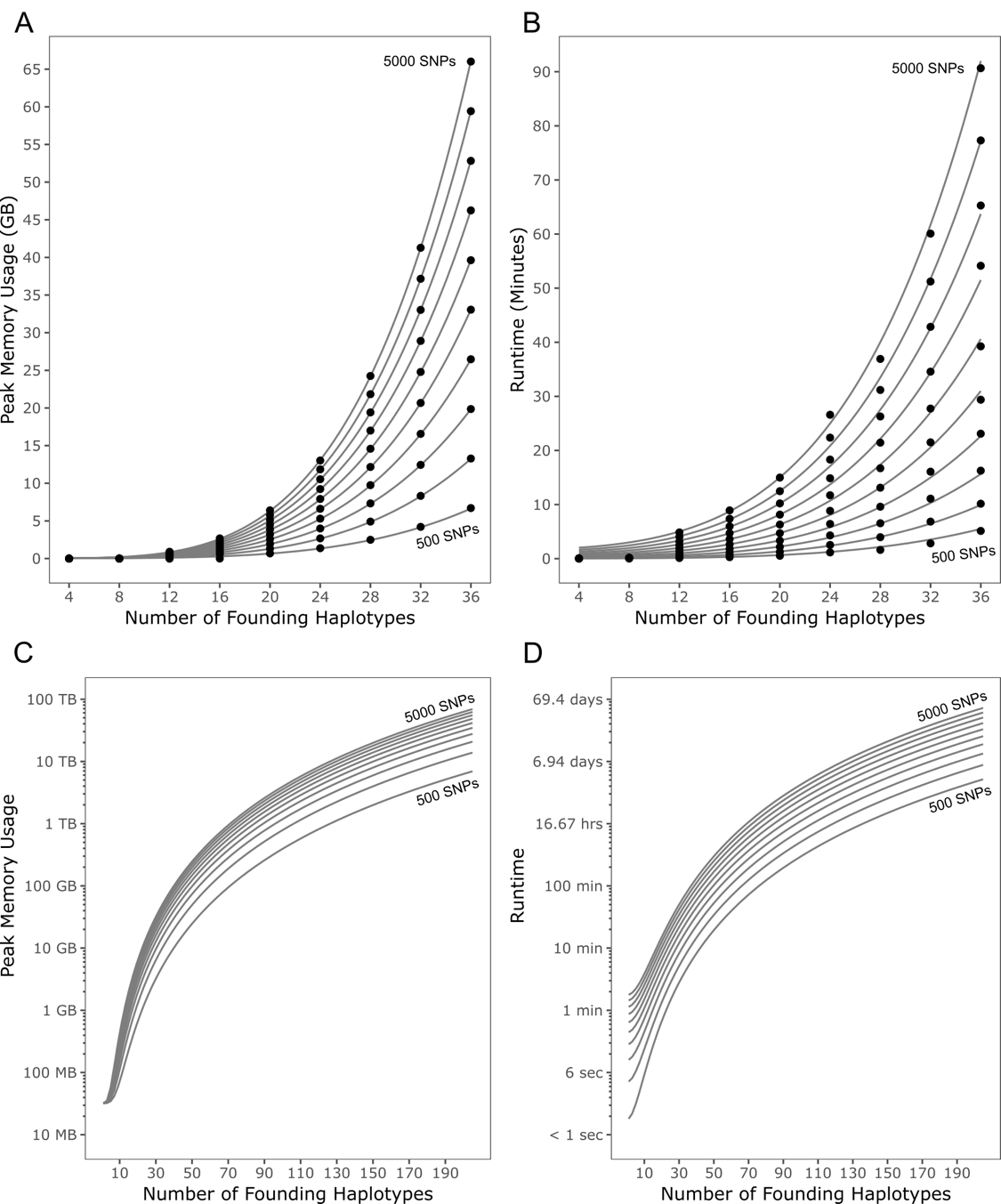


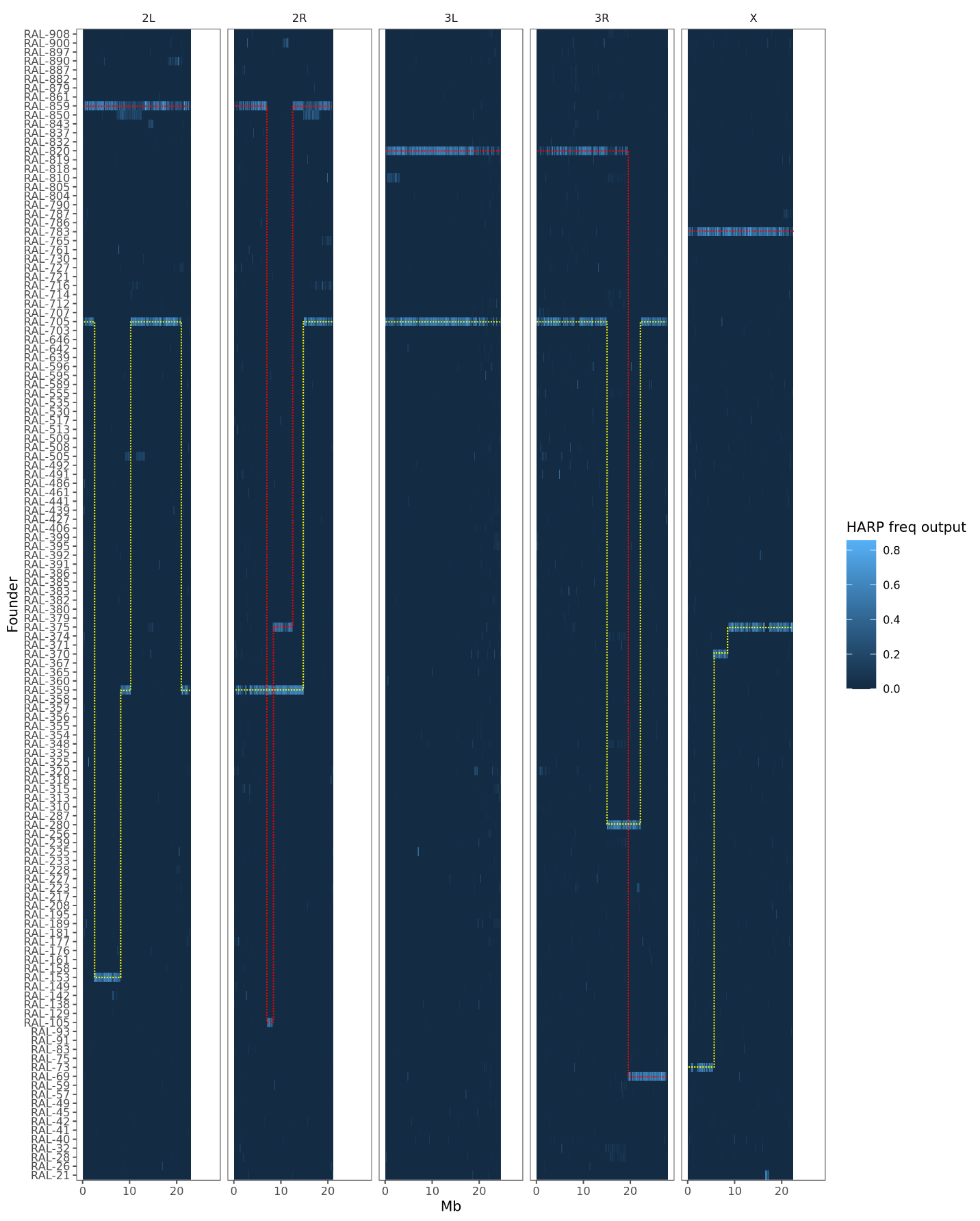
**Supplemental Figure S1**. Basic structure of the forward simulator pipeline.

