WHO-IUIS Nomenclature Subcommittee for Immunoglobulins and T cell receptors report

September 15, 2008
54th International Union of Immunological Societies Council Meeting, 4th Congress of the Federation of Immunological Societies of Asia-Oceania (FIMSA 2008), October 17-20, 2008, Tapei, Taiwan.

Standing Committee for Immunoglobulins and T cell receptors:
Marie-Paule Lefranc,
IMGT, Laboratoire d'ImmunoGénétique Moléculaire LIGM, Université Montpellier 2 and UPR CNRS, Institut de Génétique Humaine IGH, Montpellier,
141 rue de la Cardonille
34396 Montpellier cedex 5, France,
Marie-Paule.Lefranc@igh.cnrs.fr

1. The WHO-IUIS Nomenclature SubCommittee

Members of the WHO-IUIS Nomenclature Subcommitee/IMGT Nomenclature (IMGT-NC) SubCommittee for immunoglobulins and T cell receptors: Donald Capra (USA) , Max Cooper (USA) , Tasuku Honjo (Japan), Leroy Hood (USA) , Gérard Lefranc (France), Marie-Paule Lefranc (France), Fumihiko Matsuda (Japan), Hans Zachau (Germany).

Experts 2007-2008: Cynthia L. Baldwin (USA), Eva Bengtén (USA), Pierre Boudinot (France), Salvatrice Ciccarese (Italy), Deborah Dunn-Walters (UK), Jean-Pol Frippiat (France), Evelyne Jouvin-Marche (France), Serge Muyldermans (Belgium), Bettina Wagner (USA).

Since the creation of IMGT®, the international ImMunoGeneTics information system®, http://imgt.cines.fr, in 1989, at New Haven during the 10th Human Genome Mapping Workshop (HGM10), the standardized classification and nomenclature of the immunoglobulins and T cell receptors of human and other vertebrate species have been under the responsibilility of the IMGT Nomenclature Committee (IMGT-NC). In 1995, following the first demonstration on-line of the nucleotide database IMGT/LIGM-DB at the 9th International Congress of Immunology in San Francisco, the IMGT-NC SubCommittee for immunoglobulins and T cell receptors has become the WHO-IUIS Nomenclature SubCommittee.

The WHO-IUIS Nomenclature SubCommittee for immunoglobulins and T cell receptors works in close collaboration with the IMGT-NC SubCommittee for the immunoglobulin superfamily (IgSF) and the major histocompatibility complex

superfamily (MhcSF), the Human Genome Organisation (HUGO) Nomenclature Committee (HGNC), the Mouse Genomic Nomenclature Committee (MGNC), the Nomenclature Committees of newly sequenced genomes, the national and international Immunology, Immunogenetics and Genetics Societies, the editors and publishers for recommendations to Authors.

2. The IMGT-ONTOLOGY axioms and concepts

The WHO-IUIS Nomenclature SubCommittee for immunoglobulins and T cell receptors follows the rules for the nomenclatures, as described in the IMGT Scientific chart, http://imgt.cines.fr [1]. These rules are based on the axioms and concepts of IMGT-ONTOLOGY [2-4], the first ontology in immunogenetics and immunoinformatics. The main axioms and concepts of IMGT-ONTOLOGY were summarized in the 2007 report [5, 6]. Among these main axioms, the CLASSIFICATION axiom has generated the concepts of classification, which have been necessary to propose a standardized nomenclature for the immunoglobulin (IG) and T cell receptor (TR) genes [for review, see 7-10, for more details, see the nineteen "IMGT Locus in Focus" publications, freely available at http://imgt.cines.fr/textes/IMGTindex/imgtfocus.html]. The concepts of classification take into account the highly polymorphic multigenic loci and subgroups to which the IG and TR genes belong, their rearrangements and their allelic polymorphisms. These concepts are used whatever the antigen receptor (IG or TR), whatever the locus (for mammals, for example, immunoglobulin heavy IGH, immunoglobulin kappa IGK, immunoglobulin lambda IGL, T cell receptor alpha TRA, T cell receptor beta TRB, T cell receptor gamma TRG and T cell receptor delta TRD), whatever the gene configuration (germline, undefined or rearranged), and whatever the species, from fish to human. Among the concepts of classification, the "Group", "Subgroup", "Gene" and "Allele" concepts are essential and were described in the 2007 report [5, 6].

