



ELSEVIER

Contents lists available at ScienceDirect

## Journal of Theoretical Biology

journal homepage: [www.elsevier.com/locate/yjtbi](http://www.elsevier.com/locate/yjtbi)

# The effect of inbreeding constraints and offspring distribution on time to the most recent common ancestor



R.B. Campbell\*

Department of Mathematics, University of Northern Iowa, Cedar Falls, IA 50614-0506, United States

## HIGHLIGHTS

- Banning inbreeding increases time to MRCA by 4 generations per generation banned.
- Several mating structures with 2 progeny per couple have the same time to MRCA.
- Recent pedigree information provides little information on the time to the MRCA.

## ARTICLE INFO

## Article history:

Received 24 December 2014

Received in revised form

22 June 2015

Accepted 24 June 2015

Available online 2 July 2015

## Classification codes:

2008 MSC: 92D15, 92D25

## Keywords:

Coalescent

Genetic identity

Pedigrees

## ABSTRACT

The expected time to the most recent common ancestor (MRCA) of two alleles in a diploid individual is  $4N + 2$  under random mating with a Poisson progeny distribution, but  $8N - 2$  under maximum avoidance of inbreeding, which entails two progeny per mating pair. ( $N$  is the number of mating pairs, hence  $2N$  is the number of individuals, hence  $4N$  is the number of alleles.) The interrelationship of inbreeding constraints and offspring distribution is investigated by varying the level of sib mating: prohibiting sib mating increases the time to MRCA by four generations and decreases the variance of the offspring distribution by  $2/N$ . With two progeny per mating pair, the expected time to the MRCA is  $8N - 2$  under both random mating and sib mating prohibited, as well as under maximum avoidance of inbreeding, but this result does not hold for all mating structures with two progeny per mating pair.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

All alleles are ultimately identical by descent, the time to the most recent common ancestor (MRCA) governs the likelihood that they are identical by state (a mutation has not occurred since their common ancestor). The time to a common ancestor is impacted by the progeny distribution and the mating structure of the population, with increased variance of the progeny distribution and consanguineous mating reducing the time to a common ancestor. Consanguineous marriages were common in the millennia preceding the common era including extensive sib mating among the pharaohs. More recently, there has been cousin mating among the royal houses of Europe. There has also been prohibition of consanguineous marriages, even third cousin mating required dispensation in the Catholic church until 1917 (Cavalli-Sforza and Bodmer, 1971). This dichotomy remains today

with some cultures banning consanguineous marriages while other cultures favor them (Cavalli-Sforza and Bodmer, 1971; Krawczak and Barnes, 2010).

Pedigree information from parish records has allowed calculation of the frequency of consanguineous matings (Bittles and Egerbladh, 2005; Pettay et al., 2007; Wakeley et al., 2012). Some of these include information up to sixth degree cousins. A question which will be addressed is whether this is sufficient information to determine genetic identity.

The mating structure of a population is only important within the context of the population size. If the population size is infinite, the expected time since a common ancestor will be infinite as long as the pedigree grows (Lachance, 2009), because an ancestor which escapes the pedigree structure will never return. If the mating structure provides a de facto finite ancestral population size (e.g., full sib mating or double first cousin mating), then the actual population size is not relevant and the expected time to a common ancestor is finite. We want more generality than regular systems of inbreeding (Arzberger, 1988; Kimura and Crow, 1963; Wright, 1921) because they generally require too much structure, often entailing a

\* Tel.: +1 319 273 2447; fax: +1 319 273 2546.

E-mail address: [campbell@math.uni.edu](mailto:campbell@math.uni.edu)URL: <http://www.math.uni.edu/~campbell>

finite ancestral population size, and often requiring exactly two progeny per individual, which confounds the effect of the progeny distribution with the effect of the pedigree structure (Campbell, 1993). However, we shall employ regular systems of inbreeding to bound genetic behavior.

## 2. The model

This analysis is based on the observation by Kingman (1982a) that the Poisson progeny distribution is equivalent to each individual “chooses its parents at random” which introduced the coalescent into population genetics (Hein et al., 2005; Kingman, 1982a,b; Wakeley, 2008). In particular, we study the trajectory of two alleles in an individual backward in time to find the coalescent time for those two alleles, which is the time to their MRCA. We employ a dioecious diploid model with  $N$  monogamous mating pairs, hence  $2N$  individuals, hence  $4N$  alleles (at a single locus). Because this model is diploid with monogamous mating pairs, an individual chooses a mating pair for its parents rather than choosing a single individual as occurs in haploid models or two individuals as occurs under unconstrained random mating.

Specifically, the model is that every generation has  $N$  monogamous mating pairs. The  $2N$  individuals in those pairs are randomly assigned to monogamous mating pairs (their parents) the previous generation. If an assignment violates an inbreeding constraint, that individual will be randomly reassigned until no constraint is violated; inbreeding constraints may specify either that inbreeding occurs or is avoided. The inbreeding constraints will alter the progeny distribution so that it is not Poisson. The pedigree goes backward in time, but we shall employ generation 1 as the generation before generation 0, so that increasing the generation number goes back further in time. All calculations are in real time without rescaling.

## 3. Full sib mating specified

The first model specifies the amount of full sib mating with no other constraints. Random mating is a special case of this for which the time to the MRCA is  $4N+2$ .

