

Article

Assessment of Mixed *Plasmodium falciparum* *sera5* Infection in Endemic Burkitt Lymphoma: A Case-Control Study in Malawi

Nobuko Arisue ¹, George Chagaluka ², Nirianne Marie Q. Palacpac ³, W. Thomas Johnston ⁴, Nora Mutalima ^{4,5}, Sally Peprah ⁶, Kishor Bhatia ⁶, Eric Borgstein ², George N. Liomba ², Steve Kamiza ², Nyengo Mkandawire ², Collins Mitambo ⁷, James J. Goedert ⁶, Elizabeth M. Molyneux ², Robert Newton ⁴, Toshihiro Horii ³ and Sam M. Mbulaiteye ^{6,*}

¹ Research Center for Infectious Disease Control, Research Institute for Microbial Diseases, Osaka University, Osaka 565-0871, Japan; arisue@biken.osaka-u.ac.jp

² Departments of Pediatrics and Surgery, College of Medicine, University of Malawi, Private Bag 360, Chichiri, Blantyre 3, Malawi; gchagaluka@medcol.mw (G.C.); eborgstein@medcol.mw (E.B.); profngliomba@yahoo.com (G.N.L.); skamiza@medcol.mw (S.K.); nmkandawire@medcol.mw (N.M.); emmolyneux@gmail.com (E.M.M.)

³ Department of Malaria Vaccine Development, Research Institute for Microbial Diseases, Osaka University, Osaka 565-0871, Japan; nirian@biken.osaka-u.ac.jp (N.M.Q.P.); horii@biken.osaka-u.ac.jp (T.H.)

⁴ Epidemiology and Cancer Statistics Group, Department of Health Sciences, University of York, York YO10 5DD, UK; tom.johnston@york.ac.uk (W.T.J.); nora.mutalima@monashhealth.org (N.M.); Robert.Newton@york.ac.uk (R.N.)

⁵ Cancer Epidemiology Unit, University of Oxford, Oxford OX3 7LF, UK

⁶ Infections and Immunoepidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA; sally.peprah@nih.gov (S.P.); dnadoc@hotmail.com (K.B.); jamesgoedert@gmail.com (J.J.G.)

⁷ National Health Sciences Research Committee, Research Department, Ministry of Health, P.O. Box 30377, Capital City, Lilongwe 3, Malawi; collins.mitambo@health.gov.mw

* Correspondence: mbulais@mail.nih.gov; Tel: +1-(240)-276-7108; Fax: +1-(240)-276-7836

Citation: Arisue, N.; Chagaluka, G.; Palacpac, N.M.Q.; Johnston, T.; Mutalima, N.; Peprah, S.; Bhatia, K.; Borgstein, E.; Liomba, G.; Kamiza, S.; et al. Assessment of Mixed *Plasmodium falciparum* *sera5* Infection in Endemic Burkitt Lymphoma: A Case-Control Study in Malawi. *Cancers* **2021**, *13*, 1692. <https://doi.org/10.3390/cancers13071692>

Academic Editor: Chalid Assaf

Received: 20 February 2021

Accepted: 29 March 2021

Published: 2 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Simple Summary: *Plasmodium falciparum* (*Pf*) infection is a risk factor for endemic Burkitt lymphoma (eBL), the commonest childhood cancer in Africa, but the biomarkers of *Pf* infection that predict this risk are unknown. There is some evidence that the genetic complexity of *Pf* infection may be a risk factor. In 200 children with versus 140 without eBL in Malawi, this study compared variants of the malaria parasite, focusing on *Pfsera5*, a malaria protein that an infected person's antibodies target to suppress the parasite. Multiple *Pfsera5* variants, which arise when the parasite is not suppressed, were found in 41.7% of eBL children versus 24.3% of other local children, meaning that eBL risk was increased 2.4-fold with multiple *Pfsera5* variants. No specific type of variant was related to eBL risk. Research to quantify malaria parasite variants and to clarify the host immune response needed to control variant infections may yield a test to predict eBL risk.

Abstract: Background: Endemic Burkitt lymphoma (eBL) is the most common childhood cancer in Africa and is linked to *Plasmodium falciparum* (*Pf*) malaria infection, one of the most common and deadly childhood infections in Africa; however, the role of *Pf* genetic diversity is unclear. A potential role of *Pf* genetic diversity in eBL has been suggested by a correlation of age-specific patterns of eBL with the complexity of *Pf* infection in Ghana, Uganda, and Tanzania, as well as a finding of significantly higher *Pf* genetic diversity, based on a sensitive molecular barcode assay, in eBL cases than matched controls in Malawi. We examined this hypothesis by measuring diversity in *Pf*-serine repeat antigen-5 (*Pfsera5*), an antigenic target of blood-stage immunity to malaria, among 200 eBL cases and 140 controls, all *Pf* polymerase chain reaction (PCR)-positive, in Malawi. Methods: We performed *Pfsera5* PCR and sequencing (~3.3 kb over exons II–IV) to determine single or mixed *PfSERA5* infection status. The patterns of *Pfsera5* PCR positivity, mixed infection, sequence variants, and haplotypes among eBL cases, controls, and combined/pooled were analyzed using frequency tables. The association of mixed *Pfsera5* infection with eBL was evaluated using logistic regression,

controlling for age, sex, and previously measured *Pf* genetic diversity. Results: *Pf*sera5 PCR was positive in 108 eBL cases and 70 controls. Mixed *Pf*SERA5 infection was detected in 41.7% of eBL cases versus 24.3% of controls; the odds ratio (OR) was 2.18, and the 95% confidence interval (CI) was 1.12–4.26, which remained significant in adjusted results (adjusted odds ratio [aOR] of 2.40, 95% CI of 1.11–5.17). A total of 29 nucleotide variations and 96 haplotypes were identified, but these were unrelated to eBL. Conclusions: Our results increase the evidence supporting the hypothesis that infection with mixed *Pf* infection is increased with eBL and suggest that measuring *Pf* genetic diversity may provide new insights into the role of *Pf* infection in eBL.

Keywords: *Plasmodium falciparum*; Burkitt lymphoma; Epstein Barr virus; epidemiology; Africa; complexity of infection

1. Introduction

Endemic Burkitt lymphoma (eBL) is an aggressive B-cell non-Hodgkin lymphoma (NHL), first described in African children by Denis Burkitt in 1958 [1]. High incidence of eBL overlaps with holoendemic *Plasmodium falciparum* (*Pf*) infection in Africa [2], and it is one of the commonest childhood cancers [3] in countries with high malaria endemicity, including Malawi, Uganda, Tanzania, and Kenya [4]. The eBL incidence is 30-fold higher in these countries than in non-malaria endemic countries [5]. The correlation of eBL with malaria in population-based and individual-level studies, the findings of altered anti-*Pf* antibody levels in eBL cases compared to healthy children [6–8], and the reduced frequency of genetic variants that protect against malaria in eBL cases compared to healthy children [9,10] suggest that malaria infection may be related to eBL risk. Biologically, *Pf* may increase eBL risk directly by stimulating the polyclonal expansion of B lymphocytes, triggering chromosomal instability in B cells [11,12], or indirectly, by impairing immunologic control of the Epstein–Barr virus (EBV), a known carcinogen for eBL [13].

The biomarkers of *Pf* related to eBL are unknown. A correlation study of 2602 cases in Ghana, Uganda, and Tanzania with published biomarkers (parasitemia, parasite density, and genetic diversity of parasites) of *Pf* infection in the same countries [14] was the first to suggest that *Pf* genetic diversity may be related to eBL, based on a significant association between age-specific patterns of eBL and *Pf* genetic diversity. This hypothesis was evaluated among 303 eBL cases and 274 controls in Malawi by measuring *Pf* diversity using a *Pf*3D7 molecular barcode assay [15]. The results showed a significantly higher mean *Pf* genetic diversity score in the eBL cases than the controls [16]. Although these two studies represent the only evidence linking *Pf* genetic diversity with eBL, the results may be valid because they are based both on population- and individual-level designs, and an assessment of *Pf* diversity using a sensitive and specific molecular assay.

