Review Article

Human Leukocyte Antigen (HLA) System: Genetics and Association with Bacterial and Viral Infections

Sadeep Medhasi 🌱 and Narisara Chantratita 🌱

Department of Microbiology and Immunology, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

Correspondence should be addressed to Narisara Chantratita; narisara@tropmedres.ac

Received 10 January 2022; Accepted 8 May 2022; Published 26 May 2022

Academic Editor: Cinzia Ciccacci

Copyright © 2022 Sadeep Medhasi and Narisara Chantratita. 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.

The human leukocyte antigen (HLA) system is one of the most crucial host factors influencing disease progression in bacterial and viral infections. This review provides the basic concepts of the structure and function of HLA molecules in humans. Here, we highlight the main findings on the associations between HLA class I and class II alleles and susceptibility to important infectious diseases such as tuberculosis, leprosy, melioidosis, *Staphylococcus aureus* infection, human immunodeficiency virus infection, coronavirus disease 2019, hepatitis B, and hepatitis C in populations worldwide. Finally, we discuss challenges in HLA typing to predict disease outcomes in clinical implementation. Evaluation of the impact of HLA variants on the outcome of bacterial and viral infections would improve the understanding of pathogenesis and identify those at risk from infectious diseases in distinct populations and may improve the individual treatment.

1. Introduction

The cell-mediated adaptive immune response is regulated by the major histocompatibility complex (MHC) or human leukocyte antigen (HLA) in humans [1]. HLA molecules are cell surface glycoproteins whose primary function is to present endogenous and exogenous antigens to T lymphocytes for recognition and response [2]. The HLA molecules that present antigen to T lymphocytes are divided into two main classes: HLA class I and HLA class II molecules. HLA class I molecules play an essential role in the immune defense against intracellular pathogens, whereas HLA class II molecules are predominantly involved in displaying peptides from extracellular pathogens [3]. The HLA region is highly polymorphic, and polymorphisms in the HLA molecules result in variability in amino acid sequences of HLA molecules and thus affect the peptide binding specificity [4]. HLA molecules encoded by different alleles have different peptide-binding repertoires [5]. The polymorphisms in the HLA locus contribute to the genetic diversity of humans and the differences in susceptibility to diseases among genetically distinct groups, thus offering evolutionary advantages of a diverse immunological response to a wide range of infectious pathogens [6]. The associations between HLA alleles and susceptibility to or protection from infectious diseases have been well documented. However, the molecular mechanism underlying host HLA function to infection remains far from understood. Infectious disease continues to affect poor and marginalized populations; therefore, it is essential to utilize the increasing knowledge and technological advances in HLA typing to study the pathogenesis and development of novel therapeutic targets in infectious diseases of public health concerns.

Genetic variations at the loci encoding HLA genes are associated with susceptibility or protection to infectious diseases. Genetic studies have found an association between the HLA alleles or haplotypes and bacterial infectious diseases, including tuberculosis, leprosy, and melioidosis [7–9]. Identifying risk and protective HLA alleles will provide critical insights into the mechanisms that influence the pathogenesis of infections and protection. HLA typing can identify associations between HLA alleles and infections in an individual [10]. Patients exhibit different immune responses to bacterial and viral infections, and HLA molecules play an
essential role in regulating the host’s immune response. Therefore, the reported HLA alleles contributing to the susceptibility or protective effect to bacterial and viral infections will aid in elucidating the immunological mechanisms in disease outcomes [11, 12].

This review will provide the basic concepts of HLA and the current status of the HLA associations with bacterial and viral infections across world populations. This review primarily focuses on predisposing risk and protective HLA alleles among several populations in major infectious diseases, including bacterial infections (tuberculosis, leprosy, melioidosis, and Staphylococcus aureus infections) and viral infections (human immunodeficiency virus (HIV) infection, coronavirus disease 2019 (COVID-19), hepatitis B, and hepatitis C). Many studies on these infections have shown HLA-associated susceptibility in many populations, but the association with melioidosis and S. aureus infections is less characterized. We also discuss the challenges of complicating disease outcome prediction through HLA typing. A deeper understanding of the genetic basis of susceptibility to these infections will aid in understanding the pathogenesis of the disease, identify new molecular targets for prophylactic and therapeutic interventions, and develop a potential tool to identify those at risk of rapid disease progression.

2. Structure and Function of Human Leukocyte Antigen

The HLA molecule is the name for the human MHC, which orchestrates immune regulation by antigen presentation to T cells [13]. The HLA system resides in a region that spans approximately 4,000 kilobases (kb) of DNA on the short arm of chromosome 6 (6p21). This region encodes three major classes of proteins, HLA class I (HLA-A, HLA-B, and HLA-C), class II (HLA-DP, HLA-DQ, and HLA-DR), and class III (components of the complement system, 21-hydroxylase, heat shock protein, and tumor necrosis factors) (Figure 1) [14, 15].

