Identification of Genetic Variation Influencing Metformin Response in a Multi-Ancestry Genome-Wide Association Study in the Diabetes Prevention Program (DPP)

Li, Josephine H.; Perry, James A.; Jablonski, Kathleen A.; Srinivasan, Shylaja; Chen, Ling; Todd, Jennifer N.

Published in:
Diabetes

DOI:
10.2337/db22-0702

Publication date:
2023

Document Version
Peer reviewed version

Identification of genetic variation influencing metformin response in a multi-ancestry genome-wide association study in the Diabetes Prevention Program (DPP)

Running Title: GWAS for metformin response in the DPP

Authors: Josephine H. Li, MD[1-4]; James A. Perry[5]; Kathleen A. Jablonski[6]; Shylaja Srinivasan[7]; Ling Chen[1, 3]; Jennifer N. Todd[1, 3, 8†]; Maegan Harden[3]; Josep M. Mercader[1-4]; Qing Pan[6]; Adem Y. Dawed[9]; Sook Wah Yee[10]; Ewan R. Pearson[9]; Kathleen M. Giacomini[10]; Ayush Giri[11]; Adriana M Hung[12]; Shujie Xiao[13]; L. Keoki Williams[13]; Paul W. Franks[14]; Robert L. Hanson[15]; Steven E. Kahn[16]; William C. Knowler[15]; Toni I. Pollin[5]; Jose C. Florez[1-4]; and Diabetes Prevention Program Research Group

[1]Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts; [2]Diabetes Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts; [3]Programs in Metabolism and Medical & Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts; [4]Department of Medicine, Harvard Medical School, Boston, Massachusetts; [5]Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland; [6]Department of Epidemiology and Biostatistics, George Washington University Biostatistics Center, Washington, District of Columbia; [7]Division of Pediatric Endocrinology and Diabetes, Department of Pediatrics, University of California at San Francisco, San Francisco, California; [8]Division of Endocrinology, Department of Pediatrics, Boston Children's Hospital, Boston, Massachusetts; [9]Division of Population Health and Genomics, Ninewells Hospital and School of Medicine, University of Dundee, Dundee, Scotland, U.K.; [10]Department of Bioengineering and Therapeutic Sciences, University of California at San Francisco, San Francisco, California; [11]Division of Quantitative Sciences,

Department of Obstetrics and Gynecology, Vanderbilt University Medical Center, Nashville, Tennessee; [12]Division of Nephrology and Hypertension, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee; [13]Center for Individualized and Genomic Medicine Research (CIGMA), Department of Internal Medicine, Henry Ford Health System, Detroit, Michigan; [14]Genetic and Molecular Epidemiology Unit, Lund University Diabetes Centre, Lund University, Malmo, Sweden; [15]Diabetes Epidemiology and Clinical Research Section, National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, Arizona; [16]Division of Metabolism, Endocrinology and Nutrition, Department of Medicine, VA Puget Sound Health Care System and University of Washington, Seattle, Washington

[†]Currently affiliated with Division of Endocrinology, Department of Pediatrics, University of Vermont Children's Hospital, Burlington, Vermont.

Corresponding Author: Jose C. Florez, MD, PhD c/o Diabetes Prevention Program Coordinating Center, George Washington University Biostatistics Center, 6110 Executive Boulevard, Rockville, Maryland 20852

Phone: 301-881-9260; Email: dppmail@bsc.gwu.edu, jcflorez@mgh.harvard.edu

Word Count: 4,616

Number of tables/figures: 3 tables, 4 figures (limit 8)

ABSTRACT (200 words)

Genome-wide significant loci for metformin response in type 2 diabetes reported elsewhere have not replicated in the Diabetes Prevention Program (DPP). To assess pharmacogenetic interactions in pre-diabetes, we conducted a genome-wide association study (GWAS) in the DPP. Cox proportional hazards models tested associations with diabetes incidence in metformin (MET, n=876) and placebo (PBO, n=887) arms. Multiple linear regression assessed association with one-year change in metformin-related quantitative traits, adjusted for baseline trait, age, sex, and 10 ancestry principal components. We tested for gene-by-treatment interaction. No significant associations emerged for diabetes incidence. We identified four genome-wide significant variants after correcting for correlated traits ($p<9\times10^{-9}$). In MET, rs144322333 near *ENOSF1* (minor allele frequency $[MAF]_{AFR}=0.07$, $MAF_{EUR}=0.002$) was associated with an increase in % glycated hemoglobin (per minor allele β=0.39 [95% CI 0.28, 0.50], $p=2.8\times10^{-12}$). Rs145591055 near *OMSR* (MAF=0.10 in American Indians), was associated with weight loss (kg) (per G allele β=-7.55 [95% CI -9.88, -5.22], $p=3.2\times10^{-10}$) in MET. Neither variant was significant in PBO; gene-by-treatment interaction was significant for both variants ($p_{(G\times T)}<1.0\times10^{-4}$). Replication in individuals with diabetes did not yield significant findings. A GWAS for metformin response in pre-diabetes revealed novel ethnic-specific associations that require further investigation but may have implications for tailored therapy.

INTRODUCTION

Metformin is the most commonly used agent for initial treatment of type 2 diabetes (T2D) with over 100 million users worldwide. Over the years, it has remained the foundation of clinical practice guidelines for the management of diabetes globally because it is cheap, safe, effective, well-tolerated, orally administered, and does not promote weight gain (1-5). However, despite its widespread use, metformin does not work well for everyone; in the ADOPT study, which examined the glycemic durability of three oral glucose-lowering agents, each as monotherapy for T2D, 21% of individuals on metformin failed to meet glycemic goals within the first five years of treatment (6). Moreover, treatment failure rates are as high as 50% in children and adolescents as shown by the TODAY study (7).

The reasons for metformin treatment failure remain unclear and could be related to genetic, pharmacologic, pathophysiologic, or environmental factors. Using the genome-wide complex trait analysis method, the heritability of metformin response has been estimated to be between 20-34%, a value comparable to other complex phenotypes (8). Indeed, genome wide association studies (GWAS) have identified common genetic variants associated with metformin response (9-11). In the first GWAS for metformin response, a variant around the gene encoding the ataxia-telangiectasia mutated kinase (*ATM*) was found to be associated with metformin response in GoDARTS and UKPDS participants (9). Subsequently, in a meta-analysis of over 10,000 participants with a harmonized glycemic measure of metformin response, a genome-wide significant association was observed in an intron of the *SLC2A2* gene encoding the GLUT2 glucose transporter (10). However, neither of these findings was reproduced in the Diabetes Prevention Program (DPP) (10, 12). While the difference in findings could be related to statistical power, one key distinction between the DPP and other studies is that DPP participants

have pre-diabetes and not established T2D. Therefore, it is possible that the influence of genetics on drug response could vary in the pre-diabetes state compared to the more advanced T2D state.

Thus, we performed a GWAS in the DPP to uncover novel variation associated with metformin response in this distinctive multi-ethnic cohort that was selected to be at high risk of T2D by virtue of elevated 2-hour glucose, fasting glucose, and body mass index (BMI). Our objective was to identify genetic variants associated with metformin response, as measured by both diabetes incidence and change in quantitative glycemic and metabolic traits. Through this approach, we were able to examine the genetics of metformin response as it relates to both pre-diabetes and T2D development as well as glycemic response.

RESEARCH DESIGN AND METHODS

Description of participants and DPP study design

The DPP study design and baseline characteristics of the participants have been described previously (13, 14). Briefly, the DPP was a multicenter randomized control trial designed to test the effects of intensive lifestyle modification and pharmacologic intervention on preventing progression to T2D in high-risk individuals. Enrolled participants had a fasting plasma glucose ranging from 95-125 mg/dL (5.3-6.9 mmol/L) and a 2-hour plasma glucose level between 140-199 mg/dL (7.8-11.0 mmol/L) on a 75-g oral glucose tolerance test (OGTT). A total of 3,819 participants were randomized to intensive lifestyle modification (goal weight loss ≥7% and ≥150 min/week of physical activity), standard lifestyle recommendations plus metformin (850 mg twice daily), standard lifestyle recommendations plus troglitazone (400 mg daily), or standard lifestyle recommendations plus placebo. The primary endpoint was diabetes incidence, diagnosed by a fasting glucose of ≥126 mg/dL (7.0 mmol/L) or a 2-hour glucose of ≥200 mg/dL

(11.1 mmol/L) after OGTT and confirmed subsequently on a second test within 6 weeks. The primary study demonstrated that over a mean follow-up of 2.8 years, there was a 58% (95% CI 48-66%) reduction of diabetes incidence in the intensive lifestyle intervention group and a 31% (95% CI 17-43%) reduction in the metformin group compared with placebo (15).

Institutional review board approval was obtained by each participating clinical center. All participants included in this analysis provided written informed consent for the main investigation and for subsequent genetic studies.

Genome-wide genotyping and quality control

DNA was extracted from peripheral blood leukocytes. Genotyping was performed on 3,227 samples using the HumanCore Exome genome-wide array (Illumina, San Diego, CA). We excluded single nucleotide polymorphisms (SNPs) with a call rate <95% or if they failed Hardy-Weinberg Equilibrium (HWE; $p<1.0\times10^{-8}$) within each ethnic group. Samples with discrepant sex, call rate <95%, inbreeding coefficient <-1, and identity-by-state as measured by pi-hat close to 1 were discarded. As 9,730 SNPs and 3,222 samples were additionally genotyped on the Metabochip (Illumina, San Diego, CA), we performed a concordance check, excluding SNPs and samples with a concordance rate <95%. After all quality checks were performed, 3,168 samples remained (Supplemental Figure S1).

Imputation

We performed a two-stage imputation procedure, which consisted of pre-phasing the genotypes into whole chromosome haplotypes followed by imputation itself. The pre-phasing was performed using SHAPEIT2 (16) and IMPUTE2 was utilized for genotype imputation (17). GWIMp-COMPSs can incorporate the contribution of several reference panels (18), and in this work, we used 1000 Genomes Phase3 haplotypes (October 2014) (19). Supplemental Table S1

summarizes the distribution of the 3,168 samples by self-reported race/ethnicity and DPP treatment arm that underwent imputation.

Statistical analyses

Because the study was conceived as a pharmacogenetic study on the influence of genetics on metformin response, we focused our primary analyses on the standard lifestyle recommendations plus metformin (MET) group and the placebo (PBO) group. Continuous variables are presented as mean ± SD and categorical variables as frequency (%). Baseline characteristics were compared with analysis of variance (ANOVA) tests for quantitative variables and chi-square tests for qualitative variables.

Diabetes incidence

A Cox proportional hazards model tested the association between genetic variants and diabetes incidence under an additive genetic model in the metformin and placebo arms. For any genome-wide significant findings, we planned to check and ensure that the proportionality assumptions were met. We evaluated the impact of genetic variation in the metformin arm only, and in a second model, we included a formal interaction test between the metformin and placebo arms. Models were adjusted for age at randomization, sex, and 10 ancestry principal components (PCs) for population structure; secondary models were also adjusted for waist circumference or BMI. We addressed common variants (minor allele frequency [MAF]>1%), both those directly genotyped and imputed. For imputed markers, we excluded variants that had an imputation quality score <0.7, as well as variants that did not pass a filter test ($2 \times MAF \times [1\text{-}MAF] \times n.events > 75$, where n.events is the number of events), as described in the GWASTools package (20). We evaluated for genome-wide significance, defined as $p<5\times10^{-8}$.

Change in quantitative traits

We utilized a multiple linear regression model to test allelic associations, assuming additive effects, with the change in quantitative traits relevant to metformin action at one year; traits included fasting glucose, 2-hour glucose after 75-g OGTT, fasting insulin, insulin sensitivity index (ISI), hemoglobin A1c (HbA1c), and weight. One-year change in each quantitative outcome was defined as one-year minus baseline value. Non-normally distributed traits were natural log transformed. To minimize the influence of outliers, winsorization was performed (at percentiles of 0.5 and 99.5 for normally distributed traits; percentiles of 1 and 99 for natural log transformed traits) (21). Analyses were adjusted for age, sex, first 10 ancestry PCs, and the baseline value of the trait. Similar to the diabetes incidence analyses, we evaluated the impact of genetic variation in the metformin arm only for each of the six outcomes, and in a second model, we tested for a gene-by-treatment interaction for the metformin and placebo arms. We filtered our results to a study-wide MAF>1% and imputation quality ≥0.7. To account for multiple testing and the possible correlation between quantitative traits, we calculated that there were 5.47 effective outcomes, using a previously described method (22), and we set an experiment-wide significance threshold of $p<9\times10^{-9}$ ($5\times10^{-8}/5.47$). For top findings that emerged, we performed an exploratory analysis that evaluated these findings in the intensive lifestyle treatment arm, in order to understand whether the variants act through a shared pathway between metformin and lifestyle intervention.

Replication

For each of the quantitative trait outcomes, we compiled a list of the top 500 genotyped and imputed variants with a MAF>1% and imputation score ≥0.7 for imputed markers, for a total of ~6000 variants (500 markers × 6 traits × 2 models). We explored their relevance in the Metformin Genetics (MetGen) Consortium, which has data from >10,000 participants from over

12 observational studies and clinical trials, in which a pharmacogenetic meta-analysis of metformin response was performed (10). We did additional filtering of variants, excluding those with a minor allele count <10 and with imputation score <0.7 from the Pharmacogenomics of Metformin (PMET) study, which represents the largest cohort comprising MetGen participants. We then calculated the number of effective variants after accounting for correlation between genetic variants (22) and determined the replication significance threshold. In addition, we attempted to replicate our top variant in the Million Veteran Program (MVP) and the Diabetes Multi-omic Investigation of Drug Response (DIAMOND). Full details of replication cohorts and the statistical models utilized in the replication analyses are described in the Supplemental Material and in Supplemental Table S2.

Data and Resource Availability

Genetic data generated and analyzed in the current study are available in dbGap (dbGaP Study Accession: phs000681.v2.p1, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000681.v2.p1). Phenotype data for the DPP is available through the NIDDK Data Repository. No applicable resources were generated or analyzed during the current study.

RESULTS

Baseline demographics

Table 1 describes the characteristics of the subset of DPP participants in the metformin and placebo arms in whom genome-wide genotyping was available. The mean baseline age was 50.7±10.4 years, 67% were women, and approximately 44% self-reported as non-white. Baseline measurements were consistent with a population at risk of developing T2D. There were no significant differences in baseline measurements by treatment arm. The event rate for diabetes

incidence is reported for each self-reported race/ethnicity group for both the metformin and placebo arms in Supplemental Table S3.

Association of genetic variation with diabetes incidence

In the metformin arm only (n=876, Supplemental Table S4), we performed association analysis of genetic variants with diabetes incidence during the main DPP study, which had a mean follow-up of 3.2 years. While we did not observe any genome-wide significant findings (Supplemental Figure S2A), we identified a genetic locus near *HIF1AN* that reached suggestive significance. The C effect allele of rs2489017 had a frequency of 0.59 in the entire DPP cohort and was notably more common in non-white populations (White: 0.49, African American: 0.83, Asian/Pacific Islander: 0.62, Hispanic/Latino: 0.62, American Indian: 0.78). Per copy of the variant, participants had a decreased risk of progressing to diabetes (HR 0.54, 95% CI 0.43-0.67, $p=9.2\times10^{-8}$). In the placebo group, the variant did not influence diabetes incidence ($p=0.72$), and we noted the presence of an interaction with treatment arm ($p(G\times T)=5.9\times10^{-5}$).

