



Figure S7 The relationship between age of the sweep and misclassification rate when modeling soft sweeps from recurrent adaptive mutation.

(A) The fraction of simulated windows containing a hard sweep ($\alpha=1000$) classified as hard, soft (from recurrent mutation), or neutral by an SVM leveraging allele frequency information is shown according to the time in the past at which the sweep completed (in units of $2N$ generations). The most recent sweep examined in this plot completed $0.000625 \times 2N$ generations ago, and we examined older sweeps by continually doubling the time since fixation, stopping at a sweep time of $1.28 \times 2N$ generations in the past. (B) Same as panel A, but using a classifier leveraging haplotype information. (C) Results from an SVM leveraging LD information. (D) An SVM leveraging our full set of summary statistics (Methods).