Immune Characterization of *Plasmodium falciparum* Parasites with a Shared Genetic Signature in a Region of Decreasing Transmission


Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts, USA; Laboratory of Bacteriology and Virology, Le Dantec Hospital, Faculty of Medicine and Pharmacy, Cheikh Anta Diop University, Dakar, Senegal; Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA; Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA; Center for Communicable Disease Dynamics, Harvard School of Public Health, Boston, Massachusetts, USA; Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA; Simmons College, Boston, Massachusetts, USA; Laboratory of Parasitology and Mycology, Faculty of Medicine and Pharmacy, Cheikh Anta Diop University, Dakar, Senegal

As the intensity of malaria transmission has declined, *Plasmodium falciparum* parasite populations have displayed decreased clonal diversity resulting from the emergence of many parasites with common genetic signatures (CGS). We have monitored such CGS parasite clusters from 2006 to 2013 in Thiès, Senegal, using the molecular barcode. The first, and one of the largest observed clusters of CGS parasites, was present in 24% of clinical isolates in 2008, declined to 3.4% of clinical isolates in 2009, and then disappeared. To begin to explore the relationship between the immune responses of the population and the emergence and decline of specific parasite genotypes, we have determined whether antibodies to CGS parasites correlate with their prevalence. We measured (i) antibodies capable of inhibiting parasite growth in culture and (ii) antibodies recognizing the surfaces of infected erythrocytes (RBCs). IgG obtained from volunteers in 2009 showed increased reactivity to the surfaces of CGS-parasitized erythrocytes over IgG from 2008. Since *P. falciparum* EMP-1 (PFEMP-1) is a major variant surface antigen, we used var Ups quantitative reverse transcription-PCR (qRT-PCR) and sequencing with degenerate DBLα domain primers to characterize the var genes expressed by CGS parasites after short-term *in vitro* culture. CGS parasites show upregulation of UpsA var genes and 2-cysteine-containing PFEMP-1 genes and express the same dominant var transcript. Our work indicates that the CGS parasites in this cluster express similar var genes, more than would be expected by chance in the population, and that there is year-to-year variation in immune recognition of surface antigens on CGS parasite-infected erythrocytes. This study lays the groundwork for detailed investigations of the mechanisms driving the expansion or contraction of specific parasite clones in the population.

Malaria is a devastating tropical disease caused by the protozoan parasite *Plasmodium*. The most severe form of malaria is caused by *Plasmodium falciparum* and results in 200 million to 300 million people developing clinical disease annually and at least 655,000 deaths worldwide (1, 2). Malaria remains one of the leading causes of death in children under the age of 5 years in sub-Saharan Africa (3). The clinical manifestations of the disease are caused by the erythrocytic stage of the parasite’s life cycle, and the majority of the immune response that develops is directed against this stage (4–6), both against parasite proteins involved in invasion and against proteins exported to the surface of the infected erythrocyte.

With the implementation of successful malaria control measures, transmission has declined in some regions, including Senegal. With the drop in transmission intensity, the complexity of exposure, multiplicity of infection, and population-level parasite diversity have decreased; as a result, different individuals—from different households and over different temporal periods of the transmission season—are infected with the same parasite genotype (7). Such observations have been facilitated by recently developed molecular genetic tools enabling rapid tracking of individual parasite genotypes within the population and over time (7, 8). In this study, we focus on a cluster of genetically identical parasites, termed common-genetic-signature (CGS) parasites, which was the first and one of the largest to be observed; these parasites both emerged and declined rapidly, reaching 24% of clinical isolates in 2008, declining to 3.4% in 2009, and then disappearing from all subsequently monitored populations.

Here we begin to address whether the rapid expansion and contraction of this particular parasite genotype can be partially explained by the specific humoral immune responses of infected individuals to these parasites. To do this, we have utilized two different standardized *in vitro* assays that measure antibodies to erythrocytic parasites, viz., *in vitro* parasite growth inhibition as-
says (GIAs) and assays of antibodies to variant surface antigens (VSA) on infected blood red cells (9, 10).

Many of the targets of protective humoral immunity have been identified as merozoite antigens that are involved in invasion (11). Humoral immunity directed against merozoite antigens is thought to act primarily by inhibiting merozoite invasion directly (12, 13), although it has also been suggested to act at the level of enhanced opsonophagocytosis of merozoites by monocytes (14). Great efforts have been made to develop standardized assays for measuring invasion and/or growth-inhibitory antibodies (15–17), a critical tool in assessing vaccine-induced or naturally acquired antibody activity.

