

-----  
**These questions are an ungraded exercise (and were in 2007) These answers are available online.** 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?

Jon Dinman published a paper in Trends in Biochemical Sciences in 2002. The first author is Harger. ([PMID 12217519](#))

2. What **disease** is associated with OMIM entry 602421 (give the OMIM entry)?

The disease associated with OMIM entry 602421 is Cystic Fibrosis.  
([OMIM 602421](#))

3. What **gene** is associated with many "typical" cases of the disease described under OMIM entry 300322?

The disease associated with OMIM 300322 is Lesch-Nyhan syndrome, which is often caused by mutations in the gene HPRT1 ([OMIM 308000](#)).

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).

The protein is hypoxanthine phosphoribosyltransferase 1, [NP\\_000185](#)

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.

The protein is (also) hypoxanthine phosphoribosyltransferase 1, [NP\\_038584](#)

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.

The protein align throughout their length; 211/218 = 96% identical.

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

The Ensemble gene ID for FRAP1 is [ENSG00000198793](#).

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).

The Ensemble gene ID for dog FRAP1 is [ENSCAFG00000016648](#).

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.

The Ensemble gene ID for dog FRAP1 is [ENSCAFG00000016648](#).

---

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
CCAACCTTTCATGGTCCTACCTCGCTCATAGTCCACCAGGATCTTGTGGTCATCGATCGT
CATCCGATGGGCTTCATGGTAGGCATCGCGTGCGGATGATTCCGGATCGGAAGGTGATAAA
GCCATAGCCTTCAGACAGACCGGT
```

This is from *Rhysopus oryzae*. You had to search wgs (this sequence is not in the nr database)

If you got the species but say nothing about the protein, you get 4 points.

If you thought that this encodes a U1 snRNP protein, you get 8 points. If you realize that it encodes a U11 snRNP protein, you got all 10 points. Some of you didn't find the exact hit (and so did not realize that this was from *Rhysopus oryzae*) but used blastx or otherwise found the match to U11/U12 35K. These answers typically got 6 points.

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.

Nuclear hormone receptors are limited to animals. The "liver X receptor" is most similar to the ecdysone receptor of insects.