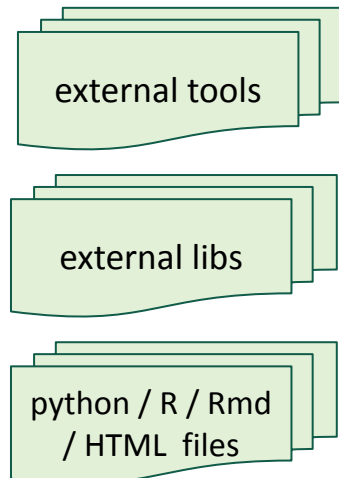


A large, stylized green frog logo is positioned on the left side of the slide. The frog is depicted in a light green color with a darker green outline. Its body is composed of a series of connected circles and lines, giving it a segmented appearance. The frog's legs are also stylized, with small circles at the joints. A DNA double helix is integrated into the frog's body, running vertically through its center. The word 'FROGS' is written in a light blue, sans-serif font across the middle of the frog's body.

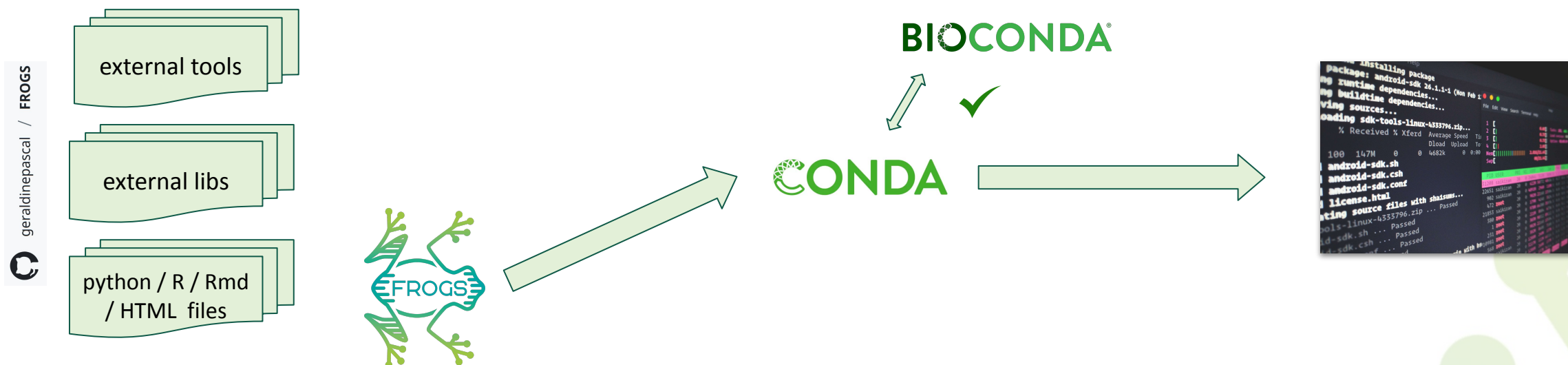
FROGS 5 - What's new

Olivier Rué, Maëlle Pomies, Maria Bernard, Lucas Auer, Gabryelle Agoutin,
Géraldine PASCAL

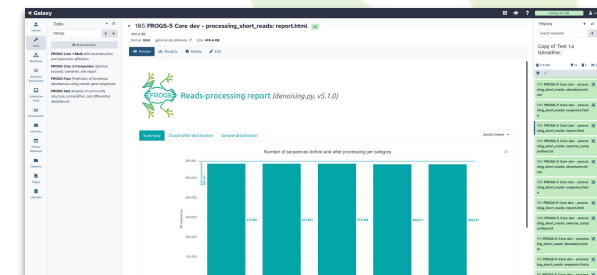
The underlying architecture of FROGS



The underlying architecture of FROGS

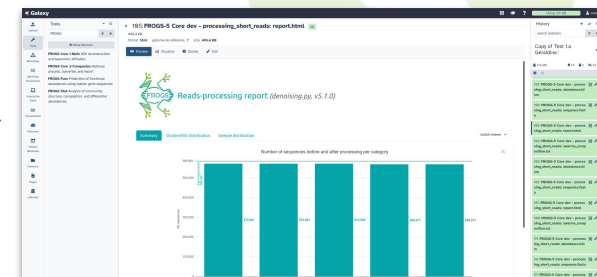
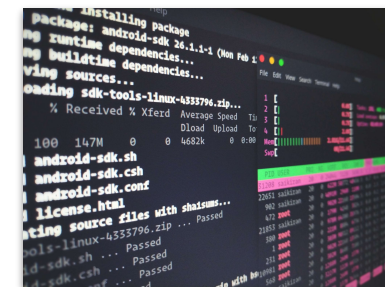


