Human Leukocyte Antigen–Disease Associations in Rheumatoid Arthritis

Vincent van Drongelen, PhD, Joseph Holoshitz, MD*

KEYWORDS

• HLA • Rheumatoid arthritis • Autoimmunity • Shared epitope

KEY POINTS

- Certain human leukocyte antigen (HLA) alleles have been found to be associated with immune mediated or autoimmune diseases, but the underlying mechanisms are largely unknown.
- Rheumatoid arthritis (RA) strongly associates with *HLA-DRB1* alleles that encode a sequence motif called shared epitope (SE), and there is variability in the strength of RA-SE association among ethnic and racial populations.
- The SE shows interaction with environmental factors (tobacco exposure) and together significantly amplify the disease risk.
- In contrast to RA risk-conferring SE-coding alleles, there are several other *DRB1* alleles that protect against the disease.
- Genome-wide association studies discovered many non-HLA RA risk loci, but their aggregate contribution to RA risk is outweighed by that of the SE.

Disclosure: The authors have nothing to disclose.
Funding Sources: Dr J. Holoshitz has been supported by the Eleanor and Larry Jackier Research Award from the UM–Israel Partnership for Research program and by grants from The National Institute of Environmental Health Sciences (R21 ES024428), The National Institute of General Medical Sciences (R01 GM088560), The National Institute of Dental and Craniofacial Research (R21 DE023845), and The National Institute of Arthritis and Musculoskeletal and Skin Diseases (R01 AR059085). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
Department of Internal Medicine, University of Michigan, 1150 West Medical Center Drive, Ann Arbor, MI 48109, USA
* Corresponding author. University of Michigan, 5520D MSRB1, SPC 5680, 1150 West Medical Center Drive, Ann Arbor, MI 48109-5680.
E-mail address: jholo@umich.edu

Rheum Dis Clin N Am 43 (2017) 363–376
http://dx.doi.org/10.1016/j.rdc.2017.04.003 **rheumatic.theclinics.com**

INTRODUCTION

Rheumatoid arthritis (RA) is a common inflammatory disease in which both genetic and environmental factors play a role in disease development. Based on twin studies, the heritability of the disease was estimated at around 60%.[1] Among all the genetic risk factors found to date, the human leukocyte antigen (HLA) locus is the most significant one. A particularly strong association between RA and *HLA-DRB1* alleles that encode an HLA-DRβ chain containing a 5 amino acid sequence motif called the shared epitope (SE) has long been documented.[2] Here the authors review salient immunogenetic, clinical, and mechanistic features of RA association with the HLA locus, focusing primarily on the SE.

Human Leukocyte Antigen Genes and Their Products

The immune system is composed of various cells that work together to protect the host against invading pathogens without harming its own tissues. Therefore, the host has to recognize which antigens are self and which are foreign. To discriminate between such self- and foreign antigens, the major histocompatibility complex (MHC) antigens, known in humans as HLA, have evolved. MHC molecules have the ability to recognize and present foreign antigens to the immune system but at the same time disregard self-antigens. This ability to discriminate between self and foreign is established through a process called "MHC restriction."[3] During the development of T cells in the thymus, T cells that react to self-antigens are eliminated, whereas those that respond to foreign antigens that are presented by a self-HLA molecule are preserved. This selection results in $CD4^+$ and $CD8^+$ T cells that only respond to foreign antigens that are presented by self-HLA molecules. Despite their ability to selectively recognize and respond to foreign antigens, many HLA alleles have been found to confer susceptibility to various diseases, most of which involve dysregulated immunity or autoimmunity.

The HLA locus is located on the short arm of chromosome 6 and covers a 7.6-Mb region that contains more than 250 highly polymorphic genes.[4] The region is organized in 3 subregions: class I, class II, and class III, which all have different functions. Both class I and II regions encode for glycoproteins that are expressed as cell surface receptors, whereas the class III region contains genes that encode for a variety of secreted immune system proteins, including complement factors and cytokines.

The class I region encodes for 3 main subsets of HLA molecules: HLA-A, HLA-B, and HLA-C. Class I HLA molecules are composed of an HLA-coded heavy α-chain and an invariant light chain, beta-2 microglobulin, which is essential for functional expression of the HLA molecule at the cell surface. The α-chain is folded to form a peptide-binding cleft that is closed and can accommodate short antigenic peptides, typically 8 to 10 amino acids long. These class I molecules are expressed on all nucleated cells and specialize in presentation of intracellular antigens, including viral antigens, to cytotoxic ($CD8^+$) T cells. Genes in the class II region encode for HLA-DR, HLA-DP, and HLA-DQ molecules as well as a few other related proteins. Class II HLA molecules are composed of an α-chain and a β-chain, both coded by the HLA class II region. Unlike the class I molecules, the peptide-binding cleft of class II molecules are open, which allows the accommodation of larger peptides of 15 to 20 amino acids long. Class II molecules are initially expressed on the cell surface of immune cells, in particular antigen-presenting cells, such as macrophages or dendritic cells, as well as B cells and activated T cells. These molecules present antigens from outside the cell to ($CD4^+$) T cells, which in turn stimulate B cells to produce antibodies toward that specific antigen, resulting in an antigen-specific immune response. After