REVIEW

Human Leukocyte Antigens and Their Correlation to Disease

Martin Gomez [1]

Lack of information on unique diseases has led to research on human components associated with disease defense such as the major histocompatibility complex (MHC), also known as the human leukocyte antigen (HLA). This genetic complex produces three classes of MHC receptors in humans, MHC/HLA I, II, and III, located on the cell membrane. MHC is a highly polymorphic gene complex with high linkage disequilibrium. Molecular changes and mutations on units of the MHC increase susceptibility for disease. These findings led to procedures which purposely induced changes on the MHC to prove its association with disease susceptibility. Analyzing and sequencing the genome of the HLA gene region is an important part of understanding the MCH, but sequencing can be difficult due to its high polymorphism trait. These issues led to faster, more efficient, and cheaper methods for sequencing not just the HLA but any gene region. Technology such as mass spectrometry is also facilitating characterization of HLAs, and finding disease correlations. With these advances in technology, a source for various diseases can be found.

The major histocompatibility complex (MHC), known as human leukocyte antigen (HLA) in humans, is a region of genes found in all mammals. This gene region expresses polypeptides present on all cell membranes which are involved in disease defense. This region of genes encompasses 7.6 Mb on chromosome 6 in humans and is the most dense gene region in the human genome (1). It is a polymorphic region, containing 252 expressed loci, which encodes all three classes of HLA (I-III), various subunits, and peptides. These expressed products of the MHC are involved in the specific binding and recognition of self and foreign antigens. HLAs are also involved in several key immune responses such as inflammation, regulation, leukocyte maturation, and stress response.

The MHC gene region contains extended sub-regions and three main classifications of HLA genes, Class I, Class II, and Class III. Class I type MHC genes express the classical I antigens, HLA-A, HLA-B, HLA-C, and several other subtypes and haplotypes. Class I types can be found on all nucleated cells and are involved in cell recognition and presentation (2). Class II type MHC genes express classical II antigens (HLD-DP, DQ, DR) found on antigen presenting cells such as macrophages, B-cells, activated T-cells, monocytes, dendritic cells, and endothelial cells. Class II antigens are mainly involved in the identification of foreign antigens (3).

[1]Senior Student, Department of Biological Sciences, Indian River State College, Fort Pierce, Florida

Class III HLAs and their subtypes are involved in expressing complementary components such as C2, C4 (C4A, C4B), heat shock proteins, inflammatory molecules, and plasma proteins, which are involved in other defensive immune functions (4).

The MHC is associated with hundreds of diseases, including most autoimmune diseases. In some cases, a mutation in a given MHC gene might be the cause of one disease and associated with a different disease (5). Class I HLA genes, HLA-A, HLA-B, and HLA-C, are all correlated with auto-immune disease. Class II MHC genes, HLA-D and its subtypes, have been associated with diseases and disorders such as rheumatoid arthritis, diabetes and Grave's Disease (6). Many other diseases have been correlated to this gene region such as psoriasis asthma, epilepsy, malignant melanoma, inflammatory bowel disease, asthma, diabetes, systemic lupus erythematosus, and celiac disease. There is more research correlating disease with the MHC gene region, but attempting to find connections between a disease and a specific MHC gene is difficult. This is due to the high polymorphism of these genes and the fact that many MHC-associated diseases are multifactorial diseases.

Diseases could be better assessed if more information and data could be obtained on the MHC, but many issues arise due to the unique characteristics of the HLA gene sequence. The HLA gene region has a high linkage disequilibrium and is highly polymorphic due to its responsibilities in recognizing highly polymorphic foreign antigens, leading to difficulty in sequencing these genes. Having better technology to identify and sequence genes could lead to finding solutions or possible treatments.

