

# Genetic Variants Associated With Glycine Metabolism and Their Role in Insulin Sensitivity and Type 2 Diabetes

Weijia Xie,<sup>1</sup> Andrew R. Wood,<sup>1</sup> Valeriya Lyssenko,<sup>2</sup> Michael N. Weedon,<sup>1</sup> Joshua W. Knowles,<sup>3</sup> Sami Alkayyali,<sup>2</sup> Themistocles L. Assimes,<sup>3</sup> Thomas Quertermous,<sup>3</sup> Fahim Abbasi,<sup>3</sup> Jussi Paananen,<sup>4</sup> Hans Häring,<sup>5</sup> Torben Hansen,<sup>6,13</sup> Oluf Pedersen,<sup>6,14,15,16</sup> Ulf Smith,<sup>7</sup> Markku Laakso,<sup>4</sup> the MAGIC Investigators, the DIAGRAM Consortium, the GENESIS Consortium, the RISC Consortium, Jacqueline M. Dekker,<sup>8</sup> John J. Nolan,<sup>9</sup> Leif Groop,<sup>2</sup> Ele Ferrannini,<sup>10</sup> Klaus-Peter Adam,<sup>11</sup> Walter E. Gall,<sup>11</sup> Timothy M. Frayling,<sup>1</sup> and Mark Walker<sup>12</sup>

Circulating metabolites associated with insulin sensitivity may represent useful biomarkers, but their causal role in insulin sensitivity and diabetes is less certain. We previously identified novel metabolites correlated with insulin sensitivity measured by the hyperinsulinemic-euglycemic clamp. The top-ranking metabolites were in the glutathione and glycine biosynthesis pathways. We aimed to identify common genetic variants associated with metabolites in these pathways and test their role in insulin sensitivity and type 2 diabetes. With 1,004 nondiabetic individuals from the RISC study, we performed a genome-wide association study (GWAS) of 14 insulin sensitivity-related metabolites and one metabolite ratio. We replicated our results in the Botnia study ( $n = 342$ ). We assessed the association of these variants with diabetes-related traits in GWAS meta-analyses (GENESIS [including RISC, EUGENE2, and Stanford], MAGIC, and DIAGRAM). We identified four associations with three metabolites—glycine (rs715 at *CPS1*), serine (rs478093 at *PHGDH*), and betaine (rs499368 at *SLC6A12*; rs17823642 at *BHMT*)—and one association signal with glycine-to-serine ratio (rs1107366 at *ALDH1L1*). There was no robust evidence for association between these variants and insulin resistance or diabetes. Genetic variants associated with genes in the glycine biosynthesis pathways do not provide consistent evidence for a role of glycine in diabetes-related traits. *Diabetes* 62:2141–2150, 2013

From <sup>1</sup>Genetics of Complex Traits, Peninsula School of Medicine, University of Exeter, Exeter, U.K.; <sup>2</sup>Lund University Diabetes Center, Department of Clinical Sciences, Diabetes and Endocrinology, Lund University, Malmö, Sweden; the <sup>3</sup>Department of Medicine, Stanford University School of Medicine, Stanford, California; the <sup>4</sup>Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland; the <sup>5</sup>Division of Endocrinology, Diabetology, Nephrology, Vascular Medicine and Clinical Chemistry, Department of Internal Medicine, University of Tübingen, Tübingen, Germany; <sup>6</sup>Novo Nordisk Foundation Center for Basic Metabolic Research Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark; the <sup>7</sup>Lundberg Laboratory for Diabetes Research, Department of Molecular and Clinical Medicine, Sahlgrenska Academy, Gothenburg, Sweden; the <sup>8</sup>Department of Epidemiology and Biostatistics, Vrije Universiteit Medical Center, Amsterdam, the Netherlands; the <sup>9</sup>Steno Diabetes Center, Gentofte, Denmark; the <sup>10</sup>Department of Internal Medicine, University of Pisa, Pisa, Italy; <sup>11</sup>Metabolon, Inc., Research Triangle Park, North Carolina; the <sup>12</sup>Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, U.K.; the <sup>13</sup>Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; <sup>14</sup>Hagedorn Research Institute, Copenhagen, Denmark; the <sup>15</sup>Faculty of Health Sciences, Institute of Biomedical Science, University of Copenhagen, Copenhagen, Denmark; and the <sup>16</sup>Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark.

Corresponding author: Timothy M. Frayling, tim.frayling@pms.ac.uk.

Received 29 June 2012 and accepted 24 January 2013.

DOI: 10.2337/db12-0876

This article contains Supplementary Data online at <http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db12-0876/-/DC1>.

A full list of investigators can be found in the Supplementary Data online.

© 2013 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

Using mass spectrometry-based metabolomic approaches, recent studies have identified associations between small molecules and insulin sensitivity and type 2 diabetes (1–6). Previous studies in the RISC (Relationship between Insulin Sensitivity and Cardiovascular disease risk) study identified novel associations between insulin sensitivity and small molecules including amino acids glycine, cysteine, isoleucine, and creatine and the organic acids  $\alpha$ -hydroxybutyrate ( $\alpha$ -HB) and  $\alpha$ -ketobutyrate ( $\alpha$ -KB). Glycine was the amino acid most strongly associated with increased insulin sensitivity (4)—a finding consistent with other studies (7–9).

While some metabolites may represent important biomarkers, the causal directions of their associations with diabetes-related traits are uncertain. It is important to understand the causal role, or otherwise, of these molecules in order to avoid an increasingly confusing picture of which biomarkers are causal and which are secondary to the diabetes disease process.

The identification of genetic variants strongly associated with metabolites may provide useful tools to help understand causal directions of correlated phenotypes. Genetic variants are unlikely to be influenced by disease processes or environmental factors and therefore provide robust tools in Mendelian randomization to assess causal directions of correlated phenotypes (10). Recently, the principle of Mendelian randomization has been used to provide evidence for a causal association between reduced B-type natriuretic peptide levels and type 2 diabetes (11) and reduced sex hormone-binding globulin levels and type 2 diabetes (12), but the approach provided no evidence for a causal relationship between raised triglycerides and increased insulin resistance (13).

In this study, we focused on the associations of glycine and glutathione biosynthesis pathways with type 2 diabetes because, apart from the strong correlations identified in the RISC study, some other recent studies have provided evidence that high glycine level is associated with increased insulin sensitivity and decreased type 2 diabetes risks (7–9). In addition, type 2 diabetic patients have unrestrained gluconeogenesis and severely deficient glutathione synthesis (14,15). Glycine supplementation can improve deficient glutathione synthesis in type 2 diabetic patients, and glutathione supplementation can improve insulin sensitivity in nondiabetic individuals (15,16). We hypothesized that glycine and glutathione pathways contribute to diabetes and insulin resistance. We aimed to

identify genetic variants influencing circulating levels of metabolites in the glycine and glutathione pathways. We tested these variants in Mendelian randomization analyses to examine the potential causal role of these metabolites in insulin resistance and type 2 diabetes.

**RESEARCH DESIGN AND METHODS**

We analyzed nondiabetic participants of European ancestry from four studies who provided DNA for genome-wide genotyping and underwent a direct measure of insulin sensitivity. The four studies include RISC (*n* = 957), Botnia (*n* = 341), EUGENE2 consortium (European network on Functional Genomics of type 2 diabetes; *n* = 577) and the Stanford Insulin Suppression Test (IST) Cohort (*n* = 263). The descriptive characteristics of the RISC participants are shown in Table 1. In brief, we excluded individuals with cryptic relatedness using PLINK pairwise identity by descent estimation (PI\_HAT >0.2). We excluded individuals with lipid disorders or diabetes, lipid medications, pregnancy, fasting plasma glucose ≥7.0 mmol/L, or 2-h plasma glucose (on a 75-g oral glucose tolerance test) ≥11.0 mmol/L. The individual study characteristics, genotyping, and phenotyping details are provided in Supplementary Data. We performed genome-wide association studies (GWAS) for metabolites in the RISC study, replicated the GWAS findings in the Botnia study, and carried out Mendelian randomization analyses in RISC, EUGENE2, and Stanford IST to test the associations of genetic variants with insulin sensitivity. **Selection and measurement of metabolites in RISC and Botnia studies.** We selected 14 metabolites for GWAS. The metabolites were selected based on the study of Gall et al. (2010) (4). We selected metabolites that were both available in the RISC study and associated with insulin sensitivity (3,5). Details are shown in Table 2.

