Report 2017

IUIS Immunoglobulins (IG), T cell Receptors (TR) and Major Histocompatibility (MH) Nomenclature SubCommittee (IMGT-NC)

66th IUIS Council Meeting, December 2, 2017, Hammamet, Tunisia, on occasion of the 10th African Congress of Immunology FAIS (December 3-7, 2017)

Chair
Marie-Paule Lefranc (Marie-Paule.Lefranc@igh.cnrs.fr)
Professor Emeritus University of Montpellier (France)
IMGT® Founder and Executive Director Emeritus
Institut de Génétique Humaine IGH, UMR9002 CNRS, University of Montpellier
Web page: www.imgt.org

List of members of the committee
Max Cooper (USA)
Tasuku Honjo (Japan)
Leroy Hood (USA)
Gérard Lefranc (France)
Marie-Paule Lefranc (France)
Fumihiko Matsuda (Japan)
Hans Zachau (Germany)
Cynthia L. Baldwin (USA)
Eva Bengtén (USA)
Pierre Boudinot (France)
Felix Breden (Canada)
Salvatrice Ciccarese (Italy)
Deborah Dunn-Walters (UK)
Jean-Pol Frippiat (France)
Véronique Giudicelli (France)
Evelyne Jouvin-Marche (France)
Sofia Kossida (France)
Véronique Laurens (France)
Ramit Mehr (Israel)
Serge Muyldermans (Belgium)
Mariano Sanchez-Lockhart (USA)
Jamie Scott (Canada)
Bettina Wagner (USA)
Corey T. Watson (USA)

Previous member
Donald Capra (USA) Founding member (1937-2015)

I. Mission

The IUIS 'Immunoglobulins (IG), T cell receptors (TR) and Major Histocompatibility (MH) Nomenclature Sub-Committee (IMGT-NC)' mission is to report to the IUIS Nomenclature committee on the standardized classification and nomenclature of the immunoglobulins (IG), T cell receptors (TR) and major histocompatibility (MH) genes and proteins of any vertebrate species with jaws (*gnathostomata*, from fishes to humans).
IMGT-NC genes and alleles are managed by IMGT®, the international ImMunoGeneTics information system® http://www.imgt.org, created in 1989 by Marie-Paule Lefranc, and the global reference in immunogenetics and immunoinformatics.

II. Members

Founding Members of the WHO-IUIS Nomenclature Subcommittee for immunoglobulins (IG), T cell receptors (TR) and major histocompatibility (MH) (IMGT NC) are Donald Capra (USA) (1937-2015) (www.jimmunol.org/content/194/12/5575.full), Max Cooper (USA), Tasuku Honjo (Japan) Leroy Hood (USA), Gérard Lefranc (France), Marie-Paule Lefranc (France), Fumihiko Matsuda (Japan) and Hans Zachau. Members of the Sub-Committee are experts and contributors to the field of immunogenetics who in their published work have promoted standardization of the IMGT IG, TR and/or MH genes and alleles. Associated to IMGT-NC are the IMGT Experts who are contributors on a case by case basis for IG, TR and/or MH loci of given species (http://www.imgt.org/IMGTindex/IMGTexperts.php).

Summary reports written by the IMGT-NC are sent on an annual basis to the chair of the WHO-IUIS Nomenclature (Michel Kazatchkine (Sept 1992-2004), Laurence Boumsell (Sept 2004-2010), Pablo Engel (Sept 2010-2016), Menno van Zelm (Sept 2016-)), for presentation at the annual IUIS Council Meetings (coincident every three years with the International Congress of Immunology (ICI). The links to the IMGT-NC events, publications and/or reports since 1989 are publicly available at http://www.imgt.org/IMGTindex/IUIS-NC.php.

III. New IG and TR genes and alleles

The biocuration of new IG and TR genes and alleles is done by the IMGT® team. This includes polymorphic genes from humans and genes of newly sequenced genomes of any species of vertebrates with jaws (*gnasthomata*) others than humans.

The complete annotation of new IG and TR genes and alleles is performed based on the IMGT-ONTOLOGY concepts and IMGT Scientific chart rules: standardized identification (keywords), description (labels), classification (genes and alleles nomenclature) and numerotation (IMGT unique numbering).