We subsequently combined the metformin and placebo arms (n=1,763, Supplemental Table S4) and utilized a Cox proportional hazards model to assess for a gene-by-treatment arm interaction for the outcome of diabetes incidence. We did not identify any genome-wide significant findings or variants that met a suggestive significance threshold of $p<1\times10^{-6}$ (Supplemental Figure S2B).

Association of genetic variation with one-year change in quantitative traits

With respect to the quantitative trait outcomes, we performed association analysis of genetic variants with one-year change in any of six quantitative traits (fasting glucose, 2-hour glucose after OGTT, fasting insulin, ISI, HbA1c, and weight) across two separate models, one examining the metformin arm only and another testing for a gene-by-treatment arm interaction.

Supplemental Table S3 presents the sample size for each model and quantitative trait outcome tested. Table 2 summarizes the 14 independent signals that met genome-wide significance ($p<5\times10^{-8}$) for at least one quantitative trait, four of which met experiment-wide significance ($p<9\times10^{-9}$): two for change in HbA1c (Supplemental Figure S3A-B) and two for change in weight (Supplemental Figure S4A-B). These results remained the same after additionally adjusting the analyses for BMI (data not shown). The effect sizes and *p*-values for the top findings, stratified by self-reported race/ethnicity, are also presented in Supplemental Table S5. Because we report the effect allele frequencies stratified by self-reported ethnicity/race, we compared these groupings with the ancestry groupings determined by principal components analysis (PCA) and illustrated that the self-identified groups tended to cluster tightly on the PCA plot (Supplemental Figure S5).

The top hit, rs144322333, which associated with change in HbA1c in the metformin treatment group, is a deletion polymorphism near *ENOSF1* (Figure 1) and was mainly found in African American participants with a MAF of 6%, whereas it is present in under 0.2% in White individuals (Table 2). Carriers have a 0.39% increase in one-year change in HbA1c per copy of the deletion, consistent with a worse response to metformin ($p=2.9\times10^{-12}$, Figure 2A). The proportion of variance in one-year change in HbA1c explained by the variant is 5.2%, compared with 10.8% for baseline HbA1c, the strongest predictor of this quantitative outcome in the model.

Since rs144322333 is more common in African American individuals, we performed a subgroup analysis and found that the effect was detectable in this group ($p=1.3\times10^{-6}$, Figure 2B). In fact, African individuals without any copies of the deletion experienced a *decrease* in HbA1c of 0.03% after a year of metformin, whereas those who carried one or two copies of the deletion

experienced *increases* of 0.2% and 0.6%. respectively. Finally, since the association was discovered in the metformin arm only, we also evaluated its impact in the placebo arm, in which there was no significant effect (p=0.32). The interaction with treatment arm was notably significant (p(G×T)=1.4×10^{-6}, Figure 2C), reflecting the differential impact of this genetic variant.

Another novel experiment-wide significant variant was rs145591055 in chromosome 5 near *OMSR* (Figure 3), a gene that encodes the Oncostatin M receptor (Table 2). Carriers of the G effect allele had a 7.6 kg greater decline in weight at one year following initiation of metformin treatment in adjusted analyses (p=3.2×10^{-10}, Figure 4A). The variant has an effect allele frequency of 10% in American Indians, calculated from the entire DPP study (Table 2). While a subgroup analysis in American Indians was limited by sample size (n=21, Figure 4B), we noted that non-carriers experienced a decrease in weight of 0.5 kg at one year, compared with a decrease of 8.6 kg in heterozygous carriers. We also evaluated the influence of this variant in the placebo arm and found that there was no significant effect on weight (per G allele beta=-0.69, p=0.50). There was again a significant gene-by-treatment interaction (p(G×T)=1.5×10^{-5}, Figure 4C).

In the gene × treatment models, we found that rs6838493, in chromosome 4 near *LINC01093* (Supplemental Figure S6), had a differential influence on change in HbA1c (p(G×T)=1.6×10^{-9}, Supplemental Figure S7), in which heterozygous carriers experienced a *rise* in HbA1c at one year in the placebo arm (beta=0.5, p=1.3×10^{-8}) in comparison to a *decline* in HbA1c in the metformin arm (beta=-0.1, p=0.03). The impact of rs148219263, located in chromosome 18 near *CDH20* (Supplemental Figure S8), on one-year change in weight was also different by treatment arm (p(G×T)=4.8×10^{-9}, Supplemental Figure S9). Here, in comparison to

their TT counterparts, TC carriers experienced weight loss in the placebo arm (beta=-7.0, p=2.4×10^{-8}) but did not achieve weight loss after receiving metformin (beta=3.9, p=4.0×10^{-3}).

In an exploratory analysis, we evaluated our top findings from Table 2 in the intensive lifestyle arm. Of the nine variants that emerged from the analysis completed in the metformin arm only, one variant (rs13401282) appeared to have a similar effect on ISI in the lifestyle arm as in the metformin arm (p=0.004, Supplementary Table S6); the other eight variants were non-significant in the lifestyle arm. For the five variants identified through the gene × treatment model that combined the metformin and placebo arms, we tested their relevance in a similar interaction model that combined the placebo and lifestyle arms. We found that four of the five variants demonstrated similar findings between the two gene × treatment models, as illustrated by the magnitude and direction of the beta estimates, as well as the p-values (Supplementary Table S6).

We searched our lead findings in the Genotype-Tissue Expression (GTex) project (https://gtexportal.org) (23) and eQTLconsortium (https://eqtlgen.org) (24) to evaluate for expression quantitative trait loci (eQTL) associations. We observed that rs17083791, which was associated with a greater decline in weight at one year following metformin, is a cis-eQTL associated with higher expression of *KIAA0825*, the nearest gene, in fibroblasts (normalized effect size [NES] 0.411, p=2.7×10^{-12}), subcutaneous adipose tissue (NES 0.285, p=7.9×10^{-7}), skeletal muscle (NES 0.241, p=5.0×10^{-5}), and visceral adipose tissue (NES 0.229, p=1.6×10^{-4}). In blood samples in eQTLconsortium, we noted that the variant was also a cis-eQTL for *MCTP1* (p=7.3×10^{-189}). We also evaluated the top variants in the pancreatic islet genotype tissue-expression resource (TIGER, http://tiger.bsc.es), which aggregates >500 human islet genomic datasets (25), but there were no notable findings. Finally, we searched top variants against

phenotypes associated with T2D using the Type 2 Diabetes Knowledge Portal (https://t2d.hugeamp.org). We found that the G allele of rs145591055, which was associated with greater one-year weight change on metformin, was marginally associated with a higher BMI-adjusted waist-hip ratio (beta=0.079, p=0.001).

Independent replication

We sought to replicate our pharmacogenetic findings in additional cohorts in which metformin response has been defined. As there are few cohorts of patients with pre-diabetes, our first approach was to explore the relevance of the variants in MetGen, in which the outcome was glycemic response, as measured by baseline minus minimum on-treatment HbA1c within 18 months after metformin initiation (10). Out of ~6000 variants, only 3050 had available genotype information in MetGen, resulting in 2610 effective variants; we thus set a significance threshold of $p<1.9\times10^{-5}$ (0.05/2610). After filtering, results were available in MetGen for only 6 of the 14 genome-wide significant variants from our DPP GWAS, none of which replicated (Table 3).

Because the study cohorts in MetGen were predominately European ancestry, we sought to replicate our top HbA1c finding in African American cohorts. We first examined 2733 participants with established T2D from MVP who self-reported as non-Hispanic African Americans and had received metformin monotherapy for up to 15 months. The top variant, rs144322333, which had a MAF of 7% and an imputation score of 0.89 in MVP, was not significantly associated with change in HbA1c (beta=0.05, p=0.15).

We also sought replication in DIAMOND, an observational population-based cohort with electronic medical record-linked clinical data on individuals with T2D on metformin monotherapy. We identified a subset of 471 individuals who self-reported as African American and were calculated to have a reliable exposure of at least 500 mg metformin daily as

monotherapy for treatment of T2D in the 120 days preceding the follow-up HbA1c measurement. In DIAMOND, rs144322333 had a MAF of 8% and had an imputation score of 0.98. Per copy of the deletion, carriers had a 0.13% reduction in HbA1c lowering over one year in DIAMOND, indicating a worse metformin response, which was similar in direction to that seen in the DPP, but smaller in magnitude and also not reaching statistical significance (p=0.19).

DISCUSSION

To date, GWAS have uncovered several genetic loci influencing the glycemic response to metformin, but these were largely performed in European populations with established T2D (9-11). When these variants were assessed in the DPP, the reported associations were not confirmed for the outcomes of diabetes incidence and other relevant physiologic parameters such as insulin sensitivity, fasting glucose, HbA1c, or oral disposition index (10, 12). One possible explanation for this inconsistency is that metformin response is defined differently in a disease cohort, in which the measure of response is the achievement of a HbA1c ≤7%, compared to a pre-diabetic cohort, in which diabetes incidence or quantitative traits are the studied outcomes. Furthermore, genetic variation may influence drug response differently in individuals with pre-diabetes, who likely have better pancreatic β-cell function.

Thus, we undertook a GWAS in the DPP, examining the outcomes of diabetes prevention as well as one-year change in six quantitative traits known to be affected by metformin therapy. The design of the DPP is unique for studying metformin pharmacogenetics, as it includes both a placebo and metformin arm, permitting the use of an interaction test to assess genetic variants that may have a differential response based on drug exposure. In assessing diabetes incidence, we did not identify any genome-wide significant variants. We acknowledge that the relatively

modest sample size of the DPP with a limited range of impaired glucose tolerance may have restricted our ability to detect an interaction effect. However, we did observe an association that met a suggestive threshold that could be investigated in the future or meta-analyzed with other cohorts.

Our quantitative trait analyses yielded several genome-wide significant associations, after correcting for multiple testing. *ENOSF1*, the closest gene to the top variant rs144322333, encodes the enolase superfamily member 1, a mitochondrial enzyme involved in the catabolism of L-fucose, a sugar found on cellular glycoproteins. Interestingly, *ENOSF1* is next to a kinase, YES1, which has been shown to regulate the activity of organic cation transporters OCT1 (*SLC22A1*) and OCT2 (*SLC22A2*) that play a critical role in metformin disposition and elimination in the kidney and liver, respectively (26, 27). Moreover, LINC01093 near rs186681623, has been shown to be related to several liver phenotypes, including fibrosis (28) and hepatocellular carcinoma (29), and may play a role in modulating metformin's action on the liver.

Many of our genome-wide significant findings were ancestry-specific (rs144322333 near *ENOSF1*, rs145591055 near *OSMR,* rs6838493 near *LINC01093*). We further illustrate this through subgroup analyses and demonstrate the persistent effect of top variants of interest after stratifying by self-reported ethnicity/race. Based on these stratified analyses, it appeared that the ancestry-specificity of our findings was mainly driven by allele frequency, since similar trends were observed across subgroups. We acknowledge that this may have been a consequence of conducting our analyses in a multi-ancestry cohort, which could have biased our results toward variants with concordant effects across subgroups. One limitation of this approach is that variants with opposite effects in different subgroups could be potentially missed. We also estimated that

our top genetic finding (rs144322333) explained 5.2% of the variance in one-year change in HbA1c, which was half of that of baseline HbA1c, supporting the notion that the contribution of genetics is modest compared to a traditional predictor of drug response.

In the metformin arm of the DPP, the average weight loss was 2.1 kg and was relatively stable throughout the study follow-up period (15). In comparison, the association of rs145591055 (near *OSMR*) variation with one-year change in weight was 7.6 kg, over three times greater in magnitude. Individuals who carry this variant, which has an effect allele frequency of 10% in American Indians, appear to derive a significant weight-loss benefit from metformin, which is especially relevant given the high incidence and prevalence of T2D in this population (30, 31). Moreover, the G allele was reported in the T2D Knowledge Portal to be marginally associated with a higher BMI-adjusted waist-hip ratio, also suggesting that this is a group of individuals that would benefit from metformin. However, the underlying mechanism responsible for this weight effect is not clear and requires further investigation. We also evaluated our top findings in the GTex project, the eQTLconsortium, and the TIGER portal. Interestingly, we observed that the variant rs17083791 was a *cis*-eQTL for the nearest gene *KIAA0825* in several relevant tissues, suggesting that changes in *KIAA0825* gene expression may modulate metformin-related weight loss. The same variant was also a strong *cis*-eQTL for *MCTP1*, which codes for a protein involved in calcium ion binding, though the significance of this finding is unknown. While these genomic resources can be useful for functional follow-up of variants, the majority of participants are of European ancestry and thus, their utility is limited for variants that are common only in specific populations.

In an exploratory analysis, we evaluated the top findings from the quantitative trait analysis using the intensive lifestyle arm of the DPP. There were five variants for which the

association with the quantitative trait was similar in the lifestyle arm as in the metformin arm. Prior studies have suggested that the intensive lifestyle intervention of the DPP worked similarly to metformin in improving insulin sensitivity (32). Therefore, it is possible that the mechanism of action of these variants may be shared between metformin and lifestyle. On the other hand, the remaining nine variants that met genome-wide significance were non-significant when evaluated in the lifestyle arm, illustrating they may be specific to metformin response.

Finally, we took a comprehensive approach to replicate findings that emerged from our GWAS, though this was a major challenge due to the lack of a suitable replication cohort of individuals with pre-diabetes who received longitudinal metformin exposure. We evaluated approximately the top 500 variants for each quantitative trait analyses in MetGen, the largest meta-analysis examining glycemic response to metformin in T2D (10). Our efforts did not yield replication, but this could be explained by the composition of MetGen as previously discussed. We subsequently identified two African American cohorts (MVP and DIAMOND) for replication of our top variant rs144322333. Though the study populations again largely reflected individuals with T2D, we constructed our models to reflect a similar length of metformin exposure as in the DPP and included similar covariates as in our quantitative trait analyses. Despite imputation on different reference panels in these cohorts, we reassuringly observed similar MAFs for rs14432233; however, neither replication produced a nominally significant result. We acknowledge the possibility that given the low study-wide MAF of our top findings and the small sample size, our findings could represent false positive results.

Our work underscores the importance of conducting multi-ancestry pharmacogenetic studies to improve upon treatment algorithms across diverse populations in the future. A prior study based on pharmacy fill data reported that African American adults appear to have a better

glycemic response to metformin compared with European Americans (33), and the TODAY study showed that African American youth with T2D on metformin had poorer glycemic response to metformin (7). Though these studies were based on self-reported race/ethnicity information, we demonstrated in the DPP that there is clustering of the self-reported ethnicity/race groups along the genetic ancestry PCs. Thus, the identification of variants in our study that are more prevalent in African American populations suggests that there may be a heritable component to these differences in response to metformin.

Strengths of our study include the diverse composition of the DPP, which includes nearly 45% non-white individuals by self-report and represents a cohort of individuals at a higher risk of developing T2D in the US. Participants were well-characterized under standardized clinical trial protocols, and the genetic data was obtained using the highest standards for genotyping and imputation at the time. As described, limitations of our study are the small sample size, which is a frequent problem in pharmacogenomic GWAS (34) and the lack of replication of our findings in patients with diabetes, though this may reflect differences in genetic modulation of metformin response between the pre-diabetes and diabetes state. However, we were exhaustive in our attempts to identify independent cohorts, including the largest African American cohorts with diabetes known to us. Future directions should include additional follow-up and meta-analyses in additional independent ancestry-specific cohorts as they arise, to advance work in this area.