We selected metabolite ratios from the 14 metabolites based on two criteria: 1) the two metabolites were linked by one-step enzymatic reactions, and 2) the ratio was associated with insulin sensitivity measured by hyperinsulinemic-euglycemic clamp (*M* value). The glycine-to-serine ratio was the only one that satisfied both criteria (Fig. 2D). In both RISC and Botnia studies, metabolites were measured using multiple-platform mass spectrometry technology (ultra-high performance liquid chromatography and gas chromatography) as previously described (17–19). Absolute quantitation was performed for the 14 metabolites (Table 2) for the RISC study samples by UHPLC-MS/MS analysis (4,49).

**GWAS of metabolites and metabolite ratios in RISC.** The plasma concentrations of metabolites were fitted in a linear regression model with adjustment for age, sex, and centers. Then, the standardized residuals were normalized by inverse-normal transformation prior to GWAS. We performed GWAS with each metabolite using MACH2QTL based on an additive genetic model (20,21).

For the glycine-to-serine ratio, we log<sub>10</sub> transformed the ratio and then adjusted for age, sex, and center in linear regression analyses. We performed GWAS using MACH2QTL as with single metabolites' concentrations.

**Candidate-region association study of metabolites in RISC.** Some of the key enzymes and transporters involved in the metabolism and transport of metabolites are known. We selected 34 genes for the fourteen metabolites, consisting of carrier-encoding or enzyme-encoding genes involved in the rate-limiting steps of the relevant biosynthetic pathways. The genes selected are listed in Table 2. We classified single nucleotide polymorphisms (SNPs) within 300 kb of these genes as candidate SNPs. To prioritize SNPs for follow-up, we corrected for multiple testing of the total number of SNPs in each candidate region (a conservative threshold, given the correlation between SNPs).

However, we still used *P* value <3 × 10<sup>-9</sup> (5 × 10<sup>-8</sup> corrected for 15 tests [14 metabolites and one ratio]) in all available studies as the final criteria for association.

**Selection of SNPs for genotyping in Botnia and meta-analysis in RISC Single metabolites.** We used two statistical thresholds to select SNPs for replication. First, we used *P* value <5 × 10<sup>-8</sup> as the standard for genome-wide significance in the context of common SNPs. Second, for the 34 candidate genes (Table 2 and RESEARCH DESIGN AND METHODS), we divided 0.05 by the total number of SNPs in the gene ±300 kb. We meta-analyzed SNP-metabolite results from RISC and Botnia using an inverse variance-weighted approach as implemented in STATA command “metan.” For Mendelian randomization analyses, we used SNPs reaching *P* value <3 × 10<sup>-9</sup> in the meta-analysis of the two studies.

**Metabolite ratio.** To validate SNPs associated with the glycine-to-serine ratio, we linked results from the recently published Cooperative Health Research in the Region of Augsburg (KORA) and UKtwins studies (22) to our GWAS results. We meta-analyzed our results with those from the KORA or UKtwins studies with a significant threshold of *P* value <3 × 10<sup>-9</sup> when including all available studies.

**Effects of associated SNPs on other metabolites in the glycine and glutathione biosynthesis pathways.** We performed further analyses for the five SNPs associated with metabolites in the glycine or glutathione biosynthesis pathways, which include glycine, serine, betaine, α-HB, α-KB, and glycine-to-serine ratio. We tested the associations of each SNP against the other metabolite traits. We performed association analyses in the linear regression model described in method section 3 in STATA (version 10.1).

**Mendelian randomization analyses**

**Association of metabolite-associated SNPs in glycine biosynthesis pathway with insulin sensitivity.** We tested the role of metabolite-associated SNPs reaching genome-wide significance with two diabetes-related traits: hyperinsulinemic-euglycemic clamp (*M* value corrected for kilograms body weight), which was a measure of whole-body insulin sensitivity, and fasting insulin. *M* value-based measures of insulin sensitivity were corrected for age, sex, and center and converted to SD units; and inverse normalized. Fasting insulin was natural log transformed; corrected for age, sex, and center; and converted to SD units. Using RISC data, we calculated two estimates of the association between metabolite SNPs and diabetes-related traits for each metabolite trait. First, we calculated an estimated expected effect if there was a causal association between metabolites and diabetes-related measures, using a triangulation approach as shown in Supplementary Fig. 1: we calculated the correlation between standardized metabolite levels and the two diabetes-related traits. We then multiplied these standardized effects by that between metabolite SNPs and metabolites to estimate an approximate expected effect size of the association between metabolite SNPs and *M* value and fasting insulin. We calculated approximate expected 95% CIs based on the observed effects and SEs using the Taylor series expansion of the ratio of two means (23). Second, we tested the observed effect between metabolite SNPs and the two diabetes-related traits. For the clamp-based measures of insulin sensitivity, we used three studies (RISC, EUGENE2, and Stanford IST) and meta-analyzed results using the program METAL (24). In EUGENE2, insulin sensitivity was measured using the same hyperinsulinemic-euglycemic clamp-based protocol as that used by RISC (25). In the Stanford study, insulin sensitivity was measured by steady-state plasma glucose method. The steady-state plasma glucose value is highly inversely correlated to *M* value (*r* = -0.93, *P* < 0.001) (26), so meta-analyses were performed between the three studies by reversing the signs of the effect sizes in Stanford.

TABLE 1  
Summary details of RISC individuals and relevant characteristics

Units	Age (years)	BMI (kg/m <sup>2</sup> )	FI (pmol/L)	<i>M</i> value (μmol/kg body wt/min)*
<i>N</i>	1,004	1,004	973	1,004
Mean	43.91	25.42	34.3	39.84
SD of mean	8.37	4.04	18.55	16.2
Median	44	24.9	30	38.33
Minimum	30	16.9	3	4.92
Maximum	61	43.9	116	114.25
Correlation with age	1			
Correlation with BMI	<i>r</i> = 0.30; <i>P</i> = 8.3 × 10 <sup>-6</sup>	1		
Correlation with FI	<i>r</i> = 0.01; <i>P</i> = 0.31	<i>r</i> = 0.11; <i>P</i> = 1.2 × 10 <sup>-68</sup>	1	
Correlation with <i>M</i> value	<i>r</i> = -0.04; <i>P</i> = 0.01	<i>r</i> = -0.12; <i>P</i> = 7.5 × 10 <sup>-58</sup>	<i>r</i> = -0.49; <i>P</i> = 6.4 × 10 <sup>-44</sup>	1

FI, fasting insulin. \**M* value for the clamp expressed per kilogram body weight.

TABLE 2  
Fourteen metabolites studied in GWAS and candidate genes in RISC ( $n = 1,004$ )

Metabolites ( $\mu\text{g/mL}$ )	Candidate genes	Mean (minimum–maximum)	SD	Median	Correlation with <i>M</i> value	
					<i>r</i>	<i>P</i>
$\alpha$ -HB	<i>LDHA, LDHB, LDHC, LDHD, <math>\alpha</math>-HBDH</i>	4.50 (1.09–13.35)	1.75	4.23	–0.35	$1.5 \times 10^{-30}$
Adrenate	—	0.20 (0.06–1.19)	0.09	0.18	–0.19	$9.0 \times 10^{-10}$
$\alpha$ -Ketoglutaric acid	—	1.09 (0.00–2.92)	0.39	1.09	–0.21	$1.4 \times 10^{-11}$
$\alpha$ -KB	<i>CBS</i>	0.38 (0.00–1.16)	0.21	0.36	–0.28	$4.4 \times 10^{-20}$
Betaine	<i>BHMT, CHDH</i>	4.26 (1.16–11.92)	1.38	4.13	0.06	$6.6 \times 10^{-11}$
Creatine	<i>CKMT1A/1B, CKMT2</i>	4.55 (1.05–14.3)	2.19	4.13	–0.18	$1.4 \times 10^{-8}$
Decanoylcarnitine	<i>CPT1C, SLC25A20</i>	0.04 (0–0.35)	0.04	0.03	0.14	$1.3 \times 10^{-5}$
Glutamate	<i>NAGS, SIRT4, GLUD1</i>	18.19 (4.99–100.26)	12.21	14.38	0.06	0.06
Glycine	<i>SHMT1/2, GLDC, GCSH</i>	17.35 (7.05–41.53)	5.26	16.12	0.24	$7.2 \times 10^{-15}$
Ketovaline	<i>BCAT2, BCKDHA, BCKDHB</i>	1.59 (0.12–2.99)	0.40	1.62	–0.23	$7.1 \times 10^{-14}$
Linoleoyl-GPC	<i>PLA2G5, PLA2G12A, PLA2G2D</i>	15.65 (5.24–39.89)	5.18	15.15	0.29	$2.4 \times 10^{-21}$
Oleate	<i>OLAH, ACSL1</i>	85.14 (11.91–569.89)	36.67	81.57	–0.17	$3.2 \times 10^{-8}$
Oleoyl-GPC	<i>OLAH, LCLAT, PLD1</i>	9.81 (3.14–22.64)	2.93	9.49	0.27	$1.2 \times 10^{-18}$
Serine	<i>PSPH, PHGDH, CBS, SDS, SHMT2</i>	10.90 (4.61–21.75)	2.23	10.70	0.14	$4.6 \times 10^{-6}$

