

# assignment 2

Mats Slik

## HMMbuild:

command:= hmmbuild H2bc\_hmm.mo MSA1.fasta

## Hmmscan

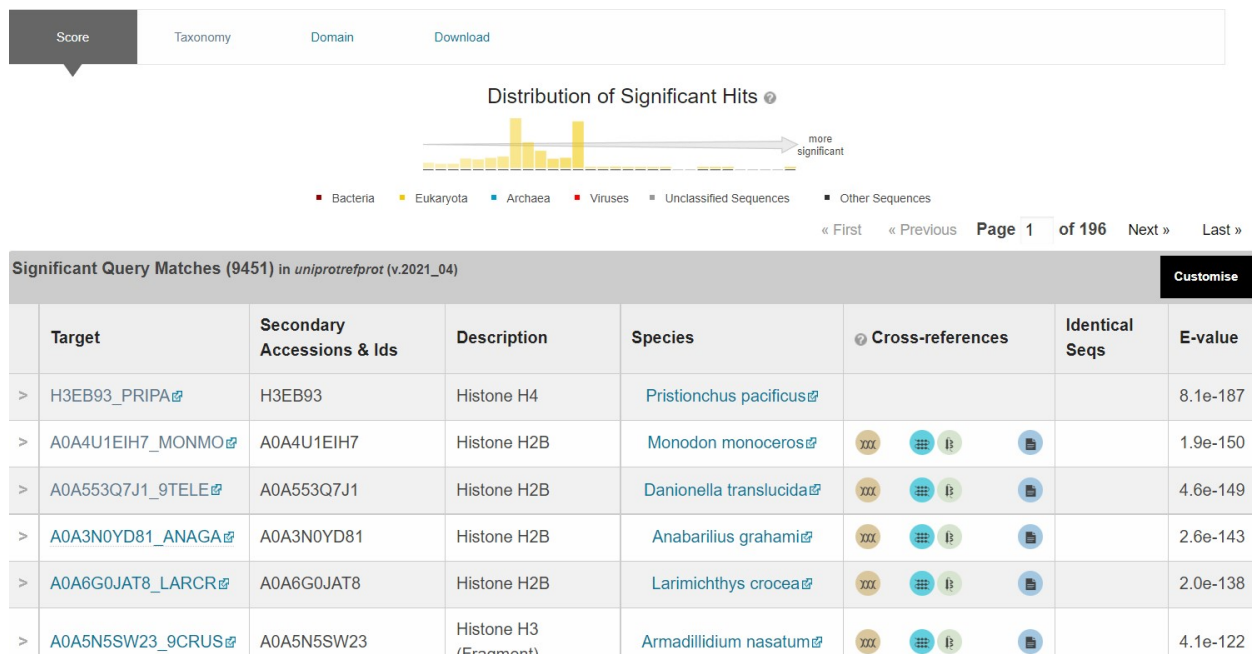
command:= hmmbuild H2bc\_hmm.mo MSA1.fasta

output created: H2bc\_hmm.mo

subsequently this file was used for the next step.

## Hmmsearch

a search was tried on the Bioinf.nl linux network bus found no connection to a refseq DB, so this was used.  
<https://www.ebi.ac.uk/Tools/hmmer/search/hmmsearch>



the Results id : 66987292-5D3C-11ED-A5D5-AC9B9F870EE0.1 results link

this gave us multiple results I selected the top three hits : 1.A0A4U1EIH7 2.A0A553Q7J1 3.A0A3N0YD81

these files were combined with the existing homologene2.fasta multi fasta file that contained the original homolog family of proteins. the new combination file is called "homologene\_new.fasta"

