

Paleofish Mitochondries

Objectif : Lancement de Eager sur référence majoritaire

CR réunion du 24 oct 2025

Présentation du projet paleofish_mt dans la forge INRAE : Forge : https://forge.inrae.fr/bios4biol/mt_paleofish

Présentation de la méthode de préparation des génomes.

Présentation des séquences consensus obtenues pour les 5 échantillons not Salmo.

A venir:

Intégrer le site web https://web-genobiinfo.toulouse.inrae.fr/~smaman/BOOTSTRAP/projet_crucial dans la forge paleofish_mt, contenu à mettre à jour pour certaines pages.

Enrichir le répertoire data/ de paleofish_mt avec l'échantillon undetermined BC4, un échantillon truite Troubat et un échantillon not salmo Troubat. Joelle nous donnera les noms de ces 2 derniers échantillons cités. Les échantillons AUD sont, pour le moment, mis de côté. Ajout de graphs résultats dans le README.

Remonter au NCBI les caractères windowsiens dans les mitogénomes.

Partage du pipeline draw.io

Publication JOSS du paleofish_mt à préparer.

Traiter les échantillons undetermined en retirant les séquences non paires.

Prochaine réunion le vendredi 7 novembre 2025 à 9h30.

Partage des scripts dans la forge INRAE

The screenshot shows a web browser window with the address bar displaying `forge.inrae.fr/bios4biol/mt_paleofish`. The browser's address bar and tabs are visible at the top. The main content area shows the repository page for `MT_Paleofish` under the `bios4biol` organization. A yellow warning banner at the top of the repository page states: "Your account is authenticated with SSO or SAML. To push and pull over HTTPS with Git using this account, you must set up a personal access token." Below this banner, the repository name `MT_Paleofish` is displayed with a lock icon. A dropdown menu shows the current branch as `main` and the repository name as `mt_paleofish`. To the right of the repository name are buttons for `+ v`, `Find file`, and `Code v`. Below the repository name, there is a section titled `Add new directory` with the text `Sarah Maman-Haddad authored 1 minute ago`. To the right of this section are buttons for `fde04733`, a copy icon, and `History`. On the left side of the browser window, a sidebar is visible with a search bar and a list of project items: `MT_Paleofish`, `Pinned`, `Issues` (0), and `Merge requests` (0). The sidebar also has a `Manage` button at the bottom.

https://forge.inrae.fr/bios4biol/mt_paleofish → CHANGER NOM / Joss ??

Site web déployé depuis la forge

Le lien générée par la forge est le suivant: <https://mt-paleofish-75c9c6.pages-forge.inrae.fr/>
Mais il est possible d'y accéder directement depuis README.

