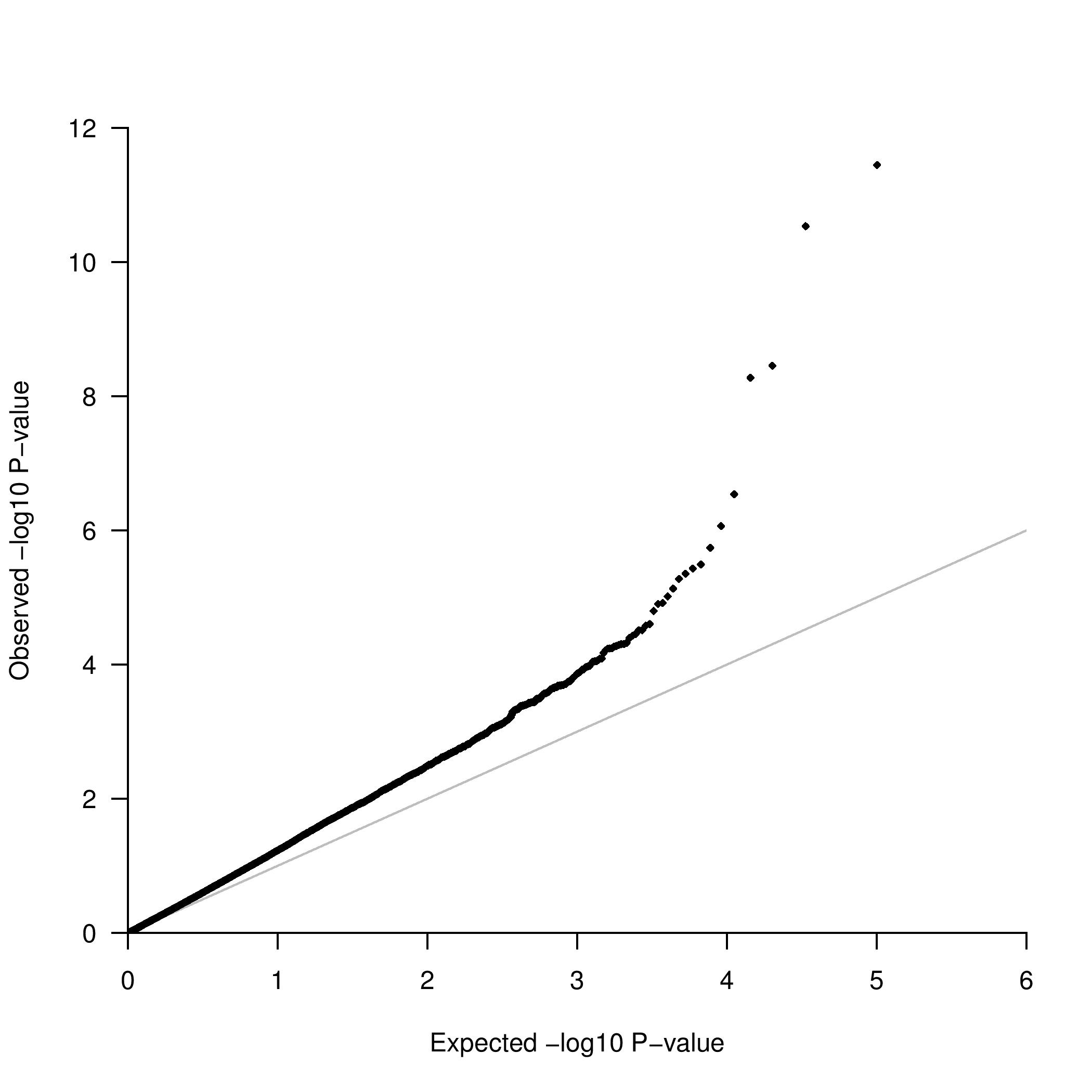
**Practical 4: Population structure in GWAS**

