

Genetic Management of Outbred Laboratory Rodent Populations

Daniel L. Hartl
Department of Organismic and Evolutionary Biology
Harvard University
Cambridge, Massachusetts

This document describes the Gold Standard populations of outbred laboratory animals that have been developed at Charles River Laboratories. Using modern techniques of population and molecular genetics, the Gold Standard populations are carefully managed in such a way as to preserve genetic variation and to provide a reliable source of outbred animals for experimental use. The Gold Standard animals are preferred for a number of applications, particularly those in which genetic variation in the response may be an issue. In the Gold Standard outbred populations, individual animals are genetically diverse, much like people are genetically diverse, so no two animals are genetically exactly alike.

Inbred Lines Are Highly Homozygous

When we picture a population of laboratory animals, particularly rodents, we think almost reflexively about inbred lines. We are accustomed to thinking about inbred lines in the context of laboratory rodents because many laboratory populations are inbred. A number of inbred strains currently in use have an ancestry that traces back nearly a century, to the years shortly after the rediscovery of Mendel's laws. Inbred strains, once formed, are maintained over succeeding generations by repeated brother-sister mating (i.e., close inbreeding), so the inbreeding is continuous and very intense.

But continual close inbreeding is unusual in natural populations of mammals. Highly inbred strains are genetically very artificial in the sense that they lack genetic variation from one individual to the next in the inbred population. Any particular inbred strain replicates only one of the multitude of genotypes that would be present in a natural population. Because genetic variation may affect any aspect of physiology, development, or behavior, and because highly inbred strains typically differ in a number of other ways from outbred populations, the utilization of outbred laboratory animals may be preferred in studies in which there is potential genetic variation affecting the response.

Outbred Populations Have High Levels of Genetic Variation

In most natural populations of animals, in particular small mammals, high levels of genetic variation are the norm. Abundant evidence of this genetic variation can be found in the extensive surveys of amino acid polymorphisms in enzyme proteins, as assessed through slight differences in electrophoretic mobility of the proteins caused by amino acid replacements. A typical study of 26 proteins is reported in Selander (1976). In this sample of protein-coding genes, about 20 percent were polymorphic, which means that two or more allelic variants leading to differences in electrophoretic mobility of the protein were found at relatively high frequency. The average animal was found to be heterozygous at about 4 percent of these protein-coding loci.

We now know, from studies of DNA polymorphisms, that electrophoretic variation in proteins underestimates the total amount of genetic variation. In mammals, only about 4 percent of the genome codes directly for the amino acid sequence of proteins. There are large numbers of nucleotide polymorphisms that are in noncoding DNA, including those in introns that interrupt coding sequences within a gene, in regulatory regions immediately upstream and downstream of coding sequences, and in the regions between genes. Because the genetic code is redundant (a particular amino acid can be specified by more than one three-nucleotide DNA sequence), there are also "silent" nucleotide polymorphisms within coding regions that do not alter the amino acid sequence. In most natural populations of mammals, the average animal is thought to be heterozygous for approximately 1 in every 1000 nucleotide sites. Given a genome size in rodents of 3 billion nucleotide pairs, this implies that the average individual is heterozygous at about 3 million nucleotide sites. How many of these polymorphisms may contribute to small differences in the physiology, development, or behavior of the organism is unknown.

Why Outbred Populations are Genetically Superior to Inbred Populations

In most species, including small mammals, outbred organisms are superior to inbred organisms in many ways. The main source of the superiority is heterozygosity. Outbred populations contain a great deal of genetic variation, as we have noted. The genetic variation includes a great many recessive mutations which, when homozygous, may be detrimental to the organism, but do little or no harm when heterozygous. Data from experimental organisms suggest that most of these detrimental mutations are not completely recessive. They do have slightly harmful effects on overall fitness even when heterozygous. On the average, they cause perhaps 2 to 3 percent of the reduction in fitness in heterozygous genotypes as in homozygous genotypes.

The presence of deleterious mutations that are almost, but not completely, recessive, accounts for much of the progressive deterioration in performance when an outbred population is subjected to close inbreeding. As the inbreeding continues, some of the detrimental recessive alleles become homozygous and are incorporated into the genotype of the inbred line. As more and more harmful alleles become homozygous, their effects on overall fitness are cumulative, and the performance of the inbred line declines.

Why do these same harmful alleles not become homozygous in an outbred population? First, in an outbred population, harmful alleles remain rare, and the more harmful they are when heterozygous, the lower their frequency of occurrence. Second, on those rare occasions when a homozygous genotype is formed, it must compete with heterozygous genotypes and genotypes that are homozygous for the nonmutant allele. A genotype that is homozygous for a harmful recessive allele will lose out in competition and be eliminated. In an inbred line, on the other hand, whenever two heterozygous genotypes are crossed, one-fourth of the progeny will be homozygous for the recessive. Since inbred lines are propagated through a relatively small number of breeding organisms, it often happens, by chance, that the breeding organisms are all homozygous for the harmful recessive, even if the genotype is detrimental. Once "fixed" in homozygous condition, the harmful allele remains in the inbred line except in the very unlikely event that a reverse mutation occurs. As an increasing number of homozygous detrimental

alleles accumulates, the performance of the inbred line progressively deteriorates. This process of deterioration in performance is known as "inbreeding depression." The decline in performance continues until all of the detrimental alleles in the original population have been either fixed or lost, and the inbred line becomes homozygous for almost all genes. Some residual heterozygosity may remain either because of rare new mutations or because of genes, such as those in the major histocompatibility complex (MHC), that are subject to strong natural selection for the maintenance of heterozygosity.

