

## Population Genetics and the Hardy-Weinberg Principle

Most genetics research focuses on the structure of genes on chromosomes, the function of genes, and the process of genetic transmission from parent to offspring.

**Population genetics** instead focuses on the overall **gene pool** in a population of interbreeding organisms - that is, the frequency of all alleles of all genes in the population - and whether the gene pool may be changing across generations in a population. Population genetics examines the relationship among genotype frequencies in a population, allele frequencies in its gene pool, and factors that can change these frequencies over time.

### Genetic Equilibrium and the Hardy-Weinberg Principle

A population is in **genetic equilibrium** when allele frequencies in the gene pool remain constant across generations. A gene pool will be in equilibrium under the following conditions:

- the population is very large
- individuals in the population mate randomly
- there is no migration into or out of the population
- natural selection does not act on any specific genotypes
- males and females have the same allele frequencies [vs. individuals are diploid and reproduce sexually]
- no mutations occur

In 1908 Godfrey Hardy and Wilhelm Weinberg, working independently, specified the relationship between genotype frequencies and allele frequencies that must occur in such an idealized population in equilibrium. This relationship, known as the **Hardy-Weinberg principle**, is important because we can use it to determine if a population is in equilibrium for a particular gene.

### Population Genotypes and Alleles

The Hardy-Weinberg principle applies to individual genes with two alleles, a dominant allele and a recessive allele. A population with such a gene can be described in terms of its **genotype numbers** - the number of individuals with each of the three resulting genotypes - or in terms of the three **genotype frequencies**. The frequency of each genotype is the number of individuals in the population with that genotype divided by the total number of individuals in the population, as displayed in Table 1. (Note that total genotype frequencies sum to 1.0.)

**Table 1. Genotype numbers and genotype frequencies in a hypothetical population.**

| Genotypes | Number with these Genotypes | Genotype Frequencies |
|-----------|-----------------------------|----------------------|
| AA        | 180                         | 0.36                 |
| Aa        | 240                         | 0.48                 |
| aa        | 80                          | 0.16                 |
| Total     | 500                         | 1.0                  |

### Calculating Allele Frequencies

We can also describe a population, somewhat more abstractly, in terms of its **allele frequencies**. The frequency of an allele is defined as the total number of copies of that allele in the population divided by the total number of copies of all alleles of the gene. We can calculate population allele frequencies from genotype numbers.

The total number of dominant A alleles in our population equals 600, which is the sum of:

- the number of AA individuals times 2 (the number of A alleles per individual) =  $180 \times 2 = 360$
  - the number of Aa individuals (times 1, the number of A alleles per individual) + 240
- 600

The total number of all alleles of the gene equals 1000, which is 2 times the number of individuals in the population (since the individuals are diploid).

By definition, the frequency of the dominant A alleles in our population equals 600/1000, or 0.60.

We can calculate the total number of a alleles in the population, and divide by the total number of alleles:

$$(2 \times 80 + 240) / 1000 = 400 / 1000 = 0.4$$

Or, since the total of all the allele frequencies sums to 1.0 and since there are only two alleles, A and a, we can calculate the allele frequency by subtracting the A allele frequency from 1.0:

$$1.0 - 0.60 = 0.40$$

### Random Mating and the Hardy-Weinberg Principle

Godfrey Hardy's and Wilhelm Weinberg's insight was that when a population is in equilibrium, genotype frequencies can be calculated from allele frequencies. The key insight is that with random mating, the probability that each parent transmits a given allele to an offspring is equal to that allele's frequency in the population.

Suppose the frequency of a dominant allele A in the population is  $p$  (0.60 in our example above)  
 the frequency of a recessive allele a in the population is  $q$  (0.40 in our example above)

Table 2 shows that probability of each of the possible offspring genotypes when the population is in equilibrium. (Note for purposes of calculation clarity, the table divides the heterozygous Aa offspring into two subsets, those who inherit the dominant allele from their mother and those who inherit the dominant allele from their father.

