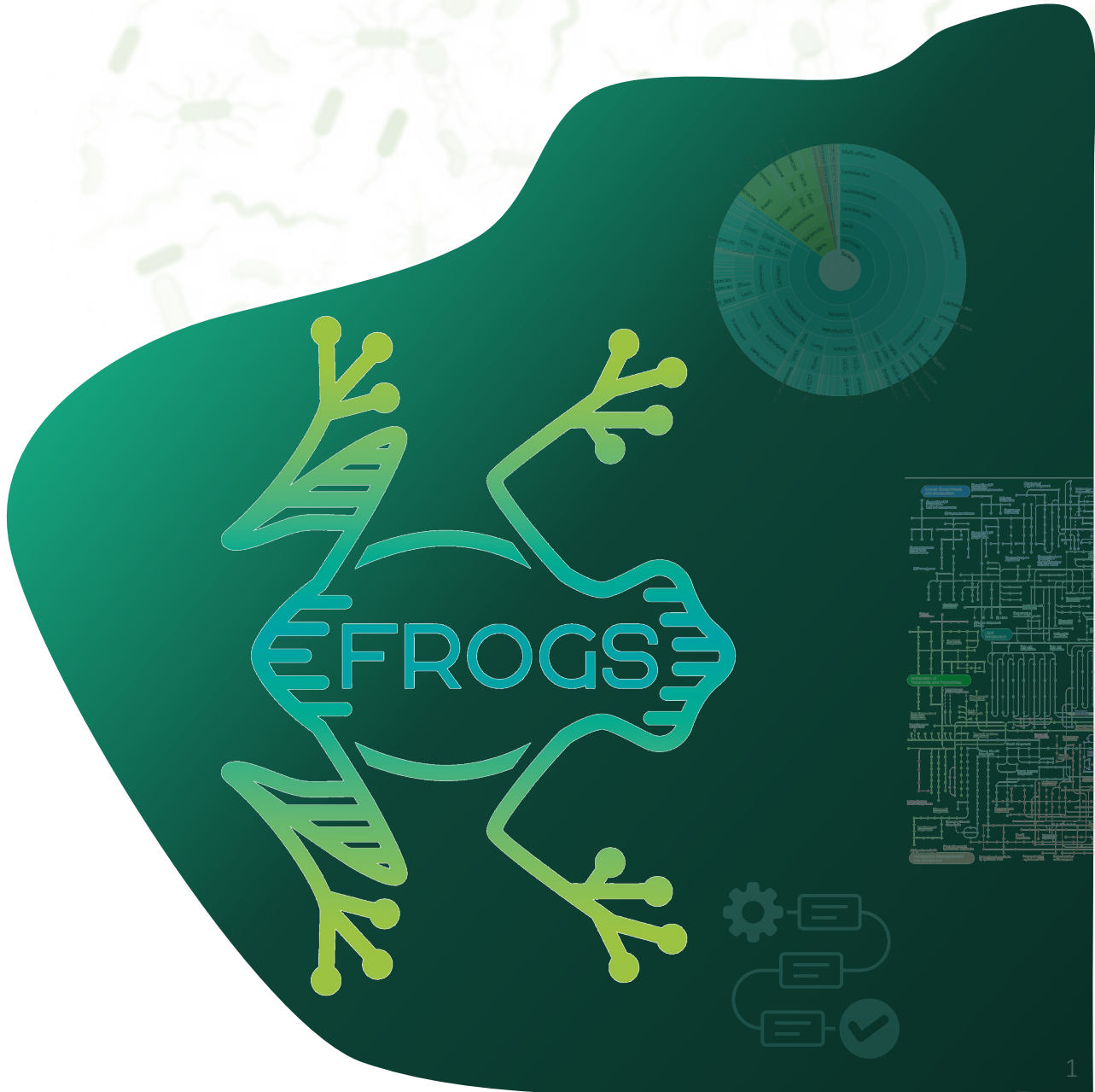
Using FROGS by itself

on long-reads data

Lucas AUER, Gabryelle Agoutin, Maria BERNARD, Géraldine PASCAL, Maëlle POMIÈS & Olivier RUÉ



Presentation of the dataset

Effect of live yeast supplementation in sow diet during gestation and lactation on sow and piglet fecal microbiota, health, and performance

