Genetic Variants in CPA6 and PRPF31 are Associated with Variation in Response to Metformin in Individuals with Type 2 Diabetes

Rotroff, Daniel M.; Yee, Sook Wah; Zhou, Kaixin; Marvel, Skylar W.; Shah, Hetal S.; Jack, John R.; Havener, Tammy M.; Hedderson, Monique M.; Kubo, Michiaki; Herman, Mark A.; Gao, He; Mychaleckyi, Josyf C.; McLeod, Howard L.; Doria, Alessandro; Giacomini, Kathleen M.; Pearson, Ewan R.; Wagner, Michael J.; Buse, John B.; Motsinger-Reif, Alison A.; MetGen Investigators and the ACCORD/ACCORDion Investigators

Published in:
Diabetes

DOI:
10.2337/db17-1164

Publication date:
2018

Document Version
Peer reviewed version

Genetic Variants in *CPA6* and *PRPF31* are Associated with Variation in Response to Metformin in Individuals with Type 2 Diabetes

Daniel M. Rotroff[1,2], Sook Wah Yee[3], Kaixin Zhou[4], Skylar W. Marvel[1], Hetal S. Shah[5], John R. Jack[1], Tammy M. Havener[6], Monique M. Hedderson[7], Michiaki Kubo[8], Mark A. Herman[9], He Gao[5], Josyf C. Mychaleckyi[10], Howard L. McLeod[11], Alessandro Doria[5], Kathleen M. Giacomini[3], Ewan R. Pearson[4], Michael J. Wagner[6], John B. Buse[12], Alison A. Motsinger-Reif[1,2#], MetGen Investigators and the ACCORD/ACCORDion Investigators.

1. Bioinformatics Research Center, North Carolina State University, Raleigh NC
2. Dept. of Statistics, North Carolina State University, Raleigh NC
3. Department of Bioengineering at Therapeutic Sciences and Institute for Human Genetics, University of California, San Francisco, CA, USA
4. School of Medicine, Ninewells Hospital and Medical School, University of Dundee, Dundee, Scotland, United Kingdom
5. Joslin Diabetes Center and Harvard Medical School, Boston, MA, USA
6. Center for Pharmacogenomics and Individualized Therapy, UNC Chapel Hill, Chapel Hill, NC, USA
7. Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA
8. RIKEN Center for Integrative Medical Science, Yokohama, Japan
9. Department of Medicine, Duke University, Durham, NC, USA
10. Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22904, USA
11. Moffitt Cancer Center, 12002 Magnolia Dr., Tampa, FL, USA
12. Division of Endocrinology, University of North Carolina School of Medicine, Chapel Hill, NC

#Corresponding Author:
Alison Motsinger-Reif, Ph.D.
NCSU Statistics Department
2311 Stinson Drive
Campus Box 7566
Raleigh, NC 27695-7566
Tel: (919) 515-3574
Email: alison_motsinger@ncsu.edu

Abstract

Metformin is the first line treatment for type 2 diabetes (T2D). Although widely prescribed, the glucose-lowering mechanism for metformin is incompletely understood. Here we used a genome-wide association approach in a diverse group of individuals with T2D from the Action to Control Cardiovascular Risk in Diabetes (ACCORD) clinical trial to identify common and rare variants associated with HbA1c response to metformin treatment, and followed up these findings in four replication cohorts. Common variants in *PRPF31* and *CPA6*, were associated with worse and better metformin response, respectively ($p<5x10^{-6}$), and meta-analysis in independent cohorts displayed similar associations with metformin response ($p=1.2x10^{-8}$ and p=0.005, respectively). Previous studies have shown that *PRPF31*(+/-) knockout mice have increased total body fat ($p=1.78x10^{-6}$) and increased fasted circulating glucose ($p=5.73x10^{-6}$). Furthermore, rare variants in *STAT3* associated with worse metformin response($q<0.1$). STAT3 is a ubiquitously expressed pleiotropic transcriptional activator that participates in the regulation of metabolism and feeding behavior. Here we provide novel evidence for associations of common and rare variants in *PRPF31, CPA6,* and *STAT3* with metformin response that may provide insight into mechanisms important for metformin efficacy in T2D.

Introduction

The incidence of Type 2 Diabetes (T2D) is increasing and the Centers for Disease Control and Prevention estimate that 11% of U.S. adults, aged 20 years or older, have diagnosed or undiagnosed T2D, and 35% of people in the same age group have prediabetes based on fasting glucose or hemoglobin A1c (HbA1c) levels(1). T2D is now considered a global epidemic with prevalence increasing from 108 million in 1980 to 422 million in 2014(2). Individuals with T2D are at an increased risk of developing blindness and kidney failure, and are at risk for lower limb amputations. Furthermore, individuals with T2D are 2-4x more likely to develop cardiovascular disease (CVD), including heart attack and stroke(3).

Metformin, a member of the biguanide class of drugs, is now considered first line therapy for individuals with T2D(4). Despite intensive investigation, the molecular mechanisms mediating metformin's beneficial effects on glycemic control remain controversial and poorly understood. Studies performed predominantly in animal models have implicated various mechanisms and molecules as participating in metformin's beneficial effects, including direct inhibitory effects on mitochondrial function, activation of hepatic AMPK, and alterations in glucagon signaling pathways(5–10). The relative importance of these mechanisms for metformin's beneficial effects on glycemic control in humans is unclear.

There is significant inter-individual variability in response to metformin(11–13). This suggests that a better understanding of the mechanisms by which metformin functions might allow for more tailored and precise treatment. Studies suggest that heritable factors contribute to this variability, which provides an opportunity to identify causal genetic contributors through genetic association studies(11). Recent GWAS have identified common variants that affected metformin response in the *ATM* locus, which has been shown to activate AMPK, and the *SLC2A2* locus, which encodes the facilitated glucose transporter, GLUT2(12,13). Here we use a GWAS approach in a large cohort of individuals with T2D in the ACCORD clinical trial to test for associations of both common and rare variant SNPs with change in HbA1c in response to metformin treatment.

The Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial followed 10,251 participants for up to 8 years at 77 clinical centers in the United States and Canada to compare the benefits and risks of treatment strategies for intensively targeting glycemia, blood pressure, and dyslipidemia versus standard targets in individuals with T2D at high risk for CVD(14–17). No overall benefit and possible harms were observed in the combined primary CVD endpoints with intensive glucose-lowering therapy. However, significant variability in treatment response was also observed, highlighting the potential for identifying genetic markers of drug response that may lead to the

development of improved and personalized treatment strategies. The results presented here point to novel mechanisms of metformin drug response and potential therapeutic targets for the treatment of T2D.

Research Design and Methods

Study Participants

The ACCORD trial (clinicaltrials.gov-NCT00000620) was a double 2x2 factorial design comparing intensive versus standard treatment approaches for controlling glycemia, blood pressure, and dyslipidemia that enrolled 10,251 patients with T2D with a history of CVD or at least two known risk factors for CVD, such as documented atherosclerosis, albuminuria, dyslipidemia, hypertension, smoking, or obesity(14). Additional details about randomization and selection criteria for the various arms of the ACCORD trial can be found in the Supplemental Material. Participants in ACCORD were given an option to provide a blood sample for future genetic studies, and over 80% of participants agreed to do so. A workflow describing the selection of subjects can be found in Figure 1. Additional demographic information can be found in Table 1.

Phenotype Definitions

Glycemic response to metformin was evaluated in subjects who began taking metformin while enrolled in ACCORD; subjects who reported taking metformin or another biguanide prior to enrollment in the trial were excluded. Some subjects were taking additional medications and were accounted for as described in the Supplemental Material and Graham and Rotroff et al.(18). Subjects were scheduled for study visits every 1 or 4 months based on randomization to the intensive or standard glycemia arms, respectively. HbA1c was recorded every 4 months. Metformin response was calculated as on-treatment HbA1c minus pre-treatment HbA1c. Pre-treatment HbA1c levels were recorded no more than 30 days

prior to the start of metformin. On-treatment HbA1c levels were defined as the first recorded measurement acquired after at least 90-days and no more than 270-days from the start of metformin treatment. This interval was chosen to allow for HbA1c levels to stabilize after starting metformin, while limiting changes in glycemia levels due to additional modifications to treatment regimens. Medication compliance and other details regarding the phenotype definition can be found in the Supplemental Material.

Genotyping

Briefly, 6,085 unique samples from ACCORD participants who consented to genetic studies conducted by any investigator, were genotyped at the University of Virginia on Illumina HumanOmniExpressExome-8 v1.0 chips (Set 1); 8,174 unique samples, including the above 6,085 samples plus 2,089 samples from ACCORD participants who consented to genetic studies only if conducted by ACCORD investigators, were genotyped at the University of North Carolina on Affymetrix Axiom Biobank1 chips (Set 2). The data were then merged, resulting in one dataset consisting of 5,971 samples genotyped at a total of 1,240,656 individual SNPs, and an additional 2,083 samples genotyped at 583,613 SNPs. Additional details can be found in the Supplementary Material and Marvel and Rotroff et al.(19).

Data Processing

Covariate Selection

We implemented a variable selection procedure to address potential confounding. Some variables were included in the model based on previous studies or expert knowledge, while other variables were selected based on a backwards selection approach and Bayesian information criteria to identify covariates specific to the ACCORD dataset. All covariate names and descriptions before variable

selection can be found in Supplemental Table 1. A substantial proportion of the cohort was taking other medications during the metformin treatment response time frame and these concomitant medications were incorporated into the model as described in the Supplementary Material. To prevent confounding due to population substructure, principal component analysis (PCA) was performed based on the genotype data using EIGENSTRAT (v4.2), and is described in the Supplemental Material. The final models after covariate selection can be found in Supplementary Tables 2-4. A workflow detailing each step of the analysis can be found in Supplemental Figure 4.

Common Variant Analysis

Associations between a phenotype, covariates, and single common variant (MAF>3%) were tested using the linear regression model. Genotyped variants were tested as an additive variable using PLINK, where $g_i \in \{0,1,2\}$ is the number of minor alleles for the i^{th} individual. Imputed variants were tested using a linear regression model in the statistical programming language, R, where $g_i = p_i(Aa) + 2p_i(aa)$ is the dosage score computed from the posterior probabilities for genotypes Aa and aa (20). For SNPs that were only genotyped in Set 1 subjects and were imputed in Set 2 subjects, association tests were calculated for each set separately and the results were combined by meta-analysis using PLINK(21,22). Tables and figures specify whether each result derives from a SNP that was genotyped in all subjects, imputed in all subjects, or represents a combination of genotyped and imputed data using the meta-analysis approach described.

Rare Variant Analysis

The rare variant analysis approach has been previously described(19). Briefly, all variants (MAF≤3%) were mapped to gene annotation from Ensemble (GRCh37.p13). A suite of five rare variant tests comprised of burden and non-burden were then used to assess associations with metformin response. We combined the set of five p-values from each test into a single p-value for each gene using

the procedure described by Dai et al.(23). Subsequently, the combined p-value was corrected for multiple comparisons with a false discovery rate approach(24,25), and a threshold of $q<0.1$ was used for statistical significance.

Replication of ACCORD Results

SNPs that associated with metformin response in the common variant analyses ($p<5x10^{-6}$) were tested for associations in the following cohorts: There was one cohort from European ancestry, which included combined samples from Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTS) and PMET1-EU (REP1) (n=6963). There were two cohorts from African American ancestry, referred to as REP2 (n=646) and REP3 (n=369). Briefly, the GoDARTS cohort in REP1 has been described before by WTCCC2(13). The REP2 cohort was collected from two main clinical sites, Kaiser Permanente Northern California and Kaiser Permanente South East, as previously described(26). The PMET1-EU in REP1 and the REP3 cohorts were collected from The Research Program on Genes, Environment, and Health (RPGEH), based at Kaiser Permanente Northern California(27). Metformin response for REP1 was defined as pre-treatment HbA1c minus post-treatment HbA1c, resulting in effects that are in the opposite direction of those in ACCORD. Because of this, the additive inverse of effect sizes in REP1 were used in the analysis with ACCORD results. For REP2 and REP3, metformin response was calculated as post–treatment HbA1c minus pre–treatment HbA1c.