3. Immunoglobulin and T cell receptor genes and alleles

3.1. Brief history

The IMGT® gene nomenclature has been approved at the international level by the Human Genome Organisation (HUGO) Nomenclature Committee (HGNC), in 1999. Two FactsBooks on the human IG and TR genes and alleles were published in 2001 [9, 10]. The IMGT-NC received the official delegation from HGNC for the IG and TR gene and allele nomenclature in 2002 [11]. The IMGT® genes were entered in Entrez Gene at the National Center for Biotechnology Information (NCBI) [12]. In order to manage the IG and TR genes and alleles, IMGT/GENE-DB, the first IMGT® genome database, was created in January 2003 [13]. The IMGT® IG and TR gene names are the official references for the genome projects and, as such, have been integrated in the MapViewer at NCBI,

on the Ensembl server [14] at the European Bioinformatics Institute (EBI) in 2006, and in the Vega [15] database at the Wellcome Trust Sanger Institute in 2008. All the mouse IMGT® gene and allele names and the corresponding IMGT reference sequences were provided to HGNC and to the Mouse Genome Informatics (MGI) Mouse Genome Database (MGD) [16] in July 2002 and were presented by IMGT® at the 19th International Mouse Genome Conference IMGC 2005, in Strasbourg, France, and entered in IMGT/GENE-DB. IMGT reference sequences have been defined for each allele of each gene based on one or, whenever possible, several of the following criteria: germline sequence, first sequence published, longest sequence, mapped sequence.

3.2. Present situation

IMGT/GENE-DB [13], the comprehensive IMGT® genome database, is the official repository of all the IG and TR genes and alleles approved by the World Health Organization (WHO)/International Union of Immunological Societies (IUIS) Nomenclature Subcommittee for IG and TR [5, 6]. IMGT/GENE-DB is freely available at the IMGT Home Page http://imgt.cines.fr. Reciprocal links exist between IMGT/GENE-DB and the HGNC database [17] and Entrez Gene [12]. IMGT-GENE-DB allows a query per gene and allele name. IMGT/GENE-DB interacts dynamically with IMGT/LIGM-DB [18], the IMGT® comprehensive nucleotide sequence database for IG and TR (126 667 sequences from 223 species, in September 2008) to download and display human, mouse and rat gene-related sequence data. This is the first example of an interaction between IMGT® databases using the concepts of classification.
In September 2008, IMGT/GENE-DB contained 1 911 IG and TR genes from human, mouse and rat, and 2 909 alleles. IMGT/GENE-DB comprises genes and alleles for the seven loci (IGH, IGK, IGL, TRA, TRB, TRG and TRD) of human (genes and alleles) and mouse (genes and alleles). A major update in 2008 has been the entry of 395 IGH and IGL genes of *Rattus norvegicus*, and 398 alleles (with a provisional nomenclature for the IGHV genes as there are still gaps in the rat IGH locus). This allowed to enter the sequences in the IMGT reference directory of IMGT/V-QUEST [19], the IMGT sequence analysis tool for repertoire analysis. These data and annotations were eagerly awaited as the rat is frequently used as an animal model. Moreover that species is widely used for the obtention of monoclonal antibodies. The switch to the definitive nomenclature will be done, as this was done previously for the mouse genes, as soon as the gaps are filled in in the rat genome.
The IG and/or TR genes from 50 other species are currently available in the IMGT Repertoire in “Gene tables” and in “Alignments of alleles”. With the completion of the genomes and the sequence annotation, these genes will be progressively entered in IMGT/GENE-DB.

4. Conclusion and perspectives

The standardization of the IG and TR genes and alleles by the IMGT-NC and the WHO-IUIS Nomenclature SubCommittee for immunoglobulins and T cell receptors represents a major breakthrough in immunogenetics. This has been translated in the IMGT® databases, tools and Web resources through the huge work performed by the IMGT® team and its constant motivation and expertise. These combined efforts contributed to make IMGT®, the international ImMunoGeneTics information system®, http://imgt.cines.fr, the global reference in immunogenetics and immunoinformatics. The IMGT® Web server at Montpellier receives more than 150,000 requests per month, from Europe, the USA and the rest of the world. IMGT® is widely used by clinicians and biological scientists from both academic and industrial laboratories, in diverse research domains: (i) fundamental and medical research (repertoire analysis of the IG antibody sites and of the TR recognition sites in normal and pathological situations such as autoimmune diseases, infectious diseases, AIDS, leukemias, lymphomas, myelomas), (ii) veterinary research (IG and TR repertoires in farm and wild life species), (iii) genome diversity and genome evolution studies of the adaptive responses, (iv) structural evolution of the IgSF and MhcSF proteins, (v) biotechnology related to antibody engineering (single chain Fragment variable (scFv), phage displays, combinatorial libraries, chimeric, humanized and human antibodies), (vi) diagnostics (clonalities, detection and follow up of residual diseases) and (vii) therapeutical approaches (grafts, immunotherapy, vaccinology).

References

[1] Lefranc, M.-P., Giudicelli, V., Kaas, Q., Duprat, E., Jabado-Michaloud, J., Scaviner, D., Ginestoux, C., Clément, O., Chaume, D. and Lefranc, G. (2005) IMGT, the international ImMunoGeneTics information system®. Nucleic Acids Res. 33, D593-D597.