For fasting insulin measures of insulin sensitivity, we used data from the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC), consisting of a meta-analysis of 23 GWAS with 27,589 individuals (27).

**Association of metabolite-associated SNPs with type 2 diabetes.** We used data from the Diabetes Genetics Replication and Meta-analysis Consortium (DIAGRAM) to assess the association of metabolite SNPs with type 2 diabetes. These data come from 8,130 patients with type 2 diabetes and 38,987 control subjects from eight GWAS (28).

**Instrumental variable analysis.** In this RISC study, where we had measures of SNPs, metabolites, and insulin sensitivity, we performed instrumental variable analyses using the two-stage least squares regression approach implemented in the STATA command “ivreg2.” Using the two-stage least squares regression approach, the instrumental variable estimator  $\beta_{IV}$  provides an estimate of the causal effects of exposure (i.e., metabolites) on outcome (i.e., insulin sensitivity) even in the presence of unmeasured confounders (10,29).

## RESULTS

**GWAS and replication of insulin sensitivity-related metabolites.** In the RISC study, we identified eight signals of interest either at genome-wide significance or reaching a locus-wide nominal level of significance around one of the candidate genes (Supplementary Table 2). One of these signals represented a widely reported association between SNPs in the *FADS* gene cluster and fatty acids (27,30,31) (in our case adrenate), and we did not pursue this association further. We successfully genotyped SNPs representing six of the remaining seven signals in the Botnia study (Table 3).

After meta-analysis of RISC and Botnia data (where available), we identified four association signals with three separate single metabolites—the amino acids glycine, serine, and betaine—at  $P$  value  $<3 \times 10^{-9}$ . For ratios of metabolites, we identified one signal for glycine-to-serine ratio that when meta-analyzed with published KORA data reached  $P$  value  $<3 \times 10^{-9}$ . Details of the associations are given in Table 3 and Fig. 1. Two SNPs associated with serine were taken forward from the RISC GWAS but did not replicate in the Botnia study (Table 3).

**Novel associations between SNPs in two loci and betaine levels.** We identified an association between rs499368 in the *SLC6A12* gene and betaine levels ( $P$  value  $1.46 \times 10^{-10}$ ) (Fig. 1A). This signal has not previously been reported with any other trait and was not captured at  $r^2 > 0.8$  in the published KORA or UKtwins data. The second association occurred between rs17823642 near a candidate

gene, *BHMT*, and betaine levels ( $P$  value  $2.3 \times 10^{-9}$ ) (Fig. 1B). This signal has not previously been reported with any other trait at genome-wide significance but was captured at  $r^2 = 1.0$  by rs7732845 in the published KORA data (but not UKtwins), and a meta-analysis of RISC, Botnia, and KORA data ( $P$  value for KORA alone  $1.98 \times 10^{-6}$ ) confirms very robust evidence of association (meta-analyzed  $P$  value  $6.07 \times 10^{-14}$ ).

**SNP in a known locus is associated with glycine levels.** We identified an association between rs715 in the 3' untranslated region of the *CPS1* gene and glycine levels at genome-wide significance ( $P$  value  $3.30 \times 10^{-50}$ ) (Fig. 1C). This signal was not captured at  $r^2 > 0.8$  in the published KORA or UKtwins data, but a SNP (rs4673558) with an  $r^2 = 0.21$  with rs715 is associated with glycine levels with  $P$  value  $4.3 \times 10^{-11}$  in the UKtwins data.

**SNP in a known locus is associated with serine levels.** We identified an association between rs478093 near the *PHGDH* gene and serine levels at genome-wide significance ( $P$  value  $1.52 \times 10^{-9}$ ) (SNP not available in Botnia study) (Fig. 1D). This signal was previously reported as associated with serine and ratios of metabolites involving serine in the published KORA and UKtwins data (based on rs477992 [ $r^2 = 0.93$ ], meta-analyzed  $P$  value  $1.94 \times 10^{-14}$ ) (22).

**SNP in a novel locus is associated with glycine-to-serine ratios.** We identified a previously unreported association between rs1107366 near the *ALDH1L1* gene and glycine-to-serine ratios ( $P$  value  $2.25 \times 10^{-6}$ ) (Fig. 1E). This signal reached genome-wide significance in combination with data from the KORA and UKtwins studies (meta-analyzed  $P$  value  $2.8 \times 10^{-12}$ ) (Table 3).

**Association between rs715 in CPS1 and glycine levels is highly sex specific.** The SNPs in the *CPS1* locus have been previously reported with sex-specific effect on glycine and homocysteine levels (rs7422339,  $r^2 = 0.92$  with rs715) (32,33). We observed a similar sex-specific association between this signal and glycine levels (Supplementary Fig. 2). The association was weak in males ( $\beta = -0.19$  [95% CI  $-0.31$  to  $-0.08$ ];  $P$  value =  $1.1 \times 10^{-3}$ ) but more than four times the effect size in females ( $\beta = -0.84$  [ $-0.98$  to  $-0.69$ ];  $P$  value  $4.5 \times 10^{-28}$ ). The  $Z$  test for the null hypothesis of no sex-specific effect was rejected at  $P$  value  $2.67 \times 10^{-13}$ . This result was consistent

TABLE 3  
Associations between SNPs and insulin sensitivity-associated metabolites in the RISC ( $n = 1,004$ ) and Botnia ( $n = 341$ ) studies

SNP	Trait	Genes	Effect allele/ other allele	RISC effect (95% CI)	RISC $P$	Botnia effect (95% CI)	Botnia $P$	Meta. effect (95% CI)	Meta. $P$	Meta. $N$
rs478093	Serine	<i>PHGDH</i>	G/A	0.30 (0.20–0.39)	$1.5 \times 10^{-9}$	NA	NA	NA	$1.9 \times 10^{-14}$	2,772
rs715	Glycine	<i>CPSI</i>	T/C	-0.58 (-0.68 to -0.48)	$5.3 \times 10^{-30}$	-0.70 (-0.54 to -0.86)	$2.0 \times 10^{-16}$	-0.61 (-0.53 to -0.69)	$3.3 \times 10^{-50}$	1,343
rs1107366	Glycine/serine*	<i>ALDH1L1</i> <i>KLF15</i>	G/A	0.02 (0.01–10.03)	$5.0 \times 10^{-6\dagger}$	0.016 (0.002–0.03)	0.024	0.019 (0.011–0.027)	$2.3 \times 10^{-6}$	1,343
rs17823642	Betaine	<i>BHMT</i> , <i>BHMT2</i>	C/T	0.35 (0.21–0.49)	$1.4 \times 10^{-6}$	0.52 (0.24–0.80)	$3.5 \times 10^{-4}$	0.38 (0.51–0.26)	$2.4 \times 10^{-9}$	1,345
rs13233754	Serine	<i>PSPH</i>	G/A	-0.61 (-0.88 to -0.33)	$2.0 \times 10^{-5}$	0.02 (-0.59 to 0.62)	0.96	-0.50 (-0.25 to -0.75)	$1.0 \times 10^{-4}$	1,343
rs4275190	Serine	<i>PSPH</i>	T/C	0.17 (0.08–0.27)	$4.0 \times 10^{-4}$	0.08 (-0.26 to 0.43)	0.63	0.16 (0.07–0.26)	$3.9 \times 10^{-4}$	1,344
rs499368	Betaine	<i>SLC6A12</i> , <i>SLC6A13</i>	A/T	-0.46 (-0.61 to -0.3)	$8.1 \times 10^{-9}$	-0.25 (-0.41 to -0.1)	$1.5 \times 10^{-3}$	-0.36 (-0.46 to -0.25)	$1.5 \times 10^{-10}$	1,342

