

## Review Article

# Human Leukocyte Antigen Diversity: A Southern African Perspective

**Mqondisi Tshabalala, Juanita Mellet, and Michael S. Pepper**

*Department of Immunology and Institute for Cellular and Molecular Medicine, Faculty of Health Sciences, University of Pretoria, P.O. Box 2034, Pretoria 0001, South Africa*

Correspondence should be addressed to Michael S. Pepper; michael.pepper@up.ac.za

Received 1 May 2015; Accepted 26 July 2015

Academic Editor: Nejat K. Egilmez

Copyright © 2015 Mqondisi Tshabalala et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Despite the increasingly well-documented evidence of high genetic, ethnic, and linguistic diversity amongst African populations, there is limited data on human leukocyte antigen (HLA) diversity in these populations. HLA is part of the host defense mechanism mediated through antigen presentation to effector cells of the immune system. With the high disease burden in southern Africa, HLA diversity data is increasingly important in the design of population-specific vaccines and the improvement of transplantation therapeutic interventions. This review highlights the paucity of HLA diversity data amongst southern African populations and defines a need for information of this kind. This information will support disease association studies, provide guidance in vaccine design, and improve transplantation outcomes.

## 1. Introduction

The human leukocyte antigen (HLA) complex on chromosome 6, also known as the major histocompatibility complex (MHC) in all mammals, consists of highly polymorphic genes whose protein products present antigens to T cells as part of an immune response to infections [1, 2]. HLA molecules also impact on the development and effectiveness of vaccines and play a determining role in the outcomes of transplantation [3–10].

The World Health Organization (WHO) indicates that there is a high burden of disease in southern Africa, especially communicable diseases such as HIV/AIDS, TB, and malaria [11]. Despite the increasingly well-documented high genetic diversity observed amongst human populations in southern Africa [12], there is limited information on HLA diversity [8]. Understanding HLA diversity in these populations will provide insight into HLA disease associations and may help in vaccine development. Transplantation as a therapeutic intervention requires strict HLA allele matching between donors and recipients to reduce rejection and the incidence of graft versus host disease (GVHD). Good clinical outcomes in transplant recipients are observed in cases of high resolution

HLA matching [13, 14], with the number of mismatches correlating with the risk of rejection and/or GVHD [15–17]. It is currently very difficult to match donor-recipient pairs in bone marrow registries in southern Africa, partly because of the great genetic diversity in this population. A recent study identified Black and Caucasian South African population-specific alleles [18], highlighting the need to investigate HLA diversity amongst southern Africans to improve global representation in the International ImMunoGeneTics information system (IMGT)/HLA database [1, 2]. HLA typing methods use the IMGT/HLA database as a reference; it is thus difficult to match individuals who have alleles which are not captured in the database.

HLA typing methods have evolved from low resolution serology typing to high resolution DNA sequencing based technologies (SBT). Despite high resolution, SBT has limitations of mostly typing certain exons within the HLA loci [19]. The antigen-binding grooves encoded by exons 2 and 3 (class I) and exon 2 (class II) are routinely sequenced in most laboratories, thereby giving partial sequences of about 10% of the reported alleles [19]. Another potential source of ambiguity in SBT HLA typing is the *cis/trans* assignment of DNA bases in a heterogeneous sample [20], yielding limited

resolution data and thereby making it difficult to assign HLA allele types. It is possible to sequence the entire HLA region with current methods, but at a very high cost and a need for expert analysis. There have been advances in the use of next generation sequencing (NGS) in HLA typing to improve coverage of the HLA gene loci by high throughput, while at the same time reducing ambiguity associated with SBT typing [19, 21, 22]. To fully appreciate the NGS HLA typing tool, there is need for a complete HLA allele database [21] highlighting the need to quantify HLA diversity in the genetically diverse southern African populations [23].

African populations have been shown to be genetically diverse [12] and are believed to be the cradle of humankind [24, 25]. In general, genetic diversity of African populations is poorly understood [8] thereby limiting our understanding of human health and susceptibility to diseases, hence the need for further analysis/evaluation to map disease association and therapeutic gene targets. Despite the general similarities in culture and shared geographical location, genetic differences exist among populations at every 1000 base pairs [26, 27]. In this review, we examine available HLA diversity data in southern Africa with a view to understanding disease burden, planning registry recruitment and donor-recipient matching, and providing insights into the evolution of the ethnic and linguistic diversity in this region. This review specifically focuses on classical HLA diversity in southern African countries (characterized by genetically, culturally, and linguistically diverse Bantu ethnicities and admixed populations [28–31]) herein defined as Zambia, Malawi, Zimbabwe, Mozambique, Angola, Namibia, Botswana, South Africa, Lesotho, and Swaziland.

## 2. HLA Diversity

There are an ever increasing number of HLA alleles, reflecting the rate of discovery of the diversity of the gene loci [1, 2]. There are currently 13412 HLA alleles described by the HLA nomenclature and included in the IMGT/HLA database (based on IMGT/HLA 3.21.0 release, 06 July 2015), with HLA-B having the highest number of alleles (3977) [32]. HLA genetic variation does not vary in an individual's lifetime, but high diversity is observed at the population level [1, 2, 33–38]. High HLA allelic diversity in humans is reflected by the high number of pseudogenes and can be explained by natural selection and coevolution with pathogens. There is an advantage of HLA diversity related to pathogen-derived peptide presentation to effector T cells: heterozygous individuals can potentially present more antigens than homozygotes for the different HLA alleles (heterozygosity advantage) [33, 39]. In nonhuman species, low MHC diversity has been observed in several species (Tasmanian devils, cheetah, and panda) and has been associated with disease susceptibility in some Tasmania devils [40], highlighting the advantage of HLA diversity in presenting many different antigens to effector cells of the immune system.

Prugnonle et al. suggested that up to 39% of observed HLA class I diversity was due to geographical distance (and consequently human migration history) from the source of

modern humans (assumed to be Ethiopia in this study), with the unaccounted source of diversity most likely being from pathogen driven selection [41]. Generally, populations exposed to a high pathogen burden show high HLA diversity, and there is a decreasing HLA diversity away from Africa (geographically measured by landmasses away from Africa) [41]. In related studies, microsatellite data has suggested that geographic distance from East Africa (probable source of modern humans) explains about 85% of a decreasing genetic diversity within human populations from the source (reviewed in [42]). Interestingly, HLA-C is less expressed on cell surfaces; hence its diversity is least likely to be driven by viral pathogens (reviewed in [41]). It is historically accepted that TB was a major selective pressure in the evolution of Western European populations [43], with malaria acting on African populations [44]. These pathogens exerted a high selective pressure mostly on genes of the immune system (particularly those involved in protective immunity).

There is growing evidence for positive selection being responsible for maintaining HLA polymorphisms, most likely due to overdominant selection (heterozygote advantage) which maintains allelic lineages for much longer periods of time than neutral polymorphisms [38, 45–47]. Globally, HLA diversity seems to be highest within populations than between populations (evidenced by major differences amongst continents) [1, 2, 35, 48]. Several studies have highlighted alternative splicing of HLA class I genes giving rise to diverse isoforms [49] which might contribute to this diversity. For example, alternative splicing to exclude exon 5 has been reported to give rise to several isoforms of HLA-A and -B [50]. Alternative splicing in other HLA class I exons has also been reported [51] including the nonclassical HLA-G gene [52].

