1. You are studying an autosomal recessive trait in the fruit fly *Drosophila* called “vestigial wings,” which is caused by the “vg” allele. A wild-type fly is “vg+” and has long wings. When answering the following parts, show all of your calculations.

A geneticist crosses wild-type males to vestigial-winged females (the P generation) to get the F1 generation. The geneticist then crosses the F1 flies together to get the F2 generation. The scientist removes the vestigial-winged flies from the F2 generation, and breeds the remaining flies together to get the F3 generation.

**(a) What are the phenotypic frequencies of the two kinds of flies in the F3 generation?**

First we will define our variables:

\[ f(vg+) = p; f(vg) = q \]

**F1 generation:**

Since all individuals are heterozygous \( p = q = 0.5 \), and \( f(vg/vg) = 1 \).

**F2 generation**

One generation of random mating will put the population in H-W equilibrium.

\[ p = 0.5; \quad q = 0.5 \]

\[ f(vg+/vg+) = p^2 = 0.25; \quad f(vg/vg) = q^2 = 0.25; \quad f(vg+/vg) = 2pq = 0.5 \]

At this point \( vg/vg \) flies are removed and we recalculate genotypic frequencies

\[ f(vg+/vg+) = (0.25/0.75) = 1/3; \quad f(vg/vg) = 0; \quad f(vg+/vg) = (0.5/0.75) = 2/3 \]

\[ p = f(vg+/vg+) + 0.5 \times (vgs/vg + vgs/vg) = 1/3 + 0.5 \times 2/3 = 2/3 \]

\[ q = f(vg/vg) + 0.5 \times f(vg+/vg) = 0 + 0.5 \times 2/3 = 1/3 \] (note that \( p + q = 1 \))

**F3 generation**

Random mating of the remaining F2 flies will yield the following genotypic classes:

\[ f(vg+/vg+) = p^2 = 4/9; \quad f(vg/vg) = 2pq = 4/9; \quad f(vg+/vg) = q^2 = 1/9 \]

Thus wild type flies will make up 8/9 of the population, while flies with vestigial wings will make up 1/9 of the population.
(b) The scientist then removes the vestigial-winged flies from the F3 generation, and breeds the remaining flies together to get the F4 generation. What are the allele frequencies of the two kinds of alleles in the F4 generation?

From part a) we know that before selection the genotypic ratios are the following:
f (vg+/vg+) = 4/9; f (vg+/vg) = 4/9 ; f (vg/vg) = 1/9
When we remove the 1/9 that are vg we must calculate new genotypic and new allele frequencies of generation F3:
f (vg+/vg+) = (4/9)/(8/9) = 1/2; f (vg+/vg) = (4/9)/(8/9) = 1/2 ; f (vg/vg) = 0
p = f (vg+/vg+) + 0.5 f (vg+/vg) = 0.5 + 0.5x0.5 = 0.75
q = f (vg/vg) + 0.5 f (vg+/vg) = 0 + 0.5X0.5 = 0.25
Random mating of the remaining F3 flies to produce the F4 flies will not change the allele frequencies across generations:
p = 0.75; q = 0.25

(c) The scientist removes the vestigial-winged flies from the F4 generation, and breeds the remaining flies together to get the F5 generation. What are the genotype frequencies of the three kinds of genotypes in the F5 generation?

From part b) we know that allele frequencies of the F3 generation after selection are:
p = 0.75; q = 0.25
Random mating of the F3 flies will give the following F4 genotypic classes:
f (vg+/vg+) = p² = 9/16; f (vg+/vg) = 2pq = 6/16; f (vg/vg) = q² = 1/16
Removal of the vestigial winged flies from the F4 will change the allele and genotypic frequencies:
f (vg+/vg+) = (9/16)/(15/16) = 9/15; f (vg+/vg) = (6/16)/(15/16) = 6/15;
f (vg/vg) = 0
p = f (vg+/vg+) + 0.5 f (vg+/vg) = 9/15 + 0.5(6/15) = 4/5
q = f (vg/vg) + 0.5 f (vg+/vg) = 0 + 0.5X6/15 = 1/5
Random mating of the remaining F4 flies will yield the following genotypes in the F5:
f (vg+/vg+) = p² = 16/25; f (vg+/vg) = 2pq = 8/25; f (vg/vg) = q² = 1/25

(d) Is the fly population described in this problem at Hardy-Weinberg equilibrium? Justify your answer in one sentence.

