

Figure 1. Simulation of selection and mutation rate bias in *Arabidopsis* mutation accumulation experiments. **a)** Overview of the simulation framework conducted using SLiM, modelled after the meristem development of the *Arabidopsis* from zygote to gamete. **b-g)** the results of these simulations, where the x-axis represents the true degree of mutation rate reduction ($\Delta\mu$) in gene bodies. The panels show the impact on measures of mutation bias and selection from various degrees of selection strength (s) and dominance (D) of *de novo* non-synonymous mutations: **b)** observed mutation rate (mutations per base pair per generation) in genic regions (scaled to *A. thaliana* genome size) **c)** observed mutation rate in intergenic regions, **d)** the perceived reduction of mutation rate in gene bodies, with the dashed line indicating the true reduction, **e)** the observed non-synonymous to synonymous (NS/S) mutation ratio, with the dashed line representing the neutral expectation, **f)** results of chi-squared tests of NS/S ratio relative to the neutral expectation, indicating the significance of deviation from neutrality ($p < 0.05$, i.e. "would we be able to detect selection?") and **g)** the estimated reduction of mutation rate in gene regions after accounting for selection on coding mutations, comparing the estimated and actual mutation rate reductions.

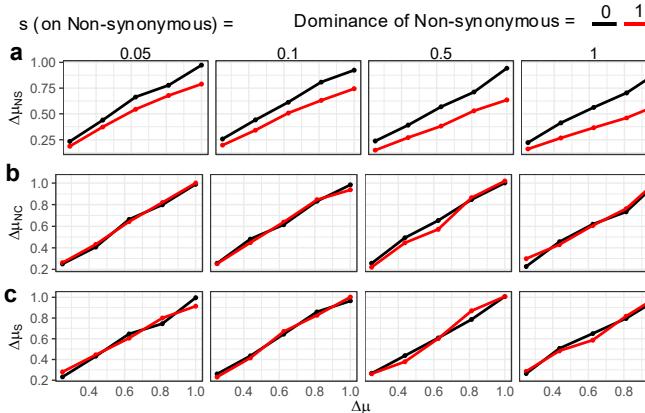


Figure 2. Measuring mutation rate reduction on individual components of genic regions.

Visualization only includes simulations with strongly deleterious non-synonymous mutations ($s > 0.03$) **a**) The reduction of non-synonymous mutations ($\Delta\mu_{NS}$) **b**) The reduction of non-coding genic (i.e. introns or untranslated regions) mutations ($\Delta\mu_{NC}$) **c**) The reduction of synonymous mutations ($\Delta\mu_S$).

Selection and mutation bias signals in *Arabidopsis* MA datasets
(bootstrapped 95% CI)

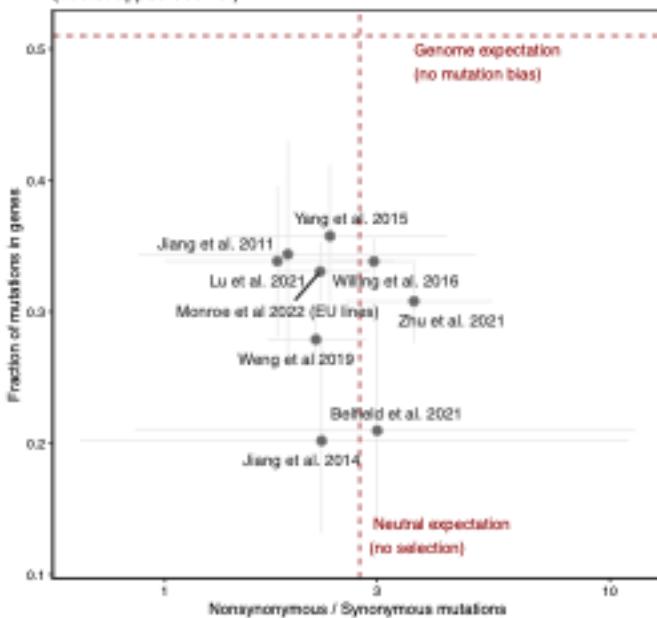


Figure 3. Genic mutation fraction vs. nonsynonymous/synonymous (NS/S) ratio for individual *Arabidopsis* MA datasets (points with 95% bootstrapped CIs, $B = 1,000$). The vertical dashed line ($NS/S = 2.71$) is the neutral expectation from spectrum \times codon use; the horizontal dashed line (0.501) is the genome-wide genic fraction. All datasets fall below the genic expectation (consistent with a genic mutation bias), but all NS/S CIs overlap the neutral expectation (i.e. we do not detect systematic purifying selection on coding mutations in any dataset). Genic fraction and NS/S were not correlated across datasets: those with a lower fraction of mutations in genes did not show a correspondingly lower NS/S ratio.

