1. SNAIL MIGRATION. Consider an island population of previously isolated snails that starts to receive migrants from a mainland population at a rate of $m = 0.2$. All the snails on the island initially are homozygous for “yellow” alleles and all the snails on the mainland are homozygous for “black” alleles.

(a, 3pts) What is the frequency of the “yellow” allele on the island after 3 generations?

(b, 3pts) About how many “yellow” individuals do we expect to see if the “yellow” allele is recessive and the island’s population size is 100,000?

(c, 2pts) What is the equilibrium frequency of the “yellow” allele in the island population?

\[
\begin{align*}
(A) \\
&P_t = P^* + (1-m)!(P_0-P^*) \\
&\text{so } P_3 = 0 + (1-0.2)^3(1-0) = 0.8^3 = 0.512 \\

(B) \\
&\text{Since the allele is recessive only } y/y \text{ homozygotes have the phenotype, the frequency of these individuals is } q^2 = 0.512^2. \text{ Multiplying by } 1,000,000 \text{ gives acceptable answers of } \\
&(0.512)^2(100,000) \text{ or } 0.262144(100,000) = 26,144 \\

(C) \\
&\text{The only equilibrium for the island is to become the same as the mainland, a frequency of zero.}
\end{align*}
\]
2. MUTATION: SELECTION BALANCE. Genetic diseases are usually considered to be due to the segregation of recessively, purely deleterious, alleles. However, sometimes these alleles can provide advantages that are not detected until we look for them - alternately they could also have additional deleterious effects that we don't fully recognize.

Consider a hypothetical genetic variant, "a", which researchers find is associated with reduced rates of infant survival compared to individuals with the wildtype "A" allele. The researchers want to know the details of the fitness consequences of this allele. Initial studies show that individuals that are homozygous for this allele have only 80% the normal chance of surviving until adulthood.

(a, 3 pts) Assuming a mutation rate of $10^{-8}$ for this genetic disorder (approximately correct for single nucleotide mutation rates), what is the predicted frequency of this allele in the population?

\[ f(a) = 0.0002236 \]

(b, 3 pts) The population of the US is approximately 300,000,000. Assuming this, what is the expected number of afflicted individuals in the US?

\[ # \text{sick} = 15.0 \]

(c, 3 pts) The population of the US is approximately 300,000,000. Assuming this, what is the expected number of heterozygous individuals, "carriers", in the US?

\[ # \text{carriers} = 134,130 \]

(d, 3 pts) If the researchers conducted a genetic testing screening and found the frequency of the deleterious allele to be 0.0002 instead of the value they did (i.e., the correct answer), what does this imply for the presence of unknown beneficial or deleterious effects caused by the gene?

\[ (A) \]
\[ q_{eq} = (10^{-8}/0.2)^{1/2} = 10^{-4} \times 5^{1/2} = 0.0002236 \]

\[ (B) \]
\[ q_{eq}^2 = ((10^{-8}/0.2)^{1/2})^2 = (10^{-8}/0.2) = 5 \times 10^{-7} \]
\[ # = (3 \times 10^8)(5 \times 10^{-7}) = 15 \]

\[ (C) \]
\[ 2q_{eq}q_{eq} = 2 \times (1-0.0002236)(0.0002236) = 4.471 \times 10^{-4} \]
\[ # = (3 \times 10^8)(4.471 \times 10^{-4}) = 134,130 \]

\[ (D) \]
This frequency is lower than 0.0002236 therefore there is likely some unknown additional deleterious effect due to the allele which is causing its frequency to be slightly lower.
3. POPULATION GENETICS. Consider a population of 2000 fish individuals that is at Hardy-Weinberg equilibrium. There is one locus with two alleles, “A” and “a”, that influence the swimming speed. The “A” allele is dominant and “AA” homozygotes are fast while “aa” individuals are slow. Fast and slow individuals are roughly as common as one another. Within the overall population, 840 of the individuals are heterozygous.

(a, 2 pt. ea) What are the allele frequencies?

\[ f(A) = \text{______} 0.3 \]
\[ f(a) = \text{______} 0.7 \]

(b, 2 pt. ea) What are the genotype frequencies?

\[ f(AA) = \text{______} 0.09 \]
\[ f(Aa) = \text{______} 0.42 \]
\[ f(aa) = \text{______} 0.49 \]

(c, 2 pt. ea) What are the phenotype frequencies?

\[ f(\text{fast}) = \text{______} 0.51 \]
\[ f(\text{slow}) = \text{______} 0.49 \]

(d, 2 pts) What proportion of the homozygous individuals are fast?

\[ \text{Prop.} = \text{______} 0.155 \]

Imagine now that an invasive species is introduced into the environment such that slower individuals are more prone to predation while faster individuals are better able to escape. This results in a 20% advantage for faster individuals compared to slower ones.

