Avsar
Egyptian Journal of Medical Human Genetics (2022) 23:92
https://doi.org/10.1186/s43042-022-00306-9

Egyptian Journal of Medical Human Genetics

RESEARCH Open Access

Analysis of missense SNPs in the *SLC47A1* and *SLC47A2* genes affecting the pharmacokinetics of metformin: Computational approach

Orcun Avsar*

Abstract

Background: Metformin as an anti-hyperglycaemic drug is commonly used for the treatment of type 2 diabetes mellitus (T2DM). The metformin response is variable due to the interindividual variation of pharmacokinetics which is based on strong genetic background. MATE1 and MATE2 proteins are significantly implicated in the pharmacokinetics of metformin. Missense SNPs with high risk of pathogenicity are expected to affect response to metformin via pharmacokinetics. Therefore, the aim of the current study is to determine the effects of missense SNPs in the *SLC47A1* and *SLC47A2* genes. The structural and functional consequences of all known *SLC47A1* and *SLC47A2* missense SNPs of the human MATE1 and MATE2 proteins were identified by various bioinformatics methods (SIFT, PhD-SNP, PolyPhen-2, PROVEAN, PMut, MUpro, I-Mutant 3.0, COACH, RaptorX Binding, ConSurf, STRING).

Results: The *SLC47A1* variants P186T, L116P and the *SLC47A2* variants I158N, L112P, V118G exhibited ΔΔG values less than − 1 kcal/mol, and these variants are considered to disrupt the structure and function of MATE1 and MATE2 proteins. *SLC47A1* R118Q and *SLC47A2* Y273C, V118G may significantly disturb protein function and transporting activities according to the analysis of ligand-binding regions.

Conclusion: It is suggested that high-risk deleterious missense SNPs may mediate the pharmacokinetics of metformin and may be associated with altered tissue distribution, renal clearance and metformin toxicity. We suppose that our results might serve as potential targets for the studies composed of the development of potential diagnostic and therapeutic strategies based on the relationship between mutations and metformin response.

Keywords: Diabetes, Metformin, SNP, Polymorphism, Mutation

Background

Metformin as an anti-hyperglycaemic drug is commonly used for the treatment of type 2 diabetes mellitus (T2DM). Metformin decreases hepatic glucose synthesis, postprandial and fasting glucose levels and enhances the utilization of peripheral glucose [1, 2]. Triglyceride and fatty acids levels are reduced by metformin treatment [3]. Furthermore, it has been reported that metformin may prevent the onset of T2DM in individuals with prediabetes [4]. The drug is not metabolized and is excreted in the urine invariably (the half-life is approximately 5 h) [2]. Metformin is widely distributed into various tissues such as kidney, liver and intestinal system by several transporter molecules in human body. Metformin is absorbed by intestinal system via plasma membrane monoamine transporter (PMAT), and OCT1 and OCT3 are implicated in the transportation of the drug. Moreover, metformin is transported into hepatic system by OCT1 and

*Correspondence: orcunavsar.gen@gmail.com

Department of Molecular Biology and Genetics, Hitit University, 19030 Corum, Turkey

OCT3. The excretion of metformin from liver is mediated by MATE1, whereas the drug is eliminated from kidney by both MATE1 and MATE2 [2, 5]. In other words, metformin transporters which play a key role in its pharmacokinetics are involved in its distribution to the body and in its elimination [6].

It is known that metformin does not work for each individual in an optimal way due to genetic variations. Also, approximately 35% of patients do not achieve the glycaemic control with metformin [7]. Genetic factors such as polymorphisms in genes encoding metformin transporters cause to variable response to metformin [8]. *SLC47A1* and *SLC47A2* genes encode multidrug and toxin extrusion (MATE) transporters. MATE transporters play a key role in the efflux of numerous hydrophilic organic cations such as metformin [9]. Metformin pharmacokinetics is mediated by MATE1 and MATE2. The human *SLC47A1* gene encodes a functional membrane transporter protein with 570 amino acids and is located on chromosome 17p11.2. *SLC47A1* is expressed in hepatocytes and kidney cells, whereas *SLC47A2* is mainly expressed in renal proximal tubules [10]. *SLC47A2* is located on chromosome 17p11.2 and involves 22 exons [11].

Genetic variations are one of the significant determinants of a patient's response to metformin, and identification of the genes and biological pathways that are associated with metformin response might also have the potential for the development of new treatment strategies of diabetes. In this regard, identification of the functional and structural effects of polymorphisms is fundamental to elucidate their impact on drug efficacy. Therefore, in this study, we aimed to determine the functional and structural consequences of all known missense SNPs of MATE1 and MATE2 transporter proteins by various bioinformatic tools.

Methods

The deleterious effects of missense SNPs on the structure, function and stability of human MATE1 and MATE2 transporter proteins were predicted by numerous bioinformatic methods.

Data collection

The missense SNPs and mutation data of *SLC47A1* and *SLC47A2* genes were obtained from the ClinVar database, the dbSNP database of the National Center for Biotechnology Information (NCBI) [12, 13] and the Human Gene Mutation Database (HGMD) [14]. The DNA and protein sequences, wild-type and mutant amino acids, SNP ID, amino acid change positions and minör allele frequency (MAF) were retrieved.

