World Marrow Donor Association Guidelines for Use of HLA Nomenclature and Its Validation in the Data Exchange Among Hematopoietic Stem Cell Donor Registries and Cord Blood Banks 2012

W Bochtler[1], M Maiers[2], JNA Bakker[3], DM Baier[4], JA Hofmann[4], J Pingel[4], HG Rist[1], M Oudshoorn[3,5], SGE Marsh[6], CR Müller[1] and CK Hurley[7], *on behalf of the Information Technology Working Group of the World Marrow Donor Association, Leiden, The Netherlands*

[1]ZKRD Zentrales Knochenmarkspender-Register für Deutschland, Ulm, Germany; [2]NMDP National Marrow Donor Program, Minneapolis, MN, USA; [3]Europdonor Foundation, Leiden, The Netherlands; [4]DKMS German Bone Marrow Donor Center, Tübingen, Germany; [5]Department of Immunohematology and Blood Transfusion, University Medical Center, Leiden, The Netherlands; [6]ANRI Anthony Nolan Research Institute and UCL Cancer Institute, Royal Free Campus, London, UK and [7]Department of Oncology, Georgetown University Medical Center, Washington, DC, USA

Introduction

The HLA Nomenclature Guidelines of the World Marrow Donor Association (WMDA) describe how a valid HLA phenotype may be presented with regard to the structure and contents of individual data fields and the cross checks applied between given fields. This set of well-defined rules is enforced for the communication between hematopoietic stem cell donor registries, donor centres and/or cord blood banks. The guidelines apply to both existing HLA data currently in databases as well as assignments to be obtained in the future. Electronically reported HLA data containing invalid codes or violating basic rules should be rejected by the addressee or at least raise appropriate warning messages for the senders to check their data. Reference materials for these validations can be found following the links listed in Table 1.

This document is an update to the WMDA HLA Nomenclature Guidelines published in 2007 [1]. Adjustments necessary due to the 2010 revision of the WHO HLA nomenclature have been made [2,3]. Moreover, grace periods were defined for HLA alleles being withdrawn or renamed; these grace periods apply also to the letter codes used to designate multiple alleles in an HLA assignment. Additional information has been added to clarify further the range and meaning of these multiple allele codes. The policy of the National Marrow Donor Program (NMDP) for creating such codes is not affected by the rules stated in this document. The present update was approved by the Board of the WMDA in May 2012.

In the following the term 'donor' refers to adult volunteer donors and cord blood units. Likewise, the term 'registry' will also refer to cord blood banks. The term 'antigen' refers to an assignment in serological nomenclature obtained with serological methods and the term 'allele' refers to an assignment in DNA nomenclature obtained by DNA-based testing methods. The two to four colon separated fields of an allelic HLA code [2,3] are denoted as 'type', 'subtype', 'silent substitution' and 'intron variant' where the last field can optionally be trailed with an 'expression level character'. All given examples are based on IMGT/HLA database release 3.9.0 [4].

1 WMDA-approved HLA typing assignments

There are three categories of WMDA-approved HLA typing assignments, (1) WHO HLA Nomenclature Committee-approved designations, (2) abbreviations of such designations and (3) additional codes:

1.1 WHO HLA Nomenclature Committee-approved designations must be used whenever possible [3,5]

1.1.1 If serologic broad, split and associated antigen assignments are used, only those assignments as listed in the current WHO report are acceptable. If a split serologic typing is submitted, the serologic broad assignment must not be included, for example, DR13, not DR13(6).

1.1.2 Only allelic assignments as listed in the current WHO report are acceptable. Inclusion of the accompanying expression level character is recommended for registry communication. These characters specify alleles that are not expressed (N) or expressed at a low level (L), alleles for which the allelic product is secreted (S) or present only in cytoplasm (C) and finally alleles for which expression is aberrant (A) or questionable (Q) [6]. HLA assignments using older forms of nomenclature that once had been approved but were subsequently altered must no longer be used i. e., only the colon delimited form introduced in April 2010 is acceptable.

