Lecture 9: Population Genetics

Plan of the lecture

I. Population Genetics: definitions
II. Hardy-Weinberg Law.
III. Factors affecting gene frequency in a population. Small populations and founder effect.
IV. Rare Alleles and Eugenics

The goal of this lecture is to make students familiar with basic models of population genetics and to acquaint students with empirical tests of these models. It will discuss the primary forces and processes involved in shaping genetic variation in natural populations (mutation, drift, selection, migration, recombination, mating patterns, population size and population subdivision).

I. Population genetics: definitions

Population – group of interbreeding individuals of the same species that are occupying a given area at a given time.

Population genetics is the study of the allele frequency distribution and change under the influence of the 4 evolutionary forces: ***natural selection, mutation, migration (gene flow), and genetic drift***. Population genetics is concerned with gene and genotype frequencies, the factors that tend to keep them constant, and the factors that tend to change them in populations.

All the genes at all loci in every member of an interbreeding population form ***gene pool***. Each gene in the genetic pool is present in two (or more) forms – ***alleles.***

Individuals of a population have same number and kinds of genes (except sex genes) and they have different combinations of alleles *(phenotypic variation).*

The applications of Mendelian genetics, chromosomal abnormalities, and multifactorial inheritance to medical practice are quite evident. Physicians work mostly with patients and families. However, as important as they may be, genes affect populations, and in the long run their effects in populations have a far more important impact on medicine than the relatively few families each physician may serve.

Why is population genetics important for human health and diseases?

(a) Human evolution has been underlain by adaptive and non-adaptive changes in gene frequencies
(b) Diseases are due to effects of alleles interacting with environments - is a spectrum from single-locus disorders to polygenic disorders

II. Hardy-Weinberg Law

In 1908 German physician *Wilhelm Weinberg* (1862-1937) and the British mathematician *Godfrey.H. Hardy* independently formulated the *Hardy-Weinberg principle*. They pointed out that under ideal conditions it could easily predict genotype frequencies from allele frequencies, at least for a diploid sexually reproducing species such as humans.

For many human autosomal recessive traits the heterozygote cannot be distinguished from the normal dominant homozygote. When this occurs the Hardy-Weinberg equilibrium is assumed to apply. These authors used different approaches but came to the same conclusions. They made *several assumptions of the population*:

1. The population follows the Laws of Mendelian genetics
2. Size of the population approaches infinity
3. There is no effect of recurrent mutation
4. There is no migration in or out of the population
5. There is no selection against any phenotype
6. Mating is totally random

Under these assumptions, Hardy and Weinberg found that the gene frequency and the genotype frequency in the population do not change from generation to generation.

Furthermore, if the frequency of the dominant allele **A** in the founding population was **p** , and the frequency of the recessive allele **a** in the founding population was **q**, then after one generation of random mating the genotype frequencies would remain fixed and would be in the ratio:

p^2 (AA) 2pq (Aa) q^2 (aa)

Since there are only two alleles in the population,

$p(A)+q(a) = 1$ *and* **$p^2(AA) + 2pq(Aa) + q^2(aa) = 1$**

Thereby, Hardy-Weinberg law is fundamental principle in population genetics stating that ***the genotype frequencies and gene frequencies of a large, randomly mating population remain constant provided migration, mutation, natural selection and genetic drift do not take place.***

Importance of "H-W" Formula

Example: If the incidence of PKU is 1 in 10,000 live births, and it is a recessive condition then:

$q^2 =$ 1/10,000 = 0.0001

$q =$ $\sqrt{}$ 0.0001 = 0.01

Where *q* is the gene frequency for PKU.

The frequency of heterozygote carriers is *2pq* as:

$p + q = 1$

$p = 1 – 0.01 = 0.99$

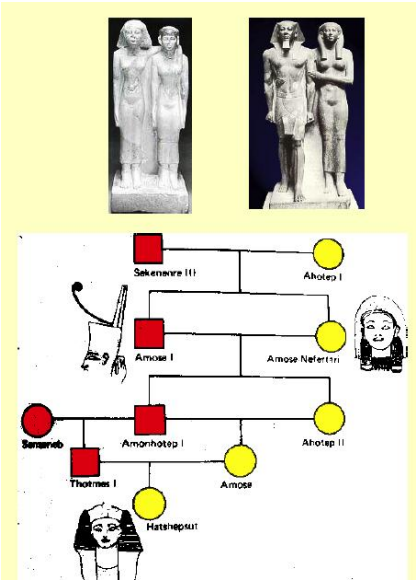
$2\,pq = 2\,(0.99 \times 0.01) = 0.0198$

III. Factors Affecting Gene Frequency in a Population:

1. Non-random mating.
2. Altered mutation rate.
3. Natural selection.
4. Small populations.
5. Migration.

1. *Non-random mating* – decreases the proportion of heterozygotes, violates assumption of Hardy-Weinberg equilibrium.