Other mechanisms of HLA diversity generation include point mutations (substitution, deletion, and insertion): gene conversion (unidirectional gene transfer) and gene cross over (bidirectional gene transfer). Gene cross over, which is a form of recombination that can be intra/inter-HLA loci during meiosis, enables exchange of genetic material linked to the generation of novel alleles in offspring as described by Carrington [53]. Other recombination events include gene conversion, a bidirectional donation of DNA between two homologous chromosomes. A recent study reports novel HLA alleles resulting from (a) nonsynonymous amino acid change (HLA B\*41:21, HLA DQB1\*02:10, HLA QAI\*01:12); (b) deletion leading to frame shift (HLA A\*01:123N); (c) intralocus gene conversion (HLA B\*35:231, HLA B\*53:31); and (d) interlocus gene conversion (HLA C\*07:294) [54]. It is important to note the low frequency of interlocus generated alleles as reported by several other studies as reviewed by Adamek et al. [54].

*2.1. HLA Diversity in Transplantation and Transfusion.* The human immune system uses HLA's uniqueness in every individual to recognize self from nonself; hence the body only mounts an immune response against foreign cells/molecules under normal conditions. Transplantation as a therapeutic

intervention matches donor and recipient HLA molecules to decrease the likelihood of rejection [33]. The likelihood of two individuals having identical HLA molecules on all loci is very low, except for siblings, who have a 25% chance of being HLA-identical as a result of HLA molecules being codominantly expressed and inherited as haplotypes from both parents. The degree of HLA matching is a predictor of clinical outcome.

GVHD is an immunocompetent donor T cell mediated response against the recipient's immune system which is common in unmatched donor-recipient pairs. Acute GVHD can be reduced by donor T cell depletion, but this increases the risk of rejection, malignant disease relapse, and impaired immune recovery [63, 64]. In addition to HLA matching, killer-cell immunoglobulin-like receptors (KIRs) have been documented to affect the clinical outcome of allogeneic transplantation [65–68]. In severe immunocompromised individuals, allogeneic transfusion with immune competent T cell-containing blood products might lead to transfusion associated GVHD. Transfusion related lung injury (TRALI) is an anti-HLA (mostly class I [69, 70]) antibody related complication which may be fatal. Anti-HLA class II antibodies induce TRALI through monocyte and subsequent neutrophil activation [69, 71]. Anti-HLA class I antibodies have been reported to be a cause of neonatal alloimmune thrombocytopenia together with platelet-derived specific antigens [72]. It is critical to know the population HLA diversity in order to improve donor-recipient matching in both transplantation and transfusion therapeutic interventions. Diversity data informs decision making in transplantation and transfusion aimed at reducing rejection while at the same time improving the outcome of the intended therapeutic intervention. Recruitment of donors from minority or underrepresented populations might help to improve HLA diversity in registries [34] which improves the chances of donor-recipient matching.

*2.2. HLA Diversity in Human Disease Associations.* The high disease burden in southern Africa [11] offers a unique opportunity to study HLA disease association [8]. Several autoimmune conditions have been directly associated with specific class I and II HLA alleles, including rheumatoid arthritis, multiple sclerosis, ankylosing spondylitis, and Grave's disease, as reviewed by Trowsdale and Knight [73]. Several alleles have been associated with varying rates of HIV disease progression [4, 39, 74–76], susceptibility, transmission, and treatment outcomes (reviewed in [76]). HLA has likewise been associated with malaria [6] and TB susceptibility and protection [77] in various populations. In another example, although not directly related to southern Africa, the HLA-B locus has been linked to fatal and nonfatal Sudanese Ebola strains. Thus, HLA-B\*67 and -B\*15 have been associated with fatal outcomes and B\*07 and B\*14 have been associated with nonfatal Ebola infections [78].

Haplotype analysis gives information on disease/condition associated alleles, which are assumed to be inherited as blocks due to strong linkage disequilibrium [79]. HLA alleles can be imputed from analyzing identity by descent (IBD) patterns within the HLA region of specific populations. This

approach leverages on the observation that chromosomes with high IBD within MHC most likely share the same alleles. Haplotype analysis or SNP-based HLA allele imputation is important for disease association studies but will not replace classical HLA typing for transplantation applications where a high degree of haplomatching is required for a good clinical outcome [80]. Currently several imputation methods are available to type HLA genes in silico and to fine-map associations within classical HLA genes [80]. Unfortunately, limited HLA diversity data from populations such as those in southern Africa make this difficult [80].

*2.3. HLA Diversity in Population Studies.* There is documented evidence of geographical distribution of human genetic variation, which helps to understand human evolution, migration and adaptation to different environments and pathogens [81]. Several efforts aimed at understanding global human genetic diversity including the Hap Map Project [82], 1000 Genomes Project [83], and recently the African Genome Variation Project [31]; however, all of these have limited information on southern African populations. Some African genetic diversity studies have focused on targeted populations like hunter gatherers [84, 85] or have had very limited sample size [86] and are therefore not representative of southern Africa. The low representation of southern African genetic data in global efforts makes it difficult to use the currently available reference panels for these populations, especially in disease association studies [31]. This suggests that targeted HLA sequencing of these diverse populations is necessary to improve their representation in reference panels.

There are marked differences in HLA diversity distribution globally, with geographically separated regions showing varying degrees of diversity [35, 41, 42, 48]. Most HLA loci show high allele numbers across populations [35, 87]. HLA DPA1 has the least number of alleles (40 as of July 2015) [88] compared to other classical HLA loci (e.g., HLA DQB1 which has 807 alleles). This is generally due to the fact that DPB1 loci are not routinely sequenced for transplantation purposes as are other HLA genes. The global distribution of HLA diversity provides insight into human migration patterns and could help understand past pathogen exposures [38]. As an example, HLA studies have been used to trace the spread of modern humans from East Africa and model for coevolution of genes and languages in Africa [89]. Interpretation of HLA in population studies can be improved by extensive knowledge of HLA diversity in these populations.

*2.4. Contemporary Studies on HLA Diversity in Southern Africa.* To highlight the paucity of HLA diversity data in southern Africa, this review used a comprehensive literature search for previously published work on HLA diversity together with the Allele Frequency Net Database (AFND) to determine the information in the public domain. The key search terms for articles were “HLA AND genetic diversity AND southern Africa”. Allele frequency data from AFND was extracted for sub-Saharan African countries, from which southern African data was compiled (Supplementary Table S1, in Supplementary Material available



online at <http://dx.doi.org/10.1155/2015/746151>). Table 1 summarizes allele frequency data from the AFND web search (<http://www.allelefrequencies.net/>) [48] used in this review. The AFND is a public global database of alleles, genotypes, and haplotype frequencies of HLA and KIRs from different studies, reports, and proceedings of international workshops in immunogenetics and histocompatibility. HLA data is generated by different typing methods but is curated in the database in accordance with the updated IMGT/HLA guidelines (this review used the 3.15.0 release, 17 January 2014) [1, 2, 35, 48]. For this review, only positive allele frequencies from all ethnic groups within sub-Saharan Africa were extracted from the database (<http://www.allelefrequencies.net/>) [35, 48]. The number of alleles reported in Mozambicans, Black South Africans, Caucasian South Africans, Tamil South Africans, Zulu South Africans, Tswana South Africans, Zambians, and Shona Zimbabweans, respectively, was 18, 33, 25, 16, 37, 15, 20, and 32 alleles for HLA-A and 25, 30, 41, 23, 45, 14, 29, and 46 alleles for HLA-B. HLA-C alleles were only reported for Black South Africans (28 alleles), Caucasian South Africans (29 alleles), Tamil South Africans (21 alleles), Zambians (12 alleles), and Shona Zimbabweans (24 alleles). All HLA class II alleles in the AFND were only reported for Shona Zimbabweans and South African Vendas as summarized in Supplementary Table S1. Tables 1 and 2 summarize the selected allele frequencies from southern African populations and the total number of classical HLA alleles reported across different global regions as defined in the AFND [35, 48], respectively.

