



A Galaxy-based framework to analyze raw sequencing data from microbiota



Bérénice Batut, Clémence Defois, Kévin Gravouil, Jean-François Brugère, Eric Peyretailade and Pierre Peyret
UMR 454 MEDIS, Clermont-Ferrand, France
Bioinformatic Group, Freiburg, Germany

New generation of sequencing platforms coupled to numerous bioinformatics tools has led to rapid technological progress in metagenomics, metatranscriptomics and complex microorganism communities' studies. Nevertheless, a combination of different bioinformatics tools remains necessary for efficient extraction of useful information in microbiota studies. Modular, accessible, sharable and user-friendly tools would greatly improve such studies.

We therefore developed ASaiM, an open-source **Galaxy-based framework** dedicated to microbiota data analyses

ASaiM provides an **expertly selected collection of tools** to explore and visualize taxonomic and functional information from raw amplicon, metagenomic or metatranscriptomic sequences. To help the analyses, **several (customizable) workflows** are included. The available workflows are supported by **tutorials** and **Galaxy interactive tours** to guide the users through the analyses. Furthermore, an effort on **documentation** of ASaiM, its tools and workflows has been made (<http://asaim.readthedocs.io/>). ASaiM is implemented as **Galaxy Docker flavour** and the associated source code is available under Apache 2 license at <https://github.com/ASaiM/framework>.

Based on the Galaxy framework, ASaiM offers sophisticated analyses to scientists without command-line knowledge. ASaiM provides a powerful framework to easily and quickly explore microbiota data in a reproducible and transparent environment.

A framework built on the shoulders of giants

To develop a modular, accessible, redistributable, sharable and user-friendly framework for scientists working with microbiota data



