2007- 53rd IUIS Council Meeting, August 27, 2007, Rio de Janeiro, Brazil
at the 13th International Congress of Immunology, Rio de Janeiro, Brazil

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

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 2006-2007: Cynthia L. Baldwin (USA), Eva Bengtén (USA), Pierre Boudinot (France), Deborah Dunn-Walters (UK), Jean-Pol Frippiat (France), Evelyne Jouvin-Marche (France), Serge Muyldermans (Belgium), Bettina Wagner (USA).

Standing Committee for Immunoglobulins and T cell receptors: Marie-Paule Lefranc, IMGT, Laboratoire d'ImmunoGénétique Moléculaire, Université Montpellier 2 and UPR CNRS, Institut de Génétique Humaine, Montpellier, France, Marie-Paule.Lefranc@igh.cnrs.fr.

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 concepts of IMGT-ONTOLOGY [2,3], the first ontology in immunogenetics and immunoinformatics. IMGT-ONTOLOGY manages the immunogenetics knowledge through diverse facets relying on seven axioms, "IDENTIFICATION", "DESCRIPTION", "CLASSIFICATION", "NUMEROTATION", "LOCALIZATION", "ORIENTATION" and "OBTENTION", that postulate that objects, processes and relations have to be identified, described, classified, numerotated, localized,

orientated, and that the way they are obtained has to be determined. These axioms constitute the Formal IMGT-ONTOLOGY, also designated as IMGT-Kaleidoscope (Fig. 1).

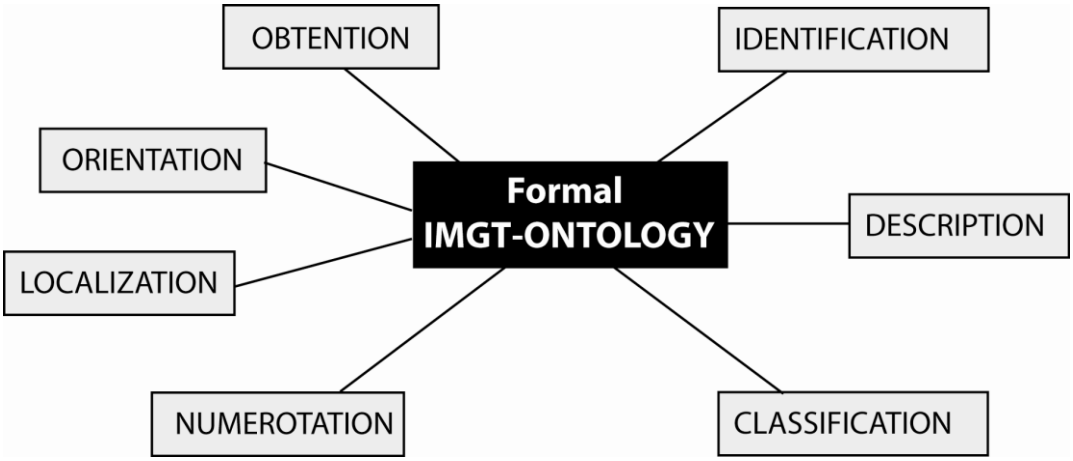

Fig. 1. The axioms of the Formal IMGT-ONTOLOGY or IMGT-Kaleidoscope.

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. These concepts take into account the highly polymorphic multigenic loci and families to which the IG and TR 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.

2.1. The "Group" and "Subgroup" concepts

The "Group" concept classifies a set of genes which belong to the same multigene family, within the same species or between different species. For the IG and TR, the set of genes is identified by an instance of the "GeneType" concept (V, D, J or C). The "Subgroup" concept classifies a subset of genes which belong to the same group, and which, in a given species, share at least 75% of identity at the nucleotide sequence level (and in the germline configuration for the V, D, and J genes) (Fig. 2).

