Variation in the Glucose Transporter gene SLC2A2 is associated with glycaemic response to metformin

Zhou, Kaixin; Yee, Sook Wah; Seiser, Eric L; van Leeuwen, Nienke; Tavendale, Roger; Bennett, Amanda J; Groves, Christopher J; Coleman, Ruth L; van der Heijden, Amber A; Beulens, Joline W; de Keyser, Catherine E; Zaharenko, Linda; Rotroff, Daniel M; Out, Mattijs; Jablonski, Kathleen A; Chen, Ling; Javorský, Martin; Židzik, Jozef; Levin, Albert M; Williams, L Keoki; Dujic, Tanja; Semiz, Sabina; Kubo, Michiaki; Chien, Huan-Chieh; Maeda, Shiro; Witte, John S; Wu, Longyang; Tká, Ivan; Kooy, Adriaan; van Schaik, Ron H N; Stehouwer, Coen D A; Logie, Lisa; Sutherland, Calum; Klovins, Janis; Pirags, Valdis; Hofman, Albert; Stricker, Bruno H; Motsinger-Reif, Alison A; Wagner, Michael J; Innocenti, Federico; Hart, Leen M 't; Holman, Rury R; McCarthy, Mark I; Hedderson, Monique M; Palmer, Colin N A; Florez, Jose C; Giacomini, Kathleen M; Pearson, Ewan R; MetGen Investigators

Published in:
Nature Genetics

DOI:
10.1038/ng.3632

Publication date:
2016

Document Version
Peer reviewed version

Title:

Variation in the Glucose Transporter gene *SLC2A2* is associated with glycaemic response to metformin

Authors:

Kaixin Zhou[1,47], Sook Wah Yee[2,47], Eric L. Seiser[3], Nienke van Leeuwen[4], Roger Tavendale[1], Amanda J Bennett[5], Christopher J. Groves[5], Ruth L. Coleman[6], Amber A. van der Heijden[7], Joline W. Beulens[8,9], Catherine E. de Keyser[10], Linda Zaharenko[11,12], Daniel M. Rotroff[13,14], Mattijs Out[15,16], Kathleen A. Jablonski[17], Ling Chen[18], Martin Javorský[19], Jozef Židzik[19], Albert M. Levin[20], L. Keoki Williams[21,22], Tanja Dujic[1,23], Sabina Semiz[23,24], Michiaki Kubo[25], Huan-Chieh Chien[2], Shiro Maeda[26,27], John S. Witte[28-31], Longyang Wu[28], Ivan Tkáč[19], Adriaan Kooy[15,16], Ron H.N. van Schaik[32], Coen D.A. Stehouwer[33], Lisa Logie[1], MetGen investigators[34], DPP investigators[35], ACCORD investigators[36], Calum Sutherland[1], Janis Klovins[11,12], Valdis Pirags[12,37,38], Albert Hofman[10], Bruno H. Stricker[10,39], Alison A. Motsinger-Reif[13], Michael J. Wagner[40], Federico Innocenti[3], Leen M. 't Hart[4,8,41], Rury R. Holman[6], Mark I. McCarthy[5,42,43], Monique M. Hedderson[44], Colin N.A. Palmer[1], Jose C. Florez[18,45,46], Kathleen M. Giacomini[2,29,48], Ewan R. Pearson[1,48]

Affiliations:

[1]School of Medicine, University of Dundee, Dundee, UK. [2]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, USA. [3]Division of Pharmacotherapy and Experimental Therapeutics, Center for Pharmacogenomics and Individualized Therapy, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA. [4]Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, The Netherlands. [5]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK. [6]Diabetes Trials Unit, Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK. [7]Department of General Practice, EMGO+ Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands. [8]Department of Epidemiology & Biostatistics, EMGO+ Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands. [9]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands. [10]department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands. [11]Latvian Genome Data Base (LGDB), Latvia. [12]Latvian Biomedical Research and Study centre, Latvia. [13]Bioinformatics Research Center, North Carolina State University, Raleigh NC, USA. [14]Department of Statistics, North Carolina State University, Raleigh NC, USA. [15]Treant Zorggroep, Location Bethesda, Hoogeveen, The Netherlands. [16]Bethesda Diabetes Research Centre, Hoogeveen, The Netherlands. [17]The Biostatistics Center, George Washington University, Rockville, MD, USA. [18]Diabetes Unit and Center for Human Genetic Research, Massachusetts General Hospital. [19]Faculty of Medicine, Šafárik University, Košice, Slovakia. [20]Department of Public Health Sciences, Henry Ford Health System, Detroit, MI, USA. [21]Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, USA. [22]Department of Internal Medicine, Henry Ford Health System, Detroit, MI, USA. [23]Faculty of Pharmacy, University of Sarajevo, Sarajevo, Bosnia and Herzegovina. [24]Faculty of Engineering and

Natural Sciences, International University of Sarajevo, Sarajevo, Bosnia and Herzegovina. [25]RIKEN Center for Integrative Medical Sciences (IMS), Yokohama City, Japan. [26]Department of Advanced Genomic and Laboratory Medicine, Graduate School of Medicine, University of the Ryukyus, Nishihara, Japan. [27]Division of Clinical Laboratory and Blood Transfusion, University of the Ryukyus Hospital, Nishihara, Japan. [28]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California. [29]Institute for Human Genetics, University of California, San Francisco, San Francisco, California. [30]Department of Urology, University of California, San Francisco, San Francisco, California. [31]UCSF Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, California. [32] Department of Clinical Chemistry, Erasmus University Medical Center, Rotterdam, The Netherlands. [33]Department of Internal Medicine and Cardiovascular Research Institute Maastricht, Maastricht University Medical Center, Maastricht, The Netherlands. [34-36]Full lists of members and affiliations appears in the Supplementary Note. [37]Faculty of Medicine, University of Latvia, Riga, Latvia. [38]Department of Endocrinology, Pauls Stradins Clinical University Hospital, Riga, Latvia. [39]Inspectorate of Healthcare, 6401 DA Heerlen, the Netherlands. [40]Center for Pharmacogenomics and Individualized Therapy, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [41]Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. [42]Wellcome Trust Centre for Human Genetics, University of Oxford. [43]Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK. [44]Division of Research, Kaiser Permanente Northern California, Oakland, California, USA. [45]Programs in Metabolism and Medical & Population Genetics, Broad Institute, Cambridge, MA 02142, USA. [46]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA. [47]These authors contributed equally to this work. [48]These authors jointly directed this work. Correspondence should be addressed to E.R.P.(e.z.pearson@dundee.ac.uk) or K.M.G.(kathy.giacomini@ucsf.edu).

