Supplemental Material

CBE—Life Sciences Education

Williams *et al*.

SUPPLEMENTAL MATERIALS

PRE-ASSESSMENT AT THE BEGINNING OF THE SEMESTER

Gender

- What is your gender? Male/Female/Other

Scientific Reasoning

- **Lawson Classroom Test of Scientific Reasoning** (LCTSR; 2000 version with 24 items, including 4 items aimed at post-formal reasoning; 24-point scoring was used).
    - Lawson AE. 1978. The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*. 15(1):11-24.
    - Lawson AE, Alkhoury S, Benford R, Clark BR, Falconer KA. 2000. What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *Journal of Research in Science Teaching*. 37(9):996-1018.

PRE-ASSESSMENT BEFORE DAY 1 OF POPULATION GENETICS INSTRUCTION

Math Anxiety (AMAS)

Notes:

- Source for questions and validity/reliability analyses in college students: Hopko, D.R., Mahadevan, R., Bare, R.L. and Hunt, M.K. 2003. The abbreviated math anxiety scale (AMAS) construction, validity, and reliability. *Assessment*, 10(2), pp.178-182.
- Answers for all questions were summed to give a final math anxiety score between 9 – 45.
- For some analyses, we categorized students into three categories: low, moderate, or high anxiety. We defined "low anxiety" as reporting an anxiety score of 18 or lower (equivalent to marking low or some anxiety on all items), "moderate anxiety" for scores of 19-27 (e.g. marking some or moderate anxiety on all items), and "high anxiety" as an anxiety score above 27 (e.g. moderate anxiety or higher on all items). The highest score we obtained was 36.

<u>*Instructions given to Students:*</u>

Please rate each item below in terms of how anxious you would feel during the event specified. (Answer Options: 1 - Low Anxiety, 2 - Some Anxiety, 3 - Moderate Anxiety, 4 - Quite a bit of Anxiety, 5 - High Anxiety).

- Having to use the tables in the back of a mathematics book. _____
- Thinking about an upcoming mathematics test one day before. _____
- Watching a teacher work an algebraic equation on the blackboard. _____
- Taking an examination in a mathematics course. _____
- Being given a homework assignment of many difficult problems due the next class meeting. _____
- Listening to a lecture in mathematics class. _____
- Listening to another student explain a mathematics formula. _____
- Being given a "pop" quiz in a mathematics class. _____
- Starting a new chapter in a mathematics book. _____

Math Skills

Notes:

- We designed the short math skills assessment to target specific mathematical abilities that directly pertain to learning population genetics, HW equilibrium, and HW equations. Justification of validity follows each question.
- Questions were open response.
- Number of correct responses (six maximum) was used as final measure.

<u>*Instructions given to Students:*</u>

For the following problems, you may not use a calculator. Please show your work:

1. If you have one six-sided die, what's the probability you will roll a 3? You may answer using a fraction or a decimal. **1/6 or 0.167**
    a. **Justification:** Because population genetics deals fundamentally with probabilistic outcomes, this basic first question simply tests whether students can identify a probability in terms of the number of events satisfying the outcome as compared to the total number of events possible.

2. If you have two identical six-sided dice and you roll them at the same time, what's the probability that you will roll two 6's? You may answer using a fraction or a decimal. **1/36 or 0.0278**

a. **Justification:** Random mating involves probabilities coming from two different sources (each parent). Thus, a key ability in understanding population genetics and HW equilibrium is to understand how two independent events function in tandem. From a mathematical skills perspective, this involves being able to understand how to take a probability for an outcome from each independent event and multiply them to produce a probability for both outcomes occurring simultaneously (Batanero & Sanchez, 2005; Shaughnessy & Ciancetta, 2002). This question, in particular, focuses on the idea of the same outcome happening for both events (e.g. each parent passing on an identical allele), represented by both dice being a 6.

3. If you have two identical six-sided dice and you roll them at the same time, what's the probability that you will roll a 6 and a 3? You may answer using a fraction or a decimal. **1/18 or 2/36 or 0.0556**
   a. **Justification:** In random mating, having the same outcome, or gamete type, from both parents is not guaranteed. Heterozygous offspring are the result of getting different outcomes, or gamete types, from each parent. Thus, heterozygous individuals can result from two separate overall outcomes (e.g., getting a dominant allele from the first parent and a recessive allele from the other, or getting a recessive allele from the first parent and a dominant allele from the other). This represents an addition mathematical skill, in recognizing that different outcomes from the independent events results in two distinct overall outcomes. However, it is known that students often fail to account for this, believing in this case that two 6's are equally likely as a 6 and a 3 (Marie-Paule Lecoutre, 1992; M-P Lecoutre & Fischbein, 1998). This question tests whether students understand the probability rules that requiring adding the separate probabilities from the two distinct outcomes together.

4. If $3x + y = 9$ and $x + y = 5$, what do x and y equal? **x=2, y=3**
   a. **Justification:** HW equilibrium includes a system of two equations that must simultaneously be true. In order to solve for various values within these equations, the students need to be fluid at solving systems of equations (for more on solving equations, see (Oktaç, 2018; Pangaribuan, 2018; Panizza, Sadovsky, & Sessa, 1999). This question assesses whether the students are able to do so for a simple pair of equations, with one equation (x+y=5) intentionally bearing close resemblance to the HW equation p+q=1.

5. If $3x – 2xy + y^2 = 3$ and $y = 2$, what does $3x – 2xy$ equal? **-1**
   a. **Justification:** Typical algebra classes teach students to solve for the values of variables in given equations. Yet, HW equations often require students to conceptualize an entire term, or group of terms, as an entity until itself. This skill is not guaranteed simply by being able to solve mathematical equations, as in the previous question (Sfard, 1991; Sfard & Linchevski, 1994). Thus, this question fundamentally determines whether students are able to conceptualize a group of terms as a single entity.

6. If $(x+1)^2 = 49$, what is x? **x=6 and x= -8**
   a. **Justification:** In HW equations, it is possible to know the probability of the overall outcome in random mating (e.g., an outcome of "pp" being 49% likely), without having been told the original probability of p. Students need to be able to solve for probabilities based on knowing the "square" of that probability. This question assesses students' ability to solve when the square is known.

REFERENCES

Batanero, C., & Sanchez, E. (2005). What is the Nature of High School Students' Conceptions and Misconceptions About Probability? In *Exploring probability in school* (pp. 241-266): Springer.

Lecoutre, M.-P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational studies in mathematics, 23*(6), 557-568.

Lecoutre, M.-P., & Fischbein, E. (1998). Évolution avec l'âge de «misconceptions» dans les intuitions probabilistes en France et en Israël. *Recherches en didactique des mathématiques (Revue), 18*(3), 311-331.

Oktaç, A. (2018). Conceptions about system of linear equations and solution. In *Challenges and strategies in teaching linear algebra* (pp. 71-101): Springer.
Pangaribuan, F. (2018). Students' abstraction in solving system of linear equations with two variables. *JPhCS, 1088*(1), 012071.
Panizza, M., Sadovsky, P., & Sessa, C. (1999). La ecuación lineal con dos variables: entre la unicidad y el infinito. *Enseņanza de las ciencias, 17*(3), 453-461.
Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational studies in mathematics, 22*(1), 1-36.
Sfard, A., & Linchevski, L. (1994). Between arithmetic and algebra: In the search of a missing link. The case of equations and inequalities. *Rendiconti del Seminario Matematico Università e Politecnico di Torino, 52*(3), 279-307.
Shaughnessy, J., & Ciancetta, M. (2002). *Students' understanding of variability in a probability environment.* Paper presented at the Proceedings of the Sixth International Conference on Teaching Statistics: Developing a Statistically Literate Society, Cape Town, South Africa.

