# POPULATION GENETICS (ABG 705) 2 UNITS

## Introduction

Most important genetic principles are derived from specific mating and genetics of an individual or individuals resulting from such matings. Here genes are thought of in terms of pairs. In population genetics, genetics as related to a group or population is considered in contrast to genetics of individuals. Hence, *population genetics* could be otherwise considered as '*genetics of populations of individuals*' which is sometimes termed statistical genetics, in which case genes are thought of in terms of pools. Certain principles are involved in population genetics that do not necessarily apply to individuals.

*Population*: Total of all individuals in a breed or species inhabiting an area. A population in genetic sense refers to:

1. The genetics constitution of one such population which is fully described by the proportion or percentage of individuals that belong to a particular genotype; and
2. The transmission of genes from one generation to the next.

E.g. Given an autosomal locus, **A** with 2 different alleles $\mathbf{A_1}$ and $\mathbf{A_2}$ present among individuals, there would be three possible genotypes ($A_1A_1$, $A_1A_2$ and $A_2A_2$). The frequency of a particular genotype being its proportion or percentage among the individuals is called *genotype frequency*. In the transmission, the genotypes of the parents are broken down and new set of genotypes are constituted in the progeny from genes transmitted in the gametes. The genes carried by the population thus have continuity from generation to generation, but the genotypes in which they appear do not. The genetic constitution of a population (genes carried) is described by array of gene frequencies. i.e. by specification of the alleles present at every locus and the numbers or proportions of the different alleles at each locus. The frequencies of all alleles at any one locus must add up to unity or 100%.

The gene frequencies at a particular locus among a group of individuals can be determined from the knowledge of genotype frequencies.

Assuming 100 individuals are classified under 3 genotypes as follows:-

| **Genotype** | **No. individuals** | **No of $A_1$ genes** | **A2 genes** | |
|---|---|---|---|---|
| $A_1A_1$ | 30 | 60 | 0 | Freq. A1 = $\frac{120}{200}$ = 0.6 or 60% |
| $A_1A_2$ | 60 | 60 | 60 | |
| $A_2A_2$ | 10 | 0 | 20 | Freq. A2 = $\frac{80}{200}$ = 0.4 or 40% |
| Total | 100 | 120 | 80 | |

*(200 i.e. 120+80 in the same sample)

*Each individual contains 2 genes, making 200 representatives of the genes at this locus.

To express the relationship in a more general form, let the frequencies of the genes and the genotypes be as follows:

| | **Genes** | | **Genotypes** | | | |
|---|---|---|---|---|---|---|
| | $A_1$ | $A_2$ | $A_1A_1$ | $A_1A_2$ | $A_2A_2$ | |
| **Freqs**: | p | q | P | H | Q | |
| **Total** | p + q =1 | | P + H + Q | = | | 1 |

Since each individual contains 2 genes, the frequency of $A_1$ genes:

= P + ½ H

$A_1$ : p = ½ (2P + H)

$A_2$: q = Q + ½ H

*The relationship between gene and genotype frequencies is of utmost importance because most of the deductions of population and quantitative genetics rest on it. If gene frequencies of 2 alleles among parents are **p** and **q**, then the genotype frequency among the progeny will be $\mathbf{p^2 + 2pq + q^2}$.

Assignment:
In a herd of short horn cattle, 50 are red, 40 are roan and 10 are white.

i. What are the frequencies of the red and white genes in the herd?
ii. If a single gene is drawn from the population at random, what would be the probability of getting a red gene?
iii. What is the probability that a sperm from the population carrying a red gene will fertilize an egg carrying a red gene?
iv. Probability of 2 individuals carrying the white genes uniting at fertilization.
v. Probability that a sperm carrying a white gene fertilizes an egg carrying a red gene.

## 1. Darwin's evolution

***Charles Darwin 1809 – 1882***

The best known proponent of the theory of evolution was an English naturalist (*Charles Darwin*). He was not the father of the idea of a gradual progressive change, but just an attending physician who brought the concept into the world. Most scholars had abandoned the notion of fixed species since their origin in the grand creation of life, long before publication of Darwin's "The origin of species" in 1859. By that time, most biologists agreed that new species arise through some process of evolution from older species. Butler said "Darwin's chief glory is not that he discovered evolution but that he made men to believe in it". He sailed the world from 1831 – 1836, studied the rich fauna and flora of South America and several islands. He was struck by the manner in which animal types shade one another and by the distinctive forms found on separate islands. He had believed in separate creations, but during the course of his voyage became convinced of mutability of species.

