Spring 2014 Genomics Exam #2
Sequence Variations and Transcriptomes

There is no time limit on this test, though I don't want you to spend too much time on it. I have tried to design an exam that will take less time than exams in the past. You do not need to read any additional papers other than the ones I send to you. There are 6 pages, including this cover sheet, for this test. You are not allowed discuss the test with anyone until all exams are turned in no later than 10:30 am on Wednesday March 26. **ELECTRONIC COPIES OF YOUR EXAM ANSWERS ARE DUE BY 10:30 am ON WEDNESDAY MARCH 26**. You may use a calculator, a ruler, your notes, the book, and the internet. You may take this exam in as many blocks of time as you want. Submit your electronic version before 10:30 am (eastern time zone).

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 (I suggest a thumb drive or other removable device). 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. 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. Support your answers with data using screen shots liberally (no permission required since your exam is a private document).

DO NOT READ or DOWNLOAD ANY PAPERS FOR THIS EXAM. RELY ON YOUR EXPERIENCE, AND YOUR SKILLS.

-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 type):

Write out the full pledge and sign (electronic signature is ideal):

How long did this exam take you to complete?

40 points

1) By now, you know how much I enjoy learning and sharing with my students. In preparing for this exam, I stumbled on a biological reality that rocked my perception of the genome and cells (Figures 1 and 2).

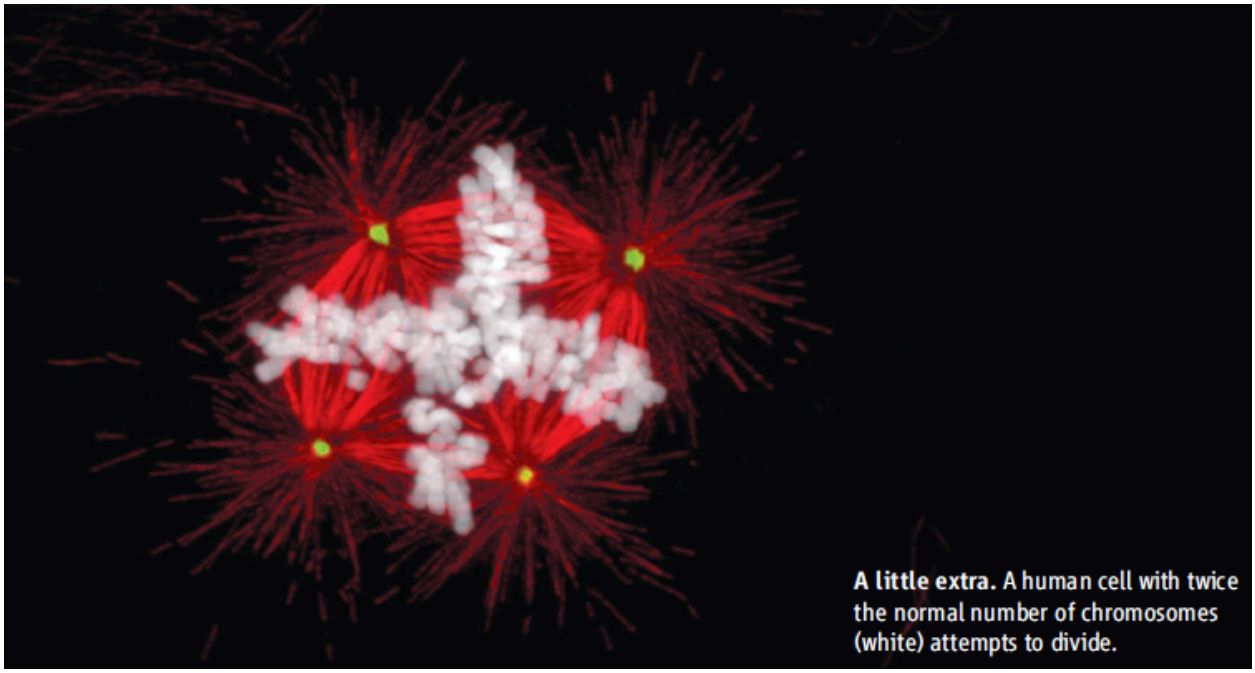
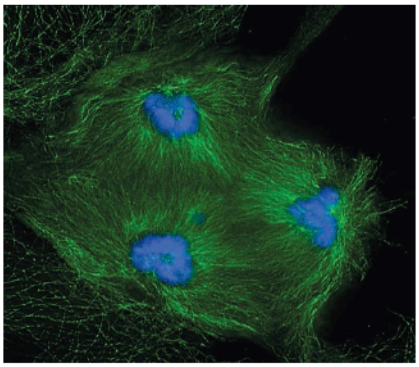


Figure 1. Image that blew Dr. C's world view.

