Hardy-Weinberg

Readings: Hamilton 2.1 – 2.4
(**Strongly**) suggested practice:
Problems 2.1 & 2.3, Interact box 2.2

1

Learning outcomes

By the end of this section you should:

- Be able to calculate expected HW genotype frequencies for autosomal and sex-linked loci
- Be able to estimate allele frequencies from phenotype frequency data (including in cases with dominance)
- Be able to perform a chi-square test to assess whether observed genotype frequencies in a population are at HW expectation
- Understand what deviations from HW expectation imply (i.e. know the assumptions underlying HW and its utility as a null model)
- Be able to apply the principles of HW when allele frequencies differ between males and females, and for sex-linked loci.

2

The Hardy-Weinberg Principle

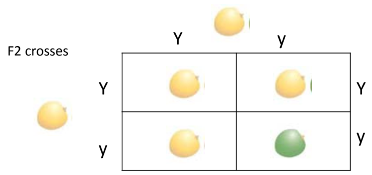
1) There is a predictable relationship between allele and genotype frequencies (expressed as the Hardy-Weinberg equation).

2) Allele and genotype frequencies remain constant across generations when certain assumptions are met (in other words, once equilibrium frequencies are obtained, the process of Mendelian transmission itself does not alter allele or genotype frequencies).

3

Hardy-Weinberg

- The formulation of the HW equation was necessitated from the criticism that, if Mendelian genetics were true, we should always see a 3:1 ratio of dominant to recessive phenotypes in a population (which we don't).



4

What is the Hardy-Weinberg equation?

- Hardy and Weinberg responded with what they thought was a mathematically obvious fact: that the genotypic ratios should depend on the allelic frequencies. In fact, they formulated the relationship between allele and genotype frequencies:

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

- Let's derive this so we understand it.

5

Hardy-Weinberg assumptions

1) Diploid organism that reproduces sexually
2) Random mating with respect to locus (and random union of gametes)
3) No natural selection at the locus
4) Mutation is negligible
5) Individuals do not move in our out of the population
6) Population is infinitely large

6

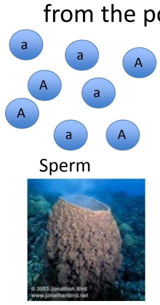
Random union of gametes

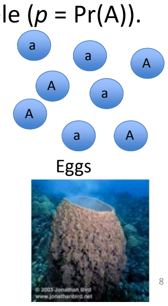
Sperm and eggs meet at random

-No opportunity for nonrandom mating as there is a large pool of gametes
-

Random union of gametes

- Remember that p (allele frequency) is the <u>probability</u> that an allele picked at random from the population is an A allele (p = Pr(A)).

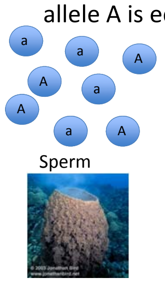


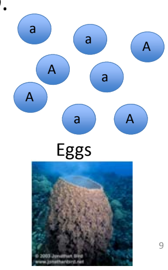


8

Random union of gametes

- Think of these sperm and eggs as populations of haploid individuals.
- The chance of choosing a sperm/egg with allele A is equal to probability p.





9

Random union of gametes

- The probability that a random sperm has allele A is the allele frequency *p*
- The probability that a random egg has allele A is the allele frequency *p*
- What is the probability that a sperm with A and an egg with A will meet?

PAA = Pr(Asperm and Aegg)
= Pr(Asperm) x Pr(Aegg)
=p x p
=p^2

PAa = Pr(Aa) = Pr(Asperm and aegg) + Pr(asperm and Aegg)
=PR(Asperm)PR(aegg) + Pr(asperm)Pr(Aegg)
= p x q + q x p
= 2pq

Random union of gametes

- Use the product rule: Pr*(A and B)* = Pr*(A)*Pr*(B)*
- What is the Pr(offspring having AA genotype)?