HLA class I molecules are present as transmembrane glycoproteins on the surface of nearly all nucleated cells. These molecules present intracellular self- or non-self-antigens to CD8⁺ cytotoxic T cell receptors and killer cell immunoglobulin-like receptors (KIR) [16]. HLA class I molecules consist of two heterodimer polypeptide chains, a heavy α chain, and a lighter β2-microglobulin chain. The α chain has three extracellular domains (α1, α2, and α3), a transmembrane region, and a C-terminal cytoplasmic tail. The two domains, α1 and α2, fold to form a peptide-binding groove and are referred to as the recognition region. The β2-microglobulin chain is primarily associated with the α3 domain and is responsible for HLA stability (Figure 2(a)) [17, 18].

Unlike HLA class I and HLA class II regions, whose functions in the immune response are well defined, the HLA class III region encodes for various inflammatory molecules, complement, and heat shock protein [11]. The HLA class III region spans 700 kb of DNA and is located between the centromeric class II (HLA-DRA) and the telomeric class I regions (MICB) (Figure 1) [20].

3. HLA Nomenclature

The WHO Nomenclature Committee for Factors of the HLA System is responsible for the formal naming of HLA alleles and has reported the names through two websites, Immuno Polymorphism Database-International ImMunoGeneTics project/HLA (IPD-IMGT/HLA) database (https://www.ebi.ac.uk/ipd/imgt/hla/) and HLA Nomenclature (http://hla.alleles.org/nomenclature/naming.html) [21, 22]. The current HLA nomenclature system uses a unique number
corresponding to up to four sets of digits separated by colons (Figure 3). The HLA-prefix signifies the human MHC gene complex. The next portion after the HLA-prefix indicates the specific HLA genomic region. The first two digits of the number (field 1) show the allele group (or allele family). The second field provides the specific HLA allele (HLA protein). The third field names the alleles that differ only by synonymous nucleotide substitutions within the coding region. The fourth field names the alleles that vary only by sequence polymorphisms in introns, 3′-untranslated regions, and 5′-untranslated regions. Last is the suffix consisting of a letter that denotes alleles with changes in expression levels of the HLA protein products. The suffix "N" is used for null alleles with no HLA protein expression. Other letters have been used to designate an allele to indicate its expression status: L: low expression, S: secreted, and Q: questionable. A standardized HLA nomenclature has contributed to the understanding of the HLA system and proved to be an essential resource to address HLA typing ambiguity in the clinical applications of HLA [23].

4. Genetic Association between HLA Loci and Infectious Diseases

The HLA family of genes is one of the most polymorphic genes in the human genome [24]. The IPD-IMGT/HLA Database is a repository for the variant sequences of HLA alleles. As of April 2022, the IPD-IMGT/HLA Database has reported 33,490 HLA alleles. Of the 24,308 HLA class I
### 4.1. HLA Associations with Tuberculosis

Tuberculosis (TB), caused by Mycobacterium tuberculosis (M. tuberculosis), is an infectious disease posing a significant public health threat primarily in low- and middle-income countries [25]. The World Health Organization reported an estimated 10.0 million TB cases, 1.2 million TB deaths among HIV-negative people, and an additional 208,000 TB deaths among HIV-positive people in 2019 [26]. M. tuberculosis can modulate the HLA class II pathway by inhibiting phagosome maturation and thus preventing the formation of bacterial peptide-MHC-II (HLA class II) complexes and subsequent T cell responses to bacterial antigens [27]. M. tuberculosis also inhibits MHC-II expression and antigen processing resulting in decreased recognition by T cells [28].

Several genetic polymorphisms of HLA have been implicated in individuals’ genetic susceptibility to tuberculosis in distinct populations (Table 2). A study on 31 pulmonary tuberculosis patients in Poland showed a higher frequency of HLA-DRB1*16 in patients when compared to the 58 healthy controls. In comparison, the frequency of the HLA-DRB1*13 allele was significantly lower in the patient group than in the healthy controls [29]. In Iranian patients with pulmonary tuberculosis, HLA-DRB1*07 and HLA-DQA1*01:01 alleles appeared to be the risk alleles, and HLA-DQA1*03:01 and HLA-DQA1*05:01 alleles were the protective alleles [30]. Wamala et al. investigated HLA class II gene polymorphisms in susceptibility to pulmonary tuberculosis in Uganda and observed that the HLA-DQB1*03:03 allele was associated with resistance to pulmonary tuberculosis [31]. In the study performed in South India, the frequencies of HLA-DRB1*15:01 and HLA-DQB1*06:01 alleles were higher in pulmonary tuberculosis patients than in the control group. In contrast, the frequency of the HLA-DPB1*04 allele was highly prevalent among the control group and was deemed to be a protective allele against pulmonary tuberculosis [32]. A study by Sveinbjörnsson et al. in Icelanders demonstrated HLA-DQA1*03 (represented by p.Ala210Thr) and a noncoding variant, rs557011, located between HLA-DQA1 and HLA-DRB1 contributing to genetic susceptibility to tuberculosis [33]. They also demonstrated the association of rs9271378 with a reduced risk of pulmonary TB in Icelanders demonstrated HLA-DQA1*03:01 and HLA-DQA1*05:01 alleles were the protective alleles [30].