In the present study, we investigated whether the presence of mixed infection at the *Pf* serine repeat antigen-5 (*Pf*sera5) locus (on chromosome 2) [17] is associated with eBL in Malawi. *Pf*sera5 has been identified as a blood-stage vaccine candidate [18,19], in part because antibodies targeting *Pf*sera5, measured using SE36—a recombinant molecule of *Pf*SERA5 [20]—were protective against malaria and eBL [7]. The potential as a vaccine target also prompted interest in the genetic diversity at this locus. *Pf*sera5 codes for a 120-kDa precursor (Figure 1A), which is critical to *Pf* blood-stage infection (and egress from the parasitophorous vacuole) and is an antigenic target of blood-stage immunity to malaria [21]. The *Pf*SERA5 precursor is processed by removing the signal peptide and the protein trimmed into three fragments: the P47, P56, and P18 kDa domains; P56 is further processed to P50 and P6 fragments [22–25]. *Pf*sera5 sequence diversity is introduced by insertions and deletions in the P47 fragment in the protein domain containing the octamer and serine repeats (where protective epitopes are located) [26], and by point mutations in a stretch without repeats [27], but the association of *Pf*sera5 diversity with eBL is unknown.

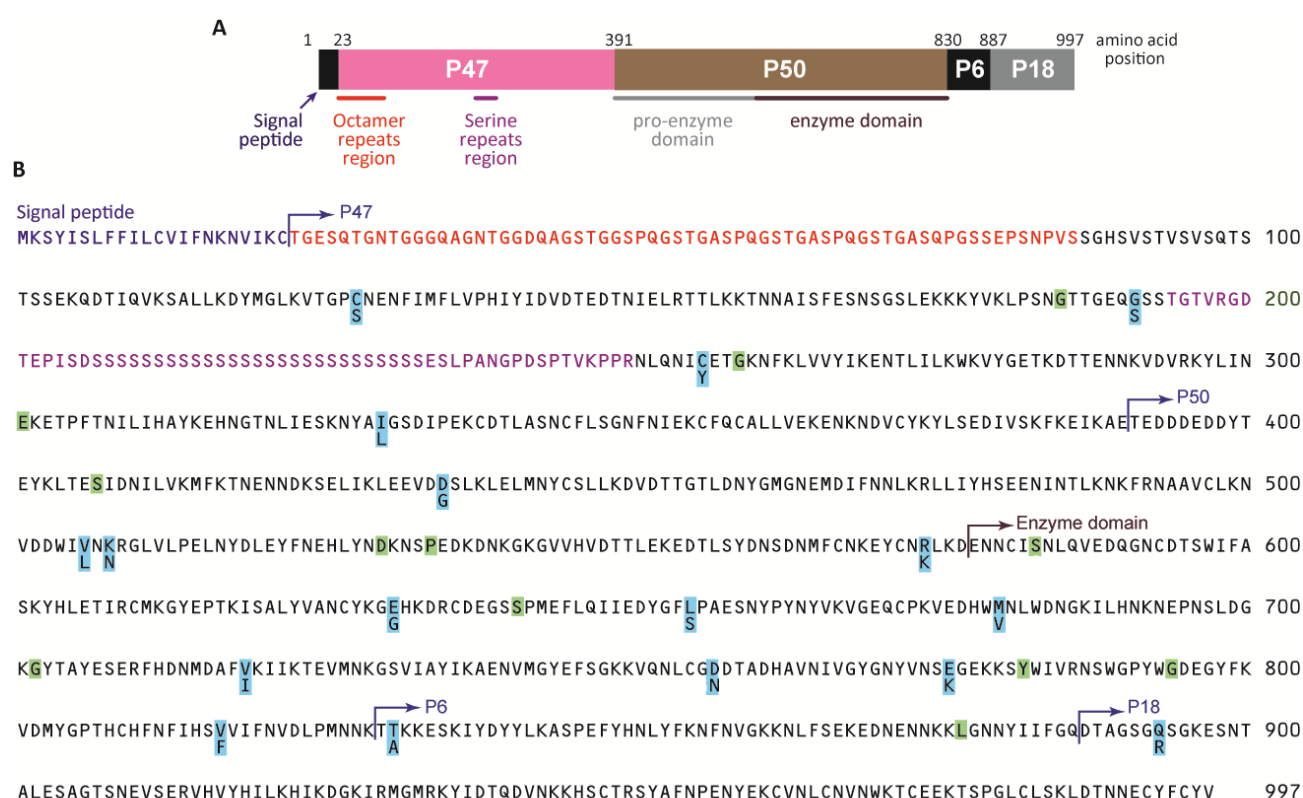


Figure 1. Structure, divergence, and polymorphism of the *P. falciparum* *sera5* gene. **(A)** Schematic structure of the *sera5* gene. Numbers are the amino acid positions of the SERA5 gene in the *P. falciparum* 3D7 strain. **(B)** Amino acid sequence of SERA5 from *P. falciparum* 3D7. Polymorphic amino acid residues are highlighted in orange. Positions with synonymous amino acid changes are highlighted in light green. The signal peptide region is shown in blue. Octamer repeats are shown in red. The serine repeat region with a 13-mer insertion/deletion region and a dimorphic 17-mer region are shown in green. This figure is similar to Figure 1 in Tanabe et al. [27] by the same authors, but the details are different, so they are appropriate for the message in this paper.

2. Materials and Methods

2.1. Study Patients

The study was conducted in the Infections and Childhood Cancer study in Malawi [28]. Briefly, children aged 0 to 15 years diagnosed with cancers (eBL, leukemia, Hodgkin lymphoma, neuroblastoma, rhabdomyosarcoma, Ewing's sarcoma, primitive neuroectodermal tumor, and Wilms' tumor) were enrolled at the Queen Elizabeth Hospital in Blantyre, Malawi between July 2005 and August 2010. All children were reviewed by one investigator (EM) to verify clinical diagnosis. Confirmation by histology, cytology, or other laboratory investigations was done when possible. Children with eBL were coded as cases and those with another diagnosis were coded as controls [16]. Children with HIV infection and those with Kaposi sarcoma were excluded. The current study included only children who were *Pf*-infected based on polymerase chain reaction (PCR) performed on samples taken at the time of diagnosis, before initiating cancer treatment targeting a 519-base pair (bp) segment of the PF07-0076 locus [16].

2.2. Ethics Review

Ethical approval was given by the Malawian College of Medicine Research and Ethics Committee (P.03/04/277R) and the Office of Human Subjects Research at the National Institutes of Health (Exempt #: 4742). The parents/guardians of the children gave written informed consent.

2.3. Laboratory Methods

2.3.1. *Pf*sera5 Gene Amplification

The *Pf*SERA5 amplification was performed on residual genomic DNA extracted from coded/masked whole blood samples taken at the time of diagnosis before initiating cancer treatment, using a QIAamp Blood DNA Kit (Qiagen, Inc., Valencia, CA, USA) at Osaka University [16]. The *Pf*SERA5 PCR amplification targeted the ~3.3 kb region in exons II–IV using the specific primers sera5-5F0 and sera5-3R0 (Table S1). The sequence before and after exon I (encoding the signal peptide) is AT-rich and contains poly-A residues, which makes it difficult to successfully sequence; thus, this region was excluded in our analysis. PCR amplification was carried out in a 25- μ L reaction mixture containing 0.4 μ M each of forward and reverse primers, 0.4 mM of deoxyribonucleotides (dNTP), 0.5 units of KOD FX Neo polymerase (TOYOBO Co., Ltd., Osaka, Japan), 12.5 μ L of 2 \times PCR buffer, and 1 μ L of genomic DNA solution. The PCR conditions were as follows: 95 $^{\circ}$ C for 2 min, 40 cycles of 95 $^{\circ}$ C for 30 sec, 59 $^{\circ}$ C for 30 sec, and 68 $^{\circ}$ C for 4 min. A 2- μ L aliquot of the PCR product was used as a template for a second PCR amplification in a 25- μ L reaction mixture using the primers sera5-5F3 and sera5-3R2 (Table S1) under the same thermocycler conditions. Samples that were not successfully amplified using this primer set were re-tested using nested PCR primers targeting half the original length of the two fragments (a 5' half and a 3' half). The nested primer set used for the 5' half fragments were sera5-5F0 and sera5-R0 for the first PCR, and sera5-5F3 and sera5-R0 for the nested PCR; and for the 3' half fragments, sera5-F1 and sera5-3R0 were used for the first PCR, followed by sera5-F2 and sera5-3R2 (Table S1). The nested PCR conditions were the same as for the full-length PCR, but the extension time was shortened to 2 min. There were 12 duplicate samples included for quality control. The results were concordant for PCR and *Pf*SERA5 single or mixed infection in nine samples at the original testing, and three samples were resolved after re-testing; two were PCR-negative and one was a single infection.