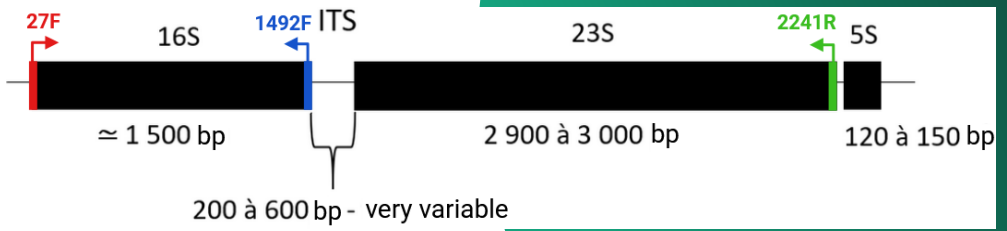
Nathalie Le Floc'h,[†,1] Caroline Stéphanie Achard,[‡] Francis Amann Eugenio,[†] Emmanuelle Apper,[‡] Sylvie Combes,[‖] and Hélène Quesnel[†]

48 sows and their piglets

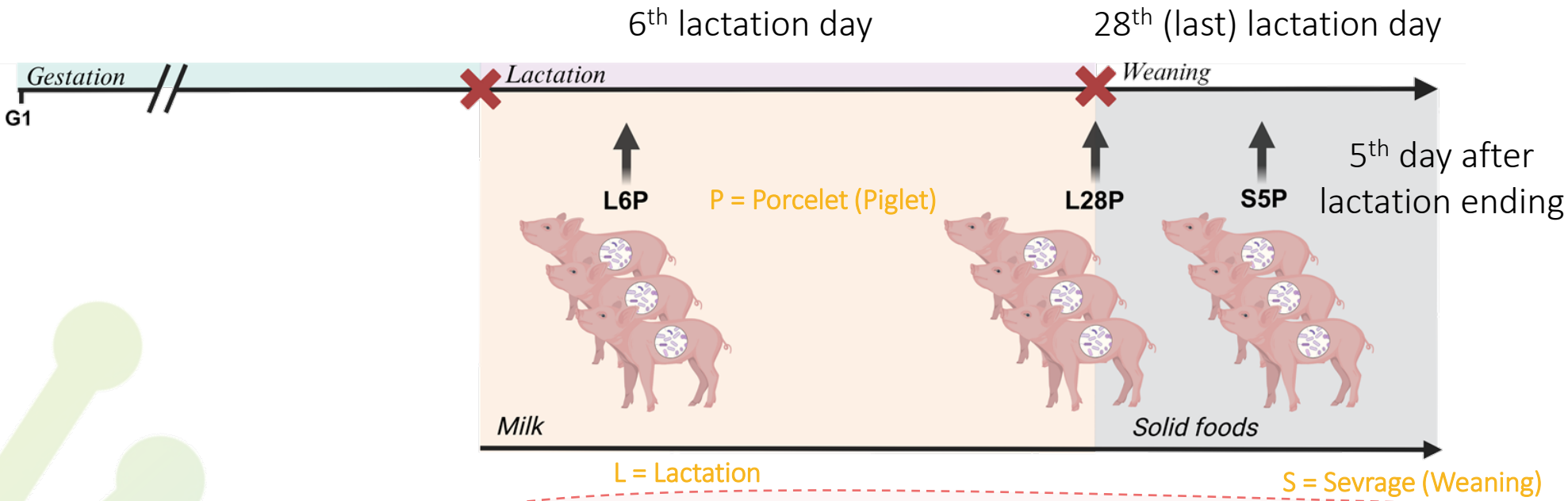
16S V3-V4 illumina

SeqOccIn

16S-ITS-23S PacBio



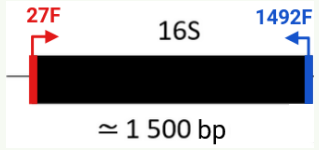
Presentation of the training subset



3 time points x 4 families x 3 piglets per family → 36 samples

L6P	107, 109 (building1)	1
L28P		2
S5P	309, 311 (building3)	3

Marker is the full 16S:



PCR primers:
27F: AGRGTTYGATYMTGGCTCAG
1492F: AAGTCGTAACAAGGTARCY
(both in 5' 3')

