

Genotype data					Phenotype data					Observed QTLs					Counts
	M_1	M_2	\dots	M_k		P_1	P_2	\dots	P_T		P_1	P_2	\dots	P_T	
S_1	AA	AB	\dots	AA	S_1	1.5	3.6	\dots	2.8	L_1		•	\dots		N_1
S_2	AA	AA	\dots	AB	S_2	4.1	1.9	\dots	2.2	L_2	•		\dots		N_2
S_3	AB	AB	\dots	AB	S_3	3.2	0.6	\dots	2.8	L_3			\dots	•	N_3
S_4	AB	AA	\dots	AA	S_4	3.3	4.8	\dots	4.2	L_4	•	•	\dots	•	N_4
S_5	AA	AA	\dots	AA	S_5	4.1	2.9	\dots	2.6	L_5		•	\dots		N_5
S_6	AB	AA	\dots	AB	S_6	0.7	1.4	\dots	2.5	L_6			\dots		N_6
S_7	AB	AB	\dots	AA	S_7	2.2	3.4	\dots	1.7	L_7		•	\dots		N_7
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
S_s	AA	AB	\dots	AA	S_s	2.3	2.6	\dots	2.9	L_l			\dots	•	N_l

Figure S1 Hotspot analysis in a genetical genomics study. The data is composed by genotypes and phenotypes on s subjects, S_1, \dots, S_s , from a segregating population. The genotype data is composed by the genotypes of k markers, M_1, \dots, M_k . The phenotype data is composed by measurements on T quantitative phenotypes, P_1, \dots, P_T . The output of the analysis is a QTL matrix, where rows represent l genomic positions, L_1, \dots, L_l , and columns the phenotypes. A significant QTL is represented by a bullet, for example, phenotype P_1 maps to QTLs located at the L_2 and L_4 genomic positions. For each genomic position, L_1, \dots, L_l , we count the number of significant QTLs, N_1, \dots, N_l . We say we detected a significant hotspot at a genomic location, L_j , when the respective count, N_j , is higher than what is expected by chance at a pre-determined genome wide error rate.