Instructor Guide for Day 1 (Equations 1st section)

Deriving and using the Hardy-Weinberg Equations

Clicker Question to start class (review probability for equation derivation):

Let's see how well you remember probability…
If you roll 2 dice at once, what is the probability you will roll a 1 and a 4?

A. 1/3
B. 2/3
C. 1/9
D. 1/18
E. 1/36

First have students try on their own.

Short class discussion (ideas generated by students as much as possible):

- *In probability, what mathematical function should be used for an "and" in the scenario? (multiplication, since creating more requirements makes the probability more restricted/decrease)*
- *In probability, what mathematical function should be used for an "or" in the scenario? (addition, since giving more possibilities makes the probability go up/increase)*
- *How many ways are there to roll a 1 and a 4 with two dice? How does this affect your scenario? (two ways, since it doesn't matter which number is on which die)*

Direct students to work with a neighbor on the problem and change their answer if needed.

Work through the problem on the board together and reveal the correct answer: 1/18

Introduction to Hardy-Weinberg Equilibrium

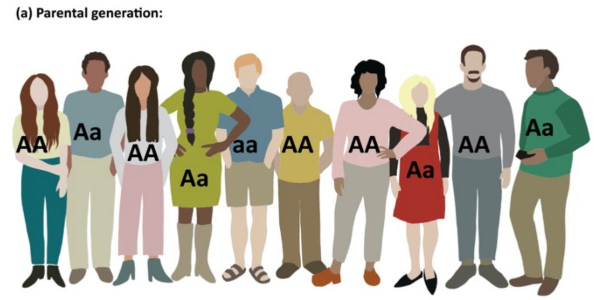
Give definition: Allele frequencies will not change in a population in the absence of evolutionary influences

Since we have already covered mechanisms of evolution, ask students to brainstorm what conditions Hardy-Weinberg Equilibrium assumes are true. If there is no evolution, there must be no… *(should be generated by students and written on the board)*

- *selection*
- *nonrandom mating*
- *gene flow*
- *genetic drift*
- *mutations*

Explain the following: RARELY are populations in perfect equilibrium (usually some type of evolution is going on), so HW model is an excellent control to compare our populations to! When we see HOW they deviate from HW, we can get an idea of what assumption is being violated

Deriving the Hardy-Weinberg Equations



Given to students: Let's consider a human population in Hardy-Weinberg Equilibrium. For a specific gene in a human population, the frequency of the dominant allele (A) is 0.7 and we will call this frequency 'p.' The frequency of the recessive allele (a) is 0.3 and we will call this frequency 'q.' What do you notice about 'p' and 'q?'

Students should come up with the first Hardy-Weinberg equation: $p + q = 1$

Why is that true? Will that always be true?

Students should come up with the idea that there are only 2 alleles in this system, so it will always add up to 1 as long as there are only 2 alleles possible.

Let's assume this population reproduces to form the next generation. Now let's calculate the PREDICTED number (all based on probability) of each genotype in the next generation. Remember, we are assuming mating is completely random! Work with your neighbor:

1. Can you calculate the predicted frequencies of each genotype (AA, Aa, aa) in the next generation?
2. You can use the frequencies of A (0.7) and a (0.3) that we had previously, but you should also try to come up with a generalizable term made up of p's and/or q's.

Students will work together and do as much of this as possible in their groups. If students struggle, offer the following hints: Let's start with the homozygous dominant genotype (AA). How would an offspring with the AA genotype be made? What's the probability of this occurring? What's the probability of getting an A from mom? From Dad? From mom AND dad?

Bring the class together to do AA together.

Students should come up with the idea (as you write on the board) that you get an AA offspring if an "A" sperm fertilizes an "A" egg (an "A" AND an "A").

- *So the probability of an AA offspring is*
    - *(probability of A) x (probability of A)*
    - *(Frequency of A) x (Frequency of A)*
    - $p \times p = p^2 = 0.7^2 =. 0.49$
    - *Thus we expect 49% of the next generation to be AA (p^2)*

Ask the students to return to pairs/groups and similarly predict the frequencies of Aa and aa offspring.

Students should come up with the idea (as you write on the board) that you get an Aa offspring if an "A" sperm fertilizes an "a" egg (an "A" AND an "a") or if an "a" sperm fertilizes an "A" egg (OR an "a" AND an "A")

- *So the probability of an Aa offspring is*
    - *(probability of A x probability of a) + (probability of a x probability of A)*
    - $(p \times q) + (p \times q) = 2pq = 2(0.7)(0.3) = 0.42$
    - *Thus we expect 42% of the next generation to be Aa (2pq)*

Students should come up with the idea (as you write on the board) that you get an aa offspring if an "a" sperm fertilizes an "a" egg (an "a" AND an "a").

- *So the probability of an aa offspring is*
    - *(probability of a) x (probability of a)*
    - *(Frequency of a) x (Frequency of a)*
    - *$q \times q = q^2 = 0.3^2 = 0.09$*
    - *Thus we expect 9% of the next generation to be aa (q^2)*

Ask students what they notice about the frequencies of AA, Aa, and aa? (in other words, what do they notice about p^2, 2pq, and q^2?)

Students should come up the second Hardy-Weinberg equation: $p^2 + 2pq + q^2 = 1$

Why is that true? Will that always be true?

Students should come up with the idea that these are the only 3 genotype frequencies possible, so it will always add up to 1 as long as there are only 2 alleles possible.

So now we have our 2nd generation: 49% AA, 42% Aa, 9% aa

Ask the students to figure out the allele frequencies of this second generation. How could we calculate the fraction "A" vs "a"? If students struggle, propose they pretend the population contains 100 people.

Students should be to come up with the following calculations (as you write on the board):

- *49 AA people= 98 A's*
- *42 Aa people= 42 A's and 42 a's*
- *9 aa people= 18 a's*
- *TOTALS: 140 A's 60 a's*
- *Frequencies (just divide # of alleles by TOTAL number of alleles, which is 200)*
- *Frequency of A: **p** = 140/200 = 0.7*
- *Frequency of a: **q** = 60/200 = 0.3*

Ask students to compare this to the original population. What do they notice? Are they surprised?

Students should notice that these are the same allele frequencies as generation 1. This shouldn't be surprising, because we followed all the assumptions that would suggest there is no evolution (mating is random, no one moved in, etc.)

Review on the board the 2 Hardy-Weinberg equations, and give students a minute to practice telling their neighbor what each variable (p, q, p^2, 2pq, q^2) represents biologically.

Formative Assessment clicker question (how to calculate allele frequencies from genotype counts)

Please calculate the allele frequencies for a population with **32 AA, 96 Aa, and 72 aa flowers.**

A. p=0.4, q=0.6
B. p=0.16, q=0.36
C. p=0.32, q=0.72
D. p=160, q=240

Have students try on their own first, then let them work with a neighbor.

