Human Immunology 80 (2019) 465–467

Contents lists available at ScienceDirect

Human Immunology

journal homepage: www.elsevier.com/locate/humimm

New HLA alleles discovered by next generation sequencing in routine histocompatibility lab work in a medium-volume laboratory

Kelly Jean Ingram[a], Hanneke Merkens[b], Elaine F. O'Shields[a], David Kiger[a], Michael D. Gautreaux[a,*]

[a] HLA/Immunogenetics and Immunodiagnostics Laboratories, Department of Pathology, Wake Forest School of Medicine, Winston-Salem, NC 27103, United States
[b] GenDx, Utrecht 3584 CM, The Netherlands

ABSTRACT

The immunogenetics research and clinical communities are undergoing a revolution in the way that Human Leukocyte Antigens (HLA) alleles are typed, thanks to the introduction and increasing acceptance of next-generation sequencing into laboratory practice. With the ability to sequence all exons of each allele, instead of the previously routine typing of exons 2 and 3 of class I and exon 2 of class II, the sequencing of previously unsequenced areas of HLA alleles is causing a host of new alleles to be discovered through the course of routine laboratory testing. In the first 4 months of routine next generation sequencing, we have identified 10 novel alleles that have been discovered through laboratory testing for all facets of HLA typing, i.e. solid organ transplantation, hematopoietic stem cell transplantation, disease association typing and pharmacogenomics testing. The advent of NGS HLA typing in routine clinical practice, and the concomitant routine typing of exons outside the norm, opens the window for rapid discovery of new HLA alleles and a potential for overwhelming the current HLA nomenclature naming conventions.

1. Introduction

The immunogenetics research and clinical communities are undergoing a revolution in the way that Human Leukocyte Antigens (HLA) alleles are typed, thanks to the introduction and increasing acceptance of next-generation sequencing into laboratory practice. The concept of resequencing, that is sequencing an area that has been previously sequenced, has allowed for HLA typing for all HLA alleles [1]. However, with the ability to sequence all exons of each allele, instead of the previously routine typing of exons 2 and 3 of class I and exon 2 of class II, the sequencing of previously unsequenced areas of HLA alleles is causing a host of new alleles to be discovered through the course of routine laboratory testing. The availability of such testing and new genomic data, coming now at increasingly greater scales, is transforming our understanding of the polymorphisms underlying the tremendous variation found in the HLA region [2].

In the first 4 months of routine next generation sequencing, we have identified 10 novel alleles that have been discovered through laboratory testing for all facets of HLA typing, i.e. solid organ transplantation, hematopoietic stem cell transplantation, disease association typing and pharmacogenomics testing. Some of these alleles appear to be the product of synonymous mutations, non-synonymous mutations, and recombination.

2. Materials and methods

Genotyping was performed by next generation sequencing (NGS) on the Illumina MiSeq next generation sequencer. Sequences were determined for HLA-A, B, C, DRB1, DRB3/4/5, DQA1, DQB1, DPA1, and DPB1 using locus-specific primers supplied by GenDx. Amplification, library preparation, and sequencing were performed according to the vendor and laboratory specifications. Sequences were analyzed using the GenDx NGS analysis software, NGSengine®.

A novel allele was suspected when a sequence result has clear and defined mismatches to all known HLA allele sequences for the HLA locus in question. When a possible novel allele was discovered, the data files were sent to GenDx for confirmation. Once the novel allele was confirmed, NGSengine created submission files to submit the sequence to GenBank. Once GenBank issued an accession number for the sequence, the files and manually entered data are submitted to the IPD-IMGT/HLA Database for naming by the WHO Nomenclature Committee. The new name was then published in the following WHO Nomenclature report following the assignation [3].

3. Results

The new alleles described below were discovered during routine HLA typing for various aspects of histocompatibility testing, such as typing for solid organ transplantation, hematopoietic stem cell

* Corresponding author.
E-mail address: mgautrea@wakehealth.edu (M.D. Gautreaux).

https://doi.org/10.1016/j.humimm.2019.03.005
Received 28 January 2019; Received in revised form 28 February 2019; Accepted 4 March 2019
Available online 05 March 2019

transplant, pharmacogenomic testing and disease association. None of the patients typed during the routine histocompatibility testing had sufficient relatives typed in order to allow for genotyping determining which haplotypes are associated with the new alleles. However, we have included the entire patient typing to allow for potential haplotypes to be approximated as more data becomes available.

