

NCBI Prelab: **See the handout on our website.**

Name \_\_\_\_\_ GSI & Sect # \_\_\_\_\_ Station # \_\_\_\_\_

Welcome to the National Center for Biotechnology Information (NCBI). This week's prelab is designed to help you feel comfortable with some of the bioinformatics tools you will use during lab. Once you are familiar with the basics, repeating this procedure for any gene or disease becomes is simple.

Before you begin the pre-lab, read the introduction, procedure, and websites referenced; having a sense of the big picture will help orient you while you are working on the pre-lab.

## Procedure

1. **Find the human SOD1 gene on the Entrez Gene database.** Use the search bar at the top of the NCBI homepage (<http://www.ncbi.nlm.nih.gov/>), selecting "Gene" from the pull-down menu, to find the gene. The first line of each hit will include the species from which the gene was sequenced. Make sure you are looking at the [*Homo sapiens*] entry! {see Figure 3}

*What is the Official Full Name of the SOD1 gene?*

*What is the identification number, or the accession number, of the SOD1 genome (the "nucleotide link" listed in black under "Genomic regions, transcripts, and products")?*

*What is the full location of the SOD1 gene (listed under "Genomic context")?*

2. **Find the OMIM entry.** Follow the OMIM link in the right-hand column of the page, under the "Links" heading. {see far right of Figure 3 for an example} Click on the link to "Amyotrophic lateral sclerosis," the disease associated with mutations in the SOD1 gene.

Read the text, description, and inheritance portions of the OMIM entry; skim the rest of the page to see the other resources OMIM can offer. If you are unfamiliar with ALS, use Google or Wikipedia to learn basic details.

*True or False: All cases of familial ALS follow an autosomal dominant pattern of inheritance.*

*Give another name for ALS.*

3. **Go to the GenBank entry.** From Entrez Gene, click the mRNA line under “Genomic regions, transcripts, and products” to pull up a sub-menu. Select GenBank under “mRNA Links” to open the GenBank page. {see Figure 6a} Hint: Most word processing programs have user-friendly word counting features.

*What is the accession number for the mRNA GenBank entry?*

*What is the PubMed ID for the research published by Parge, H.E., Hallewell, R.A. and Tainer, J.A.?*

*What two ions does SOD1 bind, and what radical does it neutralize?*

*How many nucleotides are in the mRNA?*

*How many amino acids are in the protein?*

4. **Do some translation.** Go to the ExPASy Translate tool (<http://ca.expasy.org/tools/dna.html>). {see Figure 7a} Copy and paste the nucleotide sequence from the GenBank entry for SOD1 into the box and hit “Translate sequence.” If you’ve selected “Compact” as your Output format, your results will look like Figure 7b.

*Why are six different results listed?*

*Which result is the correct protein (based on the GenBank entry)?*

5. **Mutate.** Go back to the ExPASy Translate tool and re-paste the SOD1 nucleotide sequence. This time, before you submit the sequence, delete nucleotide 181 (nucleotide 181 is a “G”). Hit “Translate sequence” and find the new protein (it should start the same way the original did).

*How many amino acids are in the mutated protein?*

*What kind(s) of mutation did you induce?*

6. **BLAST.** Enter the short sequence below into the Blast query box of a nucleotide search. The BLAST front page can be found at <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>, but you will need to open the nucleotide BLAST page to get a query box, as in the procedures for the lab. For “database,” under “choose search set,” select “Others (nr, etc).” This allows you to search for all organisms that have sequences available on NCBI. {see Figure 9b}

“cagcatgttg agccgggagc tgtgcccac”

*What is the first hit?*

*What color are the top hits in the colored distribution graphic? {see Figure 9c for an example}*

*Use your sequencing data from # 6 on your GMB I worksheet and BLAST that sequence?  
What is your first hit?*