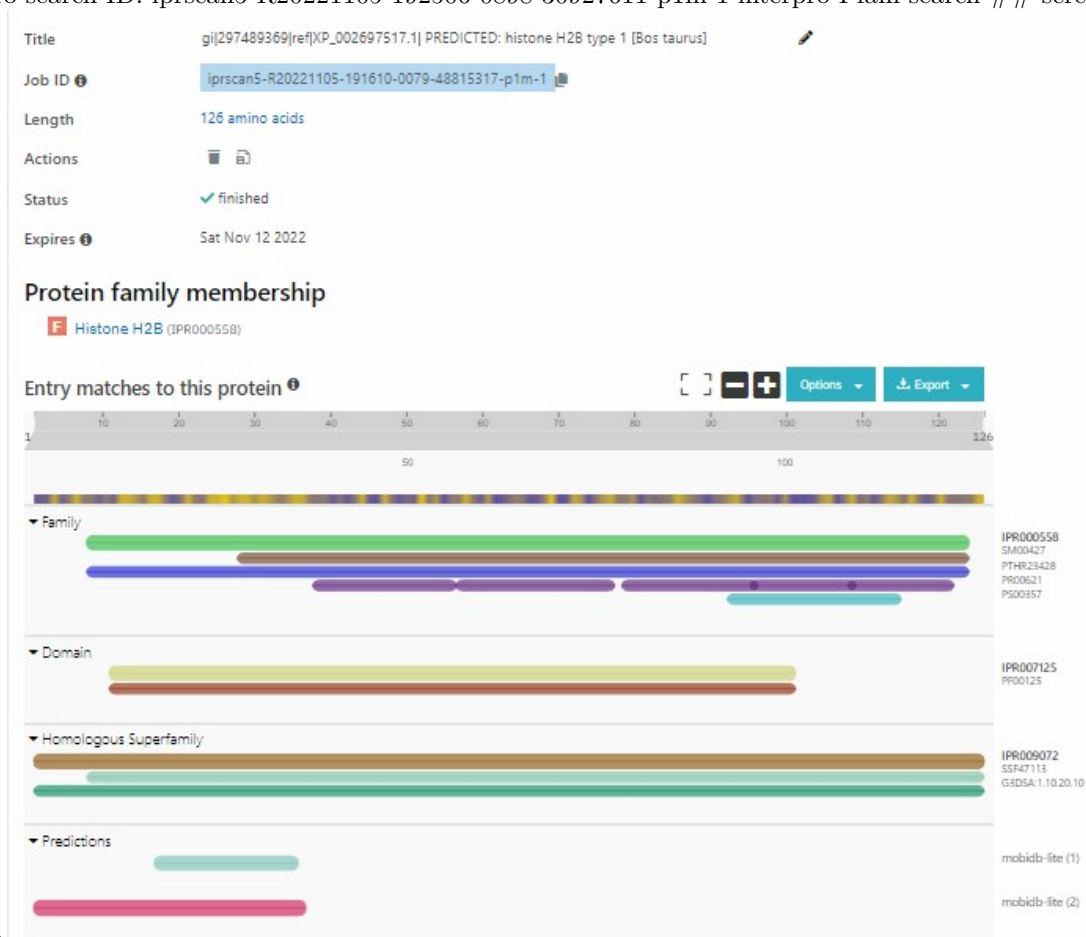
a new alignment was carried out with clustal W using this fasta, the results are stored in: 1."homologene\_new1.fast" 2."homologene\_new.aln"

## HMMbuild number 2 :

command:= hmmbuild H2bc\_hmm\_new.mo homologene\_new1.fasta



## Pfam search

interpro search ID: iprscan5-R20221105-192506-0898-30927611-p1m-1 interpro Pfam search ## screenshots






PFAM


Title tr|A0A3N0YD81|A0A3N0YD81\_ANAGA Histone H2B OS=Anabarrilius grahami OX=495550 GN=DPX16\_1745 PE=3 SV=1

