

Haplotype effects and PIC

Gibran Hemani

2016-10-26 01:31:34

If you know the causal SNP in a GWAS significant region, then $-\log_{10}$ p-values should get smaller as the R-square between causal SNP and other SNPs gets weaker, e.g.

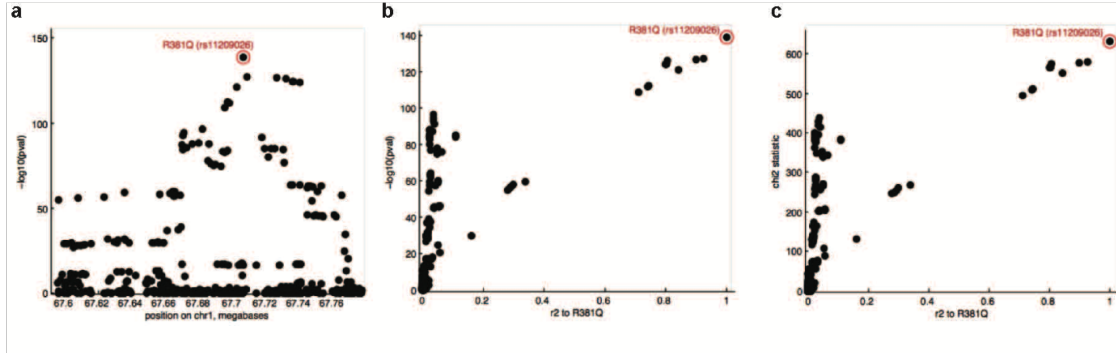
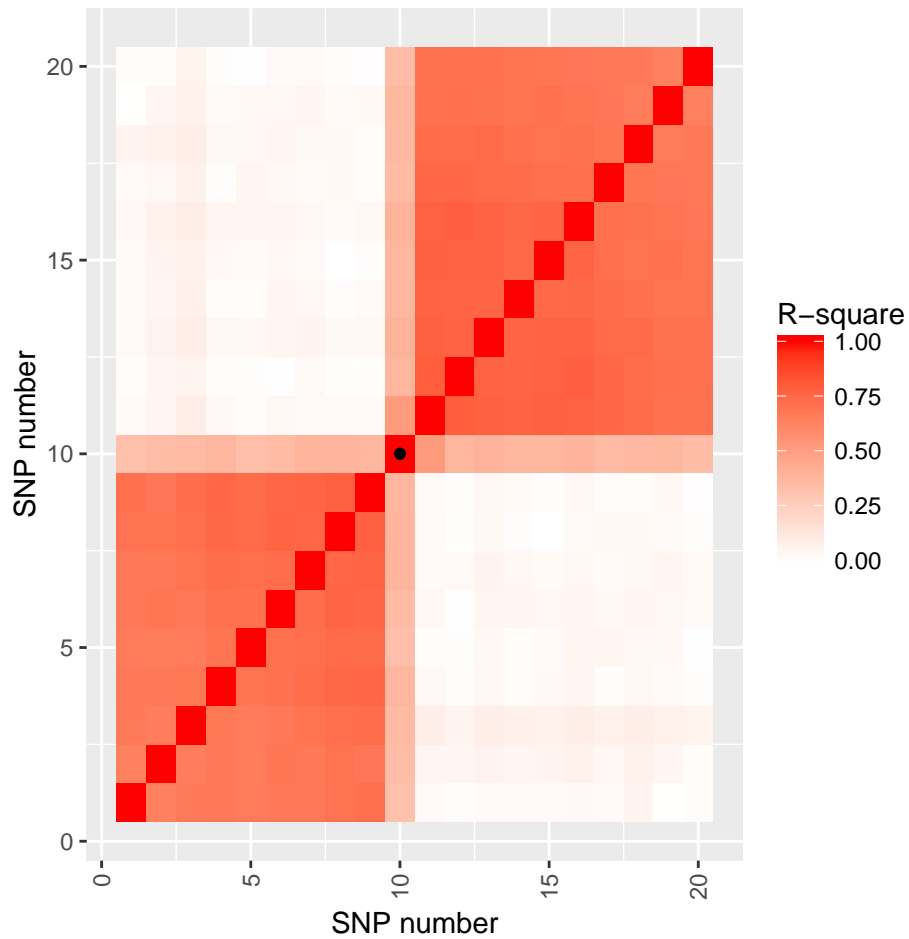