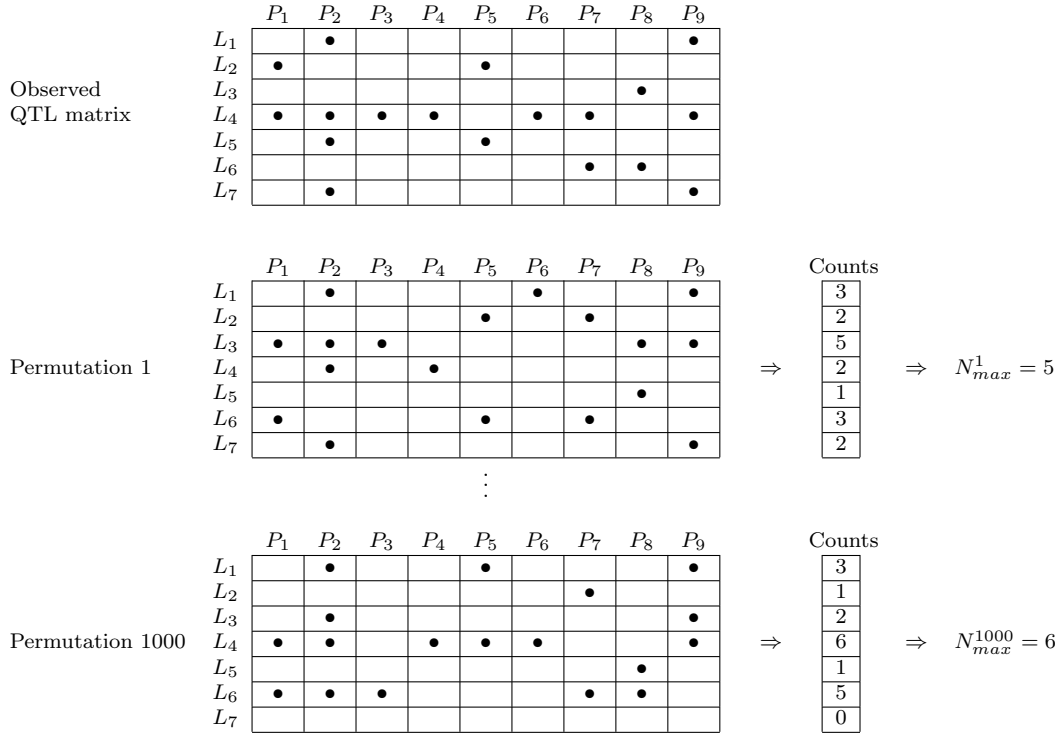


Figure S2 Permutation scheme adopted by West et al. (2007) and Wu et al. (2008). In this permutation scheme we take the QTL matrix and, for each fixed phenotype (column in the QTL matrix), we permute the QTL locations (the row cells at each fixed column). This figure depicts the result of two permutations of the observed QTL matrix. The permutation null distribution of hotspot sizes is derived as follows. For each one of the, say 1000, permutations we: (i) permute the genomic positions of the QTLs for each one of the phenotypes separately; (ii) for each genomic location we record the number of QTLs; (iii) record the maximum count N_{max}^{per} . The permutation null distribution (for the λ threshold used to derive the observed QTL matrix) is then given by the distribution of the 1,000 N_{max}^{per} values.

