Epigenetic markers associated with metformin response and intolerance in drug-naïve patients with type 2 diabetes

Authors: Sonia García-Calzón[1,2*], Alexander Perfilyev[1], Mats Martinell[3], Monta Ustinova[4] Sebastian Kalamajski[5], Paul W Franks[5], Karl Bacos[1], Ilze Elbere[4], Jussi Pihlajamäki[6,7], Petr Volkov[1], Allan Vaag[8], Leif Groop[9], Marlena Maziarz[10], Janis Klovins[4,11], Emma Ahlqvist[9] and Charlotte Ling[1*]

Affiliations:

[1]Epigenetics and Diabetes Unit, Department of Clinical Sciences, Lund University Diabetes Centre, Scania University Hospital, 214 28 Malmö, Sweden.

[2]Department of Nutrition, Food Science and Physiology, University of Navarra, 31008 Pamplona, Spain.

[3]Department of Public Health and Caring Sciences, Uppsala University, 751 22 Uppsala, Sweden.

[4]Latvian Biomedical Research and Study Centre, Rātsupītes Street 1, k-1, Riga, LV-1067, Latvia.

[5]Genetic and Molecular Epidemiology Unit, Department of Clinical Sciences, Lund University Diabetes Centre, 214 28 Malmö, Sweden.

[6]Institute of Public Health and Clinical Nutrition, Internal Medicine, University of Eastern Finland, 70211 Kuopio, Finland.

[7]Clinical Nutrition and Obesity Center, Kuopio University Hospital, 70210 Kuopio, Finland.

[8]Type 2 Diabetes Biology Research, Steno Diabetes Center, 2820 Gentofte, Denmark.

[9]Genomics, Diabetes and Endocrinology Unit, Department of Clinical Sciences, Lund University Diabetes Centre, Scania University Hospital, 214 28 Malmö, Sweden.

[10]Bioinformatics Unit, Department of Clinical Sciences, Lund University Diabetes Centre, 214 28 Malmö, Sweden.

[11]Faculty of Biology, University of Latvia, Riga, LV-1004, Latvia.

*Corresponding author. Email: sonia.garcia_calzon@med.lu.se, charlotte.ling@med.lu.se.

Overline: METABOLISM

One Sentence Summary: Blood-based epigenetic markers differentiate metformin responsiveness and tolerance in patients newly diagnosed with type 2 diabetes.

Abstract: Metformin is the first-line pharmacotherapy for managing type 2 diabetes (T2D). However, many patients with T2D do not respond to or tolerate metformin well. Currently, there are no phenotypes that successfully predict the glycemic response to, or tolerance of, metformin. We explored whether blood-based epigenetic markers could discriminate patients who respond well or poorly to metformin, and who do or do not tolerate metformin, by analyzing genome-wide DNA methylation in drug-naïve patients with T2D at the time of their diagnosis. DNA methylation of 11 and 4 sites differed between glycemic responders/non-responders and tolerant/intolerant patients, respectively, to metformin in discovery and replication cohorts. Greater methylation at these sites was associated with a higher risk of not responding to or not tolerating metformin with odds ratios between 1.43-3.09 per 1 standard deviation methylation increase. Methylation risk scores (MRS) of the 11 identified sites differed between glycemic responders and non-responders and with areas under curve (AUCs) of 0.80-0.98. MRS of the 4 sites associated with future metformin intolerance generated an AUC of 0.85-0.93. Some of these blood-based methylation markers mirrored the epigenetic pattern in adipose tissue, a key tissue in diabetes pathogenesis; and genes to which these markers are annotated to had biological functions in hepatocytes, altering metformin-related phenotypes. Overall, at diagnosis we could discriminate between glycemic responders/non-responders and subjects intolerant/tolerant to metformin by measuring blood-based epigenetic markers in drug-naïve patients with T2D. This epigenetic tool may be further developed to help patients with T2D receive an optimal therapy and may be used for personalized medicine.

Introduction

Metformin is commonly prescribed as a first-line pharmacotherapy for type 2 diabetes (T2D) (*1*). However, ~30% of patients with T2D do not respond to metformin (*2*) and ~20-30% experience intolerable side effects, including gastrointestinal symptoms that warrant discontinuation of metformin treatment in ~5% of patients (*3*). To our knowledge, there are no ways to successfully predict the glycemic response or intolerance to metformin (*4*). Genetics explain only a modest proportion of metformin response and intolerance (*4-12*). Therefore, additional studies are needed to identify markers that determine whether patients with T2D will respond to or tolerate, metformin or whether other therapies should be prioritized. We and others have demonstrated that epigenetics, specifically DNA methylation, contribute to T2D (*13-18*). We also identified blood-based epigenetic markers that mirror the methylation pattern in human islets and predict insulin secretion and T2D (*16*). Epigenetic markers could provide valuable tools for precision medicine however whether blood-based epigenetic markers associate with future drug response and intolerance in patients with T2D remains to be tested.

We aimed to investigate whether DNA methylation in blood associates with future glycemic response and intolerance to metformin therapy in multiple cohorts of drug-naïve patients with T2D from ongoing prospective studies. We further explored cross-tissue methylation patterns of sites associated with future glycemic response or intolerance to metformin in human adipose tissue (*14*). In addition, we studied whether genes to which the identified DNA methylation markers are annotated to affect phenotypes related to metformin therapy in hepatocytes.

Results

Epigenetic markers associate with future glycemic response to metformin

As part of the prospective ANDIS (All New Diabetics In Scania) study (*19*) we carried out a pharmacoepigenetic study for diabetes to identify blood-based epigenetic markers that associate with changes in glycated hemoglobin (ΔHbA1c) or future metformin response in drug-naïve patients with T2D (Fig. 1). Using an 850K array, we analyzed DNA methylation in blood of the discovery and replication cohorts for metformin response (tables S1-S2 and fig. S1-S3). We assessed if methylation status before taking metformin was associated with ΔHbA1c in the full discovery cohort after ~1.5 years of therapy, and if epigenetic markers could discriminate between non-responders and responders to metformin in a subset of patients fulfilling the American Diabetes Association (ADA) criteria for glycemic response (*20*).

First, we explored whether DNA methylation associated with change in HbA1c after ~1.5 years of metformin treatment in newly-diagnosed patients with T2D in the ANDIS discovery cohort for metformin response (Fig. 2A, table S1). Methylation of 2,583 sites was significantly associated with ΔHbA1c after ~1.5 years of metformin (False Discovery Rate [FDR]< 5%, q<0.05). Moreover, 2,577 sites remained significant (FDR<5%) after adjusting for cell composition (*21*) (table S3). Methylation of each site seemed to explain a proportion of variation in ΔHbA1c as the adjusted R-squared ranged between 0.13-0.61. Additionally, methylation of all these sites except one was associated with ΔHbA1c when adjusting regression models for fewer covariates (table S4), suggesting that these covariates did not substantially influence the association. Methylation of 499 and 48 sites was also associated with baseline HbA1c and creatinine clearance (eGFR), respectively (table S3). We proceeded with replication testing of sites associated with ΔHbA1c in a cohort of 204 newly-diagnosed subjects with T2D, the ANDIS replication cohort for metformin response (table S1). We found that methylation of 132 CpGs was also associated with ΔHbA1c (p<0.05) in the replication cohort (n=204), with beta-coefficients in the same direction as in the discovery cohort (table S5).

We next selected two well-defined groups of 26 glycemic responders (HbA1c after ~1.5 years <48-53 mmol/mol and reduction in HbA1c ≥11 mmol/mol) and 21 non-responders (HbA1c after ~1.5 years ≥48-53 mmol/mol and reduction in HbA1c <11 mmol/mol) to metformin treatment from the ANDIS discovery cohort (table S2 and fig. S1) and tested whether baseline methylation discriminated these patient groups. In this case-control set, 7,973 sites showed significant (FDR<5%) differences in methylation between glycemic responders and non-responders, and 7,916 sites remained significant (FDR<5%) when adjusting for cell composition (*21*) (Fig. 2B, table S6). Additionally, methylation of 7,542 sites was associated with glycemic response when adjusting for less covariates in regression models (table S7), suggesting that these covariates did not substantially influence the association. We then performed replication testing of sites associated with glycemic response (table S6) using two independent cohorts of 48 responders and 39 non-responders selected from the ANDIS replication cohort as well as 47 responders and 31 non-responders from the European replication cohort (table S2 and fig. S2-S3). Among the significant sites (FDR<5%) we identified in the discovery cohort, methylation of 601 and 329 sites was associated with glycemic response also in the ANDIS and European replication cohorts, respectively, with directional consistency (tables S8-S9). Furthermore, methylation of 33 sites was associated with glycemic metformin response in the discovery cohort (FDR<0.05) and in the two replication cohorts (p<0.05, table S10). In a combined meta-analysis of the discovery and replication data, 11 out of these 33 methylation markers reached epigenome-wide significance after Bonferroni correction ($p<6.1 \times 10^{-8}$, 0.05/816000) for the association with glycemic metformin response (Table 1). Higher methylation values of all 11 sites were associated with a higher risk of not responding to metformin with odds ratios (OR) ranging between 1.43 and 2.46 per 1 standard deviation (SD) increase in methylation (Fig. 3A).

We proceeded to generate combined weighted MRS (*22*) based on these 11 sites and examined if these scores could discriminate between glycemic responders and non-responders to metformin. Using the CpG-specific effect sizes (beta-coefficients from logistic models) from the ANDIS discovery cohort, we found that MRS adequately discriminated metformin responders from non-responders in the two replication

cohorts (Fig. 4). Receiver operating characteristic (ROC) curves showed that MRSs discriminated between metformin responders and non-responders with an area under the curve (AUC) of 0.80 for the ANDIS replication cohort and 0.89 for the European replication cohort (Fig. 4). We next used the CpG-specific effect sizes from ANDIS or the European replication cohorts to calculate and evaluate MRS in the other two cohorts. These MRS also allowed adequate discrimination of metformin responders and non-responders with AUCs ranging between 0.80-0.98 (fig. S4-S5). In addition, these MRS explained 68-73% of the variation in glycemic response to metformin in the ANDIS discovery cohort, 19-20% in the ANDIS replication cohort, and 38-42% in the European replication cohort (based on R-squared McFadden). These data support the notion that blood-based epigenetic markers may be useful for stratification of metformin response in drug-naïve patients with T2D.

In line with previous findings (*4*), age, body mass index (BMI), baseline HbA1c, and eGFR were not associated with future glycemic response to metformin in our cohorts (Fig. S6A). Moreover, other factors that might affect glycemic control such as ongoing treatment with lipid-lowering or antihypertensive medication, blood pressure, as well as albumin or creatinine in urine were not associated with future glycemic response either (fig. S7).

Epigenetic markers associate with future metformin intolerance

We next evaluated if methylation in blood taken before treatment could discriminate patients with T2D who experienced intolerable side effects (metformin-intolerant) from those who were able to tolerate metformin (metformin-tolerant). We analyzed the methylation of ~850,000 sites in blood samples from the discovery and replication cohorts for metformin intolerance comprising drug-naïve patients with T2D (Fig. 1, table S11).

DNA methylation of 12,579 sites was associated with intolerable side effects in the ANDIS discovery cohort (FDR<5%). (Fig. 2C). 9,676 sites remained significant after adjusting for cell composition (FDR<5%) (*21*) (table S12) and 9,673 sites were significant when adjusting for less covariates in

regression models ($p<0.05$) (table S13), suggesting that these confounders did not substantially influence the association. Most sites showed higher methylation (7,865 CpGs) in metformin-intolerant versus -tolerant patients We next performed replication testing of the sites associated with metformin intolerance (table S12) in two independent cohorts, the ANDIS and European replication cohorts. We found that methylation of 235 and 352 CpGs was associated with metformin intolerance in the ANDIS and European replication cohorts, respectively, with directional consistency (tables S14-S15). Overall, 7 methylation markers were associated with metformin intolerance in the discovery cohort ($FDR<0.05$) and in both replication cohorts ($p<0.05$) (table S16). In a combined meta-analysis of the discovery and replication data, 4 out of these 7 methylation markers reached epigenome-wide significance after Bonferroni correction ($p<6.1 \times 10^{-8}$, 0.05/816000) for association with metformin intolerance (Table 1). Higher methylation values of each of these 4 sites were associated with a higher risk of metformin intolerance with ORs ranging between 1.65 and 3.09 per 1 SD increase in methylation (Fig. 3B).

