**GENE 210: Personalized Genomics and Medicine**

**Spring 2014 Final Exam**

**Due Monday, May 26 2014 at midnight.**

Stanford University Honor Code

The Honor Code is the University’s statement on academic integrity written by students in 1921. It articulates University expectations of students and faculty in establishing and maintaining the highest standards in academic work:

• The Honor Code is an undertaking of the students, individually and collectively:

– that they will not give or receive aid in examinations; that they will not give or receive unpermitted aid in class work, in the preparation of reports, or in any other work that is to be used by the instructor as the basis of grading;

– that they will do their share and take an active part in seeing to it that others as well as themselves uphold the spirit and letter of the Honor Code.

• The faculty on its part manifests its conﬁdence in the honor of its students by refraining from proctoring examinations and from taking unusual and unreasonable precautions to prevent the forms of dishonesty mentioned

above. The faculty will also avoid, as far as practicable, academic procedures that create temptations to violate the Honor Code.

• While the faculty alone has the right and obligation to set academic requirements, the students and faculty will work together to establish optimal conditions for honorable academic work.

Signature

I attest that I have not given or received aid in this examination, and that I have done my share and taken an active part in seeing to it that others as well as myself uphold the spirit and letter of the Stanford University Honor Code.

Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ SUNet ID:\_\_\_\_\_\_\_\_\_

Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Some questions may have multiple reasonable answers: if you are unsure, provide a justification based in genetics and cite your sources (SNPedia is fine, journals are better); as long as the justification is sound, you will receive full credit.

If you are unsure which SNP(s) are associated with a trait, you may consult any reference you like.

A family of 3 (mother/father/daughter) has come to you to find out what they can learn from their genotypes. The parents were both adopted, so they do not know any of their family history. You have sent their DNA to LabCorp, which ran their genotypes on a custom 1M OmniQuad array, and they’ve returned the results at: http://www.stanford.edu/class/gene210/files/final/final\_patients.zip *(X points)*

1. A mislabeling in the lab has caused the samples to be shuffled around and they are simply labeled: ‘patient1.txt,’ ‘patient2.txt,’ and ‘patient3.txt.’ Determine which sample is the mother’s, the father’s and the daughter’s. *(15 points)*

2. What can you tell about the ancestry of the parents? *(15 points)*

3. The parents are concerned about their daughter’s chance for getting breast cancer. You investigate the genomes of the father, mother and the daughter and provide genetic counseling for the family. *(15 points total)*

1. What is the lifetime risk for breast cancer for the overall population of Europeans?
2. Does the genotype of the mother or daughter (at rs77944974) alter their risk of breast cancer? Explain briefly, providing data on the most important risk alleles and their effect on risk for breast cancer.
3. Briefly outline what advice you would give to the mother about her risk for breast cancer, based on your analysis?
4. Briefly outline what advice you would give to the daughter about her risk for breast cancer, based on your analysis?

4. Weeks later, the father (a 42 year old, 185 cm in height, 80 kg in weight, not taking any other medication) is rushed to the hospital with a stroke. What dose of warfarin would be given from a clinic that does not perform genetic testing? What dose of warfarin would be given from a clinic that does perform genetic testing? Explain the genetic basis for modifying the warfarin dose of the father given his genotype. *(8 points)*

5. In her next visit, you observe that the mother has high cholesterol. Would you prescribe simvastatin (Zocor) to the mother? Why or why not? *(7 points)*

6. You counsel the family about the risk for type 2 diabetes for their daughter. You analyze the daughter’s genome on genotation.com. You need to explain the results to the family, and how this influences the daughter’s risk for Type 2 diabetes. *(15 points total)*

1. What is the likelihood of type 2 diabetes prior to genetic testing?
2. What is the likelihood of type 2 diabetes following analysis of the daughter’s genotype using Genotation?
3. How many SNPs were used to assess the risk for type 2 diabetes?
4. How were the SNPs combined to give the overall score? Which SNP had the greatest influence on diabetes risk? Explain briefly.
5. What advice can you provide to the family to help mitigate the chance of their daughter developing type 2 diabetes?

7. The following two SNPs were shown to be associated with risk for type 2 diabetes in two GWAS studies. *(15 points total)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **snp** | **odds ratio** | **p-value** | **cases** | **controls** |
| rs4402960 | 1.14 | 8.9 x 10-16 | 14586 | 17968 |
| rs7754840 | 1.28 | 3.5x10-7 | 1921 | 1622 |

1. Which SNP has a larger effect size on risk for type 2 diabetes? Explain your answer.
2. Which SNP is most statistically significant for risk for type 2 diabetes; i.e. which SNP is most likely to have a true association?
3. Is the SNP with the biggest effect size on risk for type 2 diabetes always going to be the SNP that is most statistically significant? Why or why not?
4. rs7754840 is a SNP that lies within the CDKAL1 gene. This SNP was identified because it was contained on the Illumina Chip used for genotyping in the GWAS study. Does this result indicate that rs7754840 is the causal mutation? Does this result indicate that CDKAL1 is involved in type 2 diabetes? Explain why or why not.

8. The two parents are considering having another child. You analyze their genomes and then counsel them on their chance of having a child with one of the following diseases: hemochromatosis (rs1800562), Alzheimer’s disease (specifically, look for APOE4 status), breast cancer (BRCA1 status; rs77944974), cystic fibrosis (rs113993960) and sickle cell anemia (rs334).

For each of these five diseases, what is the chance that the child will have that disease? Briefly explain your answer. *(15 points total)*

9. Prenatal genetic diagnosis (15 points total)

A) A pregnant woman seeks non-invasive prenatal genetic testing and provides a sample of plasma. You isolate the cell-free DNA (cfDNA) from the maternal plasma and determine that 10% of it is derived from the fetus. You perform whole genome sequencing on genomic DNA samples from the mother and father. Next you perform whole genome sequencing on the cfDNA isolated from maternal plasma. For each of the sites below, you obtain 100X coverage (i.e., 100 reads for each site). Fill in the **expected** read counts in the tables below. Use the parental genotypes below and the observed allele counts for the cfDNA sequencing to infer the genotype of the fetus at each of three sites and fill them in the table.



**Site 1**

|  |  |
| --- | --- |
|  | A reads expected |
| If mother transmits A |  |
| If mother transmits G |  |

 **Site 2**

|  |  |
| --- | --- |
|  | A reads expected |
| If mother transmits A |  |
| If mother transmits G |  |

**Site 3**

|  |  |
| --- | --- |
|  | T reads expected |
| If mother transmits T |  |
| If mother transmits C |  |

**Observed reads in cell free DNA**

|  |  |
| --- | --- |
| Site 1 | 59 A reads |
| Site 2 | 52 A reads |
| Site 3 | 49 T reads |

**Infer fetal genotype:**

|  |  |  |
| --- | --- | --- |
| **Site 1** | **Site 2** | **Site 3** |
|  |  |  |

B) You worry that your call at site 2 might not be accurate. In order to improve the accuracy of your fetal genotyping, you use parental haplotype blocks. Re-evaluate your fetal genotype inference based on the maternal haplotypes below.



**Re-evaluated fetal genotype inference:**

|  |  |  |
| --- | --- | --- |
| **Site 1** | **Site 2** | **Site 3** |
|  |  |  |

10. Extra credit question available at <http://www.stanford.edu/class/gene210/web/html/extracredit.html> (*12 pts*).