



## Short Communication

## The impact of common genetic variants in the mitochondrial glycine cleavage system on relevant metabolites



Jessica O'Reilly<sup>a</sup>, Faith Pangilinan<sup>b</sup>, Karsten Hokamp<sup>c</sup>, Per M. Ueland<sup>d</sup>, John T. Brosnan<sup>e</sup>, Margaret E. Brosnan<sup>e</sup>, Lawrence C. Brody<sup>b,\*</sup>, Anne M. Molloy<sup>a</sup>

<sup>a</sup> Department of Clinical Medicine, School of Medicine, Trinity College Dublin 2, Ireland

<sup>b</sup> Genetic and Environmental Interaction Section, National Human Genome Research Institute, Bethesda, MD 20892, United States

<sup>c</sup> School of Genetics and Microbiology, Trinity College, Dublin 2, Ireland

<sup>d</sup> Section of Pharmacology, Institute of Medicine, University of Bergen and Haukeland University Hospital, 5021 Bergen, Norway

<sup>e</sup> Department of Biochemistry, Memorial University of Newfoundland, St. John's, Newfoundland, Canada

## ARTICLE INFO

## Keywords:

Glycine cleavage system (GCS)

One-carbon metabolism

Glycine

Folate

Neural tube defects

## ABSTRACT

The glycine cleavage system (GCS) is a complex of four enzymes enabling glycine to serve as a source of one-carbon units to the cell. We asked whether concentrations of glycine, dimethylglycine, formate, and serine in blood are influenced by variation within GCS genes in a sample of young, healthy individuals. Fifty-two variants tagging ( $r^2 < 0.9$ ) the four GCS genes were tested; one variant, *GLDC* rs2297442-G, was significantly associated ( $p = .0007$ ) with decreased glycine concentrations in serum.

## 1. Introduction

The glycine cleavage system (GCS) is a crucial component of mitochondrial one-carbon metabolism (OCM). The GCS involves four distinct enzymes: glycine decarboxylase (*GLDC*), aminomethyltransferase (*AMT*), glycine cleavage system protein H (*GCSH*) and dihydrolipoamide dehydrogenase (*DLD*). These enzymes act sequentially in the mitochondrion to cleave glycine obtained from the cytoplasm or from the breakdown of serine in the mitochondrion to produce *N,N*-methylene tetrahydrofolate (CH<sub>2</sub>-THF), an active folate derivative that contributes to the transfer of one-carbon units in cellular reactions [1].

Loss-of-function mutations in two GCS genes (*GLDC* and *AMT*) [2] cause non-ketotic hyperglycinemia (NKH), a rare recessive disease. Additionally, two NKH patients have been observed to have a complex rearrangement of *GCSH* [3]. Affected patients suffer from neurological impairments, seizures, and developmental delay, suggesting that the GCS is important for normal brain development and function [4]. In a study population of patients from the UK, Sweden and Japan, two of the genes encoding GCS enzymes, *AMT* and *GLDC*, were found to have missense variants associated with neural tube defects (NTDs) [5]. Similarly, variants in these genes were identified in a study of American patients with myelomeningocele, a type of NTD [6]. Finally, *GLDC* rs14742391 (p.Ser951Tyr) has been observed in a patient with NKH [2] and a patient with anencephaly, a type of NTD [7].

Although genome-wide association studies (GWASs) can identify

gene variants that influence traits, they are often underpowered to detect signals [8] from real but small effects of common alleles, or large effects of rare alleles. GWASs have been carried out for serum glycine, identifying a single nucleotide polymorphism (SNP) in carbamoylphosphate synthase 1 (rs715) [9–12]. Blood serine has been associated with a SNP in phosphoglycerate dehydrogenase (rs477992) [9, 12, 13]. To test for other genetic modifiers of OCM that may have been missed by the GWAS method, we examined whether common variants in the genes of the GCS influence relevant metabolites in a healthy population.

## 2. Materials and methods

## 2.1. The Trinity Student Study (TSS) cohort

The TSS cohort comprises a population of 2232 healthy Irish students, as described [14–16]. Informed consent and ethical approval were obtained from all participants. Briefly, blood samples were collected into EDTA and clotting tubes, processed within 3 h, and stored at  $-80^{\circ}\text{C}$  before assaying. Glycine (interassay CV = 3.3%) and serine (interassay CV = 5.7%) in serum were measured with gas-chromatography tandem mass spectrometry (GC-MS/MS [17]) and plasma dimethylglycine (interassay CV < 10%) by liquid-chromatography tandem mass spectrometry (LC-MS/MS [18]) by the laboratory of Bevit, Norway ([www.bevital.no](http://www.bevital.no)). Formate was measured using a

\* Corresponding author at: 50 South Drive, Room 5306, Bethesda, MD 20892-8004, United States.