We then generated combined MRS (*22*) based on the data from these 4 sites to assess if these scores could discriminate metformin tolerant from intolerant drug-naïve subjects with T2D. Using the CpG-specific effect sizes (beta-coefficients from logistic models) from the ANDIS discovery cohort, we calculated and evaluated MRS in the two replication cohorts, and found a separation between metformin tolerant and intolerant subjects (Fig. 5), with an AUC of 0.94 for the ANDIS replication cohort and 0.87 for the European replication cohort (Fig. 5). We next used CpG-specific effect sizes from the ANDIS replication cohort or the European replication cohort to calculate and evaluate MRS in the other two cohorts. These MRS did also give a good separation between metformin intolerant and tolerant subjects, with AUCs ranging between 0.85-0.93 (fig. S8-S9). In addition, these MRS explained 50-51% of the variation in metformin intolerance in the ANDIS discovery cohort, 51-54% in the ANDIS replication cohort, and 32-33% in the European replication cohorts (based on R-squared McFadden). In line with previous findings (*4*), age, BMI, baseline HbA1c and, eGFR were not associated with future intolerance to metformin in all subjects from our discovery and replication cohorts (fig. S6B).

Associations between genetic variants and epigenetics for discriminating metformin response and intolerance

Some studies have previously performed associations between single nucleotide polymorphisms (SNPs) and glycemic response or tolerance to metformin but the degree of confidence in the reported results varies (*5-9, 12, 23-31*). Nevertheless, we selected 26 SNPs previously associated with metformin response (*8, 9, 12, 23-30*) or intolerance (*5-7, 31*) to test if genetics together with epigenetics could better discriminate between metformin response/non-response or tolerance/intolerance. We extracted these SNPs from genome-wide Illumina array data available in the ANDIS discovery and replication cohorts.

We first assessed whether any of these SNPs were associated with DNA methylation of any of the epigenetic marks we identified as discriminating between metformin response/non-response or tolerance/intolerance (Table 1). After correcting for multiple testing, we found only one significant association between a SNP in *SCL22A1* (rs628031) and DNA methylation of cg05151280 (p_{ANOVA}= 0.001, q=0.028). Here, A/A genotype carriers had lower methylation (83.6±2.3%) compared to carriers of the G/G (85.3±1.9%, p=0.002) and G/A (85±1.8%, p=0.006) genotypes in 132 subjects from the ANDIS discovery and replication cohorts. Lower methylation of this CpG site was associated with a better glycemic response to metformin (Table 1). One previous study found a greater reduction in HbA1c in response to metformin in A/A compared to G-allele carriers of rs628031 (*30*), whereas other studies found no association between this polymorphism and metformin response (*23, 27*).

We also evaluated the extent to which these SNPs alone and in addition to our MRS discriminated between metformin responders/non-responders and tolerant/intolerant participants in ANDIS (tables S17-S18). The ability of each SNP to discriminate metformin response and intolerance was generally low, with AUCs ranging from 0.50 to 0.63. Moreover, there was no significant improvement of the AUC regarding metformin response or intolerance after adding each SNP on the top of the MRS in the ANDIS discovery and replication cohorts (p>0.05). These data support that the association between our epigenetic markers and future metformin response or intolerance occurs independently of these 26 SNPs.

We also performed regression analyses to test if any of these SNPs were associated with glycemic response or intolerance to metformin in subjects from ANDIS discovery and replication cohorts. Here, T-allele carriers of rs8192675 (*SLC2A2*) had a nominally higher risk of not responding to metformin compared with homozygous CC-carriers (CT+TT vs CC, OR=4.9(2.2), p=0.04), which is in line with previous data (*12*). We also found a nominal association between rs12208357 *(OCT1*) and metformin intolerance where T-allele carriers had a lower risk of intolerance (CT+TT vs CC, OR=0.13(2.85), p=0.05).

Cross-tissue methylation in blood and human adipose tissue

Next, we investigated whether methylation in blood of the 11 sites (8 sites available in the 450K array) associated with metformin response and the 4 sites (2 sites available in the 450K array) associated with intolerance reflected methylation in human adipose tissue. Here, we used available methylation data on the 450K array in blood and adipose tissue as for these cells, we had access to methylation data from the same subjects (*14, 32*) (Tables S19-S20). We found that DNA methylation of three sites in the blood positively correlated with methylation in adipose tissue after correcting for multiple testing (Fig. S10). These findings suggest that methylation in blood associated with metformin response and intolerance may also have a biological role in key tissues for T2D.

Functional follow-up experiments in hepatocytes cultured in vitro

We asked if genes to which the identified CpG sites associated with metformin response and intolerance (Table 1) are annotated to have a functional role in liver cells (HepG2 cells) (*33, 34*). For functional follow-up experiments, we focused on genes which have previously been related to phenotypes involved in diabetes. Based on these criteria we selected five genes (*OR4S1, SEPT11, CST1, FOXA2* and *PGM1*) (*35-41*) for functional experiments and we elucidated their effects on expression of two metformin transporters (*SLC22A1,* encoding OCT1, the main transporter for metformin uptake into hepatocytes, and *SLC47A1,* encoding MATE1, the main efflux transporter of metformin to the bile), AMP-activated protein kinase (AMPK) activity and expression of key regulators of gluconeogenesis (*PCK1* and *G6PC*) in liver

cells untreated and treated with metformin. We silenced the expression of these five genes in HepG2 cells using siRNA, which resulted in an 82 to 98% reduction in expression of all the genes (Fig. 6A) except for *OR4S1* whose mRNA expression was undetectable, probably due to low expression in this liver cell line. *OR4S1* was therefore excluded from further experiments.

As expected, metformin treatment activated AMPK and decreased *PCK1* and *G6PC* gene expression in cultured HepG2 cells, confirming the pharmacological effect of metformin in the inhibition of gluconeogenesis (*42*) (Fig. 6B). Moreover, metformin did not alter the expression of metformin transporters (*43*) (Fig. 6B).

To investigate glycemic response to metformin, we silenced two genes (*SEPT11* and *CST1*) located near CpG sites associated with metformin response in newly-diagnosed patients with T2D (cg01070242 and cg07511259, respectively). We found that *SEPT11*-deficient HepG2 cells had lower *SLC47A1* expression (Fig. 6C), which could result in lower efflux and higher metformin concentration in the hepatocytes associated with a greater pharmacologic response (*44*). Additionally, *SEPT11* deficiency resulted in lower *G6PC* expression (Fig. 6C), a mechanism previously associated with decreased gluconeogenesis and lower hepatic glucose output (*42*). *CST1* deficient HepG2 cells had increased *SLC47A1* expression (Fig. 6D), which could result in higher efflux and lower metformin concentration in the hepatocytes associated with a lower pharmacologic response (*44*). Moreover, *CST1* deficient cells had nominally decreased AMPK activity and increased expression of *PCK1* and *G6PC* (Fig. 6D), associated with increased gluconeogenesis and elevated hepatic glucose output (*42*).

Regarding intolerance to metformin, we silenced two genes, *FOXA2* and *PGM1*, near cg12356107 and cg02994863, respectively, associated with metformin intolerance in newly-diagnosed patients with T2D (Table 1). Both *FOXA2* and *PGM1* deficient HepG2 cells had higher *SLC47A1* expression (Fig. 6E-F), which could result in higher excretion and therefore lower metformin concentration in the hepatocytes (*44, 45*). Moreover, *FOXA2* and *PGM1* deficiency resulted in higher *PCK1* and *G6PC* expression (Fig. 6E-F), associated with increased gluconeogenesis, lower lactate production and therefore a better tolerance to metformin (*34*). Regarding AMPK activity, *FOXA2* deficient cells had nominally reduced AMPK

phosphorylation (Fig. 6E), which is in line with the increase in gluconeogenesis and therefore a better tolerance to metformin (*34*). However, *PGM1* deficiency did not change AMPK activity (Fig. 6F), suggesting that AMPK-independent mechanisms may be involved to activate gluconeogenesis in these cells (*42, 46*), and hence reduced lactate concentration. Overall, these experiments support that several genes annotated to CpG sites associated with response or intolerance to metformin have a functional role in liver cells where they affect metformin transporters and key regulators of gluconeogenesis.

Increased methylation in promoter and CpG island regions has been associated with decreased expression (*17*). We used a luciferase assay to study the impact of increased DNA methylation in transcriptional regulation of *SAP130*, a gene annotated to a CpG site (cg16240962) located in promoter and CpG island regions and that is associated with glycemic response to metformin (Table 1). The promoter sequence for *SAP130* was inserted into a luciferase expression plasmid and mock-methylated or methylated with the methyltransferase SssI (methylating 158 methylation sites, including cg16240962). Our data show that increased promoter methylation suppressed the transcriptional activity of *SAP130* in HepG2 cells (Fig. S11), supporting that DNA methylation of sites associated with metformin response may mediate gene regulation.

Discussion

We performed a pharmacoepigenetics study for T2D in which we identified and validated blood-based epigenetic markers associated with future glycemic response and intolerance to metformin therapy in drug-naïve subjects with T2D. Metformin non-responsive and intolerant patients with T2D should be prescribed other glycemic lowering drugs to achieve treatment goals of ADA and the European Association for the Study of Diabetes (EASD) (*1, 20*). However, there are currently no biomarkers available for identifying these patients at diagnosis (*4-12*). Our study found that DNA methylation at eleven and four specific loci was associated with future glycemic response and intolerance to metformin, respectively, in discovery and replication cohorts. Patients with higher degrees of methylation at these sites were up to 2.5 times more likely to not respond to, and up to 3 times more likely to not tolerate,

metformin due to severe side effects. Moreover, methylation at these sites used in weighted MRS was different between responders/non-responders and tolerant/intolerant to metformin. Notably, AUC for these MRS ranged between 0.80-0.98 in the cohorts for metformin response and between 0.83-0.94 in the cohorts for metformin intolerance. Although more studies are needed to validate these markers in other populations, our results support further development of epigenetic markers for stratification of non-responsive and intolerant patients to metformin already at diagnosis. Such stratification may help patients with T2D receive an optimal therapy and could be a step towards personalized medicine. Future studies in additional cohorts may optimize these MRS further and may add or replace some markers. In the current study, the MRS were slightly different depending on which CpG-specific effect sizes were estimated from one cohort and evaluated it in the other two. However, a single MRS would be useful for clinical use and therefore should be further developed and optimized in independent cohorts. Additionally, it would be useful to clinically validate epigenetic markers in a randomized clinical trial. To this end, collection of new larger cohorts should be prioritized.

There are several reasons supporting the potential use of DNA methylation markers when deciding whether to prescribe metformin therapy or not. Analyzing epigenetic markers in blood is non-invasive, safe, quick and, cost-effective. Methylation is quite stable, can persist over time and, is inherited through cell divisions (*47*). Moreover, the combination of identified methylation sites associated with response and intolerance to metformin using MRS show AUC >0.80, which is a requirement for a useful clinical discrimination (*48, 49*). Additionally, giving an optimal therapy to newly-diagnosed patients with T2D by using epigenetic markers could potentially decrease costs related to poor glycemic control, reducing visits to the doctor, sick-leave, exhaustion, and vascular complications. Of note, the average cost of vascular complications in T2D was estimated to $47,240 per patient over 30 years (*50*). This can be compared with an estimated cost of $200 for measuring DNA methylation in blood. Although this is a pharmacoepigenetic study for diabetes, there is already a commercial liver cancer test available that analyses *SEPT9* methylation (*51*) supporting the feasibility of clinical epigenetic markers.

This study has potential limitations. CpG sites were selected based on their significance in the discovery and two replication cohorts, which may result in overestimated AUCs. To mitigate this, we used CpG-specific effect sizes (beta-coefficients from logistic models) in the discovery cohorts, calculating and evaluating the MRS in the replication cohorts. This approach gave similar AUCs for all three cohorts, supporting further development of epigenetic markers for discrimination of response and intolerance to metformin. Associations were present regardless of baseline HbA1c and eGFR, providing additional support for the robustness of the findings. Also, the association was present whether adjusting for cell composition or not (*21*). Of note, whereas case-control cohorts for metformin response were balanced for basal HbA1c, eGFR, age, sex and, BMI, the full discovery cohort for metformin response included subjects with a continuum of these variables. Subsequently, one would not expect these two different analyses to give the same result. Indeed, 888 methylation sites were significant in both the case-control and full discovery cohort for metformin response, whereas some sites were only significant in one of these analyses (FDR<5%). Our study was carried out in Caucasians and validation in other ethnicities is strongly needed. However, we are not aware of any additional cohorts with blood samples available in drug-naïve newly diagnosed patients with T2D at this time point. That we had to change the inclusion criteria slightly to find subjects for replication may have reduced the possibility of replicating significant sites. However, the replicated sites were statistically robust and were found in several cohorts. Last, in line with several previous metformin studies (*8, 10, 12*), we used pharmacy registers to identify patients who were on metformin therapy. One potential limitation with this design is that we cannot examine patient medication adherence. However, for metformin intolerance, we called the patients or checked in their clinical history records for the reason why they stopped metformin therapy so we could confirm the intolerance status of these patients to metformin therapy.

