

# Mises en perspective

Analyse bioinformatique de séquences  
pour l'amélioration des plantes

# Lundi 10 février 2014

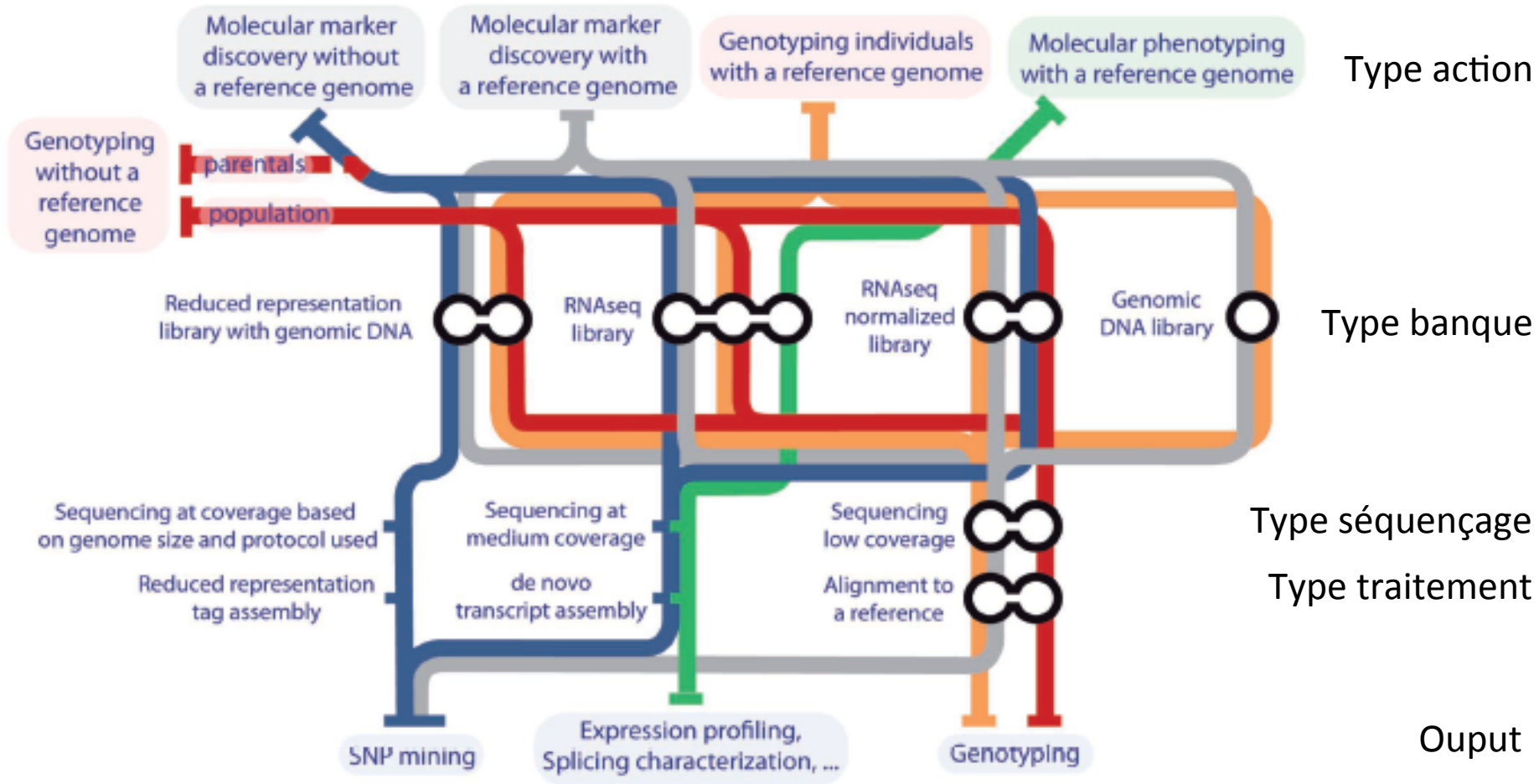
- Concepts

- Intégration de données NGS dans l'amélioration des plantes
- Gestion de données
- Utilisation d'une application générique de gestion d'analyses
- Analyses complexes avec paramétrage (workflow)
- Publication des travaux

