Human Genetic Variation:
Hardy-Weinberg Equilibrium Worksheet Key

Hardy-Weinberg Equilibrium with an autosomal recessive allele:
ΔCCR5 and HIV resistance example
"Given the impact of this mutation on the current HIV epidemic, we would like to know the frequency of this genotype."
[pause at 12:18]

From Martinson et al[1]., the following genotypic categories were quantified:

Genotype	*Phenotype*	*Martinson et al*
CCR5/ CCR5	Normal HIV infection susceptibility	647
CCR5/ ΔCCR5	Delay in progression to AIDS after HIV infection	134
ΔCCR5/ ΔCCR5	Partial HIV resistance	7
Total		788

1. Given these numbers, calculate the frequency of each genotype.

These frequencies can be calculated by dividing the number of cases with the genotype by the total number of study subjects. Those calculations are demonstrated below.

Genotype	*Phenotype*	*Martinson et al*	***Genotypic frequency***
CCR5/ CCR5	Normal HIV infection susceptibility	647	**p^2 = 647/788 = 0.821**
CCR5/ ΔCCR5	Delay in progression to AIDS after HIV infection	134	**2pq = 134/788 0.168**
ΔCCR5/ ΔCCR5	Partial HIV resistance	7	**q^2 = 7/788 = 0.011**
Total		788	

[1] Martinson JJ, Chapman NH, Rees DC, Liu YT, and Clegg JB (1997) Global distribution of the CCR5 gene 32-basepair deletion. Nat Genet. 16(1):100-3.

2. How can we also use the data in the table to calculate the frequency of each allele?

First we determine the number of each allele contributed to the population by each genotype as follows:

Genotype	*Abbreviated notation*	*Martinson et al*	***Copies of A allele***	***Copies of a allele***
CCR5/ CCR5	AA	647	**2x647 = 1294**	**0**
CCR5/ ΔCCR5	Aa	134	**134**	**134**
ΔCCR5/ ΔCCR5	aa	7	**0**	**2x7 = 14**
Total		788	**1428**	**148**

Next we calculate the frequency of each allele by taking the total number of each allele observed over the total number of alleles genotyped, as follows:

Frequency of allele A = p = [(2 x 647) + 134] / [2x788] = 1428 / 1576 = 0.906
Frequency of allele a = q = [(2 x 7) + 134] / [2x788] = 148 / 1576 = 0.094
OR q = 1 – frequency of allele A = 1 - 0.906 = 0.094

Students can check answers at time = 16:11, and if numbers don't match, they can listen to the preceding explanation or compare answers against the provided key.

Students should resume the video at 16:11 if answers match the given values.

[pause at 17:50]

3. Using Hardy-Weinberg equilibrium, calculate the expected frequencies of each genotype from the allelic frequencies.

Equations: Hardy-Weinberg equilibrium

$$p+q = 1$$
$$p^2 + 2pq + q^2 = 1$$

p = frequency of allele A
q = frequency of allele a
p^2 = frequency of genotype AA
q^2 = frequency of genotype aa
2pq = frequency of genotype Aa

Hardy-Weinberg equilibrium in autosomal recessive disease

Let's apply this to our ΔCCR5 example in a step-wise fashion. If we know from population genetics studies that the frequency of the ΔCCR5 allele "a" is 0.094 in the Western European population, what is the frequency of the individuals in this population with the ΔCCR5 heterozygote genotype?

Using the Hardy-Weinberg equations

1. *q = 0.094*
2. *As p + q = 1, p = 0.906*
3. *p^2 = genotype frequency of individual AA = 0.906 x 0.906 = 0.821*
4. *q^2 = genotype frequency of individual aa = 0.094 x 0.094 = 0.009*
5. *2pq = genotype frequency of individual Aa = 2(0.906 x 0.094) = 0.170*

Table 1 summarizes these calculations in another form.