Other possible targets for humoral immune responses to parasite-ized erythrocytes are the proteins on infected erythrocyte surfaces. These antibodies have been involved in cytoadherence, rosetting, and opsonization of infected erythrocytes and have been associated with protection from severe malaria (10, 18). Variation of some of these surface antigens allows the parasite to evade the antigen-specific partial immunity that develops with each exposure (4, 19). Traditionally, studies to measure the antibody response to VSA—encoded by multigenic families such as the var, rif, and stevor families—have used agglutination assays and correlated results with protection from disease (20, 21); increasingly, however, these experiments are being performed by flow cytometry (22–25).

As individuals are infected with diverse strains of malaria parasites, acquired immunity develops, and antibodies are a critical component of the immune response (26). However, the parasite has devised strategies to evade these protective antibody responses through clonal antigenic variation of surface-exposed antigens (27, 28). The degree of immune protection against diverse parasite infections is related to the level of exposure to different variant antigens (4, 10, 18, 21). If exposure to the diverse repertoire of parasite antigens decreases, the immune responses generated to these antigens also become more restricted in their recognition.

Because of the variability in parasite antigens and the ability of host immune responses to evolve when confronted with novel parasite antigens, we have used GIAs and VSA assays to begin to test the hypothesis that the antibody responses of the population can modulate the expansion or contraction of specific parasite genotypes. These results lay the foundation for more-detailed investigations of host immune mechanisms and other factors influencing parasite kinetics in the face of declining transmission.

**MATERIALS AND METHODS**

*P. falciparum* field parasites. This study was approved by both the Institutional Review Board of the Harvard School of Public Health (CR-16330-01) and the Ethics Committee of the Ministry of Health in Senegal (0127MSAS/DRPS/CNRES). Passive case detection was performed for patients visiting the clinic in Thiès, Senegal (Service de Lutte Anti-Parasitaire), with malaria-like symptoms, and those who were confirmed to be positive for *P. falciparum* malaria by a rapid diagnostic test (when available) and a microscopy slide were offered the opportunity to enroll in the study. Both venous blood samples (5 ml) and finger prick blood spotted onto filter paper were collected from consenting patients. Upon transport to the laboratory in Dakar, Senegal, erythrocytes and plasma were separated by centrifugation. Plasma was retained and stored at −20°C prior to IgG purification. Infected erythrocytes were frozen in Glycerol for *in vitro* culture adaptation. In this study, we focus on one cluster of identical (CGS) parasites that was present in 22 of 91 monogenic clinical isolates in 2008 (24%) and declined to 3 of 85 monogenic clinical isolates (3.4%) in 2009. (Monogenic status was determined by *Msp1* and *Msp2* typing and by 24-SNP [single nucleotide polymorphism] molecular barcoding from filter paper blood taken directly from patients.) The 25 total parasites in this CGS cluster did not demonstrate any significant difference in patient clinical parameters (age, parasitemia, temperature, and hematocrit [anemia]) from noncluster (non-CGS) parasites. Three monogenic parasite strains were successfully adapted to short-term *in vitro* culture (no more than 1 month) prior to *in vitro* immune assays: Th109.09 (CGS-1), Th153.09 (CGS-2), and Th250.08 (non-CGS). Since these three strains were adapted to *in vitro* culture, periodic periodicity was performed to confirm that there was no cross-contamination with other genotypes and to confirm the specific molecular barcode of each strain.

**Parasite genotyping**, Genomic DNA was extracted from whole-blood spots on Whatman FTA filter paper (Whatman). DNA was extracted from filter paper punches by using the manufacturer’s protocol for the Promega Maxwell DNA IQ Casework Sample kit (Promega), and a molecular barcode genotype was generated for each sample, as described previously (8). Whole-genome 17,000-SNP array hybridization, in addition to sequencing of the highly polymorphic T cell epitope of circumsporozoite protein (CSP), established that parasites with identical barcodes are identical at the level of their DNA sequences (7).

**RNA and cDNA preparation**, RNA from synchronous mid-schizont-stage parasite cultures was prepared using initial lysis with TRI reagent BD (Molecular Research Center), followed by purification on PureLink RNA minicolumns (Invitrogen). RNA was treated with Turbo DNase (Invitrogen), and cDNA was synthesized using SuperScript III and random hexamers (Invitrogen).

**Total-IgG purification**, Plasma samples from patients infected with clinically mild malaria were isolated from 5 ml venous blood collected in EDTA Vacutainers. Plasma samples were separated and were stored at −20°C prior to use. Plasma samples were randomly selected from 2008 (n = 8) and 2009 (n = 9) based on available sample volume. IgG purification was performed from 1 ml of plasma from each sample, as described in detail previously (29). After purification, the protein concentration of each total-IgG sample was determined by a NanoDrop system and was adjusted to 20 mg/ml with sterile 1× unsupplemented RPMI medium. IgG samples were stored at −80°C until GIAs were performed.

