



Feedback on a comparative metatranscriptomic analysis

Cedric Midoux, Tiago P. Delforno, Thais Z. Macedo, Gileno V. Lacerda,
Olivier Rué, Mahendra Mariadassou, Maria B.A. Varesche, Théodore
Bouchez, Ariane Bize, Valéria M. Oliveira, et al.

► To cite this version:

Cedric Midoux, Tiago P. Delforno, Thais Z. Macedo, Gileno V. Lacerda, Olivier Rué, et al.. Feedback on a comparative metatranscriptomic analysis. JOBIM 2019: Journées Ouvertes Biologie, Informatique et Mathématiques, Jul 2019, Nantes, France. , 2019, JOBIM 2019. hal-02735859

HAL Id: hal-02735859

<https://hal.inrae.fr/hal-02735859>

Submitted on 2 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.

Context

The progress of next generation sequencing favors the development of more comprehensive ecosystem studies thanks to metatranscriptomic approaches. These latter can indeed provide access to functional information at a good analysis depth. Through a study of anaerobic digesters treating wastewater contaminated by an anionic surfactant (namely the *linear alkylbenzene sulfonate*, LAS), we developed a bioinformatics pipeline to perform RNAseq analysis of the metatranscriptomics data.



Data

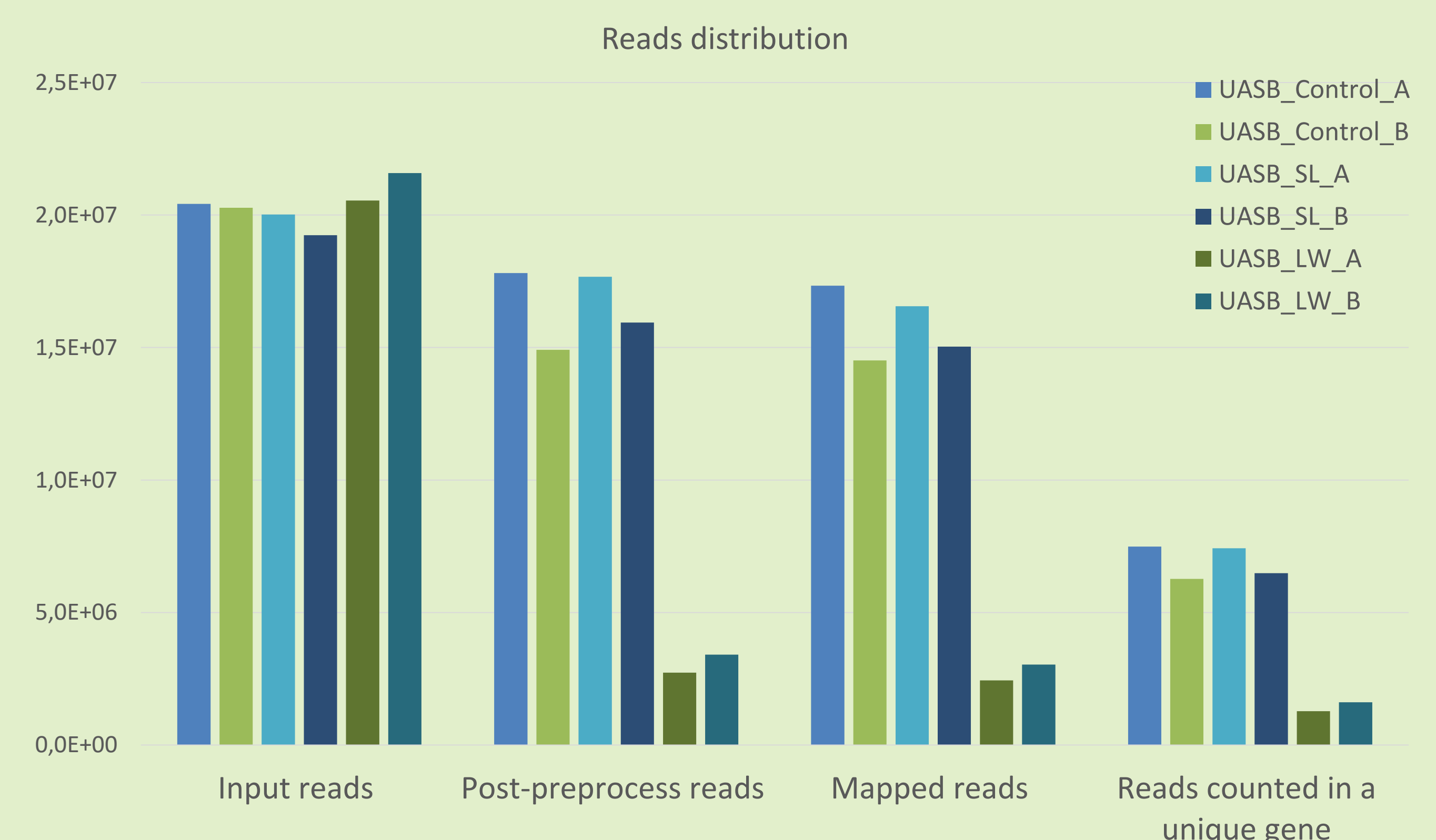
The samples originated from distinct biological reactors :

- fed with synthetic medium (UASB_Control)
- fed with synthetic medium + LAS (UASB_SL)
- fed with real laundry wastewater (UASB_LW)