To provide additional support for the role of the lead SNP, rs57081354, in modulating metformin action, we performed an association test between rs57081354 and baseline HbA1c in ACCORD individuals that were already on metformin at the start of the trial (REP4) and compared those results to the association of baseline HbA1c in individuals not on metformin at the start of the trial. All individuals in REP4 were excluded from the discovery cohort due to the lack of a recorded pre-treatment HbA1c, making it impossible to calculate a change in HbA1c upon treatment. For this analysis, an

association test was performed as described in the common variant analysis section described above with the exception of the phenotype being baseline level of HbA1c rather than change in HbA1c.

Each SNP with $p<5x10^{-8}$ in the discovery cohort, or with $p<5x10^{-6}$ in the discovery cohort and $q<0.01$ in at least one replication cohort, were meta-analyzed across available cohorts using PLINK(21,22). Meta-analysis with $p<0.05$ was considered to be statistically significant. Additional information regarding each cohort can be found in the Supplemental Material and a workflow describing how the replication was evaluated can be found in Supplemental Figure 4.

Dietary Analysis

Dietary questionnaires were assessed for 1600 participants that both consented to genotyping and had dietary records as part of the Health-Related Quality of Life Substudy within ACCORD. Scores in response to the dietary questionnaires were tested for association with the lead SNP, rs57081354, using a linear regression model and the same covariate selection process as described above. Details regarding this analysis can be found in the Supplemental Material.

Results

A total of 1312, 845, and 222 subjects were included in the common and rare variant analyses for all races combined, white, and black cohorts, respectively (Figure 1). Other racial groups were included in the all races combined group, but sample sizes were too small to perform stratified analyses. Variation was observed in HbA1c response (Figure 2). The mean change in HbA1c was -1.42% (95% CI -1.49, -1.36) for all races combined, -1.47% (95 %CI -1.55, -1.39) for white subjects only, and -1.35% (95% CI -1.50, -1.19) for black subjects only. The difference in the change in HbA1c between white and black subjects was not statistically significant($p=0.16$). Variables selected as being significantly associated with metformin response explained approximately 55% of the variation in metformin response, and were subsequently included in the common and rare variant analyses (Supplementary Tables 2-4). Briefly,

assignment to the intensive glycemia arm of the ACCORD trial, years with T2D, and pre-treatment HbA1c were significantly associated with metformin response in all tested groups.

Common Variant Analysis

A total of 316,203 genotyped, 547,639 meta-analyzed, and 7,567,403 imputed variants had MAF>3% and were included in the common variant analysis when all races were combined. A total of 1, 3, 7 SNPs were significantly associated with HbA1c response in all races combined, white subjects only, and black subjects only, respectively ($p<5x10^{-8}$). In all subjects combined, the lead SNP rs57081354 located within an intron in the *NBEA* gene associated with HbA1c response (Figure 3). 65 SNPs in 16 loci on 14 different chromosomes (in genes: *NBEA*, *LOC105374140*, *CXCL13*, *MCC*, *FAM189A1*, *LOC107984525*, *LOC105379109*, *KIAA1024*, *RBFOX1*, *LOC643339*, *UBOX5*, *NR2C1*, *BPHL*, *PRPF31*, *LOC105377165*, *ADAMTS9-AS2*) reached the threshold for suggestive significance ($p<5x10^{-6}$). For white subjects only, 3 SNPs in *KIAA1024* were associated with HbA1c response at genome-wide levels of significance ($p<5x10^{-8}$), and 23 SNPs in 7 loci (in genes: *KIAA1024*, *LOC283177*, *LOC105377165*, *ADAMTS9-AS2*, *CPA6*, *NR2C1*, *DPYD*) reached the threshold for suggestive significance ($p<5x10^{-6}$) (Supplementary Figure 1). When black subjects were analyzed separately, 2 SNPs in the *LOC102724874* locus were associated with HbA1c response at genome-wide levels ($p<5x10^{-8}$), and 53 SNPs in 8 loci (in genes: *LOC102724874*, *LOC105374308*, *LOC105369406*, *SLC35D3*, *RHPN2*, *ETS1*, *SDK2*, *GPC6*) reached the threshold for suggestive significance ($p<5x10^{-6}$) (Supplementary Figure 2). Lead SNPs associated with HbA1c response ($p<5x10^{-6}$) are presented in Tables 2-4.

Rare Variant Analysis

Rare variants (MAF$\leq$.03) in a total of 17,078 genes were tested for association with HbA1c in all subjects combined, and separately in white and black subjects. In the combined subjects' analysis, rare variants in *STAT3* were significantly associated with HbA1c response ($q<0.1$) (Supplementary Figure 3). Five rare variants were available for testing in *STAT3*: rs146620441 (monomorphic), rs140604473

(missense), rs149214040 (missense), rs114401618 (intronic), and rs17882069 (synonymous) (28). Importantly, these four SNPs only showed variation in black subjects, although there were no statistically significant rare variant associations observed in race-based analyses (q>0.2). We tested whether *STAT3* failed to yield significant associations in the stratified analysis due to the small number of black subjects (n=222) by oversampling black subjects until we had the same number of subjects as when all races were combined (n=1312). Oversampling to build a larger cohort resulted in a q=0.0016 versus a q=1 in the original cohort of black subjects, supporting the premise that this rare variant finding is relevant for black subjects even though it was detected when all subjects were combined due to an increase in statistical power.

Association with Dietary Phenotypes

NBEA has been previously reported to influence sugary food preference in animal models(29), and we further assessed whether the lead SNP in this locus, rs57081354, was associated with dietary scores. Data was available in 1553 ACCORD subjects (Table S6). The C-allele in rs57081354 that associated with worse metformin effect also associated with increased consumption of dessert (q=0.12). Subjects with the C-allele in rs57081354 were more likely to answer 'Yes' to the question, "Did you eat dessert?", within past three months (β=0.044, 95% CI= 0.012-0.075).