South Africa had the highest number of HLA data sets from the AFND compared to other southern African countries (Table 1(a)). Some southern African countries (Angola, Lesotho, Malawi, Namibia, and Swaziland) have no HLA data available (Table 1(a)). As summarized in Tables 1(b) and 1(c), HLA-A\*30 and its derivatives (A\*30:01, A\*30:02) are common in black populations (Mozambicans, Black South Africans, Zulus, Tswanas, Zambians, and Zimbabwean Shonas). Caucasians and Tamils had a completely different HLA-A allele frequency distribution compared to the other populations. HLA-A\*02:01:01 was most frequent (0.26) in South African Caucasians, as has been reported by Solberg et al. (HLA-A\*02:01) in European (27%) and white American (20%) populations [90]. This suggests that South African Caucasians have a common ancestry with the Europeans and Americans, with the A\*02:01 allele and its derivatives being restricted mostly to white populations. For the HLA-B locus, B\*58 (B\*58:02, B\*58:01) was most common in Mozambicans, Black South Africans (including Zulus and Tswanas) as highlighted in Tables 1(b) and 1(c). All HLA-B allele frequencies were less than 0.1 in Black South Africans and Shonas. All HLA-C frequencies were less than 0.2, with C\*06:02 being commonly high in Black South Africans and Tamils. Although more than ten years old (2004), the study by Cao et al. identified A\*02:02, A\*34:02, A\*36:01, A\*74:01, B\*15:03, B\*42:01, B\*53:01, B\*57:03, and B\*58:02 as unique African alleles. Recently, diverse and novel HLA alleles have been reported in sub-Saharan populations, for example, HLA class II as reviewed in Ayele et al. [91] and HLA class I as described by Paximadis et al. [18], to further

support high genetic diversity in Africans and intra-African diversity. Interestingly five new class I alleles (A\*30:01:02, A\*30:02:02, A\*68:27, B\*42:06, and B\*45:07) were reported in a recent South African study [18]. Additionally, Shepherd et al. recently reported an overrepresentation of HLA-A\*02:01, -A\*34:02, and -B\*58:02 in HIV negative controls in Zimbabwe [92] compared to the HIV positive group, which supports the earlier notion of African specific alleles.

The AFND reports very few HLA class II alleles amongst southern African populations; only Zimbabwean Shonas and Black South Africans [18] had HLA-DP data. The reported allele frequencies (Tables 1(b) and 1(c)) for the DP locus were most frequent DPB1\*01:01:01 (0.355) in Shona Zimbabweans and DPB1\*13:01 (0.148) in Black South Africans; and least frequent DPB1\*01:01:02, DPB1\*02:02, DPB1\*62:01, DPB1\*65:01, and DPB1\*80:01 (0.002) in Shona Zimbabweans. No alleles were reported for the DPA1 and DQA1 loci. The DQB1 locus was reported only in Botswana, Black South Africans, Shona Zimbabweans, and Venda South Africans. DQB1\*06 in Black South Africans was the most frequent (0.555) with DQB1\*06:15 in Shona Zimbabweans being least frequent (0.002). DRB1 alleles were reported in all the studied populations except in some South Africans (Tswana, Tamil, and Zulu). The most frequent allele was DRB1\*11 (0.366) in Black South Africans, while the least frequent ones were DRB1\*16 (0.002) in Mozambicans and DRB1\*03, DRB1\*04:04, DRB1\*12:04, DRB1\*13, and DRB1\*15:01 (all at 0.002) in Shona Zimbabweans.

The number of classical HLA alleles (Table 2) varies greatly in each geographical region, with North Africa having the highest number of AFND reported alleles globally, and sub-Saharan Africa (including southern Africa) in the top 5. In terms of HLA class II alleles, sub-Saharan Africa falls in the bottom 5 regions (with the least number of alleles, Table 2) for most of the HLA loci (DQA1, DQB1, and DRB1). The DP locus generally has fewer numbers of reported alleles globally (<http://www.allelefrequencies.net/>) [35, 48]. Interestingly, more than 50% of HLA class I alleles reported for sub-Saharan Africa are in southern Africa (Table 2), further highlighting diversity in this region. No HLA-DPA1 alleles were reported by the AFND in southern Africa, with less than 50% of the other class II alleles reported in sub-Saharan Africa coming from southern Africa.

The number of southern African HLA studies in the AFND is relatively low, reflecting the underrepresentation of this region. The data currently available is mostly low resolution with low sample numbers and is not a true reflection of HLA diversity in the southern African context. This highlights the need for continual submission of southern African HLA diversity data to centralized databases like the AFND. The few studies from southern Africa also highlight the knowledge gap on HLA diversity in this region in this era of high resolution typing. Several HLA disease association studies with allele frequency data have been reported in the region [7, 59, 93–95]; these frequencies might not be a true reflection of the general population owing to the confounding effect of the diseases. Allele frequency is highly dependent on sample size and hence might not give a clear picture of HLA diversity.

TABLE 1: Contemporary studies which provide insight into HLA diversity in southern Africa. HLA allele frequency from the studies cited was extracted from the AFND [35, 48] to assess HLA diversity in southern Africa. The AFND curated allele frequency data was generated from Mozambique, South Africa, Zambia, and Zimbabwe as shown in (a) with the most and least frequent classical HLA alleles in these populations as shown in (b) and (c).

(a) General description of studies used in this review						
Country	Year	Population	n	Typing method	Loci typed	Comments
Bots	2005		55	SSP	DRB, DQB1	55 HIV negative compared to 74 HIV positive [7]
Moza	2010	Mostly Black	202	SSOP	A, B, DRB1	91.8% Black, rest admixture. Assane et al. [35, 48, 55]
RSA	2012	Black	200	SBT, SSP	A, B, C, DRB1	Blacks from different ethnolinguistic groups in RSA. Paximadis et al. [18, 35, 48]
RSA	2012	Caucasians	102	SBT, SSP	A, B, C, DRB1	English and Afrikaner ancestry. Paximadis et al. [18, 35, 48]
RSA	2002	Tamil/Natal	51	SSOP	A, B, C	Hammond [35, 48, 56]
RSA	2000	Black Zulu/Natal	100	SSOP	A, B	Could not distinguish A*0301 from A*0303N, and B*0705 from B*0706 [35, 37, 48, 57]
RSA	2006	Black/Tswana	41	SSOP	A, B	Coetzee et al. [35, 48, 58]
RSA	2004	Black	112	SSP	DRB1, DQB1, DPB1	112 Sclerosis controls compared to cases [59]
Zam	2002	Black/Lusaka	44	SSOP	A, B, C	Alleles similar at exons 2 and 3 could not be distinguished [35, 48, 60, 61]
Zim	2002	Shona/Harare	230	SSOP	A, B, C, DPB1, DQA1, DQB1, DRB1	Louie [35, 48, 62]
(b) Most frequent alleles in different southern African populations [35, 48]						
Population	A	B	C	Loci		
Black RSA	A*30:01 (0.101)	B*42:01 (0.089), B*58:02 (0.094)	C*06:02 (0.149)	DPB1*13:01 (0.148) [59]	DQ DQB1*06 (0.555) [59] DQB1*13:01 (0.124) DQB1*16 (0.509) [7]	
Caucasian RSA	A*01:01:01 (0.2), (0.26)	B*07:02:01 (0.149)	C*07:01 (0.172), C*07:02:01 (0.137)		DRB1*03:01 (0.122) DRB1*II (0.196), DRB1*13 (0.198)	
Shona Zim	A*30:02 (0.147)	B*45:01 and B*53:01 (0.093)	C*04:01 (0.148)	DPB1*01:01:01 (0.355)	DQA1*01:02 (0.343), DQB1*05:01 (0.227), DQB1*15:03 (0.153)	
Tamil RSA	A*01:01 (0.17), A*11:01 (0.18)	B*40:06 (0.143)	C*06:02 (0.177)			