**Table 2. Offspring genotype frequencies given hypothetical allele frequencies and random mating.**

| Potential Offspring Genotype | Allele Transmitted by Mother | Allele Transmitted by Father | Probability Allele is Transmitted by Mother | Probability Allele is Transmitted by Father | Probability these Alleles are Transmitted by Mother and Father |
|------------------------------|------------------------------|------------------------------|---|---|--|
| AA                           | A                            | A                            | 0.60  | 0.60  | $0.60 \times 0.60 = 0.36$                                      |
| Aa                           | A                            | a                            | 0.60  | 0.40  | $0.60 \times 0.40 = 0.24$                                      |
| aA                           | a                            | A                            | 0.40  | 0.60  | $0.40 \times 0.60 = 0.24$                                      |
| aa                           | a                            | a                            | 0.40  | 0.40  | $0.40 \times 0.40 = 0.16$                                      |

Notice that assuming random mating, Table 2 regenerates the genotype frequencies observed in Table 1. Table 3 displays the relationship between allele frequencies and genotype frequencies more abstractly. It shows the expected genotype frequencies, assuming the frequency of the dominant allele is  $p$  and the frequency of the recessive allele is  $q$ .

**Table 3. Offspring genotype frequencies given abstract allele frequencies and random mating.**

| Potential Offspring Genotype | Allele Transmitted by Mother | Allele Transmitted by Father | Probability Allele is Transmitted by Mother | Probability Allele is Transmitted by Father | Probability these Alleles are Transmitted by Mother and Father |
|------------------------------|------------------------------|------------------------------|---|---|--|
| AA                           | A                            | A                            | $p$   | $p$   | $p \times p = p^2$   |
| Aa                           | A                            | a                            | $p$   | $q$   | $p \times q = pq$  |
| aA                           | a                            | A                            | $q$   | $p$   | $q \times p = pq$  |
| aa                           | a                            | a                            | $q$   | $q$   | $q \times q = q^2$   |

The Hardy-Weinberg principle is derived from this table. It states that if the frequency of A =  $p$  and the frequency of a =  $q$ , and the population is in equilibrium then:

$$\text{The frequency of genotype AA} = p^2 \text{ (} 0.6 \times 0.6 = 0.36 \text{)}$$

$$\text{The frequency of genotype Aa} = 2pq \text{ (} 2 \times 0.6 \times 0.4 = 0.48 \text{)}$$

$$\text{The frequency of genotype aa} = q^2 \text{ (} 0.4 \times 0.4 = 0.16 \text{)}$$

Since there are only two alleles the sum of the three possible genotype frequencies equals 1.0, it follows that

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

Table 4 demonstrates Hardy's and Weinberg's key insight, that with random mating, the probability that each parent transmits a given allele to an offspring is equal to that allele's frequency in the population.

- The first column shows three genotype frequencies for females in a population. (The same analysis applies to males).
- The second column shows the probability an offspring will inherit a dominant A allele from each one of these three maternal genotypes,.
- The third column shows the probability an offspring will inherit a recessive a allele from each one of these three maternal genotypes.
- The bottom row in columns 2 and 3 show the total probability of inheriting an A allele or an a allele from a female assuming females are randomly selected in mating. As can be seen in this row, the probability of inheriting an A allele or an a allele indeed equals the corresponding allele frequency in the population.

**Table 4. The Key Insight Underlying the Hardy-Weinberg Principle.**

| Maternal Genotype / Frequencies | Probability of offspring inheriting A allele from mother's genotype | Probability of offspring inheriting a allele from mother |
|---------------------------------|---|--|
| AA 0.36                         | 0.36  | 0  |
| Aa 0.48                         | $0.48/2 = 0.24$   | $0.48/2 = 0.24$  |
| Aa 0.16                         | 0   | 0.16   |
| TOTAL                           | 0.60  | 0.40   |

### Testing for Genetic Equilibrium

We can check if a population is in genetic equilibrium by testing if the Hardy-Weinberg principle applies, as follows:

Given the population genotype numbers,

- (1) calculate the allele frequencies from the observed population genotype numbers.
- (2) calculate the genotype frequencies from the observed genotype numbers.
- (3) apply the Hardy-Weinberg principle to calculate the **expected** genotype frequencies from the allele frequencies in the population.
- (4) If the population is in Hardy-Weinberg equilibrium the observed genotype frequencies in step 2 will be (roughly) the same as the expected frequencies in step 3. (A Chi-Square test is used to determine if the observed and expected genotype frequencies are significantly different from each other or not.)