	A	*a*
A	p^2	*pq*
a	*qp*	q^2

Hardy-Weinberg Frequencies

Genotype	AA	Aa	aa
Frequency (H-W)	$P_{AA} = p^2$	$P_{aa} = 2pq$	$P_{aa} = q^2$

$P_{AA} + P_{Aa} + P_{aa} = 1$ (they are the full set of mutually exclusive genotype frequencies – recall probability Rule #1)

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

Yet another way to think about it:

- Each union is a ‘trial’, where the father donates allele *A* with probability *p* <u>or</u> allele *a*, with probability *q*. <u>And</u>, the mother donates an allele with the same probability. Mathematically we get:

$$(pA + qa)(pA + qa)$$
$$= p^2AA + 2pqAa + q^2aa$$

Dad donates *A* <u>or</u> *a*

<u>and</u>, Mom donates *A* <u>or</u> *a*

13

$p^2 + 2pq = 3q^2$

AA Aa - aa

$p^2 = 2p(1-p) + 3(1-p)^2$

$p^2 + 2p - 2p^2 = 3 - 6p + 3p^2$

$=p^2 + 2p = 3 - 6p + 3p^2$

$2p = 3 - 6p + 4p^2$

$4p^2 - 8p + 3 = 0$

$p^2 - 2p + 3/4 = 0$

$(p-1.5)(p - 0.5) = 0$

If $p - 1.5 = 0$, $p = 1.5$

$p - 0.5 = 0$, $p = 0.5$

$(1-p)(1-p)$

$=1 - p - p + p^2$

$=1 - 2p + p^2$

In class exercise

- At what allele frequency *p* would you expect a 3:1 phenotypic ratio of dominant to recessive phenotypes?

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

14

- At what allele frequency *p* would you expect a 3:1 phenotypic ratio of dominant to recessive phenotypes?
- *p* = 0.5

$$(0.5)^2 + 2(0.5)(0.5) + (0.5)^2 = 1$$

$$0.25AA + 0.5Aa + 0.25aa = 1$$

0.75 A- (for $0.25AA + 0.5Aa$) 0.25 aa (for $0.25aa$)

15

A- = second allele doesn’t matter as it will automatically be dominant A genotype

- The HW priciple also shows that when the assumptions of HW are met, allele frequencies *p* and *q*, as well as genotypic frequencies remain the same through time.
- Let's prove this, first via an example and then in general.

16

If at gen = 0, $p = 0.6$ and $q = 0.4$

- 60% of sperm are A_1, 40% are A_2
- same for eggs.

Next generation (generation 1):

- P_{A1A1} = 0.6^2 = 0.36
- P_{A1A2} = 2(0.6)(0.4) = 0.48
- P_{A2A2} = 0.4^2 = 0.16

17

Generation 1

- $p = P_{A1A1} + ½ P_{A1A2} =$
- *p* remains unchanged from generation 0 to generation 1 when gametes interact randomly.
- We can also prove this generally:

$$p(t+1) = P_{A1A1}(t+1) + \frac{1}{2}P_{A1A2}(t+1)$$
$$p(t+1) =$$

18

pt = pt+1
pt+1 =

=pt^2 + 1/2x2pt(1-pt)
=p^2 + p - p?? Can't read his writing.

- Does Hardy-Weinberg also apply in cases of more “traditional” reproduction (i.e. that don’t actually involve a pool of randomly interacting gametes)?

19

What about when specific genotypes mate?

- When gametes interact, they are haploid, so it is simple to calculate the probabilities of different alleles interacting
- But animals that actively mate have specific genotypes that interact (AA, Aa, or aa). Does HW still work?

20

Imagine a population of animals at HWE with p = 0.4 A1A1

- What is the probability that an offspring will receive the A_1 allele from their father? This can occur if the offspring receives A_1 from an A_1A_1 father, or from an A_1A_2 father.