1. ***Assortative mating*** – non-random mating based on phenotypic similarities (in humans pair bonds established by religion, ethnic background, professional interest, etc.) **Assortative mating** (also called **assortative pairing**) takes place when sexually reproducing organisms tend to mate with individuals that are like themselves in some respect (*positive assortative mating*) or dissimilar (*negative assortative mating*). In evolution, these two types of assortative mating have the effect, respectively, of reducing and increasing the range of variation, or trait variance, when the assorting is cued on heritable traits. Positive assortative mating, therefore, results in disruptive natural selection, and negative assortative mating results in stabilizing natural selection.
2. ***Inbreeding*** – breeding between close relatives. If practiced repeatedly, it leads to an increase in homozygosity of a population. A higher frequency of recessive, deleterious traits in homozygous form in a population can, over time, result in inbreeding depression. This may occur when inbred individuals exhibit reduced health and fitness and lower levels of fertility.



Coefficient of Inbreeding is the probability that an individual being homozygous for a certain locus by receiving both alleles of this locus from ancestral source. Offspring of first-cousin marriage is homozygous at 1/16 of his loci.

Genetic consequences of inbreeding

1. Reduction in genetic variability
   a. Increase in frequency of homozygous recessives
   b. Recessive alleles exposed to natural selection
2. Social consequences
   a. Concentration of wealth and power
   b. Reduction of social ties with other groups

Example: Inbreeding in the Egyptian Royal Family (see a picture)

i. *Social advantages that sometimes outweigh its genetic cost.*
ii. *The preferred brother-sister marriages of ancient Egypt, some kings married their own daughters*, e.g., Amenhotep the Third married his

daughter Sitamun and Akhenaten married two of his daughters, Meritaten and Ankhesenpaaten.
When Akhenaten was presumably married to Kiya, mother of Tutankhamun, she was given the position of the "Wife and Great Beloved One" and not a "King's Wife."
Ramesses the Second also married two daughters, Meritamun and Bentanta.

iii. *The served to maintain political power of an elite dynasty*

3 If prevalent in a population can disturb "H-W" equilibrium by increasing the proportion of homozygotes at the expense of heterozygotes.

2. *Altered mutation rate*

Mutation (heritable changes in DNA) give rise to new alleles. Mutation rate is low: for a single locus the average frequency of mutation is about 0.0001. They may be lethal, neutral, or advantageous. Mutations are ultimate source of genetic variation. If mutation is advantageous the natural selection favours it and allele frequency changes.

Eradication of dominant disorders.

Huntington's disease (and any other dominant disorder) could in principle be eliminated in one generation by aborting every foetus carrying the gene. However, this would not prevent spontaneous mutations occurring (in Huntington's, ±1 in 100,000) unless the entire population was screened for them.

3. *Natural Selection*

Natural selection is process by which a particular genotype/phenotype combination confers an advantage to individuals in a population, thereby determining its survival and reproduction. Natural selection is often geographically-restricted.

Genetic selection acts on the individual **phenotypes** and either favours or hinders reproduction and thus the propagation of that individual's **genotype**. Natural selection acts by modifying an individual's *biological fitness*.

For an *autosomal dominant trait*, any increase in fitness, will rapidly alter the gene frequency over the next few generations to a new equilibrium.

Selection against a *recessive* genotype is less effective and results in a slow change in gene frequencies.

For *X-Linked recessive trait*, the situation is intermediate between autosomal dominant and recessive.