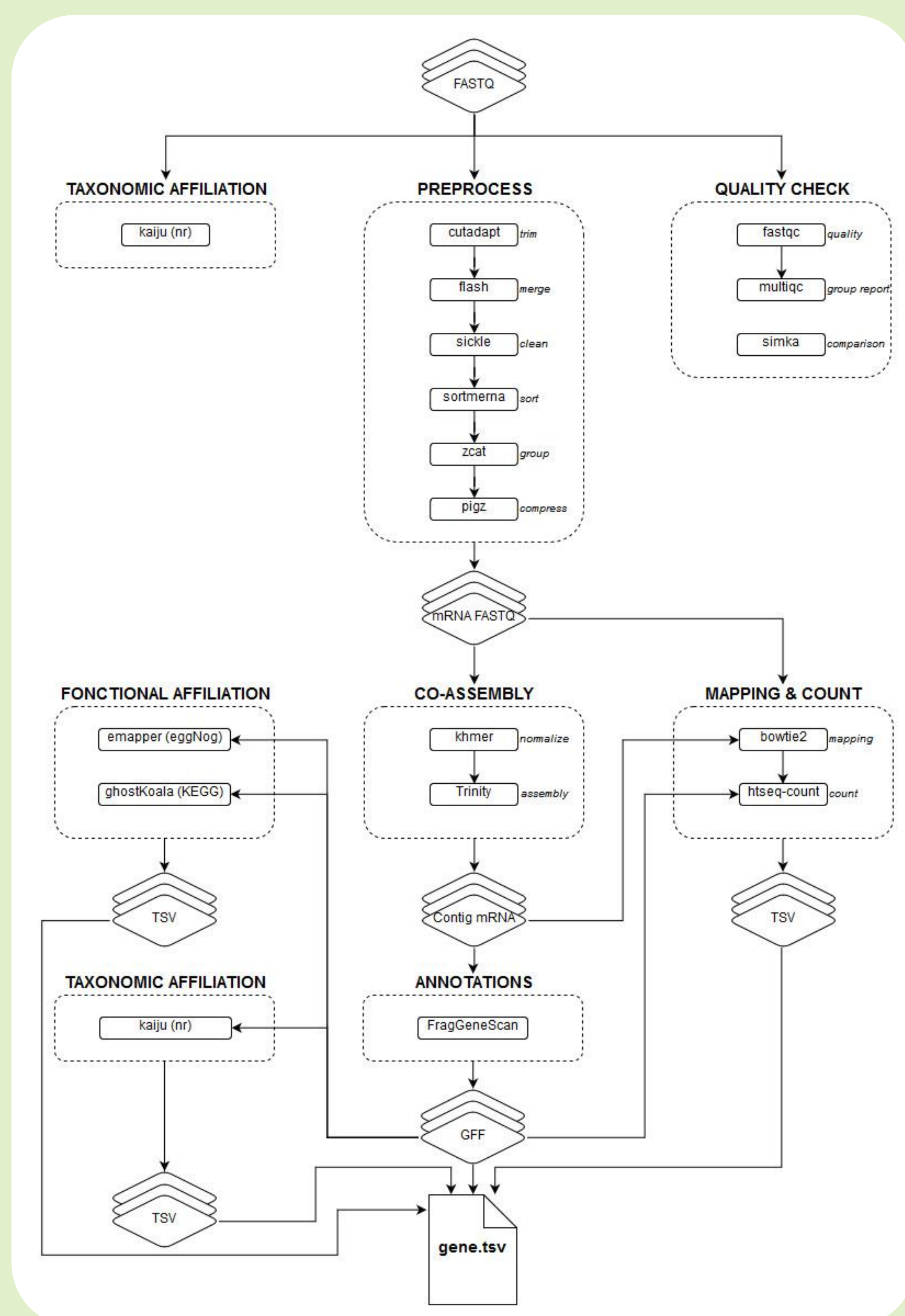
These samples were sequenced in duplicates with *Illumina HiSeq* platform (2 x 100bp ; ~20M reads/sample)



Metrics



Tools and workflow



- All results, counts and affiliation, were gathered together in a table.
- The tables and metadata were imported as R Phyloseq objects.
- Thresholds enabling to discard the genes with non-reproducible counts were determined for each sample thanks to the technical replicates, by scatterplot analysis. These genes corresponded to transcripts detected at low levels.
- Study of the biological information corresponding to genes with highly different expression levels were performed on the two-by-two comparison of the samples using the Z-distribution

Results

- Percentages of LAS removal (~50%) were similar among the three reactors.
- Rarefaction analysis revealed that two million reads were sufficient to access to the main functional capacity.
- In the first step of LAS biodegradation pathway, fumarate reductase subunit C was detected and taxonomically assigned to the genus *Syntrophobacter*.
- In the second step, many enzymes related to beta-oxidation were expressed preferentially in UASB_SL and UASB_LW. They were taxonomically related to *Smithella*, *Acinetobacter* and *Syntrophorhabdus*.
- Finally, the adenylylsulfate reductase, taxonomically related to *Desulfovibrio* and *Desulfomonile*, was observed in the desulfonation step with the highest relative abundance in UASB_LW.

Conclusions

- ➔ This study enabled detecting active genes likely involved in each step of LAS biodegradation and exploring the microbial active core related to LAS degradation (Delforno *et al.*, 2019).
- ➔ Based on this positive experience, we developed a snakemake pipeline including all this features for use in future metatranscriptomic projects.

Delforno, T.P., *et al.*, Comparative metatranscriptomic analysis of anaerobic digesters treating anionic surfactant contaminated wastewater. *Sci Total Environ*, **649**: p. 482-494 (2019).