Darwin's theories involved three principles:
1. *Principle of Variation*:
Organisms vary i.e. are not exactly like their parents. There is variation in morphology, physiology and behaviour.

2. *Principle of Heredity*:
These variations are or may be hereditary i.e. (passed onto descendants) which resemble their parents more than unrelated individuals.

3. *Principle of Selection*:
There is continual struggle for existence in nature due to the dynamics of reproduction in all species. Some forms are more successful at surviving and reproducing than other forms in a given environment. Those best adapted will be the most likely to survive and should therefore leave the most descendants. He thought that the working of the above 3 principles accounted for the great variety of forms of plants and animals that now inhabit the earth.

Darwin accepted evolution as a working hypothesis, accepted the obvious facts of variation and hereditary transmission of potentialities and adduced a tremendous amount of data that seemed to support his thesis that natural selection (the survival of the fittest) is the leading factor in the drama of survival. i.e. Evolution of the species as a whole results from the differential rates of survival and reproduction of various types, so that the relative frequencies of the types change over time. Evolution in this view is a sorting process whereby living organisms change from one form into another. Most of these changes are considered to be produced gradually over long periods of geological time.

*Evidence of evolution*
1. *Classification or Taxonomy* – The fact that the 800,000 described species of animals can be arranged into 11 phyla and the 250,000 known plant species in 4 phyla, the whole resembling a branching tree.
2. *Comparative embryology* – The fact that all embryos starts as single-celled zygotes, progress along similar lines of development. Those of higher species pass hurriedly through stages which are the end stage of organisms lower down the evolutionary scale.
3. *Comparative anatomy* - brain, heart and other organs show a progressive development from lower to higher species.
4. *Vestigial structures* – Man during his embryonic life develops a set of gill clefts and arches, a trail and a fairly heavy covering of hair, all of which disappear before birth.
5. *Paleontology* – Fossil remains of animals show a progressive development from lower to higher in ascending strata of the earth's crust.
6. *Zoogeography* – species grow increasingly divergent in form and widening circles from their point of origin.
7. *Blood tests* – Blood of closely related animals show small incompatibilities as measured by the amount of cellular agglutination whereas that of distantly related ones shows great incompatibilities.
8. *Observation and experiment* – The fact that all the breeds of farm animals have been developed by a process of selection during about the past 200 years and that many new breeds are in the process of formation.

# 2. Variation/evolutionary changes

Population genetics deals with the study of inherited variations and its modulation in time and space. It deals with both genotypic and phenotypic variation. The occurrence of hereditary variation in natural populations was the starting point of Darwin's argument for evolution by a process of natural selection. Some natural hereditary variations may be more advantageous than others for the survival and reproduction of individual carrying them. Some variants increase the chances of survival and reproduction of their carriers relative to the carriers of other variants. As a result, useful variations gradually becomes more prevalent over the generations while harmful or less useful ones are eliminated, a process called natural selection which plays a leading role in evolution. Selection against recessives is sometimes called purifying selection because it purifies the gene pool against detrimental alleles. It has been shown that the rate of increase in fitness of a population at any time is equal to its genetic variance in fitness.

The fundamental theorem applies to *allelic variation at a single gene locus*. The greater the number of variable loci and the more alleles there are at each variable locus, the greater the possibility for change in the frequency of some alleles at the expense of others.

*Evidence of genetic variation*

1. Natural populations possess a great deal of genetic variation. This has been observed through mating of close relatives or *inbreeding* with recessive genes becoming exposed, which affects fitness of organisms carrying them.

2. Variant forms of proteins have been defected by *gel electrophoresis* (allows separation of different proteins) e.g. The 3 B-globin types in humans have different banding patterns.

3. Subsequent generations of individuals *selected* for breeding for certain economically important traits change in direct of selection, indicating that original population had genetic variation with respect to the selected trait. E.g. Egg production in a flock of White Leghorn chicken increased from 125.6 eggs/hen/year in 1933 to 249.6 in 1965. Artificial selection has been successful for many economic traits in many domesticated species (cattle, pigs, sheep, goat, poultry, maize, rice, wheat etc).

4. Variation in *chromosomal structure* – has been observed among different individuals such as deletions, replications and inversions.