Students should come up with the following (write on the board as they explain)

- *32 AA plants = 64 A's*
- *96 Aa plants = 96 A's and 96 a's*
- *72 aa plants = 144 a's*
- *TOTALS: 160 A's, 240 a's*
- *Frequencies (just divide # of alleles by TOTAL, which is 400)*
- *Frequency of A: **p** = 160/400 = 0.4; Frequency of a: **q** = 240/400 = 0.6*

Formative Assessment (predicting next generation genotypes if in Hardy-Weinberg Equilibrium)

A plant population growing at 2000 feet on the east-facing slope of the Sierra Nevada Mountain range. There is a gene that determines flower color. The dominant allele (A) gives a yellow color, and the recessive allele (a) gives a white color. This plant population has 200 plants in it. We analyzed the genome of each one and found out there are **32 AA, 96 Aa, and 72 aa**. (Frequency of A = 0.4 and frequency of a = 0.6 from last problem). If this population is in Hardy-Weinberg Equilibrium, predict the genotype frequencies of the next generation.

After students try, guide them through the following solution on the board guided by their ideas:

- *We have ALLELE frequencies in this population of plants (p and q). p = 0.4 and q = 0.6*
- *Let's assume they reproduce. Now let's calculate the PREDICTED number (all based on probability) of each genotype in the next generation. Remember, we are assuming mating is completely random!*
- *What's the probability of a plant in the next generation being AA?*
  - *What's the probability this plant will get an "A" from 1st parent AND an "A" from the 2nd parent?*
  - *(probability of A) x (probability of A)*
  - *p (frequency of A allele) x p (frequency of A allele) = p*
  - *$p^2 = 0.4^2 = 0.16$*
  - *Thus we expect 16% of the next generation to be AA*
- *What's the probability a plant in the next generation will be Aa?*
  - *Receive "A" allele from 1st parent AND "a" from the 2nd parent*
  - *OR?*
  - *Receive "a" allele from 1st parent AND "A" allele from 2nd parent*
  - *(probability of A x probability of a) + (probability of a x probability of A)*
  - *(p x q) + (p x q) = 2pq*
  - *In this case, 2pq = 2(0.4)(0.6) = 0.48*
  - *Thus we expect 48% of the next generation to be Aa*
- *What's the probability a plant in the next generation will be aa?*
  - *"a" AND "a"*
  - *(probability of a) x (probability of a)*
  - *$q \times q = q^2$*
  - *In this case $q^2 = 0.6^2 = 0.36$*
  - *Thus we expect 36% of the next generation to be aa*

Review and ask for questions (Hardy-Weinberg equations/variables, Hardy-Weinberg assumptions and why the equations work only if those assumptions are met)

Manipulating the Hardy-Weinberg Equations

There is another population of this same plant species that grows at 1000 feet on the WEST-facing slope of the Sierra Nevada Mountain range that is also in Hardy-Weinberg equilibrium. Out of 200 plants, you count 8 with white flowers and 192 with yellow flowers. Can you tell me the frequencies of the yellow and white alleles (p and q)? Can you tell me how many of the 192 yellow flowers are heterozygous?

Let students try, then guide them through the following with their ideas:

- *192 yellow flowers (Aa or AA)*
  - *Since there are two genotype options for this one, we cannot just count alleles to get allele frequencies. We'll have to figure out genotype frequencies and go from there!*
- *8 white flowers (aa) ← best place to start since only one genotype possible!*
- *Total flowers = 200*
- *Frequency of aa flowers = q^2 = 8/200 = 0.04*
- *Now we can solve for q! q = square root of 0.04 = **0.2 = q***
- *Since p + q = 1, 1-q = p = 1 – 0.2 = **0.8 = p***
- *Since the population is in HW equilibrium, we expect the frequency of heterozygotes to be 2pq = 2(0.8)(0.2) = <u>0.32 = 2pq</u>*
- *0.32*200 = **64 heterozygous flowers***

END OF DAY 1

Instructor Guide for Day 1 (Punnett Square 1st section)

Using Population Punnett Squares to Model Hardy-Weinberg Equilibrium

Introduction to Hardy-Weinberg Equilibrium

Give definition: Allele frequencies will not change in a population in the absence of evolutionary influences

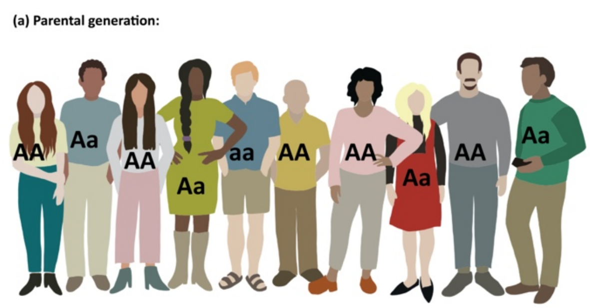
Since we have already covered mechanisms of evolution, ask students to brainstorm what conditions Hardy-Weinberg Equilibrium assumes are true. If there is no evolution, there must be no… *(should be generated by students and written on the board)*

- *selection*
- *nonrandom mating*
- *gene flow*
- *genetic drift*
- *mutations*

Explain the following: RARELY are populations in perfect equilibrium (usually some type of evolution is going on), so HW model is an excellent control to compare our populations to! When we see HOW they deviate from HW, we can get an idea of what assumption is being violated

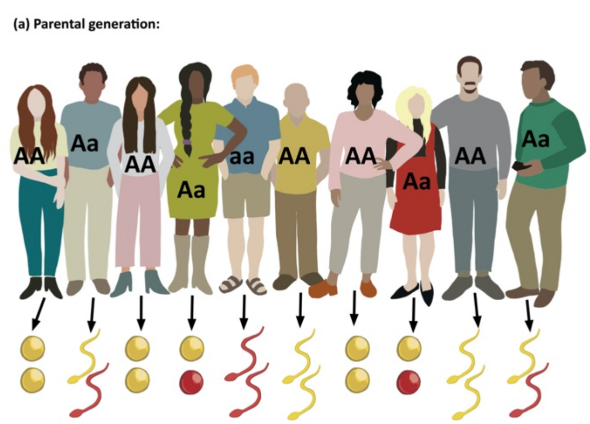
Introducing Population Punnett Squares

Given to students: Let's consider a human population in Hardy-Weinberg Equilibrium, with these 10 individuals representing the whole population (assume in the large population, male/female genotype distributions are the same). Now imagine that we wanted to predict the next generation.



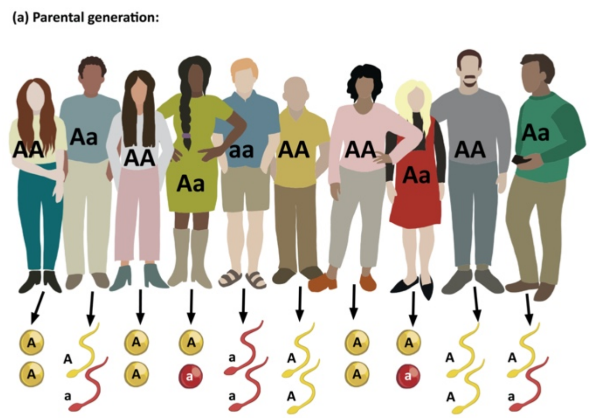
How would this scenario be similar and/or different from predicting the next generation of just a single couple like we did in our genetics unit?