(d, 2 pt. ea) After one generation of these new condition, what would you predict the new frequencies of the alleles to be?

\[ f(A) = \text{______} 0.327 \]
\[ f(a) = \text{______} 0.673 \]

(A)

\[ 2pq = 840/2000 = 0.42 \] so \[ pq=0.21 \] so \[ p \] and \[ q \] are 0.7 and 0.3.

Since the recessive phenotype is roughly 50% we know the recessive allele “a” is the 0.7 freq. (since \( 0.7^2 = 0.49 \) whereas \( 0.3^2 = 0.09 \)).

(B)

\[ p^2 = 0.32 = 0.09 \]
\[ 2pq = 2 (0.3)(0.7) = 0.42 \]
\[ q^2 = 0.72 = 0.49 \]

(C)

\[ \text{dom:} 0.09 + 0.42 = 0.51 \]
\[ \text{rec::} 0.49 \]

(D)

\[ f(\text{fast}/\text{homozygous}) = 0.09/(0.09 + 0.49) = 0.155172 \]

(E)

\[ \Delta p = pq/(w_{\text{mean}}) \times [p(1.2 - 1.2) + q(1.2-1.0)] \]
\[ \Delta p = pq/(w_{\text{mean}}) \times [p(0) + q(0.2)] \]
\[ \Delta p = (0.3)(0.7)/(w_{\text{mean}}) \times [0.2] \]
\[ \Delta p = (0.21)/(1.102) \times [0.14] \]
\[ 0.190563 \times 0.14 = 0.026679 = 0.027 \]
4. POPULATION GENETICS. Consider a population of 800 individuals that is at Hardy-Weinberg equilibrium. There are two loci, each with two additive/over-dominant alleles, that influence the "pelage" color (pelage is the fur of an animal).

► At the first locus the "CC" individuals are black, "Cc" individuals are red, and "cc" individuals are brown.
► At the second locus the "DD" individuals are dark, "Dd" individuals are medium in color, and "dd" individuals are light.

The allele frequencies are \( f(C) = 0.5 \) and \( f(D) = 0.8 \).

(a, 2 pt. ea) how many individuals of each of the following phenotypes are there? Provide these values to the nearest integer.

# dark red inds = 256
# medium black inds = 64
# light brown inds = 8

Imagine now that there is positive assortative mating based on how light or dark the individuals are.

(b, 1 pt each) For each of the following genotypes, what would the frequencies be relative to the expectations under H-W equilibrium? (write your conceptual answer in the box)

<table>
<thead>
<tr>
<th>&quot;CCDD&quot; individuals</th>
<th>&quot;CcDd&quot; individuals</th>
<th>&quot;ccdd&quot; individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher than expected. or Higher than 0.16.</td>
<td>Lower than expected. or Lower than 0.16.</td>
<td>Higher than expected. or Higher than 0.01</td>
</tr>
</tbody>
</table>

Imagine now that environmental conditions changed such that movement became more difficult, limiting the region within which individuals can disperse or search for mates.

(c, 1 pt each) For each of the following genotypes, what would the frequencies be relative to the expectations under H-W equilibrium? (write your conceptual answer in the box)

<table>
<thead>
<tr>
<th>&quot;CCDD&quot; individuals</th>
<th>&quot;CcDd&quot; individuals</th>
<th>&quot;ccdd&quot; individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher than expected. or Higher than 0.16.</td>
<td>Lower than expected. or Lower than 0.16.</td>
<td>Higher than expected. or Higher than 0.01</td>
</tr>
</tbody>
</table>
(1) Fisher’s fundamental theorem is best paraphrased by which of the following?
(A) Populations evolve the fastest when they have the highest immigration rate.
(B) Populations evolve the fastest when they have the highest mutation rate.
(C) Populations evolve the fastest when they have the most environmental variation.
(D) Populations evolve the fastest when they have the most genetic variation.
(E) Populations evolve the fastest when they have the largest population size.

