bjh research paper

Variation in the Human Leukocyte Antigen system and risk for endemic Burkitt lymphoma in northern Uganda

Samuel Kirimunda,[1] Murielle Verboom,[2] Isaac Otim,[3] Mark Ssennono,[3,4] Ismail D. Legason,[4] Hadijah Nabalende,[3] Martin D. Ogwang,[3] Patrick Kerchan,[4] Tobias Kinyera,[3] Ivan Mwebaza,[1] Moses Joloba,[1] Leona W. Ayers,[5] Steven J. Reynolds,[6] Kishor Bhatia,[7] Olusegun O. Onabajo,[8] Michael Hallensleben,[2] Robert J. Biggar,[7] Ludmila Prokunina-Olsson,[8] James J. Goedert,[7] Rainer Blasczyk[2] and Sam M. Mbulaiteye[7]

[1]College of Health Sciences, Makerere University, Kampala, Uganda, [2]Institute of Transfusion Medicine, Hannover, Germany, [3]EMBLEM Study, St. Mary's Hospital, Lacor, Gulu & African Field Epidemiology Network, [4]EMBLEM Study, Kuluva Hospital, Kuluva, Arua & African Field Epidemiology Network, Kampala, Uganda, [5]Department of Pathology, The Ohio State University, Columbus, OH, USA, [6]Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, [7]Infections and Immunoepidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health and [8]Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

Received 20 October 2019; accepted for publication 4 November 2019

Correspondence: Sam M. Mbulaiteye, Infections and Immunoepidemiology Branch, Division of Cancer Epidemiology & Genetics, National Cancer Institute, NIH, HHS, 9609 Medical Center Dr, Rm. 6E-118, MSC 3330, Bethesda, MD, 20892, USA.
E-mail: mbulaits@mail.nih.gov

Summary

Endemic Burkitt lymphoma (eBL) is an aggressive childhood B-cell lymphoma associated with *Plasmodium falciparum* (*Pf*) malaria and Epstein–Barr virus (EBV) infections. Variation in the Human Leukocyte Antigen (HLA) system is suspected to play a role, but assessments using less accurate serology-based HLA typing techniques in small studies yielded conflicting results. We studied 200 eBL cases and 400 controls aged 0–15 years enrolled in northern Uganda and typed by accurate high-resolution HLA sequencing methods. HLA results were analyzed at one- or two-field resolution. Odds ratios and 95% confidence intervals (aOR, 95% CI) for eBL risk associated with common HLA alleles *versus* alleles that were rare (<1%) or differed by <2% between the cases and controls as the reference category, were estimated using multiple logistic regression adjusting for age, sex, microgeography, region, malaria positivity and treatment history, and genetic variants associated with eBL. Compared to the controls, eBL cases had a lower frequency of *HLA-A*02* (aOR = 0·59, 95% CI 0·38–0·91), *HLA-B*41* (aOR = 0·36, 95% CI 0·13–1·00), and *HLA-B*58* alleles (aOR = 0·59, 95% CI 0·36–0·97). eBL cases had a lower frequency of *HLA-DPB1* homozygosity (aOR = 0·57, 95% CI 0·40–0·82) but a higher frequency of *HLA-DQA1* homozygosity (aOR = 2·19, 95% CI 1·42–3·37). Our results suggest that variation in HLA may be associated with eBL risk.

[Correction added on 28 February 2020, after online publication: In the Summary, Rare alleles odds ratios has been changed to Odds ratios in this version].

Keywords: Burkitt lymphoma, non-Hodgkin lymphoma, epidemiology, Epstein–Barr virus, *Plasmodium falciparum* malaria, human leukocyte antigen.

First published online 18 February 2020
doi: 10.1111/bjh.16398

Endemic Burkitt lymphoma (eBL) is a rare, aggressive childhood B-cell lymphoma that occurs nearly 50-fold more commonly in Africa than elsewhere (Burkitt, 1958). This higher risk for eBL in Africa is related to the endemicity of infection with *Plasmodium falciparum* (*Pf*) malaria(Hill *et al.*, 1992) and infection with Epstein–Barr virus (EBV) at an early age in malaria-endemic regions (de-The *et al.*, 1978). The association of eBL with these infections (Burnet, 1970) suggested the possibility that variation in the Major Histocompatibility System (MHC), known as the Human Leukocyte Antigen (HLA) in humans, may play a role in the aetiology of eBL.

However, subsequent studies assessing the role of HLA variation in eBL risk (Bodmer *et al.*, 1975b) were inconclusive (Bodmer *et al.*, 1975a; Jones *et al.*, 1980; Hall *et al.*, 1982; Jones *et al.*, 1985). These studies were limited by using less accurate serological methods capable of only low (one-field) resolution HLA typing (Terasaki & McClelland, 1964) and being of very small sample sizes. Three of the studies found no associations with *HLA-A* types investigated (Bodmer *et al.*, 1975a), one study found an association between a case group consisting of eBL and nasopharyngeal carcinoma (both EBV-associated) with *HLA-A*29* when compared to local healthy subjects (Hall *et al.*, 1982), while two studies found an association with *HLA-DR7* (Jones *et al.*, 1980; Jones *et al.*, 1985).

The characterization of HLA over the past five decades has revealed its complexity and the need for accurate typing to get reliable results (Woszczek *et al.*, 1997). HLA is a multigene family system located on the short arm of chromosome 6 that plays a crucial role in the adaptive immune system (Klein & Sato, 1998). HLA genes code for proteins that present self- and non-self- antigenic peptides to receptors on other immune cells and are categorized into two main classes: Class I (*HLA-A*, *-B*, *-C*) and Class II (*HLA-DQA1*, *-DQB1*, *-DRB1 and -DPB1*) (Thorsby, 2009). Class I proteins are α polypeptide chains expressed on the surface of nucleated cells. Class II proteins are α and β polypeptide chains expressed on the surface membrane of antigen-presenting cells (APCs), macrophages, dendritic cells, B and activated T cells (Klein & Sato, 1998). The HLA system plays a dual role of guiding the adaptive immune response against a diverse array of pathogen antigens circulating in different populations (Gilbert *et al.*, 1998), including *Pf* malaria and EBV, and suppressing the adaptive immune response against self- or mimicked antigens, which is responsible for autoimmunity (Klein & Sato, 1998).

Studies using accurate, reliable high-resolution HLA typing (Woszczek *et al.*, 1997) have shed light on the relationship between HLA variation and malaria and EBV, the co-factors of eBL. Class I *HLA-B*53* (Hill *et al.*, 1991) and class II *HLA-DRB1*04* (Osafo-Addo *et al.*, 2008) and *HLA-DPB1*17* alleles (May *et al.*, 2001) have been shown to affect a child's risk for developing severe malaria in West African populations. A recent genome-wide association study (GWAS) in Uganda reported a significant association between high antibody levels against EBV nuclear antigen-1 (EBNA1) and genetic variant rs9272371 in the Class II *HLA-DQA1* locus (Sallah *et al.*, 2017). However, their association with eBL has not been assessed.

To investigate these hypotheses, we leveraged data from the Epidemiology of Burkitt lymphoma in East African children and minors (EMBLEM) study (Legason *et al.*, 2017; Peprah *et al.*, 2020) to assess HLA associations with eBL risk in northern Uganda based on high-resolution HLA sequencing.

Materials and methods

Study population and design

The EMBLEM study has been described previously (Legason *et al.*, 2017). Briefly, the eBL cases were enrolled at St. Mary's Hospital, Lacor, in Gulu district and at Kuluva Hospital in Arua district. These hospitals have the capacity to diagnose and treat eBL and are responsible for virtually all the cases in North-Central and Northwest Uganda (Peprah *et al.*, 2020). The controls were enrolled from 100 local-area villages randomly selected from the region as previously described (Maziarz *et al.*, 2017). Participant eligibility was restricted to children aged 1–15 years who had been residing in the geographically defined study area for at least four months (Peprah *et al.*, 2020). To capture microgeographical factors that may correlate with unmeasured factors that influence malaria transmission risk (Ogwang *et al.*, 2008), villages in the study area were categorized as rural or urban and as near or far from surface water (defined as a swamp, river, or lake). Participant information, including age, sex, and inpatient and outpatient malaria treatment, was recorded using interviewer-administered questionnaires. Venous blood samples were collected before treatment in the cases. Clinical samples were immediately tested for malaria by light microscopy and rapid diagnostic tests (RDT) for malaria; research samples were separated into aliquots of blood fractions and frozen at −80°C (Maziarz *et al.*, 2017). The current study was conducted among 200 eBL cases and a subset of 400 random controls of 600 who previously were studied for association with malaria-resistance genes (*HBB* rs334, *IL10* rs1800896, *IL1A* rs2856838, and *SEMA3C* rs4461841) (Legason *et al.*, 2017), providing a 2:1 match to cases.