1) Probability of receiving A_1 from A_1A_1 father

= Probability of having an A_1A_1 father AND that this father donated an A_1 allele:

Use the Product rule (Pr*(A and B)* = Pr*(A)*Pr*(B)):*

- $Pr(A_1A_1) \times Pr(A_1A_1 \text{ donates } A_1) = p^2 \times 1 = p^2$

21

Pr(A1A1) = PA1A1

Pr(A1A1 father and A1 transmitted from father)

2) Probability of receiving A_1 from A_1A_2 father

= Probability of having an A_1A_2 father AND that this father donated an A_1 allele:

- $Pr(A_1A_2)$ x $Pr(A_1A_2$ donates $A_1)$= $2pq$ x ½
  $=pq = p(1-p) = p - p^2$

Overall probability of receiving an A_1 allele from the father:

- Apply "or" rule for mutually exclusive events Pr(A or B) = Pr(A) + Pr(B):
- $p^2 + p - p^2 = p$

22

Who mates with who not determined by genotype, but can predict genotype frequences by allele frequences

- The probabilities are the same for receiving an A_1 allele from the mother.
- So, use the product rule to calculate the probability of being A_1A_1 (i.e. of receiving A_1 from both mother AND father):

$= p \times p = p^2$

The *same answer* we got for our model of randomly interacting gametes.

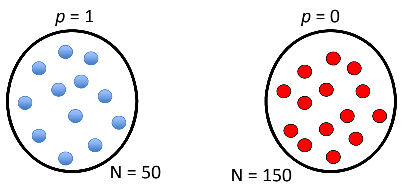
23

- So, even when considering the mating of individuals with specific genotypes, the rules we derived for the random union of gametes still apply.
- Remember that p = the probability of taking a random allele from the population and finding A_1.
- So, as long as **mating is random with respect to the loci under consideration (e.g., individuals do not choose mates in a way that is predictable by the genotypes at this locus)**, the rules of HW still apply.

24

Random mating and HW

- HW is established even after a single round of random mating.
- Say you mix two homozygous populations:



25

-Fixed = allele frequency of 1

p= # big A alleles/ total # of alleles

- In the new "population", p = 50/(50+150) = 0.25, but there are no heterozygotes so this population is obviously not at HWE; p is also the frequency of AA individuals
- AA offspring are formed by two AA parents mating = $p \times p = p^2$
- aa offspring are formed by two aa parents mating = $q \times q = q^2$
- Aa offspring are formed by one AA and one aa mating = $p \times q + q \times p = 2pq$

26

PAA = Pr(A from dad and A from mom)
= p x p
=p^2

Deviations often aren't great because they don't accumulate over generations.

Practice

- Two populations are mixed, one with p = 0.2 and one with p = 0.6. Assuming that each population is the same size, and random mating, what are the genotypic frequencies in the next generation?

27

Need to know allele frequencies in pool, but how do you know them once all together

Because equal #s, just average them

p* = 0.4

PAA = 0.4^2 = 0.16
PAa = 2pq = 0.8x0.6 = 0.48
Paa = q^2 = 0.6^2 = 0.36

Solution

- p =
- P_{AA} =
- P_{Aa} =
- P_{aa} =

What is the use of a model that is so constrained by assumptions?

1 It demonstrates clearly that the process of particulate inheritance does not itself cause changes in allele frequencies across generations.

2 Given #1, it provides a useful a null model. Deviations from H-W can be indicative of something biologically interesting going on.

What is the use of a model that is so constrained by assumptions?

3. H-W separates life history stages into gamete formation (I) and the zygote to adult stage (II). This means that models can conveniently incorporate population changes in the second stage (e.g., mortality due to differential survival, migration) and then calculate the gametes based on the change to the pop level allele frequencies.
4. H-W is approximately true in many situations where the assumptions are violated to a small degree (e.g., weak selection or low mutation rates) and if mating is random every generation (so genotype frequencies are reset and deviations do not accumulate).

Testing for H-W

- Let's find the expected number under H-W of each genotype for the data below.