In conclusion, we have identified several genome-wide significant associations with metformin response, as measured by change in quantitative traits, in a well-defined multi-ancestry clinical trial cohort. We believe that we have generated a valuable resource that can be utilized for future genetic investigation and to gain insight into the genetic underpinnings of interindividual differences in metformin response in a population at risk for developing T2D.

Finally, we illustrate the peremptory need to generate suitable pharmacogenomic and transcriptomic resources in diverse populations.

ACKNOWLEDGEMENTS

Funding. The DPP Research Group gratefully acknowledges the commitment and dedication of the participants of the DPP and DPPOS. Research reported in this publication was supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) of the National Institutes of Health (NIH) under award numbers U01 DK048489, U01 DK048339, U01 DK048377, U01 DK048349, U01 DK048381, U01 DK048468, U01 DK048434, U01 DK048485, U01 DK048375, U01 DK048514, U01 DK048437, U01 DK048413, U01 DK048411, U01 DK048406, U01 DK048380, U01 DK048397, U01 DK048412, U01 DK048404, U01 DK048387, U01 DK048407, U01 DK048443, and U01 DK048400, by providing funding during DPP and DPPOS to the clinical centers and the Coordinating Center for the design and conduct of the study, and collection, management, analysis, and interpretation of the data. Funding was also provided by the National Institute of Child Health and Human Development, the National Institute on Aging, the National Eye Institute, the National Heart Lung and Blood Institute, the National Cancer Institute, the Office of Research on Women's Health, the National Institute on Minority Health and Health Disparities, the Centers for Disease Control and Prevention, and the American Diabetes Association. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The Southwestern American Indian Centers were supported directly by the

NIDDK, including its Intramural Research Program, and the Indian Health Service. The General Clinical Research Center Program, National Center for Research Resources, and the Department of Veterans Affairs supported data collection at many of the clinical centers. Merck KGaA provided medication for DPPOS. DPP/DPPOS have also received donated materials, equipment, or medicines for concomitant conditions from Bristol-Myers Squibb, Parke-Davis, and LifeScan Inc., Health O Meter, Hoechst Marion Roussel, Inc., Merck-Medco Managed Care, Inc., Merck and Co., Nike Sports Marketing, Slim Fast Foods Co., and Quaker Oats Co. McKesson BioServices Corp., Matthews Media Group, Inc., and the Henry M. Jackson Foundation provided support services under subcontract with the Coordinating Center. This work was also supported by R01DK072041 to JCF and KAJ. JHL received individual support from T32DK007028. SS was funded by K23DK120932. AG was supported by 1K01DK120631 when this work was performed. SEK was supported in part by the Department of Veterans Affairs. The sponsor of this study was represented on the Steering Committee and played a part in study design, how the study was done, and publication. All authors in the writing group had access to all data. The opinions expressed are those of the study group and do not necessarily reflect the views of the funding agencies. A complete list of Centers, investigators, and staff can be found in the Appendix.

Duality of Interest. The authors have no conflicts of interest to report.

Author Contributions. JHL, JAP, KAJ, SS, TIP, and JCF contributed to the conception and design the study. Genotyping, quality control, and imputation of the genetic data were performed by LC, MH, and JMM. GWAS analyses were performed by JAP, KAJ, and TIP. JHL, JAP, KAJ, SS, JNT, JMM, QP, PWF, RLH, SEK, WCK, TIP, and JCF contributed to data interpretation. AYD, SWY, KMG, ERP, AG, AMH, SX, and LKW contributed to the replication analyses

described in the manuscript. JHL, JAP, KAJ, SS, and JCF contributed to manuscript writing, with editing by all authors and approval of submission of the final version. JCF is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Portions of this study were previously presented as an oral presentation at the 81st Virtual Scientific Sessions of the American Diabetes Association, 25-29 June 2021.

REFERENCES

1. 9. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes-2021. *Diabetes Care* 2021;44:S111-s124
2. Kane SP. Metformin Hydrochloride CDD, Version 21.2. ClinCalc: https://clincalc.com/DrugStats/Drugs/MetforminHydrochloride. Updated July 3, 2021. Accessed August 11, 2021.
3. Davies MJ, D'Alessio DA, Fradkin J, et al. Management of hyperglycaemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetologia* 2018;61:2461-2498
4. Chawla R, Madhu SV, Makkar BM, et al. RSSDI-ESI Clinical Practice Recommendations for the Management of Type 2 Diabetes Mellitus 2020. *Indian J Endocrinol Metab* 2020;24:1-122
5. Jia W, Weng J, Zhu D, et al. Standards of medical care for type 2 diabetes in China 2019. *Diabetes Metab Res Rev* 2019;35:e3158
6. Kahn SE, Haffner SM, Heise MA, et al. Glycemic Durability of Rosiglitazone, Metformin, or Glyburide Monotherapy. *N Engl J Med* 2006;355:2427-2443
7. Zeitler P, Hirst K, Pyle L, et al. A clinical trial to maintain glycemic control in youth with type 2 diabetes. *N Engl J Med* 2012;366:2247-2256
8. Zhou K, Donnelly L, Yang J, et al. Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. *Lancet Diabetes Endocrinol* 2014;2:481-487
9. Zhou K, Bellenguez C, Spencer CC, et al. Common variants near *ATM* are associated with glycemic response to metformin in type 2 diabetes. *Nat Genet* 2011;43:117-120
10. Zhou K, Yee SW, Seiser EL, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet* 2016;48:1055-1059
11. Rotroff DM, Yee SW, Zhou K, et al. Genetic Variants in CPA6 and PRPF31 Are Associated With Variation in Response to Metformin in Individuals With Type 2 Diabetes. *Diabetes* 2018;67:1428-1440
12. Florez JC, Jablonski KA, Taylor A, et al. The C allele of ATM rs11212617 does not associate with metformin response in the Diabetes Prevention Program. *Diabetes Care* 2012;35:1864-1867
13. The Diabetes Prevention Program Research Group. The Diabetes Prevention Program. Design and methods for a clinical trial in the prevention of type 2 diabetes. *Diabetes Care* 1999;22:623-634
14. The Diabetes Prevention Program Research Group. The Diabetes Prevention Program: baseline characteristics of the randomized cohort. The Diabetes Prevention Program Research Group. *Diabetes Care* 2000;23:1619-1629
15. Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346:393-403
16. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods* 2012;9:179-181
17. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009;5:e1000529
18. Sánchez F, Bonàs S, Guindo M, et al. GWImp-COMPSs: an integrated framework for large-scale genome-wide imputation and association testing on parallel computing platforms. Available at http://cg.bsc.es/guidance/.

19. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015;526:68-74
20. Gogarten SM, Bhangale T, Conomos MP, et al. GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. *Bioinformatics* 2012;28:3329-3331
21. Dixon WJ. Simplified Estimation from Censored Normal Samples. *The Annals of Mathematical Statistics* 1960;31:385-391
22. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)* 2005;95:221-227
23. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 2020;369:1318-1330
24. Võsa U, Claringbould A, Westra HJ, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet* 2021;53:1300-1310
25. Alonso L, Piron A, Morán I, et al. TIGER: The gene expression regulatory variation landscape of human pancreatic islets. *Cell Rep* 2021;37:109807
26. Sprowl JA, Ong SS, Gibson AA, et al. A phosphotyrosine switch regulates organic cation transporters. *Nat Commun* 2016;7:10880
27. Uddin ME, Garrison DA, Kim K, et al. Influence of YES1 Kinase and Tyrosine Phosphorylation on the Activity of OCT1. *Front Pharmacol* 2021;12
28. Tang Y, Ma N, Luo H, Chen S, Yu F. Downregulated long non-coding RNA LINC01093 in liver fibrosis promotes hepatocyte apoptosis via increasing ubiquitination of SIRT1. *J Biochem* 2020;167:525-534
29. Li L, Huang K, Lu Z, et al. Bioinformatics analysis of LINC01554 and its co-expressed genes in hepatocellular carcinoma. *Oncol Rep* 2020;44:2185-2197
30. Acton KJ, Burrows NR, Moore K, et al. Trends in diabetes prevalence among American Indian and Alaska native children, adolescents, and young adults. *Am J Public Health* 2002;92:1485-1490
31. Pavkov ME, Hanson RL, Knowler WC, et al. Changing patterns of type 2 diabetes incidence among Pima Indians. *Diabetes Care* 2007;30:1758-1763
32. Kitabchi AE, Temprosa M, Knowler WC, et al. Role of insulin secretion and sensitivity in the evolution of type 2 diabetes in the diabetes prevention program: effects of lifestyle intervention and metformin. *Diabetes* 2005;54:2404-2414
33. Williams LK, Padhukasahasram B, Ahmedani BK, et al. Differing effects of metformin on glycemic control by race-ethnicity. *J Clin Endocrinol Metab* 2014;99:3160-3168
34. McInnes G, Yee SW, Pershad Y, Altman RB. Genomewide Association Studies in Pharmacogenomics. *Clin Pharmacol Ther* 2021;110:637-648

TABLES

Table 1. Baseline characteristics of 1,768 participants in the DPP in the metformin (MET) and placebo (PBO) arms with genome-wide genotyping.

Baseline Characteristic[*]	**Total**	**PBO**	**MET**	***p*-value**[**]
n	1768	888	880	
Age in years	50.7±10.4	50.6±10.4	50.9±10.3	0.538
Men	578 (32.7)	281 (31.6)	297 (33.8)	0.345
Women	1190 (67.3)	607 (68.4)	583 (66.3)	
Self-reported Race/Ethnicity				
White	997 (56.4)	490 (55.2)	507 (57.6)	0.746
Black	364 (20.6)	186 (20.9)	178 (20.2)	
Hispanic/Latino	290 (16.4)	147 (16.6)	143 (16.3)	
Asian/Pacific Islander	69 (3.9)	38 (4.3)	31 (3.5)	
American Indian	48 (2.7)	27 (3.0)	21 (2.4)	
Weight (kg)	94.7±19.8	94.9±20.0	94.5±19.6	0.692
HbA1c (%)[†]	5.917±0.501	5.920±0.495	5.915±0.507	0.832
Fasting insulin (pmol/L)[#†]	167 (111, 236)	167 (111, 229)	167 (111, 236)	0.595[***]
Fasting glucose (mmol/L)	5.940±0.462	5.956±0.466	5.923±0.458	0.138
ISI[#†]	0.161 (0.112, 0.235)	0.161 (0.113, 0.236)	0.161 (0.110, 0.234)	0.762[***]

[*] Data are Means±SD for continuous variables and n (%) for categorical variables unless noted otherwise
[**] *p*-values from F test for continuous variables and Chi-square for categorical variables
[***] *p*-value from Kruskal-Wallis test
[#] Median (25th, 75th percentile)
†n=1766 for Total; n=887 for PBO; n=879 for MET

Table 2. 14 independent genome-wide significant findings ($p<5\times10^{-8}$) for one-year change in quantitative traits in the DPP.

rsid	Chr	Position*	Nearest gene	EA	NEA	Trait	N	Beta (95% CI)	*p*	Effect allele frequency					
										All†	AfrAm‡	AsnPI‡	Hisp‡	AI‡	White‡
Metformin only model															
rs144322333	18	705550	*ENOSF1*	C	CTGTT	HbA1c, %	818	0.39 (0.28, 0.50)	2.9×10^{-12}	0.013	0.065	0.001	0.009	0.0001	0.002
rs145591055	5	38849463	*OSMR*	G	A	Weight, kg	829	-7.55 (-9.87, -5.22)	3.2×10^{-10}	0.014	0.001	0.018	0.064	0.103	0.007
rs13401282§	2	207810690	*CPO*	A	T	ISI, ln	808	0.44 (0.29, 0.59)	1.7×10^{-8}	0.029	0.119	0.003	0.015	0.006	0.002
rs186681623	13	81546608	*LINC00377*	C	T	Weight, kg	829	-2.66 (-3.58, -1.74)	2.0×10^{-8}	0.068	0.121	0.074	0.122	0.161	0.036
rs17083791	5	93863813	*KIAA0825*	G	A	Weight, kg	829	-1.85 (-2.50, -1.20)	3.3×10^{-8}	0.149	0.052	0.194	0.210	0.367	0.137
rs9931871	16	19928315	*GPRC5B*	G	A	Fasting glucose, mmol/L	821	0.68 (0.44, 0.93)	3.5×10^{-8}	0.011	0.049	<0.0001	0.003	<0.0001	<0.0001
rs549305231	6	19497504	*LOC101928519*	A	G	HbA1c, %	818	0.35 (0.23, 0.48)	3.9×10^{-8}	0.012	0.046	<0.0001	0.005	<0.0001	0.0003
rs73944532	2	104691634	*LOC100287010*	A	G	Fasting insulin, ln	802	-0.47 (-0.63, -0.30)	4.1×10^{-8}	0.020	0.075	0.005	0.007	0.001	0.002
Gene × Treatment model‖															
rs6838493	4	185879789	*LINC01093*	A	T	HbA1c, %	1621	-0.66 (-0.87, -0.44)	1.6×10^{-9}	0.011	0.041	0.004	0.005	0.004	0.002
rs148219263	18	58808522	*CDH20*	C	T	Weight, kg	1673	10.73 (7.15, 14.30)	4.8×10^{-9}	0.012	0.007	0.011	0.011	0.0009	0.017
rs75147163	14	57626769	*EXOC5*	A	G	Fasting glucose, mmol/L	1633	-0.73 (-0.98, -0.48)	1.4×10^{-8}	0.022	0.053	0.018	0.022	0.0002	0.015
rs78075715	4	6613716	*MAN2B2*	C	T	HbA1c, %	1621	-0.51 (-0.68, -0.33)	1.4×10^{-8}	0.013	0.063	<0.0001	0.002	<0.0001	0.0004
rs12314996	12	24567467	*SOX5*	A	G	HbA1c, %	1621	-0.47 (-0.63, -0.30)	3.9×10^{-8}	0.016	0.062	<0.0001	0.007	0.001	0.002

rs143203347	7	107384199	*CBLL1*	C	G	Fasting glucose, mmol/L	1633	-0.92 (-1.24, -0.59)	4.1×10^{-8}	0.015	0.062	0.0002	0.011	<0.0001	<0.0001

EA=Effect allele; NEA=Non-effect allele; AfrAm=African American; AsnPI=Asian/Pacific Islander; Hisp=Hispanic/Latino; AI=American Indian; ISI=insulin sensitivity index. *GRCh37 assembly. †This is the effect allele frequency for the number of participants in the model (N) that was calculated based on imputation. ‡Effect allele frequency breakdown by self-reported race/ethnicity is for all 3,168 participants with genome-wide genotyping, calculated based on imputation. §rs13401282 is also associated with fasting insulin (beta [95% CI] = -0.39 [-0.53, -0.25], $p=4.65\times10^{-8}$). G×T models contain an exposure term for the treatment arm and a G×T interaction term. Results are filtered to $p(G\times T)<5\times10^{-8}$ and $p(SNP)<1\times10^{-4}$. ‖Beta estimates and confidence intervals (CI) are reported for the interaction term rather than the main effect of metformin.