Ethical approval and consent

The EMBLEM study was approved by the Uganda Virus Research Institute Research and Ethics Committee, Uganda National Council for Science and Technology (H816), and the National Cancer Institute Special Studies Institutional Review Board (10-C-N133). Guardians, typically a parent, provided written informed consent. Participants aged ≥7 years old gave their assent.

HLA typing

DNA extracted from buffy coat (or saliva, when buffy coat was not available) (Legason *et al.*, 2017) was used for high-

resolution targeted next-generation sequencing (NGS) of the HLA region (Fig 1). Multiplex polymerase chain reactions (PCRs) were performed in eight sets to partially cover 11 HLA genes: exons 1–4 of all HLA Class I genes, exons 6–7 of *HLA-C* and exons 2–3 of HLA Class II genes (Fig 2) (PROTRANS N5 HLA-NGS panel, Protrans Medical Diagnostics, Hockenheim, Germany). These exons cover the hypervariable region of the peptide-binding groove of the corresponding proteins and are sufficient for high-resolution HLA typing. The amplicons were barcoded with index adaptors (Illumina, San Diego, CA, USA) for pooling up to 96 samples into multiplex PCRs. The PCR pools were purified using Clean-NGS Beads (GC Biotech, Waddinxveen, The Netherlands) to generate one indexed library, which was quantified (Quant-IT PicoGreen ds DNA Assay Kit, Molecular Probes Inc., Eugene, OR, USA) and adjusted to 4 nM. The library was then denatured with NaOH, diluted to a final concentration of 8 pM for optimal cluster density, and 600 µl was loaded into the MiSeq reagent cartridge (v2 500 cycle kit, Illumina) to generate clusters for 2 × 250 bp paired-end reads on MiSeq sequencer (Illumina). For allele assignment, we used hlaSYSTEM software (AVALAS GmbH, Nordhausen, Germany) based on the most recent HLA database (Holdsworth *et al.*, 2009). Although typing was done at high-level resolution and HLA alleles could be scored up to five-field resolution, scoring at the one-field resolution (such as A*01, etc. serological antigenicity) was deemed adequate for our exploratory analysis. For homozygosity analyses, we used two-field resolution scoring (such as A*02:10, etc.), which is sufficient to indicate different proteins, despite similar serological antigenicity (such as *HLA-A*02:01* and *HLA-A*02:02*).

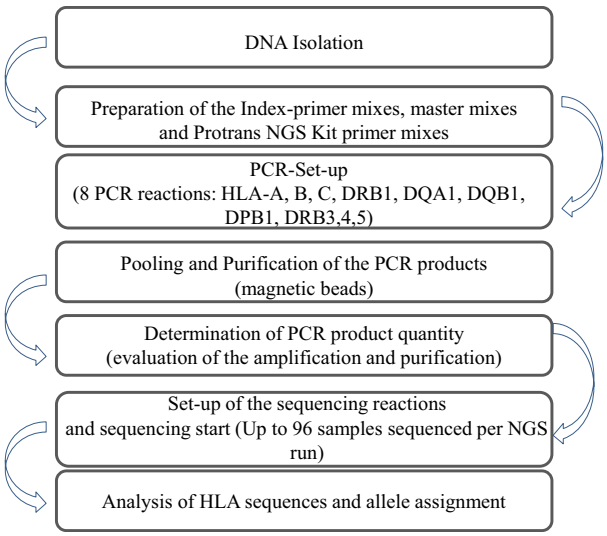

Fig 1. Workflow of the procedures followed to sequence HLA genes. [Colour figure can be viewed at wileyonlinelibrary.com]

Statistical analysis

Demographic characteristics of the eBL cases and controls were compared using frequency tables. HLA allele frequencies were tabulated for eBL cases and controls. We combined rare alleles (observed in <1% of controls) and those that differed by <2% between cases and controls into one *post-hoc* group referred to as 'other alleles'; this group was used as the referent category. We calculated odds ratios and 95% confidence intervals (OR, 95% CI) for the association of eBL risk with particular alleles using multiple logistic regression. HLA alleles previously associated with *Pf* malaria [*HLA-B*53* (Hill *et al.*, 1992), *HLA-DRB1*04* (Osafo-Addo *et al.*, 2008), *HLA-DPB1*17* (May *et al.*, 2001)], with EBV infection (Sallah *et al.*, 2017), or EBV-related Hodgkin lymphoma [*HLA-A*02* and *HLA-DQA1* (Sallah *et al.*, 2017)], or eBL [*HLA-DR*07* (Jones *et al.*, 1980)] were assessed considering these prior associations. We also tested for the association between eBL and HLA homozygosity, defined based on the two-field resolution. This analysis evaluates the hypothesis that individuals who have more diverse HLA proteins (produced by heterozygous genes) may recognize and react to a wider repertoire of epitopes than individuals whose HLA proteins are produced by homozygous genes, which may translate into better pathogen control.

To gain insights whether binding to killer cell immunoglobulin-like receptors (KIR) on natural killer cells may influence eBL risk, we also conducted exploratory analyses using Pearson's chi-squared goodness of fit test to assess associations with *HLA-C* allotypes corresponding to two major KIR epitopes defined on the basis of a dimorphism at position 80 of the α1 domain: C1 (*HLA-C*01/*03/*07/*08/*12/*14/*16*) and C2 (*HLA-C*02/*04/*05/*06/*15/*17/*18*) (Faridi & Agrawal, 2011). Group C1 (HLA-Casn80) are ligands for inhibitory KIR2DL2/3 and activating KIR2DS2, whereas group C2 (HLA-Clys80) are ligands for inhibitory KIR2DL1 and activating KIR2DS1(Faridi & Agrawal, 2011). Because the associations with HLA may be

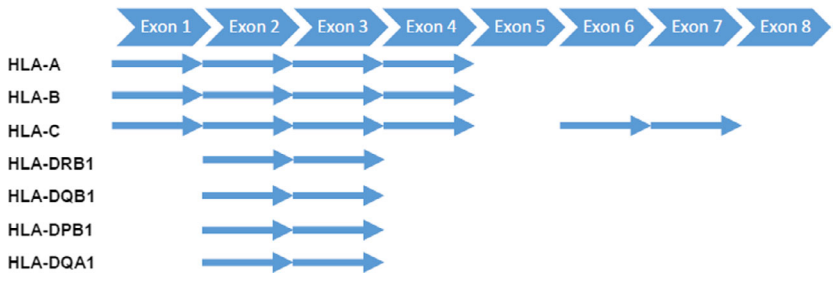

Fig 2. Schematic figure showing exons in HLA genes. Arrows show the exons targetted for sequencing in the study. [Colour figure can be viewed at wileyonlinelibrary.com]

mediated, in part, by *Pf* malaria parasite killing, we repeated the analyses restricted to controls, with malaria status as the outcome.

We adjusted the associations for demographics (sex and age), and microgeography (rural *versus* urban, near *versus* far from water) (Maziarz *et al.*, 2017). We further adjusted for the region (North-Central *versus* Northwest) because it is correlated with population genetic structure in EMBLEM (Gouveia *et al.*, 2019). We further adjusted for malaria status (positive or negative by RDT), inpatient and outpatient malaria treatment history, and the non-HLA malaria-resistance genetic polymorphisms previously associated with eBL (*HBB* rs334, *IL10* rs1800896, *IL1A* rs2856838 and *SEMA3C* rs4461841)(Legason *et al.*, 2017). We used directed acyclic graphs (DAGs) (Greenland *et al.*, 1999) to assess whether it was appropriate to include or exclude variables in the models. Using DAGs, we considered 'lifetime malaria burden' as an intermediate variable between HLA variants (exposure) and eBL (outcome). The exclusion of measures of malaria did not alter the adjusted estimates, whereas the inclusion of malaria-resistance genetic polymorphisms previously associated with eBL(Legason *et al.*, 2017) strengthened the HLA/eBL associations so these variables were included in the final models.

