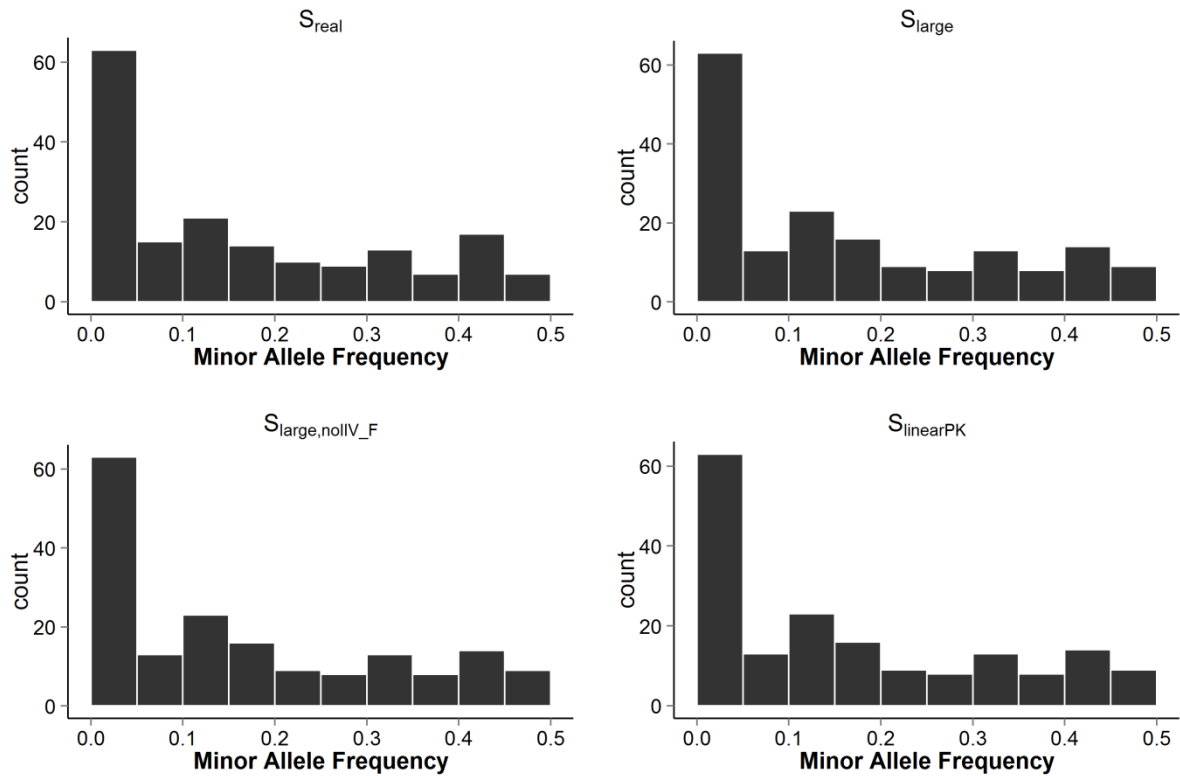


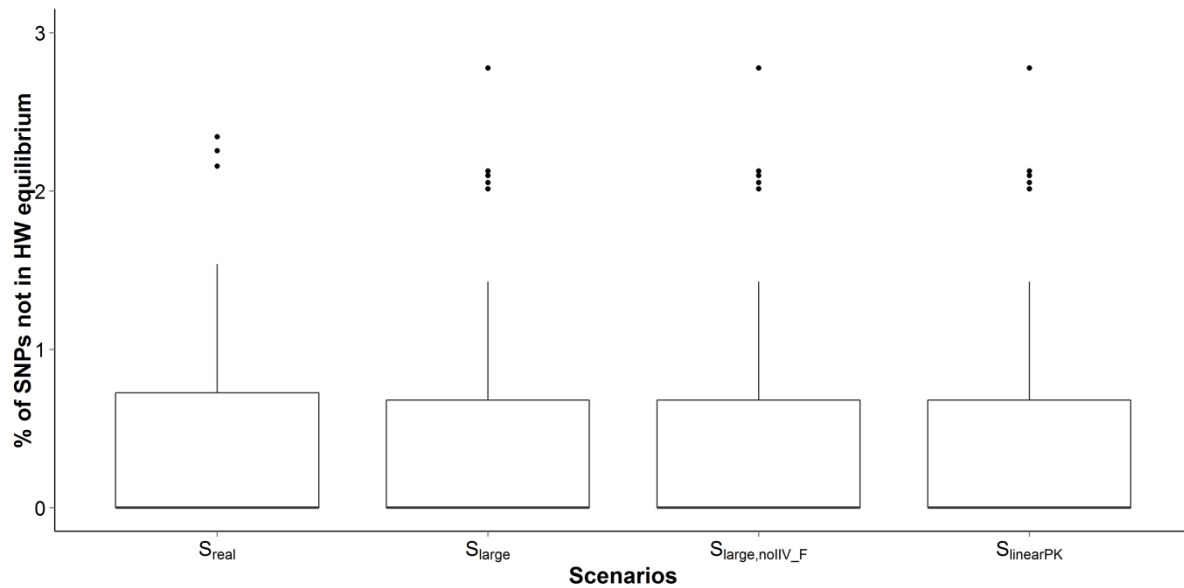
## Simulated polymorphisms information

### Minor Allele Frequencies



**Fig. Minor Allele Frequency (MAF):** the mean of frequencies of each of the 176 SNPs was computed on the 200 data sets. The distribution of mean MAF is represented for each scenario.

## Hardy-Weinberg Equilibrium



**Fig. Hardy-Weinberg (HW) equilibrium:** the percentage of SNPs not in HW equilibrium was computed in each of the 200 data sets, for each scenario. Departure from HW equilibrium was tested using a  $\chi^2$  test with a Bonferroni adjusted p-value threshold ( $0.05/\text{number of polymorphic SNPs}$ ).

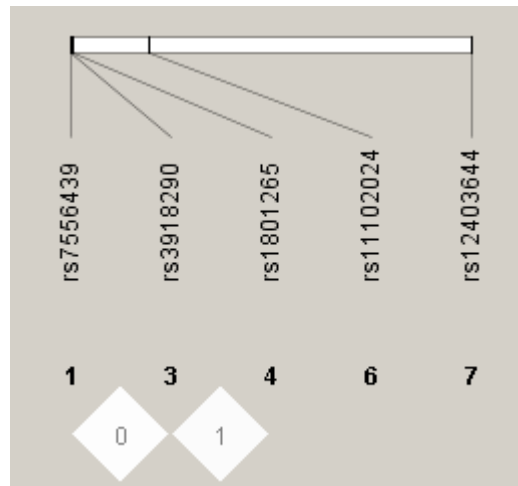