No, Hardy-Weinberg equilibrium assumes that there is no selection; however in these experiments vg/vg flies are being removed from the population before they mate. Thus selection against vg/vg flies is occurring with S=1.
2. Consider a small village in which 100 individuals are of blood group “M,” 1300 individuals are of blood group “N,” and 100 individuals are of blood group “MN.” When answering the following parts, show all of your calculations.

(a) Is this small village at Hardy-Weinberg equilibrium? Justify your answer.

First calculate genotype frequencies:

\[ f (m/m) = \frac{100}{1500} = 0.067 \]
\[ f (m/n) = \frac{100}{1500} = 0.067 \]
\[ f (n/n) = \frac{1300}{1500} = 0.867 \]

Next calculate allele frequencies:

\[ f (n) = p = f (n/n) + 0.5 f (m/n) = \frac{13}{15} + 0.5\left( \frac{1}{15} \right) = 0.9 \]
\[ f (m) = q = f (m/m) + 0.5 f (m/n) = \frac{1}{15} + 0.5\left( \frac{1}{15} \right) = 0.1 \]

Next calculate expected genotype frequencies:

\[ p^2 = 0.81; \quad 2pq = 0.18; \quad q^2 = 0.01 \]

The expected genotype frequencies do not match the observed genotype frequencies, thus the population is not in H-W equilibrium.

(b) If the M and N blood groups were not co-dominant, would it be possible to determine whether the population is in Hardy-Weinberg equilibrium? Explain why or why not?

No, because you can’t differentiate between homozygotes and heterozygotes.

(c) After one generation of random mating in this village, if the new generation has 2000 people in it, how many will be of each blood type (“M,” “N,” and “MN”)?

Random mating in the population will put the population back in HW equilibrium:

The expected genotypic frequencies for a population with the calculated allele frequencies were calculated in 2a.

\[ M = 0.01 \times 2000 = 20; \quad N = 0.81 \times 2000 = 1620; \quad MN = 0.18 \times 2000 = 360 \]

(d) In the generation of 2000 people from part (c), how many copies of each allele (“\(L^M\)” and “\(L^N\)” will be present in the population?

There are 4000 alleles in the population (2xN):

\[ f (L^M) = 0.1 \times 4000 = 400 \]
\[ f (L^N) = 0.9 \times 4000 = 3600 \]
(e) Fill in the following chart to show the number of each type of couple in the generation from part (c), where 2000 people = 1000 couples. Assume that each person mates with one (and only one) other person, and mating choice is random with respect to blood group.

<table>
<thead>
<tr>
<th>One parent</th>
<th>Other parent</th>
<th>Number of this sort of couple in the generation of 1000 couples from part (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L^M L^M + L^M L^M</td>
<td></td>
<td>1000 x 0.01^2 = 0.1 ≈ 0</td>
</tr>
<tr>
<td>L^M L^M + L^M L^N</td>
<td></td>
<td>2 x 1000 x 0.01 x 0.18 = 3.6 ≈ 4</td>
</tr>
<tr>
<td>L^M L^N + L^N L^N</td>
<td></td>
<td>2 x 1000 x 0.18 x 0.81 = 291.6 ≈ 292</td>
</tr>
<tr>
<td>L^M L^N + L^M L^N</td>
<td></td>
<td>1000 x 0.18^2 = 32.4 ≈ 32</td>
</tr>
<tr>
<td>L^M L^N + L^N L^N</td>
<td></td>
<td>2 x 1000 x 0.01 x 0.81 = 16.2 ≈ 16</td>
</tr>
<tr>
<td>L^N L^N + L^N L^N</td>
<td></td>
<td>1000 x 0.81^2 = 656.1 ≈ 656</td>
</tr>
</tbody>
</table>

There are 1000 total couples in the population.

For random mating the probability of such a pairing is simply equal to the product of the frequencies of the two individuals. We will assume that the allele frequencies are the same in the population of males and in the population of females.