Inbred founding lines (**A**) are randomly intercrossed to produce a recombinant population (**B**). Rapid generation of independent mapping populations is achieved by random down-sampling (**C**) and permutation of ancestry (**D**). Population genetic data is encoded in a highly compressed format **(E)** that references the positions of haplotype blocks instead of genotypes at every site, enabling us to generate 500 mapping populations for a given parameter combination.

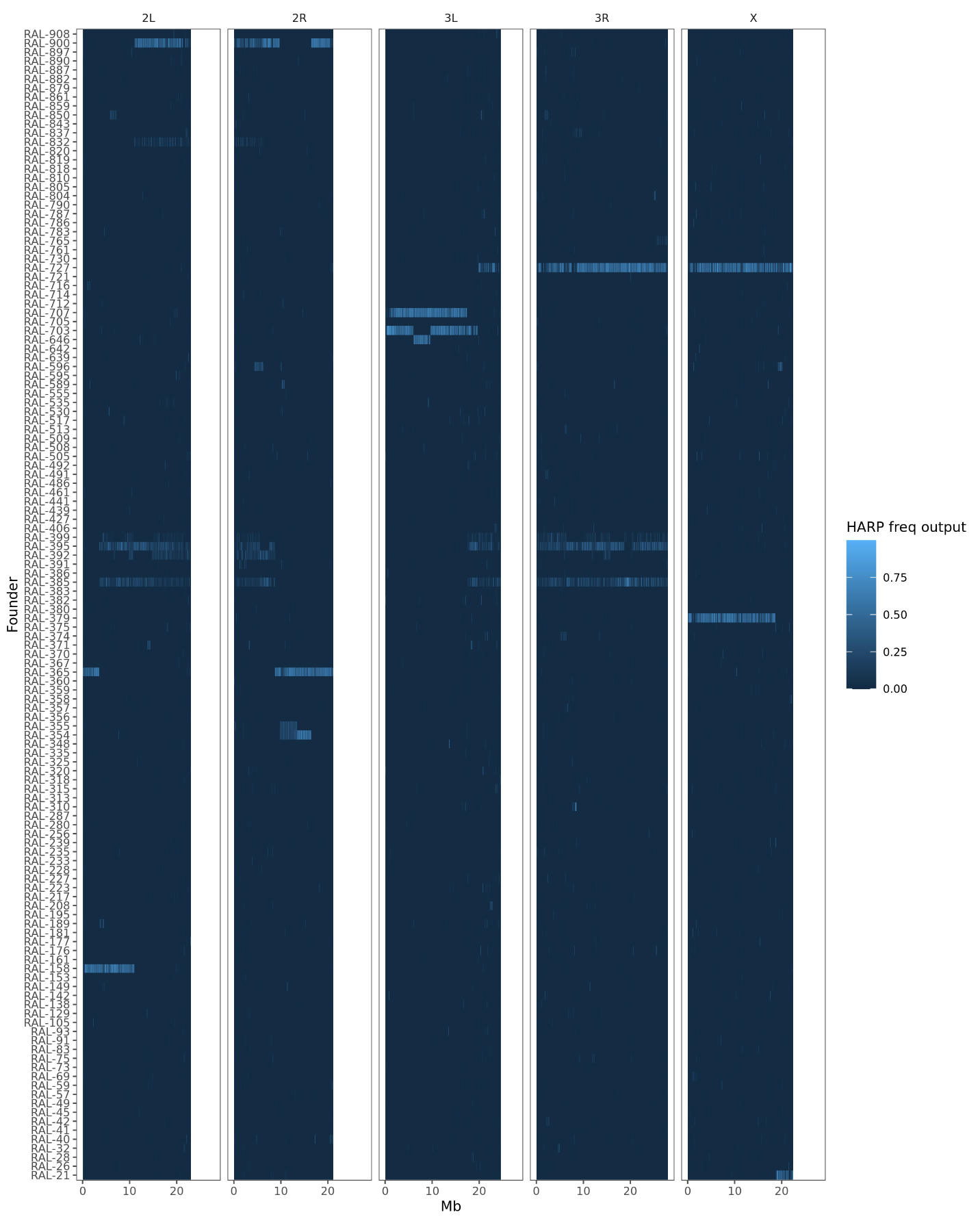


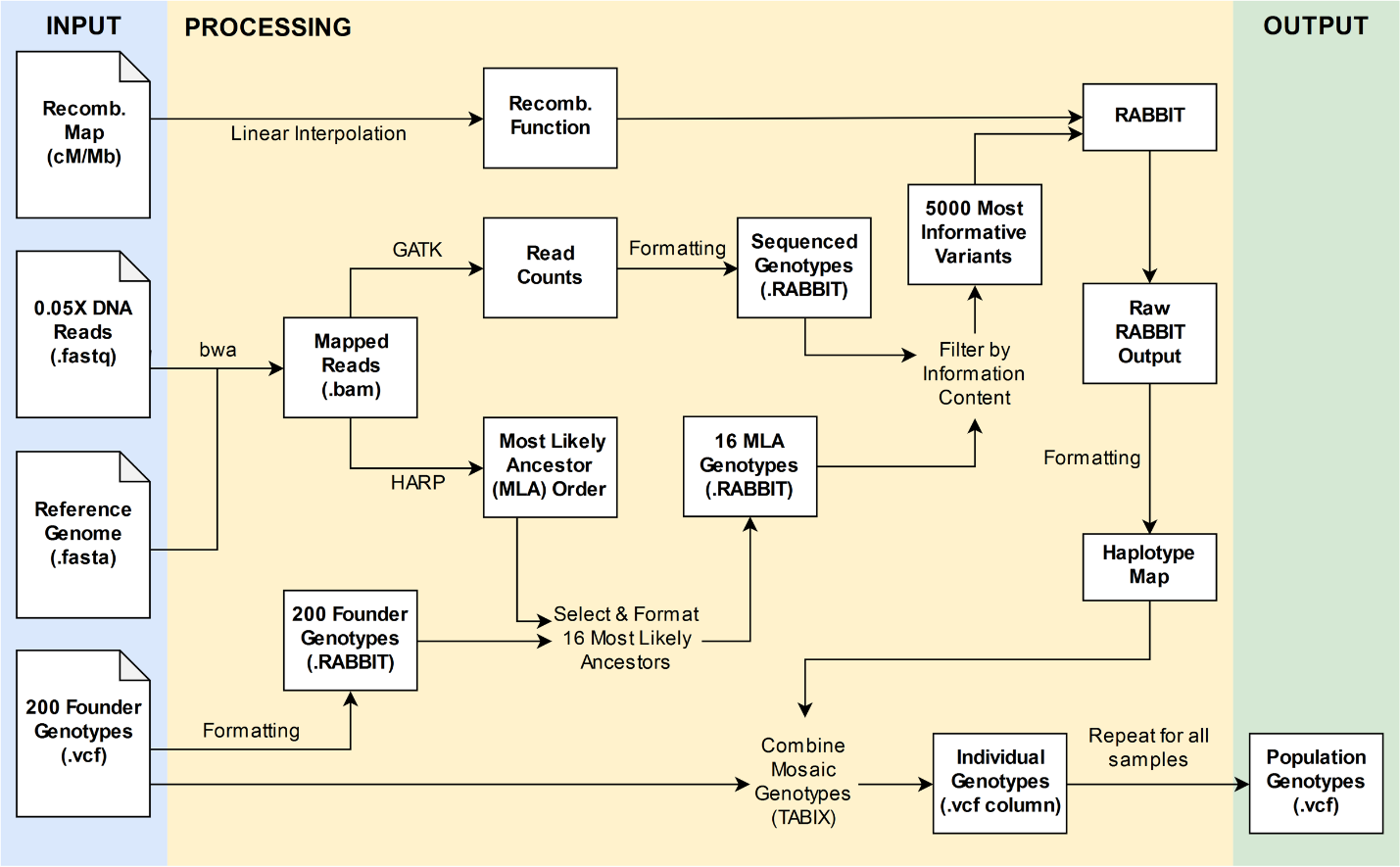
**Supplemental Figure S2**. Resource usage of RABBIT during haplotype reconstruction.

All reconstructions involve the same simulated 2L chromosome arm comprised of four haplotypes. Simulations included varied numbers of founding haplotypes (*N*) and a randomly selected set of markers (number of SNPs, *S*, incremented in steps of 500). All simulations included, at minimum, the four true haplotypes for the simulated individual. In **A** and **B**, points depict the mean of empirical values (over 10 replicates) and gray lines depict the defined regression models. Regression models for memory usage and runtime are extrapolated over a greater range for number of founding haplotypes in **C** and **D**, respectively. Peak memory grew linearly with number of SNPs used, and at a greater-than-linear rate with haplotypes (A). The runtime of RABBIT increased at a greater-than-linear rate for both number of SNPs and number of haplotypes, with the N parameter contributing most to resource cost (B). These models allowed us to estimate resource requirements at greater numbers of haplotypes (C & D) which would be unfeasible to measure empirically.