## LD plots

### Summary

**Table.** Summary on linkage disequilibrium (LD) in an example data set (from scenario  $S_{large}$ )

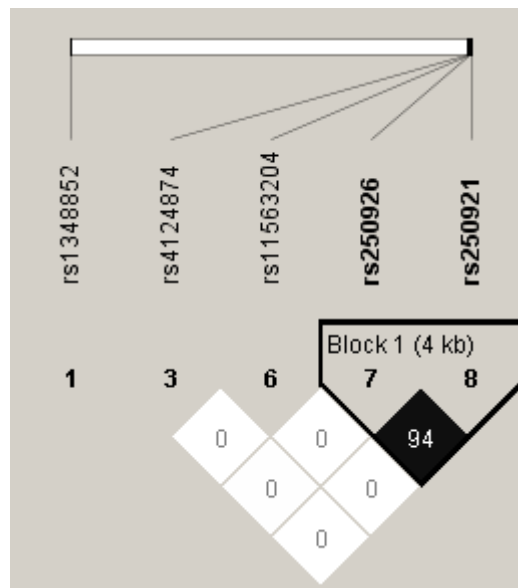
Chromosome	Blocks	Number of SNPs	Number of observed haplotypes
2	1	2	2 (0.807/0.184)
3	2	2	2 (0.667/0.333)
		3	2 (0.534/0.466)
6	3	2	3 (0.711/0.162/0.124)
		2	4 (0.690/0.269/0.030/0.011)
		4	4 (0.742/0.111/0.037/0.108)
7	2	2	3 (0.466/0.456/0.078)
		2	3 (0.454/0.402/0.143)
8	2	5	4 (0.427/0.336/0.167/0.060)
		2	3 (0.363/0.359/0.277)
10	4	4	3 (0.495/0.376/0.128)
		8	6 (0.229/0.310/0.173/0.142/0.090/0.051)
		2	3 (0.094/0.892/0.010)
		2	3 (0.406/0.522/0.072)
11	1	2	3 (0.599/0.303/0.098)
12	1	2	3 (0.465/0.374/0.161)
15	1	2	3 (0.518/0.406/0.076)
19	2	4	4 (0.454/0.306/0.229/0.010)
		3	4 (0.343/0.207/0.338/0.108)

