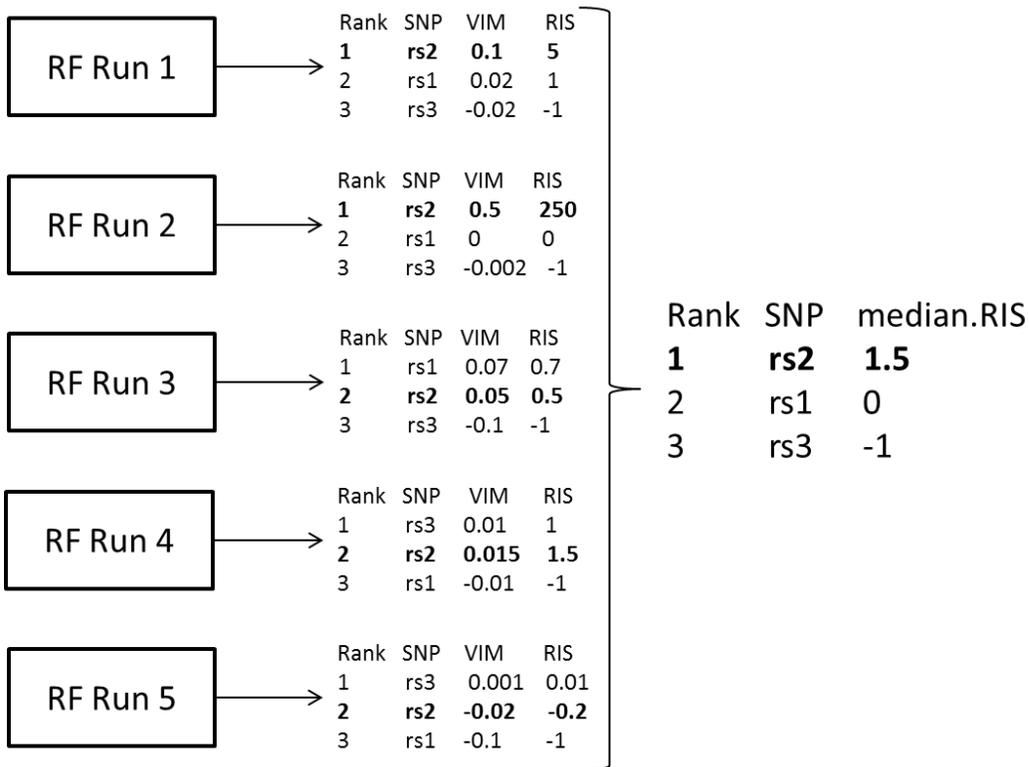
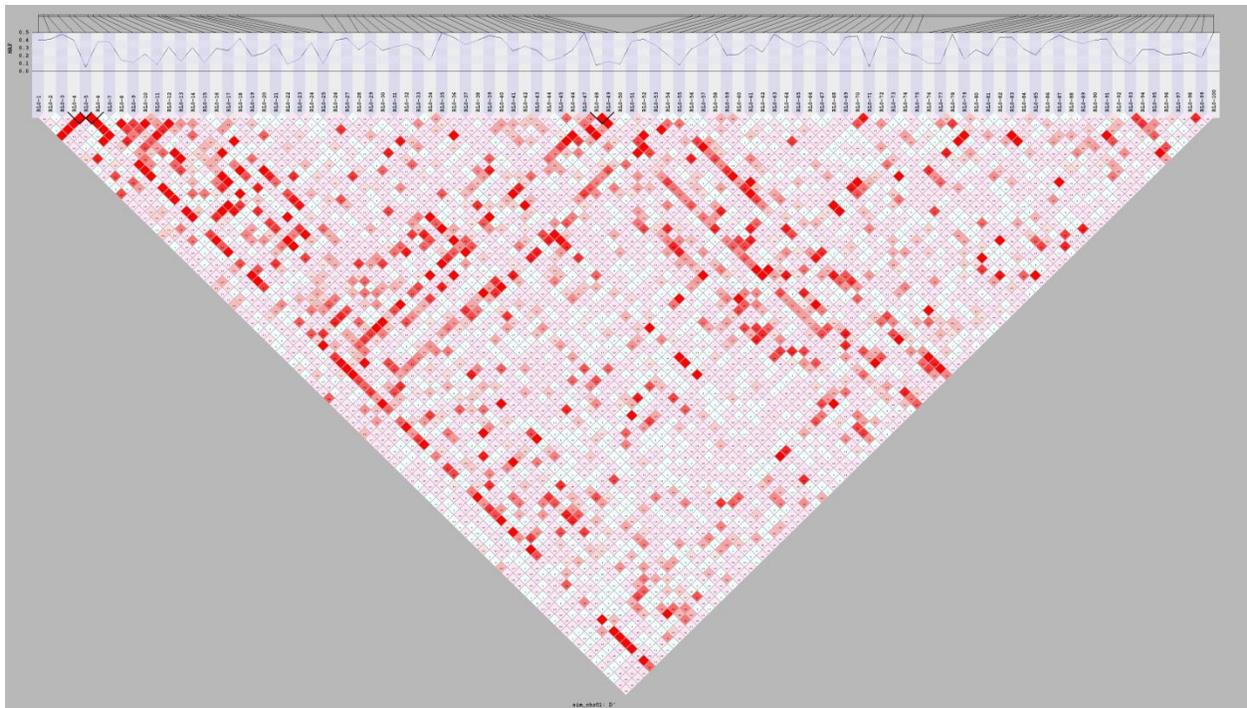


