Supplementary data for

Human pancreatic islet microRNAs implicated in diabetes and related traits by large-scale genetic analysis

Henry J. Taylor[a,b,c], Yu-Han Hung[d], Narisu Narisu[a], Michael R. Erdos[a], Matthew Kanke[d], Tingfen Yan[a], Caleb M. Grenko[a], Amy J. Swift[a], Lori L. Bonnycastle[a], Praveen Sethupathy[d,1,*], Francis S. Collins[a,1,*], D. Leland Taylor[a,1]

   a. Center for Precision Health Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA
   b. British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB2 0BB, UK
   c. Heart and Lung Research Institute, University of Cambridge, Cambridge CB2 0BB, UK
   d. Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

[1] These authors contributed equally to this work.
* To whom correspondence may be addressed. Email pr46@cornell.edu and
francis.collins@nih.gov

This repository contains supplementary data for the study entitled, "Human pancreatic islet microRNAs implicated in diabetes and related traits by large-scale genetic analysis". Further details are described in the Methods section of the study. All genomic coordinates are from the GRCh37 (hg19) assembly. As described in the Methods section, we performed all false discovery rate calculations using Storey's method. For variant associations, we report the single nucleotide polymorphism (SNP) rsid from dbSNP151. For SNPs without an rsid, we report them as `<chr>:<position>`.

1. mirna_eqtl__nominal.tsv.gz: tab-delimited file of nominal miRNA-eQTL results. This file contains results for all SNP-miRNA pairs tested.

Column	Description
var_id	SNP identifier
var_chr	SNP chromosome
var_pos	SNP position
effect_allele	Effect allele
other_allele	Non-effect allele
effect_allele_freq	Effect allele frequency
pheno_id	miRNA identifier
pheno_chr	miRNA chromosome
pheno_start	miRNA start position
pheno_end	miRNA end position
beta	Beta
std_error	Standard error
pvalue	Nominal *P*-value

2. mirna_eqtl__lead.tsv.gz: tab-delimited file of lead miRNA-eQTL results. This file contains the lead SNP-miRNA pair for each miRNA (i.e., smallest SNP-miRNA association *P*-value).

Column	Description
pheno_id	miRNA identifier
pheno_chr	miRNA chromosome
pheno_start	miRNA start position
pheno_end	miRNA end position
n_snps	Number of proximal SNPs tested
var_id	SNP identifier
var_chr	SNP chromosome
var_pos	SNP position
effect_allele	Effect allele
other_allele	Non-effect allele

effect_allele_freq	Effect allele frequency
beta	Beta
std_error	Standard error
pvalue	Nominal *P*-value
empirical_pvalue	Empirical *P*-value
empirical_qvalue	Empirical q-value

3. <u>Differential expression results</u>. We provide tab-delimited files containing summary statistics of the differential expression results for the following traits: age, sex, BMI, T2D status, and polygenic scores for fasting blood glucose, fasting insulin, blood glucose, HbA1c, and T2D. We provide a file for each trait. $\log_2$(Fold change) is denoted by LFC. In each file, missing data is denoted using NA (e.g., when an miRNA is tested in one dataset but not the other).

    a. For age, sex, BMI, and T2D status, we meta-analyze results across library preparation 1 (LP1), library preparation 2 (LP2), and Kameswaran et al. (1). The columns are:

Column	Description
miRNA	miRNA identifier
trait	Trait tested
taylor_lp1__wald_stat	LP1 analysis wald statistic
taylor_lp1__lfc	LP1 analysis LFC
taylor_lp1__std_error	LP1 analysis standard error
taylor_lp1__pvalue	LP1 analysis *P*-value
taylor_lp2__wald_stat	LP2 analysis wald statistic
taylor_lp2__lfc	LP2 analysis LFC
taylor_lp2__std_error	LP2 analysis standard error
taylor_lp2__pvalue	LP2 analysis *P*-value
kameswaran__wald_stat	Kameswaran analysis wald statistic
kameswaran__lfc	Kameswaran analysis LFC
kameswaran__std_error	Kameswaran analysis standard error
kameswaran__pvalue	Kameswaran analysis *P*-value
meta__lfc	Meta-analysis LFC
meta__std_error	Meta-analysis standard error
meta__pvalue	Meta-analysis *P*-value
meta__qvalue	Meta-analysis q-value

    b. For the fasting blood glucose, fasting insulin, blood glucose, HbA1c, and T2D polygenic scores, we meta-analyze results across African, European, and Hispanic ancestries. The columns are:

Column	Description
miRNA	miRNA identifier
trait	Trait tested
african__wald_stat	African analysis wald statistic
african__lfc	African analysis LFC
african__std_error	African analysis standard error
african__pvalue	African analysis *P*-value
european__wald_stat	European analysis wald statistic
european__lfc	European analysis LFC
european__std_error	European analysis standard error
european__pvalue	European analysis *P*-value
hispanic__wald_stat	Hispanic analysis wald statistic
hispanic__lfc	Hispanic analysis LFC
hispanic__std_error	Hispanic analysis standard error
hispanic__pvalue	Hispanic analysis *P*-value
meta__lfc	Meta-analysis LFC
meta__std_error	Meta-analysis standard error
meta__pvalue	Meta-analysis *P*-value
meta__qvalue	Meta-analysis q-value

References

1. V. Kameswaran, *et al.*, Epigenetic regulation of the DLK1-MEG3 microRNA cluster in human type 2 diabetic islets. *Cell Metab.* **19**, 135–145 (2014).