

A French Galaxy Tool Shed

to federate the national infrastructures and offering quality assessed tools



Loraine BRILLET-GUÉGUEN¹, Christophe CARON¹, Valentin LOUX² and the French Galaxy Working Group³

¹ ABIMS, FR2424 CNRS-UPMC, Station Biologique, Place Georges Teissier, 29680, Roscoff, France; ² UR1404 Mathématiques et Informatique Appliquées du Génome à l'Environnement, INRA, F-78352 Jouy-en-Josas, France; ³ Institut Français de Bioinformatique [ANR-11-INBS-0013], France Génomique [ANR-10-INBS-0009] and MetaboHUB [ANR-11-INBS-0010] - gtgalaxy@groupes.france-bioinformatique.fr

Abstract: The Galaxy environment, notably dedicated to bio-analyses, is finding a growing success in bioinformatics and biology communities. This project is coordinated by the IFB Galaxy Working Group. The IFB offers a reference repository to centralize and promote the bio-analyses tools of the French community. The scope of this repository, initially dedicated to "France-Génomique" NGS pipelines, is now extending to other national infrastructures and to national training actions.

Applicants

National infrastructures

France Génomique is the national sequencing infrastructure created thanks to the « Investment for the future » initiative. It gathers together 9 sequencing and 6 bioinformatics platforms. The main objective of the bioinformatic part of the France Génomique project is to coordinate the necessary evaluation of these methods to avoid duplication of efforts from the bioinformatics and sequencing platforms.

MetaboHUB is a national infrastructure of metabolomics and fluxomics that provides tools and services to academic research teams and industrial partners in the fields of health, nutrition, agriculture, environment and biotechnology.

Research team: we also use this Tool Shed to harmonize tools installation over several Galaxy servers (Cloud, Platform...) for training session.