Supp. Table 1. Penetrance values for each genotype combination for the two-locus model. The value corresponds to the probability and individual will be assigned a case status with the given genotype combination. The minor allele (a/b) frequency for both SNPs is 0.4. The marginal effects for the single SNP for each model are all very close to zero.

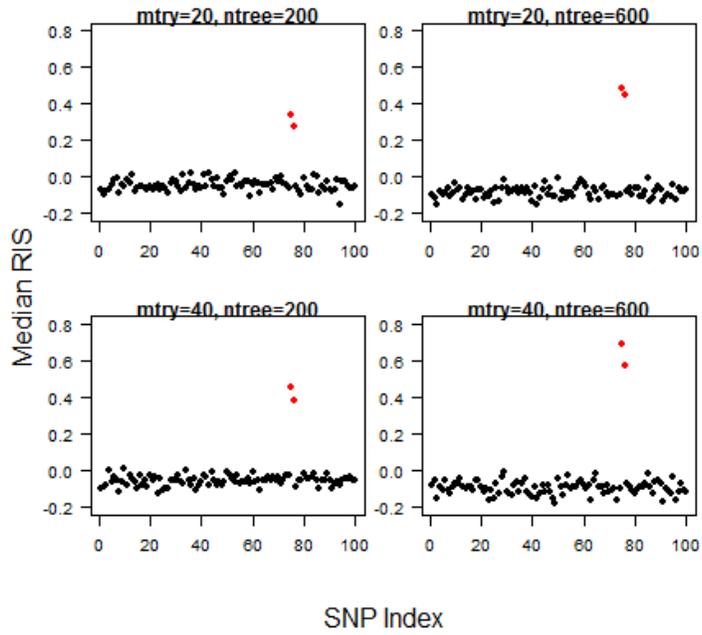
Genotype Comb.	100 SNPs / no LD	100 SNPs / LD	1000 SNPs / no LD	1000 SNPs / LD
AABB	0.305	0.387	0.329	0.352
AABb	0.382	0.259	0.318	0.299
AAbb	0.010	0.014	0.205	0.218
AaBB	0.285	0.262	0.379	0.355
AaBb	0.304	0.271	0.300	0.309
Aabb	0.290	0.240	0.142	0.153
aaBB	0.293	0.001	0.019	0.000
aaBb	0.073	0.249	0.278	0.263
aabb	0.959	0.853	0.999	0.971