2.3.2. *Pf*sera5 Nucleotide Sequencing

The PCR products were purified using a QIAquick PCR Purification kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Purified DNA fragments were eluted in 30 μ L of Tris-EDTA buffer (TE). The optical density was measured with a NanoDrop (Thermo Fisher Scientific, Wilmington, DE, USA) and the DNA concentration was adjusted to 0.026 μ g/ μ L for the 3.3 kb fragments and 0.013 μ g/ μ L for the 5'-half and 3'-half fragments using TE. At this concentration, 1 μ L was suitable for performing one sequencing reaction. *Pf*sera5 DNA sequencing was performed directly using The BigDye[®] Terminator v3.1 Cycle Sequencing Kit and 3130xI Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Sequencing primers were designed to cover target regions in both directions (Table S1). The nucleotide sequences generated during the current study are available through public database such as DDBJ and NCBI (Accession numbers: LC606291–LC606405). The sequence results without overlapping peaks on the electropherograms were interpreted as being single parasite infections; otherwise, the sample was recorded as having a mixed parasite infection. The *Pf*SERA5 sequence variations in single parasite infections were analyzed to get insights into variation in populations.

2.3.3. Sequence Alignment and *Pf* Population Genetics Analyses

The *Pf*SERA5 nucleotide sequence obtained from each sample was aligned to the *Pf*SERA5 3D7 strain using CLUSTALW implemented in GENETYX® ver. 15 (GENETYX Co., Ltd., Tokyo, Japan) and the alignment was manually inspected to ensure alignment accuracy. Analyses were done for the entire sequence and for specific protein domains: the 2562 bp nucleotide sequence includes a stretch without repeats (a non-repeat region, or NonR); the stretch of octamer repeats (OctR) and the stretch with serine repeats (SerR) correspond to amino acid positions 87–193 and 251–997, respectively [27]. The sequence information from the NonR was translated into its 854 amino acid (aa) sequences to identify nonsynonymous (variations that result in changes in an aa sequence) versus synonymous (no change in an aa sequence). To gain insights into *Pf*sera5 variations in samples from nearby Tanzania or countries with a high eBL incidence (e.g., Ghana), we accessed the *Pf*sera5 sequence from 55 asymptomatic donors in the Rufiji River Delta in Tanzania sampled in 1993, 1998, and 2003 (Accession numbers: AB634928–AB634982) [29], and from 33 children in three villages in Ghana sampled in 2004 (Accession numbers: AB634983–AB635015) [29].

Haplotype diversity (H_d) and nucleotide diversity (π , the average number of nucleotide differences per site between two sequences in all possible pairs in the sample population) were calculated using DnaSP v5.10.01 [30]. The difference between the numbers of synonymous substitutions per synonymous site (d_S) and of nonsynonymous substitutions per nonsynonymous site (d_N) was calculated by the Nei and Gojobori method [31], implemented in MEGA X, with the Jukes and Cantor correction. The statistical significance of the difference between d_N and d_S was estimated using the MEGA Z-test [32]. Higher d_S may suggest that active, synonymous amino acid substitutions that are neutral or can boost parasite fitness are more likely to be retained or accumulated, while nonsynonymous amino acid substitutions that reduce parasite fitness are removed or reduced in frequency. Wright's fixation index F_{st} was calculated to assess the genetic differentiation of *Pf*sera5 among the parasite circulating in the study population [33]. The pairwise F_{st} between parasite populations was calculated using Arlequin v3.5 [34]. A small F_{st} indicates that the parasite allele frequencies within the study populations are similar, whereas a large F_{st} indicates that the allele frequencies are different, and the study populations do not share genetic diversity.

2.3.4. Statistical Analyses

Patterns of *Pf*sera5 PCR positivity, mixed *Pf*sera5 infection, and *Pf*sera5 haplotypes and nucleotide diversity in eBL cases or controls, separately or combined/pooled, were analyzed for the entire *Pf*sera5 sequence and by domain (OctR, SerR, and NonR). The results were categorized into known haplotypes; the many “rare haplotypes” (defined among Malawi controls as SerR haplotypes observed in <10% or OctR haplotypes observed in <20% of the samples) were grouped and then compared to the common haplotypes. This grouping is post hoc, but it is useful for exploring pairwise comparisons of common versus rare haplotypes in eBL cases and controls, as well as haplotype patterns observed in Ghana and Tanzania. Differences across sample groups were evaluated using Fisher's exact test. Because age is an important predictor of exposure and immunity to malaria, the data were explored by age (<6 versus ≥ 6 years of age).

The association of mixed *Pf*SERA5 infection with eBL was assessed by calculating odds ratios and 95% confidence intervals (OR 95% CI) using logistic regression. Bivariate ORs (bORs) were obtained by adjusting separately for age, sex, and the *Pf* genetic diversity score as confounders [16], before including all these variables into one model for multivariate adjustment (aOR), but the results were not adjusted for multiple comparisons; thus, two-sided p -values < 0.05 were considered statistically significant. However, other comparisons must be interpreted cautiously for hypothesis generation.

3. Results

3.1. Characteristics of Study Subjects

The current study evaluated 341 *Pf* PCR-positive participants. One participant was >17 years and excluded from subsequent analyses. The remaining participants included 200 eBL cases and 140 controls (Table 1). Most participants were from Malawi ($n = 327$), but 13 (10 eBL cases and 3 controls) were from Mozambique, a neighboring country. Because of the small sample size of the children originating in Mozambique, the sequence data from those children were analyzed together with those from Malawi, where all children were enrolled. The three leading diagnoses among the controls were renal tumors ($n = 28$), non-eBL lymphoma and leukemia ($n = 35$, including 9 with leukemia), and soft tissue tumors ($n = 19$). The eBL cases were slightly older than the controls, but the difference was not statistically significant (7.2 years versus 6.9 years, $p = 0.44$). However, the eBL cases were more likely to be aged 6–10 years than the controls (50.5% versus 25.7%, heterogeneity <0.001). The male-to-female proportion of eBL cases and controls was not statistically different in this set (62.5% versus 58.6% in females, $p = 0.50$).

Table 1. Characteristics of Malawi and Mozambique samples used in this study.

Subject Characteristics	Sample Size	Gender	Age Group			Sera5 PCR		Sera5 Sequence				
		Male (%)	Missing age	Mean (SD)	0–5yrs	6–10yrs	11–15yrs *	Negative	Positive	Single	Mixed	Indeterminate †
Case Status												
Cases	200	125 (62.5)	8	7.2 (2.8)	69 (34.5)	101 (50.5)	22 (11.0)	92 (46.0)	108 (54.0)	63 (58.3)	45 (41.7)	-
Controls	140	82 (58.6)	5	6.9 (4.4)	66 (47.1)	36 (25.7)	33 (23.6)	70 (50.0)	70 (50.0)	52 (74.3)	17 (24.3)	1 (1.4)
Association with eBL												
Odds ratio [OR 95% CI]	-	-	-	-	-	-	-	-	-	Ref	2.18 (1.12, 4.26)	-
Adjusted OR ‡ [95% CI]	-	-	-	-	-	-	-	-	-	Ref	2.40 (1.11, 5.17)	-
Diagnose in Controls												
Leukemias	9	3 (33.3)	0	7.6 (4.5)	4 (44.4)	2 (22.2)	3 (33.3)	7 (77.8)	2 (22.2)	1 (50.0)	1 (50.0)	-
Lymphomas	26	18 (69.2)	1	10.2 (4.1)	4 (16.0)	8 (32.0)	13 (52.0)	13 (50.0)	13 (50.0)	11 (84.6)	2 (15.4)	-
Neuroblastomas	8	5 (62.5)	0	7.0 (3.9)	3 (37.5)	4 (50.0)	1 (12.5)	6 (75.0)	2 (25.0)	2 (100.0)	0 (0.0)	-
Retinoblastoma	9	3 (33.3)	0	4.0 (2.8)	8 (88.9)	0 (0.0)	1 (11.1)	5 (55.6)	4 (44.4)	2 (50.0)	2 (50.0)	-
Renal tumors	28	21 (75.0)	0	3.6 (2.9)	24 (85.7)	3 (10.7)	1 (3.6)	8 (28.6)	20 (71.4)	15 (75.0)	5 (25.0)	-
Hepatic tumors	6	3 (50.0)	0	7.6 (2.7)	2 (33.3)	3 (50.0)	1 (16.7)	2 (33.3)	4 (66.7)	3 (75.0)	1 (25.0)	-
Bone tumors	2	1 (50.0)	0	10.3 (5.4)	0 (0.0)	1 (50.0)	1 (50.0)	1 (50.0)	1 (50.0)	1 (100.0)	0 (0.0)	-
Soft tissue sarcomas	19	14 (73.7)	1	7.8 (4.2)	8 (44.4)	5 (27.8)	5 (27.8)	11 (57.9)	8 (42.1)	6 (75.0)	2 (25.0)	-
Germ cell tumors	8	3 (37.5)	0	6.8 (4.0)	4 (50.0)	2 (25.0)	2 (25.0)	3 (37.5)	5 (62.5)	3 (60.0)	2 (40.0)	-
Epithelial tumors	2	0 (0.0)	0	6.7 (7.8)	1 (50.0)	0 (0.0)	1 (50.0)	1 (50.0)	1 (50.0)	1 (100.0)	0 (0.0)	-
Other tumors	1	0 (0.0)	0	13.5 (-)	0 (0.0)	0 (0.0)	1 (100.0)	1 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	-
Non-malignancies	10	6 (60.0)	0	7.3 (4.7)	3 (30.0)	5 (50.0)	2 (20.0)	5 (50.0)	5 (50.0)	4 (80.0)	0 (0.0)	1 (20.0)
Not well specified	12	5 (41.7)	3	6.2 (4.6)	5 (55.6)	3 (33.3)	1 (11.1)	7 (58.3)	5 (41.7)	3 (60.0)	2 (40.0)	-

* Two controls between 16 and 16.2 years; † Unclear if positive infection detected was single or mixed; ‡ Odds ratio adjusted for diversity score, DNA copy number, gender, and age.