For example the number of L^M L^M + L^M L^M couples:

The frequency of L^M L^M individuals is:

f (L^M L^M) = 0.01 (from part a)

The probability of two such individuals mating is the product of the genotypic frequencies of the two individuals:

f (L^M L^M x L^M L^M) = f (L^M L^M) x F(L^M L^M) = 0.01 x 0.01 = .0001

The expected number of couples is thus:

Num = 1000 x 0.0001 = 0.1

For couples comprised of individuals of two different genotypes there are two possible combinations of individuals for each such couple. For instance, for L^M L^M + L^M L^N couples, dad could be M and mom could be MN, or mom could be M and dad could be MN. Thus a factor of 2 is used. For example the number of L^M L^M + L^M L^N couples:

The frequency of individuals is:

2 x #couples x f (L^M L^M) x (L^N L^N)
2 x 1000 x 0.01 x 0.18 = 3.6 ≈ 4
(f) Imagine a gene in the human population that has a mutation rate \( = 10^{-5} \) per generation, and the mutant form of the gene has a selective disadvantage, \( S = 0.1 \).

- If the mutation is autosomal recessive, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

  1) From lecture at steady state, for an autosomal recessive disorder
     \[ q^2 = m/S = 10^{-5}/0.1 \]
     \[ q = 10^{-2} \]
     For an autosomal recessive disorder, the frequency of affected individuals = \( f(aa) = q^2 = 10^{-4} \)

  2) The frequency of alleles added each generation by mutation is \( 10^{-5} \).
     Thus the fraction of alleles due to new mutations is:
     \[ \text{fraction new} = \frac{\text{frequency new alleles}}{\text{steady state frequency}} = \frac{m}{q} \]
     \[ \text{fraction new} = 10^{-5}/10^{-2} = 10^{-3} \]

- If the mutation is autosomal dominant, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

  1) From lecture at steady state, for an autosomal dominant disorder
     \[ q = m/S = 10^{-5}/0.1 \]
     \[ q = 10^{-4} \]
     For an autosomal dominant disorder, the frequency of affected individuals = \( f(Aa) + f(AA) = 2pq + q^2 = 2 \times 10^{-4} \)

  2) The frequency of alleles added each generation by mutation is \( 10^{-5} \).
     \[ \text{fraction new} = 10^{-5}/10^{-4} = 10^{-1} \]

- If the mutation is X-linked recessive, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

  1) From lecture at steady state, for an X-linked recessive disorder
     \[ q = 3m/S = 3 \times 10^{-5}/0.1 \]
     \[ q = 3 \times 10^{-4} \]
     For an X-linked recessive disorder, the frequency of affected individuals = \( 1/2 f(\text{affected males}) + 1/2 f(\text{affected females}) \)
     \[ = 1/2 f(X^aX^a) + 1/2 f(X^aY) = 1/2q^2 + 1/2q = 1.5 \times 10^{-4} \]

  2) The frequency of alleles added each generation by mutation is \( 10^{-5} \).
     \[ \text{fraction new} = 10^{-5}/3 \times 10^{-4} = 10^{-1}/3 \]
3. The following is a pedigree for an autosomal dominant trait. This trait is caused by the Q allele. Thus individual 1 is “qq” and individual 8 is “Qq”. You want to map the locus responsible for the trait with respect to two SSRs.

**PLEASE NOTE:** For individuals 2-7, only the alleles inherited from the father are shown. You will have to determine the alleles inherited from the mother later in the problem. Both the maternal and paternal alleles of SSR 25 and 37 are given for individuals 1, 8, and 9.

![Pedigree Diagram]

<table>
<thead>
<tr>
<th>Individual</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>paternally inherited allele at the Q locus</td>
<td>q</td>
<td>q</td>
<td>q</td>
<td>q</td>
<td>q</td>
<td>q</td>
</tr>
<tr>
<td>paternally inherited allele at SSR 25</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>paternally inherited allele at SSR 37</td>
<td>H</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>maternally inherited allele at the Q locus</td>
<td>Q</td>
<td>q</td>
<td>q</td>
<td>q</td>
<td>Q</td>
<td>q</td>
</tr>
<tr>
<td>maternally inherited allele at SSR 25</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>A</td>
<td>D</td>
</tr>
<tr>
<td>maternally inherited allele at SSR 37</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>F</td>
<td>E</td>
</tr>
</tbody>
</table>