New IG and TR genes and alleles are entered in the IMGT web resources (IMGT Repertoire Gene tables, Alignments of alleles, Protein displays, etc.), databases (IMGT/LIGM-DB, IMGT/GENE-DB, IMGT/3Dstructure-DB and IMGT/2Dstructure-DB, etc) and tools (IMGT/V-QUEST, IMGT/DomainGapAlign, etc) and are publicly and freely available to the academics.

Following the entry of a new IG or TR gene in IMGT/GENE-DB – which marks its official approval by IMGT® – the IMGT-NC chair informs the Human Genome Organization (HUGO), Gene Nomenclature Committee (HGNC), Vega Genome Browser, and the National Center for Biotechnology Information (NCBI Gene) of the new gene and of its IMGT gene name (symbol and definition, reference sequence, genome localization, functionality). Reciprocal links for individual entries are created between the four sites IMGT, HGNC, Vega and NCBI Gene.

IG and TR alleles are under the entire responsibility of IMGT® since its creation in 1989 (Human Gene Mapping 10 workshop, New Haven). They were published for the first time in the Immunoglobulin FactsBook (2001) and in the T cell receptor FactsBook (2001). The IG and TR alleles are currently managed in IMGT/GENE-DB and displayed in Alignments of alleles. There are fully annotated and entered in the IMGT® web resources, databases and tools as described above.

On November 11, 2017, IMGT/GENE-DB, the IMGT® gene database contained 4,785 IMGT genes and 6,693 alleles from 24 species.
• 706 IG and TR genes and 1,350 alleles for *Homo sapiens* (462 IG genes and 927 alleles; 244 TR genes and 423 alleles) as well as 11 RPI genes and 135 alleles
• 686 IG and TR genes and 746 alleles for *Macaca* (478 genes and 529 alleles for *Macaca mulatta*),

• 892 IG and TR genes and 1,345 alleles for *Mus* (871 genes and 1321 alleles for *Mus musculus*),
• 638 IG and TR genes and 644 alleles for *Rattus norvegicus*.
More statistics are available at http://www.imgt.org/genedb/stats.
Updates of the IMGT reference directories are reported in the Documentation of the corresponding IMGT® tools.

IV. Recent accomplishments

1) Entry of 107 *Homo sapiens* IG variable genes in UniProt.

IMGT/GENE-DB reference sequences of 107 *Homo sapiens* IG genes were provided by IMGT to UniProt and the translation of these genes entered in the database by the UniProt team. They include 37 IGHV, 38 IGKV and 32 IGLV genes. Reciprocally, cross-references were added from IMGT/GENE-DB to UniProt for these entries.
This collaborative work between IMGT and UniProt on the sequences of the IGHV, IGKV and IGLV genes is a major step, at the international level, based on the IMGT gene nomenclature (IMGT gene names approved by HGNC in 1999 and endorsed by NCBI in 2000).
The IMGT alleles of the IG genes are defined at the nucleotide levels, and allele query should be done in IMGT/GENE-DB, however for completeness and correspondence between the databases, the IMGT allele of the sequence provided by IMGT is reported in UniProt.

2) Nomenclature and next generation sequencing (NGS) of IG and TR.

IMGT/HighV-QUEST, http://www.imgt.org/HighV-QUEST/login.action, the web portal created in October/November 2010, is the first and currently the only online tool freely available for academics for the analysis of Next Generation Sequencing (NGS) antigen receptor (IG and TR) repertoires in normal and pathological situations. IMGT/HighV-QUEST analyses 500,000 sequences per batch and may compare up to one million outputs for the identification and characterization of IMGT clonotypes (AA) which include the IG and TR gene and allele names.
IMGT/StatClonotype, http://www.imgt.org/StatClonotype/ is an IMGT® tool for statistical comparison of sets from IMGT/HighV-QUEST output, on the Web since June 2016. IMGT/StatClonotype uses a generic statistical procedure for identifying significant changes in IG and TR differences of proportions of IMGT clonotypes (AA) diversity and expression, bringing an additional level of analysis and comparison between repertoires described using the IMGT IG and TR gene and allele names.

