

**Durée** : Journées de cours / TP / TD puis journées de travail en autonomie pour la réalisation d'exercices.

## **Programme :**

- Présentation des plateformes et des équipes
- Initiation à l'interface Galaxy
- Administration et wrappers Galaxy
- Pipelines de traitements (bio) et développement de wrappers (info)
- Prise en main de Nextflow nf-core
- LaTeX

**Autres sessions Genotoul disponibles** : <http://bioinfo.genotoul.fr>

## **Votre intervenante :**

Sarah Maman  
Équipe Sigenaë  
Ingénieur d'études en bioinformatique

# Présentation des plateformes et des équipes



GIS Genotoul

Créé en 2000

PF Genotoul est l'une des 13 plateformes bioinformatiques IBISA.

Depuis 2008, BioInfo Genotoul travaille en collaboration avec la plateforme génomique pour traiter d'énormes volumes de données générés par la nouvelle génération de séquenceurs.

Les fichiers FASTQ produits sont mis à disposition des biologistes.

Accès à de très hautes performances de calcul : le cluster Bioinfo Genotoul.  
Nécessité de créer un compte pour accéder à ce cluster de calcul.  
Utilisation de comptes spécifiques de formation lors de cette semaine.

Si besoin:

- Pour les exercices : penser à demander une prolongation de compte, si besoin.
- Pour les stages : ouvrir un compte avec votre adresse académique (pas de mail perso).

Un fois loggé, vous avez accès:

- \* Nombreuses banques de données biologiques
- \* Nombreux scripts et outils bioinformatiques.
- \* Accès à des sessions de formations
- \* Services publics :
  - Hébergement de sites web
  - Hébergement de machines virtuelles (VM).
  - Expertises et support scientifique (biologie / bioinformatique).



# L'équipe de la plateforme BioInfo Genotoul et l'équipe Sigenae



**Matthias Zytnicki (50% FTE)** CR INRAE / Scientific animation / Development and data analysis  
+33 (0)5 61 28 54 93  
[matthias.zytnicki@inrae.fr](mailto:matthias.zytnicki@inrae.fr)



**Claire Hoede (100% FTE)** IR INRAE / Technical animation / Development and data analysis  
+33 (0)5 61 28 53 05  
[claire.hoede@inrae.fr](mailto:claire.hoede@inrae.fr)



**Christine Gaspin (30% FTE)** DR INRAE / Development and data analysis  
+33 (0)5 61 28 52 82  
[christine.gaspin@inrae.fr](mailto:christine.gaspin@inrae.fr)



**Christophe Klopp (30% FTE)** IR INRAE / Development and data analysis  
+33 (0)5 61 28 50 30  
[christophe.klopp@inrae.fr](mailto:christophe.klopp@inrae.fr)



**Didier Laborie (100% FTE)** IE INRAE / System administrator  
+33 (0)5 61 28 54 27  
[didier.laborie@inrae.fr](mailto:didier.laborie@inrae.fr)



**Céline Noïrot (60% FTE)** IE INRAE / Development and data analysis  
+33 (0)5 61 28 57 24  
[celine.noïrot@inrae.fr](mailto:celine.noïrot@inrae.fr)



**Marie-Stéphane Trotard (80% FTE)** IE INRAE / System administrator  
+33 (0)5 61 28 52 76  
[marie-stephane.trotard@inrae.fr](mailto:marie-stephane.trotard@inrae.fr)



**Nathalie Vialaneix (25% FTE)** DR INRAE / Statistical expertise and data analysis  
+33 (0)5 61 28 55 73  
[nathalie.vialaneix@inrae.fr](mailto:nathalie.vialaneix@inrae.fr)



**Patrice Dehais (50% FTE)** IE INRAE / System administrator  
+33 (0)5 61 28 57 08  
[patrice.dehais@inrae.fr](mailto:patrice.dehais@inrae.fr)



**Philippe Bordron (100% FTE)** IE INRAE / Development and data analysis  
+33 (0)5 61 28 50 72  
[philippe.bordron@inrae.fr](mailto:philippe.bordron@inrae.fr)



**Philippe Ruiz (100% FTE)** IE INRAE / Development and data analysis  
[philippe.ruiz@inrae.fr](mailto:philippe.ruiz@inrae.fr)



**Julien Touchais (100% FTE)** CDD ANR MetRibo / Web Development  
[julien.touchais@inrae.fr](mailto:julien.touchais@inrae.fr)



**Alexis Mergez (100% FTE)** CDD PEPR BreIF / Development and data analysis  
[alexis.mergez@inrae.fr](mailto:alexis.mergez@inrae.fr)



**Florent Blaise (40% FTE)** IE INRAE / System administrator  
+33 (0)5 61 28 54 21  
[florent.blaise@inrae.fr](mailto:florent.blaise@inrae.fr)



**Adela Pouban-Couzardot (100% FTE)** CDD PEPR BreIF / Development and data analysis  
[adela.pouban-couzardot@inrae.fr](mailto:adela.pouban-couzardot@inrae.fr)