Students could offer the following ideas (guided by you):
Similarities: you would still need to come up with possible gametes and model random fertilization to predict genotypes/phenotypes of offspring
Differences: there are a lot of potential parents and you don't know exactly who will mate with who, so you need to consider all potential mating. If the population is in HW equilibrium, it's mathematically equivalent to putting all the gametes in a bag and randomly choosing a sperm and an egg (random fertilization with the whole population).



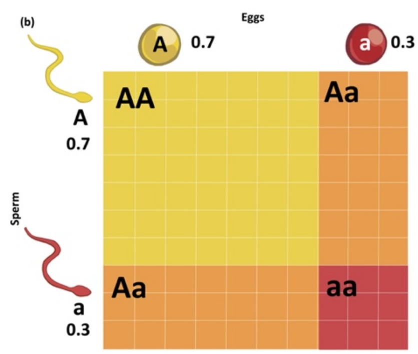
So, we definitely need to know what gametes will be generated. What percentage of the eggs/sperm will have A alleles and what percentage of the eggs/sperm will have a alleles in the population shown here? Remember, the population is much larger (these individuals just represent the overall percentages) and male and female percentages are equal (so you don't have to worry about sperm/egg being different from each other).

Students should come up with the following allele frequencies: Frequency of A = 0.7, frequency of a = 0.3.



So now we're ready to assume this population reproduces to form the next generation. Now let's calculate the PREDICTED number (all based on probability) of each genotype in the next generation. Remember, we are assuming mating is completely random! Talk to your neighbor, if we were going to make a Punnett square to represent the WHOLE population, what would it look like?

Let students talk to each other and then ask them to share. IF THEY STRUGGLE: They will likely draw a Punnett square like they're used to and not know what to do next. You can draw a Punnett square from a double heterozygote cross like they did in genetics and talk about why we have always drawn all of the boxes the same size. Why could you just count boxes to come up with probability of different offspring? Help them realize that in that case, getting an A vs an a was a 50/50 chance, so all of the boxes were equally probably. Help them arrive at the following <u>population</u> Punnett square where A and a are not equally probable. →

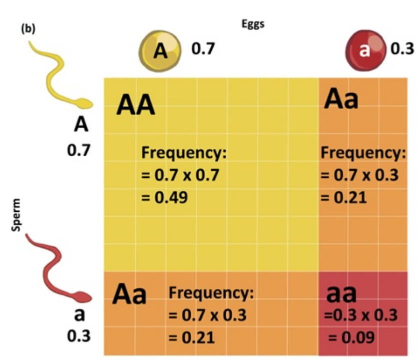


Now let's calculate the PREDICTED number (all based on probability) of each genotype in the next generation. Work with your neighbor, how could we calculate the predicted frequencies of each genotype in the next generation? In other words, what percentage of the offspring do we expect to be AA, Aa, and aa?

Students should come up with following, but if they struggle, you can do AA with them and then let them try the others. As you get student answers, fill in the Punnett square:

You get an AA offspring (top left box) if an "A" sperm fertilizes an "A" egg

- *So the probability of an AA offspring is*
  - *(probability of A) x (probability of A)*
  - *(Frequency of A) x (Frequency of A)*
  - *= 0.7^2 =. 0.49*
  - *Thus we expect 49% of the next generation to be AA*



You can get an Aa offspring (top right box) if an "A" sperm fertilizes an "a" egg

- *So the probability of an Aa offspring in the top right box is*
  - *(probability of A x probability of a)*
  - *(0.7)(0.3) = 0.21*
  - *Thus we expect 21% of the next generation to be in that top right box and be Aa*

You can get an Aa offspring (bottom left box) if an "a" sperm fertilizes an "A" egg

- *So the probability of an Aa offspring in the top right box is*
  - *(probability of a x probability of A)*
  - *(0.7)(0.3) = 0.21*
  - *Thus we expect 21% of the next generation to be in that bottom left box and be Aa*

You get an aa offspring (bottom right box) if an "a" sperm fertilizes an "a" egg

- *So the probability of an aa offspring is*
  - *(probability of a) x (probability of a)*
  - *(Frequency of a) x (Frequency of a)*
  - *$0.3^2 = 0.09$*
  - *Thus we expect 9% of the next generation to be aa*

So now we have our 2nd generation: 49% AA, 42% Aa (21% + 21%), 9% aa

Ask the students to figure out the allele frequencies of this second generation. How could we calculate the fraction "A" vs "a"? If students struggle, propose they pretend the population contains 100 people.

Students should be to come up with the following calculations (as you write on the board):

- *49 AA people= 98 A's*
- *42 Aa people= 42 A's and 42 a's*
- *9 aa people= 18 a's*
- *TOTALS: 140 A's 60 a's*
- *Frequencies (just divide # of alleles by TOTAL number of alleles, which is 200)*
- *Frequency of A: **p** = 140/200 = 0.7*
- *Frequency of a: **q** = 60/200 = 0.3*

Ask students to compare this to the original population. Are they surprised?

Students should notice that these are the same allele frequencies as generation 1. This shouldn't be surprised, because we followed all the assumptions that would suggest there is no evolution (mating is random, no one moved in, etc.)

Formative Assessment clicker question (how to calculate allele frequencies from genotype counts)

Please calculate the allele frequencies for a population with **32 AA, 96 Aa, and 72 aa flowers.**

E. p=0.4, q=0.6
F. p=0.16, q=0.36
G. p=0.32, q=0.72
H. p=160, q=240

Have students try on their own first, then let them work with a neighbor.

Students should come up with the following (write on the board as they explain)

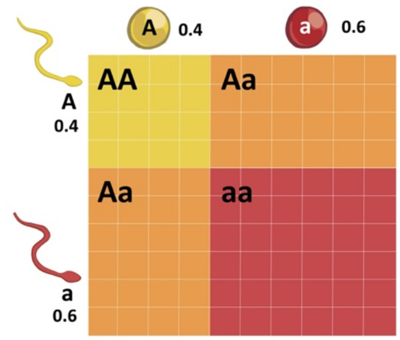
- *32 AA plants = 64 A's*
- *96 Aa plants = 96 A's and 96 a's*
- *72 aa plants = 144 a's*
- *TOTALS: 160 A's, 240 a's*
- *Frequencies (just divide # of alleles by TOTAL, which is 400)*
- *Frequency of A = 160/400 = 0.4*
- *Frequency of a = 240/400 = 0.6*

Formative Assessment (predicting next generation genotypes if in Hardy-Weinberg Equilibrium)

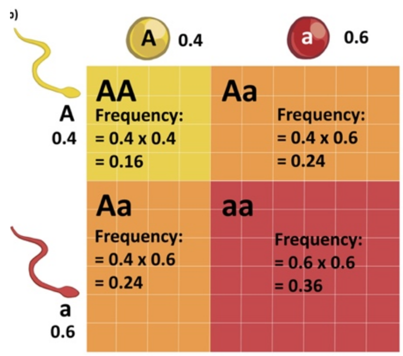
A plant population growing at 2000 feet on the east-facing slope of the Sierra Nevada Mountain range. There is a gene that determines flower color. The dominant allele (A) gives a yellow color, and the recessive allele (a) gives a white color. This plant population has 200 plants in it. We analyzed the genome of each one and found out there are **32 AA, 96 Aa, and 72 aa**. (Frequency of A = 0.4 and frequency of a = 0.6). If this population is in Hardy-Weinberg Equilibrium, predict the genotype frequencies of the next generation.