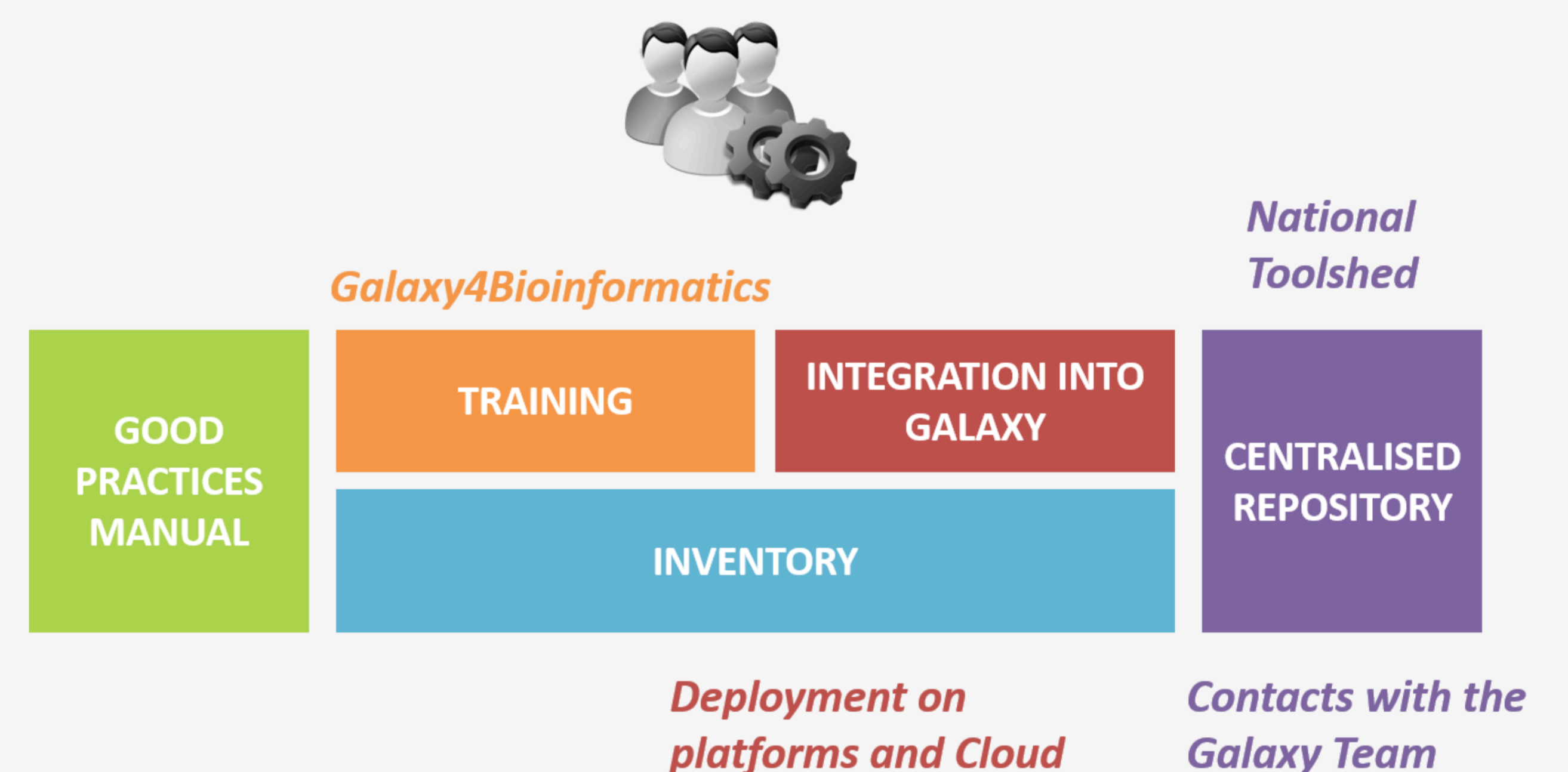
Providers

The "Institut Français de Bioinformatique" (IFB - <http://www.france-bioinformatique.fr>) is the French node of the pan-European research infrastructure ELIXIR. IFB national infrastructure provides an IT environment based on a central hub (IFB Core) and on 32 platforms distributed in 6 regional groups. This IT infrastructure is devoted to the management and analysis of biological data, in particular data generated by high-throughput technologies. IFB is also linked with others national infrastructure (MetaboHub, France Génomique...) through R&D or training activities.

IFB commissioned in 2013 a Working Group around the Galaxy platform. This group gathers several national platforms, and manages animation actions (Galaxy Day, thematic schools, etc.) and actions to structure (training, good practices guides, etc.) users and developers communities.

Strategy to federate users communities

As a part of the bioinformatics work packages funded by the "France-Génomique" project, the community has developed or evaluated many tools and set up analysis workflows. Exploitation and diffusion of these pipelines dedicated to people unfamiliar with the command line instructions now lies on using a **common platform** (Galaxy) and on creating a **shared repository** (Tool Shed).