 geraldinepascal / **FROGS**

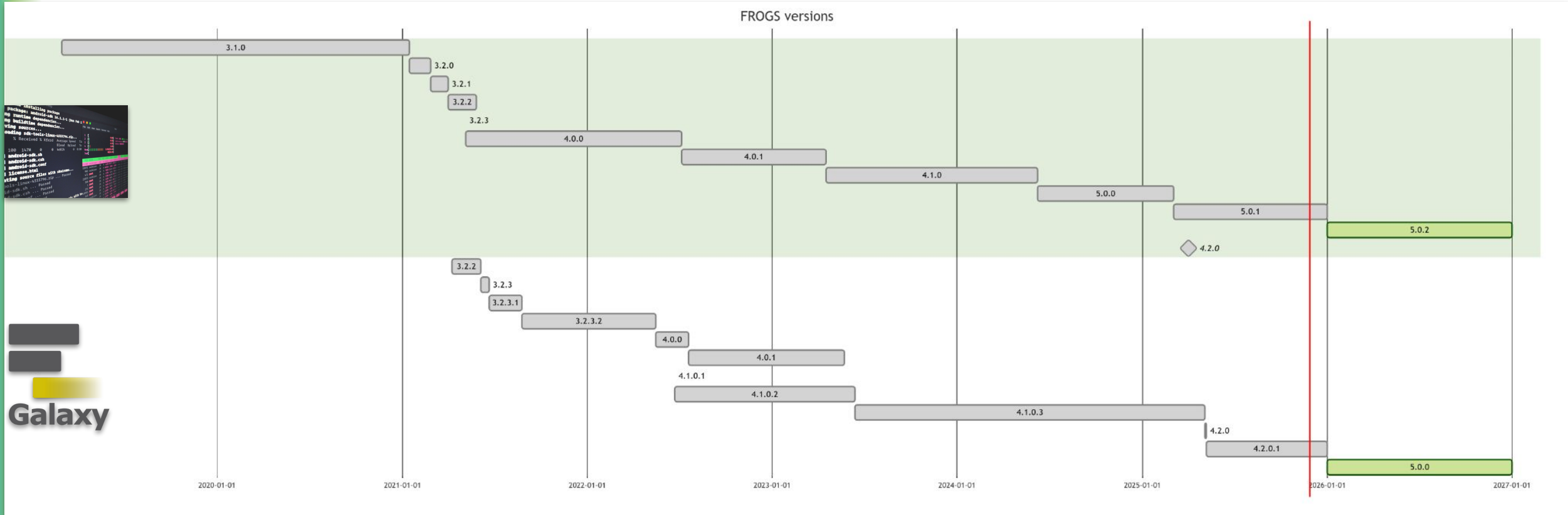


 geraldinepascal / **FROGS**

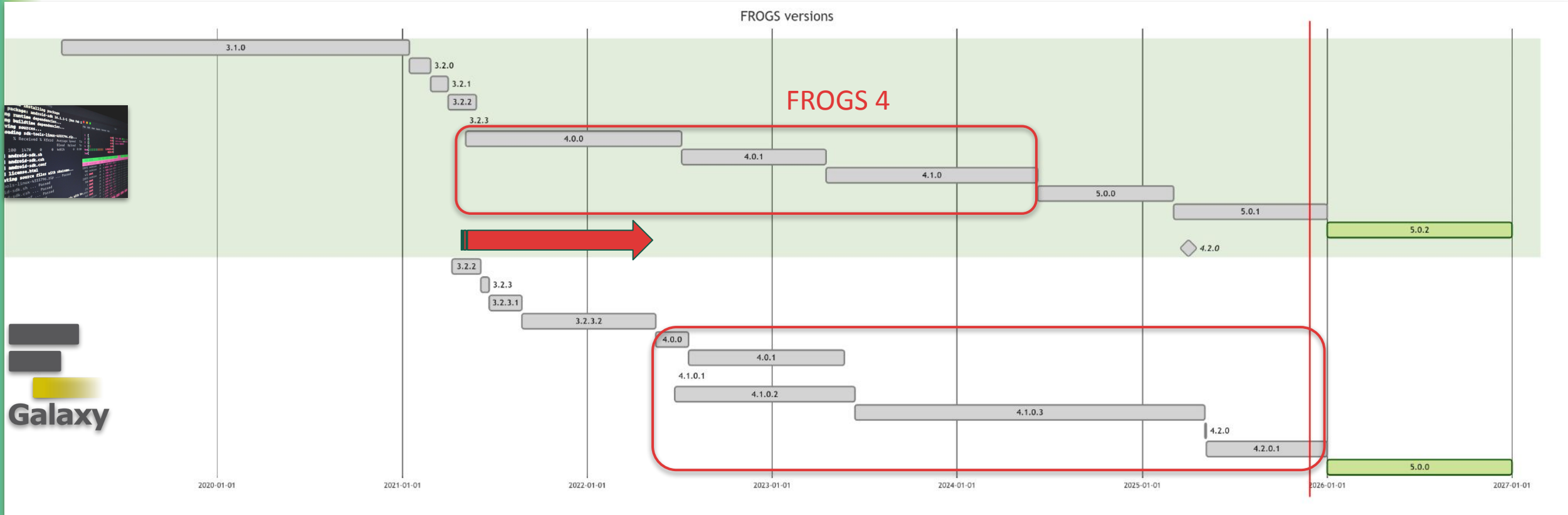
geraldinepascal / FROGS-wrappers



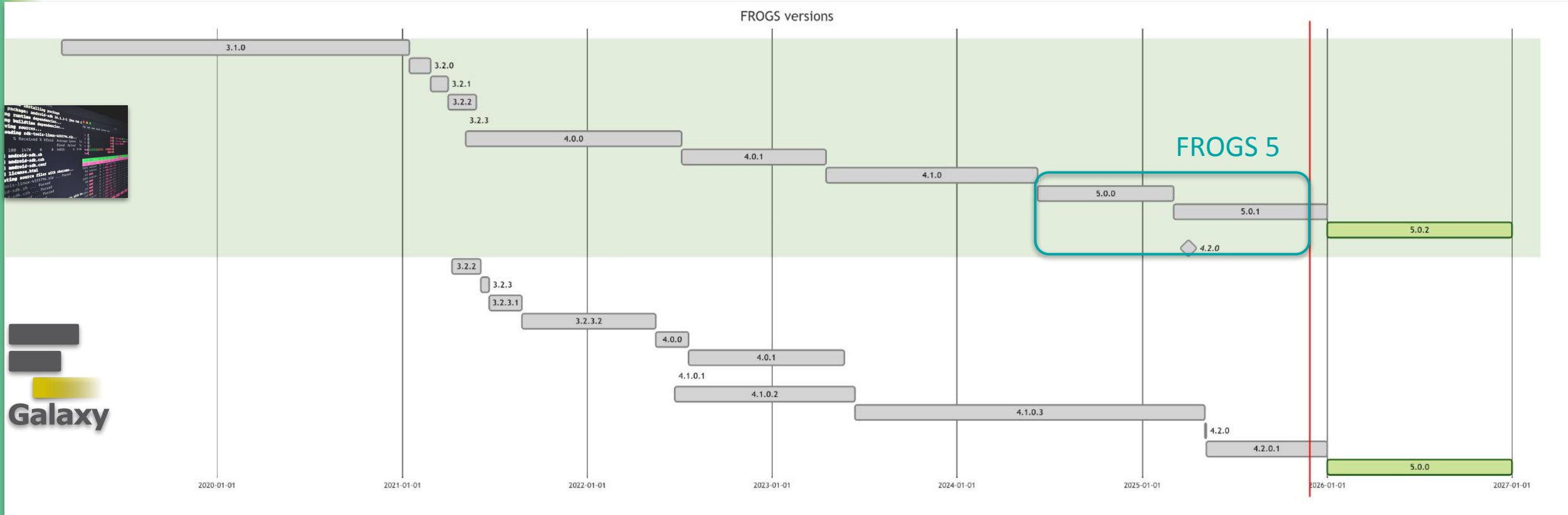
Conda vs Galaxy versions



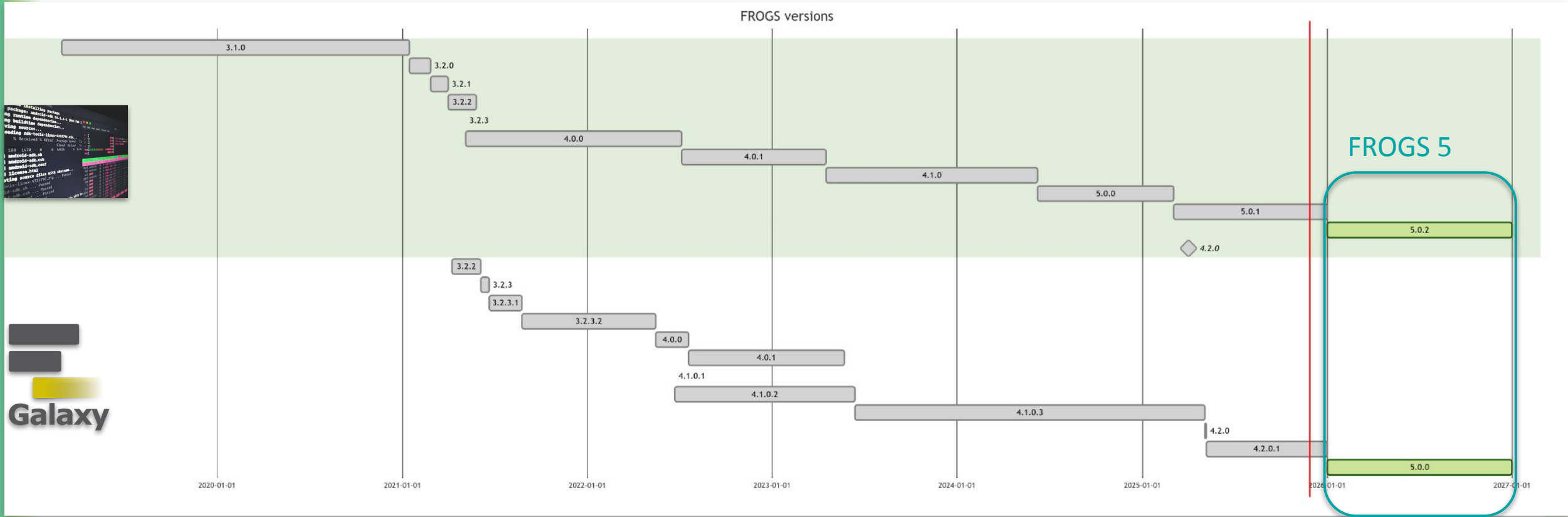
Conda vs Galaxy versions



Conda vs Galaxy versions



Conda vs Galaxy versions



Continuous changes

Within Our Control

- Evolutions
- Bugs

Beyond Our Control

- Conda and Galaxy requirements
- Galaxy administration
- Licence changes
- Deprecated functions
- External tools availability