freq., frequency; NA, genotype data not available in Botnia (this signal was previously reported as associated with serine and ratios of metabolites involving serine in the published KORA and UKtwins data, based on rs477992 [ $r^2 = 0.93$ ], meta-analyzed [Meta.]  $P = 1.94 \times 10^{-14}$ ). \*The glycine-to-serine ratio was normalized by log<sub>10</sub> transformation, and the residuals after adjusting for age, sex, and center were used as the trait in GWAS. The other traits were normalized by inverse normal transformation and standardized (i.e., in SD units). †This association was genome-wide significant when meta-analyzed with published KORA data ( $P = 3.71 \times 10^{-11}$ ;  $n = 2,687$ ) and therefore followed up in the Botnia study: the  $P$  value when combining RISC, KORA, and Botnia results is  $2.8 \times 10^{-12}$  ( $n = 3,026$ ).

with the female-specific association previously reported (32,33). There was no evidence that the association was different between pre- and postmenopausal women (premenopausal, effect  $-0.78$  [ $-0.95$  to  $-0.61$ ];  $P$  value  $1.57 \times 10^{-19}$ ; postmenopausal,  $-1.04$  [ $-1.35$  to  $-0.73$ ];  $P$  value  $4.57 \times 10^{-11}$ ). We did not observe any evidence of sex-specific effects for the other metabolite SNPs.

**Effects of metabolite-associated SNPs on other metabolites in the glycine and glutathione biosynthesis pathways.** The effects of the five confirmed signals on the other four single metabolites levels and glycine-to-serine ratio are shown in Table 4, with the last column showing the effects of metabolite–insulin sensitivity associations.

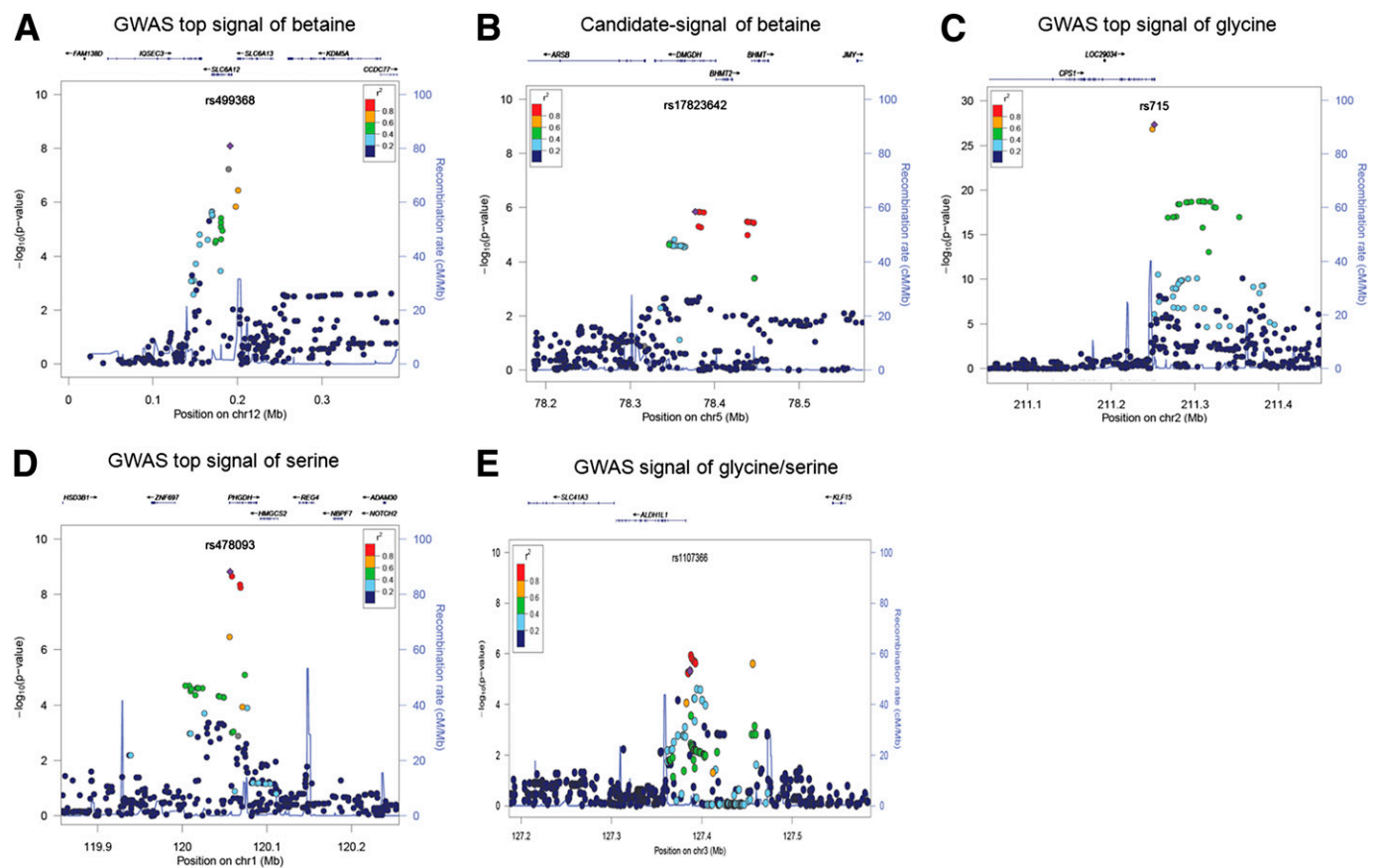
**Associations of metabolite-associated SNPs with fasting insulin-based and clamp-based measures of insulin sensitivity.** Associations between the five metabolite-associated SNPs, metabolite levels, and diabetes-related traits (fasting insulin and hyperinsulinemic-euglycemic clamp [ $M$  value]) are shown in Table 5. There were strong correlations between fasting insulin and metabolite levels ( $0.10 < r < 0.20$ ), as was expected, given that the metabolites were selected as those correlated with a measure of insulin sensitivity. These strong correlations between phenotypes meant that the MAGIC data had  $>90\%$  power to detect associations at  $P$  value  $< 0.01$  based on the estimated expected effects between metabolite SNPs and fasting insulin. However, there were no associations between metabolite SNPs and fasting insulin. The effect sizes observed in the MAGIC data were all smaller than those expected based on the triangulation calculations.

We identified a nominal association between the glycine-to-serine ratio-associated SNP rs1107366 near *ALDH1L1* and clamp-based measures of insulin sensitivity ( $\beta = 0.09$  SD [95% CI 0.03–0.15], where the allele that raises glycine-to-serine ratios increases clamp-based insulin sensitivity  $P$  value 0.005) (Table 5 and Supplementary Table 3). The observed effect size on insulin sensitivity (IS) was larger than the expected effect (expected  $\beta_{\text{SNP-IS}} = 0.03$  SD [95% CI 0.01–0.04]). Results of instrumental variable analyses in RISC were consistent with the main results: the glycine-to-serine ratio predicted from the rs1107366 genotypes is associated with clamp-based insulin sensitivity ( $\beta_{\text{IV}} = 1.00$  [0.24–1.76];  $P$  value 0.01). No other metabolite SNPs were associated with clamp-based measures of insulin sensitivity in either the triangulation analyses or the instrumental variable analyses, including the other glycine and serine signals (rs715 and rs478093).

**Association of metabolite SNPs with type 2 diabetes.** There was no evidence of association between four of the five metabolite SNPs and type 2 diabetes, based on the meta-analysis of case-control studies reported by the DIAGRAM study (Supplementary Table 4). The rs715 SNP in *CPSI* associated with glycine levels was poorly captured in the type 2 diabetes GWAS meta-analysis (28).

