# Supplementary Methods

**1. Equivalence between genotype and haplotype measures of LD**

In this section we present the standard haplotype two- and three-loci measures of LD and establish the connection between these measures and the genotype moments involved in expression [3].

***Two-loci haplotype measure of LD.***  Consider a pair of bi-allelic loci (A and B, with alleles A1/A2 and B1/B2, respectively). The haplotype linkage disequilibrium parameter is . Let when allele is present and when allele is present. Likewise, let when allele is present and when allele is present. Then , , and the covariance between and is which reduces to , thus

If the two genotypes are centered, then and

[4]

This is a haplotype analog of the 1st order measures of LD entering in expression [3].

***Three-loci haplotype measure of LD***. For a system involving three bi-allelic loci (, and , with alleles A1/A2, B1/B2, and C1/C2, respectively) a three-loci haplotype measure of LD can be defined as (Bennett 1954)

Extending the two-loci system by introduction of three binary random variables , , and , that take values when the allelic forms , and are present, respectively, and take values othewise, yields

The three terms in square brackets represent pairwise disequilibria. When the three random variables are centered their marginal expectations are zero and the expression reduces to

[5]

***Relationship with genotype measures of LD***. The disequilibria measures described above involve associations between alleles within gametes, whereas in the body of the paper, the expectations involve different genotypes. Assuming random mating, the expectations involving genotypes result in twice those involving gametes.

**2. Perfect LD between markers and QTL prevents phantom epistasis**

We demonstrate (the very intuitive result) that if a response () can be fully captured by regression on a set of predictors (), then the regression of on plus ,

, [6]

yields in the population.

***Demonstration:*** In the population, the regression coefficients of [6], are defined by the following system

Where the ’s represent covariance matrices: , and . It follows that

[7]

and

[8]

Solving [7] for yields (). Plugging this into [8] yields,

. And solving for gives

. [9]

Now if can be fully explained by regression on , that is if , with , then, , thus, and therefore, QED.

**Implication.** Let be the QTL genotype, be a vector containing the two marker genotypes and be the two-marker interaction contrast. Under perfect LD between the QTL and the markers the QTL genotype can be fully explained by linear regression on the two markers, that is. Therefore, the above result () implies , i.e., absence of phantom epistasis.

**3. Simulation Studies**

Simulations (10,000 MC replicates) were based on genotypes of unrelated Caucasians from the UK-Biobank, a cohort study consisting of about half a million participants aged between 40-69 years who were recruited in 2006-2010. The National Research Ethics Committee approved the study and informed consent was obtained from all participants. Study details are described elsewhere (Sudlow et al. 2015).

***Genotypes*** where from the Affymetrix UK BiLEVE Axiom and Affymetrix UK Biobank Axiom® arrays. Only SNPs with minor-allele-frequency greater than 0.1% and those with missing calling rate smaller than 3% were used for simulations. Furthermore, since we focused on a single locus model, we used only SNPs mapped to chromosome 1. There were SNPs mapped to chromosome , of those, passed our minor-allele frequency and calling rate inclusion criteria.

To avoid confounding due to population structure we only considered subjects whose self-reported ethnicity was Caucasian and confirmed their genetic race/ethnicity using SNP-derived principal components. From these individuals, we identified ~270,000 subjects that have pairwise genomic relationships, , smaller than 0.03. Here, and are genotypes (coded as 0, 1, 2) at the *kth* SNP of the *ith* and *jth* individual, respectively, and is the frequency of the allele counted at the *kth* locus. Genomic relationships were computed using the *getG()* function of the BGData R-package (Grueneberg and de los Campos 2017).

**Single-QTL model**: in our first simulation scneariophenotypes were generated according to a single-locus additive model (expression [1]). with the QTL explaining either 1% or 0.5% of the phenotypic variance. Markers and QTL were all from chromosome 1. In this scenario, marker was always proximal to the QTL and marker was placed at increasing base-pair distance of .

**SNPs in different chromosomes:** here the simulation was as in the previous scenario but the second marker was placed in a randomly chosen position in chromosomes 2.

**Model with infinitesimal effect**: We simulated a trait with a major QTL and infinitesimal effects. The trait heritability was 0.5, the main QTL explained 1% of the variance and the infinitesimal component the remaining 49%. The infinitesimal effect was simulated using genotypes from 500 randomly chosen SNPs from the 22 autosomal chromones with effects sampled from normal distributions. For this scenario we only run simulations for sample size equal to 50K and 250K.

Three marker-QTL pairs. In this scenario there were three marker-QTL pairs: , and . As before, the trait was strictly additive, with the genetic effect, , explaining 0.01 (i.e., 1%) of the phenotypic variance (the three QTL have identical effects, thus, on average each QTL explained about 0.01/3 of the variance). Within a pair, the marker was the SNP immediately adjacent to the QTL in the array.

***Inferences*** were based on a linear model such as that of expression [2] extended with inclusion of an intercept and the top 5 SNP-derived PCs, that is

Principal components were included to avoid any confounding that may emerge from any remaining substructure that may have been present. The PCs used in [2b] were derived using 50K SNPs evenly distributed in the entire genome. The model of expression above was fitted via least squares using the ls.fit()function of R (R Development Core Team 2012). Then for each scenario and MC replicate we saved the p-value associated to the interaction term and counted the proportion of times that was rejected when using a significance level of 0.05.

In the model involving three marker-QTL pairs, [, the last simulation scenario considered, the empirical model of the expression above was applied using the markers in the first () and third () pairs.

***Genotypic measures of LD*** between pairs of loci, , and , were computed using the squared correlation between genotypes at the two loci. This information was stored for each MC replicate of each simulated scenario. The proportion of variance of the QTL genotype explained by linear regression on the two markers, , was computed by Analysis of Variance, of a linear model where the QTL genotype was regressed, via least squares, on the two markers using a main effects additive model of the form: . The R-squared from this model was also saved for each MC replicate of each scenario and then used to analyze the relationship between this LD measure and the rate of rejection of .