FROGS 5 Workstreams

1. Improve the Galaxy user experience
2. New HTML reports
3. Reorganization and evolution of the old “preprocess” tool
4. Improve management of long reads



1. Improve the Galaxy user experience

Metabarcoding: FROGS

frogs

Show Sections

FROGS Pre-process merging, denoising and dereplication.

FROGS Tree Reconstruction of phylogenetic tree

FROGS BIOM to TSV Converts a BIOM file in TSV file.

FROGS Demultiplex reads Attribute reads to samples in function of inner barcode.

FROGS Remove chimera Remove PCR chimera in each sample.

FROGSSTAT Phyloseq Alpha Diversity with richness plot

FROGSSTAT Phyloseq Sample Clustering of samples using different linkage methods

FROGSSTAT Phyloseq Multivariate Analysis Of Variance perform Multivariate Analysis of Variance (MANOVA)

FROGSSTAT Phyloseq Structure Visualisation with heatmap plot and ordination plot

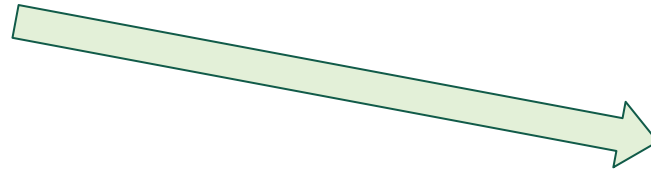
FROGS Abundance normalisation Normalise OTUs abundance.

FROGS Clusters stat Process some metrics on clusters.

FROGSSTAT Phyloseq Beta Diversity distance matrix

FROGSSTAT DESeq2 Preprocess import a Phyloseq object and prepare it for DESeq2 differential abundance analysis

FROGSSTAT Phyloseq Composition Visualisation with bar plot and



frogs

Show Sections

FROGS Core 1-Main ASV reconstruction and taxonomic affiliation

FROGS Core 2-Companion Optional process, converter, and report

FROGS Func Prediction of functional abundances using marker gene sequences

FROGS Stat Analysis of community structure, composition, and differential abundances



FROGS-core

FROGS Core 1-Main

ASV reconstruction and taxonomic affiliation
(Galaxy Version 5.0.2+galaxy0)



Run Tool

Tool Parameters

Select a tool from the FROGS Core suite to run your analysis.

☒ -- Please select a tool --

- ☐ 1.a. Reads processing of short reads
- ☐ 1.b. Reads processing of long reads
- ☐ 1.c. Reads processing of 454 reads
- ☐ 2. Remove chimera
- ☐ 3. Cluster/ASV filters
- ☐ 4. Taxonomic affiliation
- ☐ 5. Phylogenetic tree building
- ☐ ITSx

FROGS Core 2-Companion

Optional process, converter, and report
(Galaxy Version 5.0.2+galaxy0)



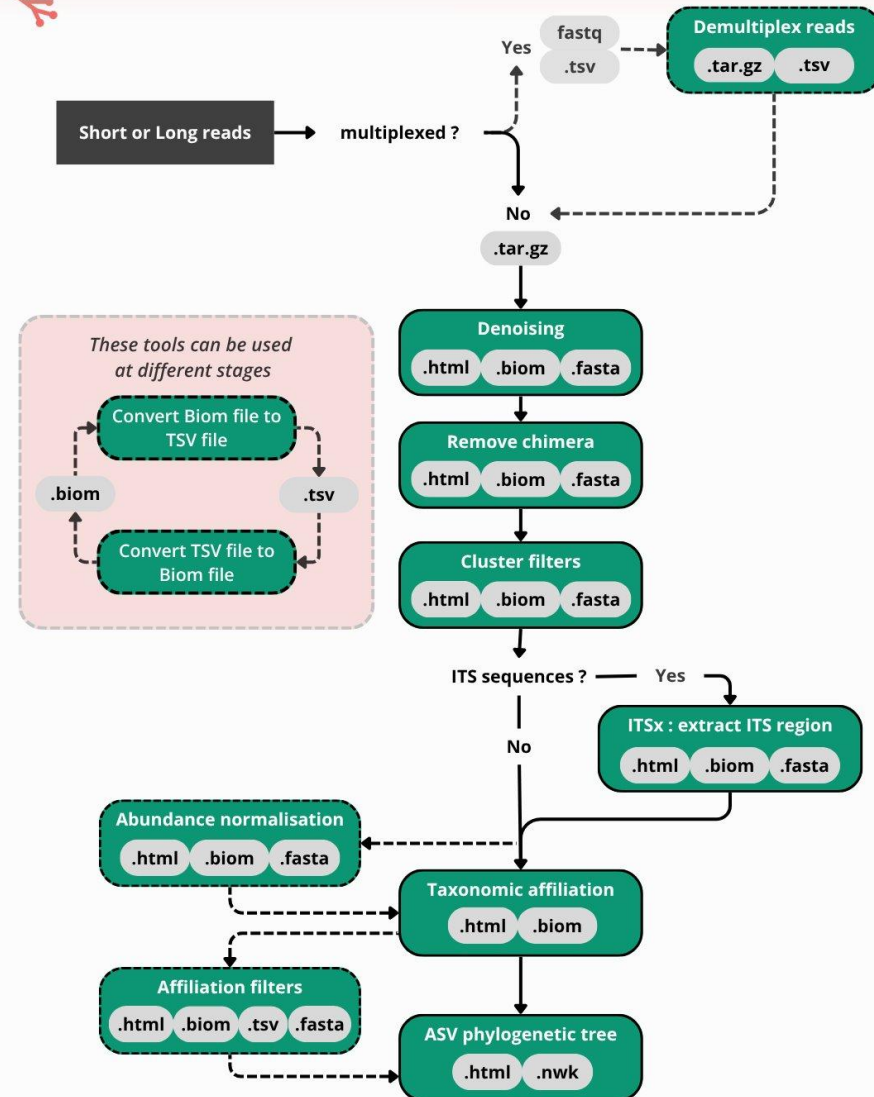
Run Tool

Tool Parameters

Select a tool from the FROGS Core suite to run your analysis.