Results

The eBL cases were slightly older than controls but did not differ by sex or the rural/urban location of their village (Table I). However, eBL cases were more likely than controls to reside in villages near *versus* far from surface water ($P = 0{\cdot}004$), in the Northwest *versus* North-Central region ($P = 0{\cdot}002$), and to have mild anaemia($P < 0{\cdot}0001$). eBL cases were more likely than controls to report inpatient and outpatient malaria treatment more than 12 months before enrollment ($P = 0{\cdot}006$ and $P < 0{\cdot}0001$, respectively), but to have a lower frequency of malaria infection detected at enrollment than controls (36·7% vs. 56·8%, $P < 0{\cdot}0001$).

All HLA amplicons were successfully generated and sequenced at high coverage of 213–1539 × coverage per region (Table SI). The average heterozygous ratio per region was close to the expected 50% (53–74%, Table SII and Table SIII), suggesting that HLA sequencing was non-preferential across all exons. Due to DNA depletion, the DRB3/4/5 results were not available for all subjects, so these incomplete data are not reported in this study.

The frequencies of Class I and Class II HLA alleles in eBL cases and controls are shown in Table II. The four most frequent HLA Class I alleles in the control samples were: *HLA-A*02, -A*30, -A*23*, and *-A*29; -B*15, -B*58, -B*53*, and *-B*08; -C*07, -C*04, -C*06*, and -C**03*. The four most frequent Class II HLA alleles were: *DPB1*04*, *DRB1*15*, *DPB1*01*, *DPB1*03* and *DPB1*02*; *DQA1*01*, *DQA1*05*, *DQA1*02*, and *DQA1*03*; *HLA-DQB1*06*, *DQB1*05*, *DQB1*03*, and *DQB1*02*; and *HLA-DRB1*13*, *DRB1*15*, *DRB1*01*, and *DRB1*03*. The distribution of these alleles in the controls was comparable to the distribution reported in two other populations from Uganda (Guech-Ongey *et al.*, 2010) and Kenya (Peterson *et al.*, 2013).

No differences were observed between eBL cases and controls for *HLA-B*53* (7·2% vs. 7·0%), for *HLA-DPB1* homozygosity (3·8% vs. 4·2%), and for *HLA-DRB1*04* (3·1% vs. 2·4%) and *HLA-DR7* (7·1% vs. 6·4%) (Table II). No differences between eBL cases and controls were observed in HLA-C

Table I. Demographic and clinical characteristics of the study population.

Variables	Cases, *n* (%)	Controls, *n* (%)	*P* value*
Characteristics			
Age, years			0·045
0–4	24 (13·3)	83 (21·3)	
5–9	93 (51·7)	197 (50·6)	
10–17	63 (35·0)	109 (28·0)	
Mean age, (standard deviation)	7·8 (3·3)	7·4 (3·4)	0·185
Sex			0·269
Male	109 (61·2)	219 (56·3)	
Female	69 (38·8)	170 (43·7)	
Rural/urban*			0·773
Rural	108 (61·7)	245 (63·0)	
Urban	67 (38·3)	144 (37·0)	
Proximity to surface water†			0·004
Near	150 (85·7)	291 (74·8)	
Far	25 (14·3)	98 (25·2)	
Study region			0·002
North-Central	115 (64·6)	292 (76·8)	
Northwest	63 (35·4)	88 (23·2)	
Anaemia‡			<0·0001
Not anaemic (Hb> 11·6 g/dL)	43 (24·0)	279 (70·9)	
Mild anaemia (Hb ≤ 11·6 g/dL)	136 (76·0)	113 (29·1)	
Inpatient malaria			0·006
No	120 (67·4)	268 (69·6)	
Past 12 months	13 (7·3)	55 (14·3)	
More than 12 months	45 (25·3)	62 (16·1)	
Outpatient malaria			<0·0001
No	58 (32·6)	82 (21·3)	
Past 12 months	79 (44·4)	264 (68·6)	
More than 12 months	41 (23·0)	39 (10·1)	
Malaria infection§			<0·0001
Negative	114 (63·3)	168 (43·2)	
Positive	66 (36·7)	221 (56·8)	

**P* value calculated using Pearson's chi-squared test, except for the comparison of mean values, for which a *t*-test was used.

†Urban/rural status and proximity to water were variables used to adjust for microgeographical factors associated with environmental risk for malaria as described in Maziarz *et al.*

‡Mild anaemia is used as a proxy for exposure to heavy malaria in children living in malaria-endemic regions.

§Malaria infection status was determined by thick film microscopy and rapid diagnostic test results as described in Maziarz *et al.* A positive test means being positive on either test or both. The children with positive malaria tests were deemed to have asymptomatic infection because <2% reported a fever.

Table II. Frequency of HLA Class I and Class II alleles in children with and without Burkitt lymphoma in the EMBLEM study

Variables	Cases, *n* (%)†	Controls, *n* (%)†
Class I loci		
HLA antigens		
HLA-A		
*01	22 (6·5)	42 (5·7)
*02	44 (12·9)	132 (18·0)
*03	14 (4·1)	42 (5·7)
*23	29 (8·5)	76 (10·4)
*24	6 (1·8)	4 (0·5)
*26	14 (4·1)	14 (1·9)
*29	26 (7·6)	56 (7·6)
*30	51 (15·0)	115 (15·7)
*31	28 (8·2)	35 (4·8)
*32	13 (3·8)	26 (3·5)
*33	11 (3·2)	30 (4·1)
*34	10 (2·9)	29 (4·0)
*36	6 (1·8)	13 (1·8)
*43	0 (0·0)	1 (0·1)
*66	13 (3·8)	16 (2·2)
*68	37 (10·9)	55 (7·5)
*74	14 (4·1)	47 (6·4)
*80	2 (0·6)	0 (0·0)
NA	6	27
HLA-B		
*07	7 (2·1)	17 (2·3)
*08	11 (3·3)	48 (6·5)
*13	13 (3·9)	29 (3·9)
*14	9 (2·7)	20 (2·7)
*15	48 (14·5)	129 (17·3)
*18	18 (5·4)	32 (4·3)
*27	4 (1·2)	5 (0·7)
*35	11 (3·1)	39 (5·2)
*37	0 (0·0)	3 (0·4)
*39	13 (3·9)	22 (3·0)
*40	4 (1·2)	10 (1·3)
*41	6 (1·8)	32 (4·3)
*42	19 (5·7)	41 (5·5)
*44	9 (2·7)	9 (1·2)
*45	25 (7·5)	38 (5·1)
*47	22 (6·6)	30 (4·0)
*49	7 (2·1)	12 (1·6)
*50	1 (0·3)	1 (0·1)
*51	13 (3·9)	12 (1·6)
*53	24 (7·2)	52 (7·0)
*57	7 (2·1)	15 (2·0)
*58	32 (9·6)	91 (12·2)
*73	0 (0·0)	1 (0·1)
*78	0 (0·0)	1 (0·1)
*81	17 (5·1)	38 (5·1)
*82	12 (3·6)	17 (2·3)
NA	6	12
HLA-C KIR epitopes‡		
*02 C2	14 (4·2)	33 (4·6)
*03 C1	23 (6·9)	77 (10·7)
*04 C2	57 (17·1)	138 (19·2)
*05 C2	2 (0·6)	5 (0·7)
*06 C2	44 (13·2)	82 (11·4)
*07 C1	87 (26·1)	215 (29·9)
*08 C1	19 (5·7)	39 (5·4)
*12 C1	11 (3·3)	15 (2·1)
*14 C1	9 (2·7)	12 (1·7)
*15 C2	6 (1·8)	5 (0·7)
*16 C1	23 (6·9)	35 (4·9)
*17 C2	22 (6·6)	43 (6·0)
*18 C2	16 (4·8)	21 (2·9)
NA	9	40
HLA Class II		
HLA-DPB1		
*01	48 (16·4)	149 (21·3)
*02	30 (10·3)	62 (8·9)
*03	32 (11)	72 (10·3)
*04	108 (37)	224 (32·1)
*09	0 (0·0)	1 (0·1)
*11	4 (1·4)	15 (2·1)
*13	25 (8·6)	59 (8·5)
*14	0 (0·0)	2 (0·3)
*15	1 (10·3)	3 (0·4)
*17	11 (3·8)	29 (4·2)
*18	8 (2·7)	12 (1·7)
*19	1 (0·3)	2 (0·3)
*26	1 (0·3))	2 (0·3)
*30	3 (1·0)	4 (0·6)
*34	1 (0·3))	2 (0·3)
*39	3 (1·0)	10 (1·4))
*40	2 (0·7)	6 (0·9)
*49	5 (1·7)	13 (1·9)
*55	0 (0·0)	7 (1·0)
*61	1 (0·3))	1 (0·1)
*80	0 (0·0)	1 (0·1)
*90	1 (0·3)	0 (0·0)
*106	0 (0·0)	2 (0·3)
*131	0 (0·0)	5 (0·7)
*133	3 (1·0)	4 (0·6)
*162	0 (0·0)	3 (0·4)
*333	0 (0·0)	1 (0·1)
*348	1 (0·3)	0 (0·0)
*370	1 (0·3)	0 (0·0)
*412	1 (0·3)	0 (0·0)
*422	1 (0·3)	7 (1·0)
NA	50	64
HLA-DQA1		
*01	106 (47·3)	290 (55·1)
*02	16 (7·1)	38 (7·2)
*03	14 (6·2)	38 (7·2)
*04	19 (8·5)	36 (6·8)
*05	69 (30·8)	124 (23·6)
NA	134	250
HLA-DQB1		
*02	40 (12·8)	92 (13·8)
*03	93 (29·7)	151 (22·6)
*04	24 (7·67)	33 (4·9)
*05	81 (25·9)	185 (27·7)
*06	75 (24·0)	207 (31·0)