Findings in Transgenic Mice

Studies have been done associating sickness and disease to HLAs. In one study, mice were genetically engineered to not express MHC class I molecules present on the cell surface. This was done through inactivation of the β 2-microglobulin gene, which codes for a protein required for the assembly and cell surface expression of the MHC class 1 molecule (7). The altered mice, along with a control group of normal mice containing MHC I peptides, were further induced by an experimental model of systemic lupus erythematous (SLE). This was done by immunizing the mice with a monoclonal anti-DNA antibody derived from an SLE patient who contained a common idiotype. The derived DNA antigens and nuclear antigens led to the production of antibodies inducing symptoms closely paralleling SLE. The findings show that the MCH I deficient mice did not respond at all to the model SLE, whereas the normal mice responded by generating antibodies to the monoclonal anti-DNA. The mice lacking MHC I were further induced with monoclonal anti-DNA derived from a mouse with SLE. This also did not elicit a response. Other findings from this study show the MHC I deficient mice were generally healthy, capable of producing antibody responses, and could combat viral infections; however, they were much

more sensitive to intracellular parasites (7).

In another study mice, were bred to be devoid of MHC class I and II (8). Through gene targeting, two sets of mice were engineered for one group of mice to not contain MHC I and the other group to not contain MHC II. These two groups were then bred together. The offspring DNA was put through screening to find offspring devoid of MHC I and II. This was done to create a type of murine model for an organism fully devoid of MHC components, and to record its effect on the immune system. Flow cytometry and Southern blotting were used to identify MHC deficient mice. These mice were later immunized with trinitrophenol-conjugated Ficoll in phosphate-buffered saline and analyzed through enzyme-linked immunosorbent assay. Even though being devoid of both types of MHC, they appeared healthy when housed under normal conditions, survived for many months, and were able to breed successfully. They were also immunocompetent, being able to produce B-cell lymphocytes and elicit antibody responses against T-dependent antigens. The MHC was thought to be a major component of the immune system, and vital for the overall health of an organism. One would assume the specimen's immune system would, at the least, be compromised. As seen here, it is obvious many other side components and pathways have a major effect on the immune system. Findings like these help obtain insight into diseases such as bare lymphocyte syndrome, where patients fail to express MHC molecules and die at an early age (8).

More procedures have been done on mice that correlate disease with the MHC. In one procedure, a gene labelled Ea gene in the MHC-II shows a correlation with lupus susceptibility in mice (9). Lupus prone mice contain different haplotypes in their MHC, but these mice with the increased susceptibility do not have a distinct Ea gene. Mice with reduced susceptibility contained the Ea gene. One copy, or one haplotype, of this gene appears to serve as a type of protection gene or an immune suppressive gene for systemic lupus.

Other earlier studies also show HLA antigen associations with malignant melanoma (10). These studies involve HLA examined through immuno-peroxidase staining of ten nevi samples and ninety-eight melanoma lesion samples. HLA Class I antigens serve an important role in antigen-peptide presentation between cytotoxic T cells and melanoma cells. The results for the study show down regulation, reduced or suppressed cellular response, of HLA Class I antigens in melanoma lesions and nevi. This down regulation led to the progression of lesion thickness, quicker disease progression, and reduced survival.

The studies on mice show a distinct connection between MHC and disease. The findings may not give extreme insight or specificity, but they do give a foundation on the MHC. They facilitated deductions on the MHC and allowed further development of techniques for analyzing the MHC.

Genetic Typing

To further understand MHC/HLA, learning how to characterize these genes through typing is an important tool in clinical applications and medical research. This has led to mapping the HLA genes, HLA single nucleotide polymorphisms (SNPs) and deletion/ insertion polymorphisms. This led to the discovery of unique tag SNPs that could be useful in future studies, which could facilitate finding patterns, functions, and possibly origins of the HLA loci (11). For example, using these unique tag SNPs led to a method to detect celiac disease predisposition, showing celiac disease is associated with unique HLA variants of the HLA-DQA region. These tag SNPs allowed for a quick, accurate, and highly specific sequencing of celiac disease HLA risk factors in celiac Disease patients (12).