Practice session

Create a new history, you may entitle it “Piglet_16S_longread” for exemple

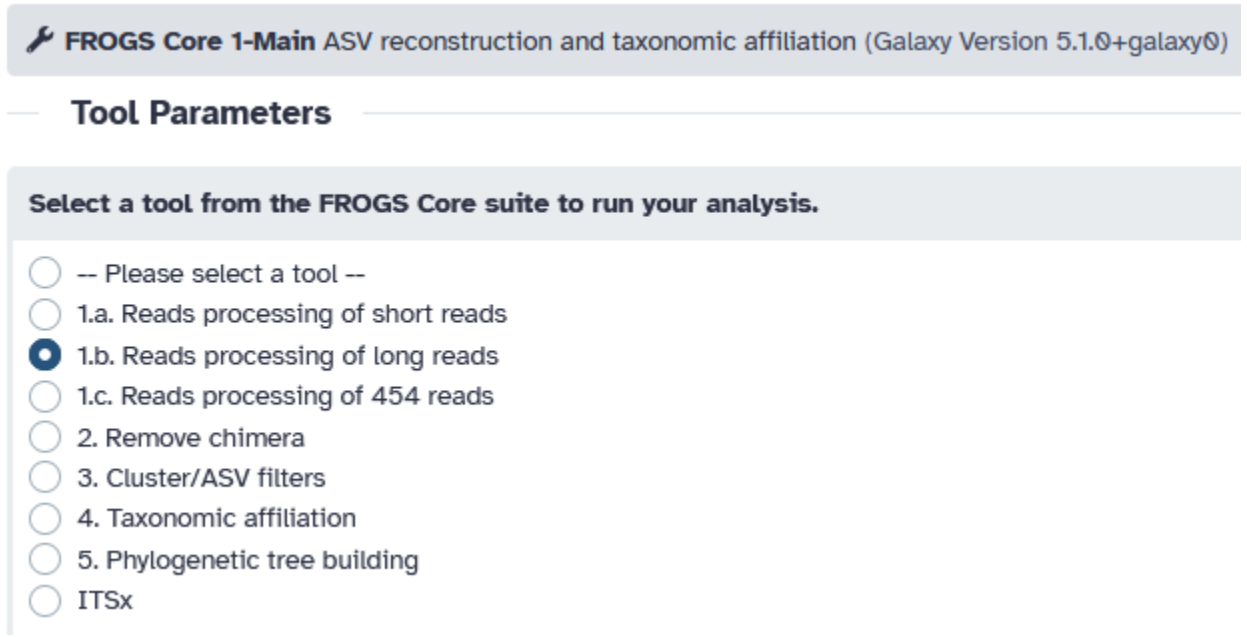
Please, upload the data and metadata files that you find here :
https://web-genobioinfo.toulouse.inrae.fr/~formation/15_FROGS/30_March2026/Piglets_16S_long.tar.gz
https://web-genobioinfo.toulouse.inrae.fr/~formation/15_FROGS/30_March2026/Piglets_metadata.tsv

Choose FROGS Core Main 1b Reads processing of long reads, use swarm to create ASVs.

Modify parameters and run the process !

Practice session

Choose FROGS Core Main 1b Reads processing of long reads



Practice session

Modify parameters and run the process !

Practice session

Do the sequences include PCR primers?

- (●) Yes
- ( ) No

Indicate whether the sequences still include PCR primers. Select “Yes” if primers are present, “No” if they have already been removed

5' primer *

AGRGTTYGATYMTGGCTCAG

Enter the 5' primer sequence. Wildcards are allowed. The sequence must be provided in 5' → 3' orientation. (--five-prim-primer)

3' primer *

AAGTCGTAACAAGGTARCY

Enter the 3' primer sequence. Wildcards are allowed. The sequence must be provided in 5' → 3' orientation. (--three-prim-primer)

Practice session

Process type

- ○ Preprocessing only
- ◉ Preprocessing and clustering with Swarm
- ○ Preprocessing and denoising with DADA2. /!\ Only Pacbio sequencers are currently supported /!\.

Select the type of process to run (--process)

Swarm distance threshold *

1

Distance threshold used by Swarm for clustering. (--distance)

Clustering refinement

- ◉ With --distance = 1, refine clusters with Swarm --fastidious option (recommended since FROGS 3.2)
- ○ With --distance > 1, perform a pre-clustering step with FROGS --pre-clustering option
- ○ No clustering refinement

(i) With --distance = 1, use the Swarm --fastidious option to refine clustering (recommended since FROGS 3.2). (ii) With --distance > 1, enable pre-clustering to reduce redundancy before final clustering step. (iii) Select this option to apply neither refinement nor pre-clustering.

Practice session

Continue to process the data by run each tool step by step.