### Chromosome 1



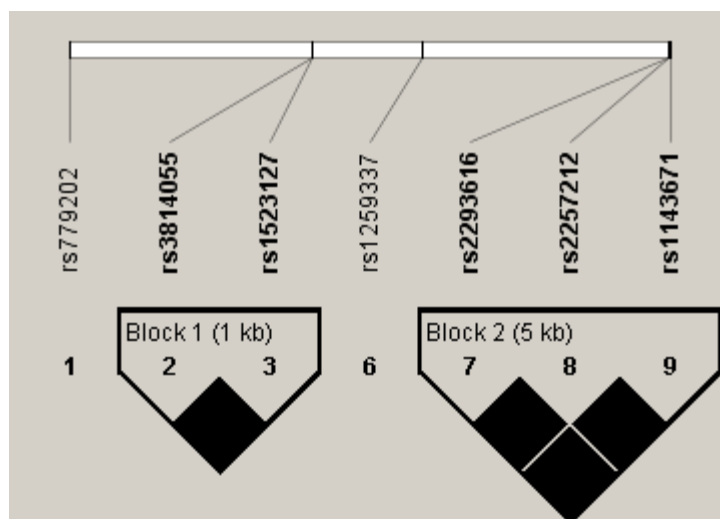
**Fig. LD plot:** Distribution of LD values in chromosome 1 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

### Chromosome 2



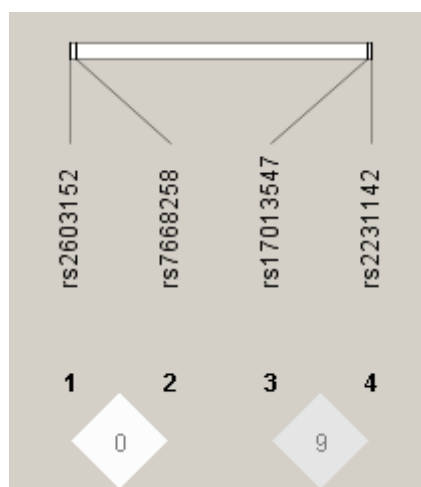
**Fig. LD plot:** Distribution of LD values in chromosome 2 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

### Chromosome 3



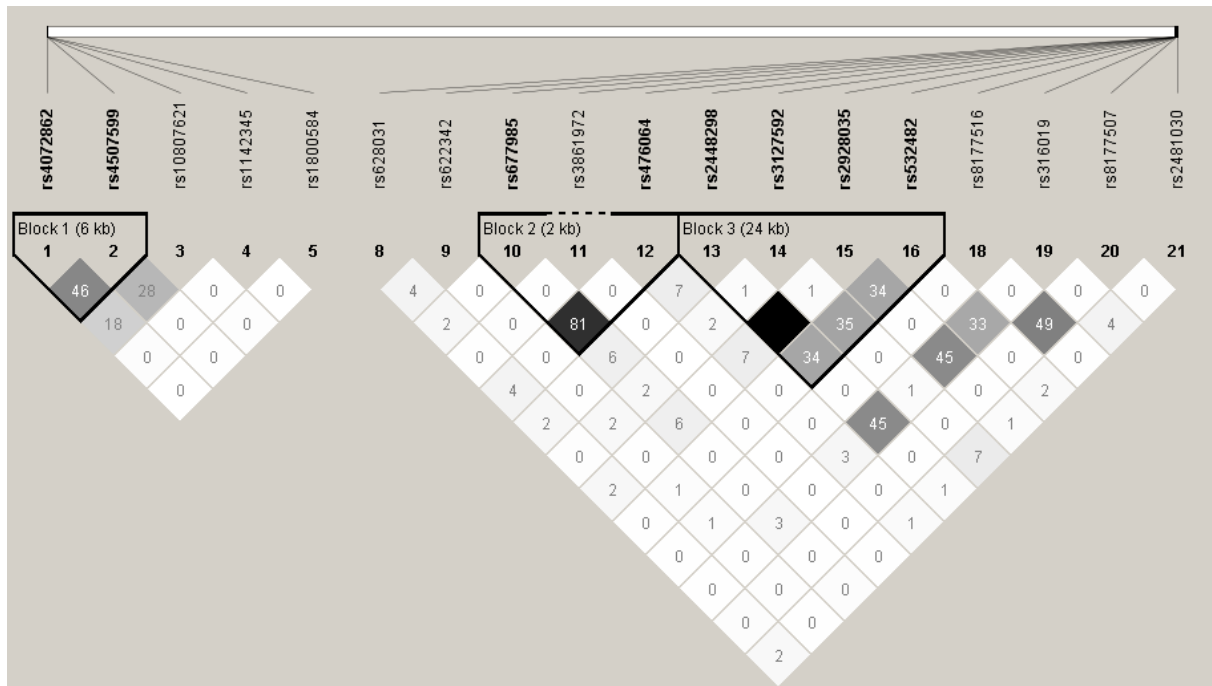
**Fig. LD plot:** Distribution of LD values in chromosome 3 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

