

TOPIC 2B. QUANTIFYING POPULATION GENETIC VARIATION

Allele frequency is the proportion of all alleles at a locus that are of a specified type, within a population or a sample of a population. Sometimes the term ‘gene frequency’ is used as a synonym.

Genotype frequency : the proportion of each genotype at a locus , within a population or a sample of a population.

2.1. WHERE DOES GENETIC VARIATION COME FROM, AND WHY DOES IT PERSIST?

Q: What is the ultimate source of all genetic variation?

A:

Q: Given that that variation arises, what happens to it over time?

A:

Example:

A population of light brown deer mice is monomorphic for genotype *bb* at a locus controlling coat color. A mutation occurs and a single deer mouse is born with dominant melanic mutation: genotype *Bb*. Let’s start a new population with the black female mouse and a single light brown male. What will happen to the new *B* allele if mice of all colors have equal fitness (e.g., there is no selection on color)?

P: $Bb \times bb$ (female is written first)

F1:

Q: What’s the allele frequency in the F1 mice? What are the genotype frequencies?

A:

Q: What happens to the allele frequency after one generation of random mating among the F1 mice?

A:

Possible matings among F1 mice: $Bb \times Bb$; $Bb \times bb$; $bb \times Bb$; $bb \times bb$, all equally likely

F2: $1/4 (Bb \times Bb) \rightarrow 1/4 BB$; $1/2 Bb$; $1/4 bb$

$1/2 (Bb \times bb) \rightarrow 1/2 Bb$; $1/2 bb$

$1/4 (bb \times bb) \rightarrow \text{all } bb$

Frequency	Phenotypes	Genotypes
$(1/4 * 1/4)=1/16$	Black	BB
$1/8+1/4=3/8=6/16$	Black	Bb
$1/16+1/4+1/4=9/16$	brown	bb

Genotype frequency = relative proportion of genotype:

Proportion of BB genotype = $1/16 = 0.0625$
 Proportion of Bb genotype = $3/8 = 0.375$
 Proportion of bb genotype = $9/16 = 0.5625$
Total = 1.0

Allele frequency = relative proportion of the allele:
 Proportion of **b** allele = $0.5625 + 0.5 \cdot 0.375 = 0.75$
 Proportion of **B** allele = $0.0625 + 0.5 \cdot 0.375 = 0.25$
Total = 1.000

Allele frequency does not change after a generation of random mating.

Q: Will the allele frequencies change after another generation of random mating? Will the genotype frequencies change?

A:

One more generation of random mating: Random mating = random union of gametes

genotypes	gamete types	gamete contribution	gamete frequency
0.5625 bb	all b	0.5625 b gametes	0.75 M gametes
0.375 Bb	1/2 b 1/2 B	0.1875 b gametes 0.1875 B gametes	
0.0625 BB	all B	0.0625 B gametes	0.25 N gametes

F2 Generation		Eggs		F3 generation
Allele freq.	Sperm	0.75 b	0.25 B	Allele freq.
(p=0.75)	0.75 b	$p^2=0.5625$ MM	$pq=0.1875$ MN	$p'=0.5625+0.5 \cdot (0.375)=0.75$
(q=0.25)	0.25 B	$pq=0.1875$ MN	$q^2=0.0625$ NN	$q'=0.0625+0.5 \cdot (0.375)=0.25$

A: Allele frequencies do not change. Genotype frequencies do not change after the first generation of random mating.

2.1.a. The Hardy-Weinberg Principle

- 1) Genotype frequencies of zygotes are predictable from gamete frequencies, assuming random mating.
- 2) Allele frequencies do not change from generation to generation under this scenario.
- 3) Genotype frequencies do not change after the first generation of random mating and are given by the following rules:

p^2	$2pq$	q^2
bb	Bb	BB

- 4) $p+q=1$
 5) $p^2+2pq+q^2=1$

2.1.b. Important Implications of HW Principle

1. When an allele is rare, there are many more heterozygotes than there are homozygotes ($2pq > q^2$ if q is small). Say A and a are alleles and freq of $A=p=0.9$ & freq of $a=q=0.1$.