1.1.3 There are two forms of 'deleted' alleles:

- alleles that have been renamed (e. g. A*01:34N; *sequence shown to be expressed at low levels and renamed to A*01:01:38L*), and
- alleles that have been removed without a replacement (e. g. A*2401; *sequence shown to contain errors*).

For data exchange between registries, it is acceptable to use the older designation for renamed alleles for a period of one year after the renaming. However, allelic assignments removed without replacement must no longer be used as an adequate treatment in HLA matching programs is not possible.

1.2 Abbreviations of allele names derived from WHO HLA Nomenclature Committee-approved allele designations may be utilized when required for practical reasons. Please note that some of the approved codes described in this section are not valid WHO nomenclature, for example B*15:01N or A*02:XX.

1.2.1 Abbreviations may include shortened forms of WHO Nomenclature Committee-approved allele designations in which only type and subtype or type, subtype and silent substitution and optionally an existing expression level character are included. No other abbreviations are allowed. Preservation of the accompanying expression level character specifying an alternatively expressed allele (e. g. N, L, S, C, A, Q) is recommended for registry communication. A shortened code without trailing expression level character includes all alleles that are prefixed by the given type, subtype and possibly silent substitution

without regard to specific expression levels. A shortened code with trailing expression level character includes all alleles that are prefixed by the fields of the shortened code and assigned the same expression level character.

Examples: B*15:01 includes B*15:01:01:01, B*15:01:01:02N, B*15:01:02, B*15:01:03, B*15:01:04, . . . , B*15:01:29. However, B*15:01N currently includes B*15:01:01:02N only.

The one-year grace period described in section 1.1.3 also applies to shortened forms derived from renamed alleles. For example, the only known non-expressed allele covered by DRB4*01:01 is DRB4*0101102N, which was renamed in 1996 to the valid allele DRB4*01:03:01:02N. Hence, the designations DRB4*01:01N and DRB4*01:01:01:01N must not be used. Strictly speaking they should never have been used as the renaming was completed long before the introduction of the colon delimited nomenclature. On the other hand DRB4*01:01 can still be used as DRB4*01:01:01:01 is a valid allele.

1.2.2 Allelic family codes derived from WHO Nomenclature Committee-approved allele designations in which the allele type is trailed with ':XX' are acceptable – for example, A*02:XX, DRB1*11:XX. Structurally similar codes based on serologic designations not found in the DNA-based WHO HLA assignments are not acceptable – for example, A*09:XX, B*05:XX, DRB1*02:XX, DQB1*01:XX. If such unacceptable codes exist within the registry database, they should be replaced with appropriate multiple allele codes as described in section 1.3.1.

1.3 WMDA-approved additional codes

1.3.1 Ambiguous typing results have to be represented by P and/or G designations of the WHO HLA nomenclature or to be reported by means of the multiple allele coding system developed and maintained by the National Marrow Donor Program (NMDP) [7]. An example for such a code is DRB1*11:AD representing the ambiguity DRB1*11:01 or DRB1*11:04. The definition of a multiple allele code is locus independent and either can be generic (e. g. AB = 01/02) or allele-specific (e. g. CRY = 01:01/01:04N). If generic, each item only refers to the subtype of WHO allele designations. If allele-specific, each item includes type and subtype.

Because multiple allele code designations are based on shortened allele names, the rules for interpretation described under section 1.2.1 apply.

Examples: The code A*01:AB encodes the ambiguity A*01:01 or A*01:02. The shortened designation A*01:01 includes all alleles with names beginning with A*01:01 and does, for example, cover A*01:01:38L and A*01:01:01:02N. The same holds for A*01:CRY which encodes the ambiguity A*01:01 or A*01:04N i. e. the alleles A*01:01:38L and A*01:01:01:02N are also included.

1.3.1.1 The set of applicable multiple allele codes is determined by the current locus independent list of code definitions provided by the NMDP with regard to the current set of WHO Nomenclature Committee-approved allele designations:

(a) For the applicability of a multiple allele code definition to a certain HLA locus, the interpretation according to section 1.2.1 and the fact that expression level characters are optional for WHO allele designations must be considered.