3.1. Synonymous mutations

HLA-A*26:01:49 was identified in a disease association patient being tested for the presence of HLA-B*57:01, which is associated with abacavir hypersensitivity. HLA-A*26:01:01:01 is the most similar known allele. A*26:01:49 differs from A*26:01:01:01 at codon 276c in exon 5. A one nucleotide change from G to A (CCG to CCA) results in a synonymous mutation coding for proline. The name A*26:01:49 has been officially assigned by the WHO Nomenclature Committee in August 2018. This follows the agreed policy that, subject to the conditions stated in the most recent Nomenclature Report [3], names will be assigned to new sequences as they are identified. Lists of such new names will be published in the following WHO Nomenclature Report. The complete typing of the patient was HLA-A*26:01:49, 33:03; HLA-B*53:01, 57:03; HLA-C*04:01, 18:02; HLA-DRB1*08:04, -; HLA-DQA1*04:01, -; HLA-DQB1*03:19, -; HLA-DPA1*02:01, -; and HLA-DPB1*01:01, -.

HLA-A*02:01:156 was identified in a potential renal transplant recipient. HLA-A*02:01:01:01 is the most similar known allele. A*02:01:156 differs from A*02:01:01 at codon 187c in exon 4. A one nucleotide change from G to A (ACG to ACA) results in a synonymous mutation coding for threonine. The name A*02:01:156 has been officially assigned by the WHO Nomenclature Committee in November 2018. The complete typing for the patient was HLA-A*02:01:156, 33:03; HLA-B*15:03, 45:01; HLA-C*02:10, 16:01; HLA-DRB1*11:02, 12:01; HLA-DRB3*02:02; HLA-DQA1*01:05, 05:05; HLA-DQB1*03:19, 05:01; HLA-DPA1*01:03, 02:07; HLA-DPB1*18:01, 85:01.

3.2. Nonsynonymous mutations

HLA-C*03:404 was identified in an NMDP donor. HLA-C*03:04:01:02 is the most similar known allele. C*03:404 differs from C*03:04:01:02 at codon 1b in exon 2. A one nucleotide change from G to A (GGC to GAC) results in a coding change from glycine to aspartic acid. A neutral, non-polar, and small amino acid (glycine) is substituted for a negatively charged, polar amino acid (aspartic acid). Exon 2 forms part of the peptide binding site and the exchange of amino acids with different physical properties could result in binding property changes. The size difference between glycine and aspartic acid may also contribute to steric effects in the binding pocket. The name C*03:404 has been officially assigned by the WHO Nomenclature Committee in August 2018. The complete typing for the donor was HLA-A*02:01, 68:01; HLA-B*40:01, 51:01; HLA-C*03:404, 15:02; HLA-DRB1*11:01, 13:01; HLA-DRB3*01:01, 02:02; HLA-DQA1*01:03, 05:05; HLA-DQB1*03:01, 05:05; HLA-DQB1*03:01, 06:03; DPA1*01:03, -; HLA-DPB1*04:01, -.

HLA-DQA1*01:18 was identified in a disease association patient with drug-induced keratoconjunctivitis being tested for the presence of HLA-B*27. HLA-DQA1*01:01:02 is the most similar known allele. DQA1*01:18 differs from DQA1*01:01:02 at codon 105b in exon 3. A one nucleotide change from C to T (CCC to CTC) results in a coding change from proline to leucine [4]. A nonpolar, hydrophobic amino acid is exchanged for another non-polar, hydrophobic amino acid. The side chain of proline is attached to the backbone at two points creating a ring structure, which can introduce a kink in a peptide chain. Exon 3 forms part of the extracellular domain and exchanging proline for another amino acid could alter the protein structure and potentially change the binding specificity. The name DQA1*01:18 has been officially assigned by the WHO Nomenclature Committee in November 2018. The complete typing for the patient was HLA-A*01:01, 33:05; HLA-B*08:01, 14:02; HLA-C*07:01, 08:02; HLA-DRB1*01:02, 03:01; HLA-DRB3*01:01; HLA-DQA1*01:18, 05:01; HLA-DQB1*02:01, 05:01; HLA-DPA1*01:03, 02:01; HLA-DPB1*04:01, 17:01.

