Genetics Spring 2022
Week 14: April 18th - 22th

Remember: the Tutoring Center offers free individual and group tutoring for this Genetics. Our Group Tutoring sessions will be Thursday from 5:15-6:15 PM at the Sid Rich basement, room 75! You can reserve a spot at https://baylor.edu/tutoring. I hope to see you there!

Keywords: Genome, Polymorphisms, Gene Flow, Hardy-Weinberg

Topic of the Week: Genomics and Proteomics (20.1-3)

Genomics: the study of the genome (the totality of genetic information) of an organism

Think back to T.H Morgan and genetic maps… Would a genetic (based on recombinant frequency) or physical map of a DNA molecule be more ‘accurate?’

Genetic Map: a map in ‘map units’ based on the recombination frequency of sets of genes → lower resolution and accuracy

Physical Map: a map in physical units (ex. pm, nm, μm) which gives the physical distance between genes on a chromosome → higher resolution and accuracy

Genome Sequencing: determining the entire set of gene sequences in an organism’s genome

This is difficult to do as only 500-700 bp can be synthesized at a time, meaning millions of overlapping fragments must be put together to determine the sequence of a genome

The Human Genome: approximately 3.2billion base pairs, or 20-25,000 genes

The Human Genome Project: the goal was to sequence the entire human genome. This was accomplished in 2003 after 13 years. This is how:

Researchers modeled the genomes of increasingly complex model organisms and built sequencing techniques
Using the tools gained from smaller genomes, they applied this to sequence the entire human genome.

Single Nucleotide Polymorphisms (SNP): single base pair variations between individuals of a given species → generally do not change phenotypes, so these are in non-coding DNA regions (arising through random, silent mutations)

Genome Wide Association Studies: using SNPs across an entire genome to find a particular gene of interest

All diagrams, tables and figures are the property of Benjamin A. Pierce; Additional sources Are credit of Pearson Education (Can
**Haplotype:** a specific set of SNP variations identifiable to an individual

**Bioinformatics:** using molecular biology and computer science to analyze and codify genes. These genes (or associated proteins) may be put into databases where they can be analyzed or annotated.

**Metagenomics:** sequencing the genome an *entire group* of organisms that share an entire environment and look at the effects on their DNA/protein expression

**Microbiome:** a metagenome that comprises all of the ‘gut bacteria’ in an animal; this collection of bacteria can be linked to health and certain cancers

**Functional Genomics:** determining the function of genes in a given organism

**Transcriptome:** the totality of an organisms’ RNA genome that may be studied by functional genomics

**Proteome:** The entire set of proteins produced by an organism, whose functions are studied

**Homology:** studying evolutionarily related genes, proteins or traits → often arise due to duplication

**Analogs:** an organism in a completely separate lineage has a gene/gene product similar to another *unrelated* species

**Orthologs:** homologous genes in two *different species* who share a [recent] common ancestor

**Paralogs:** homologous genes in the same species

---

**Highlight #1: Population Genetics and Hardy-Weinberg Equilibrium (25)**

**Population Genetics:** the study of changes in the allelic frequencies within a population of individuals (the fundamental unit of study for evolution and heredity)

**Microevolution:** small changes to *allelic frequency* with respect to time that can lead to *speciation* (the creation of a new species)

**Gene Flow:** the transmission of genetic information between two groups of a species; cutoff of gene flow between populations considered *speciation*.

**Gene Frequency:** the frequency of given genotypes of alleles

**Allelic Frequency → NOTE:** 2N is the total number of alleles present

**Dominant Allelic Frequency:** the frequency of a dominant allele (ex. A)

\[ f(A) = p = \frac{2n(AA) + n(Aa)}{2N} \]

**Multiple Alleles:** when there is more than one allele (ex A1, A2, …)

\[ f(A^3) = p = \frac{2n(AA) + n(AA1) + n(AA2)}{2N} \]

**Recessive Allelic Frequency:** the frequency of a recessive allele (ex. a)

---

All diagrams, tables and figures are the property of Benjamin A. Pierce; Genetics: A Conceptual Approach

Additional sources Are credit of Pearson Education (Campbell Biology, 2011)
Sex-Linked (X-Linked) Genes: slight variation, were the number of males and females is considered as m and f, respectively

\[ f(X) = \frac{2n(X1X1) + n(X1X2) + n(X1Y)}{2Nf + 2Nm} \]

Hardy-Weinberg Equilibrium: what about a non-evolving population? Hardy-Weinburg equilibrium describes when a population can be at “rest,” i.e. not evolving.