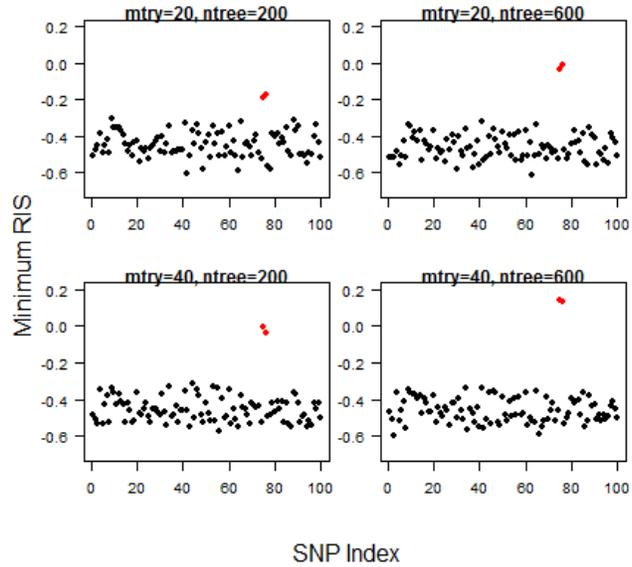
**Supp. Figure 1.** A hypothetical example to illustrate the three main components of r2VIM: (1) Estimation of VIM in RF (% change in mean squared error after variable permutation). (2) Threshold guidance using variance estimation by first calculating the relative importance score (RIS) for each run ( $RIS = VIM / \text{abs}[\text{min.VIM}]$ ). (3) Recurrent RF runs as a correction for random nature of the algorithm. We show the median RIS of the five runs as the “recurrency-corrected” RIS. In this example, the variable (rs2) is selected with a median RIS greater than 1 (i.e. greater than the variance estimate for that run).



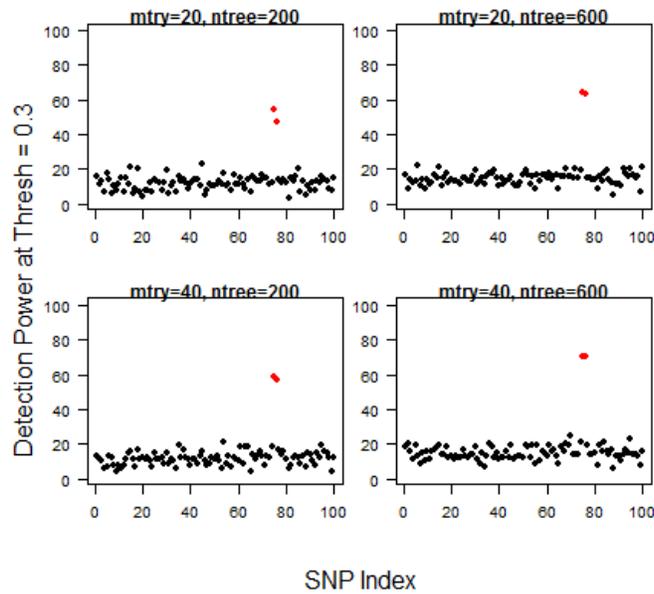
**Supp. Fig. 2.** This plot shows pairwise D-prime values for simulated LD patterns in the 100 SNP data in the pool of data from which individuals were selected to generate datasets. The 1000 SNP data had similar patterns of correlation. The functional SNPs (4 and 26) were selected to be virtually uncorrelated. The line at the top of the plot indicates minor allele frequency. This plot was created by the genomeSIMLA program.