Environmental factors such as exercise and diet as well as obesity and weight change might affect DNA methylation (*22, 52-54*). However, neither baseline BMI nor weight change had an impact on methylation of the sites associated with glycemic response or intolerance to metformin in respective cohorts in our

study. We also examined if methylation of our significant markers changed in other studies where the impact of environmental factors was investigated. Here, methylation of only 1 site (out of the 11 and 4 sites associated with metformin response and intolerance, respectively) changed in adipose tissue after 5 days of high fat diet and none of these sites changed in adipose tissue after exercise (*53, 54*). Together, these data suggest that lifestyle factors have minor effects on methylation of the sites associated with either response or intolerance to metformin.

Some of the identified blood-based epigenetic markers associated with metformin response or intolerance mirror the methylation pattern in adipose tissue, a metabolically relevant tissue for T2D (*14*). Blood-based epigenetic markers might hence reflect what it is happening in the central tissues of diabetes. In addition, silencing nearby genes (*SEPT11, CST1, FOXA2, PGM1*) annotated to these CpG sites in hepatocytes resulted in altered expression of metformin transporters and key enzymes affecting gluconeogenesis, supporting biological functions of these epigenetic markers in the pharmacological effect of metformin. For example, we found that *FOXA2* has a functional role in hepatocytes altering tolerance to metformin. Similarly, it has been shown that *FOXA2* mediates an effect of metformin on bile acid metabolism, which is a likely cause of adverse gastrointestinal effects (*55*). The genes we selected for these functional experiments have previously been shown to affect diabetes related phenotypes (*35-41*); for example, *CST1* has been proposed as a promising biomarker for both diabetic neuropathy and breast cancer (*56, 57*). Our functional data further support that some of these epigenetic markers can regulate gene transcription. Overall, we shed light on some potential biological mechanisms related to our epigenetic markers and their link to metformin response or intolerance.

In conclusion, our study provides potential blood-based epigenetic markers for stratification of newly diagnosed patients with T2D into metformin non-responsive/responsive and intolerant/tolerant. Further research is warranted to develop this panel of epigenetic markers to aid clinical decision making in T2D therapy by assigning newly diagnosed patients to receive either metformin or other glycemic-lowering medication, which may reduce suffering for patients.

Materials and Methods

Study design

This study was designed to identify blood-based epigenetic markers that could discriminate between glycemic responders/non-responders and tolerant/intolerant patients with T2D to metformin. The glycemic response of the participants to metformin treatment was based on the change in HbA1c values after ~1.5 years of therapy according to ADA and EASD guidelines, and metformin intolerance was assessed by the presence of intolerable side effects in clinical history records. Discovery and replication cohorts from ANDIS, ANDiU, and OPTIMED were included. Study size was not prespecified, and results are reported for all patients with T2D who fulfilled the criteria of being a responder or non-responder or tolerant or intolerant to metformin therapy within this population at the time of the study. Individual methylation markers associated with future metformin response or intolerance were selected based on genome-wide significance in a fixed meta-analysis and then combined in MRS to better discriminate patients into responders/non-responders and tolerant/intolerant subjects with T2D. Moreover, genetic-epigenetic interaction analyses, cross-tissue methylation patterns and functional follow-up experiments in hepatocytes in vitro were performed to better understand the role of these epigenetic markers in metformin response or intolerance. More cohort and methods details are available in the Supplementary Materials.

Study populations

All subjects included in this study were newly diagnosed drug-naïve patients with T2D with available blood samples before the start of metformin therapy. They were selected from the following cohorts: The All New Diabetics In Scania (ANDIS) (*19*) cohort, an ongoing prospective cohort, that aimed to register all new cases of diabetes in Scania for improvement of diagnosis and treatment strategies; The All New Diabetics In Uppsala County (ANDiU) cohort (*19*) (http://www.andiu.se/), an ongoing prospective cohort, that included anyone who was diagnosed with diabetes and resides in the County of Uppsala, Sweden; and The OPTIMED cohort which includes patients with T2D from Latvia (*58*). ANDIS, ANDiU, and OPTIMED were performed in accordance with the Declaration of Helsinki and written informed consent

was obtained from all participants. These cohorts were divided into discovery and replication cohorts (figs S1-S3, tables S1, S2, S11) and are further described in the Supplementary Methods.

Statistical analyses

All statistical analyses were performed using R Statistical Software. Two-sided P values were used for all the analyses. To evaluate differences between clinical variables, Mann-Whitney and $\chi 2$ tests were performed as appropriate. To assess the association between genome-wide DNA methylation and metformin response or intolerance linear regression models were fit prior transformation of those variables which were not normally distributed to achieve normality. Linear regression models were fit to identify and replicate epigenetic markers associated with metformin response and intolerance. Methylation markers were selected from the discovery stage to replication testing if FDR below 5% ($q<0.05$). Methylation markers were considered if $p<0.05$ in the replication cohorts with directional consistency. We performed combined analyses of discovery and replication data using fixed meta-analysis and here we required epigenome-wide significance after Bonferroni correction. Logistic regression models based on these methylation markers were used to assess the risk of not responding to or not tolerating metformin.

We also evaluated whether MRS could discriminate between glycemic responders and non-responders as well as between subjects tolerant and intolerant to metformin. MRS were calculated as the sum of standardized methylation values at each site associated with metformin response or intolerance, weighted by CpG-specific effect size (*22*). To validate our findings and control for overfitting, we performed three evaluations, calculating the MRS using CpG-specific effect sizes (beta-coefficients from the logistic models) estimated from one cohort (either the ANDIS discovery or replication cohort, or the European replication cohort) and evaluated in the other two using ROC curves and AUC.

Supplementary Materials

Materials and Methods

Fig. S1. Flowchart and participant selection criteria of the ANDIS discovery cohort for metformin response.

Fig. S2. Flowchart and participant selection criteria of the ANDIS replication cohort for metformin response.

Fig. S3. Flowchart and selection criteria of ANDiU and OPTIMED participants for investigation of glycemic response to metformin therapy: "The European replication cohort for metformin response".

Fig. S4. Combined MRS discriminate between glycemic responders and non-responders to metformin in drug-naïve subjects with T2D from the ANDIS discovery and the European replication cohorts.

Fig. S5. Combined MRS discriminate between responders and non-responders to metformin in drug-naïve subjects with T2D from the ANDIS discovery and the ANDIS replication cohorts.

Fig. S6. ROC curves for response (A) and intolerance (B) to metformin incorporating different clinical baseline phenotypes in all subjects from discovery and replication cohorts for metformin response and intolerance.

Fig. S7. ROC curves for response to metformin incorporating additional clinical baseline phenotypes in ANDIS subjects from the discovery and replication cohorts for metformin response combined.

Fig. S8. Combined MRS discriminate between tolerant and intolerant subjects to metformin in drug-naïve subjects with T2D from the ANDIS discovery and the European replication cohorts.

Fig. S9. Combined MRS discriminate between tolerant and intolerant subjects to metformin in drug-naïve subjects with T2D from the ANDIS discovery and the ANDIS replication cohorts.

Fig. S10. Correlations between DNA methylation in blood and DNA methylation in adipose tissue (n=28) from the same subject (Monozygotic Twin Cohort).

Fig. S11. In vitro methylation of the *SAP130* promoter resulted in decreased transcriptional activity.

Table S1. Clinical characteristics of the full discovery and replication cohorts for metformin response including drug-naïve and newly diagnosed subjects with type 2 diabetes from the ANDIS cohort.

Table S2. Clinical characteristics of case-control discovery and replication cohorts including patients who fulfil the criteria of being glycemic responders and non-responders to metformin therapy.

Table S3. CpG sites that show a significant association (FDR<5%) between DNA methylation in whole blood before taking metformin and the change in HbA1c after ~1.5 years on metformin therapy in drug-naïve discovery cohort subjects with T2D from the discovery cohort (n=63).

Table S4. Comparison of the 2577 significant CpG sites (FDR<5%) with an association between DNA methylation and the ΔHbA1c after ~1.5 years in drug-naïve subjects with T2D from the discovery cohort, with two other linear models.

Table S5. CpG sites with DNA methylation associated with the change in HbA1c in both the discovery cohort and in the ANDIS replication cohort for metformin response (ΔHbA1c).

Table S6. CpG sites exhibiting differences in DNA methylation in whole blood between glycemic responders (n=26) and non-responders (n=21) to metformin therapy in drug-naïve subjects with T2D from the discovery cohort.

Table S7. Comparison of the 7916 significant CpG sites (FDR<5%) between metformin responders and non-responders in drug-naïve subjects with T2D from the discovery cohort, with three other linear models.

Table S8. Methylated CpG sites associated with response to metformin in the discovery cohort and in the ANDIS replication cohort for metformin response.

Table S9. Methylated CpG sites associated with response to metformin in the discovery cohort and in the European replication cohort for metformin response.

Table S10. Methylated CpG sites associated with response to metformin in the discovery cohort and in both the ANDIS and the European replication cohorts for metformin response.

Table S11. Clinical characteristics of drug-naïve and newly diagnosed patients with T2D included in the metformin intolerance discovery and replication cohorts.

Table S12. CpG sites exhibiting differences in DNA methylation in whole blood between metformin intolerant (n=17) and tolerant (n=66) drug-naïve subjects with T2D from the discovery cohort.

Table S13. Comparison of the 9676 significant CpG sites (FDR<5%) between metformin tolerant and intolerant drug-naïve subjects with T2D from the discovery cohort, with two other linear models.

Table S14. CpG sites with DNA methylation associated with intolerance to metformin in the discovery cohort and in the ANDIS replication cohort for metformin intolerance.

Table S15. CpG sites with DNA methylation associated with intolerance to metformin in the discovery cohort and in the European replication cohort for metformin intolerance.

Table S16. CpG sites with DNA methylation associated with intolerance to metformin in the discovery cohort and in both the ANDIS and the European replication cohorts for metformin intolerance.

Table S17. Assessing discrimination between glycemic responders and non-responders to metformin using SNPs and MRS associated with metformin response.

Table S18. Assessing discrimination between tolerant and intolerant subjects to metformin using SNPs and MRS associated with metformin intolerance.

Table S19. Clinical characteristics of all study subjects in the monozygotic twin cohort (MZ).

Table S20. Available data from the monozygotic twin cohort used for the analyses of DNA methylation and gene expression in human tissues in the present study.

Table S21. Sequence of *SAP130* inserted into the CpG-free firefly luciferase reporter vector (pCpGL-basic) and used for luciferase experiments.

References and notes

1. S. E. Inzucchi, R. M. Bergenstal, J. B. Buse, M. Diamant, E. Ferrannini, M. Nauck, A. L. Peters, A. Tsapas, R. Wender, D. R. Matthews, Management of hyperglycaemia in type 2 diabetes, 2015: a patient-centred approach. Update to a position statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetologia* **58**, 429-442 (2015).
2. M. N. Cook, C. J. Girman, P. P. Stein, C. M. Alexander, Initial monotherapy with either metformin or sulphonylureas often fails to achieve or maintain current glycaemic goals in patients with Type 2 diabetes in UK primary care. *Diabet Med* **24**, 350-358 (2007).
3. S. E. Kahn, S. M. Haffner, M. A. Heise, W. H. Herman, R. R. Holman, N. P. Jones, B. G. Kravitz, J. M. Lachin, M. C. O'Neill, B. Zinman, G. Viberti, A. S. Group, Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *N Engl J Med* **355**, 2427-2443 (2006).
4. L. A. Donnelly, A. S. Doney, A. T. Hattersley, A. D. Morris, E. R. Pearson, The effect of obesity on glycaemic response to metformin or sulphonylureas in Type 2 diabetes. *Diabet Med* **23**, 128-133 (2006).
5. T. Dujic, A. Causevic, T. Bego, M. Malenica, Z. Velija-Asimi, E. R. Pearson, S. Semiz, Organic cation transporter 1 variants and gastrointestinal side effects of metformin in patients with Type 2 diabetes. *Diabet Med* **33**, 511-514 (2016).
6. T. Dujic, K. Zhou, L. A. Donnelly, R. Tavendale, C. N. Palmer, E. R. Pearson, Association of Organic Cation Transporter 1 With Intolerance to Metformin in Type 2 Diabetes: A GoDARTS Study. *Diabetes* **64**, 1786-1793 (2015).
7. T. Dujic, K. Zhou, R. Tavendale, C. N. Palmer, E. R. Pearson, Effect of Serotonin Transporter 5-HTTLPR Polymorphism on Gastrointestinal Intolerance to Metformin: A GoDARTS Study. *Diabetes Care* **39**, 1896-1901 (2016).
8. GoDarts, U. D. P. S. Group, C. Wellcome Trust Case Control, K. Zhou, C. Bellenguez, C. C. Spencer, A. J. Bennett, R. L. Coleman, R. Tavendale, S. A. Hawley, L. A. Donnelly, C. Schofield, C. J. Groves, L. Burch, F. Carr, A. Strange, C. Freeman, J. M. Blackwell, E. Bramon, M. A. Brown,

J. P. Casas, A. Corvin, N. Craddock, P. Deloukas, S. Dronov, A. Duncanson, S. Edkins, E. Gray, S. Hunt, J. Jankowski, C. Langford, H. S. Markus, C. G. Mathew, R. Plomin, A. Rautanen, S. J. Sawcer, N. J. Samani, R. Trembath, A. C. Viswanathan, N. W. Wood, M. investigators, L. W. Harries, A. T. Hattersley, A. S. Doney, H. Colhoun, A. D. Morris, C. Sutherland, D. G. Hardie, L. Peltonen, M. I. McCarthy, R. R. Holman, C. N. Palmer, P. Donnelly, E. R. Pearson, Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nat Genet* **43**, 117-120 (2011).