To be a “non-evolving” population, it must meet the following criteria:

- Large population
- Random mating
- No mutations
- No natural selection
- No migration

Additionally, allelic frequencies do not change and genotypic frequencies will stabilize after one generation.

When the assumptions are met, the allelic and genotypic frequencies are not changing, so there is no evolution occurring (and the opposite is true as well).

Looking at the change in a population over time is modeled as follows:

Hardy-Weinberg Equation:

\[ p^2 + 2pq + q^2 = 1 \]

\[ p + q = 1 \]

- p: dominant allelic frequency
- q: recessive allelic frequency
- \( p^2 \): frequency of homozygous dominant
- \( q^2 \): frequency of homozygous recessive
- 2pq: frequency of heterozygotes

*The significance of a change in a population’s allelic or genotypic frequencies can be modeled by - you guessed it- the \( X^2 \) test of independence!

https://docs.google.com/document/d/1wCZ2KZr8nJ0lw0WXn8qCRaprsNloxbROHgTW-6JD1A/edit?usp=sharing
Week 13 Concept Check:

1. Ankylosing spondylitis (AS), a degenerative genetic disease, occurs about once in every 1,000 Americans. Assuming that the population is in Hardy-Weinberg equilibrium, what percentage of the population would be expected to be a carrier? (think about what you are given!)
   a. 8.56%
   b. 6.12%
   c. 3.16%
   d. 96.8%

2. Alpha, beta and gamma globulin are blood proteins in humans. These evolved from a single globulin gene in humans. What type of homology is this?
   a. Analogs
   b. Orthologs
   c. Paralogs
   d. Heterologous

3. Which of these is not a type of RNA expression study? (note: this content is in the week 12 resource, but the majority of the chapter is covered here!)
   a. Reporter sequences
   b. RNA seq
   c. Western blotting
   d. Microarrays

4. True or False: Non-random mating affects allele frequencies.

5. This table models the phenotypes of a population of lions:

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golden Fur</td>
<td>120</td>
<td>122</td>
</tr>
<tr>
<td>White Fur</td>
<td>18</td>
<td>15</td>
</tr>
</tbody>
</table>

The fur white fur gene follows an *autosomal recessive* pattern of inheritance. What is the dominant allelic frequency? (assume H-W equilibrium)
   a. 0.88
   b. 0.12
   c. 0.65
   d. 0.35

6. (Using the above table) What is the frequency of *heterozygous* lions in this population?
   a. 0.12
   b. 0.83
   c. 0.46
   d. 0.91

*All diagrams, tables and figures are the property of Benjamin A. Pierce; Genetics: A Conceptual Approach*  
*Additional sources Are credit of Pearson Education (Campbell Biology, 2011)*
THINGS YOU MAY STRUGGLE WITH:

1. In SNPs, 99.9% of the human genome is shared and 0.1% is therefore variable. Thus, multiplying the variance by the genome size - 0.01x3.2billion - ~30 million bp of variation is to be completely expected.
2. SNPs may be associated with a particular disorder or can be identified to a certain individual.

CONGRATS; You made it to the end of the resource! Again, group tutoring will be every Thursday from 5:00-6:15 PM. You can reserve a spot at https://baylor.edu/tutoring. I hope to see you there!

Answers:

1. B.
   a. \( q^2 = 0.001 \rightarrow q = 0.0316; p = 0.968 \rightarrow 2pq = 0.0612 \times 100\% = 6.12\%

2. C.
3. C.
4. False
   a. Only mutation, population size (genetic drift), natural selection and migration affect allele frequencies
5. C.
6. C.