**Purpose: Explore tools available to search for genes, examine their function and find related genes.**

**There are 3 sections to this exercise.**

**Search for NCBI on the web to get to the homepage**

**(http://www.ncbi.nlm.nih.gov/)**

1. Entrez

* + Search for the “MSH2 human” gene in the Gene database (select this from pull down menu)
  + Select the first hit by clicking on MSH2
  + In the Table of Contents on the right of the page, click on Reference Sequences
  + To get the protein sequence, click on the NP\_000242.1
  + Click on FASTA to get the sequence in the correct format for BLAST
  + From this view, you can click on the Run BLAST on the right hands side of the page

2. BLAST

* You can also copy the FASTA formatted sequence and go directly to the BLAST search page
  + Scroll down and click on the words **“Algorithm parameters”** to see some options available to you including changing the number of sequences to show, the scoring matrix, etc. Try this first with the default parameters and then rerun with changing some parameters
  + Run BLAST by clicking on the big blue **BLAST button**
  + Scroll down to look at the results, in particular the alignments. Look at the information about “Identities”.
  + When sequences are very similar, it’s hard to see the differences.

Click on “**Formatting options**” at the top of the screen and select “Pairwise with dots for identities” from the Alignment View dropdown menu. Then click on the **Reformat** button. This will show you the difference between your input (Query) sequence and the database (Sbjct) sequence. What differences do you see between humans and chimps (Pan paniscus)

**Search for Ensembl on the web to get to the homepage**

**(http://www.ensembl.org/)**

3) Ensembl

* Enter MSH2 in the text box in the upper left hand side of the home page
* Click on the words “**Variation table”** to see the types and number of variations seen for this gene. You can click on “**Show**” to see details of a particular type of variation.
* On the left hand column of the page, under Comparative Genomics you can click on “**Orthologs**” to see in what other species the gene is found. (Notice that there are no known Parologs of this gene.) If time permits, there is a lot of information to explore on this page about possible homologs. For example, you can click on “Alignment (protein)” or “Alignment (cDNA)” to look at the alignment of sequences between humans and the other species.
* Explore additional information included in the list of related information under “Gene-based displays” on the left hand side of the page.