Job ID  iprscan5-R20221105-192506-0898-30927611-p1m-2 

Length 244 amino acids

Actions  






Status  finished

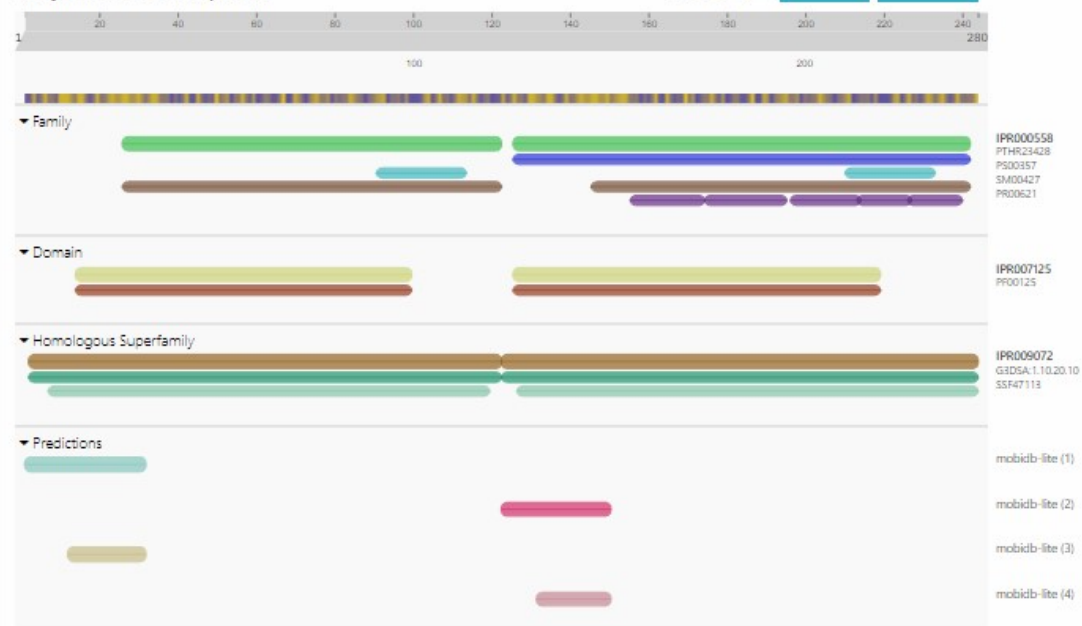
Expires  Sat Nov 12 2022

## Protein family membership

 Histone H2B (IPR000558)

### Entry matches to this protein

   Options  Export 



Title [gij109079443|ref|XP\\_001110898.1| histone H2B type 1-K \[Macaca mulatta\]](#) 

Job ID [iprsan5-R20221105-192506-0898-30927611-p1m-3](#) 

Length 126 amino acids

Actions  

Status  finished

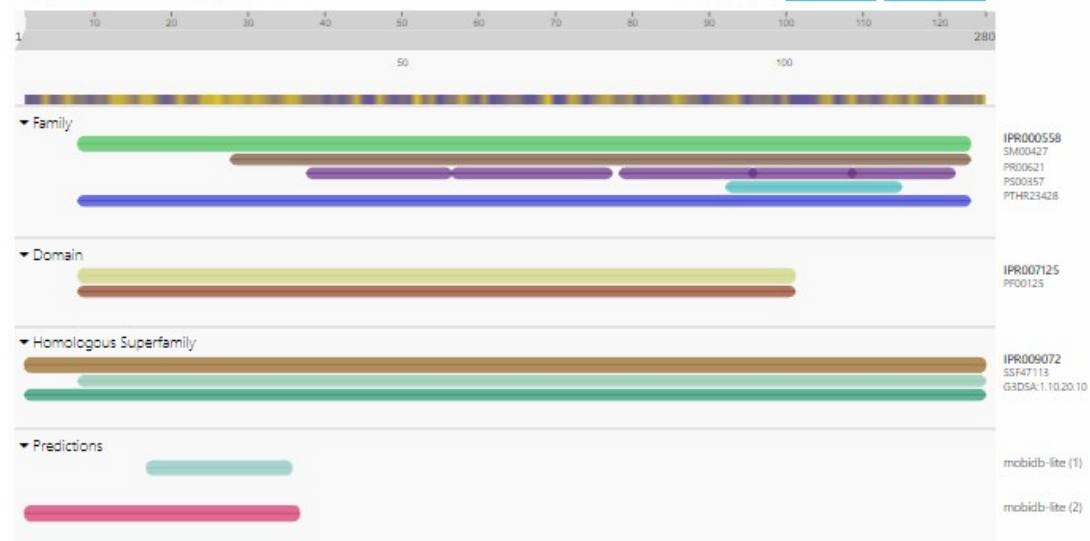
Expires  Sat Nov 12 2022

## Protein family membership



 [Histone H2B](#) (IPR000558)