☒ -- Please select a tool --

- ☐ Read demultiplexing
- ☐ Affiliation filters
- ☐ Affiliation postprocessing
- ☐ Abundance normalisation
- ☐ Convert Biom file to TSV file
- ☐ Convert TSV file to Biom file
- ☐ Cluster/ASV report
- ☐ Affiliation report



FROGS-stat

 **FROGS Stat**

Analysis of community structure, composition, and differential abundances
(Galaxy Version 5.0.2+galaxy0)

Tool Parameters

Select a tool from the FROGS Stat suite to run your analysis.

☒ -- Please select a tool --

☐ Phyloseq: Import data

☐ Phyloseq: Taxonomic composition analysis

☐ Phyloseq: Alpha diversity analysis

☐ Phyloseq: Beta diversity analysis

☐ Phyloseq: Sample clustering analysis

☐ Phyloseq: Structure analysis (based on ordination methods)

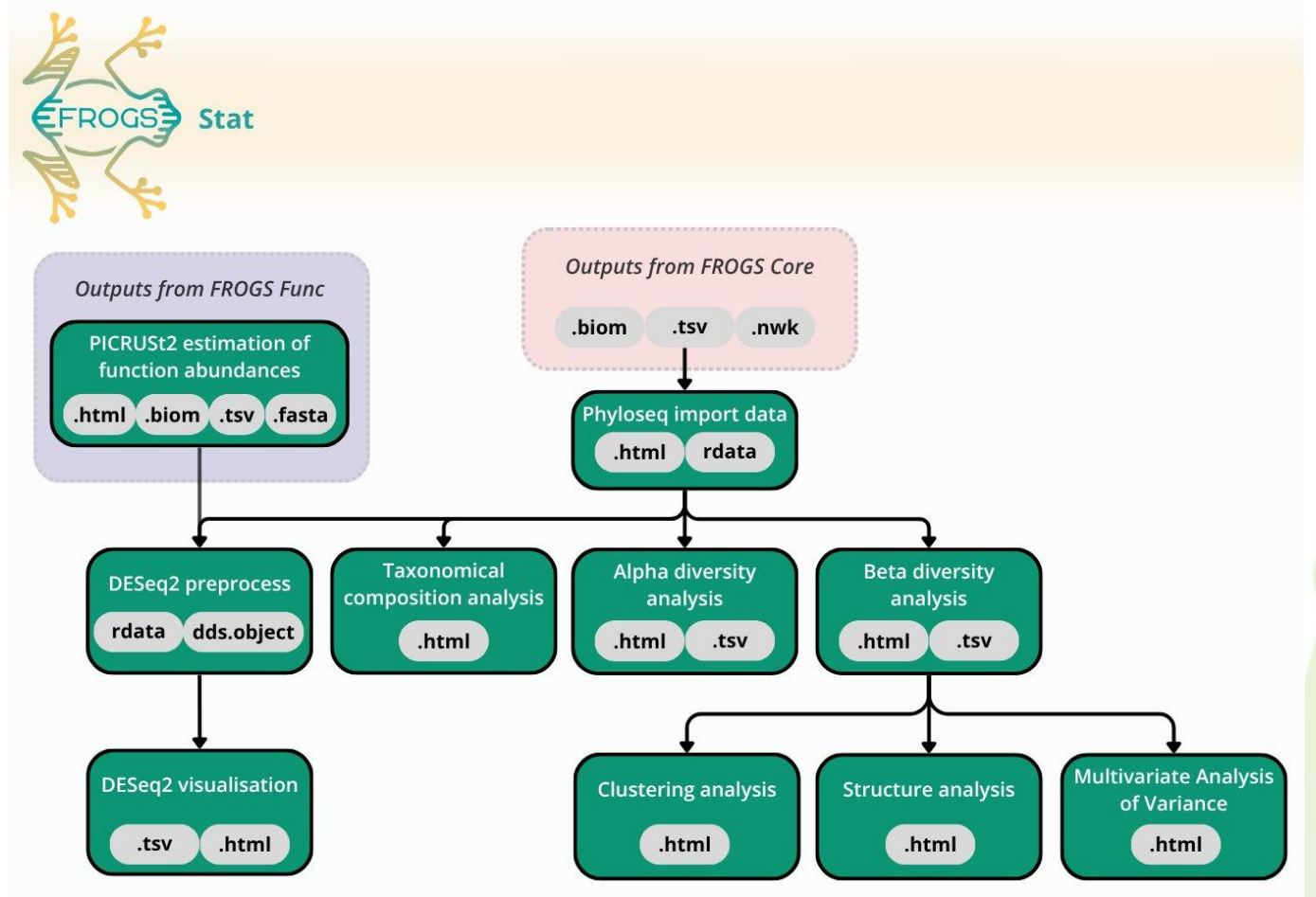
☐ Phyloseq: Multivariate Analysis of Variance

☐ DESeq2: Preprocess for differential analysis of ASV

☐ DESeq2: Preprocess for differential analysis of FUNCTION

☐ DESeq2: Visualisation of differential analysis of ASV

☐ DESeq2: Visualisation of differential analysis of FUNCTION



FROGS-func

FROGS Func

Prediction of functional abundances using marker gene sequences

(Galaxy Version 5.0.2+galaxy0)



Run Tool

Tool Parameters

Select a tool from the FROGS Func suite to run your analysis.

☒ -- Please select a tool --

- ☐ Tool 1: sequence placement and gene copy number estimation (PICRUSt2)
- ☐ Tool 2: functional abundance estimation (PICRUSt2)
- ☐ Tool 3: pathway abundance estimation (PICRUSt2)