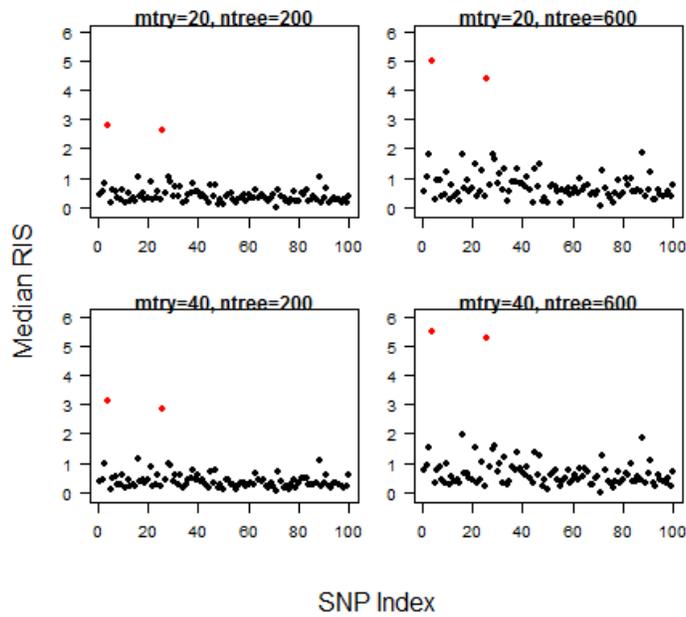
Supp. Fig. 3. Results for r2VIM analysis of datasets with 100 SNPs and no LD for all parameter settings. The plots show the median of the median RIS for each SNP across 100 dataset replicates. The functional SNPs 75 and 76 are shown in red.



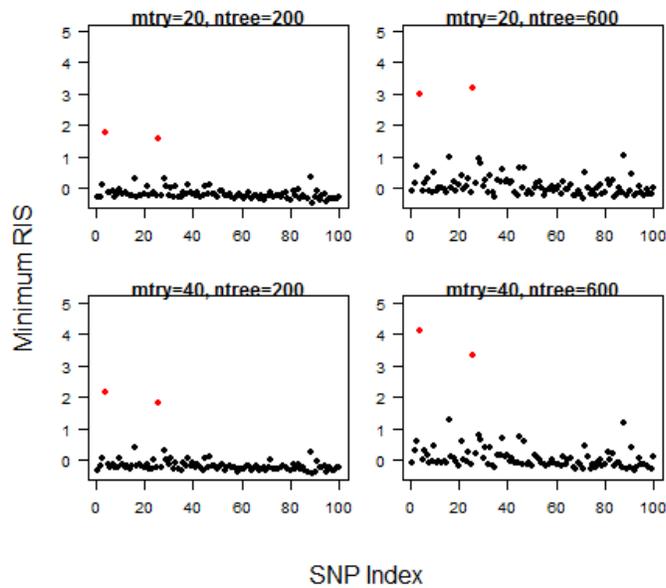
Supp. Fig. 4. Results for r2VIM analysis of datasets with 100 SNPs and no LD for all parameter settings. The plots show the median of the minimum RIS for each SNP across 100 dataset replicates. The functional SNPs 75 and 76 are shown in red.



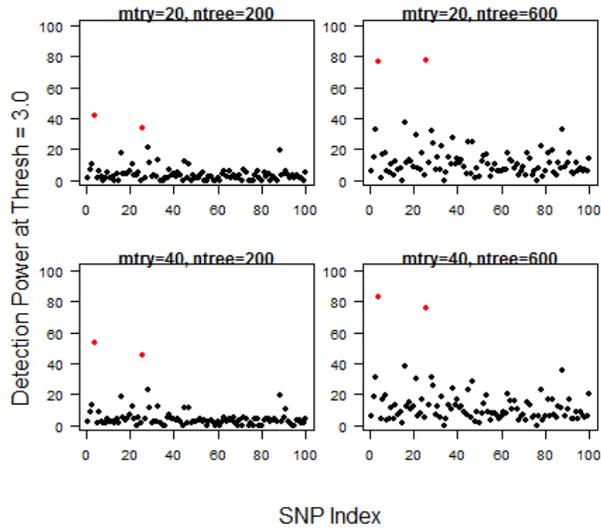
Supp. Fig. 5. Results for r2VIM analysis of datasets with 100 SNPs and no LD for all parameter settings. The plots show number of times out of 100 replicates that the median RIS was smaller than the specified significance threshold for each SNP. The functional SNPs 75 and 76 are shown in red.



