

# How to analyze raw sequencing data from microbiota? Using ASaiM!



**B  r  nice Batut**<sup>1</sup> (berenice.batut@gmail.com), Cl  mence Defois<sup>2</sup>, K  vin Gravouil<sup>2</sup>, Jean-Fran  ois Brug  re<sup>2</sup>, Eric Peyretailade<sup>2</sup> and Pierre Peyret<sup>2</sup>

New generation of sequencing platforms coupled to numerous bioinformatics tools has led to rapid technological progress in metagenomics, metatranscriptomics and complex microbial community 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 <http://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



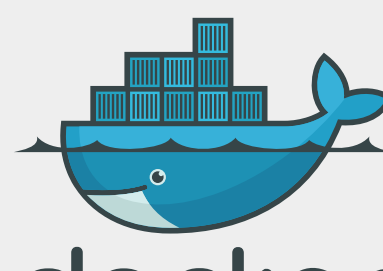
Galaxy ToolShed,  
BioBlend and Ephemeric

To install the Galaxy tools, the  
workflows and databases



Conda

To install the tools and  
their dependencies



docker

To containerize and ship  
everything



## With a comprehensive set of microbiota related tools

Tools

search tools

### FILE AND META TOOLS

[Data retrieval](#)

[Text manipulation](#)

[Sequence file manipulation](#)

[BAM/SAM file manipulation](#)

[BIOM file manipulation](#)

### GENOMICS TOOLS

[Assembly](#)

[Quality control](#)

[Clustering](#)

[Sorting](#)

[Similarity search](#)

[Mapping](#)

[Alignment](#)

### STRUCTURAL AND FUNCTIONAL ANALYSIS TOOLS

[Metagenomics data manipulation](#) VSEARCH

[Amplicon sequence processing](#) Mothur, QIIME

[Taxonomy assignment on WGS  
sequences](#) MetaPhlAn2, Kraken

[Metabolism assignment](#)

[Combination of functional and  
taxonomic results](#) HUMAnN2, PICRUST, InterProScan

[Visualization](#) Combine MetaPhlAn2 and HUMAnN2

EBISearch, ENASearch, SRA Tools

See documentation for the full list

See documentation for the full list

SAM Tools

BIOM-Format tools

FastQ joiner, FastQ-join

FastQC, PRINSEQ, Trim Galore!, Trimmomatic

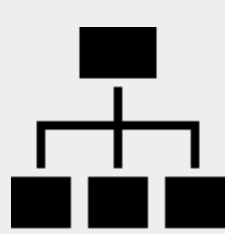
CD-Hit, Format CD-HIT outputs

SortMeRNA, FragGeneScan

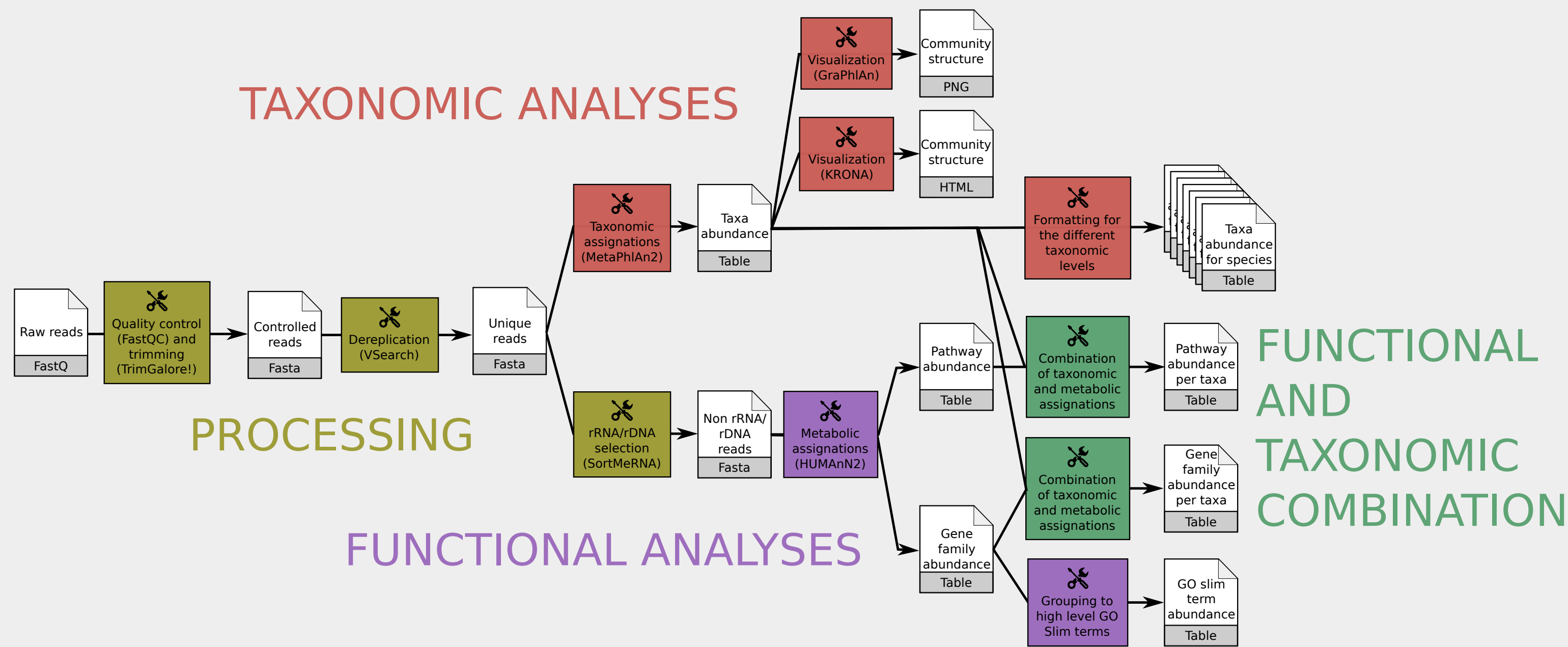
NCBI Blast+, Diamond

BWA, Bowtie

HMMER3



## 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>

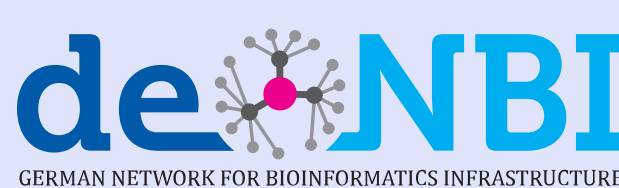
## Useful information

GitHub: <http://github.com/asaim/framework>

Doc: <http://asaim.readthedocs.io>

## Acknowledgements

CPER 2015  
Axe T2ANSH  
(Action 3, Aliments,  
Nutrition Humaine Sant  )



Check out  
our paper !



<sup>1</sup> University of Freiburg, Germany

<sup>2</sup> UMR 454 MEDIS, Clermont-Ferrand, France

Icons by The Noun Project and icon8