CALIFORNIA INSTITUTE OF TECHNOLOGY

DIVISION OF BIOLOGY

Bi188 – Human Genetics and Genomics

2004 FINAL EXAMINATION

Open book but closed internet
(the class web site is allowed, but not the links e.g. OMIM)

No collaboration at all

YOU HAVE **3 HOURS** TO COMPLETE THE EXAM

ANSWER ANY FIVE OF THE SEVEN QUESTIONS
EACH PROBLEM OF EQUAL VALUE

You can answer one extra question for extra credit
Be sure to clearly mark which question you're answering for extra credit.

DUE BY 5:00PM Thursday, June 3, 2004

The exam can be turned in electronically to Tracy (tracyt@caltech.edu) or in the "Bi 188" box in the hall outside Kerckhoff 128. We encourage typewritten submissions for legibility.

1. Shortly after the first dozen colonists reach Mars, Congress cuts off funding for the project. The settlers decline the offer to return to Earth. The population grows rapidly and a few generations later they start seeing some rather strange colonists. What advantages and disadvantages might there be in attempting to genetically map these strange phenotypes in such a population? Eventually, a new mission reaches the main Mars base from Earth and soon some of the original colonists become ill. They learn that the members of the new mission were chosen because they had not become ill during a major epidemic on Earth. Settlers at isolated bases on Mars, descended from the original colonists, need to return to the main base for supplies. Using DNA polymorphism data obtained from the new arrivals and colonists at the main base, how might you then identify individuals at the isolated bases who may be resistant to the new disease?

2. A. In your midterm projects, when studying your gene of interest, you were encouraged to use chicken as a third genome in comparative sequence analysis (human and mouse being the first two). When looking at an individual gene, what benefits are there to using a third genome? Why might using the chicken be more valuable than using rat? For what evolutionary class of features might cow or dog be a better choice than chicken?

B. If you joined a lab and were given the task of understanding the regulation of gene X, how might you use comparative genomics to help you find the answer? How would the results of your analysis guide the experiments that you would do in the lab?

3. Simple measures of DNA sequences like base composition can be useful predictors for a variety of biological processes. Describe three different areas where base composition has been used in this way.

4. Instead of sequencing the genome of an interesting mouse species, a sample of moose DNA was used by mistake. Given the substantial investment, an extensive analysis of the sequence data is being done. You have been assigned to analyze the transposon sequences present in the genome. How would you identify transposons in the data? What aspects of transposon biology would you investigate? What types of comparisons with other mammalian transposons might be of interest?

5. Given the relatively small size of bacterial genomes, it is quite common for the complete genomes of two or three related species or strains from the same species to be sequenced. What additional information (relative to a single genome) can be derived from multiple closely related genome sequences? How does this differ from the types of information one can obtain by comparing more distant species?

6. A. Upon molecular cloning of the p16/Ink4A tumor suppressor gene, a graduate of this institute was quoted in the popular press as saying that the pathway to cancer therapy via gene therapy seems remarkably straightforward. The gene predisposes people to melanoma. It is almost a decade later, and such a treatment has not been forthcoming. Why should this be - in principle- straightforward? What are biological obstacles to such a treatment being developed?

B. What are mechanistic sources of LOH and what is the importance of LOH in cancer genetics?

C. Retinoblastoma is the paradigmatic tumor suppressor gene and it is contributory to both sporadic and inherited forms of the tumor. How do you discriminate sporadic versus heritable forms of this disease clinically? How would you distinguish heritable versus sporadic forms at the DNA level?

7. A. Consider a hypothetical gene X that you are working on because your SURF advisor has noticed that a large domain of its protein coding sequence is highly similar to p53. While you are browsing its genomic sequence, you notice that it has 25 serial repeats of the sequence GAC. What unusual mode of inheritance would this lead you to suspect, and what would be diagnostic of it at the DNA level in affected individuals?

B. Assume that X, like p53, has tumor suppressor activity. Draw a hypothetical kindred that illustrates the pattern of inheritance you would look for and the basic phenotype you would seek. We seek the essential qualities of such a kindred - not a hugely elaborated version. Highlight the genotype:phenotype properties that are salient.

C. Give two molecular mechanisms by which this could account for activation of the tumor suppressor in a developing tumor.

D. If X operates normally as a transcription factor, how would you use gene microarrays to learn what its target genes are and how they are affected in tumors in which the suppressor is not normally active?