

## Bioinformatics: BLAST

---

We will be looking at different versions of BLAST in order to see how various versions can yield different results.

### Proteins

Go to the NCBI site (<http://www.ncbi.nlm.nih.gov/BLAST>).

1. Run the blastp program on the Test protein sequence on the class web site ([public.clunet.edu/~revie/bioinformatics.htm](http://public.clunet.edu/~revie/bioinformatics.htm)). Do not close the window! Read pages 225-231 to understand all the blast output.
2. Go to the EMBnet site (<http://www.ch.embnet.org/software/bBLAST.html>).
3. Select the “SwissProt” database for proteins.
4. Paste in the Test protein sequence.
5. Click on “Run blast”.
6. How do these results compare with the step 2 results?
7. Repeat steps 3 through 7, except change the database to “nonredundant”. How do the results compare with the previous blast? Why are the three blasts different?
8. Repeat steps 1 through 8, but use the tblastn program. How do the tblastn results differ from the blastp?
9. Run steps 3 through 7, except change the matrix to blosum80.
10. Run steps 3 through 7, except change the matrix to pam30.
11. Which of the three seems better or worse? Why?

### DNA

Read pages 231-235.

12. Go to the NCBI site (<http://www.ncbi.nlm.nih.gov/BLAST>).
13. Click on blastn (nucleotide-nucleotide blast).
14. Run the blast on the new Test DNA sequence (TestDNA2.txt) on the class web site. Do not close the window!
15. Repeat steps 14 through 16 using the blastx program.
16. Repeat steps 14 through 16 using the tblastx program.
17. How do the results compare?

### Iterative blasts

Go to the NCBI site (<http://www.ncbi.nlm.nih.gov/BLAST>).

18. Go through pages 240-245 in the book to see how iterative blasts work.

### Homework assignment

Use the HHV-6 protein you earlier found information about (U1 through U100). For that protein, analyze it to determine:

- a. If it is related to other proteins. How did you check?
- b. 2D and 3D structure(s). What site(s) did you use and what did you find?
- c. If it is a membrane protein. What site(s) did you use and what did you find?
- d. If it has a signal sequence, and, if so, where it is and where it is likely to be cut. What evidence indicated this?
- e. Modification sites and the likelihood of these being “real” sites.
- f. Use the iterative blast to see if the protein is in a protein family. If so, what are the functions of the proteins in that family?