**GIAs with total IgG**, GIAs were performed as described previously (17), with minor modifications. The final IgG concentration in all assays was 10 mg/ml. Due to the lower parasite multiplication rate observed for some culture-adapted field strains than for lab strains, synchronous late- trophozoite to mid-schizont-stage parasite cultures were diluted to 0.5 to 1.0× initial parasitemia rather than 0.3%. After reinvasion and growth to late-trophozoite to mid-schizont stage, assays were harvested, and results were read immediately.

**Variant surface antigen (VSA) binding of infected erythrocytes by flow cytometry**, Synchronous late-trophozoite to mid-schizont-stage parasite cultures were diluted to 1% hematocrit to prevent clumping and were washed twice with 1× PBS (phosphate-buffered saline) plus 2% FBS (fetal bovine serum). IgG surface staining was assayed as described previously (23) with minor modifications. Plasma samples from patients who were infected with clinically mild malaria were isolated from 5 ml venous blood collected in EDTA Vacutainers, separated, and stored at −20°C prior to use. The plasma samples were randomly selected from 2008 (n = 55) and 2009 (n = 55), based on available sample volume. Plasma samples were subjected to a specificity criterion of having an average IgG reactivity no greater than 5% (in general, the maximum observed background staining of unexposed plasma) to uninfected erythrocytes from two different donors. Any samples that failed to meet this criterion were excluded from analysis. Of the 110 plasma samples tested, only 1 failed to meet this criterion and was excluded, resulting in the inclusion of 109 plasma samples in the final analysis. In addition to measuring the percentage that were IgG positive, we scored plasma samples as responders or nonresponders based on a 2× cutoff (mean + 2 times the standard deviation for nine unexposed Boston plasma samples). This allowed us to compare the percentage of positive responders across all experiments.
**Var gene Ups class expression profiling.** Quantitative PCR was performed on cDNAs using iTaq SYBR Green Supermix with ROX (Bio-Rad) with an ABI Prism real-time PCR machine. The primers used in these experiments have been described previously (30), but the exact primer pairs used are listed in Table S1 in the supplemental material. cDNA (prepared with and without reverse transcriptase) was diluted 1/10 with diethyl pyrocarbonate (DEPC)-grade water, and 1 μL was used per reaction. The reaction mixture was as follows for each primer pair: 1 μL cDNA, 10 μL SYBR green, 4 μL primer mixture (equal parts of forward and reverse primers at 2.5 μM; final concentration, 0.5 μM), and 5 μL DEPC-grade water. Samples were run in triplicate except for the controls seryl-tRNA synthetase and fructose bisphosphate aldolase, which were run in sextuplicate. The reaction conditions were as follows: 95°C for 10 min (initial denaturation), followed by 40 cycles of 95°C for 15 s, 54°C for 40 s, and 60°C for 1 min, with no final extension. Data were analyzed as the change in the threshold cycle (ΔCT) relative to the C(T) for seryl-tRNA synthetase.

**Var gene DBL1αx sequencing.** DBL1αx regions were amplified from cDNAs by PCR as described previously (31) and from genomic DNA (gDNA) as a control. Briefly, three sets of primer pairs, described in Table S2 in the supplemental material, were used to amplify overlapping regions of DBL1αx: primer pair 1, nDBLF/BR (450- to 550-bp products); primer pair 2, α-ΔFBR (375- to 475-bp products); primer pair 3, nDBLF2/BR (250- to 350-bp products) (31, 33). All PCR products were amplified with the following reaction mixture: 2.5 μL 10X Platinum Taq high-fidelity buffer, 0.5 μL deoxynucleoside triphosphates (dNTPs) (10 μM stock; Denville Scientific), 1 μL 50 mM MgCl₂ (Invitrogen), 0.5 μL forward primer (10 μM), 0.5 μL reverse primer (10 μM), 0.2 μL Platinum Taq high-fidelity polymerase, 18.8 μL of DEPC water, to which 1 μL of cDNA or gDNA was added for each strand. Reactions with and without reverse transcriptase were performed for all PCR pairs. Primer pairs 1 and 3 were used with a program consisting of 95°C for 3 min, followed by 35 cycles of 94°C for 15 s, 54°C for 30 s, and 60°C for 45 s, with a final extension at 72°C for 7 min, as described previously (31). Primer pair 2 was used with a program consisting of 95°C for 5 min, 42°C for 1 min, and 60°C for 1 min, followed by 29 cycles of 94°C for 1 min, 42°C for 1 min, and 60°C for 1 min, with no final extension (32). PCR products were cloned into pCR 2.1 TOPO (Invitrogen), transformed into chemically competent Escherichia coli TOP10 or XL10-Gold cells, and plated on agar with 100 μg/mL carbenicillin. A total of 48 colonies per PCR were selected and were sequenced in both the forward and reverse directions using a Sanger dideoxy sequencing approach. Only sequences that yielded good-quality sequence in both directions were included in the analysis.

