Clinically significant findings of high-risk mutations in human SLC29A4 gene associated with diabetes mellitus type 2 in Pakistani population

Sadaf Moeeza, Sumbul Khalida, Sania Shaeena, Madiha Khalidb, Asima Ziac, Asma Guia, Rauf Niaziid and Zoya Khalide

aDepartment of Bioinformatics & Biotechnology, International Islamic University (IIUI), Islamabad, Pakistan; bPMASt Arid Agriculture University, Rawalpindi, Pakistan; cDepartment of Biosciences, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia; dDepartment of Medicine, Pakistan Institute of Medical Sciences, Shaheed Zulfiqar Ali Bhutto Medical University, Islamabad, Pakistan; eComputational Biology research lab, Department of computer science, National University of Computer and Emerging Sciences, NUCES-FAST, Islamabad, Pakistan

Communicated by Ramaswamy H. Sarma

ABSTRACT
This study conducted an in-depth analysis combining computational and experimental verifications of the deleterious missense mutations associated with the SLC29A4 protein. The functional annotation of the non-synonymous single nucleotide polymorphism (nsSNPs), followed by structure-function analysis, revealed 13 single nucleotide polymorphisms (SNP) as the most damaging. Among these, six mutants P429T/S, L144S, M108V, N86H, and V79E, were predicted as structurally and functionally damaging by protein stability analysis. Also, these variants are located at evolutionary conserved regions, either buried, contributing to the structural damage, or exposed, causing functional changes in the protein. These mutants were further taken for molecular docking studies. When verified via experimental analysis, the SNPs M108V (rs149798710), N86H (rs151039853), and V79E (rs17854505) showed an association with type 2 diabetes mellitus (T2DM). Minor allele frequency for rs149798710 (A>G) was 0.23 in controls, 0.29 in metformin responders, 0.37 in metformin non-responder, for rs151039853 (A>C) was 0.21 in controls, 0.28 in metformin responders, 0.36 in metformin non-responder and for rs17854505 (T>A) was 0.20 in controls, 0.25 in metformin responders, 0.37 in metformin non-responder. Hence, this study concludes that SLC29A4 M108V (rs149798710), N86H (rs151039853), and V79E (rs17854505) polymorphisms were associated with the increased risk of T2DM as well as with the increased risk towards the failure of metformin therapeutic response in T2DM patients of Pakistan.

Abbreviations: T2DM: Type 2 Diabetes Mellitus; SNP: Single Nucleotide Polymorphism; ADA: American Diabetes Association; PMAT: Plasma membrane monoamine transporter; OCT1: Organic cation transporter 1; OCT2: Organic cation transporter 2; MATE1: Multidrug and toxin extrusion 1; MATE2K: Multidrug and toxin extrusion 2; HbA1c: Hemoglobin A1c; HWE: Hardy-Weinberg equilibrium

INTRODUCTION

Type 2 diabetes mellitus (T2DM) is characterized when pancreatic β-cells become defective in insulin secretion or become insulin resistant (Stumvoll et al., 2005). T2DM is a metabolic disorder that is a heterogeneous and multi-factorial disease. Both environmental and genetic factors contribute to the pathogenesis of T2DM (Brunetti et al., 2014). Pakistan is the 36th largest country, and globally, Pakistan stands at position 6th in the prevalence of T2DM. The frequency of T2DM in Pakistan in the year 2000 was assessed to be 5.2 million, which was anticipated to be raised to 13.9 million by the year 2030 (Wild et al., 2004). The mean prevalence of type 2 diabetes mellitus in Pakistan was 11.77%, wherein females’ prevalence was 9.19%, and in males was 11.20% (Shu et al., 2007).

Currently, there are ten classes of orally available pharmacological agents to treat T2DM: (1) sulfonylureas, (2) meglitinides, (3) metformin (a biguanide), (4) thiazolidinediones (TZDs), (5) alpha-glucosidase inhibitors, (6) dipeptidyl peptidase IV (DPP-4) inhibitors, (7) bile acid sequestrants, (8) dopamine agonists, (9) sodium-glucose transport protein 2 (SGLT2) inhibitors and (10) oral glucagon-like peptide 1 (GLP-1) receptor agonists (Feingold et al., 2000), but metformin has been declared as the base for the treatment of type 2 diabetes mellitus all along with diet and exercise by European Association for the Study of Diabetes (EASD) and the American Diabetes Association (ADA) (Inzucchi et al., 2012).

Metformin is the most commonly used drug against diabetes, sold as Glucophage, the only biguanide drug available. Although metformin is widely used, it still has different
effects on different individuals. Clinical trial studies have shown that more than one-third of individuals getting metformin monotherapy fails to attain satisfactory control levels of fasting glucose. The fundamental cause for the lack of response to metformin in T2DM individuals may be due to changes in genes involved in drug pharmacokinetics and pharmacodynamics (Sissung et al., 2012).

It has been considered the best choice that reduces Hemoglobin A1c (HbA1c) level without causing any hypoglycemia in individuals (Nathan et al., 2006). It is a hydrophilic molecule, and the movement of metformin in the intestine, liver, and kidney is facilitated by several transporters, including plasma membrane monoamine transporter (PMAT). The passive distribution of metformin is restricted by its low lipid solubility across cell membranes (Graham et al., 2011).

Genetics plays a crucial part in the diagnosis at early stages, precise personal treatment, and how to monitor complex diseases, including T2DM (Ali, 2013). Genetic polymorphisms in the genes that encode their transporters directly affect the metformin pharmacokinetics and its therapeutic response (Vitarani et al., 2019). Besides, metformin’s pleiotropic effects give a new perspective for both descriptions of its inter-individual differences in response and the explanation of novel drug targets.

Researchers have lately identified significant associations between single nucleotide polymorphisms (SNPs) and the incidence of T2DM. Some studies have explored the effects of variations in genes organic cation transporter 1 (OCT1), organic cation transporter 2 (OCT2), human multidrug and toxin extrusion (MATE1) (Tarasova et al., 2012), human multidrug and toxin extrusion (MATE2) (Raj et al., 2018), and plasma membrane monoamine transporter (PMAT) (Dawed et al., 2019), on the patients of T2DM treated with metformin. Genetic polymorphisms in PMAT exhibited an association with decreased metformin absorption (Zhou et al., 2007). PMAT (SLC29A4) was first recognized from public genomic databanks as the 4th member in the mammalian SLC29 gene family, which encodes equilibrative nucleoside transporters (ENTs) molecularly and functionally different from the OCTs. PMAT is a poly-specific organic cation transporter, and its function is to transport many xenobiotic cations and bio- genic amines, including Metformin. SLC29A4 gene encodes PMAT, is located at chromosome 7p22.1. It has 12 exons and has 530 amino acid residues with 10-12 transmembrane segments. In the intestine, the absorption of metformin is generally carried out by PMAT. This transporter is expressed on the luminal side of enterocytes. However, no in-vivo studies have been performed on PMAT function and metformin’s pharmacological effect (Gong et al., 2012; Semiz et al., 2013).