Example: The definition AD = 01/04 is applicable to the allele type HLA-A*01 because A*01:AD encodes the valid ambiguity A*01:01 or A*01:04. Note that the code A*01:04 covers the non-expressed allele A*01:04N only.

Although not ideal, sometimes there are several coding options for a given ambiguity. For example the codes AD, CRY, BMFN, PECY and PECX are defined as follows:

AD = 01/04
CRY = 01:01/01:04N
BMFN = 01:01/01:01N/01:04N
PECY = 01:01/01:01L/01:04N
PECX = 01:01/01:01L/01:01N/01:04N

In the context of the allele type HLA-A*01, these five codes represent the same set of alleles which is A*01:01 or 01:01L or 01:01N or 01:04N according to section 1.2.1. The generic code AD has a unique meaning for type HLA-A*01 and hence is acceptable for registry communication. However, in this case it may be preferable to avoid the generic code as its definition conceals the information that A*01:04 only exists as a null allele.

To meet the practical needs of search coordinators, it is recommended that search reports show (e. g., in tooltips and/or legends) the current interpretation (including all covered expression codes) of a multiple allele code and not only the pure definition string. For example, A*01:AD, A*01:CRY, A*01:BMFN, A*01:PECY and A*01:PECX should all be described by the string A*01:01/01:01L/01:01N/01:04N.

(b) For allele-specific definitions that contain more than one allele type (crosscodes), the following designation rule for the leading digits of the multiple allele code must be used:

The type with the highest number of items in the definition string must be selected. Example: A*01:UTT (UTT = 01:01/01:02/02:02), as there are two items from type A*01 and only one item from type A*02 (i. e., A*02:UTT is invalid).

If there is a tie in the number of items from each type, the type with the lowest numeric value is selected. Example: A*01:GWXM (GWXM = 01:14/01:49/36:01/36:05) as there are two items from both types A*01 and A*36, and 01 is numerically lower than 36 (i. e., A*36:GWXM is invalid).

(c) In the interest of patient welfare, all formally valid and interpretable multiple allele codes must be allowed for communication between registries. Local policies should not hinder the transmission of HLA typing results.

1.3.1.2 Multiple allele codes can become deprecated due to removal and/or renaming of encoded alleles. Such codes must be replaced within one year by suitable replacement codes. After this period of time, deprecated multiple allele codes must no longer be used for data exchange between registries. For example, the code B*08:MNF (MNF = 08:01/08:06/08:07/08:08N/08:10) is deprecated since July 2010 when the allele B*08:06 was renamed to B*08:20. Hence, either the allele-specific code B*08:GYAP (GYAP = 08:01/08:07/08:08N/08:10/08:20) or the generic alternative B*08:PFRJ (PFRJ = 01/07/08/10/20) must be used as replacement.

1.3.1.3 Multiple allele codes including removed or renamed alleles must include alleles that correspond to one single release of the WHO HLA nomenclature report only. For example, the code B*08:DVKZ (DVKZ = 06/68) is regarded as invalid, as the allele B*08:06 was renamed to B*08:20 in IMGT/HLA database release 3.1.0 but the allele B*08:68 was introduced later in the subsequent release 3.3.0.

1.3.2 Special codes allowed include: XXXX for individuals tested positive for DRB3, DRB4 and/or DRB5 as in section 2.9.3, NNNN to be used for individuals tested negative for DRB3, DRB4 and/or DRB5 as in section 2.9.4, UUUU for alleles that currently have not been tested but for which an earlier typing exists as in section 2.8.5. Examples are DRB3*XXXX or DRB4*NNNN or DRB5*UUUU. The special code NEW must be used temporarily for an allele that has not yet been given an official name as in section 2.5. To avoid confusion with multiple allele code definitions as in section 1.3.1, HLA assignments for potentially new alleles must not take the form B*15:NEW, for example, but instead B*NEW.