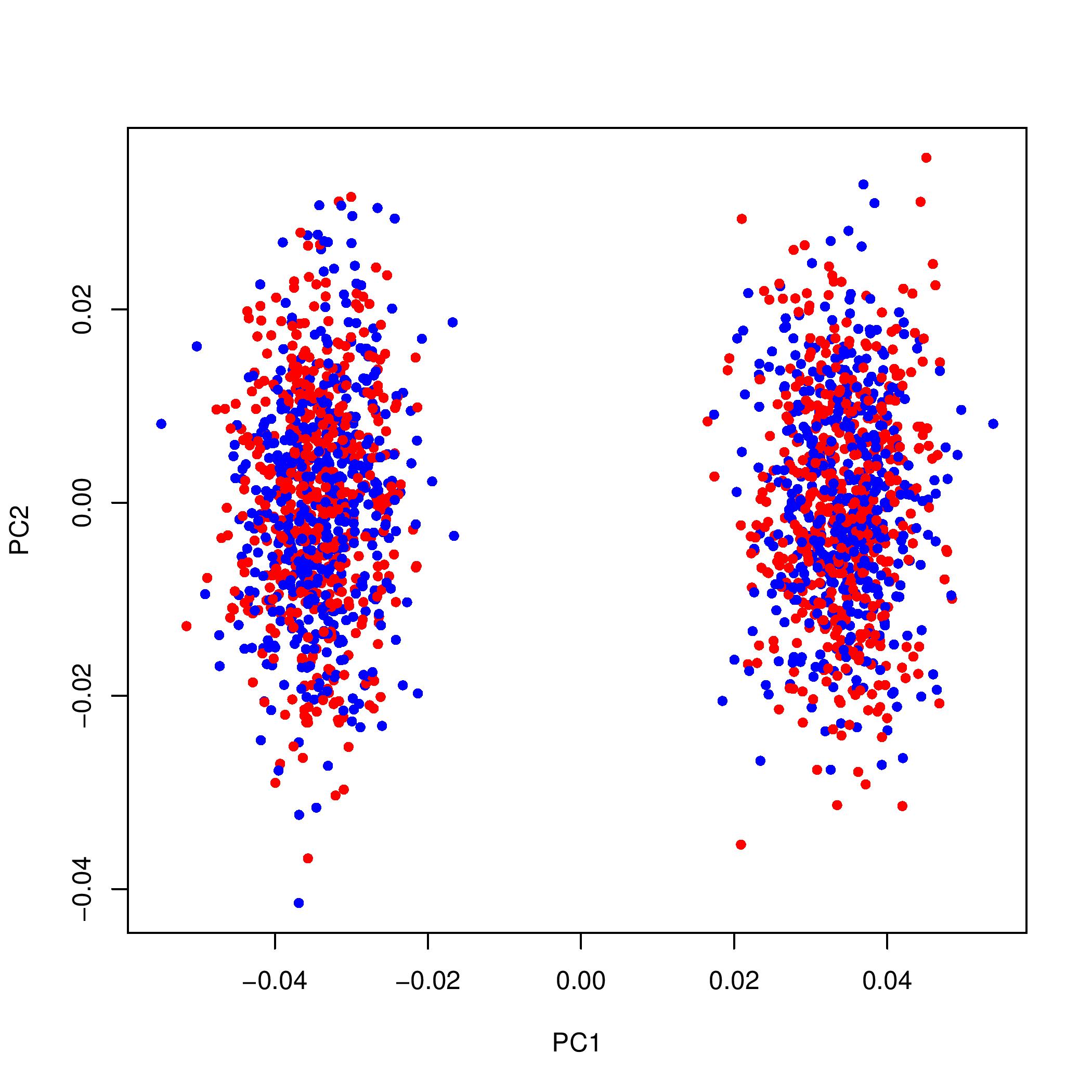
**Brief solutions**

Without adjustment for population structure, there is evidence for inflation in the QQ plot:

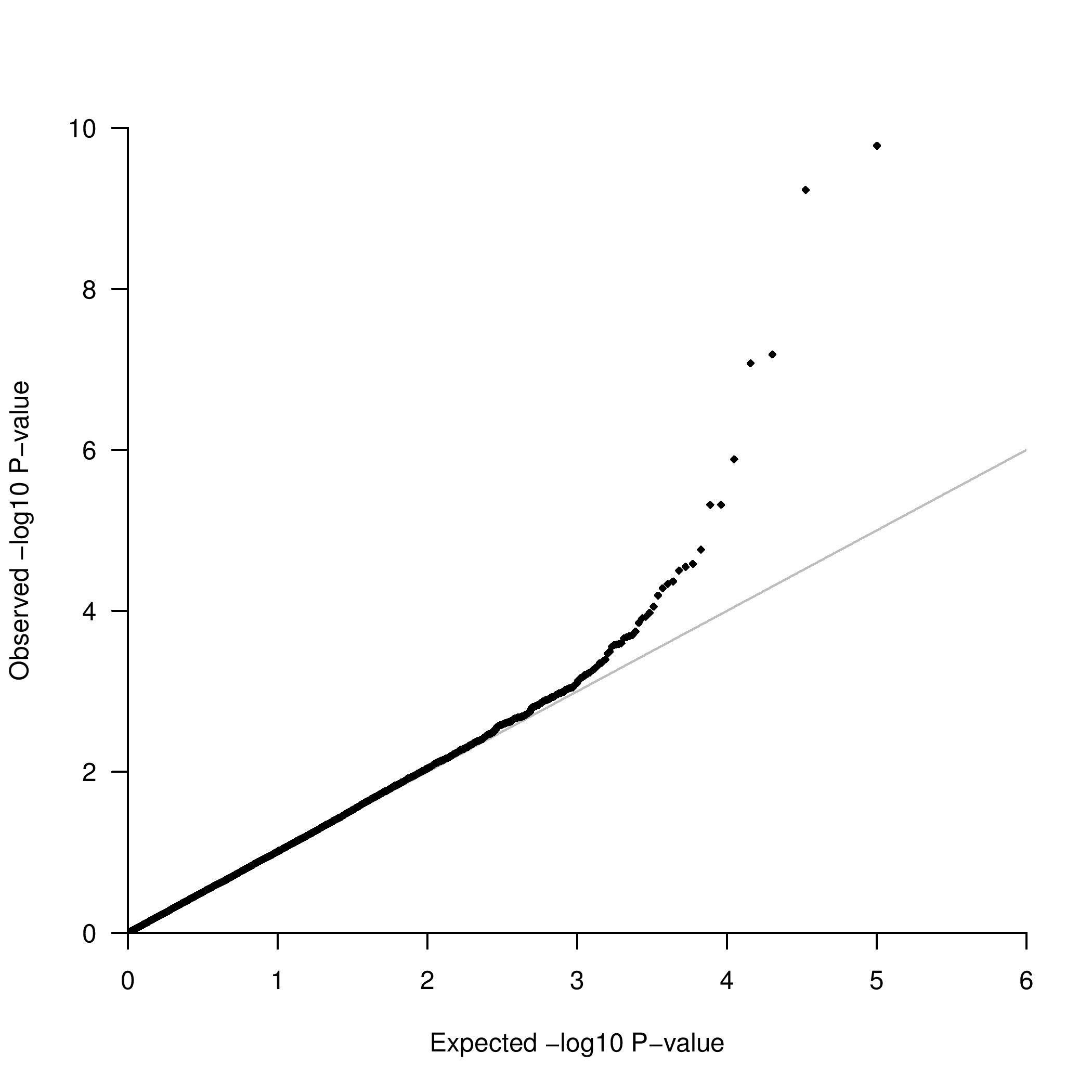


The genomic control inflation factor is 1.31, suggesting that there is structure that has not been accounted for in the analysis.

The plot of the first two eigenvectors from multi-dimensional scaling should look like:



After adjusting for the first two eigenvectors from multi-dimensional scaling, there is no evidence for inflation in the QQ plot:



The genomic control inflation factor is 1.019, suggesting that any population structure has been accounted for through adjustment for eigenvectors from multi-dimensional scaling.

In the analysis without correction for population structure, recall that the lead SNP from the additive model was rs12129523, with p=3.6x10-12. After adjusting for the first two eigenvectors from multi-dimensional scaling, the p-value from the additive model is less impressive, p=1.7x10-10, but remains genome-wide significant. Both eigenvector C1 and C2 are associated with disease status (p=0.0010 and p=0.26, respectively).