As the base



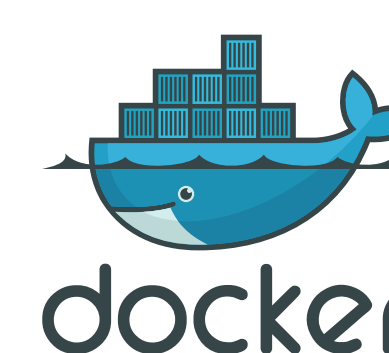
To install the Galaxy tools



To install the tools and their dependencies

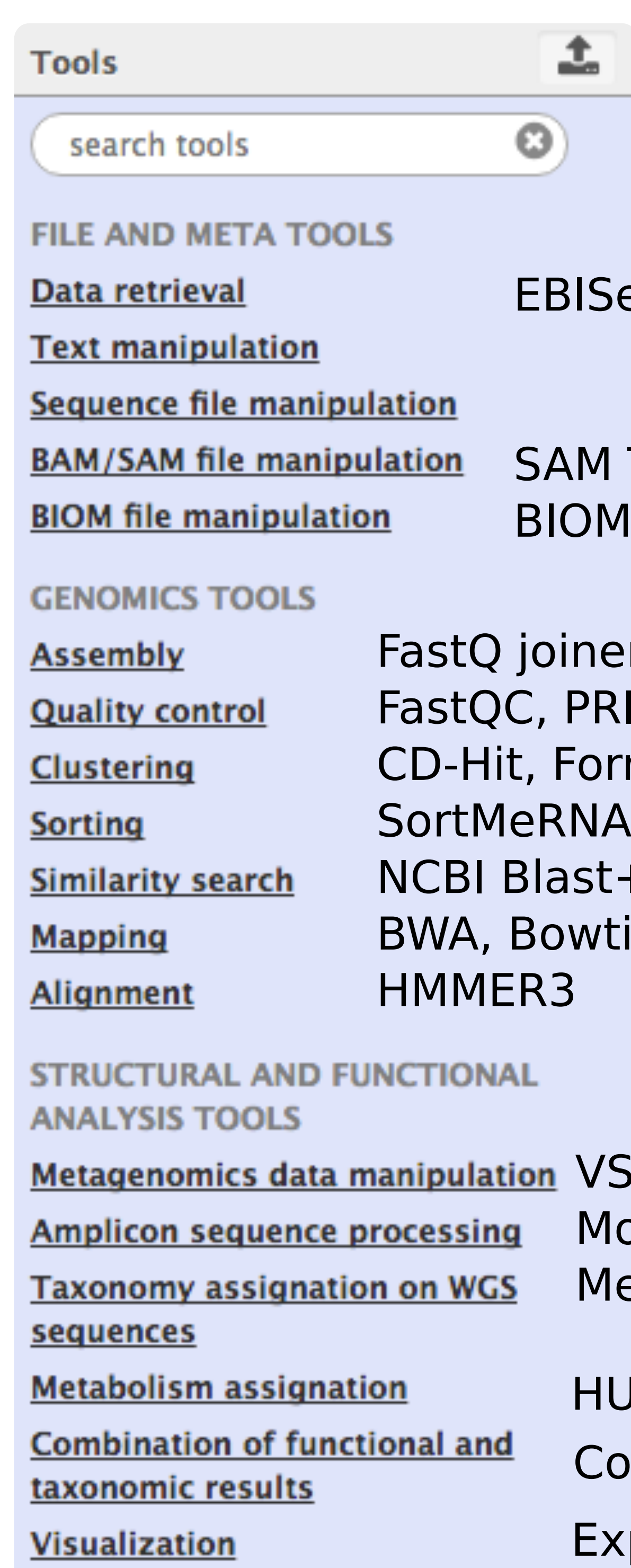


To install the workflows and databases



To containerize and ship everything

With a comprehensive set of microbiota related tools



EBISearch, ENASearch, SRA Tools

SAM Tools

BIOM-Format tools

FastQ joiner, FastQ-join

FastQC, PRINSEQ, Trim Galore!, Trimmomatic

CD-Hit, Format CD-HIT outputs

SortMeRNA

NCBI Blast+, Diamond

BWA, Bowtie

HMMER3

VSEARCH

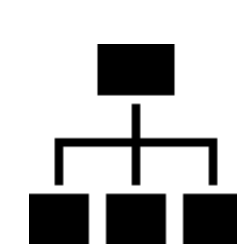
Mothur, QIIME

MetaPhlAn2, Kraken

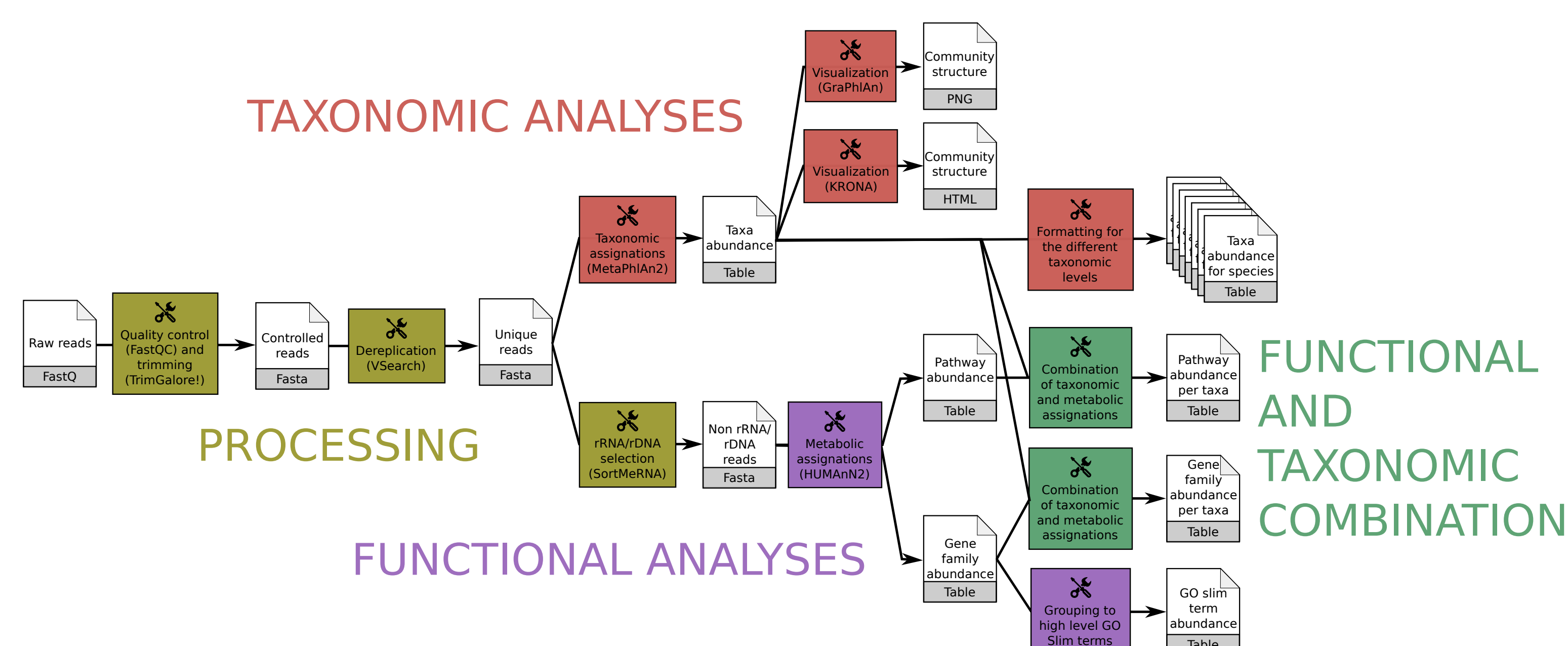
HUMAN2, PICRUST, InterProScan

Combine MetaPhlAn2 and HUMAN2

Export2graphlan, GraPhlan, Krona



With predefined and tested workflows



Main workflow to analyze raw shotgun metagenomics data



With documentation and training

- Online documentation: <http://asaim.readthedocs.io>
how to use and deploy ASaiM
which tools are integrated with some documentation about them
which workflows are integrated and how to use them
- Galaxy Interactive Tours inside ASaiM
- Tutorials
Developped with the help of the Galaxy Training Network (GTN)
Hosted in the GTN website
<http://galaxyproject.github.io/training-material/metagenomics>

Useful information

GitHub: <http://github.com/asaim/framework>
Doc: <http://asaim.readthedocs.io>

Acknowledgments



Contact

Bérénice Batut
berenice.batut@gmail.com