After students try on their own and compare with neighbors, guide them through the following solution on the board guided by their ideas:

- *First, we can draw a population Punnett square with 0.4 A and 0.6 a. →*



- *We can then predict the frequency of offspring in each box by multiplying the frequency of the corresponding gamete on the side to get this →*
- *Thus we would predict 16% AA, 48% Aa (24% + 24%), and 36% aa in the next generation*



Review and ask for questions (population Punnett squares, Hardy-Weinberg assumptions and why these calculations only work if those assumptions are met)

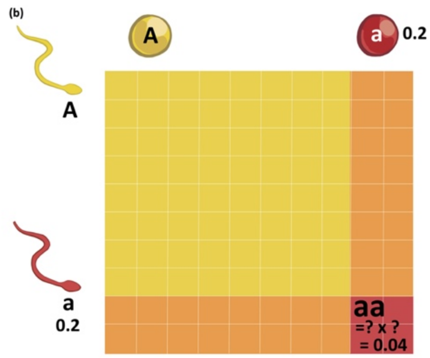
Manipulating the Hardy-Weinberg Equations

There is another population of this same plant species that grows at 1000 feet on the WEST-facing slope of the Sierra Nevada Mountain range that is also in Hardy-Weinberg equilibrium. Out of 200 plants, you count 8 with white flowers and 192 with yellow flowers. Can you tell me the frequencies of the yellow and white alleles? Can you tell me how many of the 192 yellow flowers are heterozygous?

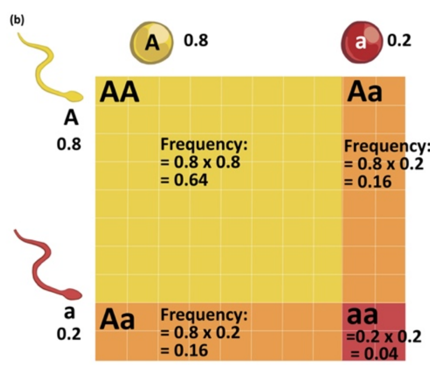
Let students try, then guide them through the following with their ideas:

- *Total flowers = 200*
- *192 yellow flowers (Aa or AA)*
  - *Since there are two genotype options for this one, we cannot just count alleles to get allele frequencies. We'll have to figure out genotype frequencies and go from there!*
- *8 white flowers (aa) ← best place to start since only one genotype possible!*

- *Frequency of aa flowers (BOTTOM RIGHT BOX) = 8/200 = 0.04 →*
- *Now we can figure out the frequency of the a allele. We know it's the frequency of the a allele times itself that got us 0.04. To go backwards = square root of 0.04 = **0.2***



- *Since there are only 2 alleles (A and a), what's the frequency of A? 1 – 0.2 = **0.8***
- *If we complete the square, we expect the frequency of heterozygotes **= 0.32 (0.16 + 0.16)***
- *0.32*200 flowers = **<u>64 heterozygous flowers</u>***



END OF DAY 1

MID-ASSESSMENT (with work type examples)

After Day 1 of population genetics instruction

Hardy-Weinberg Calculations

Instructions:

For the following problems, mark the correct answer **and** show all your work:

1. The ability to taste a chemical called PTC (phenylthiocarbamide) is controlled by a single gene. The dominant allele of this gene allows a person to detect PTC's bitter taste while the recessive allele leads to the non-taster phenotype. If 0.2 (or 20%) of the US population CANNOT taste PTC, what's the frequency of the dominant allele? Assume the US population is in Hardy-Weinberg equilibrium.
    - A. 0.200 or 20.0 %
    - B. 0.306 or 30.6 %
    - C. 0.447 or 44.7 %
    - D. 0.494 or 49.4 %
    - E. **0.553 or 55.3 %**
    - F. 0.800 or 80 %

$\boxed{\text{EQ}}$

$q^2 = 0.2 \quad p = ?$

$q = \sqrt{q^2} = \sqrt{0.2}$

$q = 0.447$

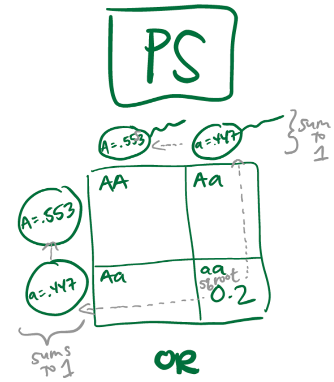
$p = 1 - q = 1 - 0.447$

$p = 0.553$

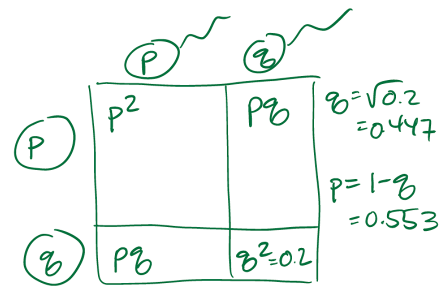
OR

$\sqrt{0.2} = 0.447$

$1 - 0.447 = 0.553$



OR



2. Polydactyly (being born with more than 5 fingers or toes) is caused by a dominant allele of a single gene. If the frequency of the recessive allele is 0.9 (or 90%) in a certain population, what percentage of the population would you expect to be heterozygotes? Assume this population is in Hardy-Weinberg equilibrium. (Note: In reality, there are other causes of polydactyly when it is accompanied by other disorders, so this problem is an over-simplification.)
    - A. 0.01 or 1 %
    - B. 0.10 or 10 %
    - C. **0.18 or 18 %**
    - D. 0.19 or 19 %
    - E. 0.81 or 81 %
    - F. 0.90 or 90 %

$\boxed{\text{EQ}}$

$q = 0.9 \quad 2pq = ?$

$p = 1 - q = 1 - 0.9$

$p = 0.1$

$2pq = 2(0.1)(0.9)$

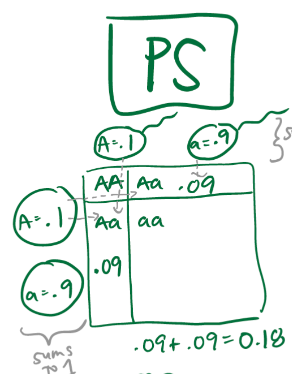
$= 0.18$

OR

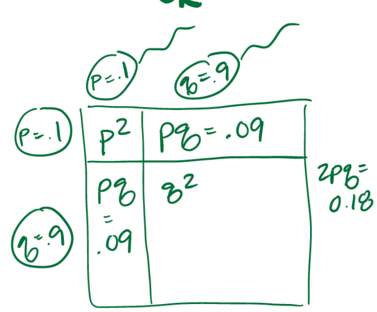
$1 - 0.9 = 0.1$

$2(.9)(.1)$

$= 0.18$



OR



3. Maple Syrup Urine Disease (MSUD) is a disorder in which patients are unable to break down certain amino acids. Patients end up with dangerously high levels of these amino acids in their blood, causing the rapid degeneration of brain cells and, if left untreated, even death. The disorder is caused by a recessive allele of the BCKDHA gene. If 0.95 (or 95%) of a population has the normal phenotype, what is the frequency of the recessive allele? Assume this population is in Hardy-Weinberg equilibrium.

