

# Paleofish Mitochondries

Objectif : Lancement de Eager sur référence majoritaire

# CR 17 sept 2025

Voici les points abordés:

Présentation des résultats du multimapping et de Eager pour *S. salar* et *S. trutta*.

Il a été décidé:

- \* Retirer MB1 et MB2 car ces échantillons ont déjà été traités dans le projet Paleotrota.
- \* Lancer les 5 échantillons ni *S. salar* ni *S. trutta*, sur un nouveau multifasta (composé de plusieurs mito génomes intra genres) fournit par Joelle.
- \* Pour les échantillons ayant un nombre de reads proches *salar* et *trutta*, générer le consensus avec les 2 génomes mito de référence pour comparer les consensus. Les traitements seront dans le répertoire 02\_eager\_consensus\_comparison/.

# 5 échantillons non S. trutta ni S. salar

smaman@genobioinfo1 /work/project/crucial/PALEOFISH/01\_Multimapping \$ grep 'NC\_006531.1' RESULTATS\_MAJORITAIRES  
**L11\_7\_SCO\_1354\_S70:NC\_006531.1:182488** → **Anguille**

smaman@genobioinfo1 /work/project/crucial/PALEOFISH/01\_Multimapping \$ grep 'NC\_012928.1' RESULTATS\_MAJORITAIRES  
**L0\_4\_EG9\_S84:NC\_012928.1:19566** → **Thymallus thymallus**

smaman@genobioinfo1 /work/project/crucial/PALEOFISH/01\_Multimapping \$ grep 'NC\_018341.1' RESULTATS\_MAJORITAIRES  
**L2\_4\_CD2\_S74:NC\_018341.1:5799** → **Brachymystax**

smaman@genobioinfo1 /work/project/crucial/PALEOFISH/01\_Multimapping \$ grep 'NC\_020762.1' RESULTATS\_MAJORITAIRES  
**L0\_5\_EG10\_S85:NC\_020762.1:38793** → **Coregonus**  
**L7\_14\_22d\_S48:NC\_020762.1:1255682**

=> Joelle : liste des autres génomes mito pour lancer un multimapping sur ce nouveau multifasta.

4 Undetermined

5 Autres que S. salar et S.trutta

32 S. salar

47 S. trutta

---

88 échantillons

# Échantillons écartés

Echantillons avec gaps donc écartés, déjà traités dans Paleotrutta avec l'option no gap de Eager:

**MB1**

**MB2**

Autres :