The flip side of inbreeding depression is "heterosis," which means the superior performance of the progeny of crosses between inbred lines. Heterosis is the genetic basis of the success of hybrid corn and other hybrid crops as well as the increase in fitness or performance of hybrid rodents. The principle is that two independently derived inbred lines are unlikely to be homozygous for exactly the same set of detrimental alleles, and so, when the inbred lines are crossed, the near-dominance of the favorable allele of any gene inherited from one parent compensates for the detrimental recessive allele inherited from the other parent. Some plant breeders refer to this genetic mutual covering up as "nicking." An analogy that is often used is that of a person with holes in his socks who puts two socks on each foot in the hope that the holes will be in different places. In addition to nicking, there may be some genes for which the heterozygous genotype is superior to either homozygous genotype, but only a small proportion of heterosis, if any, is thought to be due to such genes.

There is another important respect in which outbred populations are superior to inbred populations. The organisms in inbred populations are generally less capable of dealing with environmental effects on physiology, development, and behavior. The result is that inbred populations show more variation among organisms than do outbred populations. This effect is the opposite of what would be expected from considerations of genotype. Since a highly inbred population consists of organisms, all of whom have the same homozygous genotype, one would expect a smaller variation among organisms than in an outbred population, in which the genotype differs from one organism to the next. The finding of a larger variance in inbred populations than in outbred populations has been called "genetic homeostasis." The biological basis of the phenomenon is not known, but it is well documented. Many examples, including numerous examples in rodents and other small mammals, are summarized in Phelan and Austad (1994).

Individual Identity Versus Group Identity

The contrast between inbred and outbred populations challenges our concept of "identity." The contrast may be illustrated by comparing Table 1 with Table 2. Each table shows a three-gene genotype of each of nine organisms in each of four populations. In Table 1, the populations are inbred lines, in Table 2 they are outbred populations. In each of the inbred lines, every organism has the same homozygous genotype as every other. There is no genetic variation from one organism to the next. On the other hand, the genotype found in one inbred line differs from that in another inbred line. Each inbred line is genetically different. By "identity" with regard to inbred lines, we mean identity in genotype among individual organisms within any one inbred line. The identity does not carry over from one population to the next, because each inbred line is different.

With respect to outbred subpopulations, the situation is entirely the opposite. (The term "subpopulation" refers to more or less genetically isolated subgroups of an outbred population.) As shown in Table 2, within any of the outbred subpopulations, there is a variety of genotypes.

Table 1: Identity within, differences between, inbred lines

Organism	Inbred line			
	1	2	3	4
1	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>
2	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>
3	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>
4	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>
5	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>
6	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>
7	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>
8	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>
9	<i>AA BB CC</i>	<i>AA bb cc</i>	<i>aa bb CC</i>	<i>aa BB cc</i>

There is genetic variation within each subpopulation, and no two organisms have the same genotype. In comparing different outbred subpopulations, however, we see that, while each subpopulation is genetically variable, each outbred subpopulation contains the same genotypes as any other, though perhaps differing somewhat in the frequencies in which they occur.

Table 2: Differences within, identity between, outbred populations

Organism	Outbred subpopulation			
	1	2	3	4
1	<i>AA Bb Cc</i>	<i>aa bb cc</i>	<i>aa Bb CC</i>	<i>Aa Bb CC</i>
2	<i>Aa BB Cc</i>	<i>AA bb Cc</i>	<i>AA Bb Cc</i>	<i>AA Bb cc</i>
3	<i>aa Bb CC</i>	<i>aa Bb CC</i>	<i>Aa Bb CC</i>	<i>Aa BB Cc</i>
4	<i>AA Bb cc</i>	<i>Aa BB CC</i>	<i>AA Bb cc</i>	<i>aa Bb CC</i>
5	<i>Aa BB CC</i>	<i>Aa Bb CC</i>	<i>aa bb cc</i>	<i>aa Bb Cc</i>
6	<i>aa Bb Cc</i>	<i>aa Bb Cc</i>	<i>Aa BB Cc</i>	<i>aa bb cc</i>
7	<i>AA bb Cc</i>	<i>AA Bb cc</i>	<i>AA bb Cc</i>	<i>Aa BB CC</i>
8	<i>Aa Bb CC</i>	<i>Aa BB Cc</i>	<i>Aa BB CC</i>	<i>AA Bb Cc</i>
9	<i>aa bb cc</i>	<i>AA Bb Cc</i>	<i>aa Bb Cc</i>	<i>AA bb Cc</i>

With regard to outbred subpopulations, the concept of "identity" does not mean the genetic identity of different organisms. It means that different outbred subpopulations are statistically identical in the sense that they contain the same genotypes in approximately the same frequencies.

To summarize, the concept of identity with regard to inbred lines is genetic uniformity within lines but genetic variation between lines. With regard to outbred subpopulations, the concept of identity means genetic variation within each subpopulation but genetic identity, in a

statistical sense, between subpopulations. The contrast is one of individual identity with group differences versus individual differences with group identity.

Genetic Divergence in Outbred Subpopulations

A number of processes militate against the maintenance of group identity among outbred subpopulations over the long term, particularly if the outbred subpopulations are small. These include mutation, natural selection, unconscious selection, and random genetic drift. Each of these processes warrants a brief discussion.

New mutations are apparently a source of divergence in group identity. The effect is more apparent than real, however. In the first place, new mutations are rare. If the rate of mutation is, for example, 1 nucleotide substitution per 10^9 nucleotide pairs per generation, and the genome size is 10^9 base pairs, there is, on the average, one new mutation per gamete per generation. Since only about 4 percent of the mammalian genome codes for proteins, only 1 in 25 gametes is expected to carry a new mutation in a coding region. A significant fraction of these will be "silent" mutations that do not cause a change in the amino acid sequence. Another significant fraction will be so detrimental that they will be eliminated by natural selection very quickly. Among new mutations that have a chance to be retained in the population, the prospects are very bleak. Even for a new silent mutation, the probability that it will ultimately be lost from the population by chance alone is $1 - (1/2N)$, where N is the population size (actually, the "effective population size" discussed in a later section). If N is reasonable large, a very large fraction of new silent mutations are lost by chance, most in the early generations. If a mutation is detrimental its chances are even worse; and even if a mutation is favorable its chances are not much better. (Favorable mutations are generally regarded as very rare in the first place.)

