IMGT®, the international ImMunoGeneTics information system®

MARIE-PAULE LEFRANC
Marie-Paule.Lefranc@igh.cnrs.fr

Marie-Paule Lefranc
Emeritus Professor,
University Montpellier 2

Gérard Lefranc,
Emeritus Professor
University Montpellier 2

Patrick Duroux,
Research Engineer CNRS

Véronique Giudicelli,
Research Engineer UM2

Géraldine Folch,
Engineer CNRS

Joumana Jabado-Michaloud,
Engineer CNRS

Chantal Ginestoux,
Technician CNRS

Souphatta Sasorith,
Post-doctoral Fellow

Eltaf Alamyar,
PhD student

Engineers :
Fatena Bellahcene,
Emilie Carillon,
Nelly Jouffre,
Amandine Lacan,
Denis Moreno,
Claire Poiron,
Mansour Saljoqi,
Saida Saljoqi,
Caroline Tournier

Our research activities are focused on molecular immunogenetics, immunoinformatics, bioinformatics and rare human genetic diseases, more particularly Primary ImmunoDeficiencies (PID). We are studying the genetics, structures, functions, as receptors and initiators of signaling pathways, and repertoires of the immunoglobulins (IG) of B lymphocytes and plasmocytes, and of the T cell receptors (TR) on T lymphocytes, which are essential components of the adaptive (specific) immunity in humans and other vertebrates.

In 1989, we created IMGT®, the international ImMunoGeneTics information system® (Montpellier 2 University and CNRS). IMGT®, a CNRS registered trademark (EU, Canada and USA), is the global reference in immunogenetics and immunoinformatics.

This high-quality integrated knowledge resource is specialized in the IG, TR and major histocompatibility (MH) proteins of vertebrate species, and in the immunoglobulin superfamily (IgSF), MH superfamily (MhSF) and related proteins of the immune system (RPI) of any species. IMGT® provides a common access to expertly annotated nucleotide and protein sequences, structural data and genetic information. IMGT® includes seven databases (IMGT/LIGM-DB, a comprehensive database of more than 168,000 IG and TR sequences from human and 300 other vertebrate species in July 2012; IMGT/GENE-DB, IMGT/CLL-DB, IMGT/PRIMER-DB, IMGT/2Dstructure-DB, IMGT/3Dstructure-DB and IMGT/mAb-DB), seventeen interactive tools and more than 10,000 pages of web resources. IMGT/HighV-QUEST analyses Next-Generation Sequencing (NGS) High Throughput Sequencing (HTS)/Next-Generation Sequencing (NGS) data of IG and TR by batch of up to 150,000 sequences.

IMGT/DomainGapAlign is widely used for antibody engineering and design of humanized antibodies as it allows the precise definition of FR-IMGT and CDR-IMGT and the easy comparison of amino acid sequences between the nonhuman (mouse, rat…) V domains and the closest human germline genes. Since July 1995, IMGT® is available on the web at http://www.imgt.org. IMGT® is used by academic and industrial scientists involved in fundamental research, medical research (autoimmune and infectious diseases, AIDS, leukemia, lymphoma, myeloma), veterinary research, genomics (genome diversity and evolution of the adaptive immune system), biotechnology related to antibody engineering for humanization of therapeutic antibodies, diagnostics (detection of minimal residual diseases) and therapeutic approaches (grafts, immunotherapy, vaccinology).

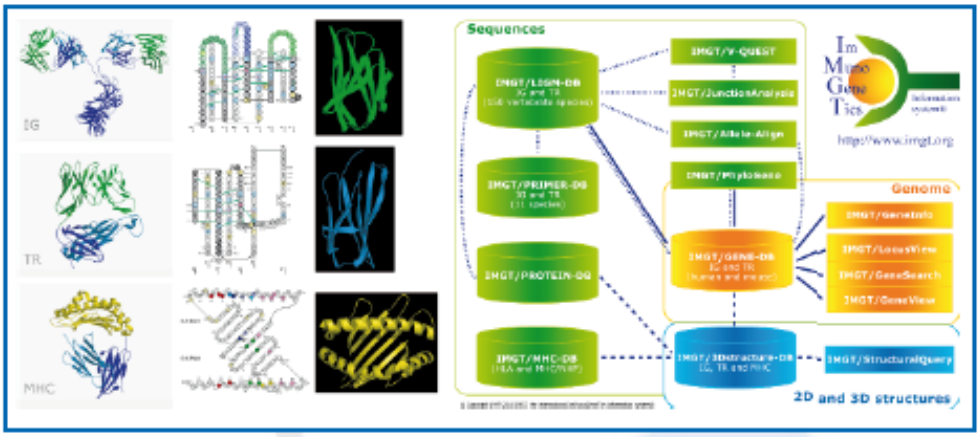
- Alamyar, E., Giudicelli, V., Shuo, L., Duroux, P., Lefranc, M-P. (2012) IMGT/HighV-QUEST: the IMGT® web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep sequencing. **Immunome Res.**, *8*, 1, 26.