Types of Natural Selection

1. *Directional Selection* - shift in the variation in a consistent direction within the phenotypic range (e.g., antibiotic resistance in bacteria)
2. *Stabilizing Selection* - loss of extreme forms with stabilization of an intermediate form **Example:** *stabilizing selection* on human birth weight - limitation on infant size is the size of the birth canal. Infants greater or less than 3.4 kg (7.5 lbs) have increased mortality.
3. *Disruptive Selection* favors individuals at the extremes with a reduction of intermediate forms (e.g., gender in human)

Fitness. A fundamental concept in evolutionary theory is "***fitness***", which can be defined as the ability to survive and reproduce. Reproduction is a key: to be evolutionarily fit, an organism must pass its genes on to future generations. Basic idea behind evolution by natural selection: *the more fit individuals contribute more to future generations than less fit individuals*. Thus, the genes found in more fit individuals ultimately take over the population.

Natural selection requires 3 basic conditions:

i. there must be inherited traits.
ii. there must be variation in these traits among members of the species.
iii. some inherited traits must affect fitness

Heterozygosity and Polymorphism. Sickle cell anemia - red blood cell protein polymorphism

The existence in a population of more than one form of a particular trait or suite of related traits is known as *polymorphism*.

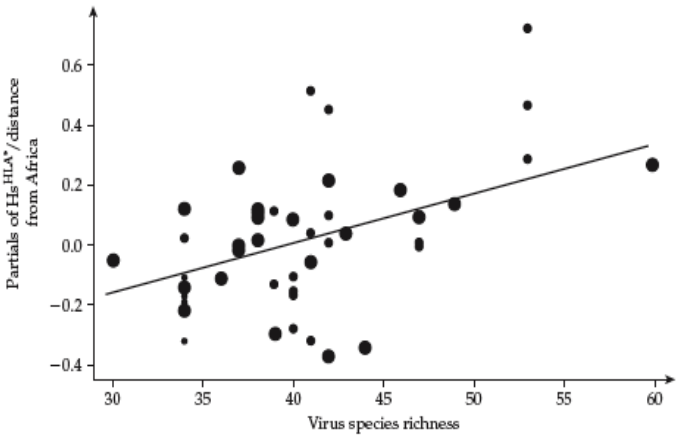
In population genetics, this usually refers to the different phenotypes resulting from different alleles at a particular locus. The simplest description of populational variation at a single locus is *relative genotype frequency.*

Polymorphism (Greek: *poly* = many, and *morph* = form) occurs when two or more clearly different phenotypes exist in the same population of a species – in other words, the occurrence of more than one *form* or *morph*.

A gene is called "*polymorphic*" if there is more than 1 allele present in at least 1% of the population. Genes with only 1 allele in the population are called "*monomorphic*". Some genes have 2 alleles: they are "*dimorphic*"

An example of a polymorphic gene is a gene, which determines the blood types of ABO system.

The most-polymorphic loci known in humans are *HLA loci*, which are involved in immune responses to pathogens. There is a positive correlation (see a figure), among human populations, between *HLA heterozygosity levels and virus species richness*, suggesting that viruses impose selection for maintenance of genetic variation at immune system loci.



Heterozygosity is the percentage of heterozygotes in a population.

Some genes exist at a rather high frequency in the population because the heterozygote is more fit than either homozygote.

The only documented example of this is sickle-cell anemia in Western Africa. There are three major genotypes for the sickle cell locus, each producing a different phenotype, in West Africans:

1. ***AA***, or normal individuals,
2. ***AS*** or heterozygote individuals (often called carriers),
3. ***SS*** individuals who will have sickle-cell anemia.

Without medical intervention, ***SS*** individuals will have fitness less than 1. In the tropical malarial environment of West Africa (i.e., where the unicellular parasite *Plasmodium falciparum* inhabits*)*, ***AA*** and ***AS*** individuals get malaria, but ***AS*** individuals usually have much milder cases of the disease and usually survive while ***AA*** individuals are less likely to do so. The heterozygote is the most fit phenotype of the three and ***S*** allele is only favored in malarial area.

Other red blood cell genes show similar patterns of heterozygote advantage (see a table below).

