An Update on the Genetic Polymorphism of \textit{HLA-B*27} With 213 Alleles Encompassing 160 Subtypes (and Still Counting)

Muhammad A. Khan

Abstract

\textit{Purpose of Review} This publication updates an earlier review on the ever increasing knowledge about genetic polymorphism of \textit{HLA-B*27} and discusses its clinical relevance.

\textit{Recent Findings} As of January 1, 2017, there are 213 known alleles of \textit{HLA-B*27} at nucleotide sequence level, while at the translated protein level, there are 160 known subtypes based on one or more amino acid sequence differences. Some of these subtypes exhibit differential association with ankylosing spondylitis, and there may even be some level of hierarchy in this regard. On the other hand, \textit{HLA-B*27} has a protective effect against HCV, and this effect is also influenced by some of the subtypes of \textit{HLA-B*27}. This may have important implications for designing anti-viral vaccines for global population and also for developing individualized treatments and vaccines.

\textit{Summary} Disease association and disease protective roles of \textit{HLA-B*27} suggest a common ground, i.e., promoting a more pronounced immune/inflammatory response for effective clearance of some pathogens, but that might, on the other hand, lead to autoimmunity and tissue injury in some circumstances.

\textbf{Keywords} \textit{HLA-B27} · \textit{HLA-B*27} · \textit{HLA-B*27:05} · HLA alleles · Subtypes · Polymorphism · Ankylosing spondylitis · Spondyloarthritis · Homozygosity · HCV · HIV

Human leucocyte antigens (HLA) represent a group of highly polymorphic genes that reside in the major histocompatibility complex (MHC) that is located within the \textit{6p21.3} region on the short arm of chromosome 6, and encodes many of the proteins of the immune system [1]. These include HLA-class I genes that are co-dominantly expressed on the cell surface presenting intracellularly derived peptides to CD8 positive T cells. HLA-\textit{B*27} belongs to a family of closely related cell surface proteins encoded in the HLA-B locus. It was discovered as a serological specificity in 1969 and found to be distributed worldwide but with variable prevalence. Four years later it was found to show a remarkable association with ankylosing spondylitis (AS) and related forms of spondyloarthritis (SpA) [2, 3]. It was later observed that the strength of this association varies not only for different forms of SpA but also among some of the ethnic and racial groups in the world [2, 3, 4, 5]. Substantial evidence exists that strongly favors its direct role in genetic susceptibility to AS/SpA, although the underlying molecular basis has yet to be identified [2, 4, 6]. But it is not a prerequisite for occurrence of AS since this disease also affects individuals who lack this gene, and there are now 48 genetic loci known to be associated with increased risk for AS [2, 4, 5, 6, 7].

\textit{HLA-B*27} gene shows a high degree of polymorphism as it encompasses an ever increasing number of alleles as well as “subtypes” (i.e., specific HLA proteins). Please note that the name of a gene or allele is written in italics, while normal text is used when using it for mentioning its encoded protein. I hereby provide an update on my last review on this subject [2]. But before discussing it further, it is imperative to first explain the current HLA nomenclature, especially for those readers who may be unfamiliar or are somewhat confused by this revised nomenclature.
The HLA Nomenclature

The naming of HLA genes, allele sequences, and their quality control is the responsibility of the Word Health Organization (WHO) Nomenclature Committee for Factors of the HLA System. It is now an established procedure for authors to submit their report of new HLA sequences directly to the IPD-IMGT/HLA Database for checking and assignment of an official name prior to publication [1]. This has resolved the old problems associated with renaming published sequences and the confusion of multiple names for the same sequence.

All HLA alleles are initially identified by a letter(s) indicating its genetic location in the MHC region, such as locus A, B, and C, followed by a number that is assigned in order of discovery, for example HLA-A1, HLA-A2, HLA-A9, HLA-B7, HLA-B8, and HLA-B27. Such HLA nomenclature was used years ago when typing was done by serologic method. HLA-B27 is thus a serologic specificity, but serologic methods resolve only a tiny fraction of all known alleles. Therefore, HLA typing at the DNA level has become the method of choice, and it is also more robust and reproducible.

