Supplementary Materials

Evaluation of the genetic instruments for the MR studies

We computed the proportion of the variance of the respective protein level explained by the *cis*-pQTL (R^2) using the following formula: $R^2 \approx 2\beta^2 f(1-f)$, where β and f denote the effect estimate and the effect allele frequency of the allele on a standardized phenotype respectively (1). We also calculated the F-statistic of each *cis*-pQTL using the following formula: $F = (R^2/k)/([1-R^2]/[n-k-1])$, where R^2 is the proportion of the variance of the respective protein level explained by the *cis*-pQTL, k is the number of instruments used in the model (in this case k=1 since there was a single *cis*-pQTL per protein) and n is the GWAS sample size (2).

Multi-instrument MR and Sensitivity Analyses

We queried for *trans*-pQTL (for the proteins prioritized by our main MR analysis using only *cis*-pQTL) in the *Sun et al* GWAS. Whenever available, these additional SNPs were included as instruments in an MR analysis, and their effects were meta-analyzed to generate an inverse variance weighted MR estimates (IVW), along with estimates from pleiotropy-robust MR methods , such as the MR-Egger (3), the weighted median (4), the mode-based method (all three implemented in the TwoSampleMR R package) and the Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO) method (5) (global, outlier and distortion tests implemented in the MR-PRESSO R package, version 1.0) .

ELISA validation

For each assay, the following metrics were assessed: linearity, spike-recovery and intra-assay precision. To evaluate linearity, serum samples were serially diluted with a standard diluent or phosphate-buffered saline to produce samples with values within the dynamic range of the assay. The recovery was defined as the ratio of observed concentration at dilution on the expected concentration after dilution. Recovery rates between 80%-120% indicate that the assay provides flexibility to measure samples with different levels of protein.

Spike-Recovery evaluation was conducted by adding a known amount of each protein to the sample matrix and standard diluents, and the two sets of responses were compared based on values calculated from a standard curve. The Spike-Recovery was defined as the ratio of the observed spiked sample value minus the unspiked sample value on the actual amount spiked in a sample. Recovery rates between 80%-120% suggest the ELISA method is compatible with the tested sample matrix. Finally, to assess intra-assay precision, a single sample was tested multiple times on one assay plate.

Co-localization analyses

The results of our co-localization analyses as computed by the coloc R package were interpreted as follows : each colocalization analysis provided posterior probabilities for H0 (no association of the genomic locus with either trait), H1 (association with type 1 diabetes but not with the protein level), H2 (association with the protein level but not with type 1 diabetes), H3 (association with type 1 diabetes and the protein level through two different SNPs), and H4 (association with type 1 diabetes and the protein level through one shared SNP).

References

1. Park JH, Wacholder S, Gail MH, Peters U, Jacobs KB, Chanock SJ, Chatterjee N. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. Nat Genet 2010;42:570-575

2. Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, Davey Smith G, Sterne JA. Using multiple genetic variants as instrumental variables for modifiable risk factors. Stat Methods Med Res 2012;21:223-242

3. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512-525

4. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016;40:304-314

5. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018;50:693-698