2 WMDA-approved phenotype-specific requirements

2.1 Laboratories must assign DNA nomenclature to results obtained using DNA-based methods and serologic nomenclature to results obtained using antibody reagents.

2.1.1 The practice of converting alleles to antigens for technical reasons related to matching programs should be abandoned as soon as possible.

2.1.2 Only actual typing assignments and no derived values should be exchanged.

2.2 Antigen and allele assignments must be distinguished from one another. This might be accomplished by information technology systems that provide separate fields for the two assignments, for example, for DR and DRB1. If so, either one or both fields may be filled with typing information.

2.3 A maximum of two antigens can be provided for each locus and/or a maximum of two allele assignments can be provided for each locus. Exceptions are known to occur, but they are extremely rare. These exceptions may be noted in a remarks field.

2.4 If the identity of an antigen is unknown, further testing using a DNA-based method is recommended to clarify the assignment. The code X may not be used to indicate unclear serologic assignments, i. e. a testing result like A2, A'X' will not be accepted.

2.5 If the typing suggests that a previously unknown allele is probably present, the result should be reported as the special code NEW. Once the allele has been characterized and an allele assignment received from the WHO HLA Nomenclature Committee, the typing should be updated.

2.6 Corresponding antigen and allele assignments, if submitted together, must be consistent with WHO assignments or other WMDA-approved resources.

2.7 If corresponding antigen and allele assignments are available, the number of entries must not differ except in one of the following cases:

2.7.1 One of the two alleles is expressed at a low level or not at all, for example, B*07:04, B*15:26N corresponds to B7 only.

2.7.2 One of the two alleles does not have a corresponding known antigen; for example, C*01:03, C*12:05 corresponds to Cw1 only.

2.7.3 Both alleles correspond to the same antigen; for example, DRB1*11:01, DRB1*11:04 corresponds to DR11 only.

2.7.4 A single code is an appropriate multi-allele assignment for both antigens; for example, B62, B75 is consistent with B*15:XX only.

2.8 Individuals tested as carrying a single HLA antigen or allele must be listed as the single assignment only, unless family segregation or an appropriate testing method have been used to underpin homozygosity.

2.8.1 The single assignment must be accompanied by an empty field. Any other symbol or letter code as a placeholder to indicate the single assignment is not allowed, for example, A2 only and not A2, A2

2.8.2 The empty field should be treated either as a blank (meaning unknown type) or as a second copy of the type entered in the first field (i. e., a homozygote).

2.8.3 It is acceptable to convert the HLA assignments of volunteer donors with two identical assignments at a locus (homozygous) and already in the registry database back to one single assignment. It is unlikely that segregation information is available on these donors and matching programs are unlikely to be influenced by this conversion. This should not be done if the two assignments are listed with multiple allele codes (e. g. DRB1*12:CVT, DRB1*12:CVT).

2.8.4 HLA assignments of patients with two identical assignments at a locus (homozygous) should not be converted back to one single assignment, as this is regarded as a loss of information. For a patient, it is likely that homozygosity was determined by family segregation.

2.8.5 A result of a verification typing including one untested allele assignment must be reported using the special code UUUU. Such a result will only be accepted when a complete and consistent low-resolution allele assignment is known beforehand; for example, DRB1*01:XX, DRB1*04:XX before DRB1*UUUU, DRB1*04:01. This code must not be used when reporting donor phenotypes for matching purposes.

2.9 Concerning secondary DRB loci, that is, DRB3, DRB4, DRB5

2.9.1 Secondary DRB loci assignments are not required.

2.9.2 Typing one secondary DRB locus does not imply that the other secondary DRB loci have been tested.

2.9.3 Typing a secondary DRB locus as present with no allele subset designated can be reported using the special code XXXX.

2.9.4 Typing a secondary DRB locus that shows the absence of any allele can be reported using the special code NNNN.

2.9.5 More than two alleles present in one or more of the three secondary DRB loci should be regarded as implausible and should generate a warning.