Genotype	Number	Expected
SS	141	
SF	111	
FF	28	
Total	280	

31

#S alleles / total # alleles
= 141(2) + 111 / 280(2)

Genotype	Number	Expected
SS	141	
SF	110	
FF	29	
Total	280	

$$p = \frac{141 + \frac{1}{2}(110)}{280} = 0.7$$

Or $$p = \frac{2(141) + 110}{2(280)} = 0.7$$

32

Expected numbers

- $q = 1 - p = 0.3$

- $SS_{exp} = p^2N =$
- $SF_{exp} = 2pqN =$
- $FF_{exp} = q^2N =$

N - need to multiply by population size

33

Genotype	Number	Expected
SS	141	137.2
SF	110	117.6
FF	29	25.2
Total	280	280

- Is this population in HWE?
- How confident are you in your conclusion?

34

Genotype	Number	Expected
SS	141	137.2
SF	110	117.6
FF	29	25.2
Total	280	280

- Is this population in HW equilibrium?
- How confident are you in your conclusion?

- How do you formally test HW?

35

Box 3.1 - Calculating the Chi-Square Test Statistic

Whether populations are significantly different than the conditions expected under HWE can be tested using the chi-square test.

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

36

Statistical tests: review

- We are always assessing whether we can reject the null hypothesis (H_0)
- If we can reject H_0, then we conclude that there is a difference

H_0 = There is no difference

H_1 = There is a difference

- In this class, we are testing whether we can reject the hypothesis that a population is in HW equilibrium.

37

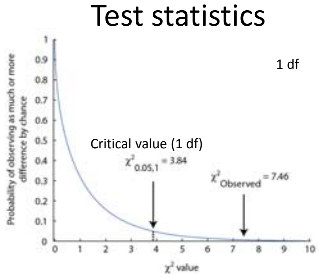
Test statistics

- For a chi-square test, the test statistic (χ^2) is a measure of how different our observed data is from what is predicted by HWE. It is calculated using the observed and expected NUMBERS OF INDIVIDUALS (NOT FREQUENCIES):

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

38

Test statistics



- The higher the χ^2 value, the lower the chance that the difference between observed and expected is due to chance. is more likely to be found by random chance.

39

Degrees of freedom

- df is a measure of how much information was used in estimating the chi-square value.
- In general, the df is the number of independent observations (i.e. data points) that contribute to the estimate minus the number of parameters that had to be estimated

40

Degrees of freedom

For chi-square, degrees of freedom (df) = Number of data classes - Number of parameters estimated from the data – 1

E.g. For two alleles:

df = 3 observed counts of genotypes – one allele frequency estimated - 1 = 1

41

Significance values (P-values)

- P-value is the probability of rejecting H_0 despite it being true (termed a type I error, or false positive)
- By convention, we reject H_0 when there is a 5% or less chance making a type I error (i.e. a P-value ≤ 0.05).
- We use the test statistic and df to find a *p*-value, and we use the P-value to make a conclusion.

42

P-values

- Table 2.5 in text gives χ^2 critical values for df of 1 to 5. Normal convention is to use P-value = 0.05 as the cutoff.

Table 2.5 χ^2 values and associated cumulative probabilities in the right-hand tail of the distribution for 1–5 df.

df	Probability 0.5	0.25	0.1	0.05	0.01	0.001
1	0.4549	1.3233	2.7055	3.8415	6.6349	10.8276
2	1.3863	2.7726	4.6052	5.9915	9.2103	13.8155
3	2.3660	4.1083	6.2514	7.8147	11.3449	16.2662
4	3.3567	5.3853	7.7794	9.4877	13.2767	18.4668
5	4.3515	6.6257	9.2364	11.0705	15.0863	20.5150

43

Genotype	Number	Expected
SS	141	137.2
SF	110	117.6
FF	29	25.2
Total	280	280

- Let's test whether this population is in HWE.

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

44

Genotype	Number	Expected	(Observed-Expected)^2/Expected
SS	141	137.2	0.1052
SF	110	117.6	0.4912
FF	29	25.2	0.573
Total	280	280	1.1694

- df = 3 genotypes – 1 allele frequency - 1 = 1

45

- Where does χ^2 = 1.169 with 1 df fall on the Chi-squared table?

df	Probability 0.5	0.25	0.1	0.05	0.01	0.001
1	0.4549	1.3233	2.7055	3.8415	6.6349	10.8276

- P-value > 0.25
- What conclusion do we make?