Table 3. Attempt at replication of genome-wide significant findings from the DPP in the Metformin Genetics Consortium.

DPP Model	Trait	rsid	DPP Beta*	MetGen (n)	MetGen Beta*	MetGen *p*-value
MET	Weight, kg	rs186681623	-2.66 (greater weight loss)	10519	-0.05 (decreased HbA1c reduction)	0.10
MET	Weight, kg	rs17083791	-1.85 (greater weight loss)	12578	0.004 (greater HbA1c reduction)	0.79
MET	Fasting insulin, ln	rs73944532	-0.47 (greater decrease in fasting insulin)	7048	0.03 (greater HbA1c reduction)	0.84
G×T	Weight, kg	rs148219263	10.73 (reduced weight loss)	10519	-0.04 (decreased HbA1c reduction)	0.39
G×T	Fasting glucose, mmol/L	rs75147163	-0.73 (greater decrease in fasting glucose)	10288	-0.02 (decreased HbA1c reduction)	0.70
G×T	HbA1c, %	rs12314996	-0.47 (greater decrease in HbA1c)	7048	0.19 (greater HbA1c reduction)	0.24

*In the DPP, the beta estimates represent one-year change in the quantitative trait as calculated by follow-up value minus baseline value. In MetGen, the outcome was change in HbA1c (%), defined as baseline value minus follow-up value within 18 months (will have opposite signs of beta estimates for change in HbA1c compared to that in the DPP). For each analysis and cohort, the impact of the variant on the metformin response outcome is indicated in parentheses.

FIGURE LEGEND

Figure 1. Regional association plot of rs144322333 for one-year change in HbA1c. The red line indicates the experiment-wide significance threshold of $p<9\times10^{-9}$ and the blue line indicates a suggestive significance threshold of $p<1\times10^{-6}$.

Figure 2. (A) Box plot illustrating the mean change in HbA1c (one-year minus baseline) by rs144322333 genotype in the metformin only arm, n=818. (B) Stratified analyses by self-reported race/ethnicity. *P*-value was calculated for the subgroup of African American individuals, n=166. (C) Comparison of the influence of rs144322333 genotype on the mean change in HbA1c in all subjects in the metformin (MET, n=818) and placebo (PBO, n=803) arms. The interaction *p*-value is reported. For the purposes of generating the box plots, fractional alleles were converted to "hard calls" by rounding to the nearest integer.

Figure 3. Regional association plot of rs145591055 for one-year change in weight. The red line indicates the experiment-wide significance threshold of $p<9\times10^{-9}$ and the blue line indicates a suggestive significance threshold of $p<1\times10^{-6}$.

Figure 4. (A) Box plot illustrating the mean change in weight (one-year minus baseline) by rs145591055 genotype in the metformin only arm, n=829. (B) Box plot stratified by self-reported race/ethnicity. (C) Comparison of the influence of rs144322333 genotype on the mean change in weight in all subjects in the metformin (MET, n=829) and placebo (PBO, n=844) arms. The interaction *p*-value is reported. For the purposes of generating the box plots, fractional alleles were converted to "hard calls" by rounding to the nearest integer.

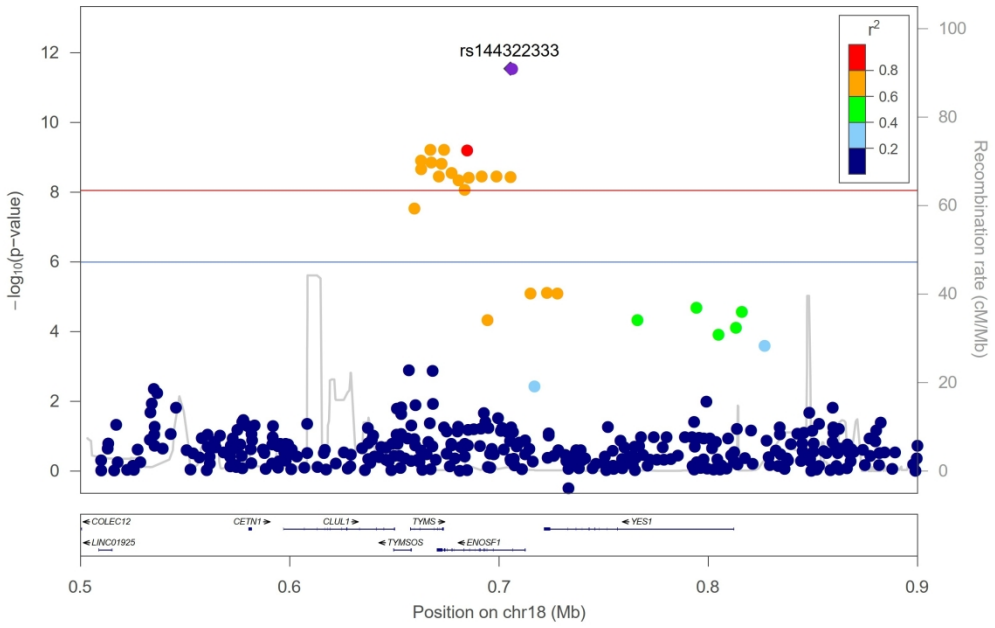


Figure 1. Regional association plot of rs144322333 for one-year change in HbA1c. The red line indicates the experiment-wide significance threshold of p<9×10-9 and the blue line indicates a suggestive significance threshold of p<1×10-6.

242x152mm (330 x 330 DPI)

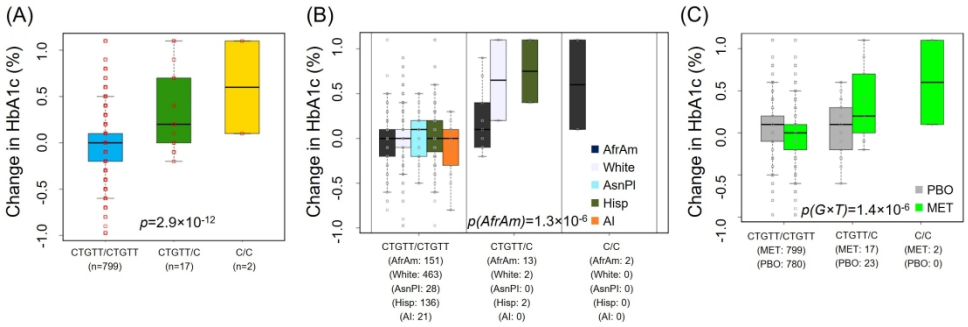


Figure 2. (A) Box plot illustrating the mean change in HbA1c (one-year minus baseline) by rs144322333 genotype in the metformin only arm, n=818. (B) Stratified analyses by self-reported race/ethnicity. P-value was calculated for the subgroup of African American individuals, n=166. (C) Comparison of the influence of rs144322333 genotype on the mean change in HbA1c in all subjects in the metformin (MET, n=818) and placebo (PBO, n=803) arms. The interaction p-value is reported. For the purposes of generating the box plots, fractional alleles were converted to "hard calls" by rounding to the nearest integer.

238x79mm (330 x 330 DPI)

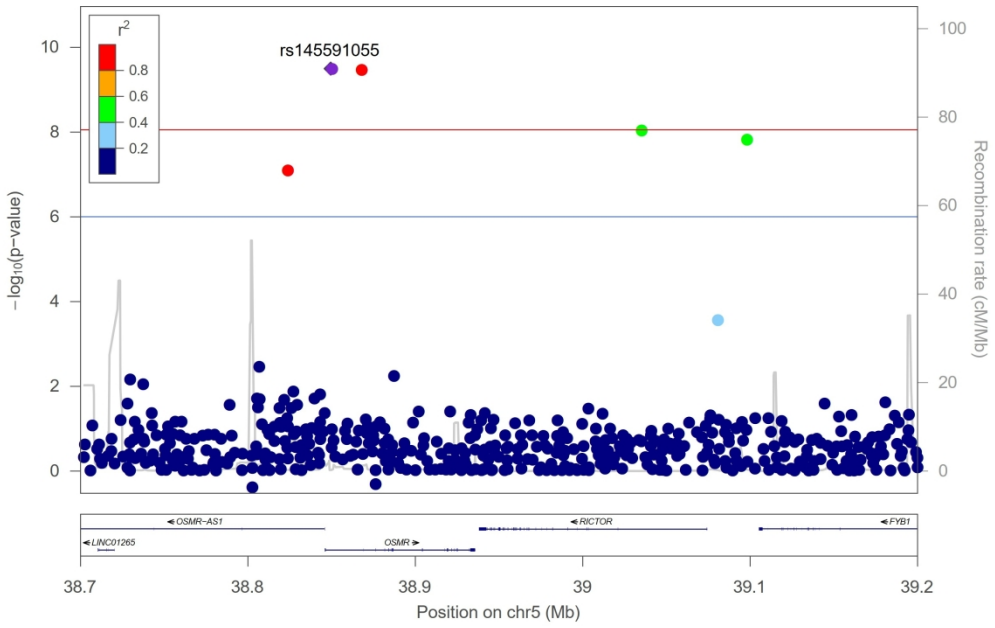


Figure 3. Regional association plot of rs145591055 for one-year change in weight. The red line indicates the experiment-wide significance threshold of p<9×10-9 and the blue line indicates a suggestive significance threshold of p<1×10-6.

242x152mm (330 x 330 DPI)

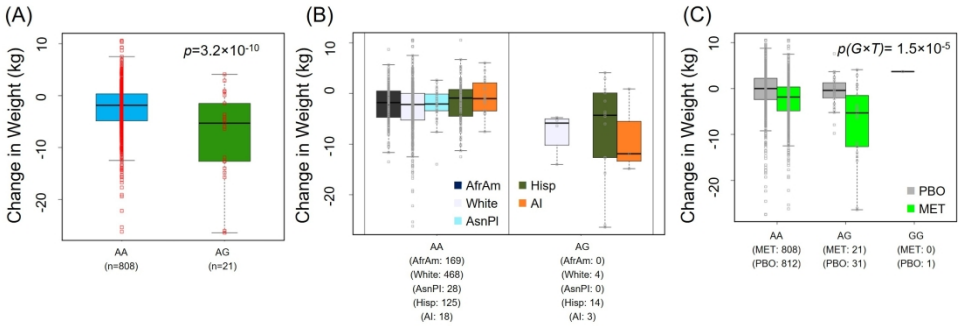


Figure 4. (A) Box plot illustrating the mean change in weight (one-year minus baseline) by rs145591055 genotype in the metformin only arm, n=829. (B) Box plot stratified by self-reported race/ethnicity. (C) Comparison of the influence of rs144322333 genotype on the mean change in weight in all subjects in the metformin (MET, n=829) and placebo (PBO, n=844) arms. The interaction p-value is reported. For the purposes of generating the box plots, fractional alleles were converted to "hard calls" by rounding to the nearest integer.

237x79mm (330 x 330 DPI)

DPP, DPPOS 1, DPPOS 2 & DPPOS 3A Research Group

(1996-2021)

Pennington Biomedical Research Center (Baton Rouge, LA)
George A. Bray, MD*
Kishore M. Gadde, MD*
Iris W. Culbert, BSN, RN, CCRC**
Jennifer Arceneaux RN, BSN**
Annie Chatellier, RN, CCRC**
Amber Dragg RD, LDN**
Catherine M. Champagne, PhD, RD
Crystal Duncan, LPN
Barbara Eberhardt, RD, LDN
Frank Greenway, MD
Fonda G. Guillory, LPN
April A. Herbert, RD
Michael L. Jeffirs, LPN
Betty M. Kennedy, MPA
Erma Levy, RD
Monica Lockett, LPN
Jennifer C. Lovejoy, PhD
Laura H. Morris, BS
Lee E. Melancon, BA, BS
Donna H. Ryan, MD
Deborah A. Sanford, LPN
Kenneth G. Smith, BS, MT
Lisa L. Smith, BS
Julia A. St.Amant, RTR
Richard T. Tulley, PhD
Paula C. Vicknair, MS, RD
Donald Williamson, PhD
Jeffery J. Zachwieja, PhD

University of Chicago (Chicago, IL)
Kenneth S. Polonsky, MD*
Janet Tobian, MD, PhD*
David A. Ehrmann, MD*
Margaret J. Matulik, RN, BSN**
Karla A. Temple, PhD, RDN, LDN**
Bart Clark, MD
Kirsten Czech, MS
Catherine DeSandre, BA
Brittnie Dotson, MS
Ruthanne Hilbrich, RD
Wylie McNabb, EdD
Ann R. Semenske, MS, RD

Jefferson Medical College (Philadelphia, PA)
Jose F. Caro, MD*
Kevin Furlong, DO*
Barry J. Goldstein, MD, PhD*
Pamela G. Watson, RN, ScD*
Kellie A. Smith, RN, MSN**
Jewel Mendoza, RN, BSN**
Marsha Simmons, CCRP**
Wendi Wildman, RN**
Renee Liberoni, MPH
John Spandorfer, MD
Constance Pepe, MS, RD

University of Miami (Miami, FL)
Richard P. Donahue, PhD*
Ronald B. Goldberg, MD*
Ronald Prineas, MD, PhD*
Jeanette Calles, MSEd**
Anna Giannella, RD, MS**
Patricia Rowe, MPA**
Juliet Sanguily, RN**
Paul Cassanova-Romero, MD
Sumaya Castillo-Florez, MPH
Hermes J. Florez, MD
Rajesh Garg, MD
Lascelles Kirby, MS
Olga Lara
Carmen Larreal
Valerie McLymont, RN
Jadell Mendez
Arlette Perry, PhD
Patrice Saab, PhD
Bertha Veciana

The University of Texas Health Science Center (San Antonio, TX)
Steven M. Haffner, MD, MPH*
Helen P. Hazuda, PhD*
Maria G. Montez, RN, MSHP, CDE**
Kathy Hattaway, RD, MS
Juan Isaac, RN, BSN
Carlos Lorenzo, MD, PhD
Arlene Martinez, RN, BSN, CDE
Monica Salazar
Tatiana Walker, RD, MS, CDE

University of Colorado (Denver, CO)
Dana Dabelea, MD, PhD*
Richard F. Hamman, MD, DrPH*
Patricia V. Nash, MS**
Sheila C. Steinke, MS**
Lisa Testaverde, MS**
Jennifer Truong, MPH**
Denise R. Anderson, RN, BSN
Larry B. Ballonoff, MD
Alexis Bouffard, MA, RN, BSN
Brian Bucca OD, FAOO
B. Ned Calonge, MD, MPH
Lynne Delve
Martha Farago, RN
James O. Hill, PhD
Shelley R. Hoyer, BS

* denotes Principal Investigator
** denotes Program Coordinator

DPP, DPPOS 1, DPPOS 2 & DPPOS 3A Research Group
(1996-2021)

Tonya Jenkins, RD, CDE
Bonnie T. Jortberg, MS, RD, CDE
Dione Lenz, RN, BSN, CDE
Marsha Miller, MS, RD
Thomas Nilan, BS
Leigh Perreault, MD
David W. Price, MD
Judith G. Regensteiner, PhD
Emily B. Schroeder, MD
Helen Seagle, MS, RD
Carissa M. Smith, BS
Brent VanDorsten, PhD