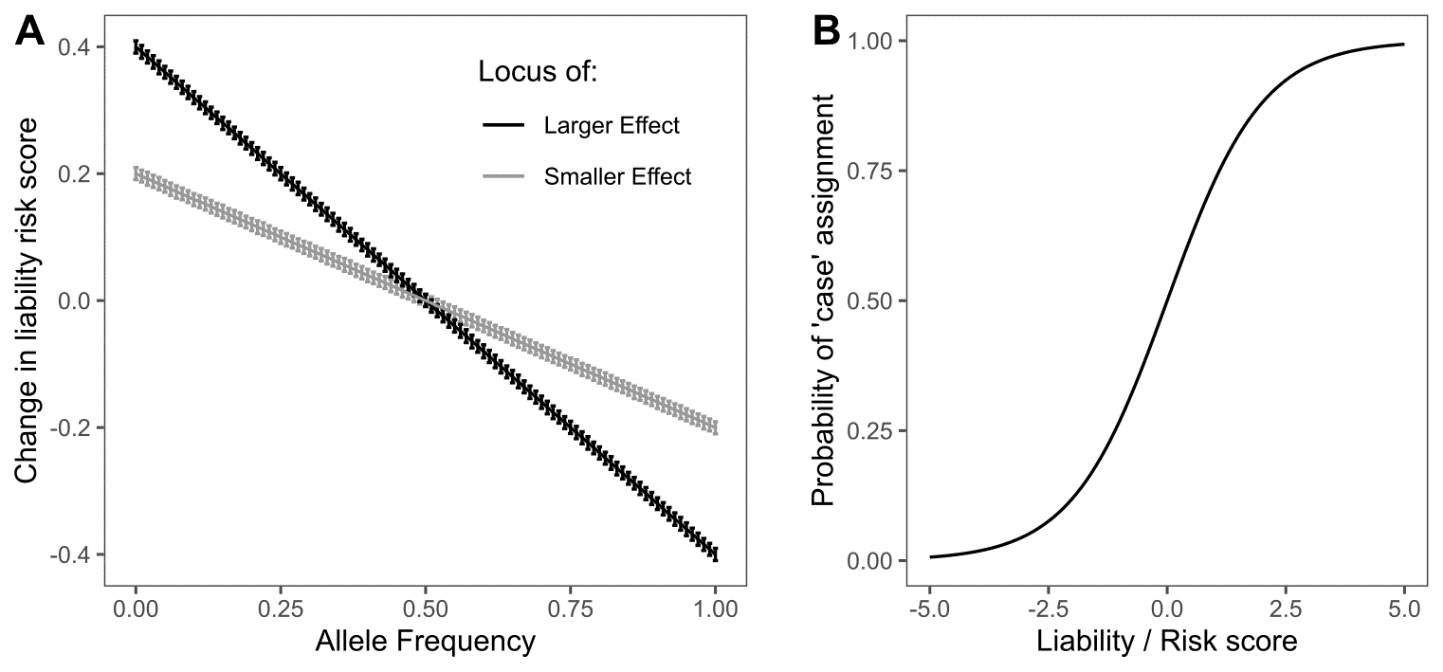


**Supplemental Figure S3.** Example of unambiguous haplotype inference from HARP. For each chromosome, each haplotype is indicated by red and yellow dotted lines.

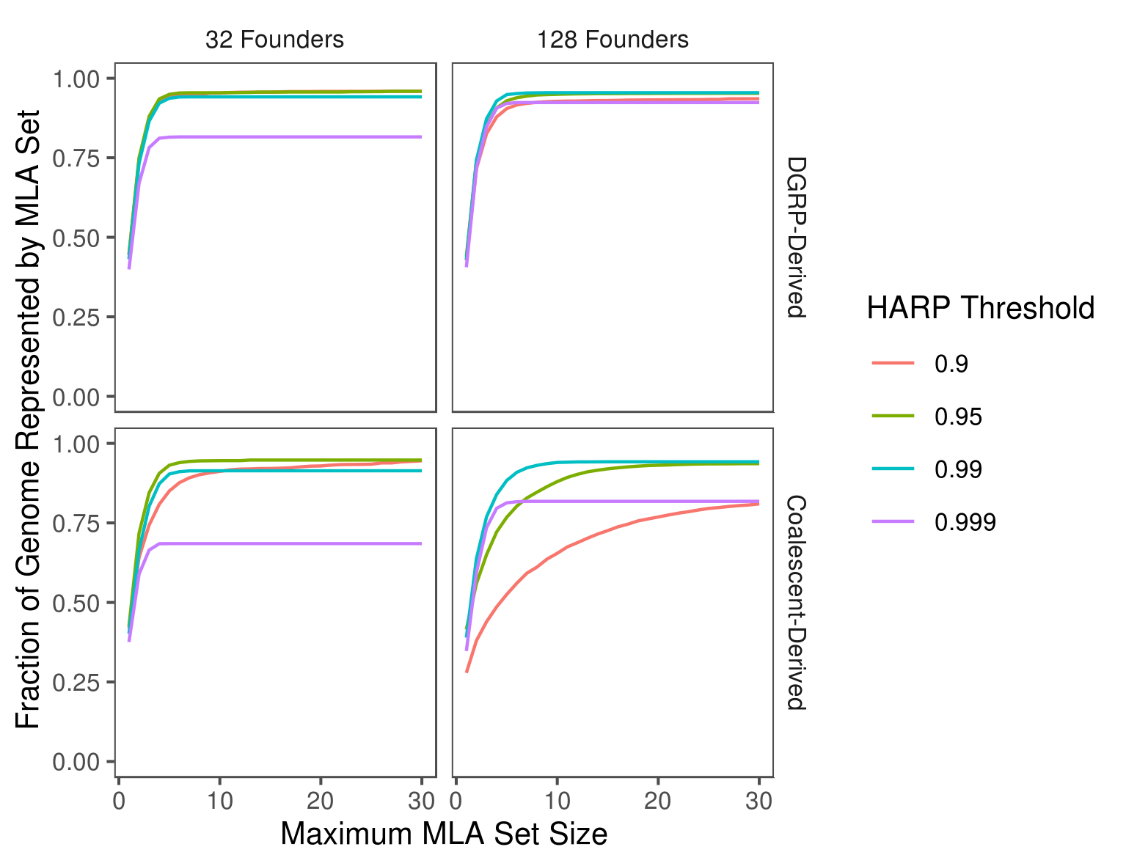
**Supplemental Figure S4.** Example of HARP freq output with ambiguous haplotype inference. For each chromosome, each haplotype is indicated by red and yellow dotted lines. For this individual’s autosomes, at least one haplotype cannot be unambiguously resolved.



**Supplemental Figure S5**. Workflow for reconstructing whole genomes from ultra-low coverage sequencing data. Three required inputs are sequencing reads (.fastq), a reference genome (.fasta), and homozygous (or phased) founder haplotypes (.vcf file). A recombination map can optionally be provided for use with RABBIT to inform variation in recombination rates across chromosomes. Flowchart designed in draw.io.

