#### M2 CHPS

# **CHPS1001**

Analyse de séquence

BATISTE Quentin COGNÉ Romain CORTINHAL Simao 2024/2025

#### Introduction

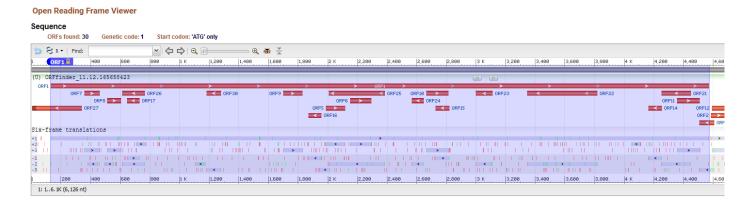
Dans ce second travail pratique de bioinformatique, nous aborderons l'analyse de séquences et l'utilisation des bases de données biologiques pour des applications scientifiques concrètes.

L'étude de cas se concentre sur la mucoviscidose, une maladie génétique fréquente due à l'instabilité du gène CFTR. La protéine CFTR, altérée chez les patients atteints de mucoviscidose, joue un rôle crucial dans la régulation des échanges d'ions chlorure au niveau des muqueuses, et son dysfonctionnement engendre une accumulation de mucus épais, source d'inflammation et de complications sévères.

## 1)Traduction de la séquence d'ADN en protéine

Il y a 6 lignes de résultats car il y a trois cadres dans un sens de lecture, ainsi que 3 dans l'autre sens.

Pour trouver le Codon STOP de l'ORF correspondant à la protéine CFTR, on sélectionne le meilleur candidat dans le sequence viewer, puis au lieu d'afficher la séquence de la protéine, on affiche la séquence nucléotidique, puis on note les trois dernières lettres : « TAG ».



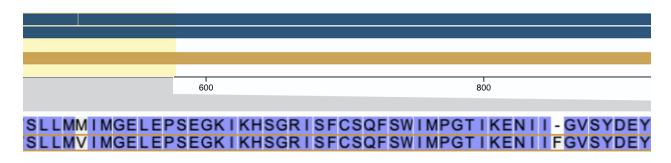
# 2) Recherche de la protéine à laquelle correspond cette séquence via un BLAST

Après avoir effectué une recherche BLAST de la séquence donnée, nous avons pu rechercher la mutation associée à cette dernière.

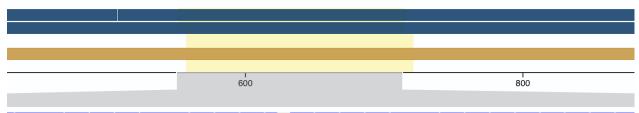
Code d'entrée de la séquence : P13569 (CFTR\_HUMAN)

En analysant la séquence, nous avons pu remarquer différentes mutations, les voici :

• Emplacement des mutations: 470, 508

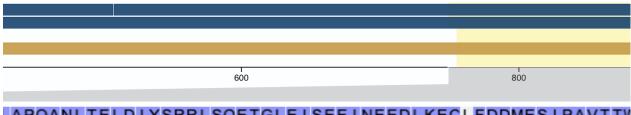


• Emplacement: 620



#### (TRILVTSKMEHLKKADKILIL<mark>N</mark>EGSSYFYGTFSELQNLQPDFSSKLMGCI (TRILVTSKMEHLKKADKILILHEGSSYFYGTFSELQNLQPDFSSKLMGCI

• Emplacement: 833



LAPQANLTELD | YSRRLSQETGLE | SEE | NEEDLKECLFDDMES | PAVTTV LAPQANLTELD | YSRRLSQETGLE | SEE | NEEDLKECFFDDMES | PAVTTV

#### 3) Lecture d'une entrée GenBank

Après avoir réalisé un blastn à partir de la séquence, voici les résultats obtenus : Code d'entrée GenBank : NM 000492.4

On observe quatre différentes mutations aux emplacements suivants :