This study has identified the high-risk SNPs in PMAT associated with Type 2 diabetes mellitus by utilizing both in-silico and experimental methods. The deleterious missense SNPs that are disrupting the structure and function of the protein were studied. The protein stability changes and analysis of amylogenic regions were explored in particular. This study facilitates finding the effect of mutations on the structure and function of PMAT. To the best of our knowledge, this is the first study that conducted a comprehensive analysis of non-synonymous SNPs of PMAT protein to find a suitable treatment for diabetic patients.

Materials and methods

High-Risk single nucleotide polymorphisms dataset

The amino acid sequence of PMAT (Q7RTT9) was downloaded from UniProt. Total 514 single nucleotide polymorphism (SNPs) were collected from dbSNP (Short Genetic Variations) https://www.ncbi.nlm.nih.gov/projects/SNP/ (Dayem Ullah et al., 2012). First, the non-synonymous SNPs were separated from synonymous ones by utilizing SNPnexus that outputs the combinatorial results from SIFT and PolyPhen. These variants were further subjected to high-risk SNP analysis.

PROVEAN

PROVEAN (Protein Variation Effect Analyzer) is a sequence-based tool that predicts variants’ effect on protein function. The threshold of –0.5 is set, if the variants are below the threshold, the SNPs are categorized as deleterious, and if it’s above the threshold, the variants are labeled as neutral. The tool is accessible at http://provean.jcvi.org/index.php (Choi & Chan, 2015).

Prediction of evolutionary conserved residues

The evolutionary conservation of amino acids was predicted using the ConSurf server https://consurf.tau.ac.il/ (Glaser et al., 2003). The server generates multiple sequence alignment of the queried protein sequence and computes the Bayesian algorithm’s conservation score.

Prediction of Disease-Related SNPs

To predict the SNPs as disease-associated or neutral, the Meta-SNP predictor was used. This tool combines four predictors, namely PANTHER, PhD-SNP, SIFT, and SNAP. The tool is accessible at http://snps.biofold.org/meta-snp/index.html (Manning et al., 2011). The majority voting of the four tools will label a mutation as diseased or neutral.

3D Structure modeling of PMAT

The 3D structure of PMAT was predicted by using homology modeling server I-TASSER. It takes in the FASTA sequence as input and predicts the 3D structure of the protein. The server can be accessed here, https://zhanglab.ccmb.med.umich.edu/I-TASSER/ (Yang et al., 2015). Further, the mutant structures were also generated from the PYMOL mutagenesis wizard. Each mutant is superimposed with the wild-type structure and is visualized using USCF Chimera. Further, the energy minimization of the mutant models was generated using the ModRefiner. The respective RMSDs were generated to analyze how well the mutant models deviate from the wild type.
Protein stability changes upon mutation

For analyzing the effect of protein stability changes, we used the Mutation Cutoff Scanning Matrix (mCSM), which relies on a graph-based structure. The server takes in protein structure and mutations list as input and outputs the mutations as stabilizing or destabilizing based on the scores. The tool can be accessed at http://biosig.unimelb.edu.au/mcsm/ (Pires et al., 2014). This is a graph-based method that computes distance patterns around the mutant residues. The method calculates the difference of Gibbs free energy, which then labels a mutation as destabilizing or neutral.

Prediction of amyloidogenic regions

Aggregation of incorrectly folded proteins is the cause of various diseases. The missense variants result in the formation of the amyloid aggregates and disrupt the protein function. We used the FoldAmyloid server to predict part of the sequence that is involved in aggregation. The server is available at http://bioinfo.protres.ru/fold-amyloid/ (Garbuzynskiy et al., 2010).

Predicting Protein-Drug interactions using ClusPro and AutoDock Vina

To predict the PMAT- metformin interactions computationally, we have employed the ClusPro web server with the default settings. The wild-type PMAT structure and the mutant models were docked with metformin to compute the difference in docking energies. ClusPro generated 10 different docked poses with the scores for each experiment that determines the binding affinities of the docked molecules. The server can be accessible at https://cluspro.bu.edu/ (Ngan et al., 2012).

Further, to validate our docking protocol, we have utilized AutoDock Vina (Kozakov et al., 2017). The preprocessing steps include preparing ligand and receptor files in pdbqt format via the graphical user interface program AutoDock Tools (ADT). ADT adds Kollman charges polar hydrogens, solvation parameters, and fragmental volumes to the protein. The AutoGrid was used to set the grid size, which we fix to 60 × 60 × 60 xyz points in this case. The protein and ligand were considered rigid while performing the docking. Further, the docked poses were selected based on the lowest binding energy or binding affinity.

Molecular dynamics simulations using CABS-Flex webserver

The molecular dynamics simulations provide a detailed analysis of the conformational changes of the protein by plotting Root Mean Square Fluctuations (RMSF). CABS-Flex uses CABS coarse-grained model for simulations. The input protein structure is represented in a coarse-grained representation with a single residue that can be represented by four different atoms. The wild-type protein and the mutant models were provided as an input to the CABS-Flex server, and RMSF was analyzed to observe the difference in the models in terms of conformational flexibility. The server can be accessed at http://biocomp.chem.uw.edu.pl/CABSflex2/ (Trott & Olson, 2010).

Experimental validation of in silico results by ARMS-PCR

We have designed an experimental analysis to provide sufficient justification for the results obtained from in-silico analysis. The protocol is described below.

Ethics statement

The endocrinology department of PIMS-Pakistan approved the present study. All the patients were informed before blood sample collection, and written consent was obtained from all the participants. All the clinical information gathered from the participants was kept confidential.

Study population and design

This was a case-control study. A total of 1200 unrelated individuals, including 800 clinically diagnosed T2DM individuals and 400 ethnically matched healthy individuals, were enrolled in this study as per their permission. The sample size was calculated by using an online calculator (Kuriata et al., 2018) by considering confidence level 95% and confidence interval (CI) in line with the literature (Creative Research Systems, 2017). Observed mean ages of healthy controls and T2DM patients were 43.55 ± 14.00 and 56.04 ± 12.86. There were 61.5% males in healthy controls and 58.63% in T2DM, whereas 38.5% females in healthy controls and 41.37% in T2DM. The T2DM patients were further categorized into two groups of 400 T2DM patients based on HbA1c level. T2DM patients with HbA1c level ≤7.0% were considered metformin responders and T2DM patients with HbA1c level >7.0% were considered metformin non-responders. American Diabetes Association (ADA) guidelines were followed to categorize T2DM patients into metformin responders and metformin non-responders. All T2DM patients were clinically diagnosed by the diabetologist of Pakistan Institute of Medical Sciences (PIMS) hospital, Pakistan-Islamabad. Individuals who failed to match the drugs criteria were disqualified from the study.