**DISCUSSION**

Using a genome-wide approach, we have identified five associations between genetic variants and circulating levels of three metabolites and one metabolite ratio. These metabolites occur in pathways strongly correlated with the gold standard measure of insulin sensitivity (hyperinsulinemic-euglycemic clamp) in the RISC study—primarily, the glycine biosynthesis pathway. Three of these associations



**FIG. 1.** Regional association plots of the five SNP-metabolite associations in the RISC cohort. In each plot, the *top panel* shows the name and location of genes in the UCSC Genome Browser. The  $-\log_{10}$  of  $P$  values of the imputed SNPs are plotted on the  $y$ -axis against genomic position (NCBI Build 36) on the  $x$ -axis. The top signal is represented by a diamond. Estimated recombination rates (taken from HapMap) are plotted to reflect the local linkage disequilibrium structure around the associated SNPs and their correlated proxies (according to a gray scale from  $r^2 = 0$  to 1, based on pairwise  $r^2$  values from HapMap Phase II CEU). chr, chromosome. (A high-quality color representation of this figure is available in the online issue.)

have been previously identified at genome-wide levels of significance. We have tested the association of these variants with insulin sensitivity using the largest collective set of studies with these measures, including RISC, EUGENE2, and Stanford.

**Amino acid-associated SNPs and diabetes-related traits.** Glycine and serine are glycogenic amino acids involved in hepatic gluconeogenesis and glutathione biosynthesis, which are potentially important pathways in diabetes and insulin resistance. In the RISC study, glycine, serine, and betaine were positively correlated with clamp-based measures of insulin sensitivity and negatively correlated with fasting insulin. These associations are in line with previous findings (7–9,34).

Using Mendelian randomization analyses, we assessed whether these amino acids play a causal role in insulin sensitivity and type 2 diabetes risks. Our Mendelian randomization analyses do not support a causal association between genetically changed glycine, serine, and betaine levels and insulin sensitivity levels as measured by fasting insulin or type 2 diabetes. However, for the clamp-based measures of insulin sensitivity we observed a suggestive association of a glycine-to-serine ratio-associated SNP, rs1107366 (near the *ALDH1L1* gene). The allele that raises glycine-to-serine ratios increases clamp-based insulin sensitivity. The rs1107366–insulin sensitivity association needs further replication in a larger sample size, especially

given that other glycine or serine-associated SNPs (e.g., rs715 and rs478093) are not associated with insulin sensitivity. The rs715 variant in *CPS1*, for example, explains a greater proportion of the variance in glycine levels (~13% compared with ~2% for rs1107366). It is also possible that rs1107366 could influence insulin sensitivity via non-glycine-mediated (pleiotropic) effects.

In an insulin-resistant state, the increase of hepatic gluconeogenesis would result in greater consumption of glycogenic amino acids, which may be accentuated in individuals with genetically influenced lower levels of these molecules and is consistent with the hypothesis of reverse causality (35). However, causal mechanisms in both directions remain plausible because gluconeogenesis is controlled in many different ways.

**Biology of metabolite levels.** Our data highlight some candidate genes and protein products important in controlling circulating metabolite levels. At each locus, there is a clear candidate gene, although we cannot be certain which gene is affected by the associated SNP.

**Betaine and serine signals are in or near functionally relevant genes.** The *SLC6A12* gene is a highly plausible candidate for influencing betaine levels. Betaine is an osmolyte used by cells for protection against hyperosmotic environments (36), and *SLC6A12* encodes a highly conserved osmoregulator, which controls cellular volume by extrusion of betaine (37,38). Previous

TABLE 4  
Associations between SNPs and five metabolites in the glycine and glutathione biosynthesis pathway in the RISC study

Metabolites	Glycine: rs715		Glycine/serine: rs1107366		Serine: rs478093		Betaine: rs17823642		Betaine: rs499368		Insulin sensitivity (M)	
	Effect	P	Effect	P	Effect	P	Effect	P	Effect	P	Effect	P
Single metabolites												
Glycine	-0.58	$5.3 \times 10^{-30}$	-0.17	$7.5 \times 10^{-5}$	-0.10	0.05	0.07	0.35	-0.18	0.02	0.23	$8.4 \times 10^{-14}$
Serine	-0.24	$1.9 \times 10^{-6}$	0.03	0.57	-0.30	$1.5 \times 10^{-9}$	0.07	0.31	-0.14	0.09	0.18	$2.1 \times 10^{-8}$
Betaine	0.24	$1.3 \times 10^{-6}$	-0.08	0.07	0.09	0.06	0.35	$1.4 \times 10^{-6}$	-0.46	$8.1 \times 10^{-9}$	0.20	$7.0 \times 10^{-11}$
$\alpha$ -HB	-0.07	0.20	0.09	0.03	0.09	0.07	-0.15	0.04	0.19	0.02	-0.31	$3.3 \times 10^{-23}$
$\alpha$ -KB	0.00	0.97	0.05	0.22	0.06	0.23	0.00	0.96	0.08	0.34	-0.20	$9.6 \times 10^{-11}$
Metabolite ratios												
Glycine to serine	-0.05	$9.8 \times 10^{-20}$	-0.02	$3.45 \times 10^{-6}$	0.01	$4.7 \times 10^{-3}$	0.00	0.75	-0.01	0.12	0.2	$1.1 \times 10^{-4}$

The single metabolites and glycine-to-serine ratio were log<sub>10</sub> transformed, and all associations were fitted in a linear regression model with adjustment for age, sex, and centers. The effect sizes of SNPs on single metabolites are not comparable with those in Tables 3 and 5, where the metabolite traits are in SD units and under inverse-normal transformation.

studies have shown that hyperosmolarity could induce insulin resistance by impairing insulin receptor substrate (IRS)-1 tyrosine phosphorylation and degradation of IRS1 and IRS2 in adipocytes (39,40). This connection may partly explain the association of betaine with insulin sensitivity.

The *BHMT* gene encodes a cytosolic enzyme that catalyzes the conversion of betaine and homocysteine to dimethylglycine and methionine, respectively. The rs17823642 SNP is highly correlated with three SNPs associated with *BHMT* enzyme activity and protein level (rs41272270,  $r^2 = 1$ ; rs16876512,  $r^2 = 0.925$ ; and rs6875201,  $r^2 = 0.925$ ) (41).

The previously described serine GWAS signal rs478093 is in *PHGDH* (also known as *PGDH* [see 3-*PGDH* in Fig. 2A]), the gene product of which catalyzes the first and rate-limiting step in the phosphorylation pathway of serine biosynthesis.

**Effect of variation at the *CPS1* locus on glycine level.** The rs715 SNP represents the same association signal as that identified between the SNP rs2216405 near *CPS1* ( $r^2 = 0.47$  with rs715) and glycine levels (31), but the variance in glycine levels explained by rs715 (12.87%) in our study compared with the variance explained by the rs2216405 SNP (8.64%) suggests that rs715 is a better marker for the causal variant. A recent study reported an association between rs715 and glycine levels specific to females, consistent with our results (32).

The enzyme encoded by *CPS1* catalyzes synthesis of carbamoyl phosphate from ammonia and bicarbonate (Fig. 2B). Patients with defects in the function or expression of *CPS1* suffer from life-threatening hyperammonemia (42). It is possible that variants at this locus perturb the conversion of ammonia and bicarbonate to carbamoyl phosphate. We hypothesize that *CPS1* variants may cause excess ammonia, which may then lead to increased production of glycine and tetrahydrofolate in the glycine cleavage system (Figs. 2D and 3).

**ALDH1L1 as a candidate enzyme involved in glycine metabolism.** Glycine is a key component of the folate pathway (Fig. 3). The protein product of *ALDH1L1* catalyzes the conversion of 10-formyltetrahydrofolate, NADP, and water to tetrahydrofolate, NADPH, and carbon dioxide (Fig. 2C). Our association between a SNP near the *ALDH1L1* gene and glycine-to-serine ratio implicates this enzyme in glycine/serine conversion rate. This is in accordance with the knowledge that the glycine cleavage system, which accounts for ~41% of whole-body glycine flux, is tightly linked with tetrahydrofolate in folate metabolism (43) (Fig. 2D and 3).