Replication

A total of 134 unique SNPs were associated with metformin response ($p<5x10^{-6}$) when all races were combined, in white subjects only, or in black subjects only. Of these SNPs, 81, 64, and 61 were available in REP1, REP2, and REP3 cohorts, respectively. The most significant finding in the discovery cohort was rs57081354 located in the *NBEA* gene ($p=4.02x10^{-8}$, β=0.323). This finding did not replicate in REP1 (p>0.05), and was not available for testing in REP2 or REP3. Out of the 26 SNPs in *NBEA* with $p<1x10^{-6}$, only rs1337379 was tested in REP2 and REP3, but was not associated with metformin response (p>0.05). SNPs in *NBEA* were not significantly associated with metformin response in ACCORD when

black or white subjects were tested separately, only when all races were combined. Interestingly, in the set of subjects who were already on metformin when they enrolled in ACCORD (REP4), rs57081354, was marginally associated with higher baseline HbA1c (p=0.09), whereas rs57081354 was not associated with baseline variation in HbA1c in those not on metformin (p=0.63). Although the phenotypes for REP1 and REP4 are not identical, meta-analysis of the ACCORD discovery cohort with REP1 and REP4 produced an association representing higher HbA1c on metformin (p=0.009)(Figure 3).

In ACCORD, rs254271, located in an intron in Pre-MRNA Processing Factor 31 (*PRPF31*), was associated with worse metformin response when all races were combined ($p=3.79x10^{-6}$). In the REP1 cohort, rs254271 was associated with worse metformin response when corrected for baseline HbA1c ($p=6.21x10^{-5}$) and without adjusting for baseline HbA1c($p=1.57x10^{-2}$). Genotypes for rs254271 were not available for analysis in the African American replication cohorts, REP2 and REP3. Meta-analysis of ACCORD with REP1 produced a significant association for rs254271 with worse metformin response ($p=1.2x10^{-8}$) (Figure 4).

In white subjects in ACCORD, rs2162145, located in Carboxypeptidase A6 (*CPA6*), was associated with better metformin response ($p=4.04x10^{-6}$). This finding was consistent with findings in African American subjects in REP2 (p=0.006). Although, rs2162145 was not significantly associated with metformin response in black subjects in ACCORD (p=0.14), it is possible that we were underpowered to detect an association due to the relatively small number of black subjects available (n=222), and that meta-analysis could increase statistical power through the inclusion of additional subjects. Meta-analysis of black subjects in ACCORD and those in REP2 and REP3 demonstrated the most significant association of rs2162145 with better metformin response (p=0.005) compared to meta-analysis of all white subjects (ACCORD white subjects and REP1) and all cohorts (ACCORD, REP1-3) with p=0.22 and p=0.09, respectively (Figure 4).

Discussion

T2D adversely affects the quality of life for millions of individuals and places a significant burden on the health care system in the United States and globally (30,31). Although many treatment options for T2D are available, metformin has remained the first-line treatment for T2D for decades. Nevertheless, the mechanisms by which metformin lowers blood glucose are not well-understood, and factors that may predict which patients will have optimal responses to metformin are also not well-understood. Here we use a GWAS approach in a large cohort of individuals with T2D in the ACCORD clinical trial to test for both common and rare variant SNPs that associate with change in HbA1c level in response to metformin treatment. To support our initial findings, we explored these associations in multiple independent cohorts.

A previous study conducted in a large cohort using electronic medical records (EMR) found that African-American patients on metformin had a larger reduction in HbA1c than European Americans(32). We failed to find a significant difference between metformin response in white subjects and black subjects in ACCORD with changes in HbA1c of -1.47% and -1.35%, respectively (p=0.16). It is possible that enrollment in a more structured clinical trial impacts these outcomes differently than the far less structured care reflected in EMR data. Additional research will be needed to definitely characterize the response of metformin across ethnicities. Previous GWAS have identified rs11212617 in the *ATM* locus, and rs8192675 in *SLC2A2* as being significantly associated with metformin response(12,13). However, neither SNP was significantly associated with metformin response here (p>0.1). These SNPS may have failed to replicate in the current study due to cohort differences and in particular, the sample size of the ACCORD cohort and the use of concomitant medications. Information regarding the prescribed dose for metformin was not recorded in ACCORD and may have impacted our findings. However, as described in the methods, we sought a minimal time window to assess the response of metformin while limiting the

likelihood of other medication changes. Additional details about how these studies compare to ACCORD can be found in the Supplemental Material.

In the analysis of subjects from all racial groups combined (n=1312) (Figure 1), those with the C-allele of rs254271, an intronic variant in *PRPF31*, had a worse metformin response at suggestive significance levels ($p=3.79x10^{-6}$, $\beta=0.16$). The effect of this variant was also observed in REP1, consisting of subjects of European ancestry ($p=6.21x10^{-5}$, $\beta= -0.073$). Interestingly, rs254271, is an expression quantitative trait locus in several tissues for PRPF31 and nearby genes, including NADH:Ubiquinone oxidoreductase, NDUFA3, which codes for a subunit of the Complex I of the respiratory chain(33–35). Metformin is known to inhibit Complex I of the respiratory chain to cause AMPK activation(5). In addition, previous studies have shown that *PRPF31*(+/-) knockout mice have increased total body fat ($p=1.78x10^{-6}$) and increased fasted circulating glucose ($p=5.73x10^{-6}$)(36).

Genotypes for rs254271 were not available for analysis in the African American replication cohorts, REP2 and REP3. The phenotype in REP1 represents HbA1c reduction, so the direction of the effect for rs254271 is consistent between ACCORD and REP1, and represents worse HbA1c response in individuals with the C allele (Figure 4). Mutations in *PRPF31* have been associated with an autosomal dominant rare Mendelian disease, retinitis pigmentosa(37). *PRPF31* was previously identified as a candidate gene for diabetic retinopathy, in a genome-wide linkage study in Mexican-Americans with T2D(38). We tested rs254271 for association with eye disorders in ACCORD, however, no significant associations were found ($p>0.05$)(data not shown). Further research will be needed to determine what role SNPs in *PRPF31* may have in metformin drug response or in diabetic retinopathy.