POLYPLOID CELL TYPES IN MAMMALS

CELL TYPE	LOCATION	FUNCTION	NUMBER OF GENOME COPIES
Megakaryocyte	Bone marrow	Producing blood clotting platelets	Up to 128
Hepatocyte	Liver	Detoxification, metabolism	Typically 4 to 16
Trophoblast giant cell	Embryo	Promote implantation	Up to 1000
Cardiomyocyte	Heart	Contraction	Typically 4



Bonus DNA. The polyploid cells in mammalian bodies differ in their location, function, and number of chromosome sets (table). In a liver cell (*right*), the multiple chromosome copies (blue) have sorted into three clusters in preparation for cell division.

Figure 2. Information that made Dr. C. realize he had been lied to for decades about humans as diploids organisms.

5 points

a) Find a human gene that has experimental evidence supporting its role in normal (non-cancerous) polyploidy formation in human cells. You must provide two independent sources that verify your gene is involved.

5 points

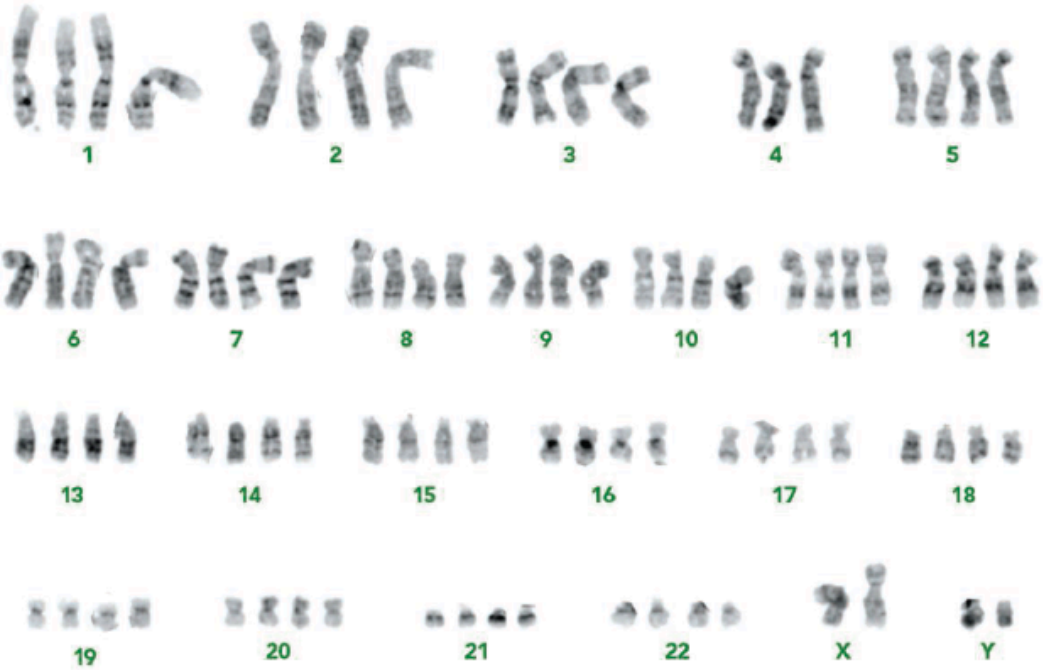
b) Determine the expression of your gene in a wide range of human tissues. Show your data and list the top 3 expressing healthy tissues.

10 points

c) For the gene you found in part (a) above, find evidence of two categories of sequence variations in humans. One source must show mutations that lead to disease states. The second source must show variations that provide population frequency information for a range of ethnically diverse people but these mutations cause no known negative phenotypes.

20 points

d) Look at this karyotype from a non-pathogenic human liver cell.



Double down. The chromosome copies from a polyploid liver cell arranged by size, showing that the cell carries four copies of almost every one.

Design an experiment to determine what percentage of liver cells are tetraploid. Your experimental method should be able to detect when one or more chromosomes are not fully tetraploid as you can see for chromosome 4 above. **Using an outline format**, tell me what you would do starting with a 5 gram piece of human liver biopsy. What methods would you use and what sorts of data would you collect? Provide some theoretical data that you would expect to get from your method.

I.

II.