- 1408 (GTG -> ATG)
- 1522 (Délétion de TTT)
- 1855 (CAT -> AAT)
- 2495 (TTT -> CTT)

Query	1321	CTGAAAGATATTAATTTCAAGATAGAAAGAGGACAGTTGTTGGCGGTTGCTGGATCCACT	1380
Sbjct	1391		1450
Query	1381	GGAGCAGGCAAGACTTCACTTCTAATGATGATTATGGGAGAACTGGAGCCTTCAGAGGGTG.	1440
<b>Sbjct</b>	1451		1510
Query	1441	AAAATTAAGCACAGTGGAAGAATTTCATTCTGTTCTCAGTTTTCCTGGATTATGCCTGGC	1500
Sbjct	1511		1570
Query	1501	ACCATTAAAGAAAATATCATCGGTGTTTCCTATGATGAATATAGATACAGAAGCGTC	1557
<b>Sbjct</b>	1571		1630
Query	1558	ATCAAAGCATGCCAACTAGAAGAGGACATCTCCAAGTTTGCAGAGAAAGACAATATAGTT	1617
Sbjct	1631		1690
Query	1618	CTTGGAGAAGGTGGAATCACACTGAGTGGAGGTCAACGAGCAAGAATTTCTTTAGCAAGA	1677
Sbjct	1691		1750
Query	1678	GCAGTATACAAAGATGCTGATTTGTATTTATTAGACTCTCCTTTTGGATACCTAGATGTT	1737
Sbjct	1751		1810
Query	1738	TTAACAGAAAAAGAAATATTTGAAAGCTGTGTCTGTAAACTGATGGCTAACAAAACTAGG	1797
Sbjct	1811		1870
Query	1798	${\tt ATTTTGGTCACTTCTAAAATGGAACATTTAAAGAAAGCTGACAAAATATTAATTTTGAAT}{\tt C}.$	1857
<b>Sbjct</b>	1871		1930
Query	1858	GAAGGTAGCAGCTATTTTTATGGGACATTTTCAGAACTCCAAAATCTACAGCCAGACTTT	1917
Sbjct	1931		1990
Query	1918	AGCTCAAAACTCATGGGATGTGATTCTTTCGACCAATTTAGTGCAGAAAGAA	1977
Sbjct	1991		2050

#### 4) Lecture d'une entrée Uniprot

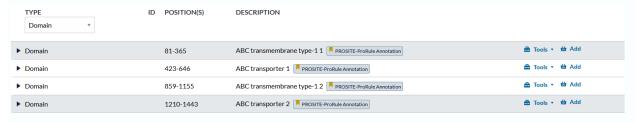
**Fonction**: epithelial ion channel that plays an important role in the regulation of epithelial ion and water transport and fluid homeostasis

Fonction catalytique: ATP + H2O + closed Cl- channel = ADP + phosphate + open Cl- channel.

Mutation: VAR\_000171

**Effet de la mutation**: causes local changes to the surface that mediates interactions between domains; decreases frequency of channel opening in vitro; binds to the cytokeratin-8 and through this binding is primed for the degradation pathway that ends in the proteasome, thus impairing trafficking; impairs maturation and trafficking to the cell membrane; impairs recycling to the cell membrane after endocytosis; abolishes MARCHF2-mediated degradation

**Nombre de domaines :** Il y a 4 domaines : 81-365 ABC transmembrane type 1-1, 423-646 ABC Transporter 1, 859-1155 ABC transmembrane type 1-2, 1210-1443 ABC Transporter 2.