A allele
B allele
C allele
D allele

E allele
F allele
G allele
H allele

Schematic of electrophoretic gel:

```
IIIIII

IIIIII

IIIIII
```

<table>
<thead>
<tr>
<th>SSR 25</th>
<th>SSR 37</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A allele</th>
<th>B allele</th>
<th>C allele</th>
<th>D allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>E allele</td>
<td>F allele</td>
<td>G allele</td>
<td>H allele</td>
</tr>
</tbody>
</table>

(a) to be filled in for part (a)

c) to be filled in for part (c)

g) to be filled in for part (g)
(a) Fill in the chart on page 4 with the paternally inherited alleles at all three loci.

The father is qq, thus he must give each of his children a q allele. For SSR25 he is heterozygous with a B-C genotype; the SSR25 allele given to each child can be determined from the gel. For SSR37 he is heterozygous with a G-H genotype; the SSR25 allele given to each child can be determined from the gel.

(b) Draw all phases of individual 8 with respect to the Q locus and SSR 25 that are possible given everything you know about her. Make sure to draw phases using the proper notation:

Based on the genotype of individual 8 (Q/q; A/D) there are two possible phases:

<table>
<thead>
<tr>
<th>Phase</th>
<th>Q or q</th>
<th>A, B, C, or D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Q</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>D</td>
</tr>
<tr>
<td>2</td>
<td>Q</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>q</td>
<td>A</td>
</tr>
</tbody>
</table>

However, we know the genotype of individual 9 (q/q; D/D). Individual 8 must have inherited a q-D chromosome from individual 9 and a Q-A chromosome from her father. **Thus the correct phase is phase 1.**

(c) A human geneticist now tells you that the total probability of seeing the maternally inherited alleles of the Q and SSR 25 loci in individuals 2-7 equals \((0.5)^5 \times (0.0)^0\), given that the Q and SSR 25 loci are linked at \(\theta = 0\). Given this information, fill in chart above and draw in the maternally inherited bands in the above gel. (You have the SSR data for all the children, so fill all columns in.)

For the Q locus we know that affected individuals must have received a Q allele from their mother since they can only inherit a q allele from their father. Unaffected individuals must be qq, therefore they must have inherited a q from their mother.

If the Q locus and SSR25 are linked at \(\theta = 0\) then no recombination between the two loci will take place and the maternal allele combinations (QA & qD) cannot be shuffled. Thus, all affected individuals will inherit allele A and all normal individuals will inherit allele D.
(d) Calculate the LOD score at $\theta = 0$ for SSR 25 and the Q locus. Show all calculations.

$$\text{LOD}_{\theta = 0} = \log \frac{p \text{ (data having arisen given that SSR25 and Q are linked at } \theta = 0)}{p \text{ (data having arisen given that SSR25 and Q are not linked)}}$$

$P \text{ data having arisen if SSR25 and Q are unlinked} = 0.25^5$

Probability of getting any combination of alleles by chance $= 0.25$, for five individuals this is $0.25^5$. We’re ignoring one because of the twins.

$P \text{ data having arisen given linked at } 0 = 0.5^5 x 0^0$

Probability of getting either QA or qD is 0.5 since the maternal allele combinations aren’t shuffled. Probability of getting a recombinant is 0.

We have five parental types and no recombinants, therefore the probability of seeing these combinations is $0.5^5 x 0^0$.

$$\text{LOD}_{\theta = 0} = \log \left( \frac{0.5^5 x 0^0}{0.25^5} \right) = 1.5$$

(e) Individuals 1 and 8 had six children and the human geneticist has taken the data for all six into account correctly in part (b). Knowing this, give a brief explanation for why the human geneticist only included five meioses in the calculation of the LOD score shown above. Assume NO new mutations. Go back and make sure your data table and gel are consistent with your explanation.

One possibility is that we are only looking at five meiotic events rather than six. This could happen if individuals 3 and 4 are identical twins, thus both were produced by a single meiosis.

(f) You now switch to mapping the two SSRs with respect to each other. Given everything you know about the mother (individual 8), draw all possible phases of the mother with respect to SSR 37 and SSR 25. Make sure to draw the phases using the proper notation.