**Var sequence analysis.** Sequence reads were analyzed individually. Contigs were generated using SeqMan (Lasergene 10) and were translated using MacVector (version 12.7). Translated consensus sequences were scanned for the RSFADIG motif described previously (31). For sequences containing the RSFADIG motif, alignments of consensus sequences were performed using ClustalW. PREMP-1 motifs associated with disease states have been described previously (31), and motif logos were created using WebLogo (UC Berkeley).

**Var transcript dominance.** The dominance of the var transcript for each patient was scored according to the method described by Normark et al. by using the equation \( \frac{r_i}{n_i} \) (where \( r_i \) is the number of sequences for patient \( i \) in cluster \( j \) and \( n_i \) is the total number of reads from each patient \( i \)) for each primer pair (31). The \( \alpha-\Delta \)BR primer set (primer pair 2) was prioritized in the determination of the 3 dominant contigs for each patient isolate, as described previously (31). var sequences identified by BLAST as PFE1640w (varCOMMON or var1CSA) were excluded from dominance analysis.

**Var DBL1α network analysis.** We used a network approach to analyze the relationships between var sequences. We generated consensus sequences from the three primer sets (VARSeg1 to -3 [see Fig. 4A]), yielding 126, 131, and 124 PCR products for non-CGS, CGS-1, and CGS-2 strains, respectively. These were then trimmed to subdomains II and III and were examined for 100% identity at the amino acid level, collapsing, respectively, to 36, 44, and 43 unique sequences for both gDNA and cDNA; 26, 30, and 26 unique sequences for gDNA alone; and 16, 22, and 21 unique sequences for cDNA alone. We refer to the number of redundant sequences that collapse to each unique sequence as the number of corroborating sequences. The percentage of identity between sequences was measured pairwise by (i) aligning sequences using MUSCLE (33) and (ii) calculating identity as the ratio of the number of residues of agreement in the alignment to the total possible, i.e.,

\[
\% \text{identity} = \frac{\text{aligned length} (x, y)}{\text{Hamming distance} (x, y)} \times 100.
\]

The 80%-identity network was formed by linking two vertices (genes) if they were identical at a level of 80% or higher.

To investigate the amount of var gene repertoire and expressed var transcript overlap at 80% identity that one would expect by chance, we conducted five randomization tests using an augmented data set consisting of CGS-1, CGS-2, and non-CGS parasites, as well as seven parasites whose var repertoires were published from whole-genome sequencing (34): 3D7, HB3, IT4, D2D, PFCLIN, IGH, and RAJ116. In our randomized tests, two parasites are selected—which two depends on which of the five tests—and we select sequences from each parasite’s repertoire, without replacement. We create an 80%-identity network between the two sets of sequences and measure the Jaccard distance between them, where the size of the union is the total number of network components and the size of the intersection is the number of mixed components, connecting both parasites. The five randomized tests were as follows: (i) compare the repertoires of CGS-1 and CGS-2, (ii) compare the repertoires of 3D7 and 3D7, (iii) compare 2 of the 10 parasites’ repertoires, chosen at random, provided that the 2 parasites are not the same and are not CGS-1 and CGS-2, (iv) compare the repertoires of the CGS-1 and non-CGS parasites, and (v) compare the repertoires of CGS-2 and non-CGS parasites. A total of 10,000 replicates were carried out for each test, giving a high degree of discriminating power. These five tests serve to measure the distributions of var repertoire overlap (under PCR-like subsampling) between (i) CGS-1 and CGS-2 parasites, (ii) two parasites known to be identical, (iii) two parasites believed to be different, and (iv and v) CGS and non-CGS parasites with the same subsampling procedure as comparisons i, ii, and iii. Subsampling was done with 15 sequences for cDNA analysis alone because the smallest expressed var transcript repertoire in the data set comprised 16 sequences. Subsampling was done with 25 sequences for gDNA analysis alone because the smallest var gene repertoire in the data set comprised 26 sequences.

Separately, a recombination network was generated using the same augmented set of CGS-1, CGS-2, and non-CGS parasites and the seven whole-genome repertoires (“global” [34]) and applying the network construction technique of Larremore et al. (35). Briefly, highly variable regions of sequences are extracted and compared pairwise at the amino acid level, with two sequences linked when they share an exact match of 6 amino acids or longer within the highly variable region.

