

Chapter 1

Overview

1.1 What is a muller diagram?

Muller diagrams excel at visualizing evolutionary dynamics of a population over time. A common use case, and the intended use case during development, describes the abundance and succession of genotypes within a population derived from sequencing data from samples at specified timepoints. This provides a useful method of quickly discerning the evolutionary relationship between genotypes seen in a population and how this changes over time.

1.2 Some Terminology

trajectory/mutational trajectory A set of frequency measurements of a specific mutation at each sampled time point over the course of an evolution experiment.

genotype A group of trajectories which follow a common path.

clustering The process of grouping trajectories together to form a genotype. There are two basic types of clustering. Agglomerative clustering describes the case where trajectories or small clusters are grouped together to form a larger cluster. Divisive clustering occurs when an existing cluster is split up into smaller clusters. These scripts use both types of clustering.

uncertainty A parameter provided by the user which indicates the uncertainty in the frequency measurements. The default value of 0.03 was chosen based on the performance of Breseq.

background When a mutation arises in a population, it is said to be in the background of the mutations that appeared in the same population prior to its detection. A core purpose of these scripts is to describe how any given mutation is related to those that arose before its detection (i.e. its background).

genotype fixing/sweeping When a genotype fixes, it removes all pre-existing variation other than itself. All subsequent mutations arise in the background of this genotype.

1.3 Limitations of the scripts