In subpopulations of outbreeding organisms in a state of nature, natural selection can play an important role in genetic divergence among the subpopulations. Natural selection plays this role most effectively if the subpopulations live in rather different environments, so genes that confer an adaptive advantage may differ from one subpopulation to the next. As time and mutation goes on, natural selection will act in such a way that, as each subpopulation becomes progressively more adapted to its own local environment, its genetic composition will diverge from those of the other subpopulations.

Natural selection for genetic diversity plays a limited role in laboratory populations, especially if the rearing practices are similar in all subpopulations. Natural selection does continue to operate on such traits as mating success, fertility, and survivorship, but it acts on these traits uniformly across all subpopulations.

One type of selection that can result in genetic divergence among outbred laboratory subpopulations is "unconscious selection." This term was first used by Charles Darwin to refer to the situation in which animal breeders select unwittingly, but by unconscious preference, for animals with a certain temperament, body conformation, color markings, and so on. Some degree of unconscious selection is unavoidable even in the most carefully controlled breeding programs. Unconscious selection for large litter size, lack of runt or deformed offspring, a short interval between successive litters, large body size of mother, rapid growth of offspring,

docile temperament, as well as other traits that pose economic advantage could all influence the breeding of rodents for research.

By far the most important source of genetic divergence among outbred subpopulations is random genetic drift, especially in small populations. Random genetic drift results from a sort of genetic lottery in which each organism in the parental generation gambles, not in money but in gametes, for representation among the organisms in the progeny generation. Not every organism in the parental generation contributes exactly the same number of gametes to the progeny generation, and so the allele frequency in the progeny generation is not necessarily that same as it was in the parental generation. Even if each organism in the parental generation did contribute equally to the progeny generation, there would still be random genetic drift because of Mendelian segregation. Any heterozygous genotype produces an equal proportion of gametes carrying each of the two types of alleles. If every parental organism contributes exactly two gametes to the next generation, which is the number required to keep the population size constant from one generation to the next, a heterozygous organism has only a 50 percent chance of contributing one allele of each type to the next generation. Because Mendelian segregation is also a lottery, the allele frequencies in a population may change from one generation to the next. This effect is further magnified if only a small portion of the reproductively fit organisms are permitted to breed which is the usual case in laboratory reared populations where many of the progeny are used for experimentation.

An example of random genetic drift and its causes is illustrated in Table 3, which shows the genotypes of the organisms in the parental generation of a very small population ($N = 8$), the gametes they produce, and the genotypes of the organisms in the progeny generation. Each parental organism produces a potentially infinite pool of gametes, of which only a small number are lucky enough to make it to the next generation. The lucky ones are indicated by the underlined alleles, chosen at random from among the entire pool of gametes and then joined randomly in pairs to form the progeny genotypes. The effect of random genetic drift in this

Table 3: An example of random genetic drift

Parental Generation	Gamets	Offspring generation
AA	A <u>A</u> A A A <u>A</u> A A	AA
AA	A A A A A A A A	Aa
Aa	A A <u>A</u> A a <u>a</u> <u>a</u> a	Aa
Aa	A <u>A</u> A A <u>a</u> a a a	Aa
Aa	<u>A</u> A A A a a a <u>a</u>	aa
Aa	A A A A a a a <u>a</u>	aa
Aa	a <u>a</u> a <u>a</u> <u>a</u> a a a	aa
Aa	a a <u>a</u> <u>a</u> a a a a	aa

example is evident from the change in allele frequency. Among the parents, the allele frequency of A is 8/16 whereas, among the progeny, it is 5/16.

If the example in Table 3 were carried out for additional generations, the allele frequency would continue to drift, sometimes increasing, sometimes decreasing, until eventually, purely by chance, only one allele (either A or a) is represented in the progeny generation. From this time forward the population will remain fixed for the remaining allele either A or a (whichever is the only allele represented), and the other allele would be lost until some very distant future date in which it may arise again by a new mutation.

Random genetic drift causes genetic divergence between subpopulations because it occurs independently from one subpopulation to the next. If the allele frequency of A increases in one subpopulation, it may decrease in another. The situation for two subpopulations is shown in Figure 1. In this example, each subpopulation has the size $N = 10$, and each begins the process with 10 A and 10 a alleles. As time goes on, the allele frequencies in the subpopulations drift apart. In other words, they undergo genetic divergence with respect to the A and a pair of alleles. In this example, after 20 generations, the A allele has a frequency of 17/20 in Subpopulation 1, but has been lost in Subpopulation 2.