3) Nomenclature and IG and TR amino acid sequences and three-dimensional structures

IMGT/DomainGapAlign analyses the amino acid sequences of IG (or antibodies) and TR. The IMGT IG and TR gene and allele names and the use of the same rules (numbering, CDR-IMGT, FR-IMGT) bridge the gap between sequences and three-dimensional structures, which can be both visualized as IMGT Collier de Perles graphical representations. These standards have been used for the last ten years for the monoclonal antibodies definitions published in the proposed and recommended lists of the WHO International Nonproprietary Names (INN) programme. They can describe any novel format resulting from antibody engineering. The long-term collaboration for standardization started in 1992 has recently been formalized between IUIS and WHO.

4) The following publication in Frontiers in Immunology, at the invitation of the WHO IUIS Nomenclature Committee (Proceedings of ICI Milan 2013)

Lefranc M-P. Immunoglobulin (IG) and T cell receptor genes (TR): IMGT® and the birth and rise of immunoinformatics. Front Immunol. 2014 Feb 05;5:22
https://www.frontiersin.org/articles/10.3389/fimmu.2014.00022/full' has reached 9,474 views on November 22, 2017. UniProt has made a link to that publication in the abstract of the files of the 107 IG variable genes and HGNC has quoted this reference in the two introductory pages on the IG and TR.

V. Ongoing projects

1) Collaboration with HGNC, Vega, Ensembl, NCBI, UniProt.
2) Collaboration with the European Research Initiative on CLL (ERIC).
3) Collaboration with the EuroClonality-NGS consortium.
4) Collaboration with the Adaptive Immune Receptor Repertoire (AAIR) Consortium.
5) Reciprocal links to IUPHAR-DB and to IEDB.
6) Diffusion of the IMGT booklet (11 papers, 144 pages) edited by Cold Spring Harbor Protocols (CSHP). The content of this booklet was detailed in the 2012 Sub-Committee report. CSHP specifically edited the IMGT booklet for educational purposes and authorized IMGT® to have it freely available on the IMGT® site http://www.imgt.org (available in 'IMGT References'). CSHP also authorized that the IMGT booklet be printed and distributed freely. IMGT® databases and tools described in these chapters use the IUIS/IMGT nomenclature approved by Human Genome Organization (HUGO) Nomenclature Committee (HGNC).

VI. Challenges and opportunities in 2018

Large scale genome sequencing

1) IMGT/GENE-DB biocuration and nomenclature of IG and TR genes and alleles of species from newly sequenced genomes.
A landmark is represented by the approved IMGT nomenclature of the IG and TR loci of *Canis lupus familiaris* (Martin et al. 2017) which will be used as a benchmark for other species. The *Canis lupus familiaris* reference sequences have been entered in IMGT/HighV-QUEST reference directories allowing NGS repertoire analysis

2) IMGT-NC is currently developing the concept of localization to annotate and manage the copy number variations (CNV) and polymorphisms by insertion/deletion in the *Homo sapiens* IG and TR loci.

3) A working group (WG) within AAIR intends to analyse the criteria for defining inferred alleles from NGS (the procedure would include a submission of inferred alleles validated by the WG to a generalist database, before submission to the IMGT-NC). If IMGT data quality requirement is preserved, inferred alleles could be considered in the IMGT reference directories of IMGT/V-QUEST and IMGT/HighV-QUEST.

Pespectives

IG, TR and MH standardized nomenclature based on the IMGT-ONTOLOGY concepts of identification (standardized keywords), description (standardized labels), classification (gene and allele nomenclature) and numerotation (IMGT unique numbering and IMGT Collier de Perles) have been crucial in the development of immunoinformatics since its creation in 1989.

These concepts are necessary more than ever in large scale genome sequencing, immune repertoire NGS studies and antigen receptor biotechnology for immunotherapy. Future directions consist in promoting IUIS/IMGT/HGNC nomenclature for new data originating from genome analysis of animal models, veterinary and wild life species, repertoire next generation sequencing and antibody engineering.