In the analysis of only white subjects in ACCORD, rs2162145, in the *CPA6* gene, was associated with metformin response at suggestive significance levels ($p=4.04x10^{-6}$, $\beta= -0.197$). This same SNP was associated with metformin response in the REP2 cohort ($p=0.006$, $\beta= -0.127$), but not in REP1 or REP3

(p>0.05) (Figure 4). To our knowledge, SNPs in *CPA6* have not previously been associated with metformin response, however, rs2162145 has a weak significant association with fasting glucose (p=0.02)(39). In addition, a meta-analysis of T2D in 17 GWAS comprising 8,284 cases and 15,543 controls in African Americans identified rs7003257, an intron variant in *CPA6*, as being associated with T2D ($p=1.17x10^{-6}$)(40) and this SNP has also been previously reported to affect 2 hour glucose levels(41). We were underpowered to detect an association in black subjects. However, meta-analyzing all available black subjects in ACCORD, REP2, and REP3 improved the statistical power so an association was detectable (p=0.005). It will be necessary for additional studies to investigate the extent to which *CPA6* is involved with metformin drug response.

A genotyped SNP in the *NBEA* gene on chr13, rs57081354, was associated with change in HbA1c levels at genome-wide significance ($p=4.02x10^{-8}$, $\beta=0.323$) in our discovery cohort (Figure 3). SNPs in *NBEA* did not replicate in REP1-3 cohorts. Rs57081354 was marginally associated with baseline HbA1c levels in individuals who were taking metformin at the time they enrolled in ACCORD, with the C-allele displaying higher HbA1c values (p=0.09) (Figure 3c). NBEA is a protein kinase A anchor protein that is expressed at low levels ubiquitously and at higher levels in the brain, pituitary, and beta cells in the pancreas(42). Interestingly, haplotype-insufficient *Nbea*(+/-) mice prefer sugary foods (glucose and fructose) more than their wild-type controls, gain more weight per unit of food than wild type mice, and have higher insulin and leptin levels(29). Subsequently, we tested for associations between rs57081354 and 19 dietary scores recorded for a subset of 1553 ACCORD subjects and found that rs57081354 was associated with an increase in dessert consumption ($q<0.2$) in those with C-allele (Table S6). An expanded summary of the literature regarding *NBEA* can be found in the Supplementary Material.

Additionally, we tested rare variants for associations with change in HbA1c, and *STAT3* when all races were combined, was the only gene that achieved statistical significance for metformin treatment response ($q<0.1$). However, as described in the results, black subjects were the only population with

rare variation in *STAT3* and *STAT3* was not detected in the stratified analysis likely due to the small number of subjects available for analysis (n=222). In this same cohort, rare variants in *STAT3* were not associated with baseline HbA1c levels(q>0.2). *STAT3* encodes a ubiquitously expressed transcription factor with broad activities in metabolism, immunity, and cell growth(43). STAT3 is an essential signal transducer for the key metabolic hormone leptin and the pleiotropic cytokine interleukin 6 with broad effects on feeding behavior and fuel homeostasis. Interestingly, metformin has been shown to interact with STAT3 signaling in chemotherapeutic applications and in an inflammatory bowel disease model(44). To our knowledge, interactions between metformin and STAT3 signaling have not been investigated with regards to glucose homeostasis.

Here we present evidence for *PRPF31, CPA6*, and *STAT3* as being involved in novel glucose-lowering mechanisms for metformin. Additional investigation will also be required to confirm the clinical impact of the identified variants with regards to metformin therapy in diabetic subjects. Although prior studies suggest roles for *NBEA* and *STAT3* genes in glucose metabolism, additional research will be required to investigate mechanisms by which they might interact with metformin. Such studies are likely to provide new insight into mechanisms regulating glucose metabolism and may point the way towards novel therapeutic targets for more precise interventions in T2D.

Acknowledgments

Author contributions

D.M.R., S.W.Y., M.A.H., H.L.M., M.J.W., J.B.B., and A.M.R. wrote the article. D.M.R., K.M.G., E.R.P., M.M.H, M.K., A.D., J.C.M., H.L.M., M.J.W., J.B.B., and A.M.R. designed the research. D.M.R., S.W.M., T.M.H., S.W.Y., K.Z., H.S.H., H.G., M.J.W., and A.M.R. performed the research. D.M.R., J.R.J., S.W.M., S.W.Y., and K.Z. analyzed the data.

Funding

This research was supported by an NHLBI grant to the University of North Carolina at Chapel Hill (5R01HL110380-04) and R01 HL110400 to the Joslin Diabetes Center and the University of Virginia.

S.W.Y and K.M.G. were supported by GM61390 and GM117163. E.R.P holds a Wellcome Trust Investigator Award (102820/Z/13/Z).

Conflict of Interest

The authors declare no conflict of interest.

Prior Presentation

Parts of this study were presented in abstract form at the American Society for Clinical Pharmacology and Therapeutics Annual Meeting. Washington DC. March 15, 2017.

References

1. Centers for Disease Control and Prevention. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014. Atlanta, GA: US Department of Health and Human Services; 2014.

2. World Health Organization. WHO Global report on diabetes [Internet]. [cited 2017 Feb 1]. Available from: http://www.who.int/diabetes/global-report/en/

3. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart Disease and Stroke Statistics—2016 Update A Report From the American Heart Association. Circulation. 2015;CIR–0000000000000350.

4. Rena G, Pearson ER, Sakamoto K. Molecular mechanism of action of metformin: old or new insights? Diabetologia. 2013 Sep 1;56(9):1898–906.

5. Zhou G, Myers R, Li Y, Chen Y, Shen X, Fenyk-Melody J, et al. Role of AMP-activated protein kinase in mechanism of metformin action. J Clin Invest. 2001 Oct;108(8):1167–74.

6. Foretz M, Hébrard S, Leclerc J, Zarrinpashneh E, Soty M, Mithieux G, et al. Metformin inhibits hepatic gluconeogenesis in mice independently of the LKB1/AMPK pathway via a decrease in hepatic energy state. J Clin Invest. 2010 Jul 1;120(7):2355–69.

7. Buse JB, DeFronzo RA, Rosenstock J, Kim T, Burns C, Skare S, et al. The primary glucose-lowering effect of metformin resides in the gut, not the circulation: results from short-term pharmacokinetic and 12-week dose-ranging studies. Diabetes Care. 2016;39(2):198–205.