Table 2.1 Examples of red cell genes involved in malaria resistance, and which polymorphism worldwide may have been partly determined by the presence/absence of malaria

Cell component	Variant	Gene	Protein and function	Effect on malaria	Main distribution
Haemoglobin	Hb S	HBB	β-globin (haemoglobin component)	Protects against severe malaria	Africa, Middle East, India, Mediterranean
	Hb C	HBB	β-globin (haemoglobin component)	Protects against severe malaria	Africa
	Hb E	HBB	β-globin (haemoglobin component)	Reduces parasite invasion	Southeast Asia
	α-thalassemia	HBA	β-globin (haemoglobin component)	Protects against severe malaria	Africa, Mediterranean India, Southeast Asia
	β-thalassemia	HBB	α-globin (haemoglobin component)	Protects against severe malaria	Africa, Mediterranean, India, Southeast Asia, Melanesia
Red cell enzymes	G6PD deficiency	G6PD	Glucose-6-phosphate deshydrogenase (protects against oxidative stress)	Protects against severe malaria	Africa, Mediterranean, India, Southeast Asia
Red cell membrane	FY*O	FY	Duffy antigen (Chemokine receptor)	Protects against Plasmodium vivax[a]	Africa

[a] *Plasmodium vivax* is one of the agents of the human malaria. The others are *Plasmodium falciparum*, *Plasmodium malariae*, and *Plasmodium ovale*. The deadliest is *P. falciparum*.

Here, we see 'fit' between alleles and environments, and variation is maintained locally by heterozygote advantage.

There are other examples of infectious diseases in which the severity of infection can be directly linked to blood phenotype. It has been showed the frequencies of blood types A and B differ between

populations. It may have adaptive significance through susceptibility or resistance to infectious disease due to similarity of blood cell protein structure between pathogens (or vectors) and the human host. Antibodies created against the bacteria could react against blood cells of host. Antigenic similarity exist between the blood types and variety of bacterial, rickettsial and helminthic species, including typhoid, streptococci (blood type A), staphylococci (blood type O), bacteria *Shigella* causing dysentery and *Proteus* responsible for urinary and septic infections (see a picture).

Example: relatively high distribution of blood group A in Europe in the wake of Bubonic plague because O type is more susceptible.

Genotype		Adaptive significance?
AA	Homozygote	More susceptible to smallpox
BB	Homozygote	Resistant to syphilis, susceptible to infant diarrhea
OO	Homozygote	Resistant to syphilis, susceptible to Bubonic plague

Eradicating (as best possible) recessive inherited disorders in genetic isolates

In 2006, the researcher G.Cochran and his colleagues at the University of Utah's anthropology department (USA) published the scientific article "*Natural history of Ashkenazi intelligence"*.

Ashkenazi Jews (Ashkenazim) are descended from the Jewish communities of Germany, Austria, Poland, and Eastern Europe that date back to the X^{th} century. Today they make up around 80% of the world's Jews.

Ashkenazim have the highest average IQ of any ethnic group, scoring 12 to 15 points above the European average. They are also strongly represented in fields and occupations requiring high cognitive ability. For instance, European-origin Jews account for 27% U.S. Nobel science prize winners but make up only about 3% of the U.S. population.

But the group is also associated with neurological disorders, including Tay-Sachs, Gaucher's, and Niemann-Pick diseases.

Researchers at the University of Utah's investigated a possible link between these genetic illnesses and above-average intelligence in Ashkenazi Jews. They suggest both are *the result of natural selection* for enhanced brainpower. Tay-Sachs and other Ashekenazi disease-causing genes ***may*** promote IQ when heterozygous (by affecting brain development).

The high rate of a disease *thalassemia* in people of Mediterranean origin, the high rate of *sickle cell anemia* in people of West African descent, the high rate of *cystic fibrosis* in people from Western Europe, and the high rate of *Tay-Sachs disease* in Jewish ethnic groups from Eastern Europe may all owe their origin to environmental factors that cause changes in gene frequencies in large populations by giving some advantage to heterozygotes who carry a deleterious allele. Although one may never use the calculations of population genetics in medical practice, the underlying principles should be understood.

4. *Small Populations*

For religious, geographical (spatial, e.g., due to islands, mountains, and glaciers that isolate local populations), tribal or other reasons a small group of individuals may become genetically isolated from the rest of the population **(genetic isolates).**

By chance one allele may fail to be passed on to the next generation and so disappear **(extinction)** leaving only the alternative allele at that locus **(fixation).**

Gene frequencies in small isolate populations do not reflect those of the larger founding population from which they were derived because of two factors, ***founder effect*** and ***random genetic drift***.

