# The multiple-genomes browser of the IFB cloud

Marie-Laure Franchinard[1], Frédéric Sapet[2], Sandra Derozier[1], Franck Samson[1], Valentin Loux[1], Jean-François Gibrat[1]

1 INRA, UR1404 Mathématiques et Informatique Appliquées du Génome à l'Environnement, F-78352 Jouy-en-Josas, France
2 BIOGEMMA - Upstream Genomics Group, Centre de Recherche de Chappes, CS 90126, 63720 Chappes, France

Auteur à contacter: marie-laure.franchinard@jouy.inra.fr

## Motivation

IGV (Integrative Genomics Viewer) [1] is a very efficient genome browser written in Java, allowing users to visualize and explore a large variety of genomic data types, but limited to a single genome. However, to perform comparative genomic studies, it is very useful to be able to observe different types of data simultaneously on several genotypes.

As part of the BioDataCloud project, which received funding from the call for proposals "Cloud computing 3 - Big data" of Investments for the Future Initiative program, a collaboration between the INRA Migale platform and the Biogemma company was established to tackle this issue.

[1] Helga Thorvaldsdóttir, James T. Robinson, Jill P. Mesirov. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics 14, 178-192 (2013).

## Development

According to the technical specifications set by Biogemma, a new feature has been added to IGV that allows users to jump to a new genotype from different types of data (genes, regions in genomic sequence, genetic markers) selected by the user on the reference genome.

This jump results in the opening of a new IGV window on these data, if they are available for the new genotype. This window retains all IGV features and synchronizes simultaneously with the main window.

All jumps can be saved in an IGV session file allowing users to quickly restore already used genotypes and data or to share them with other. With this new feature, the user can now compare different genotypes with the reference genome and navigate between them synchronously while keeping the IGV performance.



## Deployment on the IFB cloud

The deployment of this new version of IGV in the "BioDataCloud-IGV" virtual machine (appliance) allows users to benefit from:
✓ Large material resources in terms of CPU, RAM, storage, network
✓ User-friendly environment for users as well as developers
✓ On-demand and flexible self service
The software is thus readily available to the whole life sciences community

## Browser limits

The number of jumps achievable and therefore the number of simultaneously observable genotypes depends only upon the available hardware capabilities and the availability of the corresponding resources:
➢ IGV is by default RAM guzzling and this new feature requires around 4NGo where N is the number of genomes to be compared
➢ Large storage capacities and efficient network are required to explore large genomic data