HLA-DQA1*01:19 was identified in a bone marrow donor. HLA-DQA1*01:02:01:01 is the most similar known allele. DQA1*01:19 differs from DQA1*01:02:01:01 at codon 221b in exon 4. A one nucleotide change from G to A (CGT to CAT) results in a coding change from arginine to histidine. A basic, polar amino acid is exchanged for another amino acid with the same properties. However, the pK_a of arginine and histidine vary greatly at 12.5 and 6, respectively. At physiological pH, arginine is protonated and positively charged, while histidine remains neutral. Exon 4 codes for the transmembrane protein and cytoplasmic tail. The exchange of amino acids with different charges could impact protein folding and hydrogen bond formation. The name DQA1*01:19 has been officially assigned by the WHO Nomenclature Committee in November 2018. The complete typing for the donor was HLA-A*02:01, 25:01; HLA-B*07:02, 18:01; HLA-C*07:02, 07:02; HLA-DRB1*04:01, 15:01; HLA-DRB4*01:03; HLA-DRB5*01:01; HLA- DQA1*01:19, 03:01; HLA-DQB1*03:02, 06:01; HLA-DPA1*01:03, -; HLA-DPB1*04:01, -.

HLA-B*38:82 was identified in a disease association patient with low back pain being tested for the presence of HLA-B*27. HLA-B*38:02:01 is the most similar known allele. B*38:82 differs from B*38:02:01 at codon 324b in exon 6. A one nucleotide change from C to T (GCG to GTG) results in a coding change from alanine to valine. Alanine and valine are non-polar and hydrophobic. Amino acids with these properties can be readily substituted for one another. Exon 6 codes for the cytoplasmic tail, so this mutation is not associated with peptide binding specificity. There are likely no intracellular effects of this amino acid exchange. The name B*38:82 has been officially assigned by the WHO Nomenclature Committee in November 2018. The complete typing for the patient was HLA-A*11:01, 30:02; HLA-B*14:02, 38:82; HLA-C*07:02, 08:02; HLA-DRB1*08:06, 15:02; HLA-DRB5*01:08N; HLA-DQA1*01:01, 01:02; HLA-DQB1*05:01, 06:02; HLA-DPA1*01:03, 02:02; HLA-DPB1*05:01, 34:01.

DQA1*05:13 was identified in a disease association patient with arthralgia being tested for the presence of HLA-B*27. HLA-DQA1*05:05:01:01 is the most similar known allele. DQA1*05:13 differs from DQA1*05:05:01:01 at codon -11a in exon 1. A one nucleotide change from G to T (GCC to TCC) results in a coding change from alanine to serine. Alanine is a non-polar, hydrophobic amino acid and serine is polar and hydrophilic. Exon 1 encodes the signal peptide, which typically consists of hydrophobic amino acids. The name DQA1*05:13 has been officially assigned by the WHO Nomenclature Committee in December 2018. The complete typing for the patient was HLA-A*01:01, 32:01; HLA-B*08:01, 27:05; HLA-C*01:02, 07:01; HLA-DRB1*11:01, 14:54; HLA-DRB3*02:02; HLA-DQA1*01:04, 05:13; HLA-DQB1*03:01, 05:03; HLA-DPA1*01:03, 02:01; HLA-DPB1*01:01, 04:01.

HLA-DQA1*01:20 was identified in a potential renal transplant recipient. HLA-DQA1*01:02:01:01 is the closest known allele. DQA1*01:20 differs from DQA1*01:02:01:01 at codon 71a in exon 2. A one nucleotide change from C to T (CAC to TAC) results in a coding change from histidine to tyrosine. Histidine is a polar, hydrophilic amino acid and tyrosine is hydrophobic. Exon 2 encodes the peptide binding site and this exchange can impact protein folding, which could alter peptide binding specificity. The name DQA1*01:20 has been officially assigned by the WHO Nomenclature Committee in November 2018. The complete typing for the patient was HLA-A*01:01, 23:01; HLA-B*08:01, 44:03; HLA-C*04:01, 07:01; HLA-DRB1*03:01, 15:01; HLA-DRB3*01:01; HLA-DRB5*01:01; HLA-DQA1*01:20, 05:01; HLA-DQB1*02:01, 06:02; HLA-DPA1*01:03, 02:02; HLA-DPB1*04:01, 05:01.

