

Fall 2005 Genomics Exam #2
Genomic Variation and Microarrays

There is no time limit on this test, though I don't want you to spend too much time on this. There are three pages for this test, including this cover sheet. You are not allowed discuss the test with anyone until all exams are turned in at 11:30 am on Friday November 4. **EXAMS ARE DUE AT CLASS TIME ON FRIDAY NOVEMBER 4.** You may use a calculator, a computer, but only the web pages that appear in this exam. You are NOT allowed to explore the internet to take this exam. This is a new policy and is required if I am to shorten the length of the exams. You may take it in as many blocks of time as you need to. NOTE: I leave town on November 4 and I want to take the tests with me to grade. Submit your paper and electronic versions before 11:30 am so I can take them with me along with paper versions.

The **answers to the questions must be typed in a Word file and emailed to me as an attachment.** Be sure to backup your test answers just in case. You will need to capture screen images as a part of your answers which you may do without seeking permission since your test answers will not be in the public domain. Print this test but make sure the screen shots are big enough to be read easily. Remember to explain your thoughts in your own words and use screen shots to support your answers. **Screen shots without your words are worth very few points.**

THIS IS A CLOSED BOOK EXAM TO HELP SHORTEN THE TEST.

-3 pts if you do not follow this direction.

Please do not write or type your name on any page other than this cover page.

Staple all your pages (INCLUDING THE TEST PAGES) together when finished with the exam.

Name (please print):

Write out the full pledge and sign:

How long did this exam take you to complete (excluding typing)?

20 pts.

1) Genome Variations question:

“The demonstration of association between common genetic variants and chronic human diseases such as obesity could have profound implications for the prediction, prevention, and treatment of these conditions. Unequivocal proof of such an association, however, requires independent replication of initial positive findings. Recently, three (rs2236418, rs928197, and rs992990) single nucleotide polymorphisms (SNPs) within glutamate decarboxylase 2 (GAD2) were found to be associated with class III obesity (body mass index $>40 \text{ kg/m}^2$). The association was observed among 188 families (612 individuals) segregating the condition, and a case-control study of 575 cases and 646 lean controls....We found no evidence for a relationship between the three GAD2 SNPs and obesity.”

You may use this web site only: <http://www.ncbi.nlm.nih.gov/SNP/> and the pages your searching of this site directly leads you to.

- Give me the DNA sequences for these 3 mutations. Provide the sequences in a readable screen shot. Copy and pasting the sequence is not acceptable.
- What is the frequency for each SNP? Use a screen shot to show me your data.
- Describe any differences of frequency between populations for each of these SNPs? Support your answer with data from this web site.
- What evidence is there to validate these 3 SNPs? Use text to support your answer.
- Do any of these 3 SNPs alter the protein primary structure? Support your answer with data from this web site.

Now go to <http://www.hapmap.org/cgi-perl/gbrowse/gbrowse/hapmap/> and answer two more questions.

- Would you expect these 3 SNPs to be in linkage disequilibrium in any population or populations? Support your answer with data from this site.
- Find a SNP for which this is no variation. Support your answer with data from this site (even though this question sounds like an oxymoron).

20 pts.

2) Use the attached Figure 1 PDF file to answer this question. Interpret figure 1 as completely as you can. Interpret the data and tell me what you can deduce about the biology being revealed. Principle components analysis is a way to objectively identify the portions of the data that are responsible for the most amount of inter-sample variation.

20 pts.

- 3) Use only this web site to answer the following questions: <http://www.ncbi.nlm.nih.gov/geo/> . Search for this gene: NFKB1. (Read question #4 too so you will not have to redo any of this question.) Use screen shots to show one microarray example when this human gene was:
- Strongly induced in one condition but not another. What were the conditions?
 - Repressed in both conditions. What were the conditions?

- c) Strongly repressed in only one condition. What were the conditions?
- d) What are the meanings of the red and the blue symbols? Explain your answer in terms a Bio111 student could understand.
- e) What is the value to knowing the answer to part d above?

20 pts.

- 4) Read all the parts to this question before you begin.
- a) Use your answer to question 3 above that had a fold change the furthest from 1. Tell me which condition you chose, and supply me with a screen shot of the one you have chosen.
- b) Tell me the fold change for your chosen gene and the experimental conditions.
- c) Convert the fold change to a ratio of two numbers that is consistent with your data.
- d) If control is green and experimental is red, what color spot would you see on the microarray, assuming this is not an Affy chip? To answer this question, you must draw the circle and color in the spot here →

- e) Draw an arrow on this color scale to indicate the color you'd choose for your example's ratio:

**20 pts.**

- 5) Use <http://db.yeastgenome.org/cgi-bin/expression/expressionConnection.pl> to answer the following questions concerning this list of yeast genes: *Rad26*, *Rad51*, *Rad52*, *Rad54*, *Rad55*, *Rad57*, and *Rad59*.
- a) Are these genes transcribed in a coordinated fashion when exposed to environmental stresses? Support your answer with data from this web site only.
- b) Are these genes transcribed in a coordinated fashion when the genome ploidy is altered? Support your answer with data from this web site only.
- c) Use data on this web site only to support the claim that the expression profiles for these 7 genes under the two conditions above (parts a and b) accurately represents independent gene regulation and not either of two common microarray artifacts. Name each artifact then show and describe data that demonstrate each artifact is not in play for these 7 genes.
- d) One artifact cannot be argued away with these genes. What artifact is this and what information do you need in order to evaluate its presence or absence?