(b) Continued.

Population	Loci				
	A	B	C	DP	DRBI
Tswana RSA	A*02 (0.146), A*30 (0.159)	B*58 (0.22)			DRBI*II (0.184)
Venda RSA	A*30:02 (0.233)	B*42:01 (0.148)	C*17:01 (0.156)		DQB1*06 (0.437)
Zam	A*30 (0.195)	B*15 (0.15), B*58 (0.145)			

(c) Least frequent alleles in different southern African populations [35, 48]

Population	Loci				
	A	B	C	DP	DRBI
Bots				DQB1*02 (0.127) [7]	DRBI*10 and DRBI*12 (0.074) [7]
Caucasian RSA	A*02:05, A*02:17, A*11:12, A*24:07, A*25:01:01, A*33:03:01 and A*69:01 (0.005)	B*07:06, B*14:01, B*15:02, B*15:03, B*15:10, B*15:13, B*15:16, B*15:24, B*27:02, B*35:05, B*40:06:01, B*41:01, B*44:04, B*44:27, B*45:01, B*49:01, B*50:01 and B*58:02 (0.005)	C*02:05, C*03:16, C*04:08, C*04:09N, C*06:11, C*07:22, C*08:01, C*14:04 and C*17:01 (0.005)		DRBI*03:02, DRBI*04:08, DRBI*12:02, DRBI*14:04 and DRBI*15:07 (0.005)
Moza	A*32 (0.002)	B*27, B*37, B*73 and B*82 (0.002)			DRBI*16 (0.002)
Shona Zim	A*02:17, A*32:02, A*34:01, A*80:01, A*66:02, A*66:03 and A*74 (0.002)	B*07:12, B*13:04, B*14:04, B*15:17, B*15:18, B*35:02, B*39:10, B*40:01, B*40:16, B*50:02 and B*73:01 (0.002)	C*03:04:01, C*07:08, C*12:04:02 and C*15:05 (0.02)	DPBI*01:01:02, DPBI*02:02, DPBI*62:01, DPBI*65:01 and DPBI*80:01 (0.002)	DRBI*03, DRBI*04:04, DRBI*12:04, DRBI*13 and DRBI*15:01 (0.002)
Tamil RSA	A*02:01, A*02:03, A*03:02, A*24:07, A*30:01 and A*32:01 (0.001)	B*15:25, B*27:05, B*44:07, B*50:01 and B*56:01 (0.01)	C*02:02:01, C*12:03, C*15:02 and C*16:01 (0.01)		
Tswana RSA	A*01, A*31, A*32, A*36 and A*80 (0.012)	B*35, B*40, B*50 and B*53 (0.012)			
Venda RSA				DQB1*04 (0.094)	DRBI*10:01 (0.004)
Zam	A*02:06, A*02:14, A*26:01, A*33:01, A*34:02, A*43:01 and A*66:01 (0.012)	B*07:05, B*13:02, B*15:18, B*18:03, B*41:01, B*44:05, B*47:01, B*49:01 and B*57:01 (0.011)	C*03:03 and C*07:04 (0.022)		
Zulu RSA	A*31, A*31:01:02, A*33 and A*33:03 (0.005)	B*15:01, B*15:16, B*41:01, B*41:02, B*67, B*67:01, B*82 and B*82:01 (0.005)			

n = sample size, Bots = Botswana, Moza = Mozambique, RSA = Republic of South Africa, Zam = Zambia, Zim = Zimbabwe, SSP = sequence specific primers, SBT = sequence based typing, SSOP = sequence specific oligonucleotide primers, and (number) is allele frequency in the population stated. Blanks indicate no alleles reported in the population or ethnicity not defined or typing method not specified.

TABLE 2: Number of classical HLA alleles reported in each geographical region. Sub-Saharan Africa (including southern Africa) generally has a high number of class I alleles (ranked in the top 5 regions) with a low number of class II alleles (ranked in the bottom 5 regions). More than half of the reported class I alleles in sub-Saharan Africa come from southern Africa, with less than half of the reported class II alleles in the sub-Saharan region coming from southern Africa, data from AFND [35, 48].

Region	HLA loci							
	A	B	C	DPA1	DPB1	DQA1	DPB1	DRB1
Australia	49	95	33	*	20	12	17	40
Europe	714	1121	387	16	137	47	89	602
N. Africa	982	1559	600	*	32	30	89	269
N. America	721	1166	390	7	74	29	93	574
N.E. Asia	262	477	131	12	78	47	57	318
S. Central America	121	288	59	12	78	28	60	549
S./S.E. Asia	407	731	227	10	99	21	64	280
Sub-Saharan Africa	154	313	94	12	87	23	48	220
W. Asia	215	366	167	*	29	21	57	138
Oceania	163	256	85	16	84	10	48	93
Southern Africa	131	291	54	*	21 <sup>a</sup>	8 <sup>a</sup>	20	58

N. Africa = North Africa, N. America = North America, N.E. Asia = North East Asia, S. Central America = South and Central America, S./S.E. Asia = South and South East Asia, and W. Asia = West Asia. \*No loci specific alleles were reported in this region in the AFND. <sup>a</sup>Alleles only reported in Zimbabwean black Shona population in the AFND.

### 3. Concluding Remarks

There is limited data on HLA diversity in southern Africa, with most having been generated from disease association studies and which is therefore not a true reflection of the general population. It is often difficult to assign causality of a specific HLA allele to an infection/condition, because of linkage disequilibrium and other factors such as selection pressure, which are dependent on the condition/infection and the other arms of the immune system which are HLA independent [96]. As evidenced by the HIV example, several HLA-B alleles have been associated with control of viremia [4, 97, 98] yet some individuals with these protective alleles develop AIDS (fail to control the virus) [99]. Recently Chen et al. showed that HLA B\*27 restricted CD8 T cells had variable viral replication inhibition capabilities in HIV controllers *versus* progressors due to a modulation by specific T cell receptor clonotypes [5]. There are few high resolution HLA datasets from southern African populations [1, 2, 35, 48] despite growing advancement in NGS HLA typing.