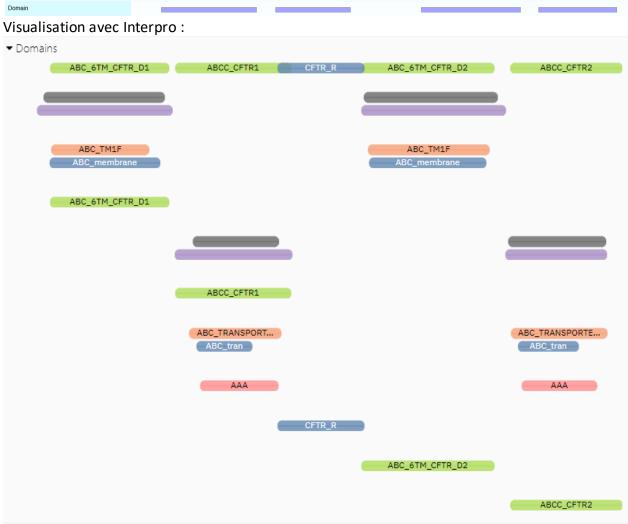


Il y a 2 domaines transmembranaires avec 6 segments chacun.

Transmembrane

Domain

Visualisation avec Interpro:



Il existe aussi des structures résolues expérimentalement pour cette protéine, par différentes Technologies (Xray, EM, NMR, prédiction avec alphafold).

- En Xray, il est difficile d'avoir des structures résolues pour les domaines transmembranaires (et sur une petite partie de la protéines)
- En EM, il est plus facile d'avoir la molécule entière à une résolution correcte (par crystalisation)

On observe les structures tridimensionnelles des domaines suivants :

- ABC transporter 1 (389-678)
- ABC transporter 2 (1210-1443)

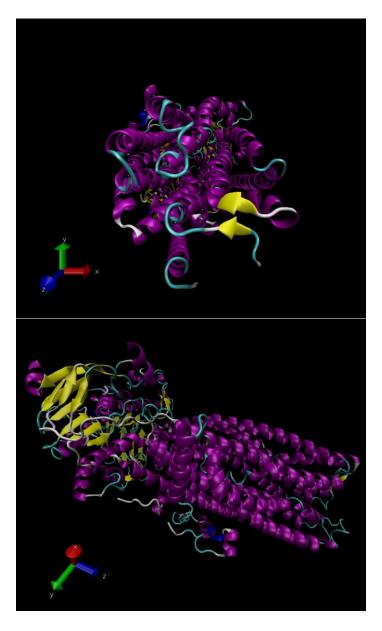
Parmi les différentes molécules, nous avons choisi la molécule **7SVD** qui est la molécule avec la résolution la plus précise et obtenue avec un microscope électronique (par cristallisation de la protéine).

#### On observe les motifs suivants:

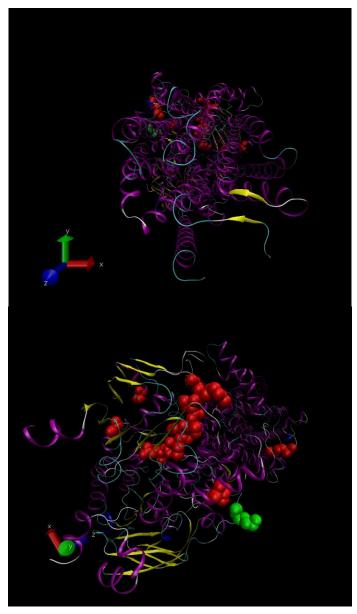
- Des hélices pour les segments des domaines transmembranaires
- Des brins pour les autres domaines (transporteurs, régulateurs)

## 5) Visualisation de la structure de cette protéine avec VMD

lci nous affichons les structures secondaires colorées en fonctions des structures secondaires.



Photos des résidus fixant l'ATP (en rouge) et du résidu muté (en vert) sur la protéine entière.



#### Conclusion

Ce second TP de bioinformatique nous a permis d'approfondir notre compréhension des traitements de séquences et de l'utilisation des bases de données biologiques, appliqués à l'étude de la mucoviscidose et de la protéine CFTR. En traduisant la séquence d'ADN, en identifiant les mutations via BLAST et en analysant les entrées dans GenBank et UniProt, nous avons pu localiser les altérations génétiques associées à cette maladie.

Également, avec le logiciel de visualisation VMD, nous avons pu visualiser la molécule en 3D et mettre en évidence les ATP et le résidu muté.