SNP data analysis also led to distinguishing a correlation for HLA genes and rheumatoid arthritis (RA). The common association with RA risk is known to be with HLA-DRBA1 alleles. Yet controversy persists on these data, which has led to continued research on the topic. Genome wide-SNP data of five thousand seropositive cases, and fifteen thousand controls were genotyped across the MHC for HLA- (A, B, C, DPA1, DQA1, DQB1, DRB1) and three thousand SNPs. Results show HLA-DRB1, HLA-B, and HLA-DPB1 (peptide-binding groove genes) associate with increased RA (13).

HLA genes share a vast number of similarities and sometimes may differ by just one nucleotide, making some current HLA DNA typing methods such as sequenced-based typing (SBT) yield unclear results due to phase ambiguity. Ambiguity results from several alleles producing the same identical sequences, leaving them unable to be identified. Attempting to resolve this issue is too time consuming and costly making these old methods unsuitable. To overcome this, the super-high-resolution single-molecule sequenced-based-typing (SS-SBT) method was used to produce HLA allele sequences at the 8-digit level. This method can determine an HLA allele sequence from any DNA sample even if it includes enhancers, promoters, exons, introns, and UTRs. It uses long range PCR with high fidelity polymerase with next generation sequence platforms (Roche GS Junior system), and compares allele sequences using an IMGT-HLA database. This led to the discovery of polymorphisms across the whole gene and solved the phase ambiguity issue with HLA- (A,-B,-C,-DRB1,-DQB1) which were assigned to single HLA alleles. This method is an effective HLA DNA typing method to detect new HLA alleles and null alleles without ambiguity (14). Developments in DNA sequencing have led to a higher resolution, more cost effective method to type HLA genes by sequencing. This newer process, which builds on top of SS-SBT, uses long range amplification, long-range PCR on HLA-(A,-B,-C ,–DRB1), reference cell line samples, clinical samples, and then sequences the samples using high-throughput platforms (HiSeq, Miseq). This newer technique has the capacity to sequence thousands of samples at low cost and at high throughput speed. This method is also specific and is able to sequence multiple samples of different origins simultaneously, which could allow insight into associated diseases (15).

Continued developments in DNA sequencing have led to the most effective method, the bead-based normalization for uniform sequencing (BENUS) depth protocol. This technique is a high-throughput, cost effective sequence method that involves PCR amplification of HLA samples, using a MiSeq sequencer. This HLA sequencing method is faster, more cost effective, and especially suited for multi-samples. This technique is also very specific, and has the ability to show allelic sequence at the 8-digit level, making this a very effective method for HLA sequencing (16).

Development in genetic sequencing/ typing of the MHC has led to a connection between MHC and inflammatory bowel disease (IBD). IBD, also known as Crohn's disease (CD) and ulcerative colitis (UC), has previously been shown to have a strong correlation with the MHC. For example, the MHC-class I chain related gene-A (MICA), found in HLA-B gene region, is associated with UC. The MICA gene solely expresses its products in the gastrointestinal epithelium. Once it was sequenced, it was shown to have a unique triple repeat polymorphism. Samples were taken from patients with UC and from random controls with no UC and sequenced for the MICA gene. Results show UC patients contain higher amounts of MICA alleles and phenotype frequencies, where the controls without UC have no MICA alleles (17). Continued studies indicate multiple associations with HLA and IBD, with the most consistent involving HLA-DRB1 and HLA-DQB1. Attempts were done to differentiate HLA correlations between CD and UC. Meta-analysis of genome-wide association studies show that CD and UC share the majority of HLA genetic risk factors. High-density SNP typing of the MHC was then performed on over thirty-two thousand patients with IBD to attempt to differentiate between the two. Results show a genetic variation between CD and UC in Class I HLA- (B, C, A) and Class II HLA-(DRB1, DRB3, DRB4, DRB5, DQA1, DQB1, DPA1, DPB1) genes. A unique result shows a special association between ulcerative colitis and the HLA-DRB1*01:03 allele (18).