HLA diversity data forms the cornerstone of population-specific vaccine development, and taking into consideration the high disease burden in southern Africa, information of this nature is particularly important in this region [11]. This review highlights the paucity of information on HLA genotypic data and documents the extent of HLA diversity data from the southern African perspective based on the limited data available. This underpins an urgent need for HLA data from the general populations in this region and for studies which elucidate the extent of this diversity. There is a need to build an HLA diversity resource for southern Africa (or Africa as a whole) such as, for example, the HLA-net (a European network) [100] which focuses on HLA diversity and its applications in histocompatibility, transplantation, epidemiology, and population genetics. This network has

developed analysis pipelines and guidelines for HLA diversity data for mostly European populations [100, 101]. It is thus possible to build such a resource for the genetically diverse and disease burdened African continent to be used as a guideline for future studies including donor recruitment strategies [34], population studies [38, 89, 101], and disease association studies [6, 8, 77, 78].

Furthermore, advancement in HLA typing methods such as NGS will help to finely investigate HLA diversity, as previous strategies have targeted a few exons per locus thereby missing some of medically important variants outside the typed regions.

An understanding of HLA diversity will provide insight into allele frequency dependent selection fitness which varies between populations. This might help understand the high disease burden (especially with regard to HIV) and form the basis of vaccine development for the many infectious diseases as well as in the planning of vaccine clinical trials in the region. The paucity of HLA data from this region is a major hurdle in vaccine design [7]. Brumme et al. highlight, for example, the need to elucidate HLA-restricted CTL responses in HIV vaccine design [102]. HLA class II antigens presented to CD4<sup>+</sup> T cells induce B cells leading to an antigen specific humoral immune response [103]. HLA class II alleles have been associated with humoral immune response inducing vaccines for malaria [104], active anticancer immunotherapy [105], and HIV [106]. The combined use of HLA class II T helper (Th) epitopes with CD8+ CTL epitopes theoretically generates a high efficacy vaccine as reviewed by Xu et al. [105]. HLA diversity data might be useful in predicting the relative population coverage of a specific vaccine, adding knowledge on epitope targets for vaccines [107], mechanisms of immune evasion [108, 109], and evaluation of drug efficacy [110]. Posteraro et al. reviewed the significance of HLA diversity in efficacy of vaccination, highlighting the need to further



understand the link between genetic variation and immune responses [11].

It is generally easier to match donor-recipient pairs from populations with known HLA genotypes than in areas with information gaps [3], highlighting the need to understand population HLA diversity in order to improve donor-recipient matching. It is generally difficult to find a donor HLA match for patients of African descent owing to the paucity of Africans in global registries together with the occurrence of African specific alleles and or haplotypes and the high genetic diversity in these populations [60].

It is thus important to fully understand HLA diversity in the southern African context, to establish HLA-disease associations, to use this data for the informed design of population-specific vaccines against the many diseases, and to improve donor-recipient matching.

## Conflict of Interests

The authors declare no conflict of interests.

## Acknowledgments

This research as well as the publication thereof is the result of funding provided by the Medical Research Council of South Africa in terms of MRC's Flagships Awards Project SAMRC-RFA-UFSP-01-2013/STEM CELLS, the Institute for Cellular and Molecular Medicine of the University of Pretoria, and the National Research Foundation of South Africa.

## References

- [1] J. Robinson, J. A. Halliwell, H. McWilliam, R. Lopez, P. Parham, and S. G. E. Marsh, "The IMGT/HLA database," *Nucleic Acids Research*, vol. 41, no. 1, pp. D1222–D1227, 2013.
- [2] J. Robinson, J. A. Halliwell, J. D. Hayhurst, P. Flicek, P. Parham, and S. G. E. Marsh, "The IPD and IMGT/HLA database: allele variant databases," *Nucleic Acids Research*, vol. 43, no. 1, pp. D423–D431, 2015.
- [3] P. G. Beatty, K. M. Boucher, M. Mori, and E. L. Milford, "Probability of finding HLA-mismatched related or unrelated marrow or cord blood donors," *Human Immunology*, vol. 61, no. 8, pp. 834–840, 2000.
- [4] M. Carrington and S. J. O'Brien, "The influence of HLA genotype on AIDS," *Annual Review of Medicine*, vol. 54, pp. 535–551, 2003.
- [5] H. Chen, Z. M. Ndhlovu, D. Liu et al., "TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection," *Nature Immunology*, vol. 13, no. 7, pp. 691–700, 2012.
- [6] L. Z. Garamszegi, "Global distribution of malaria-resistant MHC-HLA alleles: the number and frequencies of alleles and malaria risk," *Malaria Journal*, vol. 13, article 349, pp. 1475–2875, 2014.
- [7] T. Ndung'u, S. Gaseitsiwe, E. Sepako et al., "Major histocompatibility complex class II (HLA-DRB and -DQB) allele frequencies in Botswana: association with human immunodeficiency virus type 1 infection," *Clinical and Diagnostic Laboratory Immunology*, vol. 12, no. 9, pp. 1020–1028, 2005.
- [8] M. Ramsay, "Africa: continent of genome contrasts with implications for biomedical research and health," *FEBS Letters*, vol. 586, no. 18, pp. 2813–2819, 2012.
- [9] C. Brander, N. Frahm, and B. D. Walker, "The challenges of host and viral diversity in HIV vaccine design (impedes of vaccine development owing to incomplete HLA information)," *Current Opinion in Immunology*, vol. 18, no. 4, pp. 430–437, 2006.
- [10] I. G. Ovsyannikova and G. A. Poland, "Vaccinomics: current findings, challenges and novel approaches for vaccine development," *The AAPS Journal*, vol. 13, no. 3, pp. 438–444, 2011.
- [11] WHO, *Global Health Report*, WHO, Geneva, Switzerland, 2013.
- [12] T. R. Disotell, "Archaic human genomics," *American Journal of Physical Anthropology*, vol. 55, pp. 24–39, 2012.
- [13] H. Kunze-Schumacher, R. Blasczyk, and C. Bade-Doeding, "Soluble HLA technology as a strategy to evaluate the impact of HLA mismatches," *Journal of Immunology Research*, vol. 2014, Article ID 246171, 8 pages, 2014.
- [14] S. J. Lee, J. Klein, M. Haagenson et al., "High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation," *Blood*, vol. 110, no. 13, pp. 4576–4583, 2007.
- [15] E. W. Petersdorf, "Optimal HLA matching in hematopoietic cell transplantation," *Current Opinion in Immunology*, vol. 20, no. 5, pp. 588–593, 2008.
- [16] D. Fürst, C. Müller, V. Vucinic et al., "High-resolution HLA matching in hematopoietic stem cell transplantation: a retrospective collaborative analysis," *Blood*, vol. 122, no. 18, pp. 3220–3229, 2013.
- [17] J. Pidala, T. Wang, M. Haagenson et al., "Amino acid substitution at peptide-binding pockets of HLA class I molecules increases risk of severe acute GVHD and mortality," *Blood*, vol. 122, no. 22, pp. 3651–3658, 2013.
- [18] M. Paximadis, T. Y. Mathebula, N. L. Gentle et al., "Human leukocyte antigen class I (A, B, C) and II (DRB1) diversity in the black and Caucasian South African population," *Human Immunology*, vol. 73, no. 1, pp. 80–92, 2012.
- [19] D. De Santis, D. Dinauer, J. Duke et al., "16th IHIW: review of HLA typing by NGS," *International Journal of Immunogenetics*, vol. 40, no. 1, pp. 72–76, 2013.
- [20] C. Lind, D. Ferriola, K. Mackiewicz et al., "Next-generation sequencing: the solution for high-resolution, unambiguous human leukocyte antigen typing," *Human Immunology*, vol. 71, no. 10, pp. 1033–1042, 2010.
- [21] C. Gabriel, D. Fürst, I. Faé et al., "HLA typing by next-generation sequencing—getting closer to reality," *Tissue Antigens*, vol. 83, no. 2, pp. 65–75, 2014.
- [22] H. Erlich, "HLA DNA typing: past, present, and future," *Tissue Antigens*, vol. 80, no. 1, pp. 1–11, 2012.
- [23] P. Parham and T. Ohta, "Population biology of antigen presentation by MHC class I molecules," *Science*, vol. 272, no. 5258, pp. 67–74, 1996.
- [24] J. R. Stewart and C. B. Stringer, "Human evolution out of Africa: the role of refugia and climate change," *Science*, vol. 335, no. 6074, pp. 1317–1321, 2012.
- [25] J. H. Relethford, "Genetic evidence and the modern human origins debate," *Heredity*, vol. 100, no. 6, pp. 555–563, 2008.
- [26] <http://www.broadinstitute.org>.
- [27] E. M. S. Belle and G. Barbujani, "Worldwide analysis of multiple microsatellites: language diversity has a detectable influence on DNA diversity," *The American Journal of Physical Anthropology*, vol. 133, no. 4, pp. 1137–1146, 2007.