Blood collection and DNA extraction

Venous blood of all the involved patients and controls individuals were collected in 5 mL EDTA (ethylenediaminetetra-acetic acid) vacutainers. The blood was collected from the individuals from the year 2015 to 2016.

A standard phenol-chloroform method was used with minor alterations to extract DNA from WBCs (Kirby et al., 2002). 750 µL of blood was added into a tube with solution A. After the incubation of 10 min, tubes were centrifuged. 400 µL of solution A was added again, and tubes were centrifuged. Resuspension of the nuclear pellet was obtained by adding solution B, 20% sodium dodecyl sulphate (SDS), and proteinase K. The tube was then kept in a shaking incubator.
at 37°C for overnight. The next day fresh mixture of equal volumes (1:1) of solution C and solution D was added into the tube. After centrifugation, the upper aqueous phase was collected in a labelled tube. Followed by the addition of solution D to the aqueous layer, it was again subjected to centrifugation. The aqueous layer was collected in a properly labelled fresh tube. Sodium acetate and ethanol were added to precipitate DNA by inverting tubes several times. Once again, the tube was centrifuged, and the supernatant was wasted. To the DNA pellet, 70% ethanol was added. The mixture was centrifuged, and ethanol was discarded. The centrifuge tube was blot dried and then air-dried for 2-3 min. 100 μL of distilled water was added to the DNA pellet. The extracted DNA was successively examined on 1% agarose gel, as shown in Figure S1.

**Figure 1.** Evolutionary conservation profile generated from ConSurf web server that predicts the functional and structural residues of PMAT.

Haplotype and linkage disequilibrium analysis

The haplotype frequency was estimated using SNPStats, which uses the expectation-maximization (EM) algorithm for three SNPs. R software and genetics packages were used for linkage disequilibrium analysis.

Statistical analysis

Data were analyzed using IBM SPSS Statistics 20.0. The direct gene count method was used to calculate the genotypic and allelic frequencies. Hardy-Weinberg equilibrium test was used to compare the actual genotypes with the expected number based on the Hardy-Weinberg equilibrium theory (p = allele frequency, q = 1-p, p^2 + q^2 = 1) in controls. The Hardy-Weinberg principle states that a population’s allele and genotype frequencies will remain constant in the absence of evolutionary mechanisms. According to the Hardy-Weinberg principle, the variable p often represents the frequency of a dominant allele, and the variable q represents the recessive allele frequency. The sum of the frequencies must add up to 1, or 100 percent.

We used descriptive statistics with the mean ± SD in groups. To compare the differences between continuous variables, the student’s t-test or the Mann-Whitney test was done. The difference in genotypic and allelic frequencies of the polymorphisms was analyzed between healthy individuals and T2DM patients, metformin responder, and non-responder individuals using Fisher’s exact test. With the adjustment of age and gender, multinomial logistic
regression was applied to calculate the odds ratios (ORs) and 95% confidence intervals (CIs). To analyze the SNPs, five multinomial logistic regression models (co-dominant, dominant, recessive, over-dominant, and additive) were also used. To find % difference, the mean ± SD was calculated by the student’s t-test and variables significance was checked by applying the ordered Probit econometric technique because the dependent variables of the model were categorical. We estimate it by using Stata software. Haplotype analysis were generated from the genotyped data. The haplotype frequency was calculated using SNPStats, which uses the expectation-maximization (EM) algorithm for two SNPs. R software and Genetics package was used for Linkage disequilibrium analysis. $p < 0.05$ was considered significant.

## Results

### High-Risk SNP dataset

The non-synonymous substitutions were filtered from synonymous ones from SNPNexus. Out of 516 SNPs, 21 were categorized as missense variants. These variants were subjected to PROVEAN, which further classified 14 SNPs as deleterious. Next, these variants were further passed through Meta-SNP to predict if these SNPs were disease-associated or not. The tool predicted 12 SNPs as disease-causing. The method takes the average voting of 4 in-built tools, namely PHD-SNP, PANTHER, SIFT, and SNAP, to label a SNP as disease-related or neutral. The results are tabulated in Table 1.

### Evolutionary conservation analysis

The ConSurf web server was used to analyze evolutionarily conserved residues. The mutants, N123S, N86H, F234I, P429T, and P429S, were predicted as highly conserved and exposed at the protein’s surface, while two mutants G461S and A471V, were predicted as highly conserved and buried in the core of the protein. The buried amino acids usually contribute structurally to the protein; thus, a mutation might cause structural damage to the protein. While the mutants predicted as exposed are mostly functional therefore will disturb the function of the protein, which in turn leads to disease formation. Figure 1 highlights the mutant amino acids with their structural and functional impact on the protein.

### 3D Structure prediction using I-TASSER

The 3D structure of PMAT was generated from the i-TASSER homology modeling tool. The model is generated by using the top 10 threading templates with an average sequence coverage of 67%. The details of the templates and alignment are provided in Figure S2. The model was analyzed using phi and psi distributions of the Ramachandran plot, which showed that the highly preferred observations are 85.9 (Figure S3). Further, the PyMol was utilized to generate mutant models of PMAT, which were then superimposed and analyzed by the UCSF Chimera 1.11 visualization tool (Mahjabeen et al., 2013). All the mutant models were superimposed with the native structure, and the deviation was calculated using RMSD as figured in Figure S4. ModRefiner was used to perform the energy minimization of the modeled structures. The respective RMSDs calculated for wild-type and mutant structures are tabulated in Table1; only those mutants have tabulated whose calculated RMSD's are significantly different from the wild-type. The wild-type 3D structure and its mutant models are figured in Figure 2.

### Analyzing protein stability changes

The protein stability changes were predicted from mCSM. The difference of Gibbs's free energy between the folded and unfolded state of the proteins indicates the stability of the protein. More positive this value means more stable the protein is. The predicted scores of the SNPs determine the SNP as destabilizing or neutral. All the negative DDG scores produced from mCSM were labelled as destabilizing, while the positive values were labelled as stabilizing proteins. Considering this, 13 mutants are classified as destabilizing, while M108V, L144S, P429T, and P429S were predicted as highly destabilizing. The results are tabulated in Table 1.