Figure 1: pics

The suggestion here is that there are two regression lines, suggesting possibly two real causal variants. This could be induced by two causal variants, but it struck me that this could also just be due to a haplotype effect.

Simulating haplotype effect

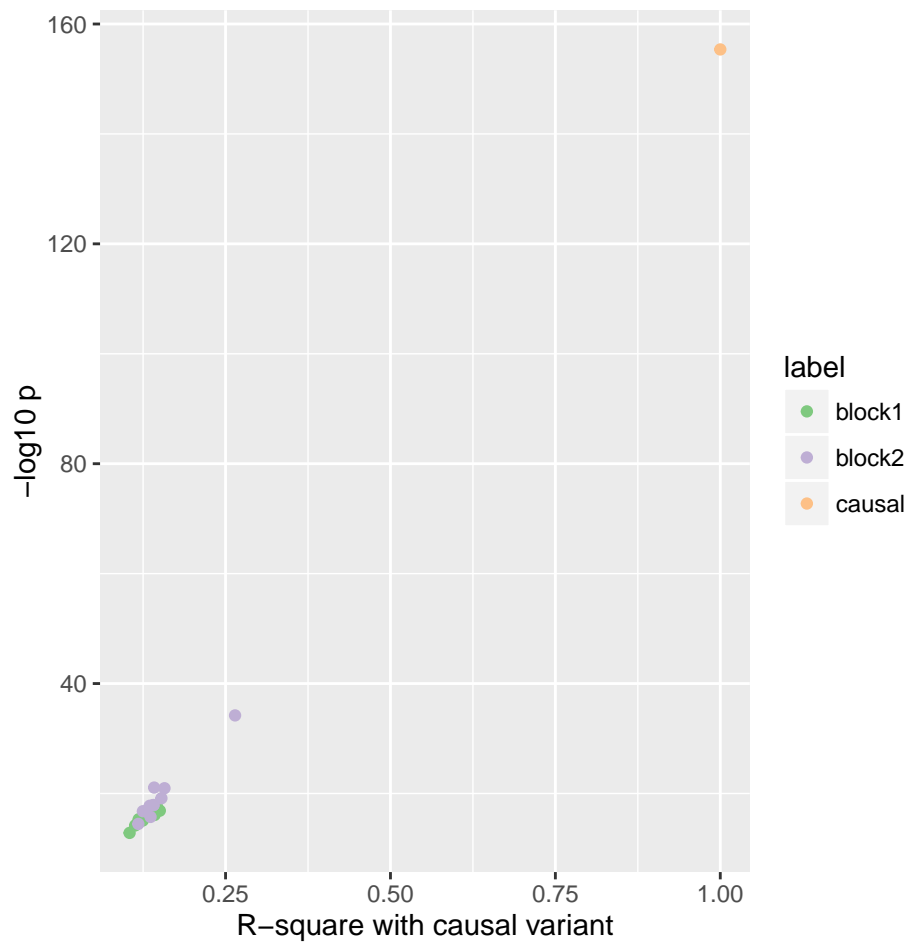
Imagine that there is a causal SNP, and it sits between two haplotype blocks. SNPs in block 1 are not correlated with SNPs in block 2. But SNPs in block 1 and 2 are all correlated with the causal SNP. For example an LD map might look something like this:



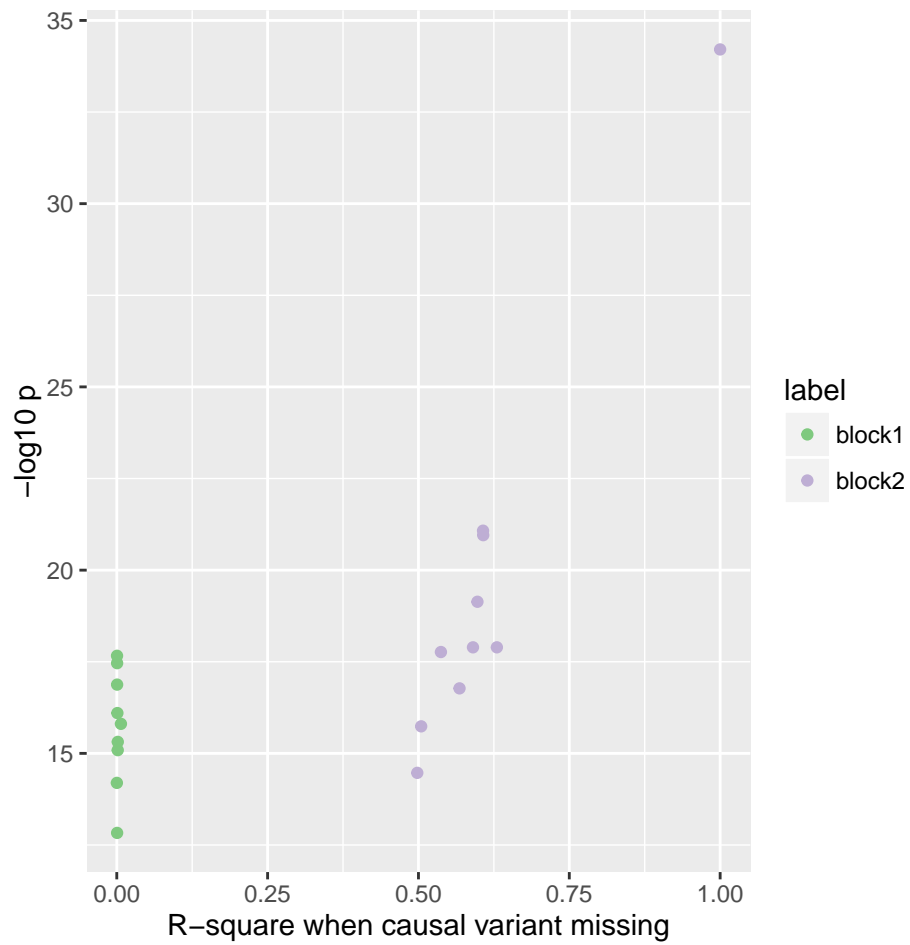
This is a depiction of a correlation matrix, that SNP that sits in between the two large correlated blocks is our causal variant, marked by the black spot.

Simulate a trait where the causal SNP has an effect on it, calculate the p-values of each SNP against the trait.

If the causal variant is in the data, we can make the PIC plot - rsq against $-\log_{10}$ pval:



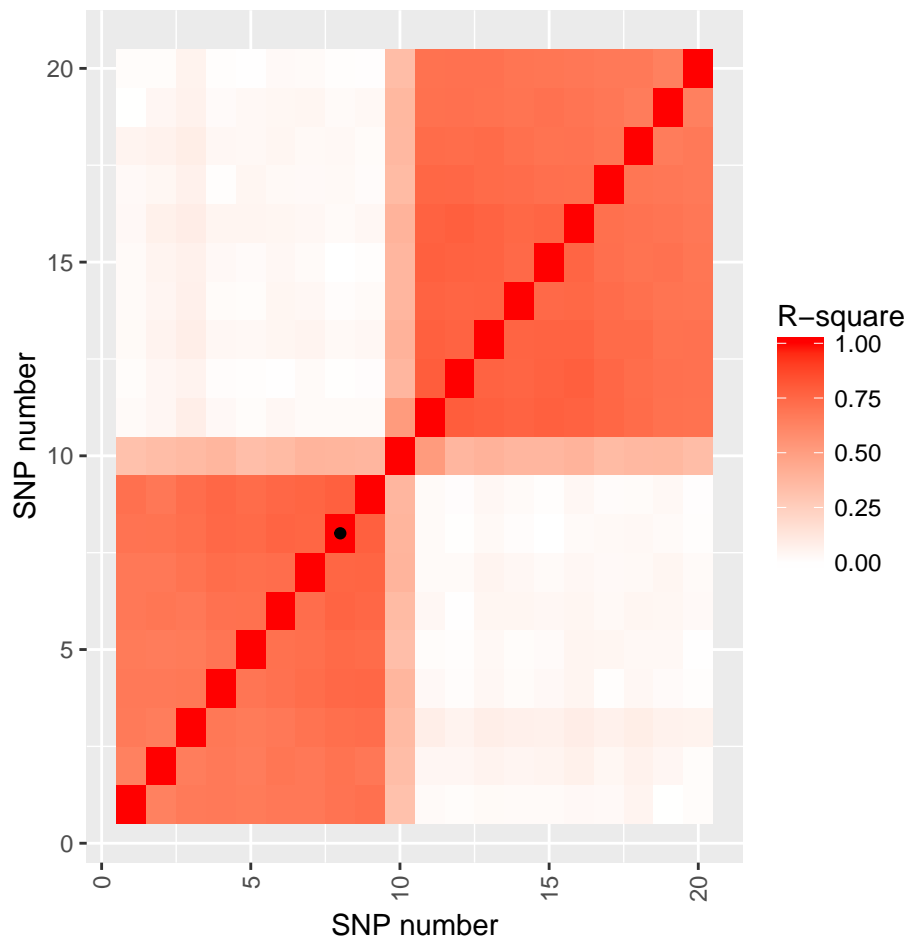
There is the expected linear relationship. If we didn't have that causal variant in the data, and we calculate the R-square against the best SNP in the data instead, we get the following relationship:



This is a similar pattern to that seen in the PIC paper (fig 1).

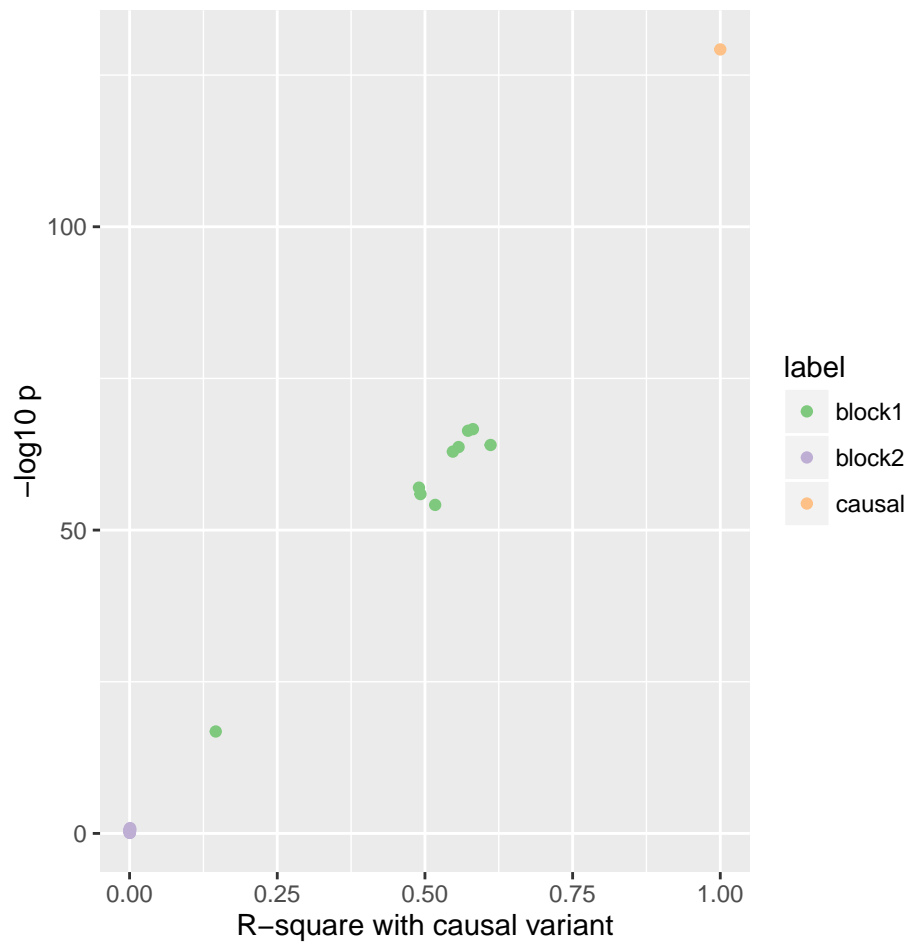
Single causal variant inside haplotype block