### Table 2: Associations between HLA and tuberculosis.

<table>
<thead>
<tr>
<th>Population</th>
<th>Study design</th>
<th>Sample size</th>
<th>Serotype, allele, SNP, or haplotype</th>
<th>Type of association</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polish</td>
<td>Case-control</td>
<td>31 pulmonary TB patients and 58 healthy controls</td>
<td>HLA-DRB1<em>16, HLA-DRB1</em>13, HLA-DRB1<em>07 and HLA-DQA1</em>01:01, HLA-DQA1<em>03:01 and HLA-DQA1</em>05:01</td>
<td>Susceptibility, Protection</td>
<td>[29]</td>
</tr>
<tr>
<td>Iranian</td>
<td>Case-control</td>
<td>40 pulmonary TB patients and 100 healthy controls</td>
<td>HLA-DQB1*03:03</td>
<td>Protection</td>
<td>[31]</td>
</tr>
<tr>
<td>Uganda</td>
<td>Case-control</td>
<td>43 pulmonary TB patients and 42 healthy controls</td>
<td>HLA-DRB1<em>15:01 and HLA-DQB1</em>06:01, HLA-DPB1<em>04, rs557011[T] located between HLA-DQA1 and HLA-DRB1, HLA-DQA1</em>03, rs9271378[G] located between HLA-DQA1 and HLA-DRB1</td>
<td>Susceptibility to pulmonary TB and M. tuberculosis infection, Reduced risk of pulmonary TB</td>
<td>[32]</td>
</tr>
<tr>
<td>Icelander</td>
<td>Case-control</td>
<td>3,686 pulmonary TB patients, 14,723 patients with M. tuberculosis infection, 8,162 patients with any other forms of TB, and 277,643 healthy controls</td>
<td>HLA-DRB1<em>16, HLA-DPB1</em>04, rs557011[T] located between HLA-DQA1 and HLA-DRB1, HLA-DQA1*03, rs9271378[G] located between HLA-DQA1 and HLA-DRB1</td>
<td>Susceptibility to pulmonary TB and M. tuberculosis infection</td>
<td>[33]</td>
</tr>
<tr>
<td>Han Chinese</td>
<td>Case-control</td>
<td>4,310 TB patients and 6,386 healthy controls</td>
<td>HLA-DRB5*rs41553512</td>
<td>Susceptibility</td>
<td>[34]</td>
</tr>
<tr>
<td>Thai</td>
<td>Case-control</td>
<td>682 TB patients and 836 healthy controls</td>
<td>HLA-DRB1<em>09:01 and HLA-DQB1</em>03:03</td>
<td>Susceptibility to TB caused by modern M. tuberculosis strains</td>
<td>[35]</td>
</tr>
</tbody>
</table>
Strain-based association analysis between HLA class II genes and tuberculosis in the Thai population identified a significant association of HLA-DRB1*09:01 and HLA-DQB1*03:03 with a modern strain of M. tuberculosis (absence of M. tuberculosis-specific deletion 1 (TbD1) region) [35].

4.2. HLA Associations with Leprosy. Leprosy is a chronic infectious disease caused by Mycobacterium leprae. Based on clinical, histopathological, microbiological, and immunological features, Ridley and Jopling classified the leprosy spectrum into five groups: tuberculoid (TT), borderline-tuberculoid (BT), borderline-borderline (BB), borderline-lepromatous (BL), and lepromatous (LL) [36].

Both HLA class I and class II genes have been implicated in susceptibility to leprosy and its subtypes in different populations (Table 3). In a study in India, the frequencies of HLA-A*02:06, HLA-A*11:02, HLA-B*40:16, HLA-B*51:10, HLA-Cw*04:07, and HLA-Cw*07:03 alleles were significantly higher in leprosy patients compared to healthy controls, while the frequencies of HLA-A*0101, HLA-Cw*04011, and HLA-Cw*06:02 alleles were markedly lower in leprosy patients compared to healthy controls [37]. Another study of genetic susceptibility to leprosy in India found rs1071630 located in HLA-DQA1 and rs9270650 in HLA-DRB1 associated with susceptibility to leprosy [38]. A genome-wide association study in 706 patients with leprosy and 1225 unaffected controls in Han Chinese found a single-nucleotide polymorphism (SNP) rs602875 at the HLA-DR-DQ locus associated with susceptibility to leprosy [39]. HLA-DR molecules activate T cells by presenting M. leprae peptide antigens to CD4+ T cells and activate various pathways. Anomalies in those pathways could cause HLA-associated leprosy [39]. The HLA-DRB1*15 allele was associated with leprosy, while HLA-DRB1*09 was significantly protective against leprosy in Han Chinese [40]. A meta-analysis by Zhang et al. identified HLA-DQA1*03:03 and HLA-C*07:06 as causal variants to leprosy susceptibility in the Han Chinese population [41]. In the association study between HLA-DRB1 and leprosy among Brazilian and Vietnamese people, the HLA-DRB1*04 allele was associated with protection against leprosy, and the HLA-DRB1*10 allele was found to be associated with susceptibility to leprosy [42]. A recent study was conducted to investigate the association of HLA class I and II genes with leprosy in a Brazilian population. The study identified the association of HLA-C*12 and HLA-DPB1*105 with susceptibility to leprosy, while HLA-C*08, HLA-DPB1*04 and HLA-DPB1*18 were protective against multibacillary leprosy.