### Analysis of amylogenic regions

The protein aggregates analysis showed eight mutants N86H, M108V, N123S, I218T, F234I, R256C, P429T/S, and L430F that were found in the aggregation profile. This implies that mutations are evidently disrupting the helix formation, which ultimately results in the aggregation of beta-sheets. Amyloids are associated with various diseases, including neurodegenerative disorders, diabetes, and prion diseases. Figure 3.
shows the graphical representation of amylogenic regions predicted from the FoldAmyloid server. Few of the mutations are also occurring in the amylogenic region, as figured in Figure S5.

Predicting protein-drug interactions using ClusPro and AutoDock vina

To determine the structural effect of the mutations, we have carried out blind docking of PMAT with Metformin using ClusPro. The interactions between the protein and ligand were analyzed by computing the buried surface area (BSA), which we considered as a strength of a bound complex. The BSA was computed for wild-type interactions and for each mutant model interaction with the ligand. The protein-ligand bound complex with the highest BSA is considered as the most stable compound. Among wildtype and all the mutant models, the wild-type interactions showed the highest BSA hence expected to be stable with the mutations effecting the stability of the complex. The results are tabulated in Table 2 and figured in Figure 4.

Further, we have employed Auto Dock Vina for validating the docking interactions obtained from ClusPro. The docking routine returns 10 docked poses, out of which we have selected the best-docked poses based on lower binding affinities. To analyze the interaction patterns between PMAT and Metformin, LigPlot + was utilized (Laskowski & Swindells, 2011). The 2D plots were drawn, and the binding residues were observed in the wild-type complex and the mutant complexes. The 2D plot of the wild-type illuminates the presence of hydrogen bonds at a distance of 3.13 Å whereas, the spiked atoms show the hydrophobic contacts. The binding pocket shows the presence of different residues in the wild-type compared to the mutants; hence might be affecting the protein-ligand interactions. The results are figured in Figure 4 and in supplementary Table S2 that shows the 2D interaction profile of the protein-ligand interaction and the 3D interaction plots which we have generated from the PLIP server (Pettersen et al., 2004). PLIP is a fully automated tool that determines seven contacts of protein-ligand interactions including hydrogen bonds, hydrophobic contacts, pi-stacking, pi-cation interactions, salt bridges, water bridges, and halogen bonds.

Molecular simulations using CABS-Flex 2.0

For the current study, we have utilized the CABS-Flex server to carry out simulations of the wild-type docked complex and the mutant model docked complexes to all atom dynamics molecular algorithms. The RMSF plots generated are figured in Figure 5. The plots show the RMSF range in the wild-type protein model from 0.02 nm to 2.44 nm. The high fluctuating residues can be seen around the residues 302-352. The center residues of all the protein structures show the highest fluctuations. Compared with the mutant models, the RMSF plot shows high fluctuating residues in all mutant complexes, especially in N86H, V79E, M108V, and L144S. This indicates that the wild-type structure is more stable as compared to the mutant ones. Thus, these mutations decreases the stability of the protein and can be correlated with the mCSM results.
The genotype frequencies of all three SNPs were further analyzed by four genetic models: dominant, over-dominant, recessive model, and additive model, as shown in Table S3.

Comparison between patients groups (metformin responders and non-responders)
For first SNP A > G (rs149798710), frequency of heterozygote AG (47.0% vs. 40.0%; OR = 1.51; 95%CI = 1.1284 to 2.0396; *p* = 0.0058) and mutant homozygote GG (13.5% vs. 9.0%; OR = 1.93; 95%CI = 1.2103 to 3.0990; *p* = 0.0059) were observed significantly higher in metformin non-responders (OR = 1.30; 95%CI 0.99 to 1.71; *p* = 0.05) compared to metformin responders. Remarkably, T2DM individuals that were homozygotes for wild allele ‘A’ (AA) had approximately 2 times more probability to respond metformin monotherapy. Same pattern was detected when assessed under other genetic models, (60.5% vs. 49.0%; OR = 1.59; 95%CI = 1.2044 to 2.1101; *p* = 0.0011) for dominant, (13.5% vs. 9.0%; OR = 1.57; 95%CI = 1.0096 to 2.4666; *p* = 0.0453) for recessive, (47.0% vs. 40.0%; OR = 1.33; 95%CI = 1.0050 to 1.7606; *p* = 0.0460) for over-dominant and 37.0% vs. 29.0%; OR = 1.43; 95%CI = 1.1660 to 1.7731; *p* = 0.0007) for additive as shown in Table S3.

For second SNP A > C (rs151039853) of SLC29A4, individuals with one mutant allele ‘A’ (AC) (42.5% vs. 48.5%; OR = 1.2103 to 3.0990; *p* = 0.0054) of rs149798710 were fewer among metformin responders than those who were unsuccessful to respond. T2DM individuals with AA genotype had 2 times more chances to respond metformin monotherapy. A similar trend was observed when evaluated under other penetrance models, (49.5% vs 60.0%; OR = 1.53; 95%CI = 1.1565 to 2.0249; *p* = 0.0029) for dominant, (7.0% vs. 11.5%; OR = 1.72; 95%CI = 1.0557 to 2.8232; *p* = 0.0296) for recessive, (28.5% vs. 35.75%; OR = 1.41; 95%CI = 1.1442 to 1.7455; *p* = 0.0013) for additive. No significant association was found for over-dominant (p > 0.05) (Table S3).

For third SNP of SLC29A4 T > A (rs17854505), the heterozygous TA genotype (37.0% vs. 48.0%; OR = 1.89; 95%CI = 1.4107 to 2.5484; *p* < 0.0001) and homozygous AA genotype (6.0% vs. 13.0%; OR = 3.16; 95%CI = 1.8736 to 5.3521; *p* = 0.0001) of rs17854505 were significantly less among metformin responders than those who failed to respond. Patients with T2DM that were homozygous for wild allele ‘T’
(TT) had 3 times more probability to respond metformin monotherapy. Same pattern was noticed when studied under other penetrance models (43.0% vs. 61.0%; OR = 2.07; 95%CI = 1.5638 to 2.7489; p < 0.0001) for dominant, (6.0% vs. 13.0%; OR = 2.34; 95%CI = 1.4126 to 3.8797; p < 0.0001) for recessive, (37.0% vs. 48.0%; OR = 1.57; 95%CI = 1.1852 to 2.0843; p = 0.0017) for over-dominant and (24.5% vs. 37.0%; OR = 1.80; 95%CI = 1.4586 to 2.2457; p < 0.0001) for additive model (Table S3).

Figure 4. Docked poses of protein-ligand with 2D and 3D interaction figures highlighting the most active residues (UNK0 is Metformin). (a) Wild type PMAT-metformin interaction. (b) Mutant Model N86H with Metformin, (c) Mutant model M108V with Metformin (d) Mutant model V79E with Metformin (e) Mutant model L144S with Metformin (f) Mutant model P429S with Metformin (g) Mutant model P429T with Metformin.