- [28] M. Alessandrini, S. Asfaha, T. M. Dodgen, L. Warnich, and M. S. Pepper, "Cytochrome P450 pharmacogenetics in African populations," *Drug Metabolism Reviews*, vol. 45, no. 2, pp. 253–275, 2013.
- [29] E. de Wit, W. Delport, C. E. Rugamika et al., "Genome-wide analysis of the structure of the South African Coloured Population in the Western Cape," *Human Genetics*, vol. 128, no. 2, pp. 145–153, 2010.
- [30] [https://en.wikipedia.org/wiki/Bantu\\_peoples](https://en.wikipedia.org/wiki/Bantu_peoples).
- [31] D. Gurdasani, T. Carstensen, F. Tekola-Ayele et al., "The African Genome Variation Project shapes medical genetics in Africa," *Nature*, vol. 517, no. 7534, pp. 327–332, 2015.
- [32] <http://www.ebi.ac.uk/ipd/imgt/hla/stats.html>.
- [33] H. Chapel, M. Haeney, S. Misbah, and N. Snowden, *Essentials of Clinical Immunology*, edited by: John Wiley & Sons, Wiley-Blackwell, West Sussex, UK, 6th edition, 2014.
- [34] S. Y. Choo, "The HLA system: genetics, immunology, clinical testing, and clinical implications," *Yonsei Medical Journal*, vol. 48, no. 1, pp. 11–23, 2007.
- [35] F. F. Gonzalez-Galarza, S. Christmas, D. Middleton, and A. R. Jones, "Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations," *Nucleic Acids Research*, vol. 39, supplement 1, pp. D913–D919, 2011.
- [36] J. A. N. Klein and A. Sato, "The HLA system: first of two parts," *The New England Journal of Medicine*, vol. 343, no. 10, pp. 702–709, 2000.
- [37] D. Middleton, F. Williams, A. Meenagh et al., "Analysis of the distribution of HLA-A alleles in populations from five continents," *Human Immunology*, vol. 61, no. 10, pp. 1048–1052, 2000.
- [38] A. Sanchez-Mazas, M. Fernandez-Viña, D. Middleton et al., "Immunogenetics as a tool in anthropological studies," *Immunology*, vol. 133, no. 2, pp. 143–164, 2011.
- [39] M. Carrington, G. W. Nelson, M. P. Martin et al., "HLA and HIV-1: heterozygote advantage and B\*35-Cw\*04 disadvantage," *Science*, vol. 283, no. 5408, pp. 1748–1752, 1999.
- [40] H. V. Siddle, A. Kreiss, M. D. B. Eldridge et al., "Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 41, pp. 16221–16226, 2007.
- [41] F. Prugnolle, A. Manica, M. Charpentier, J. F. Guégan, V. Guernier, and F. Balloux, "Pathogen-driven selection and worldwide HLA class I diversity," *Current Biology*, vol. 15, no. 11, pp. 1022–1027, 2005.
- [42] L. J. L. Handley, A. Manica, J. Goudet, and F. Balloux, "Going the distance: human population genetics in a clinal world," *Trends in Genetics*, vol. 23, no. 9, pp. 432–439, 2007.
- [43] G. S. Cooke and A. V. S. Hill, "Genetics of susceptibility to human infectious disease," *Nature Reviews Genetics*, vol. 2, no. 12, pp. 967–977, 2001.
- [44] L. H. Miller, "Impact of malaria on genetic polymorphism and genetic diseases in Africans and African Americans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 7, pp. 2415–2419, 1994.
- [45] A. L. Hughes and M. Nei, "Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, no. 3, pp. 958–962, 1989.
- [46] M. Yeager and A. L. Hughes, "Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution," *Immunological Reviews*, vol. 167, pp. 45–58, 1999.
- [47] A. L. Hughes, T. Ota, and M. Nei, "Positive Darwinian selection promotes charge profile diversity in the antigen-binding cleft of class I major-histocompatibility-complex molecules," *Molecular Biology and Evolution*, vol. 7, no. 6, pp. 515–524, 1990.
- [48] F. F. Gonzalez-Galarza, L. Y. C. Takeshita, E. J. M. Santos et al., "Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations," *Nucleic Acids Research*, vol. 43, no. 1, pp. D784–D788, 2015.
- [49] A. Belicha-Villanueva, J. Blickwedehl, S. McEvoy, M. Golding, S. O. Gollnick, and N. Bangia, "What is the role of alternate splicing in antigen presentation by major histocompatibility complex class I molecules?" *Immunologic Research*, vol. 46, no. 1–3, pp. 32–44, 2010.
- [50] M. S. Krangel, "Secretion of HLA-A and -B antigens via an alternative RNA splicing pathway," *The Journal of Experimental Medicine*, vol. 163, no. 5, pp. 1173–1190, 1986.
- [51] L. Norgaard, L. Fugger, H. O. Madsen, and A. Svejgaard, "Identification of 4 different alternatively spliced HLA-A transcripts," *Tissue Antigens*, vol. 54, no. 4, pp. 370–378, 1999.
- [52] P. Paul, F. Adrian Cabestre, E. C. Ibrahim et al., "Identification of HLA-G7 as a new splice variant of the HLA-G mRNA and expression of soluble HLA-G5, -G6, and -G7 transcripts in human transfected cells," *Human Immunology*, vol. 61, no. 11, pp. 1138–1149, 2000.
- [53] M. Carrington, "Recombination within the human MHC," *Immunological Reviews*, vol. 167, pp. 245–256, 1999.
- [54] M. Adamek, C. Klages, M. Bauer et al., "Seven novel HLA alleles reflect different mechanisms involved in the evolution of HLA diversity: description of the new alleles and review of the literature," *Human Immunology*, vol. 76, no. 1, pp. 30–35, 2015.
- [55] A. A. A. Assane, G. M. Fabricio-Silva, J. Cardoso-Oliveira et al., "Human leukocyte antigen-A, -B, and -DRB1 allele and haplotype frequencies in the Mozambican population: a blood donor-based population study," *Human Immunology*, vol. 71, no. 10, pp. 1027–1032, 2010.
- [56] M. G. Hammond and D. Anley, "Tamil from Natal Province, South Africa," in *Proceedings of the 13th International Histocompatibility Workshop*, International Histocompatibility Working Group Press, Seattle, Wash, USA, 2006.
- [57] M. G. Hammond, D. Middleton, and D. Anley, "Zulu from Natal province, South Africa," in *Proceedings of the 13th International Histocompatibility Workshop and Conference*, IHWG Press, Seattle, Wash, USA, 2006.
- [58] V. Coetzee, L. Barrett, J. M. Greeff, S. P. Henzi, D. I. Perrett, and A. A. Wadee, "Common HLA alleles associated with health, but not with facial attractiveness," *PLoS ONE*, vol. 2, no. 7, article e640, 2007.
- [59] M. Tikly, A. Rands, N. McHugh, P. Wordsworth, and K. Welsh, "Human leukocyte antigen class II associations with systemic sclerosis in South Africans," *Tissue Antigens*, vol. 63, no. 5, pp. 487–490, 2004.
- [60] K. Cao, A. M. Moormann, K. E. Lyke et al., "Differentiation between African populations is evidenced by the diversity of alleles and haplotypes of HLA class I loci," *Tissue Antigens*, vol. 63, no. 4, pp. 293–325, 2004.
- [61] K. Cao, C. Masaberg, J. Yu, D. L. Mann, and M. A. Fernández-Viña, "Zambians from Lusaka, Zambia," in *Proceedings of the 13th International Histocompatibility Workshop and Conference: Immunobiology of the Human MHC*, J. A. Hansen, Ed., vol. 1, pp. 589–590, IHWG Press, Seattle, Wash, USA, 2006.

