

## File S1

### Supporting Information

**Power and False Positive Rate.** In order to evaluate the power of SFselect and XP-SFselect to detect positive selection as compared to other neutrality tests, we applied these tests to several datasets simulated under different model parameters. For a given test on a given dataset, the power at 5% false positive rate (FPR) was estimated as the fraction of test-statistic values exceeding a set threshold when applied to the selected samples. The threshold was set to the top 5% of the null distribution, obtained by applying the test to neutral samples. For cross-populations tests (including XP-SFselect) we used the same procedure, only applying the test to selected vs. neutral samples, while the null was obtained by applying the test to neutral1 vs. neutral2 samples.

**SVM implementation details.** We used a linear (dot product) kernel function SVM. Linear kernels have two important advantages. First, because feature-weights learned by a linear SVM represent a maximum-margin separating hyperplane of the training data in the problem space (rather than in a higher dimensional space), they correspond to the relative importance of features in separating the training data, making the trained SVM easily interpretable. Secondly, normalization of the training and testing data is done in the input space, without the need for complicated normalization of the kernel function itself (Graf *et al.* 2003).

The SVM implementation we used was from the LIBSVM library (Chang and Lin 2011), packaged in the python library scikit-learn (Pedregosa *et al.* 2011). For the parameter-specific SVMs, where we lacked sufficient simulated data to hold the test data out of training, we report power as the mean over 50-fold cross validation. For the general two-stage SVM (SFselect and XP-SFselect), testing and training were done on completely separate datasets.