**Evaluation of clinical features between metformin responders and non-responders’ groups (baseline and after treatment)**

Table S5 presents the alterations from baseline in the clinical features of metformin responders and non-responders’ groups after successive metformin treatment for about...
6 months. The mean (SD) of metformin daily dose required in metformin responders and non-responders in the starting 6 months was 1000 mg.

No significant change was observed in age and gender between responders and non-responders. In metformin responders and non-responders, significant mean % variation change was observed in BMI (-2.32 ± 0.35 vs 0.01 ± 0.08, \(p < 0.001\)), fasting plasma glucose (-70.31 ± 0.13 vs -33.76 ± 0.43, \(p < 0.001\)), 2HPP blood glucose (-97.35 ± 0.25 vs -31.17 ± 1.95, \(p < 0.001\)), HbA1c (-2.44 ± 0.01 vs. 0.2 ± 0.03,

### Table 2.

| Bound Complex | Buried SA (Å²) | Binding Affinities (kcal/mol) |
|---------------|---------------|------------------------------|
| Wild-type     | 244.3654      | -4.9                         |
| N86H          | 212.0629      | -4.6                         |
| M108V         | 219.3676      | -4.7                         |
| V79E          | 210.0825      | -4.7                         |
| V79G          | 298.765       | -4.10                        |
| P429S         | 244.3655      | -4.9                         |
| P429T         | 244.3677      | -4.9                         |
| F234I         | 245.6789      | -4.9                         |
| L144S         | 220.0532      | -4.9                         |
| P429S         | 244.3677      | -4.9                         |

Figure 5. Superimposed 10 Protein 3D structures simulated with RMSF plots drawn for wild-type PMAT and mutant models N86H, V79E, M108V, L144S, P249T, and P249S.

6 months. The mean (SD) of metformin daily dose required in metformin responders and non-responders in the starting 6 months was 1000 mg.

No significant change was observed in age and gender between responders and non-responders. In metformin responders and non-responders, significant mean % variation change was observed in BMI (-2.32 ± 0.35 vs 0.01 ± 0.08, \(p < 0.001\)), fasting plasma glucose (-70.31 ± 0.13 vs -33.76 ± 0.43, \(p < 0.001\)), 2HPP blood glucose (-97.35 ± 0.25 vs -31.17 ± 1.95, \(p < 0.001\)), HbA1c (-2.44 ± 0.01 vs. 0.2 ± 0.03,
Table 3. Summary of a case-control study examining haplotypes of SLC29A4 SNPs (rs149798710, rs151039853, and rs17854505) and T2DM risk in metformin responders.

| Genotypes of SLC29A4 (rs149798710) | Genotypes of SLC47A2 (rs151039853) | Genotypes of SLC47A2 (rs17854505) | Frequency | OR (95% CI) | P-value |
|-----------------------------------|-----------------------------------|-----------------------------------|-----------|-------------|---------|
| 1 A                               | A                                 | T                                 | 0.7421    | 1.00        | —       |
| 2 G                               | C                                 | A                                 | 0.2141    | 1.47 (1.15 — 1.89) | p = 0.0026 |
| 3 G                               | C                                 | T                                 | 0.0255    | 1.49 (0.80 — 2.76) | p = 0.21 |
| 4 A                               | A                                 | A                                 | 0.0255    | 1.27 (0.10 — 1.99) | p = 0.29 |
| rare                              | *                                 | *                                 | 0.0183    | 1.19 (0.56 — 2.56) | p = 0.65 |

Abbreviations: G = Guanine; A = Adenine; T = Thymine; C = Cytosine, bp = base pairs, OR = odd ratios, CI = Confidence Interval. P-value < 0.05 was considered as significant.

Table 4. Summary of case-control study examining haplotypes of SLC29A4 SNPs (rs149798710, rs151039853 and rs17854505) and T2DM risk in metformin non-responders.

| Genotypes of SLC29A4 (rs149798710) | Genotypes of SLC47A2 (rs151039853) | Genotypes of SLC47A2 (rs17854505) | Frequency | OR (95% CI) | P-value |
|-----------------------------------|-----------------------------------|-----------------------------------|-----------|-------------|---------|
| 1 A                               | A                                 | T                                 | 0.6869    | 1.00        | —       |
| 2 G                               | C                                 | A                                 | 0.2696    | 2.73 (2.12 — 3.52) | p < 0.0001 |
| 3 G                               | C                                 | T                                 | 0.0157    | 0.51 (0.21 — 1.24) | p = 0.14 |
| 4 A                               | A                                 | A                                 | 0.0127    | 1.59 (0.34 — 186.55) | p = 0.002 |
| rare                              | *                                 | *                                 | 0.0151    | 0.86 (0.37 — 2.00) | p = 0.72 |

Where G = Guanine; A = Adenine; T = Thymine; C = Cytosine, bp = base pairs, OR = odd ratios, CI = Confidence Interval. P-value < 0.05 was considered as significant.

Table 5. Summary of a case-control study examining haplotypes of SLC29A4 SNPs (rs149798710, rs151039853, and rs17854505) and risk of the failure of metformin therapy.

| Genotypes of SLC29A4 (rs149798710) | Genotypes of SLC47A2 (rs151039853) | Genotypes of SLC47A2 (rs17854505) | Frequency | OR (95% CI) | P-value |
|-----------------------------------|-----------------------------------|-----------------------------------|-----------|-------------|---------|
| 1 A                               | A                                 | T                                 | 0.6511    | 1.00        | —       |
| 2 G                               | C                                 | A                                 | 0.2998    | 1.77 (1.41 — 2.23) | p < 0.0001 |
| 3 G                               | C                                 | T                                 | 0.0191    | 0.36 (0.15 — 0.85) | p = 0.2 |
| 4 A                               | A                                 | A                                 | 0.0124    | 0.17 (0.19-0.99) | p = 0.39 |
| rare                              | *                                 | *                                 | 0.0177    | 0.78 (0.35 — 1.74) | p = 0.55 |

Where G = Guanine; A = Adenine; T = Thymine; C = Cytosine, bp = base pairs, OR = odd ratios, CI = Confidence Interval. P-value < 0.05 was considered as significant.

Change in the study parameters from baseline to 6 months of metformin therapy according to the genotypes

Comparisons of differential values in metformin responders and non-responders with different SLC29A4 variants rs149798710, rs151039853, and rs17854505 genotypes before and after metformin treatment are presented in Table S4.

Haplotype analysis

Haplotypes were generated using 3 SNP’s of SLC29A4 i.e. rs149798710, rs151039853 and rs17854505 in healthy controls and T2DM patients (metformin responders). Haplotype G-C-A (OR = 1.47; 95%CI = 1.15 to 1.89; p = 0.0026) was linked significantly with the increased risk of T2DM (Table 3).

