Supplementary Material

1. Obtaining normally distributed statistics

Many GWAS studies report χ^2 statistics with one degree of freedom (df) along with odds ratios (or beta regression coefficients) for the reference (tested) alleles. If only p-values are available, these can be transformed into χ^2 statistics using the inverse cumulative distribution function of the χ^2 statistics with one df. These χ^2 statistics need to be converted into normally distributed statistics (Z-scores). This can be accomplished, for instance, by taking i) the absolute value of the Z-score to be the square root of the χ^2 statistic and ii) the sign of the Z-score to be a) the sign of reported log odds ratio for binary traits or b) the sign of the reported slope coefficients for quantitative traits.

2. Simulation model

While the performance of tested methods should be assessed via simulations based on a real data set, phenotype simulations and univariate statistics calculation can be extremely computationally intensive, even when using fast analysis tools like PLINK (Purcell, Neale, Todd-Brown, Thomas, Ferreira, Bender, Maller, Sklar, de Bakker, Daly et al. 2007). Consequently, we decided to simulate the GWAS statistics directly by i) assuming that the density of SNPs in our GWAS is 1 SNP/Kbp, ii) assuming that SNPs more than 1 Mbp apart are uncorrelated and iii) employing a time series approach similar in spirit to Roeder et al.(Roeder, Bacanu, Wasserman, and Devlin 2006). To do so, we first estimated R^2 of SNP genotypes as a function of the distance between SNPs using reference Caucasian haplotypes available from Mach(Li, Willer, Ding, Scheet, and Abecasis 2010). Then we determined which Autoregressive Moving Average (ARMA) model better fits i) the estimated average cumulative R^2 between the genotype of a SNP and the genotypes of SNPs within 1 Mbp and ii) the estimated average R^2 between genotypes of SNPs situated {1,2,3,...,1000} Kbp apart. We deemed an ARMA (3,4) having an AR vector of {0.8716, 0.9782, -0.851} and an MA vector of {-0.6652, -0.9976, 0.6594, 0.0252} to be the most desirable model. This model provided the best approximation to the estimated average cumulative R^2 between the statistic at a certain SNP and the statistics at SNPs within 1 Mbp of it, even though it somewhat overestimates the R^2 between SNPs separated by intermediate distances (~100-300Kbp).

The number, effect size and genomic position of causal loci was modeled on the $m = 180$ significant findings from a mega-analysis of human height (Lango, Estrada, Lettre, Berndt, Weedon, Rivadeneira, Willer, Jackson, Vedantam, Raychaudhuri et al. 2010). We assumed that the phenotype under investigation has m_1 causal loci which represent a fraction $\gamma_c = \{0.25, 0.5, 1\}$ of the number of significant loci in height study

(Table I), i.e. $m_1 = \gamma_c\, m$. When $\gamma_c < 1$, the m_1 causal loci are chosen at random from the $m = 180$ significant loci in the height study. We simulated sample sizes equaling a fraction $\gamma_s = \{0.125, 0.25, 1\}$ of the height meta-analysis sample size $(n \approx 180{,}000)$. A sample size $n_1 = 0.125\, n$ is similar to the sample size of the discovery phase of Psychiatric Genetics Consortium (PGC) schizophrenia meta-analysis (Ripke, Sanders, Kendler, Levinson, Sklar, Holmans, Lin, Duan, Ophoff, Andreassen et al. 2011) and a sample size of $n_2 = 0.25\, n$ is similar to type 2 diabetes meta-analysis (Voight, Scott, Steinthorsdottir, Morris, Dina, Welch, Zeggini, Huth, Aulchenko, Thorleifsson et al. 2010) (and close to the sample size of both replication and discovery phases of PGC).

With these assumptions we can compute the mean of univariate statistics for each SNP and use them to simulate GWAS univariate statistics. To do so, we denote by $\boldsymbol{\theta}$ the vector of means of univariate statistics at the m_1 causal loci as estimated in the height meta-analysis paper (Lango, Estrada, Lettre, Berndt, Weedon, Rivadeneira, Willer, Jackson, Vedantam, Raychaudhuri et al. 2010). It follows that the means of univariate statistics at these causal loci for the simulated sample size, $n_1 = \gamma_s\, n$, is $\boldsymbol{\theta}_1 = \sqrt{\gamma_s}\, \boldsymbol{\theta}$. Based on the means of univariate statistics at causal loci, $\boldsymbol{\theta}_1$, and the ARMA LD structure, we use conditional expectation to compute the vector, $\boldsymbol{\mu}$, of means for univariate statistics at SNPs on a chromosome. To ease the computational burden, we set to zero the means of statistics at SNPs more than 1 Mbp away from the closest causal locus. Based on $\boldsymbol{\mu}$, we simulate, for each chromosome, SNP statistics as $\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ is the vector of ARMA(3,4) residuals having unit variance.