Abstract

Metformin is the first-line antidiabetic drug with over 100 million users worldwide, yet its mechanism of action remains unclear[1]. Here the Metformin Genetics (MetGen) Consortium reports a three-stage genome wide association study (GWAS), consisting of 13,123 participants of different ancestries. The C-allele of rs8192675 in the intron of *SLC2A2*, which encodes the facilitated glucose transporter GLUT2, was associated with a 0.17% ($p=6.6x10^{-14}$) greater metformin induced HbA1c reduction in 10,577 participants of European ancestry. rs8192675 is the top cis-eQTL for *SLC2A2* in 1,226 human liver samples, suggesting a key role for hepatic GLUT2 in regulation of metformin action. In obese individuals C-allele homozygotes at rs8192675 had a 0.33% (3.6mmol/mol) greater absolute HbA1c reduction than T-allele homozygotes.This is about half the effect seen with the addition of a DPP-4 inhibitor, and equates to a dose difference of 550mg of metformin, suggesting rs8192675 as a potential biomarker for stratified medicine.

Main text

Metformin was commercialized before the modern era of target-based drug discovery. It typically reduces HbA1c by 1~1.5% (11~16mmol/mol) and has an excellent safety record, but considerable variation exists in how well patients respond to metformin[2,3]. We have recently established that genetic factors influence glycaemic response to metformin, with many common variants across the genome together explaining a significant proportion of the variation, ranging from 21% to 34%, depending on how glycaemic response was measured[4]. Hypothesis-driven studies of pharmacokinetic variants have shown no consistent results[5-10]. The only GWAS published to date revealed an association with rs11212617 near the ATM locus, which has been further replicated[11,12].

Here we extended the previous GWAS by an additional 345 samples to a screening set of 1,373 participants. As in our previous report[12], rs11212617 remained the top signal with no other genome-wide significant hit (Supplementary Figure 1). A systematic three-stage replication was undertaken, with the work flow shown in Supplementary Figure 2. Only rs8192675 in the intron of *SLC2A2* was replicated through the first two stages with a combined $p=1x10^{-7}$ derived from 3,456 participants (Supplementary Data and Supplementary Table 1).

The final replication of rs8192675 was performed as a meta-analysis by the MetGen Consortium. Measures of glycaemic response to metformin were aligned across the cohorts as the absolute HbA1c reduction (expressed as reduction in %HbA1c). Within each cohort, associations with rs8192675 were tested with two multiple linear models with or without the adjustment of baseline HbA1c, in addition to other available clinical covariates (Supplementary Table 2). In the meta-analysis of 10,557 participants of European ancestry (Figure 1), each copy of the C-allele was associated with a greater HbA1c reduction of 0.07% ($p=2x10^{-8}$, $p_{het}=0.35$) when adjusting for baseline HbA1c; whilst without adjustment the allelic effect of C-allele was 0.17% ($p=6.6x10^{-14}$, $p_{het}=0.52$). There was no effect of rs8192675 on the efficacy of metformin in delaying progression to diabetes, or on metformin efficacy in a small insulin treated cohort (Supplementary table 3).

We tested the pharmacogenetic effect of rs8192675 in 2,566 participants of non-European ancestries (Supplementary Table 4). The meta-analysis showed the C-allele was associated with a 0.08% greater HbA1c reduction ($p=0.006$, $p_{het}=0.63$) when adjusting for baseline HbA1c; whilst the allelic effect of the C-allele was 0.15% ($p=0.005$, $p_{het}=0.95$) without the baseline adjustment. In the meta-analysis of 13,123 participants of any ancestry (data not shown), no genetic heterogeneity ($p_{het}>0.29$) was observed between different ethnic groups despite the C-allele frequency ranging from 24% in Latino to around 70% in African Americans.

We examined whether rs8192675 had an impact on baseline HbA1c, because the effect sizes of its association with glycaemic response to metformin differed depending on whether adjusting for the baseline HbA1c. In the 10,557 participants of European ancestry, the C-allele was associated with a 0.13% ($p=2.6x10^{-8}$) higher baseline HbA1c but a 0.04% (p=0.007) lower on-treatment HbA1c, which together contributed to the observed 0.17% ($p=6.6x10^{-14}$) pharmacogenetic impact on HbA1c reduction in the model without baseline adjustment (Supplementary Figure 3).

Given the association of rs8192675 with HbA1c prior to treatment with metformin, we assessed whether this variant was marking a general ability to respond to any antihyperglycaemic treatment. Therefore we studied the pharmacogenetic impact of rs8192675 in 2,654 participants treated with sulfonylureas (Supplementary Table 5), another commonly used class of antidiabetic drug[13,14]. As in metformin users, the C-allele was also associated with a higher baseline HbA1c in these sulfonylureas users (beta=0.15%, $p=3.1x10^{-4}$). However, in contrast to metformin, the C-allele remained associated with a higher on-treatment HbA1c (beta=0.09%, p=0.006) in these sulfonylureas users, which resulted in no net pharmacogenetic impact (beta=0.04%, p=0.44) on sulfonylurea induced HbA1c reduction. These data suggest that rs8192675 is marking a genetic defect in glucose metabolism in type 2 diabetes that is ameliorated by metformin treatment but not by sulfonylurea treatment. The fact that rs8192675 is not associated with sulfonylurea response strongly supports a specific role for this variant on glycaemic response to metformin, rather than simply reflecting the higher pre-treatment (baseline) HbA1c seen within carriers of this C-allele. In addition, the association with metformin induced HbA1c reduction remain significant after adjustment for baseline HbA1c, corroborating a specific effect on response beyond its effect on baseline glycaemia.

Metformin is particularly recommended for the treatment of diabetes in obese individuals due to its beneficial effect on body weight[15-17]. Therefore, we explored whether the pharmacogenetic impact of rs8192675 varied by BMI in the MetGen cohorts (n=7581). BMI is associated with HbA1c reduction (beta=-0.01%; $p=1.7x10^{-4}$) but not rs8192675 genotype (p=0.52). Adjusting for BMI does not attenuate the observed pharmacogenetic effect of rs8192675 (Supplementary Table 6). When participants were stratified into non-obese (BMI<30 kg/m^2) and obese groups (BMI≥30kg/m^2), there was a significant (p=0.02) gene by BMI group interaction (Figure 2). The pharmacogenetic effect size of the C-allele was 0.13% (SE=0.04%, p=0.001) in the non-obese participants as compared to that of 0.24% (SE=0.04%, $p=5.0x10^{-11}$) in the obese participants.

We performed a locus-wise meta-analysis to narrow down the candidate causal gene and variant list. Variant rs8192675 and its proxies showed the strongest association with HbA1c reduction (Figure 3).

The linkage disequilibrium block covers three genes, of which *SLC2A2* encodes the facilitated glucose transporter GLUT2, whilst *EIF5A2* and *RPL22L1* have little known functionality. Previous GWAS studies showed the nonsynonymous rs5400 in *SLC2A2* is the main variant associated with glycaemic traits such as fasting glucose and HbA1c[18,19]. Because rs8192675 and rs5400 are in partial LD (D'=1; r^2=0.35), here rs5400 was also associated with metformin response (beta=0.13%, p=5.2x10^{-4}). However, when conditioning on rs5400, rs8192675 remains strongly associated with metformin response (beta=0.21%, SE=0.04%, p=2.3x10^{-9}); when conditioning on rs8192675, rs5400 is non-significant (p=0.29). These results suggest the pharmacogenetic impact of rs8192675 is unlikely to be via the amino acid change of GLUT2 at rs5400.

