Nos séquences de *Syphacia* sont alignées avec Mega X et des séquences de référence des différentes espèces de *Syphacia* (cf. page suivante). Une matrice des distances et un arbre phylogénétique a été créé, les valeurs de bootstrap ont été calculés.

Les clusters et les localisations dans les tubes digestifs ont permis de déterminer les espèces. L’attribution des séquences au séquences a été fait via R

**S. obvelata**

Séquence 1

>GQ260137.1 Syphacia obvelata isolate PMZ5 cytochrome oxidase subunit 1 (cox1) gene, partial cds; mitochondrial

<https://www.ncbi.nlm.nih.gov/nuccore/GQ260137.1>

Article en rapport avec analyse du génome mitochondriale

Même si non publié semble fiable

Séquence 2

>MG386204.1 Syphacia obvelata isolate HUP6-15 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial

<https://www.ncbi.nlm.nih.gov/nuccore/MG386204.1>

Article d’Afrique du Sud qui semble fiable sur la phylogénie des helminthes

Séquences 3

>HM204792.1 Syphacia obvelata isolate PMZ2b cytochrome oxidase subunit 1 (cox1) gene, partial cds; mitochondrial

<https://www.ncbi.nlm.nih.gov/nuccore/HM204792.1>

Même article que 1 donc peut être clone mais pas le même nombre de pb

S. stroma

Séquence 4

>MF142428.1 Syphacia stroma isolate WALES-GWYN-12AS235Ss cytochrome c oxidase subunit 1 (Cox1) gene, partial cds; mitochondrial

<https://www.ncbi.nlm.nih.gov/nuccore/MF142428.1>

Behnke en dernier auteur

Article sur les syphacia des Muridae en particulier S. frederici

Séquence 5

>LC038091.1 Syphacia stroma mitochondrial COI gene for cytochrome c oxidase subunit 1, partial cds, isolate: CzHam1

<https://www.ncbi.nlm.nih.gov/nuccore/LC038091.1>

Thèse Japon, sur nématodes parasites de indonesian murines avec ref spé aux syphacia

Séquence 6

>MF142420.1 Syphacia stroma isolate FRANCE-12Apo11Ss cytochrome c oxidase subunit 1 (Cox1) gene, partial cds; mitochondrial

<https://www.ncbi.nlm.nih.gov/nuccore/MF142420.1>

Même article que sequence 4 mais mulot vient de France

S. frederici

Séquence 7

>MF142429.1 Syphacia frederici isolate WALES-GWYN-15As01Sff cytochrome c oxidase subunit 1 (Cox1) gene, partial cds; mitochondrial

<https://www.ncbi.nlm.nih.gov/nuccore/MF142429.1>

Même article que séquence 4 et 6

Séquence 8

>MF142423.1 Syphacia frederici isolate NOTTINGHAM-12As74Sff cytochrome c oxidase subunit 1 (Cox1) gene, partial cds; mitochondrial

<https://www.ncbi.nlm.nih.gov/nuccore/MF142423.1>

Même article que 4 mais seul avec frederici

Séquence 9

>MF142426.1 Syphacia frederici isolate PORTUGAL-13As18Sf cytochrome c oxidase subunit 1 (Cox1) gene, partial cds; mitochondrial

<https://www.ncbi.nlm.nih.gov/nuccore/MF142426.1>

Même article

Explication de la méthode de comparaison

**🧪 CONTEXTE RAPIDE**

Tu compares des séquences de différentes espèces de *Syphacia*, donc :

* Ce sont **des espèces proches** → les différences sont petites mais significatives.
* Le but est de **voir la similarité/distance génétique** → tu veux un résultat précis, mais pas faussé par les parties non alignées ou très divergentes.

**✅ RÉGLAGES À CHOISIR**

**1. Model / Method**

* ✅ **p-distance** *(déjà choisi)*

**Pourquoi ?**  
Le p-distance est simple : il calcule le **pourcentage de différences** (nombre de sites différents ÷ nombre total de sites comparés). Très utilisé pour des comparaisons de base entre séquences proches.

**2. Substitutions to include**

* ✅ **Transitions + Transversions** *(choisir cette option)*

**Pourquoi ?**  
Cela inclut **toutes les substitutions possibles**, pas seulement les transitions (changements purine ↔ purine ou pyrimidine ↔ pyrimidine). Comme tu ne veux **rien ignorer**, on choisit les deux.

