Supplementary Methods

Expected numbers of IDP replications

To calculate the chance of replicating an individual SNP-IDP association in the replication set of 930 subjects we extracted the following information for each SNP detailed in Supplementary Tables S1 and S2

p = non-reference allele frequency
β = estimated effect size

If we let Y denote a phenotype and G denote the genotypes at a SNP with frequency p and effect size β, then consider the regression of Y on G. The variance of the SNP effect will be $\beta^2 p(1-p)$. If we assume that the phenotype has been mean centered and scaled to have variance 1, and that the SNP is likely to account for a very small fraction of the phenotype variance, then the estimated residual variance in the regression will be close to 1. This implies that the non-centrality parameter of the F-test for a non-zero effect size at the SNP, when Y is regressed onto G, will be approximately

$$\eta = N\beta^2 p(1-p)$$

where N is the sample size.

Therefore, if F is a random variable with F-distribution with 1 and N degrees of freedom and non-centrality parameter η, then the probability that the SNP will show an association at the 5% significance level is approximately equal to probability that F is greater than f, the 95% point of an F-distribution with 1 and N degrees of freedom and non-centrality parameter 0.

These replication probabilities were calculated for each SNP in Supplementary Tables S1 and S2. The probabilities were then summed to give the expected number of replications.