**Martin Racoupeau (100% FTE)** CDD ANR Symway / Development and data analysis  
[martin.racoupeau@inrae.fr](mailto:martin.racoupeau@inrae.fr)



**Christophe Klopp**  
PROJECT COORDINATOR

*Genomic, Transcriptomic, Assembly  
Data processing, AWK, Perl, Python, R.*



**Philippe Bardou**  
BIOINFORMATICS ENGINEER

*DNaseq, Variant calling, sRNAseq (miRNA)  
Perl, Python, GIT, HTML5, CSS3, Javascript.*



**Maria Bernard**  
BIOINFORMATICS ENGINEER

*Variant calling, RADseq, Metagenomic, RNAseq, ASE  
Python, Snakemake, GIT, Rmarkdown, Galaxy.*



**Cedric Cabau**  
BIOINFORMATICS ENGINEER

*RNAseq, Assembly, Annotation  
Perl, Awk, Python, GIT, HTML/CSS/JS.*



**Mathieu Charles**  
BIOINFORMATICS ENGINEER

*Variant calling, CNV, RNAseq  
Perl, Python, Java, C++, Snakemake, Jupyter, R.*



**Patrice Dehais**  
SYSADMIN AND B.E.

*Sysadmin HPC Virtualization  
Parallelization/Optimization  
applied to bioinformatics  
software.*



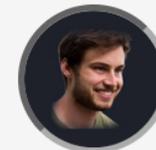
**Cervin Guyomar**  
BIOINFORMATICS ENGINEER

*Genomics, metagenomics, functional genomics  
R, Python, Nextflow.*



**Sarah Maman**  
BIOINFORMATICS ENGINEER

*RNAseq, Galaxy, elearning  
Perl, Python, Cheetah, Nextflow.*



**Quentin Boone**  
BIOINFORMATICS ENGINEER

*Web developer  
GIT, HTML5, CSS3, Javascript*

## DEVELOPMENT PROJECT

The Sigenae team develops or is involved in the development of many tools, databases and workflows.  
See below some examples of significant achievements.

- ALL
- TOOLS
- DATABASES
- WORKFLOWS




**Galaxy**

Tools

search tools

MANAGE YOUR DATA FILES

- Get Data
- Download Data
- Jobs statistics

FILES MANIPULATION

- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- GFF
- BED Tools
- Convert Formats
- Fetch Sequences
- Statistics
- Graph/Display Data

SEQUENCES MANIPULATION

- FASTA manipulation
- FastQC: fastq/sam/bam
- Illumina fastq
- Generic FASTQ manipulation
- FASTX-Toolkit for FASTQ data

SAM/BAM MANIPULATION : PICARD (BETA)

- Conversion
- QC/Metrics for sam/bam
- BAM/SAM Cleaning
- SAM/BAM manipulation: SAMtools
- Sequences Queries
- VCF Tools

SGS MAPPING

- BLAST
- BWA - Bowtie
- Indel Analysis
- Variant calling
- SNP annotation
- RNAseq Alignment
- RNAseq Raw Expression
- RNAseq Cufflinks



## How to cite Galaxy workbench ?

Depending on the help provided you can cite us in acknowledgements, references or both.

Examples :

Research teams can thank the Toulouse Midi-Pyrenees bioinformatics platform and Sigeneae group, using in their publications the following sentence : "We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrenees and Sigeneae group for providing help and/or computing and/or storage resources thanks to Galaxy instance <http://sigeneae-workbench.toulouse.inra.fr/>".

In cases of collaboration, you can directly quote the person who participated to the project : Name, Sigeneae group, GenPhySE, INRA Auzeville CS 52627 31326 Castanet Tolosan cedex.

References  
X. SIGENAE [<http://www.sigeneae.org/>]

History

search datasets

**TESTIS-PIG**  
7 shown, 110 deleted

78.98 GB

109: /work/robic/Reads_ext_ernes_2020/TESTIS/Pig/ERR3418012_2.fastq.gz			
6: ERR3418012_1.fastq.gz			
5: ERR3417952_2.fastq.gz			
4: ERR3417952_1.fastq.gz			
3: ERR3417904_2.fastq.gz			
2: ERR3417904_1.fastq.gz			
1: Upload			



### Ready to use Workflows

- 1/ FROGS - Find, Rapidly, Otus with Galaxy Solution
  - 2/ Tax4Fun workflow
- 3/ Function Table for Tax4Fun matrix workflow



### Galaxy News

- 1/ How to use datasets collections ?
- 2/ How to upload several files in Galaxy ?
- 3/ New tools : Tax4Fun (manual), VIP (manual), Salmon, multiQC.
- 4/ Tools availables for virologie analyses
  - 5/ Virology tools
- 6/ How to use Function Table tool ?
- 7/ How to export and import your histories ?



### E-learning

<https://inra.classila.com/Login>

2 e-learning availables : Galaxy initiation and Quality sequences analysis

Training resources : [BioInfo Genotoul / Sigeneae](#)

# Merci pour votre écoute

-

## Remerciements



Fonds Européen  
de Développement Régional

