**Bioinformatics**

## Assembly

Grapevine datasets PacBio Hifi : Trio bining using parental sequences (See Figure 5)

* Magdeleine franc (Parent)
* Cabernet (Parent)
* Merlot

Publication

https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-023-09453-8

Raw data

<https://www.ebi.ac.uk/ena/browser/view/PRJEB59893>

Final Assembly using hifiasm (https://github.com/chhylp123/hifiasm)

<https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.57745/OJ07SN>

Workflow Snakemake for assembly (using Hifasm) and quality control (Kat, GenomeScope, Busco)

<https://forgemia.inra.fr/asm4pg/GenomAsm4pg>

Software

* Hifiasm
* Verkko
* HICANU

Tutorial

<https://genomicsaotearoa.github.io/long-read-assembly/>

Benchmark and assembly quality check

Busco, Quast, …

## Structural annotation of coding gene models

In 2014, we published the first genome assembly of Coffea canephora

See <https://www.science.org/doi/10.1126/science.1255274>

The assembly and annotation can be downloaded for the Coffea Genome Hub.

This dataset serves as a valuable reference for evaluating various annotation strategies.

Assembly

<http://www.coffee-genome.org/sites/coffee-genome.org/files/download/coffea_canephora.assembly.fna>

GFF3

<http://www.coffee-genome.org/sites/coffee-genome.org/files/download/coffea_canephora.gff3.gz>

List of software that can be used

* EuGene (<http://eugene.toulouse.inra.fr/>)
* Funannotate (<https://funannotate.readthedocs.io/en/latest/index.html>)Braker3 (<https://github.com/Gaius-Augustus/BRAKER>) : Transcriptomics + Proteins
* Galba (<https://github.com/Gaius-Augustus/GALBA>) : Only proteins. Uses the evidence from protein to genome alignment during gene prediction. Train augustus model with these evidence.
* Maker2
* EVM/Pasa
* …

## Pangenome Approach

* Minigraph-cactus

https://github.com/ComparativeGenomicsToolkit/cactus/blob/master/doc/pangenome.md

* Pggb

<https://github.com/pangenome/pggb>

* Pantools

https://pantool**s.readthedocs.io/en/stable/**