

## File S2

### Simulation studies

Here we provide further details on the simulation studies presented in the main text.

## File S3

### Pilot simulation study

We conducted a total of 10 simulation studies, generating data from the five models described in Figure 2 in the main text using sample sizes 112 and 1,000 (the choice 112 was motivated by the sample size in our real data example). For each model, we simulated 1,000 backcrosses composed with 3 chromosomes of length 100cM containing 101 unequally spaced markers per chromosome. For each one of the simulated backcrosses, the additive and dominance genetic effects were sampled, respectively, from the  $U[-0.75, 0.75]$  and  $U[0, 0.75]$  distributions, where  $U[a, b]$  represents the uniform distribution on the interval  $[a, b]$ . Residual error rates were sampled from  $U[0.5, 1.5]$ , and the phenotype to phenotype regression coefficients in Figures 2 A, B and C were sampled from  $U[-1, 1]$ . The hidden-variable to phenotype regression coefficients on Figures 2 B and E were sampled from  $U[-1, 1]$  and  $U[0.5, 1]$ , respectively. This choice of parameters ensured that approximately 99% of the  $R^2$  coefficients between phenotypes and QTL ranged between 0.08 and 0.32 for the simulations based on sample size of 112 subjects (see Figure SI.2a, and the axis scales on Figures S3-S7) and between 0.01 to 0.20 for the simulations based on 1,000 subjects (see Figure SI.2b, and the axis scales on Figures S8-S12).

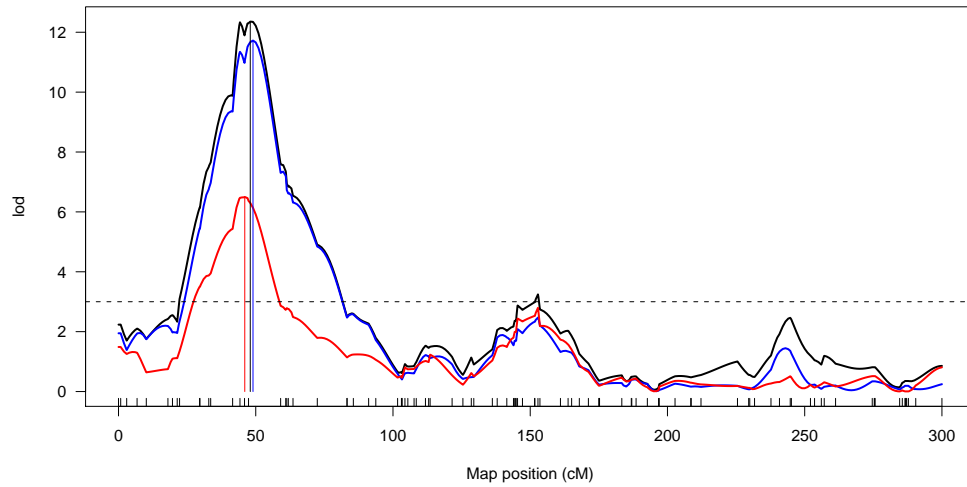
The backcross simulations and the QTL mapping analyses were performed using the R/qtl software (Broman et al. 2003). We performed Haley-Knott regression (Haley and Knott 1992) and adopted Haldane's map function, genotype error rate of 0.0001, and set

the maximum distance between positions at which genotype probabilities were calculated to 2cM. We used a permutation LOD threshold (Churchill and Doerge 1994) of 2.24 for the QTL mapping analysis, aiming to control the genome wide error rate of falsely detecting a QTL at a 5% rate.

Often times the phenotypes map to nearby but not precisely the same QTL, and we need to decide which QTL to use as the causal anchor. When testing expression traits against clinical traits, Millstein et al. (2009) and Schadt et al. (2005) suggest using the clinical trait QTL as the anchor. We adopt a different approach. When the phenotypes map to distinct regions that are less than 2cM apart we determine the QTL position using both phenotypes, jointly, as follows. For each pair of phenotypes ( $Y_1, Y_2$ ) we perform unconditional mapping analysis for  $Y_1$  and  $Y_2$  and conditional mapping analysis for  $Y_2$  given  $Y_1$ . Let  $LOD_1$  represent a LOD score for the mapping analysis of  $Y_1$ , and  $LOD_{2|1}$  for the mapping analysis of  $Y_2$  given  $Y_1$ . Since

$$\log_{10} \left\{ \frac{f(y_1, y_2 | q)}{f(y_1, y_2)} \right\} = \log_{10} \left\{ \frac{f(y_1 | q)}{f(y_1)} \right\} + \log_{10} \left\{ \frac{f(y_2 | y_1, q)}{f(y_2 | y_1)} \right\}, \quad (1)$$

we compute the joint LOD score of ( $Y_1, Y_2$ ) as  $LOD_{1,2} = LOD_1 + LOD_{2|1}$  (or equivalently as  $LOD_{1,2} = LOD_2 + LOD_{1|2}$ ). We determine the peak QTL position,  $\lambda$ , using the  $LOD_{1,2}$  scores profile and assign the QTL to  $Y_1$  and  $Y_2$  if  $LOD_1$  and  $LOD_2$  are greater than the mapping threshold at the  $\lambda$  position. Figure SI1 illustrates our approach. When both phenotypes co-map to more than one QTL we select the QTL with the highest joint mapping peak.



**Figure SI.1** We simulated data from a model  $Q \rightarrow Y_1 \rightarrow Y_2$ , with a QTL,  $Q$ , at 50cM. The blue and red curves show the (unconditional) LOD profiles of phenotypes  $Y_1$  and  $Y_2$ , respectively. The black curve depicts the joint LOD curve, and the peak QTL position  $\lambda$  is given by the black vertical line. Instead of having to perform an arbitrary choice between the QTLs given by the red and blue vertical lines we use the QTL given by the black line. The dashed line shows the QTL mapping threshold.