Pfsera5 PCR positivity was detected in 108 (54.0%) eBL cases and 70 (50.0%) controls (Table 1). The PCR product was insufficient in one control sample and cannot be conclusively ascertained if the sample is a single or mixed *Pfsera5* infection. Likewise, not all *Pf* positive samples were suitable for amplification of the near full-length *Pfsera5*, due mainly to the quality of samples available for the study.

3.2. Association of Mixed *Pfsera5* Infection with eBL

Pfsera5 sequences revealed mixed infection in 45 (41.7%) eBL cases and 17 (24.3%) controls. Mixed *Pfsera5* infection was associated with eBL risk in crude analysis (cOR of 2.18, 95% CI of 1.12–4.26; Table 1; Table S2) and in bivariate analyses adjusting for gender (bOR of 2.18, 95% CI of 1.12–4.26), *Pf* diversity score (bOR of 2.23, 95% CI of 1.10–4.53), *Pf* DNA copy number (bOR of 1.95, 95% CI of 0.96–3.95), and age group (<6, 6–10, 11+ years) (bOR of 2.44, 95% CI of 1.12–4.26). The association of mixed *Pfsera5* infection with eBL remained in models adjusting for all the variables above (aOR of 2.40, 95% CI of 1.11–5.17).

3.3. *Pfsera5* Haplotype Diversity in eBL Cases and Controls in Malawi

Analysis of Hd and π were calculated for 115 samples in Malawi and 88 individuals in Tanzania and Ghana with single *Pfsera5* infection. As shown in Figure 1B, the parasites harbored variants at 29 nucleotide positions in the sequenced 2562 bp NonR region; 17 led to nonsynonymous amino acid changes (highlighted orange) and 12 led to synonymous amino acid changes (highlighted green). The numbers of haplotypes are shown in Table 2. Analysis of the entire *Pfsera5* sequence yielded 96 haplotypes (88 haplotypes at the amino acid sequence level), 34 in the OctR, 42 in the SerR, and 26 in the NonR regions (Table 2, detailed in Figures S1 and S2). The Hd for the entire *Pfsera5* sequence was 0.993 in eBL cases, 0.996 in the controls, and 0.991 and 0.996 in the samples from Tanzania and Ghana, respectively (Table 2). When Hd was considered by *Pfsera5* regions, it was lowest for the NonR region and moderately low for the OctR region; Hd was similar for the entire *Pfsera5* and the SerR region, which bears the protective epitopes of *PfSERA5* [26]. Hd in the NonR region was low in Tanzania (0.362) and Ghana (0.472) compared to that in the eBL cases and controls combined (0.545) (Malawi and Mozambique).

Table 2. Haplotype diversity of the entire sequence, and the octamer repeat (OctR), serine repeat (SerR), and non-repeat (NonR) 2562 base pair (bp) regions of the *P. falciparum* *SERA5* gene.

Country	n	Entire Sequence		OctR Region		SerR Region		NonR 2562 bp Region	
		No. of Haplotypes	Hd \pm SD	No. of Haplotypes	Hd \pm SD	No. of Haplotypes	Hd \pm SD	No. of Haplotypes	Hd \pm SD
Malawi and Mozambique	115	96 (88)	0.994 \pm 0.003	34 (34)	0.809 \pm 0.032	42 (39)	0.935 \pm 0.011	26 (16)	0.545 \pm 0.011
eBL case	63	54 (50)	0.993 \pm 0.005	19 (19)	0.786 \pm 0.047	28 (27)	0.928 \pm 0.019	16 (9)	0.551 \pm 0.075
Control	52	47 (45)	0.996 \pm 0.002	21 (21)	0.839 \pm 0.040	26 (23)	0.945 \pm 0.016	14 (9)	0.544 \pm 0.083
Tanzania	55	44 (42)	0.991 \pm 0.006	21(21)	0.804 \pm 0.054	30 (30)	0.968 \pm 0.010	11 (9)	0.362 \pm 0.084
Ghana	33	31 (29)	0.996 \pm 0.009	16 (16)	0.873 \pm 0.048	23 (23)	0.968 \pm 0.018	8 (7)	0.472 \pm 0.106

The numbers of amino acid sequence variations are shown in parenthesis. eBL: endemic Burkitt Lymphoma, Hd: Haplotype diversity, SD: standard deviation

3.4. Patterns of Common Versus Rare *Pfsera5* Haplotypes in eBL Cases and Controls

In analyses restricted to those with a single *Pfsera5* infection, the frequency of the common haplotypes in the OctR and SerR regions was similar in children with eBL and controls. The common haplotypes in the OctR and SerR regions were more frequent in the combined eBL cases and the controls in Malawi and Mozambique than in Ghana and Tanzania (Figure 2, Figures S1 and S2, $p < 0.05$).