9. K. A. Jablonski, J. B. McAteer, P. I. de Bakker, P. W. Franks, T. I. Pollin, R. L. Hanson, R. Saxena, S. Fowler, A. R. Shuldiner, W. C. Knowler, D. Altshuler, J. C. Florez, G. Diabetes Prevention Program Research, Common variants in 40 genes assessed for diabetes incidence and response to metformin and lifestyle intervention in the diabetes prevention program. *Diabetes* **59**, 2672-2681 (2010).

10. D. M. Rotroff, S. W. Yee, K. Zhou, S. W. Marvel, H. S. Shah, J. R. Jack, T. M. Havener, M. M. Hedderson, M. Kubo, M. A. Herman, H. Gao, J. C. Mychaleckyi, H. L. McLeod, A. Doria, K. M. Giacomini, E. R. Pearson, M. J. Wagner, J. B. Buse, A. A. Motsinger-Reif, I. MetGen, A. C. A. Investigators, Genetic Variants in CPA6 and PRPF31 Are Associated With Variation in Response to Metformin in Individuals With Type 2 Diabetes. *Diabetes* **67**, 1428-1440 (2018).

11. K. Zhou, L. Donnelly, J. Yang, M. Li, H. Deshmukh, N. Van Zuydam, E. Ahlqvist, C. Wellcome Trust Case Control, C. C. Spencer, L. Groop, A. D. Morris, H. M. Colhoun, P. C. Sham, M. I. McCarthy, C. N. Palmer, E. R. Pearson, Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. *Lancet Diabetes Endocrinol* **2**, 481-487 (2014).

12. K. Zhou, S. W. Yee, E. L. Seiser, N. van Leeuwen, R. Tavendale, A. J. Bennett, C. J. Groves, R. L. Coleman, A. A. van der Heijden, J. W. Beulens, C. E. de Keyser, L. Zaharenko, D. M. Rotroff, M. Out, K. A. Jablonski, L. Chen, M. Javorsky, J. Zidzik, A. M. Levin, L. K. Williams, T. Dujic, S. Semiz, M. Kubo, H. C. Chien, S. Maeda, J. S. Witte, L. Wu, I. Tkac, A. Kooy, R. H. van Schaik, C. D. Stehouwer, L. Logie, I. MetGen, D. P. P. Investigators, A. Investigators, C. Sutherland, J.

Klovins, V. Pirags, A. Hofman, B. H. Stricker, A. A. Motsinger-Reif, M. J. Wagner, F. Innocenti, L. M. Hart, R. R. Holman, M. I. McCarthy, M. M. Hedderson, C. N. Palmer, J. C. Florez, K. M. Giacomini, E. R. Pearson, Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet*, (2016).

13. T. Dayeh, P. Volkov, S. Salo, E. Hall, E. Nilsson, A. H. Olsson, C. L. Kirkpatrick, C. B. Wollheim, L. Eliasson, T. Ronn, K. Bacos, C. Ling, Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. *PLoS Genet* **10**, e1004160 (2014).

14. E. Nilsson, P. A. Jansson, A. Perfilyev, P. Volkov, M. Pedersen, M. K. Svensson, P. Poulsen, R. Ribel-Madsen, N. L. Pedersen, P. Almgren, J. Fadista, T. Ronn, B. Klarlund Pedersen, C. Scheele, A. Vaag, C. Ling, Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. *Diabetes* **63**, 2962-2976 (2014).

15. E. Nilsson, A. Matte, A. Perfilyev, V. D. de Mello, P. Kakela, J. Pihlajamaki, C. Ling, Epigenetic Alterations in Human Liver From Subjects With Type 2 Diabetes in Parallel With Reduced Folate Levels. *J Clin Endocrinol Metab* **100**, E1491-1501 (2015).

16. K. Bacos, L. Gillberg, P. Volkov, A. H. Olsson, T. Hansen, O. Pedersen, A. P. Gjesing, H. Eiberg, T. Tuomi, P. Almgren, L. Groop, L. Eliasson, A. Vaag, T. Dayeh, C. Ling, Blood-based biomarkers of age-associated epigenetic changes in human islets associate with insulin secretion and diabetes. *Nat Commun* **7**, 11089 (2016).

17. P. Volkov, K. Bacos, J. K. Ofori, J. L. Esguerra, L. Eliasson, T. Ronn, C. Ling, Whole-Genome Bisulfite Sequencing of Human Pancreatic Islets Reveals Novel Differentially Methylated Regions in Type 2 Diabetes Pathogenesis. *Diabetes* **66**, 1074-1085 (2017).

18. J. C. Chambers, M. Loh, B. Lehne, A. Drong, J. Kriebel, V. Motta, S. Wahl, H. R. Elliott, F. Rota, W. R. Scott, W. Zhang, S. T. Tan, G. Campanella, M. Chadeau-Hyam, L. Yengo, R. C. Richmond, M. Adamowicz-Brice, U. Afzal, K. Bozaoglu, Z. Y. Mok, H. K. Ng, F. Pattou, H. Prokisch, M. A.

Rozario, L. Tarantini, J. Abbott, M. Ala-Korpela, B. Albetti, O. Ammerpohl, P. A. Bertazzi, C. Blancher, R. Caiazzo, J. Danesh, T. R. Gaunt, S. de Lusignan, C. Gieger, T. Illig, S. Jha, S. Jones, J. Jowett, A. J. Kangas, A. Kasturiratne, N. Kato, N. Kotea, S. Kowlessur, J. Pitkaniemi, P. Punjabi, D. Saleheen, C. Schafmayer, P. Soininen, E. S. Tai, B. Thorand, J. Tuomilehto, A. R. Wickremasinghe, S. A. Kyrtopoulos, T. J. Aitman, C. Herder, J. Hampe, S. Cauchi, C. L. Relton, P. Froguel, R. Soong, P. Vineis, M. R. Jarvelin, J. Scott, H. Grallert, V. Bollati, P. Elliott, M. I. McCarthy, J. S. Kooner, Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes Endocrinol* **3**, 526-534 (2015).

19. E. Ahlqvist, P. Storm, A. Karajamaki, M. Martinell, M. Dorkhan, A. Carlsson, P. Vikman, R. B. Prasad, D. M. Aly, P. Almgren, Y. Wessman, N. Shaat, P. Spegel, H. Mulder, E. Lindholm, O. Melander, O. Hansson, U. Malmqvist, A. Lernmark, K. Lahti, T. Forsen, T. Tuomi, A. H. Rosengren, L. Groop, Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol*, (2018).

20. A. American Diabetes, 6. Glycemic Targets: Standards of Medical Care in Diabetes-2018. *Diabetes Care* **41**, S55-S64 (2018).

21. E. A. Houseman, J. Molitor, C. J. Marsit, Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics* **30**, 1431-1439 (2014).

22. S. Wahl, A. Drong, B. Lehne, M. Loh, W. R. Scott, S. Kunze, P. C. Tsai, J. S. Ried, W. Zhang, Y. Yang, S. Tan, G. Fiorito, L. Franke, S. Guarrera, S. Kasela, J. Kriebel, R. C. Richmond, M. Adamo, U. Afzal, M. Ala-Korpela, B. Albetti, O. Ammerpohl, J. F. Apperley, M. Beekman, P. A. Bertazzi, S. L. Black, C. Blancher, M. J. Bonder, M. Brosch, M. Carstensen-Kirberg, A. J. de Craen, S. de Lusignan, A. Dehghan, M. Elkalaawy, K. Fischer, O. H. Franco, T. R. Gaunt, J. Hampe, M. Hashemi, A. Isaacs, A. Jenkinson, S. Jha, N. Kato, V. Krogh, M. Laffan, C. Meisinger, T. Meitinger, Z. Y. Mok, V. Motta, H. K. Ng, Z. Nikolakopoulou, G. Nteliopoulos, S. Panico, N. Pervjakova, H. Prokisch, W. Rathmann, M. Roden, F. Rota, M. A. Rozario, J. K. Sandling, C. Schafmayer, K.

Schramm, R. Siebert, P. E. Slagboom, P. Soininen, L. Stolk, K. Strauch, E. S. Tai, L. Tarantini, B. Thorand, E. F. Tigchelaar, R. Tumino, A. G. Uitterlinden, C. van Duijn, J. B. van Meurs, P. Vineis, A. R. Wickremasinghe, C. Wijmenga, T. P. Yang, W. Yuan, A. Zhernakova, R. L. Batterham, G. D. Smith, P. Deloukas, B. T. Heijmans, C. Herder, A. Hofman, C. M. Lindgren, L. Milani, P. van der Harst, A. Peters, T. Illig, C. L. Relton, M. Waldenberger, M. R. Jarvelin, V. Bollati, R. Soong, T. D. Spector, J. Scott, M. I. McCarthy, P. Elliott, J. T. Bell, G. Matullo, C. Gieger, J. S. Kooner, H. Grallert, J. C. Chambers, Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* **541**, 81-86 (2017).

23. M. L. Becker, L. E. Visser, R. H. van Schaik, A. Hofman, A. G. Uitterlinden, B. H. Stricker, Genetic variation in the organic cation transporter 1 is associated with metformin response in patients with diabetes mellitus. *Pharmacogenomics J* **9**, 242-247 (2009).

24. M. M. Christensen, C. Brasch-Andersen, H. Green, F. Nielsen, P. Damkier, H. Beck-Nielsen, K. Brosen, The pharmacogenetics of metformin and its impact on plasma metformin steady-state levels and glycosylated hemoglobin A1c. *Pharmacogenet Genomics* **21**, 837-850 (2011).

25. T. Dujic, T. Bego, M. Malenica, Z. Velija-Asimi, E. Ahlqvist, L. Groop, E. R. Pearson, A. Causevic, S. Semiz, Effects of TCF7L2 rs7903146 variant on metformin response in patients with type 2 diabetes. *Bosn J Basic Med Sci* **19**, 368-374 (2019).

26. T. Dujic, K. Zhou, S. W. Yee, N. van Leeuwen, C. E. de Keyser, M. Javorsky, S. Goswami, L. Zaharenko, M. M. Hougaard Christensen, M. Out, R. Tavendale, M. Kubo, M. M. Hedderson, A. A. van der Heijden, L. Klimcakova, V. Pirags, A. Kooy, K. Brosen, J. Klovins, S. Semiz, I. Tkac, B. H. Stricker, C. Palmer, L. M. t Hart, K. M. Giacomini, E. R. Pearson, Variants in Pharmacokinetic Transporters and Glycemic Response to Metformin: A Metgen Meta-Analysis. *Clin Pharmacol Ther* **101**, 763-772 (2017).

27. E. Shikata, R. Yamamoto, H. Takane, C. Shigemasa, T. Ikeda, K. Otsubo, I. Ieiri, Human organic cation transporter (OCT1 and OCT2) gene polymorphisms and therapeutic effects of metformin. *J Hum Genet* **52**, 117-122 (2007).

