

Significant Relationships

by John Armstrong

How would you determine the impact of a famous Champion on his breed?

A dog who has won many shows and earned many titles may have been quite popular as a stud and may have sired more winning progeny than other contemporary males. However, that does not guarantee that he will have more impact five or ten generations down the line than another dog who was bred only two or three times.

Percent Contribution

If sufficient data is available, one way of determining the significance of an ancestor is to calculate his percent contribution to the current dogs. The % contribution (aka **percentage of blood**) is determined by the way genes are passed from the parents to the progeny. An individual inherits one set of chromosomes, and the genes they carry, from his or her sire and a second, homologous (equivalent) set from the dam. Thus, each parent makes a 50% contribution. As the parents in any generation always contribute 50% of their genes to their progeny, it seems reasonable to expect that 25% will come from each grandparent, 12.5% from each great-grandparent, and so on. However, once we are past the parents, we are dealing in probabilities, not certainties. This is not like mixing paint! When dad passes you one set of his chromosomes, they will include a selection of ones inherited from both his parents, but there is no guarantee that the selection will be exactly equal. There is even a small chance (very small) that he will pass on those from only one of his parents.

By the time we get back 10 generations, the contribution from each of the 1024 ancestors would, in theory, amount to slightly less than 0.1%. However, in the pedigree of the average purebred dog, there are seldom more than 100-200 different names and some appear 50 times or more. These are the significant ancestors that make the major genetic contributions.

If you have a pedigree, you can calculate % contribution of any repeats simply by multiplying the number of times each ancestor appears in any generation by the appropriate percentage for that generation and then add together all of the calculated percentage of contributions from each generation. The table listed below shows the percentage of blood inherited from each ancestor at the given generation levels. Generation "1" is the parents.

Genetic Contribution of Ancestors										
Generation	1	2	3	4	5	6	7	8	9	10
% Contribution	50.0	25.0	12.5	6.25	3.125	1.563	0.781	0.391	0.195	0.098

You should get a number between 0 and 1; multiply by 100% to get the % contribution.

Databases exist for many breeds that will contain the data enabling you to extend a pedigree to 10 generations or more. Manual computation, though tedious, is still possible, but hardly convenient. Several pedigree programs (e.g. [CompuPed](#)) will quickly calculate % contribution for selected ancestors or all ancestors for a specified number of generations, providing you with information on which dogs have been most influential.



Inbreeding Coefficients

While most breeders recognize that a mating between half-sibs or cousins represents inbreeding, the majority probably have no idea which is the closer relationship. This is not helped by the non-standard definition of inbreeding in some books (e.g. Onstott's "Breeding Better Dogs").

The standard definition of inbreeding is that it is any scheme which results in the sire and the dam having common ancestors. Many breeders use the term "inbreeding" for close relatives and "linebreeding" for more distantly related individuals, but there is no fundamental difference.

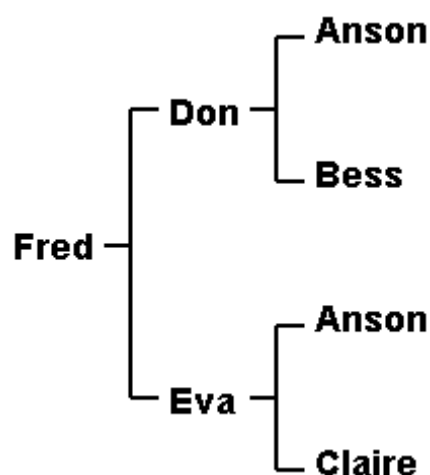
The parameter used to express this common heritage is called the **inbreeding coefficient** and was first proposed by Sewell Wright in 1922. Designated **F** by Wright (but more commonly **IC** or **COI** by breeders), it can theoretically range from 0 to 100%, and indicates the probability that the two alleles for any gene are **identical by descent**.