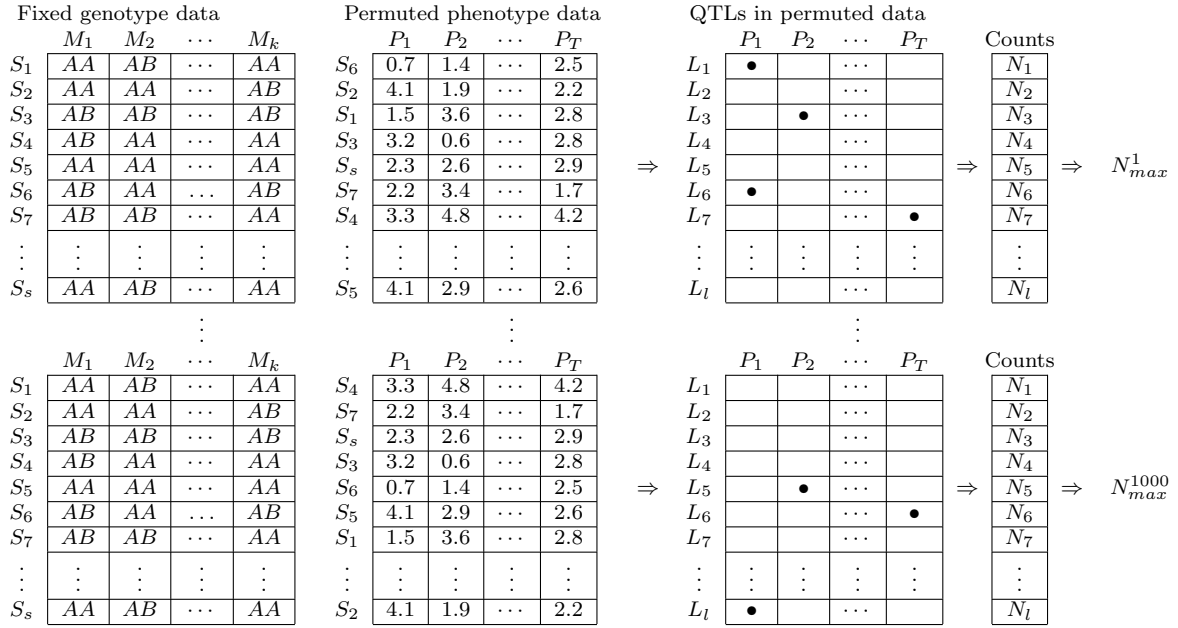


Figure S3 A permutation scheme that preserves the correlation among the phenotypes. Breitling et al. (2008) proposed a permutation scheme where the rows of the phenotype data matrix are permuted, while the genotype data matrix is kept intact. The idea is to break the connection between the genotype and phenotype data, but to preserve the correlation structure among the phenotypes. The permutation null distribution of hotspot sizes is derived as follows. For each one of the, say 1000, permutations we: (i) permute the rows of the phenotype data matrix, while keeping the genotype data intact (note the different row orderings of the permuted phenotype data matrices in relation to original phenotype matrix in Figure 1); (ii) perform mapping analysis of the T phenotypes, using a predetermined LOD threshold, λ , to determine a new QTL matrix (note that all QTLs detected with the permuted data are false positives); (iii) for each genomic location L_1, \dots, L_l we record the number of QTLs, N_1, \dots, N_l ; (iv) we record the maximum count $N_{max}^{per} = \max \{N_1, \dots, N_l\}$. The permutation null distribution for the chosen λ threshold is then given by the distribution of the 1,000 N_{max}^{per} values.