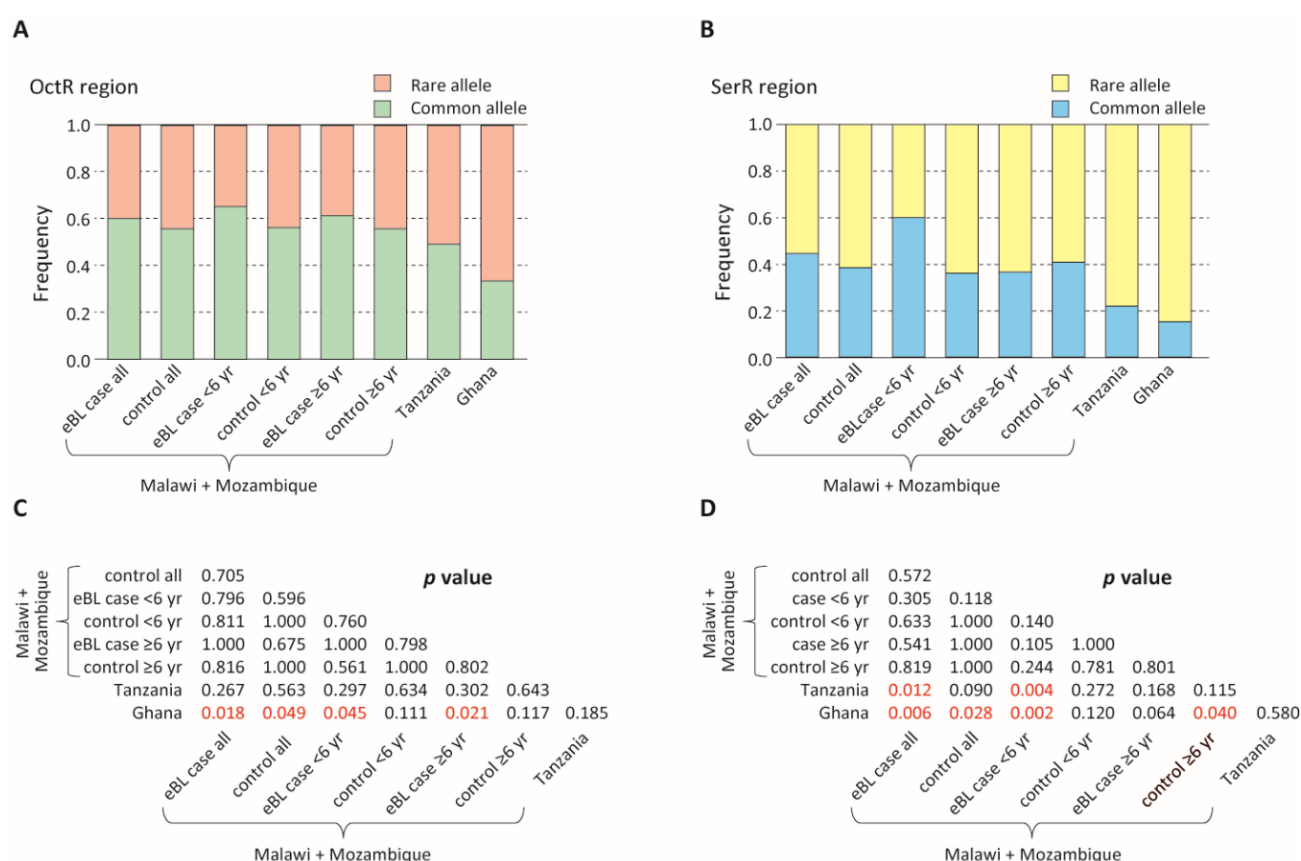


Figure 2. Comparison of the proportion of rare and common alleles among endemic Burkitt lymphoma (eBL) cases and controls. **(A)** Comparison of the proportion of the rare and common alleles in the octamer repeat (OctR) region. Alleles with a frequency >0.2 are defined as common alleles, while the rest are defined as rare alleles. **(B)** Comparison of the proportion of the rare and common alleles in the serine repeat (SerR) region. Alleles with a frequency >0.1 are defined as common alleles, while the rest are defined as rare alleles. **(C,D)** p -values of the Fisher's exact test (C: OctR region; D: SerR region). $p < 0.05$ are shown in red.

3.5. Genetic Differentiation between Parasite Populations

Table 3 shows nucleotide diversity (π) in the 2562 bp NonR and Table 4 shows the location of the 29 polymorphic sites (12 synonymous and 17 non-synonymous). There are 23 new variations, that is, variations not previously reported, while 6 (5 non-synonymous and 1 synonymous) have previously been reported [21], but the specific variations were not different between the cases and controls. The eBL cases, controls, and individuals in Tanzania and Ghana had similar nucleotide diversity (π). However, the current eBL cases and controls (combined, Malawi and Mozambique) had $dS > dN$ while the opposite pattern (i.e., $dN > dS$) was observed in the samples from Tanzania ($p < 0.06$) and Ghana ($p = 0.0034$) (Table 3). *Fst* analysis suggested genetic differentiation ($p < 0.05$) in the OctR and SerR, but not the NonR regions for parasites in the eBL cases and controls (combined) and in individuals in Ghana (Table 5). However, *Fst* analysis did not support evidence of genetic differentiation between parasites in Malawi and Mozambique versus those in Tanzania and Ghana.

Table 3. Nucleotide divergence and polymorphism in the non-repeat 2562 bp sequence regions.

Country	<i>n</i>	No. of Polymorphic Sites	No. of Substitutions		Nucleotide Diversity	Codon-Based Evolutionary Divergence		
			Synonymous	Non-Synonymous	$\pi \pm SD$	dS \pm SE	dN \pm SE	<i>p</i> Value
Malawi and Mozambique	115	29	12	17	0.00029 \pm 0.00004	0.00050 \pm 0.00012	0.00024 \pm 0.00008	dS > dN: 0.0595
eBL case	63	17	8	9	0.00028 \pm 0.00005	0.00055 \pm 0.00017	0.00021 \pm 0.00009	dS > dN: 0.0450
Control	52	16	6	10	0.00031 \pm 0.00006	0.00044 \pm 0.00017	0.00027 \pm 0.00011	dS > dN: 0.1943
Tanzania	55	11	1	10	0.00021 \pm 0.00007	0.00007 \pm 0.00007	0.00025 \pm 0.00008	dN > dS: 0.0549
Ghana	33	10	1	10	0.00032 \pm 0.00011	0.00000 \pm 0.00000	0.00040 \pm 0.00015	dN > dS: 0.0034

π : nucleotide diversity, SD: standard deviation, dS: the numbers of synonymous substitutions per synonymous site, SE: standard error, dN: nonsynonymous substitutions per nonsynonymous site, eBL: endemic Burkitt Lymphoma.

Table 4. Sequence variation in *P. falciparum* *SERA5* in the non-repeat 2562 bp region. Nucleotide and amino acid positions are numbered after the *P. falciparum* 3D7 sequence. The number of isolates with variations are shown. Variations shared with previously reported *P. falciparum* *sera5* sequences in PlasmoDB (<https://plasmodb.org/plasmo/> accessed on 23 October 2020) and NCBI (<https://www.ncbi.nlm.nih.gov/> accessed on 23 October 2020) databases are shown in bold font.

(A) Non-synonymous change.																												
Nucleotide Position	3	4	5	5	5	7	7	7	7	9	3	5	5	6	7	8	9	0	0	1	2	3	3	4	4	6	8	8
	8	7	6	6	7	6	8	9	8	0	1	2	1	2	9	6	3	6	5	6	2	2	4	9	7	0	3	
	3	5	2	6	1	7	2	2	8	4	6	4	3	1	2	4	8	5	5	9	5	6	9	1	8	1	3	
Common	G	K	G	A	G	G	G	G	A	A	G	G	A	G	A	T	A	T	G	G	C	G	G	A	A	G	A	A
Variation	C	G	A	C	A	A	A	T	T	G	C	T	T	A	G	C	G	G	A	A	A	A	T	G	G	A	G	G
Amino Acid Position	1	1	1	1	1	2	2	2	3	4	5	5	5	5	6	6	6	6	7	7	7	7	8	8	8	9	9	
	2	5	8	8	9	5	6	6	3	3	0	0	3	7	3	5	8	8	1	5	7	7	1	3	9	3	4	
	8	9	8	9	1	6	1	4	0	5	6	8	8	4	1	5	0	9	9	7	5	6	7	1	3	4	5	
Common	C	K	G	E	G	C	N	L	I	D	V	K	D	R	E	L	M	L	V	D	S	E	V	T	Q	R	K	
Variation	S	E	S	A	S	Y	S	F	L	G	L	N	V	K	G	S	V	V	I	N	R	K	F	A	R	H	E	
Malawi and Mozambique, n = 115	1	-	-	-	1	1	-	-	1	1	1	1	-	1	1	1	12	-	1	1	-	2	1	1	1	-	-	
eBL case, n = 63	1	-	-	-	1	-	-	-	-	1	1	-	-	-	-	1	6	-	-	-	-	1	-	1	1	-	-	
Control, n = 52	-	-	-	-	-	1	-	-	1	-	-	1	-	1	1	-	6	-	1	1	-	1	1	-	-	-	-	
Tanzania, n = 55	-	1	1	-	2	-	1	-	1	-	-	-	2	-	-	-	3	-	-	-	1	-	-	-	-	1	1	
Ghana, n = 33	-	1	1	1	1	-	-	1	1	-	-	-	-	-	-	-	4	1	-	-	-	-	-	-	-	1	1	
(B) Synonymous change.																												
Nucleotide Position	5	7	7	9	2	5	6	7	9	1	1	6	7	9	1	3	8	3	6	1	3	8	3	6	1	3	6	
	6	7	9	0	2	9	0	4	2	0	4	2	0	4	8	3	6	1	6	1	6	2	1	6	1	3	6	
	4	7	2	3	1	0	2	9	3	6	6	2	1	6	2	1	6	1	6	1	6	2	1	6	1	3	6	
Common	T	A	T	A	T	A	T	T	A	T	T	A	T	T	A	T	T	A	T	T	A	T	T	A	T	T	A	
Variation	G	G	C	G	C	C	G	G	C	C	C	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
Amino Acid Position	1	2	2	3	4	5	5	5	6	7	7	7	7	8	9	9	9	9	9	9	9	9	9	9	9	9	9	
	8	5	6	0	0	3	3	8	4	0	8	0	3	8	9	8	9	8	9	8	9	8	9	8	9	8		
	8	9	4	1	7	0	4	3	1	2	2	4	3	1	2	2	4	3	1	2	2	4	3	1	2	2		
Amino Acid	G	G	L	E	S	D	P	S	S	G	Y	G	L	S	G	Y	G	L	S	G	Y	G	L	S	G	L		
Malawi and Mozambique, n = 115	2	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
eBL Case, n = 63	2	1	-	1	1	1	1	-	1	1	-	1	-	1	-	-	1	-	-	1	-	1	-	-	-	-	-	
Control, n = 52	-	-	-	-	-	-	-	1	-	-	1	-	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	
Tanzania, n = 55	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	
Ghana, n = 33	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

Table 5. Genetic differentiation (*Fst*) of the *P. falciparum* *sera5*.