### Chromosome 4



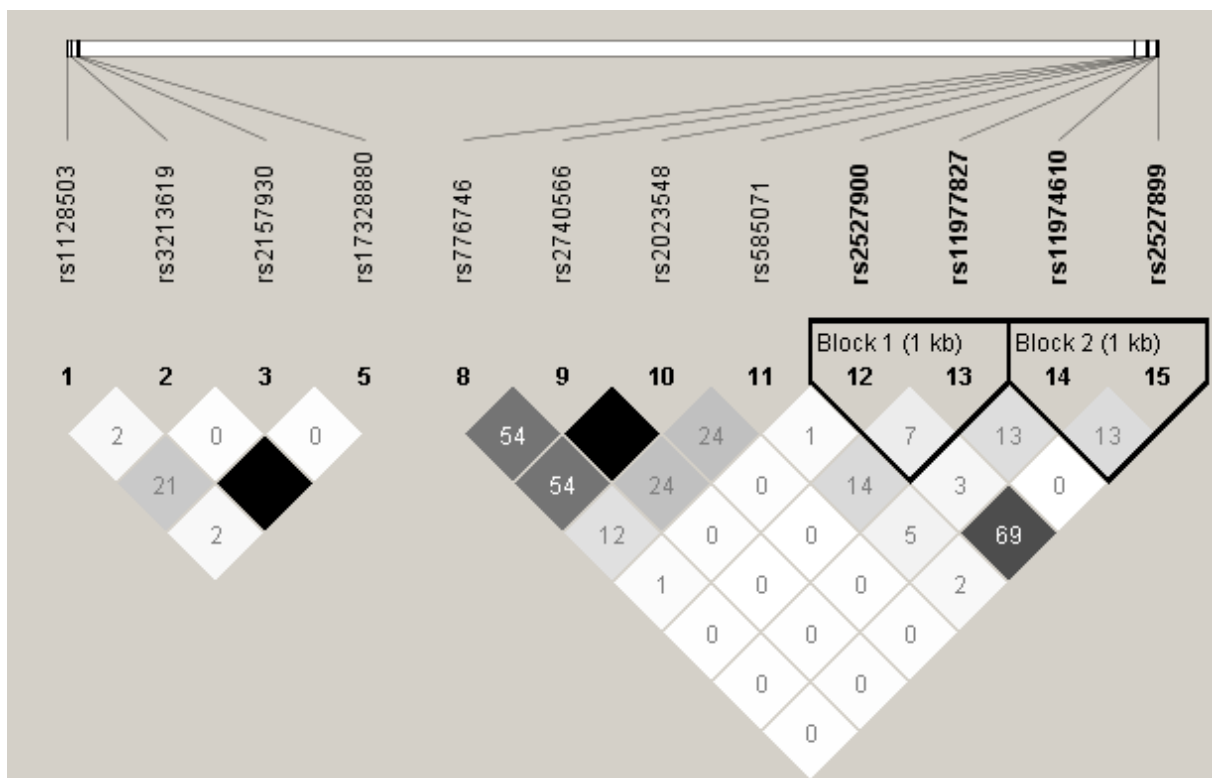
**Fig. LD plot:** Distribution of LD values in chromosome 4 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