Haplotypes were generated using 3 SNP’s of SLC29A4 i.e. rs149798710, rs151039853 and rs17854505 in healthy controls and T2DM patients (metformin responders). Haplotype G-C-A (OR = 1.47; 95%CI = 1.15 to 1.89; p = 0.0026) was linked significantly with the increased risk of T2DM (Table 3).

Linkage disequilibrium analysis

All three SNPs are located at the same chromosome in proximity; the likelihood of linkage disequilibrium (LD) between the SNPs (rs149798710, rs151039853, and rs17854505) was studied. In LD, the D’ represents the coinheritance of the rare allele; when it is present, it is always inherited with one particular allele of the 50% polymorphism, whereas the r2
represents if it is a rare allele. The LD value close to $> 0.8$ indicates that SNPs are in strong LD and coinherited. $D$ (measures the deviation of haplotype frequencies from expected values based on gene frequencies). The value of $D$ ($D$ value $= -0.25$ to $0.25$, and $D = 0$ means no LD) is close to $0.2$ suggesting strong LD. The value of $r^2$ represents a rare allele; the $r^2$ value shows that the mutant allele of these SNPs's is rare.

**Discussion**

T2DM is one of the most common chronic metabolic diseases within developing countries (Olokoba et al., 2012). Predominantly, the drug "Metformin" is the most common glucose-reducing drug administered orally (Scarpello & Howlett, 2008; Viollet et al., 2012). Despite the emergence of new drug targets and augmented therapies for the treatment of diabetic patients, patient care still remains insufficient and not so promising because of the various side effects of the drugs used to treat the disease symptoms or disease, in general. To date, there has been the identification of nine classes of drugs employed for diabetes treatment in T2DM patients explicitly. Of all these, metformin has been considered the first line of treatment globally (Hu et al., 2012; Thule, 2012). In this study, the mutations in PMAT were systematically analyzed by combining sequence and structure based computational methods with further experimental verification analysis. Firstly, the SNPs were functionally annotated by various bioinformatics tools; out of 516 SNPs obtained from dbSNP, 13 SNPs were predicted as damaging/deleterious. These were further passed to evolutionary conservation analysis, which showed that the variants N86H, F234I, G461S, P429T/S as highly conserved and exposed on the surface of the protein have a functional role. Further, the protein stability analysis showed the substitutions N86H, V79E, L144S, M108V, and P429T/S as highly destabilizing and can affect the protein’s structure and function. Moreover, these mutants were also found in the aggregation profile, which shows that the mutations affect the alpha-helices formation, which results in the aggregation of beta-sheets. The aggregation of proteins is associated with a wide range of diseases, including diabetes (Cheng et al., 2012; Mukherjee et al., 2015). Further, the mutant models of PMAT were also generated using PyMOL and subjected to RMSD analysis to identify how deviated the mutants are from the wild-type structure. The results revealed L144S, P429T/S, M108V, N86H, and V79E as highly deviated with the RMSD values $>1$. These mutations were then analyzed by docking studies using the ClusPro web server and later validated with AutoDock Vina software. The computed buried surface area of the mutant models N86H, M108V, and P429T/S showed a significant decrease in BSA compared to the wild-type complex hence contributing to affecting the protein-ligand interactions. The 2D and 3D plots generated from LigPlot and PLIP servers also confirmed the activity of different residues in the binding pocket of wild-type PMAT and the mutant models. Later, the simulation analysis carried out by CABS-Flex determined the root mean square fluctuations with mutant models showing the highest fluctuations compared to the wild type hence contributing to disrupting the conformational flexibility of PMAT. Earlier, few studies have reported this fluctuation difference in protein native and mutant structures, which can shed light on the stability
analysis of the protein (Kumar et al., 2014; Pandey et al., 2016; Pereira et al., 2019).

These mutations were also further analyzed by the HOPE web server, which provides an in-detail description of the mutations affecting the structure-function of PMAT. For P429T/S, the wild-type residue is more hydrophobic than the mutant residue. These differences in hydrophobicity can affect the hydrophobic interactions with the membrane lipids. Further, the wild-type residue is proline. Prolines are known to be very rigid and therefore induce a unique backbone conformation, which might be required at this position. The mutation can disturb this special conformation. In L144S, the size and the hydrophobicity difference between the wild-type and mutant residue can affect the hydrophobic interactions with the membrane lipids.

The results of V79E showed a difference in charge between the wild-type and mutant amino acid. The mutation introduces a charge; this can cause the repulsion of ligands or other residues with the same charge. The hydrophobicity of the wild-type and mutant residue differs. Hydrophobic interactions, either in the core of the protein or on the surface, will be lost. As was observed in the docking studies, which further adds to our findings. The analysis of N86H showed that the mutant residue is located at a highly conserved position. The size difference between the two can affect the contact with the lipid membrane. For variant M108V, the mutants are located near a highly conserved position. The mutant residue is also smaller; this might lead to loss of interactions (Venselaar et al., 2010).

In light of these results, it can be inferred that with 13 SNPs predicted as deleterious among them, six mutants P429T/S, L144S, M108V, N86H, and V79E were found as most damaging to the structure and function of the protein with L144S, M108V, N86H, and V79E disrupting the protein-ligand bound complex. These mutants were seen as highly deviated from their native structures and predicted as destabilizing to the protein structure. The mutants P429T/S were predicted as highly conserved residues; hence mutation in these positions can cause damage to the function of the protein. M108V was found in the aggregation profile. Further, V79E and N86H were also decreasing the stability of PMAT as predicted by mCSM. The N86H was also predicted as highly evolutionary conserved and is also associated with beta-sheet aggregation. Hence, it could be one of the causative agents of diabetes. Its role should be explored further for exploring its association with disease in detail.

The results obtained from in-silico analysis were further verified experimentally. This was a case-control study. Polymorphisms in the gene SLC29A4 could result in loss-of-function or gain-of-function mutations resulting in the efflux transporters’ altered function. PMAT (SLC29A4) was first identified as the 4th member of the mammalian SLC29 gene family from public genomic databases, which encodes molecular and functionally specific equilibrium nucleoside transporters (ENTs) from OCTs family of transporters. This transporter is located on the enterocyte’s luminal side. Genetic polymorphisms in PMAT are associated with reduced absorption of metformin (Zhou et al., 2007). Currently, however, there have been no reports published about the pharmacological impact of metformin on the function of PMAT. In our present study, we have analyzed three variants and established a significant association with SLC29A4 variants M108V (rs149798710), N86H (rs151039853), and V79E (rs17854505), and Metformin clinical efficacy.