E-mail address: [lbrody@mail.nih.gov](mailto:lbrody@mail.nih.gov) (L.C. Brody).

newly developed GC–MS method [19]. A single control sample was frozen in aliquots for daily analysis with formate assays of participant serum samples. The day-to-day variation for this sample was 7.4% ( $n = 38$ ). Genome-wide SNP genotyping was conducted at the Center for Inherited Disease Research (CIDR, USA) using Illumina 1 M HumanOmni1-Quad\_v1-0\_B chips.

## 2.2. Metabolite association analyses

Conservative ( $r^2 < 0.9$ ) tagging variants in the four GCS genes including 10 kb flanking regions were selected to test for association with levels of serum glycine, plasma dimethylglycine, serum serine, and serum formate.  $\log_{10}$  transformations of metabolite concentrations were used to meet normality assumptions. Linkage disequilibrium (LD) analyses were performed with Haploview; a total of 52 tag SNPs covering 93 variants were selected for association testing [20, 21]. For each SNP, association was tested using linear regression with a 1-df test (additive genetic model) or a 2-df test (genotypic model) (R statistical language and environment (version 3.3.2) [22], R package `snpStats` [23]). Bonferroni correction for 52 tests was applied to the nominal significance threshold ( $p < .05$ ) to obtain a corrected significance threshold ( $p < .00096$ ). Associations between  $\log_{10}$  transformed metabolite data were assessed using Pearson's correlation coefficient ( $r$ ).

## 2.3. GLDC expression association analyses

The public Gene-Tissue Expression resource (<https://www.gtexportal.org/home>) was used to search for associations between SNPs and *GLDC* mRNA levels (mRNA expression quantitative trait locus, eQTLs).

## 3. Results

The 52 tag SNPs in the four glycine cleavage genes (*GLDC*, *AMT*, *GCSH*, and *DLSD*) were tested for association with blood levels of glycine, dimethylglycine, formate, and serine in 2232 young healthy Irish adults (Table 1). Serum glycine concentrations were significantly correlated with all the other metabolites measured. The strength of these associations varied from  $r = 0.52$ ;  $p < .001$  with serine to  $-0.07$ ;

**Table 1**  
Metabolite characteristics in the TSS and by genotype group of its most associated GCS variant.

	Serine ( $\mu\text{M}$ )	Glycine ( $\mu\text{M}$ )	Dimethylglycine ( $\mu\text{M}$ )	Formate ( $\mu\text{M}$ )
Mean +/-	147.2 +/-	293.7 +/-	4.17 +/- 1.22	25.9 +/- 7.8
SD	23.9	63.9		
Median	145.6	288.1	3.99	24.8
Number*	2210	2210	2227	1535
Top SNP	rs2297442	rs2297442	rs16924717	rs7031325
p-value**	0.0225	0.0007	0.0047	0.0133
Genotype (No.)	AA (1198)	AA (1198)	AA (2130)	TT (647)
Mean +/-	148.1 +/-	297.6 +/-	4.2 +/- 1.2	26.1 +/- 8.2
SD	24.5	65.8		
Genotype (No.)	AG (848)	AG (848)	AG (95)	TC (701)
Mean +/-	146.9 +/-	290.6 +/-	4.6 +/- 2	25.7 +/- 7.1
SD	23.7	63.2		
Genotype (No.)	GG (164)	GG (164)	GG (2)	CC (187)
Mean +/-	143.1 +/-	281.1 +/-	3.8 +/- 2.7	25.9 +/- 9
SD	23.4	53.7		

\* Number of participants with metabolite and genotype values.

\*\* The Bonferroni-corrected threshold for significance is  $p < .00096$ .

$p = .004$  with formate. Only one GCS SNP is significantly associated with a metabolite. *GLDC* rs2297442 returns a  $p$ -value of 0.0007 following 1-df association testing with  $\log_{10}$  transformed glycine concentration. The 2-df test of association is also significant ( $p = .0031$ ), nominally.

We then sought to replicate the association between rs2297442-G and decreased levels of circulating glycine by examining results from two recent GWASs in which glycine was measured. Our observation was replicated in two studies of Finnish individuals ( $n = 16,506$ ,  $\beta = -0.039$ ,  $p = .0014$  [24];  $n = 8545$ ,  $\beta = -0.039$ ,  $p = .028$  [25]).