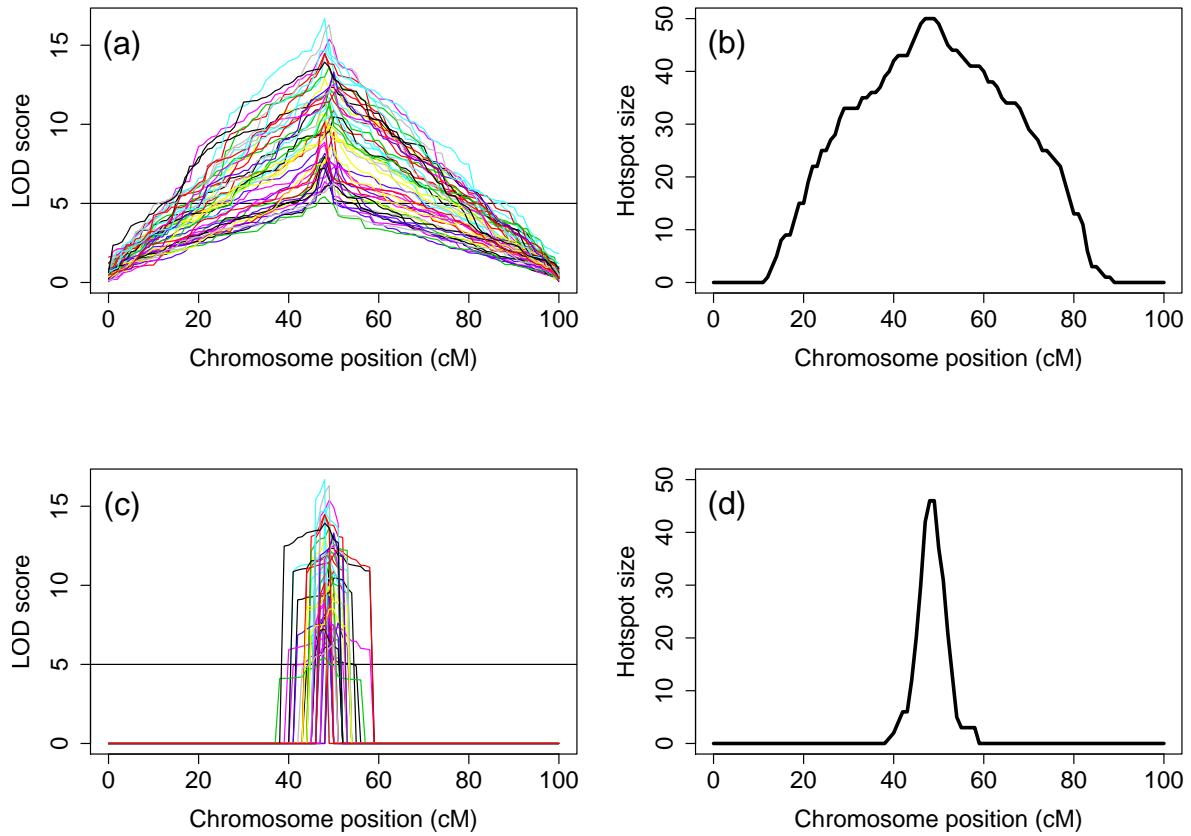


Figure S4 Using LOD support intervals to reduce the spread of QTL hotspots. Panel (a) shows the LOD profile curves of 50 traits showing peaks around 50cM. Panel (b) shows how many traits have LOD score above the LOD threshold 5 (horizontal line) for each genomic location. Panel (c) shows the processed LOD curves where, for each trait, we computed the 1.5 LOD support interval and set the LOD scores outside the interval to zero. Panel (d) shows the counts based on the processed LOD profiles. Note how the spread of the hotspot location is drastically reduced from panel (b) to panel (d).

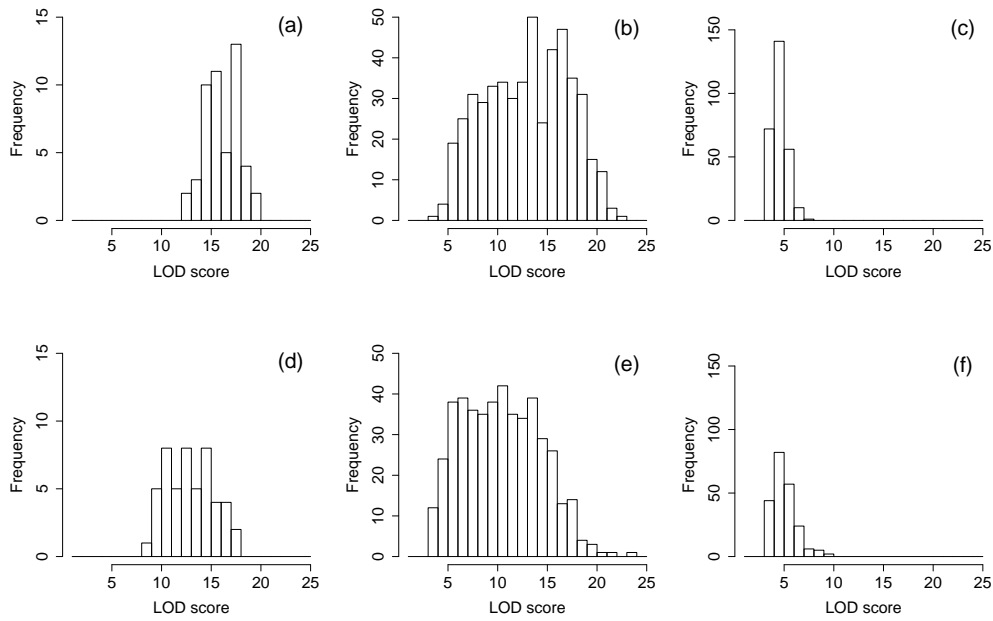


Figure S5 Hotspot LOD score distributions for simulated examples 1 and 2. Panels (a) and (d) show the LOD score distribution for the hotspot on chromosome 3 for simulated example 1 and 2, respectively. Panels (b) and (e) show the LOD score distribution for the hotspot on chromosome 7 for simulated example 1 and 2, respectively. Panels (c) and (f) show the LOD score distribution for the hotspot on chromosome 15 for simulated example 1 and 2, respectively. The histograms show the distribution of the LOD scores of the traits composing the hotspot at the hotspot peak location.