It was decided to insert an asterisk (*) after the letter that assigns the locus and followed by two digit specificity of the molecule in order to indicate that the typing done at the DNA level. So, as an example, HLA-A2, HLA-B7, and HLA-B27 have become HLA-A*02, HLA-B*07, and HLA-B*27, respectively. This is defined as a “low resolution typing.”

An increasing number of HLA alleles are now being discovered, and therefore, the “HLA Extended Allele Nomenclature” system is now used in which a second set of digits (preceded by a colon that is used as a field separator) is added to identify them, such as HLA-B*27:05 and HLA-B*27:161. This second set of digits is assigned in the order in which the DNA sequences have been determined. Such alleles containing the two sets of digits (that are separated by a colon) must have one or more nucleotide substitutions that change the amino acid sequence of their encoded proteins. Therefore, the second of these two sets of digits indicates specific HLA proteins or subtypes. However, a very few of the alleles are not expressed at all (i.e., they do not produce any protein) and are thus called “null” alleles and carry the suffix “N”; for example HLA-B*27:59.N.

Alleles that differ only by synonymous nucleotide substitutions (also called “silent” or “non-coding” substitutions) within the coding sequence are distinguished by assigning a third set of two digits. Thus HLA-B*27:05:32 is one such example where: HLA denotes the human MHC, and it is followed by a hyphen (−) and then the letter B that denotes the HLA-B locus. It is followed by an asterisk (*) that denotes molecular typing, and then the first set of two digits (27) to indicate the antigen (or allele) family. The second and the third sets of digits are separated by colons (that act as field separators); the second set :05 indicates an allele level “subtype” (i.e., a specific HLA protein), and the third set :32 indicates a silent or “non-coding” substitution within the coding region.

Those alleles that only differ by sequence differences that are located outside of the coding region are distinguished by adding a fourth set of digits. An example of such an allele is HLA-B*27:02:01:02 where :01 indicates a silent nucleotide substitution within the coding region, and the last set :02 indicates a difference in the non-coding region.

A suffix “S” indicates that the allele specifies a protein which is expressed as a soluble “Secreted” molecule that is not expressed on the cell surface, e.g., HLA-B*44:02:01:02S. Another suffix “L” is used to indicate an allele that has “Low” level of expression on cell surface, e.g., HLA-A*30:14 L. A “Q” suffix is used when the expression of an allele is “Questionable.” A suffix “A” indicates an “Aberrant” expression where there is some doubt as to whether a protein is actually expressed, and a “C” suffix is assigned to alleles that produce proteins that are present in the “Cytoplasm” and not on the cell surface.

HLA-B*27 Polymorphism

As of January 1, 2017, there are 213 known alleles of HLA-B*27 that have been defined at the level of nucleotide sequence. This list of alleles is obtained from the IPD-IMGT/HLA Database (The Immuno Polymorphism Database (IPD), International ImMunoGeneTics (IMGT) HLA Database) Release 3.27.0 (2017-01) [1]. Many of the mutations are located within introns and thus are silent or they occur in exons but do not cause amino acid changes. Therefore, at the translated protein level, there are 160 known subtypes of HLA-B*27 based on one or more amino acid sequence differences. These 160 subtypes can be encompassed by the numbering system HLA-B*27:01 to HLA-B*27:161 because one of the assignments—HLA-B*27:22—was subsequently withdrawn when it was found to be identical to HLA-B*27:06.