**Sex-specific effect at *CPS1* locus on glycine and homocysteine levels.** We observed a sex-specific association of *CPS1* variants on glycine levels. This is consistent with the findings of Mittelstrass et al. (2011) (32). The variants at the *CPS1* locus also have female-specific effects on homocysteine levels (33), and the glycine-raising allele is associated with raised homocysteine (44).

**Links between glycine, serine, homocysteine, and betaine in folate and homocysteine metabolism.** Glycine, serine, and betaine are linked to homocysteine and folate metabolism (44,45) (Fig. 3). The *CPS1* variant rs715 is strongly correlated with rs7422339 ( $r^2 = 0.92$ ), which was previously reported to be associated with homocysteine and folate levels (33). House, O'Connor, and Guenter (44) demonstrated that the plasma concentrations of homocysteine, glycine, and serine were all elevated in folate-deficient rats. From the link between betaine and



**TABLE 5**  
**Summary of associations between SNPs, insulin sensitivity-associated metabolites, and diabetes-related traits (fasting insulin and insulin sensitivity)**

SNP	Effect allele/ other allele	Metabolite	SNP-net		Met-trait		Expected SNP trait effect, SD (95% CI)	Observed SNP trait effect, SD (95% CI) <sup>†</sup>	Observed SNP trait <i>P</i>	<i>N</i>
			effect, SD (95% CI)	<i>P</i> *	effect, SD (95% CI)	<i>P</i>				
<b>Fasting insulin</b>										
rs478093	G/A	Serine	0.31 (0.22-0.40)	$1.9 \times 10^{-10}$	-0.19 (-0.25 to -0.13)	$2.0 \times 10^{-9}$	-0.06 (-0.08 to -0.03)	$1.1 \times 10^{-3}$ (-0.007 to 0.009)	0.79	37,795
rs715	T/C	Glycine	-0.58 (-0.67 to -0.48)	$5.0 \times 10^{-32}$	-0.17 (-0.23 to -0.11)	$1.0 \times 10^{-7}$	0.10 (0.06-0.14)	0.01 (-0.005 to 0.02)	0.21	14,381
rs1107366	G/A	Glycine/ serine	0.21 (0.12-0.29) <sup>‡</sup>	$4.1 \times 10^{-6}$	-0.05 (-0.11 to 0.01)	0.12	NA	NA	NA	NA
rs17823642	C/T	Betaine	0.35 (0.21-0.49)	$1.3 \times 10^{-6}$	-0.13 (-0.2 to -0.07)	$2.4 \times 10^{-5}$	-0.05 (-0.08 to -0.02)	$1.9 \times 10^{-3}$ (-0.01 to 0.01)	0.76	38,227
rs499368	A/T	Betaine	-0.46 (-0.61 to -0.30)	$6.4 \times 10^{-9}$	-0.13 (-0.2 to -0.07)	$2.4 \times 10^{-5}$	0.06 (0.03-0.10)	$4.0 \times 10^{-4}$ (-0.009 to 0.009)	0.93	27,174
<b>Hyperinsulinemic-euglycemic clamp</b>										
rs478093	G/A	Serine	0.31 (0.22-0.40)	$1.9 \times 10^{-10}$	0.18 (0.11-0.24)	$2.1 \times 10^{-8}$	0.05 (0.03-0.08)	-0.02 (-0.09 to 0.05)	0.64	1,884
rs715	T/C	Glycine	-0.58 (-0.67 to -0.48)	$5.0 \times 10^{-32}$	0.23 (0.17-0.29)	$8.4 \times 10^{-14}$	-0.13 (-0.18 to -0.09)	0.02 (-0.05 to 0.09)	0.58	1,884
rs1107366	G/A	Glycine/ serine	0.21 (0.12-0.29) <sup>‡</sup>	$4.1 \times 10^{-6}$	0.13 (0.07-0.19)	$3.6 \times 10^{-5}$	0.03 (0.01-0.04)	0.09 (0.03-0.15)	$5.5 \times 10^{-3}$	1,884
rs17823642	C/T	Betaine	0.35 (0.21-0.49)	$1.3 \times 10^{-6}$	0.20 (0.14-0.27)	$7.0 \times 10^{-11}$	0.07 (0.04-0.11)	-0.05 (-0.16 to 0.06)	0.36	1,884
rs499368	A/T	Betaine	-0.46 (-0.61 to -0.30)	$6.4 \times 10^{-9}$	0.20 (0.14-0.27)	$7.0 \times 10^{-11}$	-0.09 (-0.14 to -0.05)	0.03 (0.05-0.12)	0.45	1,884

SNP-metabolite, metabolite-trait, and expected SNP trait effect sizes are in SD units and based on the RISC study. Observed SNP trait values are in SD units and are taken from a meta-analysis of the RISC, EUGENE2, and Stanford studies. Expected SNP trait effect, the approximate expected effect sizes of SNPs on fasting insulin and clamp-based measures of insulin sensitivity, based on the product of SNP-net effect and met-trait effect; Met-trait, the associations between metabolites and diabetes-related continuous traits; NA, the expected and observed correlations between rs1107366 and fasting insulin are not shown because glycine-to-serine ratios are not associated with fasting insulin ( $P = 0.12$ ); Observed SNP trait effect, the observed effects of SNPs on fasting insulin and clamp-based measures of insulin sensitivity; SNP-net, the associations between the five metabolite-associated SNPs and metabolites. \**P* values for SNP-metabolite associations differ very slightly from those in Table 3 because we used a different software package (MACH2QTLD). †Observed SNP fasting insulin values (pmol/L) are normalized by natural log transformation in the MAGIC group and so are not directly comparable with the expected trait values; an approximate conversion factor from these units to SDs is  $1/0.5 = 2.0$  for natural log insulin (given that the median SD for natural log insulin across the MAGIC studies is 0.5). ‡The effect direction is different from that in Table 4 because the effects were reported on allele A in Table 4, and in Table 5 the glycine-to-serine ratio was inverse-normal transformed for consistency with the other two associations in the triangulation analysis.

