

GENOTYPE, PHENOTYPE AND GENE FREQUENCIES

Prof. Arjumand S. Warsy
Department of Biochemistry
College of Science, King Saud University, Riyadh

Introduction

Genotype is the genetic makeup of an individual at a particular locus.

Phenotype is the observed biochemical, physiological and morphological characteristics of an individual, as determined by his or her genotype and the environment in which it is expressed.

- Generally, it is easy to determine the phenotype in a population (without applying the molecular biology technology).
- From the incidence rates it is possible to calculate individual allele frequencies.
- This require knowledge of:
 - gene in a population
 - and - their transmission from generation to generation
 - e.g. single gene traits segregate in families and tend to occur in fixed and predicted proportion.

e.g. consider the blood group MN which has two codominantly expressed alleles. M & N, and give rise to 3 genotypes MM, MN and NN and the three corresponding phenotypes M, N and MN.

A specific mating of an MN male and MN female would result in three types of distinguishable offspring (M, MN and N) in the expected proportion $\frac{1}{4}M$, $\frac{1}{2}MN$ and $\frac{1}{4}N$.

This principle can be extended to mating within the population and the number can be converted to relative frequencies by dividing them all by the total number of observations.

e.g. in a population of 1419, the following phenotypes were observed:

Phenotype	Genotype	No. of people	Relative frequency (%)
M	MM	392	27.62
N	NN	320	22.55
MN	MN	707	49.82

Gene frequencies can be determined from the genotype and phenotype frequencies. Since each autosomal genotype consists of two alleles the observed frequency of the M allele in the above population is

$$\frac{(2 \times 392) + 707}{1419 \times 2} = 0.53$$

The frequency of N can be calculated in the same way:

$$\frac{(2 \times 320) + 707}{1419 \times 2} = 0.47$$

This can also be done, by subtracting: 0.53 from 1.0, since the relative frequencies of the two alleles must add up to 1.

A 'gene pool' contains all of the alleles at a particular locus for the entire population.

If the relative frequencies of the two alleles, in the 'gene pool' are known, and are : 0.53 for M and 0.47 for N and the population size is 1419. The relative proportion of the three genotypes can be obtained:

$$\begin{aligned} 0.53 \times 0.53 &= 0.281 \quad \text{for MM} \\ 0.47 \times 0.47 &= 0.221 \quad \text{for NN} \\ (0.53 \times 0.47) + (0.47 \times 0.53) &= 0.498 \quad \text{for MN} \end{aligned}$$

If the locus has three alleles, q, p and r the genotypic distribution can be determined from $(p + q + r)^3$.

Hardy-Weinberg Law and factors affecting the Hardy-Weinberg Equilibrium

Introduction

In an 'ideal' population (neglecting mutation, selection, gene flow and genetic drifts) the gene frequency and genotype frequencies remain constant from generation to generation. This is the Hardy-Weinberg Law, named after George Hardy, an English Mathematician and Wilhelm Weinberg, a German Physician, who independently formulated this law in 1908.

A population that demonstrates the basic features of the Hardy-Weinberg (HW) Law is said to be in Hardy-Weinberg equilibrium. The HW Law can be used to calculate the expected number of homozygotes using heterozygote frequencies.

The questions are: Why doesn't a dominant allele automatically increase in frequency to replace its recessive counterpart? Why doesn't deleterious recessive allele fall in frequency until it is eliminated from the population?

The Hardy-Weinberg Law explains the basis for the constant gene frequencies as an application of the binomial theory (first put forward by Sir Isaac Newton).

$$1 - p = q$$

$$p + q = 1$$

The frequencies of the possible combinations of $p + q$ in a series of 'n' trials are given by the expansion of $(p + q)^n$.

Boys and girls in 3 child families:

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

p^3 = __ of sibships contain 3 boys
 $3p^2q$ = __ contain 2 boys and 1 girl
 $3pq^2$ = __ contain 2 girls and 1 boy
 q^3 = __ contain 3 girls

Each new sib has equal probability of being male and female

Frequency of autosomal genes and genotypes

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

$$(p+q+r)^2 = p^2 + 2pq + 2pr + 2qr + q^2 + r^2$$

Frequencies of AD genes can also be determined from incidence data. In general, the AD homozygotes are very rare and can be ignored for the sake of calculations.

As almost all patients are heterozygous with one normal and one abnormal allele, the frequency of the mutant allele (q) is approx. half the disease incidence (2 pq or almost 2q since p is very nearly 1) e.g. familial hypercholesterolaemia (FH) occurs at an incidence of 1/500. The frequency of mutant FH gene is about half of this or 0.001.

Frequency of X-linked genes and genotypes

In the male since there is only 1 X chromosome, the frequency of the abnormal gene on the X chromosome can be calculated by dividing the No. of abnormal male with total male investigated. e.g. for colour blindness X^{cb} , male can be either X or X^{cb} . Female can be $X^{cb}X$, XX , $X^{cb}X^{cb}$. Frequency of colour blindness is much less in females compared to the males though the allele frequency is the same. (Table 2)

Factors affecting Hardy-Weinberg equilibrium

Several assumptions are made by HW law. These include:

1. Population investigated is characterised by random mating.
2. Locus under consideration exhibits constant mutant rate.
 - There is no selection for or against a particular phenotype.
 - Population is sufficiently large and so there is no random fluctuation of frequencies.
 - There is no change in the population structure by migration (this can change genes by increasing or decreasing the number of individuals).

If one or more of these assumptions does not hold true, the genotypes in that population may not be in Hardy-Weinberg equilibrium.

