W116: Managing and Exploring Large Genotyping Data with Gigwa

Monday, January 14, 2019
04:45 PM - 05:05 PM
 Golden West

With the decreasing cost of genome sequencing, many laboratories are increasingly adopting genotyping technologies as a routine component of their analytical workflows, generating large datasets (e.g. VCF files) of genotyping information. Nevertheless, manipulating such large datasets remains a challenge for many scientists. In this context, we developed Gigwa (Genotype Investigator for Genome-Wide Analyses) with the aim of providing a user-friendly system to meet the requirements of scientists who need to filter large datasets and export them into various formats for subsequent analyses.

Gigwa is species-agnostic, cross-platform, scalable and easy to deploy. It can be configured to run on a local computer or setup across servers to act as a data portal. It may be used to share data with collaborators while providing means to seek variants of interest based on location, functional annotations or genotype patterns.

Based on NoSQL technology, it supports very large datasets (up to tens of millions of genotypes) when configured on suitable hardware. Its most attractive features are: ergonomic interface including user management, numerous import and export formats, powerful filtering engine, interoperability via REST APIs and connection to online or standalone tools.

Gigwa now in version 2 (http://gigwa.southgreen.fr), is developed within the scope of South Green bioinformatics, a cross-institute platform and community dedicated to genetics and genomics of tropical and Mediterranean plants, based in Montpellier, France.

Authors

Guilhem Sempéré
CIRAD, UMR Intertryp

Julien Frouin
CIRAD - UMR AGAP

Adrien Petel
IRD

Yann Hueber
Bioversity International

Mathieu Rouard
Bioversity International

Fabien De Bellis
CIRAD

Pierre Larmande   *IRD - UMR DIADE*

Find Similar

View Related Events

Day: Monday, January 14, 2019