I have summarized this long list into a concise table (Table 1). The subtypes of HLA-B*27 are listed in the left column and that listing is continued on to the third column from the left. The second and fourth columns list alleles belonging to some of these subtypes; eg, HLA-B*27:05 encompasses 31 alleles: HLA-B*27:05:02 to HLA-B*27:05:32. Because the protein encoded by all of them is identical to the one encoded by HLA-B*27:05. It is worth mentioning that HLA-B*27:05:02 is the most widely distributed HLA-B*27 allele, and is probably the ancestral allele from which the others evolved. Please note that HLA-B*27:05:01 is not listed as it was withdrawn when subsequently found to be identical to HLA-B*27:05.
Differential Association of HLA-B*27 Subtypes With AS

The HLA-B*27 subtypes show an extremely varied racial and ethnic prevalence throughout the world [4, 5]. HLA-B*27:05 is the most widely distributed subtype and has been the subject of most of the clinical studies dealing with its association with AS. Among the other relatively common disease-associated subtypes are HLA-B*27:02 (Mediterranean populations) and HLA-B*27:04 (Chinese and other Asian populations).

### Table 1: Remarkable polymorphism of HLA-B*27

<table>
<thead>
<tr>
<th>Subtypes</th>
<th>Alleles</th>
<th>Subtypes</th>
<th>Alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>B*27:01</td>
<td>B*27:02:01</td>
<td>B*27:05</td>
<td>B*27:05:01</td>
</tr>
<tr>
<td>B*27:02</td>
<td>B*27:02:01:02</td>
<td>B*27:50</td>
<td>B*27:50:01</td>
</tr>
<tr>
<td></td>
<td>B*27:02:02</td>
<td>B*27:50:02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B*27:02:03</td>
<td>B*27:51</td>
<td></td>
</tr>
<tr>
<td>B*27:03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B*27:04</td>
<td>B*27:04:01</td>
<td>B*27:90</td>
<td>B*27:90:01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B*27:90:04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B*27:90:01</td>
</tr>
<tr>
<td>B*27:05</td>
<td>B*27:05:02</td>
<td>B*27:91</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B*27:96:02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B*27:96:01</td>
</tr>
<tr>
<td>B*27:06</td>
<td></td>
<td></td>
<td>B*27:96:02</td>
</tr>
<tr>
<td>B*27:07</td>
<td>B*27:07:01</td>
<td>B*27:97</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B*27:97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B*27:161</td>
</tr>
<tr>
<td>B*27:08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B*27:21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table is derived from the list of HLA-B*27 alleles maintained by the IPD-IMGT/HLA Database Release 3.27.0 (2017-01). The 160 subtypes: HLA-B*27:01 to HLA-B*27:161, are listed in the left column and are continued on to the third column from the left. Please note that there is no HLA-B*27:22 because it was withdrawn when it was found to be identical to HLA-B*27:06. The second and fourth columns list alleles belonging to some of the subtypes; eg, HLA-B*27:05 encompasses 31 alleles: HLA-B*27:05:02 to HLA-B*27:05:32. Please note that HLA-B*27:05:01 is not listed as it was withdrawn when it was subsequently found to be identical to HLA-B*27:05, HLA-B*27:13 differs from HLA-B*27:05:02 only in the leader segment of the gene, which is not part of the expressed product at the cell surface. So the HLA-B*27 molecule encoded by these two alleles is identical [Khan MA, Ball EJ. Genetic aspects of ankylosing spondylitis. Best Pract Res Clin Rheumatol. 2002;16:675–90.]. The “null” alleles are HLA-B*27:59 N, HLA-B*27:64 N, HLA-B*27:65 N, HLA-B*27:66 N, and HLA-B*27:94 N. Updated, with permission, from: Khan MA: ANKYLOSING SPONDYLITIS-AXIAL SPONDYLOARTHRITIS. PCI (Professional Communications, Inc). West Islip, NY. 2016. pp. 1–333

**Differential Association of HLA-B*27 Subtypes With AS**

The HLA-B*27 subtypes show an extremely varied racial and ethnic prevalence throughout the world [4, 5].
HLA-B*27:06, a subtype prevalent in Southeast Asian populations, has been reported to lack association with AS [8–11]. It differs from the disease-associated subtype HLA-B*27:04 only at residues at position 114 (His to Asp) and 116 (Asp to Tyr) that form part of the F pocket at the bottom of the groove where the bound antigenic peptide sits [2••].