Outputs from FROGS Core

.fasta .biom

PICRUSt2 placement in a phylogenetic tree and estimation of gene copy number

.html .biom .tsv .fasta
.txt .nwk

PICRUSt2 estimation of function abundances

.html .biom .tsv .fasta

Inputs for FROGS Stat

DESeq2 preprocess

rdata dds.object

PICRUSt2 estimation of pathway abundances

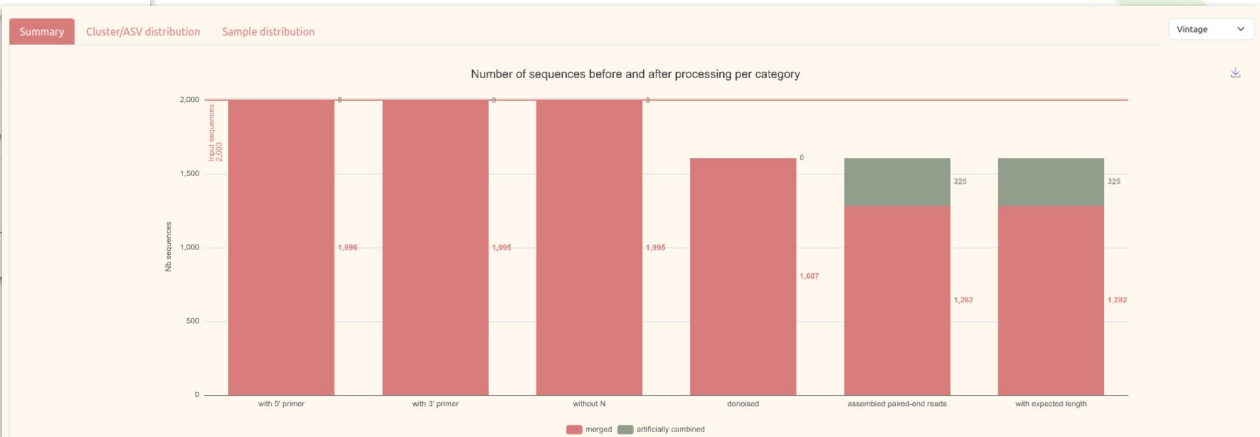
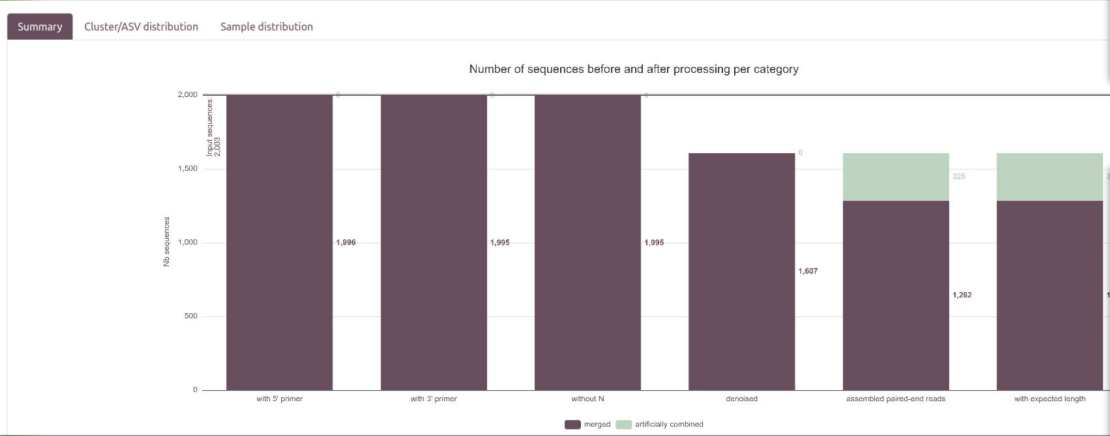
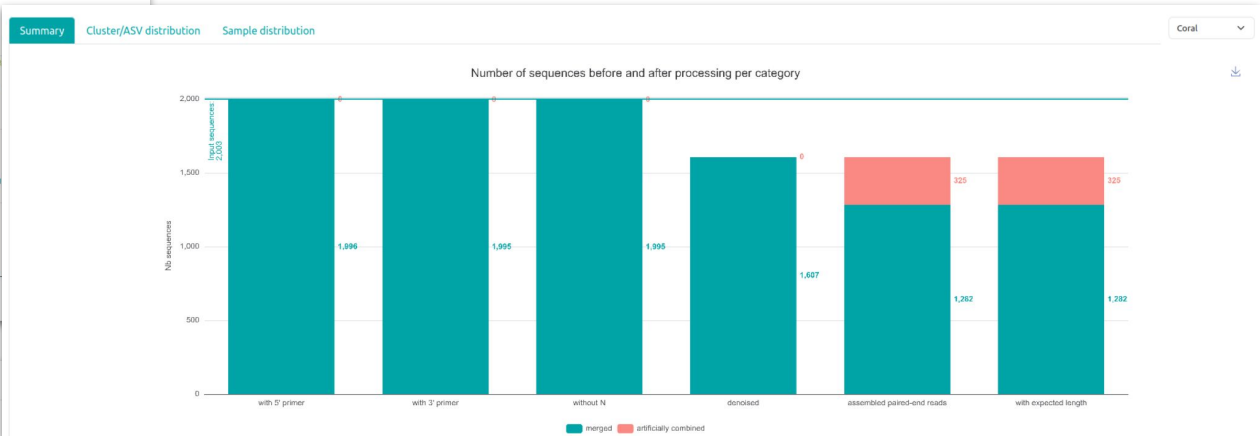
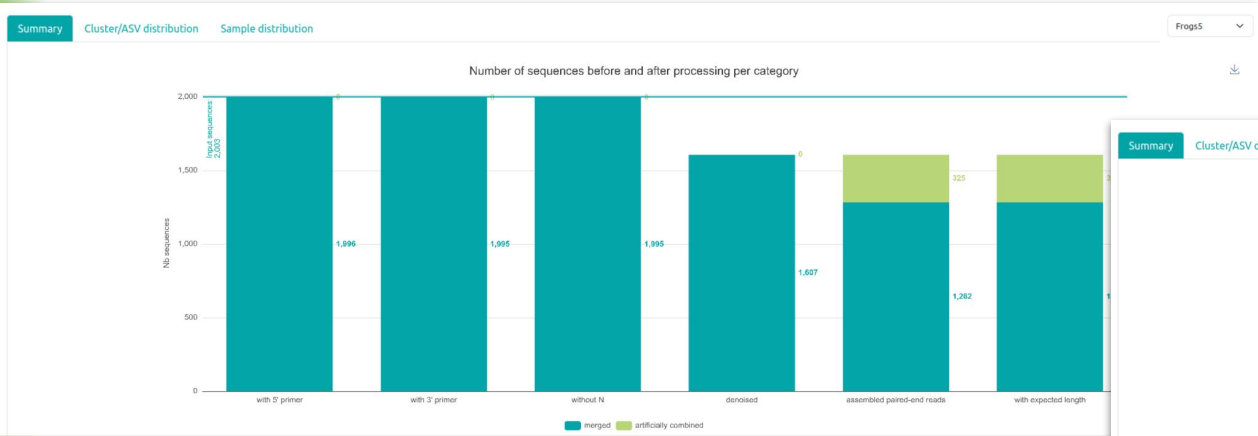
.html .tsv