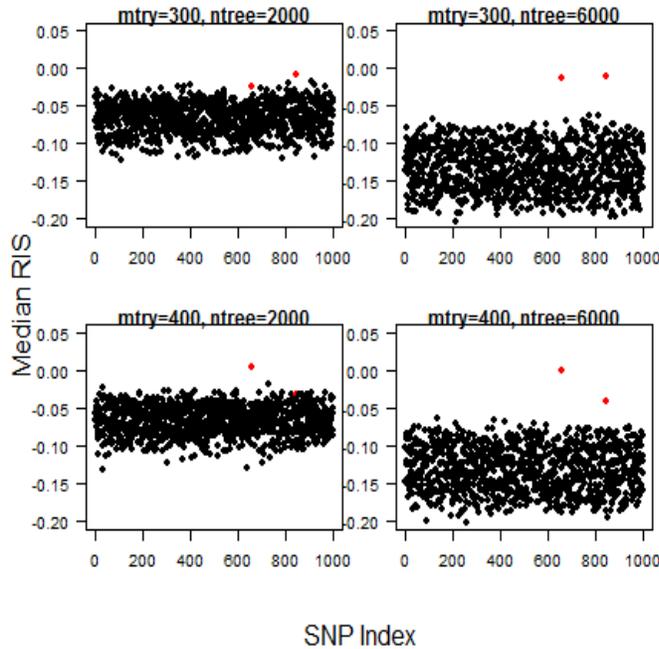
Supp. Fig. 6. Results for r2VIM analysis of datasets with 100 SNPs and LD for all parameter settings. The plots show the median of the median RIS for each SNP across 100 dataset replicates. The functional SNPs 4 and 26 are shown in red.



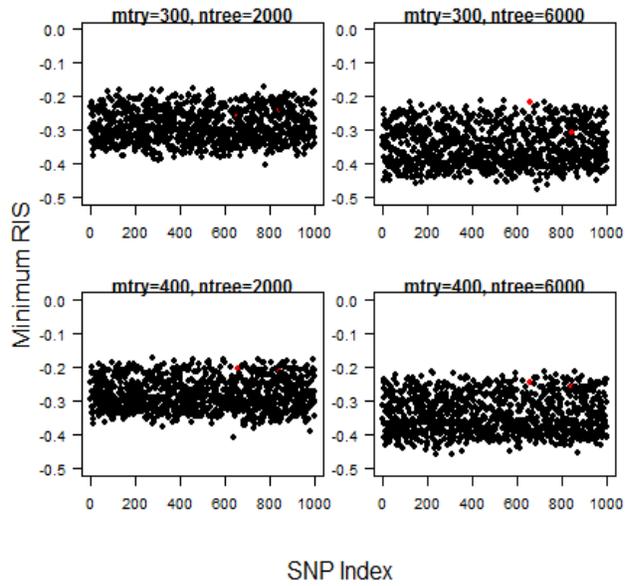
Supp. Fig. 7. Results for r2VIM analysis of datasets with 100 SNPs and LD for all parameter settings. The plots show the median of the minimum RIS for each SNP across 100 dataset replicates. The functional SNPs 4 and 26 are shown in red.