We employ  $s$  for the fraction of matings which are between (full) sibs (since we are assuming monogamous mating pairs, there are no half-sibs). The model entails  $N$  mating pairs each generation,  $sN$  of them are randomly chosen to be randomly assigned with replacement a single mating pair as parents, while the rest are assigned with replacement two distinct mating pairs as parents (i.e., one mating pair the previous generation for each individual in the mating pair). Going forward in time, this entails  $sN$  mating pairs chosen (with replacement) to parent mating pairs (i.e., full sib matings the next generation), and then mating pairs are chosen (with replacement) to parent the remaining individuals subject to the constraint that no mating pair is the parent of both individuals in a mating pair. An alternative model is presented in the Appendix.

The mating pairs assigned two sets of parents must be assigned two different sets of parents, but they may be assigned mating pairs chosen by individuals in other mating pairs, including individuals in full sib mating pairs. This process is iterated through the ancestral generations. This process can be analyzed using three quantities:  $T$ , the expected time to a common ancestor of two alleles in an individual;  $U$ , the expected time to a common ancestor of two alleles in two individuals within a mating pair; and  $V$ , the expected time to a common ancestor of two alleles in two different mating pairs (see Fig. 1).

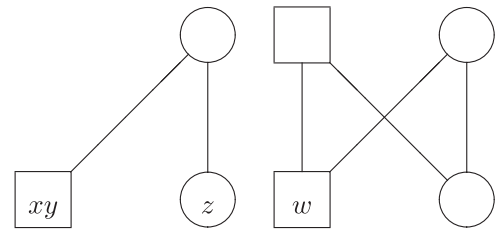


Fig. 1. Location of alleles. Alleles  $x$  and  $y$  are in the same individual, alleles  $x$  and  $z$  are in a mating pair, but not in the same individual, and alleles  $x$  and  $w$  are in two different mating pairs. Squares represent males, circles represent females, and lines extend upward to progeny.

This provides a system of three equations:

$$T = U + 1 \tag{1}$$

because if two alleles are in an individual, one was in each parent the previous generation – two alleles in an individual in generation 0 were two alleles in two individuals within a mating pair in generation 1.

$$U = s(.25 + .25(1+T) + .5(1+U)) + (1-s)(1+V) \tag{2}$$

If that mating pair in generation 0 were sibs (shared parents in generation 1), the alleles would be copies of the same allele .25 of the time, copies of two alleles in an individual .25 of the time, and copies of alleles in two individuals in a mating pair .5 of the time. If that mating pair were not sibs, their ancestral alleles would have been in two different mating pairs in generation 1.

$$V = (1/N)(.25 + .25(1+T) + .5(1+U)) + (1-1/N)(1+V) \tag{3}$$

If the alleles were in two different mating pairs in generation 0,  $1/N$  of the time they would have been in the same mating pair in generation 1 by random choice of parents, and  $(1-1/N)$  of the time they would have been in different mating pairs in generation 1. (Even if one or both individuals in generation 0 were the result of sib mating, they could have chosen the same parents.)

Eqs. (1)–(3) have the solution  $T = 6 + (1-s)4N$ . If  $s=1$  (mandatory sib mating),  $T$  is equal to 6; this can be easily shown directly since mandatory full sib mating provides each individual an ancestral population of size 2. If  $s=0$  (sib mating precluded),  $T = 4N+6$ . The expected time to a common ancestor decreases linearly as  $s$  increases. Random mating is provided by  $s = 1/N$ , which yields  $T = 4N+2$ ; this can be derived directly by noting that (2) and (3) for  $U$  and  $V$  above are the same if  $s = 1/N$ , so combining them with (1) provides the single equation  $T = 1 + 1/(4N) + (1+T)/(4N) + T(1-1/(2N))$ .

The assumption of constant population size requires that the expected number of progeny per mating pair is two, but inbreeding constraints will cause the variance of the progeny distribution to change. With random mating (individuals randomly choose parents), each of the  $N$  mating pairs will have probability  $1/N$  of parenting each of the  $2N$  progeny, and the binomial distribution provides that the variance of the number of progeny per mating pair is  $2N(1/N)(1-1/N)$ . If sib mating is precluded, each mating pair will have probability  $2/N$  of having exactly one progeny in each of the  $N$  mating pairs the next generation, and by the binomial distribution the variance of the number of progeny will be  $N(2/N)(1-2/N)$ . With mandatory sib mating, each of the  $N$  mating pairs will have probability  $1/N$  of parenting each of the mating pairs (i.e., both progeny in the mating pair) the following generation, and by the binomial distribution the variance of the number of progeny will be  $4 \times N(1/N)(1-1/N)$ . The initial 4 is present because each parental mating pair has 2 or 0 progeny in a mating pair the next generation rather than 1 or 0. Random mating entails that the fraction of sib-mating is  $s = 1/N$ , and  $2N(1/N)(1-1/N) = (1/N) \times 4N(1/N)(1-1/N) + (1-1/N) \times N(2/N)$

Download English Version:

<https://daneshyari.com/en/article/6369514>

Download Persian Version:

<https://daneshyari.com/article/6369514>

[Daneshyari.com](https://daneshyari.com)