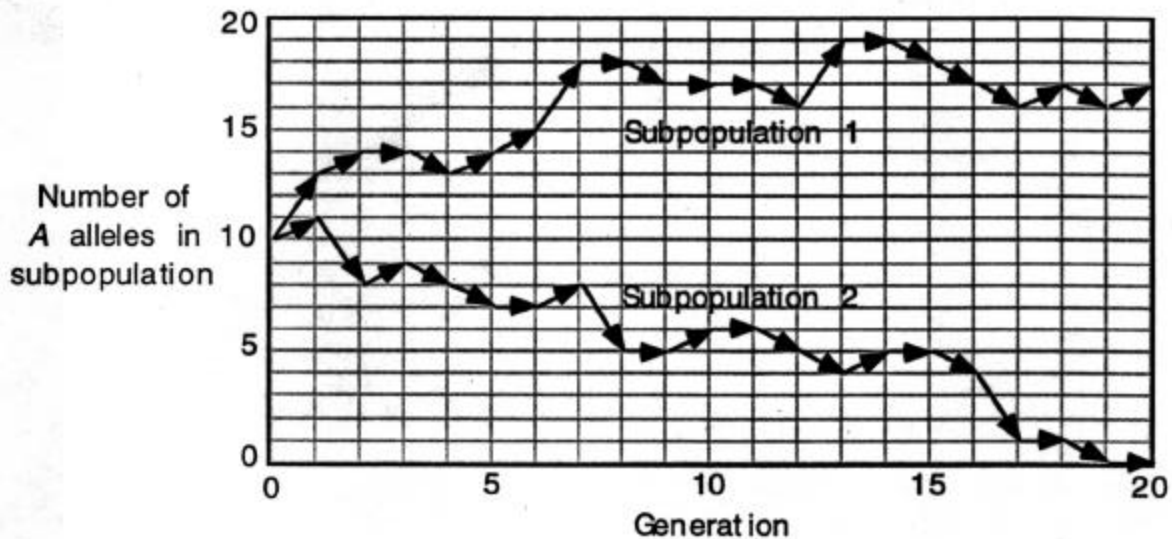


Figure 1. Random genetic drift in two subpopulations

Progress of Random Genetic Drift

In a set of subpopulations undergoing random genetic drift, the degree of genetic divergence among the subpopulations can be measured by tracking the frequency of heterozygous genotypes, averaged across the whole set of subpopulations. Within each subpopulation mating is random, and the frequency of heterozygous genotypes is equal to two times the product of the allele frequencies. This is known as the Hardy-Weinberg Principle. The heterozygosity within a subpopulation changes through time as the allele frequencies change, from a maximum of 50 percent when the allele frequencies are equal to a minimum of 0 when one or the other allele is fixed. If there were no random genetic drift, the proportion of heterozygous genotypes would remain constant at two times the product of the allele frequencies present in the initial population. In a set of subpopulations, as random drift progresses, the

average heterozygosity across the subpopulations changes as the allele frequencies change willy-nilly from one generation to the next. If every subpopulation began with the same allele frequencies, then the average heterozygosity must inevitably decrease over a long period of time from its initial value until each subpopulation has been fixed for one allele or another. Then the heterozygosity will be 0 because there are no heterozygotes in any of the subpopulations.

A convenient measure of the progress of random genetic drift in a set of subpopulations is the fixation index, symbolized F_{ST} , which is the reduction in average heterozygosity across the subpopulations, measured relative to a situation in which the subpopulations were fused into a single random-mating population so large that the magnitude of random genetic drift would be negligible. The subscripts S and T stand for average heterozygosity in the Subpopulations relative to that in the Total (as if not subdivided) population.

F_{ST} is a convenient measure of the progress of random genetic drift in part because it has been very well studied theoretically (reviewed in Hartl and Clark 1998). The fixation index ranges from a minimum value of 0, when all subpopulations have the same allele frequencies, to a maximum value of 1, when each subpopulation is fixed for a different allele. As might be expected, the progress of random genetic drift is strongly dependent on the size of each subpopulation, N . The pattern is illustrated in Figure 2. Subpopulations of size 50 or less are highly vulnerable to random genetic drift, whereas the effect is mitigated in larger populations consisting of several hundred individuals.

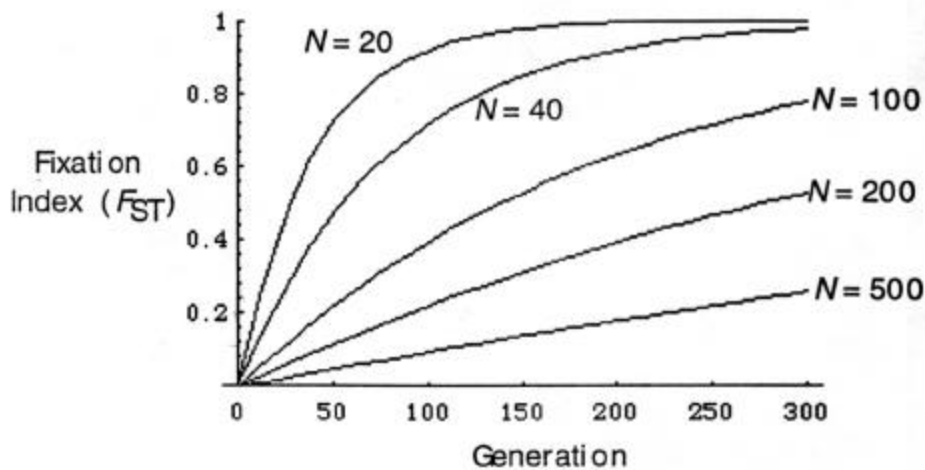


Figure 2. Progress of random genetic drift as a function of population size

Values of F_{ST} approaching 1 are rarely found in nature because the degree of genetic divergence represented by such high values is very extreme. Guidelines to the interpretation of F_{ST} have been given by Wright (1978), who invented the concept:

- F_{ST} in the range 0 to 0.05 indicates little genetic divergence
- F_{ST} in the range 0.05 to 0.15 indicates moderate genetic divergence
- F_{ST} in the range 0.15 to 0.25 indicates great genetic divergence
- F_{ST} above 0.25 indicates very great genetic divergence

These guidelines to interpretation dramatize the effect shown in Figure 2, that increasing population size minimizes the consequence of random genetic drift.