28. Y. Shu, S. A. Sheardown, C. Brown, R. P. Owen, S. Zhang, R. A. Castro, A. G. Ianculescu, L. Yue, J. C. Lo, E. G. Burchard, C. M. Brett, K. M. Giacomini, Effect of genetic variation in the organic cation transporter 1 (OCT1) on metformin action. *J Clin Invest* **117**, 1422-1431 (2007).

29. J. N. Todd, J. C. Florez, An update on the pharmacogenomics of metformin: progress, problems and potential. *Pharmacogenomics* **15**, 529-539 (2014).

30. Y. Zhou, W. Ye, Y. Wang, Z. Jiang, X. Meng, Q. Xiao, Q. Zhao, J. Yan, Genetic variants of OCT1 influence glycemic response to metformin in Han Chinese patients with type-2 diabetes mellitus in Shanghai. *Int J Clin Exp Pathol* **8**, 9533-9542 (2015).

31. L. Tarasova, I. Kalnina, K. Geldnere, A. Bumbure, R. Ritenberga, L. Nikitina-Zake, D. Fridmanis, I. Vaivade, V. Pirags, J. Klovins, Association of genetic variation in the organic cation transporters OCT1, OCT2 and multidrug and toxin extrusion 1 transporter protein genes with the gastrointestinal side effects and lower BMI in metformin-treated type 2 diabetes patients. *Pharmacogenet Genomics* **22**, 659-666 (2012).

32. M. D. Nitert, T. Dayeh, P. Volkov, T. Elgzyri, E. Hall, E. Nilsson, B. T. Yang, S. Lang, H. Parikh, Y. Wessman, H. Weishaupt, J. Attema, M. Abels, N. Wierup, P. Almgren, P. A. Jansson, T. Ronn, O. Hansson, K. F. Eriksson, L. Groop, C. Ling, Impact of an exercise intervention on DNA methylation in skeletal muscle from first-degree relatives of patients with type 2 diabetes. *Diabetes* **61**, 3322-3332 (2012).

33. M. R. Luizon, W. L. Eckalbar, Y. Wang, S. L. Jones, R. P. Smith, M. Laurance, L. Lin, P. J. Gallins, A. S. Etheridge, F. Wright, Y. Zhou, C. Molony, F. Innocenti, S. W. Yee, K. M. Giacomini, N. Ahituv, Genomic Characterization of Metformin Hepatic Response. *PLoS Genet* **12**, e1006449 (2016).

34. L. J. McCreight, C. J. Bailey, E. R. Pearson, Metformin and the gastrointestinal tract. *Diabetologia* **59**, 426-435 (2016).

35. N. Moreno-Castellanos, A. Rodriguez, Y. Rabanal-Ruiz, A. Fernandez-Vega, J. Lopez-Miranda, R. Vazquez-Martinez, G. Fruhbeck, M. M. Malagon, The cytoskeletal protein septin 11 is associated

with human obesity and is involved in adipocyte lipid storage and metabolism. *Diabetologia* **60**, 324-335 (2017).

36. C. Wolfrum, E. Asilmaz, E. Luca, J. M. Friedman, M. Stoffel, Foxa2 regulates lipid metabolism and ketogenesis in the liver during fasting and in diabetes. *Nature* **432**, 1027-1032 (2004).

37. S. Y. Wong, T. Gadomski, M. van Scherpenzeel, T. Honzik, H. Hansikova, K. S. B. Holmefjord, M. Mork, F. Bowling, J. Sykut-Cegielska, D. Koch, J. Hertecant, G. Preston, J. Jaeken, N. Peeters, S. Perez, D. D. Nguyen, K. Crivelly, T. Emmerzaal, K. M. Gibson, K. Raymond, N. Abu Bakar, F. Foulquier, G. Poschet, A. M. Ackermann, M. He, D. J. Lefeber, C. Thiel, T. Kozicz, E. Morava, Oral D-galactose supplementation in PGM1-CDG. *Genet Med* **19**, 1226-1235 (2017).

38. J. Y. Hwang, X. Sim, Y. Wu, J. Liang, Y. Tabara, C. Hu, K. Hara, C. H. Tam, Q. Cai, Q. Zhao, S. Jee, F. Takeuchi, M. J. Go, R. T. Ong, T. Ohkubo, Y. J. Kim, R. Zhang, T. Yamauchi, W. Y. So, J. Long, D. Gu, N. R. Lee, S. Kim, T. Katsuya, J. H. Oh, J. Liu, S. Umemura, Y. J. Kim, F. Jiang, S. Maeda, J. C. Chan, W. Lu, J. E. Hixson, L. S. Adair, K. J. Jung, T. Nabika, J. B. Bae, M. H. Lee, M. Seielstad, T. L. Young, Y. Y. Teo, Y. Kita, N. Takashima, H. Osawa, S. H. Lee, M. H. Shin, D. H. Shin, B. Y. Choi, J. Shi, Y. T. Gao, Y. B. Xiang, W. Zheng, N. Kato, M. Yoon, J. He, X. O. Shu, R. C. Ma, T. Kadowaki, W. Jia, T. Miki, L. Qi, E. S. Tai, K. L. Mohlke, B. G. Han, Y. S. Cho, B. J. Kim, Genome-wide association meta-analysis identifies novel variants associated with fasting plasma glucose in East Asians. *Diabetes* **64**, 291-298 (2015).

39. A. K. Manning, M. F. Hivert, R. A. Scott, J. L. Grimsby, N. Bouatia-Naji, H. Chen, D. Rybin, C. T. Liu, L. F. Bielak, I. Prokopenko, N. Amin, D. Barnes, G. Cadby, J. J. Hottenga, E. Ingelsson, A. U. Jackson, T. Johnson, S. Kanoni, C. Ladenvall, V. Lagou, J. Lahti, C. Lecoeur, Y. Liu, M. T. Martinez-Larrad, M. E. Montasser, P. Navarro, J. R. Perry, L. J. Rasmussen-Torvik, P. Salo, N. Sattar, D. Shungin, R. J. Strawbridge, T. Tanaka, C. M. van Duijn, P. An, M. de Andrade, J. S. Andrews, T. Aspelund, M. Atalay, Y. Aulchenko, B. Balkau, S. Bandinelli, J. S. Beckmann, J. P. Beilby, C. Bellis, R. N. Bergman, J. Blangero, M. Boban, M. Boehnke, E. Boerwinkle, L. L. Bonnycastle, D. I. Boomsma, I. B. Borecki, Y. Bottcher, C. Bouchard, E. Brunner, D. Budimir, H.

Campbell, O. Carlson, P. S. Chines, R. Clarke, F. S. Collins, A. Corbaton-Anchuelo, D. Couper, U. de Faire, G. V. Dedoussis, P. Deloukas, M. Dimitriou, J. M. Egan, G. Eiriksdottir, M. R. Erdos, J. G. Eriksson, E. Eury, L. Ferrucci, I. Ford, N. G. Forouhi, C. S. Fox, M. G. Franzosi, P. W. Franks, T. M. Frayling, P. Froguel, P. Galan, E. de Geus, B. Gigante, N. L. Glazer, A. Goel, L. Groop, V. Gudnason, G. Hallmans, A. Hamsten, O. Hansson, T. B. Harris, C. Hayward, S. Heath, S. Hercberg, A. A. Hicks, A. Hingorani, A. Hofman, J. Hui, J. Hung, M. R. Jarvelin, M. A. Jhun, P. C. Johnson, J. W. Jukema, A. Jula, W. H. Kao, J. Kaprio, S. L. Kardia, S. Keinanen-Kiukaanniemi, M. Kivimaki, I. Kolcic, P. Kovacs, M. Kumari, J. Kuusisto, K. O. Kyvik, M. Laakso, T. Lakka, L. Lannfelt, G. M. Lathrop, L. J. Launer, K. Leander, G. Li, L. Lind, J. Lindstrom, S. Lobbens, R. J. Loos, J. Luan, V. Lyssenko, R. Magi, P. K. Magnusson, M. Marmot, P. Meneton, K. L. Mohlke, V. Mooser, M. A. Morken, I. Miljkovic, N. Narisu, J. O'Connell, K. K. Ong, B. A. Oostra, L. J. Palmer, A. Palotie, J. S. Pankow, J. F. Peden, N. L. Pedersen, M. Pehlic, L. Peltonen, B. Penninx, M. Pericic, M. Perola, L. Perusse, P. A. Peyser, O. Polasek, P. P. Pramstaller, M. A. Province, K. Raikkonen, R. Rauramaa, E. Rehnberg, K. Rice, J. I. Rotter, I. Rudan, A. Ruokonen, T. Saaristo, M. Sabater-Lleal, V. Salomaa, D. B. Savage, R. Saxena, P. Schwarz, U. Seedorf, B. Sennblad, M. Serrano-Rios, A. R. Shuldiner, E. J. Sijbrands, D. S. Siscovick, J. H. Smit, K. S. Small, N. L. Smith, A. V. Smith, A. Stancakova, K. Stirrups, M. Stumvoll, Y. V. Sun, A. J. Swift, A. Tonjes, J. Tuomilehto, S. Trompet, A. G. Uitterlinden, M. Uusitupa, M. Vikstrom, V. Vitart, M. C. Vohl, B. F. Voight, P. Vollenweider, G. Waeber, D. M. Waterworth, H. Watkins, E. Wheeler, E. Widen, S. H. Wild, S. M. Willems, G. Willemsen, J. F. Wilson, J. C. Witteman, A. F. Wright, H. Yaghootkar, D. Zelenika, T. Zemunik, L. Zgaga, D. I. G. Replication, C. Meta-analysis, C. Multiple Tissue Human Expression Resource, N. J. Wareham, M. I. McCarthy, I. Barroso, R. M. Watanabe, J. C. Florez, J. Dupuis, J. B. Meigs, C. Langenberg, A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet* **44**, 659-669 (2012).

40. C. Xing, J. C. Cohen, E. Boerwinkle, A weighted false discovery rate control procedure reveals alleles at FOXA2 that influence fasting glucose levels. *Am J Hum Genet* **86**, 440-446 (2010).

41. Z. Zhu, Y. Lin, X. Li, J. A. Driver, L. Liang, Shared genetic architecture between metabolic traits and Alzheimer's disease: a large-scale genome-wide cross-trait analysis. *Hum Genet* **138**, 271-285 (2019).

42. G. Rena, D. G. Hardie, E. R. Pearson, The mechanisms of action of metformin. *Diabetologia* **60**, 1577-1585 (2017).

43. A. Y. Dawed, A. Ali, K. Zhou, E. R. Pearson, P. W. Franks, Evidence-based prioritisation and enrichment of genes interacting with metformin in type 2 diabetes. *Diabetologia* **60**, 2231-2239 (2017).

44. S. L. Stocker, K. M. Morrissey, S. W. Yee, R. A. Castro, L. Xu, A. Dahlin, A. H. Ramirez, D. M. Roden, R. A. Wilke, C. A. McCarty, R. L. Davis, C. M. Brett, K. M. Giacomini, The effect of novel promoter variants in MATE1 and MATE2 on the pharmacokinetics and pharmacodynamics of metformin. *Clin Pharmacol Ther* **93**, 186-194 (2013).

45. K. Toyama, A. Yonezawa, S. Masuda, R. Osawa, M. Hosokawa, S. Fujimoto, N. Inagaki, K. Inui, T. Katsura, Loss of multidrug and toxin extrusion 1 (MATE1) is associated with metformin-induced lactic acidosis. *Br J Pharmacol* **166**, 1183-1191 (2012).

46. M. Foretz, S. Hebrard, J. Leclerc, E. Zarrinpashneh, M. Soty, G. Mithieux, K. Sakamoto, F. Andreelli, B. Viollet, Metformin inhibits hepatic gluconeogenesis in mice independently of the LKB1/AMPK pathway via a decrease in hepatic energy state. *J Clin Invest* **120**, 2355-2369 (2010).

47. C. Ling, L. Groop, Epigenetics: a molecular link between environmental factors and type 2 diabetes. *Diabetes* **58**, 2718-2725 (2009).

48. D. W. Hosmer, S. Lemeshow, *Applied logistic regression*. Wiley series in probability and statistics Texts and references section (Wiley, New York, ed. 2nd, 2000), pp. xii, 373 pages.

49. J. N. Mandrekar, Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* **5**, 1315-1316 (2010).

50. J. J. Caro, A. J. Ward, J. A. O'Brien, Lifetime costs of complications resulting from type 2 diabetes in the U.S. *Diabetes Care* **25**, 476-481 (2002).