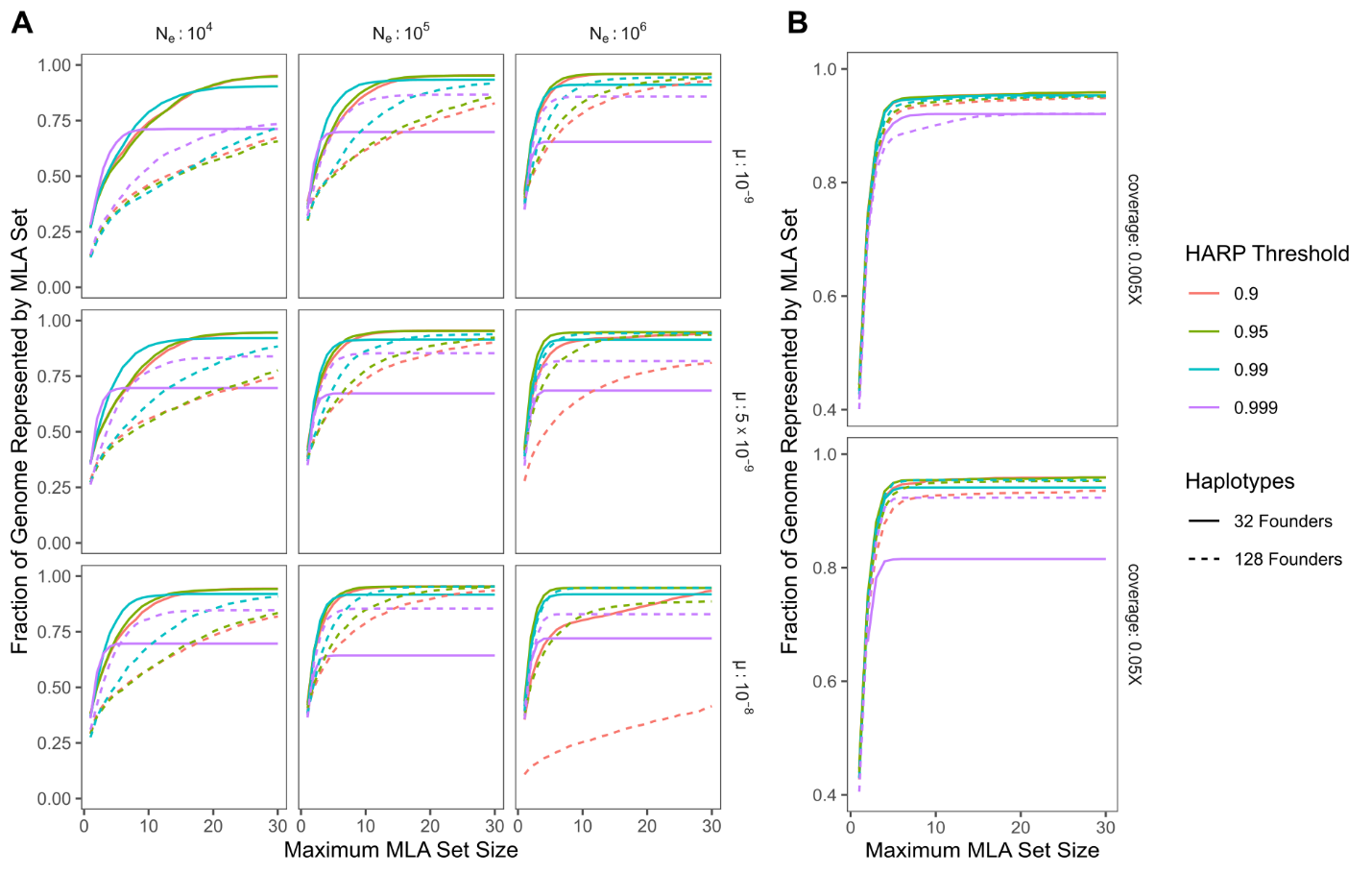


**Supplemental Figure S6.** Liability phenotype model for GWAS simulations. for transforming additive risk scores to probability of assignment to the case population, i.e. probability of developing a disease. Rare alleles increase risk score, which in turns increases probability of being assigned a case phenotype. **A**: The relationship between allele frequency and effect on liability risk score. Simulations were modeled as larger or smaller effect with a user-defined effect size coefficient. Error bars describe the standard deviation of a gaussian sample which constitutes random phenotypic noise. **B**: Logistic transformation of the form y=(1+*e*-x)-1 translates the liability score into probability of being assigned to case population.

