214, 1081–1084.

67. Bozdech Z, Mok S, Hu G, et al. The transcriptome of Plasmodium vivax reveals divergence and diversity of transcriptional regulation in malaria parasites. *Proc Natl Acad Sci U S A*. 2008;105(42):16290-16295.

68. Mayor A, Alano P. Bone marrow reticulocytes: a Plasmodium vivax affair? *Blood*. 2015;125(8):1203-1205.

69. Loy DE, Plenderleith LJ, Sundararaman SA, Liu W, Gruszczyk J, Chen YJ, Trimboli S, Learn GH, MacLean OA, Morgan ALK, Li Y, Avitto AN, Giles J, Calvignac-Spencer S, Sachse A, Leendertz FH, Speede S, Ayouba A, Peeters M, Rayner JC, Tham WH, Sharp PM, Hahn BH. Evolutionary history of human *Plasmodium vivax* revealed by

genome-wide analyses of related ape parasites. *Proc Natl Acad Sci USA* 2018;115:E8450-8459.

70. Gilabert A, Otto TD, Rutledge GG, Franzon B, Ollomo B, Arnathau C, Durand P, Moukodoum ND, Okouga AP, Ngoubangoye B, Makanga B, Boundenga L, Paupy C, Renaud F, Prugnolle F, Rougeron V. *Plasmodium vivax*-like genome sequences shed new insights into *Plasmodium vivax* biology and evolution. *PLoS Biol* 2018;16: e2006035.

71. Liu W, Li Y, Shaw KS, Learn GH, Plenderleith LJ, Malenke JA, Sundararaman SA, Ramirez MA, Crystal PA, Smith AG, Bibollet-Ruche F, Ayouba A, Locatelli S, Esteban A, Mouacha F, Guichet E, Butel C, Ahuka-Mundeke S, Inogwabini B, Ndjango JB, Speede S, Sanz CM, Morgan DB, Gonder MK, Kranzusch PJ, Walsh PD, Georgiev AV, Muller MN, Piel AK, Stewart FA, Wilson ML, Pusey AE, Cui L, Wang Z, Färnert A, Sutherland CJ, Nolder D, Hart JA, Hart TB, Bertolani P, Gillis A, LeBreton M, Tafon B, Kiyang J, Djoko CF, Schneider BS, Wolfe ND, Mpoudi-Ngole E, Delaporte E, Carter R, Culleton RL, Shaw GM, Rayner JC, Peeters M, Hahn BH, Sharp PM. African origin of the malaria parasite *Plasmodium vivax*. *Nat Commun* 2014;5:3346.

72. Kaiser M, Löwa a, Ulrich M, Ellerbrok H, Goffe AS, Blasse A, Zommers Z, Couacy-Hymann E, Babweteera F, Zuberbühler K, Metzger S, Geidel S, Boesch C, Gillespie TR, Leendertz FH. Wild chimpanzees infected with 5 *Plasmodium* species. *Emerg Infect Dis* 2010;16:1956-1959.

73. Prugnolle F, Rougeron V, Becquarta P, Berryd A, Makangab B, Raholaa N, Arnathaua C, Ngoubangoye B, Menardd S, Willaumeg E, Ayalah FJ , Fontenille D, Ollomob B ,

Duranda P, Paupy C, Renaud F. Diversity, host switching and evolution of *Plasmodium vivax* infecting African great apes. *Proc Natl Acad Sci USA* 2013;110:8123-8128.

74. Maddison WP, Knowles LL. Inferring phylogeny despite incomplete lineage sorting. *Syst Biol* 2006;55:21-30.

75. Liu L, Yu L, Kubatko L, Pearl DK, Edwards SV. Coalescent methods for estimating phylogenetic trees. *Mol Phylogenet Evol* 2009;53:320-8.

76. Smith SA, Moore MJ, Brown JW, Yang Y. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evol Biol* 2015;15:150.

77. Baker DA. Malaria gametocytogenesis. *Mol Biochem Parasitol* 2010;172:57-65.

78. Liu Z, Miao J, Cui L. Gametocytogenesis in malaria parasite: commitment, development and regulation. *Future Microbiol* 2011;6:1351-69.

79. Dixon MWA, Thompson J, Gardiner DL, Trenholme KR. Sex in *Plasmodium*: a sign of commitment. *Trends Parasitol* 2008;24:168-75.

80. Reece SE, Ramiro RS, Nussey DH. Plastic parasites: sophisticated strategies for survival and reproduction? *Evol Appl* 2009;2:11-23.

81. Nacher M, Carrara VI, McGready R, Ashley E, Nguyen JV, Thwai KL, Looareesuwan S, Nosten F. Seasonal fluctuations in the carriage of *Plasmodium vivax* gametocytes in Thailand. *Ann Trop Med Parasitol* 2004;98:115-120.