Factors affecting Hardy -Weinberg equilibrium

Several factors can disturb the Hardy-Weinberg equilibrium by influencing either the distribution of genes in the population or by altering the gene frequencies. These include the following:

1. Non Random Mating

- Random mating (Panmixis) is the selection of a partner irrespective of his/her genotype.
- Non-random mating can result in an increase in the frequency of homozygous state i.e. by
 - assortative mating - the tendency to choose partners who share similar characteristics of heights, racial origin.
 - Consanguinity - marriage between blood relatives who have at least one common ancestor. Consanguinity in a community results in a relative increase in frequency of affected homozygotes with a relative decrease in the heterozygote frequency.

2. Mutation rate:

If a locus shows a higher mutation rate, then the H-W equilibrium will be destroyed and there will be a steady increase in mutant allele. In a population found in H-W equilibrium fitness of affected individuals is balanced by the occurrence of new mutations. The spontaneous mutation rate (μ) varies with different loci.

3. Selection

In an 'ideal' population there is no selection for or against any particular genotype. Selection is a powerful tool for altering gene frequency. It may reduce (negative selection) or increase (positive selection) a particular phenotype and hence its genotype.

Genetic selection acts on the individual phenotype and either hinders or favours reproduction and thus propagation of the individuals genotype.

4. Small population size

In a large population the number of children produced by individuals with different genotypes, assuming no alterations in fitness for any particular genotype, will tend to balance out, so the gene frequency will remain stable.

However, in small populations, one allele could be transmitted to a higher proportion of offspring by chance, thus resulting in marked changes in allele frequency from one generation to the next, thus disturbing HW equilibrium.

This phenomenon is called "random genetic drift". If one allele is lost it is said to be

"extinguished" and the other allele is described as having become "fixed".

Small population may become "genetic isolates" due to religious, geographical or other reasons. In such populations due to admixing, recessive and late-onset dominant traits increase in frequency, then in the general population. Table 2 gives examples of some genetic isolates with high frequencies of certain autosomal disorders.

5. Gene flow (Migration)

- When new alleles are introduced into a population as a consequence of migration, and subsequent inter-marriages, this alters allele frequencies.
- The slow diffusion of alleles across a racial or geographical boundary is known as gene flow e.g. Blood group B is thought to have originated in Asia and slowly spread to other populations, particularly westward due to invasions.

$$\begin{aligned} \text{Allele frequency in Asia} &= 0.3 \\ \text{Allele frequency in W. Europe} &= 0.06 \end{aligned}$$

Validity of Hardy-Weinberg equilibrium

By applying HW equilibrium it is possible to calculate the expected number of genotypes in a population from the allele frequency. If the expected number is close to the observed number then the population is in HW equilibrium.

Example 1

Two allele system, A and a. Among 1000 individuals the following genotypes were observed:

$$\begin{aligned} \text{AA} &= 800 \\ \text{Aa/Aa} &= 185 \\ \text{aa} &= 15 \end{aligned}$$

Find out if the population is in HW equilibrium.

$$\text{Frequency of A (p)} = \frac{(2 \times 800) + 185}{2000} = 0.8925$$

$$\text{Frequency a (q)} = \frac{(2 \times 15) + 185}{2000} = 0.1075$$

Now calculate the expected number:

	<u>Expected</u>	<u>Observed</u>
AA = (p ² x 1000)	796.5	800
Aa/Aa = (p ² qx1000)	192	185
aa (q ² x 1000)	11.5	15

These expected and observed values are close and x² analysis can be applied to confirm whether the difference is statistically significant or not. The population is in HW equilibrium.

Example 2

$$\begin{array}{l} \text{Total No.} \qquad \qquad \qquad 1000 \\ \text{BB} \qquad \qquad \qquad = \qquad 430 \\ \text{Bb/Bb} \qquad \qquad \qquad = \qquad 540 \\ \text{bb} \qquad \qquad \qquad = \qquad 30 \\ \text{Frequency of B:} \qquad \frac{(2 \times 430) + 540}{2000} = 0.7 \end{array}$$

$$\text{Frequency of b:} \qquad \frac{(2 \times 30) + 540}{2000} = 0.3$$

<u>Genotype</u>	<u>Expected</u>	<u>Observed</u>
BB	490 ($p^2 \times 1000$)	430
Bb/Bb	420 ($2pq \times 1000$)	540
bb	90 ($q^2 \times 1000$)	30

This population has increased number of heterozygote and of BB homozygotes and decreased number of bb homozygotes. So the population is not in HW equilibrium and it is necessary to look for possible factors that alter this equilibrium.

Table 1: Frequency of colour blindness in a male and female population

Sex	Genotype	Phenotype	Incidence (Approx.)
Male	X^+ X^{cb}	Normal Colour blind	$p = 0.92$ $q = 0.08$
Female	X^+/X^+ X^+/X^{cb} X^{cb}/X^{cb}	Normal (<u>homo</u>) Normal hetero. Normal (<u>total</u>) Colour blind	$p^2 = (0.92)^2 = 0.846$ $p^2q = (0.92)(0.08) = 0.147$ $p^2 + p^2q = 0.993$ $q^2 = (0.08)^2 = 0.0064$

Table 2: Genetic isolates with high frequencies of certain disorders

Isolate	Disorder
Hopi Indians	Albinism
Cuna Indians of Panama	Albinism
Finns	Congenital nephrotic syndrome
Eskimos	Congenital adrenal hyperplasia
Afrikaners	Huntington disease; lipoid proteinuria, Hypercholesterolaemia, Variegated porphyria
Amish	Chondroectodermal dysplasia