Joslin Diabetes Center (Boston, MA)
Edward S. Horton, MD*
Medha Munshi, MD*
Kathleen E. Lawton, RN**
Sharon D. Jackson,CCRC,MS, RD, CDE**
Catherine S. Poirier, RN, BSN**
Kati Swift, RN, BSN**
Ronald A. Arky, MD
Marybeth Bryant
Jacqueline P. Burke, BSN
Enrique Caballero, MD
Karen M. Callaphan, BA
Barbara Fargnoli, RD
Therese Franklin
Om P. Ganda, MD
Ashley Guidi, BS
Mathew Guido, BA
Alan M. Jacobsen, MD
Lyn M. Kula, RD
Margaret Kocal, RN, CDE
Lori Lambert, MS, RD, LD
Kathleen E. Lawton, RN
Sarah Ledbury, Med, RD
Maureen A. Malloy, BS
Roeland J.W. Middelbeek, MD
Maryanne Nicosia, MS, RD
Cathryn F. Oldmixon, RN
Jocelyn Pan, BS, MPH
Marizel Quitingon
Riley Rainville, BS
Stacy Rubtchinsky, BS
Ellen W. Seely, MD
Jessica Sansoucy, BS
Dana Schweizer, BSN
Donald Simonson, MD
Fannie Smith, MD
Caren G. Solomon, MD, MPH
Jeanne Spellman, RD
James Warram, MD

VA Puget Sound Health Care System and University of Washington (Seattle, WA)
Steven E. Kahn, MB, ChB*
Basma Fattaleh, BA **
Brenda K. Montgomery, RN, BSN, CDE**
Celeste Colegrove, BS
Wilfred Fujimoto, MD
Robert H. Knopp, MD
Edward W. Lipkin, MD
Michelle Marr, BA
Ivy Morgan-Taggart
Anne Murillo, BS
Kayla O'Neal, BS
Dace Trence, MD
Lonnese Taylor, RN, BS
April Thomas, RD, MPH, CDE
Elaine C. Tsai, MD, MPH

University of Tennessee (Memphis, TN)
Samuel Dagogo-Jack, MD, MSc, FRCP, FACP*
Abbas E. Kitabchi, PhD, MD, FACP*
Mary E. Murphy, RN, MS, CDE, MBA**
Laura Taylor, RN, BSN, CDE**
Jennifer Dolgoff, RN, BSN**
William B. Applegate, MD, MPH
Michael Bryer-Ash, MD
Debra Clark, LPN
Sandra L. Frieson, RN
Uzoma Ibebuogu, MD
Raed Imseis, MD
Helen Lambeth, RN, BSN
Lynne C. Lichtermann, RN, BSN
Hooman Oktaei, MD
Harriet Ricks
Lily M.K. Rutledge, RN, BSN
Amy R. Sherman, RD, LD
Clara M. Smith, RD, MHP, LDN
Judith E. Soberman, MD
Beverly Williams-Cleaves, MD
Avnisha Patel, MLT
Ebenezer A. Nyenwe, MD, FACP
Ethel Faye Hampton, R.N.

Northwestern University's Feinberg School of Medicine (Chicago, IL)
Boyd E. Metzger, MD*
Mark E. Molitch, MD*
Mariana K. Johnson, MS, RN**
Daphne T. Adelman, MBA, RN
Catherine Behrends
Michelle Cook, MS
Marian Fitzgibbon, PhD
Mimi M. Giles, MS, RD

* denotes Principal Investigator
** denotes Program Coordinator

DPP, DPPOS 1, DPPOS 2 & DPPOS 3A Research Group

(1996-2021)

Deloris Heard, MA
Cheryl K.H. Johnson, MS, RN
Diane Larsen, BS
Anne Lowe, BS
Megan Lyman, BS
David McPherson, MD
Samsam C. Penn, BA
Thomas Pitts, MD
Renee Reinhart, RN, MS
Susan Roston, RN, RD
Pamela A. Schinleber, RN, MS
Amisha Wallia, MD

Massachusetts General Hospital (Boston, MA)

David M. Nathan, MD*
Charles McKitrick, BSN**
Heather Turgeon, BSN**
Mary Larkin, MSN**
Marielle Mugford, BA**
Kathy Abbott
Ellen Anderson, MS, RD
Laurie Bissett, MS, RD
Kristy Bondi, BS
Enrico Cagliero, MD
Jose C. Florez, MD, PhD+
Linda Delahanty, MS, RD
Valerie Goldman, MS, RD
Elaine Grassa
Lindsery Gurry BSN, RN, CDE
Kali D'Anna
Fernelle Leandre BS
Peter Lou, MD
Alexandra Poulos
Elyse Raymond, BS
Valerie Ripley, BS
Christine Stevens, RN
Beverly Tseng

University of California-San Diego (La Jolla, CA)

Jerrold M. Olefsky, MD*
Elizabeth Barrett-Connor, MD*
Sunder Mudaliar, MD*
Maria Rosario Araneta, PhD*
Mary Lou Carrion-Petersen, RN, BSN**
Karen Vejvoda, RN, BSN, CDE, CCRC**
Sarah Bassiouni, MPH
Madeline Beltran, RN, BSN, CDE
Lauren N. Claravall, BS
Jonalle M. Dowden, BS
Steven V. Edelman, MD
Pranav Garimella, MBBS
Robert R. Henry, MD
Javiva Horne, RD
Marycie Lamkin, RN
Simona Szerdi Janesch, BA
Diana Leos
William Polonsky, PhD
Rosa Ruiz
Jean Smith, RN
Jennifer Torio-Hurley

Columbia University (New York, NY)

F. Xavier Pi-Sunyer, MD*
Jane E. Lee, MS**
Susan Hagamen, MS, RN, CDE**
David B. Allison, PhD
Nnenna Agharanya
Nancy J. Aronoff, MS, RD
Maria Baldo
Jill P. Crandall, MD
Sandra T. Foo, MD
Jose A. Luchsinger, MD, MPH
Carmen Pal, MD
Kathy Parkes, RN
Mary Beth Pena, RN
Ellen S. Rooney, BA
Gretchen E.H. Van Wye, MA
Kristine A. Viscovich, ANP

Indiana University (Indianapolis, IN)

Mary de Groot, PhD*
David G. Marrero, PhD*
Kieren J. Mather, MD*
Melvin J. Prince, MD*
Susie M. Kelly, RN, CDE**
Marcia A. Jackson**
Gina McAtee**
Paula Putenney, RN**
Ronald T. Ackermann, MD
Carolyn M. Cantrell
Yolanda F. Dotson, BS
Edwin S. Fineberg, MD
Megan Fultz
John C. Guare, PhD
Angela Hadden
James M. Ignaut, MA
Marion S. Kirkman, MD
Erin O'Kelly Phillips
Kisha L Pinner
Beverly D. Porter, MSN
Paris J. Roach, MD
Nancy D. Rowland, BS, MS
Madelyn L. Wheeler, RD

Medstar Research Institute (Washington, DC)

Vanita Aroda, MD*
Michelle Magee, MD*

* denotes Principal Investigator
** denotes Program Coordinator

DPP, DPPOS 1, DPPOS 2 & DPPOS 3A Research Group

(1996-2021)

Robert E. Ratner, MD*
Gretchen Youssef, RD, CDE**
Sue Shapiro, RN, BSN, CCRC**
Natalie Andon, RN
Catherine Bavido-Arrage, MS, RD, LD
Geraldine Boggs, MSN, RN
Marjorie Bronsord, MS, RD, CDE
Ernestine Brown
Holly Love Burkott, RN
Wayman W. Cheatham, MD
Susan Cola
Cindy Evans
Peggy Gibbs
Tracy Kellum, MS, RD, CDE
Lilia Leon
Milvia Lagarda
Claresa Levatan, MD
Milajurine Lindsay
Asha K. Nair, BS
Jean Park, MD
Maureen Passaro, MD
Angela Silverman
Gabriel Uwaifo, MD
Debra Wells-Thayer, NP, CDE
Renee Wiggins, RD

University of Southern California/UCLA Research Center (Alhambra, CA)

Mohammed F. Saad, MD*
Karol Watson, MD*
Maria Budget**
Sujata Jinagouda, MD**
Medhat Botrous, MD**
Anthony Sosa**
Sameh Tadros**
Khan Akbar, MD
Claudia Conzues
Perpetua Magpuri
Kathy Ngo
Amer Rassam, MD
Debra Waters
Kathy Xapthalamous

Washington University (St. Louis, MO)

Julio V. Santiago, MD*
Samuel Dagogo-Jack, MD, MSc, FRCP, FACP*
Neil H. White, MD, CDE*
Angela L. Brown, MD*
Samia Das, MS, MBA, RD, LD**
Prajakta Khare-Ranade, MSc, RDN, LD**
Tamara Stich, RN, MSN, CDE**
Ana Santiago, RN
Edwin Fisher, PhD
Emma Hurt, RN
Tracy Jones, RN
Michelle Kerr, RD
Lucy Ryder, RN
Cormarie Wernimont, RD, LD

Johns Hopkins School of Medicine (Baltimore, MD)

Sherita Hill Golden, MD, MHS, FAHA*
Christopher D. Saudek, MD*
Vanessa Bradley, BA**
Emily Sullivan, MEd, RN**
Tracy Whittington, BS**
Caroline Abbas
Adrienne Allen
Frederick L. Brancati, MD, MHS
Sharon Cappelli
Jeanne M. Clark, MD
Jeanne B. Charleston, RN, MSN
Janice Freel
Katherine Horak, RD
Alicia Greene
Dawn Jiggetts
Deloris Johnson
Hope Joseph
Kimberly Loman
Nestoras Mathioudakis, MD, MHS
Henry Mosley
John Reusing
Richard R. Rubin, PhD
Alafia Samuels, MD
Thomas Shields
Shawne Stephens
Kerry J. Stewart, EdD
LeeLana Thomas
Evonne Utsey
Paula Williamson

University of New Mexico (Albuquerque, NM)

David S. Schade, MD*
Karwyn S. Adams, RN, MSN**
Janene L. Canady, RN, CDE**
Carolyn Johannes, RN, CDE**
Claire Hemphill, RN, BSN**
Penny Hyde, RN, BSN**
Leslie F. Atler, PhD
Patrick J. Boyle, MD
Mark R. Burge, MD
Lisa Chai, RN
Kathleen Colleran, MD
Ateka Fondino
Ysela Gonzales
Doris A. Hernandez-McGinnis

* denotes Principal Investigator
** denotes Program Coordinator

DPP, DPPOS 1, DPPOS 2 & DPPOS 3A Research Group
(1996-2021)

Patricia Katz, LPN
Carolyn King, Med
Julia Middendorf, RN
Amer Rassam, MD
Sofya Rubinchik, MD
Willette Senter, RD
Debra Waters, PhD

Albert Einstein College of Medicine (Bronx, NY)
Jill Crandall, MD*
Harry Shamoon, MD*
Janet O. Brown, RN, MPH, MSN**
Gilda Trandafirescu, MD**
Danielle Powell, MPH
Elsie Adorno, BS
Liane Cox, MS, RD
Helena Duffy, MS, C-ANP
Samuel Engel, MD
Allison Friedler, BS
Angela Goldstein, FNP-C, NPP, CSW
Crystal J. Howard-Century, MA
Jennifer Lukin, BA
Stacey Kloiber, RN
Nadege Longchamp, LPN
Helen Martinez, RN, MSN, FNP-C
Dorothy Pompi, BA
Jonathan Scheindlin, MD
Elissa Violino, RD, MS
Elizabeth A. Walker PhD, RN
Judith Wylie-Rosett, EdD, RD
Elise Zimmerman, RD, MS
Joel Zonszein, MD

University of Pittsburgh (Pittsburgh, PA)
Trevor Orchard, MD*
Elizabeth Venditti, PhD*
Rena R. Wing, PhD*
Susan Jeffries, RN, MSN**
Gaye Koenning, MS, RD**
M. Kaye Kramer, BSN, MPH**
Marie Smith, RN, BSN**
Susan Barr, BS
Catherine Benchoff
Miriam Boraz, PhD
Lisa Clifford, BS
Rebecca Culyba, BS
Marlene Frazier
Ryan Gilligan, BS
Stephanie Guimond, BS
Susan Harrier, MLT
Louann Harris, RN
Andrea Kriska, PhD
Qurashia Manjoo, MD
Monica Mullen, MHP, RD
Alicia Noel, BS
Amy Otto, PhD
Jessica Pettigrew, CMA
Bonny Rockette-Wagner, PhD
Debra Rubinstein, MD
Linda Semler, MS, RD
Cheryl F. Smith, PhD
Valarie Weinzierl, MPH
Katherine V. Williams, MD, MPH
Tara Wilson, BA

University of Hawaii (Honolulu, HI)
Marjorie K. Mau, MD*
Narleen K. Baker-Ladao, BS**
John S. Melish, MD
Richard F. Arakaki, MD*
Renee W. Latimer, BSN, MPH**
Mae K. Isonaga, RD, MPH**
Ralph Beddow, MD
Nina E. Bermudez, MS
Lorna Dias, AA
Jillian Inouye, RN, PhD
Kathy Mikami, BS, RD
Pharis Mohideen, MD
Sharon K. Odom, RD, MPH
Raynette U. Perry, AA
Robin E. Yamamoto, CDE, RD

Southwest American Indian Centers (Phoenix, AZ; Shiprock, NM; Zuni, NM)
William C. Knowler, MD, DrPH*+
Harelda Anderson, LMSW**
Norman Cooeyate**
Charlotte Dodge**
Mary A. Hoskin, RD, MS**
Carol A. Percy, RN, MS**
Alvera Enote**
Camille Natewa**
Kelly J. Acton, MD, MPH
Vickie L. Andre, RN, FNP
Rosalyn Barber
Shandiin Begay, MPH
Peter H. Bennett, MB, FRCP
Mary Beth Benson, RN, BSN
Evelyn C. Bird, RD, MPH
Brenda A. Broussard, RD, MPH, MBA, CDE
Brian C. Bucca, OD, FAAO
Marcella Chavez, RN, AS
Sherron Cook
Jeff Curtis, MD
Tara Dacawyma
Matthew S. Doughty, MD

* denotes Principal Investigator
** denotes Program Coordinator

DPP, DPPOS 1, DPPOS 2 & DPPOS 3A Research Group
(1996-2021)

Roberta Duncan, RD
Cyndy Edgerton, RD
Jacqueline M. Ghahate
Justin Glass, MD
Martia Glass, MD
Dorothy Gohdes, MD
Wendy Grant, MD
Robert L. Hanson, MD, MPH
Ellie Horse
Louise E. Ingraham, MS, RD, LN
Merry Jackson
Priscilla Jay
Roylen S. Kaskalla
Karen Kavena, ANP
David Kessler, MD
Kathleen M. Kobus, RNC-ANP
Jonathan Krakoff, MD
Jason Kurland, MD
Catherine Manus, LPN
Cherie McCabe
Sara Michaels, MD
Tina Morgan
Yolanda Nashboo
Julie A. Nelson, RD
Steven Poirier, MD
Evette Polczynski, MD
Christopher Piromalli, DO
Mike Reidy, MD
Jeanine Roumain, MD, MPH
Debra Rowse, MD
Robert J. Roy
Sandra Sangster, RD
Janet Sewenemewa
Miranda Smart
Chelsea Spencer
Darryl Tonemah, PhD
Rachel Williams, FNP
Charlton Wilson, MD
Michelle Yazzie