Table II. (*Continued*)

Variables	Cases, *n* (%)†	Controls, *n* (%)†
NA	27	92
HLA- DRB1		
*01	36 (11·1)	61 (8·6)
*03	28 (8·7)	60 (8·5)
*04	10 (3·1)	17 (2·4)
*07	23 (7·1)	45 (6·4)
*08	16 (5·0)	20 (2·8)
*09	5 (1·5)	16 (2·3)
*10	20 (6·2)	49 (6·9)
*11	107 (33·1)	233 (33·0)
*12	3 (0·9)	11 (1·6)
*13	39 (12·1)	92 (13·0)
*14	1 (0·3)	7 (1·0)
*15	35 (10·8)	90 (12·7)
*16	0 (0·0)	5 (0·7)
NA	21	52

†HLA alleles are defined using information from the first field of sequence results (e.g. A*01 or A*02), which specify the HLA allele with similar serological activity (see *Materials and methods*). The frequencies in the table reflect the number of alleles which is twice the number of the subjects studied.

‡When considering HLA-C molecules as ligands for killer cell immunoglobulin-like receptors (KIR) receptors on natural killer cells, all HLA-C allotypes can be grouped into two major KIR epitopes: C1 (HLA-C*01/*03/*07/*08/*12/*14/*16) and C2 (HLA-C*02/*04/*05/*06/*15/*17/*18), on the basis of a dimorphism at position 80 of the α1 domain. Group C1 (HLA-Casn80) are ligands for inhibitory KIR2DL2/3 and activating KIR2DS2, and group C2 (HLA-Clys80) are ligands for inhibitory KIR2DL1 and activating KIR2DS1.

allotypes corresponding to the two major KIR epitopes (HLA-C2: 48·3% vs. 45·4%).

HLA alleles were not associated with malaria positivity in the controls (data not shown). Table III shows the HLA allele frequencies and the crude and adjusted ORs and 95% CIs of association with eBL from models mutually adjusted for all significant variables and *a-priori*-selected variables. The adjusted odds of eBL were decreased in those with *HLA-A*02* (aOR = 0·59, 95% CI 0·38–0·91), *HLA-A*74* (aOR = 0·52, 95% CI 0·26–1·01), *HLA-B*41* (aOR = 0·36, 95% CI 0·13–1·00), and *HLA-B*58* (aOR = 0·59, 95% CI 0·36–0·97). When homozygosity was considered, the adjusted odds of eBL were decreased in those with *HLA-DPB1* homozygosity (aOR = 0·57, 95% CI 0·40–0·82) and were elevated in those with *HLA-DQA1* homozygosity (aOR = 2·19, 95% CI 1·42–3·37).

Discussion

Leveraging data from the EMBLEM study (Legason *et al.*, 2017), we present novel results about associations between eBL and HLA variation in northern Uganda. We observed significant decreased odds of eBL with carriage of *HLA-A*02*, *-B*41*, and *-B*58*, and *HLA-DPB1* homozygosity, and increased odds of eBL with *HLA-DQA1* homozygosity. These results support the hypothesis that variation in the HLA may influence eBL risk.

Our findings of elevated odds of eBL in those with *HLA-DQA1* homozygosity and *HLA-A*02* antigenicity may be clues about HLA's role in the ability to control for EBV infection (de-The *et al.*, 1978). A recent GWAS conducted in Uganda reported significant association between genetic variant rs9272371 in the Class II *HLA-DQA1* region with high antibody titres against EBV EBNA1 protein (Sallah *et al.*, 2017). Carriage of *HLA-A*02* is associated with strong immunogenic responses to EBV-derived peptides (Bollard *et al.*, 2004), with strong EBV-directed $CD8^+$ cytotoxicity (Murray *et al.*, 1992), and with decreased odds of EBV-positive Hodgkin lymphoma risk (Niens *et al.*, 2007). Our findings of lower frequency of *HLA-A*02* in eBL cases (12·9%) than in controls (18·0%) are consistent with the notion that eBL cases poorly control EBV infection (Coghill *et al.*, 2020). Together, our results that *HLA-DQA1* locus homozygosity and carriage of *HLA-A*02* allele are associated with eBL risk warrant further follow-up.

Our findings that eBL risk may be associated with *HLA-B*41*, *HLA-B*58*, and *HLA-DPB1* homozygosity are new and unexplained. We suggest that they may point to a role of autoimmunity and/or allergy in eBL. *HLA-B*41* has been linked with Henoch–Schönlein purpura among Spaniards (Lopez-Mejias *et al.*, 2015); *HLA-B*58* has been linked with severe cutaneous adverse reactions during treatment with allopurinol, particularly in Asians (Hung *et al.*, 2005); *HLA-DPB1* variants (homozygosity found in 50% of our controls) have been associated with myasthenia gravis (Horiki *et al.*, 1994), granulomatosis with polyangiitis, immune reactions to beryllium (Potolicchio *et al.*, 1997), and antigenic response after transplantation (Shaw *et al.*, 2007). A role of autoimmunity and/or allergy in BL was suggested by findings that the odds of sporadic BL under age 50 years were decreased in those with a history of allergy or asthma in the International Lymphoma consortium (Mbulaiteye *et al.*, 2014). It is also suggested by molecular data showing that BL tumours disproportionately use a small set of immunoglobulin heavy-chain variable (IGHV) gene segments involved in autoreactivity (Bhat *et al.*, 1993), such as *IGHV4-34* (Lombardo *et al.*, 2017; Grande *et al.*, 2019), and that some BL tumours secrete proteins against self-antigens (Ng *et al.*, 1994; Riboldi *et al.*, 1994).

We did not find any association between *HLA-B*53* (Hill *et al.*, 1991), *HLA-DRB1*04* and *HLA-DPB1*17* (Osafo-Addo *et al.*, 2008) and *HLA-DR7* (Jones *et al.*, 1980; Jones *et al.*, 1985), previously associated with decreased risk for severe malaria and eBL in West Africa. These null results were surprising at first but not unexpected given that the malaria/HLA associations discovered in West Africa have not been replicated in studies conducted in East Africa (Hill, 1998). This non-replication may be explained by geographical differences in HLA allele frequency and its impact on power to replicate associations. The frequencies of *HLA-B53* and *HLA-DR7* are lower in East Africa (Guech-Ongey *et al.*, 2010; Peterson *et al.*, 2013) than in West African populations (Jones *et al.*, 1985; Hill *et al.*, 1991). The non-replication may be due to non-HLA genetic

Table III. Frequency of HLA Class I and Class II loci alleles in children with and without Burkitt lymphoma in the EMBLEM study.

