**HBx STRAIN SAMPLES**

>sp|P03165|X\_HBVD3 Protein X OS=Hepatitis B virus genotype D subtype ayw (isolate France/Tiollais/1979) OX=490133 GN=X PE=1 SV=2

MAARLCCQLDPARDVLCLRPVGAESRGRPFSGSLGTLSSPSPSAVPTDHGAHLSLRGLPV

CAFSSAGPCALRFTSARRMETTVNAHQILPKVLHKRTLGLSAMSTTDLEAYFKDCLFKDW

EELGEEIRLKVFVLGGCRHKLVCAPAPCNFFTSA

>sp|Q69027|X\_HBVCJ Protein X OS=Hepatitis B virus genotype C subtype ayr (isolate Human/Japan/Okamoto/-) OX=928302 GN=X PE=1 SV=1

MAARLCCQLDPARDVLCLRPVGAESRGRPVSGPFGPLPSPSSSAVPADHGAHLSLRGLPV

CAFSSAGPCALRFTSARSMETTVNAHQVLPKVLHKRTLGLSAMSTTDLEAYFKDCLFKDW

EELGEEIRLKVFVLGGCRHKLVCSPAPCNFFPSA

**Host Virus**

>sp|Q7Z434|MAVS\_HUMAN Mitochondrial antiviral-signaling protein OS=Homo sapiens OX=9606 GN=MAVS PE=1 SV=2

MPFAEDKTYKYICRNFSNFCNVDVVEILPYLPCLTARDQDRLRATCTLSGNRDTLWHLFN

TLQRRPGWVEYFIAALRGCELVDLADEVASVYQSYQPRTSDRPPDPLEPPSLPAERPGPP

TPAAAHSIPYNSCREKEPSYPMPVQETQAPESPGENSEQALQTLSPRAIPRNPDGGPLES

SSDLAALSPLTSSGHQEQDTELGSTHTAGATSSLTPSRGPVSPSVSFQPLARSTPRASRL

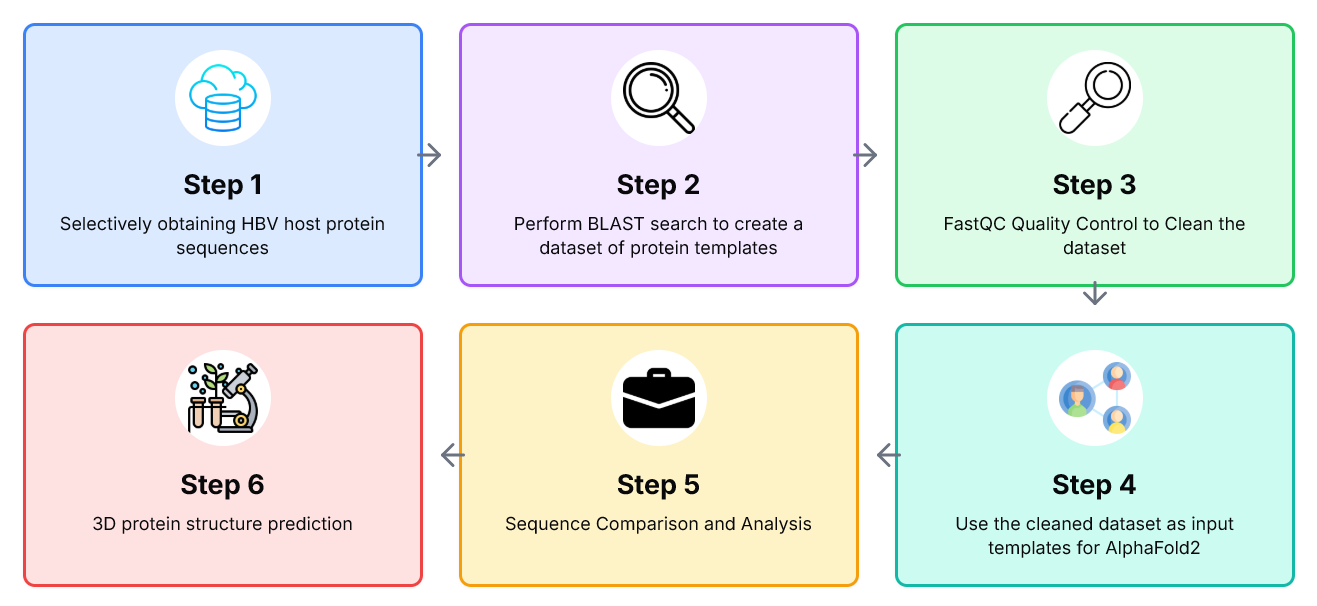
PGPTGSVVSTGTSFSSSSPGLASAGAAEGKQGAESDQAEPIICSSGAEAPANSLPSKVPT

TLMPVNTVALKVPANPASVSTVPSKLPTSSKPPGAVPSNALTNPAPSKLPINSTRAGMVP

SKVPTSMVLTKVSASTVPTDGSSRNEETPAAPTPAGATGGSSAWLDSSSENRGLGSELSK

PGVLASQVDSPFSGCFEDLAISASTSLGMGPCHGPEENEYKSEGTFGIHVAENPSIQLLE

GNPGPPADPDGGPRPQADRKFQEREVPCHRPSPGALWLQVAVTGVLVVTLLVVLYRRRLH

****

**I am not a bioinformatics engineer but I am making a simple research report for my assignment so using the information I’ve provided below,help me edit the case study overview and objectives in chapter 2 without making a lot of changes and also use the information below to create for me chapter three by writing it following the heading needed in the previous result**

**Use The first refered to papers and anymore relevant Papers, new used papers should be cited and the new papers should be updated to the most previous list of reference**

**i want to use Bioinformatics for hepatitits B virus case study in a way that i check whether ones protein sequence has any HBV mutation career main or sidechains**

**how- i obtain the person’s protein sequence and then use the alphafold2 to match the sequence against our created dataset which has templates that we obtained from BLAST, and then visualise the result to check for match chains that can confirm the existence of persistent HBsAg gene that maybe carried from mother to child**

**2. For conceptual model, step 1,we obtained three HBV host proteinsmin fasta sequence format one from French origin, Genotype D and its of length 154, second is of Japanese origin, genotype D and is of length 154 and the last one is a Ugandan sequence of length 540**

**Step 2: we created blast on each of the three sequences to create our dataset of protein templates**

**step 3: Talk about trimming, FastQC we cleaned the dataset to remove duplicates and too long and too short sequences**

**Side Chains and MainChains are derived from sequencing**

**step 4:Talk about alignment and matching plus sequencing, we used our custom dataset from blast as input templates for the alphafold 2 model in colab to create a predicted structure of the HBV side chains and mainchains**

**step 5: we now obtain any human's protein sequence and now check for presence of sidechains and mainchains and intereprete confidence results for indication of possible career mutations**

**Tools -Protein Databank, NCBI GenBank as a database for blast to obayin templates**

**Uniprot to obtain the first two sequences and Uganda Genome resource for the local host sequence**

**Blast tool for Blast operation to create a dataset**

**Colab\_Alphafold for prediction and creation of main and side chains**

**Chimera x for visualization of the 3D protein structures**