



**HAL**  
open science

## **ANOMALY: AmplicoN wOrkflow for Microbial community AnaLYsis**

Sébastien Theil, Etienne Rifa

### ► **To cite this version:**

Sébastien Theil, Etienne Rifa. ANOMALY: AmplicoN wOrkflow for Microbial community AnaLYsis. 5. Colloque de Génomique Environnementale, Oct 2019, La Rochelle, France. <hal-02340484>

**HAL Id: hal-02340484**

**<https://hal.science/hal-02340484v1>**

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# ANOMALY : AMPLICON WORKFLOW FOR MICROBIAL COMMUNITY ANALYSIS

Sebastien Theil<sup>1</sup>, Etienne Rifa<sup>1</sup>

<sup>1</sup>Université Clermont Auvergne, INRA, VetAgro Sup, UMR Unité Mixte de Recherche sur le Fromage, 20 côte de Reyne, 15000 Aurillac, France

## Context

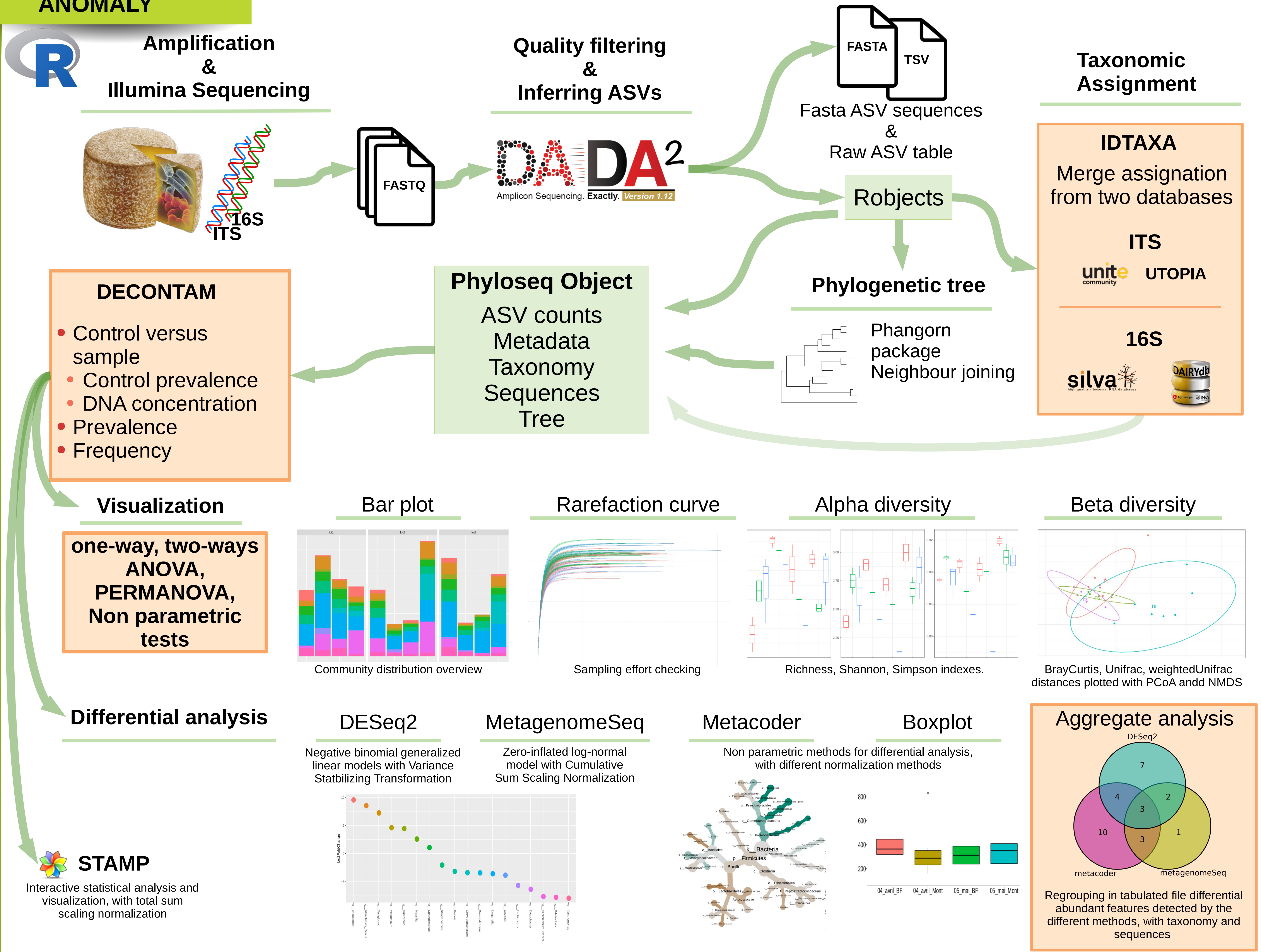
Bioinformatic tools for amplicon sequencing data analysis are continuously and rapidly evolving, thus integrating most recent techniques and tools is challenging.

We present ANOMALY, an R workflow for 16S and ITS amplicons sequencing. This workflow is based on several Rscripts using mainly DADA2 and phyloseq packages amongst others. This workflow allows to perform an analysis from raw fastq sequence files to final statistical analysis. The objective was to automate bioinformatic analyses to ensure reproducibility between projects and trying to be versatile and simple enough to integrate new bioinformatic tools or statistical techniques.

## Features

- All integrated analyse 16S & ITS in R
- DADA2 Amplicon Sequence Variant
- Decontamination from control samples (decontam)
- Filtering : frequency, prevalence
- Taxonomic assignment from 2 references with validation
- Integrated statistics and visualization
- Multiple differential analyse
- Versatility (easy to integrate new features)
- Use well established methods

## ANOMALY



## Conclusions

ANOMALY uses latest approved methods like :

- **Amplicon sequence variant** allows more resolution in taxonomic unit detection
- **Double taxonomic assignment** with environment dedicated and general database increases accuracy
- **Contamination control** reduces interpretation errors
- **Integrated statistical analysis** allowing reproducibility, all global analysis **in one tool**.
- **Constant development** with regular updates.

Available on  **git** repository at :  
<https://forgemia.inra.fr/umrf/anomaly>