Mass Spectrometry

Mass spectrometry has led to findings regarding the characteristics of the MHC/HLA gene region by analyzing the expressed amino acids and peptides. One type, micro-capillary high performance liquid chromatography (HPLC) electro-spray ionization tandem mass spectrometry, was used to sequence peptides isolated from MHC molecules. A previous method, HPCL-mass spectrometry was only able to be applied to a few fractions that contained one or two dominant peptides. This new technique can, within hours, determine the molecular mass, the peptide length, and the number of each individual peptide (19).

A different high-throughput mass-spectrometry method was used to produce accurate identification of thousands of unique peptides and binding motifs. This method took cell lines and purified HLA complexes and peptidomes from each cell line and separated the peptides from the hydrophobic HLA heavy chains through elution. Peptides were further purified and concentrated for mass spectrometry analysis. Peptides were individually separated by a Nanoflow HPLC (Proxeon Biosystems) and analyzed with a nano-electrospray ion source (Q Exactive) mass spectrometer. This approach takes advantage of the new developments in mass spectrometry technology and HLA purification techniques. This method only requires small sample sizes, reduces measurement time, and enables in-depth accurate identification of HLA peptidomes (20).

The use of mass spectrometry for the analysis of the MHC can enhance knowledge of peptide-MHC interactions. Further development can enhance the speed and accuracy of the analysis of these peptides. It is important to understand the expressed products of the MHC, considering many of the associated disease with the MHC are multi-factorial. Being multi-factorial multiplies the gravity of the disease because associations can be in different DNA regions or in different peptide regions. Continued development in mass spectrometry is important, so that it can facilitate future technology developments and possibly treatments.

Discussion

As has been seen, the MHC/HLA gene series is a unique sequence of genes which have a tremendous effect on health and the immune system. It is associated with many diseases such as systemic lupus erythematosus, epilepsy, diabetes, Grave's disease, psoriasis, asthma, rheumatoid arthritis, inflammatory bowel disease, and celiac disease.

Earlier procedures done to associate MHC with disease predisposition led to studies done on transgenic mice to correlate SLE with MHC genes. In these procedures, mice were genetically altered to be devoid of MHC I molecules. In another procedure, mice were bred to be devoid of both MHC I and II molecules. Results were similar for both procedures showing the mice did not respond to induced SLE, had an overall healthy immune system, healthy reproductive system, were capable of producing antibody responses, and were able to survive the disease.

Other studies on mice show results with a clear connection between HLA-Ea and SLE. Findings show HLA-Ea gene containing mice have a reduced susceptibility to systemic lupus, whereas mice not containing HLA-Ea gene had an increased susceptibility. Other findings show a connection between HLA genes and malignant melanoma. Through staining procedures, results show that there is a clear case of down regulation of HLA antigens in malignant melanoma which can lead to increased lesions and disease progression of malignant melanoma. These findings in HLA connection with disease may not give great insight into the functions of the MHC but they give researchers a foundation to continue studies into MHC disease correlation.

Issues arise from attempts to identify MHC gene sequences due to its high rate of polymorphism and linkage disequilibrium rate. Other issues with sequencing HLA genes are due to the high phase ambiguity problem where several HLA alleles may produce the same sequence due to HLA genes sharing a large amount of similarities with some differing by only one nucleotide. These issues drove the need to develop new techniques in HLA sequencing and

typing. This pursuit led to the mapping of SNPs and HLA genes across the MHC. The SNP mapping led to the discovery of unique SNP tags which have already determined correlations between HLA genes and diseases.

Continued research into MHC and disease association has led to the development of efficient HLA gene sequencing technology. Techniques, such as SS-SBT, have resolved phase ambiguity issues, are able to discover polymorphisms across the whole MHC gene, are more cost effective, and more time effective. SS-SBT also demonstrates high specificity, producing sequences at the 8-digit level. This technology allows for large samples and multi-sample sequencing, which could be helpful in analyzing multi-factorial diseases. The newest development in sequencing, BENUS, is capable of sequencing multi-samples, has high-throughput, is very specific (8-bit digit level), faster, and has shown to be even more effective than previous methods. These methods have led to discoveries showing correlations between IBD and the MHC and have differentiated Crohn's disease and ulcerative colitis found in IBD.