<table>
<thead>
<tr>
<th>Population</th>
<th>Study design</th>
<th>Sample size</th>
<th>Serotype, allele, SNP, or haplotype</th>
<th>Type of association</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian</td>
<td>Case-control</td>
<td>32 leprosy patients and 67 healthy controls</td>
<td>HLA-A<em>02:06, HLA-A</em>11:02, HLA-B<em>40:16, HLA-B</em>51:10, and HLA-Cw<em>04011, and HLA-Cw</em>06:02</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Indian</td>
<td>Case-control and family-based</td>
<td>258 leprosy patients, 161 families, and 300 healthy controls</td>
<td>HLA-DQA1 rs1071630 and HLA-DRB1 rs9270650</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Han Chinese</td>
<td>Case-control</td>
<td>3,254 leprosy patients and 5,955 healthy controls</td>
<td>HLA-DR-DQ rs602875</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Han Chinese</td>
<td>Case-control</td>
<td>305 leprosy patients and 527 healthy controls</td>
<td>HLA-DRB1<em>15, HLA-DRB1</em>09</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Han Chinese</td>
<td>Meta-analysis</td>
<td>Four imputed data sets</td>
<td>HLA-DQA1<em>03:03 and HLA-C</em>08:01</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Brazilian</td>
<td>Case-control</td>
<td>578 leprosy patients and 691 healthy controls</td>
<td>HLA-DRB1<em>10, HLA-DRB1</em>04</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Brazilian</td>
<td>Case-control</td>
<td>411 leprosy patients and 415 healthy controls</td>
<td>HLA-C<em>12 and HLA-DPB1</em>105</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Vietnamese</td>
<td>Family-based</td>
<td>194 families</td>
<td>HLA-DRB1<em>10, HLA-DRB1</em>04</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Vietnamese</td>
<td>Case-control</td>
<td>687 leprosy patients and 468 healthy controls</td>
<td>HLA-DQA1<em>01:05 and HLA-DRB1</em>10:01</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Argentinian</td>
<td>Case-control</td>
<td>142 leprosy patients and 162 healthy controls</td>
<td>HLA-DRB1<em>14:01 and HLA-DRB1</em>14:06</td>
<td>Susceptibility</td>
<td></td>
</tr>
<tr>
<td>Taiwanese</td>
<td>Case-control</td>
<td>65 multibacillary leprosy patients and 190 healthy controls</td>
<td>HLA-DRB1*04:05</td>
<td>Protection against multibacillary leprosy</td>
<td></td>
</tr>
</tbody>
</table>
were protective against leprosy [43]. Dallmann-Sauer et al. performed next-generation sequencing to genotype three HLA class I and eight class II genes in 1,155 individuals from a Vietnamese leprosy case-control sample. The HLA-DQA1 * 01:05 and HLA-DRB1 * 10:01 alleles in complete linkage disequilibrium (LD) were associated with leprosy, whereas the HLA-C * 07:06 allele was shown to be protective against leprosy in the Vietnamese population [44]. A study of the association of HLA-DRB1 alleles in 71 leprosy patients and 81 healthy controls in Argentina found a higher frequency of HLA-DRB1 * 14:01 and HLA-DRB1 * 14:06 alleles in leprosy patients compared to controls. In contrast, the frequency of HLA-DRB1 * 08:08 and HLA-DRB1 * 11:03 was highly prevalent among the healthy controls compared to the leprosy patients hence indicating resistance to leprosy [45]. Interestingly, a study in Taiwan assessing the leprosy association with HLA class I and class II alleles found a protective effect of HLA-DRB1 * 04:05 on multibacillary leprosy [46].

4.3. HLA Associations with Melioidosis. Melioidosis is an infectious disease caused by the Gram-negative bacillus *Burkholderia pseudomallei*. Melioidosis is widely endemic in Southeast Asia, especially in Thailand, and northern Australia [47]. The disease is highly seasonal, and the organism is commonly found in soil and water in the endemic areas [48]. Risk factors include diabetes mellitus, chronic kidney disease, chronic lung disease, alcohol abuse, and steroid therapy [49, 50]. Diabetes mellitus is the major underlying risk factor occurring in 60-75% of patients diagnosed with melioidosis [50–52]. Clues to the mechanisms involved between diabetes mellitus and melioidosis might be explained by the role of HLA alleles in both diabetes mellitus and melioidosis. HLA class II alleles have been documented to have prominent effects on diabetes mellitus in distinct populations [53–57].

Studies in Thailand have reported the risk of HLA alleles associated with melioidosis (Table 4). In 1998, Dharakul et al. investigated the associations between HLA class II alleles and melioidosis in 79 melioidosis patients and 105 healthy controls in Northeast Thailand [8]. The study demonstrated a significant association between the DRB1 * 16:02 allele and the susceptibility to melioidosis in the Thai population. In addition, associations were observed with the DRB1 * 16:02 allele for severe melioidosis and septicemic melioidosis when various clinical groups of melioidosis patients were compared with healthy controls. In another study in Northeast Thailand, HLA-B * 46 and HLA-C * 01 (HLA class I alleles) were associated with increased mortality from acute melioidosis compared to the survived patients from acute melioidosis [58].