Given that liver is the most established site of metformin action, we examined whether rs8192675 is an eQTL in 1,226 liver samples of European ancestry. Figure 3 shows rs8192675 as the top cis-eQTL for *SLC2A2*, with the C-allele associated with decreased (p=4.2x10^{-12}) expression level. In the 48 tissues examined by GTEx, *SLC2A2* was sufficiently expressed in 7 tissues (Supplementary Table 7). rs8192675 showed a significant (p=5.7 x10^{-4}) impact on *SLC2A2* expression in the 271 transformed fibroblasts samples, but no other significant associations[20]. Beyond GTEx, we sought additional eQTL evidence for other tissues that have been implicated in metformin action or glucose homeostasis. Directionally consistent and supportive evidence of rs8192675 or its proxies being *SLC2A2* cis-eQTLs was found in 118 islets (rs8192675, p = 0.0025)[21], 173 intestinal samples (rs5398, p = 0.007)[22], and 44 kidney samples (rs1905505, p = 0.04) (Supplementary Table 7).

Patients with Fanconi-Bickel Syndrome (OMIM#227810), who carry rare loss-of-function variants of GLUT2, can provide useful insight into the role of GLUT2 in glucose homeostasis and into the differing impact of common GLUT2 variants in different physiological states (Figure 4). Patients with Fanconi-Bickel syndrome exhibit low fasting glucose but high post-prandial glucose[23,24]. In parallel, the C-allele of rs8192675 that is associated with reduced *SLC2A2* expression is associated with *lower* fasting glucose and HbA1c among individuals of normal glycaemia[18,19]. Here we report that in patients with type 2 diabetes the expression-decreasing C-allele of rs8192675 was associated with a *higher* HbA1c prior to treatment with either metformin or sulfonylureas. This deleterious genetic effect of rs8192675 on HbA1c was reversed with metformin treatment (C-allele associated with lower on-treatment HbA1c and therefore better response to metformin), but not by sulfonylurea treatment.

In humans, GLUT2 is a facilitative glucose transporter highly expressed in the liver, kidney, small intestine and islets, and to a lesser extent in certain brain regions and other tissues. Genetic defects in GLUT2 could potentially alter glucose homeostasis at any or all of these sites[25]. Metformin's main

site of action is widely believed to be the liver, primarily acting to suppress hepatic glucose production[1,26-28]. In mice with *Glut2* inactivation, glucose and glucose-6-phosphate accumulated in the cytoplasm due to reduced glucose efflux, resulting in increased expression levels of nuclear ChREBP, L-pyruvate kinase and lipogenic genes[29]. Our eQTL data in liver samples (Figure 3) and corresponding reporter assays (Supplementary Figure 4) showed that the C-allele at rs8192675 is associated with lower expression levels of *SLC2A2*. This suggests that the variant may lead to similar effects on hepatic gene expression in humans, which will be potentially modulated by metformin's well-described effect on hepatic glucose production and lipogenesis[30,31]. An alternative explanation could be that reduced SLC2A2 expression due to rs8192675 is associated with reduced glucose mediated glucose clearance (glucose effectiveness) due to a decreased ability for glucose to enter the liver. This is seen in mice lacking *Glut2* in the liver, and is an effect that is improved by metformin treatment[32], although the mechanism for this is not understood.

Metformin is also increasingly believed to exert some of its beneficial effects by acting on the intestines to increase gut glucose uptake and non-oxidative glucose disposal, as well as increasing bile acid reabsorption, GLP-1 secretion and altering the microbiome[33]. In ob/ob mice, metformin has been shown to increase translocation of Glut2 to the apical surface resulting in improved glucose homeostasis[34]. Interestingly, in light of the interaction we report between rs8192675 and BMI on metformin response, obese humans are reported to have altered GLUT2 localisation in the fasting state compared to non-obese humans[34], suggestive of dysregulation of glucose sensing and transport in obese individuals. If reduced *SLC2A2* expression due to rs819265 were to result in reduced apical GLUT2, metformin could potentially overcome this by restoring GLUT2 transport in the enterocytes and improving glucose homeostasis.

Finally, given that metformin is transported into different tissues by several organic cation transporters, including OCTs, MATEs and THTR2[35], we examined whether GLUT2 is able to transport metformin in X. laevis oocytes. Our results suggest that metformin is not a substrate or an inhibitor of GLUT2 (Supplementary Figure 5). Detailed human physiological studies, as well as functional exploration in animal and cellular model systems, are required to fully elucidate the role of GLUT2 in metformin response, and whether this is mediated via a hepatic, intestinal or other mechanism.

We examined the potential clinical impact of rs8192675. An unbiased (from the non-discovery cohorts) estimate of its allelic effect is a 0.15% absolute reduction in %HbA1c. This is equivalent to the pharmacological impact of taking 250mg extra metformin per day, which is 26% of the average daily dose. More clinical potential is seen in obese patients as the C-allele homozygote carriers had a 0.33% (SE=0.09%, $p=6.6\times10^{-4}$) greater reduction in %HbA1c than those carrying the T-allele

211 homozygotes; this equates to 24% of the average glycaemic reduction seen with metformin
212 treatment in the MetGen cohorts and is equivalent to the impact of 550mg extra metformin. Given
213 that newer agents such as DPP-4 inhibitors only reduce HbA1c by 0.6-0.8% on average[36], this genetic
214 effect is large and has potential to be of clinical utility. C-allele homozygotes could be treated with
215 lower doses, and be exposed to less side effects; conversely T-allele carriers could be treated with
216 doses higher than normally recommended to achieve a response. This may be of particular
217 importance in African Americans where 49% of the population are C-allele homozygotes, in contrast
218 to only 9% in European Americans. Stratified clinical trials, in different ethnic groups, are required to
219 evaluate the potential for this pharmacogenetic variant to impact on clinical care.

220 In conclusion, we have established a robust association between rs8192675 and metformin-induced
221 HbA1c reduction with a large multi-ethnic cohort. rs8192675 was the top cis-eQTL for *SLC2A2* in the
222 liver and potentially islets, kidney and intestine. Reduced *SLC2A2* expression resulted in a defect in
223 glucose homeostasis in type 2 diabetes before initiation of therapy, which could be ameliorated by
224 metformin. The clinically appreciable impact in obese patients suggests rs8192675 has the potential
225 to be a biomarker for stratified medicine.

Acknowledgement

We would like to acknowledge Dr. Graeme I. Bell (University of Chicago) for providing the expression vector for SLC2A2 (pSP64T-SLC2A2) and Dr. Daniel L. Minor and Dr. Felix Findeisen (University of California San Francisco) for their guidance to perform oocytes injection and preparing cRNA. For full acknowledgments, see the Supplementary Notes.

