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).
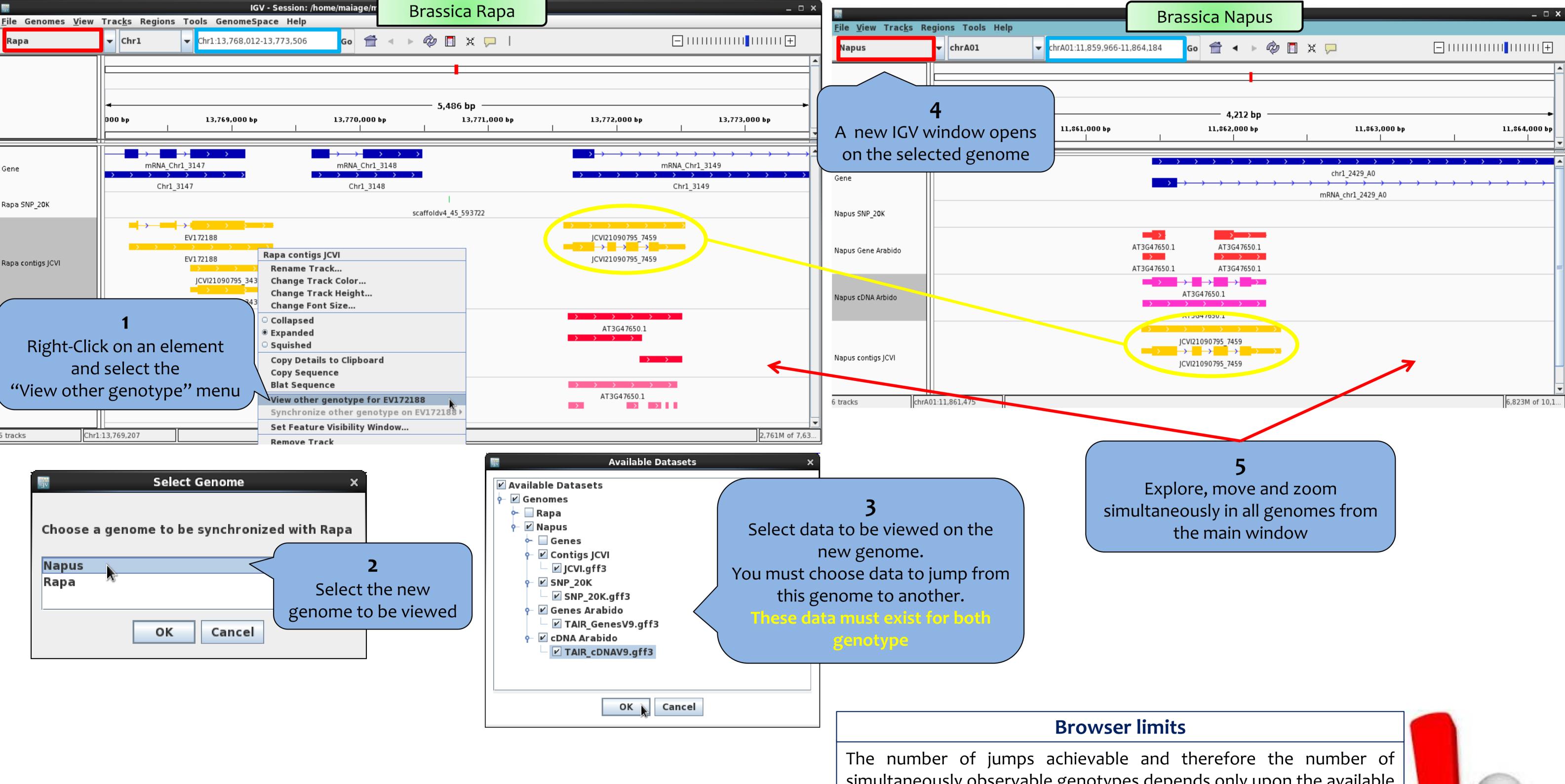
Development

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.

Brassica Rapa

1 Right-Click on an element and select the "View other genotype" menu

Select Genome
Choose a genome to be synchronized with Rapa
Napus
Rapa
OK   Cancel

2 Select the new genome to be viewed

Available Datasets
- ☑ Available Datasets
  - ☑ Genomes
    - ☐ Rapa
    - ☑ Napus
      - ☐ Genes
      - ☑ Contigs JCVI
        - ☑ JCVI.gff3
      - ☑ SNP_20K
        - ☑ SNP_20K.gff3
      - ☑ Genes Arabido
        - ☑ TAIR_GenesV9.gff3
      - ☑ cDNA Arabido
        - ☑ TAIR_cDNAV9.gff3

OK   Cancel

3 Select data to be viewed on the new genome. You must choose data to jump from this genome to another. These data must exist for both genotype

Brassica Napus

4 A new IGV window opens on the selected genome

5 Explore, move and zoom simultaneously in all genomes from the main window

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
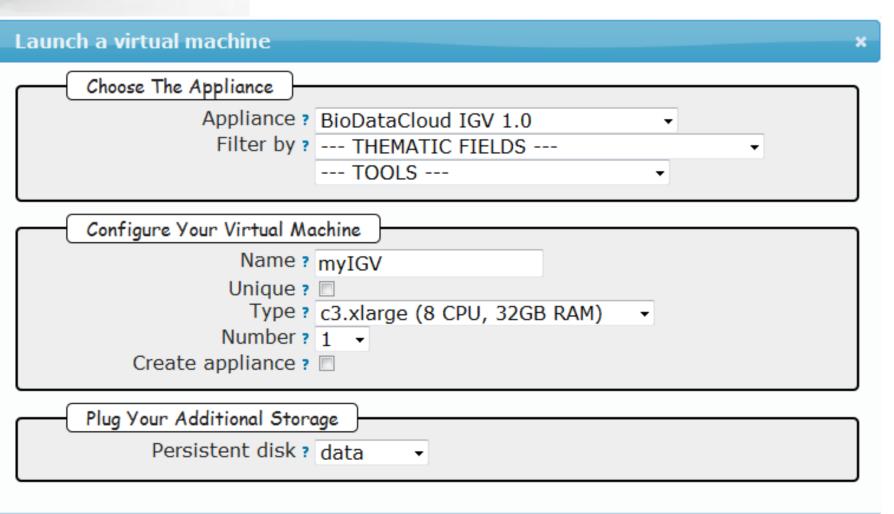
Deployment on the IFB cloud

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

Launch a virtual machine
- Choose The Appliance
  - Appliance: BioDataCloud IGV 1.0
  - Filter by: --- THEMATIC FIELDS ---
  - --- TOOLS ---
- Configure Your Virtual Machine
  - Name: myIGV
  - Unique:
  - Type: c3.xlarge (8 CPU, 32GB RAM)
  - Number: 1
  - Create appliance:
- Plug Your Additional Storage
  - Persistent disk: data

Run   Cancel

IFB BIOINFORMATICS CLOUD
YOU ARE SIGNED IN AS MLFRANCHINARD
NEWS | DASHBOARD | MONITOR | SETTINGS | HELP | SIGN OUT

DASHBOARD

Hosted at   Powered by

Shutdown   Go   Get IPs   Rename
Showing 1 to 1 of 1 entries
New Instance   New vDisk   Show Instances   Show vDisks   Show Appliances
Search:

ID	Name	Appliance	CPU%	CPU	Mem.	#Storage	Access
5558	myIGV	BioDataCloud IGV 1.0	0%	8	0	1	ssh