Knowing about the MHC/HLA at the genetic level is important, but analyzing their expressed products is also important. This has led to developments in mass spectrometry for analyzing the MHC peptides, amino acids, and overall peptidome. One method, micro-capillary high performance liquid electrospray ionization-tandem mass spectrometry, is able to determine the molecular mass, peptide length, and number of unique peptides within hours. This nano-electrospray mass spectrometry is faster, more efficient, and needs less sample.

The continued development of techniques in analyzing genetic components has, in recent years, become faster, cheaper, and more efficient. The growth in technology is always, increasing and the time periods between developments of new techniques are shortening as well. This increase in speed of developing technology can be attributed to the increased accessibility of data, which has allowed researchers to use and evolve older procedures. This is important because new technology has already led to important connections between MHC and disease. Most unique diseases are multi-factorial, but as technology becomes more efficient, characterizing these various factors will become easier, and once a disease can be characterized treatment can begin.

REFERENCES AND NOTES

1. C. Roucard, F. Garban, N A. Mooney, D.J. Charron, M.L. Ericson. Conformation of Human Leukocyte Antigen Class II Molecules, Evidence for Superdimers and Empty Molecules on Human Antigen Presenting Cells. *J. Bio Chem.* **271(24)**.
2. M. Steinmetz, L. Hood. Genes of the Major Histocompatibility Complex in Mouse and Man. *Science*. **222(4625).** 727-733 (1983).
3. S. Beck, D. Geraghty, H. Inoko, L. Rowen, B. Aguado, S. Bahram, R.D. Campbell, S.A. Forbes, T. Guillaudeux, L. Hood, R. Horton, M. Janer, C. Jasoni, A. Madan, S. Milne, M. Neville, A. Oka, S. Qin, G. Ribas-Despuig, J. Rogers, T. Shiina, T. Spies, G. Tamiya, H. Tashiro, J. Trowsdale, Q. Vu, L. Williams, M. Yamazaki. Complete sequence and gene map of a human major histocompatibility complex. *Nature*. **401(6756)**. 921-923. (1999).
4. A.J. Mungall, S.A. Palmer, S.K. Sims, C.A. Edwards, J.L. Ashurst, L. Wilming, M.C. Jones, R. Horton, S.E. Hunt, C.E. Scott, J.G.R. Gilbert, M.E. Clamp, G. Bethel, S. Milne, R. Ainscough, J P. Almeida, K.D. Ambrose, TD. Andrews, R.I.S. Ashwell, A.K. Babbage, C.L. Bagguley, J. Bailey, R. Banerjee, D.J. Barker, K.F. Barlow, K. Bates, D.M. Beare, H. Beasley, O. Beasley, C.P. Bird, S. Blakey, S. Bray-Allen, J. Brook, A.J. Brown, J.Y. Brown, D.C. Burford, W. Burrill, J. Burton, C. Carder, N.P. Carter, J.C. Chapman, S.Y. Clark, G. Clark, C.M. Clee, S. Clegg, V. Cobley, R.E. Collier, J.E. Collins, L.K. Colman, N.R. Corby, G.J. Coville, K.M. Culley, P. Dhami, J. Davies, M. Dunn, M.E. Earthrowl, A.E. Ellington, K.A. Evans, L. Faulkner, M.D. Francis, A. Frankish, J. Frankland, L. French, P. Garner, J. Garnett, M.J.R. Ghori, L.M. Gilby, C.J. Gillson, R.J. Glithero, D.V. Grafham, M. Grant, S. Gribble, C. Griffiths, M. Griffiths, R. Hall, K.S. Halls, S. Hammond, J.L. Harley, E