- Compétences

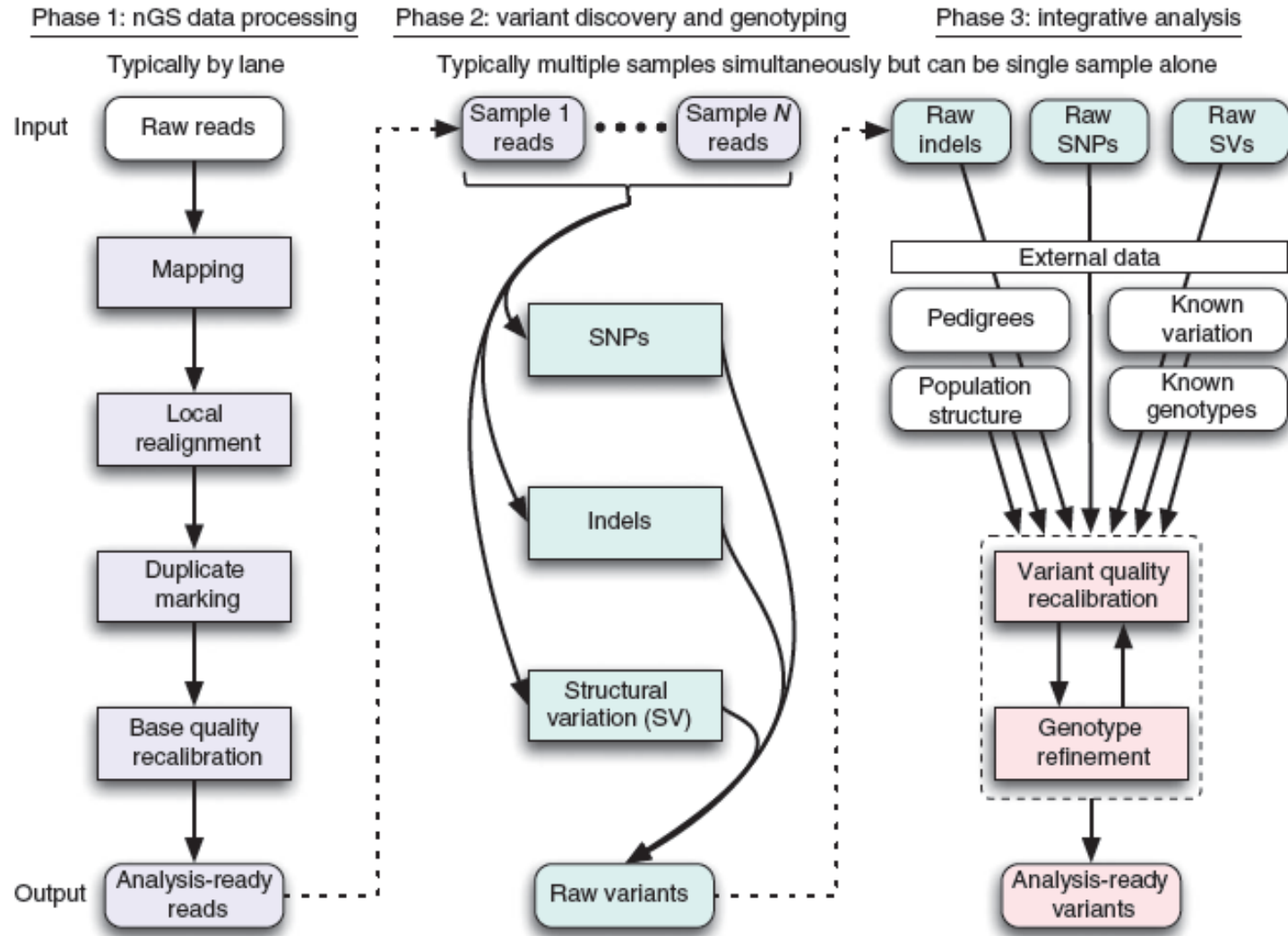
- Savoir manipuler l'interface Galaxy
- Maîtriser la qualité des données NGS
- Connaitre les bonnes pratique de l'utilisation de Galaxy
- Etre capable de suivre la suite de la formation avec un poids en moins 😊

# Mise en perspective: workflows et traitement des données NGS



Pipeline pour l'identification et le génotypage des SNP, ainsi que le phénotypage moléculaire, dans un objectif de recherche de QTL chez les plantes cultivées (Jimenez-Gomez, 2011, Frontiers in Plant Science)

# Mise en perspective: workflows et traitement des données NGS



# Mise en perspective: workflows et traitement des données NGS



Data intensive biology *for everyone.*

Galaxy is an open, web-based platform for data intensive biomedical research. Whether on the free public server or your own instance, you can perform, reproduce, and share complete analyses.