A. 0.050 or 5.0 %
B. 0.224 or 22.4 %
C. 0.347 or 34.7 %
D. 0.603 or 60.3 %
E. 0.776 or 77.6 %
F. 0.950 or 95.0 %

$\boxed{\text{EQ}}$

$p^2+2pq = 0.95 \quad q = ?$

$q^2 = 1-(p^2+2pq)$

$q^2 = 1-0.95$

$\sqrt{q^2} = \sqrt{0.05}$

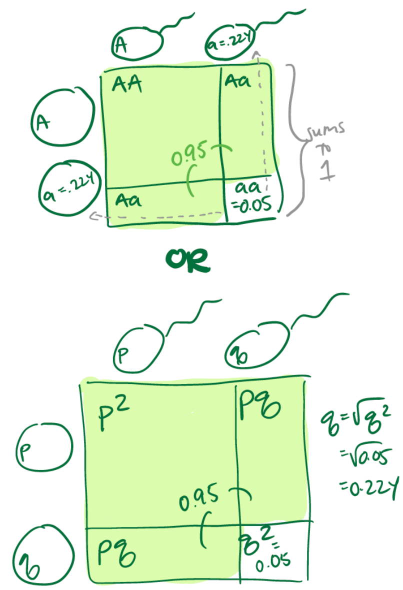
$q = 0.224$

OR

$1-0.95 = 0.05$

$\sqrt{0.05} = 0.224$

$\boxed{\text{PS}}$



Self-efficacy

4. How confident are you in your ability to answer Hardy-Weinberg calculation questions correctly? Circle your answer.

A. No confidence
B. Slight confidence
C. Moderate confidence
D. High confidence
E. Complete confidence

Instructor Guide for Day 2 (Equations 1st section)

Using Population Punnett Squares to Model Hardy-Weinberg Equilibrium

Formative Assessment Clicker Question (review from last class)

In a hypothetical bug population, color is determined by a single gene and black is dominant over white. The frequency of the black allele is 0.1. If this population is in Hardy-Weinberg Equilibrium, what percentage of the bugs would we expect to be white?

A. 9%
B. 10%
C. 18%
D. 81%
E. 90%

Have students click in. Depending on whether most of the class is correct or not, you can let them try again with a neighbor and/or go through the answer slowly together on the board:

$p = 0.1$
Since $p + q = 1$, $q = 0.9$
$q^2 = 0.81$

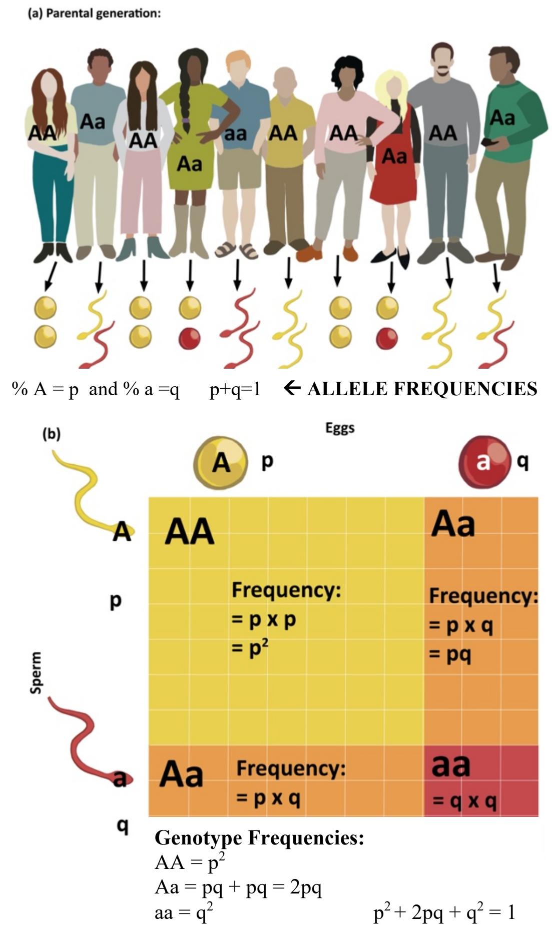
Introducing Population-level Punnett Squares

Ask students to do the following with a neighbor…

- Consider the BIOLOGY behind our Hardy-Weinberg equations:
  - $p + q = 1$
  - $p^2 + 2pq + q^2 = 1$
- Based on what these algebraic terms MEAN, how could you use a population Punnett square to represent these equations graphically
  - What would be different about this Punnett square compared to the ones we used for our genetics unit?
  - Where would you put p, q, p^2, $2pq$, q^2, etc.?

Wander around the room and help students by asking them questions about what each term represents biologically and where that would go on a Punnett square.

Then bring class together and get the following ideas up on the board (generated by students as much as possible) →

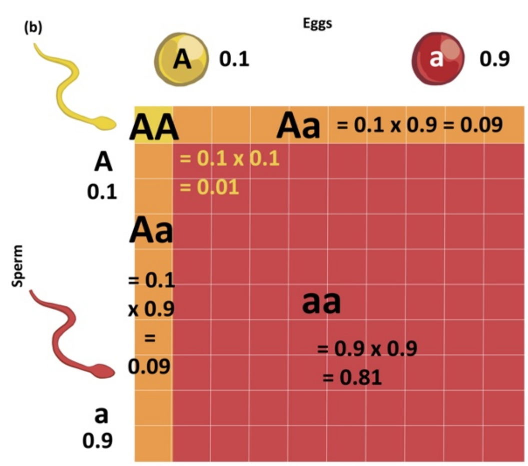


If the frequency of the dominant allele is p, then p goes on the outside of the Punnett square with the gamete that contains the dominant allele. Along the same lines, q will go on the outside of the Punnett square with the gamete that contains the recessive allele. How are p and q related to each other? They still add up to 1 (because in our scenario, there is no other option that can go on the outside of the Punnett square)

The square is just another way to represent probability, this egg AND this sperm coming together is represented by the yellow box ($p \times p = p^2$). Same thing with q^2 and the recessive phenotype zygotes in the red box. Why are there two boxes for Aa? There are two ways to get a heterozygote, each one is pq (that's where the 2 comes from, those 2 boxes). How are the numbers in all of the zygote boxes related to each other? They still add up to 1 (because in our scenario, there is no other zygote genotype possible)

Ask students to go back to the original clicker question and re-solve it using a Punnett square:

After giving them a chance to try on their own, write the following on the board generated with student ideas: →



Practice

Give students the following prompt:

Cystic Fibrosis Example (Use the equations OR the population Punnett square. You choose.)

- Cystic fibrosis (CF) is a disabling and often fatal disease inherited as an autosomal (not sex-linked) recessive characteristic. CF affects cell membrane function and as a consequence causes problems with the glands that produce mucus, digestive enzymes and sweat. Affected individuals often have digestive, respiratory, and reproductive problems. CF occurs in about 1 of every 2000 births in the US white population.

- From this data, what would you estimate to be the frequency of the CF allele in the US white population given that we are in Hardy-Weinberg equilibrium in regards to this gene?