A. Hart, P.D. Heath, R. Heathcott, S.J. Holmes, P.J. Howden, K.L. Howe, G.R. Howell, E. Huckle, S.J. Humphray, M.D. Humphries, A.R. Hunt, C.M. Johnson, A.A. Joy, M. Kay, S.J. Keenan, A.M. Kimberley, A. King, G K. Laird, C. Langford, S. Lawlor, D.A. Leongamornlert, M. Leversha, C.R. Lloyd, D.M. Lloyd, J.E. Loveland, J. Lovell, S. Martin, M. Mashreghi-Mohammadi, G.L. Maslen, L. Matthews, O.T. McCann, S.J. McLaren, K. McLay, A. McMurray, M.J.F. Moore, J.C. Mullikin, D. Niblett, T. Nickerson, K.L. Novik1, K. Oliver, E.K. Overton-Larty, A. Parker, R. Patel, A.V. Pearce, A.I. Peck, B. Phillimore, S. Phillips, R.W. Plumb, K.M. Porter, Y. Ramsey, S.A. Ranby, C.M. Rice, M.T. Ross, S.M. Searle, H.K. Sehra, E. Sheridan, C.D. Skuce, S. Smith, M. Smith, L. Spraggon, S.L. Squares, C.A. Steward, N. Sycamore, G. Tamlyn-Hall, J. Tester, A.J. Theaker, D.W. Thomas, A. Thorpe, A. Tracey, A. Tromans, B. Tubby, M. Wall, J.M. Wallis, A.P. West, S.S. White, S.L. Whitehead, H. Whittaker, A. Wild, D.J. Willey, T.E. Wilmer, J.M. Wood, P.W. Wray, J.C. Wyatt, L. Young, R.M. Younger, D.R. Bentley, A. Coulson, R. Durbin, T. Hubbard, J.E. Sulston, I. Dunham, J. Rogers, S. Beck. The DNA sequence and analysis of human chromosome 6. *Nature*. **425**. 805-811 (2003).
5. J. Dausset, The major histocompatibility complex in man. *Science*. **213(4515)**. 1469-1474 (1981).
6. C.A. Janeway, P. Travers, M. Walport. *Immunobiology: The Immune System in Health and Disease.* 5th edition. New York: Garland Science. (2001).
7. E. Mozes, L.D. Kohn, F. Hakim, D.S Singer. Resistance of MHC class I-deficient mice to experimental systemic lupus erythematosus. *Science*. **261**, pp91 (1993).
8. J. Grusby, H. Auchincloss., R. Lee, R.S. Johnson, J.P. Spencer, M. Ziljstra, R. Jaenisch, V.E. Papaioannous, L.H. Glimcher, Mice Lacking Major Histocompatibility Complex Class I and Class II Molecules. *PNAS*. **90(9).** 3913-3917 (1993).
9. N. Ibnou-Zekri, M. Iwamoto, L. Fossati, P.J. McConahey, S. Izui. Role of the Major Histocompatibility Complex Class II Ea Gene in Lupus Susceptibility in Mice. *PNAS*. **94(26)**. 14654-14659 (1997).
10. T. Kageshita, S. Hirai, T. Ono, D.J. Hicklin, S. Ferron., Down-Regulation of HLA Class I Antigen-Processing Molecules in Malignant Melanoma Association with Disease Progression. *Am J Pathol.* **154(3)**. 745–754 (1999).
11. P.I.W. de Bakker, G. Mcvean, P.C. Sabeti, M.M. Miretti, T. Green, J.

Marchini, X. Ke, A.J. Monsuur, P. Whittaker, M. Delgado, J. Morrison, A. Richardson, E.C. Walsh, X. Gao, L. Galver, J. Hart, D.A. Hafler, M. Pericak-Vance, J.A. Todd, M.J. Daly, J. Trowsdale, C. Wijmenga, T.J. Vyse, S. Beck, S.S Murray, M. Carrington, S. Gregory, P. Deloukas, J.D. Rioux. A high resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet.* **38(10)**. 1166–1172 (2006).

