**Practical 2 – JALVIEW exercises**

**Name: Jialin Kang**

**Email: jkang58@jhu.edu**

The P53 gene (and its protein product TP53) is a transcription factor and a master regulator of transcription**. In this exercise, you will explore the evolutionary history of the TP53 protein using Jalview.**

To get started, read about P53 in Wikipedia

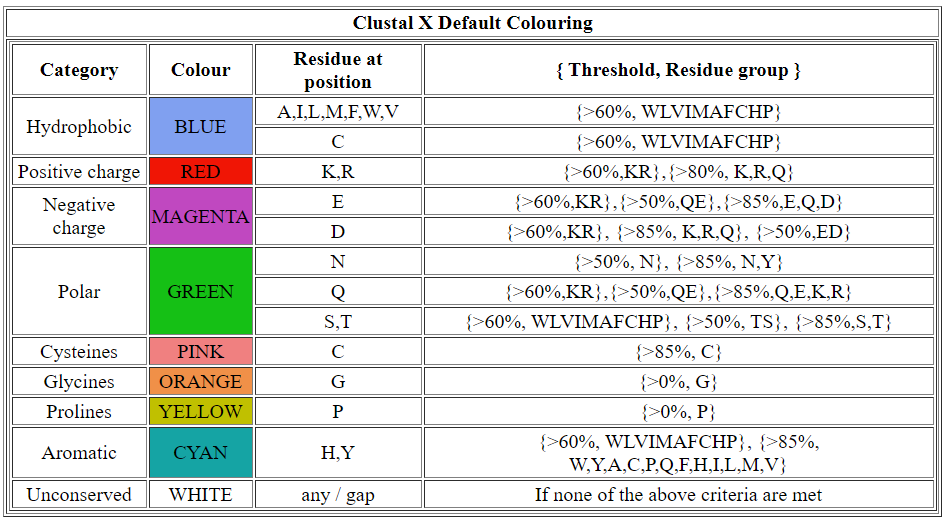
<https://en.wikipedia.org/wiki/P53>

1. In Jalview load the PFAM seed alignment of the P53 family. You will use the Fetch Sequences command. The accession ID of P53 is PF00870. You will find additional information needed for these exercises on the UniProtKB web page for P53\_human <https://www.uniprot.org/uniprot/P04637>

2. Move the P53\_Human protein to the first position in the alignment. Click to select, then use the up arrow key.

3. Color the alignment with “ClustalX” coloring. What amino acid residue properties do the colors represent?

<http://www.jalview.org/help/html/colourSchemes/clustal.html>



4. To see the amino acid residue number of a residue, mouse over. The text is displayed in the bar at the bottom of the window. What is the residue number of the S in P53\_human in the first column of the alignment.

99

5. do you see examples of columns where a property of the amino aicd residues (such as hydrophobicity) is completely conserved, but the column contains more than a single amino acid residue. Do you see columns in which a single amino acid residue is completely conserved?

Yes, followings are examples:

More than a single amino acid residue(hydrophobicity completely conserved):

The 11th column of the alignment (109 residue number of F in P53\_human)

The 64th column of the alignment (133 residue number of M in P53\_human)

The 65th column of the alignment (109 residue number of F in P53\_human)

The 66th column of the alignment (109 residue number of C in P53\_human)

Single amino acid residue (positive charge):

The 115th column of the alignment (176 residue number of C in P53\_human)

The 235th column of the alignment (238 residue number of C in P53\_human)

6. Locate the columns for the amino acid residues where each of the four zinc-binding sites occurs. What are the amino acids in those positions? Are they completely conserved, almost completely conserved? If there are columns that are almost completely conserved, what species harbor the amino acid that is different from human?

The position of the four zinc-binding sites is

176(C:cysteine) positive charge completely conserved,

179(H:histidne) aromatic almost completely conserved, W5JCA5\_ANODA, Q7QBX6\_ANOGA and A0A084W6W1\_ANOSI are different from human.

238(C:cysteine) positive charge completely conserved,

242(C:cysteine) positive charge completely conserved.

7. Select five missense mutations that are associated with Li-Fraumeni syndrome. Find the position of each one in the alignment and look at the column. Do you see the mutated amino acid in any of the species in that column? If yes, which species?

Position:132 K->E (63rd column) do not see the mutated amino acid in other species

Position:133 M->T (64th column) do not see the mutated amino acid in other species

Position:138 A->P (69th column) do not see the mutated amino acid in other species

Position:141 C->Y (72nd column) do not see the mutated amino acid in other species

Position:151 P->T (82nd column) N6UG92\_DENPD, Q171M1\_AEDAE, A0A0N4UDB8\_DRAME

8. Select all the sequences in the alignment. Then use the Calculate command to calculate “Tree or PCA”.

a. Select Tree/Neighbor Joining and the BLOSUM62 scoring matrix. Use the output tree to answer these questions:

1. What is the closest related P53 protein to human in another species?

P53\_DANRE

1. Why is P73\_Human more distant from P53\_Human than P53\_Danre?

P73\_Human and P53\_Human are in human, but with different original. P53\_Human and P53\_Danre begins at the same gene. So P73\_Human is more distant from P53\_Human than P53\_Danre.

3) Click on the tree and move the red line to identify clades of related P53 proteins

4) What species harbors the P53-like protein that is the most distant from human P53 on the tree?

ADA183RBZ3\_9TREM

b. Select PCA and the BLOSUM62 scoring matrix. Use the output plot to answer these questions.

1) View the first three principal components.

2) Find human P53 in the PCA space (HINT: the color should match the P53\_HUMAN highlight color in the alignment window).

3) Based on PCA, how many P53 clades are represented in this alignment?

Maybe 6 clades.

Save your Jalview project.

What to turn in: Write the answers to all questions in this document. Email to [fcbb2homework@gmail.com](mailto:fcbb2homework@gmail.com) by 11:59PM Feb. 14, 2020