**RNA-seq analysis.** Two hundred nanograms of total RNA was sent to BGI (Tai Po, Hong Kong) for RNA-seq (transcriptome sequencing) data generation. cDNA transcripts were enriched using poly(A) selection, and raw read data were analyzed at the Broad Institute (Cambridge, MA). Reads were trimmed with Trimmomatic, ends of reads were removed if phred quality was less than 15 (36), reads were aligned to the 3D7 reference genome with TopHat (37), transcript expression was measured with CuffLinks (38), and samples were compared with CuffDiff2 (39). Statistical analysis was performed using FPKM (fragments per kilobase per million) metric values for each locus. Additionally, de novo transcriptome assembly (without a 3D7 consensus) was performed for each of the sequenced samples using the Trinity platform (40). Dominant var genes were identified in the transcriptome assemblies by using the RSEM utility (41) to map reads back onto the assembled contigs, and calculated FPKM values were compared with those identified by cloning and sequencing using BLAST (42).
RESULTS

Parasites with a common genetic signature are equally inhibited by IgG. To determine whether parasites with a common genetic signature were equally or differentially recognized by growth-inhibitory antibodies, we performed standardized GIA using IgG purified from randomly selected plasma samples collected from Senegalese malaria patients in 2008 (n = 8) and 2009 (n = 9) and CGS parasite strains adapted to short-term in vitro culture. We observed a strong correlation between GIA inhibition levels for the CGS-1 and CGS-2 parasite strains, indicating that these strains are antigenically similar, at least in terms of their growth-inhibitory repertoires (Fig. 1A). Further, when IgG was stratified by year, we observed no difference in the levels of inhibition by IgG from 2008 and IgG from 2009, whether cumulatively for both CGS parasites (Fig. 1B) or individually for the CGS-1 (Fig. 1C) or CGS-2 (Fig. 1D) parasite.

Parasites with a common genetic signature express similar var Ups types, marked by upregulation of var UpsA. Various multigene families, such as the var, rif, and stevor multigene families, encode antigens expressed on the infected-erythrocyte surface that are recognized by anti-VSA antibodies. However, because PfEMP-1 is a dominant component of the VSA response (43), we characterized the var genes expressed by CGS parasites by var Ups quantitative reverse transcription-PCR (qRT-PCR) (Fig. 3). This approach allows us to determine the classes of var genes, defined by their promoters, expressed in each parasite strain. We observed that the CGS parasites expressed similar var genes (data not shown), we observed a strong positive correlation between the VSA recognition by CGS-1 and that by CGS-2, implying that the CGS parasites are antigenically similar at the level of VSA recognition (Fig. 2A). When all 109 plasma samples analyzed were classified as positive or negative responders by using a cutoff, the non-CGS parasite showed equivalent reactivities to 2008 and 2009 plasma samples, whereas both CGS parasites exhibited dramatically higher surface recognition of 2009 plasma samples than of 2008 plasma samples by a Z-test for two population proportions (Fig. 2B).

Parasites with a common genetic signature express similar var Ups types, marked by upregulation of var UpsA. Various multigene families, such as the var, rif, and stevor multigene families, encode antigens expressed on the infected-erythrocyte surface that are recognized by anti-VSA antibodies. However, because PfEMP-1 is a dominant component of the VSA response (43), we characterized the var genes expressed by CGS parasites by var Ups quantitative reverse transcription-PCR (qRT-PCR) (Fig. 3). This approach allows us to determine the classes of var genes, defined by their promoters, expressed in each parasite strain. We observed that the CGS parasites expressed similar var Ups classes (Fig. 3A), marked by striking upregulation of UpsA var genes (between 5-fold and 4,000-fold), specifically var UpsA3 (upregulated 4,790-fold in CGS-1 and 200-fold in CGS-2) (Fig. 3B), relative to expression by the non-CGS parasite.
Parasites with a common genetic signature express different yet similar dominant var genes and are marked by a predominance of 2-cysteine-containing (2-Cys) var genes. In addition to Ups class, we wanted to determine the dominant var genes expressed in these parasites and to identify any sequence motifs or characteristics that might distinguish the groups functionally. We characterized the var gene DBL1α domains, which are involved in rosetting and endothelial binding, by sequencing using degenerate DBL1α domain primers, as described previously (31, 32) (Fig. 4A). The 3 dominant var transcripts expressed for each PCR segment are described in Table 1. The dominant var gene transcribed for both CGS parasites (CGS-1 and CGS-2) was PF3D7_1100200 (GenBank accession no. ABC95935.1), an UpsA var gene, whereas the dominant var gene expressed for the non-CGS parasite had GenBank accession no. ABC95935.1 and was most similar by homology to a var gene of an uncompiled isolate from Mali (Table 1). These data were further corroborated by RNA-seq transcriptome assembly results, which identified the same dominant var transcript for each strain: PF3D7_1100200 for CGS-1 and CGS-2 and GenBank accession no. ABC95935.1 for the non-CGS parasite (data not shown). Further, the transcriptional profiles of the CGS parasites were highly similar, with no significant differences observed (see Fig. S2A in the supplemental material), whereas comparisons between the CGS-1 or CGS-2 strain and the non-CGS strain (see Fig. S2B and C in the supplemental material, respectively) yielded 4 and 6 significant differences, respectively, in the transcribed var profiles for 77 var genes and pseudogenes.