Proportion of heterozygotes= $2pq=2(0.9)*(0.1)$

Proportion of aa homozygotes= $q^2=(0.01)^2=0.01$

2. If the assumptions are met, then allele frequencies stay the same from generation to generation. That is, there is **no evolution!**

Assumptions:

- i) **Mating is random** wrt the gene of interest.
- ii) Allele frequencies are the **same in females and males**.
- iii) All genotypes have equal viability and fertility (**no selection**).
- iv) Mutation does not occur, or is so rare it can be ignored (**no mutation**).
- v) Migration into the population can be ignored (**no migration**).
- vi) The population is large enough that the allele frequencies do not change from generation to generation due to chance (**no random genetic drift**).

Deviations from these assumptions lead to allele frequency change=EVOLUTION!

Therefore, deviations from H-W tells us that one or more evolutionary forces is acting on the population

Q: How do we know if deviations from H-W are real, or if they are just due to chance?

Sample of blood types from British blood donors:

Genotypes	MM	MN	NN
	298	489	213
Alleles	M		N
	$(298+244.5)/1000$		$(213+244.5)/1000$

freq M = $p = 0.54$

freq N = $q = 0.46$

Is the locus at HW equilibrium?

Expected genotypic frequencies under H-W?

MM	MN	NN
p^2	$2pq$	q^2

χ^2 test for goodness of fit:

O	E	(O-E)²	(O-E)²/E
	d.f. = 1		

2.1.c. Sex-linked Traits

Males: **genotype frequencies = allele frequency.**

Genotype	Allele frequency	Genotype frequency
$X^A Y$	p	p
$X^a Y$	q	q

If 8% of men are color blind => allele frequency $q=0.08$. What proportion of women are color blind?

Genotype	Expected Frequencies
$X^A X^A$	p^2
$X^A X^a$	$2pq$
$X^a X^a$	q^2
Total	1.0000

So fewer than 1 out of 100 women are color blind.

2.1.d. More than two alleles

<u>AA</u>	<u>Aa</u>	<u>Aa'</u>	<u>aa</u>	<u>aa'</u>	<u>a'a'</u>
8	38	121	27	252	401

Allele frequencies:

$$p_i = P_{ii} + \frac{1}{2} \sum_{j=1}^n P_{ij}; \text{ where } P_{ii} \text{ is frequency of homozygote; } P_{ij} \text{ is freq. of heterozygote}$$

Expected genotype frequencies: p_{ii}^2 for homozygotes and $2p_{ij}$ for heterozygotes.

Q: Calculate allele frequencies and expected genotype frequencies from the following observed genotype frequencies:

<u>AA</u>	<u>Aa</u>	<u>Aa'</u>	<u>aa</u>	<u>aa'</u>	<u>a'a'</u>
8	38	121	27	252	401

Is the population in H-W equilibrium? Do this before next class period.

To get started: $8/847 + 1/2(38/847) + 1/2(121/847) =$

2.1.e. Different allelic frequencies in males and females (Grads only)

Read handout pp. 56-62 and do problems 3 and 4 from the handout.

2.1.f. For a challenge, read Section III in Ch.2 of the handout (Grads only)

2.2. MEASURES OF GENETIC VARIATION FOR DISCRETE PHENOTYPES

Discrete variation (visible, allozyme, RFLP, microsatellite, etc.)

Polymorphism refers to multiple loci. It is the proportion of all genetic loci that exist in more than one allelic form. Humans, $P=0.32$, *D. melanogaster*, $H=0.42$.

Heterozygosity refers to a single locus. It is the proportion of individuals that are heterozygous (as opposed to homozygous) for alternate alleles at a specific locus. Humans: average $H=0.06$, *D. melanogaster*, average $H=0.14$

$$\text{Observed Heterozygosity} = H_o = \frac{\text{Observed \# Heterozygotes}}{\text{Total \# of Individuals}}$$

Another measure of variation within a population is the **Expected (Hardy-Weinberg)**

$$\text{Heterozygosity} = H_E = 1 - \sum_i p_i^2 \text{ (one minus expected homozygosity).}$$

H_E is maximal when all alleles equally frequent.

Note for graduate students: If we are dealing with small sample sizes ($N < 50$), we need to use a correction:

$$H_E = \frac{2N}{2N-1} \left(1 - \sum_i \hat{p}_i^2 \right)$$