We also asked if *GLDC* mRNA expression levels are influenced by rs2297442 genotype. A Gene-Tissue Expression (GTEx) query of rs2297442 did not reveal a strong or consistent effect on *GLDC* mRNA levels across multiple tissues. In contrast, several other *GLDC* SNPs strongly influenced *GLDC* mRNA levels in several tissues ( $p < 1 \times 10^{-40}$ ). For example, in thyroid tissue, *GLDC* rs35374927 is estimated to have a large impact on *GLDC* mRNA levels compared to *GLDC* rs2297442 (Fig. 1). These eQTLs of strong effect reside in LD blocks near the 5' end of the gene, as opposed to the SNPs in the LD block at the 3' end of the gene where rs2297442 is located.

## 4. Discussion

The results of these association tests confirm that variation in *GLDC* can influence serum glycine concentrations. The variant reported here, *GLDC* rs2297442, is a noncoding polymorphism found in the seventh of the ten introns in this gene. The GTEx data revealed that other, unlinked SNPs in the gene have a larger impact on *GLDC* mRNA levels, albeit in limited, specific tissues. We conclude that rs2297442 does not influence serum glycine levels via modulation of *GLDC* transcript levels. It is possible that rs2297442 is simply linked to the causal SNP. A search for a linked, exonic SNP in a population of European ancestry [26] failed to identify any strongly related variants ( $r^2 < 0.2$ ).

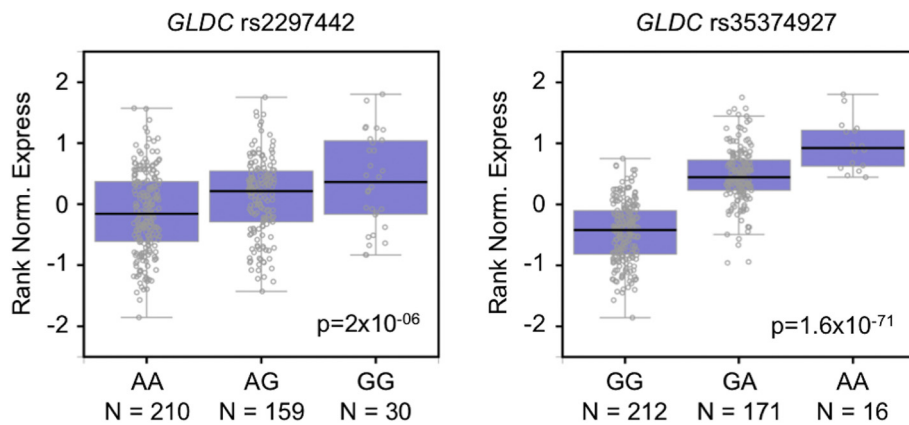
These results are consistent with the presence of a common SNP in *GLDC* that may increase the activity or alter the tissue distributions of this enzyme, causing reduced levels of circulating glycine. We predict that this effect on serum glycine concentrations may not be clinically relevant in this healthy population; however, it may contribute to diseases with a complex mode of inheritance, where combinations of many environmental and genetic factors of small individual effect must converge. The known link between folic acid supplementation and neural tube defect (NTD) prevention implicates perturbations of OCM as a key risk factor. In a study of 258 NTD patients, 27 single-base substitutions were discovered in *GLDC*, six of which influenced enzyme activity in an in vitro system [5]. The association between *GLDC* rs2297442 and serum glycine levels in the healthy populations in this study and others identifies a candidate locus for studies investigating the genetic basis of NKH, NTDs or any pathophysiology involving OCM.

## Acknowledgements

The authors acknowledge the contributions made by the study participants in the Trinity Student Study and the replication cohorts. The authors are grateful to Narisu Narisu and colleagues for providing summary statistics for the METSIM study. The authors acknowledge genotyping performed by the Center for Inherited Disease Research (CIDR) under contract HHSN268201200008I.

## Data statement

The data upon which these analyses are based have been deposited in and are available from the dbGaP database under dbGaP accession phs000789.v1.p1.



**Fig. 1.** Association of rs2297442 (left) or rs35374927 (right) with *GLDC* mRNA levels in human thyroid. In these graphs generated by GTEX, participants are grouped by genotype and plotted for ranked normalized expression of *GLDC* mRNA in thyroid. Unadjusted *p*-values for linear regression analyses are shown. *GLDC* rs2297442 effect size (i.e., slope of the linear regression) = 0.3. *GLDC* rs35374927 effect size = 1.0.

## Funding

This work was supported by funding from the Intramural Research Programs of the National Human Genome Research Institute.

## Conflict of interest statement

The authors have no conflicts of interest to declare.