Author Contributions

Conception and design of the study: E.R.P. and K.M.G.; Data analysis: K.Z., S.W.Y, E.L.S., N.V.L., A.V.H., J.W.B., C.E.K., L.Z., D.M.R., M.O., K.A.J., L.C., M.J., A.M.L., L.K.W., T.D., A.A.M.; data collection and genotyping: S.W.Y., C.S., R.T.,A.J.B.,C.J.G., R.L.C., L.L., L.K.W., T.D., S.S., M.K., M.M.H., H.C.C., F.I., S.M., J.S.W., L.W., J.Z., I.T., A.K., R.H.S., C.D.S., J.K., V.P., A.H., B.H.S., M.J.W., L.M.H., J.C.F., R.R.H., M.I.M. and C.N.A.P.; Manuscript writing: E.R.P., K.Z., S.W.Y., K.M.G. with contributions from all authors on the final version.

Competing financial interests

The authors have declared that no competing interests exist.

References

1. Madiraju, A.K. *et al.* Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase. *Nature* **510**, 542-6 (2014).
2. DeFronzo, R.A. & Goodman, A.M. Efficacy of metformin in patients with non-insulin-dependent diabetes mellitus. The Multicenter Metformin Study Group. *N Engl J Med* **333**, 541-9 (1995).
3. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). UK Prospective Diabetes Study (UKPDS) Group. *Lancet* **352**, 854-65 (1998).
4. Zhou, K. *et al.* Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. *Lancet Diabetes Endocrinol* **2**, 481-7 (2014).
5. Pawlyk, A.C., Giacomini, K.M., McKeon, C., Shuldiner, A.R. & Florez, J.C. Metformin pharmacogenomics: current status and future directions. *Diabetes* **63**, 2590-9 (2014).
6. Tkac, I. *et al.* Pharmacogenomic association between a variant in SLC47A1 gene and therapeutic response to metformin in type 2 diabetes. *Diabetes Obes Metab* **15**, 189-91 (2013).
7. Stocker, S.L. *et al.* The effect of novel promoter variants in MATE1 and MATE2 on the pharmacokinetics and pharmacodynamics of metformin. *Clin Pharmacol Ther* **93**, 186-94 (2013).
8. Jablonski, K.A. *et al.* Common variants in 40 genes assessed for diabetes incidence and response to metformin and lifestyle intervention in the diabetes prevention program. *Diabetes* **59**, 2672-81 (2010).
9. Shu, Y. *et al.* Effect of genetic variation in the organic cation transporter 1, OCT1, on metformin pharmacokinetics. *Clin Pharmacol Ther* **83**, 273-80 (2008).
10. Zhou, K. *et al.* Reduced-function SLC22A1 polymorphisms encoding organic cation transporter 1 and glycemic response to metformin: a GoDARTS study. *Diabetes* **58**, 1434-9 (2009).
11. van Leeuwen, N. *et al.* A gene variant near ATM is significantly associated with metformin treatment response in type 2 diabetes: a replication and meta-analysis of five cohorts. *Diabetologia* **55**, 1971-7 (2012).
12. Zhou, K. *et al.* Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nat Genet* **43**, 117-20 (2011).
13. Zhou, K. *et al.* Loss-of-function CYP2C9 variants improve therapeutic response to sulfonylureas in type 2 diabetes: a Go-DARTS study. *Clin Pharmacol Ther* **87**, 52-6 (2010).
14. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. *Lancet* **352**, 837-53 (1998).
15. Diabetes Prevention Program Research, G. Long-term safety, tolerability, and weight loss associated with metformin in the Diabetes Prevention Program Outcomes Study. *Diabetes Care* **35**, 731-7 (2012).
16. Diabetes Prevention Program Research, G. *et al.* 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. *Lancet* **374**, 1677-86 (2009).
17. Kahn, S.E. *et al.* Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *N Engl J Med* **355**, 2427-43 (2006).
18. Scott, R.A. *et al.* Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet* **44**, 991-1005 (2012).
19. Soranzo, N. *et al.* Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes* **59**, 3229-39 (2010).

20. Consortium, G.T. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-60 (2015).
21. van de Bunt, M. *et al.* Transcript Expression Data from Human Islets Links Regulatory Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic Traits to Their Downstream Effectors. *PLoS Genet* **11**, e1005694 (2015).
22. Kabakchiev, B. & Silverberg, M.S. Expression quantitative trait loci analysis identifies associations between genotype and gene expression in human intestine. *Gastroenterology* **144**, 1488-96, 1496 e1-3 (2013).
23. Manz, F. *et al.* Fanconi-Bickel syndrome. *Pediatr Nephrol* **1**, 509-18 (1987).
24. Fanconi, G. & Bickel, H. [Chronic aminoaciduria (amino acid diabetes or nephrotic-glucosuric dwarfism) in glycogen storage and cystine disease]. *Helv Paediatr Acta* **4**, 359-96 (1949).
25. Zhou, K. *et al.* Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. *Lancet Diabetes Endocrinol* (In press).
26. Miller, R.A. *et al.* Biguanides suppress hepatic glucagon signalling by decreasing production of cyclic AMP. *Nature* **494**, 256-60 (2013).
27. Hosokawa, M. & Thorens, B. Glucose release from GLUT2-null hepatocytes: characterization of a major and a minor pathway. *Am J Physiol Endocrinol Metab* **282**, E794-801 (2002).
28. Burcelin, R., Dolci, W. & Thorens, B. Glucose sensing by the hepatoportal sensor is GLUT2-dependent: in vivo analysis in GLUT2-null mice. *Diabetes* **49**, 1643-8 (2000).
29. Seyer, P. *et al.* Hepatic glucose sensing is required to preserve beta cell glucose competence. *J Clin Invest* **123**, 1662-76 (2013).
30. Hundal, R.S. *et al.* Mechanism by which metformin reduces glucose production in type 2 diabetes. *Diabetes* **49**, 2063-9 (2000).
31. Fullerton, M.D. *et al.* Single phosphorylation sites in Acc1 and Acc2 regulate lipid homeostasis and the insulin-sensitizing effects of metformin. *Nat Med* **19**, 1649-54 (2013).
32. Pau, C.T., Keefe, C., Duran, J. & Welt, C.K. Metformin improves glucose effectiveness, not insulin sensitivity: predicting treatment response in women with polycystic ovary syndrome in an open-label, interventional study. *J Clin Endocrinol Metab* **99**, 1870-8 (2014).
33. McCreight, L.J., Bailey, C.J. & Pearson, E.R. Metformin and the gastrointestinal tract. *Diabetologia* (2016).
34. Ait-Omar, A. *et al.* GLUT2 accumulation in enterocyte apical and intracellular membranes: a study in morbidly obese human subjects and ob/ob and high fat-fed mice. *Diabetes* **60**, 2598-607 (2011).
35. Gong, L., Goswami, S., Giacomini, K.M., Altman, R.B. & Klein, T.E. Metformin pathways: pharmacokinetics and pharmacodynamics. *Pharmacogenet Genomics* **22**, 820-7 (2012).
36. Bailey, C.J. The Current Drug Treatment Landscape for Diabetes and Perspectives for the Future. *Clin Pharmacol Ther* **98**, 170-84 (2015).