George Washington University Biostatistics Center (DPP Coordinating Center Rockville, MD)

Raymond Bain, PhD*
Sarah Fowler, PhD*
Marinella Temprosa, PhD*
Michael D. Larsen, PhD*
Tina Brenneman**
Sharon L. Edelstein, ScM**
Solome Abebe, MS
Julie Bamdad, MS
Melanie Barkalow
Joel Bethepu
Tsedenia Bezabeh
Anna Bowers
Nicole Butler
Jackie Callaghan
Caitlin E. Carter
Costas Christophi, PhD
Gregory M. Dwyer, MPH
Mary Foulkes, PhD
Yuping Gao
Robert Gooding
Adrienne Gottlieb
Kristina L. Grimes
Nisha Grover-Fairchild, MPH
Lori Haffner, MS
Heather Hoffman, PhD
Kathleen Jablonski, PhD
Steve Jones
Tara L. Jones
Richard Katz, MD
Preethy Kolinjivadi, MS
John M. Lachin, ScD
Yong Ma, PhD
Pamela Mucik
Robert Orlosky
Qing Pan, PhD
Susan Reamer
James Rochon, PhD
Alla Sapozhnikova
Hanna Sherif, MS
Charlotte Stimpson
Ashley Hogan Tjaden, MPH
Fredricka Walker-Murray

Lifestyle Resource Core

Elizabeth M. Venditti, PhD*
Andrea M. Kriska, PhD
Linda Semler, MS, RD, LDN
Valerie Weinzierl, MPH

Central Biochemistry Laboratory (Seattle, WA)

Santica Marcovina, PhD, ScD*
F. Alan Aldrich**
Jessica Harting**
John Albers, PhD
Greg Strylewicz, PhD

NIH/NIDDK (Bethesda, MD)

R. Eastman, MD
Judith Fradkin, MD
Sanford Garfield, PhD
Christine Lee, MD, MS

Centers for Disease Control & Prevention (Atlanta, GA)

* denotes Principal Investigator
** denotes Program Coordinator

DPP, DPPOS 1, DPPOS 2 & DPPOS 3A Research Group
(1996-2021)

Edward Gregg, PhD
Ping Zhang, PhD

Carotid Ultrasound
Dan O'Leary, MD*
Gregory Evans

Coronary Artery Calcification Reading Center
Matthew Budoff, MD
Chris Dailing

CT Scan Reading Center
Elizabeth Stamm, MD*

Dual Energy X-ray Absorptiometry Reading Center (San Francisco, CA)
Ann Schwartz, PhD
Caroline Navy
Lisa Palermo, MS

Epidemiological Cardiology Research Center-Epicare (Winston-Salem, NC)
Pentti Rautaharju, MD, PhD*
Ronald J. Prineas, MD, PhD**
Teresa Alexander
Charles Campbell, MS
Sharon Hall
Yabing Li, MD
Margaret Mills
Nancy Pemberton, MS
Farida Rautaharju, PhD
Zhuming Zhang, MD
Elsayed Z. Soliman, MD*
Julie Hu, MSc
Susan Hensley, BS
Lisa Keasler
Tonya Taylor

Fundus Photo Reading Center (Madison, WI)
Barbara Blodi, MD*
Ronald Danis, MD*
Matthew Davis, MD*
Larry Hubbard*
Ryan Endres**
Deborah Elsas**
Samantha Johnson**
Dawn Myers**
Nancy Barrett
Heather Baumhauer
Wendy Benz
Holly Cohn
Ellie Corkery
Kristi Dohm
Amitha Domalpally, MD, PhD
Vonnie Gama
Anne Goulding
Andy Ewen
Cynthia Hurtenbach
Daniel Lawrence
Kyle McDaniel
Jeong Pak
James Reimers
Ruth Shaw
Maria Swift
Pamela Vargo, CRA
Sheila Watson

Neurocognitive Assessment Group
Jose A. Luchsinger, MD, MPH
Jennifer Manly, PhD

Nutrition Coding Center (Columbia, SC)
Elizabeth Mayer-Davis, PhD*
Robert R. Moran, PhD**

Quality of Well-Being Center (La Jolla, CA)
Ted Ganiats, MD*
Kristin David, MHP*
Andrew J. Sarkin, PhD*
Erik Groessl, PhD
Naomi Katzir
Helen Chong, MA

University of Michigan (Ann Arbor, MI)
William H. Herman, MD, MPH
Michael Brändle, MD, MS
Morton B. Brown, PhD

+Genetics Working Group
Jose C. Florez, MD, PhD[1, 2]
David Altshuler, MD, PhD[1, 2]
Liana K. Billings, MD[1]
Ling Chen, MS[1]
Maegan Harden, BS[2]
Robert L. Hanson, MD, MPH[3]
William C. Knowler, MD, DrPH[3]
Toni I. Pollin, PhD[4]
Alan R. Shuldiner, MD[4]
Kathleen Jablonski, PhD[5]
Paul W. Franks, PhD, MPhil, MS[6, 7, 8]
Marie-France Hivert, MD[9]

1=Massachusetts General Hospital
2=Broad Institute
3=NIDDK
4=University of Maryland
5=Coordinating Center
6=Lund University, Sweden
7=Umeå University, Sweden
8=Harvard School of Public Health
9=Université de Sherbrooke

* denotes Principal Investigator
** denotes Program Coordinator

SUPPLEMENTAL MATERIAL

Supplemental Table S1. Distribution of 3,168 samples with genomic data by self-reported race/ethnicity and DPP treatment arm that underwent imputation.

	Treatment arm				
Self-reported race/ethnicity	**MET**	**ILS**	**PBO**	**TROG**	**Total**
AfrAm	178	173	186	94	631
AsnPI	31	49	38	20	138
Hisp	143	161	147	89	540
AI	21	28	27	3	79
White	507	479	490	304	1780
Total	880	890	888	510	3168

AfrAm=African American; AsnPI=Asian/Pacific Islander; Hisp=Hispanic/Latino; AI=American Indian; MET=metformin; ILS=intensive lifestyle modification; PBO=placebo; TROG=troglitazone

Descriptions of replication cohorts:

The Metformin Genetics (MetGen) Consortium: MetGen consists of 20 research institutions from Europe and the United States with available data for studies of metformin pharmacogenetics from population observational studies and clinical trials. This resource has an excess of 10,000 multi-ethnic individuals with established type 2 diabetes (T2D) across over 12 cohorts in whom a pharmacogenetic meta-analysis of metformin has been performed using a harmonized measure of glycemic response. Additional details regarding the methodology of the meta-GWAS can be found in Zhou *et al.* (1).

The Million Veteran Program (MVP): The MVP is a mega-biobank in the Department of Veteran Affairs (VA) healthcare system that combines survey data, clinical electronic health record information, and biospecimens to enable genomic research. Recently, 318 novel genetic risk loci for T2D were identified in a multi-ethnic analysis that incorporated over 100,000 cases of T2D from MVP (2). For this replication analysis, we included 2733 individuals who self-identified as non-Hispanic African Americans with a baseline HbA1c ≥6.5% and received metformin monotherapy for up to 15 months.

Diabetes Multi-omic Investigation of Drug Response (DIAMOND): DIAMOND is an NIH-funded study that aimed to enroll 5500 individuals with T2D from the metropolitan Detroit and the surrounding areas of southeast Michigan for the purposes of understanding the genetic causes of diabetes, diabetes-related traits, and medication response. Individuals with more than two diabetes diagnoses and at least two historical HbA1c tests drawn over 4 months apart were identified from the electronic medical record in the Henry Ford Health System. For this replication analysis, we assessed a subset of 471 individuals who self-reported as African American and had an average daily exposure of ≥500 mg metformin in the 120 days prior to the follow-up HbA1c measurement. The follow-up period was between 4-18 months.

Supplemental Table S2. Genotyping platforms and models utilized in each replication analysis.

Cohort	Genotyping and Imputation	Outcome	Covariates
MetGen	See Supplementary Table 2 of Zhou *et al.* (1) for details of the genotyping platforms and imputation methods used for the meta-GWAS	Change in HbA1c (baseline minus follow-up within 18 months)	Baseline HbA1c, adherence, metformin dose, treatment group (metformin monotherapy vs. add-on to sulfonylurea), 10 ancestry PCs Some covariates were included to varying degrees – see Supplementary Table 2 of Zhou *et al.* (1) for details of the study-specific models
MVP	Custom Affymetrix Axiom array (MVP 1.0) Imputation using 1000 Genomes Project Phase 3 reference panel	Change in HbA1c (baseline minus follow-up within 15 months)	Baseline HbA1c, age, BMI, metformin dose, number of HbA1c measurements, 10 ancestry PCs
DIAMOND	Axiom Precision Medicine Diversity Array Imputation using the TOPMed reference panel (version R2 on GRC38)	Change in HbA1c (follow-up minus baseline within 18 months)	Baseline HbA1c, age, sex, 10 ancestry PCs

PC=principal components

Supplemental Table S3. Event rate for diabetes incidence by DPP treatment arm and self-reported race/ethnicity information.

	MET		PBO	
Self-reported race/ethnicity	**N**	**# Events (%)**	**N**	**# Events (%)**
AfrAm	178	36 (20.2)	186	61 (32.8)
AsnPI	31	8 (25.8)	38	12 (32.4)
Hisp	143	34 (23.8)	147	39 (26.5)
AI	21	4 (20.0)	27	8 (29.6)
White	507	101 (25.1)	490	133 (27.1)
Total	880	183 (20.9)	888	253 (28.5)

AfrAm=African American; AsnPI=Asian/Pacific Islander; Hisp=Hispanic/Latino; AI=American Indian; MET=metformin; PBO=placebo.

Supplemental Table S4. Sample sizes for each individual model and outcome tested.

Outcome	**MET*** **(n=880)**	**PBO*** **(n=888)**	**G×T*** **(n=1768)**
Diabetes incidence	876	887	1763
HbA1c	818	803	1621
Weight	829	844	1673
Fasting glucose	821	812	1633
2-hr glucose	817	810	1627
Insulin sensitivity index	808	818	1626
Fasting insulin	802	792	1594

MET=metformin; PBO=placebo; G×T=gene by treatment. *The decrease in sample size from the counts in the column headings is due to data missingness within in each model (i.e., follow-up measurement for the quantitative trait was unavailable, so the value for one-year change could not be calculated).

Supplemental Table S5. Effect size and *p*-values, stratified by self-reported race/ethnicity, for the 14 independent genome-wide significant findings for one-year change in quantitative traits in the DPP.

rsid	Chr	Position*	Nearest gene	EA	NEA	Trait	EAF†	Self-reported race/ethnicity	N	Beta	SE	*p*
Metformin only model												
rs144322333	18	705550	*ENOSF1*	C	CTGTT	HbA1c, %	0.013	All	818	0.39	0.06	2.9E-12
							0.065	AfrAm	166	0.32	0.06	1.3E-06
							0.009	Hisp	138	1.05	0.22	3.8E-06
rs145591055	5	38849463	*OSMR*	G	A	Weight, kg	0.014	All	829	-7.55	1.19	3.2E-10
							0.018	AsnPI	28	6.18	18.75	7.5E-01
							0.064	Hisp	139	-5.56	1.60	6.8E-04
							0.103	AI	21	-3.20	2.79	2.9E-01
							0.007	White	473	-10.71	3.02	4.4E-04
rs13401282	2	207810690	*CPO*	A	T	ISI, ln	0.029	All	808	0.44	0.08	1.7E-08
							0.119	AfrAm	167	0.42	0.08	2.2E-07
							0.015	Hisp	134	-0.25	0.36	4.9E-01
rs186681623	13	81546608	*LINC00377*	C	T	Weight, kg	0.068	All	829	-2.66	0.47	2.0E-08
							0.121	AfrAm	169	-1.98	0.61	1.5E-03
							0.074	AsnPI	28	-7.65	3.11	2.9E-02
							0.122	Hisp	139	-3.19	1.13	5.4E-03
							0.161	AI	21	-5.24	2.68	9.8E-02
							0.036	White	473	-3.83	0.87	1.4E-05
rs17083791	5	93863813	*KIAA0825*	G	A	Weight, kg	0.149	All	829	-1.85	0.33	3.3E-08
							0.052	AfrAm	169	-1.65	0.95	8.3E-02
							0.194	AsnPI	28	0.42	1.53	7.9E-01
							0.210	Hisp	139	-2.08	0.83	1.3E-02
							0.367	AI	21	-3.19	2.39	2.3E-01
							0.137	White	473	-1.71	0.44	1.2E-04
rs9931871	16	19928315	*GPRC5B*	G	A	Fasting glucose, mmol/L	0.011	All	821	0.68	0.12	3.5E-08
							0.049	AfrAm	168	0.62	0.13	5.9E-06
rs549305231	6	19497504	*LOC101928519*	A	G	HbA1c, %	0.012	All	818	0.35	0.06	3.9E-08
							0.046	AfrAm	166	0.40	0.07	4.9E-08
							0.005	Hisp	138	-1.22	1.24	3.3E-01
rs73944532	2	104691634	*LOC100287010*	A	G	Fasting insulin, ln	0.020	All	802	-0.47	0.08	4.1E-08
							0.075	AfrAm	165	-0.50	0.09	1.4E-07
							0.007	Hisp	134	-0.13	0.30	6.6E-01

