

Exploring Tests of Selection in Pooled Data

Keywords: Genetic – Model organisms – Population genetics –

Abstract: For organisms with fast breeding cycles, artificial selection can be used to create sub-populations with different phenotypic traits. Genetic tests can be employed to identify causal markers for the phenotypes, as a precursor to engineering strains with a combination of traits. Traditional approaches involve analyzing crosses of inbred strains to test for co-segregation with genetic markers. Here we take advantage of cheaper next-generation sequencing techniques to identify genetic signatures of adaptation to the selection constraints. Obtaining individual sequencing data is often unrealistic due to cost and sample issues, so we focus on pooled genomic data. In this paper, we develop a statistical test for selection using pooled case and control populations. Extensive simulations are used to show that the approach works well for a wide range of population divergence times and strong selective pressures. It is also robust to reasonable variations in several different parameters, including window size, basecalling error rate, and sequencing coverage, and has similar power to statistics such as Tajima's D applied to underlying haplotypes. We then demonstrate the viability of this method in two independent *Drosophila melanogaster* populations bred under selection for hypoxia tolerance and accelerated development. Testing for extreme hypoxia tolerance showed clear signals of selection, pointing to loci that are important for hypoxia adaptation.

Authors:

- Nitin Udpa nudpa@ucsd.edu United States UCSD
- Dan Zhou d2zhou@ucsd.edu United States UCSD
- Gabriel Haddad ghaddad@ucsd.edu United States UCSD
- Vineet Bafna vbafna@cs.ucsd.edu United States UCSD

This abstract was generated automatically from the easychair submission page.