Genotype Investigator for Genome-Wide Analyses V2

Guilhem SEMPERE, UMR Intertryp, CIRAD

Application purpose

Providing a **graphical interface** allowing non-bioinformaticians to:

- **Host / share** large amounts of **genotyping data**, possibly enriched with metadata

- Applying **real-time filters** on this data

- **Browse** selected variants and genotypes

- Re-**export** thus defined **subsets** of data into various formats / **push into external tools**

Main characteristics

- Supports datasets of tens of **billions of genotypes**
- Single instance handling **multiples databases**
- 2 setup modes: **local** instance / data **portal**
- 3 import formats / **7 export formats**
- **2 REST APIs** implemented: GA4GH & BrAPI
- Variant filtering based on:
  – marker features, functional **annotations**
  – **genotype** patterns, data **quality**

Appealing aspects

- Support for **haploid**, diploid, **polyploid** data
- Query **progress display**
- Ability to **abort running queries**
- No loss of **phasing information** when provided
- Ability to search for variants **discriminating 2 groups of individuals**
- Instant variant **density graph** display
- Easy **connection with external tools** (Galaxy, IGV, online genome browsers, …)

Technical details

- Web-application in **Java** / HTML / Javascript
- Data hosted in **MongoDB**
- **Multi-platform** (Linux, Windows, MacOS)

More on Gigwa homepage

http://southgreen.fr/content/gigwa

(github URL, live instance list, installation bundles, etc...)
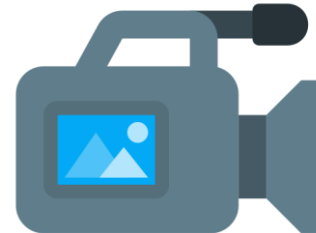
Learning how to use it

Online documentation embedded in the application:

https://gigwa.southgreen.fr/gigwa/docs/gigwa_docs.html

Tutorial videos provided on project home page

http://southgreen.fr/content/gigwa

Publications / How to cite

Gigwa - Genotype investigator for genome-wide analyses

Guilhem Sempéré, Florian Philippe, Alexis Dereeper, Manuel Ruiz, Gautier Sarah and Pierre Larmande

GigaScience 2016 5:25, https://doi.org/10.1186/s13742-016-0131-8

Gigwa v2 - Extended and improved genotype investigator*

Guilhem Sempéré; Adrien Pétel; Mathieu Rouard; Julien Frouin; Yann Hueber; F De Bellis; Pierre Larmande
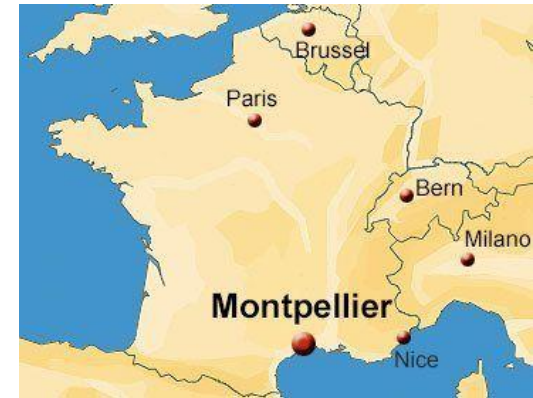
*Currently in press @ GigaScience

NB: HOW_TO_CITE.txt included in all exported archives

Development context

http://www.southgreen.fr

Local (Montpellier, France) network of bioinformaticians

- Multi-institute
- Focus: genetics / genomics of tropical & mediterranean plants
- Provides HPC + training & support to biologists
- Collaborative development: analysis pipelines & toolkits, web apps (Galaxy + in-house) & portals (genome hubs)