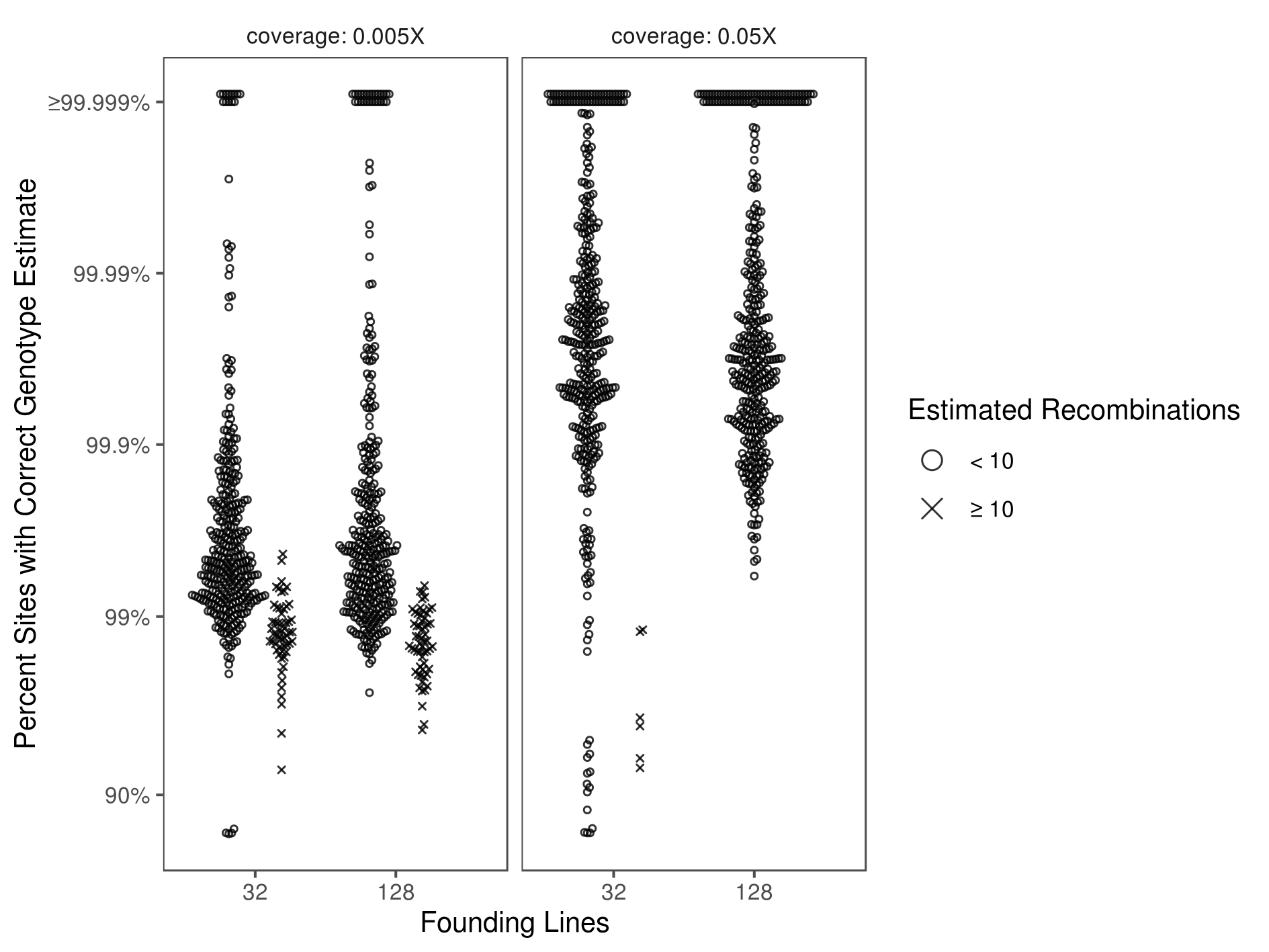


**Supplemental Figure S7**. Optimization curves for Most-Likely-Ancestor (MLA) selection.

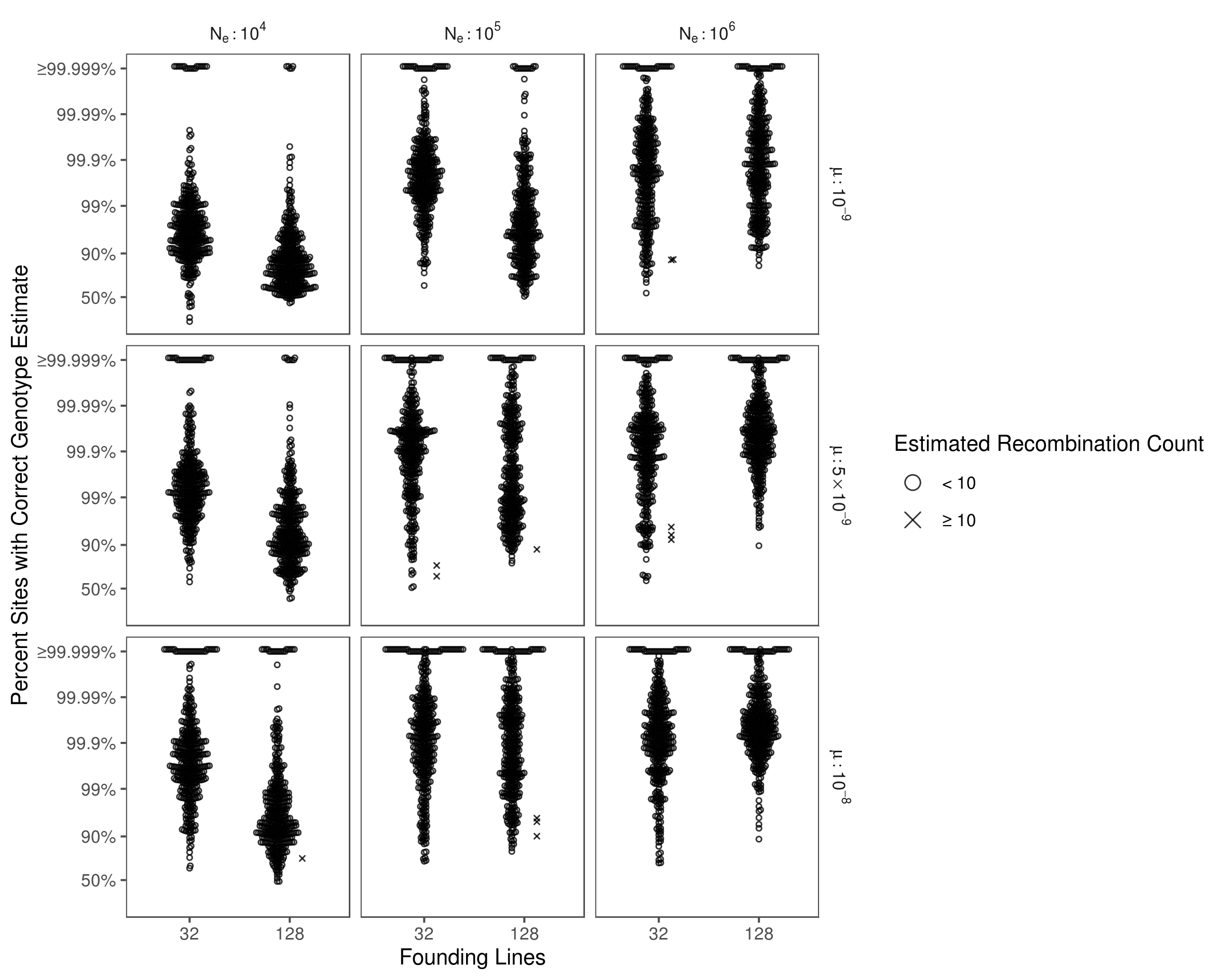
To ensure reconstruction was both accurate and computationally feasible, we selected the smallest set of most-likely-ancestors that still represent the greatest fraction of to-be-reconstructed genomes. Increasing the upper limit for the number of MLAs chosen improves the fraction of genome represented with diminishing returns. The HARP threshold of 0.99 performed best for 128-founder populations, for both DGRP-derived and coalescent-derived simulations. Conversely, 32-founder populations performed best with a HARP threshold of 0.95. Data shown reports means across 400 replicates made up of 100 simulated individuals (4 autosomes each for coalescent simulations, 4 autosome arms each for DGRP simulations) per parameter combination.  Coalescent-derived populations described here wer1e simulated with Ne=1×106 and mutation rate μ=5×10-9.