Variables	Cases, *n* (col. %)	Controls, *n* (col. %)	Crude OR (95% CI)	Adjusted OR (95% CI)*
Class I Loci				
HLA-A				
Other alleles†	240 (69·4)	505 (66·5%)	Ref.	Ref.
*02	44 (12·7)	132 (17·4)	0·70 (0·48–1·01)	0.59 (0·38–0·91)
*26	14 (4·1)	14 (1·8)	2·10 (0·99–4·48)	2·10 (0·90–4·87)
*31	28 (8·1)	35 (4·6)	1·68 (1·00–2·83)	1·62 (0·88–2·98)
*74	14 (4·1)	47 (6·2)	0·63 (0·34–1·16)	0·52 (0·26–1·01)
NA	6 (1·7)	27 (3·6)	0·47 (0·19–1·15)	0·34 (0·13–0·94)
P value			0·0034	0·0009
HLA-A homozygosity‡				
No	222 (66·8)	472 (66·9)	Ref.	–
Yes	110 (33·1)	234 (33·1)	1·00 (0·76–1·31)	–
LR X^2 value (*P* value)			0·0 (1·0)	–
HLA-B				
Other alleles	222 (65·7)	441 (58·3)	Ref.	Ref.
*15	48 (14·2)	129 (17·1)	0·74 (0·51–1·07)	0·70 (0·46–1·06)
*35	11 (3·3)	39 (5·2)	0·56 (0·28–1·12)	0·53 (0·25–1·14)
*41	6 (1·8)	32 (4·2)	0·37 (0·15–0·90)	0·36 (0·13–1·00)
*51	13 (3·9)	12 (1·6)	2·15 (0·97–4·79)	1·34 (0·55–3·28)
*58	32 (9·5)	91 (12·0)	0·70 (0·45–1·08)	0·59 (0·36–0·97)
NA	6 (1·8)	12 (1·6)	0·99 (0·37– 2·68)	0·83 (0·27– 2·60)
P value			0·0133	0·0509
HLA-B homozygosity‡				
No	226 (69·3)	498 (68·0)	Ref.	–
Yes	100 (30·7)	234 (32·0)	0·94 (0·71–1·25)	–
P value			0·676	–
HLA-C KIR epitopes§				
Other alleles	201 (58·8)	402 (52·9)	Ref.	Ref.
*03 C1	23 (6·7)	77 (10·1)	0·60 (0·36–0·98)	0·74 (0·43–1·28)
*07 C1	87 (25·4)	215 (28·3)	0·81 (0·60–1·09)	0·77 (0·55–1·09)
*15 C2	6 (1·7)	5 (0·7)	2·40 (0·72–7·96)	2·22 (0·59–8·37)
*18 C2	16 (4·7)	21 (2·8)	1·52 (0·78– 2·98)	1·20 (0·55– 2·58)
NA	9 (2·6)	40 (5·3)	0·45 (0·21–0·95)	0·35 (0·15–0·82)
P value			0·013	0·0546
HLA-C homozygosity‡				
No	240 (74·1)	486 (71·7)	Ref.	–
Yes	84 (25·9)	192 (28·3)	0·89 (0·66–1·19)	–
P value			0·426	–
HLA Class II				
HLA-DPB1				
Other alleles	74 (21·6)	191 (25·1)	Ref.	Ref.
*01	48 (14·0)	149 (19·6)	0·83 (0·54–1·27)	0·89 (0·55–1·43)
*02	30 (8·8)	62 (8·1)	1·25 (0·75–2·08)	1·30 (0·72–2·36)
*03	32 (9·4)	72 (9·5)	1·14 (0·70–1·88)	1·30 (0·74–2·31)
*04	108 (31·6)	224 (29·4)	1·24 (0·87–1·77)	1·19 (0·80–1·79)
NA	50 (14·6)	64 (8·4)	2·02 (1·28–3·18)	1·79 (1·07–3·03)
P value			0·014	0·192
HLA- DPB1 homozygosity‡				
No	162 (63·8)	314 (49·7)	Ref.	Ref.
Yes	92 (36·2)	318 (50·3)	0·56 (0·42–0·76)	0·57 (0·40–0·82)
P value			0·0001	0·0018
HLA-DQA1				
Other alleles	49 (13·7)	112 (14·4)	Ref.	Ref.
*01	106 (29·6)	290 (37·4)	0·84 (0·56–1·25)	0·75 (0·47–1·20)
*05	69 (19·3)	124 (16·0)	1·27 (0·81–1·99)	1·17 (0·70–1·95)
NA	134 (37·4)	250 (32·2)	1·22 (0·82–1·82)	1·31 (0·83–2·08)
P value			0·0479	0·0176

Table III. (*Continued*)

Variables	Cases, *n* (col. %)	Controls, *n* (col. %)	Crude OR (95% CI)	Adjusted OR (95% CI)*
HLA- DQA1 homozygosity‡				
No	98 (51·0)	266 (63·3)	Ref.	Ref.
Yes	94 (49·0)	154 (36·7)	1·66 (1·17–2·34)	2·19 (1·42–3·37)
P value			0·0042	0·0003
HLA- DQB1				
Other alleles	145 (42·6)	310 (40·8)	Ref.	Ref.
*03	93 (27·4)	151 (19·9)	1·32 (0·95–1·82)	1·23 (0·85–1·78)
*06	75 (22·1)	207 (27·2)	0·77 (0·56–1·08)	0·81 (0·55–1·18)
NA	27 (7·9)	92 (12·1)	0·63 (0·39–1·01)	0·68 (0·40–1·15)
P value			0·0061	0·1093
HLA- DQB1 homozygosity‡				
No	212 (74·1)	422 (73·3)	Ref.	–
Yes	74 (25·9)	154 (26·7)	0·96 (0·69–1·32)	–
P value			0·787	–
HLA- DRB1				
Other alleles	271 (78·8)	625 (82·5)	Ref.	Ref.
*01	36 (10·5)	61 (8·1)	1·36 (0·88–2·10)	1·44 (0·88–2·35)
*08	16 (4·6)	20 (2·6)	1·84 (0·94–3·61)	1·90 (0·87–4·13)
NA	21 (6·1)	52 (6·9)	0·93 (0·55–1·58)	1·17 (0·64–2·14)
P value			0·180	0·2131
HLA- DRB1 homozygosity‡				
No	214 (70·9)	464 (71·0)	Ref.	–
Yes	88 (29·1)	190 (29·0)	1·00 (0·74–1·36)	–
P value			0·978	–

n denotes the number of subjects in the cell; Col.%, column percentages; OR (95% CI), odds ratio (95% Confidence Interval); LR, Log-likelihood ratio test.

*Adjustment includes all the variables with results shown (mutual adjustment) as well as sex, age as a continuous variable in single years, malaria status (positive or negative), rural or urban status, proximity of village to surface water (lake, river or swamp), inpatient and outpatient malaria and region, and genotypes previously found to be associated with eBL (*HBB* rs334, *IL10* rs1800896, *IL1A* rs2856838 and *SEMA3C* rs4461841). Adjusted ORs and 95% CIs were not estimated for variables that did not improve significance of the model, so these fields are marked by a hyphen (-). Adjusted *P* values are based on the log-likelihood ratio test comparing the models including and excluding the variable (see *Materials and methods*).

†Other alleles defined by grouping all rare alleles as well as alleles that differed by 2% or less between cases and controls as shown in Table II.

‡HLA homozygosity was defined from the two-field resolution of the HLA sequence results (such as A*02:101 and A*02:102). The two-field indicates the protein coded for by the allele and similar two-field numbers indicate homozygosity, while different numbers indicate heterozygous proteins (see *Materials and methods*).

§When considering HLA-C molecules as ligands for killer cell immunoglobulin-like receptors (KIR) on NK cells, all HLA-C allotypes can be grouped into two major KIR epitopes: C1 (HLA-C*01/*03/*07/*08/*12/*14/*16) and C2 (HLA-C*02/*04/*05/*06/*15/*17/*18), on the basis of a dimorphism at position 80 of the α1 domain. Group C1 (HLA-Casn80) are ligands for inhibitory KIR2DL2/3 and activating KIR2DS2, and group C2 (HLA-Clys80) are ligands for inhibitory KIR2DL1 and activating KIR2DS1.

variation, such as in *HBB*, *ABO*, *DARC* and *G6PD* genes (Karlsson *et al.*, 2014) (Malaria Genomic Epidemiology & Malaria Genomic Epidemiology, 2014) in East *versus* West African populations due to ancestrally distinct populations in West *versus* East Africa (Gouveia *et al.*, 2019). West African populations carry a predominantly West–Central African ancestry, while the Ugandan East African populations carry distinct predominantly Nilotic or admixed Nilotic/Southern Bantu ancestry (Gouveia *et al.*, 2019). Moreover, we also noted distinct malaria-driven positive selection in the *ATP2B4* region (rs10900588) (Gouveia *et al.*, 2019), which is implicated in malaria resistance (Timmann *et al.*, 2012), in northern Ugandan populations but not West African populations. Taken together, these results suggest that genetic findings in West Africa may differ from those in East Africa and underscore the importance of replicating findings in ancestrally matched populations (Wojcik *et al.*, 2019).