3. FIQT R script

```
FIQT<-function(z=z, min.p=10^-300){
        pvals<-2*pnorm(abs(z),low=F)
### Very low p-values bounded below to be able to compute (inverse) normal cdf for z's
        pvals[pvals<min.p]<- min.p
        adj.pvals<-p.adjust(pvals,method="fdr")
        mu.z<-sign(z)*qnorm(adj.pvals/2,low=F)
### Do not use the above adjustment for extremely large Z-scores (above around 37…)
### because their p-values were conservatively raised to min.p
        mu.z[abs(z)>qnorm(min.p/2,low=F)]<-z[abs(z)>qnorm(min.p/2,low=F)]
        mu.z
}
```

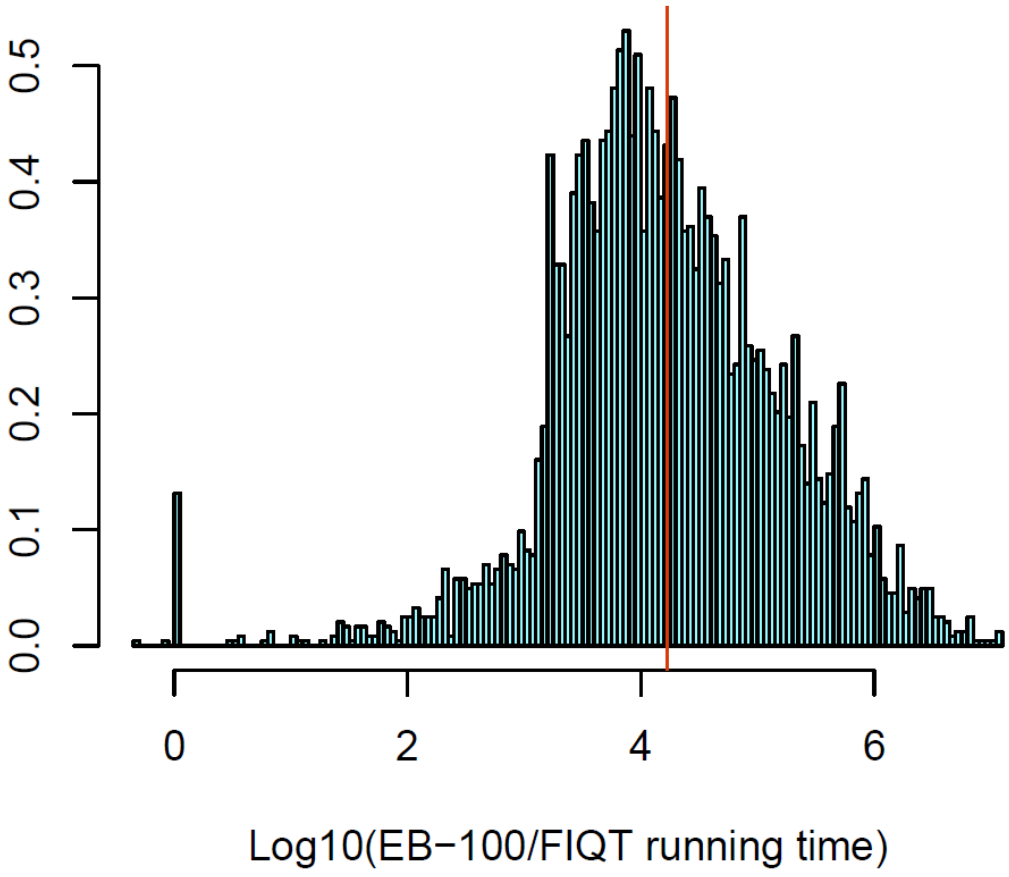

Figure S1. Log10 of the quotient between the running times of EB-100 and FIQT. The red vertical line denotes the mean.

