The trans-ancestral genomic architecture of glycemic traits

LifeLines Cohort Study; Chen, Ji; Spracklen, Cassandra N.; Marenne, Gaëlle; Varshney, Arushi

Published in:
Nature Genetics

DOI:
10.1038/s41588-021-00852-9

Publication date:
2021

Document Version
Peer reviewed version

Extended Data Figure Legends

Extended Data Figure 1. Flow diagram of this study. The figure shows the data, key methods and main analyses included in this effort.

Extended Data Figure 2. Locus diagram. Trans-ancestry locus A contains a trans-ancestry lead variant for one glycemic trait represented by the blue diamond, and another single-ancestry index variant for another glycemic trait represented by the orange triangle. Single-ancestry locus B contains a single-ancestry lead variant represented by the purple square. The orange, blue and purple bars represent a +/- 500Kb window around the orange, blue, and purple variants, respectively. The black bars indicate the full locus window where trans-ancestry locus A contains trans-ancestry lead and single-ancestry index variants for two traits and single-ancestry locus B has a single-ancestry lead variant for a single trait.

Extended Data Figure 3. Venn diagram. Overlap of TA loci between traits.

Extended Data Figure 4. Allele frequency versus effect size. Allele frequency versus effect size for all signals detected through the trans-ancestry meta-analyses, for each of the four traits. Frequency and effect size are from the European meta-analyses. The power curves were computed based on the European sample size for each trait, and the mean (m) and standard deviation (sd) computed on the FENLAND study: FG, m=4.83 mmol/l, sd=0.68; FI, m=3.69 mmol/l, sd=0.60; 2hGlu, m=5.30 mmol/l, sd=1.74; HbA1c, m=5.55%, sd=0.48.

Extended Data Figure 5. EAF correlation and heterogeneity test. Pearson correlation of EAF on the lower tri-angle and p-value of one-side heterogeneity test without multiple testing corrections on the upper tri-angle of the trans-ancestry lead variants associated with each trait between ancestries. Correlations > 0.7 are in bold.

Extended Data Figure 6. Forest plot of T2D GRS from HbA1c variants. The p-value on the right side is from the two-side test without multiple testing corrections. Vertical points of each diamond represent the point estimate of the odds ratio. The horizontal points of each diamond represent the 95% confidence interval of the odds ratio. Figure shows the association results between HbA1c-associated variants built into a GRS for T2D by taking each HbA1c-associated variant and using a weight that corresponds to its T2D effect size (logOR) based on analysis by the DIAGRAM consortium. The overall GRS is subsequently partitioned according to the HbA1c signal classification. The overall and partitioned GRS were tested for association with T2D based on data from UK biobank.

Extended Data Figure 7. Enrichment of glycemic trait associated GWAS variants to overlap genomic annotations using GREGOR. Figure shows enrichment for 59 total static and stretch enhancer annotations considered. One-side test significance (red) is determined after Bonferroni correction to account for 59 total annotations tested for each trait; nominal significance ($P<0.05$) is indicated in yellow.

Extended Data Figure 8. Enrichment of glycemic trait associated GWAS variants to overlap genomic annotations using fGWAS. Figure shows log2(Fold Enrichment) of GWAS variants to overlap 59 static and stretch enhancer annotations calculated. Significant enrichment (red) is considered if the 95% confidence intervals (shown by the error bars) do not overlap 0.

Extended Data Figure 9. Enrichment of glycemic trait associated GWAS variants to overlap genomic annotations using GARFIELD. Figure shows the beta or effect size (log odds ratio) for GWAS variants to overlap 59 static and stretch enhancer annotations. GWAS variants were included at two significance thresholds, 1e-05 (A) and 1e-08 (B). One-side test significance (red) is determined after Bonferroni correction to account for effective annotations tested for each trait reported by GARFIELD (see supplementary note); nominal significance ($P<0.05$) is indicated in yellow. The 95% confidence intervals are shown by the error bars.