Figures legends

Figure 1. Pharmacogenetic impact of rs8192675 on metformin response in participants of European ancestry. The forest plot shows meta-analyses of association test results for metformin induced change in HbA1c in a total number of 10,557 participants from 10 MetGen cohorts. The two panels present the results from linear regression models with (left) and without (right) adjustment for baseline HbA1c respectively. HbA1c was measured in percentage.

Figure 2. HbA1c reduction by BMI group and rs8192675 genotype. Participants were stratified into obese (BMI≥30 kg/m^2) and non-obese groups (BMI<30 kg/m^2). The error bars are for the standard error of the mean HbA1c reduction.

Figure 3. Regional plots of *SLC2A2* locus. SNPs are plotted by position on the chromosome 3 against association with meta-analysis of HbA1c reduction without baseline adjustment ($-\log_{10}P$) in 7,223 participants (left panel) and meta-analysis of SLC2A2 expression ($-\log_{10}P$) in 1,226 liver samples (right panel). In both plots rs8192675 (purple circle) and its proxies are the top signals. The non-synonymous SNP rs5400 (pointed by arrow) is also nominally associated with HbA1c reduction. Estimated recombination rates (cM/Mb) are plotted in blue to reflect the local LD structure. The SNPs surrounding the most significant SNP, rs8192675, are color coded to reflect their LD with this SNP. This LD was taken from pairwise r^2 values from the HapMap CEU data. Genes, the position of exons and the direction of transcription from the UCSC genome browser are noted.

Figure 4. Genetic impact of GLUT2 variants on glucose homeostasis in different physiological and pharmacologic states. In patients with the monogenic Fanconi-Bickel Syndrome (FBS), the loss-of-function variants led to lower fasting glucose but higher post-prandial glucose; the reduced expression C-allele at rs8192675 was associated with lower HbA1c in normal glycaemia state but higher HbA1c in hyperglycaemia state (before pharmacological treatment was indicated in patients with type 2 diabetes); metformin, but not sulfonylurea treatment reverses the genetic impact on HbA1c.

METHODS

Studies and Samples

Both GWAS screening and the first-stage replication analysed participants with type 2 diabetes of European ancestry from the GoDARTS cohort. The current GWAS screening used 1,373 participants, which included data from 345 samples released after our initial GWAS report on 1,028 participants[12]. The first-stage replication included up to 1,473 from the remaining GoDARTS participants depending on the call rate and genotyping assay. The second-stage replication consisted of 1,223 participants of European ancestry from the UKPDS study. The final replication and meta-analysis was conducted within the MetGen Consortium which included an extra 6,488 participants of European ancestry and 2,566 participants of non-European ancestry. Detailed information on the MetGen participants is provided in Supplementary Table 2. Of note, about 50% of the MetGen cohort is from PMT, which represents ethnically diverse U.S. populations. These cohorts were used extensively in our multi-ethnic analysis for replication purposes. Participants from the largest PMT cohort, PMT2, were selected from the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort, a subsample of the Kaiser Permanente Research Program on Genes, Environment, and Health (RPEGH) [37]. Three MetGen cohorts, GoDARTS, UKPDS and DCS also provided data on response to sulfonylureas. All human research was approved by the relevant institutional review boards, and all participants provided written informed consent.

Genotyping and quality control

Genotyping for the GWAS screening and the first-stage CardioMetabochip replication in GoDARTS cohort has been described before by WTCCC2 and DIAGRAM[12,38]. Standard quality control procedures were applied to both data sets to filter SNPs with minor allele frequency (MAF)<1% or call rate <98% or Hardy-Weinberg Equilibrium (HWE) deviation ($p<10^{-4}$). Samples with call rate <98% or extra heterozygosity (more than 3 standard deviation away from the mean) or correlated with another sample (identity by descent [IBD]>0.125) were filtered out. In-house genotyping of the GoDARTS samples in the first-stage replication were performed with Sequenom MassArray for 66 SNPs and TaqMan based Allelic Discrimination assays for 9 SNPs. Details of the SNP selection procedure is described in Supplementary Data. All 75 SNPs had call rate >90% and no deviation from HWE (p>0.005). The second-stage genotyping of the UKPDS sample was carried out in duplicate runs using standard TaqMan assays. All the SNPs were in HWE (p>0.05) and only samples with concordant genotypes from both runs were analysed. The third-stage replication used high quality genotypes from either TaqMan assay or GWAS imputed data on rs8192675 (Supplementary Table 2).

Assessment of glycaemic response to metformin and sulfonylureas

As with our previous GWAS[12], two correlated measures of glycaemic response to metformin were used in the current GWAS screening and the first-stage replication. A quantitative measure of HbA1c reduction (baseline minus on-treatment HbA1c) and a categorical measure of whether achieving a target of treatment HbA1c≤7% were used for genetic association tests. Therefore only participants with type 2 diabetes and a baseline HbA1c>7% were included. Baseline HbA1c was measured within 6 months prior to metformin start whilst on-treatment HbA1c was taken as the minimum achieved within 18 months after metformin start.

In the second-stage replication and the meta-analysis in the third-stage replication, we opted to maximize the sample size by synchronizing the measurement of metformin efficacy in a wider spectrum of participants with type 2 diabetes (including those with baseline HbA1c<7%) across the MetGen. Therefore only the quantitative outcome of HbA1c reduction was used to assess the glycaemic response to metformin. To maintain relative clinical homogeneity, only participants with type 2 diabetes on metformin monotherapy or using metformin as an add-on therapy to another oral agent were included.

Data from two MetGen cohorts, which used alternative measures of glycaemic response, were not included in the current meta-analyses, but the results are shown in Supplementary Table 4. In the DPP cohort of pre-diabetes participants, Cox proportional hazards regression was used to evaluate the genetic impact on the time to diabetes incidence[8]. In the HOME cohort, a multiple linear regression was used to test the genetic association with the difference in daily dose of insulin because metformin was used in conjunction with insulin in these participants[39].

Assessment of glycaemic response to sulfonylureas adopted a similar approach as the quantitative outcome of metformin response in the MetGen. Baseline HbA1c and on-treatment HbA1c were captured in a similar manner as those in defining metformin response. Only participants with type 2 diabetes who were on sulfonylureas monotherapy or using sulfonylureas as an add-on therapy to metformin were included. All participants had a baseline HbA1c>7%.

Statistical Analysis

In the GWAS screening and first-stage replication, each SNP was tested for association with the continuous measure and categorical measure of glycaemic response to metformin separately with PLINK software using linear and logistic regression respectively[40]. Baseline HbA1c, adherence, metformin dose, creatinine clearance and treatment scheme (whether on metformin monotherapy or dual therapy of metformin add-on to sulfonylureas) and the first 10 principle component from EIGENSTRAT were used as covariates[41]. Statistical evidence of the two associations at each SNP was

averaged by taking the geometric mean of the two p-values in cases in which the direction of effect was consistent (for example more HbA1c reduction and more likely to achieve the treatment target both indicate better response).