🔗 MT_Paleofish

Name

Project PALEOFISH mitochondries

Web site link

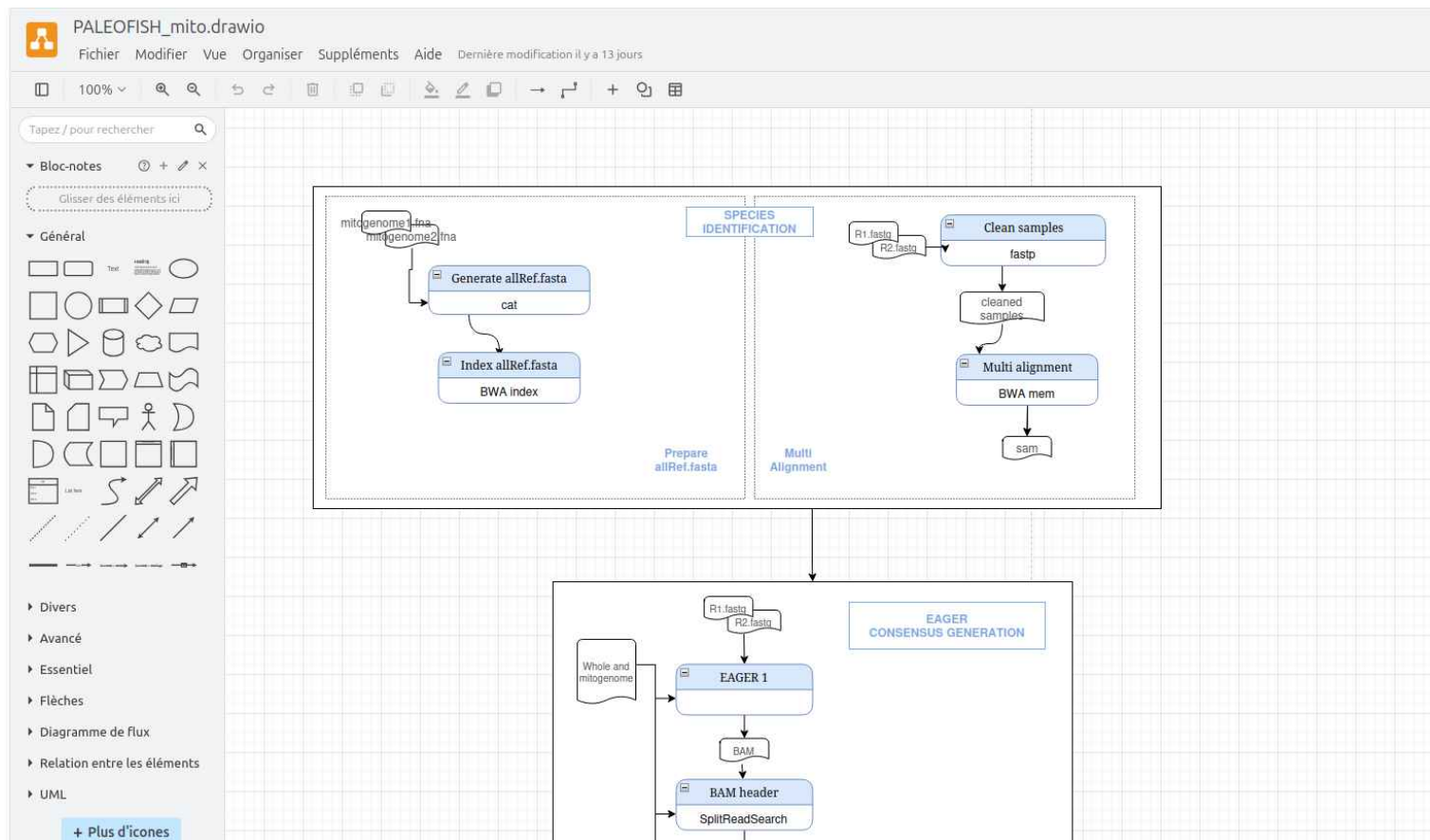


The screenshot displays the MT_Paleofish website. At the top, the 'Crédits' section lists the contributors: Stefanie WAGNER, Christophe KLOPP, and Sarah MAMAN. Below this, the 'STATISTIQUES' section is titled 'SEQUENCES ANALYSIS'. It contains several interactive buttons with icons and text:

- FLASH histograms**: Click here to see all FLASH histograms results. (Icon: pink book)
- Sequences statistiques**: Click here to see all sequences statistiques. (Icon: yellow book)
- 1st alignement on multiFasta**: References results comparisons. (Icon: green list)
- 2nd alignement on majoritary reference**: Click here to see alignments results and statistics. (Icon: green book)
- mapDamage & underterminated FASTQ files**: mapDamage results and underterminated sequences detection. (Icon: yellow globe)
- References reconstructions**: Path to references constructions and VCF stats. (Icon: blue clock)
- Eager results files**: Eager pipeline. (Icon: green globe)

Partage draw.io

- Flowmaker en ligne
- Mail de partage ?