The strengths of our study include using NGS methodology to accurately type subjects and having a well-characterized large dataset with multiple variables to adjust for confounding. The weaknesses of our study include relying on *post-hoc* groups of HLA reference allele and not adjusting for multiple comparisons. Despite its large size, our study is still small, so the findings need to be confirmed in a larger sample. Our results may also be misclassified because they are based on EBV-defined BL phenotypes (Grande *et al.*, 2019). Our adjustment for measures of malaria exposure (RDT, treatment history) may be overly conservative because our results are potentially biased to the null. Our observation of strengthening of the HLA associations after adjustment for non-HLA genetic associations raises the

possibility of epistatic interactions, which should be explored in future analyses.

In conclusion, our results suggest novel associations of decreased eBL risk with *HLA-A*02*, *HLA-B*41*, and *-B*58* alleles, and *HLA-DPB1* homozygosity and increased eBL risk with *HLA-DQA1* homozygosity. These findings support the hypothesis that variation in HLA may affect eBL risk. We did not find eBL associations with malaria-associated HLA types previously shown to affect risk of severe malaria in children in West Africa, perhaps because of population genetic or HLA substructures in populations in East *versus* West Africa.

Acknowledgements

We thank Ms. Janet Lawler-Heavner at Westat Inc. (Rockville, MD, USA) and Mr. Erisa Sunday at the African Field Epidemiology Network (Kampala, Uganda) for managing the study. We are grateful to Mr. Wilson Nyegenye at the Uganda Bureau of Statistics (Kampala, Uganda) for training EMBLEM staff in field survey methods. We thank Ms. Laurie Buck, Dr. Carol Giffen, and Mr. Greg Rydzak at Information Management Services Inc. (Calverton, MD, USA) for coordinating data and preparing data analysis files. The work was funded by contracts from the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI) (Contracts HHSN261201100063C and HHSN261201100007I) and, in part, by the Intramural Research Program, National Institute of Allergy and Infectious Diseases (SJR), National Institutes of Health, Department of Health and Human Services, and by a Beginning Investigator Grant for Catalytic Research (BIG Cat) from the National Cancer Institute for African Scientists (US Civilian Research and Development Foundation [CRDF] Global award DAA3-16-62702 to SK) through a collaborative partnership between Makerere University's College of Health Sciences and the Hannover Medical School's Institute of Transfusion Medicine. The content of this manuscript is the sole responsibility of the authors, and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. The sponsors had no role in the study design, data collection, analysis, interpretation, writing of the manuscript, or the decision to submit the paper for publication. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Author Contributions

SK and SMM designed the study; MV, MH and RB performed experiments; SK and SMM performed data analysis; IO, IDL, HN, MDO, PK and TK performed field work; MJ, LWA and SJR supervised field work; SK, MV, MH, RB, MJ, OOO, LWA, SJR, LP-O, JJG, RJB and SMM interpreted the data; SK and SMM wrote the manuscript with input from all co-authors.

Conflict of interest

The authors declare no conflict of interest.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. The average number of sequencing reads per HLA locus.

Table S2. The average ratio (%) of sequencing reads for heterozygous samples indicating similar technical detection of both alleles for each HLA locus.

Table S3. The number of sequenced regions detected as hemizygous or heterozygous per HLA locus

References

Bhat, N.M., Bieber, M.M., Chapman, C.J., Stevenson, F.K. & Teng, N.N. (1993) Human antilipid A monoclonal antibodies bind to human B cells and the i antigen on cord red blood cells. *The Journal of Immunology*, **151**, 5011–5021.

Bodmer, J.G., Bodmer, W.F., Pickbourne, P., Degos, L., Dausset, J. & Dick, H.M. (1975a) Combined analysis of three studies of patients with Burkitt's lymphoma. *Tissue Antigens*, **5**, 63–68. https://doi.org/10.1111/j.1399-0039.1975.tb00529.x

Bodmer, J.G., Bodmer, W.F., Ziegler, J. & Magrath, I.T. (1975b) HL-A and Burkitt's tumour–a study in Uganda. *Tissue Antigens*, **5**, 59–62. https://doi.org/10.1111/j.1399-0039.1975.tb00528.x

Bollard, C.M., Straathof, K.C., Huls, M.H., Leen, A., Lacuesta, K., Davis, A., Gottschalk, S., Brenner, M.K., Heslop, H.E. & Rooney, C.M. (2004) The generation and characterization of LMP2-specific CTLs for use as adoptive transfer from patients with relapsed EBV-positive Hodgkin disease. *Journal of Immunotherapy*, **27**, 317–327. https://doi.org/10.1097/00002371-200407000-00008

Burkitt, D. (1958) A sarcoma involving the jaws in African children. *British Journal of Surgery*, **46**, 218–223. https://doi.org/10.1002/bjs.18004619704

Burnet, F.M. (1970) A certain symmetry: histocompatibility antigens compared with immunocyte receptors. *Nature*, **226**, 123–126. https://doi.org/10.1038/226123a0

Coghill, A.E., Proietti, C., Liu, Z., Krause, L., Bethony, J., Prokunina-Olsson, L., Obajemu, A., Nkrumah, F., Biggar, R.J., Bhatia, K., Hildesheim, A., Doolan, D.L. & Mbulaiteye, S.M. (2020) The association between the comprehensive Epstein-Barr virus serological profile and endemic Burkitt lymphoma. *Cancer Epidemiology, Biomarkers & Prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, **29**, 57–62. https://doi.org/10.1158/1055-9965.EPI-19-0551

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., & Henle, W. (1978) Epidemiological evidence for causal relationship between Epstein-Barr virus and Burkitt's lymphoma from Ugandan prospective study. *Nature*, **274**, 756–761. https://doi.org/10.1038/274756a0

Faridi, R.M. & Agrawal, S. (2011) Killer immunoglobulin-like receptors (KIRs) and HLA-C allorecognition patterns implicative of dominant activation of natural killer cells contribute to recurrent miscarriages. *Human Reproduction*, **26**, 491–497. https://doi.org/10.1093/humrep/deq341

Gilbert, S.C., Plebanski, M., Gupta, S., Morris, J., Cox, M., Aidoo, M., Kwiatkowski, D., Greenwood, B.M., Whittle, H.C. & Hill, A.V. (1998) Association of malaria parasite population structure, HLA, and immunological antagonism. *Science*, **279**, 1173–1177. https://doi.org/10.1126/science.279.5354.1173

Gouveia, M.H., Bergen, A.W., Borda, V., Nunes, K., Leal, T.P., Ogwang, M.D., Yeboah, E.D., Mensah, J.E., Kinyera, T., Otim, I., Nabalende, H., Legason,

I.D., Mpoloka, S.W., Mokone, G.G., Kerchan, P., Bhatia, K., Reynolds, S.J., Birtwum, R.B., Adjei, A.A., Tettey, Y., Tay, E., Hoover, R., Pfeiffer, R.M., Biggar, R.J., Goedert, J.J., Prokunina-Olsson, L., Dean, M., Yeager, M., Lima-Costa, M.F., Hsing, A.W., Tishkoff, S.A., Chanock, S.J., Tarazona-Santos, E. & Mbulaiteye, S.M. (2019) Genetic signatures of gene flow and malaria-driven natural selection in sub-Saharan populations of the "endemic Burkitt Lymphoma belt". *PLoS Genetics*, **15**, e1008027. https://doi.org/10.1371/journal.pgen.1008027

Grande, B.M., Gerhard, D.S., Jiang, A., Griner, N.B., Abramson, J.S., Alexander, T.B., Allen, H., Ayers, L.W., Bethony, J.M., Bhatia, K., Bowen, J., Casper, C., Choi, J.K., Culibrk, L., Davidsen, T.M., Dyer, M.A., Gastier-Foster, J.M., Gesuwan, P., Greiner, T.C., Gross, T.G., Hanf, B., Harris, N.L., He, Y., Irvin, J.D., Jaffe, E.S., Jones, S.J.M., Kerchan, P., Knoetze, N., Leal, F.E., Lichtenberg, T.M., Ma, Y., Martin, J.P., Martin, M.R., Mbulaiteye, S.M., Mullighan, C.G., Mungall, A.J., Namirembe, C., Novik, K., Noy, A., Ogwang, M.D., Omoding, A., Orem, J., Reynolds, S.J., Rushton, C.K., Sandlund, J.T., Schmitz, R., Taylor, C., Wilson, W.H., Wright, G.W., Zhao, E.Y., Marra, M.A., Morin, R.D. & Staudt, L.M. (2019) Genome-wide discovery of somatic coding and noncoding mutations in pediatric endemic and sporadic Burkitt lymphoma. *Blood*, **133**, 1313–1324. https://doi.org/10.1182/blood-2018-09-871418