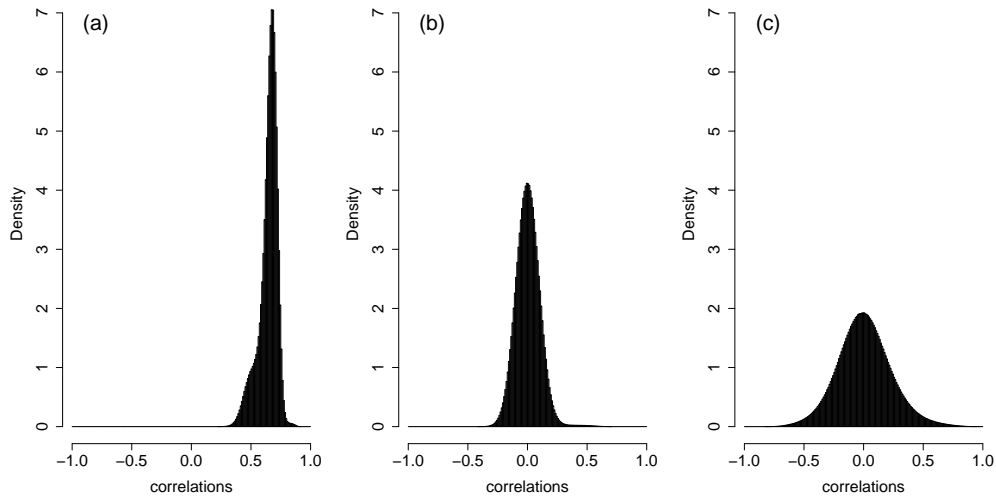


Figure S6 Pairwise correlations among phenotypes. Panels (a) and (b) show the results for simulated examples 1 and 2, respectively. Panel (c) shows the distribution of the pairwise correlations for the yeast data-set.

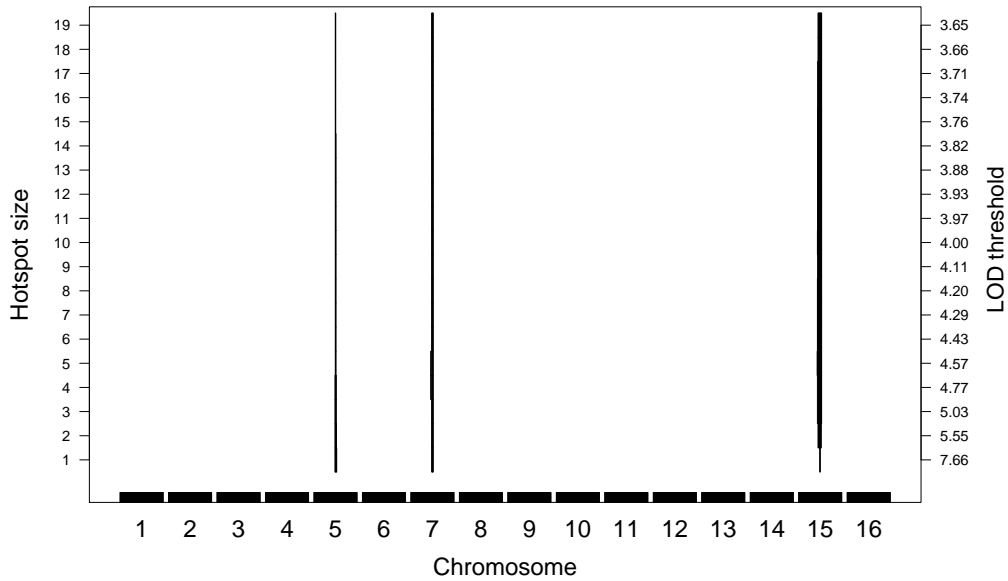


Figure S7 Hotspot size significance profile derived with the *NL*-method for simulated example 2. For each genomic location (i.e., x-axis position) this figure shows the hotspot sizes at which the hotspot was significant, that is, at which the hotspot locus had more traits mapping to it with a LOD score higher than the threshold on the right, than expected by chance. The scale in the left shows the range of spurious hotspot sizes investigated by our approach. The scale in the right shows the respective LOD thresholds associated with the spurious hotspot sizes in the left. The range is from 7.66, the conservative empirical LOD threshold associated with a spurious “hotspot of size 1”, to 3.65, the single trait empirical threshold, associated with a spurious hotspot of size 19. All permutation thresholds were computed targeting $\text{GWER} \leq 0.05$, for $n = 1, \dots, 19$. Note that this figure does not show the actual size of the hotspots. For instance, at the single trait LOD threshold of 3.65 this plot indicates that more traits mapped to the middle regions of chromosomes 5, 7 and 15 than what would be expected by chance (19 traits), but do not show the actual hotspot sizes (50, 464 and 220, respectively, as depicted in Figure 4a).