A019 · Characterizing Ancestry-Related Heterogeneity Between Additive and Recessive GWAS Models for Type 2 Diabetes

Presenter: Christelle Moise — Broad Institute, Broad Summer Scholars Program (BSSP)

Authors: Christelle Moïse, Samyak Maharjan, Maheak Vora, Katherine Taylor, Alicia Huerta-Chagoya, Josep M. Mercader

Session: Day 2 · Fri Oct 2, 2026 · 2:15–4:15 PM

Genome-wide association studies (GWAS) have identified numerous loci associated with type 2 diabetes (T2D), but genetic effects may vary across ancestral populations. We integrated publicly available multi-ancestry GWAS summary statistics from Suzuki et al. with an internal recessive GWAS meta-analysis to investigate ancestry-related heterogeneity. Variants were evaluated using heterogeneity analyses and prioritized based on recessive association significance, recessive-to-additive effect size ratios, and allele frequency differences across ancestries. Five loci showed strong evidence of ancestry-dependent recessive effects, with ancestry-specific analyses revealing variation in both effect sizes and allele frequencies. These findings emphasize the importance of diverse populations and recessive models for understanding T2D genetics.

Broad Summer SCHOLARS PROGRAM

Characterizing Ancestry-Related Heterogenei Between Additive and Recessive GWAS Models for Type 2 Diabetes

Christelle Moïse[1], Samyak Maharjan[2], Maheak Vora[3], Katherine Taylor[3], Alicia Huerta-Chagoya[3], Josep M. Me

[1] Belmont High School, [2] Revere High School, [3] Programs in Metabolism and Medical & Population Genetics, B Institute of Harvard and MIT, Center for Genomic Medicine, Massachusetts General Hospital Diabetes Unit Cardiovascular Research Center

ABSTRACT

Genome-wide association studies (GWAS) have identified numerous loci associated with type 2 diabetes (T2D), but genetic effects may vary across ancestral populations. We integrated publicly available multi-ancestry GWAS summary statistics from Suzuki et al. with an internal recessive GWAS meta-analysis to investigate ancestry-related heterogeneity. Variants were evaluated using heterogeneity analyses and prioritized based on recessive association significance, recessive-to-additive effect size ratios, and allele frequency differences across ancestries. Five loci showed strong evidence of ancestry-dependent recessive effects, with ancestry-specific analyses revealing variation in both effect sizes and allele frequencies. These findings emphasize the importance of diverse populations and recessive models for understanding T2D genetics.

BACKGROUND

Most type 2 diabetes (T2D) GWAS rely on additive models, missing variants that operate under recessive inheritance. Furthermore, genetic associations vary across ancestral populations due to differences in allele frequencies and effect sizes.

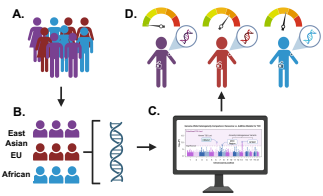

Figure 1. (A) Multi-ancestry T2D cohort aggregation. **(B)** Ancestral group stratification and genomic variant profiling. **(C)** Genome-wide additive vs. recessive model testing to identify heterogeneous loci. **(D)** Mapping prioritized loci back to ancestral risk profiles.

To address this, we investigate whether target T2D variants are both enriched in a specific ancestry and significant under a recessive model. Concordant signals across independent multi-ancestry datasets provide strong evidence for population-specific recessive risk.

METHODS & RESULTS

METHODS: MODEL FRAMEWORK

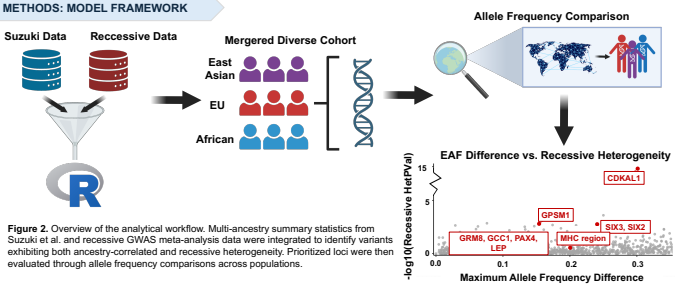

Figure 2. Overview of the analytical workflow. Multi-ancestry summary statistics from Suzuki et al. and recessive GWAS meta-analysis data were integrated to identify variants exhibiting both ancestry-correlated and recessive heterogeneity. Prioritized loci were then evaluated through allele frequency comparisons across populations.

RESULTS: PRIORITIZED LOCI

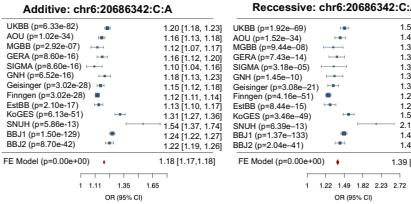

Figure 3. Side-by-side forest plots demonstrating cohort-specific effect sizes (Odds Ratio, 95% Confidence Interval), effect allele frequencies (EAF), and P-values before and after updating the dataset with UK Biobank summary statistics. Individual cohort estimates across major multi-ancestry biobanks illustrate effect magnitude consistency, while the overall pooled meta-analysis estimate (red summary diamond at the base) highlights a robust, statistically significant recessive risk association.

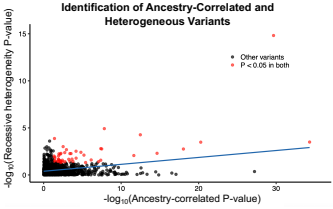

Figure 4. Ancestry-specific odds ratios for the CDKAL1 recessive variant across six ancestry groups. Error bars represent 95% confidence intervals; EAF and association p-values are shown for each ancestry.

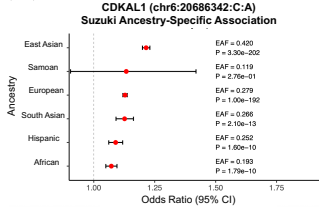

Figure 5. Comparison of effect sizes (OR, 95% CI), EAFs, and P-values across six major ancestry groups. Results demonstrate broad population risk consistency alongside notable allele frequency enrichment in East Asian cohorts.