

Übungszettel 5

Benjamin Sae-Chew

2) ersten Basenpaare des Gene 1 des Genoms des „*Human T-lymphotropic virus 1*“

```
ATGGGCCAAATCTTTTCCCGTAGCGCTAGCCCTATTCCGCGGCCGCCCGGGGGCTGGCCGCTCATCACT
GGCTTAACTTCCTCCAGGCGGCATATCGCC
```

Quelle: NCBI

3) ersten Aminosäuren 30 Aminosäuren des translatierten Gene 1 des „*Human T-lymphotropic virus 1*“ des 1. 5'3' Frame

```
MGQIFSRASPIRPPRGLAAHHLNQLQA
```

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

a) Der genetische Code ist degeneriert, da mehrere Basentriplets für die selbe Aminosäure codieren. Der Vergleich der Basen kann dazu führen, dass weniger konservierte Bereiche sichtbar werden, obwohl die Mutationen keine Auswirkung auf Proteinebene hätte (Silent Mutation). Daher ist es besser die Aminosäuresequenz zu nehmen, da diese Silent Mutation keinen Einfluss haben

b) Mit allen drei Leserastern ist es möglich alle verschiedenen Startpunkte bezüglich der Translation zu erfassen.

ACGTATT besitzt 3 Startpunkte welche mit dem Leseraster zu erfassen sind










1. Startcodon ACG
2. Startcodon CGT
3. Startcodon GTA

Zusätzlich ist nicht klar definiert ob die mRNA ein Transkript des gegebenen oder komplementären Strang ist. Daraus resultiert der 6-Reading Frame

4) Ergebnisse der Suche mit Pfam

Pfam Matches

Advanced

| Family | | Clan | Description | Cross-references | Start | End | Domain E-values | | |
|--------|-------------------------|----------------------------|------------------------|--|--|-----|-----------------|---------|---------|
| Id | Accession | | | | | | Ind. | Cond. | |
| > | Gag_p24 | PF00607.19 | CL0148 | gag gene protein p24 (core nucleocapsid protein) |    | 147 | 344 | 1.2e-64 | 2.2e-68 |
| > | Gag_p19 | PF02228.15 | CL0074 | Major core protein p19 |    | 1 | 92 | 7.3e-57 | 1.3e-60 |
| > | zf-CCHC | PF00098.22 | CL0511 | Zinc knuckle |    | 355 | 372 | 8.5e-07 | 1.5e-10 |

Your search took: 0.1 secs

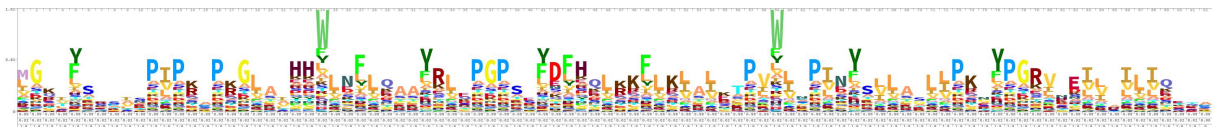
HMM-Logo des „gag gene protein p24 (core nucleocapsid protein)“



Quelle: <http://pfam.xfam.org/family/PF02228.15#tabview=tab4>

Beurteilung: Die Übereinstimmungsregion liegt zwischen Aminosäure 147-344. Einige konservierte Aminosäuren sind identisch mit den Aminosäuren der gegebenen Sequenz an der richtigen Position. Es gibt sehr viele unkonservierte Bereiche in HMM Logo. Die Cysteine im letzten Drittel scheinen stark konserviert zu sein.