Processes That Oppose Genetic Divergence

Large population size Random genetic drift can be counteracted quite effectively in a number of ways. One relates to the effect of population size illustrated in Figure 2. The larger the population, the smaller the effects of random drift, because a larger population reduces the chance variation in allele frequency from one generation to the next. A second ameliorating factor is the maintenance of a nearly equal ratio of breeding females and males. Because females and males must make equal genetic contributions to each progeny generation, whatever the sex ratio, a highly biased sex ratio increases the effect of random genetic drift. The sex that is in the minority still contributes half the genes to the progeny generation. For this reason the minority sex represents a sort of genetic "bottleneck" through which half the genes must pass, and so the variance in allele frequency is increased. Sometimes for production efficiency, as well as conservation of housing space, polygamous matings must be used. A third factor that minimizes random genetic drift is an approximately equal contribution of progeny from each breeding adult. In effect, a large variance in progeny number means that a few parents contribute a disproportionate number of genes to the progeny generation. This has the same effect as a reduction in the total number of breeding adults, which increases the chance variation in allele frequency that is the cause of random genetic drift. Absolute equality, when each breeding adult contributes the same number of progeny as any other, is not achievable in practice because some animals are sterile or are undesirable as parents for other reasons. The best that can be achieved in a practical sense is a Poisson distribution of progeny number, which means that each breeding adult has an equal and random chance of leaving progeny as any other. Even this ideal is difficult to achieve in practice because of the large natural variance in progeny number from one mating pair to the next, but grossly unequal contributions can be minimized.

When a population geneticist uses the term "effective population size" with reference to some actual population, it means the population size of a theoretically ideal population that would have the same rate of random genetic drift as found in the actual population. A theoretically ideal population has an equal number of breeding males and females, a Poisson distribution of progeny, and is constant in size through time. The effective size of an actual population is almost always smaller than the total number of individuals in the population, sometimes considerably smaller if the sex ratio is highly skewed or if there is a high variance in the distribution of progeny number among the mating pairs. The closer the actual population approximates the assumptions of an equal sex ratio and a Poisson distribution of progeny number, the more nearly the actual population size approximates the effective size. Quantitative aspects of effective population size can be found in Hartl (2000) and Hartl and Clark (1997).

Migration among subpopulations When an outbred population is subdivided into a number of subpopulations, the effect of random genetic drift can also be mitigated by migration among the subpopulations. Migration means that some individuals who are reared in one subpopulation become breeding members of another subpopulation. It is quite clear intuitively that migration will tend to counteract genetic divergence, because the organisms that migrate between subpopulations carry alleles from one subpopulation to another. Not only does fixation become less likely, the whole process of genetic divergence is slowed down.

What is not so obvious intuitively is that, when there are a large number of subpopulations, migration and random drift eventually reach an equilibrium stage at which the additional divergence attributable to random genetic drift in any generation is exactly offset by the homogenizing effects of the migration. Furthermore, the amount of migration needed to counteract random drift and dramatically reduce the equilibrium fixation index is remarkably small. This effect is shown in Figure 3, in which the symbol m stands for the proportion of

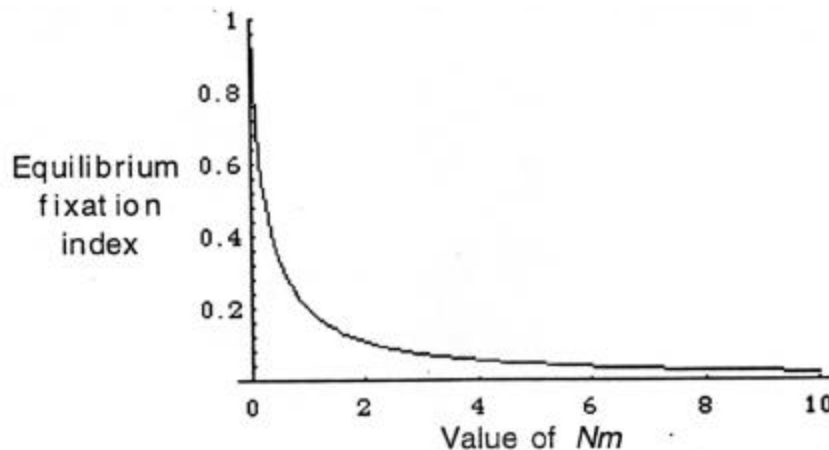


Figure 3. Migration counteracts random genetic drift

individuals in each subpopulation in each generation that are migrants. The effect of migration depends on the product Nm (the effective population size times the migration rate), but this product has a simple biological interpretation. Because N is the population size, and m is the migration rate (proportion of migrant individuals), the product Nm is the actual number of migrant individuals into each subpopulation in each generation. Figure 3 indicates that only a small number of migrants per generation is required to reduce the equilibrium fixation index to values of 0.05 or less.

Stabilizing selection We have seen how unconscious selection is another process that potentially reduces genetic variation in outbred populations by allowing certain unconsciously favored genotypes to contribute disproportionately to the pool of breeding individuals. Some of this type of selection is necessary and desirable. Animals that are sterile cannot reproduce, and those that are morphologically, physiologically, or neurologically defective should not reproduce. On the other hand, some kinds of unconscious selection can and should be controlled. Among these are unconscious selection for rapid and multiple mating in males and for excessive offspring number in females. Such effects can be counteracted by deliberate selection favoring the average animals at the expense of the extremes, which is known as stabilizing selection.

Multiple matings by a few sexually aggressive males in animal populations caged together allows a small number of males to contribute disproportionately to future generations, and the genes for rapid and multiple mating therefore increase in frequency. At the same time, the large variance in mating success among males reduces the effective population size. The unconscious selection for sexual aggressiveness, as well as the reduction in effective population number, can be counteracted by stabilizing selection in which males are housed individually with

females so that each male gets an equal chance to mate. This tends to equalize the sex ratio among breeding pairs, which maximizes the effective population number.

Similarly, there should be stabilizing selection for litter size by culling females with excessively large or small litters. The selection against small litters is necessary to maintain productivity and eliminate, insofar as possible, detrimental alleles that decrease fecundity. The selection against excessively large litters is desirable to offset unconscious selection for offspring number, at the expense of other desirable life-history traits, and also to reduce the variance in offspring number among females. Stabilizing selection for moderate litter size to the extent that it is practical tends to even out the contribution of each female to future generations, which maximizes the effective population number.