In melioidosis, suppressed HLA-DR expression on classical monocytes was associated with poor outcomes [59]. A transcriptomic analysis of changes in gene expression of nonsurvivors from melioidosis in Northeast Thailand found the downregulation of HLA class II genes, including HLA-DPB1, HLA-TRA, HLA-DOA, and HLA-DOB [60]. The role of HLA in the pathogenesis and poor prognosis of melioidosis is not fully understood. Reynolds et al. demonstrated a strong binding affinity of alkyl hydroperoxide reductase (AhpC), a highly dominant *B. pseudomallei* antigen, with HLA-DR alleles, HLA-DR1, HLA-DR3, HLA-DR4, HLA-DR7, HLA-DR9, HLA-DR11, HLA-DR13, HLA-DR15:01, and HLA-DR15:02, and the HLA-DQ alleles, HLA-DQB1 * 06:02 and HLA-DQB1 * 03:02 [61]. In addition, the study also reported that among patients with acute melioidosis in Northeast Thailand, survival was associated with a strong HLA class II-restricted T cell response to AhpC.

4.4. HLA Associations with Staphylococcus aureus Infections. *S. aureus* is an opportunistic pathogen and a leading cause of morbidity and mortality in hospital and community settings [62]. The *S. aureus* superantigens, toxic shock syndrome toxin-1 (TSST-1) and *S. aureus* enterotoxin B (SEB), bind to the HLA class II molecule HLA-DR1 [63]. Genetic variations within the host are associated with susceptibility to *S. aureus* infections suggesting why one-third of humans are known to be colonized with *S. aureus* in their anterior nares, but most avoid clinically significant *S. aureus* infections [64].

Several HLA alleles are proposed as susceptibility factors to *S. aureus* infection (Table 5). A GWAS conducted to identify specific genetic variants that underlie susceptibility to infections caused by *S. aureus* in white subjects reported three SNPs, rs4321864 located in the HLA-DRA gene and rs115231074 and rs35079132 located in HLA-DRB1 genes, associated with *S. aureus* infection. The study also found an association between HLA-DRB1 * 04 serotype and *S. aureus* infection [65]. Cyr et al. evaluated the role of genetic variation on susceptibility to *S. aureus* bacteremia in African Americans. They found the genetic association of one region on chromosome 6 in the HLA class II region with susceptibility *S. aureus* bacteremia [66].

4.5. HLA Associations with HIV Infection. HIV infection is a major global public health issue with a devastating impact on social and economic indicators [67–69]. HLAs play a complex role in immunomodulation during HIV infection, and variations at the HLA class I locus have been linked to the efficiency of CD8+ T cell control of viremia [70].

Polymorphisms within HLA class I and II loci have been identified as the host genetic modifier of HIV disease progression in several populations (Table 6). In the Argentinian population, the frequency of the HLA-B * 39 allele was significantly higher in HIV-1-positive subjects than in controls, whereas the HLA-B * 44 allele was absent among the HIV-1-positive subjects [71]. Claiborne et al. identified four HLA class I alleles (B * 14:01, B * 57, B * 58:01, and B * 81) and two HLA class II alleles (DQB1 * 02 and DRB1 * 15) associated with the protection from rapid CD4+ T cell decline without controlling early plasma viral load in a Zambian early infection cohort [72]. A GWAS in HIV-1 infected Caucasian subjects showed HLA-B * 57:01 rs2395029 and HLA-C rs9264942 associated with HIV-1 disease progression [73]. Analysis of the HLA-B allele frequencies among HIV-1-infected individuals classified as rapid progressors (RPs), typical progressors (TPs), and long-term nonprogressors (LTNPs) in the Brazilian population revealed the higher frequency of the HLA-B * 52 allele in the LTNP group than...
in either the TP group or the RP group, and thus, the presence of the HLA-B*52 allele is favorable to slow AIDS progression [74]. A study involving treatment-naïve patients with chronic HIV-1 infection from (i) Warsaw, Poland; (ii) Athens, Greece; (iii) Mexico City, Mexico; (iv) Bonn, Germany; (v) Boston, MA; (vi) Barcelona, Spain; and (vii) Thames Valley, UK, suggested that HLA-B*27:02 was associated with slower progression to HIV disease [75]. In a study among the HIV clade B-infected ART-naïve individuals from Mexico and Central America, several HLA alleles were identified as protective (A*03:01, B*15:39, B*27:05, B*39:02, B*57:01/02/03, and B*58:01) and risk (A*68:03/05, B*15:30, B*35:02, B*35:12/14, B*39:01/06, B*39:05, and B*40:01) factors for disease progression [76].

4.6. HLA Associations with COVID-19. COVID-19 is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It is a worldwide pandemic with 198,778,175 confirmed cases, including 4,235,559 deaths (as of 3 August 2021) [77]. Patients with severe COVID-19 have been found to exhibit immune dysregulation characterized by IL-6-mediated low HLA-DR expression [78].

