ELSEVIER

Common and Well-Documented HLA Alleles

Report of the Ad-Hoc Committee of the American Society for Histocompatiblity and Immunogenetics

Pedro Cano, William Klitz, Steven J. Mack, Martin Maiers, Steven G.E. Marsh, Harriet Noreen, Elaine F. Reed, David Senitzer, Michelle Setterholm, Anajane Smith, and Marcelo Fernández-Viña

ABSTRACT: In histocompatibility testing some genotype ambiguities are almost always resolved into the genotype with the most common alleles. To achieve unambiguous assignments additional unwieldy tests are performed. The American Society for Histocompatibility and Immunogenetics formed a committee to define what human leukocyte antigen (HLA) genotypes do not need to be resolved in external proficiency testing. The tasks included detailed analysis of large datasets of high-resolution typing and thorough review of the pertinent scientific literature. Strict criteria were used to create a catalogue of common and well-documented (CWD) alleles. In total, 130, 245, 81, and 143 of the highly polymorphic HLA-A, -B, -C, and DRB1 loci fell into the CWD category; these represent 27%–30% of all alleles recognized. For the loci DRB3/4/5, DQA1, DQB1, and DPB1, a total of 29, 16, 26, and 52 CWD alleles were identified. A recommendation indicated that an acceptable report should only include one possible genotype; multiple genotypes can only be reported if only one of these includes two alleles of the CWD group. Exceptions in which resolution is not necessary are ambiguities involving functional alleles with identical sequences in the antigen recognition site. The criteria were established for proficiency testing, which could be a valuable tool when making clinical histocompatibility decisions. *Human Immunology 68, 392–417 (2007).* © American Society for Histocompatibility and Immunogenetics, 2007. Published by Elsevier Inc.

KEYWORDS: HLA; common allele; null allele; well-documented allele; proficiency testing; antigen recognition site; clinical histocompatibility; genotype ambiguity

ABBREVIATIONS

ARS	antigen recognition site
ASHI	American Society for Histocompatibility and Immunogenetics
CWD	common and well-documented
GF	gene frequency
HLA	human leukocyte antigen
WHO	World Health Organization

From the University of Texas M. D. Anderson Cancer Center (P.C., M.F.-V.), HLA Typing Laboratory, Houston, Texas; Public Health Institute (W.K.), Oakland, and University of California–Berkeley, Berkeley, California; Children's Hospital–Oakland Research Institute (S.J.M.), Oakland, California, and Roche Molecular Systems Department of Human Genetics, Alameda, California; National Marrow Donor Program (M.M., M.S.), Minneapolis, Minnesota; Anthony Nolan Research Institute (S.M.), Royal Free Hospital, London, United Kingdom; Fairview–University of Minnesota Medical Center (H.M.), Immunology Laboratory, Minneapolis, Minnesota; UCLA Immunogenetics Center (E.F.R.), Department of Pathology, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, California; City of Hope National Medical Center (D.S.), Duarte, California; Seattle Cancer Care Alliance (A.S.), Clinical Immunogenetics Laboratory, Seattle, Washington.

Address reprint requests to: Marcelo Fernández-Viña, Ph.D, University of Texas M. D. Anderson Cancer Center, HLA Typing Laboratory, 8515 Fannin, NA01.091C, Houston, TX 77054; Tel: (713) 792-2658; Fax: (713) 794-4773; E-mail: MFernand@mdanderson.org.

Received January 11, 2007; accepted January 17, 2007.

0198-8859/07/$–see front matter
doi:10.1016/j.humimm.2007.01.014

INTRODUCTION

The growing number of human leukocyte antigen (HLA) alleles with high sequence homology, and the extensive sharing of the distinguishing sequences with other alleles of the same locus, result in genotype ambiguities when performing molecular tests for HLA typing [1].

The results of many individual high-resolution molecular typing methods, in many instances, correspond to more than one possible genotype, including different allele combinations. Additional tiered tests are performed to exclude some allele combinations and to identify unambiguously, the actual genotype for a given locus [2]. The additional tests are expensive and add a significant burden to the histocompatibility laboratory.

Of all possible genotypes corresponding to ambiguous results, the combinations in which both alleles are frequent are more likely to be the actual genotype than genotypes in which both alleles are rare. The likelihood ratio for a given genotype, compared with the occurrence of other possible genotypes, can be estimated based on knowledge of gene frequencies of the alleles included in each possible genotype in the same population from which the individual comes. These estimations apply only to individuals whose parents are not genetically closely related; the accuracy of the likelihood estimates correlates directly with the accuracy of the estimation of gene frequency values.

The American Society for Histocompatibility and Immunogenetics (ASHI) formed an ad-hoc committee to summarize the result of studies that analyzed the genetic diversity of the HLA system in various populations, with the objective of determining criteria to establish what HLA genotypes do not need to be resolved when reporting results for external proficiency testing.

The accurate estimation of frequencies may be difficult for rare alleles and haplotypes. It has been estimated that at least three examples of a particular allele are needed to obtain an accurate frequency estimate [3].

In clinical histocompatibility practice, it is therefore important to recognize alleles that are observed with a significant frequency, and to distinguish them from alleles of the same locus that may have been observed only in one or a few instances and are not likely to be found again.

The Committee proposes to categorize HLA alleles on the basis of their gene frequencies. The distribution of alleles in various populations is different for different HLA loci. For example, the distributions of alleles at DRB1 and DPB1 loci are contrasting in most populations; typically up to 17 DRB1 alleles account for 90% of the alleles in almost all outbred populations. In comparison to this, only eight to ten DPB1 alleles are needed to cover the same proportion of alleles. Therefore, it is easy to distinguish the common alleles from the rare alleles in DPB1; in contrast, the distribution of DRB1 alleles seems to be more a continuum of alleles with intermediate frequencies. It is therefore difficult to determine a cut-off frequency to determine which alleles are rare.

MATERIALS AND METHODS

Some operational definitions were established. The committee proposed to define three categories of alleles:

1. *Common alleles.* Common alleles are those that appear with gene frequencies greater than 0.001 in any reference population. As mentioned above, accurate frequency estimations can be made for alleles observed three or more times; therefore, population studies with a sample size of 1,500 unrelated subjects ($2n$ = 3,000 chromosomes) could detect with accuracy only allele frequencies of 0.001 or greater.
2. *Well-documented alleles.* A second category includes those alleles that have been observed in at least three independent unrelated individuals; however, their gene frequencies have not been estimated accurately. Alleles that are found only in isolated populations may have low frequency in any outbred population, and their frequencies are likely to be lower overall than those of most common alleles.
3. *Rare Alleles.* The remaining alleles are considered rare alleles.These alleles have extremely low frequencies and are not likely to be found again in a significant number of unrelated subjects.

Classifying every allele this way may be difficult in some cases, given the lack of systematic studies of all populations, despite large numbers of subjects having been included. In addition, alleles found with intermediate or high frequency in some isolated populations may appear as well-documented or rare alleles in outbred populations with gene flow from those isolated populations. The studies conducted so far vary in sample size and typing resolution.

In the United States, the Histocompatibility laboratories serve a diverse population that includes patients and donors from virtually all world populations. This ad-hoc committee proposed to identify the "common" and "well-documented" alleles in all world populations.

The committee noted that some databases and studies might include allele level assignments that resulted from intermediate- or low-resolution tests. In these datasets, some of the allele level assignments may be erroneous and could have resulted from spurious reactions of some reagents. Therefore, the determination of allele frequen-