## Chromosome 6



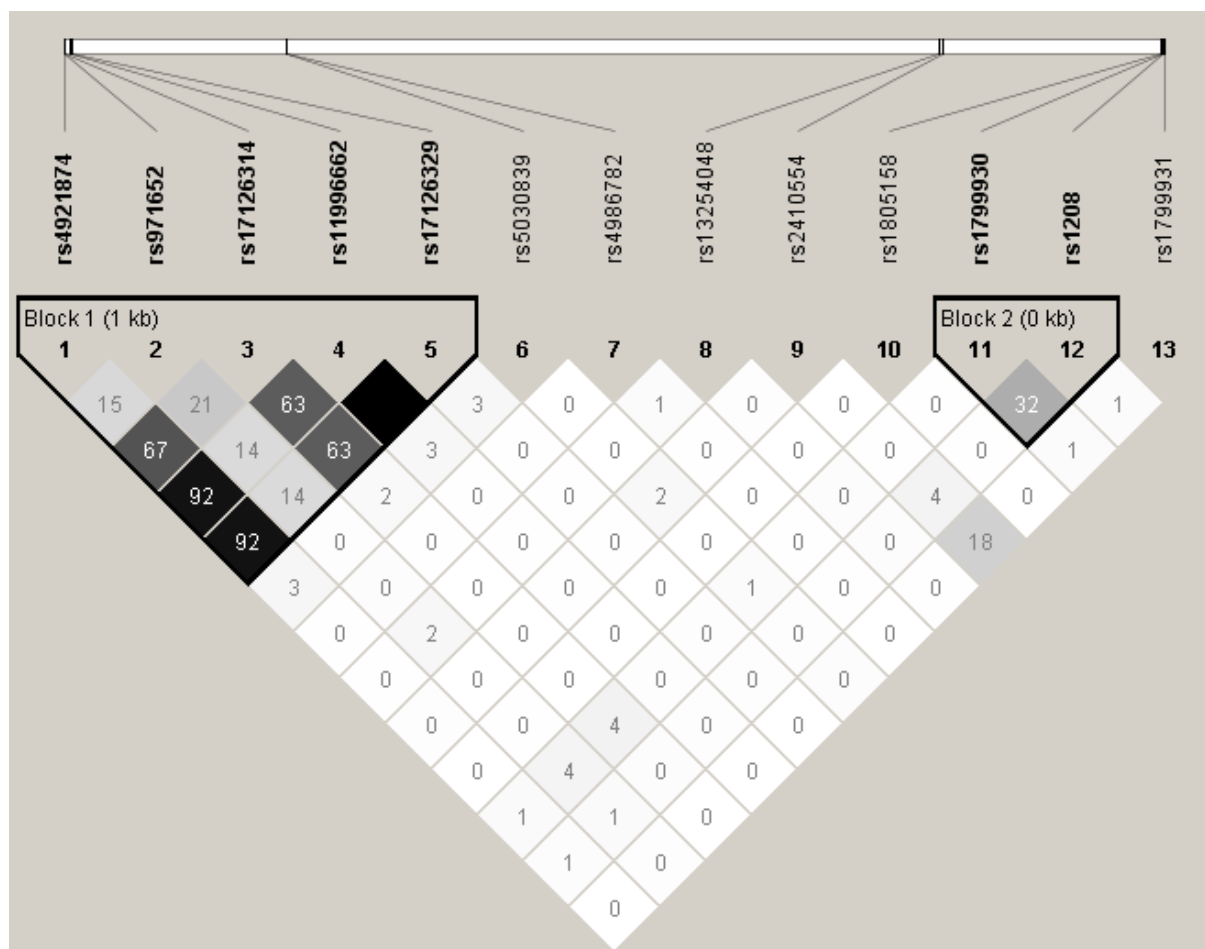
**Fig. LD plot:** Distribution of LD values in chromosome 6 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