51. A. Oussalah, S. Rischer, M. Bensenane, G. Conroy, P. Filhine-Tresarrieu, R. Debard, D. Forest-Tramoy, T. Josse, D. Reinicke, M. Garcia, A. Luc, C. Baumann, A. Ayav, V. Laurent, M. Hollenbach, C. Ripoll, R. M. Gueant-Rodriguez, F. Namour, A. Zipprich, M. Fleischhacker, J. P. Bronowicki, J. L. Gueant, Plasma mSEPT9: A Novel Circulating Cell-free DNA-Based Epigenetic Biomarker to Diagnose Hepatocellular Carcinoma. *EBioMedicine* **30**, 138-147 (2018).

52. M. C. Benton, A. Johnstone, D. Eccles, B. Harmon, M. T. Hayes, R. A. Lea, L. Griffiths, E. P. Hoffman, R. S. Stubbs, D. Macartney-Coxson, An analysis of DNA methylation in human adipose tissue reveals differential modification of obesity genes before and after gastric bypass and weight loss. *Genome Biol* **16**, 8 (2015).

53. L. Gillberg, A. Perfilyev, C. Brons, M. Thomasen, L. G. Grunnet, P. Volkov, F. Rosqvist, D. Iggman, I. Dahlman, U. Riserus, T. Ronn, E. Nilsson, A. Vaag, C. Ling, Adipose tissue transcriptomics and epigenomics in low birthweight men and controls: role of high-fat overfeeding. *Diabetologia* **59**, 799-812 (2016).

54. T. Ronn, P. Volkov, C. Davegardh, T. Dayeh, E. Hall, A. H. Olsson, E. Nilsson, A. Tornberg, M. Dekker Nitert, K. F. Eriksson, H. A. Jones, L. Groop, C. Ling, A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. *PLoS Genet* **9**, e1003572 (2013).

55. Q. Chen, X. Yang, H. Zhang, X. Kong, L. Yao, X. Cui, Y. Zou, F. Fang, J. Yang, Y. Chang, Metformin impairs systemic bile acid homeostasis through regulating SIRT1 protein levels. *Biochim Biophys Acta Mol Cell Res* **1864**, 101-112 (2017).

56. J. Hur, K. A. Sullivan, M. Pande, Y. Hong, A. A. Sima, H. V. Jagadish, M. Kretzler, E. L. Feldman, The identification of gene expression profiles associated with progression of human diabetic neuropathy. *Brain* **134**, 3222-3235 (2011).

57. D. N. Dai, Y. Li, B. Chen, Y. Du, S. B. Li, S. X. Lu, Z. P. Zhao, A. J. Zhou, N. Xue, T. L. Xia, M. S. Zeng, Q. Zhong, W. D. Wei, Elevated expression of CST1 promotes breast cancer progression and predicts a poor prognosis. *J Mol Med (Berl)* **95**, 873-886 (2017).

58. L. Zaharenko, I. Kalnina, K. Geldnere, I. Konrade, S. Grinberga, J. Zidzik, M. Javorsky, A. Lejnieks, L. Nikitina-Zake, D. Fridmanis, R. Peculis, I. Radovica-Spalvina, D. Hartmane, O. Pugovics, I. Tkac, L. Klimcakova, V. Pirags, J. Klovins, Single nucleotide polymorphisms in the intergenic region between metformin transporter OCT2 and OCT3 coding genes are associated with short-term response to metformin monotherapy in type 2 diabetes mellitus patients. *Eur J Endocrinol* **175**, 531-540 (2016).

59. V. Rovite, Y. Wolff-Sagi, L. Zaharenko, L. Nikitina-Zake, E. Grens, J. Klovins, Genome Database of the Latvian Population (LGDB): Design, Goals, and Primary Results. *J Epidemiol* **28**, 353-360 (2018).

60. A. American Diabetes, 8. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes-2018. *Diabetes Care* **41**, S73-S85 (2018).

61. G. Control, F. M. Turnbull, C. Abraira, R. J. Anderson, R. P. Byington, J. P. Chalmers, W. C. Duckworth, G. W. Evans, H. C. Gerstein, R. R. Holman, T. E. Moritz, B. C. Neal, T. Ninomiya, A. A. Patel, S. K. Paul, F. Travert, M. Woodward, Intensive glucose control and macrovascular outcomes in type 2 diabetes. *Diabetologia* **52**, 2288-2298 (2009).

62. I. M. Stratton, A. I. Adler, H. A. Neil, D. R. Matthews, S. E. Manley, C. A. Cull, D. Hadden, R. C. Turner, R. R. Holman, Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. *BMJ* **321**, 405-412 (2000).

63. S. Moran, C. Arribas, M. Esteller, Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389-399 (2016).

64. R. C. Gentleman, V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang, J. Zhang, Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80 (2004).

65. D. L. McCartney, R. M. Walker, S. W. Morris, A. M. McIntosh, D. J. Porteous, K. L. Evans, Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genom Data* **9**, 22-24 (2016).

66. P. Du, X. Zhang, C. C. Huang, N. Jafari, W. A. Kibbe, L. Hou, S. M. Lin, Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* **11**, 587 (2010).

67. P. Du, W. A. Kibbe, S. M. Lin, lumi: a pipeline for processing Illumina microarray. *Bioinformatics* **24**, 1547-1548 (2008).

68. W. E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118-127 (2007).

69. A. E. Teschendorff, F. Marabita, M. Lechner, T. Bartlett, J. Tegner, D. Gomez-Cabrero, S. Beck, A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189-196 (2013).

70. X. Stephenne, M. Foretz, N. Taleux, G. C. van der Zon, E. Sokal, L. Hue, B. Viollet, B. Guigas, Metformin activates AMP-activated protein kinase in primary human hepatocytes by decreasing cellular energy status. *Diabetologia* **54**, 3101-3110 (2011).

Acknowledgments: We thank Ylva Wessman, Maria Sterner, Emma Nilsson, the SCIBLU genomics facility at Lund University and the Genome Database of the Latvian Population for providing biological material and data for the OPTIMED cohort. We thank Prof. Ewan Pearson for his help in the design of statistical models. **Funding:** This work was supported by grants from the Novonordisk foundation, Swedish Research Council, Region Skåne (ALF), ERC-Co Grant (PAINTBOX, No 725840), H2020-Marie Skłodowska-Curie grant agreement No 706081 (EpiHope), The Swedish Heart Lung Foundation, EFSD, Exodiab, Swedish Foundation for Strategic Research for IRC15-0067, Swedish Diabetes Foundation, Albert Påhlsson Foundation, ERC-CoG-2015_681742_NASCENT. The group of Swedish twins was recruited from the Swedish Twin Registry, which is supported by grants from the Swedish Research Council. The funders had no role in study design, data collection, analysis and interpretation, decision to publish, or preparation of the manuscript. **Author contributions:** SG, LG, EA, and CL contributed to the conception of the work. SG, SK, PWF, MM, MU, IE, JP, AV, LG, JK, EA, and CL contributed to the data collection. SG, AP, MM, SK, KB, EA, and PV contributed to the data analysis. SG, SK and KB performed experiments. SG and CL drafted the article. All authors contributed to the interpretation of data and critical revision of the article. All authors gave final approval of the version to be published. Dr Sonia García-Calzón and Prof Charlotte Ling had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. SG and CL are guarantors. **Competing interests:** SG, AP and CL have a pending patent ("DNA methylation level of specific CpG sites for prediction of glycemic response and tolerance to metformin treatment in type 2 diabetic patients", P5066SE00) on using the degree of DNA methylation of specific sites for prediction of glycemic response and tolerance to metformin treatment in patients with T2D. PWF has received research funding from Boehringer Ingelheim, Eli Lilly, Janssen, Novo Nordisk A/S, Sanofi Aventis and Servier, received consulting fees from Eli Lilly, Novo Nordisk and Zoe Global Ltd and has stock options in Zoe Global Ltd. PV works now at the pharma company AstraZeneca. The other authors declare that they have no conflicts of interest. **Data and material availability:** Data associated with this study are available in the main text or the supplementary materials. DNA methylation data with accession numbers;

LUDC2020.08.1 (Discovery cohort for metformin response), LUDC2020.08.2 (ANDIS replication cohort for metformin response), LUDC2020.08.3 (Discovery cohort for metformin response case-control), LUDC2020.08.4 (ANDIS replication cohort for metformin response case-control), LUDC2020.08.5 (European replication cohort for metformin response case-control), LUDC2020.08.6 (Discovery cohort for metformin intolerance), LUDC2020.08.7 (ANDIS replication cohort for metformin intolerance) and LUDC2020.08.8 (European replication cohort for metformin intolerance) are deposited in the LUDC repository (https://www.ludc.lu.se/resources/repository) and are available upon on request.

Figure captions

Fig. 1. Study design. DNA methylation was analyzed genome-wide to identify blood-based epigenetic markers that could associate with change in HbA1c and discriminate future glycemic response and intolerance to metformin therapy. Discovery and replication cohorts from ANDIS, ANDIU and OPTIMED were included. A fixed meta-analysis was performed to select individual methylation markers associated with future metformin response or intolerance. Methylation risk scores (MRS) were calculated and used to stratify patients with T2D into glycemic responders/non-responders and metformin intolerant/tolerant. We then assessed if these blood-based epigenetic markers mirror DNA methylation in human adipose tissue, a central tissue of diabetes. Last, functional in vitro follow-up experiments in hepatocytes tested if genes annotated to the identified epigenetic markers might influence phenotypes related to metformin therapy such as expression of metformin transporters and regulators of gluconeogenesis as well as AMPK activity.

Fig. 2. Associations between DNA methylation and metformin response and intolerance in the ANDIS discovery cohorts. A) Associations between methylation and ΔHbA1c in 63 drug-naïve subjects with T2D after adjusting for basal HbA1c, eGFR, and time gaps (between baseline HbA1c and methylation measurements and the start of metformin) displaying 2,583 significant CpG sites (FDR < 5%, q-value<0.05). **B)** Associations between methylation and metformin response in 21 non-responders and 26 responders after adjusting for basal HbA1c, eGFR and, time gaps in baseline HbA1c and methylation displaying 7,973 significant CpG sites (FDR below 5%, q-value<0.05). Beta-coefficients in the Volcano plot are shown when comparing glycemic non-responders vs. responders to metformin. **C)** Associations between methylation and metformin intolerance in 66 tolerant and 17 intolerant subjects with T2D after adjusting for basal HbA1c, eGFR and, time gap in baseline methylation displaying 12,579 significant CpG sites (FDR below 5%, q-value<0.05). Beta-coefficients in the Volcano plot are shown when comparing metformin intolerant vs. tolerant subjects. Blue dashed lines and red lines indicate methylome-wide significance (q-value<0.05).

Fig. 3. Risk for not responding to or not tolerating metformin of methylation markers associated with glycemic response or intolerance. Logistic models were performed for each CpG site in all subjects with T2D included in the discovery and replication cohorts for metformin response (n=212) (**A**) and in all subjects included in the discovery and replication cohorts for metformin intolerance (n=151) (**B**). Odds ratios (OR) are shown per 1 SD increase in methylation for each CpG site.

Fig. 4. Combined MRS discriminate between glycemic responders and non-responders to metformin in drug-naïve subjects with T2D**.** The MRS include the 11 CpG sites associated with future metformin response (see Table 1). CpG-specific effect sizes (beta-coefficients from logistic models) from the ANDIS discovery cohort for metformin response (n=47) (**A**) were used to calculate and evaluate the MRS in the ANDIS (n=87) (**B-D**) and European (n=78) (**E-G**) replication cohorts for metformin response. Boxplots show significantly different MRS between glycemic responders and non-responders to metformin in both the ANDIS replication (P for U Mann-Whitney= 6.6×10^{-7}) (**B**) and the European replication (P for U Mann-Whitney= 1.6×10^{-10}) (**E**) cohorts. Histogram plots show distributions of the MRS stratified by response to metformin in the ANDIS (**C**) and the European (**F**) replication cohorts. Red bars represent non-responders, yellow bars represent responders to metformin. The ROC curves show the discrimination between responders/non-responders based on MRS. The AUC for metformin response was 0.80 in the ANDIS replication cohort (**D**) and 0.89 in the European replication cohort (**G**).

Fig. 5. Combined MRS discriminate metformin tolerance and intolerance in drug-naïve subjects with T2D**.** The MRS include the 4 CpG sites associated with future metformin intolerance (see Table 1). CpG-specific effect sizes (beta-coefficients from logistic models) from the ANDIS discovery cohort for metformin intolerance (n=83) (**A**) were used to calculate and evaluate the MRS in the ANDIS (n=48) (**B-D**) and European (n=20) (**E-G**) replication cohorts for metformin intolerance. Boxplots show significantly different MRS between tolerant and intolerant subjects to metformin in both the ANDIS replication (P for U Mann-Whitney= 4.5×10^{-7}) (**B**) and the European replication (P for U Mann-Whitney= 1.5×10^{-2}) (**E**) cohorts. Histogram plots show distributions of the MRS stratified by intolerance to metformin in the

ANDIS (**C**) and the European (**F**) replication cohorts. Red bars represent intolerant subjects, yellow bars represent tolerant subjects to metformin. The ROC curves show the discrimination between tolerant/intolerant subjects based on MRS. The AUC (for metformin intolerance was 0.94 in the ANDIS replication cohort (**D**) and 0.87 in the European replication cohort (**G**).