Another subtype—HLA-B*27:09—that occurs mostly among Italians that live on the island of Sardinia has also been reported to lack association with AS (12). And, most interestingly, it differs from the most frequent and disease-associated HLA-B*27:05 subtype by only a single substitution (His vs. Asp) at position 116 [2••]. Many investigators have been exploring the effects of these sequence variations on peptide-binding specificities of HLA-B*27:06 and HLA-B*27:09 in the hope to uncover the underlying molecular basis of disease association [12, 13, 14•, 15, 16, 17•].

Schittenhelm et al. [17••] have reported 26 peptides that are presented in lower abundance by HLA-B*27:06 and HLA-B*27:09 compared with disease-associated HLA-B*27 subtypes, in their pursuit of the putative “arthritogenic peptide(s).” They indicate that the differences in charge and size that accompany the His-to-Asp substitution at residue 114, and Asp to Tyr substitution at residue 116 (in case of HLA-B*27:06), exclude the acceptance of the sought-after “arthritogenic peptide.” Such studies, it is hoped, will also help define which alleles, and therefore which polymorphic determinants, predispose to the disease. Dimerization of HLA-B*27 heavy chain during assembly process has been implicated in disease pathogenesis [6••, 18]. Guiliano et al. [19•] have just reported that the above mentioned differences in the residues within the F pocket between the disease-associated and non-disease-associated subtypes of HLA-B*27 influence this assembly process and heavy chain dimerization.

Differences among some of the HLA-B*27 subtypes have been reported when it comes to their strength of disease association; and there may even be some level of hierarchy in this regard among some of these subtypes. For example, HLA-B*27:05 and HLA-B*27:02, the relatively common subtypes among Caucasian populations, seem to confer equal susceptibility to AS. On the other hand, among Chinese populations, HLA-B*27:04 to Thr in HLA-B*27:15), seems to be associated with a much younger age of onset than that observed among patients possessing HLA-B*27:04 or HLA-B*27:05 [34, 42]. Few studies have examined the relationship between HLA-B*27 homozygotes and risk for AS and association with clinical manifestations. Khan et al. [43] first reported increased risk for AS, and it was finally confirmed 28 years later [44]. HLA-B*27 homozygosity, however, has no influence on clinical manifestations, radiographic damage and functional disability in AS [43, 45, 46].

**Protective Role of HLA-B*27 in HCV and HIV Infections**

HLA-B*27 is associated with high rates of spontaneous CD8+ T cell-mediated clearance of hepatitis C virus (HCV) genotype1 infection, and it also has a protective role in human immunodeficiency virus (HIV) infection [47, 48]. These are two of the most highly variable viral pathogens for which effective vaccines are still lacking [49••, 50]. In HIV infection, HLA-B*27 is associated with a low viral load, slow CD4+ decline, and the delayed onset of AIDS [49••, 50].

This efficacy of HLA-B*27 in responding to the HIV infection is associated with its specificity for a highly conserved CD8+ epitope of Gag protein of the virus and which requires a complicated pathway of viral escape to evade this protective response. However, a compensatory mutation of that conserved epitope ultimately leads to viral escape and the onset of AIDS in HLA-B*27+ individuals [49••, 50•].
The protective effect against HCV is linked to dominant HLA-B*27 restricted epitopes of the viral protein NS5B. HLA-B*27-restricted CD8+ T cells target three NS5B epitopes; and one of them is NS5B2820. This epitope is immunodominant exclusively in the context of HLA-B*27:02, but not in the context of HLA-B*27:05, the most common subtype worldwide [4•, 52••]. HLA-B*27:02 is mostly prevalent in the Middle East and North Africa, but has only 2 to 10% prevalence among northern European populations [4•, 5, 53, 54]. This differential epitope targeting and immunodominance between these two common HLA-B*27 subtypes has important implications for designing anti-viral vaccines for global population and also for developing individualized treatments and vaccines. In light of these data pertaining to HCV infection, it is possible that a similar differential viral epitope targeting and immunodominance may also exist with regard to HIV infection among some of the HLA-B*27 subtypes, and therefore, it should be further investigated.