Founder effect refers to the loss of genetic variation when a new colony is established by a very small number of individuals from a larger population, i.e. the population grew from a few founding individuals.

A few individuals cannot represent all of the genomes of the founding population. Each of us is carrying from 1 to 8 mutant genes in the heterozygous state, even though we are normal. When the founding population is small, intermarriage must result even though steps are taken to avoid it. The mutations carried by the founders are in higher frequency than they would be in the general population from which the founders came. Island populations founded by pirates or shipwreck that

were isolated for several generations tend to have different gene and genotype frequencies because of founder effect (read about the examples below).

Similarly, religious isolates, where marriage outside the religion is forbidden, also have founder effects.

As a result of the loss of genetic variation, the new population may be distinctively different, both genetically and phenotypically, from the parent population from which it is derived. In extreme cases, the founder effect is thought to lead to the speciation and subsequent evolution of new species.

Even if the founders of small isolate populations had exactly the same genotypes and gene frequencies of the original parent population, gene and genotype frequencies would change because of random ***genetic drift***. (***Genetic drift*** – random (by chance) change of allele frequencies from generation to generation). Random genetic drift occurs because a small population cannot maintain randomness.

Examples of geographical isolation:

1. Pingelap: Island of the Colorblind

Pingelap is an atoll in the Pacific Ocean, part of Pohnpei state of the Federated States of Micronesia, consisting of three islands: Pingelap Island, Sukoru and Daekae, linked by a reef system and surrounding a central lagoon, although only Pingelap Island is inhabited.

In 1775, a catastrophic typhoon swept through the island, killing 90% of the inhabitants and leaving only approximately 20 people. It is believed that one of the survivors, namely Nahnmwarki Mwanenised (the ruler at that time), was a carrier for complete achromatopsia (known on the island as maskun, meaning literally "not see" in Pingelapese), a recessive genetic disorder which causes total colour-blindness in sufferers[1]. All Achromats on this island nowadays can trace their ancestry to this male survivor. However, the Achromatopsia disorder did not appear until the fourth generation after the typhoon, where 2.70% of the Pingelapese were affected. By generation 6, the incidence rose to approximately 4.92%. These statistics can be accounted for by inbreeding and two related concepts: so called the bottleneck effect and genetic drift. In the case of Achromatopsia on the Pingelap Island, the Achromatopsia mutation fluctuated immensely from generation 3 to generation 4 under an extreme form of genetic drift. This type of genetic drift occurs only when the population is extremely small (20 survivors after typhoon) and is also known as the founder effect. Of course, both concepts occur due to inbreeding.

2. Tristan da Cunha

It is a group of remote islands in the south Atlantic Ocean, inbetween South Africa and South America (but closest to the former). It is a colony of Saint Helena which is 2173 kilometres to its north. This territory contains mostly Tristan da Cunha (the largest island in the territory - which is inhabited), as well as quite a few islands which aren't lived on. Unsurprisingly, the island was never inhabited prior to European discovery. It was discovered in 1506 by the Portuguese navigator, Tristão d'Acunha: he was unable to land, and named the island after himself.

A French survey was made of the island in the mid-18th century, and it was at this point that fresh water was discovered on the island, giving the island some value as a waypoint in Atlantic crossings. In the early-19th century an American settled the island, claiming it as his own and naming it the Island of Refreshment, before dying only a few years after he arrived. Shortly after his death the War of 1812 broke out between America and Britain, and America used Tristan da Cunha as a naval base to attack British ships on their way to the United States.

The island is one of the least inhabited places in the world, with a population of around 270 people. This small population is further complicated by the incredible isolation of the island. With only 80 families making up the entire social group of the island, many young people leave the islands to find spouses, and eventually hope to return.

Examples of Founder Effect

1. Amish.

The Amish are a group descended from 30 Swiss founders who renounced technological progress. Most Amish mate within the group. One of the founders had ***Ellis-van Crevald syndrome***, which causes short stature, polydactily (extra fingers and toes), and heart defects. Today about 1 in 200 Amish are homozygous for this syndrome, which is very rare in the larger US population.

Note the effect inbreeding has here: the problem comes from this recessive condition becoming homozygous due to the mating of closely related people!!!