## Use Galaxy



Use project's free server or other public servers

## Get Galaxy



Install locally or in the cloud or get Galaxy on SlipStream

## Learn Galaxy



Screencasts, Galaxy 101, ...

## Get Involved



Mailing lists, Tool Shed, wiki

[Search all resources](#)

The Galaxy Team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

<http://galaxyproject.org/>

# Mise en perspective: workflows et traitement des données NGS

The screenshot shows the Galaxy web interface at <https://usegalaxy.org>. The search bar contains '3000 rice genomes'. The main content area features a central tutorial titled 'Running Your Own Galaxy: Understanding how Galaxy works' with the subtitle 'An in-depth tutorial'. To the right, there is a 'Tweets' section with three tweets from the Galaxy Project (@galaxyproject). The left sidebar contains a 'Tools' section with a search bar and a list of tool categories such as 'Get Data', 'Text Manipulation', 'FASTA manipulation', etc. The right sidebar shows a 'History' section with 'Unnamed history' and 0 bytes.

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#).

## Running Your Own Galaxy

Understanding how Galaxy works

An in-depth tutorial

Tweets

- Galaxy Project @galaxyproject 12h  
MT @UofCr4kids: In Calgary? next generation sequencing, begets need for more bioinformatics. Learn Galaxy THIS WEEK [bit.ly/1bXtjqs](http://bit.ly/1bXtjqs)
- Galaxy Project @galaxyproject 3 Feb  
GlobusWorld abstract deadline is Feb 15 Includes new Biosciences/Genomics Program [bit.ly/globusworld2014](http://bit.ly/globusworld2014) (and [bit.ly/gxyGlobusGenom...](http://bit.ly/gxyGlobusGenom...)) Expand
- Galaxy Project @galaxyproject 31 Jan  
The February 2014 Galaxy Update

Tools

search tools

Get Data

Lift-Over

Text Manipulation

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and Group

Extract Features

Fetch Sequences

Fetch Alignments

Get Genomic Scores

Operate on Genomic Intervals

Statistics

Graph/Display Data

Regional Variation

Multiple regression

Multivariate Analysis

Evolution

Motif Tools

Multiple Alignments

Metagenomic analyses

Genome Diversity

NGS TOOLBOX BETA

Phenotype Association

NGS: QC and manipulation

NGS: Mapping

NGS: SAM Tools

NGS: GATK Tools (beta)

NGS: Peak Calling

NGS: RNA Analysis

NGS: Picard (beta)

NGS: Variant Detection

snpEff

BEDTools

EMBOSS

PENNSTATE

JOHNS HOPKINS UNIVERSITY

TACC

iPlant Collaborative

The Galaxy Team is a part of the [Center for Comparative Genomics and Bioinformatics](#) at Penn State, and the [Department of Biology](#) and at [Johns Hopkins University](#).

This instance of Galaxy is utilizing infrastructure generously provided by the [iPlant Collaborative](#) at the [Texas Advanced Computing Center](#), with support from the [National Science Foundation](#).

The Galaxy Project is supported in part by NSF, NHGRI, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Johns Hopkins University.

# Bibliographie

- J. Jimenez-Gomez. **Next generation quantitative genetics in plants.** Frontiers in plant science (2011) 2, 77 doi: 10.3389/fpis.2011.00077
- A.M. Pérez-de-Castro et al. **Application of genomic tools in Plant breeding.** Current genomics (2012) 13, 179-195
- R.J. Henry et al. **Application of large-scale sequencing to marker discovery in plants.** J. Biosci. (2012) 37(5) 829-841
- M.A. DePristo et al. **A framework for variation discovery and genotyping using next-generation DNA sequencing data.** Nature Genetics (2011) 43. 491-498
- J. Goecks et al. **Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences.** Genome Biology (2010) 11,:R86