Why

Strong focus on quality tools & validation

- Functional tests
- Test workflows

Intranet access before public releases

Distribute France Génomique analysis workflows

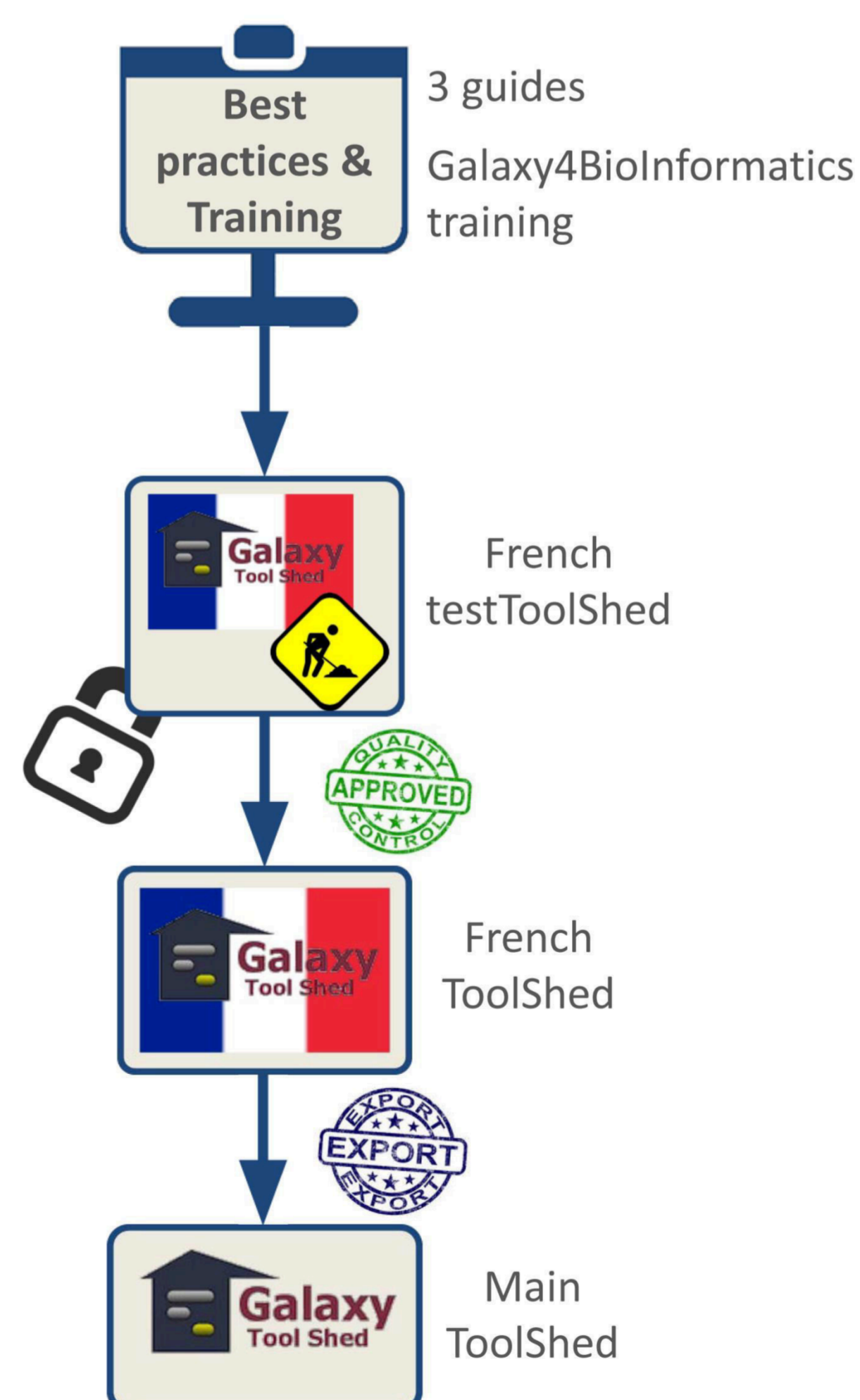
Visibility to justify our grants

IFB
France Génomique, MetaboHUB...

Automatic deployment

Virtual Machine
Cloud

Tool Shed As A Bridge



Results

Best practices guides

Quick start, Advanced and Toolshed
<http://www.france-bioinformatique.fr/galaxy-working-group>

Training :

Galaxy4Bioinformatics

Roscoff (11/2014), Nantes (03/2015), Toulouse (11/2015)
Used by national summer school : NGS AVIESAN (09/2015)

SARTools

H, Varet, J.-Y, Copee and M.-A, Dillies, *SARTools: a DESeq2 and edgeR-based R pipeline for comprehensive differential analysis of RNA-seq data*, 2015 (submitted)

W4M tools : 29 modules

...

Road Map

Planemo evaluation (11/2015)

Quality : automatic test specification, tools reviewing ...

[1] Goecks, J, Nekrutenko, A, Taylor, J and The Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* 2010 Aug 25;11(8):R86.
[2] Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. "Galaxy: a web-based genome analysis tool for experimentalists". *Current Protocols in Molecular Biology.* 2010 Jan; Chapter 19:Unit 19.10.1-21.
[3] Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A. "Galaxy: a platform for interactive large-scale genome analysis." *Genome Research.* 2005 Oct; 15(10):1451-5.
[4] Daniel Blankenberg, Gregory Von Kuster, Emil Bouvier, Dannon Baker, Enis Afgan, Nicholas Stoler, the Galaxy Team, James Taylor and Anton Nekrutenko, "Dissemination of scientific software with Galaxy ToolShed," in *Genome Biology* 2014, 15:403, doi:10.1186/gb4161