<i>P. falciparum</i> SERA5 region / Country	Subject Type	Malawi and Mozambique			Tanzania	Ghana
		All	eBL Case	Control	All	All
(A) Octamer repeat region						
Malawi and Mozambique	Combined	-	0.9910	0.9910	0.0721	0.0360
	eBL case	-0.0094	-	0.4595	0.1171	0.0090
	Control	-0.0097	-0.0030	-	0.0811	0.0270
Tanzania	All subjects	0.0113	0.0085	0.0130	-	0.0811
Ghana	All subjects	0.0291	0.0271	0.0286	0.0104	-
(B) Serine repeat region						
Malawi and Mozambique	Combined	-	0.9910	0.9910	0.0811	0.0000
	eBL case	-0.0093	-	0.6216	0.0541	0.0000
	Control	-0.0096	-0.0027	-	0.3604	0.0451
Tanzania	All subjects	0.0069	0.0100	0.0013	-	0.2072
Ghana	All subjects	0.0202	0.0204	0.0181	0.0039	-
(C) Non-repeat 2562 bp sequence						
Malawi and Mozambique	Combined	-	0.9910	0.9910	0.2973	0.5676
	eBL case	-0.0099	-	0.9279	0.4234	0.5946
	Control	-0.0105	-0.0056	-	0.2342	0.8018
Tanzania	All subjects	0.0011	-0.0002	0.0023	-	0.5586
Ghana	All subjects	-0.0033	-0.0037	-0.0067	-0.0044	-

Fst values are shown in the bottom left and *p*-values are shown in the upper right. *p* < 0.05 are shown in bold. eBL: endemic Burkitt Lymphoma.

4. Discussion

We investigated the hypothesis that *Pf* genetic diversity, based on the detection of mixed *Pfsera5* infection, is associated with eBL in Malawi [16]. Our findings of 2.40-fold higher odds of mixed *PfSERA5* infection in eBL cases than controls, which remained statistically significant after the adjustment of confounders, including gender, age group, *Pf* DNA copy, and *Pf* genetic diversity [16], led us to reject the null hypothesis. These results strengthen the evidence that *Pf* genetic diversity may be related to eBL, as highlighted in our earlier ecological study in Uganda, Tanzania, and Ghana [14] and in our case-control in Malawi using a sensitive molecular assay [16].

There is accumulating evidence that mixed *Pf* infection may be a molecular surrogate of chronic asymptomatic *Pf* infection in high malaria transmission areas [35]. The consequences for malaria risk appear similar to earlier studies, that is, the risk for malaria (infection and symptoms) is high in children below age 5 [36,37] but low in children above age 5 [38], consistent with the level of acquired immunity against malaria at those ages [39]. Thus, in older children, mixed infection appears to be a molecular marker of asymptomatic infection [35], incomplete clearance of parasites [40], clinical and parasite immunity [41], and, according to our results, eBL risk. A recent finding of a median of the multiplicity of infection of six haplotypes in infective mosquitoes suggests that mixed infections may arise from co-transmission of genetically diverse oocysts from a single mosquito bite [35]. Super-infection, that is, the inoculation of different parasites from different mosquitoes, is also possible [42] in partially immune people living in malarious, non-sheltered environments [43].

Chronic, asymptomatic, low-density parasite infection has been reported in eBL cases and may precede the disease [40]. Asymptomatic infection is due, in part, to antibodies that target parasite proteins, particularly those that are secreted and embedded in infected red blood cell (iRBC) membranes, thereby blocking sequestration and promoting splenic clearance of iRBCs [44–46], and thus suppressing parasite density and clinical symptoms. However, because asymptomatic infections are sub-clinical, they remain untreated for

prolonged periods; the associated chronic immune response may lead to eBL as a rare complication in children unsheltered from malaria [8]. While a strong immune response may explain the reduced frequency of symptomatic malaria and lower parasite densities in eBL cases than other children in the same region [40], we speculate that *Pfsera5* may contribute to the molecular camouflage that facilitates a tolerance of low-grade infection and sets the stage for progression to cancer. *PfSERA5* N-terminal domain decorates the surface of the merozoite where it tightly binds to a host serum protein, vitronectin, which in turn binds to another host protein that camouflages the parasite from the host immune system [47], potentially reducing pressure for mutations in the *Pfsera5* gene. However, this immune response is not completely effective, resulting in immune tolerance, high antibody titers against *PfSERA5*, and low-grade parasitemia [20]. There is still a dearth of understanding in the chronic nature of malaria infection and how it amplifies or down-regulates the immune response, although the continuous insult by low-grade, often genetically complex variants (mean of 4) [35] likely increases the risk of genetic instability in B cells and their progression to eBL [11]. In agreement with this hypothesis, the spleen is frequently enlarged in children with asymptomatic *Pf* infection [48], whom we hypothesize are the population at risk of eBL [40]. Plausibly, a robust B-cell response promotes survival against malaria [49], on one hand by suppressing *Pf* positivity in eBL [40,43,50], while promoting eBL risk in children with asymptomatic infection, on the other hand.

The use of *Pfsera5* molecular data expands our knowledge of variants in *Pf* field isolates in countries where eBL incidence is high [4]. Consistent with previous reports [27], we observed that *Pfsera5* is highly conserved with limited evidence of genetic differentiation, consistent with the frequent gene flow in Africa. We also observed that *Pfsera5* diversity was not unique to Malawi and Mozambique or to eBL cases. These results are reassuring because our comparisons are based on samples obtained using varied, potentially biased methods. The samples from Ghana and Tanzania are from specific communities living in relatively small geographical areas, whereas the Malawi and Mozambique samples are from a much larger geographical area. Thus, the observations that no specific *Pf* variants or haplotypes are associated with eBL and that both eBL cases and controls in Malawi harbor common haplotypes are consistent with both groups being exposed to similarly high environmental pressure of *Pf* infection [35]. However, we note that rare haplotypes may be harder to detect because they exist at lower submicroscopic levels. Our findings of a lower frequency of the common haplotypes in Tanzania and Ghana than in Malawi and Mozambique may be due to geographic confounding, as noted above.

Our research suggests that molecular markers of *Pf* genetic diversity may complement the search for biomarkers of *Pf* infection related to eBL risk. The general finding that *PfSERA5* does not show antigenic variation [20] suggests that *Pfsera5* might be a good tool to characterize the genetic complexity of parasite populations in eBL cases and controls. However, this should be balanced against the likelihood that the sample size will shrink because amplifying about 3.2 Kb of *Pfsera5* requires higher quality gDNA than amplifying the 519 bp that we used to confirm *Pf* infection in all our samples. This explains why our sample size shrunk by 50%. Other limitations of our study include the use of cross-sectional samples from a case-control study, and the testing of one sample from a single time point, which underestimates parasite diversity [51]. These limitations are balanced by our hypothesis-driven approach, our focus on diversity at a locus that is an antigenic target of blood-stage immunity to malaria, and the use of cancer controls, in whom referral bias and reverse-causality biases are minimized. Our study included only children who were confirmed to have *Pf* infection by PCR [16], which increases the validity of our assumption that all children were exposed to infection.

5. Conclusions

Our study provides new evidence for the association of mixed *Pf* infection with eBL risk in Malawi. Further research utilizing molecular markers of *Pf* infection may lead to the discovery of biomarkers of eBL, severe malaria, and/or asymptomatic infection.

Supplementary Materials: The following are available online at www.mdpi.com/2072-6694/13/7/1692/s1: Figure S1. Allele distribution in *sera5* octamer repeat (OctR) region, Figure S2. Allele distribution in SERA5 serine repeat (SerR) region, Table S1. Primers for PCR amplification and sequencing, Table S2. Odds ratio (OR) and 95% confidence interval (CI) for eBL case status in univariate and multivariate logistic regression models.

Author Contributions: N.A., N.M.Q.P., T.H., and S.M.M. conceived, designed, and supervised the study, interpreted data, and critically edited the manuscript. N.A. performed the laboratory experiments, interpreted the data, and drafted the manuscript. W.T.J. prepared the epidemiological analysis datasets. R.N., N.M. (Nora Mutalima), and E.M.M. are the principal investigators of the Infections and Childhood Cancer study in Malawi. E.B., G.N.L., S.K., and N.M. (Nyengo Mkandawire) performed the fieldwork. S.P. performed the analyses. C.M., K.B., G.C., and J.J.G. interpreted the data. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute at the National Institutes of Health, Department of Health and Human Services (Support Services Contract NO2-CP-31003 and HHSN261201100007I). The study was also funded by The Biomedical Kansai project supported by the Regional Innovation Cluster Program of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. The original study in Malawi was funded by Cancer Research UK and by the Rhodes Trust (Fellowship awarded to Nora Mutalima).