- [62] L. Louie, K. Mather, D. Meyer et al., "Shona from Harare, Zimbabwe," in *Proceedings of the 13th International Histocompatibility Workshop and Conference: Immunobiology of the Human MHC*, J. A. Hansen, Ed., vol. 1, pp. 587–589, IHWG Press, Seattle, Wash, USA, 2006.
- [63] P. J. Martin, "The role of donor lymphoid cells in allogeneic marrow engraftment," *Bone Marrow Transplantation*, vol. 6, no. 5, pp. 283–289, 1990.
- [64] J. J. Cornelissen and B. Löwenberg, "Developments in T-cell depletion of allogeneic stem cell grafts," *Current Opinion in Hematology*, vol. 7, no. 6, pp. 348–352, 2000.
- [65] L. Moretta, F. Locatelli, D. Pende, E. Marcenaro, M. C. Mingari, and A. Moretta, "Killer Ig-like receptor-mediated control of natural killer cell alloreactivity in haploidentical hematopoietic stem cell transplantation," *Blood*, vol. 117, no. 3, pp. 764–771, 2011.
- [66] A. Bishara, D. De Santis, C. C. Witt et al., "The beneficial role of inhibitory KIR genes of HLA class I NK epitopes in haploidentically mismatched stem cell allografts may be masked by residual donor-alloreactive T cells causing GVHD," *Tissue Antigens*, vol. 63, no. 3, pp. 204–211, 2004.
- [67] K. C. Hsu, T. Gooley, M. Malkki et al., "KIR ligands and prediction of relapse after unrelated donor hematopoietic cell transplantation for hematologic malignancy," *Biology of Blood and Marrow Transplantation*, vol. 12, no. 8, pp. 828–836, 2006.
- [68] J. M. Venstrom, T. A. Gooley, S. Spellman et al., "Donor activating KIR3DS1 is associated with decreased acute GVHD in unrelated allogeneic hematopoietic stem cell transplantation," *Blood*, vol. 115, no. 15, pp. 3162–3165, 2010.
- [69] M. R. Kelher, T. Masuno, E. E. Moore et al., "Plasma from stored packed red blood cells and MHC class I antibodies causes acute lung injury in a 2-event in vivo rat model," *Blood*, vol. 113, no. 9, pp. 2079–2087, 2009.
- [70] A. Davoren, O. P. Smith, C. A. Barnes, E. Lawlor, R. G. Evans, and G. F. Lucas, "Case report: four donors with granulocyte-specific or HLA class I antibodies implicated in a case of transfusion-related acute lung injury (TRALI)," *Immunohematology*, vol. 17, no. 4, pp. 117–121, 2001.
- [71] U. J. H. Sachs, W. Wasel, B. Bayat et al., "Mechanism of transfusion-related acute lung injury induced by HLA class II antibodies," *Blood*, vol. 117, no. 2, pp. 669–677, 2011.
- [72] S. Saito, M. Ota, Y. Komatsu et al., "Serologic analysis of three cases of neonatal alloimmune thrombocytopenia associated with HLA antibodies," *Transfusion*, vol. 43, no. 7, pp. 908–917, 2003.
- [73] J. Trowsdale and J. C. Knight, "Major histocompatibility complex genomics and human disease," *Annual Review of Genomics and Human Genetics*, vol. 14, pp. 301–323, 2013.
- [74] H. Hendel, S. Caillat-Zucman, H. Lebuane et al., "New class I and II HLA alleles strongly associated with opposite patterns of progression to AIDS," *Journal of Immunology*, vol. 162, no. 11, pp. 6942–6946, 1999.
- [75] C. Rohowsky-Kochan, J. Skurnick, D. Molinaro, and D. Louria, "HLA antigens associated with susceptibility/resistance to HIV-1 infection," *Human Immunology*, vol. 59, no. 12, pp. 802–815, 1998.
- [76] M. Tshabalala, G. D. Morse, and L. S. Zijenah, "HLA genetic polymorphisms: role in HIV-1 susceptibility, disease progression and treatment outcomes," *Retrovirology: Research and Treatment*, vol. 5, pp. 1–8, 2013.
- [77] J.-J. Yim and P. Selvaraj, "Genetic susceptibility in tuberculosis," *Respirology*, vol. 15, no. 2, pp. 241–256, 2010.
- [78] A. Sanchez, K. E. Wagoner, and P. E. Rollin, "Sequence-based human leukocyte antigen-B typing of patients infected with Ebola virus in Uganda in 2000: identification of alleles associated with fatal and nonfatal disease outcomes," *Journal of Infectious Diseases*, vol. 196, no. 2, pp. S329–S336, 2007.
- [79] L. Kauppi, A. Sajantila, and A. J. Jeffreys, "Recombination hotspots rather than population history dominate linkage disequilibrium in the MHC class II region," *Human Molecular Genetics*, vol. 12, no. 1, pp. 33–40, 2003.
- [80] P. I. W. de Bakker and S. Raychaudhuri, "Interrogating the major histocompatibility complex with high-throughput genomics," *Human Molecular Genetics*, vol. 21, no. R1, pp. R29–R36, 2012.
- [81] L. R. Botigue, B. Henn, S. Gravel et al., "Gene flow from North Africa contributes to differential human genetic diversity in southern Europe," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 29, pp. 11791–11796, 2013.
- [82] The International HapMap Consortium, "The international HapMap project," *Nature*, vol. 426, no. 6968, pp. 789–796, 2003.
- [83] G. R. Abecasis, A. Auton, L. D. Brooks et al., "An integrated map of genetic variation from 1,092 human genomes," *Nature*, vol. 491, no. 7422, pp. 56–65, 2012.
- [84] J. P. Jarvis, L. B. Scheinfeldt, S. Soi et al., "Patterns of ancestry, signatures of natural selection, and genetic association with stature in Western African pygmies," *PLoS Genetics*, vol. 8, no. 4, Article ID e1002641, 2012.
- [85] C. M. Schlebusch, P. Skoglund, P. Sjödin et al., "Genomic variation in seven Khoe-San groups reveals adaptation and complex African history," *Science*, vol. 338, no. 6105, pp. 374–379, 2012.
- [86] S. A. Tishkoff, F. A. Reed, F. R. Friedlaender et al., "The genetic structure and history of Africans and African Americans," *Science*, vol. 324, no. 5930, pp. 1035–1044, 2009.
- [87] S. Buhler and A. Sanchez-Mazas, "HLA DNA sequence variation among human populations: molecular signatures of demographic and selective events," *PLoS ONE*, vol. 6, no. 2, Article ID e14643, 2011.
- [88] <http://hla.alleles.org/nomenclature/stats.html>.
- [89] A. Sanchez-Mazas and E. Thorsby, "HLA in anthropology: the enigma of Easter Island," *Clinical Transplants*, pp. 167–173, 2013.
- [90] O. D. Solberg, S. J. Mack, A. K. Lancaster et al., "Balancing selection and heterogeneity across the classical human leukocyte antigen loci: a meta-analytic review of 497 population studies," *Human Immunology*, vol. 69, no. 7, pp. 443–464, 2008.
- [91] F. T. Ayele, E. Hailu, C. Finan et al., "Prediction of HLA class II alleles using SNPs in an African population," *PLoS ONE*, vol. 7, no. 6, Article ID e40206, 2012.
- [92] B. L. Shepherd, R. Ferrand, S. Munyati et al., "HLA correlates of long-term survival in vertically infected HIV-1-positive adolescents in Harare, Zimbabwe," *AIDS Research and Human Retroviruses*, vol. 31, no. 5, pp. 504–507, 2015.
- [93] K. R. Alkharsah, M. Dedicoat, R. Blasczyk, R. Newton, and T. F. Schulz, "Influence of HLA alleles on shedding of kaposi sarcoma-associated herpesvirus in saliva in an African population," *Journal of Infectious Diseases*, vol. 195, no. 6, pp. 809–816, 2007.
- [94] D. F. Carr, M. Chaponda, A. L. Jorgensen et al., "Association of human leukocyte antigen alleles and nevirapine hypersensitivity in a Malawian HIV-infected population," *Clinical Infectious Diseases*, vol. 56, no. 9, pp. 1330–1339, 2013.

