**Transplant – Elixir meeting**
**1$^{st}$ – 2$^{nd}$ July 2014**

confidential

# Biogemma, a plant Biotech company

- **Mainly funded by seed companies**
- **Strong expertise in**
  - Genetics
  - Genetic transformation
  - **Genomics**
  - **Bioinformatics**

- **Species of interest: Maize, Wheat, Rapeseed, Sunflower**
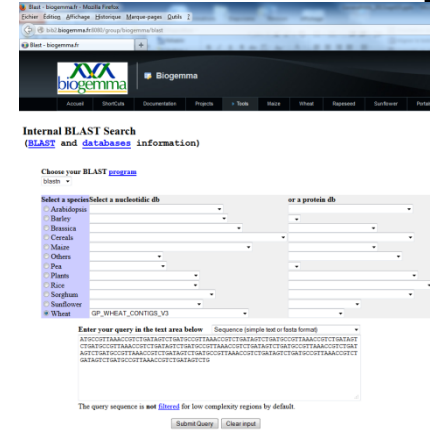
- **Ressources**
  - Sequencing facility (miseq, Hiseq data)
  - Genotyping facility (Kaspar -> Axiom)
  - Informatics
    - SGI UV1000, 264 cores, 2.5Tb RAM
      => allows to perform NGS analysis (mapping, whole genome assemblies) on large genomes
    - 100To storage
  - We ARE
    - the main facility
    - the main interface between private and academics
  - for sequencing and bioinformatics for our shareholders

# Bioinformatics as a platform

## Data management

- Databases (genome, genetic map, SNP)
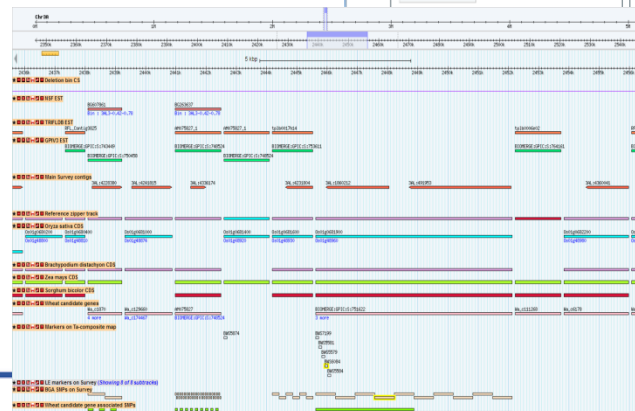- Data formatting
- Data integration

## Analysis tools

- Sequence analysis tools (blast, primer design…)
- Custom script developement for biologists
- Galaxy: workflows for biologists

## Species dedicated web portal

- Physical and genetic data
- SNP database
- Integrated view of all available data
- Private portal per shareholder

# Bioinformatics Research activities

## NGS data analysis

- De novo assembly / mapping
- Polymorphism discovery and genotyping
- Structural variation analysis
- Epigenetics and methylation profiling
- FST sequencing (GM events)

## Genome integration

- Get a genome for all our species
  - Integrate genomic sequences and genetic maps
- Pangenomics

# Involvement in Transplant

- Structural variation detection
- Pangenome viewer
- Standards in phenotyping

*Pied de page*

# What we could do

- **Tools specification**
- **User oriented**
- **Validation on private datasets**
- **Access to significant amount of data to stress tools**

- **What we do not share**
  - Our data: confidentiality / seed companies
  - Our calculation ressources

- **What we are interested in**
  - Multi-reference genomes
  - GWAS on large datasets
  - Standards
  - User oriented databases

confidential

# Positioning / community

## French Plant Bioinformatics community

- Historical partnership with URGI
- Collaborative project with IFB
- Lead/participation in Bioinformatics WP of species large research projects (IAP) in Maize, wheat, sunflower, rapeseed and pea

## International projects

- Transplant
  - Tools evaluation
  - Specification
  - standards
- Genomics of Sunflower consortium funding
- IWGSC
  - Executive Comittee
  - Participation to the development and application of a whole genome SNP detection pipeline in collaboration with GDEC (INRA Clermont-fd): >4M SNP

## Promote projects on our species of interest
## Be connected to the academic bioinformatics community for tool developement