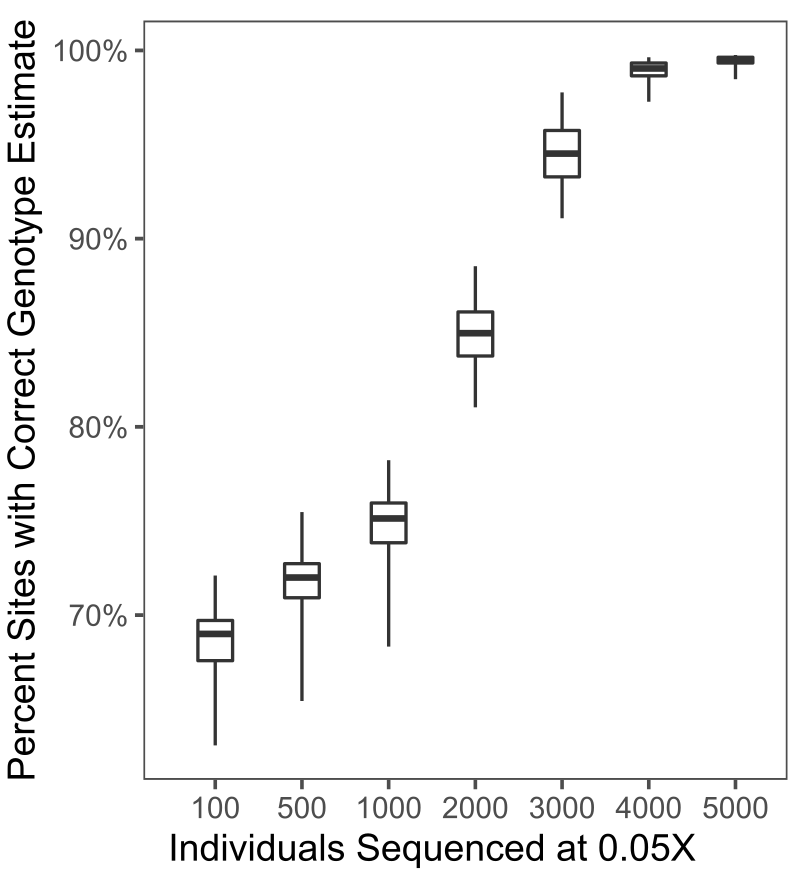
**Supplemental Figure S8**. Optimization curves for Most-Likely-Ancestor inclusion for additional parameter combinations. A: SCRM (coalescent) derived populations across different parameters; B: accuracy is similar for both 0.05X and 0.005X sequencing coverage in populations simulated with DGRP genomes.



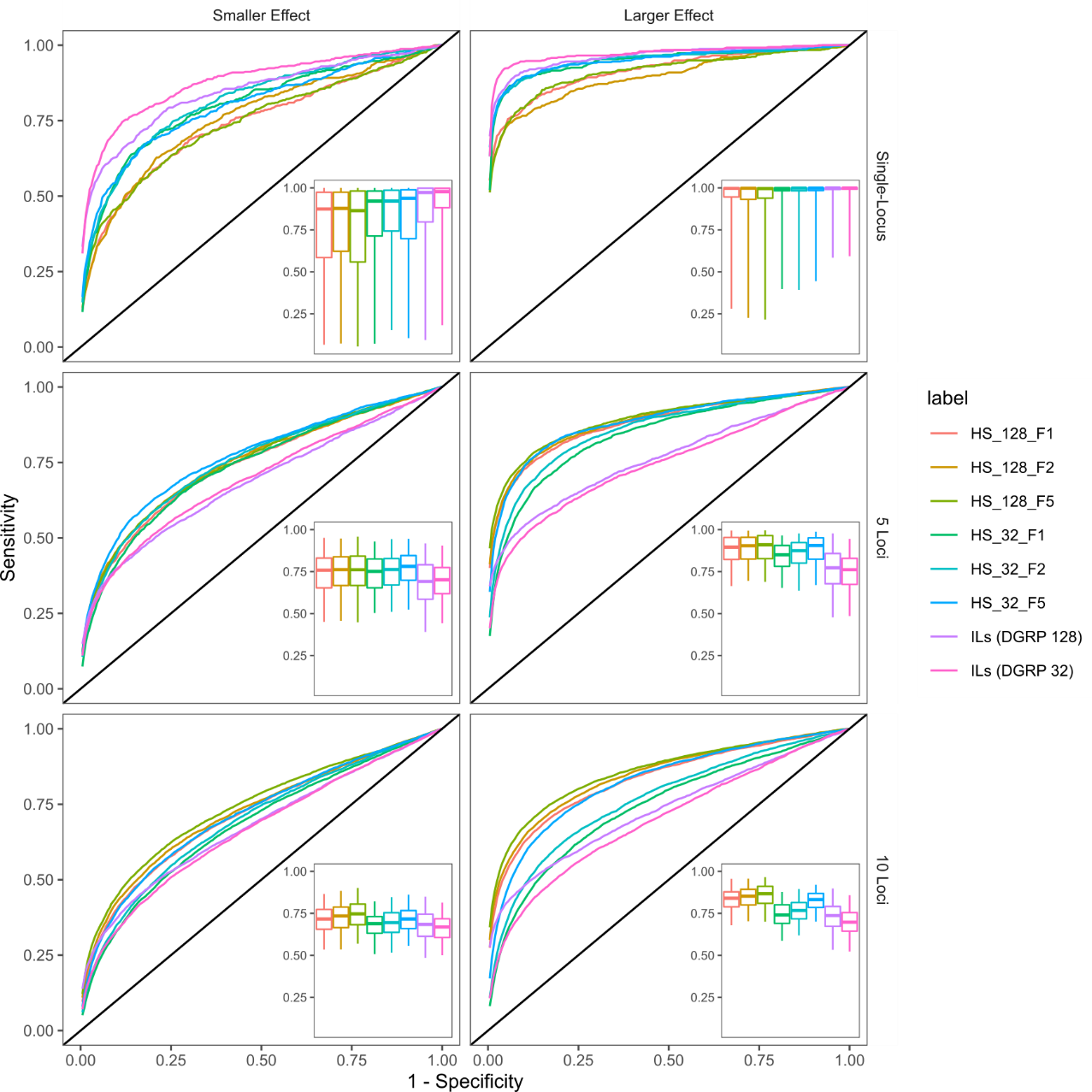
**Supplemental Figure S9.** Accuracy of genome reconstruction for simulated, DGRP-derived F5 hybrid swarm individuals. Reconstructions were performed for populations simulated as being founded by either 32 or 128 inbred lines for two levels of ultra-low sequencing coverage. Accuracy is represented on a logit scale, as most points occur above 90%. Reconstructed chromosomes estimated to have >10 recombination events are denoted by an ×, offset from the bulk of the distribution. Each parameter combination includes 400 reconstructed chromosomes (from 100 simulated individuals).