Greenland, S., Pearl, J. & Robins, J.M. (1999) Causal diagrams for epidemiologic research. *Epidemiology*, **10**, 37–48. https://doi.org/10.1097/00001648-199901000-00008

Guech-Ongey, M., Verboom, M., Pfeiffer, R.M., Schulz, T.F., Ndugwa, C.M., Owor, A.M., Bakaki, P.M., Bhatia, K., Figueiredo, C., Eiz-Vesper, B., Blasczyk, R. & Mbulaiteye, S.M. (2010) HLA polymorphisms and detection of kaposi sarcoma-associated herpesvirus DNA in saliva and peripheral blood among children and their mothers in the uganda sickle cell anemia KSHV Study. *Infectious Agents and Cancer*, **5**, 21. https://doi.org/10.1186/1750-9378-5-21

Hall, P.J., Levin, A.G., Entwistle, C.C., Knight, S.C., Wasunna, A., Kung'u, A. & Brubaker, G. (1982) HLA antigens in East African Black patients with Burkitt's lymphoma or nasopharyngeal carcinoma and in controls: a pilot study. *Human Immunology*, **5**, 91–105. https://doi.org/10.1016/0198-8859(82)90055-6

Hill, A.V. (1998) The immunogenetics of human infectious diseases. *Annual Review of Immunology*, **16**, 593–617. https://doi.org/10.1146/annurev.immunol.16.1.593

Hill, A.V., Allsopp, C.E., Kwiatkowski, D., Anstey, N.M., Twumasi, P., Rowe, P.A., Bennett, S., Brewster, D., McMichael, A.J. & Greenwood, B.M. (1991) Common west African HLA antigens are associated with protection from severe malaria. *Nature*, **352**, 595–600. https://doi.org/10.1038/352595a0

Hill, A.V., Elvin, J., Willis, A.C., Aidoo, M., Allsopp, C.E., Gotch, F.M., Gao, X.M., Takiguchi, M., Greenwood, B.M., Townsend, A.R., McMichael, A. J. & Whittle, H. C. (1992) Molecular analysis of the association of HLA-B53 and resistance to severe malaria. *Nature*, **360**, 434–439. https://doi.org/10.1038/360434a0

Holdsworth, R., Hurley, C.K., Marsh, S.G., Lau, M., Noreen, H.J., Kempenich, J.H., Setterholm, M. & Maiers, M. (2009) The HLA dictionary 2008: a summary of HLA-A, -B, -C, -DRB1/3/4/5, and -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR, and -DQ antigens. *Tissue Antigens*, **73**, 95–170. https://doi.org/10.1111/j.1399-0039.2008.01183.x

Horiki, T., Inoko, H., Moriuchi, J., Ichikawa, Y. & Arimori, S. (1994) Combinations of HLA-DPB1 and HLA-DQB1 alleles determine susceptibility to early-onset myasthenia gravis in Japan. *Autoimmunity*, **19**, 49–54.

Hung, S.I., Chung, W.H., Liou, L.B., Chu, C.C., Lin, M., Huang, H.P., Lin, Y.L., Lan, J.L., Yang, L.C., Hong, H.S., Chen, M.J., Lai, P.C., Wu, M.S., Chu, C.Y., Wang, K.H., Chen, C.H., Fann, C.S., Wu, J.Y. & Chen, Y.T. (2005) HLA-B*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. *Proceedings of the National Academy of Sciences USA*, **102**, 4134–4139. https://doi.org/10.1073/pnas.0409500102

Jones, E.H., Biggar, R.J., Nkrumah, F.K. & Lawler, S.D. (1980) Study of the HLA system in Burkitt's lymphoma. *Human immunology*, **1**, 207–210. https://doi.org/10.1016/0198-8859(80)90015-4

Jones, E.H., Biggar, R.J., Nkrumah, F.K. & Lawler, S.D. (1985) HLA-DR7 association with African Burkitt's lymphoma. *Human immunology*, **13**, 211–217. https://doi.org/10.1016/0198-8859(85)90013-8

Karlsson, E.K., Kwiatkowski, D.P. & Sabeti, P.C. (2014) Natural selection and infectious disease in human populations. *Nature Reviews. Genetics*, **15**, 379–393. https://doi.org/10.1038/nrg3734

Klein, J. & Sato, A. (1998) Birth of the major histocompatibility complex. *Scandinavian Journal of immunology*, **47**, 199–209. https://doi.org/10.1046/j.1365-3083.1998.00292.x

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., Bayanjargal, A., Emmanuel, B., Kagwa, P., Talisuna, A.O., Bhatia, K., Yeager, M., Biggar, R.J., Ayers, L.W., Reynolds, S.J., Goedert, J.J., Ogwang, M.D., Fraumeni, J.F. Jr, Prokunina-Olsson, L. & Mbulaiteye, S.M. (2017) Evaluating the causal link between malaria infection and endemic burkitt lymphoma in Northern Uganda: A mendelian randomization study. *EBioMedicine*, **25**, 58–65. https://doi.org/10.1016/j.ebiom.2017.09.037

Lombardo, K.A., Coffey, D.G., Morales, A.J., Carlson, C.S., Towlerton, A.M.H., Gerdts, S.E., Nkrumah, F.K., Neequaye, J., Biggar, R.J., Orem, J., Casper, C., Mbulaiteye, S.M., Bhatia, K.G. & Warren, E.H. (2017) High-throughput sequencing of the B-cell receptor in African Burkitt lymphoma reveals clues to pathogenesis. *Blood Advances*, **1**, 535–544. https://doi.org/10.1182/bloodadvances.2016000794

Lopez-Mejias, R., Genre, F., Perez, B.S., Castaneda, S., Ortego-Centeno, N., Llorca, J., Ubilla, B., Remuzgo-Martinez, S., Mijares, V., Pina, T., Calvo-Rio, V., Marquez, A., Miranda-Filloy, J.A., Parejo, A.N., Conde-Jaldon, M., Ortiz-Fernandez, L., Argila, D., Aragues, M., Rubio, E., Luque, M.L., Blanco-Madrigal, J.M., Galindez-Aguirregoikoa, E., Gonzalez Escribano, F., Ocejo-Vinyals, J.G., Martin, J., Blanco, R. & Gonzalez-Gay, M.A. (2015) Association of HLA-B*41:02 with Henoch-Schonlein Purpura (IgA Vasculitis) in Spanish individuals irrespective of the HLA-DRB1 status. *Arthritis Research & Therapy*, **17**, 102. https://doi.org/10.1186/s13075-015-0622-5

Malaria Genomic Epidemiology, N. & Malaria Genomic Epidemiology, N. (2014) Reappraisal of known malaria resistance loci in a large multicenter study. *Nature Genetics*, **46**, 1197–1204. https://doi.org/10.1038/ng.3107

May, J., Lell, B., Luty, A.J., Meyer, C.G. & Kremsner, P.G. (2001) HLA-DQB1*0501-restricted Th1 type immune responses to Plasmodium falciparum liver stage antigen 1 protect against malaria anemia and reinfections. *The Journal of Infectious Diseases*, **183**, 168–172.