8. McCreight LJ, Bailey CJ, Pearson ER. Metformin and the gastrointestinal tract. Diabetologia. 2016;59(3):426–435.

9. Baur JA, Birnbaum MJ. Control of gluconeogenesis by metformin: does redox trump energy charge? Cell Metab. 2014;20(2):197–199.

10. Ferrannini E. The target of metformin in type 2 diabetes. N Engl J Med. 2014;371(16):1547–1548.

11. Zhou K, Donnelly L, Yang J, Li M, Deshmukh H, Van Zuydam N, et al. Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. Lancet Diabetes Endocrinol. 2014 Jun;2(6):481–7.

12. Zhou K, Yee SW, Seiser EL, van Leeuwen N, Tavendale R, Bennett AJ, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. Nat Genet. 2016 Sep;48(9):1055–9.

13. GoDARTS and UKPDS Diabetes Pharmacogenetics Study Group, Wellcome Trust Case Control Consortium 2, Zhou K, Bellenguez C, Spencer CCA, Bennett AJ, et al. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. Nat Genet. 2011 Feb;43(2):117–20.

14. Buse JB. Action to Control Cardiovascular Risk in Diabetes (ACCORD) Trial: Design and Methods. Am J Cardiol. 2007 Jun 18;99(12, Supplement):S21–33.

15. Goff DC, Gerstein HC, Ginsberg HN, Cushman WC, Margolis KL, Byington RP, et al. Prevention of cardiovascular disease in persons with type 2 diabetes mellitus: current knowledge and rationale for the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial. Am J Cardiol. 2007;99(12):S4–S20.

16. Gerstein HC, Riddle MC, Kendall DM, Cohen RM, Goland R, Feinglos MN, et al. Glycemia treatment strategies in the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial. Am J Cardiol. 2007;99(12):S34–S43.

17. Ginsberg HN, Bonds DE, Lovato LC, Crouse JR, Elam MB, Linz PE, et al. Evolution of the lipid trial protocol of the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial. Am J Cardiol. 2007;99(12):S56–S67.

18. Graham* HT, Rotroff* DM, Marvel SW, Buse JB, Havener TM, Wilson AG, et al. Incorporating Concomitant Medications into Genome-Wide Analyses for the Study of Complex Disease and Drug Response. Front Genet. 2016;7:138.

19. Marvel* SW, Rotroff* DM, Wagner MJ, Buse JB, Havener TM, McLeod HL, et al. Common and rare genetic markers of lipid variation in subjects with type 2 diabetes from the ACCORD clinical trial. PeerJ. 2017 May 2;5:e3187.

20. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/. [Internet]. R Foundation for Statistical Computing, Vienna, Austria; 2014. Available from: http://www.R-project.org

21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–575.

22. Shaun Purcell. PLINK v1.07 [Internet]. 2009. Available from: http://pngu.mgh.harvard.edu/purcell/plink/

23. Dai H, Leeder JS, Cui Y. A modified generalized Fisher method for combining probabilities from dependent tests. Front Genet [Internet]. 2014 Feb 20 [cited 2014 Oct 1];5. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3929847/

24. Dabney A, Storey JD, Warnes G. Q-value estimation for false discovery rate control. Medicine (Baltimore). 2004;344:539–548.

25. Storey JD. A direct approach to false discovery rates. J R Stat Soc Ser B Stat Methodol. 2002;64(3):479–498.

26. Goswami S, Yee SW, Xu F, Sridhar SB, Mosley JD, Takahashi A, et al. A Longitudinal HbA1c Model Elucidates Genes Linked to Disease Progression on Metformin. Clin Pharmacol Ther. 2016 Nov;100(5):537–47.

27. Banda Y, Kvale MN, Hoffmann TJ, Hesselson SE, Ranatunga D, Tang H, et al. Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. Genetics. 2015 Aug;200(4):1285–95.

28. National Center for Biotechnology Information, National Library of Medicine. Database of Single Nucleotide Polymorphisms (dbSNP) [Internet]. [cited 2017 May 16]. Available from: http://www.ncbi.nlm.nih.gov/SNP/

29. Olszewski PK, Rozman J, Jacobsson JA, Rathkolb B, Strömberg S, Hans W, et al. Neurobeachin, a Regulator of Synaptic Protein Targeting, Is Associated with Body Fat Mass and Feeding Behavior in Mice and Body-Mass Index in Humans. PLOS Genet. 2012 Mar 15;8(3):e1002568.

30. da Rocha Fernandes J, Ogurtsova K, Linnenkamp U, Guariguata L, Seuring T, Zhang P, et al. IDF Diabetes Atlas estimates of 2014 global health expenditures on diabetes. Diabetes Res Clin Pract. 2016 Jul;117:48–54.

31. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. The Lancet. 2011 Jul 8;378(9785):31–40.

32. Williams LK, Padhukasahasram B, Ahmedani BK, Peterson EL, Wells KE, González Burchard E, et al. Differing Effects of Metformin on Glycemic Control by Race-Ethnicity. J Clin Endocrinol Metab. 2014 Sep 1;99(9):3160–8.

33. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, et al. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. PloS One. 2010 May 18;5(5):e10693.

34. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013 Jun;45(6):580–5.

35. Owen MR, Doran E, Halestrap AP. Evidence that metformin exerts its anti-diabetic effects through inhibition of complex 1 of the mitochondrial respiratory chain. Biochem J. 2000 Jun 15;348 Pt 3:607–14.

36. Brown SDM, Moore MW. The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping. Mamm Genome. 2012 Oct 1;23(9–10):632–40.

37. Waseem NH, Vaclavik V, Webster A, Jenkins SA, Bird AC, Bhattacharya SS. Mutations in the Gene Coding for the Pre-mRNA Splicing Factor, PRPF31, in Patients with Autosomal Dominant Retinitis Pigmentosa. Invest Ophthalmol Vis Sci. 2007 Mar 1;48(3):1330–4.