46

Calculate P value, don't need table, use P calculation on Excel

Know how to do chi-squared tests

- Because P-value > 0.05, we fail to reject H_0 that the population is in HWE.

- Know how to perform chi-square tests to determine whether populations are in HWE, as it may reappear on assignments and tests!

47

H-W with more than 2 alleles

- Say we are interested in a locus with 3 loci (p_1, p_2, and p_3).
- What are the expected genotypic frequencies for this locus under HW?

$p_1 + p_2 + p_3 = 1$

48

More than 2 alleles

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

	p_1	p_2	p_3
p_1	p_1^2	p_1p_2	p_1p_3
p_2	p_1p_2	p_2^2	p_2p_3
p_3	p_1p_3	p_2p_3	p_3^2

- $(p_1 + p_2 + p_3)^2$

$= p_1^2 + p_2^2 + p_3^2 + 2p_1p_2 + 2p_1p_3 + 2p_2p_3 = 1$

49

More than 2 alleles

- The HW genotypic frequencies for any number of alleles are

$(p_1 + p_2 + \ldots + p_i)^2 = 1$

- Where ***i*** is the total number of alleles.
- You should be able to expand any function like this.

50

In Class Exercise: Testing for Hardy-Weinberg.

Calculate the Expected Hardy-Weinberg Frequencies

Genotype	Number	Expected
SS	141	
SF	111	
FF	28	
SI	32	
FI	15	
II	5	
Total	332	

S = F = I =

51

Answers

Genotype	Observed Number	Expected Frequency	Expected Number
SS	141	0.4096	136
SF	111	0.3507	116
FF	28	0.0751	25
SI	32	0.1101	37
FI	15	0.0471	16
II	5	0.0074	2
Total	332	1.0000	332
χ^2			

- What are the degrees of freedom for 3 alleles?

52

Chi – square test for fit of HWE

Genotype	Number	Expected Number	(Observed-Expected)^2/Expected
SS	141	136	0.1838
SF	111	116	0.2155
FF	28	25	0.36
SI	32	37	0.6757
FI	15	16	0.0625
II	5	2	4.5
Total	332	332	**5.9975**

53

First, find the χ^2-crit, or, the χ^2 value above which the probability of the deviation from HWE occurring by chance is < 5%.

Table 2.5 χ^2 values and associated cumulative probabilities in the right-hand tail of the distribution for 1–5 df.

	Probability					
df	0.5	0.25	0.1	0.05	0.01	0.001
1	0.4549	1.3233	2.7055	3.8415	6.6349	10.8276
2	1.3863	2.7726	4.6052	5.9915	9.2103	13.8155
3	2.3660	4.1083	6.2514	7.8147	11.3449	16.2662
4	3.3567	5.3853	7.7794	9.4877	13.2767	18.4668
5	4.3515	6.6257	9.2364	11.0705	15.0863	20.5150

Critical value

Observed value of 5.9975 falls somewhere between 0.25 < P-value < 0.1,
so we say P-value > 0.1 and we therefore do not reject the null hypothesis (i.e. HWE)

54

Interpretation

- Actual P-value is approximately 0.112 with 3 df (you can calculate this using the chidist function in Excel)
- Therefore, there is a 15% probability that chance of getting a value this large despite H_0 being true, which strengthens our confidence in the fit of the HWE model to the data.

NOTE: **a P-value is the probability of obtaining a test statistic (in this case a χ^2 value) at least as extreme as the one that was actually observed, assuming that the null hypothesis is true. In other words, the probability of rejecting H_0 despite it being true.**

55

H-W Recap

- Under certain assumptions, there is a predictable relationship between allele and genotype frequencies.
- Once H-W is established, these frequencies do not change from one generation to another under these assumptions.
- We can test statistically whether genotype frequencies at a locus meet H-W expectation.