Fig. 6. Silencing of genes associated with metformin response (*SEPT11* and *CST1*) or intolerance (*FOXA2* and *PGM1*) in hepatocytes affects expression of metformin transporters, AMPK activity, and expression of key regulators of gluconeogenesis. A) Quantification of siRNA-mediated knockdown of *SEPT11*, *CST1*, *FOXA2*, and *PGM1* (siSEPT11, siCST1, siFOXA2, siPGM1) compared with negative control siRNA (siNC) in Hep2G cells. **B)** AMPK activity and expression of key regulators of gluconeogenesis (*PCK1* and *G6PC*) and metformin transporters (*SLC22A1*, *SLC47A1*) in metformin-treated (exposed to 5 mM and 2.5mM metformin, respectively) compared to non-metformin exposed HepG2 cells. **C-F)** mRNA expression of metformin transporters (*SLC22A1*, *SLC47A1)*, AMPK activity, and mRNA expression of key regulators of gluconeogenesis (*PCK1* and *G6PC*) in Hep2G cells deficient for *SEPT11* (**C**), *CST1* (**D**), *FOXA2* (**E**) or *PGM1* (**F**) expression compared to cells transfected with siNC after non-metformin exposure or metformin treatment overnight. For all panels, data are mean ± SEM of four independent experiments performed in different passages of Hep2G cells, with two technical replicates for each condition. *P* values were calculated using paired t-tests, #*P*=0.056-0.075, **P* <0.05, ***P*<0.01, ****P*<0.001. Two-sided *P* values were calculated using paired t-tests of logged values for all the analyses.

Table 1. Differentially methylated CpGs associated with future metformin response or intolerance in the combined meta-analysis of discovery and replication cohorts ($p<6.1 \times 10^{-8}$).

				ANDIS discovery cohort			ANDIS replication cohort			European replication cohort			Meta-analysis		
CpG sites	**Gene**	**Gene region**	**CpG Island region**	**Beta-coeff**	**SEM**	**p-value**	**Beta-coeff**	**SEM**	**p-value**	**Beta-coeff**	**SEM**	**p-value**	**Beta-coeff**	**SEM**	**p-value**
Metformin glycemic response															
cg00153082		Intergenic	Open sea	0.55	0.11	6.43E-06	0.22	0.09	1.60E-02	0.30	0.10	2.21E-03	0.34	0.06	1.13E-09
cg03529510	*CFAP58*	Body	Open sea	0.32	0.07	2.87E-05	0.25	0.10	1.60E-02	0.25	0.08	4.13E-03	0.28	0.05	1.47E-09
cg05402062	*OR4S1*	TSS1500	Open sea	0.34	0.08	2.86E-04	0.22	0.10	3.40E-02	0.29	0.07	1.56E-04	0.29	0.05	1.98E-09
cg16704073	*GPHA2*	Body	Open sea	0.41	0.08	9.06E-06	0.18	0.08	3.40E-02	0.25	0.09	7.62E-03	0.28	0.05	5.63E-09
cg01894192		Intergenic	Open sea	0.25	0.05	1.38E-05	0.15	0.07	2.70E-02	0.14	0.06	2.85E-02	0.19	0.03	1.14E-08
cg16240962	*SAP130*	TSS1500	Island	0.14	0.03	1.95E-04	0.13	0.04	3.00E-03	0.10	0.04	1.41E-02	0.12	0.02	1.16E-08
cg01070242	*SEPT11*	5'UTR;Body	S_Shelf	0.43	0.09	1.40E-05	0.25	0.11	2.20E-02	0.23	0.10	1.89E-02	0.31	0.06	1.21E-08
cg08713722		Intergenic	S_Shore	0.21	0.05	6.32E-05	0.12	0.05	1.20E-02	0.15	0.05	4.96E-03	0.16	0.03	1.26E-08
cg05151280	*LRRN2*	5'UTR	Open sea	0.26	0.06	5.63E-05	0.14	0.05	4.00E-03	0.12	0.05	1.37E-02	0.16	0.03	1.62E-08

cg07511259	*CSTT*	TSS1500	Open sea	0.25	0.06	2.43E-04	0.14	0.05	1.30E-02	0.21	0.07	2.32E-03	0.19	0.03	2.56E-08
cg01282725		Intergenic	Open sea	0.23	0.06	2.35E-04	0.15	0.06	8.00E-03	0.21	0.07	4.67E-03	0.19	0.04	2.88E-08
Metformin Intolerance															
cg27553780	*SCYL1*	Body	S_Shelf	0.29	0.08	3.33E-04	0.29	0.09	2.91E-03	0.31	0.09	4.90E-03	0.30	0.05	2.75E-09
cg12356107	*FOXA2*	TSS1500	Island	0.36	0.08	2.93E-05	0.37	0.13	6.80E-03	0.34	0.14	3.03E-02	0.36	0.06	4.91E-09
cg02994863	*PGM1*	1stExon	Island	0.49	0.1	6.15E-06	0.2	0.08	2.21E-02	0.36	0.13	1.37E-02	0.32	0.06	1.08E-08
cg08148545	*FAM107A*	TSS200;Body	S_Shore	0.29	0.08	2.41E-04	0.46	0.12	2.59E-04	0.42	0.19	4.41E-02	0.35	0.06	2.35E-08

Beta-coefficients (SEM) are estimated from a linear model in the given cohort with either non-responders vs. responders to metformin (top panel) or intolerant vs. tolerant to metformin (bottom panel).
Models were adjusted for basal HbA1c, eGFR, and time gaps (time gap in baseline HbA1c and time methylation gap) for metformin response and eGFR and time methylation gap for metformin intolerance. Time gap in baseline HbA1c was defined as the number of days between the measurement of baseline HbA1c and the start of metformin therapy, whereas time gap in baseline methylation was defined as the number of days between the measurement of DNA methylation in blood and the start of metformin therapy. Inverse variance fixed meta-analysis of discovery and replication cohorts was performed and Bonferroni-corrected. $P<6.1 \times 10^{-8}$ was considered significant.

