

View and synchronize several genotypes with IGV

Marie-Laure FRANCHINARD¹, Frédéric SAPET², Sandra DEROZIER¹, Franck SAMSON¹ and Jean-François GIBRAT¹

¹ INRA, UR1404 Mathématiques et Informatique Appliquées du Génome à l'Environnement, 78352 Jouy-en-Josas, France

² BIOGEMMA - Upstream Genomics Group, Centre de Recherche de Chappes, CS 90126, 63720 Chappes, France

Corresponding author: Marie-Laure.Franchinard@jouy.inra.fr

Résumé

IGV (Integrative Genomics Viewer) [1] est un outil graphique écrit en Java très efficace pour visualiser et explorer facilement une très grande variété de données génomiques, mais sur un seul génome à la fois. Or pour des besoins de génomique comparative, il peut être très utile d'observer différents types de données simultanément sur plusieurs génotypes.

Dans le cadre du projet BioDataCloud (Programme des Investissements d'Avenir), une collaboration entre la plateforme INRA Migale et l'entreprise Biogemma a été mise en place pour répondre à ce besoin.

Conformément aux cahiers des charges établies par Biogemma, nous avons ajouté une nouvelle fonctionnalité à IGV permettant un "saut" vers un nouveau génotype à partir de différents types de données (gènes, régions dans la séquence génomique, marqueurs génétiques) sélectionnés par l'utilisateur sur le génome de référence. Ce saut se traduit par l'ouverture d'une nouvelle fenêtre IGV sur les données sélectionnées à condition qu'elles soient disponibles sur le nouveau génotype. Cette fenêtre conserve toutes les fonctionnalités d'IGV et se synchronise simultanément avec la fenêtre principale. Le nombre de sauts réalisables et donc de génotypes observables simultanément est illimité et ne dépend que des capacités matérielles utilisées et de la disponibilité des données correspondantes. Tous les sauts peuvent être sauvegardés dans un fichier session d'IGV ce qui permet de restaurer rapidement les génotypes et les données utilisés ou de les partager avec d'autres utilisateurs.

Grâce à cette nouvelle fonctionnalité, l'utilisateur peut désormais comparer différents génotypes avec le génome de référence et naviguer entre eux de façon synchronisée tout en conservant les performances d'IGV.

Summary

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 (Investments for the Future Initiative), a collaboration between the INRA Migale platform and the Biogemma company was established to tackle this issue. 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. The number of jumps achievable and therefore the number of simultaneously observable genotypes is unlimited and depends only upon the available hardware capabilities and the availability of the corresponding data. 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.

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