## References

- [1] A.S. Tibbetts, D.R. Appling, Compartmentalization of mammalian folate-mediated one-carbon metabolism, *Annu. Rev. Nutr.* 30 (2010) 57–81.
- [2] C.R. Coughlin II et al., The genetic basis of classic nonketotic hyperglycinemia due to mutations in *GLDC* and *AMT*, *Genet. Med.* 19 (1) (2017) 104–111.
- [3] H. Koyata, K. Hiraga, The glycine cleavage system: structure of a cDNA encoding human H-protein, and partial characterization of its gene in patients with hyperglycinemias, *Am. J. Hum. Genet.* 48 (2) (1991) 351–361.
- [4] Y.J. Pai, et al., Glycine decarboxylase deficiency causes neural tube defects and features of non-ketotic hyperglycinemia in mice, *Nat. Commun.* 6 (2015) 6388.
- [5] A. Narisawa, et al., Mutations in genes encoding the glycine cleavage system predispose to neural tube defects in mice and humans, *Hum. Mol. Genet.* 21 (7) (2012) 1496–1503.
- [6] R.H. Shah, et al., Genetic association of the glycine cleavage system genes and myelomeningocele, *Birth Defects Res. A Clin. Mol. Teratol.* 106 (10) (2016) 847–853.
- [7] M. Ishida, et al., A targeted sequencing panel identifies rare damaging variants in multiple genes in the cranial neural tube defect, anencephaly, *Clin. Genet.* 93 (4) (2018) 870–879.
- [8] W.Y. Wang, et al., Genome-wide association studies: theoretical and practical concerns, *Nat. Rev. Genet.* 6 (2) (2005) 109–118.
- [9] H.H.M. Draisma, et al., Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels, *Nat. Commun.* 6 (2015) 7208.
- [10] J.A. Hartiala, et al., Genome-wide association study and targeted metabolomics identifies sex-specific association of *CPS1* with coronary artery disease, *Nat. Commun.* 7 (2016) 10558.
- [11] S.Y. Shin, et al., An atlas of genetic influences on human blood metabolites, *Nat. Genet.* 46 (6) (2014) 543–550.
- [12] W. Xie, et al., Genetic variants associated with glycine metabolism and their role in insulin sensitivity and type 2 diabetes, *Diabetes* 62 (6) (2013) 2141–2150.
- [13] K. Suhre, et al., Human metabolic individuality in biomedical and pharmaceutical research, *Nature* 477 (7362) (2011) 54–60.
- [14] A.M. Molloy, et al., A common polymorphism in *HIBCH* influences methylmalonic acid concentrations in blood independently of cobalamin, *Am. J. Hum. Genet.* 98 (5) (2016) 869–882.
- [15] N. Stone, et al., Bioinformatic and genetic association analysis of microRNA target sites in one-carbon metabolism genes, *PLoS One* 6 (7) (2011) e21851.
- [16] J.L. Mills, et al., Do high blood folate concentrations exacerbate metabolic abnormalities in people with low vitamin B-12 status? *Am. J. Clin. Nutr.* 94 (2) (2011) 495–500.
- [17] O. Midttun, et al., Combined measurement of 6 fat-soluble vitamins and 26 water-soluble functional vitamin markers and amino acids in 50 µL of serum or plasma by high-throughput mass spectrometry, *Anal. Chem.* 88 (21) (2016) 10427–10436.
- [18] O. Midttun, G. Kvalheim, P.M. Ueland, High-throughput, low-volume, multianalyte quantification of plasma metabolites related to one-carbon metabolism using HPLC-MS/MS, *Anal. Bioanal. Chem.* 405 (6) (2013) 2009–2017.
- [19] S.G. Lamarre, et al., An isotope-dilution, GC-MS assay for formate and its application to human and animal metabolism, *Amino Acids* 46 (8) (2014) 1885–1891.
- [20] J.C. Barrett, et al., Haploview: analysis and visualization of LD and haplotype maps, *Bioinformatics* 21 (2) (2005) 263–265.
- [21] P.I. de Bakker, et al., Efficiency and power in genetic association studies, *Nat. Genet.* 37 (11) (2005) 1217–1223.
- [22] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2016.
- [23] D. Clayton, *snpStats: SnpMatrix and XSNPMatrix Classes and Methods*, (2015).
- [24] T.M. Teslovich, et al., Identification of seven novel loci associated with amino acid levels using single variant and gene-based tests in 8,545 Finnish men from the METSIM study, *Hum. Mol. Genet.* 27 (9) (2018) 1664–1674.
- [25] J. Kettunen, et al., Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA, *Nat. Commun.* 7 (2016) 11122.
- [26] C. Genomes Project, et al., A global reference for human genetic variation, *Nature* 526 (7571) (2015) 68–74.