2. Pitcairn Islanders

The **Pitcairn Islands** are a group of four volcanic islands in the southern Pacific Ocean, the second largest and measuring about 3.6 km (2.2 mi) from east to west, is inhabited. The islands are inhabited by the descendants of the mutineers from the ship *Bounty,* commanded by Captain William Bligh, and the Tahitians (or Polynesians) who accompanied them. Now a genetically unique population of Pitcairn numbers about 67 inhabitants (9 families) and only 46-48 people are permanent residents. According to population census of 2011, it is the least populous jurisdiction in the world.

3. The Dunkers.

It is a religious isolate in the founded in by about 50 families from Germany who settled in Pennsylvania (USA) in 1719. They have frequencies of alleles for ABO blood type bent little fingers, hitchhikers thumb and other traits differ from those in place of origin or area of Pennsylvania where they now live. These differences may also be explained by *founder effect*!

5. Migration

Until recently most people stayed in the place they were born and married someone nearby. The last half of 20th century was characterizes by breakups of local population isolates. This major change in breeding patterns has resulted in increased in genetic heterogeneity.

Migration *is the movement of individuals* ***in*** *or* ***out*** *of a population*.

Migration is necessary to keep a species from fragmenting into several different species. Even as low a level as one individual per generation moving between populations is enough to keep a species unified. Migration can be thought of as combining two populations with different allele frequencies and different numbers together into a single population. After one generation of random mating, the combined population will once again be in H-W equilibrium.

Gene flow – movement of alleles into and out of populations by migration of individuals. It balances genetic differences that arise through mutation. Genetic flow also helps keep separated populations genetically similar (ie. the same species.

Genetic drift is random (by chance) change of allele frequencies from generation to generation. It leads to homozygous condition – loss of genetic diversity. The effect of genetic drift is greater in small than large populations. Migrant individuals will modify the gene pool of their descendents.

As population size increases the effects of genetic drift decrease. For example, the chance of completely loosing an allele (fixation) by chance is great in a population of 25 individuals. In a population of 500 individuals, the probability that one of the alleles will be lost by chance is low and, in the absence of other influences (natural selection etc.) they tend to be maintained in the population at constant relative frequency.

IV. Rare Alleles and Eugenics

Population genetics is also the most widely misused area of human genetics, sometimes bordering on "vigilante genetics," a term coined by Newton Morton. Persons have mistakenly applied population genetics to "prove" race superiority for intelligence and aptitudes, and have misused it in eugenics.

A popular idea early in the XXth century was “Eugenics”, improving the human population through selective breeding. The first thorough exposition of eugenics was made by British scientist *Francis Galton*, who in his book “*Hereditary Genius”* (1869) proposed that a system of arranged marriages between men of distinction and women of wealth would eventually produce a gifted race.

Eugenics is an attempt to apply genetic knowledge directly for the improvement of human existence. It is the study of or belief in the possibility of improving the qualities of the human species or a human population, especially by such means as discouraging reproduction by persons having genetic defects or presumed to have inheritable undesirable traits **(negative eugenics)** or encouraging reproduction by persons presumed to have inheritable desirable traits **(positive eugenics).**

The assumptions of eugenicists came under sharp criticism beginning in the 1930s and were discredited after the German Nazis used eugenics to support the extermination of Jews, blacks, and homosexuals during World War II. We no longer force “genetically defective” people to be sterilized. However, note that positive eugenics (encouraging people to breed with superior partners) is still practiced in places.

The problem with sterilizing “defectives” is that most genes, which produce the notable genetic diseases, are recessive and the genes remain unexpressed in heterozygotes. If you only sterilize the homozygous individuals who manifest the genetic “defects”, you are missing the vast majority of people who carry the allele.

For example, assume that the frequency of a gene for a recessive genetic disease is 0.001, a very typical figure.

Thus $p(A)$ = 0.999 and $q(a)$ = 0.001, and p^2 = 0.998, 2pq = 0.002, and q^2 = 0.000001.

The ratio of heterozygotes (undetected carriers) to homozygotes (people with the disease) is 2000 to 1: you are sterilizing only 1/2000 of the people who carry the defective allele.

This is simply not a workable strategy for improving the gene pool.

Population Genetics: Importance

Patterns of genetic variation shed light on recombination, demography, admixture, and evolutionary selection in the human population. In turn, knowledge about human population history helps inform studies in Medical Genetics.