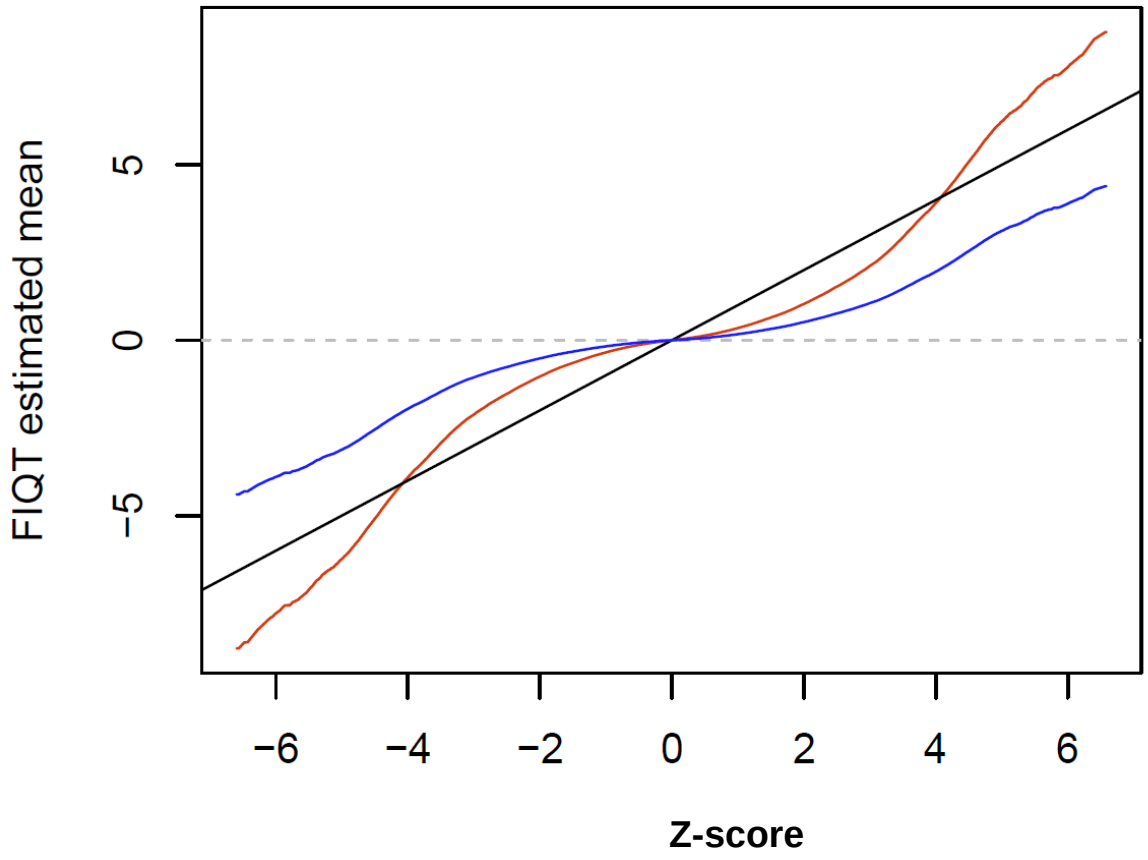

Figure S2. FIQT estimated means for PGC1 (blue) and PGC2 (red). First bisector (black) added for reference.

Reference List

1. Lango, A.H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S., et al. 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**:832-838

2. Li, Y., Willer, C.J., Ding, J., Scheet, P., and Abecasis, G.R. 2010. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol.* **34**:816-834

3. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum.Genet.* **81**:559-575

4. Ripke, S., Sanders, A.R., Kendler, K.S., Levinson, D.F., Sklar, P., Holmans, P.A., Lin, D.Y., Duan, J., Ophoff, R.A., Andreassen, O.A., et al. 2011. Genome-

wide association study identifies five new schizophrenia loci. *Nat.Genet.* **43**:969-976

5. Roeder, K., Bacanu, S.A., Wasserman, L., and Devlin, B. 2006. Using linkage genome scans to improve power of association in genome scans. *Am J Hum.Genet.* **78**:243-252

6. Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G., et al. 2010. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat.Genet.* **42**:579-589