

# Détection de virus dans l'ADN

Projet de fin du cours Informatique et Programmation

École EEEA – 2022

Jean-Jacques Lévy

4 Janvier 2023

**Avertissement** Soyez clairs, précis et concis.

L'acide désoxyribonucléique (ADN) est une séquence de plusieurs milliards de bases adénine (A), guanine (G), cytosine (C), and thymine (T). Voici un exemple (`ex1`) de bout d'ADN humain :

```
TATTTACCATATCAGATTCACATTCAGTCCTCAGCAAAATGAAGGGCTCCATTTTCACTCTGTTTTTATT
TATTTACCATATCAGATTCACATTCAGTCCTCAGCAAAATGAAGGGCTCCATTTTCACTCTGTTTTTATT
CTCTGTCCTATTTGCCATCTCAGAAGTGCGGAGCAAGGAGTCTGTGAGACTCTGTGGGCTAGAATACATA
CGGACAGTCATCTATATCTGTGCTAGCTCCAGGTGGAGAAGGCATCAGGAGGGGATCCCTCAAGCTCAGC
AAGCTGAGACAGGAAACTCCTTCCAGCTCCCACATAAACGTGAGTTTTCTGAGGAAAATCCAGCGCAAAA
CCTTCCGAAGGTGGATGCCTCAGGGGAAGACCGTCTTTGGGGTGGACAGATGCCCACTGAAGAGCTTTGG
AAGTCAAAGAAGCATTTCAGTGATGTCAAGACAAGATTTACAACTTTGTGTTGCACTGATGGCTGTTCCA
TGACTGATTTGAGTGCTCTTTGCTAAGACAAGAGCAAATACCCAATGGGTGGCAGAGCTTTATCACATGT
TTAATTACAGTGTTTTACTGCCTGGTAGAACACTAATATTGTGTTATTAATAATGATGGCTTTTGGGTAGG
CAAAACTTCTTTTCTAAAAGGTATAGCTGAGCGGTTGAAACCACAGTGATCTCTATTTTCTCCCTTTGCC
TAAAAATGCTATAAACCA
```

En Python, on représentera les bouts d'ADN par des chaînes de caractères 'A', 'T', 'G', 'C'. Par exemple `s = 'TATTTACCATATCAGAT'`. Rappel : une expression Python sur plusieurs lignes s'écrit en finissant chaque ligne (sauf la dernière) par le caractère `\`.

\*\*\*\*

**Question 1** Écrire la fonction `nb_base(c, s)` qui compte le nombre de base `c` dans le bout d'ADN `s`.

**Question 2** Quels sont les nombres de base A, C, T, G dans l'exemple `ex1` d'ADN.

\*\*\*\*

On recherche les nombres de motifs dans des bouts d'ADN. Par exemple, on compte 10 motifs CTA dans l'exemple `ex1` d'ADN.

**Question 3** Écrire la fonction `eq_seq_at(w, i, s)` qui répond vrai si le motif `w` figure en position `i` dans `s`, et faux s'il n'y figure pas.

**Question 4** Écrire la fonction `nb_seq(w, s)` qui compte le nombre d'occurrences du motif `w` dans le bout d'ADN `s`.

**Question 5** Quels sont les nombres des motifs GCA et CATCTCAGA dans l'exemple `ex1` d'ADN.