Genetic Management of International Genetic Standard (IGS) Outbred Populations

Charles River Laboratories has implemented a genetic management program for outbred animals referred to as the International Genetic Standard (IGS) program. It takes advantage of the principles of population genetics to maintain genetic variation over time. The protocol consists of a master foundation colony housed in isolators and connected through genetic migration to a set of "production" colonies maintained at facilities throughout the world. A schematic of the relationship is shown in Figure 4, and we will discuss several important issues relating to the

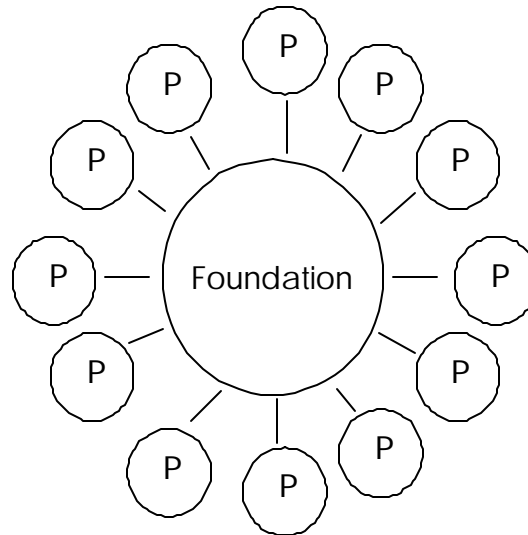


Figure 4. Nucleus and production colonies

nucleus and production colonies in turn. In Figure 4, the number of production colonies is set at 12 only for convenience in representation and concreteness in discussion.

Management of the Nucleus Colony

Much consideration has been given to the management of the foundation colony because it is the principal outbred colony whose genetic variation must be maintained through time. One important feature of the foundation colony is its size. It is maintained at a stable population of

250 mating pairs, or 500 individual organisms. It can be seen from Figure 2 that this size is sufficient to counteract most of the random changes in allele frequency due solely to the effects of random genetic drift. Because the variance in allele frequency due to random genetic drift decreases according to the reciprocal of N , there are diminishing returns as the population size is increased. Maintaining a population size much above 500 generates little additional benefit in minimizing the effects of random genetic drift.

The N in Figure 2 is the effective population size. Two procedures have been implemented in order to have the effective population size approximate the actual size as close as feasible. First, the sex ratio in the foundation colony is maintained at an equal number of females and males. This avoids the "genetic bottleneck" effect of a highly skewed sex ratio. Second, in order to reduce the variance in offspring number among the breeding animals due to differential mating success, breeding is monogamous with each breeding pair being caged separately.

In any population of finite size, even one as large as 500 individuals, there is a chance that the ancestry of any mating pair can be traced back to one or more common ancestors. This is a form of inbreeding due to mating between remote relatives that is less familiar than the usual type of inbreeding between close relatives, such as parent-offspring, brother-sister, or first-cousin mating. Nevertheless, in the foundation colony, the effects of mating between remote relatives is minimized by a procedure of maximum avoidance of inbreeding based on a circular pair mating system. This system relies on a fixed rotation of breeders to lessen the chance of inbreeding. This system also minimizes the number of transfers of animals between the 20 semirigid gnotobiotic isolators in which the colony is maintained, thereby minimizing the potential of an undetected microbiological contamination being moved between isolators.

The foundation colony is also subjected to stabilizing selection to counteract two undesirable forms of unconscious selection. In order to ensure continuing high levels of male mating success and male fertility, the caged mating pairs are subjected to a time limit for mating. If, during this period, no mating takes place, then pair is retired and replaced with a newly constituted pair. On the female side of the equation, litter size is subjected to stabilizing selection by culling females with extraordinarily large (greater than 16) or unusually small (less than 6) litters. This procedure not only maintains female fertility, it ensures that other desirable life-history traits are not sacrificed for the sake of ever increasing litter sizes.

Out-migration

Now we can specify in more detail how the nucleus colony and the production colonies in Figure 4 are genetically connected. The two issues are out-migration and in-migration. These are illustrated by the direction of the arrows in Figure 5. Consider first out-migration: How are the genes present in the foundation colony distributed to the production colonies? There is actually an optimum level of out-migration. Too little out-migration runs the risk that the production colonies may become genetically divergent from the foundation colony and also

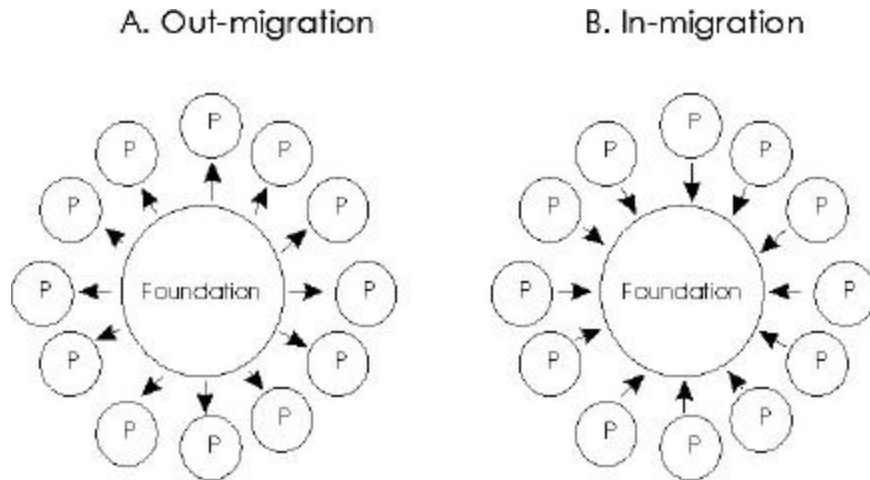


Figure 5. Out-migration (A) and In-migration (B)

genetically divergent from each other. The ideal of group identity among outbred subpopulations therefore becomes compromised. On the other hand, too much out-migration is disruptive to the management routine of the production colonies without yielding a commensurate increase in group identity.