VII. Selected recent publications derived from the work of the committee

Rubelt F, Busse CE, Bukhari SAC, Bürckert JP, Mariotti-Ferrandiz E, Cowell LG., Watson CT, Marthandan N, Faison WJ, Hershberg U, Laserson U, Corrie BD, Davis MM, Peters B, Lefranc M-P, Scott JK, Breden F; AIRR Community, Luning Prak ET, Kleinstein SH
Adaptive Immune Receptor Repertoire Community recommendations for sharing immune-repertoire sequencing data
Nat Immunol. 2017 Nov 16; 18(12):1274–1278. doi:10.1038/ni.3873.
PMID: 29144493

Breden F, Luning Prak ET, Peters B, Rubelt F, Schramm CA, Busse CE, Vander Heiden JA, Christley S, Bukhari SAC, Thorogood A, Matsen IV FA, Wine Y, Laserson U, Klatzmann D, Douek DC, Lefranc M-P, Collins AM, Bubela T, Kleinstein SH, Watson CT, Cowell LG, Scott JK, Kepler TB
Reproducibility and reuse of Adaptive Immune Receptor Repertoire data
Front Immunol doi: 10.3389/fimmu.2017.01418

Martin J, Ponstingl H, Lefranc M-P, Archer J, Sargan D, Bradley A.
Comprehensive annotation and evolutionary insights into the canine (*Canis lupus familiaris*) antigen receptor loci.
Immunogenetics. 2017 Sep 19. doi: 10.1007/s00251-017-1028-0. [Epub ahead of print]
PMID: 28924718

Giudicelli V, Duroux P, Kossida S, Lefranc M-P.
IG and TR single chain Fragment variable (scFv) sequence analysis: a new advanced functionality of IMGT/V-QUEST and IMGT/HighV-QUEST
BMC Immunol. 2017 Jun 26;18(1):35. doi: 10.1186/s12865-017-0218-8.
PMID: 28651553

Prieur A, Cappellini M, Habif G, Lefranc M-P, Mazard T, Morency E, Pascussi J-M, Flacelière M, Cahuzac N, Vire B, Dubuc B, Durochat A, Liaud P, Ollier J, Pfeiffer C, Poupeau S, Saywell V, Planque C, Assenat E, Bibeau F, Bourgaux J-F, Pujol P, Sézeur A, Ychou M, Joubert D.
Targeting the Wnt pathway and cancer stem cells with anti-progastrin humanized antibodies: a major breakthrough for K-RAS mutated colorectal cancer treatment
Clinical Cancer Res. 2017 Sep 1;23(17):5267-5280. doi: 10.1158/1078-0432.CCR-17-0533. Epub 2017 Jun 9.
PMID: 28600477

Xochelli A, Baliakas P, Kavakiotis I, Agathangelidis A, Sutton L-A, Minga E, Ntoufa S, Tausch E, Yan XJ, Shanafelt TD, Plevova K, Boudjogra M, Rossi D, Davis Z, Navarro A, Sandberg Y, Vojdeman FJ, Scarfò L, Stavroyianni N, Sudarikov A, Veronese S, Tzenou T, Karan Djurasevic T, Catherwood MA, Kienle D, Chatzouli M, Facco M, Bahlo J, Pott C, Pedersen LB, Mansouri L, Smedby KE, Chu CC, Giudicelli V, Lefranc M-P, Panagiotidis P, Juliusson G, Anagnostopoulos A, Vlahavas I, Antic D, Trentin L, Montillo M, Niemann CU, Dohner H, Langerak AW, Pospisilova S, Hallek M, Campo E, Chiorazzi N, Maglaveras N, Oscier D, Gaidano G, Jelinek DF, Stilgenbauer S, Chouvarda I, Darzentas N, Belessi C, Davi F, Hadzidimitriou A, Rosenquist R, Ghia P, Stamatopoulos K.
Chronic lymphocytic leukemia with mutated IGHV4-34 receptors: shared and distinct immunogenetic features and clinical outcomes
Clin Cancer Res. 2017 Sep 1;23(17):5292-5301. doi: 10.1158/1078-0432.CCR-16-3100. Epub 2017 May 23.
PMID: 28536306

Rosenquist R, Ghia P, Hadzidimitriou A, Sutton L-A, Agathangelidis A, Baliakas P, Darzentas N, Giudicelli V, Lefranc M-P, Langerak AW, Belessi C, Davi F, Stamatopoulos K, on behalf of ERIC, the European Research Initiative on CLL
Immunoglobulin gene sequence analysis in chronic lymphocytic leukemia: updated ERIC recommendations
Leukemia. 2017 Jul;31(7):1477-1481. doi: 10.1038/leu.2017.125. Epub 2017 Apr 25.
PMID: 28439111