We observed that the CGS parasites expressed predominantly 2-cysteine-containing PIEMP-1 molecules, in contrast to the non-CGS parasite, which expressed mostly 4-cysteine-containing PIEMP-1 molecules (Fig. 4B). In addition to the difference in the number of cysteine in PIEMP-1 molecules, we observed motifs previously associated with disease severity (31) in the CGS parasites, in contrast to largely mild motifs and low-rosetting motifs in the non-CGS parasite (Fig. 4C).

Senegalese isolates exhibit a level of diversity comparable to that of global var gene repertoires, yet var gene repertoires and expressed var transcripts of parasites with a common genetic signature are highly overlapping. To determine the relationships between var DBL1α sequences, we employed a network approach described previously (35). While standard phylogenetic analyses are appropriate in many applications, the extreme diversity of var sequences, generated by frequent recombination, is better analyzed using networks. First, by comparison to other global populations, we sought to determine whether the overall diversity of the Senegalese sequences obtained in this study was in some way unique. We compared CGS-1, CGS-2, and non-CGS parasite se-
quences to those from 3D7, HB3, IT4, DD2, PFC01, IGH, and RAI116 (34). The Senegalese vertices did not cluster together but instead were spread throughout the network, indicating within-isolate diversity on the same order as total global diversity (35) (Fig. 4D). Moreover, the Senegalese vertices were distributed into all major network communities, indicating full var gene repertoire diversity and structure similar to those of the global population.

To directly investigate similarities between expressed var cDNA transcripts, we generated an additional var sequence network at 80% identity (described in detail in Materials and Methods). This network shows that the expressed var transcripts of CGS-1 and CGS-2 overlap substantially, while the expressed var transcripts of the non-CGS parasite are almost unique (Fig. 4E). At 80% sequence identity, the first link from the non-CGS parasite
TABLE 1 DBL1α-sequenced dominant var sequencesa