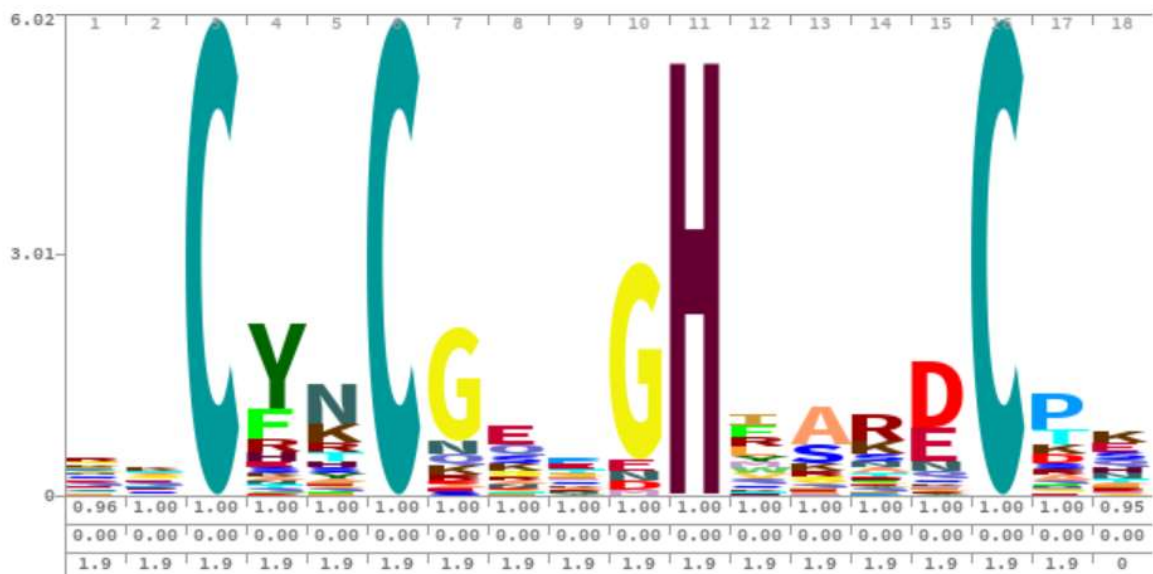
HMM-Logo des „Major core protein p19“



Quelle : <http://pfam.xfam.org/family/PF02228.15#tabview=tab4>

Beurteilung: Die Übereinstimmungsregion liegt zwischen Aminosäure 1-92. Es scheint keine stark konservierte Aminosäure zu geben.

HMM-Logo des „Zinc Knuckle



Quelle: <http://pfam.xfam.org/family/PF00098.22#tabview=tab4>

Das HMM-Logo zeigt die Entropie der einzelnen Positionen bezüglich Aminosäure an.

Beurteilung: Die Übereinstimmung ist von Aminosäure 355 bis 372 und zeigt, dass es einige sehr konservierte Aminosäuren gibt. Die Cysteine sind sehr stark konserviert und weisen eine sehr geringe Entropie auf. Histidin ist ebenfalls stark konserviert. Allgemein gibt es relativ viele leicht konservierte Aminosäuren. Die ersten beiden Aminosäuren scheinen sehr variabel, dementsprechend ist die Entropie dieser beiden Positionen hoch. Der Vergleich mit dem HMM-Logo zeigt, dass die stark konservierten mit den Aminosäuren der gegebenen Sequenz übereinstimmen.

5) ersten 100 Basenpaare des Gene 1 des Genoms „*Halorubrum pleomorphic virus 1*“

```
ATGTCACACAACGATACCCCCGCAAATAACACTAACGGCATAGCCGACGTTTACGCTGACGAGAATCCAT
GTGAAATAGCCGGGGAAGATAACCCCCCAC
```

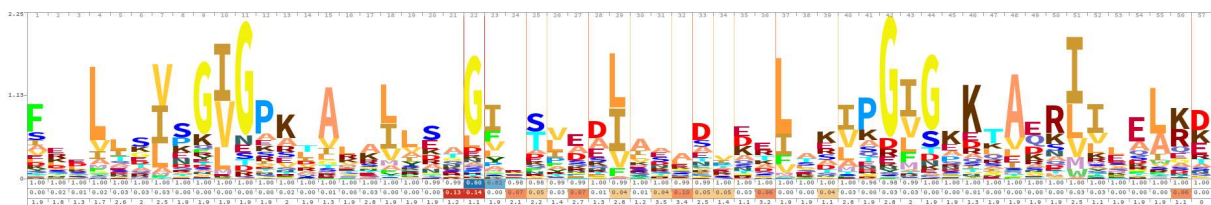
Quelle: https://www.ncbi.nlm.nih.gov/nucore/NC_012558.1?from=1&to=1503&report=fasta

Ersten 30 Aminosäuren des translatierten Gene 1 des „*Halorubrum pleomorphic virus 1*“ des 1. 5'3' Frame

```
MSHNDTPANNTNGIADVYADENPCEIAGED
```

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

HMM-Logo des Gene 1 des „*Halorubrum pleomorphic virus 1*“



Quelle : <http://pfam.xfam.org/family/PF14520.5#tabview=tab4>

Beurteilung: Die Übereinstimmung ist von Aminosäure 408-460. Es gibt sehr viele stark konservierte Regionen, beispielsweise im ersten und im letzten Drittel. Manche Positionen werden von 2-3 Aminosäuren dominiert und sind innerhalb dieser Position konserviert. Einige konservierte Aminosäuren stimmen mit der gegebenen Sequenz überein.