



Plateforme de BioInformatique - INRA Jouy en Josas

View and synchronize several genotypes using IGV

Marie-Laure Franchinard
D emo JOBIM 2015





❖ Ce projet a bénéficié des fonds de l'appel à propositions « Cloud computing n°3 – Big Data » du programme des Investissements d'Avenir.

❖ **Objectif** : Conception et mise en œuvre d'un système de type « cloud computing » adapté à une production efficace de données à haute valeur ajoutée dans le domaine de la biotechnologie végétale.

❖ **Défis applicatifs en bioinformatique** :

- Assemblage de novo
- RNAseq
- **Visualisation pour la comparaison simultanée de plusieurs génotypes**

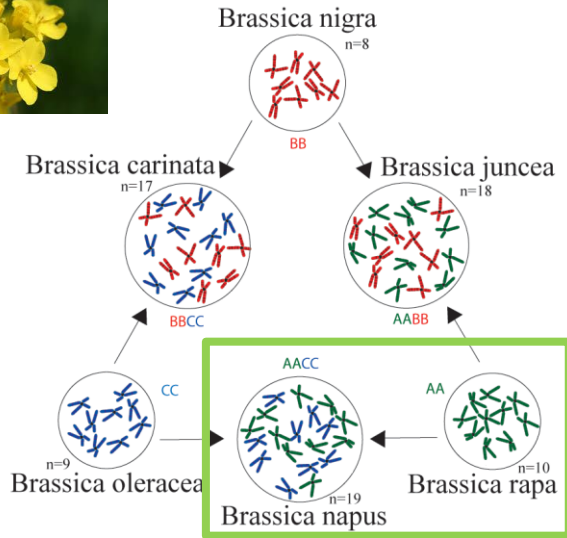


Plateforme de BioInformatique - INRA Jouy en Josas



Le besoin

✓ Plusieurs génomes/génotypes: Brassica Rapa et Brassica Napus



✓ Différents types de données pour chaque génome : gènes, régions de la séquence génomique, marqueurs génétiques:

Rapa	Napus
CN727916 Chr1:13810504-13811484	CN727916 chrC01:20792323-20793283
EE470447 Chr5:11500665 -1505687	EE470447 chrA09:28153493-28153620 chrC04:12345789-12346332
AT1G73445.1 Chr9:7830310-7830915	∅

- Visualiser des données sur plusieurs génomes simultanément
- Synchroniser la navigation dans ces génomes

Etat de l'art

Besoins	Jbrowse	IGV	FLAGdb++	GenomeView	Ensembl	GBrowse
Vues empilables	X	X			X	X
Idéogramme	X	X			X	X
Modifications des pistes	X	X	X	X	X	X
Popup sur features	X	X	X			X
Zoom sémantique	X					X
Saut entre génomes						X
Filtrage à la volée		X	X			
Edition annotation				X		X
IntersectBed		X				
User-Friendly		X	X	X		
Score sur 10	5	7	4	3	3	7



- ✓ Déjà très performant pour visualiser et explorer facilement une très grande variété de données génomiques
- ✓ User-Friendly
- ✓ JAVA

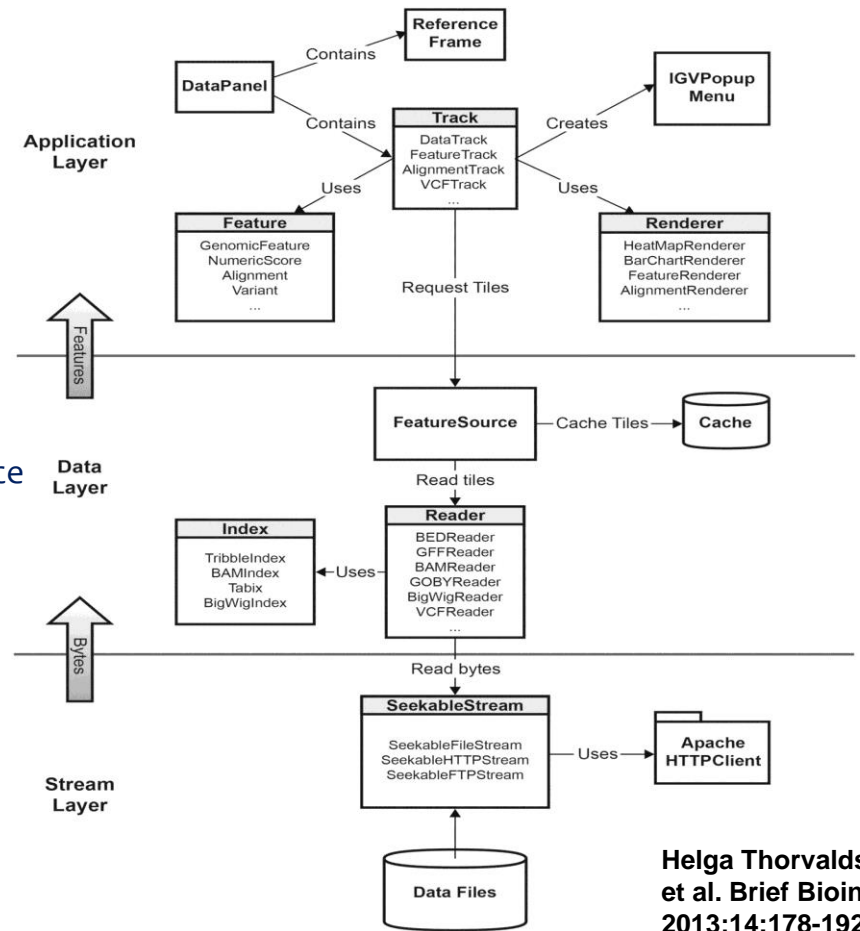


➤ Interface graphique JAVA

➤ 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).

➤ <http://www.broadinstitute.org/igv/home>

➤ Code source dense: +2000 classes mais bien écrit



Helga Thorvaldsdóttir et al. Brief Bioinform 2013;14:178-192



Projet en cours...

- Autres fonctionnalités à développer:
 - Zoom sémantique
 - Popup-menu configurable
 - Édition d'annotation

- Saut dépendant des capacités matérielles utilisées et de la disponibilité des données :
 - mémoire : proportionnelle au nombre de sauts/génomomes et au volume des données
 - $\sim 6NG$ avec N le nombre de génomes à visualiser
 - Stockage et réseau : accès aux données volumineuses locales ou distantes

- Création d'une appliance **BioDataCloud-IGV** sur le cloud de l'IFB :
 - mutualiser les ressources
 - nombre de sauts/génotypes observables plus élevé



SIGN IN



Username

Password

Login

[Lost password](#) | [Request account](#)

IFB acknowledges funding by the call "Infrastructures in Biology and Health" in the framework of the French "Investments for the Future" initiative



IFB is the French ELIXIR node



<http://cloud.france-bioinformatique.fr>

Launch a virtual machine

Choose The Appliance

Appliance ? BioDataCloud IGV 0.1

Filter by ? --- THEMATIC FIELDS ---

--- TOOLS ---

Configure Your Virtual Machine

Name ? myIGV

Unique ?

Type ? c3.large (4 CPU, 16GB RAM)

Number ? 1

Create appliance ?

Plug Your Additional Storage

Persistent disk ? Data

Run Cancel



Merci de votre attention