CR réunion du 3 novembre

- Joelle :
 - * Nom du pipeline/outil plutôt que nom du projet / JOSS.
 - * Cibler les fichiers à traiter en exemple dans DATA/ ? Autres échantillons publics ancien et moderne ??
- Cf. CR réunion du 24/10 : Enrichir le répertoire data/ de paleofish_mt avec l'échantillon undetermined BC4, un échantillon truite Troubat et un échantillon not salmo Troubat. Joelle nous donnera les noms de ces 2 derniers échantillons cités. Les échantillons AUD sont, pour le moment, mis de côté.*

```
smaman@genobioinfo2 /work/project/crucial/PALEOFISH/FASTQ_MITO/troubat $ ls
```

```
L3_10_MB13_S5_R1_001.fastq.gz L3_12_MB6_S7_R2_001.fastq.gz L3_15_MB12_S10_R1_001.fastq.gz L3_6_MB3_S3_R2_001.fastq.gz
L4_3_MB16_S13_R1_001.fastq.gz L4_7_MB1_S17_R2_001.fastq.gz
L3_10_MB13_S5_R2_001.fastq.gz L3_13_MB11_S8_R1_001.fastq.gz L3_15_MB12_S10_R2_001.fastq.gz L3_9_MB17_S4_R1_001.fastq.gz
L4_3_MB16_S13_R2_001.fastq.gz
L3_11_MB18_S6_R1_001.fastq.gz L3_13_MB11_S8_R2_001.fastq.gz L3_1_MB15_S1_R1_001.fastq.gz L3_9_MB17_S4_R2_001.fastq.gz
L4_6_MB2_S16_R1_001.fastq.gz
L3_11_MB18_S6_R2_001.fastq.gz L3_14_MB10_S9_R1_001.fastq.gz L3_1_MB15_S1_R2_001.fastq.gz L4_1_MB14_S11_R1_001.fastq.gz
L4_6_MB2_S16_R2_001.fastq.gz
L3_12_MB6_S7_R1_001.fastq.gz L3_14_MB10_S9_R2_001.fastq.gz L3_6_MB3_S3_R1_001.fastq.gz L4_1_MB14_S11_R2_001.fastq.gz
L4_7_MB1_S17_R1_001.fastq.gz
```

CR réunion du 3 novembre

Choisir 2 MB selon Salmo salar (NC_001960.)/trutta(NC_024032.1) et selon le nb de séquences alignées (https://mt-paleofish-75c9c6.pages-forge.inrae.fr/align_stats.html)

Saumon : L4_13_MB5_S23 NC_001960.1 2 537 533 **ou** L4_4_MB9_S14 NC_001960.1 2309563 **ou** L4_8_MB4_S18 NC_001960.1 2530104 **ou** L5_6_MB8_S31 NC_001960.1 1327909

Trutta : L3_1_MB15_S1 NC_024032.1 40851 **ou** L3_10_MB13_S5 NC_024032.1 4 036 626 **ou** L3_11_MB18_S6 NC_024032.1 2049395 **ou** L3_12_MB6_S7 NC_024032.1 2324679 **ou**

L3_13_MB11_S8 NC_024032.1 628834

L3_14_MB10_S9 NC_024032.1 3487773

L3_15_MB12_S10 NC_024032.1 2599352

L3_6_MB3_S3 NC_024032.1 3741776

L3_9_MB17_S4 NC_024032.1 1499826

L4_1_MB14_S11 NC_024032.1 704764

L4_3_MB16_S13 NC_024032.1 500638

L4_4_MB9_S14 NC_001960.1 2309563

L4_6_MB2_S16 NC_024032.1 2021408

L4_7_MB1_S17 NC_024032.1 2805539

L4_8_MB4_S18 NC_001960.1 2530104

CR réunion du 3 novembre

- * Mise à jour du site web.
- * Retours sur le pipeline draw.io
- Sarah :
 - * Remonter au NCBI les caractères windowsiens dans les mitogénomes.
 - * Traiter les échantillons undetermined en retirant les séquences non pairées.
 - * Ajout de graphs résultats dans le README.
 - * IGV sur les séquences consensus not Salmo.
- Prochaine réunion le