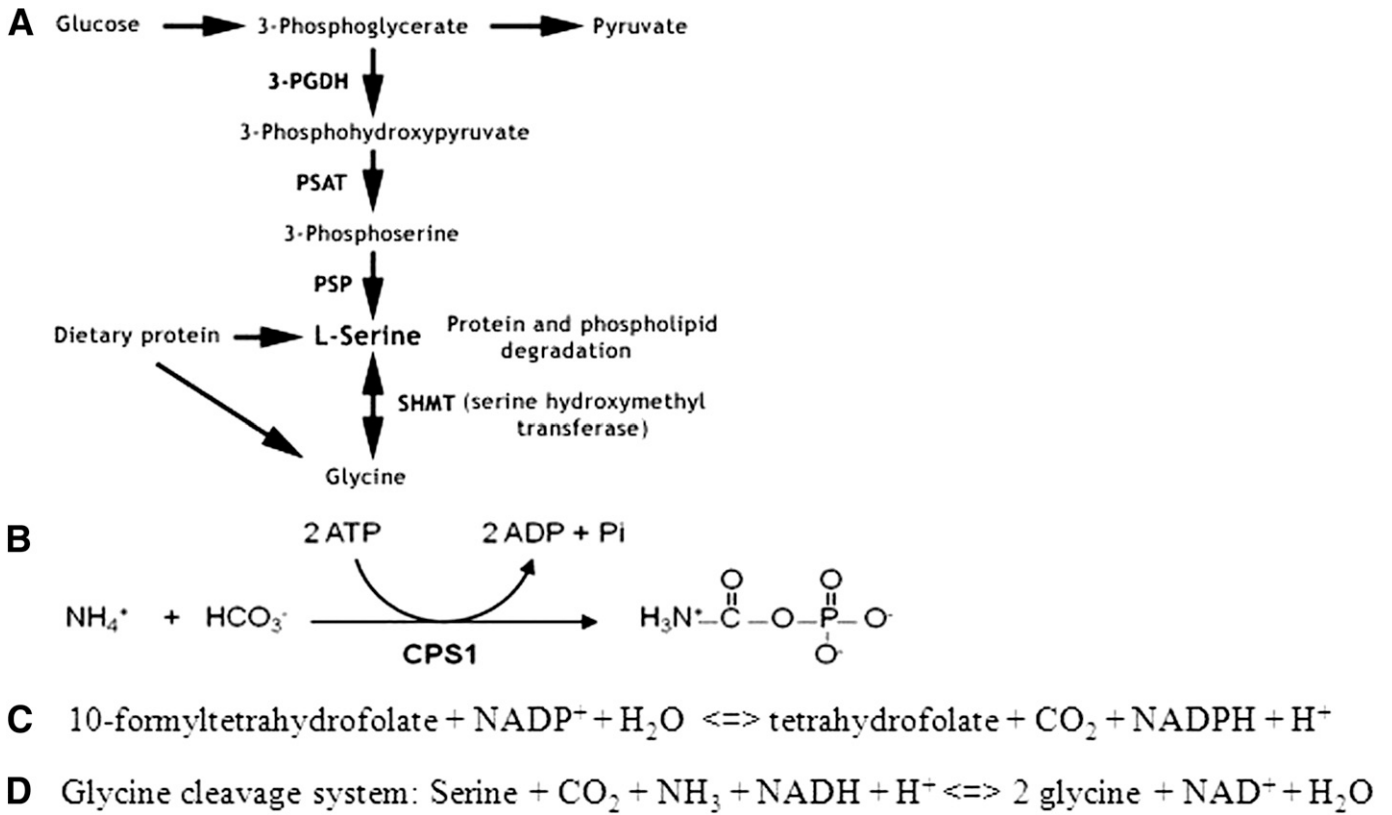


FIG. 2. Schematics of metabolic pathways relevant to SNP-metabolite associations. PSAT, phosphoserine aminotransferase.

homocysteine in the reaction catalyzed by BHMT, we hypothesized that if rs17823642 is associated with betaine level through reduced functioning of BHMT, then it would result in elevation of not only betaine but also homocysteine levels (46). We assessed the effect of rs17823642 on homocysteine levels in an independent European study (Invecchiare in Chianti, aging in the Chianti area [InCHIANTI]). We observed a nominal association in females ( $n = 575$ ;  $\beta = 0.26$

[95% CI 0.07–0.44];  $P$  value  $5.9 \times 10^{-3}$ ) but not in males ( $n = 458$ ;  $\beta = -0.02$  [–0.25 to 0.21];  $P$  value 0.87), where the betaine-raising allele also correlated with increased homocysteine levels. However, this association requires further confirmation with a larger sample size.

**Limitations.** There are a number of limitations in our study. First, the triangulation approach for estimating expected effects does not take into account the complicated

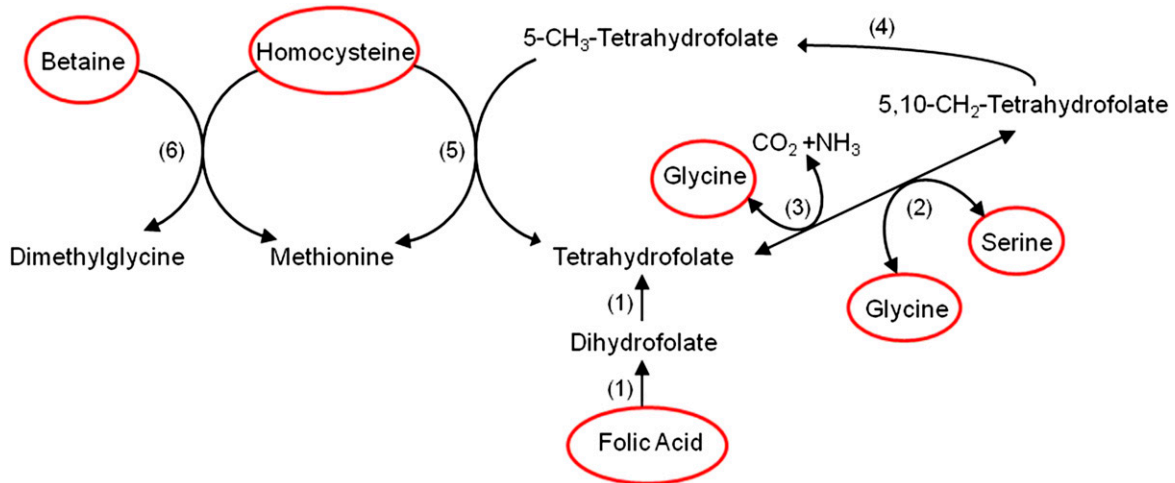


FIG. 3. Links between glycine, serine, folate, homocysteine, and betaine in folate metabolism and homocysteine metabolism. Enzymes (1): dihydrofolate reductase (2), serine hydroxymethyltransferase (3), glycine synthase (also called glycine cleavage enzyme) (4), methylenetetrahydrofolate reductase (5), methionine synthase (the other name of 5-methyltetrahydrofolate-homocysteine methyltransferase) (6), and betaine-homocysteine methyltransferase. Modified from House et al. (44) and Van Tellingen et al. (45). (A high-quality color representation of this figure is available in the online issue.)



feedback mechanisms and interactions involved in controlling metabolite levels. The SNPs that we identified are associated with several metabolites (Table 4), which means that they do not provide specific instruments for one metabolite. However, the associations for any one SNP are all with metabolites closely connected in well-annotated pathways and therefore provide an instrument to test the relationship between alterations in those pathways and diabetes-related outcomes.

Second, our estimated effects are approximate, with metabolite levels only measured in RISC and observed estimates coming from separate studies. Nevertheless, the size of the MAGIC study provided very good power to see the expected very small effects on fasting insulin levels.

Finally, we have only been able to assess some of the many metabolites associated with insulin sensitivity. The metabolites selected in our study were those most strongly associated with clamp-measured insulin sensitivity in the RISC study (4), and we focused efforts on these metabolite traits. Some recent studies have reported other metabolites associated with dysglycemia (e.g., branched chain and aromatic amino acids) (3,5,47,48), but these metabolites have not been measured in RISC study.

In conclusion, our study provides novel insight into the genetic regulation of metabolite levels, particularly those involved in the glycine-related pathways closely correlated to insulin sensitivity. Genetic variants associated with metabolite levels provide an important approach to helping unravel the functional role of the metabolic pathways that influence diabetes-related traits.

#### ACKNOWLEDGMENTS

This work was supported by grants from the Wellcome Trust (083270/Z/07/Z) and Medical Research Council (G0601261). The RISC study is partly supported by European Union Grant QL61-CT-2001-01252. The EUGENE2 study is supported by a grant from the European Community's FP6 EUGENE2 (LSHM-CT-2004-512013). J.W.K. is supported by American Heart Association National Fellowship to Faculty Transition Award 10FTF3360005. The RISC study is also supported by AstraZeneca.

W.E.G. is an employee of Metabolon, Inc. (Research Triangle Park, NC), as a Director in the Diagnostics division. Metabolon is a diagnostics and services company offering the industry's leading biochemical profiling platform. No other potential conflicts of interest relevant to this article were reported.

W.X. designed the study, researched data, performed analyses, and cowrote the initial draft of the article. A.R.W. analyzed data. V.L. contributed data and reviewed and edited the manuscript. M.N.W. analyzed data. J.W.K. contributed data, contributed to discussion, and reviewed and edited the manuscript. S.A., T.L.A., T.Q., F.A., and J.P. contributed data and reviewed the manuscript. H.H. and T.H. provided samples and data from individual studies, contributed to discussion, and reviewed the manuscript. O.P. and U.S. provided samples and data from individual studies and reviewed the manuscript. M.L. contributed data and reviewed and edited the manuscript. J.M.D. provided samples and data from individual studies and reviewed the manuscript. J.J.N. performed phenotyping in individual studies and reviewed the manuscript. L.G. contributed to discussion and reviewed and edited the manuscript. E.F. and K.-P.A. provided samples and data from individual studies and reviewed the manuscript. W.E.G.

performed phenotyping in individual studies, contributed to discussion, and reviewed the manuscript. T.M.F. designed the study, researched data, performed analyses, and cowrote the initial draft of the article. M.W. contributed data and reviewed and edited the manuscript. T.M.F. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

The authors thank Merck Research Laboratories for conducting DNA genotyping on the RISC samples and Stanford samples.

