**Day 10: PFAM & KEGG**

*Complete the activity individually while working together in the group. Submit it Canvas individually.*

**PFAM**

*Which domains/families are found in the aromatic amino acid hydroxylases from human (I know, you may be tired of these by now, but let’s use them one final time)?*

1. Fill out the table with the information for the PFAM hits for human PAH (P00439). Clicking on the domain/family takes you to its entry.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Family (code)** | **Entry type (Domain or family)** | **# of architectures** | **# of seq in seed** | **# of seq in full** |
| **1** |  |  |  |  |  |
| **2** |  |  |  |  |  |

1. In which species are you likely find domain/family 1 and 2, respectively?
2. What is the main function of the first family/domain according to the Pfam or Interpro description found under the Summary tab?
3. For the first domain/family, which amino acid is most common in position 28 in the HMM logo? What are shared characteristics of the 3 most common amino acids at this site?
4. What is the difference between a PFAM domain and a PFAM family? See [Pfam: Help (xfam.org)](http://pfam.xfam.org/help)

**KEGG**

**Go to KEGG:** [KEGG: Kyoto Encyclopedia of Genes and Genomes](https://www.genome.jp/kegg/)

Click on **KEGG PATHWAY**

Click on **Environmental Information Processing**.

Under Signal Transduction, find the link to **TGF-beta signaling pathway**

The upper left corner should look like this:

****

***Subcellular location***

Focus on the four locations shown in red (this is an example from a different pathway than the one you are working with).



1. Where is ERK found?
2. Where is BMP found?
3. Where is the complex of TGFβ RI and TGFβ RII found?
4. Where is c-Myc found?

***Regulation***

Use the Help tab to decode what the different symbols and lines mean.



1. How is BMP regulated by Noggin?
2. How does BMP regulate BMPRII?
3. How does Noggin indirectly regulate BMPRII?
4. How is SMAD1/5/8 activated by BMP RI?
5. Does SMAD2/3 inhibit SMAD6/7 or vice versa?

***What’s behind the button?***

Click on it an find familiar information such as GeneID, accession number, sequence, GO term, and more links to other pathways that this protein is participating in etc. Explore, click on the boxes, BLAST if needed.

1. There are several SMADs in the TGFβ pathway. Assuming they are related, are they isoforms or paralogs?
2. What proteins in the refseq\_protein database are identical in sequence and have with 100% query cover to NodalRI and NodalRII, respectively? (name and accession)
3. Are BMPRI, BMPRII, TGFβRI, TGFβRII, ActivinRI, ActivinRII, NodalRI, and NodalRII in human homologs (use BLAST to answer this question)?