82. Albsheer MMA, Pestana K, Ahmed S, Elfaki M, Gamil E, Ahmed SM, Ibrahim ME, Musa AM, Lo EYY, Hamid MMA. Distribution of Duffy phenotypes among *Plasmodium vivax* infections in Sudan. *Genes* 2019;10:437.

83. de Oliveira TC, Rodrigues PT, Menezes MJ, Gonc¸alves-Lopes RM, Bastos MS, Lima NF, Barbosa S, Gerber AL, Loss de Morais G, Berna L, Phelan J, Robello C, de Vasconcelos ATR, Alves JMP, Ferreira MU. Genome-wide diversity and differentiation in New World populations of the human malaria parasite *Plasmodium vivax*. *PLoS Negl Trop Dis* 2017;11:e0005824.

84. Cowell AN, Valdivia HO, Bishop DK, Winzeler EA. Exploration of *Plasmodium vivax* transmission dynamics and recurrent infections in the Peruvian Amazon using whole genome sequencing. *Genome Med* 2018;10:52.

85. Jennison C, Arnott A, Tessier N, Tavul L, Koepfli C, Felger I, Siba PM, Reeder JC, Bahlo M, Mueller I, Barry AE. *Plasmodium vivax* populations are more genetically diverse and less structured than sympatric *Plasmodium falciparum* populations. *PLoS Negl Trop Dis* 2015;9:e0003634.

86. Lo EYY, Lam N, Hemming-Schroeder E, Nguyen J, Zhou G, Lee MC, Yang Z, Cui L, Yan G. Frequent spread of *Plasmodium vivax* malaria maintains high genetic diversity at the Myanmar-China Border, without distance and landscape barriers. *J Infect Dis* 2017;216:1254-1263.

87. Koepfli C, Rodrigues PT, Antao T, Orjuela-Sánchez P, Van den Eede P, Gamboa D, Gamboa D, van Hong N, Bendezu J, Erhart A, Barnadas C, Ratsimbasoa A, Menard D, Severini C, Menegon M, Nour BY, Karunaweera N, Mueller I, Ferreira MU, Felgner

I. *Plasmodium vivax* diversity and population structure across four continents. *PLoS Negl Trop Dis* 2015;9:e0003872.

88. Lo EYY, Hemming-Schroeder E, Yewhalaw D, Ngyuyen J, Kebede E, Zemene E, Getachew S, Tushune K, Zhong D, Zhou G, Petros B, Yan G. Transmission dynamics of co-endemic *Plasmodium vivax* and *P. falciparum* in Ethiopia and prevalence of antimalarial resistant genotypes. *PLoS Negl Trop Dis* 2017;11:e5806.

89. White MT, Shirreff G, Karl S, Ghani AC, Mueller I. Variation in relapse frequency and the transmission potential of *Plasmodium vivax* malaria. *Proc Biol Sci* 2016;283:20160048.

90. Battle KE, Karhunen MS, Bhatt S, Gething PW, Howes RE, Golding N, Van Boeckel TP, Messina JP, Shanks D, Smith DL, Baird JK, Hay SI. Geographical variation in *Plasmodium vivax* relapse. *Malar J* 2014;13:144.

Tables and Figures

Table 1 Comparison of *P. vivax* infection rate in Duffy-negative populations across different study sites in Cameroon, Botswana, Ethiopia, and Sudan.

Figure 1. Map showing the distribution of study sites and the Duffy status of febrile patients included in the present study.

Figure 2. Bantu expansion and admixture according to two hypotheses [17].

Figure 3. Comparison of *P. vivax* parasitemia based on quantitative PCR assays between Duffy negative and Duffy positive symptomatic infections among different geographical regions in Africa. Variations in parasitemia among samples were presented as boxplots showing the median and interquartile range values.

Figure 4. Phylogeny based on *PvDBP* sequences showing multiple source/origin of Duffy negative *P. vivax* in Africa. The reference *P. vivax* strain PVP01 isolated from an Indonesian patient was used as an outgroup. No clear differentiation was observed between the Duffy negative and Duffy positive *P. vivax* but nested within one another, suggestive of similar DBP haplotypes.

Figure 5. Hypothetical models illustrating the genetic origin of *P. vivax* in Duffy-negative Africans in a phylogenetic context.

Supplementary Files

Supplementary Table 1. Genbank accession number and geographical location of *PvDBP* sequences included in the present study.

Supplementary Table 2. Duffy genotype based on Sanger sequencing and parasite density based on quantitative PCR assay of Duffy negative *P. vivax* samples.