2. Remove Highcharts library on all our reports

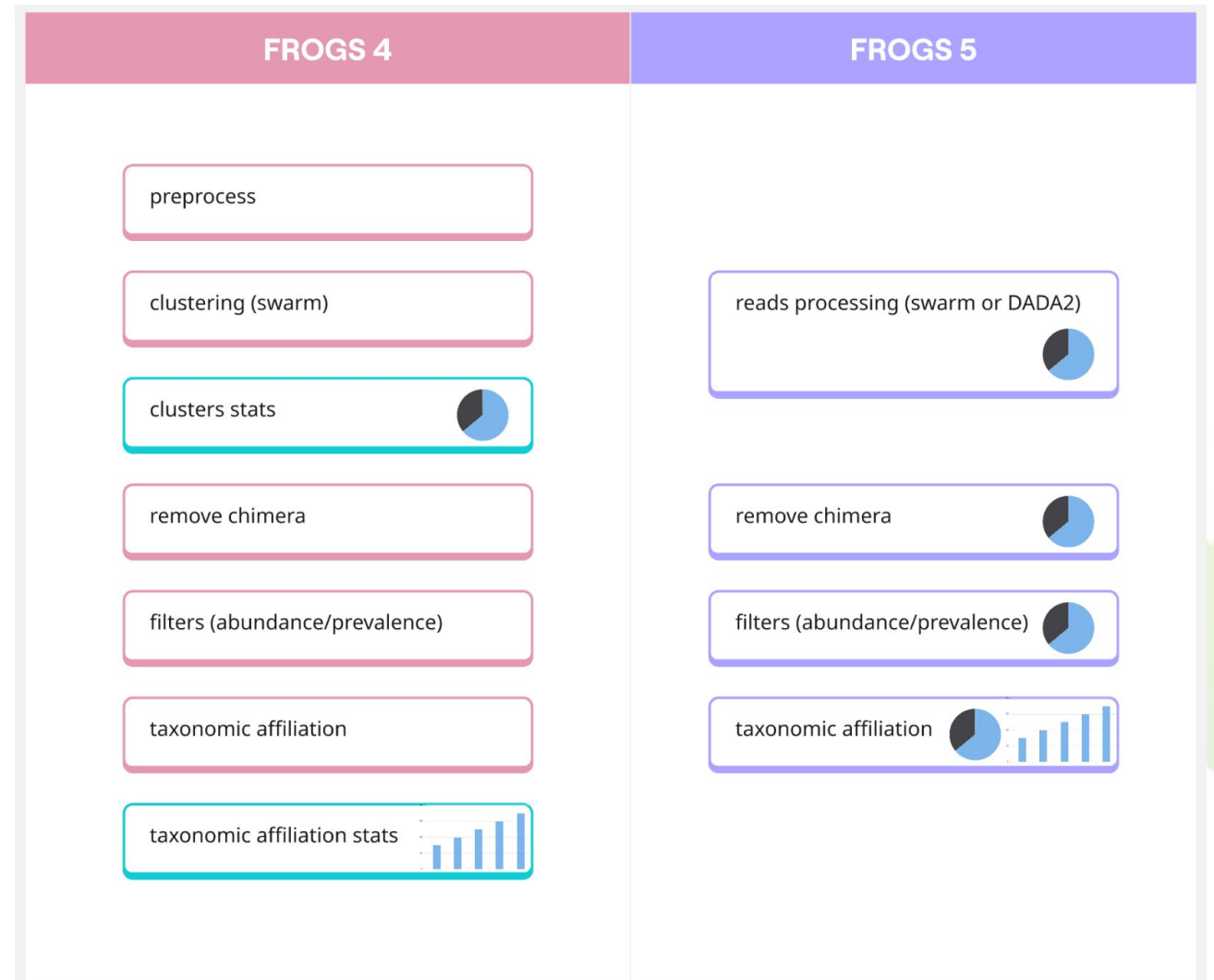


Theme switch



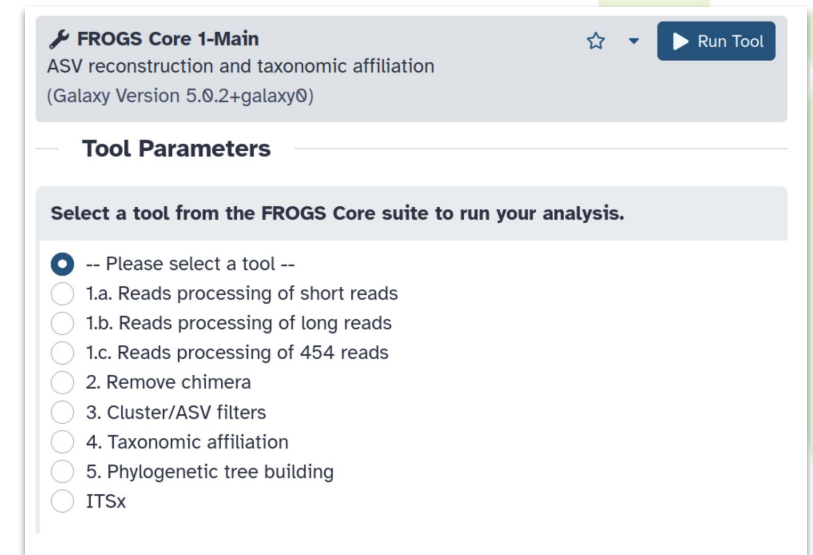
3. Reorganization and evolution of the old preprocess tool

- ❑ Now called “reads processing”
- ❑ *Clustering* and *cluster_stats* included



3. Reorganization and evolution of the old preprocess tool

- ❑ Splitted by reads category in Galaxy
 - ❑ 1.a Reads processing of *short* reads
 - ❑ 1.b Reads processing of *long* reads
 - ❑ 1.c Reads processing of *454* reads



FROGS Core 1-Main ☆ ▾ ▶ Run Tool
ASV reconstruction and taxonomic affiliation
(Galaxy Version 5.0.2+galaxy0)

Tool Parameters

Select a tool from the FROGS Core suite to run your analysis.

- ☒ -- Please select a tool --
- ☐ 1.a. Reads processing of short reads
- ☐ 1.b. Reads processing of long reads
- ☐ 1.c. Reads processing of 454 reads
- ☐ 2. Remove chimera
- ☐ 3. Cluster/ASV filters
- ☐ 4. Taxonomic affiliation
- ☐ 5. Phylogenetic tree building
- ☐ ITSx

DADA2 as an alternative to Swarm

Pros

- Able to distinguish ASVs with 1 difference
- Useful for longitudinal studies (merge based on ASV sequence)

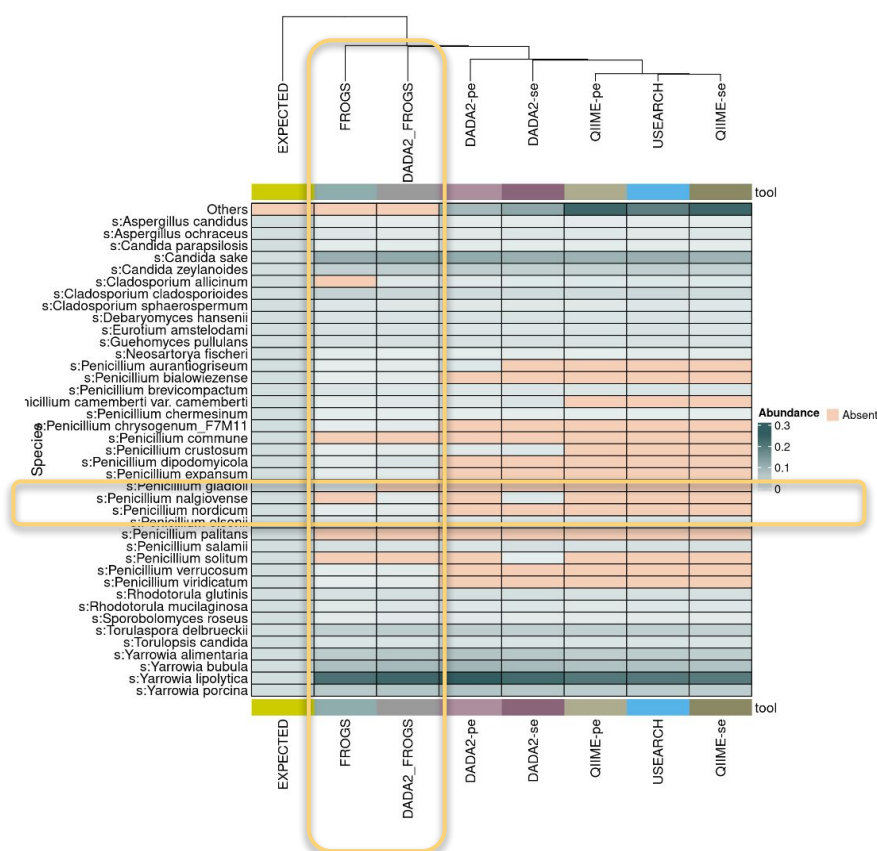
Cons

- Needs a minimal amount of reads (diversity of qualities)
- Technical biases considered as biological variations
- No error model available for all sequencing technologies