***Measures of genetic variations***

1. *Proportion of polymorphic loci:*

Using the allelic frequencies calculated for a given gene, a gene can be defined as polymorphic (having many forms – 2 or more alleles in substantial frequency) or monomorphic (having only one allele in high frequency). The simplest measure of genetic variation is to categorize loci as polymorphic or monomorphic and then to calculate the proportion of all polymorphic loci. A gene is considered monomorphic if

the frequency of the most common allele is 0.99 or greater (if all other alleles combined have a frequency of 0.01 or less). E.g. Sickle cell gene in a sample of 400 individuals from West Africa is polymorphic for 2 alleles because the frequency of HbA < 0.99.

Ex. Given a sample of 400 individuals, 320 with normal genotype (HbA HbA), 72 carriers (HbA HbS) and 8 (mostly young individuals) had sickle cell disease (HbS HbS).

**P** = $\frac{320}{400}$ = 0.8; H = 0.18, Q = 0.02

**p** = 0.8 + ½ (0.18) = 0.89; **q** = 0.02 + ½ (0.18) = 0.11

If 30 loci are examined (by electrophoresis) and 15 of the loci in a population are polymorphic, then the degree of polymorphism in that population = 15/30 = 0.5. Estimates from several samples or sub populations can then be averaged to obtain a measure of variation in the overall population.

2. *Heterozygosity of the population:*
A second approach of measuring genetic variation is to calculate average heterozygosity over all loci. That is, average frequency of heterozygous individuals overall loci in a population. This takes into account all levels of genetic variation rather than just classify loci into 2 categories. Average heterozygosity for a sample of *n* loci is given as:

$H = 1/n (H_A + H_B + H_C + \ldots +)$ where $H_A$ is heterozygosity for gene A, and so on.

If $H_A = 0.5$, $H_B = 0.0$, $H_C = 0.1$, $H_D = 0.0$, $H_E = 0.2$ and n =5,
Then, $H = 1/5 (0.5 + \ldots + 0.2) = 0.16$ i.e. 16% of the loci in a given individual are heterozygous for the variants.
**Note**: 3 of the above loci are polymorphic (A, C & E) i.e. have the most common allele with a frequency of 0.99 or less.

Proportion of polymorphic loci = 3/5 = 0.6

*Heterozygosity is a good measure because it estimates the probability that 2 alleles taken at random from the proportion are different.

Surveys of the genetic variations in various species demonstrate that up to 20% of the genes examined by electrophoresis are polymorphic and that 10% of these genes are heterozygous in a given individual.

# 3. Hardy-Weinberg Equilibrum

Several factors are knwn to have the potential for profoundly changing the genetic structure of he population. The factors are violations of Hardy –Weinberg condition and include migration, mutation, selection and genetic drift. The first three factors are systematic processes and tend to change gene frequency in a systematic manner both in amount and in direction. Genetic drift is a dispersive process and is predictable in amount but nto indirection. Other , factors include population size (gene frequencies are subject to sampling variation between successive generations) and mating system (gene frequencies may change as a result of assortative mating).
When Hardy –Weinberg condition exist the population remains in equilibrium and no change is possible.

Hardy –Weinberg law
It states that in a large random mating population with no migration, mutation, selection and genetic drift, the gene frequencies and genotype frequencies remain constant from generation to generation, and furter more, there is a simple relationship between gene frequencies and genotype frequencies.

If the gene frequencies of 2 alleles (A, a) among the parents are p and q respectively, then the genotype frequencies among the progeny are P2 + 2pq + q2. This shows that Hardy – Weinberg proportions are really binomial proportions and that genotypic frequencies sum to unity. This relationship refers to autosomal genes that segregate normally in gametogeneis and the gene frequencies are the same in males and females.

To illustrate the effect of different allelic frequencies on genotypic frequencies, the figure below given frequencies of the 3 different genotypes, assuming H –W proportions for the total range of allelic frequencies. Heterozygote is the most common genotype for intermediate allelic frequencies, while one of the homozygotes principle: Genotypic frequencies for a gene in 2 different alleles is a binomial function of the allelic frequencies.

When Hardy – Weinberg conditions exist, the population remains in equilibrium and no change is possible. Several factors are known to have the potential for profoundly changing the genetic structure of the population. The factors are violations of the H – W conditions and include migration, mutation, selection and genetic drift.

First 3 factors (migration, mutation and selection) are systematic processes which tend to the fgs in a predictable manner both in amount and in direction. The latter factor (genetic drift) is a dispersive process and is predicable in amount but not in direction.