Note that it doesn't work if the causal variant resides neatly inside a single haplotype block. For example, imagine we change the causal variant:

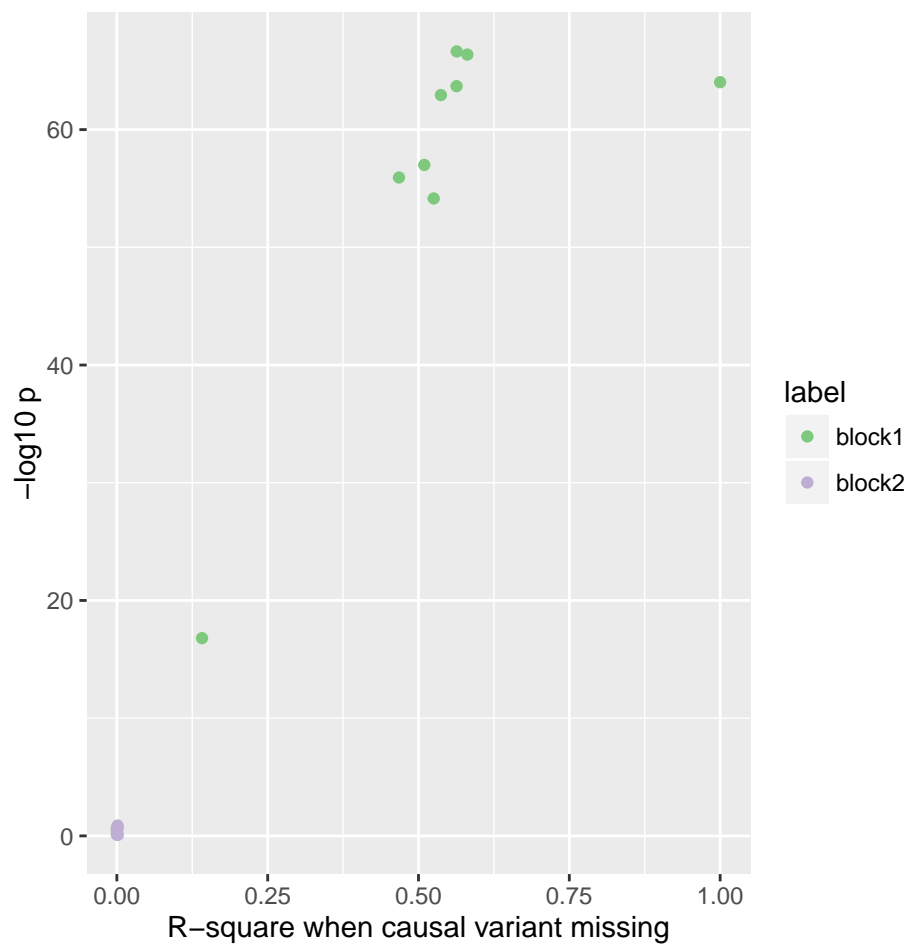


and run the simulation again:

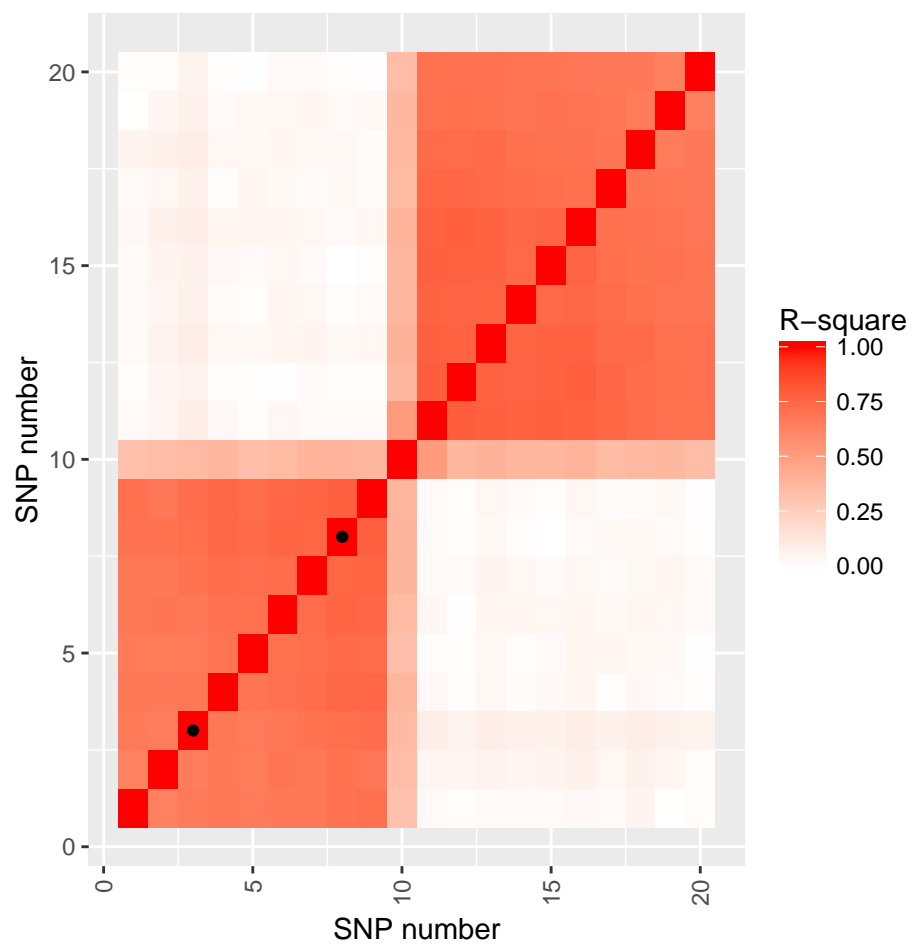
This plot shows what happens if the true causal variant is in the data:

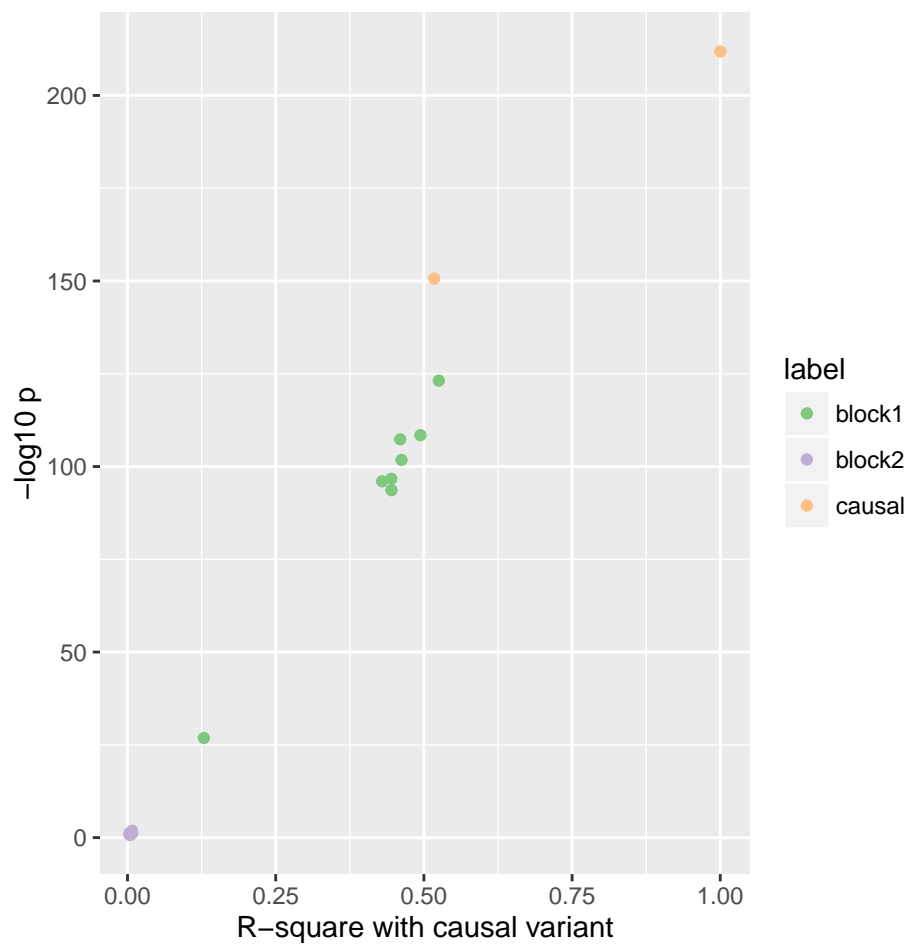


and this one shows what happens if that causal variant isn't present in the data:

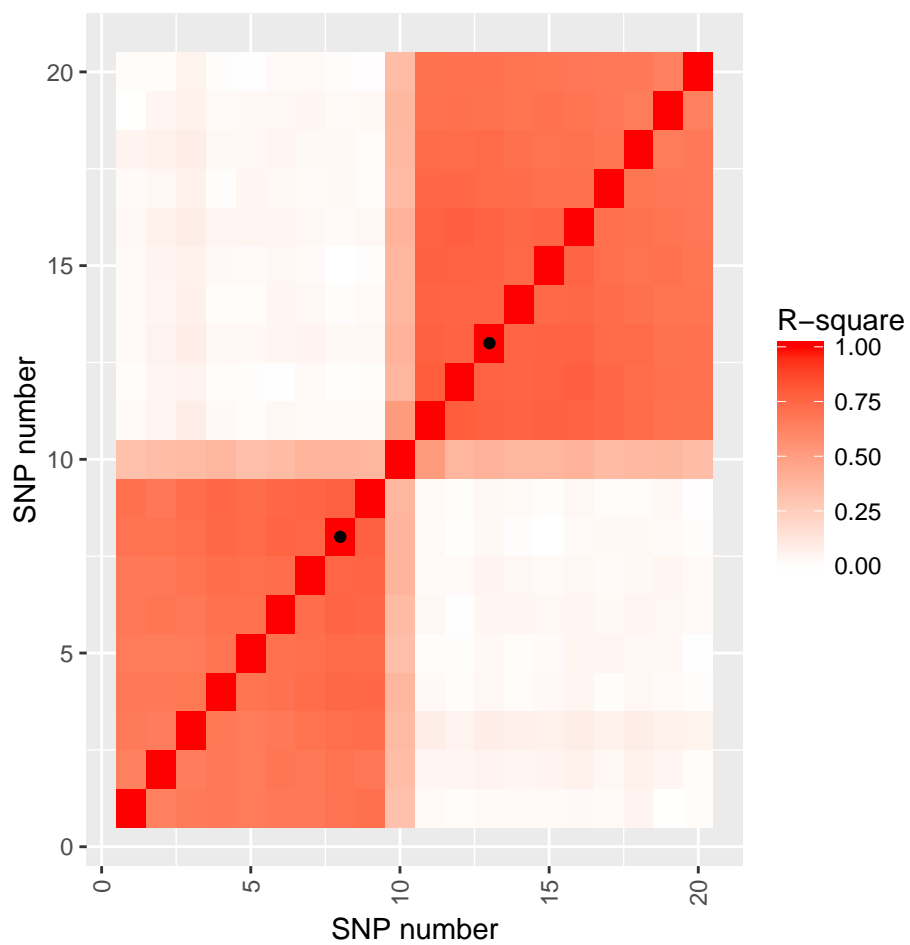


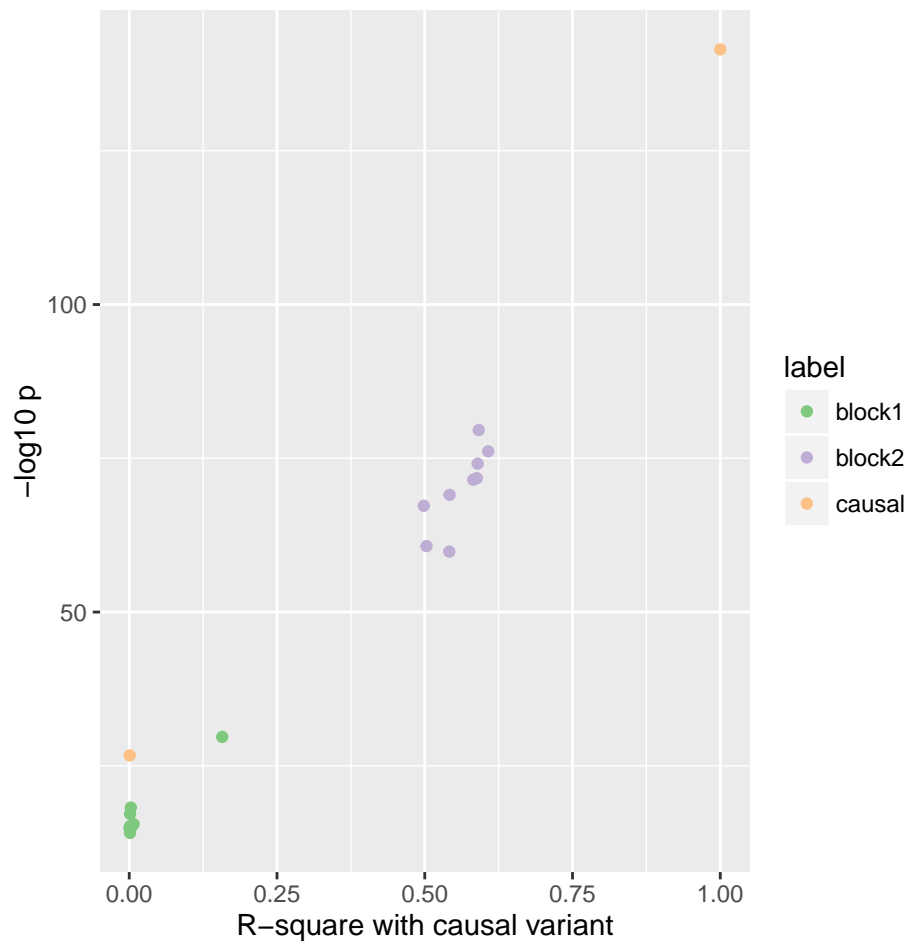
Two causal variants in one haplotype block





Two causal variants in different haplotype blocks





This recapitulates somewhat the pattern seen in PIC paper.

Conclusions

A causal variant that sits between haplotype blocks can induce the pattern seen in the PIC paper without requiring there to be two causal variants.