The primary consequence of inbreeding is to increase homozygosity. However, the IC is not a direct measure of homozygosity because the two alleles passed down from different ancestors may be functionally the same. Furthermore, some proportion of all the genes will be the homozygous because there is only one allele. The IC serves as an indicator of what proportion of the remainder may have been made homozygous by inbreeding.

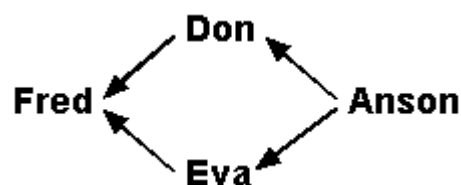
The inbreeding coefficient is a function of the number and location of the common ancestors in a pedigree. It is **not** a function, except indirectly, of the inbreeding of

the parents. Thus, one can mate two highly inbred individuals who share little common ancestry and produce a litter with a very low IC. (Because the potential number of ancestors doubles every generation, eventually you reach a point where the number of ancestors exceeds the number of individuals alive at that time. You are, therefore, bound to find some common ancestors if you go back far enough.) Conversely, it is possible to mate two closely related dogs, both of which have low ICs, and boost the IC substantially.

The most widely used approach for calculating inbreeding coefficients is Wright's "paths" method ([see note](#)), best illustrated by a simple example. Suppose we mate half-sibs, the common ancestor, Anson, being the father. Don is the son of Anson and Bess; Eva the daughter of Anson and Claire. Fred is one of their progeny.



To simplify, we don't show the ancestors that aren't shared:



Now we consider a gene for which Anson carries two different alleles, $a1$ and $a2$. There is a 50% probability of the allele Anson passed to Don being passed on to Fred. There is also a 50% probability that the same allele will be passed from Anson to Eva, and a 50% probability of it being passed from Eva to Fred, if Eva got it. When dealing with events that are contingent (this *and* that must happen), we multiply the probabilities - in this case $0.5 \times 0.5 \times 0.5 = 0.125$ (12.5%). This final number is the probability that Fred will be homozygous for either $a1$ or $a2$ because of the common grandfather.

In general, Wright's method is to determine the path from Fred to the common ancestor, Anson, and back again on the other side of the pedigree (Fred-Don-Anson-Eva-Fred), count the number of individuals in the path, excluding Fred (there are 3,

Don-Anson-Eva) and then calculate $\frac{1}{2}^n$, where n is that number. So, in the present case, we have $(\frac{1}{2})^3$ or $(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}) = 1/8$, or 12.5%. If this were the only common ancestor, the inbreeding coefficient for Fred would be 12.5%.

Now, suppose the common ancestor was one of the grandfathers of the parents (i.e. a great-grandfather of the litter). This adds an individual on each side of the pedigree, so that we will get a path of the type Fred-X-Don-Anson-Eva-Y-Fred, and the inbreeding on Anson will be $(\frac{1}{2})^5$ or 1/32 (3.125%).

Like many other genetic calculations, the IC is based on probabilities, not certainties. An individual may be more or less highly inbred than the number computed.

If we had only a single common ancestor to deal with, life would be relatively simple. However, there are two complications to deal with. The first is that there will be more than one common ancestor. Let's consider the case of first cousins. In human populations such a pairing is prohibited in some societies but allowed in others. We have already calculated the inbreeding for a single shared grandparent. First cousins have two shared grandparents, and we simply add the inbreeding coefficient for each to get 6.25%.

The second complication is that the common ancestor may be inbred. If so, his or her inbreeding coefficient will have to be calculated. To account for this we have to multiply the inbreeding coefficient calculated for Fred by $(1 + F_A)$, where F_A is the inbreeding coefficient calculated for Anson. For example, if Anson is the product of a mating of first cousins, the total inbreeding for Fred will be $0.125 \times 1.0625 = 0.133$ (13.3%) if there are no other shared ancestors in the pedigree.