## Chromosome 7



**Fig. LD plot:** Distribution of LD values in chromosome 7 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

### Chromosome 8



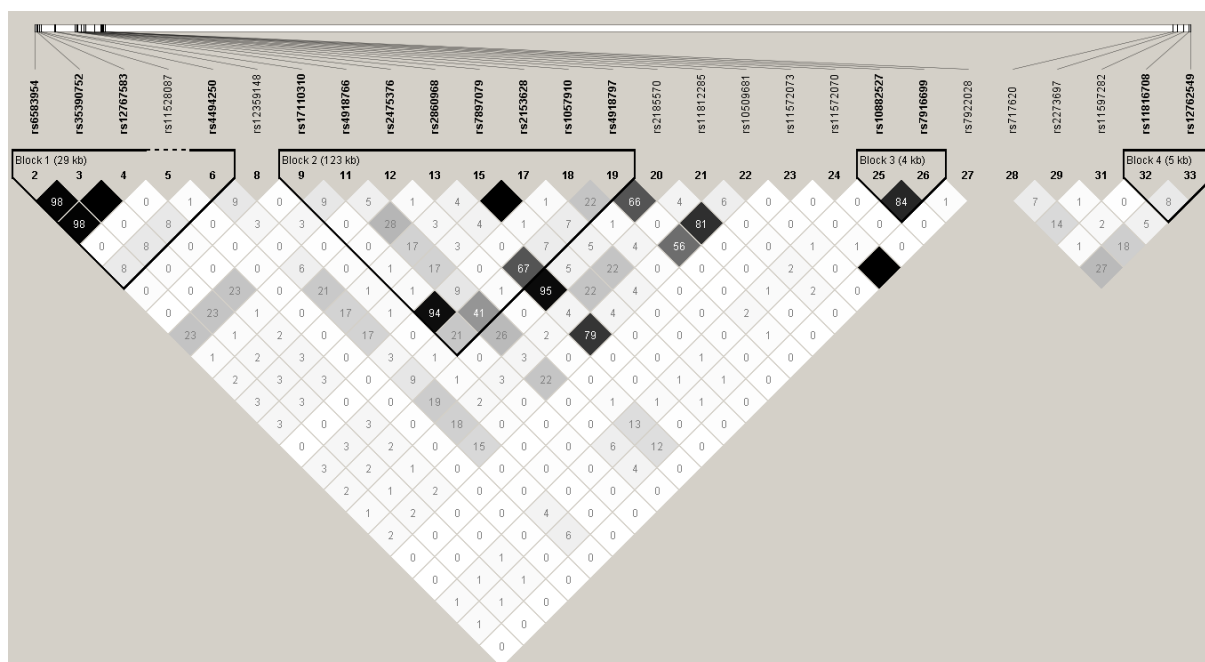
**Fig. LD plot:** Distribution of LD values in chromosome 8 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

## Chromosome 9



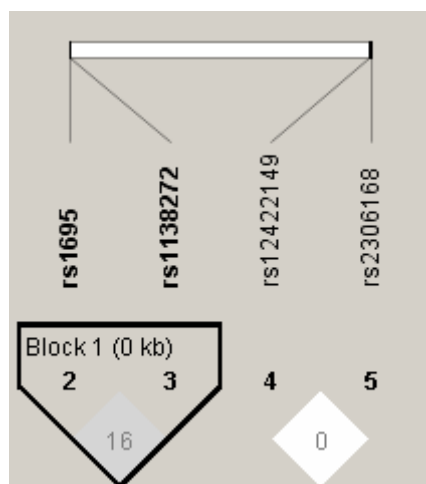
**Fig. LD plot:** Distribution of LD values in chromosome 9 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