<table>
<thead>
<tr>
<th>Segment</th>
<th>Strain</th>
<th>Proportion of reads amplified (%)</th>
<th>PSI BLAST GenBank accession no.</th>
<th>% identity</th>
<th>Annotation of similar sequence identified by BLAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>VARseg1</td>
<td>CGS-1</td>
<td>7/15 (46.6)</td>
<td>ACH53749.1</td>
<td>64</td>
<td>Var UpsC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4/15 (26.6)</td>
<td>ABD92262.1</td>
<td>97</td>
<td>Var UpsA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2/15 (13.3)</td>
<td>ADX61680.1</td>
<td>77</td>
<td>PIK11 (Pikine, Senegal)</td>
</tr>
<tr>
<td>CGS-2</td>
<td>5/12 (41.6)</td>
<td>ADX68755.1</td>
<td>65</td>
<td>Isolate IT4/25/5 (Pikine, Senegal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/15 (16.6)</td>
<td>ADX61988.1</td>
<td>73</td>
<td>PIK45 (Pikine, Senegal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2/12 (16.6)</td>
<td>ADX61557.1</td>
<td>99</td>
<td>PIK05 (Pikine, Senegal)</td>
</tr>
<tr>
<td>Non-CGS</td>
<td>7/14 (50)</td>
<td>ADX60873.1</td>
<td>73</td>
<td>KIL04 (Kilifi, Kenya)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/14 (14.3)</td>
<td>ACH53720.1</td>
<td>77</td>
<td>Var UpsC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/14 (14.3)</td>
<td>ADX61012.1</td>
<td>70</td>
<td>KIL31 (Kilifi, Kenya)</td>
<td></td>
</tr>
<tr>
<td>VARseg2</td>
<td>CGS-1</td>
<td>7/33 (21.2)</td>
<td>ABD92262.1</td>
<td>99</td>
<td>Var UpsA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5/33 (15.2)</td>
<td>CAJ39106.1</td>
<td>86</td>
<td>Patient 4140, sequence group 2, Kilifi, Kenya (2-Cys)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4/33 (12.1)</td>
<td>CAJ39782.1</td>
<td>81</td>
<td>Patient 4130, sequence group 3, Kilifi, Kenya (2-Cys)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4/33 (12.1)</td>
<td>ABD92262.1</td>
<td>100</td>
<td>Var UpsA</td>
</tr>
<tr>
<td>CGS-2</td>
<td>10/32 (31.3)</td>
<td>ABD92262.1</td>
<td>99</td>
<td>Var UpsA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>7/32 (21.9)</td>
<td>ACH53883.1</td>
<td>67</td>
<td>Isolate u24D1639.1 (Pikine, Senegal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4/32 (12.5)</td>
<td>ACH53674.1</td>
<td>77</td>
<td>Isolate u19B74125.3 (Pikine, Senegal)</td>
</tr>
<tr>
<td>Non-CGS</td>
<td>16/27 (59.3)</td>
<td>ABC95935.1</td>
<td>75</td>
<td>Mali, uncomplicated isolate</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2/27 (7.4)a</td>
<td>ABC95940.1</td>
<td>91</td>
<td>Mali, uncomplicated isolate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2/27 (7.4)a</td>
<td>CAJ40536.1</td>
<td>100</td>
<td>Patient 4187, sequence group 4, Kilifi, Kenya (4-Cys)</td>
</tr>
<tr>
<td>VARseg3</td>
<td>CGS-1</td>
<td>6/14 (42.9)</td>
<td>ADX61607.1</td>
<td>86</td>
<td>PIK07 (Pikine, Senegal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5/14 (35.7)</td>
<td>ADX61607.1</td>
<td>87</td>
<td>PIK07 (Pikine, Senegal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2/14 (14.3)</td>
<td>ADX61607.1</td>
<td>87</td>
<td>PIK07 (Pikine, Senegal)</td>
</tr>
<tr>
<td>CGS-2</td>
<td>3/11 (27.3)</td>
<td>ADX60838.1</td>
<td>71</td>
<td>KIL23 (Kilifi, Kenya)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3/11 (27.3)</td>
<td>ADX60703.1</td>
<td>69</td>
<td>KIL09 (Kilifi, Kenya)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1/11 (9.0)a</td>
<td>ADX61607.1</td>
<td>87</td>
<td>PIK07 (Pikine, Senegal)</td>
<td></td>
</tr>
<tr>
<td>Non-CGS</td>
<td>15/22 (68.2)</td>
<td>ADX61664.1</td>
<td>66</td>
<td>PIK10 (Pikine, Senegal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3/22 (13.6)</td>
<td>ADX61664.1</td>
<td>66</td>
<td>PIK10 (Pikine, Senegal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3/22 (13.6)</td>
<td>ACH53748.1</td>
<td>70</td>
<td>Isolate s29C32r25.2 (Pikine, Senegal)</td>
<td></td>
</tr>
</tbody>
</table>

a The three rows represent the three dominantly expressed var sequences for each parasite strain (CGS-1, CGS-2, and non-CGS) based on the results from 48 clones. Boldface sequences represent the dominant var expressed for each parasite isolate.

b The percentage of reads amplified falls below the 10% cutoff for a dominant transcript.

to either of the CGS parasites appears, while the CGS-1 and CGS-2 parasites are linked at 100% transcript sequence identity.

To investigate the amount of var gene or transcript repertoire overlap at 80% identity that one would expect by chance, we conducted five randomization tests using the CGS-1, CGS-2, and non-CGS parasites, as well as seven parasite strains whose var repertoires have been published previously (34). Null model calculations (described in detail in Materials and Methods) confirm that the non-CGS parasite represents a distinct var repertoire, while those of the CGS-1 and CGS-2 parasites, though statistically different by a Wilcoxon rank-sum test of 10,000 iterations (all P values are <0.0001), are not substantially different, showing high similarity both at the gDNA level (see Fig. S3A and B in the supplemental material) and at the cDNA level (see Fig. S3C and D in the supplemental material). We can conclude with high confidence that the CGS parasites are not unrelated—the “different” and “similar” parasite distributions are completely nonoverlapping, meaning there is a high degree of discriminating power. (Some uncertainty arises in the comparison at the gDNA level due to subsampling of the total ~60 var genes of each isolate, as 48 clones were sampled per isolate.) The CGS-1 and CGS-2 parasites were similar in both their gDNA var gene repertoires and their expressed var transcripts nearly at the level of replicates of an individual strain (see Fig. S3A and C).