Unfortunately, in the average pedigree, there are a large number of shared ancestors. Therefore, the total inbreeding for a dog cannot generally be calculated manually and appropriate software must be used (e.g. [CompuPed](#)). Calculating inbreeding for only the first few generations is not particularly useful. If there are more than one or two common ancestors in four or five generation pedigree, the inbreeding is probably already higher than desirable. Unfortunately, having none is no guarantee that common ancestors will not occur in abundance further back, and some pedigrees of this type still achieve moderately high inbreeding coefficients. Neither can the number of shared ancestors be used as a reliable guide, as the inbreeding coefficient is very sensitive to when and where they occur in a pedigree.

Is there a quick way of determining how genetically similar two dogs are?

Suppose a breeder has two bitches (A and B) she wants to mate to different males. After careful research she has identified three potentially suitable males (C, D and E), all of which look equally good. She hopes to get a male puppy from one litter and a female from the other, and would like to eventually breed them to each other. The

objective could be to pick the combination that will minimize the potential inbreeding. Alternatively, she may be looking for two dogs that are not close relatives yet have similar heritage.

One approach would be to produce hypothetical litters for all combinations: AC, AD, AE, BC, BD and BE. Then we would have to look at the possibilities for the second generation. There will be six if we don't permit shared grandparents, and 36 if there are no restrictions. These potential litters could then be evaluated for inbreeding or the % contribution of significant ancestors. This will certainly provide the data, but is unnecessarily tedious.

The Coefficient of Relationship

The coefficient of relationship (RC) provides a way of objectively assessing the similarity of two pedigrees by giving a number that is a direct measure of shared ancestry. In most human populations, two individuals picked at random would likely have a RC of 0, a brother and sister 50% and identical twins 100%. Other relationships would fall between 0 and 50%.

The number generated may be viewed as analogous to the % composition, except that you are comparing two dogs instead of looking at one. A brother and sister will give a value of 50% as long as an ancestor is not repeated. Once ancestors start to repeat, the individuals no longer have an inbreeding coefficient of zero. Two sibs from a highly inbred line may have an RC of 80% or more, and two dogs that are not sibs may have an RC above 50%.

The formula for the RC is:

$$R_{AB} = 2f_{AB} \div [(1 + F_A)(1 + F_B)]^{1/2}$$

where f_{AB} is the inbreeding coefficient of a hypothetical litter between A and B, and F_A and F_B are the inbreeding coefficients for the two individuals, A and B.

A simpler approach to the breeder's problem would be to compute the RCs for C vs D and E, and D vs E. This is not a pencil and paper calculation. However, presented with just such a problem, it took me about 2 minutes to obtain the three RCs with the latest version of [CompuPed](#). My results were R_{CD} 10.4%, R_{CE} 13.4%, R_{DE} 17.2%.

As D and E share the most common ancestry, so would the progeny from their two prospective litters, while C and D share the least. To minimize inbreeding and maximize diversity, they would be my choice, all else being equal. (These values actually all fall below the average for the breed, which is ~ 23%.)

The Kinship Coefficient

The f_{AB} term in the RC equation is sometimes called the "kinship coefficient" and may

also be used as a measure of the relationship between two individuals. It's computation is the same as that of an inbreeding coefficient for a hypothetical litter between the two dogs. (It doesn't matter if they are the same sex.)

The [mean kinship](#) (mk_i) for individual i is the average of the kinship coefficients (f_{ij}) between i and all the other breedable individuals in the population:

$$mk_i = \frac{\sum_{j=1}^N f_{ij}}{N}$$

A conservation biologist would consider the individual with the lowest mean kinship to be the most genetically valuable in terms of maintaining diversity in the population, and would try to favor that individual in a breeding program.

Note: An alternative approach, often referred to as the "tabular" method, calculates inbreeding from the earliest ancestor forward to the current dog (or dogs).

© John B. Armstrong, 1998, 1999