The missense mutation, polymorphism M108V A>G (rs149798710), arises in exon 4, allowing methionine to be transformed into valine at amino acid position 108. To date, no literature analysis on this SNP has been published so that the data generated is novel in its orientation. SLC29A4 rs149798710 polymorphism decreased Hba1c levels upon metformin treatment such that patients with the SLC29A4 rs149798710 AA genotype showed more significant reduction compared to those having AG or GG genotype. It might be due to the fact that the AA genotype enhances response to metformin which, in turn, causes a decrease in Hba1c values in these patients. Likewise, after 6 months of metformin treatment, the decrease in Hba1c was more significant in AA homozygotes than patients with the AG + GG genotype. The effect of SLC29A4 rs149798710 polymorphism on FBG and 2HPP also that AA genotype patients had better glucose lowering effect as well as improved lipid profiles after 6 months follow-up. Nonetheless, replication studies in a relatively large population in a multi-ethnic diabetic cohort are needed to confirm the SLC29A4 rs149798710 polymorphism effect on metformin efficacy both in glycaemic control and lipid profiles.

The N86H A>C (rs151039853) missense mutation occurs in exon 3, causing asparagine to be transformed into histidine at amino acid position 86. No previous data were published in the literature on this SNP; therefore, the produced information is novel in their orientation. SLC29A4 rs151039853 polymorphism has been found to reduce Hba1c upon treatment with metformin. This reduction was more in patients with the SLC29A4 rs151039853 AA genotype than those with AC or CC genotype. Furthermore, after 6 months of metformin therapy, the decrease in Hba1c was greater in AA homozygotes compared to patients with the AC + CC genotype. The effect of SLC29A4 rs151039853 polymorphism on FBG and 2HPP showed that AA genotype patients had better glucose lowering effect and improved lipid profile after 6 months follow-up in T2DM. Replication studies in a relatively large population in a multi-ethnic diabetic cohort are needed to confirm the SLC29A4 rs151039853 polymorphism effect on metformin efficacy both in glycaemic control and lipid profiles.

The missense mutation, polymorphism V79E T>A (rs17854505), arises in exon 3, triggering the transformation of valine to glutamic acid at amino acid position 79. To date, no literature on this SNP has been published so that the data generated is novel in its orientation. It was observed that SLC29A4 rs17854505 polymorphism decreased Hba1c values upon treatment with metformin. This reduction, however, was more significant in individuals exhibiting SLC29A4 rs17854505 TT genotype than those carrying TA or AA genotype. Comparably, after 6 months of metformin treatment, this decrease in Hba1c was greater in AA homozygotes.
compared with patients having TA + AA genotype. In contrast, TT genotype patients had better glucose lowering effects and improved lipid profiles. Replication studies in a relatively large population in a multi-ethnic diabetic cohort are needed to confirm the SLC29A4 rs17854505 polymorphism effect on metformin efficacy both in glycaemic control and lipid profiles. In the current study, these three SNPs of SLC29A4 were associated with an increased risk of T2DM and failure to respond to metformin therapy.

The results of the current study might have a practical implication in future personalized treatment of T2DM patients. However, the small sample size of the present study can be considered as a limitation of the study; therefore, more research in different ethnic groups with larger sample sizes is required to elucidate the role of M108V (rs149798710), N86H (rs151039853), and V79E (rs17854505) variants in Metformin response.

Conclusions

To the best of our knowledge, this is the first study from the Pakistani population that aims to find significant deleterious SNPs associated with PMAT. The current study has first conducted the in-silico analysis that employs various computational tools to determine the structural and functional damaging mutations, which were further validated through experimental analysis. The experimental method confirms the presence of three mutations M108V (rs149798710), N86H (rs151039853), and V79E (rs17854505) in patients. Hence, these three are more crucial in causing a structure-function impact on the PMAT protein. The mutant M108V was found in the aggregation profile, which means forming beta-sheet aggregates, which are most likely associated with Diabetes. These three SNPs are found as strong candidates in causing diseases associated with mutations in PMAT. Therefore, it can be explored further for determining its possible effects on drug discovery.

Acknowledgements

The entire group is grateful to all the members who agreed to be a part of this study. We acknowledge International Islamic University for providing lab facilities. We also acknowledge PIMS hospital for allowing us to collect blood samples from T2DM patients.

Disclosure statement

The authors have declared that they have no competing interests.

Author’s contributions

SM participated in the design of the study, carried out the experiments, performed the statistical analysis, and drafted the manuscript. SK and AG provided lab facilities to conduct experiments. MK helped in statistical analysis. AZ helped in haplotype and LD analysis. RN helped in designing blood collection. ZK carried out the computational analysis and also helped to draft the manuscript. All authors read and approved the final manuscript.

References

Ali, O. (2013). Genetics of type 2 diabetes. World Journal of Diabetes, 4(4), 114–123. https://doi.org/10.4239/wjd.v4i4.114
Brunetti, A., Chieffari, E., & Foti, D. (2014). Recent advances in the molecular genetics of type 2 diabetes mellitus. World Journal of Diabetes, 5(2), 128–140.
Cheng, P. N., Liu, C., Zhao, M., Eisenberg, D., & Nowick, J. S. (2012). Amyloid β-sheet mimics that antagonize protein aggregation and reduce amyloid toxicity. Nature Chemistry, 4(11), 927–933. https://doi.org/10.1038/nchem.1433
Choi, Y., & Chan, A. P. (2015). PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics (Oxford, England), 31(16), 2745–2747.

Creative Research Systems. (2017). Creative Research Systems. (online)- cited 2017 Jun 12. Available from: URL:www.systemsv.com

Dawed, A. Y., Zhou, K., van Leeuwen, N., Mahajan, A., Robertson, N., Koivula, R., Elders, P. J. M., Rauh, S. P., Jones, A. G., Holl, R. W., Stingl, J. C., Franks, P. W., McCarthy, M. I., T Hart, L. M., & Pearson, E. R. (2019). Variation in the plasma membrane monoamine transporter (PMAT, encoded in SLC29A4) and organic cation transporter 1 (OCT1, Encoded in SLC22A1) and gastrointestinal intolerance to metformin in type 2 diabetes: An IMI DIRECT study. Diabetes Care, 42(6), 1027–1033. https://doi.org/10.2337/dc18-2182

Dayem Ullah, A. Z., Lemoine, N. R., & Chelala, C. (2012). SNPnexus: A web server for functional annotation of novel and publicly known genetic variants (2012 update). Nucleic Acids Research, 40(Web Server issue), W65–W70.

Feingold KR, Anawalt B, Boyce A, (Eds.). (2000). Feingold's textbook of endocrinology, 5th ed. London: Elsevier Science Publishing Co.

