2004 BI 188 Midterm Paper

Due: Monday, May 10 before class
Extra 3points if you turn it in by 5pm on Friday, May 7
The paper will be graded out of 100 points

Decide which genetic disorder you'll be working on and tell Tracy (tracyt@caltech.edu) by Monday, April 26.

Assignment:

Select a human genetic disorder, by which we mean a disease or heritable predisposition to a disease condition. It can be a single gene trait, or a polygenic one. In the case of a polygenic disorder, you need only do the genomic analysis for one of the involved genes. Also, you can and should feature that selected gene with greatest detail in describing the evidence that it is part of causation and reviewing the underlying genetic evidence for the claim. Some disorders are not polygenic traits, but you will find that more than one gene can cause or predispose different individuals to the disease (BRCA1 and BRCA2 for heritable predisposition to breast cancer, for example). In a case like this, you would introduce both genes as causing the predisposition, and then select one to focus on in detail for the rest of the paper.

I. A. What is the phenotype of the disorder?

B What are the genetics of the disorder in humans and what is the evidence?

C. The genetic epidemiology of the disorder.

II. Give a thorough but succinct account of the evidence that this gene is causal (or partly causal) for the genetic disorder, including highlights of the most important evidence. Say why this particular evidence is powerful and convincing (if it is). Do you think the assignment of this gene to the disorder could be incorrect? Be critical in your account.

III. A. Go to Genbank and/or use both the human genome browsers to find the locus in human and mouse genomes. See if you can find additional vertebrate orthologs (check recent release of chicken, using blast – it is not yet annotated with genes by name).

B. Find out (or identify for yourself) if your gene has suspected orthologous genes in other vertebrate species (which ones?); invertebrate model organisms? Yeast? How sure are you that a candidate ortholog is indeed orthologous and why? Does your gene have candidate paralogous genes in man? What does the literature on orthologs of your gene from other model organisms say, if anything, about the function and evolution of function of your gene? You might want to go to Wormbase and Flybase as a complement to information in the literature and in the human browser annotations.

IV. Set up a Family Relations analysis that covers ~30kb of your gene. If it is small enough, this should be the entire coding region plus as much 5' and 3' flanking sequence as you can get within the 30kb. The minimum comparison would be mouse/human 2-way (inputs will have to be repeat masked – Tracy can provide more info on this). Examine the comparison at various thresholds of similarity and give your main conclusions about conserved coding and noncoding domains. We expect that for most genes you will be able to do a second two-way or a three-way – you can select between these. A nice study would be mouse/human/fugu or mouse/human/chicken. Sequence availability may mean that mouse/human/rat is the only three way you can do, but because we expect the relative closeness of rat and mouse to give you only a little bang for the comparative buck, this is the least desirable three-way.

V. A. Conclude with a brief discussion of what you think is most important to find out next about this disorder and the gene(s) that cause it, and discuss why.

B. Outline how you would approach this future research, assuming that lab money and manpower are not your limitations.

VI. Bibliography. Primary literature references from the journals are desirable when they are really the source: however when you are taking information from a review, text or website that – in turn – cites the primary source, you should cite the review (unless you really go see what the primary research paper says). Over-documenting where the credit comes from for an idea or fact is far better than under-documenting, but we do not want or need inflated bibliographies. If you use a web site, please be sure to credit the site.