

Lab exercise 1

BLAST provides convenient ways to seek for DNA and protein sequences of genes (see <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>). Features such as blastn, blastp, and blastx are particularly useful. It also provides links to search for homologs among species. The practice here is:

- I. Use a DNA sequence from *Drosophila* (see below) to find:
 - a. What is the gene encoded by this DNA?
 - b. What is the amino acid sequence of the protein encoded by this gene?
 - c. What is the amino acid sequence of the *C. elegans* ortholog of this gene?
 - d. What is/are the amino acid sequence of the mouse ortholog(s) of this gene?
 - e. What is/are the amino acid sequence of the human ortholog(s) of this gene?
 - f. By the information provided in your search result and those from associated links, please answer the following questions
 - i. Describe the structure of this protein
 - ii. What is the known function of this protein?
- II. Explain your strategy of working out the questions above in a simple scheme and tell what softwares/databases you use

The unknown sequence from *Drosophila*:

```
TACAATTTACATATAAGTACGAAAAACTTTCTGGGCCGGAACAGATTCCG
AAAACAAAGGAGCAAACACCCACATATTTGAAGAGGATTAATCATGGACATC
TCAAAGGCACCAAATCCGCGAAAACCTGGAGCTGTGTGCGCAAATACTTCTTTG
CTGGCTTTGCATTTCTGCCCTTTGTGTGGGCCATTAACGTTTGCTGGTTTTT
CACGGAGGCCTTCCATAAGCCACCATTTTTCGGAGCAGAGCCAAATAAAGAG
ATATGTTATATACTCTGCAGTGGGGACTCTATTCTGGCTGATAGTACTAACT
GCCTGGATAATAATATTCCAGACAAATCGCACAGCCTGGGGCGCCACAGCG
GACTATATGAGCTTCATCATAACCCCTAGGCAGTGCATAGACATAACTAGATT
AATTCGTTAGCAACTAATGATATTAATAAAGACTTCATTCCTAAACAAAATTA
ACGTTTATTATTAATCCA
```