(2) Consider a situation in which a diploid population has 3000 individuals and a generation time of 5 years. Suddenly, a new neutral allele arises by mutation. Which of the following is closest to the probability that it will fix and be the wildtype allele in the future?
(A) 0.0002
(B) 0.0004
(C) 0.0008
(D) 0.0012
(E) 0.0015

(3) Consider a situation in which a diploid population has 3000 individuals and a generation time of 5 years. Suddenly, a new neutral allele arises by mutation. If this novel alleles fixes and becomes the new wildtype allele, which of the following is closest to how long would this take?
(A) 750 years
(B) 3,750 years
(C) 12,000 years
(D) 15,000 years
(E) 60,000 years

For the next two questions consider a cross between an AABbcc and an aaBbCC individual where the alleles at the three loci influence a quantitative trait.

(4) How many different genotypes can result from the cross?
(A) 1
(B) 2
(C) 3
(D) 4
(E) 5

(5) Assume that the effects of the different alleles at each locus are identical with regard to increasing or decreasing the trait and that capital letters indicate alleles dominant to lower case ones. How many different phenotypes can result from the cross?
(A) 1
(B) 2
(C) 3
(D) 4
(E) 5

For the next two questions consider a cross between an AaBbCc and an AaBbCc individual where the alleles at the three loci influence a quantitative trait.

(6) How many different genotypes can result from the cross?
(A) 3
(B) 6
(C) 9
(D) 18
(E) 27

(7) Assume that the effects of the different alleles at each locus are identical with regard to increasing or decreasing the trait and that capital letters indicate alleles dominant to lower case ones. How many different phenotypes can result from the cross?
(A) 1
(B) 2
(C) 3
(D) 4
(E) 5

(8) Which of the following is closest to the census population size of a population that has 500 males and 300 females?
(A) 400
(B) 600
(C) 700
(D) 750
(E) 800

(9) Which of the following is closest to the effective population size of a population that has 500 males and 300 females?
(A) 400
(B) 600
(C) 700
(D) 750
(E) 800
For the next 5 questions consider these 60 nucleotide aligned DNA sequences for six hypothetical individuals within a population (the vertical numbers tell you the position from 1 to 60):

```
Ave   ACTTCGATAGCTAGAATCGATAGCTAGAACCAGATAGCCAGAGAACTAGACTAGGCCAGA
Bas   ...A.........................G...G...G........
Cry   .............A....T......C..................G........
Der   ..................C.............C...................T...
Eve   ...A.............C........C...................G........T...
Fry   ............C........C.............T................
```

(10) What is the allele frequency of the adenine single nucleotide polymorphism (SNP) at position 3?
(A) 0.166  (B) 0.333  (C) 0.500  (D) 0.666  (E) 0.833

(11) How many transitions appear to have occurred during the history of these taxa?
(A) 3  (B) 4  (C) 5  (D) 6  (E) 7

(12) How many transversions appear to have occurred during the history of these taxa?
(A) 3  (B) 4  (C) 5  (D) 6  (E) 7

(13) The SNPs at which two sites appear to exhibit the most linkage disequilibrium?
(A) 13 & 22  (B) 13 & 54  (C) 22 & 36  (D) 22 & 54  (E) 36 & 54

(14) If the difference between having an adenine or a cytosine at position 28 equates to having either an asparagine (medium sized polar amino acid coded for by the AAC codon) or histidine (large basic amino acid coded for by the CAC codon), which of the following would we use to describe it?
(A) a nonsynonymous polymorphism for a radical amino acid change.
(B) a nonsynonymous substitution for a conservative amino acid change.
(C) a nonsynonymous substitution for a radical amino acid change.
(D) a synonymous polymorphism for a conservative amino acid change.
(E) a synonymous polymorphism for a radical amino acid change.

(15) The proposal that selection does not play a role in most substitutions was called which of the following?
(A) The effective population size  (D) Fisher's fundamental theorem
(B) The Haldane-Muller principle  (E) Wright's adaptive landscape
(C) The neutral theory

(16) Which of the following is equivalent to the narrow sense heritability?
(A) The proportion of overall environmental variance contributed by additive genetic factors.
(B) The proportion of overall environmental variance contributed by dominant genetic factors.
(C) The proportion of overall phenotypic variance contributed by additive genetic factors.
(D) The proportion of overall phenotypic variance contributed by dominant genetic factors.
(E) The proportion of overall phenotypic variance contributed by epistatic genetic factors.