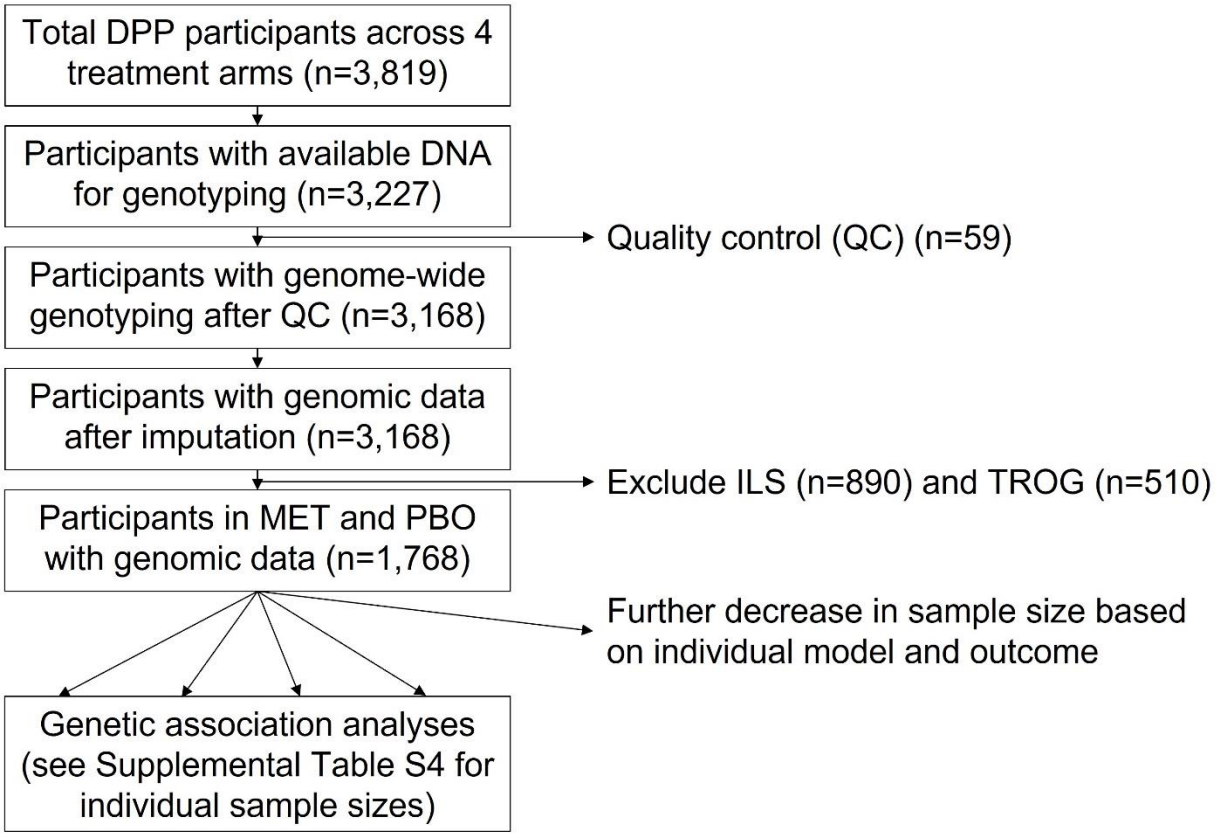
rsid	Chr	Position*	Nearest gene	EA	NEA	Trait	EAF†	Self-reported race/ethnicity	N	Beta	SE	*p*
Gene × Treatment model‖												
rs6838493	4	185879789	*LINC01093*	A	T	HbA1c, %	0.011	All	1621	-0.66	0.11	1.6E-09
							0.041	AfrAm	330	-0.80	0.14	6.0E-08
rs148219263	18	58808522	*CDH20*	C	T	Weight, kg	0.012	All	1673	-6.92	1.22	1.9E-08
							0.007	AfrAm	342	-4.09	6.50	5.3E-01
							0.011	Hisp	275	-0.37	2.98	9.0E-01
							0.017	White	947	-8.81	1.50	5.6E-09
rs75147163	14	57626769	*EXOC5*	A	G	Fasting glucose, mmol/L	0.022	All	1633	-0.73	0.13	1.4E-08
							0.053	AfrAm	334	-1.05	0.25	4.2E-05
							0.022	Hisp	267	-0.36	0.39	3.6E-01
							0.015	White	925	-0.53	0.18	3.0E-03
rs78075715	4	6613716	*MAN2B2*	C	T	HbA1c, %	0.013	All	1621	-0.51	0.09	1.4E-08
							0.063	AfrAm	330	-0.53	0.11	1.4E-06
rs12314996	12	24567467	*SOX5*	A	G	HbA1c, %	0.016	All	1621	-0.47	0.08	3.9E-08
							0.062	AfrAm	330	-0.51	0.11	4.5E-06
							0.007	Hisp	266	-0.24	0.60	6.9E-01
rs143203347	7	107384199	*CBLL1*	C	G	Fasting glucose, mmol/L	0.015	All	1633	-0.92	0.17	4.1E-08
							0.062	AfrAm	334	-0.92	0.21	1.6E-05
							0.011	Hisp	267	-0.42	0.62	5.0E-01

EA=Effect allele; NEA=Non-effect allele; EAF=effect allele frequency. AfrAm=African American; AsnPI=Asian/Pacific Islander; Hisp=Hispanic/Latino; AI=American Indian; ISI=insulin sensitivity index. *GRCh37 assembly. †This is the effect allele frequency taken from Table 2. For each variant, we only included stratified analyses when the effect allele frequency was greater than 0.005 in the self-reported race/ethnicity group. ‖Beta estimates and standard errors (SE) are reported for the interaction term rather than the main effect of metformin.

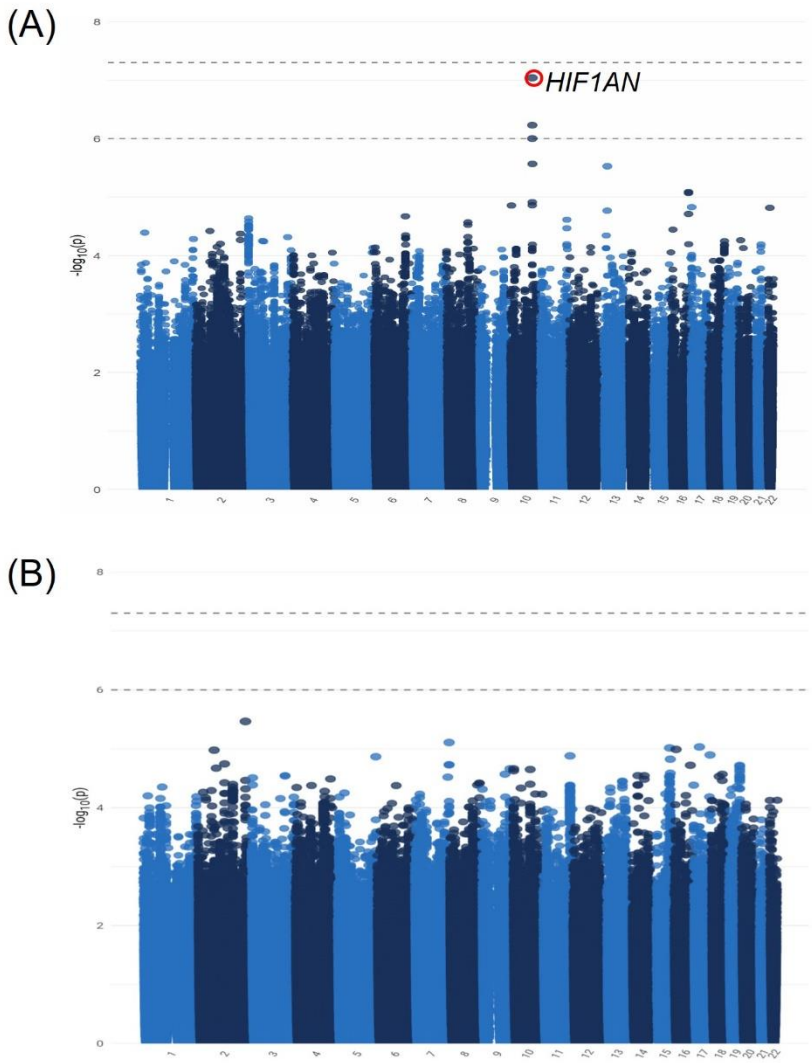
Supplemental Table S6. Evaluation of the 14 independent genome-wide significant findings from Table 2 in the lifestyle treatment arm of the DPP.

rsid	Chr	Position*	Nearest gene	EA	NEA	Trait	EAF†	Metformin only model			Lifestyle only model		
								Beta	SE	*p*	Beta	SE	*p*
rs144322333	18	705550	*ENOSF1*	C	CTGTT	HbA1c, %	0.013	0.39	0.06	2.9E-12	0.04	0.06	5.3E-01
rs145591055	5	38849463	*OSMR*	G	A	Weight, kg	0.014	-7.55	1.19	3.2E-10	1.08	1.31	4.1E-01
rs13401282	2	207810690	*CPO*	A	T	ISI, ln	0.029	0.44	0.08	1.7E-08	0.26	0.09	4.1E-03
rs186681623	13	81546608	*LINC00377*	C	T	Weight, kg	0.068	-2.66	0.47	2.0E-08	0.79	0.61	2.0E-01
rs17083791	5	93863813	*KIAA0825*	G	A	Weight, kg	0.149	-1.85	0.33	3.3E-08	0.03	0.47	9.5E-01
rs9931871	16	19928315	*GPRC5B*	G	A	Fasting glucose, mmol/L	0.011	0.68	0.12	3.5E-08	-0.07	0.12	5.4E-01
rs549305231	6	19497504	*LOC101928519*	A	G	HbA1c, %	0.012	0.35	0.06	3.9E-08	-0.12	0.11	2.8E-01
rs73944532	2	104691634	*LOC100287010*	A	G	Fasting insulin, ln	0.020	-0.47	0.08	4.1E-08	-0.15	0.10	1.3E-01
rsid	**Chr**	**Position***	**Nearest gene**	**EA**	**NEA**	**Trait**	**EAF†**	**Gene × Treatment model‖ (metformin and placebo)**			**Gene × Treatment model‖ (lifestyle and placebo)**		
								Beta	**SE**	***p***	**Beta**	**SE**	***p***
rs6838493	4	185879789	*LINC01093*	A	T	HbA1c, %	0.011	-0.66	0.11	1.6E-09	-0.54	0.12	5.9E-06
rs148219263	18	58808522	*CDH20*	C	T	Weight, kg	0.012	-6.92	1.22	1.9E-08	6.31	1.96	1.3E-03
rs75147163	14	57626769	*EXOC5*	A	G	Fasting glucose, mmol/L	0.022	-0.73	0.13	1.4E-08	-0.75	0.12	1.1E-09
rs78075715	4	6613716	*MAN2B2*	C	T	HbA1c, %	0.013	-0.51	0.09	1.4E-08	-0.40	0.09	7.9E-06
rs12314996	12	24567467	*SOX5*	A	G	HbA1c, %	0.016	-0.47	0.08	3.9E-08	-0.35	0.10	2.8E-04
rs143203347	7	107384199	*CBLL1*	C	G	Fasting glucose, mmol/L	0.015	-0.92	0.17	4.1E-08	-0.78	0.19	5.0E-05

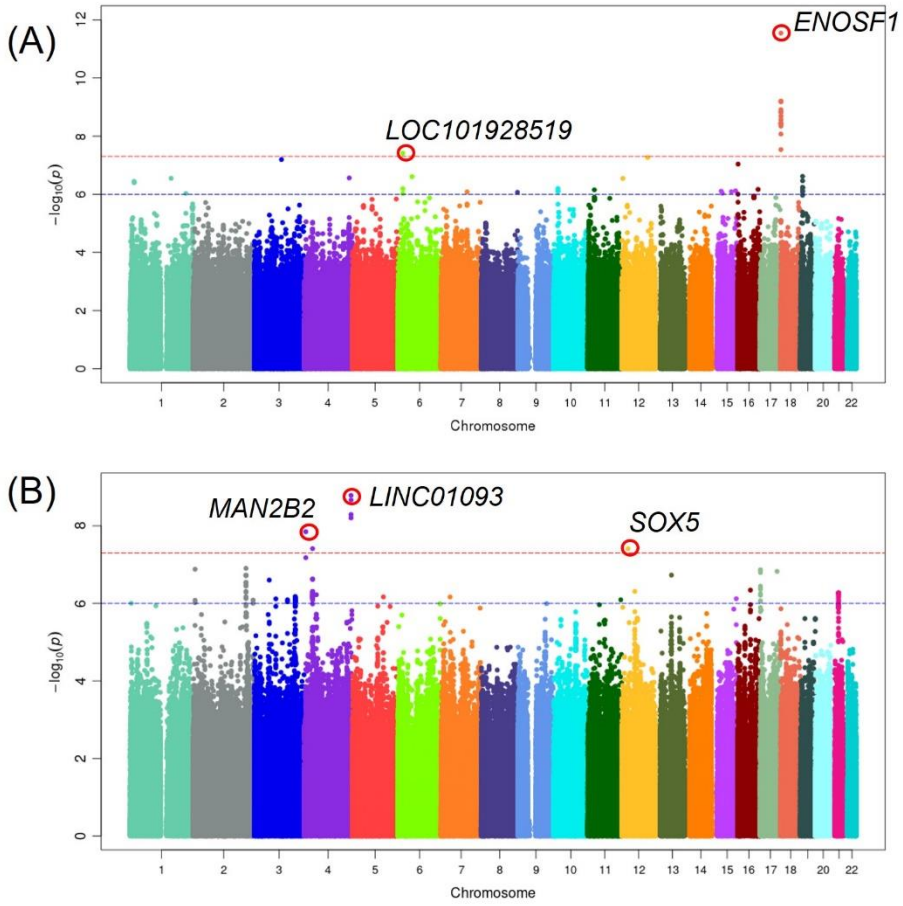
EA=Effect allele; NEA=Non-effect allele; EAF=effect allele frequency. ISI=insulin sensitivity index. *GRCh37 assembly. †This is the effect allele frequency taken from "All" participants in Table 2. ‖Beta estimates and standard errors (SE) are reported for the interaction term rather than the main effect of metformin or lifestyle.



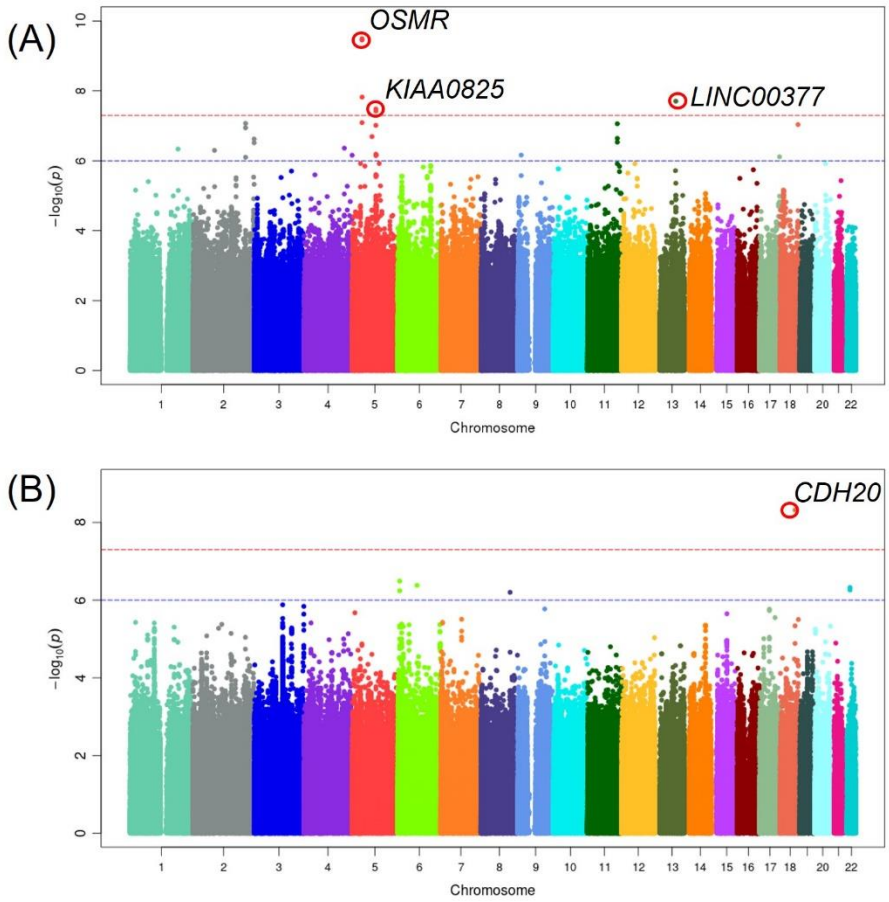
Supplemental Figure S1. Flowchart illustrating the sample size reduction in the DPP leading up to genetic association analyses. MET=metformin; ILS=intensive lifestyle modification; PBO=placebo; TROG=troglitazone.