38. Hallman DM, Boerwinkle E, Gonzalez VH, Klein BEK, Klein R, Hanis CL. A Genome-Wide Linkage Scan for Diabetic Retinopathy Susceptibility Genes in Mexican Americans With Type 2 Diabetes From Starr County, Texas. Diabetes. 2007 Apr 1;56(4):1167–73.

39. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 2010 Feb;42(2):105–16.

40. Ng MCY, Shriner D, Chen BH, Li J, Chen W-M, Guo X, et al. Meta-Analysis of Genome-Wide Association Studies in African Americans Provides Insights into the Genetic Architecture of Type 2 Diabetes. PLOS Genet. 2014 Aug 7;10(8):e1004517.

41. Saxena R, Hivert M-F, Langenberg C, Tanaka T, Pankow JS, Vollenweider P, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat Genet. 2010 Feb;42(2):142–8.

42. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419.

43. Johnston PA, Grandis JR. STAT3 SIGNALING: Anticancer Strategies and Challenges. Mol Interv. 2011 Feb 1;11(1):18.

44. Deng X-S, Wang S, Deng A, Liu B, Edgerton SM, Lind SE, et al. Metformin targets Stat3 to inhibit cell growth and induce apoptosis in triple-negative breast cancers. Cell Cycle Georget Tex. 2012 Jan 15;11(2):367–76.

45. Feng Y, Ke C, Tang Q, Dong H, Zheng X, Lin W, et al. Metformin promotes autophagy and apoptosis in esophageal squamous cell carcinoma by downregulating Stat3 signaling. Cell Death Dis. 2014 Feb 27;5(2):e1088.

46. Tan X-L, Bhattacharyya KK, Dutta SK, Bamlet WR, Rabe KG, Wang E, et al. Metformin suppresses pancreatic tumor growth with inhibition of NFκB/STAT3 inflammatory signaling. Pancreas. 2015 May;44(4):636–47.

47. Lee S-Y, Lee SH, Yang E-J, Kim E-K, Kim J-K, Shin D-Y, et al. Metformin Ameliorates Inflammatory Bowel Disease by Suppression of the STAT3 Signaling Pathway and Regulation of the between Th17/Treg Balance. PLOS ONE. 2015 Sep 11;10(9):e0135858.

48. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26(18):2336–2337.

Tables

Table 1. Cohort Demographics and Concomitant Medications

	All Races Combined (N=1312)	*White Subjects Only (N=845)*	*Black Subjects Only (N=222)*
Intensive glycemia arm (%)	56.33	59.41	53.15
Standard glycemia arm (%)	43.67	40.59	46.85
Intensive blood pressure arm (%)	23.86	23.67	34.23
Standard blood pressure arm (%)	23.78	21.89	33.33
Fibrate lipid treatment arm (%)	25.99	25.68	14.41
Placebo lipid treatment arm (%)	26.37	28.76	18.01
Female (%)	39.48	36.33	52.70
Age (mean [95% CI])	62.72 [62.36-63.10]	63.14 [62.70-63.58]	62.25 [61.37-63.12]
BMI (mean [95% CI])	32.43 [32.13-32.73]	33.18 [32.82-33.54]	32.67 [31.94-33.39]
Years with T2D (mean [95% CI])	10.10 [9.67-10.51]	9.73 [9.23-10.24]	10.82 [9.76-11.87]
History of cardiovascular disease (%)	33.54	35.27	29.28
Concomitant Medications[*]			
Angiotensin type 2 antagonists (%)	16.39	15.98	16.67
Ace inhibitors (%)	46.19	46.98	52.25
Alpha-glucosidase inhibitors (%)	0.53	0.47	0.45
Cholesterol absorption inhibitors (%)	2.13	2.49	1.80
Statin (%)	49.70	51.83	30.18
Lisinopril (%)	16.46	16.45	25.23
Loop diuretics (%)	8.23	8.99	10.36
Meglitinides (%)	2.97	2.37	3.15
Nitrates (%)	5.72	6.51	3.60
Other Diabetes Treatments (%)	2.67	2.84	4.05
Sulfonylureas (%)	52.44	51.83	45.05
Thiazolidinediones (%)	31.63	33.02	29.28
Insulin (%)	38.11	38.11	45.95

[*]Tabulated percentages represent the percentage of subjects with at least one record of taking the medication, or a medication in the drug class, during the metformin treatment window. This is not an

exhaustive list of medications recorded in ACCORD, rather a representative list of the commonly used medications in ACCORD.

Table 1. Lead SNPs Associated with HbA1c response in Common Variant Analysis in All Races Combined ($p < 5x10^{-6}$)

SNP	Chromosome	Position	Gene	Type	MAF	Beta	P Value
rs7371352	2	205916609	NA	IMPU	0.461	0.166	1.29E-06
rs12694529	2	220685415	NA	META	0.131	0.232	1.06E-06
rs2371651	3	64899735	*ADAMTS9-AS2*	IMPU	0.346	0.171	4.89E-06
rs36050186	3	73884807	*LOC105377165*	IMPU	0.061	0.315	4.46E-06
rs28535480	3	144238665	*LOC105374140*	IMPU	0.447	0.184	4.12E-07
rs7669299	4	77525989	*CXCL13*	GENO	0.106	-0.348	5.88E-07
rs79827403	4	140248372	NA	IMPU	0.046	-0.386	2.19E-06
rs6884548	5	104706194	*LOC105379109*	IMPU	0.038	0.439	2.64E-06
rs56061109	5	113100738	*MCC*	IMPU	0.048	0.394	6.00E-07
rs6903843	6	3125135	*BPHL*	IMPU	0.031	0.587	3.36E-06
rs79573377	6	91226278	NA	IMPU	0.033	0.437	3.19E-06
rs59506474	6	136927217	NA	IMPU	0.066	-0.355	1.57E-06
rs2578120	10	78453110	NA	IMPU	0.358	-0.160	2.53E-06
rs7307035	12	53778761	*LOC107984525*	META	0.033	-0.440	1.77E-06
rs12300320	12	93077216	*LOC643339*	IMPU	0.053	-0.386	3.01E-06
rs12827634	12	95026750	*NR2C1*	IMPU	0.362	0.154	3.18E-06
rs57081354	13	35202456	*NBEA*	GENO	0.077	0.326	4.02E-08
rs4144603	14	96729576	NA	META	0.405	0.159	1.36E-06
rs7173199	15	29424654	*FAM189A1*	IMPU	0.046	0.419	1.77E-06
rs182384419	15	79438983	*KIAA1024*	IMPU	0.043	0.394	2.66E-06
rs74007109	16	6852072	*RBFOX1*	IMPU	0.040	0.407	2.92E-06
rs2529698	17	12350355	NA	META	0.061	0.331	9.24E-07
rs254271	19	54630757	*PRPF31*	IMPU	0.330	0.163	3.79E-06
rs6139020	20	3156571	*UBOX5*	IMPU	0.102	0.257	3.08E-06

