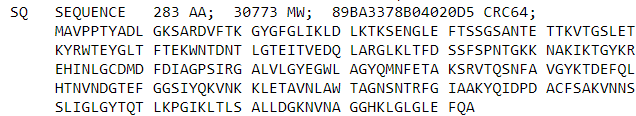
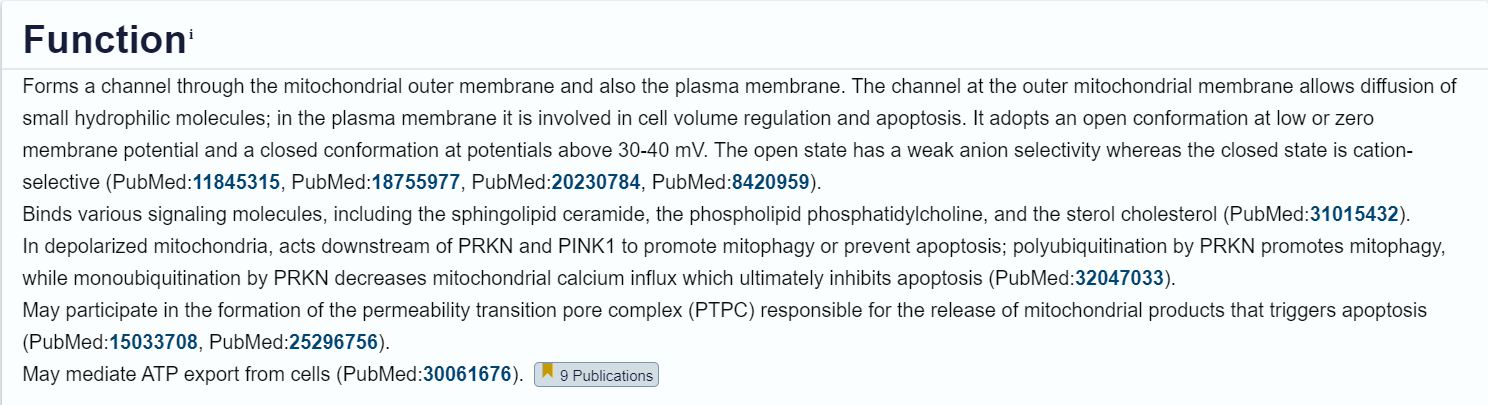
Shree

## Q1)

### A1) Sequence of human mitochondrial beta barallel membrane protein VDAC



### A2) Function:



### B) Number of transmembrane sequence present in the protein = **19**



### Q2) Obtain the sequences of “transcription factors” with 50% sequence identity in FASTA format. List the count of sequences and count of clusters. <http://www.uniprot.org/uniprot/>

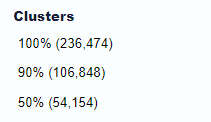
### 

Count of Sequences: **4352**

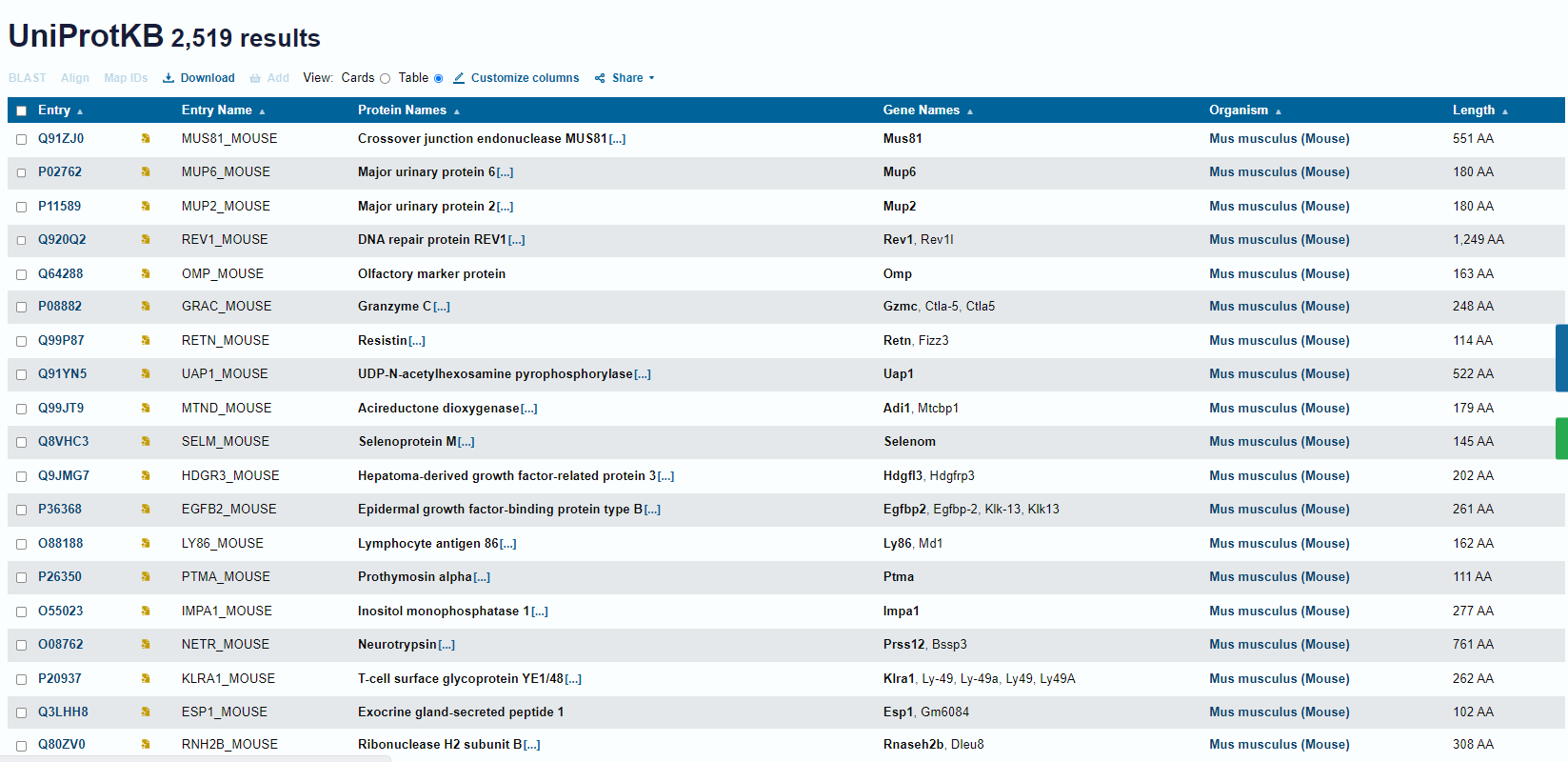
Count of Clusters: **235**

The sequences’ FASTA file is in the folder named ‘Q2\_ transcription\_factors’

### Q3) How many protein sequences from Homo sapiens are obtained at identity cutoff of 100%, 90% and 50% sequence identity?



### Q4) . In UniProt, how many mouse (Mus musculus) protein sequences are manually annotated? And how many of these manually annotated protein sequences are associated with PDB (3D structures)?

1. Manually annotated sequences are **17807**, under the Swiss-Prot Database.
2. There are **2519 sequences** 

### Q5)

### Q6)