Several HLA polymorphisms are associated with susceptibility and severity to COVID-19 in different populations (Table 7). In a study among individuals of European descent experiencing variable clinical outcomes following COVID-19 infection, the frequency of HLA-DRB1*04:01 was higher among asymptomatic COVID-19 patients than the severe COVID-19 patients and suggested the protective effects of the HLA-DRB1*04:01 allele against developing severe complications from COVID-19 [79]. The DRB1*09:01 allele was associated with risk for severe COVID-19 in Japanese [80]. Novelli et al. analyzed the HLA allele frequency distribution in Italian COVID-19 patients to identify potential markers of susceptibility to the disease and observed that

### Table 4: Associations between HLA and melioidosis.

<table>
<thead>
<tr>
<th>Population</th>
<th>Study design</th>
<th>Sample size</th>
<th>Serotype, allele, SNP, or haplotype</th>
<th>Type of association</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thai</td>
<td>Case-control</td>
<td>79 melioidosis patients and 105 healthy controls</td>
<td>HLA-DRB1 * 16:02</td>
<td>Susceptibility and poor prognosis</td>
<td>[8]</td>
</tr>
<tr>
<td>Thai</td>
<td>Case-control</td>
<td>183 acute melioidosis patients and 21 healthy controls</td>
<td>HLA-B * 46 and HLA-C * 01</td>
<td>Increased mortality</td>
<td>[58]</td>
</tr>
</tbody>
</table>

### Table 5: Associations between HLA and S. aureus infections.

<table>
<thead>
<tr>
<th>Population</th>
<th>Study design</th>
<th>Sample size</th>
<th>Serotype, allele, SNP, or haplotype</th>
<th>Type of association</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>Case-control</td>
<td>4,701 culture-confirmed S. aureus cases and 45,344 healthy controls</td>
<td>HLA-DRA, rs4321864, HLA-DRB1, rs115231074 and rs35079132, and HLA-DRB1 * 04</td>
<td>Susceptibility</td>
<td>[65]</td>
</tr>
<tr>
<td>African American</td>
<td>Case-control</td>
<td>390 cases and 175 healthy controls</td>
<td>HLA-B * 57, B * 58:01 and B * 81, and HLA-DQB1 * 02 and DRB1 * 15</td>
<td>Susceptibility</td>
<td>[66]</td>
</tr>
</tbody>
</table>

### Table 6: Associations between HLA and HIV infection.

<table>
<thead>
<tr>
<th>Population</th>
<th>Study design</th>
<th>Sample size</th>
<th>Serotype, allele, SNP, or haplotype</th>
<th>Type of association</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentinian</td>
<td>Case-control</td>
<td>56 HIV-1-positive patients and 56 healthy individuals</td>
<td>HLA-B * 39 and HLA-B * 44</td>
<td>Susceptibility Protection</td>
<td>[71]</td>
</tr>
<tr>
<td>Zambian</td>
<td>Longitudinal</td>
<td>127 subjects with acute HIV-1 infections</td>
<td>HLA-B * 14:01, B * 57, B * 58:01 and B * 81, and HLA-DQB1 * 02 and DRB1 * 15</td>
<td>Slow disease progression</td>
<td>[72]</td>
</tr>
<tr>
<td>Caucasian</td>
<td>Longitudinal</td>
<td>2,554 HIV-1 infected subjects</td>
<td>HLA-B * 5701, rs2395029 and HLA-C, rs9264942</td>
<td>Accelerated disease progression</td>
<td>[73]</td>
</tr>
<tr>
<td>Brazilian</td>
<td>Retrospective observational</td>
<td>218 HIV-1 infected subjects</td>
<td>HLA-B * 52</td>
<td>Slow disease progression</td>
<td>[74]</td>
</tr>
<tr>
<td>Mexican and Central American</td>
<td>Multicenter cross-sectional</td>
<td>3,213 HIV clade B-infected patients</td>
<td>HLA-A * 68:03/05, HLA-B * 15:30, B * 35:02, B * 35:12/14, B * 39:01/06, B * 39:05, and B * 40:01</td>
<td>Accelerated disease progression</td>
<td>[76]</td>
</tr>
</tbody>
</table>

References:

[71] [72] [73] [74] [75] [76] [77] [78] [79] [80]
the frequencies of HLA-DRB1 *15:01, HLA-DQB1 *06:02, and HLA-B *27:07 alleles were higher among the severe affected COVID-19 patients compared to healthy controls [81]. In a comprehensive in silico analysis, HLA-B*46:01 had the fewest predicted binding peptides for SARS-CoV-2, indicating that the individuals with this allele may be particularly vulnerable to COVID-19 [82]. Another in silico analysis found the association between HLA-A *02:01 and an increased risk for COVID-19, and HLA-A *02:01 was predicted to present the lower SARS-CoV-2 antigens and subsequent lower T cell-mediated antiviral responses compared to HLA-A *11:01 or HLA-A *24:02 alleles [83]. In a study conducted by Shkurnikov et al. in deceased patients with COVID-19 in Russia, HLA-A *01:01 was associated with early COVID-19 deaths among the high-risk patients, and HLA-A *02:01 and HLA-A *03:01 alleles were associated with early COVID-19 deaths among the low-risk patients [84].