## Chromosome 10



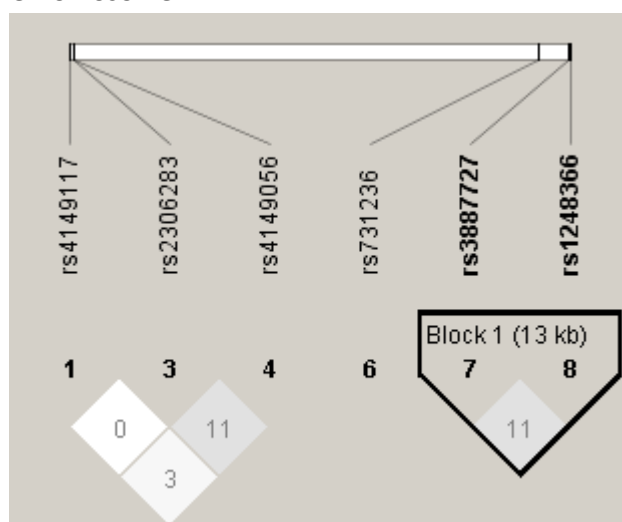
**Fig. LD plot:** Distribution of LD values in chromosome 10 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

## Chromosome 11



**Fig. LD plot:** Distribution of LD values in chromosome 11 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

## Chromosome 12



**Fig. LD plot:** Distribution of LD values in chromosome 12 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD



rs1048943 rs762551 rs3764244 rs7164654 rs4886841 rs7183961 rs17380507 rs12899543

1 3 4 6 7 8 9 10

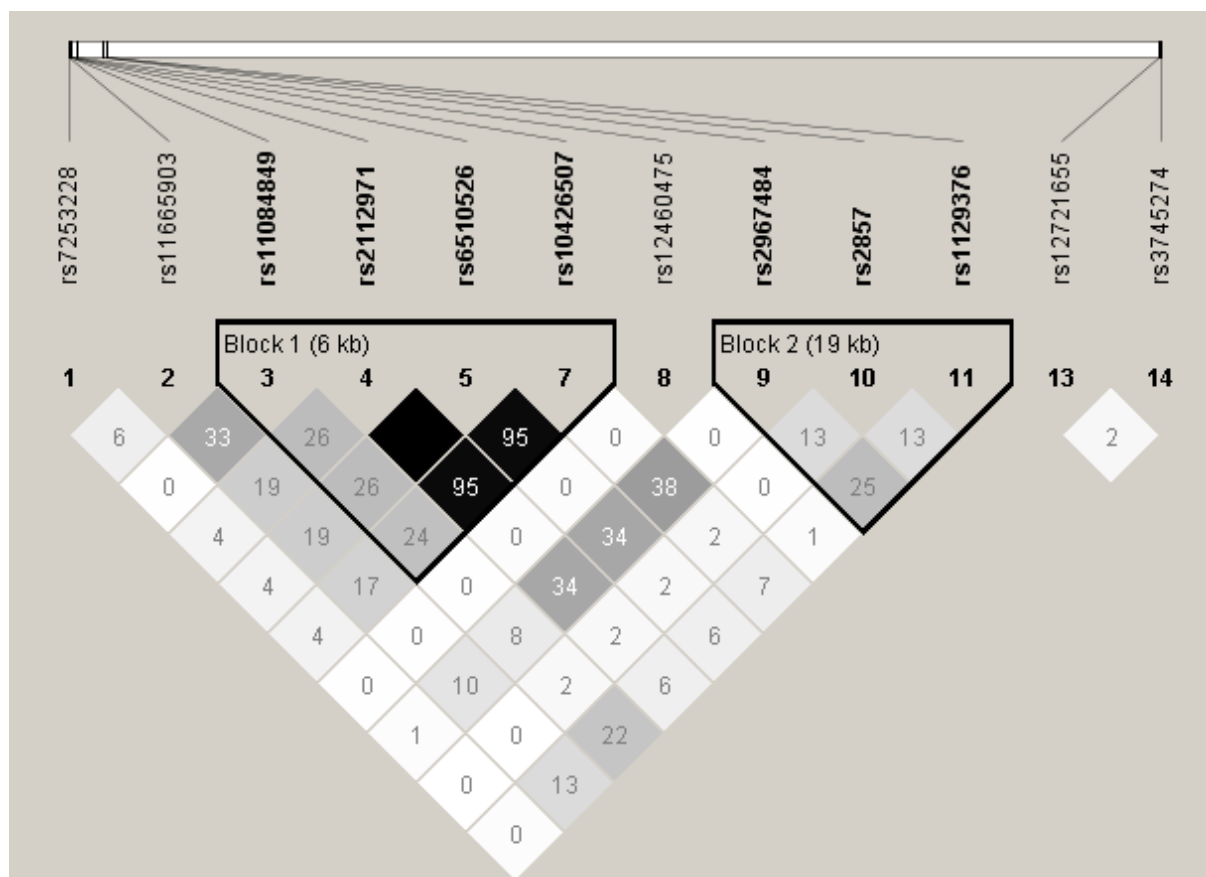
Block 1 (16 kb)