Table 2. Lead SNPs Associated with HbA1c response in Common Variant Analysis in White Subjects Only ($p < 5x10^{-6}$)

SNP	Chromosome	Position	Gene	Type	MAF	Beta	P Value
rs12047072	1	97401162	*DPYD*	META	0.184	-0.215	4.66E-06
rs13034779	2	234091216	NA	IMPU	0.305	0.239	6.24E-07
rs2371651	3	64899735	*ADAMTS9-AS2*	IMPU	0.431	0.192	3.71E-06
rs36050186	3	73884807	*LOC105377165*	IMPU	0.082	0.332	2.47E-06
rs71140472	3	144248959	NA	IMPU	0.461	0.177	3.30E-06
rs2929535	8	67754375	*CPA6*[a]	META	0.307	-0.188	3.47E-06
rs111976192	8	83716899	NA	IMPU	0.267	0.195	4.99E-06
rs10904150	10	3970713	NA	IMPU	0.072	-0.336	4.66E-06
rs4567482	11	134474052	*LOC283177*	META	0.123	-0.286	1.66E-07
rs7298631	12	95063515	*NR2C1*	IMPU	0.397	0.176	4.51E-06
rs4775474	15	62150069	NA	META	0.355	0.189	2.07E-06
rs182384419	15	79438983	*KIAA1024*	IMPU	0.054	0.503	1.53E-08

[a] rs2929535 does not fall in *CPA6*, but rs2162145 ($p=4.04x10^{-6}$) is in the same peak and is located in *CPA6*.

Table 3. Lead SNPs Associated with HbA1c response in Common Variant Analysis in Black Subjects Only ($p < 5x10^{-6}$)

SNP	Chromosome	Position	Gene	Type	MAF	Beta	P Value
rs7573365	2	183278246	NA	META	0.325	-0.379	3.67E-06
rs12496540	3	32062976	NA	IMPU	0.352	0.388	4.98E-06
rs62284968	3	197429996	*LOC105374308*	IMPU	0.040	1.066	1.20E-07
rs9295608	6	23977533	NA	GENO	0.208	-0.462	1.35E-06
rs111792250	6	136927930	*SLC35D3*[a]	IMPU	0.192	-0.528	7.97E-07
rs4413694	7	142430649	NA	IMPU	0.458	-0.346	4.43E-06
rs201739756	8	73149128	NA	IMPU	0.054	0.875	2.04E-06
rs145554029	8	77485920	*LOC102724874*[b]	IMPU	0.035	-1.221	8.76E-09
rs10897557	11	80135250	*LOC105369406*	IMPU	0.319	-0.432	6.73E-07
rs7130041	11	128552747	*ETS1*	IMPU	0.348	-0.406	3.02E-06
rs76853142	13	93580057	*GPC6*	IMPU	0.041	-0.940	4.54E-06
rs1449836	14	42994192	NA	META	0.471	-0.353	1.43E-06
rs62072150	17	73412120	*SDK2*	IMPU	0.451	-0.423	4.12E-06
rs185341538	19	32981426	*RHPN2*	IMPU	0.291	-0.428	2.99E-06
rs34600526	19	33504770	NA	IMPU	0.295	-0.395	4.49E-06

[a]rs111792250 does not fall in *SLC35D3*, but rs7763578 ($p=1.47x10^{-6}$) is in the same peak and is located in *SLC35D3*. [b] rs145554029 does not fall in *LOC102724874*, but rs75121529 ($p=1.22x10^{-8}$) is in the same peak and is located in *LOC102724874*.

Figure Legends

Figure 1. Study design and quality control workflow. Boxes represent the step taken in the workflow. The number of remaining subjects after a quality control step can be found in the relevant attached circle.

Figure 2. Distributions of HbA1c response to metformin treatment. A) HbA1c response when all races are combined. B) HbA1c response in only white participants. C) HbA1c response in only black participants.

Figure 3. HbA1C response to metformin treatment is diminished by SNPs in *NBEA*. A) Manhattan plot of HbA1c response to metformin treatment when all races were combined (N=1312). The horizontal dashed lines represent suggestive statistical significance ($p=5x10^{-6}$) and genome-wide significance ($p=5x10^{-8}$). Each point represents a single SNP. Genotyped SNPs are represented with black dots, imputed SNPs are represented by blue dots, and SNPs meta-analyzed from subjects genotyped in Set 1 subjects and imputed in Set 2 subjects are represented by pink dots. B) A LocusZoom (48) plot demonstrating the lead SNP, rs57081354, association with change in HbA1c when all races are combined. Linkage disequilibrium was calculated using the 1000 genomes reference panel using the hg19/1000 Genomes Nov 2014 EUR selection, matching the panel used to perform imputation in ACCORD. C) Forest plot demonstrating associations of rs57081354 in available replication cohorts. The beta values for REP1 were changed to the additive inverse to maintain a consistent direction of effect. A larger beta value means a diminished metformin response. Data for rs57081354 were not available for REP2 and REP3Bars represent the 95% confidence interval. D) Change in HbA1c by rs57081354 genotype as represented by violin plots.

Figure 4. Forest plot of associations with replicated SNPs in REP1, REP2, and REP3. A) Discovery and replication results for rs2162145, located in the *CPA6* gene. B) Discovery and replication results for rs254271, an intronic variant in *PRPF31*. A larger beta value means a diminished metformin response. The beta values for REP1 were changed to the additive inverse to maintain a consistent direction of effect. Bars represent the 95% confidence interval.