- What is the frequency of heterozygous carriers of the CF allele who are most often parents of children with CF?

Let students work with their neighbor on these questions. After you've wandered around helping for a bit, pause the students to give them the following hints (only if needed).

> *Remember, look specifically for what frequency is given and what frequency is requested: Given = "CF occurs in about 1 of every 2000 births", asks for "frequency of the CF allele" and "frequency of heterozygotes"*

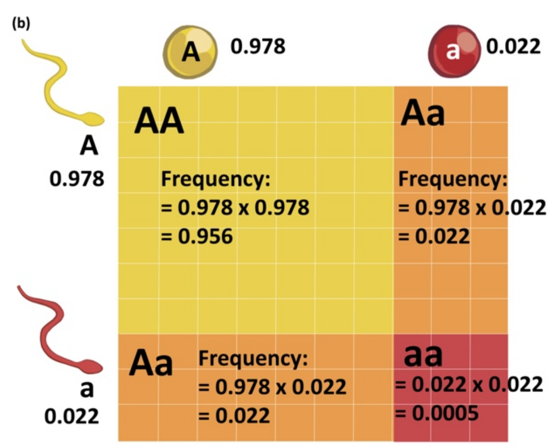
Finally, when they are ready, work through the problem on the board (using BOTH methods), with ideas generated by students.

"CF occurs in about 1 of every 2000 births" = recessive phenotype = aa = q^2=1/2000 = 0.0005

q = square root of q^2 ***= 0.022 (answer #1)***

since p + q = 1, p = 0.978

2pq = 2(0.022)(0.978) ***= 0.044 (answer #2)***



Instructor Guide for Day 2 (Punnett square 1st section)

Deriving and using the Hardy-Weinberg Equations

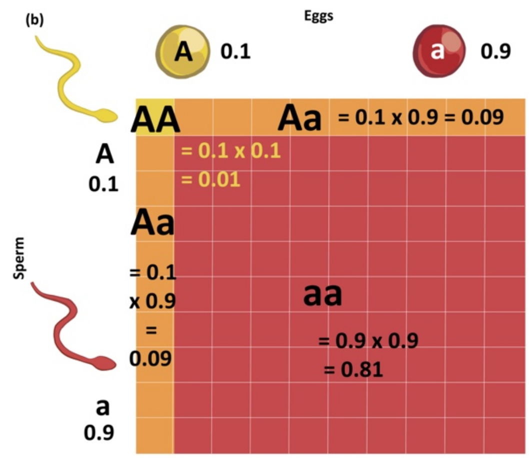
Formative Assessment Clicker Question (review from last class)

In a hypothetical bug population, color is determined by a single gene and black is dominant over white. The frequency of the black allele is 0.1. If this population is in Hardy-Weinberg Equilibrium, what percentage of the bugs would we expect to be white?

F. 9%
G. 10%
H. 18%
I. 81%
J. 90%

Have students click in. Depending on whether most of the class is correct or not, you can let them try again with a neighbor and/or go through the answer slowly together on the board →

81%



Deriving the Hardy-Weinberg Equations

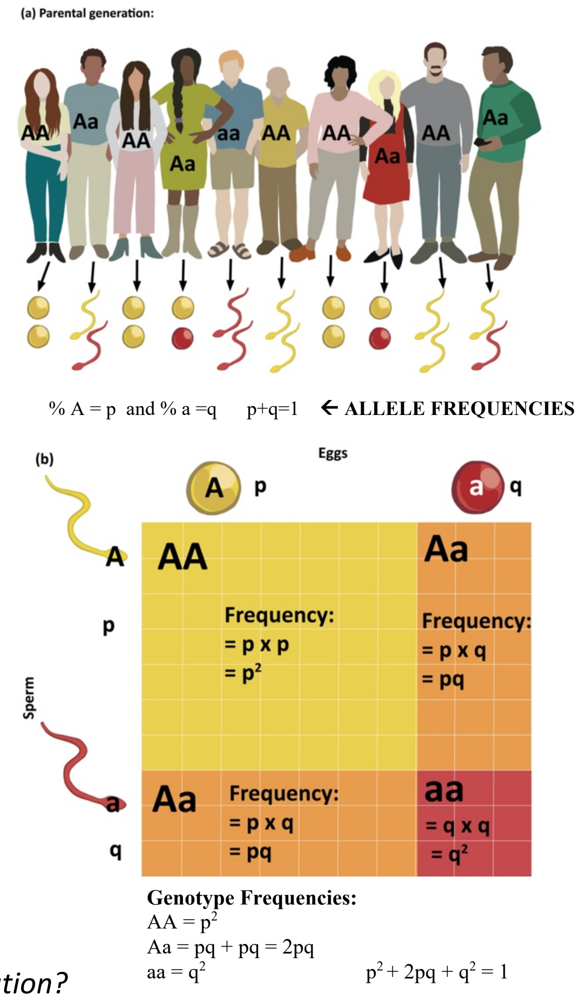
Ask students to do the following with a neighbor…

1. Consider the MATH that you are really doing as you use a population Punnett square. Do you notice any patterns (i.e. that you are doing similar calculations each problem, although perhaps in different orders)?
2. Can you come up with **two equations** to represent all the math you are doing when you use a population Punnett square?
   - Each equation will add up to 1
   - We will call the frequency of the dominant allele "p" and the frequency of the recessive allele "q"
   - How would you represent the frequency of AA, Aa, and aa using p's and q's?

Wander around the room and help students by asking them questions about the patterns they noticed while doing calculation problems with the Punnett square.

Then bring class together and get the following ideas up on the board (generated by students as much as possible) →

If the frequency of the dominant allele is p, then p goes on the outside of the Punnett square with the gamete that contains the dominant allele. Along the same lines, q will go on the outside of the Punnett square with the gamete that contains the recessive allele. How are p and q related to each other? They always add up to 1 (because in our scenario, there is no other option that can go on the outside of the Punnett square) So how could we write that as an equation? ***Allele frequency equation: $p + q = 1$***



Since we would always multiply the frequency of A by the frequency of A to get the frequency of the top left box (AA), that would be represented by p x p or p^2. (Because it's the probability you have that egg AND that sperm, AND means multiply) Same thing with the frequency of aa (q^2). Each heterozygote box would be pq. How are the numbers in all of the zygote boxes related to each other? They always add up to 1 (because in our scenario, there is no other zygote genotype possible) So how could we write that as an equation? ***Genotype frequency equation: $p^2 + pq + pq + q^2 = 1$ or $p^2 + 2pq + q^2 = 1$***

Write each equation on the board plus each term of each equation with its definition (e.g. p = frequency of the dominant allele (A), 2pq = frequency of Aa)

Ask students to go back to the original clicker question and re-solve it using our new equations:

After giving them a chance to try on their own, write the following on the board generated with student ideas:

$p = 0.1$
Since $p + q = 1$, $q = 0.9$
$q^2 = 0.81$

Practice

Give students the following prompt:

Cystic Fibrosis Example (Use the equations OR the population Punnett square. You choose.)

