

Detecting Selection on Segregating Gene Duplicates in a Population

Tristan Stark¹, Rebecca Kaufman¹, Maria Maltepes¹, and David Liberles¹

¹Temple University

December 9, 2020

Abstract

Gene duplication is a fundamental process that has the potential to drive phenotypic differences between populations and species. While evolutionarily neutral changes have the potential to affect phenotypes, detecting selection acting on gene duplicates can uncover cases of adaptive diversification. Existing methods to detect selection on duplicates work mostly inter-specifically and are based upon selection on coding sequence changes, here we present a method to detect selection directly on a copy number variant segregating in a population. The method relies upon expected relationships between allele (new duplication) age and frequency in the population dependent upon the effective population size. Using both a haploid and a diploid population with a Moran Model under several population sizes, the neutral baseline for copy number variants is established. The ability of the method to reject neutrality for duplicates with known age (measured in pairwise d_S value) and frequency in the population is established through mathematical analysis and through simulations. Power is particularly good in the diploid case and with larger effective population sizes, as expected. With extension of this method to larger population sizes, this is a tool to analyze selection on copy number variants in any natural or experimentally evolving population.

Hosted file

mol_ecol_1220_continuous_line_numbers.pdf available at <https://authorea.com/users/381943/articles/497870-detecting-selection-on-segregating-gene-duplicates-in-a-population>