



Région  
**PAYS DE LA LOIRE**

# GALAXY & LE GRAND OUEST

---

- Un groupe de travail : GUGGO
- Plusieurs instances :
  - PCIM Brest (existante)
  - ABiMS Roscoff (existante)
  - GenOuest Rennes (existante)
  - BIRD Nantes (en test)
  - IRHS Angers (en projet)

Yvan Le Bras, Cyril Monjeaud

CNRS UMR 6074 IRISA-INRIA Rennes

@yvan2935

Yvan.le\_bras@irisa.fr

Cyril.Monjeaud@irisa.fr





Prétentieux!

---



Claudia Heriveau

Ingénieure

Cherche emploi  
Bio-informatique  
orientée  
Bio-analyse



# LE GROUPE GUGGO

---

Fédérer les activités autour de Galaxy en Bretagne et Pays de la Loire

# Objectifs

- Fédérer les activités autour de Galaxy
  - Communiquer sur la mise en place
    - d'instances
    - de formations
  - Obtenir les retours des utilisateurs
  - Réflexion sur les bonnes pratiques
    - Périmètres des instances
    - Administration système, développement, utilisation
  - Faciliter le partage
    - d'outils au sein de la communauté Biogenouest
    - d'expérience (administration système, développement, utilisation)
  - Accélérer transfert entre Bio-info recherche et service, dev et utilisation

# Historique

- Projet fédérateur Biogenouest Biologie Intégrative 2010-2012
  - Demandes insistantes des scientifiques
  - Demandes
    - d'instances de Galaxy
    - de formations
  - On a finit par céder
    - En ajoutant l'aspect échange

# Historique



Le Groupe de travail des Utilisateurs de Galaxy du Grand Ouest

Est né!

- Depuis
  - 3 réunions (avril 2012, juin et novembre 2013)
  - 5 sites (Plouzané, Roscoff, Rennes, Nantes, Angers)
  - 1 Toolshed
  - 1 instance de contribution d'outil
  - > 12 Formations données
  - 1 espace d'échange via le HUB eBGO



# Réunions

- Les instances

- PCIM Plouzané : [site du PCIM](#)
  - Laure Quintric
- ABiMS Roscoff : [galaxy.sb-roscoff.fr](http://galaxy.sb-roscoff.fr)
  - Christophe Caron
- GenOuest Rennes : [galaxy.genouest.org](http://galaxy.genouest.org)
  - Yvan Le Bras
- BIPAA INRA: [site de BIPAA](#)
  - Anthony Bretaudeau
- BIRD Nantes : [site de la plateforme](#)
  - Audrey Bihouée
- IRHS Angers : [site de l'équipe](#)
  - Sandra Pelletier

- Les utilisateurs

- Ifremer, CNRS, UR1, Université Nantes, INRA, INRIA, INSERM

# Le Toolshed GUGGO

- Toolshed hébergé par GenOuest : <http://toolshed.genouest.org>
- Portail de « tools » Galaxy
  - Dépôt / téléchargement public
  - Téléchargement et installation facilité pour n'importe quelle instance
  - Intégration des dépendances, des datatypes, des workflows...
- Transfert accéléré
  - Entre recherche en bio-informatique et service / développement et test
- Mise à disposition
  - Pour les intégrateurs
  - Pour « reviewers »
- Quelques statistiques
  - 9 catégories
  - 45 « repositories »
  - 60 outils
  - 1467 clones le 28 Novembre -> 1610 soit 143 téléchargements en 5 jours

# Le Toolshed GUGGO

- Visibilité wiki Galaxy project
- Wiki pour le développement
  - Utilisation du Toolshed
- Divers outils intégrés
  - Symbiose
    - NGS : discoSNP, compareads, minia, mapsembler2, Gasst
    - Text manipulation : spreadsheet2csv
    - Decompress an archive
      - Modification code Galaxy
      - Ajout datatypes
      - Ajout outils de décompression
  - Logiciels externes
    - Métagénomique : Mothur, ngsfilter, pairend consensus, Usearch
    - Génétique des populations : Bayescan, Carthagène, Instruct
    - NGS : idxstats, Samtools sort
    - En cours : Stacks

## ToolShed Contributions

### New Public ToolSheds

This is the first time we've highlighted publicly accessible ToolSheds. We will continue to advertise this as we become aware of them.



### GenOuest ToolShed

The GenOuest public ToolShed contains over 60 valid tools in 36 repositories in 9 categories.