Institutional Review Board Statement: Ethical approval to conduct the Infections and Childhood Cancer study was given by the Malawian College of Medicine Research and Ethics Committee (P.03/04/277R date of approval: December 7th, 2011) and the Office of Human Subjects Research at the National Institutes of Health (Exempt #: 4742).

Informed Consent Statement: Written informed consent was obtained from the parents or guardians of the children.

Data Availability Statement: The *Pfsera5* sequence data has been uploaded to DDBJ (Accession numbers: LC606291-LC606405). The remaining clinical data and the code used for these analyses are available upon reasonable request from the corresponding authors.

Acknowledgments: We are grateful to the participants and to all the local doctors in Malawi who contributed to this study, and to the College of Medicine Research and Ethics Committee at the University of Malawi and Malawi National Health Sciences Research Committee for approving the work to be done.

Conflicts of Interest: The authors declare that they have no conflict of interests.

References

1. Burkitt, D. A sarcoma involving the jaws in african children. *Br. J. Surg.* **1958**, *46*, 218–223.
2. Bouvard, V.; Baan, R.A.; Grosse, Y.; Lauby-Secretan, B.; El Ghissassi, F.; Benbrahim-Tallaa, L.; Guha, N.; Straif, K. Carcinogenicity of malaria and of some polyomaviruses. *Lancet Oncol.* **2012**, *13*, 339–340.
3. Stiller, C.A.; Parkin, D.M. International variations in the incidence of childhood lymphomas. *Paediatr. Perinat. Epidemiol.* **1990**, *4*, 303–324.
4. Hämmerl, L.; Colombet, M.; Rochford, R.; Ogwang, D.M.; Parkin, D.M. The burden of Burkitt lymphoma in Africa. *Infect. Agents Cancer* **2019**, *14*, 1–6, doi:10.1186/s13027-019-0236-7.
5. Linet, M.S.; Brown, L.M.; Mbulaiteye, S.M.; Check, D.; Ostroumova, E.; Landgren, A.; Devesa, S.S. International long-term trends and recent patterns in the incidence of leukemias and lymphomas among children and adolescents ages 0–19 years. *Int. J. Cancer* **2016**, *138*, 1862–1874.
6. Carpenter, L.M.; Newton, R.; Casabonne, D.; Ziegler, J.; Mbulaiteye, S.; Mbidde, E.; Wabinga, H.; Jaffe, H.; Beral, V. Antibodies against malaria and epstein-barr virus in childhood burkitt lymphoma: A case-control study in uganda. *Int. J. Cancer* **2008**, *122*, 1319–1323.
7. Aka, P.; Vila, M.C.; Jariwala, A.; Nkrumah, F.; Emmanuel, B.; Yagi, M.; Palacpac, N.M.; Periago, M.V.; Neequaye, J.; Kiruthu, C.; et al. Endemic burkitt lymphoma is associated with strength and diversity of plasmodium falciparum malaria stage-specific antigen antibody response. *Blood* **2013**, *122*, 629–635.
8. Derkach, A.; Otim, I.; Pfeiffer, R.M.; Onabajo, O.O.; Legason, I.D.; Nabalende, H.; Ogwang, M.D.; Kerchan, P.; Talisuna, A.O.; Ayers, L.W.; et al. Associations between igg reactivity to plasmodium falciparum erythrocyte membrane protein 1 (pfemp1) antigens and burkitt lymphoma in ghana and uganda case-control studies. *EBioMedicine* **2019**, *39*, 358–368.

9. Legason, I.D.; Pfeiffer, R.M.; Udquim, K.I.; Bergen, A.W.; Gouveia, M.H.; Kirimunda, S.; Otim, I.; Karlins, E.; Kerchan, P.; Nabalende, H.; et al. Evaluating the causal link between malaria infection and endemic burkitt lymphoma in northern uganda: A mendelian randomization study. *EBioMedicine* **2017**, *25*, 58–65.
10. Williams, A.O. Haemoglobin genotypes, abo blood groups, and burkitt's tumour. *J. Med. Genet.* **1966**, *3*, 177–179.
11. Robbiani, D.F.; Deroubaix, S.; Feldhahn, N.; Oliveira, T.Y.; Callen, E.; Wang, Q.; Jankovic, M.; Silva, I.T.; Rommel, P.C.; Bosque, D.; et al. Plasmodium infection promotes genomic instability and aid-dependent b cell lymphoma. *Cell* **2015**, *162*, 727–737.
12. Torgbor, C.; Awuah, P.; Deitsch, K.; Kalantari, P.; Duca, K.A.; Thorley-Lawson, D.A. A multifactorial role for p. Falciparum malaria in endemic burkitt's lymphoma pathogenesis. *PLoS Pathog.* **2014**, *10*, e1004170.
13. Proceedings of the IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Epstein-Barr Virus and Kaposi's Sarcoma Herpesvirus/Human Herpesvirus 8. Lyon, France, 17–24 June 1997. *IARC Monogr. Eval. Carcinog. Risks Hum.* **1997**, *70*, 1–492.
14. Emmanuel, B.; Kawira, E.; Ogwang, M.D.; Wabinga, H.; Magatti, J.; Nkrumah, F.; Neequaye, J.; Bhatia, K.; Brubaker, G.; Biggar, R.J.; et al. African burkitt lymphoma: Age-specific risk and correlations with malaria biomarkers. *Am. J. Trop. Med. Hyg.* **2011**, *84*, 397–401.
15. Daniels, R.; Volkman, S.K.; Milner, D.A.; Mahesh, N.; Neafsey, D.E.; Park, D.J.; Rosen, D.; Angelino, E.; Sabeti, P.C.; Wirth, D.F.; et al. A general snp-based molecular barcode for plasmodium falciparum identification and tracking. *Malar. J.* **2008**, *7*, 223.
16. Johnston, W.T.; Mutalima, N.; Sun, D.; Emmanuel, B.; Bhatia, K.; Aka, P.; Wu, X.; Borgstein, E.; Liomba, G.N.; Kamiza, S.; et al. Relationship between plasmodium falciparum malaria prevalence, genetic diversity and endemic burkitt lymphoma in malawi. *Sci. Rep.* **2014**, *4*, 3741.
17. Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; Shallom, S.; Mason, T.; Yu, K.; Fujii, C.; et al. Chromosome 2 sequence of the human malaria parasite plasmodium falciparum. *Science* **1998**, *282*, 1126–1132.
18. Okech, B.; Mujuzi, G.; Ogwal, A.; Shirai, H.; Horii, T.; Egwang, T.G. High titers of igg antibodies against plasmodium falciparum serine repeat antigen 5 (sera5) are associated with protection against severe malaria in ugandan children. *Am. J. Trop. Med. Hyg.* **2006**, *74*, 191–197.
19. Palacpac, N.M.; Ntege, E.; Yeka, A.; Balikagala, B.; Suzuki, N.; Shirai, H.; Yagi, M.; Ito, K.; Fukushima, W.; Hirota, Y.; et al. Phase 1b randomized trial and follow-up study in uganda of the blood-stage malaria vaccine candidate bk-se36. *PLoS ONE* **2013**, *8*, e64073.
20. Palacpac, N.M.; Arisue, N.; Tougan, T.; Ishii, K.J.; Horii, T. Plasmodium falciparum serine repeat antigen 5 (se36) as a malaria vaccine candidate. *Vaccine* **2011**, *29*, 5837–5845.
21. Arisue, N.; Palacpac, N.M.Q.; Tougan, T.; Horii, T. Characteristic features of the sera multigene family in the malaria parasite. *Parasites Vectors* **2020**, *13*, 170.
22. Debrabant, A.; Maes, P.; Delplace, P.; Dubremetz, J.F.; Tartar, A.; Camus, D. Intramolecular mapping of plasmodium falciparum p126 proteolytic fragments by n-terminal amino acid sequencing. *Mol. Biochem. Parasitol.* **1992**, *53*, 89–95.
23. Li, J.; Matsuoka, H.; Mitamura, T.; Horii, T. Characterization of proteases involved in the processing of plasmodium falciparum serine repeat antigen (sera). *Mol. Biochem. Parasitol.* **2002**, *120*, 177–186.
24. Yeoh, S.; O'Donnell, R.A.; Koussis, K.; Dluzewski, A.R.; Ansell, K.H.; Osborne, S.A.; Hackett, F.; Withers-Martinez, C.; Mitchell, G.H.; Bannister, L.H.; et al. Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. *Cell* **2007**, *131*, 1072–1083.
25. Stallmach, R.; Kavishwar, M.; Withers-Martinez, C.; Hackett, F.; Collins, C.R.; Howell, S.A.; Yeoh, S.; Knuepfer, E.; Atid, A.J.; Holder, A.A.; et al. Plasmodium falciparum sera5 plays a non-enzymatic role in the malarial asexual blood-stage lifecycle. *Mol. Microbiol.* **2015**, *96*, 368–387.
26. Yagi, M.; Bang, G.; Tougan, T.; Palacpac, N.M.; Arisue, N.; Aoshi, T.; Matsumoto, Y.; Ishii, K.J.; Egwang, T.G.; Druilhe, P.; et al. Protective epitopes of the plasmodium falciparum sera5 malaria vaccine reside in intrinsically unstructured n-terminal repetitive sequences. *PLoS ONE* **2014**, *9*, e98460.
27. Tanabe, K.; Arisue, N.; Palacpac, N.M.; Yagi, M.; Tougan, T.; Honma, H.; Ferreira, M.U.; Farnert, A.; Bjorkman, A.; Kaneko, A.; et al. Geographic differentiation of polymorphism in the plasmodium falciparum malaria vaccine candidate gene sera5. *Vaccine* **2012**, *30*, 1583–1593.
28. Mutalima, N.; Molyneux, E.M.; Johnston, W.T.; Jaffe, H.W.; Kamiza, S.; Borgstein, E.; Mkandawire, N.; Liomba, G.N.; Batumba, M.; Carpenter, L.M.; et al. Impact of infection with human immunodeficiency virus-1 (hiv) on the risk of cancer among children in malawi—Preliminary findings. *Infect. Agent Cancer* **2010**, *5*, 5.
29. Tanabe, K.; Mita, T.; Jombart, T.; Eriksson, A.; Horibe, S.; Palacpac, N.; Ranford-Cartwright, L.; Sawai, H.; Sakihama, N.; Ohmae, H.; et al. Plasmodium falciparum accompanied the human expansion out of africa. *Curr. Biol.* **2010**, *20*, 1283–1289.
30. Librado, P.; Rozas, J. Dnasp v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452.
31. Nei, M.; Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **1986**, *3*, 418–426.
32. Kumar, S.; Stecher, G.; Li, M.; Nkayaz, C.; Tamura, K. Mega x: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549.
33. Wright, S. The interpretation of population structure by f-statistics with special regard to systems of mating. *Evolution* **1965**, *19*, 395–420.