Fig. 1

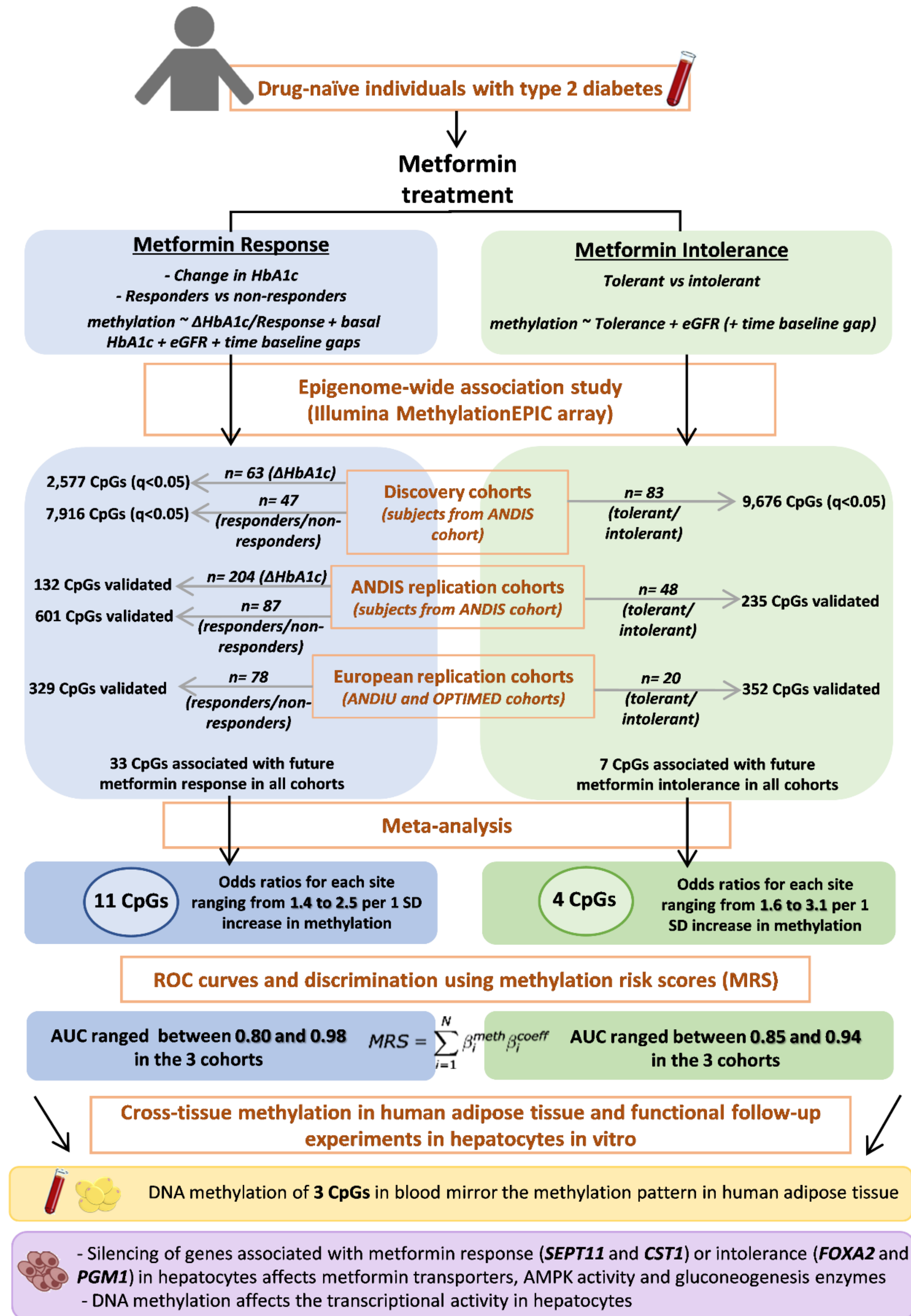


Fig. 2

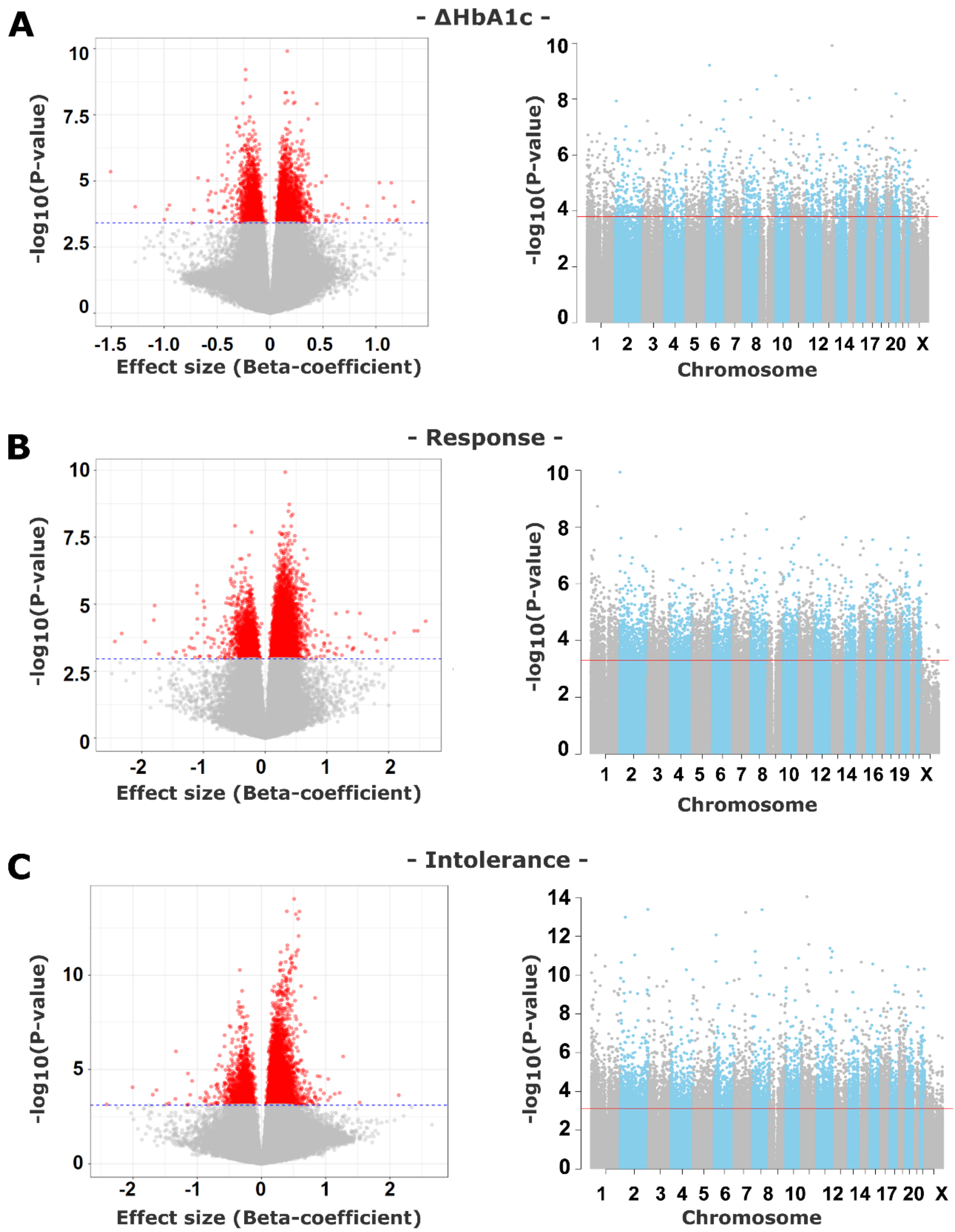
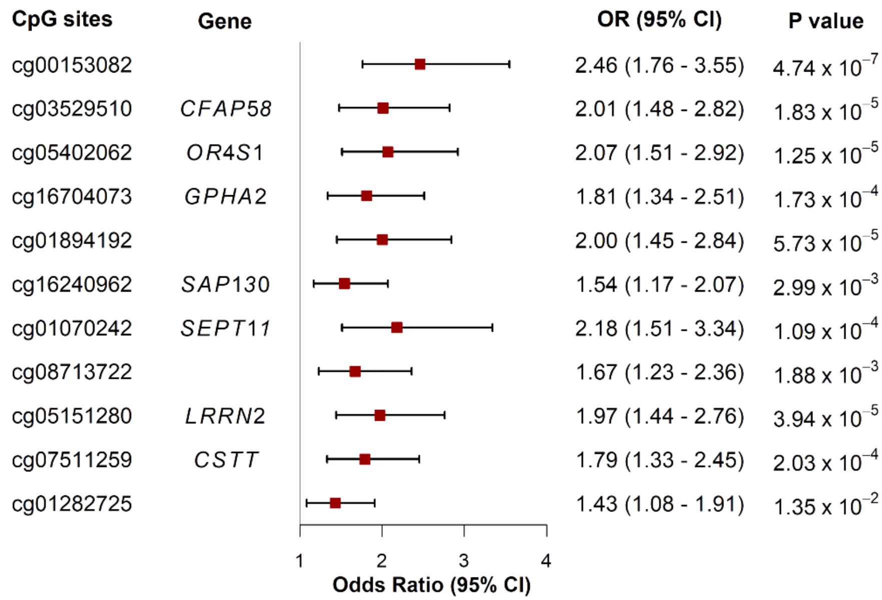


Fig 3.

A



B

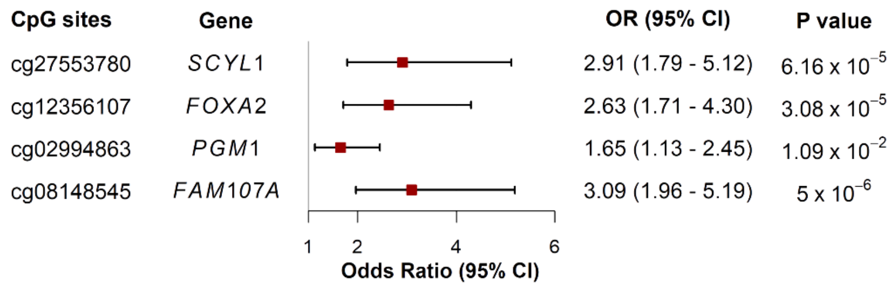
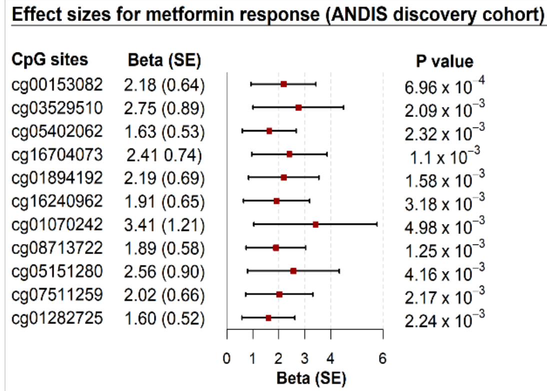
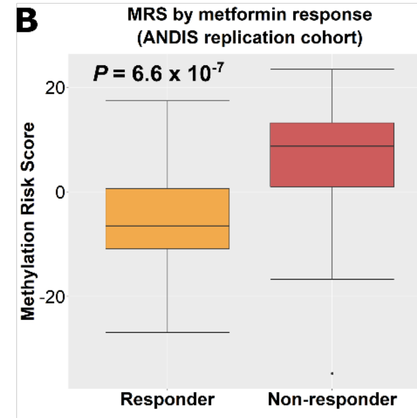
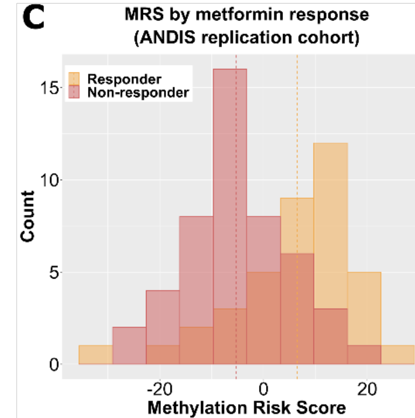
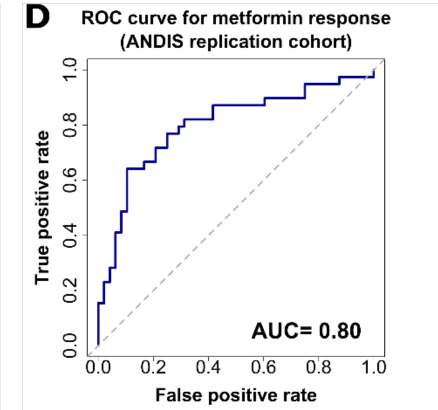


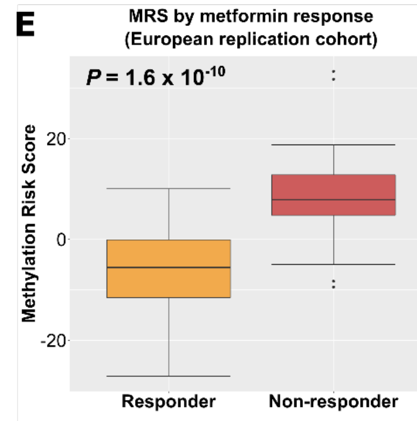
Fig 4.

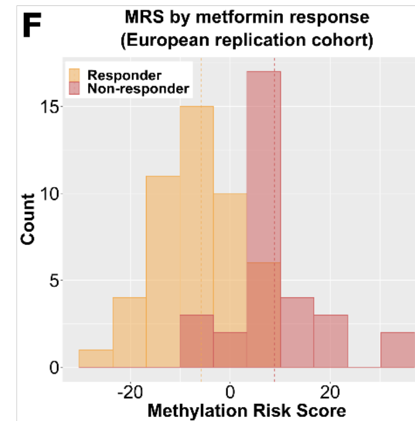












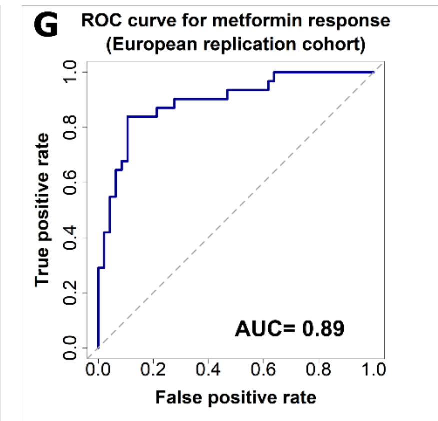
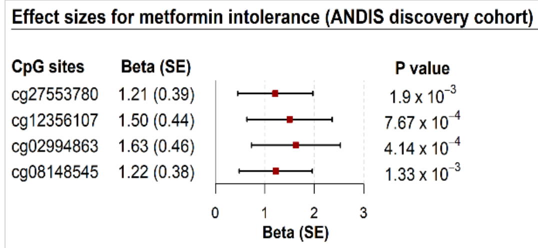
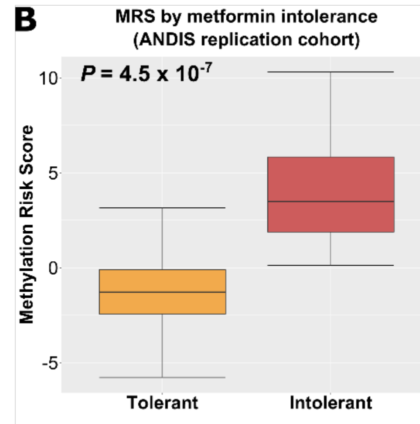
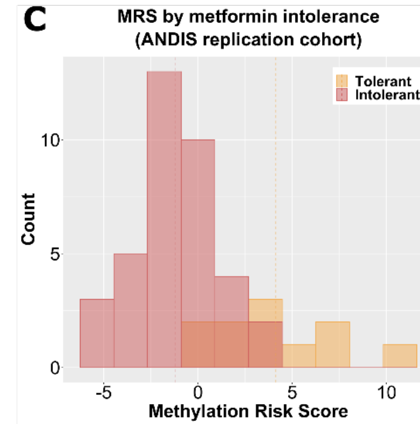
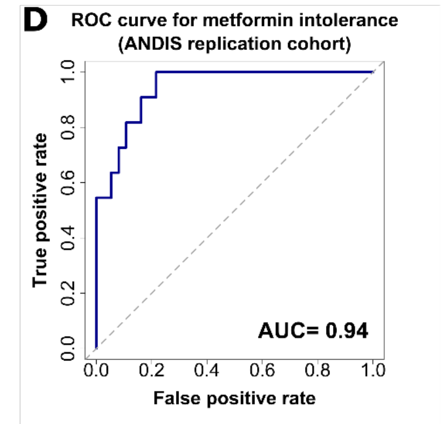


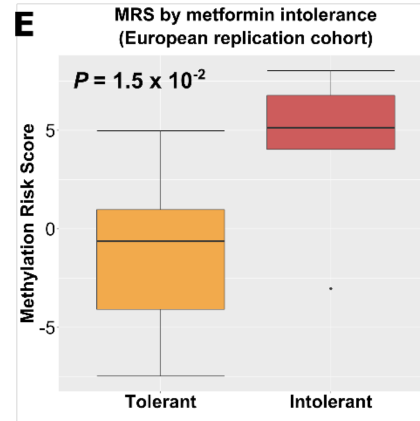
Fig 5.

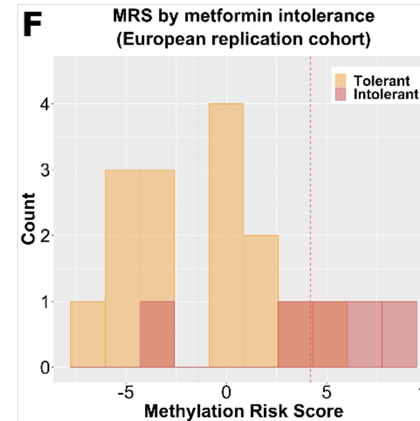












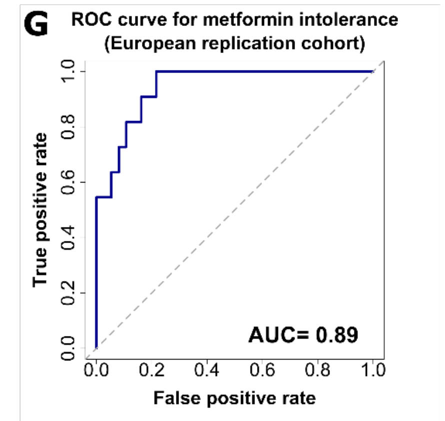


Fig 6.