Gong, L., Goswami, S., Giacomini, K. M., Altman, R. B., & Klein, T. E. (2012). Metformin pathways: Pharmacokinetics and pharmacodynamics. PharmacoGenetics and Pharmacogenomics, 22(11), 820–827.

Graham, G. G., Punt, J., Arora, M., Day, R. O., Doogue, M. P., Duong, J. K., Furlong, T. J., Greenfield, J. R., Greenup, L. C., Kirkpatrick, C. M., Ray, J. E., Timmins, P., & Williams, K. M. (2011). Clinical pharmacokinetics of metformin. Clinical Pharmacokinetics, 50(2), 81–98. https://doi.org/10.2165/11534750-000000000-00000

Hu, L., Zhuo, W., He, Y.-J., Zhou, H.-H., & Fan, L. (2012). Pharmacogenetics of P450 oxidoreductase: implications in drug metabolism and therapy. Pharmacogenetics and Genomics, 22(11), 812–819. https://doi.org/10.1097/FPG.0b013e3283589d2b

Inzucchi, S. E., Bergenstal, R. M., Buse, J. B., Diamant, M., Ferrannini, E., Nauck, M., Peters, A. L., Tsapas, A., Wender, R., & Matthews, D. R. (2012). Management of hyperglycemia in type 2 diabetes: A patient-centered approach: Position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Spectrum, 25(3), 154–171. https://doi.org/10.2337/diaspect.25.3.154

Kirby, A., Gebski, V., & Keech, A. C. (2002). Determining the sample size in a clinical trial. The Medical Journal of Australia, 177(5), 256–257. https://doi.org/10.5694/j.1326-5377.2002.tb04759.x

Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhony, D., Yueh, C., Beglov, D., & Vajda, S. (2017). The ClusPro web server for protein-protein docking. Nature Protocols, 12(2), 255–278. https://doi.org/10.1038/nprot.2016.169
Kumar, C. V., Svetha, R. G., Anbarasu, A., & Ramaiah, S. (2014). Computational analysis reveals the association of threonine 118 methionine mutation in PMP22 resulting in CMT-1A. Advances in Bioinformatics, 2014, 502618. https://doi.org/10.1155/2014/502618

Kuriata, A., Gierut, A. M., Oleniecki, T., Giemmy, M. P., Kolinski, A., Kurcinski, M., & Kniecik, S. (2018). CABS-flex 2.0: A web server for fast simulations of flexibility of protein structures. Nucleic Acids Research, 46(W1), W338–43. https://doi.org/10.1093/nar/gky356

Laskowski, R. A., & Swindells, M. B. (2011). LigPlot ++: multiple ligand-protein interaction diagrams for drug discovery. Journal of Chemical Information and Modeling, 51(10), 2778–2786. https://doi.org/10.1021/ cic200227u

Mahjabeen, I., Baig, R. M., Masood, N., Sabir, M., Inayat, U., Malik, F. A., & Raj, M. G., Mathaiyan, J., Mukta, V., & Priyadarshini, R. (2018). Lack of methionine mutation in PMP22 resulting in CMT-1A. Advances in Physiology Education, 32(4), 175–185. https://doi.org/10.1515/dmpet-2018-0030

Scarpello, J. H., & Howlett, H. C. (2008). Metformin therapy and clinical uses. Diabetes & Vascular Disease Research, 5(3), 157–167. https://doi.org/10.3132/dvdr.2008.027

Semiz, S., DJuc, T., & Causevic, A. (2013). Pharmacogenetics and personalized treatment of type 2 diabetes. Biochimia Medica, 23(2), 154–171.

Shu, Y., Sheard, S. A., Brown, C., Owen, R. P., Zhang, S., Castro, R. A., Ianculescu, A. G., Yue, L., Lo, J. C., Burchard, E. G., Brett, C. M., & Giacomini, K. M. (2007). Effect of genetic variation in the organic cation transporter 1 (OCT1) on metformin action. The Journal of Clinical Investigation, 117(5), 1422–1431.

Sissing, T. M., Troutman, S. M., Campbell, T. J., Pressler, H. M., Sung, H., Bates, S. E., & Figg, W. D. (2012). Transporter polymorphisms affect normal physiology, diseases, and pharmacotherapy. Discovery Medicine, 13(68), 19–34.

Stumvoll, M., Goldstein, B. J., & Van Haften, T. W. (2005). Type 2 diabetes: Principles of pathogenesis and therapy. The Lancet, 365(9467), 1333–1346. https://doi.org/10.1016/S0140-6736(05)61032-X

Tarasova, L., Kalina, I., Gendner, K., Bumbare, A., Ritenbera, R., Nikitina-Zake, L., Fridmanis, D., Vaiivade, I., Pirags, V., & Klovins, J. (2012). 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. Pharmacogenetics and Genomics, 22(9), 659–666. https://doi.org/10.1097/PFC. 0b013e3283561666

Thule, P. M. (2012). Mechanisms of current therapies for diabetes mellitus type 2. Advances in Physiology Education, 36(4), 275–283.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multi threading. Journal of Computational Chemistry, 31(2), 455–461. https://doi.org/10.1002/jcc.21334

Venselaar, H., Te Beek, T. A., Kuipers, R. K., Heekelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics, 11(1), 548. https://doi.org/10.1186/1471-2105-11-5

Violett, B., Guigas, B., Garcia, N. S., Leclerc, J., Foretz, M., & Andreelli, F. (2012). Cellular and molecular mechanisms of metformin: An overview. Clinical Science (London, England : 1979), 122(6), 253–270. https://doi.org/10.4042/CS20110386

Vitarani, D. A., Ningrum, R. I., & Rhea, F. (2019). Allele Frequency of SLC22A1 Met420del Metformin Main Transporter Encoding Gene among Javanese-Indonesian Population. Open Access Macedonian Journal of Medical Sciences, 7(3), 378–383. https://doi.org/10.3889/oamjms.2019.087

Wild, S., Roglic, G., Green, A., Sicree, R., & King, H. (2004). Global prevalence of diabetes: Estimates for the year 2000 and projections for 2030. Diabetes Care, 27(5), 1047–1053. https://doi.org/10.2337/diacare. 27.5.1047

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: Protein structure and function prediction. Nature Methods, 12(1), 7–8. https://doi.org/10.1038/nmeth.3213

Zhou, M., Xia, L., & Wang, J. (2007). Metformin transport by a newly cloned proton-stimulated organic cation transporter (plasma membrane monoamine transporter) expressed in human intestine. Drug Metabolism and Disposition: The Biological Fate of Chemicals, 35(10), 1956–1962. https://doi.org/10.1124/dmd.107.015495