

**These questions are an ungraded exercise (and were in 2007) These answers are available online. These are being used in 2008 as a guide to your study.** I encourage you to explore these sites with others, following my tutorial handout or using the help pages.

1. Jon Dinman published a paper in one of the Trends journals in 2002. Who was the first author?
2. What **disease** is associated with OMIM entry 602421 (give the OMIM entry)?
3. What **gene** is associated with many "typical" cases of the disease described under OMIM entry 300322?
4. What **protein** is the product of this gene (your answer to 3, the gene associated with OMIM entry 300322)? Answer with the **refseq accession number** for the **protein** (see [www.ncbi.nlm.nih.gov/RefSeq/key.html#accessions](http://www.ncbi.nlm.nih.gov/RefSeq/key.html#accessions) for explanation of refseq).
5. What mouse protein is most closely related to your answer to question 4? Once again, give the **refseq** accession number and a name for the mouse protein.
6. What is the extent of amino acid **identity** between the human and mouse proteins (your answers to 4 and 5) as determined by blast ("blast 2 sequences")? **Do not filter.** Report the number of identical amino acids / length of match.
7. What is the Ensemble Gene ID number for the human FRAP1 protein (this is the HGNC Symbol ID; the refseq accession number is NP\_004949)? Your answer should be something like ENSG0000012345
8. What **chromosome** in the **dog** genome contains the homolog of FRAP1? (hint: the UCSC browser is very useful for questions 8 and 9, but you can answer this at Ensembl or at NCBI).
9. You have been assigned a gene using the yeast (*Saccharomyces cerevisiae*) refseq accession number (on WebCT). What is the best human match (based on blastp score)? Report the gene name (HGNC Symbol ID), refseq accession number and Ensemble gene ID.

Homework questions: (10 points each). **These questions will be graded; do them alone.**

- 1) Where is the following sequence from (what organism?) Does it encode a protein? If so, what is the likely function of that protein?

```
>MysterySequence1
CCAACCTTTCATGGTCTACCTCGCTCATAGTCCACCAGGATCTTGTGGTCATCGATCGT
CATCCGATGGGCTTCATGGTAGGCATCGCGTGCGGATGATTCGGATCGGAAGGTGATAAA
GCCATAGCCTTCAGACAGACCGGT
```

- 2) When in evolution did the LXR gene arise? Different functions are attributed to the LXR gene, so be specific about your answer (you might say something like "if you consider [this feature], then the gene is shared by members of [this group], but genes with [that feature] are found throughout [that group].). Give support for your argument (e.g. cite accession numbers of proteins that you think are likely to share function, and explain why). Your answer will involve reading the literature **and** sequence analysis. You will have to form an informed opinion about the relationship of function to sequence for this protein.