Conclusion

As of January 1, 2017, there are 160 known subtypes of HLA-B*27, but an overwhelming majority of them are very rare and have not been studied for disease association. However, a very few of the relatively common subtypes exhibit differential association with AS, and there may even be some level of hierarchy in this regard. It has been hypothesized that the pathogenic role of HLA-B*27 in AS and its protective role in HCV and HIV infections suggest a common ground, i.e., promoting a more pronounced immune/inflammatory response for effective clearance of some pathogens but that can, on the other hand, lead to autoimmunity and tissue injury in some circumstances [49••, 55]. This may involve not only efficient antigen presenting properties of HLA-B*27 but also co-inheritance of other gene variants, thymic selection of CD8+ T-cell precursors, and/or specific T cell receptor repertoires [6••, 7••, 14•, 17••, 19•, 55, 56].

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
•• Of major importance

1. Robinson J, Halliwell JA, Hayhurst JD, et al. The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res. 2015;43(Database issue):D423–31. doi:10.1093/nar/gku1161. This publication explains the Immuno Polymorphism Database (IPD) that was developed in 2003 to provide a centralized system for the study of polymorphism in genes of the immune system. The IPD project was established by the HLA Informatics Group.


7. Ellingham D, Jostin L, Spain SL, et al. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nat Genet. 2016;48(5):S10–S0. The authors report combined assessments of Immunochip genotyping datasets from 52,262 patients with five closely associated diseases (AS, Crohn’s disease, psoriasis, primary sclerosing cholangitis and ulcerative colitis; all seronegative inflammatory diseases) and 34,213 healthy controls. They were able to delineate the genetic overlap between these conditions, identify 27 new associations, and highlights disease-specific patterns at shared loci.


17. Schittenhelm RB, Sivaneswaran S, Lim Kam Sian TC, et al. Human leukocyte antigen (HLA) B27 allelo-type-specific binding and candidate arthritogenic peptides revealed through heuristic clustering of Data-independent Acquisition Mass Spectrometry (DIA-MS) data. Mol Cell Proteomics. 2016;15(6):1867–76. The authors have identified 26 peptides, which are presented in
lower abundance by HLA-B*27:06 and HLA-B*27:09 compared with disease-associated HLA-B27 subtypes. They point out that "although these differences were observed to be very subtle, these 26 peptides might encompass the sought-after arthropogenic peptide(s)."


49. Neumann-Haefelin C. HLA-B27-mediated protection in HIV and hepatitis C virus infection and pathogenesis in spondyloarthritis: two sides of the same coin? Curr Opin Rheumatol. 2013;25(4):426–33. The author reviews the multiple virological and immunological mechanisms that have been suggested to contribute to HLA-B27-mediated protection in HIV and HCV infection. It is suggested that some of these mechanisms may also be involved in HLA-B27-associated pathogenesis in AS and associated forms of spondyloarthritis.


51. Hecht L, Dormer A. The evolutionary significance of certain amino acid substitutions and their consequences for HIV-1 immunogenicity toward HLA’s A*0201 and B*27. Bioinformation. 2013;9(6):315–
This study shows that the vaccines based upon the envelope glycoprotein of HIV are unlikely to be effective in eliciting an immune response in vivo in HLA-B*27+ individuals because the predominant 9-mers IRIGPGQAF of gp120 are significantly less immunogenic toward HLA-B*27 than to HLA-A*02.

Nitschke K, Barriga A, Schmidt J, et al. HLA-B*27 subtype specificity determines targeting and viral evolution of a hepatitis C virus-specific CD8+ T cell epitope. J Hepatol. 2014;60(1):22-9. This important study shows that a hepatitis C virus-specific CD8+ T cell epitope NS5B2820 is immunodominant in the context of HLA-B*27:02 (that is frequent in the Mediterranean populations), but is not restricted by HLA-B*27:05 (the most common subtype worldwide). These findings may have important implications for the selection of epitopes for global vaccines.


