perm.pub/dsi:D9qSdCY6GPrxthT3ZnFouEU35ow/0.1
Additional formats and editions available online.

Edition 0.1

Obsolete

Author date: 2022-03-31

Citation:
E. Castedo Ellerman (2022) "Lineal admixture time" perm.pub https://perm.pub/dsi:D9qSdCY6GPrxthT3ZnFouEU35ow/0.1

Lineal admixture time

E. Castedo Ellerman (castedo@castedo.com)

Abstract

This discussion document defines lineal admixture time, a key construct under mathematical study by ECE. Desired future outcomes are mathematical and computation tools, based on lineal admixture time, that become useful to downstream empirical research, specifically the estimation of the timing and extent of interbreeding between previously separated populations. Feedback is greatly appreciated, especially regarding the relevance to downstream empirical research.

> ***WORKING DRAFT***

Introduction

Genealogies are a link between genetic evidence and population events of the past. By genealogies we mean the parent-offspring relations of a group of individuals and their ancestors. In contrast to estimates of purely genetic quantities, estimates of quantities about genealogies can have closer correspondence to non-genetic lines of evidence, such as archaeological, geological, and cultural evidence.

Definition of lineal admixture time

We define *lineal admixture time* per lineage. By (organismal) lineage we mean a linear chain of parent-offspring pairs descending through a genealogy. The following definition assumes a past point in time when all individuals are classified into separate (non-admixed) subpopulations. Lineages do not have a lineal admixture time *point* if they consist of individuals from only one non-admixed subpopulation. For all other lineages, the lineal admixture time *point* is the fertilization time of the first individual with parents from two distinct non-admixed subpopulations. From an observation time at the end of a lineage, the lineal admixture time *duration* is the time since the linear admixture time *point*. If this time point does not exist, then the time duration is zero.

The *average lineal admixture time* of an individual is averaged across all lineages ending in that individual. At every merging of lineages, through the mother and father of offspring, equal weighting is given to the two respective sets of lineages. The *average lineal admixture time* of a group of individuals at an observation time point, is the average lineal admixture time across all those individuals.

Advantages of lineal admixture time

Lineal admixture time has a number of advantages as a quantity for research.

Firstly, it is independent of any particular model of an admixture process. With enough genealogical information one could calculate lineal admixture times of a real group of individuals. There is one true set of lineal admixture times for a real population, even though realistically, we can only hope to estimate those numbers based on evidence and models. But the definition exists independent of any particular model used for estimation.

Secondly, the interpretation of lineal admixture time does not require fluency in probability theory. Researchers with interests in empirical evidence and not mathematics can make use of estimates of lineal admixture times.

Thirdly, we conjecture that *distributions* of lineal admixture times across individuals, groups of individuals, and chromosomes will prove to be a useful mathematical tool in the timing of admixture. This conjecture is based on not-yet-documented mathematical work by ECE. Lineal admixture time is conveniently representable as a random variable from a stochastic process in which lineages are random objects.

Mendel's hybrid generation numbers

Lineal admixture time can be interpreted as a generalization of Mendel's hybrid generation numbers. In his famous experiments with peas, Mendel [1] assigns ordinal numbers (first, second, third, ...) to successive generations of hybrids. Today, the first and second hybrid generations are often referred to as F_1 and F_2 [2]. An example of using these generation numbers to infer admixture timing is in the study of admixture in Greenland [3].

Mendel's hybrid generation numbers correspond to a simple special case with the following assumptions:

- generation numbers are assigned to mature individuals at non-negative integer time points
- after time zero, all mature individuals are admixed
- parents are mature and reproducing for only one time point
- the fertilization time of offspring is when the parents are reproducing
- maturation time follows exactly one time unit after fertilization

With these assumptions, lineal admixture time durations observed at positive integer time points are exactly the hybrid generation numbers.

Fin

This document presents a definition of *lineal admixture time*. This quantity is the basis for current mathematical studies of ECE. Feedback and input is greatly appreciated, especially regarding the relevance to downstream empirical research. In particular, the following questions are of particular interest:

1. Is *lineal admixture time* easily understood and interpretable without fluency in probability theory?
2. Does *lineal admixture time* benefit from being comparable to non-genetic lines of evidence (i.e archaeological, geological, linguistic, historical, cultural)?
3. What are existing terms or literature for this specific definition of admixture time?

4. Are there reasons for confusion with these names chosen in this document
    - lineal admixture time
    - lineage
    - genealogy

   or how they are used?

References

1. Abbott S, Fairbanks DJ. Experiments on Plant Hybrids by Gregor Mendel. Genetics. 2016;204: 407–422. doi:10.1534/genetics.116.195198
2. Hartl DL, Jones EW. Essential genetics: A genomics perspective. 4th ed. Boston: Jones; Bartlett Publishers; 2006.
3. Waples RK, Hauptmann AL, Seiding I, Jørsboe E, Jørgensen ME, Grarup N, et al. The genetic history of Greenlandic-European contact. Current biology: CB. 2021;31: 2214–2219.e4. doi: 10.1016/j.cub.2021.02.041