- Cystic fibrosis (CF) is a disabling and often fatal disease inherited as an autosomal (not sex-linked) recessive characteristic. CF affects cell membrane function and as a consequence causes problems with the glands that produce mucus, digestive enzymes and sweat. Affected individuals often have digestive, respiratory, and reproductive problems. CF occurs in about 1 of every 2000 births in the US white population.

- From this data, what would you estimate to be the frequency of the CF allele in the US white population given that we are in Hardy-Weinberg equilibrium in regards to this gene?

- What is the frequency of heterozygous carriers of the CF allele who are most often parents of children with CF?

Let students work with their neighbor on these questions. After you've wandered around helping for a bit, pause the students to give them the following hints (only if needed).

Remember, look specifically for what frequency is given and what frequency is requested: Given = "CF occurs in about 1 of every 2000 births", asks for "frequency of the CF allele" and "frequency of heterozygotes"

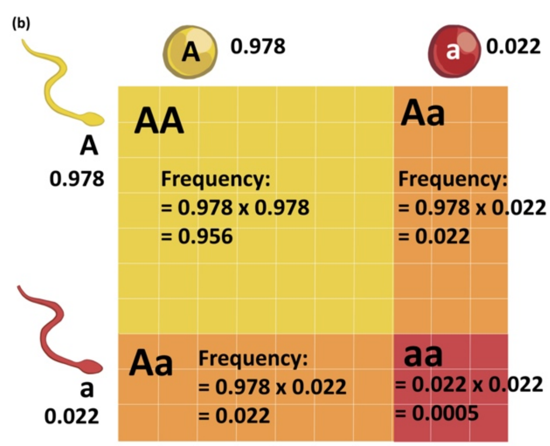
Finally, when they are ready, work through the problem on the board (using BOTH methods), with ideas generated by students.

"CF occurs in about 1 of every 2000 births" = recessive phenotype = aa = q^2=1/2000 = 0.0005

q = square root of q^2 ***= 0.022 (answer #1)***

since p + q = 1, p = 0.978

2pq = 2(0.022)(0.978) ***= 0.044 (answer #2)***



POST-ASSESSMENT

After Day 2 of population genetics instruction

Hardy-Weinberg Calculations

Instructions:

For the following problems, mark the correct answer **and** show all your work:

The ability to taste a chemical called PTC (phenylthiocarbamide) is controlled by a single gene. The dominant allele of this gene allows a person to detect PTC's bitter taste while the recessive allele leads to the non-taster phenotype. If the frequency of the dominant allele is 0.45 (or 45%) in a population that is in Hardy-Weinberg equilibrium, what percentage of the population do you expect to be non-tasters?

A. 0.203 or 20.3 %
B. **0.303 or 30.3 %**
C. 0.450 or 45.0 %
D. 0.495 or 49.5 %
E. 0.550 or 55.0 %
F. 0.698 or 69.8 %

Polydactyly (being born with more than 5 fingers or toes) is caused by a dominant allele of a single gene. In the US, the polydactyly trait has a frequency of 0.001 (or 0.1%). If the US population is in Hardy-Weinberg equilibrium, what is the frequency of the polydactyly-causing allele in this population? (Note: In reality, there are other causes of polydactyly, so this problem is an over-simplification.)

A. 0.00000025 or 0.000025 %
B. 0.00050013 or 0.050013 %
C. 0.00100000 or 0.100000 %
D. 0.00099975 or 0.099975 %
E. 0.99900000 or 99.900000 %
F. 0.99949987 or 99.949987 %

Maple Syrup Urine Disease (MSUD) is a disorder in which patients are unable to break down certain amino acids, causing these amino acids to build up in their blood. This causes the rapid degeneration of brain cells and, if left untreated, even death. The disorder is caused by a recessive allele of the BCKDHA gene. Among Mennonites in Pennsylvania, 1 out of every 176 babies is born with this disorder (which is really high compared to the rest of the US, where 1 in every 180,000 babies has this disease). If the Mennonite population is in Hardy-Weinberg equilibrium, what proportion of the population do you expect to be heterozygotes?

A. 0.0057 or 0.57 %
B. 0.0754 or 7.54 %
C. 0.1394 or 13.94 %
D. 0.8549 or 85.49 %
E. 0.9246 or 92.46 %
F. 0.9943 or 99.43 %

Self-Efficacy

How confident are you in your ability to answer Hardy-Weinberg calculation questions correctly?

A. No confidence
B. Slight confidence
C. Moderate confidence
D. High confidence
E. Complete confidence

Derivation of more complex equations

If there were THREE possible alleles for a gene of interest instead of just two, what would the two Hardy-Weinberg equations be? Use p, q, and r instead of just p and q. (Note: This is not something we covered in class. Rather, we want to see if you understood Hardy-Weinberg well enough to derive a new equation you've never seen.) Show all your work in the space below. Even if you don't know how to finish, at least do as much as you can.

$p + q + r = 1$
$p^2 + q^2 + r^2 + 2pq + 2pr + 2qr = 1$

Understanding of Hardy-Weinberg

Choose one of the assumptions of Hardy-Weinberg Equilibrium (random mating, no natural selection, no mutation, no genetic drift, or no gene flow) and explain why the Hardy-Weinberg equation would not hold if that assumption were violated. Include both biological and mathematical reasons.

Open Response

Please indicated what each of the following Hardy-Weinberg variables represent biologically:
p = **the frequency of the dominant allele in a population**
q = **the frequency of the recessive allele in a population**
p^2 = **the frequency of the homozygous dominant genotype in a population**
$2pq$ = **the frequency of the heterozygous genotype in a population**
q^2 = **the frequency of the homozygous recessive genotype (and recessive phenotype) in a population**
$p^2 + 2pq$ = **the frequency of the dominant phenotype in a population**

Instruction Preference

Which day of Hardy-Weinberg instruction in class was most helpful? Circle your answer.

A. The day we learned the actual equations
B. The day we learned how to use population Punnett squares

Math Anxiety (AMAS) Repeated

- Hopko, D.R., Mahadevan, R., Bare, R.L. and Hunt, M.K. 2003. The abbreviated math anxiety scale (AMAS) construction, validity, and reliability. Assessment, 10(2), pp.178-182.

Supplemental Data

Table S1. Correct equation term definitions by section.

Term	Derived Equations 1st		Punnett Square 1st		Fisher's exact test result (p)
	Incorrect N	**Correct N**	**Incorrect N**	**Correct N**	
p	6	61	8	61	0.78
q	6	61	7	62	1.00
p^2	4	63	10	59	0.16
$2pq$	0	67	2	67	0.50
q^2	9	57	11	58	0.81
$p^2 + 2pq$	12	54	15	53	0.67

Table S2. Pre- and post-math anxiety for all students who took both the pre- and post-assessments.

Variable	Derived Equations 1st			Punnett Square 1st		
	Mean	**SD**	**N**	**Mean**	**SD**	**N**
Pre-anxiety	19.55	5.47	67	20.40	6.59	70
Post-anxiety	19.51	5.72	67	20.82	6.69	70

*Two-way mixed ANOVA to investigate the effects of time (within-subjects) and section (between subjects): time $p = 0.45$, $\eta_p^2 = 0.004$; section $p = 0.29$, $\eta_p^2 = 0.008$; time*section $p = 0.35$, $\eta_p^2 = 0.006$.*

Figure S1. Success on Question 1 of the mid-assessment by treatment and major or anxiety.