2.9.6 Rare DRB1–DRB3 (or DRB4 or DRB5) combinations, e. g. DRB1*01:01, DRB5*01:01, may be flagged, but should not be rejected. The overall frequency of such uncommon combinations should be monitored as part of quality assurance measures

2.9.7 Secondary DRB loci assignments may be used, when present, as a check for the two-digit DR or DRB1 assignment. For example, a DR52- (or DRB3-) positive typing should include at least one DR3, 5, 6, 11, 12, 13, 14, 1403, 1404, 17 or 18 or their comparable DNA-based assignment. It should be noted that exceptions to these associations do exist (e. g., DRB1*15 without DRB5, and vice versa) so violations of the known DRB locus associations should be used only to generate warnings but not to flag errors.

Continuously updated lists of assignments and correspondences approved in the context of these guidelines are maintained on two WMDA-specified reference web sites: the WHO web site for HLA nomenclature [5] and the US National Marrow Donor Program web site for codes that include multiple allele alternatives [7]. The latter site includes contact information for registry requests of new multiple allele codes. All entries on the reference web sites are assigned a first and, if applicable, a last validity date. Changes will be announced on the reference web sites and will become valid 48 h later.

Table 1: Reference web sites [3,4,5,7,8,9]

Web address		Information provided
hla.alleles.org/wmda/index.html		
→ Files:	hla_nom.txt	Current, renamed and deleted HLA antigens and alleles
	hla_nom_g.txt	Current G group definitions
	hla_nom_p.txt	Current P group definitions
	rel_ser_ser.txt	Relationship between all current serologically defined HLA antigens
	rel_dna_ser.txt	All current HLA alleles and serologically equivalent antigens
bioinformatics.nmdp.org/HLA/Allele_Codes/Allele_Codes.aspx		List of letter codes for encoding of allelic ambiguities
www.ebi.ac.uk/ipd/imgt/hla/		IMGT/HLA database
www.ebi.ac.uk/ipd/imgt/hla/dictionary.html		Serologic information available for each HLA allelic product
www.ebi.ac.uk/ipd/imgt/hla/searchdet.html		Search determinants used by several registries

References

[1] Bochtler W, Maiers M, Oudshoorn M, Marsh SGE, Raffoux C, Mueller C *et al.* World Marrow Donor Association guidelines for use of HLA nomenclature and its validation in the data exchange among hematopoietic stem cell donor registries and cord blood banks. *Bone Marrow Transplant* 2007; **39**(12): 737–741.

[2] Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA *et al.* An update to HLA Nomenclature, 2010. *Bone Marrow Transplant* 2010; **45**(5): 846-848.

[3] Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA *et al.* Nomenclature for factors of the HLA system, 2010. *Tissue Antigens* 2010; **75**(4): 291-455.

[4] Robinson J, Mistry K, McWilliam H, Lopez R, Parham P, Marsh SGE. The IMGT/HLA database. *Nucleic Acids Res* 2011; **39**(Database issue): D1171-1176.

[5] WMDA reference website:
http://hla.alleles.org/wmda/index.html (Accessed on 01 March 2013).

[6] Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA *et al.* Nomenclature for factors of the HLA system, 2002. *Tissue Antigens* 2002; **60**(5): 407-464.

[7] WMDA reference website:
http://bioinformatics.nmdp.org/HLA/Allele_Codes/Allele_Codes.aspx (Accessed on 01 March 2013).

[8] Holdsworth R, Hurley CK, Marsh SGE, Lau M, Noreen HJ, Kempenich JH *et al.* The HLA Dictionary 2008: a summary of HLA-A, -B, -C, -DRB1/3/4/5, -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR and -DQ antigens. *Tissue Antigens* 2009; **73**(2): 95-170.

[9] Hurley CK, Setterholm M, Lau M, Pollack MS, Noreen H, Howard A *et al.* Hematopoietic stem cell donor registry strategies for assigning search determinants and matching relationships. *Bone Marrow Transplant* 2004; **33**(4): 443–450.