Course Overview

02-223 Personalized Medicine: Understanding Your Own Genome
Fall 2014
Overview

• Course website: http://www.cs.cmu.edu/~sssykim/teaching/f14/f14.html

• All lab reports and term paper should be submitted to blackboard by midnight on the due date

• There is no pre-requisite for this course
Overview

• Lab reports
  – We do not require you to be able to program. However, we will use Matlab to perform simple analyses of genome data
  
  – We will not do an extensive Matlab programming, but use simple operations/routines available in Matlab
  
  – CMU students can download Matlab from the following CMU software website for free
    • https://www.cmu.edu/computing/software/all/

• Term paper
  – Team of up to two students can work together for a report

• Late submission policy
  – 80% for one day late, 50% for two day late, 0 afterwards
Genomes

• Genomes as DNA sequence: a sequence of A, T, C, G’s

• Human Genome
  – 23 pairs of chromosomes
    • 22 pairs of autosomes
    • One pair of sex chromosomes
      – Males: X and Y chromosomes
      – Females: two X chromosomes
2001: Human Genome Sequencing Project

2011: 1000 Genome Project
# Genetic Variants, Genetic Polymorphisms

## The ABO Blood System

<table>
<thead>
<tr>
<th>Blood Type (genotype)</th>
<th>Type A (AA, AO)</th>
<th>Type B (BB, BO)</th>
<th>Type AB (AB)</th>
<th>Type O (OO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Blood Cell Surface Proteins (phenotype)</td>
<td><img src="image1.png" alt="Type A Red Blood Cell Surface Proteins" /></td>
<td><img src="image2.png" alt="Type B Red Blood Cell Surface Proteins" /></td>
<td><img src="image3.png" alt="Type AB Red Blood Cell Surface Proteins" /></td>
<td><img src="image4.png" alt="Type O Red Blood Cell Surface Proteins" /></td>
</tr>
<tr>
<td>A agglutinogens only</td>
<td>B agglutinogens only</td>
<td>A and B agglutinogens</td>
<td>No agglutinogens</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasma Antibodies (phenotype)</th>
<th><img src="image5.png" alt="Type A Plasma Antibodies" /></th>
<th><img src="image6.png" alt="Type B Plasma Antibodies" /></th>
<th><img src="image7.png" alt="Type AB Plasma Antibodies" /></th>
<th><img src="image8.png" alt="Type O Plasma Antibodies" /></th>
</tr>
</thead>
<tbody>
<tr>
<td>b agglutinin only</td>
<td>a agglutinin only</td>
<td>No agglutinin</td>
<td>a and b agglutinin</td>
<td></td>
</tr>
</tbody>
</table>

Percent of population that has the B allele:

- 0-5
- 5-10
- 10-15
- 15-20
- 20-25
- 25-30
Single Nucleotide Polymorphisms (SNPs)

Polymorphism rate: number of letter changes between two different members of a species
Humans: ~1/1,000
Disease-related Genetic Variants

Healthy

Cancer

Genetic variants  Disease causing genetic variants

ACTCGTACGTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
TCTCGTACGTAGACGTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
ACTCGAACCTAGACCTAGCATTACGCAATAATGCGA
Disease-related Genetic Variants

Disease causing genetic variants

Biological mechanism

Healthy

Diabetic
Gene\textsuperscript{c} Variants for Lung Cancer

- Lung cancer, all types
- Adenocarcinoma
- Squamous cell carcinoma
Genetic Variant for Blond Hair Color

Regulatory region of KITLG controls hair pigmentation of human and mice

Kingsley et al. Nature Genetics, 2014
Gene?c	
  
  Variant	
  for	
  Blond	
  Hair	
  Color

G	
  allele	
  (associated	
  with	
  blond	
  hair)	
  is	
  more	
  prevalent	
  in	
  Europe

Frequency of A and G alleles in different parts of the world

G allele (associated with blond hair) is more prevalent in Europe
Genetic Variant for Blond Hair Color

Mutations in the corresponding (orthologous) mouse genome region results in light coat color
A Little Bit of History

• 2001: The first draft of a human genome sequence became available

• 2001: The International SNP Map Working Group publishes a SNP Map of 1.42 million SNPs that contained all SNPs identified so far
A Little Bit of History

• 2005: HapMap Phase I
  – Genotype at least one single-nucleotide polymorphisms (SNPs) every 5kb across 270 individuals
  – Geographic diversity
    • 30 trios (mother, father, child) from Yoruba in Ibadan, Nigeria (YRI)
    • 30 trios of European ancestry living in Utah (CEPH)
    • 45 unrelated Han Chinese in Beijing (CHB)
    • 45 unrelated Japanese (JPT)
  – 1.3 million SNPs

Hapmap.org
A Little Bit of History

• 2007: HapMap Phase II
  – Genotype additional 2.1 million SNPs
  – Estimated to contain 25-35% of all 9-10 million common SNPs in human genome.