DISCUSSION

In this population-based study, we characterize the immune response to genetically identical parasites isolated from different infected individuals from Thiès, Senegal. While these identical parasites have been observed within a transmission season, they have also been observed across multiple years. These clusters of identical parasites both grow and decline, and some remain stable in their frequency over the years. Such observations could have important implications for the diversity of the immune responses generated over time within the population, and therefore for patterns of herd immunity, but this will depend on the strength of association between antigenic determinants and the loci that are usually tracked by genotyping techniques, which tend to be neutral. In pathogens with high rates of recombination, this association may be weak or variable. Many questions remain as to the forces driving the frequencies of these genetically identical parasites, and we have begun to address the role of host immune responses in the expansion or contraction of specific parasite genotypes in the population.

Here we characterized the immune response to CGS parasites from the first large cluster to be observed that expanded in 2008 and contracted in 2009 before disappearing from subsequent populations. Since the dynamics of the emergence and decline of this
particular parasite population were so dramatic, it might be possible to isolate responses to this CGS group from general year-to-year variations. We utilized two different immunologic measurements that might reflect functional antibodies to asexual parasites, viz., an in vitro parasite growth inhibition assay and a VSA assay measuring antibodies to parasite antigens on the surfaces of infected erythrocytes. The former assay evaluates the ability of antibodies to inhibit the invasion of erythrocytes by merozoites, and the latter measures antibodies that could be involved in the clearance of infected red cells.

We found that the two representative CGS parasites from this large cluster were recognized to similar degrees by IgG in both GIAs and VSA assays, showing that they are antigenically similar. To determine whether associations exist between the level of immune reactivity and the frequency of the parasite genotype in the population, we compared the level of GIA activity and VSA staining over time, in 2008 and 2009. While there was no difference in the level of growth inhibition by GIA between 2008 and 2009, there was a dramatic difference in VSA surface recognition between these years. Since VSA recognition has been correlated with protective immunity (10, 18), the increase in population-level recognition in 2009 could possibly help explain why this parasite genotype declined from 24% of clinical isolates to 3.4% in this period before disappearing from the population. One possible mechanism is that a higher level of VSA recognition could lead to increased antibody-mediated effector function against parasitized cells, thus reducing the frequency of this genotype in the population. However, to prove causality would require looking at clusters that span multiple years and ideally looking longitudinally within individuals to measure the development of genotype-specific immunity over time and the potential of this immunity to neutralize reinfesting strains of the same genotype.

While the CGS parasites appeared to be antigenically similar in functional immune assays, we wanted to determine the level of similarity in their variant surface antigen repertoires. While not the only determinant, one of the major determinants of the infected erythrocyte surface immune response is PfEMP-1 (43), encoded by approximately 60 var genes. Depending on the domain architecture of a particular PfEMP-1 molecule, these antigens can be involved in pathogenic processes such as cytoadherence and rosetting (44–46). Immune responses directed against PfEMP-1 have been shown to protect against severe malaria (10, 18). Here we used var Ups qRT-PCR and sequencing with degenerate DBL1α domain primers to characterize the var genes expressed by CGS parasites. We observed that the CGS parasites expressed the same var Ups classes, marked by a striking upregulation of UpsA var genes, while this was not observed for the non-CGS control. This was an unusual result, because all the parasites in this study have been shown to protect against severe malaria (10, 18). Here we used var Ups qRT-PCR and sequencing with degenerate DBL1α domain primers to characterize the var genes expressed by CGS parasites. We observed that the CGS parasites expressed the same var Ups classes, marked by a striking upregulation of UpsA var genes, while this was not observed for the non-CGS control.

In order to definitively link the immune response to variant surface antigens and population-level immune selection, further studies examining more parasites from more clusters that span multiple years, both expanding and contracting clusters, are needed. Additionally, longitudinal measurement of the development of immunity to a specific genotype in order to determine if this immunity can protect against reinfection was done using the rodent malaria model. However, this low-transmission environment might be the ideal setting for examination of the impact of transmission of a single genotype to multiple individual human hosts and the role of the individual immune response in shaping the variant antigen repertoire.

ACKNOWLEDGMENTS

We acknowledge Daria Van Tyne and Kate Fernandez for assistance with adapting Senegalese isolates to culture, Samuel E. Moretz for assistance with IgG purification, and Kim Brolin for useful discussions regarding Ups qPCR. We further acknowledge all the study participants in Thiès, Senegal, for agreeing to participate in this work, as well as Aminata Mbaye, Baba Diaye, Yaye Die Ndiaye, Amanda Lukens, Papa Diogoye Sene, Younous Diedhiou, Lamine Ndiaye, Amadou Macar Mbaye, and Ngayo Sy for assistance with sample collection and processing.
REFERENCES