1								
	0							
		0						
			0					
				0				
					0			
						0		
							0	
								0

## Chromosome 16

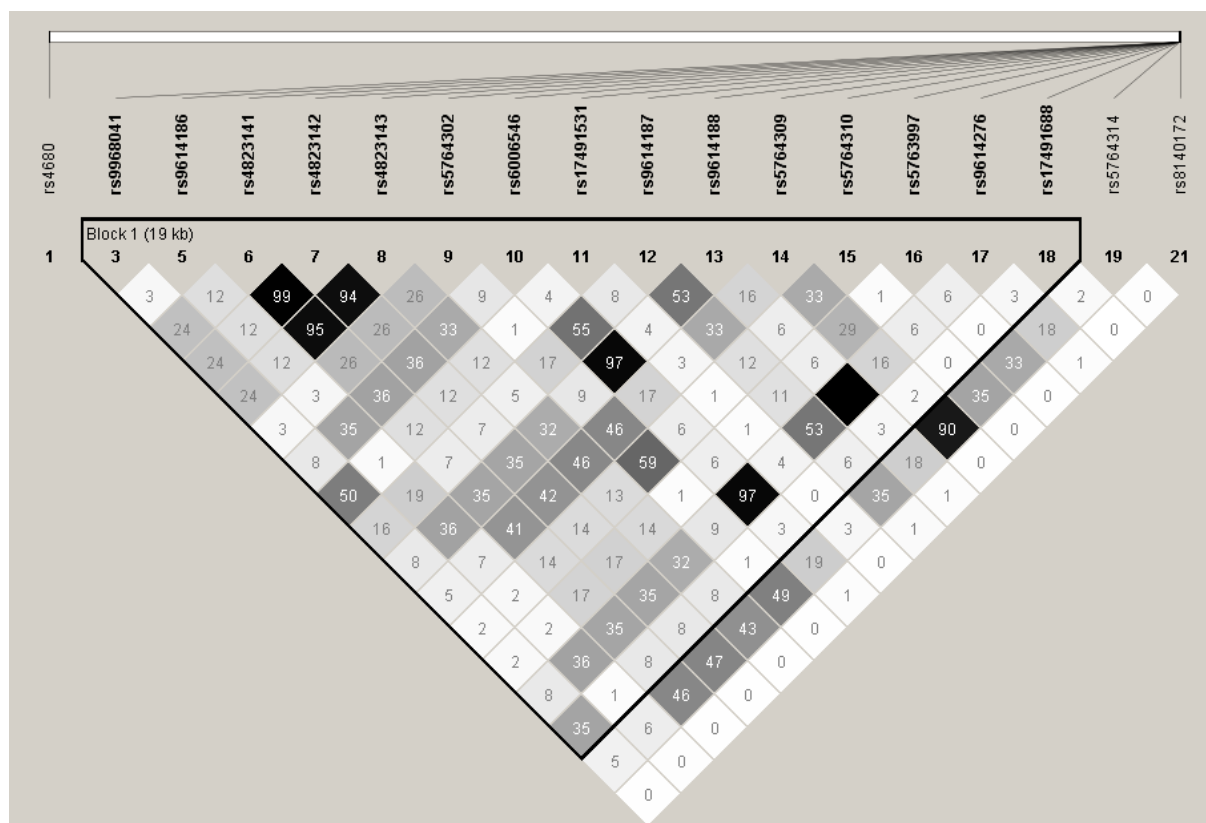
Neither of the two SNPs simulated for this chromosome were polymorphic in the example data set

## Chromosome 19



**Fig. LD plot:** Distribution of LD values in chromosome 19 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD

## Chromosome 22



**Fig. LD plot:** Distribution of LD values in chromosome 22 for an example data set (from scenario  $S_{large}$ ). Darker shades indicate higher levels of LD (as measured by  $R^2$ ) and the lighter shades indicate lower levels of LD