1) A | E
   D | F

2) A | F
   D | E

Based on the genotype of individual 8 (A/D; E/F) there are two possible phases:
However, we know the genotype of individual 9 (D/D; E/H). Individual 8 must have inherited a D-E chromosome from individual 9 and an A-F chromosome from her father.

Thus the correct phase is phase 2
(g) The human geneticist now tells you that the total probability of seeing the maternally inherited alleles of the SSR 25 and SSR 37 loci in individuals 2-7 equals \((0.45)^4 \times (0.05)^1\), given that the two SSRs are linked at \(\theta = 0.1\). Knowing this, draw in the maternally inherited bands of SSR 37 in the above gel. (You have the SSR data for all the children, so fill all columns in. Make sure you fill in the gel in a way that is consistent with the fact that the human geneticist only included 5 meioses in the calculation.)

One of the meioses produced a recombinant gamete (i.e. A-E or D-F). We have no way of knowing which child is recombinant \(a\ priori\), so we will just randomly assign one as a recombinant (the chart is filled out as if 2 is a recombinant). Thus there are five equally good solutions to this part.

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 recombinant:</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>F</td>
<td>E</td>
</tr>
<tr>
<td>3 &amp; 4 recombinant:</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>E</td>
<td>F</td>
<td>E</td>
</tr>
<tr>
<td>5 recombinant:</td>
<td>F</td>
<td>E</td>
<td>E</td>
<td>F</td>
<td>F</td>
<td>E</td>
</tr>
<tr>
<td>6 recombinant:</td>
<td>F</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>7 recombinant:</td>
<td>F</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

The bands are drawn on the gel consistent with the table (2 is a recombinant).

(h) Given everything you know about the father (individual 1), draw all possible phases of the father with respect to SSR 37 and SSR 25. Make sure to draw the phases using the proper notation.

1) \[ B \quad G \]
    \[ C \quad H \]

2) \[ B \quad H \]
    \[ C \quad G \]

Based on the genotype of individual 1 (B/C; G/H) there are two possible phases. Since we don’t have any information about the configuration of alleles in the parents of individual 1 we cannot eliminate one of the phases.
(i) Calculate the LOD score at $\theta = 0.1$ for the two SSRs. Make sure to incorporate the data from both the mother and the father, because both are informative parents regarding the inheritance of alleles at these two loci. Show all calculations.

We are looking at meioses in two parents, the mother and the father, so we will calculate two separate LOD scores that will be additive to give us a final LOD score.

Maternal alleles:

P data having arisen because the two SSRs are unlinked $= 0.25^5$

Probability of getting any combination of alleles by chance $= 0.25$, for five individuals this is $0.25^5$. We’re ignoring one because of the twins.

P data having arisen given that the two SSRS are linked at $\theta = 0.1$

$= 0.45^4 \times 0.05^1$

Given that the alleles are linked at $\theta = 0.1$, the probability of getting a parental type is 0.9 and the probability of getting a recombinant type is 0.1. Thus the probability of getting A-F is 0.45, and 0.45 for D-E. The probability of getting A-E is 0.05, and 0.0.5 for D-F. We see 1 recombinant and 4 parentals therefore the maternal probability is $(0.45)^4 \times (0.05)^1$.

$\text{LOD}_{\theta = 0.1} = \log \left( \frac{0.45^4 \times 0.05^1}{0.25^5} \right) = 0.322$

Paternal alleles:

P data having arisen because the two SSRs are unlinked $= 0.25^5$

P data having arisen given that the two SSRS are linked at $\theta = 0.1$

$= 1/2(0.45^3 \times 0.05^2) + 1/2 (0.45^2 \times 0.05^3)$

We can’t determine the phase for the father thus we must take both into account. A priori the probability of each phase is 1/2. For phase one there are 2 recombinants and 3 parentals, for phase two there are 3 recombinants and 2 parentals.

$\text{LOD}_{\theta = 0.1} = \log \left( \frac{1/2(0.45^3 \times 0.05^2) + 1/2 (0.45^2 \times 0.05^3)}{0.25^5} \right) = -0.887$

To calculate the total LOD score, sum the LOD score from the father and the LOD score from the mother.

$\text{LOD}_{\theta = 0.1} = \text{LOD}_{\theta = 0.1} \text{ maternal} + \text{LOD}_{\theta = 0.1} \text{ paternal} = -0.565$