34. Excoffier, L.; Lischer, H.E. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under linux and windows. *Mol. Ecol. Resour.* **2010**, *10*, 564–567.
35. Sumner, K.M.; Freedman, E.; Abel, L.; Obala, A.; Pence, B.W.; Wesolowski, A.; Meshnick, S.R.; Prudhomme-O'Meara, W.; Taylor, S.M. Genotyping cognate plasmodium falciparum in humans and mosquitoes to estimate onward transmission of asymptomatic infections. *Nat. Commun.* **2021**, *12*, 909.
36. Buchwald, A.G.; Sixpence, A.; Chimanya, M.; Damson, M.; Sorkin, J.D.; Wilson, M.L.; Seydel, K.; Hochman, S.; Mathanga, D.P.; Taylor, T.E.; et al. Clinical implications of asymptomatic plasmodium falciparum infections in malawi. *Clin. Infect. Dis.* **2019**, *68*, 106–112.
37. Henning, L.; Schellenberg, D.; Smith, T.; Henning, D.; Alonso, P.; Tanner, M.; Mshinda, H.; Beck, H.P.; Felger, I. A prospective study of plasmodium falciparum multiplicity of infection and morbidity in tanzanian children. *Trans. R. Soc. Trop. Med. Hyg.* **2004**, *98*, 687–694.
38. Eldh, M.; Hammar, U.; Arnot, D.; Beck, H.P.; Garcia, A.; Liljander, A.; Mercereau-Puijalon, O.; Migot-Nabias, F.; Mueller, I.; Ntoumi, F.; et al. Multiplicity of asymptomatic plasmodium falciparum infections and risk of clinical malaria: A systematic review and pooled analysis of individual participant data. *J. Infect. Dis.* **2020**, *221*, 775–785.
39. Langhorne, J.; Ndungu, F.M.; Sponaas, A.M.; Marsh, K. Immunity to malaria: More questions than answers. *Nat. Immunol.* **2008**, *9*, 725–732.
40. Redmond, L.S.; Ogwang, M.D.; Kerchan, P.; Reynolds, S.J.; Tenge, C.N.; Were, P.A.; Kuremu, R.T.; Masalu, N.; Kawira, E.; Otim, I.; et al. Endemic burkitt lymphoma: A complication of asymptomatic malaria in sub-saharan africa based on published literature and primary data from uganda, tanzania, and kenya. *Malar. J.* **2020**, *19*, 239.
41. Smith, T.; Beck, H.P.; Kitua, A.; Mwankusye, S.; Felger, I.; Fraser-Hurt, N.; Irion, A.; Alonso, P.; Teuscher, T.; Tanner, M. Age dependence of the multiplicity of plasmodium falciparum infections and of other malariological indices in an area of high endemicity. *Trans. R. Soc. Trop. Med. Hyg.* **1999**, *93* (Suppl. 1), 15–20.
42. Nkhoma, S.C.; Trevino, S.G.; Gorena, K.M.; Nair, S.; Khoswe, S.; Jett, C.; Garcia, R.; Daniel, B.; Dia, A.; Terlouw, D.J.; et al. Co-transmission of related malaria parasite lineages shapes within-host parasite diversity. *Cell Host Microbe* **2020**, *27*, 93–103.e4.
43. Peprah, S.; Ogwang, M.D.; Kerchan, P.; Reynolds, S.J.; Tenge, C.N.; Were, P.A.; Kuremu, R.T.; Wekesa, W.N.; Sumba, P.O.; Masalu, N.; et al. Risk factors for burkitt lymphoma in east african children and minors: A case-control study in malaria-endemic regions in uganda, tanzania and kenya. *Int. J. Cancer* **2020**, *146*, 953–969.
44. Turner, L.; Lavtsen, T.; Berger, S.S.; Wang, C.W.; Petersen, J.E.; Avril, M.; Brazier, A.J.; Freeth, J.; Jespersen, J.S.; Nielsen, M.A.; et al. Severe malaria is associated with parasite binding to endothelial protein c receptor. *Nature* **2013**, *498*, 502–505.
45. Magallon-Tejada, A.; Machevo, S.; Cistero, P.; Lavtsen, T.; Aide, P.; Rubio, M.; Jimenez, A.; Turner, L.; Valmaseda, A.; Gupta, H.; et al. Cytoadhesion to gc1qr through plasmodium falciparum erythrocyte membrane protein 1 in severe malaria. *PLoS Pathog.* **2016**, *12*, e1006011.
46. Joergensen, L.; Bengtsson, D.C.; Bengtsson, A.; Ronander, E.; Berger, S.S.; Turner, L.; Dalgaard, M.B.; Cham, G.K.; Victor, M.E.; Lavtsen, T.; et al. Surface co-expression of two different pfemp1 antigens on single plasmodium falciparum-infected erythrocytes facilitates binding to icam1 and pecam1. *PLoS Pathog.* **2010**, *6*, e1001083.
47. Tougan, T.; Emdin, J.R.; Takashima, E.; Morita, M.; Shinohara, M.; Shinohara, A.; Tsuboi, T.; Horii, T. Molecular camouflage of plasmodium falciparum merozoites by binding of host vitronectin to p47 fragment of sera5. *Sci. Rep.* **2018**, *8*, 5052.
48. Ziegler, J.L.; Bluming, A.Z.; Templeton, A.C. Burkitt's lymphoma and tropical splenomegaly syndrome. *Lancet* **1971**, *2*, 317.
49. Engwerda, C.R.; Beattie, L.; Amante, F.H. The importance of the spleen in malaria. *Trends Parasitol.* **2005**, *21*, 75–80.
50. de-The, G.; Geser, A.; Day, N.E.; Tukei, P.M.; Williams, E.H.; Beri, D.P.; Smith, P.G.; Dean, A.G.; Bronkamm, G.W.; Feorino, P.; et al. Epidemiological evidence for causal relationship between epstein-barr virus and burkitt's lymphoma from ugandan prospective study. *Nature* **1978**, *274*, 756–761.
51. Nkhoma, S.C.; Banda, R.L.; Khoswe, S.; Dzoole-Mwale, T.J.; Ward, S.A. Intra-host dynamics of co-infecting parasite genotypes in asymptomatic malaria patients. *Infect. Genet. Evol.* **2018**, *65*, 414–424.