HLA-DQA1*01:21 was identified in a potential renal transplant recipient. HLA-DQA1*01:02:01:04 is the closest known allele. DQA1*01:21 differs from DQA1*01:02:01:04 at codon 187b in exon 4.

A one nucleotide change from C to T (GCC to GTC) results in a coding change from alanine to valine. Alanine and valine are both hydrophobic and non-polar. These amino acids can be readily substituted for one another. The name DQA1*01:21 has been officially assigned by the WHO Nomenclature Committee in December 2018. The complete typing for the patient was HLA-A*02:01, 33:03; HLA-B*27:05, 58:01; HLA-C*02:02, 03:02; HLA-DRB1*08:01, 13:02; HLA-DRB3*03:01; HLA-DQA1*01:21, 04:01; HLA-DQB1*04:02, 06:09; HLA-DPA1*01:03, 02:01; HLA-DPB1*01:01, 03:01.

3.3. Recombination

DPB1*835:01 was identified in a potential renal transplant patient. HLA-DPB1*34:01 is the most similar known allele. DPB1*835:01 matches DPB1*34:01 at exon 1 and 2, but differs at seven nucleotide positions in exon 3 and one in exon 4. A recombination event most likely occurred between exon 2 and 3. This is significant because exon 2 and 3 code for the extracellular domain. At codon 96b, a one nucleotide change from A to G (AAG to AGG) results in a coding change from lysine to arginine. Lysine and arginine are both basic, polar amino acids and can be substituted for one another with minimal effects under most circumstances. At codon 98c, a one nucleotide change from C to T (AAC to AAT) results in a synonymous mutation coding for asparagine. At codon 107a, a one nucleotide change from C to T (CTG to TTG) results in a synonymous mutation coding for leucine. At codon 118c, a one nucleotide change from A to G (ACA to ACG) results in a synonymous mutation coding for threonine. At codon 167c, a one nucleotide change from C to T (GAC to GAT) results in a synonymous mutation coding for aspartic acid. At codon 170b, a one nucleotide change from T to C (ATC to ACC) results in a coding change from isoleucine to threonine. Isoleucine is a non-polar, hydrophobic amino acid and threonine is a polar, hydrophilic amino acid. This amino acid exchange can affect protein folding and structure in the extracellular, peptide-binding domain. At codon 179c, a one nucleotide change from C to T (GAC to GAT) results in a synonymous mutation coding for aspartic acid. At codon 205a, in exon 4, a one nucleotide change from A to G (ATG to GTG) results in a coding change from methionine to valine. Both amino acids are non-polar and hydrophobic. Amino acids with these properties can be readily substituted for one another. The name DPB1*835:01 has been officially assigned by the WHO Nomenclature Committee in October 2018. The complete typing for the patient was HLA-A*03:01, 33:03; HLA-B*15:10, 35:01; HLA-C*04:01, -; HLA-DRB1*10:01, 13:03; HLA-DRB3*01:01; HLA-DQA1*01:05, 05:05; HLA-DQB1*03:01, 05:01; HLA-DPA1*01:03, 02:02; HLA-DPB1*01:01, DPB1*835:01.

4. Discussion

Due to their role in immune function, HLA genes are highly polymorphic. It is evolutionarily advantageous for the human population to have a diverse set of genes responsible for initiating an immune response [5]. However, the nature of HLA also has disadvantages when it comes to autoimmunity and transplantation [6]. Given the role that HLA plays in immunity, disease and transplantation, there is no surprise that HLA typing has large presence in research and clinical enterprises. While individual new genes are given merely a new number by the nomenclature committee, the fact remains that there are numerous potential reasons for the differences between two otherwise similar HLA alleles. Many individual alleles are in fact complex multi-SNP haplotypes [7]. We have shown at least three of the multiple mechanisms for generating HLA diversity in this manuscript. Doubtless, many more will be coming soon.