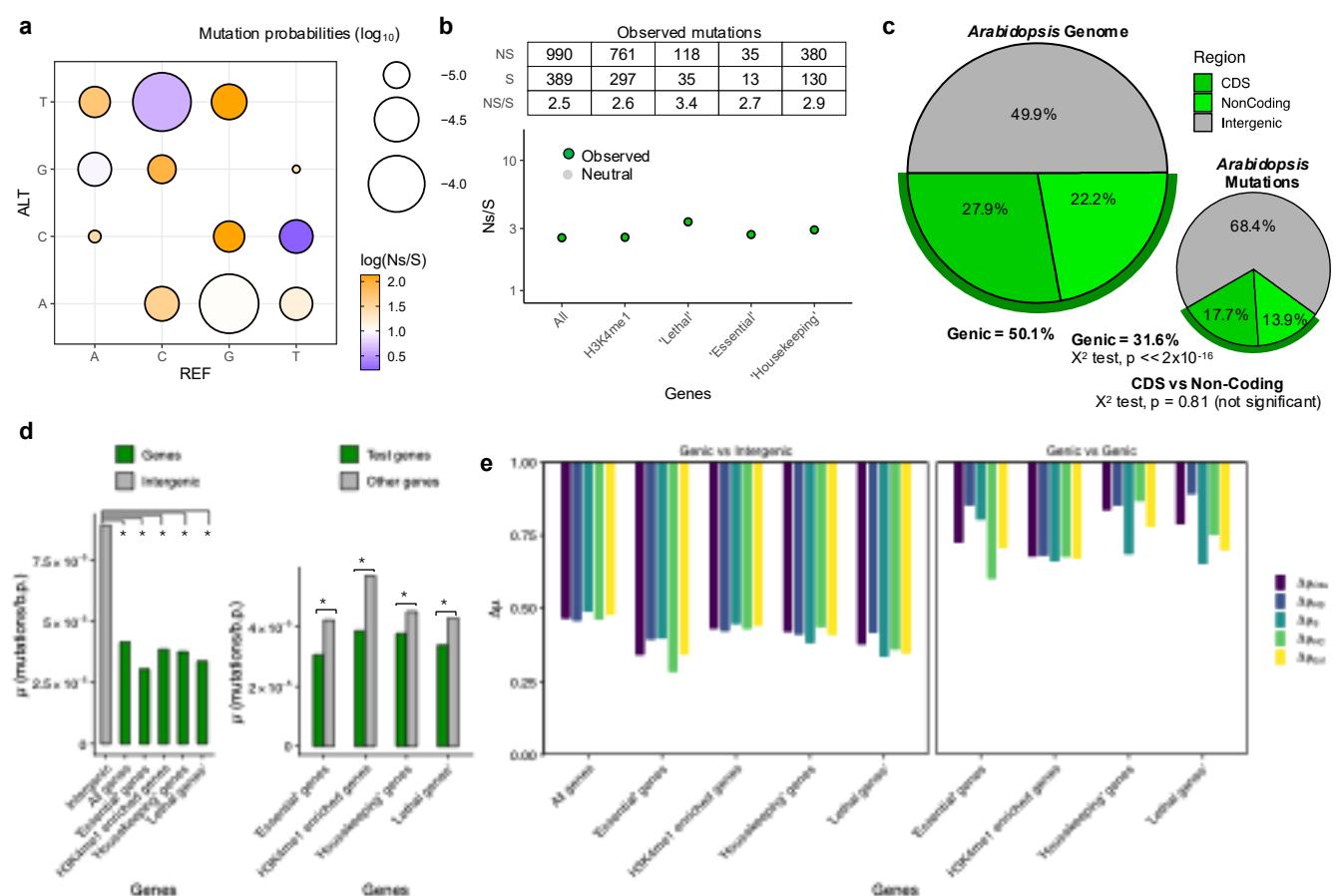


Figure 4. Selection and mutation bias in *Arabidopsis* mutation datasets. **a)** Projection of mutation spectra (REF > ALT) from non-coding regions onto *Arabidopsis* coding regions for estimating neutral Ns/S ratios. **b)** Random sampling of coding sequence mutations based on mutation probabilities (1000 iterations), with mutation number samples equal to the number of coding mutations found in all genes (n=1385 mutations), ‘essential’ genes (n=48), ‘lethal’ genes (n=154), H3K4me1-enriched genes (n=1062), and ‘housekeeping’ genes (n=510). **c)** Genome of *Arabidopsis* (TAIR10) (upper pie chart) and the proportion of single base substitution (SBS) mutations in each region (lower pie chart). Mutation datasets are from Weng et al. 2019, Belfield et al. 2021, etc. Total mutations, n=7836. **d)** Same gene groups as in (b), comparing mutation rates in gene body (CDS + Non-Coding) with intergenic regions. The left panel shows the comparison of mutation rates in specific gene groups with other genes (i.e. ‘Essential’ genes vs non-essential genes). Asterisks indicate p -value < 0.001 for chi-squared test. **e)** Measures of mutation rate reduction of individual components. For genic vs intergenic comparisons, the mean of sampled neutral Ns/S for all genes from (c) was used (left). The right panel shows genic comparisons of individual gene groups, with neutral Ns/S in these cases indicating the Ns/S observed in other genes for each comparison.