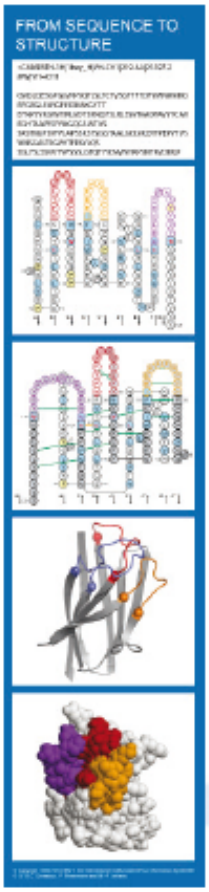
- Giudicelli, V., Lefranc, M-P. (2012) IMGT-ONTOLOGY 2012. Frontiers in Bioinformatics and Computational Biology. **Front. Genet.**, *3*, 79.

- Jabara, HH., McDonald, DR., Janssen, E., Massaad, MJ., Ramesh, N., Borzutzky, A., Rauter, I., Benson, H., Schneider, L., Baxi, S., Recher, M., Notarangelo, LD., Wakim, R., Dbaibo, G., Dasouki, M., Al-Herz, W., Barlan, I., Baris, S., Kutukculer, N., Ochs, HD., Plebani, A., Kanariou, M., Lefranc, G., Reisli, I., Fitzgerald, KA., Golenbock, D., Manis, J., Keles, S., Ceja, R., Chatila, TA., Geha, RS. (2012) DOCK8 functions as an adaptor that links TLR-MyD88 signaling to B cell activation. **Nat Immunol.**, *13*, 6, 612-620.

- Lana, E., Mégarbané, A., Tourrière, H., Sarda, P., Lefranc, G., Claustres, M., De Sario, A. (2012) DNA replication is altered in Immunodeficiency Centromeric instability Facial anomalies (ICF) cells carrying DNMT3B mutations. **Eur J Hum Genet**. *20*, 10, 1044-1050.

- Lefranc, M-P., Lefranc, G. (2012) Human Gm, Km and Am allotypes and their molecular characterization: a remarkable demonstration of polymorphism. **Methods Mol Biol.**, *882*, 635-680.





The IMGT® web server at Montpellier is accessed by more than 80,000 sites per year. IMGT® has an exceptional response with more than 150,000 requests per month. The IMGT® platform has been certified ISO 9001:2008 by LRQA France, in October 2010.

Antibodies represent a large number of the pharmaceutical substances submitted to the World Health Organization/International Nonproprietary Names (WHO/INN) Programme. The INN definition of antibodies is based on the IMGT-ONTOLOGY concepts. Since 2008, amino acid sequences of monoclonal antibodies (mAb, INN suffix -mab), of fusion proteins for immune applications (FPIA, INN suffix -cept) and composite proteins for clinical applications (CPCA) from WHO/INN have been entered into IMGT®. These therapeutic applications emphasize the importance of the IMGT-ONTOLOGY concepts in bridging the gap between antibody sequences and 2D and 3D structures.

Another great research interest, in collaboration with the Unit of Medical Genetics, St-Joseph University, Beirut (Pr A. Mégarbané), concerns very rare autosomal recessive genetic diseases in consanguineous Lebanese families. Patients are autozygous (homozygous by descent) for a very rare mutated gene present in the common ancestor of their parents who are cousins. These pathologies, almost unknown in panmictic populations, are invaluable starting points from which to identify unknown genes, their products and functions as well as links, unsuspected in populations without consanguinity, with cell physiology. For examples, the ICF (Immunodeficiency, Centromeric region instability and Facial anomalies) syndrome results from mutations either in the DNA methyltransferase 3B (DNMT3B) gene in most cases (type 1) or in the ZBTB24 gene (type 2); there are two forms of the Hyper IgE syndrome: a sporadic-dominant form due to dominant negative mutations of STAT3 and a recessive form due to mutations in the Dedicator Of CytoKinesis 8 (DOCK8) gene; mutations of SP110 are responsible for the hepatic veno-occlusive disease with immunodeficiency (VODI-ID); many candidate genes for adaptive and innate immunodeficiencies have been investigated; recessive infantile osteopetrosis, a bone disease with neural involvement in the most severe form, results from mutations of the TCIRG1 (Atp6a3), CLCN7 or OSTM1 (grey lethal) genes. The genome evolution (Alu sequences, mtDNA, Y chromosome) is analyzed in Lebanon and in Tunisia, along the paths of human expansion out of Africa. We study also markers of positive selection or, conversely, of susceptibility towards infectious diseases. In these cases also, as well as in complex genetic diseases, consanguineous families are powerful and time-saving sources of information.