35 points

2) There are many aspects of biology that fascinate me. One is the creativity of investigators who ask questions that never occurred to me. Here is some text from a recent abstract that poses a question that would have never occured to me.

> Humans traveling in space might allow us unprecedented exploration and discovery, but we need to more fully understand the consequences of long-term exposure to spaceflight. Microgravity (μg) is constant in space, and hypergravity (hyper g) is experienced during launch and landing. Immune dysfunction in both μg and hyper g has been independently documented multiple times. The human immune system is weakened in prolonged exposure to space travel which results in increased vulnerability to opportunistic infections. To understand the human immune system when exposed to space exploration, we need to better understand how the metazoan immune systems response to space flight. We used *Drosophila* as a model of innate immunity to see how space flight affects a normal immune response to pathogens as simulated by exposure to common pathogens.

5 points

a) Summarize the conclusion from Figure 2a. Support the main conclusion using data in this figure. A fly gene called *yuri* is responsible for sensing gravity in flies. Mutnant fly strain *yuri* has a deleted *yuri* gene while the rescued fly UAS contains the same deleted *yuri* allele as well as a transgene (UAS) of *yuri* that complements the missing gene and encoded protein. 1 g = earth gravity; 4 g = hyper g; minus sign indicates non-infected flies; + indicates flies infected by a fungus. Error bars = SEM for 3 experiments.

5 points

b) From Figure 2a, is it possible to distinguish between two possibilities: either 1) the fungus has an altered virulence at hyper g or 2) the fly has an altered immune response to hyper g? Explain how you reached your conclusion.

5 points

c) List the signature gene set(s) that are induced or repressed only in a 1g environment with a P value of at least 10^{-6} in Figure 2b. Fly DNA microarrays were used to survey every gene's activity comparing flies on earth vs flies on the space station. The number of genes induced or repressed in 1g flies only (Earth) or μg flies only (Space) or in both (overlap) are listed in the Venn diagrams. Signature genes set are shown on the right side of the figure. Figures show the number of genes in each signature set (bar graphs; upper y-axis), and the P values showing statistical over-representation of each signature set (circles; lower y-axis).

5 points

d) Figure 2c shows microarray data for 7 genes involved in innate immunity. Affymetrix microarrays were used to generate the data. * = $p < 0.01$; ** = $p < 0.005$; *** = $p < 0.001$.

Produce a Table that has the name of each gene at the top of columns and the rows are labled by the experimental condition. Then fill in the table by typing "ind" for every gene that is induced or "rep" for every gene that is repressed on earth compared to in space. U = uninfected flies; B = bacterially infected flies; F = fungally infected flies. Leave table cells blank if they are not significantly altered.

5 points

e) Look at Figure 2d. As you did for question 2c above, which signature gene set(s) show significan differences? U = uninfected; E = earth; S = space station; 1 – 3 represent the triplicate batches of flies. Color scale uses the same trends that Botstein and Brown developed for their microarrays.

5 points

f) All of the stress response genes in Figugre 2d are heat shock protein (hsp) paralogs. Look up what heat shock proteins do with regards to protein folding. Speculate why hsp gene regulation was an idication of what was going wrong with fly innate immunity with regards to gravity. Support your speculation with data presented in this exam as well as online resources.

5 points

g) What was the value of using DNA microarrays in this study?

25 points

3)

5 points

a) Summarize the genomic and evolutionary history of the bonobo and chimpanzee. Your answer cannot exceed 75 words. Cite your source of information (citation is not part of the word count). You may want to refer to Figure 3map for an understanding of geography.

5 points

b) The investigators sequenced the mtDNA as well as 15 different autosomal loci. Figure 3a shows you the evolutionary tree from a subset of these data. Describe the overall genome variation for the five different populations whose DNA were sequenced. Red = bonobos; green for western chimps; grey = central; blue = eastern; yellow = Nigerian/Cameroonian (aka vellerosus).

5 points

c) Go online to find out what pairwise F_{st} values are and how they are used. Apply what you learn to Figure 3b. Summarize your interpretation of this comparative table.

5 points

d) Look at Figure 3c which uses principle component analysis (just like Dr. Peek did). What can you conclude from these data? Are these data consistent or inconsistent with what you have learned so far in Question 3?

5 points

e) Change of pace: Some folks at Carnegie Mellon University developed a meta-analysis tool (ExpressionBlast) for transcriptome data. Choose their sample project to find a human gene that is down regulated when exposed to 25 μM resveratrol and show me the screen shot of that gene.
http://www.expression.cs.cmu.edu/index.html