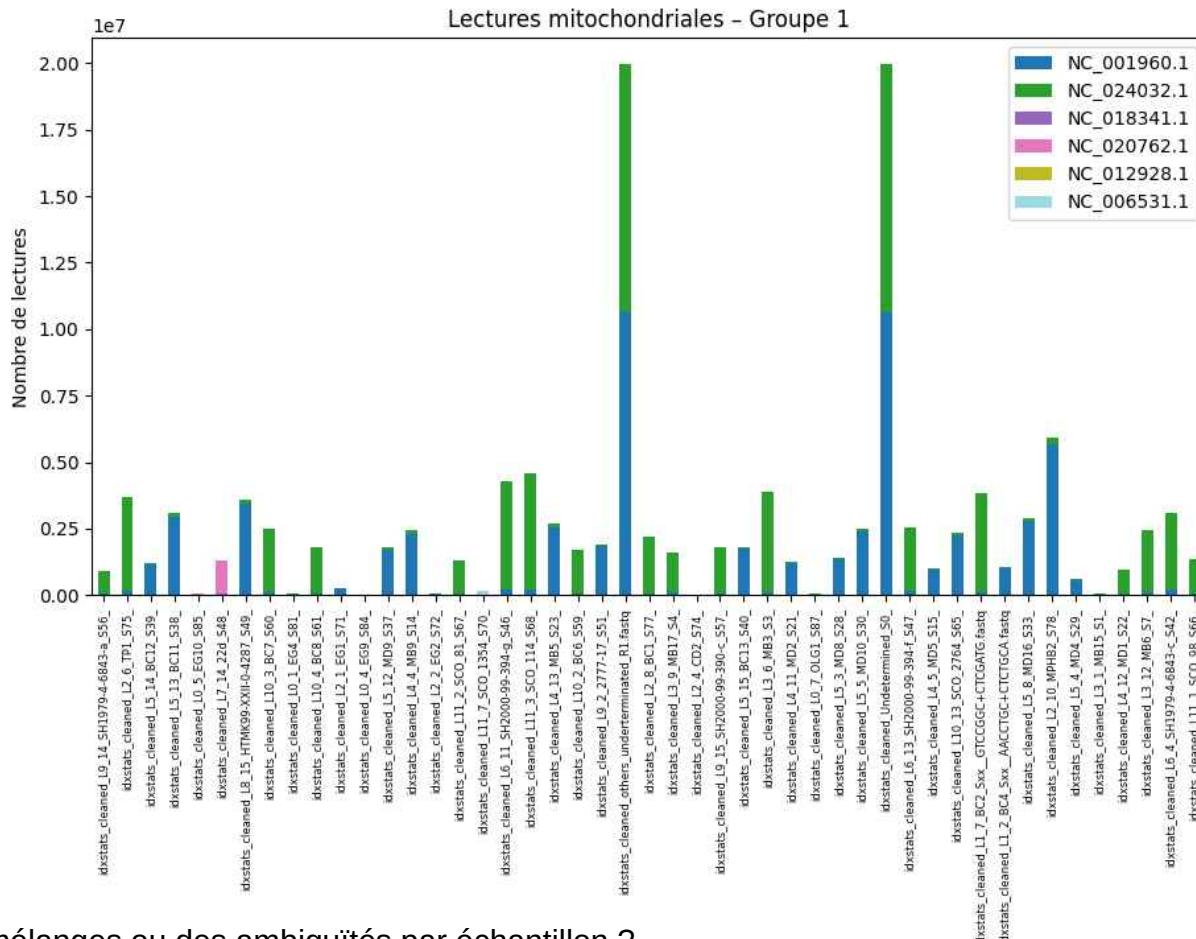
# Résultats du pipeline pour S. salar et S. trutta

```
/work/project/crucial/PALEOFISH/02_eager_39_mito_genomes/crucial_nextflow_VersionPaleofish.config

/work/project/crucial/PALEOFISH/02_eager_39_mito_genomes/NC_001960.1:
01_script_eager_part1_NC_001960.1.sh 03_eager_only_genotyping documentation input_NC_001960.1.tsv pipeline_info reference_genome slurm-24635376.out
02bis-bam_tsv.sh adapterremoval endorspy mapdamage pipeline_trace.txt reheader_bam slurm-24635526.out
02_rehead_bam_files_v2_NC_001960.1.sh crucial_nextflow_VersionPaleofish.config FastP mapping preseq samtools work
02ter_tsv_modif.sh deduplication fastqc multiqc qualimap slurm-24562555.out

/work/project/crucial/PALEOFISH/02_eager_39_mito_genomes/NC_024032.1:
01_script_eager_part1_NC_024032.1.sh 03_eager_only_genotyping FastP L4_6_MB2_slurm-24562820.out pipeline_trace.txt work
02bis-bam_tsv.sh adapterremoval fastqc OLD_with_Sxx_input_NC_024032.1.tsv reference_genome
02_rehead_bam_files_v2_NC_024032.1.sh documentation input_NC_024032.1.tsv pipeline_info slurm-24635219.out
smaman@genobioinfo1 /work/project/crucial/PALEOFISH/02_eager_39_mito_genomes/NC_024032.1 $ █
```

# Multimapping BWA 1/1



Question : Avons-nous des mélanges ou des ambiguïtés par échantillon ?  
 /work/project/crucial/PALEOFISH/01\_detect\_species/STATS/graph2.py  
 Pas d'ambiguïté sauf pour les fichiers Undertermined.

# Multimapping BWA 2/2

Lectures mitochondriales - Groupe 2