#### REFERENCES

- Newsholme P, Bender K, Kiely A, Brennan L. Amino acid metabolism, insulin secretion and diabetes. *Biochem Soc Trans* 2007;35:1180–1186
- Wang TJ, Larson MG, Vasani RS, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med* 2011;17:448–453
- Newgard CB, An J, Bain JR, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab* 2009;9:311–326
- Gall WE, Beebe K, Lawton KA, et al.; RISC Study Group. alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS ONE* 2010;5:e10883
- Huffman KM, Shah SH, Stevens RD, et al. Relationships between circulating metabolic intermediates and insulin action in overweight to obese, inactive men and women. *Diabetes Care* 2009;32:1678–1683
- Fiehn O, Garvey WT, Newman JW, Lok KH, Hoppel CL, Adams SH. Plasma metabolomic profiles reflective of glucose homeostasis in non-diabetic and type 2 diabetic obese African-American women. *PLoS ONE* 2010;5:e15234
- Floegel A, Stefan N, Yu Z, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* 2013;62:639–648
- Wang-Sattler R, Yu Z, Herder C, et al. Novel biomarkers for pre-diabetes identified by metabolomics. *Mol Syst Biol* 2012;8:615
- Michaliszyn SF, Sjaarda LA, Mihalik SJ, et al. Metabolomic profiling of amino acids and  $\beta$ -cell function relative to insulin sensitivity in youth. *J Clin Endocrinol Metab* 2012;97:E2119–E2124
- Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;27:1133–1163
- Pfister R, Sharp S, Luben R, et al. Mendelian randomization study of B-type natriuretic peptide and type 2 diabetes: evidence of causal association from population studies. *PLoS Med* 2011;8:e1001112
- Perry JR, Weedon MN, Langenberg C, et al.; MAGIC. Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. *Hum Mol Genet* 2010;19:535–544
- De Silva NM, Freathy RM, Palmer TM, et al. Mendelian randomization studies do not support a role for raised circulating triglyceride levels influencing type 2 diabetes, glucose levels, or insulin resistance. *Diabetes* 2011;60:1008–1018
- Altmajer E, Ramsay SL, Graber A, Mewes HW, Weinberger KM, Suhre K. Bioinformatics analysis of targeted metabolomics—uncovering old and new tales of diabetic mice under medication. *Endocrinology* 2008;149:3478–3489
- Sekhar RV, McKay SV, Patel SG, et al. Glutathione synthesis is diminished in patients with uncontrolled diabetes and restored by dietary supplementation with cysteine and glycine. *Diabetes Care* 2011;34:162–167
- Paolisso G, Giugliano D, Pizza G, et al. Glutathione infusion potentiates glucose-induced insulin secretion in aged patients with impaired glucose tolerance. *Diabetes Care* 1992;15:1–7
- Boudonck KJ, Mitchell MW, Nemet L, et al. Discovery of metabolomics biomarkers for early detection of nephrotoxicity. *Toxicol Pathol* 2009;37:280–292
- Evans AM, DeHaven CD, Barrett T, Mitchell M, Milgram E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal Chem* 2009;81:6656–6667
- Lawton KA, Berger A, Mitchell M, et al. Analysis of the adult human plasma metabolome. *Pharmacogenomics* 2008;9:383–397
- Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet* 2009;10:387–406

21. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010;34:816–834
22. Suhre K, Shin SY, Petersen AK, et al.; CARDIoGRAM. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 2011;477:54–60
23. Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in non-genetic observational studies using “Mendelian triangulation” by Bautista et al. *Ann Epidemiol* 2007;17:511–513
24. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26:2190–2191
25. Boesgaard TW, Zilinskaite J, Vanttinen M, et al.; EUGENE 2 Consortium. The common SLC30A8 Arg325Trp variant is associated with reduced first-phase insulin release in 846 non-diabetic offspring of type 2 diabetes patients—the EUGENE2 study. *Diabetologia* 2008;51:816–820
26. Greenfield MS, Doberne L, Kraemer F, Tobey T, Reaven G. Assessment of insulin resistance with the insulin suppression test and the euglycemic clamp. *Diabetes* 1981;30:387–392
27. Dupuis J, Langenberg C, Prokopenko I, et al.; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; Anders Hamsten on behalf of Procardis Consortium; MAGIC investigators. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010;42:105–116
28. Voight BF, Scott LJ, Steinthorsdottir V, et al.; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 2010;42:579–589
29. Hausman JA. Specification Tests in Econometrics. *Econometrica* 1978;46:1251–1271
30. Tanaka T, Shen J, Abecasis GR, et al. Genome-wide association study of plasma polyunsaturated fatty acids in the InCHIANTI Study. *PLoS Genet* 2009;5:e1000338
31. Illig T, Gieger C, Zhai G, et al. A genome-wide perspective of genetic variation in human metabolism. *Nat Genet* 2010;42:137–141
32. Mittelstrass K, Ried JS, Yu Z, et al. Discovery of sexual dimorphisms in metabolic and genetic biomarkers. *PLoS Genet* 2011;7:e1002215
33. Lange LA, Croteau-Chonka DC, Marvelle AF, et al. Genome-wide association study of homocysteine levels in Filipinos provides evidence for CPS1 in women and a stronger MTHFR effect in young adults. *Hum Mol Genet* 2010;19:2050–2058
34. Perseghin G, Ghosh S, Gerow K, Shulman GI. Metabolic defects in lean nondiabetic offspring of NIDDM parents: a cross-sectional study. *Diabetes* 1997;46:1001–1009
35. Wahren J, Felig P, Cerasi E, Luft R. Splanchnic and peripheral glucose and amino acid metabolism in diabetes mellitus. *J Clin Invest* 1972;51:1870–1878
36. Craig SA. Betaine in human nutrition. *Am J Clin Nutr* 2004;80:539–549
37. Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN. Living with water stress: evolution of osmolyte systems. *Science* 1982;217:1214–1222
38. Deo RC, Hunter L, Lewis GD, et al. Interpreting metabolomic profiles using unbiased pathway models. *PLOS Comput Biol* 2010;6:e1000692
39. Bratusch-Marrain PR, DeFronzo RA. Impairment of insulin-mediated glucose metabolism by hyperosmolality in man. *Diabetes* 1983;32:1028–1034
40. Gual P, Gonzalez T, Grémeaux T, Barres R, Le Marchand-Brustel Y, Tanti JF. Hyperosmotic stress inhibits insulin receptor substrate-1 function by distinct mechanisms in 3T3-L1 adipocytes. *J Biol Chem* 2003;278:26550–26557
41. Feng Q, Kalari K, Fridley BL, et al. Betaine-homocysteine methyltransferase: human liver genotype-phenotype correlation. *Mol Genet Metab* 2011;102:126–133
42. Summar ML. Molecular genetic research into carbamoyl-phosphate synthase I: molecular defects and linkage markers. *J Inher Metab Dis* 1998;21 (Suppl. 1):30–39
43. Lamers Y, Williamson J, Gilbert LR, Stacpoole PW, Gregory JF 3rd. Glycine turnover and decarboxylation rate quantified in healthy men and women using primed, constant infusions of [1,2-(13)C2]glycine and [(2)H3]leucine. *J Nutr* 2007;137:2647–2652
44. House JD, O'Connor CP, Guenter W. Plasma homocysteine and glycine are sensitive indices of folate status in a rodent model of folate depletion and repletion. *J Agric Food Chem* 2003;51:4461–4467
45. Van Tellingen A, Grooteman MP, Bartels PC, et al. Long-term reduction of plasma homocysteine levels by super-flux dialyzers in hemodialysis patients. *Kidney Int* 2001;59:342–347
46. Heil SG, Lievers KJ, Boers GH, et al. Betaine-homocysteine methyltransferase (BHMT): genomic sequencing and relevance to hyperhomocysteinemia and vascular disease in humans. *Mol Genet Metab* 2000;71:511–519
47. Huffman KM, Slentz CA, Bateman LA, et al. Exercise-induced changes in metabolic intermediates, hormones, and inflammatory markers associated with improvements in insulin sensitivity. *Diabetes Care* 2011;34:174–176
48. Shaham O, Wei R, Wang TJ, et al. Metabolic profiling of the human response to a glucose challenge reveals distinct axes of insulin sensitivity. *Mol Syst Biol* 2008;4:214
49. Cobb J, Gall W, Adam KP, et al. A novel fasting blood test for insulin resistance and prediabetes. *J Diabetes Sci Technol* 2013;1:100–110