# Le Toolshed GUGGO

**Galaxy Tool Shed** Repositories Admin Help User

62 valid tools on Dec 03, 2013

**Search**

- Search for valid tools
- Search for workflows

**Valid Galaxy Utilities**

- Tools
- Custom datatypes
- Repository dependency definitions
- Tool dependency definitions

**All Repositories**

- Browse by category

**Repositories I Can Change**

- Repositories I own
- Deprecated repositories I own
- My writable repositories
- Reset metadata on my repositories
- Latest revision: missing tool tests
- Latest revision: installation errors
- Latest revision: failing tool tests
- Latest revision: skip tool tests
- Latest revision: all tool tests pass
- Latest revision: invalid tools

**Available Actions**

- Create new repository

## Category Next Generation Sequencing

search repository name, description

Name ↓	Synopsis	Type	Metadata Revisions	Tools Verified	Owner
<a href="#">compareads</a>	metagenomics raw reads sets comparisons	Unrestricted	9:bb8eb7cce66f	no	<a href="#">vlebras</a>
<a href="#">discosnp</a>	discoSnp is an efficient tool for detecting SNPs without a reference genome	Unrestricted	5:49d94657de46	no	<a href="#">cmonieau</a>
<a href="#">discosnp_dependencies</a>	discoSnp dependencies	Unrestricted	0:641a10aee5dd	n/a	<a href="#">cmonieau</a>
<a href="#">gassst</a>	Global Alignment Short Sequence Search Tool	Unrestricted	11:0722ba08995c	no	<a href="#">vlebras</a>
<a href="#">gassst_dependencies</a>	Gassst dependencies	Unrestricted	1:d94a4ea0cfd4	n/a	<a href="#">cmonieau</a>
<a href="#">idxstats_from_samtools</a>	provides simple stats on RNA-seq BAM files	Unrestricted	26:14caacce6237	no	<a href="#">vlebras</a>
<a href="#">idxstats_workflow</a>	workflow using samtools sort, bam to bai and idxstats	Unrestricted	0:1502409ee1f6	n/a	<a href="#">vlebras</a>
<a href="#">mapsembler2</a>	Mapsembler2 is a targeted assembly software	Unrestricted	3:572d42c76b26	no	<a href="#">cmonieau</a>
<a href="#">reinject_duplicate_seq</a>	Reinjects redundancy stored in a 'names' file into a cluster file	Unrestricted	6:0b5f448dd348	no	<a href="#">mathieu-bahin</a>
<a href="#">samtools_sort</a>	sort bam files	Unrestricted	4:515de62990d1	no	<a href="#">vlebras</a>
<a href="#">sff_to_fasta</a>	Turns a list of sff and metadata files into a clean fasta and a groups file	Unrestricted	9:aaf08b099af4	no	<a href="#">mathieu-bahin</a>
<a href="#">usearch</a>	Usearch software suite	Unrestricted	3:516e654f8bb8	no	<a href="#">mathieu-bahin</a>

# Le Galaxy contribution GUGGO

- Instance pour la création d'outils Galaxy à partir de scripts <http://galaxy-contrib.genouest.org>
- Multi-language : bash, perl, python, R
- Instance expérimentale pour le moment
- Basée sur l'outil Toolfactory

# Le Galaxy contribution GUGGO

The screenshot displays the Galaxy web interface. At the top, the navigation bar includes "Galaxy by GenQuest" and menu items for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". The user's memory usage is shown as "Using 15.9 MB".

The left sidebar contains navigation links: "Tools" (with a search box), "Get Data", "Send Data", "Text Manipulation", "Filter and Sort", "Join, Subtract and Group", "TOOL FACTORY", "Create your tool" (with sub-links for "Tool Factory" and "Tool Factory for R-3.0.1"), and "Workflows" (with a link for "All workflows").

The main content area is titled "Tool Factory (version 0.10)" and contains the following configuration sections:

- Select an input file from your history:** A dropdown menu set to "Selection is Optional". A note states: "Most scripts will need an input - your script MUST be ready for whatever format you choose".
- New tool ID and title for outputs:** A text input field containing "My dynamic script". A note states: "This will become the toolshed repository name so please choose thoughtfully to avoid namespace dashes with other tool writers".
- Create a tar.gz file ready for local toolshed entry:** A dropdown menu set to "No. Just run the script please". A note states: "Ready to deploy securely!".
- Create an HTML report with links to all output files and thumbnail links to PDF images:** A dropdown menu set to "No, no HTML output file thanks". A note states: "Recommended for presenting complex outputs in an accessible manner. Turn off for simple tools so they just create one output".
- Create a new (default tabular) history output:** A dropdown menu set to "My script writes to a new history output". A note states: "This is useful if your script creates a single new tabular file you want to appear in the history after the tool executes".
- Galaxy datatype for your tool's output file:** A dropdown menu set to "Tabular". A note states: "You may need to edit the xml to extend this list".
- Select the interpreter for your code. This must be available on the path of the execution host:** A dropdown menu set to "Rscript".
- Cut and paste the script to be executed here:** A large text area for pasting the script. A note below states: "Script must deal with two command line parameters: Path to input tabular file path (or 'None' if none selected) and path to output tabular history file (or 'None').".