Maziarz, M., Kinyera, T., Otim, I., Kagwa, P., Nabalende, H., Legason, I.D., Ogwang, M.D., Kirimunda, S., Emmanuel, B., Reynolds, S.J., Kerchan, P., Joloba, M.M., Bergen, A.W., Bhatia, K., Talisuna, A.O., Biggar, R.J., Goedert, J.J., Pfeiffer, R.M. & Mbulaiteye, S.M. (2017) Age and geographic patterns of Plasmodium falciparum malaria infection in a representative sample of children living in Burkitt lymphoma-endemic areas of northern Uganda. *Malaria Journal*, **16**, 124. https://doi.org/10.1186/s12936-017-1778-z

Mbulaiteye, S.M., Morton, L.M., Sampson, J.N., Chang, E.T., Costas, L., de Sanjose, S., Lightfoot, T., Kelly, J., Friedberg, J.W., Cozen, W., Marcos-Gragera, R., Slager, S.L., Birmann, B.M. & Weisenburger, D.D. (2014) Medical history, lifestyle, family history, and occupational risk factors for sporadic Burkitt lymphoma/leukemia: the Interlymph Non-Hodgkin Lymphoma Subtypes Project. *Journal of the National Cancer Institute Monographs*, **2014**, 106–114. https://doi.org/10.1093/jncimonographs/lgu003

Murray, R.J., Kurilla, M.G., Brooks, J.M., Thomas, W.A., Rowe, M., Kieff, E. & Rickinson, A.B. (1992) Identification of target antigens for the human cytotoxic T cell response to Epstein-Barr virus (EBV): implications for the immune control of EBV-positive malignancies. *The Journal of Experimental Medicine*, **176**, 157–168. https://doi.org/10.1084/jem.176.1.157

Ng, V.L., Hurt, M.H., Fein, C.L., Khayam-Bashi, F., Marsh, J., Nunes, W.M., McPhaul, L.W., Feigal, E., Nelson, P., & Herndier, B.G. (1994) IgMs produced by two acquired immune deficiency syndrome lymphoma cell lines: Ig binding specificity and VH-gene putative somatic mutation analysis. *Blood*, **83**, 1067–1078.

Niens, M., Jarrett, R.F., Hepkema, B., Nolte, I.M., Diepstra, A., Platteel, M., Kouprie, N., Delury, C.P., Gallagher, A., Visser, L., Poppema, S., te Meerman, G.J. & van den Berg, A. (2007) HLA-A*02 is associated with a reduced risk and HLA-A*01 with an increased risk of developing EBV+ Hodgkin lymphoma. *Blood*, **110**, 3310–3315. https://doi.org/10.1182/blood-2007-05-086934

Ogwang, M.D., Bhatia, K., Biggar, R.J. & Mbulaiteye, S.M. (2008) Incidence and geographic distribution

of endemic Burkitt lymphoma in northern Uganda revisited. *International Journal of Cancer. Journal International du Cancer*, **123**, 2658–2663. https://doi.org/10.1002/ijc.23800

Osafo-Addo, A.D., Koram, K.A., Oduro, A.R., Wilson, M., Hodgson, A. & Rogers, W.O. (2008) HLA-DRB1*04 allele is associated with severe malaria in northern Ghana. *The American Journal of Tropical Medicine and Hygiene*, **78**, 251–255. https://doi.org/10.4269/ajtmh.2008.78.251

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., Kawira, E., Magatti, J., Kinyera, T., Otim, I., Legason, I.D., Nabalende, H., Dhudha, H., Ally, H., Genga, I.O., Mumia, M., Ayers, L.W., Pfeiffer, R.M., Biggar, R.J., Bhatia, K., Goedert, J.J. & Mbulaiteye, S.M. (2020) Risk factors for Burkitt lymphoma in East African children and minors: A case-control study in malaria-endemic regions in Uganda, Tanzania and Kenya. *International Journal of Cancer. Journal International du Cancer*, **146**, 953–969. https://doi.org/10.1002/ijc.32390

Peterson, T.A., Kimani, J., Wachihi, C., Bielawny, T., Mendoza, L., Thavaneswaran, S., Narayansingh, M.J., Kariri, T., Liang, B., Ball, T.B., Ngugi, E.N., Plummer, F.A. & Luo, M. (2013) HLA class I associations with rates of HIV-1 seroconversion and disease progression in the Pumwani Sex Worker Cohort. *Tissue Antigens*, **81**, 93–107. https://doi.org/10.1111/tan.12051

Potolicchio, I., Mosconi, G., Forni, A., Nemery, B., Seghizzi, P. & Sorrentino, R. (1997) Susceptibility to hard metal lung disease is strongly associated with the presence of glutamate 69 in HLA-DP beta chain. *European Journal of Immunology*, **27**, 2741–2743.

Riboldi, P., Gaidano, G., Schettino, E.W., Steger, T.G., Knowles, D.M., Dalla-Favera, R. & Casali, P. (1994) Two acquired immunodeficiency syndrome-associated Burkitt's lymphomas produce specific anti-i IgM cold agglutinins using somatically mutated VH4-21 segments. *Blood*, **83**, 2952–2961. https://doi.org/10.1182/blood.V83.10.2952.2952

Sallah, N., Carstensen, T., Wakeham, K., Bagni, R., Labo, N., Pollard, M.O., Gurdasani, D., Ekoru, K., Pomilla, C., Young, E.H., Fatumo, S., Asiki, G., Kamali, A., Sandhu, M., Kellam, P., Whitby, D., Barroso, I. & Newton, R. (2017) Whole-genome association study of antibody response to Epstein-Barr virus in an African population: A pilot. *Glob Health Epidemiol Genom*, **2**, e18. https://doi.org/10.1017/gheg.2017.16

Shaw, B.E., Gooley, T.A., Malkki, M., Madrigal, J.A., Begovich, A.B., Horowitz, M.M., Gratwohl, A., Ringden, O., Marsh, S.G. & Petersdorf, E.W. (2007) The importance of HLA-DPB1 in unrelated donor hematopoietic cell transplantation. *Blood*, **110**, 4560–4566. https://doi.org/10.1182/blood-2007-06-095265

Terasaki, P.I. & McClelland, J.D. (1964) Microdroplet assay of human serum cytotoxins. *Nature*, **204**, 998–1000. https://doi.org/10.1038/204998b0

Thorsby, E. (2009) A short history of HLA. *Tissue Antigens*, **74**, 101–116. https://doi.org/10.1111/j.1399-0039.2009.01291.x

Timmann, C., Thye, T., Vens, M., Evans, J., May, J., Ehmen, C., Sievertsen, J., Muntau, B., Ruge, G., Loag, W., Ansong, D., Antwi, S., Asafo-Adjei, E., Nguah, S.B., Kwakye, K.O., Akoto, A.O., Sylverken, J., Brendel, M., Schuldt, K., Loley, C., Franke, A., Meyer, C.G., Agbenyega, T., Ziegler, A. & Horstmann, R.D. (2012) Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature*, **489**, 443–446. https://doi.org/10.1038/nature11334

Wojcik, G.L., Graff, M., Nishimura, K.K., Tao, R., Haessler, J., Gignoux, C.R., Highland, H.M., Patel, Y.M., Sorokin, E.P., Avery, C.L., Belbin, G.M., Bien, S.A., Cheng, I., Cullina, S., Hodonsky, C.J., Hu, Y., Huckins, L.M., Jeff, J., Justice, A.E., Kocarnik, J.M., Lim, U., Lin, B.M., Lu, Y., Nelson, S.C., Park, S.L., Poisner, H., Preuss, M.H., Richard, M.A., Schurmann, C., Setiawan, V.W., Sockell, A., Vahi, K., Verbanck, M., Vishnu, A., Walker, R.W., Young, K.L., Zubair, N., Acuna-Alonso, V., Ambite, J.L., Barnes, K.C., Boerwinkle, E., Bottinger, E.P., Bustamante, C.D., Caberto, C., Canizales-Quinteros, S., Conomos, M.P., Deelman, E., Do, R., Doheny, K., Fernandez-Rhodes, L., Fornage, M., Hailu, B., Heiss, G., Henn, B.M., Hindorff, L.A., Jackson, R.D., Laurie, C.A., Laurie, C.C., Li, Y., Lin, D.Y., Moreno-Estrada, A., Nadkarni, G., Norman, P.J., Pooler, L.C., Reiner, A.P., Romm, J., Sabatti, C., Sandoval, K., Sheng, X., Stahl, E.A., Stram, D.O., Thornton, T.A., Wassel, C.L., Wilkens, L.R., Winkler, C.A., Yoneyama, S., Buyske, S., Haiman, C.A., Kooperberg, C., Le Marchand, L., Loos, R.J.F., Matise, T.C., North, K.E., Peters, U., Kenny, E.E. & Carlson, C.S. (2019) Genetic analyses of diverse populations improves discovery for complex traits. *Nature*, **570**, 514–518. https://doi.org/10.1038/s41586-019-1310-4

Woszczek, G., Borowiec, M., Mis, M., Gorska, M. & Kowalski, M.L. (1997) Comparison of serological and molecular (PCR-SSP) techniques of HLA-DR typing in clinical laboratory routine. *Ann Transplant*, **2**, 39–42.