

Bioinformatik Übung 5

Bericht zum Human T-cell leukemia virus type I

Aufgabe 2)

Erste Hundert Basen des kompletten Genoms vom humanen T-lymphotropischen Virus 1

```
1  ggctcgcatc tctccttcac gcgcccgcgc ccttacctga ggccgccatc cacgccggtt
61  gagtcgcggt ctgccgcctc ccgcctgtgg tgccctctga
```

Quelle: <https://www.ncbi.nlm.nih.gov/nuccore/9626453>

Aufgabe 3)

Erste 30 Aminosäuren der ersten translatierten Region vom humanen T-lymphotropischen Virus 1

5'3' Frame 1

MGQIFSRASPIPRPPRGLAAHHWLNFLQA..

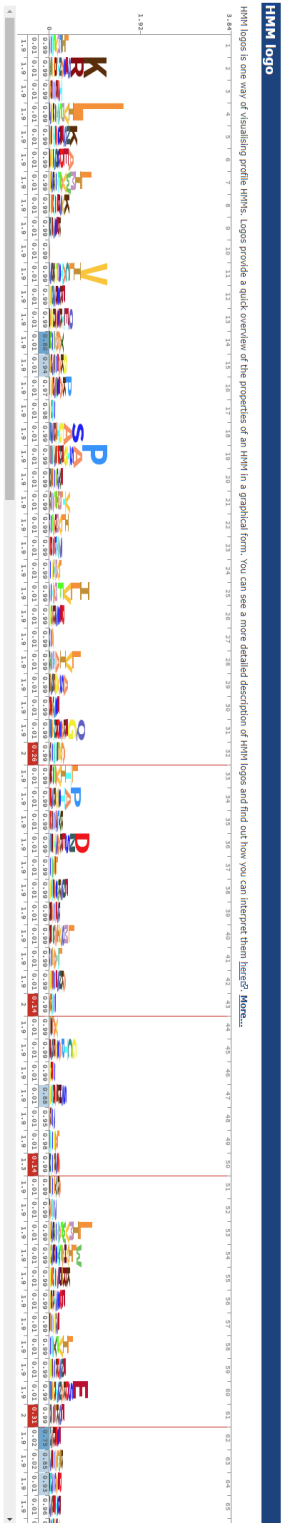
Mit <https://web.expasy.org/cgi-bin/translate/dna2aa.cgi> erstellt

- Gensequenzen sind groß und unübersichtlich, an Hand der Aminosäuresequenz lassen sich die Auswirkungen von Mutationen viel schneller erkennen.
- Die von einem Algorithmus erstellten Frames unterscheiden sich grundlegend voneinander; Erstens in der Leserichtung und dann jeweils nochmal mit zwei Leserasterverschiebungen, da jeweils ein Triplet eine Aminosäure codiert.

Aufgabe 4)

Die folgende Abbildung 1 zeigt das Pfam-HMM Logo vom gag gene protein p24 (core nucleocapsid protein). Erster Treffer der AS-Sequenz der ersten CDS vom humanen T-lymphotropischen Virus 1 über das HMM-Profil-Such-Tool in der Pfam-Datenbank.

Der Domain E-Value dieses Treffers beläuft sich auf 2.4×10^{-65} und ist damit sehr klein, somit liegt eine sehr hohe Übereinstimmung zwischen dem Logo und der Suchsequenz vor



Aufgabe 5)

Gewähltes Genom: Feline foamy Virus (NC_002201)

Erste 100 bp:

```
1  tggatatcaga gccatattta gatcaatcct cggttgactg aacaatgagc tcaagccatg
61 agaatgaagc tgacgcttct caacacatat cttctaacga
```

Erste 30 AS im ersten 5'3' Frame des ersten Gens:

MSSSHENEADASQHISSNEEYIFSNEENENG..

Eine Pfam-HMM-Suche ergab nur einen Treffer, das Virale movement protein (MP, Accession-No. PF01107.17) mit einem Domain E-Value von $4.4e-51$, also hoher Übereinstimmung