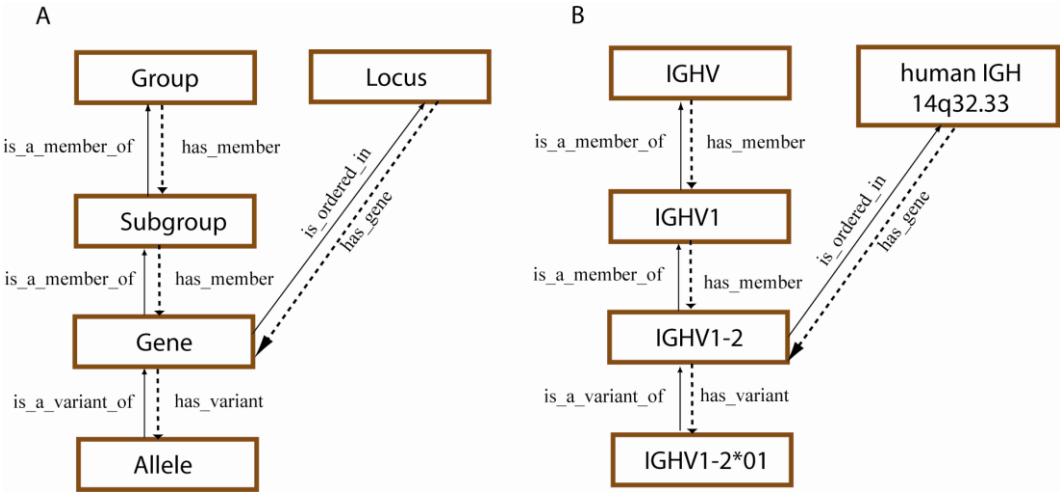

Fig. 2. Concepts of classification for gene and allele nomenclature (CLASSIFICATION axiom). (A) Hierarchy of the concepts of classification and their relations. (B) Examples of concept instances for each concept of classification. The concept instances are associated to an instance of the "Taxon" concept, and more precisely for the "Gene" and "Allele" concepts to an instance of the "Species" concept (here, *Homo sapiens*). The "Locus" concept is a concept of localization (LOCALIZATION axiom).

2.2. The "Gene" and "Allele" concepts

The "Gene" concept classifies a unit of DNA sequence that can be potentially transcribed and/or translated (this definition includes the regulatory elements in 5' and 3', and the introns, if present). The instances of the "Gene" concept are gene names. In IMGT-ONTOLOGY, a gene name is composed of the name of the species (instance of the Taxon "Species" concept) and of the international HGNC/IMGT gene symbol, for example, *Homo sapiens* IGHV1-2 (human immunoglobulin heavy variable 1-2) (Fig. 2). By extension, orphons and pseudogenes are also instances of the "Gene" concept. The "Allele" concept classifies a polymorphic variant of a gene. The instances of the "Allele" concept are allele names. Alleles identified by the mutations of the nucleotide sequence are classified by reference to allele *01. Full description of mutations and allele name designations are currently recorded for the core sequences (V-REGION, D-REGION, J-REGION, C-REGION).

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. The IMGT® genes were entered in the Genome Database (GDB) and in LocusLink at the National Center for Biotechnology Information (NCBI) in 1999-2000, and in Entrez Gene (NCBI) [4], when this gene database superseded LocusLink. Two FactsBooks on the human IG and TR genes and alleles were published in 2001 [5,6]. The IMGT-NC received the official delegation from HGNC for the IG and TR gene and allele nomenclature in 2002 [7]. In order to manage the IG and TR genes and alleles, IMGT/GENE-DB, the first IMGT genome database, was created in January 2003 [8]. 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 and, in 2006, in the Ensembl server at the European Bioinformatics Institute (EBI).

3.2. Present situation

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 [9]. 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).

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

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

[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. IMGT-ONTOLOGY for Immunogenetics and Immunoinformatics, http://imgt.cines.fr. Epub 2003, 0004, 22 Nov 2003, *In Silico* Biol. 4, 17-29 (2004).

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

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

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

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

[8] Giudicelli V., Chaume D., 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).

[9] Duroux P., Giudicelli V., Kaas Q., Jabado-Michaloud J., Folch G., Ginestoux C., Brochet X., Lane J., Régnier L., Wu Y., Garapati V.P., Bellahcene F., Servier E.-J., Ehrenmann F., Lefranc G. and Lefranc M.-P. IMGT®, the international ImMunoGeneTics information system®, the reference in immunogenetics and immunoinformatics. In: Proceedings First International Congress on Macromolecular Biochemistry and Genetics (Ed. Fattoum A.), Gafsa, Tunisia, 12-16 April 2007 (in press).