12. A.J. Monsuur, P.I.W de Bakker, A. Zhernakova, D. Pinto, W. Verduijn, J. Romanos, R. Auricchio, A. Lopez, D. A. van Heel, J. B. A. Crusius, C. Wijmenga, Effective Detection of Human Leukocyte Antigen Risk Alleles in Celiac Disease Using Tag Single Nucleotide Polymorphisms. *PLOS ONE*. **4(5)**. (2009).
13. S. Raychaudhuri, C. Sandor, E.A. Stahl, J. Freudenberg, H. Lee, X. Jia, L. Alfredsson, L. Padyukov, L. Klareskog, J. Worthington, K.A. Siminovitch, S. Bae, R.M. Plenge, P.K. Gregersen, P.I.W. de Bakker, Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet*. **44(3)**. 291–29 (2012).
14. T. Shiina, S. Suzuki, Y. Ozaki, H. Taira, E. Kikkawa, A. Shigenari, A. Oka , T. Umemura, S. Joshita, O. Takahashi, Y. Hayashi, M. Paumen, Y. Katsuyama, S. Mitsunaga, M. Ota , J.K. Kulski, H. Inoko. Super high resolution for single molecule-sequence-based typing of classical HLA loci at the 8-digit level using next generation sequencer. *Tissue Antigens*. **80(4).** 305-316 (2012).
15. C. Wang, S. Krishnakumar, J. Wilhelmy, F. Babrzadeh , L. Stepanyan, L.F. Su, D. Levinson , M.A. Fernandez-Viña, R.W. Davis, M.M Davis, M. Mindrinos, High-throughput, high-fidelity HLA genotyping with deep sequencing. *Natl Acad Sci USA*. **109(22).** 8676-8681 (2012).
16. K. Hosomichi, S. Mitsunaga, H. Nagasaki, I. Inoue. A Bead-Based Normalization for Uniform Sequencing depth (BeNUS) protocol for multi-samples sequencing exemplified by HLA-B. *BMC Genomics*. **15(645).** (2014).
17. K. Sugimura; M. Ota, J. Matsuzawa, Y. Katsuyama, K. Ishizuka, T. Mochizuki, N. Mizuki, S.S. Seki, T. Honma, H. Inoko, H. Asakura. A close relationship of triplet repeat polymorphism in MHC class I chain-related gene A (MICA) to the disease susceptibility and behavior in ulcerative colitis. *Tissue Antigens*. **57(1)**. 9-14 (2001).
18. P. Goyette, G. Boucher, D. Mallon, E. Ellinghaus, L. Jostins, H. Huang, S. Ripke, E.S. Gusareva, V. Annese, S. L Hauser, J.R. Oksenberg, I. Thomsen, S. Leslie, International IBD Genetics Consortium16, M.J. Daly, K.V. Steen, R.H. Duerr, J.C. Barrett, D.P.B McGovern, L.P. Schumm, J.A. Traherne, M.N. Carrington, V. Kosmoliaptsis, T.H. Karlsen, A. Franke, J.D. Rioux. High density mapping of the MHC identifies a shared role for HLA-DRB1*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. *Nature Genetics*. **47(2).** 172-179 (2014).
19. D.F. Hunt, R.A. Henderson, J. Shabanowttz, K. Sakaguchi, H. Michel, N. Sevilir, A. L. Cox, E. Appella, V. H. Engelhard. Characterization of Peptides Bound to the Class I MHC Molecule HLA-A2.1 by Mass Spectrometry. *J Immunol*. **179(5).** 2669–2671 (2007).
20. M. Bassani-Sternberg, S. Pletscher-Frankild, L.J. Jensen, M Mann. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol Cell Proteomics*. **14(3).** 658-673 (2015).

ACKNOWLEDGEMENTS

The author thanks the IRSC librarians and IRSC Biology faculty for guidance in completing this work.