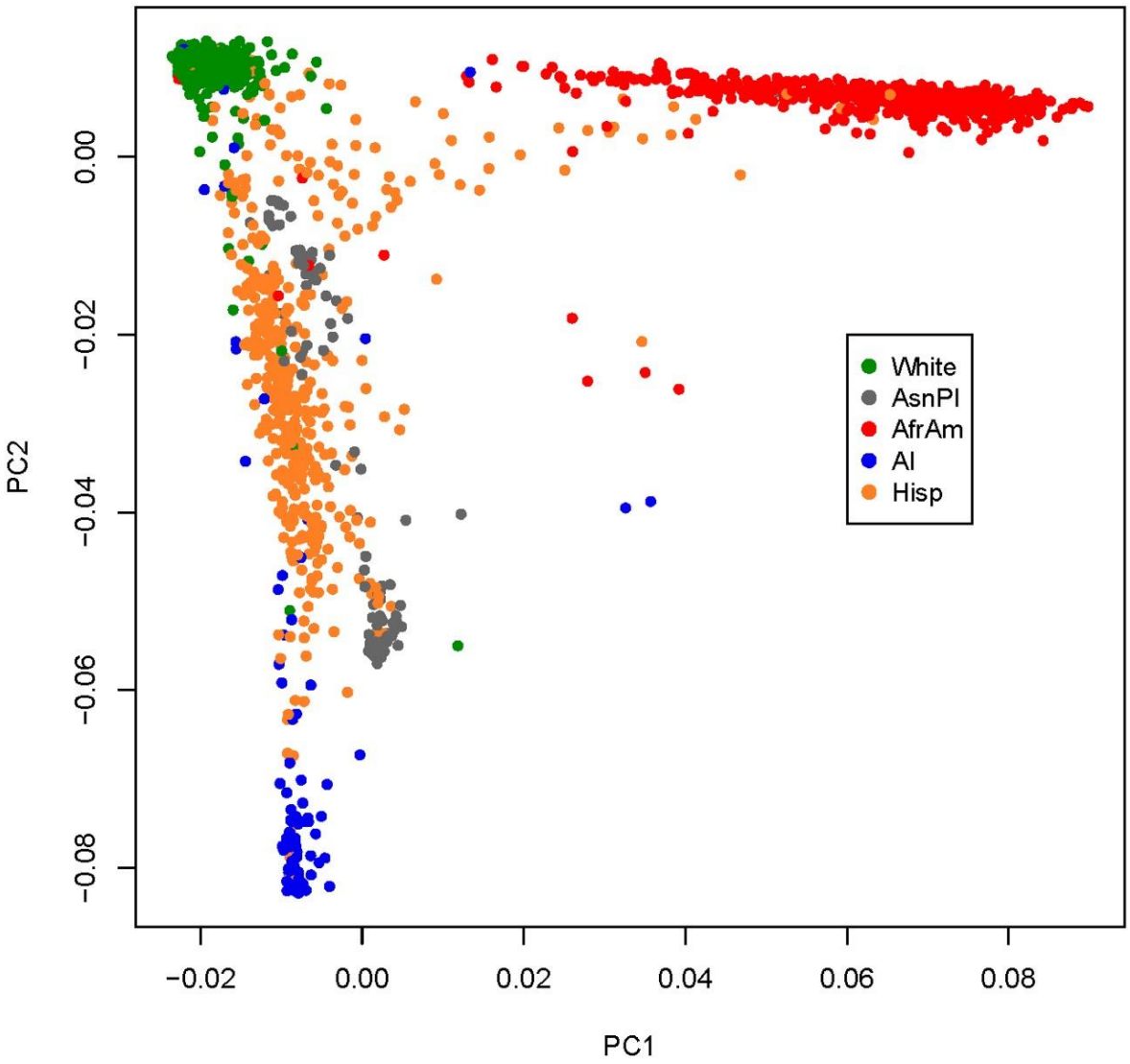
Supplemental Figure S2. Manhattan plot of genome-wide results from single marker association with diabetes incidence (A) using an additive genetic model in 876 individuals in the metformin arm only and (B) testing a gene × treatment interaction in the placebo and metformin arms in 1,763 individuals. For each panel, the two dotted lines indicate the suggestive significance ($p<1\times10^{-6}$) and genome-wide significance ($p<5\times10^{-8}$). Top loci are circled in red with the closest gene labeled.



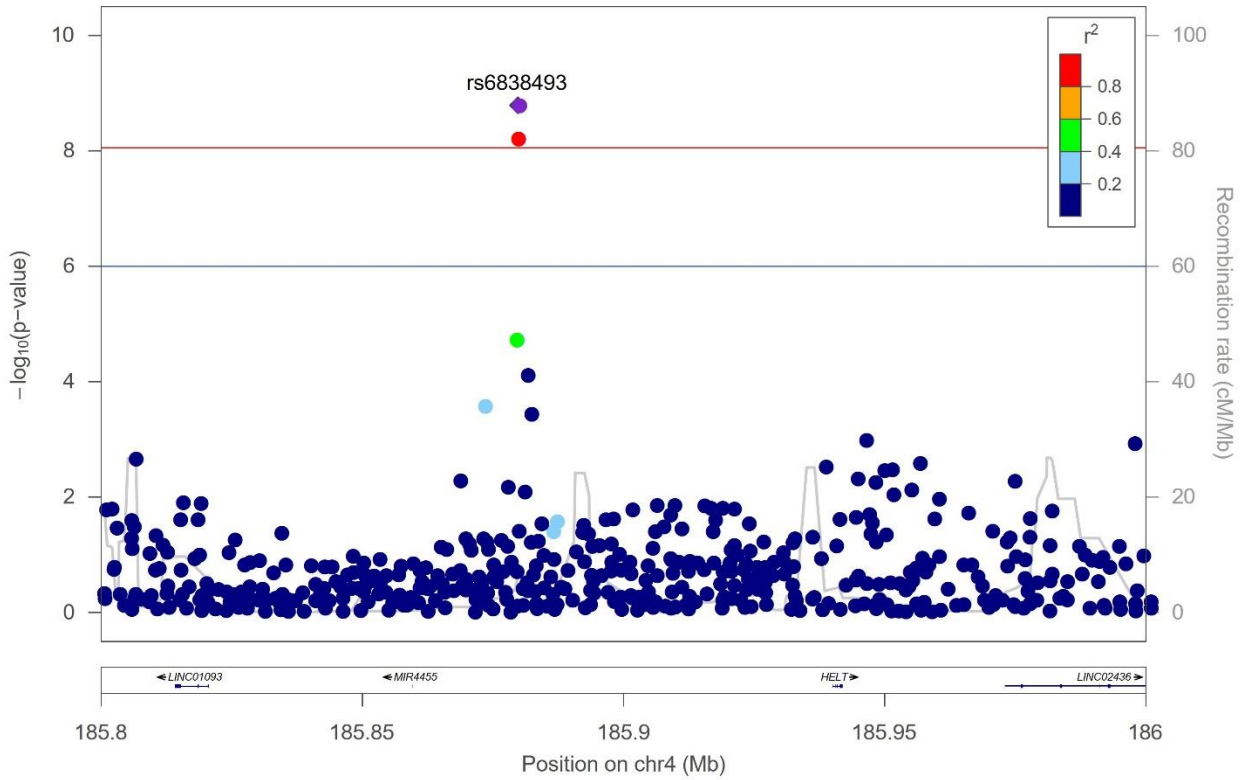
Supplemental Figure S3. Manhattan plots of genome-wide results from single marker association with one-year change in HbA1c (A) using an additive genetic model in 818 individuals in the metformin arm only and (B) testing a gene × treatment interaction in the placebo and metformin arms in 1,621 individuals. The blue line indicates suggestive significance ($p<1\times10^{-6}$) and the red line indicates genome-wide significance ($p<5\times10^{-8}$). Top loci are circled in red with the closest gene labeled.



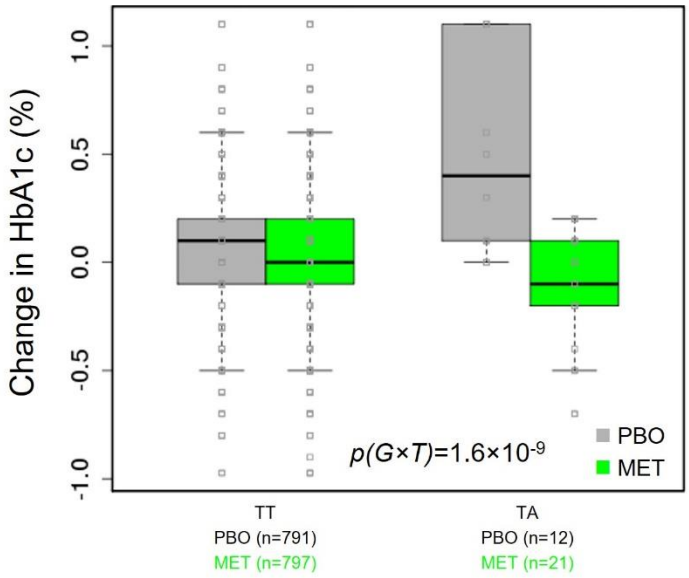
Supplemental Figure S4. Manhattan plots of genome-wide results from single marker association with one-year change in weight (A) using an additive genetic model in 829 individuals in the metformin arm only and (B) testing a gene × drug interaction in the placebo and metformin arms in 1,673 individuals. The blue line indicates suggestive significance ($p<1\times10^{-6}$) and the red line indicates genome-wide significance ($p<5\times10^{-8}$). Top loci are circled in red with the closest gene labeled.



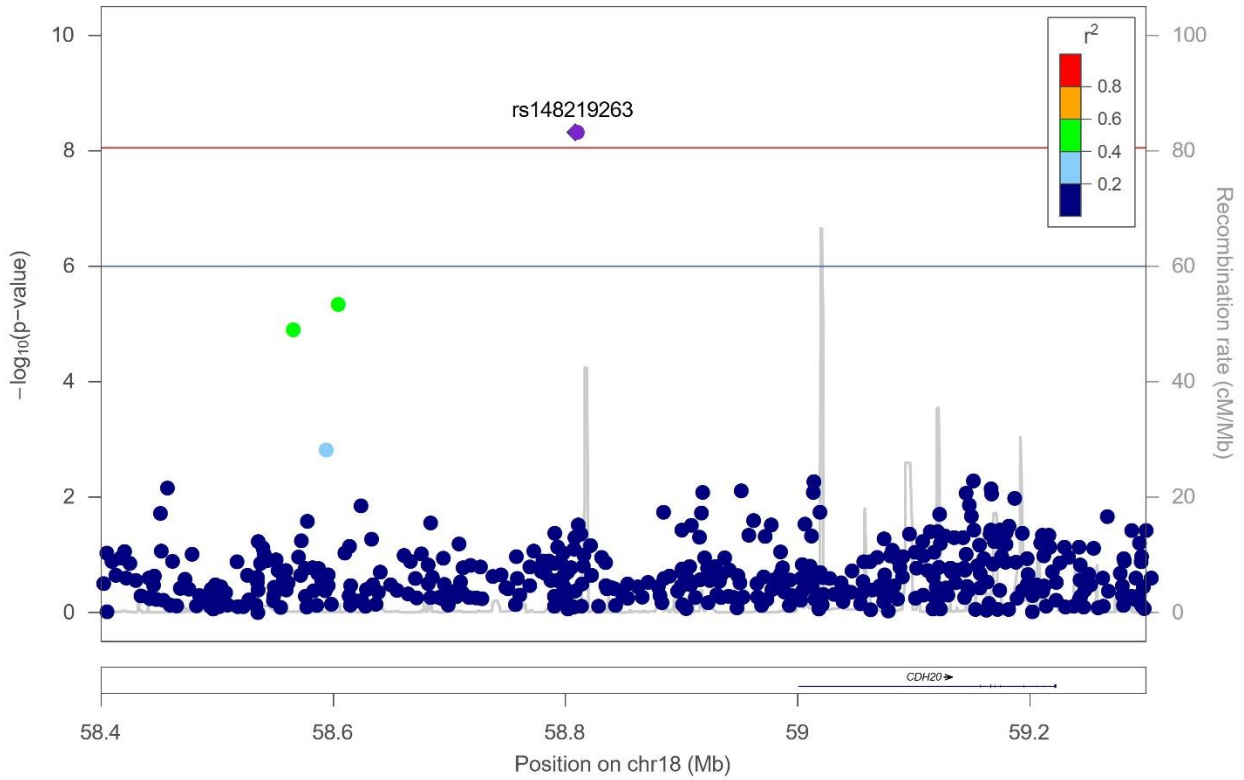
Supplemental Figure S5. Self-reported race/ethnicity (colored circles) displayed on a background plot of the first two genetic ancestry principal components.



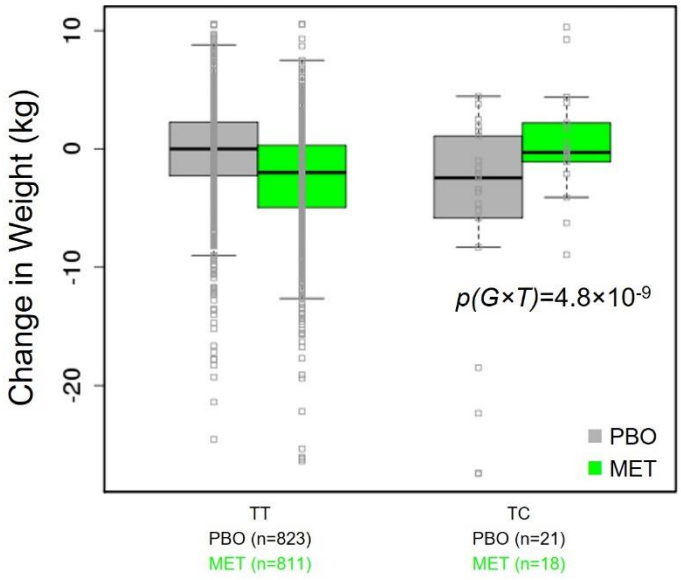
Supplemental Figure S6. Regional association plot of rs6838493 for one-year change in HbA1c. The red line indicates the experiment-wide significance threshold of $p<9\times10^{-9}$ and the blue line indicates a suggestive significance threshold of $p<1\times10^{-6}$.



Supplemental Figure S7. Comparison of the influence of rs6838493 genotype on the mean change in HbA1c (one-year minus baseline) in the metformin (MET, n=818) and placebo (PBO, n=803) arms. The interaction *p*-value is reported. For the purposes of generating the box plots, fractional alleles were converted to "hard calls" by rounding to the nearest integer.



Supplemental Figure S8. Regional association plot of rs148219263 for one-year change in weight. The red line indicates the experiment-wide significance threshold of $p<9\times10^{-9}$ and the blue line indicates a suggestive significance threshold of $p<1\times10^{-6}$.



Supplemental Figure S9. Comparison of the influence of rs148219263 genotype on the mean change in weight (one-year minus baseline) in the metformin (MET, n=829) and placebo (PBO, n=844) arms. The interaction *p*-value is reported. For the purposes of generating the box plots, fractional alleles were converted to "hard calls" by rounding to the nearest integer.

References

1. Zhou K, Yee SW, Seiser EL, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet* 2016;48:1055-1059
2. Vujkovic M, Keaton JM, Lynch JA, et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat Genet* 2020;52:680-691