Structure of populations, genetic drift, importance for evolution

Population structure (or “population subdivision”) – instead of a single, simple population, populations are subdivided in some way.

Metapopulation - the overall "population of populations"

Subpopulations (local populations, or demes) - the individual component populations (amongst which gene flow is restricted)

Exemple: i) European Caucasians would be a (meta)population, with groups within them defined by nationality or religion would be the subpopulations, ii) the citizens of India may be defined as a population, and the different caste groups as its subpopulations.

In fact, in many real populations, there may not be any obvious substructure at all, and the **populations are continuous**. However, even in effectively continuous populations, different areas can have different allelic frequencies, because the **whole metapopulation is not panmictic**. Among humans, many **populations are structured**, but continuously, in space.

Population are structured when they have deviations from H-W proportions, or deviations from panmixia. If there is inbreeding, or selection, or if migration is important, then populations can be said to be structured in some way. Gene flow (migration) between subpopulations retards the process of genetic differentiation between populations.

Random genetic drift

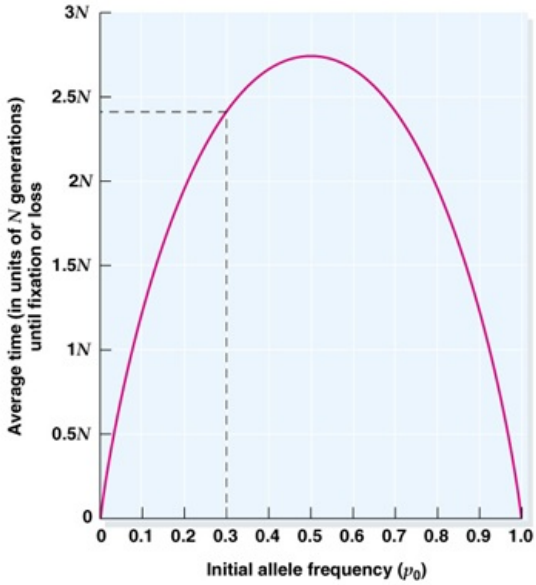
Gene pool / gametic pool / zygotic pool In a population, some individuals may by chance not pass on their alleles to next generation, others by chance pass on more than their “fair share”.

Formation of zygotic pool is a random process allelic frequencies may change/vary in the course of generations (from generation to generation).

Gene(tic) drift is the change of allelic frequencies in the gene pool produced by random causes. It depends on the size of population – changes are unpredictable; more visible/ particularly pronounced in small populations.

Extreme limit – stabilization/fixation of one allele in the small population and elimination of second alternative allele genetic homozygotization. Given enough time, any allele frequency can drift to 1 (fixation) or 0 (extinction). Speed (time T) of fixation is dependent on the size of population and original allelic frequencies.

$T_{fixation} = - 4N \cdot (p \cdot \ln p + q \cdot \ln q)$



Alleles (mutations) may be lost or fixed within a (small) population (N, number of individuals in this population)

Bottleneck

Many populations go through “bottlenecks” where the size of population is reduced (migration, disease, famine, climate). A small sample from a population may have a non-random distribution of alleles.

When the population grows, it will have different allele frequencies from the population before bottleneck.

Founder effect

A few individuals colonising a new region can cause a “founder effect” (f.e.) whereby some genes are more common in the colony than the population they came from. F. e. explains extremely high frequencies of some diseases in subpopulations.

Examples: i) myotonic dystrophy (inherited muscular disease) is much more common in a region of French Canadian immigrants than in Europe, because some of the original settlers were carrying the gene. ii) Tay-Sachs disease (AR inborn error of metabolism with neurologic symptomatology) has high frequency in Ashkenazi Jews (originated from East Europe)

The possibility of equilibrium between drift and migration

The former tends to cause differentiation, the latter to homogenize gene frequencies.

- Wright's „Island model“ of population structure (unrealistic, much abused)
- Wright's „Isolation by distance model“ of population structure

Measurement of population substructure using F_{ST}

Subdivision into populations with distinct allelic frequencies could even create a heterozygote deficit (this can be considered a sort of inbreeding), F:

$F=1-(2p_{av}q_{av}/0.5)$

This sort of inbreeding coefficient is called F_{st} after Fixation index in the Subpopulation relative to the Total population (p_{av}, q_{av} – average allelic frequencies of A and a alleles over all subpopulations in a metapopulation). F_{st}can also be shown to be equal to the standardized variance of gene fequency (also called the Wahlund variance) in the k subpopulations, divided by (standardized by) the maximum possible variance, pavqav. The maximum possible variance is the variance when different populations are fixed for A or a.

where

$$F_{st}=\frac{V_p}{p_{av}q_{av}} \quad \text{where } V_p=\frac{1}{k}\sum_{i=1}^{k}p_i^2-p_{av}$$

F_{st} is a fraction of total variance, the proportion of genetic variation found between as opposed to within each populations. Thus, (1-F_{st}) is the proportion of the total metapopulation genetic variation found within as opposed to between populations. If there is a lot of local fixation or inbreeding, F_{st} will be near 1; if very little, F_{st} will be low (0.05, for instance, might be a typical value).

Evolutionary aspects

Drift is the major cause of genetic differences between subpopulations. Since the drift is quantifiably effective in small populations only, it had to play the principal role in early stages of human evolution when our populations were small.

If populations are subdivided, they can evolve apart, somewhat independently. Population structure allows populations to diversify. This is the reason why population structure is a very important part of evolutionary genetics.

The term “population structure” usually refers to the patterns in neutral genetic variation that result from the past or present departure from panmixia of a population. Understanding past population structure is of interest to evolutionary biologists because it can point to, e.g., potential environmental factors such as climate changes as driving evolutionary forces.

Gene flow (migration) homogenizes gene frequencies and can reduce local adaptation, by preventing divergence.There could be a complete absence of gene flow between subpopulations because of social, geographic, ecological, or even biological barriers. In such cases, evolutionary changes between subpopulations occur under complete isolation.

So, new mutations arising in certain subpopulations remain "private", and genetic differentiation between subpopulations occur with a speed governed by mutation rate and the breeding size of subpopulations.

Population structure has important biomedical consequences either when a number of subpopulational groups is locally adapted to particular environmental conditions (and maladapted when exposed to new environments) or represents a confounding factor in the study of the statistical association between genetic variants and phenotypic traits.