In the second and third stage replications, association with HbA1c reduction was tested with multiple linear regression. Within each cohort, two linear models were fitted either with or without adjustment for baseline HbA1c. Baseline HbA1c has been shown as the strongest predictor of metformin induced HbA1c reduction in pharmaco-epidemiological studies[42]. Adjusting for baseline HbA1c could reduce the confounding of measurement error in baseline HbA1c and increase the statistical power for pharmacogenetic studies[43]. However, if a variant is associated with baseline HbA1c, adjusting for baseline HbA1c would lead to a reduced estimate of its pharmacogenetic effect compared to a model that did not adjust for the baseline HbA1c. Therefore we presented both models in the current study. Other clinical factors such as creatinine clearance (or other measurement of kidney function) and treatment scheme were included as covariates where available (Supplementary Table 2). Combining the association results from individual cohort was conducted by a fixed-effect inverse-variance–weighted meta-analysis as applied in GWAMA [44]. Cochran's heterogeneity statistic's p-value was reported as p_{het}.

For the genetic association tests with response to sulfonylureas, multiple linear regression was used to assess the association between rs8192675 and baseline HbA1c, on-treatment HbA1c, HbA1c reduction and baseline adjusted HbA1c reduction. Treatment scheme (whether on sulfonylureas monotherapy or using sulfonylureas as add-on treatment to metformin) was included as a covariate when modelling sulfonylureas induced HbA1c reduction. Association test results from the three cohorts were combined with fixed-effect inverse-variance–weighted meta-analysis in GWAMA.

Locus-wise association was performed with GWAS imputed data of 7,223 participants available in the GoDARTS and PMT2-EU. Software IMPUTE2 was used to impute the post quality control GWAS data at 1Mb flank of rs8192675 against the 1000 Genomes reference panel[45]. Only SNPs with high imputation quality (info>0.9 and MAF>0.02) in both cohorts were tested for association with SNPTEST [46]. Summary statistics from GoDARTS and PMT2-EU were combined with fixed-effect inverse-variance–weighted meta-analysis in GWAMA.

To evaluate the translational potential of rs8192675, we derived an unbiased estimate of its allelic effect by excluding the discovery cohort in the meta-analysis. This effect size was aligned to the clinical impact observed in the PMT2-EU which was the biggest replication cohort and used the median average daily dose in the MetGen. The average daily dose and dosing impact in PMT2-EU were 962mg/day and an extra 0.6% HbA1c reduction per gram metformin respectively. The

evaluation of rs8192675 genotype by BMI group interaction was performed with linear regression by adjusting for treatment group, sex and study cohort.

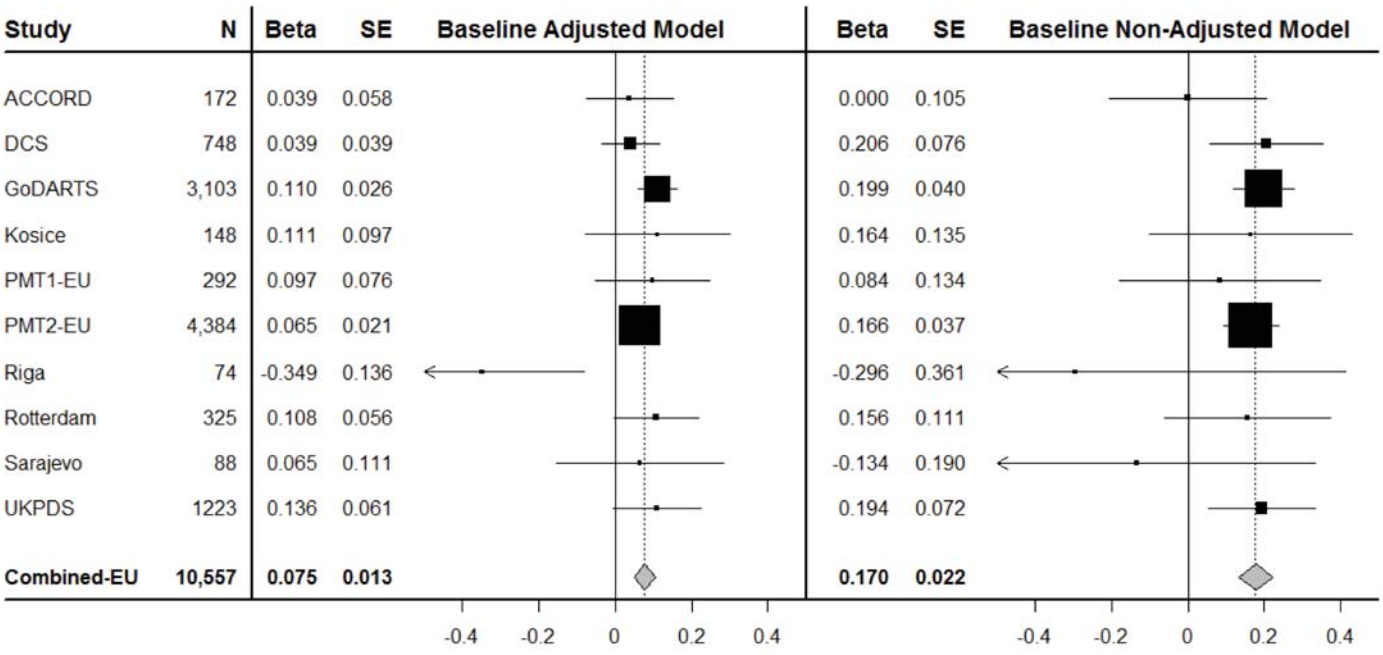
Expression quantitative trait locus (eQTL) analyses.