56

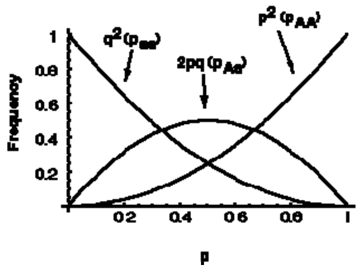
Hardy-Weinberg and Rarity

- Suppose the A_2 allele is *rare*, or in other words that q (= 1 – p) is quite small.
- We might want to know if A_2 alleles are more likely to be found in homozygotes or heterozygotes.
- The ratio of A_2 found in hetero *vs.* homozygotes is:

$$\frac{2pq}{q^2} = \frac{2p}{q} \approx \frac{2}{q}$$

57

Graph of genotype Frequencies under HWE



58

- Relationship between allele and genotype frequencies has implications for the transmission of rare recessive alleles
- Many rare human genetic disorders, such as PKU (the inability to metabolize phenylalanine) are recessive and are therefore expressed only in individuals that receive two copies of the allele
- Even though the disease itself may be rare, the disease allele can be hidden and passed on through heterozygotes

59

Using H-W to estimate allele frequencies

- If we have all allele or genotype frequencies, we can assess whether a locus in a population is at HW equilibrium.
- But can we do the reverse – assume HW and use this to estimate missing information?

60

Dominance

- Dominance makes it **impossible to distinguish homozygous dominant and heterozygous individuals from their phenotype alone (i.e. their genotype is unknown)**.
- If we assume H-W equilibrium, then *q* can be estimated from the frequency of homozygous recessives: $\hat{q} = \sqrt{R}$

61

Dominance

- By assuming HWE to calculate *q*, we are actually estimating *q*, hence the use of $\hat{q}$.

62

Dominance

- By assuming H-W to calculate *q*, we are actually estimating *q*, hence the use of $\hat{q}$.
- Note that in the case where *p* and *q* are estimated by assuming H-W, it is not possible to perform the χ^2 goodness of fit test (because our estimated allele frequencies were calculated by assuming it's true ; or equivalently, *df* = # of data classes - # of parameters estimated – 1 = 3 -2 -1 = 0).

63

Exercise

- A population in China has a low frequency of Rh deletion, a recessive allele. If 20 people out of a population of 1000 have the recessive Rh negative phenotype, what are the frequency of the two alleles? What proportion of heterozygous individuals (who exhibit the dominant phenotype) are expected? Assume allele frequencies are at Hardy-Weinberg Equilibrium.

64

Exercise

$$\hat{q} = \sqrt{\frac{20}{1000}} = \sqrt{0.02} = 0.14$$

$$\hat{p} = 1 - 0.14 = 0.86$$

$$2\hat{p}\hat{q} = 2(0.14)(0.86) = 0.24$$

65

Extending HWE to >2 alleles

- Can you assume HWE to find allele and genotype frequencies for loci with more than two alleles?
- Yes, you just need more information.
- Remember that $\sum p_i = 1$
- For 3 alleles, you need to be able to calculate 2 allele frequencies.
- For 4 alleles, you need to be able to calculate 3 allele frequencies.

66

Exercise

- A locus with three alleles (A1, A2, A3) in fruit flies controls eye colour. Wild-type red eyes are seen if A1 is present, as well as for the heterozygote A2A3. Homozygous A2 individuals have brown eyes, and homozygous A3 individuals have orange eyes. A population has 600 wild-type flies, 50 brown eyed flies, and 20 orange eyed flies. Assuming HW, find the estimated number of individuals with each genotype.

67

Step 1

- Estimate all the allelic frequencies you can.

$$\hat{p}_2 =$$

$$\hat{p}_3 =$$

68

Step 2

- Use calculated allelic frequency estimates to find the last allelic frequency.

$$\hat{p}_1 =$$

69

Step 3

- Use the estimated allelic frequencies to calculate the expected phenotypic frequencies.

