

Genome first 121 pairs:

```
1  tgacaatgac catgagcccc aatatcccc cgggggctta gagcctccca gtgaaaaaca
61  tttccgcgaa acagaagtct gaaaagggtca gggcccagac taaggctctg acgtctcccc
121 ccggagggac agctcagcac cggctcaggc taggcctga cgtgtccccc tgaagacaaa
```

source: <https://www.ncbi.nlm.nih.gov/nuccore/LC183873.1>

Aminosäuren des 1. 5'3' Frames :

-Q-P-APNIPRGLRASQ-KTFPRNRS�KRS GPRLRL-RLPPEGQLSTGSG-ALTCPPEDKS-
AQTSGKPPEPPISSPCLSSRPQALTTTPHLKKLFMARIWLNKLTGVYKSVETVQEGARISPSR
ARRPT-GRHPRRLSRVLPPPACGAS-TASAV-
VSLELRSRPGLCPALPWSLPRLSRFSTLCLTLLAQLCVFVSFSVLRRYRSKVPPLSLSFTTDCR
LGPRPSTGDSVGSEPATAHSIALSRRET-YTVGGSSGIRAPLYSLGN

Vermutungsteil

Aufgabe a:

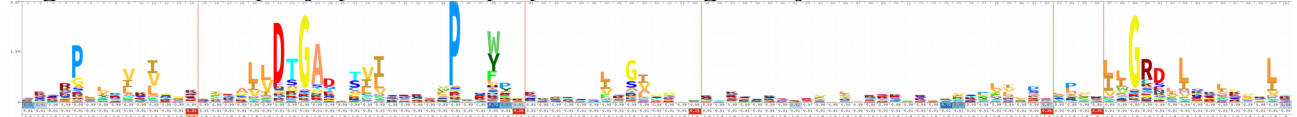
Da wir auf Basis von Buchstaben arbeiten, bedeutet weniger Buchstaben weniger Rechenaufwand und weniger Speicherbedarf. Bei weiterer Überlegung ist das aber nicht unbedingt der Fall, wenn die Buchstaben als Zahlen dargestellt werden wie bei fasterfrequentwords und die Basis bei Aminosäuren damit um einiges höher wäre.

Ein anderer Aspekt ist, dass die Merkmale nach denen wir suchen in der DNA (aus informationstechnischer Sicht) noch komprimiert sind.

Aufgabe b:

Es ist sinnvoll alle Frames zu untersuchen, da man so nichts übersieht und feststellen kann wie sich die Frames unterscheiden und ob Proteine vielleicht sogar in mehreren Auftreten.

Logo: Retroviral aspartyl protease <http://pfam.xfam.org/family/PF00077.19#tabview=tab4>



Weiterhin sollen sie eine kurze Einschätzung abgeben, wie ähnlich das Logo zu ihrer Suchsequenz ist

Ich hab keinen blassen Schimmer, da alle online-logo-erstell-Eingabemasken die ich ergoogeln konnte mir kein LOGO für meine Suchsequenz erstellen wollten.

Malaria Mücken Virus Genom: https://www.ncbi.nlm.nih.gov/nuccore/NC_011317.1

```
1  agtattattc cgtgatacgg atactgtaag atacagatgt aatacatcgt ttctaataga
61  aactgtatct tacagtatcc gtatcacgga ataatacttt ttatatggat tatggactta
121 tatcaaattc ctatatggat cactggaggt ggaaaataag agaaaaacat aagggtggaaa
```

AS-sequenz

SIIP-YGYCKIQM-YIVSNRNCILQYPYHGILFIWIMDL YQIPIWITGGGK-
EKNIRWKLSLFSTHKYLNLFHYHMHVPHYGSLEVENKKKT-GGN-
AYSPHTNTTLLISTTTTWSTPI-GVQKERRIE—

kann kein HMM profil finden. <https://www.youtube.com/watch?v=Fy3rjQGc6lA>