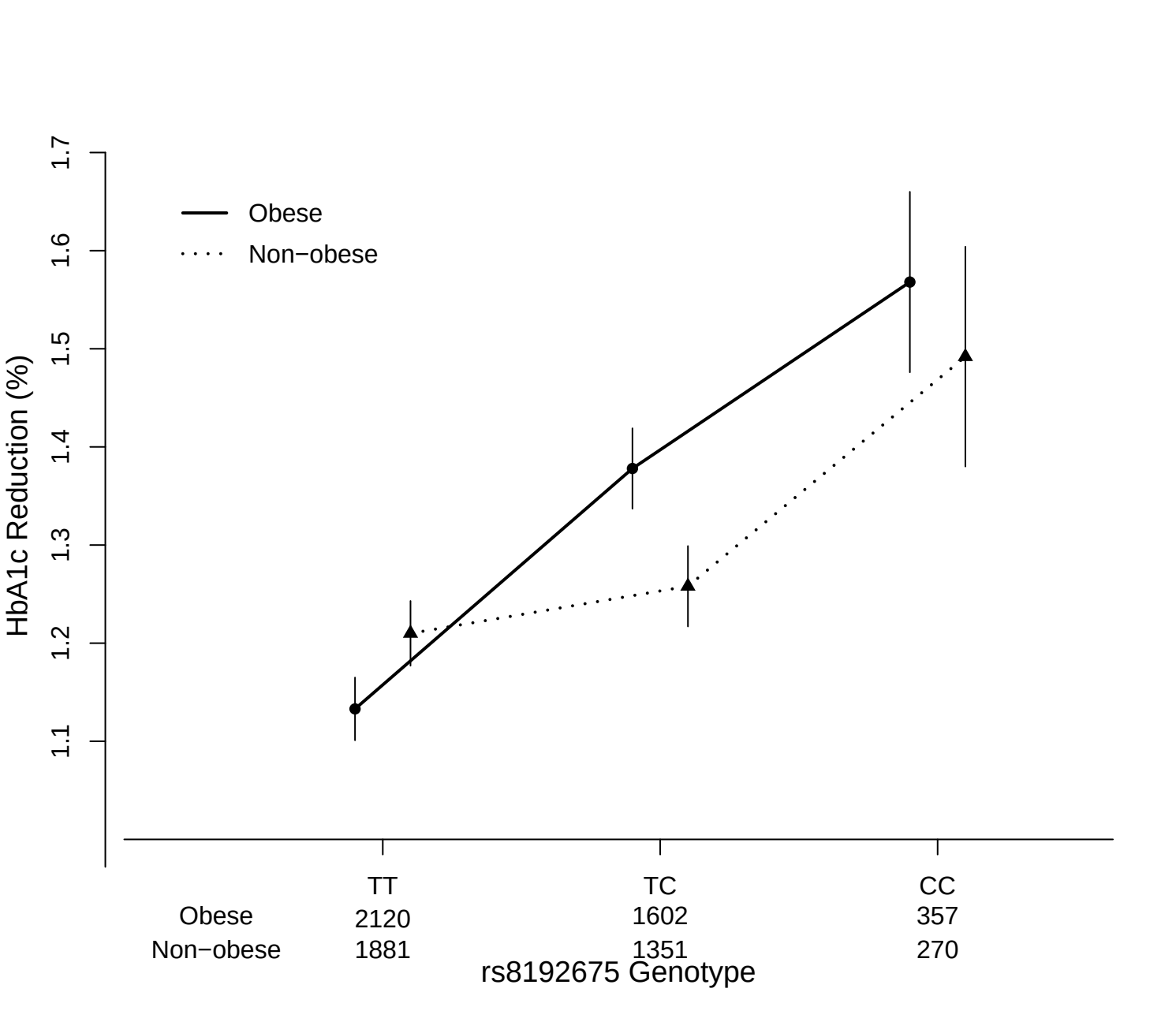
We used four liver eQTL datasets comprising a total number of 1,226 livers samples from individuals of European ancestry (Supplementary Table 8). Tissue procurement, gene expression analysis, genotyping and eQTL analyses have been described previously for three of the datasets[47-49]. The forth dataset was contributed by Dr. Eric Schadt (unpublished data by Schadt, Molony, Chudin, Hao, Yang *et al.*). Genotypes were imputed to the 1000 Genome reference panel with IMPUTE2. Expression probe sequences were mapped to ENSEMBL genes and only the common genes across all datasets were included for subsequent analyses. Within each dataset, the genome-wide eQTL analysis was run with an additive genetic model including dataset specific covariates to examine *cis*-associations within a 100kb flanking window. Results from the four datasets were then combined with a modified meta test statistic which was calculated using the following approach: $t_{meta}=(\sum w_i t_i)/\sqrt{(\sum w_i^2)}$, $w=\sqrt{(n-(\#covariates)-1)}$ where i=data sets 1-4 and n=sample size[50]. This method Generation of p-values was accomplished by assuming the meta test statistics were normally distributed; a Benjamini-Hochberg multiple testing correction was applied to the p-values. For the current study, we extended the cis-association tests to all SNPs within 1Mb window of *SLC2A2* and report the locus-wise p-values of the meta test statistic.

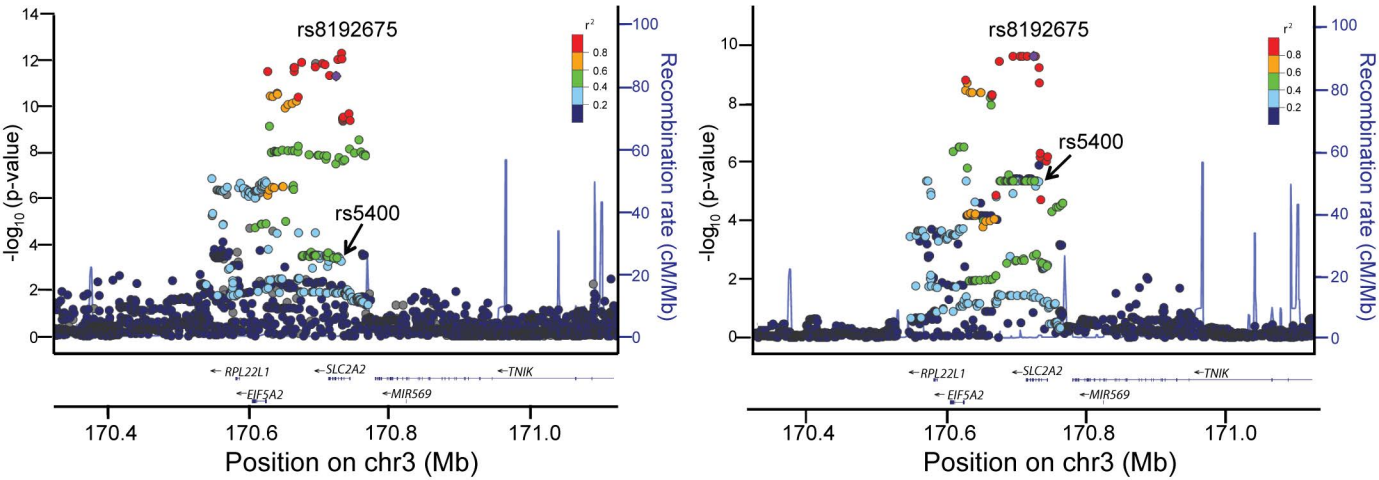
We investigated whether rs8192675 is a cis-QTL in other tissues in the GTEx data release V6. Due to the sample size limitation, rs8192675 is not a genomewide significant cis-eQTL for SLC2A2 in any of tissues examined. However, given the strong evidence of the variant being a cis-eQTL in the large liver samples reported in this study, we considered a directionally consistent association with $p<0.05$ as supportive evidence. The eQTL data for islet and intestine were acquired through contacting the authors of the original publications. The eQTL data for kidney were obtained by quantitative real-time PCR of 44 kidney samples genotyped with the Affymetrix Axiom array. Sample acquirement and tissue preparation was described previously[51]. The transcript levels of *SLC2A2* were determined using TaqMan probe (ID Hs01096908_m1). The relative expression level of *SLC2A2* transcript was calculated by the comparative method (ΔΔCt) normalized to the housekeeping gene GAPDH, as described previously[52].