The manner in which these offsetting considerations are handled in the IGS outbred populations is shown in Figure 6. The production colonies are served in rotation on a three-year cycle, one-third of the production colonies in any year. This pattern is indicated by the shading of the production colonies in Figure 6. The first year serves the production colonies shaded black, the second year those that are crosshatched, and the third year those that are stippled. In any one

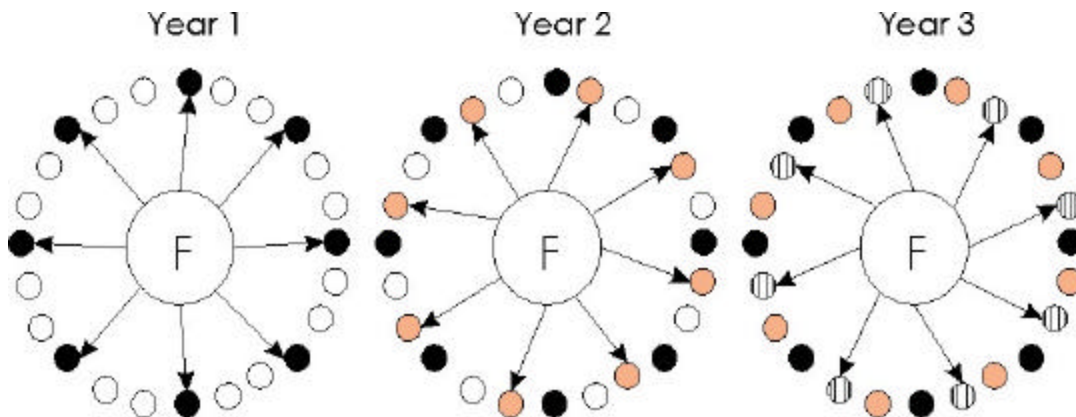


Figure 6. Out-migration protocol

year, when a production colony is being served, $\frac{1}{4}$ of the males are replaced. The overall rate at which the germ plasm in the production colonies is replaced is, therefore, 24 years. The rationale behind this number is that, with one-third of the males replaced every three years, all of the males are replaced every 12 years. Since males contribute half of the genes to any generation of offspring, it requires 12 years to replace half the germ plasm or, on average, 24 years to replace all of it.

The choice of males for the out-migration is one of convenience. The sex of the migrants makes little practical difference because the migration of males is almost equivalent in its genetic effect as the migration of females. Two exceptions are found in the mitochondrial DNA, which is transmitted exclusively through the female, and in the X chromosome, which is present in two copies in females but one in males. Because of the XX-XY difference between females and males, the rate of replacement of the X chromosome is a little less than that of other chromosomes, 36 years versus 24 years. Although the out-migration does not include mitochondrial DNA, this minor component of the genome is nevertheless subject to migration because, as we shall see, the in-migration from the production colonies to the nucleus colony does include females.

In-migration

In regard to in-migration, the compromise is between creating a genetic bridge between the production colonies and the foundation colony without the rate of in-migration being so excessive as to swamp the genotypes present in the foundation colony. It is the foundation colony that is maintained at a large size with numerous quality-control provisions governing sex ratio, distribution of offspring number, avoidance of inbreeding, and stabilizing selection. This is the colony whose genotypic composition is to be kept as constant as feasible through time. It is the source of renewed germ plasm in the production colonies. Yet some in-migration from the production colonies is desirable to help maintain group identity between the foundation colony and the production colonies as well as to increase the effective population size of the foundation colony by allowing the return of some germ plasm from the production colonies.

The pattern of in-migration that has been established in the IGS outbred colonies is illustrated in Figure 7. Each production colony serves as a source of migrants in turn.

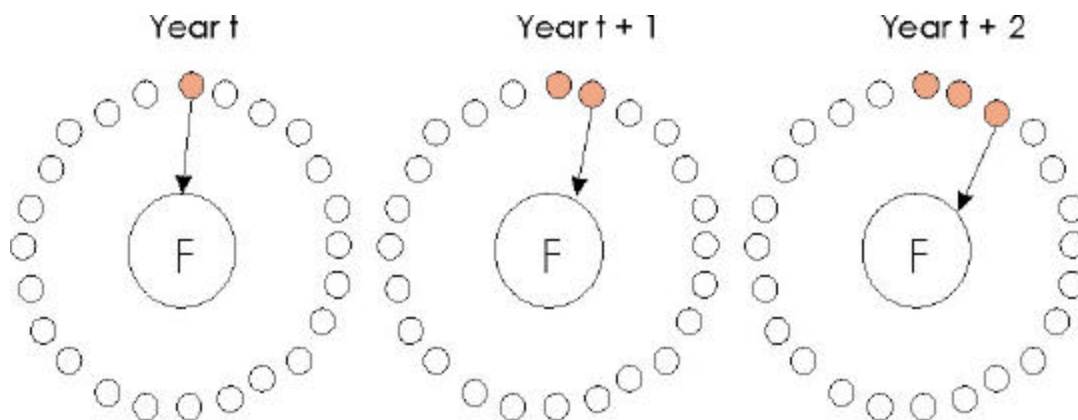


Figure 7. In-migration protocol

Figure 7 shows the rotation starting with the production colony at the top. The next year the migrants come from the production colony next in line, and so forth. With the protocol, one complete rotation around the circle is accomplished in as many years as there are production colonies.

The in-migrants in any year consist of 3 mating pairs. This rate equals 6 individuals per year or, assuming approximately two generations per year, 3 individuals per generation. Judged in terms of Figure 3, a rate of 3 migrants per generation has a substantial effect on minimizing genetic divergence as measured by the fixation index F_{ST} .

The rate of in-migration can also be examined in another way. The replacement of 3 mating pairs per year represents 1.2 percent of the population. One complete replacement of the germ plasm in the foundation colony therefore requires, on the average, 83 years. This rate is fast enough to maintain a genetic cohesion between the production colonies and the foundation colony. At the same time, it is slow enough that the heterogeneous genetic composition of the foundation colony can be maintained through time by the implementation of the stringent IGS genetic management protocol.

