



**HAL**  
open science

## Plateforme bioinformatique MIGALE

Mouhamadou Ba, Damien Berry, David Christiany, H el ene Chiapello, Olivier Inizan, Valentin Loux, Mahendra Mariadassou, V eronique Martin, C edric Midoux, Olivier Ru e, et al.

► **To cite this version:**

Mouhamadou Ba, Damien Berry, David Christiany, H el ene Chiapello, Olivier Inizan, et al.. Plateforme bioinformatique MIGALE. Journ ees Ouvertes de Biologie, Informatique et Math ematique (JOBIM), Jun 2020, Montpellier, France. 2020. hal-04176782

**HAL Id: hal-04176782**

**<https://hal.inrae.fr/hal-04176782v1>**

Submitted on 3 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

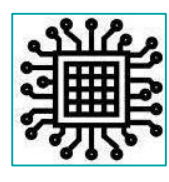
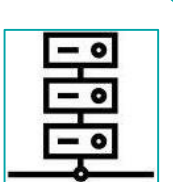


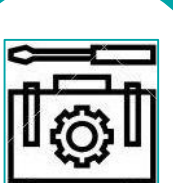

Mouhamadou Ba  
 Damien Berry  
 David Christiany  
 Hélène Chiapello  
 Olivier Inizan  
 Valentin Loux  
 Mahendra Mariadassou  
 Véronique Martin  
 Cédric Midoux  
 Olivier Rué  
 Sophie Schbath  
 Valérie Vidal

## Plateforme bioinformatique MIGALE

Migale est une infrastructure scientifique collective INRAE ouverte à tous. Elle est membre de l'Institut Français de Bioinformatique et France-Génomique. Ses activités sont certifiées ISO9001 depuis avril 2011.



### Nos missions

-  Déploiement d'une **infrastructure de calcul** scientifique
-  Mise à disposition des **collections de données publiques** issues de travaux de génomique
-  **Analyse de données génomiques**
-  **Déploiement** sur la plateforme de **logiciels d'analyses bioinformatiques** et statistiques
-  **Conception et développement** d'applications
-  **Formation** des utilisateurs aux méthodologies de la bioinformatique et aux logiciels d'analyse

→ <https://migale.inrae.fr/>

### Infrastructure de calcul scientifique

- Un **serveur frontal** (migale) de 40 cœurs et de 196 Go de mémoire
- Un **cluster de calcul** de **1016 cœurs**, 3 serveurs dédiés aux **bases de données**
- 22 serveurs physiques, 15 serveurs virtuels dédiés à des projets collaboratifs ou des outils
- Plus de **300 To** d'espace de stockage

### Données et Outils

- ~ 300 **outils** mis à disposition
- 90 **banques de données**
- **Mises à jour régulières** des banques et outils
- Portail Galaxy : <https://galaxy.migale.inrae.fr/> pour une utilisation conviviale
- **Packaging et intégration d'outils** en protéomique, statistiques, ...



### Analyse de données

#### Domaines d'expertise:

- Assemblage & annotation de génomes microbiens
- Génomique et génomique comparée
- Métagénomique, Métatranscriptomique
- Statistiques
- Fouille de texte et données

#### Exemples de projets en metagénomique :

- **MetaPDO-Cheese** : Étude de la biodiversité des fromages AOP
- Ingénierie écologique de **communautés microbiennes** pour la valorisation de bio-déchets
- Test de différents **gènes marqueurs** pour caractériser les communautés dans les écosystèmes alimentaires

### Conception et développement

- **Florilège** : base de données de phénotypes microbiens
- **Easy16S** : interface de visualisation de jeux de données métagénomiques 16S
- **ProteoRE** : interprétation et exploration de données protéomiques

### Formations

- ~120 **personnes formées**/an
- 14 **modules** différents, dont :
  - Python, Python avancé
  - Initiation à R, ggplot2
  - Galaxy
  - Traitement primaire de données NGS (Galaxy / ligne de commande)
  - Annotation automatique de génomes bactériens
  - RNA-seq : bioinfo et statistiques
  - Métagénomique 16S, Métagénomique shotgun
  - Modélisation 3D de protéines

Centre  
 Île-de-France – Jouy-en-Josas-Antony

Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France  
 Université Paris-Saclay, INRAE, BioinfOmics, Migale, 78350, Jouy-en-Josas, France