
Homework (Put your answers, in order, on another sheet of paper). 2 points each.

1. What is the *Arabidopsis* version of your gene? Give me both a locus designation (e.g. At1g66340 for *ETR1*) and a refseq accession number for a protein (e.g. NP_176808 for *ETR1*). If there are paralogs (see questions 4 and 5), you may want to give me the refseq accession number of each, but I also need you to **pick one** to use for questions 2 and 3.
2. How similar is this *Arabidopsis* protein to its human homolog? Your answer should include the length of each protein and the length and percent identity for each pair involving the *Arabidopsis* protein (based on blast2 alignments **without** filtering).
3. How similar is this *Arabidopsis* protein to its yeast (*S. cerevisiae*) homolog? Your answer should include the length of each protein and the length and percent identity for each pair involving the *Arabidopsis* protein (based on blast2 alignments **without** filtering).
4. Are there paralogs of your gene in *Arabidopsis*? How many? (Your answer can be based on a simple neighbor-joining approximation, *i.e.* How many *Arabidopsis* proteins more similar to each other than any is to the yeast protein)?
5. Are there more paralogs in *Arabidopsis* or in the mouse (once again, count those paralogs more similar to each other than to your yeast gene)?
6. Go to the Salk T-DNA insertion site (The "T-DNAexpress database;" you can get there from the course page or from TAIR or by navigating within <http://signal.salk.edu/>). How many T-DNA insertions are there in your gene (as before, if you have multiple genes, use the one you picked for question 3)? Give me the name (something like SALK_136291) and location (intron, coding exon, or noncoding exon [specify 5' UTR or 3'UTR]) of one insertion. Pick the one you would prefer to use for a reverse genetic analysis of your gene. You will need to figure out how to read the data (please read the FAQ page).

Once again, I encourage you to find the page for your gene on each of the *Arabidopsis* sites (TAIR, MIPS and TIGR) and explore a bit.

7. You obtain a SALK line T-DNA insertion line. You grow up a single plant, which (unbeknownst to you) is heterozygous for each of two distinct T-DNA insertions, both of which confer kanamycin resistance. If these two insertions are unlinked, what proportion of the progeny seedlings would be expected to show kanamycin resistance.
8. What if the two insertions are linked, but 40 cM. apart? (neglect interference)