- [95] O. O. Yang, M. J. Lewis, E. F. Reed et al., "Human leukocyte antigen class I haplotypes of human immunodeficiency virus-1-infected persons on Likoma Island, Malawi," *Human Immunology*, vol. 72, no. 10, pp. 877–880, 2011.
- [96] J. P. A. Ioannidis, G. Thomas, and M. J. Daly, "Validating, augmenting and refining genome-wide association signals," *Nature Reviews Genetics*, vol. 10, no. 5, pp. 318–329, 2009.
- [97] R. A. Kaslow, M. Carrington, R. Apple et al., "Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection," *Nature Medicine*, vol. 2, no. 4, pp. 405–411, 1996.
- [98] S. A. Migueles, M. S. Sabbaghian, W. L. Shupert et al., "HLA B\*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 6, pp. 2709–2714, 2000.
- [99] F. Pereyra, M. M. Addo, D. E. Kaufmann et al., "Genetic and immunologic heterogeneity among persons who control HIV infection in the absence of therapy," *Journal of Infectious Diseases*, vol. 197, no. 4, pp. 563–571, 2008.
- [100] J. M. Nunes, S. Buhler, D. Roessli, and A. Sanchez-Mazas, "The HLA-net GENE[RATE] pipeline for effective HLA data analysis and its application to 145 population samples from Europe and neighbouring areas," *Tissue antigens*, vol. 83, no. 5, pp. 307–323, 2014.
- [101] A. Sanchez-Mazas, B. Vidan-Jeras, J. M. Nunes et al., "Strategies to work with HLA data in human populations for histocompatibility, clinical transplantation, epidemiology and population genetics: HLA-NET methodological recommendations," *International Journal of Immunogenetics*, vol. 39, no. 6, pp. 459–476, 2012.
- [102] Z. L. Brumme, D. R. Chopera, and M. A. Brockman, "Modulation of HIV reservoirs by host HLA: bridging the gap between vaccine and cure," *Current Opinion in Virology*, vol. 2, no. 5, pp. 599–605, 2012.
- [103] P. J. Delves and I. M. Roitt, "The Immune system—second of two parts," *The New England Journal of Medicine*, vol. 343, no. 2, pp. 108–117, 2000.
- [104] H. A. F. Stephens, A. E. Brown, D. Chandanayingyong et al., "The presence of the HLA class II allele DPB1\*0501 in ethnic Thais correlates with an enhanced vaccine-induced antibody response to a malaria sporozoite antigen," *European Journal of Immunology*, vol. 25, no. 11, pp. 3142–3147, 1995.
- [105] M. Xu, N. L. Kallinteris, and E. von Hofe, "CD4+ T-cell activation for immunotherapy of malignancies using Ii-Key/MHC class II epitope hybrid vaccines," *Vaccine*, vol. 30, no. 18, pp. 2805–2810, 2012.
- [106] R. Paris, S. Bejrachandra, P. Thongcharoen et al., "HLA class II restriction of HIV-1 clade-specific neutralizing antibody responses in ethnic Thai recipients of the RV144 prime-boost vaccine combination of ALVAC-HIV and AIDSVAX B/E," *Vaccine*, vol. 30, no. 5, pp. 832–836, 2012.
- [107] L. Zhao, M. Zhang, and H. Cong, "Advances in the study of HLA-restricted epitope vaccines," *Human Vaccines and Immunotherapeutics*, vol. 9, no. 12, pp. 2566–2577, 2013.
- [108] Y. Yagita, N. Kuse, K. Kuroki et al., "Distinct HIV-1 escape patterns selected by cytotoxic T cells with identical epitope specificity," *Journal of Virology*, vol. 87, no. 4, pp. 2253–2263, 2013.
- [109] J. M. Carlson, A. Q. Le, A. Shahid, and Z. L. Brumme, "HIV-1 adaptation to HLA: a window into virus–host immune interactions," *Trends in Microbiology*, vol. 23, no. 4, pp. 212–224, 2015.
- [110] S. Paul, R. V. Kolla, J. Sidney et al., "Evaluating the immunogenicity of protein drugs by applying in vitro MHC binding data and the immune epitope database and analysis resource," *Clinical and Developmental Immunology*, vol. 2013, Article ID 467852, 7 pages, 2013.
- [111] B. Posteraro, R. Pastorino, P. Di Giannantonio et al., "The link between genetic variation and variability in vaccine responses: systematic review and meta-analyses," *Vaccine*, vol. 32, no. 15, pp. 1661–1669, 2014.