### Entry matches to this protein


   Options  Export 





Title trjA0A553Q7J1|A0A553Q7J1\_9TELE Histone H2B OS=Danionella translucida OX=623744 GN=DNTS\_011926 PE=3 SV=1

Job ID  iprscan5-R20221105-192506-0898-30927611-p1m-4 

Length 256 amino acids

Actions  

Status  finished

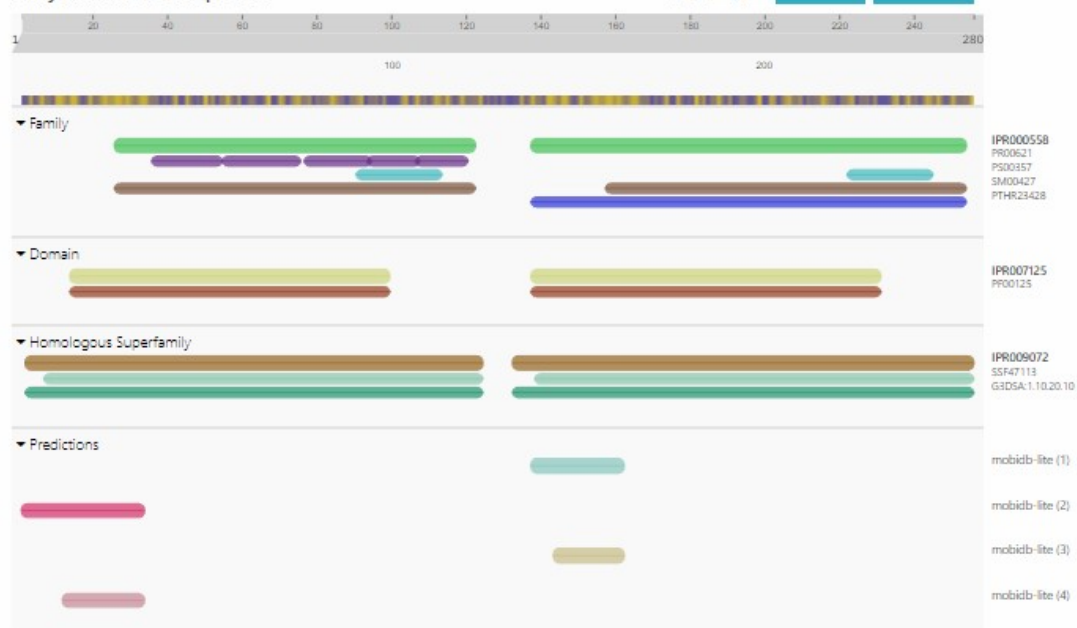
Expires  Sat Nov 12 2022

## Protein family membership

 Histone H2B (IPR000558)

### Entry matches to this protein



   Options  Export 





Title tr[A0A4U1EIH7]A0A4U1EIH7\_MONMO Histone H2B OS=Monodon monoceros OX=40151 GN=EI555\_001635 PE=3 SV=

Job ID [iprscan5-R20221105-192506-0898-30927611-p1m-5](#)

Length 280 amino acids

Actions  

Status  finished

Expires  Sat Nov 12 2022

## Protein family membership

 **Histone H2B** (IPR000558)

### Entry matches to this protein



[Home](#) / [Result](#) / [InterProScan](#) / [Iprscan5-R20221105-192506-0898-30927611-P1m-5](#) / [Entry](#) / [InterPro](#)

## InterProScan Search Result

Overview **Entries 3** Sequence

This protein matches these entries:

1 - 3 of 3 entries matching InterPro  			
ACCESSION	NAME	SOURCE DATABASE	MATCHES
IPR000558	Histone_H2B	InterPro	
IPR009072	Histone-fold	InterPro	
IPR007125	Histone_H2A/H2B/H3	InterPro	

Show  results Previous **1** Next