$\hat{p}_1^{\,2} =$

$2\hat{p}_1\hat{p}_2 =$

$2\hat{p}_1\hat{p}_3 =$

$2\hat{p}_2\hat{p}_3 =$

70

Caution

- Reminder: because we assumed H-W to calculate the genotype frequencies, we can not use a chi-square test to determine if the population is in HWE. This would be circular.

71

Different Allele Frequencies in Males and Females

- What happens if allele frequency differs in males and females?
- E.g. a bunch of males (with different allele frequencies) are introduced into a population)

72

Random union of gametes

- We had already considered the probability of alleles coming from sperm or eggs.
- Because allele frequencies were the same for the sexes, p_m and p_f were simply replaced with p

	p	q
p	p^2	pq
q	qp	q^2

73

Random union of gametes

- When the allele frequencies differ between the sexes, the same principles still apply, we just have to consider p_m and p_f separately.

	p_m	q_m
p_f	$p_f p_m$	$p_f q_m$
q_f	$q_f p_m$	$q_f q_m$

74

- After mating, the frequencies of genotypes in the next generation will be:

$$P_{11}(t+1) = p_m p_f$$
$$P_{12}(t+1) = p_m q_f + q_m p_f$$
$$P_{22}(t+1) = q_m q_f$$

75

- From these genotype frequencies, we can calculate the allele frequencies in the next generation:

$$\begin{aligned} p(t+1) &= P_{11}(t+1)+\frac{1}{2}P_{12}(t+1) \\ &= p_m p_f + \frac{1}{2}\left(p_m q_f + q_m p_f\right) \\ &= p_m p_f + \frac{1}{2}\left(p_m(1-p_f)+(1-p_m)p_f\right) \\ &= p_m p_f + \frac{1}{2}\left((p_m - p_f p_m)+(p_f - p_f p_m)\right) \\ &= p_m p_f + \frac{1}{2}\left(p_m - 2p_f p_m + p_f\right) \\ &= \frac{1}{2}\left(p_m + p_f\right) = \bar{p} \end{aligned}$$

Thus, after one generation, the allele frequency is the same for males and females and will be the average of what is was in males and females of the parental generation.

76

Question

- However, while allele frequency will be the same in the male and female offspring, genotype frequencies will not necessarily be at HW
- If HW assumptions hold, how long will it take for a population with unequal p_m and p_f to reach HW genotype frequencies?

77

Sex chromosomes

In chromosomal sex determination, one sex has two copies of one sex chromosome (the homogametic sex) and the other sex has one copy of each of two different sex chromosome (the heterogametic sex):

	males	females
Mammals	XY	XX
Birds/Lepidoptera	ZZ	ZW

Predicting genotype frequencies under HW at a locus on the sex chromosomes requires a bit more care.

78

Under random mating:

			Male gametes		
			X-bearing		Y-bearing
			X^A	X^a	Y
	Allele	Freq.	*p*	*q*	
Female gametes	X^A	*p*	X^AX^A p^2	X^AX^a *pq*	X^AY *p*
	X^a	*q*	X^AX^a *qp*	X^aX^a q^2	X^aY *q*

Genotype frequencies in females equal the HW frequencies (p^2, $2pq$, and q^2), while in males they equal the allele frequencies (*p* and *q*).

79

Sex chromosomes consequences

Phenotypes resulting from a recessive allele for an X-linked gene will be more common in males than in females

Why? Because *q* is the frequency of affect males, while q^2 is the frequency of affected females, and $q > q^2$.

80

Exercise

Green-type colour blindness is X-linked, with q ≈ 0.05 in Western Europeans. What is the ratio of affected males to affected females?

81

Summary

- Hardy-Weinberg equation allows us to determine predicted genotype frequencies given allele frequencies
- HW is a useful null hypothesis for testing for the presence of a variety of evolutionary processes
- Differences in allele ratios in males and females and X-linkage can lead to deviations from HW – however, if the assumptions are met, HW will return once random mating resumes.

82