

**Mating prescription, proscription, and the time since a common ancestor, a coalescent approach**

**Abstract:** The expected time since a common ancestor of two genes in a diploid individual which is  $4N+2$  (where  $N$  is the number of mating pairs, hence  $2N$  is the number of individuals, hence  $4N$  is the number of genes) is initially increased by approximately 4 for each generation that inbreeding is proscribed. This approximation breaks down as the the number of generations inbreeding is proscribed approaches the number for maximum avoidance of inbreeding. The expected time since a common ancestor under maximum avoidance of inbreeding (which entails two progeny per mating pair) is  $8N-2$ , which is also the expected time for random mating with two progeny per mating pair. Hence progeny distribution is a significant component of mating structure for determining the expected time since a common ancestor of two genes in an individual.

**Introduction:** Mating with relatives has been favored or considered taboo in various cultures. This study considers the effect of controlling the relatedness of mates on the time since a common ancestor of two genes in an individual. A diploid population with  $N$  faithful mating pairs (hence  $2N$  individuals, hence  $4N$  genes) is assumed, so that pairs of genes can be characterized as being in the same individual, being in the same mating pair, but different individuals, or being in different mating pairs. The coalescent model with individuals choosing their parents is employed. Although randomly choosing parents produces the Poisson progeny distribution, the inbreeding constraints will cause the progeny distribution to deviate from Poisson.

**Zero to mandatory full sib mating with randomly choosing parents:** Let  $s$  be the fraction of matings which are between full sibs, with other parents randomly chosen subject to the constraint that members of a mating pair do not choose the same parents. Let  $T$  be the expected time since a common ancestor of two genes in an individual,  $U$  the expected time since a common ancestor of two genes in different individuals within a mating pair, and  $V$  the expected time since a common ancestor of two genes in different mating pairs. Then

$$\begin{aligned} T &= U + 1 \\ U &= s(.25 + .25(1 + T) + .5(1 + U)) + (1 - s)(1 + V) \\ V &= (1/N)(.25 + .25(1 + T) + .5(1 + U)) + (1 - 1/N)(1 + V). \end{aligned}$$

These equations have the solution  $T = 6 + (1 - s)4N$  which specializes to 6 for mandatory sib mating ( $s = 1$ ),  $4N + 2$  for random mating ( $s = 1/N$ ), and  $4N + 6$  for sib mating precluded ( $s = 0$ ).

**Zero to mandatory full sib mating with two progeny per mating pair:** With the constraint of two progeny per mating pair, the above equations become

$$\begin{aligned} T &= U + 1 \\ U &= s(.25 + .25(1 + T) + .5(1 + U)) + (1 - s)(1 + V) \\ V &= (((1 - s)N((1 - s)N - 1))/(N(N - 1)(2(1 - s)N - 2)))(.25 + .25(1 + T) + .5(1 + U)) + \\ &\quad (1 - (((1 - s)N((1 - s)N - 1))/(N(N - 1)(2(1 - s)N - 2))))(1 + V). \end{aligned}$$

$[((1 - s)N((1 - s)N - 1))/(N(N - 1))]$  is the fraction of pairs of mating pairs where neither is a result of sib mating and  $1/(2(1 - s)N - 2)$  is the probability two individuals (in distinct mating pairs) from such pairs are siblings.] For the above cases the solutions are  $T = 6$  for mandatory sib mating ( $s = 1$ ),  $T = 8N - 2$  for random mating ( $s = 1/(2N - 1)$ ), and  $T = 8N - 2$  for sib mating precluded ( $s = 0$ ).

**Proscribing  $n$  generations of inbreeding with randomly choosing parents:** If inbreeding is precluded for  $n$  generations, that entails an initial addition of  $n$  generations before a common ancestor can occur by random mating for two genes in an individual. But thrice as often as a common ancestor occurs, the ancestors

of the two genes in an individual will be in a mating pair, but not the same gene (two genes in an individual or one gene in each member of the mating pair), which will ‘reset the clock’, hence provide an  $n$  generation delay for each occurrence, hence the total delay in the expected time since a common ancestor will be  $4n$  generations. However, this is a heuristic argument which assumes ancestral matings are precluded solely based on the two initial genes, if common ancestors are precluded for more than one generation, other gene pairs may prevent an ancestral mating pair where the two original genes could find a common ancestor, thereby altering the time since a common ancestor. Thus the increment of 4 generations is the increment from random mating to full sib mating precluded as noted above, but the increment per generation inbreeding is precluded is only approximately 4 generations for more remote inbreeding precluded, with the approximation becoming poorer with more remote inbreeding precluded.

**Proscribing  $n$  generations of inbreeding with two progeny per mating pair:** The heuristic argument of an increment of 4 generations in expected time to a common ancestor for each generation inbreeding preclusion is extended does not even hold for precluding full sib mating if there are two progeny per mating pair, because the proscription of full sib mating (as opposed to randomly choosing parents) increases the probability that two genes in different mating pairs came from the same mating pair the previous generation. As noted above, the expected time to a common ancestor with two progeny per mating pair is the same for random mating and for full sib mating precluded.

**Maximum avoidance of inbreeding:** Inbreeding can only be avoided for  $n$  generations in a population of  $N = 2^n$  mating pairs, at which time all mating pairs are  $n + 1$  fold  $n^{\text{th}}$  cousins ( $n = 0$  provides full sib mating,  $n = 1$  provides double first cousin mating). This avoidance of inbreeding requires two progeny per mating pair. The expected time since a common ancestor for two genes in an individual under such mating structures is  $8N - 2$  generations, which was found for random mating and full sib mating precluded with two progeny per mating pair. This is not always the expected time since a common ancestor with two progeny per mating pair, the expected time was only  $6 = 8 \times 2^0 - 1$  if there was full sib mating no matter how large the population was, and other examples could be given. But this provides a bound on the expected time since a common ancestor. With individuals randomly choosing their parents subject to constraints on inbreeding, as the number of generations inbreeding is proscribed increases, the constraints alter the Poisson progeny distribution into two progeny per mating pair, and the  $8N - 2$  generation expected time since a common ancestor is obtained.

#### References:

- A. H. Bittles and I. Egerbladh, *The Influence of Past Endogamy and Consanguinity on Genetic Disorders in Northern Sweden*, *Annals of Human Genetics* **69** (2005), 549–558.
- R. B. Campbell, *The Coalescent Time in the Presence of Background Fertility Selection*, *Theoretical Population Biology* **55** (1999), 260–269.
- L. L. Cavalli-Sforza and W. F. Bodmer, *The Genetics of Human Populations*, W. H. Freeman, San Francisco, 1971.//
- M. Kimura and J. F. Crow, *On the Maximum Avoidance of Inbreeding*, *Genetical Research, Camb.* **4** (1963), 399–415.
- J. F. C. Kingman, *On the Genealogy of Large Populations*, *Journal of Applied Probability*, **19A** (1982), 27–43.
- J. E. Pettay, S. Helle, J. Jokela, and V. Lummaa, *Natural Selection of Female Life-History Traits in Relation to Socio-Economic Class in Pre-Industrial Human Populations*, *PLoS ONE* **2**(7): e603, doi:10.1371/journal.pone.0000606, (2007).
- J. Wakeley, L. King, B. S. Low, and S. Ramachandran, *Gene Genealogies within a Fixed Pedigree, and the Robustness of Kingman’s Coalescent*, *Genetics* **190** (2012), 1433–1445.
- S. Wright, *Systems of Mating. II. The Effects of Inbreeding on the Genetic Composition of a Population*, *Genetics* **6**(1921), 124–143.