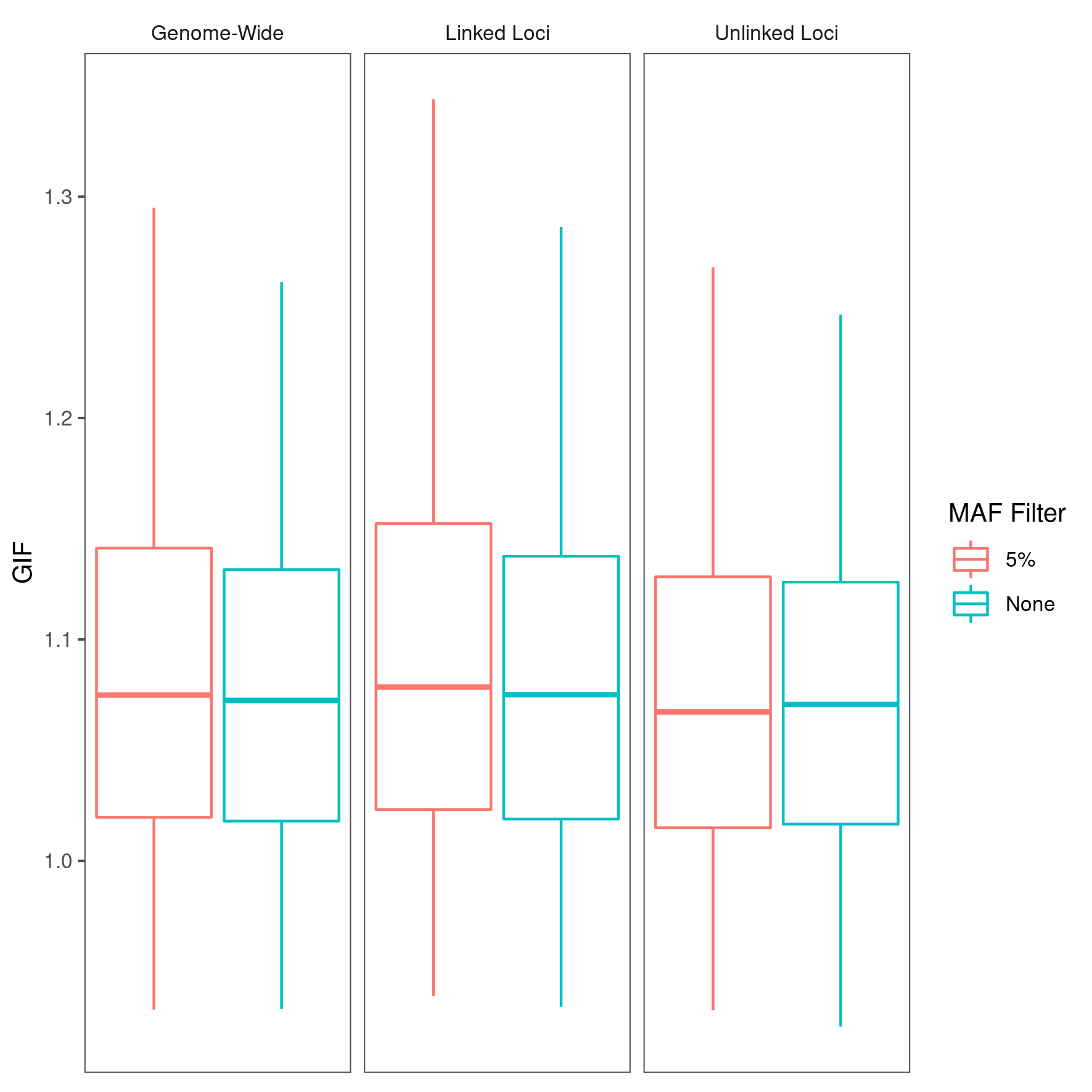
**Supplemental Figure S10.** Accuracy of genome reconstruction for simulated, coalescent-derived F5 hybrid swarm individuals. Reconstructions were performed for populations simulated as being founded by either 32 or 128 inbred lines for various effective population sizes (Ne) and mutation rates (μ). Accuracy is represented on a logit scale, as most points occur above 90%. Reconstructed chromosomes estimated to have >10 recombination events are denoted by an ×, offset from the bulk of the distributions. Each parameter combination includes 400 reconstructed chromosomes (from 100 simulated individuals).



**Supplemental Figure S11**. Genotyping accuracy of STITCH for simulated 32-founder F5 hybrid swarm populations sequenced at 0.05X coverage. Genotype accuracy improves with greater numbers of sequenced individuals, as STITCH infers missing genotypes from other haplotypes in the population. Boxes represent the median and interquartile range; whiskers extending to the lower and upper bounds of the 95% quantiles. See methods for parameters used.



**Supplemental Figure S12.** Receiver Operator Characteristic (ROC) curves for additional GWAS simulations.To generate single representative ROC curves across 500 GWAS simulations per population, we stepped through all specificity values and calculated the mean sensitivity. Inset boxplots display the median and interquartile range of Area Under the Curve (AUC) distributions, with whiskers spanning the middle 95% of data. **ILs**: inbred lines. **HS**: Hybrid Swarm populations founded by 32 or 128 lines.



**Supplemental Figure S13.** Genomic Inflation Factor with and without low-frequency filtering. Excluding alleles with a minor allele frequency below 5% does not reduce genomic inflation factor in inbred lines. Box depicts the middle 50% quantiles with whiskers extending to the middle 95% of data points. Traits were modeled as a single locus of large effect.