The advent of next-generation sequencing-based HLA typing in routine clinical practice, and the concomitant routine typing of exons outside the “normal” exonic regions, opens the window for rapid discovery of new HLA alleles and a potential for overwhelming the current HLA nomenclature naming conventions. Also, since most NGS systems currently sequence all of the introns of class I, and many of the introns of class II, it is likely that intronic regions hold a rich cache of previously undiscovered polymorphisms. However, in this manuscript, intronic differences were not apparent in any of the 10 presented polymoprhisms described.

Our laboratory, at the time of the writing of this manuscript, has been performing HLA typing by NGS for 4 months and has discovered, and had named, 10 novel alleles and have at least 4 more awaiting naming by the WHO nomenclature committee. The current system takes weeks or even months to assess and name a novel allele and relies too heavily on human hands to move the process forward. Although it may seem to be “Back to the Future”, the possibility of allowing for a temporary designation (e.g. “potentially novel allele – PNA” akin to the serological “workshop – w” designation) would allow a clinical laboratory, making a discovery of a new allele, to report a patient result without the delay necessitated by the current naming process.

The process of submitting to GenBank is almost automatic with the GenDX software, requiring very little manual data entry, and an accession number is generated within 1–2 days. IPD-IMGT/HLA Database requires a great deal of manual data entry. This is not to say that the data is unnecessary, just that the time required to fully compile and enter the data may be a strain for a busy clinical laboratory. If there was a way to have the GenBank data simply go straight into the IMGT, and have an intermediate designation provided, a lab can then make a quick report of a patient’s typing with a comment that the patient has a potential novel allele. The clinical report can be amended after final naming occurs.

Based on these findings, we expect to encounter a steady influx of additional novel alleles as we continue to sequence more individuals over the next few years. Our patient population accounts for an extremely small subset of the world population. If we were able to identify many new alleles in our patients, there are countless alleles unaccounted for in the remainder of the population. This shows the importance of molecular sequencing in a field with such a polymorphic and rapidly evolving set of genes. With the future implementation of NGS in other laboratories, the HLA community will gain a more comprehensive view of these polymorphic genes.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humimm.2019.03.005.

References

[1] S. Goodwin, J.D. McPherson, W.R.M. McCombie, Coming of age: ten years of next-generation sequencing technologies, Nat. Rev. Genet. 17 (2016) 333.
[2] D. Meyer, V.R.C. Aguilar, B.D. Bitarello, D.Y.C. Brandt, K.A. Numes, Genomic perspective on HLA evolution, Immunogenetics 70 (2018) 5–27.
[3] S.G.E. Marsh, E.D. Albert, W.F. Bodmer, R.E. Bontrop, B. Dupont, H.A. Erlich, M. Fernandez-Vina, D.E. Geraghty, R. Holdsworth, C.K. Hurley, M. Lau, K.W. Lee, B. Mach, M. Maiers, W.R. Mayr, C.R. M¨uller, P. Parham, E.W. Petersdorf, T. Sasazuki, J.L. Strominger, A. Svejgaard, P.I. Terasaki, J.M. Tiercy, J. Trowsdale, Nomenclature for factors of the HLA system, Tissue Antigens 75 (2010) 291–455.
[4] Chapter 4. Amino acids. In: Hardin, C.C., Knopp, J.A., (Eds.), Biochemistry – Essential concepts, (2013) Oxford University Press, New York, NY.
[5] P. Jin, E. Wang, Polymorphism in clinical immunology - From HLA typing to immunogenetic profiling, J. Transl. Med. 1 (1) (2003) 8, https://doi.org/10.1186/1479-5876-1-8.
[6] T.M. Williams, Human leukocyte antigen gene polymorphism and the histocompatibility laboratory, J. Mol. Diagn.: JMD 3 (3) (2001) 98–104.
[7] Richard J.N. Allcock, The major histocompatibility complex: a paradigm for studies of the human genome, in: Frank T Christiansen, Brian D. Tait (Eds.), Immunogenetics: Methods and Applications in Clinical Practice, Methods in Molecular Biology, © Springer Science + Business Media, New York, 2012, , https://doi.org/10.1007/978-1-61779-842-9_1.