At the bottom of the main content area is an "Execute" button.

The right sidebar shows the "History" panel, which lists recent jobs:

- Bio-imaging & R (8.6 MB)
- 55: [Marksubjectsinimages.html](#)
- 54: [Decreaseimagequality.html](#)
- 53: [Decreaseimagequality.toolshed.gz](#)
- 52: [Decreaseimagequality.html](#)
- 1: [RGB\\_stack2.jpg](#)

# Formations

- ABIMS

- 14 – 18 janvier : Ecole Aviesan Initiation au traitement des données de génomique obtenues par NGS
- 17 – 22 novembre : Ecole Aviesan Initiation au traitement des données de génomique obtenues par NGS
- 13 septembre : Initiation Galaxy
- 17 septembre : Initiation Galaxy
- 18-19 Septembre 2013 : RNAseq de novo/avec référence & cleaning

- GenOuest

- 24 Juin 2013 : Initiation Galaxy doctorant
- 9 Octobre 2013 : Initiation Galaxy INRA
- 6 Novembre 2013 : Atelier Workflow colloque génomique environnementale à Rennes
- 11 Décembre 2013 : Initiation Galaxy
- 13 Décembre 2013 : Initiation Galaxy Brest

- PCIM

- Formations en cours

- BIPAA

- 22 Octobre 2013 : Formation Montpellier
- 28 Octobre 2013 : Formation interne pour le CATI

# Groupe GUGGO sur le HUB eBGO



eBGO

Discover e-Infrastructure Community About Support



Group Manager

- Overview
- Welcome to the Galaxy User Group Grand Ouest!
- Instance presentation : Galaxy@genouest
- Members
- Wiki
- Messages
- Resources
- Discussion
- Blog
- Wish List
- Projects
- Calendar

Discoverability:  
Visible  
Join Policy:  
Restricted  
Created:  
08 Apr, 2013

## Galaxy User Group Grand Ouest Overview

- Re-introduction
- Data analysis
- Data histories
- Data backups
- Data management
- Data sharing
- Genetics
- Genomics
- Proteomics
- Training
- Web Platform
- Workflows

### Bienvenue dans le groupe GUGGO!

Ce groupe du HUB eBGO a été créé afin de proposer un espace d'échange et de collaboration entre scientifiques intéressés par la plateforme web d'analyse de données Galaxy et les instances proposées dans le grand ouest. Ce groupe a pour but de permettre la concentration en un lieu unique des infos utiles à la communauté GUGGO.

Une fois rejoint le groupe GUGGO, vous pourrez consulter et/ou

- échanger avec les différents [membres](#)
- réviser, éditer le [GUGGO wiki](#)
- partager des [ressources](#) d'intérêt pour la communauté GUGGO
- proposer des développements via la gestion d'une [wish list](#)
- créer
  - des [blogs](#)
  - des [projets](#)

### Les différentes instances de Galaxy répertoriées dans le grand ouest

Quatre instances sont pour l'instant référencées :

- Plateforme [GenOuest](#), Rennes : [galaxy.genouest.org](#)
- Plateforme [ABIMS](#), Roscoff : [galaxy.ab-roscoff.fr](#)
- Plateforme [Bird](#), Nantes : [Site de la plateforme](#)
- Pôle de Calcul Intensif pour la Mer, Plouzané : [Site du PCIM](#)

Chacune d'entre elle semble accessible après demande d'ouverture de compte pour accéder aux ressources de calcul de la structure.

### Les différentes formations sur Galaxy répertoriées dans le grand ouest

Initiation à Galaxy

- Plateforme [ABIMS](#), Roscoff
- Plateforme [GenOuest](#), Rennes

Analyse de données RNAseq de novo via Galaxy

- Plateforme [ABIMS](#), Roscoff

Analyse de données de micro ARNs via Galaxy

- Plateforme [ABIMS](#), Roscoff



# Merci de votre attention

La plate-forme Bio-informatique GenOuest



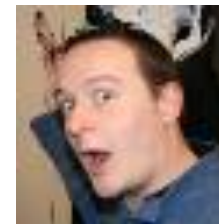
Le groupe Symbiose IRISA/INRIA  
*GenOuest-Dyliss-Genscale*



eBGO HUB (*collaboration*)

<http://e-biogenouest.org>

Cyril Monjeaud



EMME portal (*data management*) <http://emme.genouest.org>

Olivier Collin



Galaxy instances (*data analysis*) <http://galaxy.genouest.org>

<http://toolshed.genouest.org> <http://galaxy-contrib.genouest.org>

GO4Bioinformatics (*education*) <http://go4bioinformatics.genouest.org>