6. KINSHIP, RELATIONSHIP AND INBREEDING

ANNA K. SONESSON, JOHN A. WOOLLIAMS AND THEO H. E. MEUWISSEN

6.1 Introduction

Inbreeding is a result of mating of related individuals. These related individuals often form a more or less closed population such as a managed breeding population or a wild population that has become isolated from others with little or no migration. In this chapter, we will deal with basic concepts like coefficients of kinship, relationship and inbreeding, which will clarify why relationships are related to the identity of genes, before dealing with the genetic processes that breeding and isolated populations go through.

6.2 Concepts

Malecot (1848) quantified the relationship between two animals U and W by the ‘coefficient de parenté’, which is translated in English as the coefficient of coancestry or the coefficient of kinship. The coefficient of kinship is defined as the probability that a randomly sampled allele out of the two homologous alleles of animal U is Identical By Descent (IBD) to a randomly sampled allele of animal W. Two alleles are IBD if they are two copies of the same allele of a common ancestor. If the paternal and maternal alleles of animal U are denoted by \( u_p \) and \( u_m \) (i.e. those inherited from its sire and dam, respectively) and those of animal W by \( w_p \) and \( w_m \), the coefficient of kinship \( f_{UW} \) is:

\[
f_{UW} = \frac{1}{4} \left[ \text{Prob}(u_p = w_p) + \text{Prob}(u_p = w_m) + \text{Prob}(u_m = w_p) + \text{Prob}(u_m = w_m) \right]
\]

where \( \text{Prob}(u_p = w_p) \) is the probability that allele \( u_p \) equals \( w_p \).

We can define the kinship of an animal with itself, i.e. \( f_{UU} \), and see that the first and last probabilities in equation (6.1) will be 1.0. In equation form, we can then see that the kinship of animal U with itself is

\[
f_{UU} = \frac{1}{4} \left[ \text{Prob}(u_p = u_p) + \text{Prob}(u_p = u_m) + \text{Prob}(u_m = u_p) + \text{Prob}(u_m = u_m) \right]
= \frac{1}{2} \left[ 1 + \text{Prob}(u_p = u_m) \right]
\]

We can write \( f_{UU} = \frac{1}{2} [1 + F_U] \), where \( F_U \) is the inbreeding coefficient for U, which is defined as the probability that the alleles at a randomly chosen neutral locus are identical by descent (IBD). This is the same probability as the probability that a randomly sampled allele from U’s sire is IBD to a randomly sampled allele from U’s dam. Therefore, it is clear that the inbreeding coefficient of U is the kinship coefficient between its parents i.e. \( F_U = f_{SD} \), where S and D are the sire and dam of U.

Probability, which always varies between 0 and 1, is used in these definitions, because

© 2005 Springer. Printed in the Netherlands.*
in the absence of complete molecular genetic information we cannot see whether \( u_0 = w_0 \) and have to work with the probabilities. More generally, it is important to recognize that the random sampling of alleles described in the definition of \( f_{IBD} \) above is done with replacement, which means that even if a parental allele has been sampled once, it is not removed from the pool of parental alleles, but can be sampled again. Random mating, which is one of the assumptions made in deriving many of the standard results in quantitative genetics, also assumes that parents are sampled with replacement. Strictly, random mating should allow selfing, but here we make the assumption that there is no selfing, i.e. the sire and the dam must be different individuals, and that there are two sexes.

Until now, we have looked at IBD at the locus level. However, in real populations, we are generally interested in the average IBD probability over the entire genome. Therefore, all the results will be expectations over all possible loci and populations. From now on, we will focus on this ‘quantitative genetic’ approach.

An individual can only be inbred \( (F > 0) \) if it traces back through both its parents to a common ancestor so, when there is no selfing, it must have a loop in the pedigree. Inbreeding is inevitable in any closed population over a period of time. To avoid having \( F > 0 \), you need two different parents, four different grand parents, eight different great grand parents, i.e. with the number of distinct ancestors doubled each generation, which is clearly impossible for any finite population. Therefore any finite population will have inbred individuals.

The initial population that is used is often called the base population. Similar to the base population, the initial generation is often called the base generation, whose individuals are assumed unrelated and non-inbred and from which individuals descend. All the alleles carried by individuals in the base generation are considered to be distinct. If two alleles of an individual at a locus have descended from the same allele in the base generation, i.e. they are copies of alleles in the base generation, then these two alleles are IBD as defined earlier. It may be possible to group the alleles in the base generation into forms that are chemically identical in some defined way, e.g. they may share the same sequence of base pairs, or code for the same sequence of amino acids, or some other chemical property. If two alleles in generation \( t \) are distinct copies of two base generation alleles, which are identical in form, the alleles are called alike in state (AIS). Being AIS is therefore different from being IBD, since IBD alleles must be AIS (ignoring mutations after the base), but AIS alleles need not be IBD.

For breeding purposes, the base population can be made up of the populations of wild fish that are used to start the breeding program (see Chapter 16 for a discussion on different ways to set up a base population in practice) and the base generation can be the first generation of recorded pedigree. However, in each of the two cases, it is probably not true that all individuals are totally unrelated, i.e. \( F = 0 \), yet any other coefficient of kinship between individuals is probably also not true. Therefore, the choice of the base generation is quite arbitrary but it is a necessary reference point. It implies that absolute measures of relationship, and thus of inbreeding, within a population only have meaning when the base population is defined, and comparisons between populations are arbitrary.