**3. Gaps / Missing Data Treatment**

Tu as **3 options** :

* **Pairwise deletion** ✅ → **à choisir**
* Complete deletion
* Use all sites

**Choisis : Pairwise deletion**

**Pourquoi ?**

* Les alignements contiennent souvent des **gaps** (indels ou manques).
* Avec **Pairwise deletion**, MEGA calcule chaque distance **en ne prenant en compte que les positions sans gap entre les 2 séquences concernées**.
* C’est plus **souple** que “complete deletion” (qui supprimerait complètement les colonnes avec des gaps dans *n’importe* quelle séquence, ce qui peut jeter beaucoup d'information).

**📸 Résumé rapide des bons choix dans MEGA X :**

| **Option** | **Choix recommandé** | **Pourquoi** |
| --- | --- | --- |
| **Model** | p-distance | Simple, direct |
| **Substitutions to include** | Transitions + Transversions ✅ | Inclut toutes les différences |
| **Gaps / Missing Data** | Pairwise deletion ✅ | Optimise l’info utilisée sans trop perdre de données |

Explication entière :

**ALIGNEMENT DES SÉQUENCES**

**2. Ouvrir MEGA X**

**Lance le logiciel**.

**3. Créer un alignement**

**Dans MEGA X :**

* Clique sur **Align > Edit/Build Alignment > Create a new alignment > DNA** (ou RNA si c’est de l’ARN).

**Pourquoi ?**  
Pour pouvoir comparer les séquences, il faut les **aligner**, c’est-à-dire les mettre en correspondance base par base.

**4. Importer les séquences**

* Dans la fenêtre d’alignement, clique sur **File > Retrieve sequences from file**, puis sélectionne ton fichier FASTA.

**Pourquoi ?**  
Cela charge toutes tes séquences pour les préparer à l’analyse.

**📐 ALIGNEMENT MULTIPLE**

**5. Faire l’alignement multiple**

* Dans la fenêtre d’alignement, clique sur **Alignment > Align by ClustalW** (ou **MUSCLE**, au choix).

**Pourquoi ?**  
L’alignement multiple ajuste les séquences pour que les positions homologues (mêmes positions évolutives) soient alignées. C’est **essentiel** pour comparer des séquences, notamment pour les arbres phylogénétiques.

**⚠️ Vérifie à la fin que les séquences ont à peu près la même longueur et qu’il n’y a pas trop de “trous” (gaps).**

**💾 ENREGISTRER L’ALIGNEMENT**

**6. Sauvegarder l’alignement**

* Dans la fenêtre d’alignement, fais **File > Save session** ou **Export alignment > MEGA format (.meg)**.

**Pourquoi ?**  
Tu vas utiliser ce fichier aligné pour les analyses suivantes (phylogénie, distances, etc.).

**🌳 COMPARAISON ET ANALYSES**

Tu as plusieurs options maintenant :

**✅ Pour comparer les séquences directement (distance génétique) :**

1. Va dans **MEGA X > Analyze > Compute Pairwise Distances**.
2. Choisis ton fichier .meg d’alignement.
3. Choisis un modèle (par exemple **p-distance**, qui compte le nombre de différences).
4. Clique sur **Compute**.

**Pourquoi ?**  
Cela te donne une **matrice de distances** entre toutes les séquences. Si deux séquences sont très proches (distance faible), elles sont probablement de la même espèce ou très apparentées.

**✅ Pour faire un arbre phylogénétique :**

1. MEGA X > **Phylogeny > Construct/Test Neighbor-Joining Tree** (ou Maximum Likelihood).
2. Sélectionne ton fichier aligné.
3. Choisis les options (garde les options par défaut si tu n’es pas sûr).
4. Clique sur **Compute**.

**Pourquoi ?**  
L’arbre montre comment les séquences sont reliées entre elles évolutivement. Si tes échantillons sont proches d’une séquence de référence, cela indique qu’ils appartiennent probablement à la même espèce.

Pour arbre phylogénétique je test la robustesse des nœuds à l’aide de bootsrtap

Une valeur supérieure à 70% c’est bien

 Ouvre MEGA X.

 Va dans **"Phylogeny" > "Construct/Test Neighbor-Joining Tree"** (ou Maximum Likelihood, selon ta méthode).

 Dans la fenêtre de configuration, coche l’option **"Test of Phylogeny"**.

 Choisis **Bootstrap method** et entre un nombre de répétitions (par ex. **1000**).

 Lance la construction de l’arbre.