

# New tools and workflows on W4M, the Galaxy metabolomic infrastructure

Marion LANDI<sup>1</sup>, Claire LOPEZ<sup>1</sup>, Marie TREMBLAY-FRANCO<sup>2</sup>, Cécile CANLET<sup>2</sup>, Yann GUITTON<sup>5-6</sup>, Daniel JACOB<sup>3</sup>, Gildas LE CORGUILLE<sup>7</sup>, Jean-François MARTIN<sup>2</sup>, Mishar MONSOOR<sup>7</sup>, Nils PAULHE<sup>1</sup>, Mélanie PETERA<sup>1</sup>, Camille VACQUIE<sup>7</sup>, Franck GIACOMONI<sup>1</sup>, Estelle PUJOS-GUILLOT<sup>1</sup>, Christophe CARON<sup>7</sup>, Etienne THEVENOT<sup>4</sup>

1 PFEM, UMR1019, INRA, Centre Clermont-Ferrand-Theix, 63122 Saint Genès Champanelle, France

2 PF MetaToul-AXIOM, UMR 1331 Toxalim, INRA, 180 chemin de Tournefeuille, F-31027 Toulouse, France

3 IBVM, INRA, Metabolome Facility of Bordeaux Functional Genomics Center, 33140 Villenave d'Ornon, France

4 LIST/LADIS, CEA, Saclay Center, F-91191 Gif-sur-Yvette, France

5 IRISA, UMR 6074, CNRS, Campus de Beaulieu, 35042 Rennes, France

6 LINA, UMR 6241, Université de Nantes, 44322 Nantes, France

7 ABiMS, FR2424, CNRS-UPMC, Station Biologique, Place Georges Teissier, 29680 Roscoff, France

Galaxy is a free and open-source science web-based platform, which provides tools and data processing pipelines (“workflow”) mainly dedicated to genomics (<http://galaxyproject.org>). Our platform based on this Galaxy environment <http://workflow4metabolomics.org> (W4M) is dedicated to metabolomic. W4M is available thanks to strong collaboration between software developers (statisticians and bioinformaticians) and chemists for their feedbacks, all coming from six research units.

Several tools were initially developed to process data coming from liquid chromatography coupled to mass spectrometry (LCMS). These tools cover all metabolomic data treatment steps, from preprocessing to annotation, and including statistical analysis. In the last twelve months, new and innovative modules dedicated to 1D NMR spectra processing have been developed and integrated in Galaxy. These tools are focused on NMR specific process, such as construction of a binned data from spectra (bucketing), several normalisation methods, and annotation. As for LCMS workflow, statistical analysis of NMR data can be performed by using univariate or multivariate tools (ANOVA, PLS, OPLS...).

Until recently, different software were necessary for LCMS and NMR data pre-treatment, statistical analysis and metabolite assignment. Now the present work within Galaxy has opened the door to the possibility of creating workflows including all the data treatment steps in metabolomic for each analytical technique. In this context, our perspectives are to enrich Galaxy with workflows for gas chromatography coupled to mass spectrometry.

In conclusion, the Galaxy environment and the Workflow4Metabolomics project allow users to execute tools in a stand-alone mode or to create easily their own workflows through the graphical interface, set the parameters, and visualize the results. All this work can be accomplished in a user-friendly environment without programming skills.

**Mots clés :** Galaxy, metabolomic, workflow