4.7. HLA Associations with Hepatitis B. Hepatitis B is a significant public health problem putting people at high risk of death from cirrhosis and liver cancer [85]. Hepatitis B is caused by the hepatitis B virus (HBV). HBV-specific CD8+ cytotoxic T lymphocytes play a critical role in viral clearance and liver injury, and HLA polymorphisms have been reported to alter CD8+ cytotoxic T lymphocyte responses [86].

Multiple population association studies have provided evidence of an association between HLA locus variations and hepatitis B virus infection (Table 8). A study comparing the distribution of HLA alleles between persistent and transient HBV infection in children and adults in the Gambia found HLA-DRB1 *13:02 associated with protection against persistent HBV infection among children and adults [87]. A Chinese study by Fan et al. showed an association between the HLA-DQ rs9275319C allele and decreased HBV infection risk and an increased HBV clearance [88]. Another Chinese study showed that HLA-DQB1 *06:03 protected against HBV infection [89]. An association analysis performed among the Turkish population revealed the association of the HLA-DPB1 rs9277555A allele with the risk of persistent HBV infection [90]. A Caucasian study showed that HLA-A *03:01 was associated with viral clearance, and HLA-B *8 was associated with viral persistence [91]. Al-Qahtani et al. demonstrated an association of HLA-DQ alleles (rs2856718A and rs9275572A) and HLA-DP alleles (rs3077G and rs9277535G) with HBV infection in Saudi Arabian patients [92]. Nishida et al. showed that HLA-DQB1 *06:01 was associated with chronic HBV infection in Japanese patients [93].

4.8. HLA Associations with Hepatitis C. Hepatitis C virus (HCV) infection is a significant cause of acute and chronic hepatitis. Chronic hepatitis leads to liver cirrhosis and hepatocellular carcinoma (HCC) [94, 95]. HCV persistence or clearance is proposed to depend on the response of the HLA class I-restricted HCV-specific CD8+ cytotoxic T cell-mediated lysis of virus-infected host cells [96].

Depending on ethnicity, a significant association has been suggested between HLA alleles and HCV persistence or spontaneous clearance (Table 9). In the Thai population, the frequency of HLA-DRB1 *03:01 and HLA-DQB1 *02:01 was higher in the persistent HCV infection group than in the transient HCV infection group, revealing their susceptibility effect on persistent HCV infection [97]. Genotyping a large multiracial cohort of US women to evaluate associations between HLA alleles and HCV viremia indicated some HLA alleles (B *57:01, B *57:03, Cw *01:02, and DRB1 *01:01) were associated with the absence of HCV RNA. At the same time, the presence of HCV RNA was observed for HLA-DRB1 *03:01 [98]. Huang et al. reported the association of HLA-A *02:01 and HLA-DRB1 *11:01 with HCV spontaneous clearance in the Chinese population [99]. In Egyptian HCV patients and their families or close household contacts, HLA-DRB1 allele associations with HCV were reported (DRB1 *03:01:01 and DRB1 *13:01:01 alleles and the risk of progression to chronic hepatitis C infection and DRB1 *04:01:01, DRB1 *04:05:01, DRB1 *07:01:01, and DRB1 *11:01:01 and protection against HCV infection) [100].

5. Challenges in HLA Typing to Predict Disease Outcomes

5.1. HLA Diversity. The genetic diversity of HLA within each population can be explained or measured by allelic richness (ar) and the expected heterozygosity (H). The allelic richness of a population at a particular locus is the expected number of alleles present in the population at that locus [101]. The expected heterozygosity is defined as the average proportion
of heterozygotes per locus in a randomly mating population [102]. Pathogen richness is the number of pathogens within a defined geographical region [103]. Sanchez-Mazas et al. reported a significant positive correlation between genetic diversity and pathogen richness at HLA-A and HLA-B and a significant negative correlation at HLA-DQB1 [104].

Identifying the most clinically relevant HLA variant is necessary for facilitating improvements in the diagnosis and treatment of human disease. The identifiable HLA variants provide opportunities to refine medical management to optimize patient health and medical outcomes. However, genetic diversity within and between populations poses a

<table>
<thead>
<tr>
<th>Table 8: Associations between HLA and hepatitis B.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
</tr>
<tr>
<td>Gambian</td>
</tr>
<tr>
<td>Chinese</td>
</tr>
<tr>
<td>Chinese</td>
</tr>
<tr>
<td>Turkish</td>
</tr>
<tr>
<td>Caucasian</td>
</tr>
<tr>
<td>Saudi Arabian</td>
</tr>
<tr>
<td>Japanese</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 9: Associations between HLA and hepatitis C.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
</tr>
<tr>
<td>Thai</td>
</tr>
<tr>
<td>Multiracial US women</td>
</tr>
<tr>
<td>Chinese</td>
</tr>
<tr>
<td>Egyptian</td>
</tr>
</tbody>
</table>
challenge in HLA genomics to become a standard component of health care. Many rare HLA variants are identified, and they are likely to contribute to interindividual differences in risk or protection to disease. HLA data from various populations have been collected, but populations of African and Asian descent have limited representation to provide insight into HLA disease associations [105]. HLA genetic diversity among Europeans is well documented [106–108]. Hurley et al. recently reported global frequencies of common, intermediate, and well-documented HLA alleles and highlighted the HLA diversity in world populations [109].

