

### Opdracht 3

Als eerste is de opdracht een HMM te maken, dit doe ik via de bin servers en ik gebruik hetzelfde bestand als opdracht 2 een alignment genaamd alignment.txt dit bevat een fasta bestand met TNF sequenties van mensen.

deze heb ik geupload naar de bin servers en de command {hmmbuild hmmoutput1.hmm alignment.txt} gebruikt:

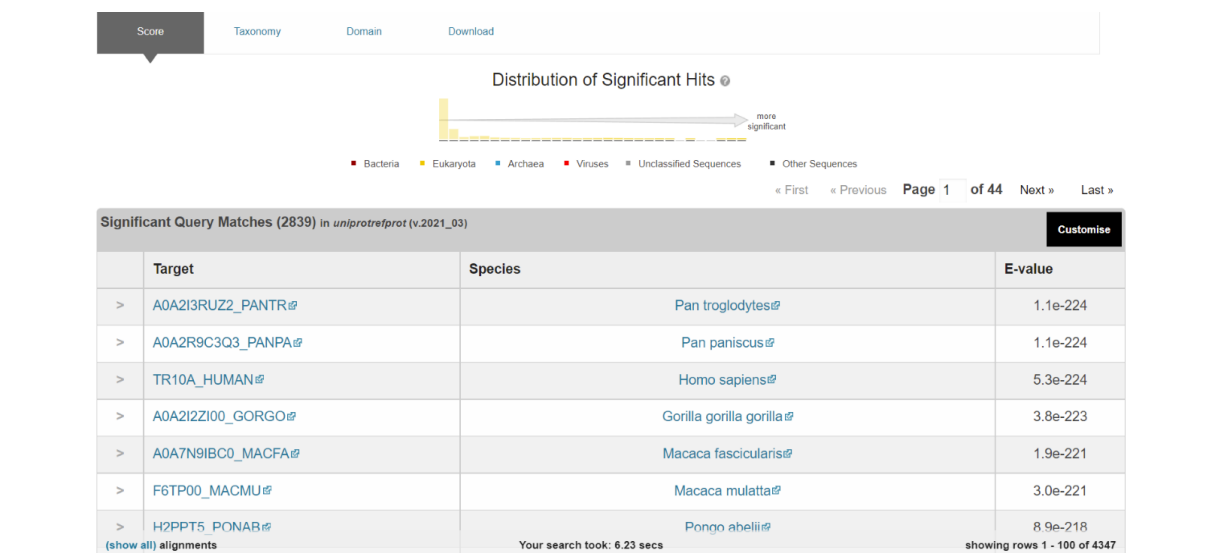
```
# CPU time: 0.24u 0.00s 00:00:00.24 Elapsed: 00:00:00.24
jotolhuis@bin203:~/Documents/Thema09/Bio_informatica03$ hmmbuild output2.hmm alignment.txt
# hmmbuild :: profile HMM construction from multiple sequence alignments
# HMMER 3.3.2 (Nov 2020); http://hmmerr.org/
# Copyright (C) 2020 Howard Hughes Medical Institute.
# Freely distributed under the BSD open source license.
# -----
# input alignment file:          alignment.txt
# output HMM file:              output2.hmm
# -----
# idx name                      nseq  alen  mlen  eff_nseq  re/pos  description
#-----
1      alignment                7    502   410    1.00   0.591
# CPU time: 0.24u 0.00s 00:00:00.24 Elapsed: 00:00:00.24
jotolhuis@bin203:~/Documents/Thema09/Bio_informatica03$
```

deze heb ik vervolgens geupload naar de website:

<https://www.ebi.ac.uk/Tools/hmmer/search/hmmsearch>

Om te kijken voor vergelijkbare sequenties.

Dit gaf een paar resultaten:



Het gevonden model van meest voorkomende letters per positie ziet er zo uit, dit vertelt ook iets over de conservering van bepaalde gebieden.

[illegible]

Door op deze sequenties te zoeken op in de Pfam database, vinden we de sequenties van een mens. Wat bewijst dat de hits betrouwbaar zijn.

## Sequence search results

[Show](#) the detailed description of this results page.

We found **4** Pfam-A matches to your search sequence (**3** significant and **1** insignificant)



**Show** the search options and sequence that you submitted.

[Return](#) to the search form to look for Pfam domains on a new sequence.