Genetic Monitoring: Validation and Verification

Validation and verification refer to periodic genetic screening of representative samples of animals taken from IGS outbred populations. These tests ensure that the levels of genetic variation are maintained through time and allow comparison of heterozygosity between colonies.

Traditional approaches to genetic screening make use of genetic variation in antigenic determinants, such as those associated with blood groups or histocompatibility; or they make use of genetically determined variation in the electrophoretic mobility of protein molecules, such as in the enzymes of intermediary metabolism. Both of these types of genetic markers are very useful, but they have the limitation that the number of polymorphic genetic markers is relatively small. An additional practical limitation is that each of the polymorphic genes requires a somewhat different serological or enzymatic assay, so the implementation of large-scale screening can be logistically quite cumbersome.

Modern molecular genetics has provided a dense genetic map of highly polymorphic markers that can be used for purposes of validation and verification. These genetic markers correspond to a number of different types of genetic variation, typically the presence or absence of a cleavage site for a restriction endonuclease in a particular fragment of DNA, or the number of repetitions of a short nucleotide sequence that is repeated in tandem within a particular larger fragment of DNA. Genetic markers based on restriction-site differences are called RFLPs (restriction fragment length polymorphisms), whereas those based on the number of short tandem repeats are called VNTRs (variable number of tandem repeats) or SSRs (simple sequence repeat polymorphisms). The main difference between VNTRs and SSRs is that the repeating unit in SSRs is shorter, often consisting of a repeating dinucleotide, such as 5'-TGTGTGTGTG ...-3'

SSRs are widely distributed throughout the mammalian genome. They are also usually highly polymorphic in having a large number of alleles at moderate allele frequencies, so that a high proportion of an outbred population, in some cases 70% or more, is heterozygous for any particular SSR marker. These features have made SSRs the genetic marker of choice for high-resolution genetic mapping in the mouse (Copeland et al. 1993), human beings (Comprehensive Human Genetic Linkage Center 1994), and many other organisms of genetic interest.

The features of SSRs that are favorable for genetic mapping-genome-wide abundance and high heterozygosity-also make them ideal markers for genetic screening. They also bring a practical advantage to the screening protocol. Each SSR marker is assayed through DNA amplification by means of the polymerase chain reaction (PCR), and the products of the amplification are identified by size, such as through separation by gel electrophoresis. Each SSR assay differs only according to the oligonucleotides that are used as specific DNA-replication primers in the PCR amplification reaction and in the expected size of the resulting amplified fragments. The procedures and assays are otherwise identical. Not only are they identical, they are relatively simple and so have been largely automated to allow high-throughput analysis.

Summary: The IGS Outbred Animal

Outbred populations differ from inbred lines in that they are genetically heterogeneous, like natural populations. The genetic variation promotes heterozygosity for recessive genes that, in an inbred line, may become homozygous. The genetic variation also confers other advantages associated with outbreeding, such as a smaller variance than inbred lines with respect to most quantitative characteristics. The IGS outbred populations are managed in accordance with the principles of population genetics in order to maintain their genetic variation through time as well as to promote group identity between the foundation colony and the production colonies. The maintenance of genetic variation is promoted through measures that maximize the effective population size of the foundation colony. Deliberate stabilizing selection is practiced on mating and productivity traits to counteract unconscious selection for reproductive success compromising other performance characteristics. A regular program of migration from the foundation colony to the production colonies (out-migration) and from the production colonies to the foundation colony (in-migration) is followed in order to maintain genetic variation and group identity between the foundation colony and the production colony, as well as to maintain an even larger effective population size in the foundation colony. Periodic genetic monitoring of a representative samples of the foundation and production colony animals enables verification of the level of genetic variation maintained in the IGS outbred populations and an assessment of the similarity of colonies.

References

Comprehensive Human Genetic Linkage Center: J. C. Murray, K. H. Buetow, J. L. Weber, S. Ludwigsen, T. Scherpbier-Heddema, F. Manion, J. Quillen, V. C. Schffield, S. Sunden, G. M. Duyk; Généthon: J. Weissenbach, G. Gyapay, C. Dib, J. Morrissette, G. M. Lathrop, A. Vignal; University of Utah: R. White, N. Matsunami, S. Gerken, R. Melis, H. Albertson, R. Plaetke, S. Odelberg; Yale University: D. Ward; Centre d'Etude du Polymorphisme Humain (CEPH): J. Dausset, D. Cohen, H. Cann. 1994. A comprehensive human linkage map with centimorgan density. *Science* 265: 2049-2054.

Copeland, N. G., N. A. Jenkins, D. J. Gilbert, J. T. Eppig, L. J. Maltais, J. C. Miller, W. F. Dietrich, A. Weaver, S. E. Lincoln, R. G. Steen, L. D. Stein, J. H. Nadeau and E. S. Lander. 1993. A genetic linkage map of the mouse: Current applications and future prospects. *Science* 262: 57-66.

Hartl, D. L. 2000. *A Primer of Population Genetics*. Third edition. Sinauer Associates, Sunderland, MA.

Hartl, D. L. and A. C. Clark. 1997. *Principles of Population Genetics*. Third edition. Sinauer Associates, Sunderland, MA.

Phelan, J. P. and S. N. Austad. 1994. Selecting animal models of human aging: Inbred strains often exhibit less biological uniformity than F1 hybrids. *Journal of Gerontology* 49: B1-11.

Selander, R. K. 1976. Genetic variation in natural populations. pp. 21-45 in F. J. Ayala (editor). *Molecular Evolution*. Sinauer Associates, Sunderland, MA.

Wright, S. 1978. *Evolution and the Genetics of Populations*. Volume 4. *Variability Within and Among Natural Populations*. University of Chicago Press, Chicago.