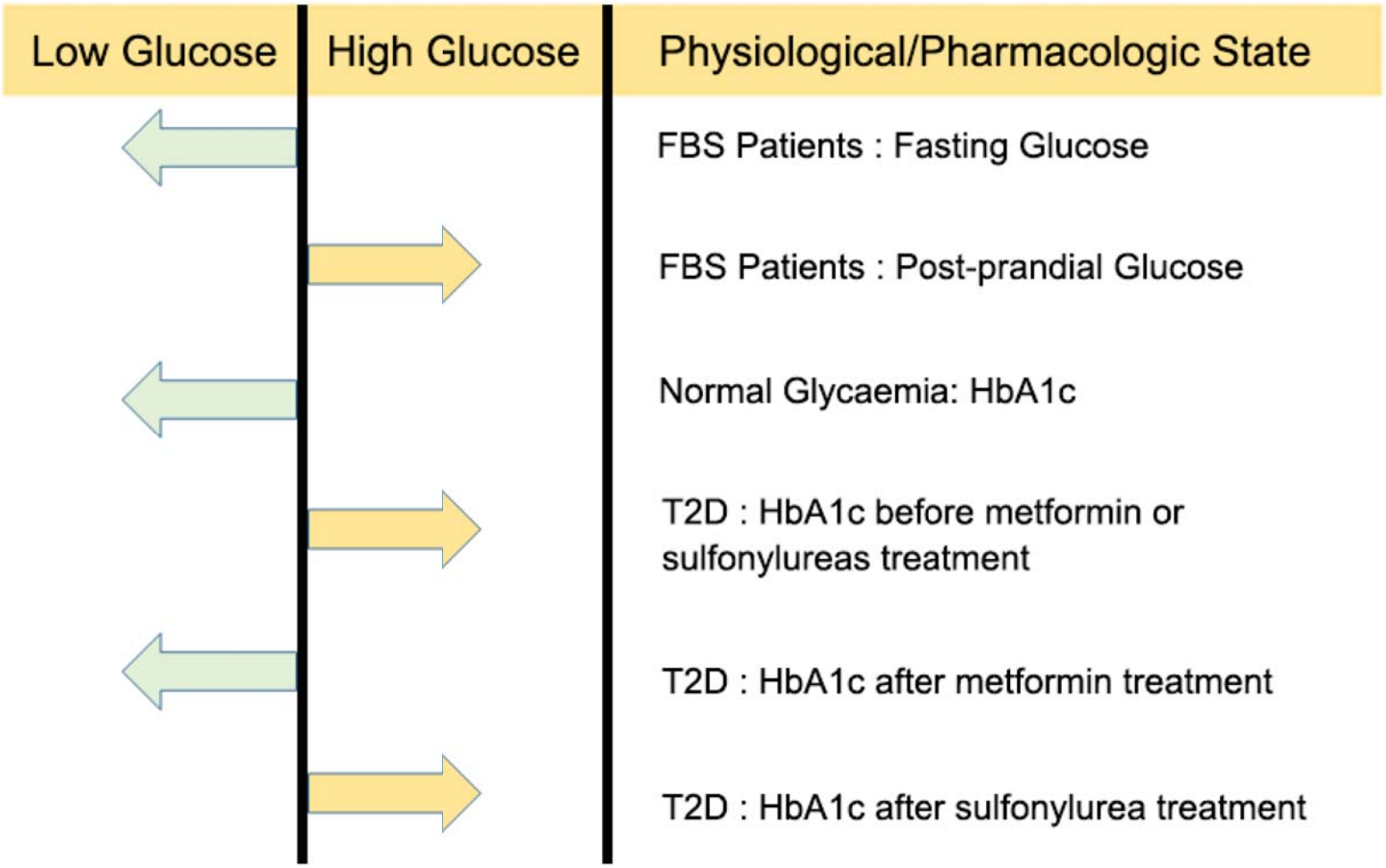
Methods-only references

37. Banda, Y. *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* **200**, 1285-95 (2015).
38. Morris, A.P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* **44**, 981-90 (2012).
39. Kooy, A. *et al.* Long-term effects of metformin on metabolism and microvascular and macrovascular disease in patients with type 2 diabetes mellitus. *Arch Intern Med* **169**, 616-25 (2009).
40. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
41. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
42. Sherifali, D., Nerenberg, K., Pullenayegum, E., Cheng, J.E. & Gerstein, H.C. The effect of oral antidiabetic agents on A1C levels: a systematic review and meta-analysis. *Diabetes Care* **33**, 1859-64 (2010).
43. Postmus, I. *et al.* Pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins. *Nat Commun* **5**, 5068 (2014).
44. Magi, R. & Morris, A.P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* **11**, 288 (2010).
45. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
46. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).
47. Innocenti, F. *et al.* Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. *PLoS Genet* **7**, e1002078 (2011).
48. Schroder, A. *et al.* Genomics of ADME gene expression: mapping expression quantitative trait loci relevant for absorption, distribution, metabolism and excretion of drugs in human liver. *Pharmacogenomics J* **13**, 12-20 (2013).
49. Schadt, E.E. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* **6**, e107 (2008).
50. Xia, K. *et al.* seeQTL: a searchable database for human eQTLs. *Bioinformatics* **28**, 451-2 (2012).
51. Dahlin, A. *et al.* Gene expression profiling of transporters in the solute carrier and ATP-binding cassette superfamilies in human eye substructures. *Mol Pharm* **10**, 650-63 (2013).
52. Liang, X. *et al.* Metformin Is a Substrate and Inhibitor of the Human Thiamine Transporter, THTR-2 (SLC19A3). *Mol Pharm* **12**, 4301-10 (2015).


Study
N
Beta
SE
Baseline Adjusted Model
Beta
SE
Baseline Non-Adjusted Model
ACCORD 172 0.039 0.058 0.000 0.105
DCS 748 0.039 0.039 0.206 0.076
GoDARTS 3,103 0.110 0.026 0.199 0.040
Kosice 148 0.111 0.097 0.164 0.135
PMT1-EU 292 0.097 0.076 0.084 0.134
PMT2-EU 4,384 0.065 0.021 0.166 0.037
Riga 74 -0.349 0.136 -0.296 0.361
Rotterdam 325 0.108 0.056 0.156 0.111
Sarajevo 88 0.065 0.111 -0.134 0.190
UKPDS 1223 0.136 0.061 0.194 0.072
Combined-EU 10,557 0.075 0.013 0.170 0.022
-0.4
-0.2
0
0.2
0.4
-0.4
-0.2
0
0.2
0.4


Obese
Non−obese
HbA1c Reduction (%)
1.1
1.2
1.3
1.4
1.5
1.6
1.7
TT
TC
CC
Obese
2120
1602
357
Non−obese
1881
1351
270
rs8192675 Genotype


rs8192675
rs5400
r²
0.8
0.6
0.4
0.2
-log10 (p-value)
Recombination rate (cM/Mb)
RPL22L1
SLC2A2
TNIK
EIF5A2
MIR569
170.4
170.6
170.8
171.0
Position on chr3 (Mb)


Low Glucose
High Glucose
Physiological/Pharmacologic State
FBS Patients : Fasting Glucose
FBS Patients : Post-prandial Glucose
Normal Glycaemia: HbA1c
T2D : HbA1c before metformin or sulfonylureas treatment
T2D : HbA1c after metformin treatment
T2D : HbA1c after sulfonylurea treatment