### Significant Pfam-A Matches

[Show](#) or [hide](#) all alignments.

| Family                  | Description                    | Entry type | Clan                   | Envelope |     | Alignment |     | HMM      |           | HMM length | Bit score | E-value | Predicted active sites | Show/hide alignment  |
|-------------------------|--------------------------------|------------|------------------------|----------|-----|-----------|-----|----------|-----------|------------|-----------|---------|------------------------|----------------------|
|                         |                                |            |                        | Start    | End | Start     | End | From     | To        |            |           |         |                        |                      |
| <a href="#">Death</a>   | Death domain                   | Domain     | <a href="#">CL0041</a> | 368      | 449 | 369       | 447 | <b>4</b> | <b>84</b> | 86         | 64.7      | 6.7e-18 | n/a                    | <a href="#">Show</a> |
| <a href="#">TNFR_c6</a> | TNFR/NGFR cysteine-rich region | Domain     | <a href="#">CL0607</a> | 148      | 188 | 148       | 188 | 1        | 39        | 39         | 33.5      | 4.4e-08 | n/a                    | <a href="#">Show</a> |
| <a href="#">TNFR_c6</a> | TNFR/NGFR cysteine-rich region | Domain     | <a href="#">CL0607</a> | 190      | 229 | 190       | 229 | 1        | 39        | 39         | 34.3      | 2.5e-08 | n/a                    | <a href="#">Show</a> |

### Insignificant Pfam-A Matches

Show or hide all alignments.

| Family | Description  | Entry type | Clan | Envelope |     | Alignment |     | HMM      |           | HMM length | Bit score | E-value | Predicted active sites | Show/hide alignment  |
|--------|--------------|------------|------|----------|-----|-----------|-----|----------|-----------|------------|-----------|---------|------------------------|----------------------|
|        |              |            |      | Start    | End | Start     | End | From     | To        |            |           |         |                        |                      |
| TMIF   | TMIF protein | Family     | n/a  | 230      | 267 | 236       | 265 | <b>7</b> | <b>33</b> | 85         | 8.0       | 2.6     | n/a                    | <a href="#">Show</a> |



Pfam is part of the ELIXIR infrastructure

Rfam is an Elixir service [Read more](#)

Comments or questions on the site? Send a mail to [pfam-help@ebi.ac.uk](mailto:pfam-help@ebi.ac.uk).

De beste hits zijn inderdaad van het TNF domein. Met de volgende alignment:

### Significant Pfam-A Matches

[Show](#) or [hide](#) all alignments.

| Family         | Description   | Entry type | Clan                   | Envelope |     | Alignment |     | HMM      |           | HMM length | Bit score | E-value | Predicted active sites | Show/hide alignment  |
|----------------|---|------------|------------------------|----------|-----|-----------|-----|----------|-----------|------------|-----------|---------|------------------------|----------------------|
|                |   |            |                        | Start    | End | Start     | End | From     | To        |            |           |         |                        |                      |
| <b>Death</b>   | Death domain  | Domain     | <a href="#">CL0041</a> | 368      | 449 | 369       | 447 | <b>4</b> | <b>84</b> | 86         | 64.7      | 6.7e-18 | n/a                    | <a href="#">Hide</a> |
| #HMM           | <a href="#">Iarllldhpdklgrkrlarklglsseerdrlekensksqsyallqlwedgenatvtgtltsarkrgldraekles</a> |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #MATCH         | +++++ ++ +W+l+r+l+l+ eId ++ + +y+a+l +w+++g:n+a++tLl+a++ ++ ek++                            |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #PP            | 57899999999*****g7  |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #SEQ           | <a href="#">DQFAHVHTVPSNDQLPRLDLTKNEIDVRAGTAGPGDALVAHKMKVNKTRGRASHTLLDALEREPEERHAKECTG</a>  |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| <b>TNFR_c6</b> | TNFR/NGFR cysteine-rich region  | Domain     | <a href="#">CL0607</a> | 148      | 188 | 148       | 188 | 1        | 39        | 39         | 33.5      | 4.4e-08 | n/a                    | <a href="#">Hide</a> |
| #HMM           | <a href="#">Cpeg.tytdense.eClpctr.cpggav.lrpctptsdtvc</a>                                   |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #MATCH         | C eg Ytl+ ++ clpct C+++++ +pctltst+ C   |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #PP            | 999966*****gg9  |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #SEQ           | <a href="#">ETE-VQITASHMLTACLTACKSDEFEFSPTTTSHTAG</a>                                       |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| <b>TNFR_c6</b> | TNFR/NGFR cysteine-rich region  | Domain     | <a href="#">CL0607</a> | 190      | 229 | 190       | 229 | 1        | 39        | 39         | 34.3      | 2.5e-08 | n/a                    | <a href="#">Hide</a> |
| #HMM           | <a href="#">Cpegty.tdense.eClpctr.Cpggav.lrpctptsdtvc</a>                                   |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #MATCH         | c+gt+ d++ ++ C++C+ CP gVH ++Ctp sd +C   |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #PP            | *****55555666*****g9  |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |
| #SEQ           | <a href="#">XPGLTHINDNSAENLRKKSTDCPRGMVKYCDTPMSDIE</a>                                      |            |                        |          |     |           |     |          |           |            |           |         |                        |                      |

De clan is te vinden op <http://pfam.xfam.org/clan/CL0041>. De family is te vinden op <http://pfam.xfam.org/family/PF00531.24>