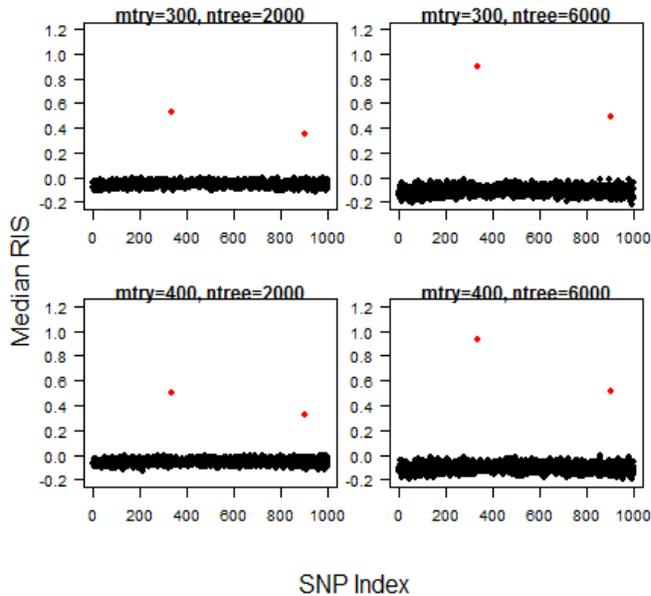
Supp. Fig. 8. Results for r2VIM analysis of datasets with 100 SNPs and LD for all parameter settings. The plots show number of times out of 100 replicates that the median RIS was smaller than the specified significance threshold for each SNP. The functional SNPs 4 and 26



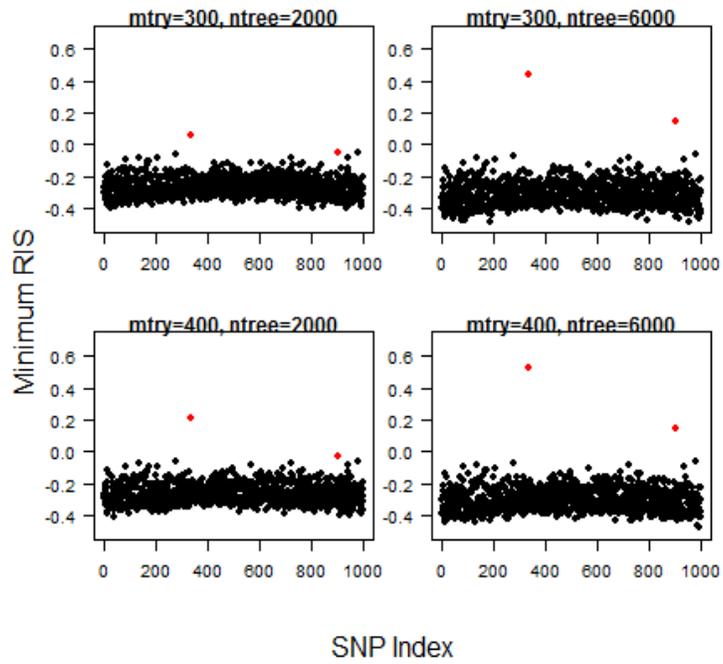
Supp. Fig. 9. Results for r2VIM analysis of datasets with 1000 SNPs and LD for all parameter settings. The plots show the median of the median RIS for each SNP across 100 dataset replicates. The functional SNPs 657 and 844 are shown in red.



Supp. Fig. 10. Results for r2VIM analysis of datasets with 1000 SNPs and LD for all parameter settings. The plots show the median of the minimum RIS for each SNP across 100 dataset replicates. The functional SNPs 657 and 844 are shown in red.



Supp. Fig. 11. Results for r2VIM analysis of datasets with 1000 SNPs and LD for all parameter settings. The plots show the median of the median RIS for each SNP across 100 dataset replicates. The functional SNPs 336 and 903 are shown in red.



Supp. Fig. 12. Results for r2VIM analysis of datasets with 1000 SNPs and LD for all parameter settings. The plots show the median of the minimum RIS for each SNP across 100 dataset replicates. The functional SNPs 336 and 903 are shown in red.