Sex	*Genotype*	*Phenotype*	*Incidence (approx)*
Male or Female	*AA*	*Normal (homozygote)*	*p^2 = 0.821*
Male or Female	*Aa*	*Normal (heterozygote)*	*2pq = 0.170*
Male or Female	*aa*	*Partial HIV resistance*	*q^2 = 0.009*

4. Do these frequencies match those ascertained by Martinson et al.?

Genotype	*Phenotype*	*Martinson et al.*	*Genotypic frequency*
CCR5/ CCR5	Normal HIV infection susceptibility	647	= 647/788 = 0.821
CCR5/ ΔCCR5	Delay in progression to AIDS after HIV infection	134	= 134/788 = 0.170
ΔCCR5/ ΔCCR5	Partial HIV resistance	7	= 7/788 = 0.009
Total		788	

Yes, these match the frequencies predicted by Hardy-Weinberg equilibrium calculations!

[Resume play at 22:26 – pause at 23:22]

Hardy-Weinberg Equilibrium with an autosomal dominant allele:
Marfan Syndrome example

5. Marfan's syndrome is caused by an autosomal dominant mutation in the fibrillin-1 gene. Given that the incidence of Marfan's syndrome in a particular population is 1 in 100,000 individuals, and that individuals homozygous for this dominant allele are, for all intents and purposes, non-existent in the population, what is the allelic frequency of mutated fibrillin-1 in this population? Does this allele frequency predict the observed absence of individuals with the homozygous genotype in the population?

In autosomal dominant disease, the components of Hardy-Weinberg equilibrium are a little different. 2pq (Aa) is the incidence of an autosomal dominant condition, which includes only heterozygotes = 1 in 100,000 or 0.00001. The allelic frequency of the diseased gene A (p) is usually very small, thus the allelic frequency of normal gene a (q) approximates 1.

Using the Hardy-Weinberg equations

1. *q = ~1*
2. *2pq = incidence of Marfan's syndrome = 0.00001*
3. *As 2pq = 0.00001 and q = ~1, then p = (0.00001)/2 and p = 0.000005*
4. *p^2 = genotype frequency of individual AA, which is lethal prior to reproductive age = 2.5 X 10^{-11} = ~0*
5. *q^2 = genotype frequency of individual aa, homozygous normal = 1 – 2pq – p^2 = 1 – 0.00001 - ~0 = 0.99999 = ~ 1*

Table 2 summarizes these calculations in another form.

Sex	*Genotype*	*Phenotype*	*Incidence (approx)*
Male or Female	*AA*	*Embryonically lethal*	*p^2 = 2.5 X 10^{-11} = ~0*
Male or Female	*Aa*	*Marfan's syndrome*	*2pq = 0.00001*
Male or Female	*aa*	*Normal (homozygote)*	*q^2 = 0.99999 = ~1*

[Resume play at 25:14 – pause at 25:50]

Hardy-Weinberg Equilibrium with an X-linked recessive allele:
Red-green color blindness example

Protanopia is one type of red-green color blindness inherited in an X-linked recessive fashion. In a certain population, the prevalence of protanopic males is 1 in 100.

6. What is the frequency of protanopic females?

As males are hemizygous for the X chromosome, a male individual only has only copy of each trait, indicating that the frequency of affected males is equal to the allele frequency. Thus q =

0.01 and p = 0.99.

For females to be affected by this X-linked recessive disorder, they must be homozygous for the recessive allele. Using the Hardy-Weinberg equations, this can be calculated as follows:

1. *frequency of X^aY protanopic males = q = 0.01*
2. *frequency of X^aX^a protanopic females = q^2 = 0.0001*
3. *p = 0.99*
4. *p^2 = 0.9801*
5. *2pq = 0.0198*

Table 3 constructs these calculations in another form.

Sex	*Genotype*	*Phenotype*	*Incidence (approx)*
Male	*X/Y*	*Normal hemizygous*	*p = 0.99*
Male	*X^a/Y*	*Color blind*	*q = 0.01*
Female	*X/X*	*Normal homozygous*	*p^2 = 0.9801*
Female	*X^a/X*	*Normal heterozygote*	*2pq = 0.0198*
Female	*X^a/X^a*	*Color blind*	*q^2 = 0.0001*

An affected female would have two affected copies of the allele – thus the frequency would be 0.0001. This confirms the observation that X-linked recessive disorders are more prevalent among males in a population.

[Resume play 27:08]