5.2. HLA Genotyping. Different DNA-based molecular techniques represent the modern methods used for HLA typing in clinical applications. Depending on their power to discriminate between HLA alleles, DNA-based HLA typing methods are characterized by low resolution (result at the level of the digits composing the first field in the HLA nomenclature) and high resolution (result at the level of four digits) [110]. The most widely used DNA-based methods in conjunction with PCR for HLA typing include sequence-specific oligonucleotides (SSO), sequence-specific primers (SSP), and sequence-based typing (SBT) [111]. In PCR-SSO, PCR products are hybridized into sequence-specific oligonucleotide probes. In PCR-SSP, primers complementary to particular HLA allele sequences are used, and amplification with sequence-specific primers yields only a product if the target sequences are present in the DNA sample. In SBT, HLA genetic regions are amplified by PCR using locus-specific primers, followed by direct sequencing of the PCR products [112].

Although SSO and SSP methods are widely used, they are not practical and capable of detecting all known HLA polymorphisms and novel HLA alleles [113]. SSO and SSP typing methods struggle to resolve the major allele groups [114]. The SBT provides high-resolution HLA genotyping and can identify new alleles. While SBT allows for a detailed interpretation of HLA alleles, it has limitations, including time-consuming protocols, low throughput, and ambiguities in HLA typing results [115–117]. SNP-based HLA typing on microarray produces a high-resolution HLA type but has not been used in clinical typing due to its tendency to miss several HLA variants [117].

Next-generation sequencing- (NGS-) based HLA typing methods allow high-throughput sequencing, massively parallel analysis, and high-resolution HLA typing with minimal ambiguity [118, 119]. NGS-based HLA typing has been implemented with better accuracy compared to traditional HLA typing methods [120]. NGS-based HLA typing approaches are promising but are not yet ready to be implemented in routine clinical care settings due to the higher cost and complex protocol [111].

5.3. Implementing Therapeutic Approaches Using Genomic Knowledge of Specific Targets and Their Roles in Disease. Utilization of HLA typing can identify alleles associated with disease risks and improve clinical outcomes. However, genomic literacy among clinicians is a low to moderate level which presents a challenge in adopting genomic services by clinicians in clinical practices [121]. To overcome this challenge, high-quality results in HLA association studies must be disseminated among the health workforce, including the policymakers and the personnel on the ground. Awareness about the implications of HLA typing into mainstream clinical practice must be raised by educating the health workforce. There is a need for more reporting on the clinical validity and clinical utility of genetic testing used for screening of risk and protective HLA alleles in diseases [122, 123]. Integrating genomic services utilizing HLA-specific tests face challenges and barriers to widespread adoption. These include the lack of a single standard approach to achieve HLA typing by NGS data, integrating electronic health records (EHR) of genomic results and clinical decision support (CDS), ensuring confidentiality for patients and families and lack of reimbursement [124–127].

6. Conclusions

This article summarized the findings from association studies of HLA variants with bacterial and viral infections. It is important to note that there were no overlaps in the HLA variants associated with susceptibility or protection to infectious diseases in populations worldwide. Despite the evidence of association of HLA variants with disease susceptibility in our review, a consistent genetic HLA locus has not been demonstrated within the population.

The global pathogens will expand to new geographical locations with genetic mutations in the future [128]. New sequencing techniques that allow faster, cheaper, and less intensive sequencing need to be developed to effectively implement HLA typing in routine clinical care. Advancements in HLA typing technologies are enabling a more accurate linking of HLA genotypes to disease outcomes. NGS technologies will provide a deeper insight into disease mechanisms and biological processes of HLA. Combining HLA sequencing with the expression levels of HLA genes can provide a clearer picture of the role of HLA genes in the pathogenesis of diseases. To understand HLA evolution, HLA expression levels must be integrated with information on HLA genetic variation in diverse populations. At the same time, large data sets need to be generated to reinterpret information to understand pathogen spread in the future [129].

In light of the current evidence, the goal of HLA typing should aid in the realization of “precision medicine” that will benefit patients in diverse populations. Significant research on HLA and infectious diseases is needed in subjects of all ethnic origins to achieve optimum therapeutic outcomes for broader clinical implementation.

Data Availability

The data availability is not declared. Our manuscript is a review paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.
Acknowledgments

This work was supported by the Mahidol University Post-doctoral Fellowship 2021.

References


[76] H. Valenzuela-Ponce, S. Alva-Hernández, D. Garrido-Rodríguez et al., “Novel HLA class I associations with HIV-1 con-


requires HLA genotyping-based association analysis,” *Scientific Reports*, vol. 6, no. 1, p. 24767, 2016.


[103] L. A. Baxter-Lowe, “The changing landscape of HLA typing: understanding how and when HLA typing data can be used with confidence from bench to bedside,” *Human Immunology*, vol. 82, no. 7, pp. 466–477, 2021.