Our tests on several datasets show similar results at Genus level

Illustration

MEAT ITS1 DNA MEAT ITS1 PCR MEAT ITS2 DNA MEAT ITS2 PCR MEAT D1/D2 DNA
MEAT D1/D2 PCR MEAT RPB2 DNA MEAT RPB2 PCR



1 10 20 30 40 50 60
Penicillium_nalgiove
Penicillium_nordicum
Consensus
GAAATGCGATACGTAATGTGAATTGCAAAATTCAGTGAATCATCGAGCTTTGAACGCACA
GAAATGCGATACGTAATGTGAATTGCAAAATTCAGTGAATCATCGAGCTTTGAACGCACA
GAAATGCGATACGTAATGTGAATTGCAAAATTCAGTGAATCATCGAGCTTTGAACGCACA

61 70 80 90 100 110 120
Penicillium_nalgiove
Penicillium_nordicum
Consensus
TTGCGCCCCCTGGTATTCGGGGGGCATGCTGTCCGAGCGTCATTGCTGCCCTCAAGCC
TTGCGCCCCCTGGTATTCGGGGGGCATGCTGTCCGAGCGTCATTGCTGCCCTCAAGCC
TTGCGCCCCCTGGTATTCGGGGGGCATGCTGTCCGAGCGTCATTGCTGCCCTCAAGCC

121 130 140 150 160 170 180
Penicillium_nalgiove
Penicillium_nordicum
Consensus
CGGCTTGTTGTTGGGCCCCGTCCTCCGATCCGGGGGACGGGCCGAAAGGCAGCGGCG
CGGCTTGTTGTTGGGCCCCGTCCTCCGATCCGGGGGACGGGCCGAAAGGCAGCGGCG
CGGCTTGTTGTTGGGCCCCGTCCTCCGATCCGGGGGACGGGCCGAAAGGCAGCGGCG

181 190 200 210 220 230 240
Penicillium_nalgiove
Penicillium_nordicum
Consensus
GCACCGCGTCCGGTCTCGAGCGTATGGGGCTTTGTACCCGCTCTGTAGGCCCGGCCGG
GCACCGCGTCCGGTCTCGAGCGTATGGGGCTTTGTACCCGCTCTGTAGGCCCGGCCGG
GCACCGCGTCCGGTCTCGAGCGTATGGGGCTTTGTACCCGCTCTGTAGGCCCGGCCGG

241 250 260 270 280 290 300
Penicillium_nalgiove
Penicillium_nordicum
Consensus
CGCTTGCCGATCAACCCAAATTTTATCCAGGTTGACCTCGGATCAGGTAGGGATACCCG
CGCTTGCCGATCAACCCAAATTTTATCCAGGTTGACCTCGGATCAGGTAGGGATACCCG
CGCTTGCCGATCAACCCAAATTTTATCCAGGTTGACCTCGGATCAGGTAGGGATACCCG

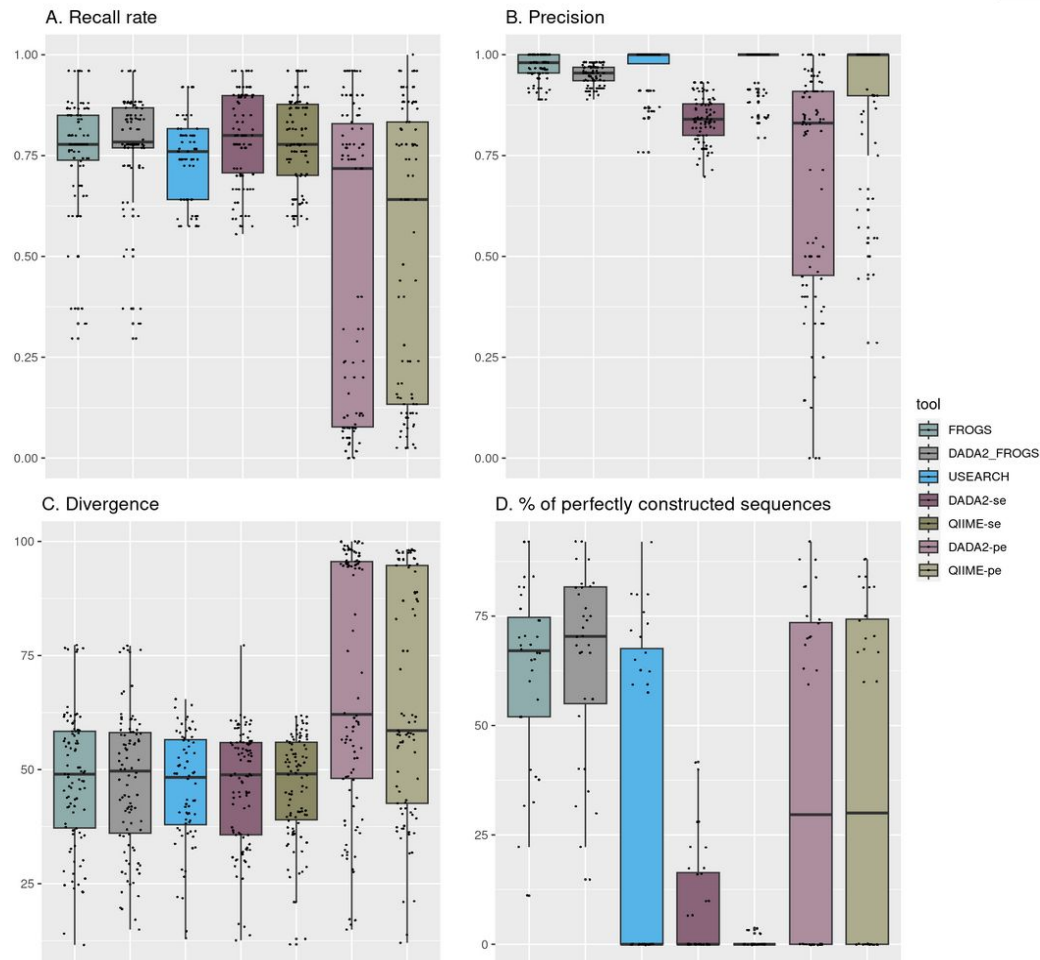
301 310
Penicillium_nalgiove
Penicillium_nordicum
Consensus
CTGAACCTTA
CTGAACCTTA
CTGAACCTTA

RESEARCH ARTICLE
Published
2023-10-05
Cite as
Olivier Rué, Monika Coton, Eric Dugat-Bony, Kate Howell, Françoise Irlinger, Jean-Luc Legras, Valentin Loux, Elisa Michel, Jérôme Mounier, Cécile Neuveglise and Delphine Sicard (2023) Comparison of metabarcoding taxonomic markers to describe fungal communities in fermented foods, Peer Community Journal, 3: e97.
Correspondence
delphine.sicard@univ-lorraine.fr
Peer-review
https://doi.org/10.24072/pcjournal.321

Comparison of metabarcoding taxonomic markers to describe fungal communities in fermented foods
Olivier Rué^{1,2}, Monika Coton³, Eric Dugat-Bony⁴, Kate Howell⁵, Françoise Irlinger⁶, Jean-Luc Legras⁷, Valentin Loux^{8,12}, Elisa Michel⁹, Jérôme Mounier¹⁰, Cécile Neuveglise¹¹, and Delphine Sicard¹²
Volume 3 (2023), article e97
https://doi.org/10.24072/pcjournal.321



Why not DADA2 only?