[2] Giudicelli, V. and Lefranc, M.-P. (1999) Ontology for Immunogenetics: IMGT-ONTOLOGY. Bioinformatics, 15, 1047-1054.

[3] Lefranc, M.-P., Giudicelli, V., Ginestoux, C., Bosc, N., Folch, G., Guiraudou, D., Jabado-Michaloud, J., Magris, S., Scaviner, D., Thouvenin, V., Combres, K., Girod, D., Jeanjean, S., Protat, C., Yousfi Monod, M., Duprat, E., Kaas, Q., Pommié, C., Chaume, D. and Lefranc, G. (2004) IMGT-ONTOLOGY for Immunogenetics and Immunoinformatics, http://imgt.cines.fr. Epub 2003, 0004, 22 Nov 2003. In Silico Biol. 4, 17-29.

[4] Duroux, P., Kaas, Q., Brochet, X., Lane, J., Ginestoux, C., Lefranc, M.-P. and Giudicelli, V. (2008) IMGT-Kaleidoscope, the Formal IMGT-ONTOLOGY paradigm. Biochimie, 90, 570-583.

[5] Lefranc, M.-P. (2008) WHO-IUIS Nomenclature Subcommittee for Immunoglobulins and T cell receptors report Immunoglobulins and T cell receptors report August 2007, 13th International Congress of Immunology, Rio de Janeiro, Brazil. Dev. Comp. Immunol., 32, 461-463.

[6] Lefranc, M.-P. (2007) WHO-IUIS Nomenclature Subcommittee for Immunoglobulins and T cell receptors report. Immunogenetics, 59, 899-902.

[7] Lefranc, M.-P. (2000) Nomenclature of the human immunoglobulin genes. In Coligan, J.E., Bierer, B.E., Margulies, D.E., Shevach, E.M. and Strober, W. (eds.), Current Protocols in Immunology, John Wiley & Sons, Inc., NJ, Hoboken, pp. A.1P.1-A.1P.37.

[8] Lefranc, M.-P. (2000) Nomenclature of the human T cell receptor genes. In Coligan, J.E., Bierer, B.E., Margulies, D.E., Shevach, E.M. and Strober, W. (eds.), Current Protocols in Immunology, John Wiley & Sons, Inc., NJ, Hoboken, pp. A.1O.1-A.1O.23.

[9] Lefranc, M.-P. and Lefranc, G. (2001) The Immunoglobulin FactsBook, Academic Press, 458 pages.

[10] Lefranc, M.-P. and Lefranc, G. (2001) The T cell receptor FactsBook, Academic Press, 398 pages.

[11] Wain, H.M., Bruford, E.A., Lovering, R.C., Lush, M.J., Wright, M.W. and Povey S. (2002) Guidelines for human gene nomenclature, Genomics 79, 464-470.

[12] Maglott, D., Ostell, J., Pruitt, K.D. and Tatusova, T. (2007) Entrez Gene: gene-centered information at NCBI. Nucl. Acids Res. 35, D26-31.

[13] Giudicelli, V., Chaume, D. and Lefranc, M.-P. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. Nucl. Acids Res. 33, D256-261 (2005).

[14] Cochrane, G., Akhtar, R., Aldebert, P., Althorpe, N., Baldwin, A., Bates, K., Bhattacharyya, S., Bonfield, J., Bower, L., Browne, P. *et al*. (2008) Priorities for nucleotide trace, sequence and annotation data capture at the Ensembl Trace Archive and the EMBL Nucleotide Sequence Database**.** Nucleic Acids Res., 36, D5-D12.

[15] Wilming, L.G., Gilbert, J.G.R., Howe, K., Trevanion, S., Hubbard, T. and Harrow, J.L. (2008) The vertebrate genome annotation (Vega) database. Nucleic Acids Res. 36, D753-D760.

[16] Bult, C.J., Eppig, J.T., Kadin, J.A., Richardson, J.E., Blake, J.A., and the Mouse Genome Database Group. (2008) The Mouse Genome Database (MGD): mouse biology and model systems. Nucleic Acids Res., 36, D724-D728.

[17] Bruford, E.A., Lush, M.J., Wright, M.W., Sneddon, T.P., Povey, S.,and Birney, E. (2008) The HGNC Database in 2008: a resource for the human genome. Nucleic Acids Res., 36, D445-D448.

[18] Giudicelli, V. Duroux, P., Ginestoux, C., Folch, G., Jabado-Michaloud, J., Chaume, D. and Lefranc, M.-P. (2006) IMGT/LIGM-DB, the IMGT®

comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. Nucleic Acids Res. 34, D781-D784.

[19] Brochet, X., Lefranc, M.-P. and Giudicelli,V. (2008) IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. Nucleic Acids Res, 36, W503-508.