Langerak AW, Brüggemann M, Davi F, Darzentas N, van Dongen JJM, Gonzalez D, Cazzaniga G, Giudicelli V, Lefranc M-P, Giraud M, Macintyre EA, Hummel M, Pott C, Groenen PJTA, Stamatopoulos K; EuroClonality-NGS consortium.
High-throughput immunogenetics for clinical and research applications in immunohematology: potential and challenges.
J Immunol. 2017 May 15;198(10):3765-3774. doi: 10.4049/jimmunol.1602050. Epub 2017 Apr 17.
PMID: 28416603

Dambrun M, Dechavanne C, Emmanuel A, Aussenac F, Leduc M, Giangrande C, Vinh J, Dugoujon JM, Lefranc M-P, Guillonneau F, Migot-Nabias F.
Human immunoglobulin heavy gamma chain polymorphisms: molecular confirmation of proteomic assessment.
Mol Cell Proteomics. 2017 May;16(5):824-839. doi: 10.1074/mcp.M116.064733. Epub 2017 Mar 6.
PMID: 28265047

Marillet S., Lefranc M-P., Boudinot P., and Cazals F.
Novel structural parameters of IG - Ag complexes yield a quantitative description of interaction specificity and binding affinity
Front. Immunol. 2017, 8:34. doi: 10.3389/fimmu.2017.00034
PMID: 28232828
http://journal.frontiersin.org/article/10.3389/fimmu.2017.00034/abstract

Aouinti S, Giudicelli V, Duroux P, Malouche D, Kossida S, Lefranc M-P.
IMGT/StatClonotype for Pairwise Evaluation and Visualization of NGS IG and TR IMGT Clonotype (AA) Diversity or Expression from IMGT/HighV-QUEST.
Front Immunol. 2016 Sep 9;7:339. doi: 10.3389/fimmu.2016.00339. eCollection 2016.
PMID: 27667992

Besbes S, Hamadou WS, Boulland ML, Lefranc M-P, Ben Youssef Y, Achour B, Khelif A, Fest T, Soua Z.
Combined IKZF1 and IG markers as new tools for diagnosis and minimal residual disease assessment in Tunisian B-ALL.
Bull Cancer. 2016 Oct;103(10):822-828. doi:10.1016/j.bulcan.2016.07.008.
PMID: 27614734

Linguiti G, Antonacci R, Tasco G, Grande F, Casadio R, Massari S, Castelli V, Consiglio A, Lefranc M-P, Ciccarese S.
Genomic and expression analyses of *Tursiops truncatus* T cell receptor gamma (TRG) and alpha/delta (TRA/TRD) loci reveal a similar basic public γδ repertoire in dolphin and human.
BMC Genomics. 2016 Aug 15;17(1):634. doi: 10.1186/s12864-016-2841-9.
PMID: 27528257

Yu GY, Mate S, Garcia K, Ward MD, Brueggemann E, Hall M, Kenny T, Sanchez-Lockhart M, Lefranc M-P, Palacios G.
Cynomolgus macaque (*Macaca fascicularis*) immunoglobulin heavy chain locus description.
Immunogenetics. 2016 Jul;68(6-7):417-28. doi: 10.1007/s00251-016-0921-2. Epub 2016 May 27.
PMID: 27233955

Boudinot P, Mondot S, Jouneau L, Teyton L, Lefranc M-P, Lantz O.
Restricting nonclassical MHC genes coevolve with TRAV genes used by innate-like T cells in mammals.
Proc Natl Acad Sci U S A. 2016 May 24;113(21):E2983-92. doi: 10.1073/pnas.1600674113. Epub 2016 May 11.
PMID: 27170188

Aouinti, S., Malouche D, Giudicelli V, Kossida, S., Lefranc, M-P
IMGT/HighV-QUEST statistical significance of IMGT clonotype (AA) diversity per gene for standardized comparisons of next generation sequencing immunoprofiles of immunoglobulins and T cell receptors
PLoS One. 2015 Nov 5;10(11):e0142353. doi: 10.1371/journal.pone.0142353. eCollection 2015. PMID: 26540440.
Correction: PLoS ONE 2016 Jan 5;11(1): e0146702. doi: 10.1371/journal.pone.0146702
PMID: 26731095

Montpellier, November 22, 2017