- ❑ The combination of DADA2 and FROGS shows better results, especially for retrieving exact sequences [5]
- ❑ FROGS and DADA2+FROGS are best at finding exact expected sequences (D)
- ❑ DADA2+FROGS offers efficient tools for merging pairs (vsearch, pear), remove chimera (vsearch and FROGS sample cross-validation)
- ❑ Similar results between DADA2+FROGS and FROGS for most datasets
- ❑ We offer FROGS users the opportunity to choose and/or compare the approach they prefer

DADA2 as an alternative to Swarm

- ❑ Works well for:
 - ❑ Illumina
 - ❑ AVITI
 - ❑ PacBio
- ❑ Not really suitable for:
 - ❑ 454 and IonTorrent (not implemented yet)
 - ❑ Oxford Nanopore Technology

4. Improve management of long reads

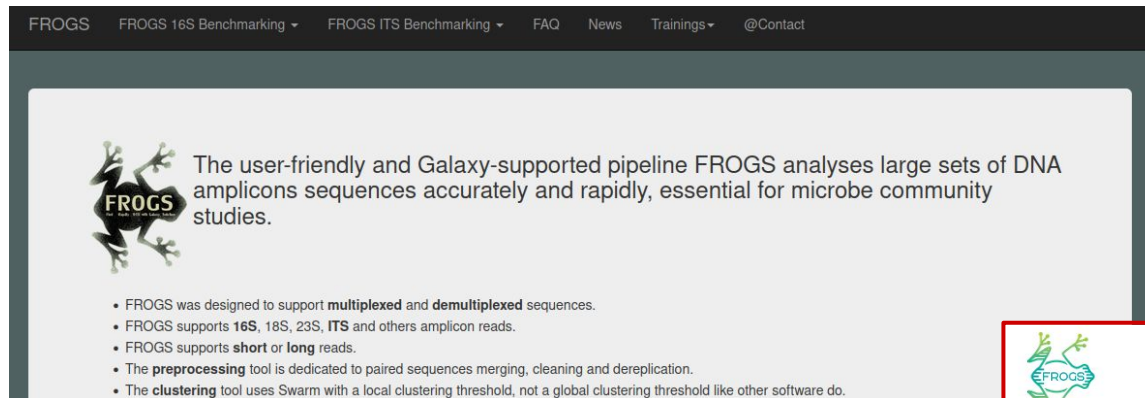
- ❑ Reads processing
 - ❑ add PacBio error function for DADA2
- ❑ Remove chimera
 - ❑ “chimera_denovo” vsearch algorithm
- ❑ Taxonomic affiliation - Blast metrics
 - ❑ Add subject length / coverage

Tutorials to guide users through the change

- ❑ Command line and Galaxy
- ❑ Best practices / Standard Operational Procedure
- ❑ and...

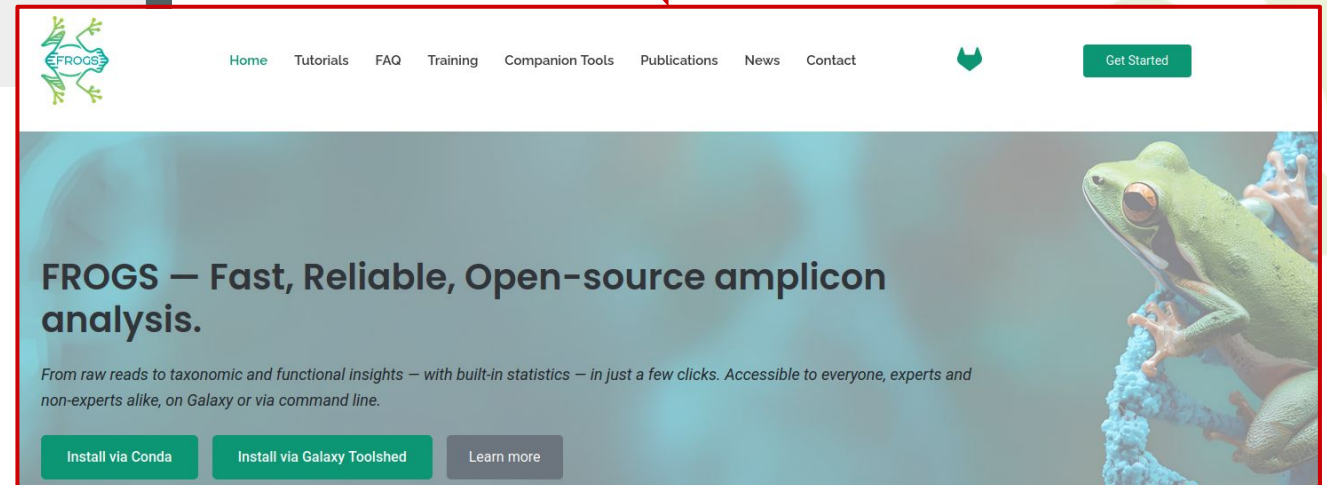
New FROGS website

<https://frogs.inrae.fr/>



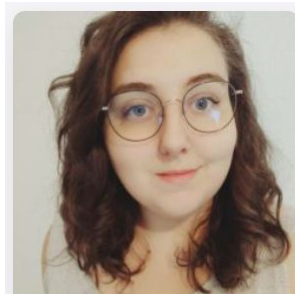
The screenshot shows the old FROGS website interface. It has a dark header with navigation links: FROGS, FROGS 16S Benchmarking, FROGS ITS Benchmarking, FAQ, News, Trainings, and @Contact. The main content area features the FROGS logo (a green frog) and a description: "The user-friendly and Galaxy-supported pipeline FROGS analyses large sets of DNA amplicons sequences accurately and rapidly, essential for microbe community studies." Below this, there is a list of bullet points:

- FROGS was designed to support **multiplexed** and **demultiplexed** sequences.
- FROGS supports **16S**, 18S, 23S, **ITS** and others amplicon reads.
- FROGS supports **short** or **long** reads.
- The **preprocessing** tool is dedicated to paired sequences merging, cleaning and dereplication.
- The **clustering** tool uses Swarm with a local clustering threshold, not a global clustering threshold like other software do.



The screenshot shows the new FROGS website interface. It has a light header with navigation links: Home, Tutorials, FAQ, Training, Companion Tools, Publications, News, and Contact. There is a GitHub icon and a "Get Started" button. The main content area features a large image of a green frog and the text: "FROGS — Fast, Reliable, Open-source amplicon analysis." Below this, there is a description: "From raw reads to taxonomic and functional insights — with built-in statistics — in just a few clicks. Accessible to everyone, experts and non-experts alike, on Galaxy or via command line." At the bottom, there are three buttons: "Install via Conda", "Install via Galaxy Toolshed", and "Learn more". A red arrow points from the old website to the new one.

Thank you for your attention



Gabryelle Agoutin

UMR GenPhySE, INRAE
Occitanie-Toulouse



Lucas Auer

UMR IAM, INRAE Nancy



Maria Bernard

UMR GABI, Sigenae, INRAE Jouy-
en-Josas



Maëlle Pomiès

UMR IAM, INRAE Nancy



Olivier Rué

UR MaIAGE, Migale, INRAE Jouy-
en-Josas



Géraldine Pascal

UMR GenPhySE, INRAE
Occitanie-Toulouse