• 2010: 1000 Genome Pilot Project
  – A more complete characterization of human genetic variations
GENETIC TESTING AND POLICY ISSUES
Learn hundreds of things about your health.

- Reports on 240+ health conditions and traits
- Discover your lineage, find relatives and more
- Get updates on your DNA as science advances

23andMe DNA Spit Kit

order now $99
Learn hundreds of things about your health.

- Reports on 240+ health conditions and traits
- Discover your lineage, find relatives and more
- Get updates on your DNA as science advances

Genetic risk for developing diseases:
- Do I have the BRCA mutation that is known to increase breast cancer susceptibility?
- How likely is it that I will develop type 2 diabetes?
- How is my body going to respond to a particular drug?
- What is the appropriate prevention strategy for me?
Carrier status
Find out if your children are at risk for inherited conditions, so you can plan for the health of your family.

Health risks
Understand your genetic health risks. Change what you can, manage what you can't.

Drug response
Arm your doctor with information on how you might respond to certain medications.

Health tools
Document your family health history, track inherited conditions, and share the knowledge.

Inherited traits
Explore your genetic traits for everything from lactose intolerance to male pattern baldness.

Scientific advances
Keep receiving updates on your DNA as discoveries are made, so your knowledge grows as you do.

www.23andme.com
I am Asian. Do I have African or Caucasian ancestors? If so, to what extent?
To what extent were Neanderthals my ancestors?
Learn hundreds of things about your health.

- Reports on 240+ health conditions and traits
- Discover your lineage, find relatives and more
- Get updates on your DNA as science advances

ordernow $99
• Participate in studies of conditions and traits you care about
• Join research communities for
  • Parkinson’s disease
  • Sarcoma
  • Myeloproliferative Neoplasma

www.23andme.com
23andMe: Policy/Legal Issues

• Does the test provide scientific information or medical information?

• Direct-to-consumer model: genetic test results are available directly to consumers without going through medical professionals
Policy/Legal Issues

• Health-related genetic tests were suspended by FDA, beginning November 2013

• Only genome-based ancestry service is provided now

• 23andMe is working with FDA for a resolution
Find out what your DNA says about you and your family.

- Learn what percent of your DNA is from populations around the world
- Contact your DNA relatives across continents or across the street
- Build your family tree and enhance your experience with relatives

order now $99
The largest DNA ancestry service in the world.

With over 700,000 genotyped members, you'll experience more matches, more data and more discoveries.

23 pairs of chromosomes. One unique you.

Find out what percent of your DNA comes from populations around the world, ranging from East Asia, Sub-Saharan Africa, Europe, and more. Break European ancestry down into distinct regions such as the British Isles, Scandinavia and Italy. People with mixed ancestry, African Americans, Latinos, and Native Americans will also get a detailed breakdown.
Find relatives across continents or across the street.

Ever wonder who’s related to you? You’ll likely discover dozens or even hundreds of people who share DNA and ancestors. The matches you’ll get can range from close family to distant cousins.

Neanderthal DNA lives on in us.

Even though Neanderthals vanished about 30,000 years ago, their DNA lives on in us. Based on research, they interbred with humans between 37,000 and 86,000 years ago and we are able to tell you what percentage of your genome came from Neanderthals.

We estimate how much of your DNA is from Neanderthals and compare it to other users.

- **Modern Human**
  - Higher brow
  - Narrower shoulders
  - Slightly taller

- **Neanderthals**
  - Heavy eyebrow ridge
  - Long, low, bigger skull
  - Prominent nose with developed nasal chambers for cold-air protection
GENETICS AND STATISTICS
How to Analyze Your Own Genome

• Genotype/sequence your genome
• Annotate your own genome
  – By comparing your genome with the reference genome
  – By comparing your genome with the genomes of the rest of the population

• Statistics!
Genetics and Statistics

• Population genetics
  – Study of genome data collected for a large number of individuals
  – Statistics for genetics

• Population genetics before genome sequencing?

R.A. Fisher (1890-1962): geneticist & statistician
James Watson’s and J. Craig Venter’s Genome

• Demonstrates the full genome sequencing for individualized genomics

• The two genomes were made available to the public despite the risk of divulging personal details
  – Ethical concerns!
What You Should Know

• Genetic variants: single nucleotide polymorphisms

• What can you learn from genetic variants?
  – Your disease risk
  – Your ancestry

• Ethical/legal concerns regarding genome information