Estimation of deleterious missense SNPs

In the current study, we used five bioinformatics programs for the prediction of functional effects of missense SNPs. The Sorting Intolerant from Tolerant (SIFT) algorithm estimates the impacts of variants on the function of protein. SIFT is one of the significant web tools to characterize nonsynonymous SNPs. A nonsynonymous SNP is estimated to be deleterious if the score is < 0.05. On the other hand, a missense variant with SIFT score ≥ 0.05 is classified as benign [15]. Predictor of human Deleterious Single Nucleotide Polymorphisms (PhD-SNP) is an online web tool and estimates whether mutant phenotype is disease-related or neutral (not disease-related). PhD-SNP program presents a reliability index score [16]. Polymorphism Phenotyping v2 (PolyPhen-2) is an online web server and determines the possible effects of amino acid changes on the function and stability of proteins. The software categorizes missense mutations as "benign", "possibly damaging" and "probably damaging" [17]. Protein Variation Effect Analyzer (PROVEAN) predicts the impact of amino acid changes, deletions, and insertions. Delta alignment score analyzes the effect of a mutation. Lower delta scores are predicted as mutations with negative impacts on protein function, whereas higher scores are categorized as mutations with neutral effects [18]. PMut which is freely accessible algorithmic program predicts the pathological effects of mutations on protein. The estimation scores of the web tool are within the range of 0 and 1 (0.5–1: pathological; 0–0.5: neutral) [19].

Predictions of the effects of missense SNPs on the stability of MATE1 and MATE2 transporter proteins

MUpro and I-Mutant 3.0 were used to analyzes the protein stability alterations of MATE1 and MATE2 induced by the high-risk missense SNPs of the human *SLC47A1* and *SLC47A2* genes, respectively. MUpro is used to estimate the effects of missense mutations on the stability of proteins. MUpro may estimate protein stability alteration by applying protein sequence and tertiary structure information [20]. I-Mutant 3.0 predicts protein stability alterations caused by point mutations by using protein structure and sequence information. The prediction of the stability change is calculated by $\Delta\Delta G$ value (kcal/mol). The $\Delta\Delta G$ value lower than "0" shows that this mutation leads to decrease of stability, whereas the $\Delta\Delta G$ value higher than "0" shows that this mutation increases protein stability [21].

Estimation of ligand-binding regions

The ligand-binding regions in MATE1 and MATE2 were estimated by the COACH server and the RaptorX Binding web portal. COACH server predicts the

ligand-binding regions of proteins by the use of S-SITE and TM-SITE. The estimated C-scores are in the range of 0 and 1, and the scores increase with reliability [22]. RaptorX Binding tool predicts the binding regions of a protein sequence according to 3D model. This web portal estimates protein disordered sites, secondary and tertiary structures, solvent accessibility, binding regions. uGDT (GDT) and uSeqID (SeqID) are used for the prediction of binding sites [23].

Analysis of evolutionary conservation

The phylogenetic conservation analysis of amino acid residues of MATE1 and MATE2 was performed by ConSurf bioinformatics tool. The conservation scores are in the range of 1 and 9, where 9 shows conserved regions and 1 shows rapidly evolving positions. Buried regions with high scores are considered to be structural, whereas exposed regions with high scores are considered to be functional [24].

Analysis of protein–protein interactions

The STRING database was applied to obtain functional interactants of MATE1 and MATE2 transporter proteins [25].

Results

SNP annotation

We obtained *SLC47A1* and *SLC47A2* SNPs applying NCBI dbSNP database that involved total 11,675 SNPs and total 10,923 SNPs in *SLC47A1* and *SLC47A2* genes, respectively. 428 SNPs among a total of 11,675 SNPs in the human *SLC47A1* gene were missense SNPs, and 195 SNPs were synonymous SNPs. Out of the SNPs, 446 SNPs were missense and 219 SNPs were synonymous SNPs in the human *SLC47A2* gene. In the current study, only missense SNPs were selected for further in silico analysis.

Identification of deleterious missense SNPs

Five different in silico bioinformatic programs (SIFT, PolyPhen-2, PROVEAN, PhD-SNP, PMut) were used to reveal the deleterious SNPs that may significantly affect the structure and/or function of MATE1 and MATE2 proteins. Based on the SIFT scores, the missense SNPs with the value <0.05 were classified as deleterious and selected for further analysis, whereas the missense SNPs with the value >0.05 were categorized as tolerant and excluded from the study. PolyPhen-2 results that are in the range of 0 and 1 are divided into three groups: Benign, possibly damaging and probably damaging. In the current study, only the missense SNPs which had the nearest score to 1 (probably damaging) and consequently most significant deleterious effect on protein were selected. PROVEAN cut-off score is -2.5, and the polymorphisms with the values ≤ -2.5 were estimated to be damaging to the protein. The mutation with a PhD-SNP score >0.5 was predicted as deleterious, and the RI values are between 0 and 10 where 10 shows the highest reliability. PMut scores are in the range of 0 and 1, and the mutation with PMut score >0.5 was evaluated as pathogenic.

In studies, the missense SNPs which are predicted as deleterious by two, three or four bioinformatic tools at the same time are approved as deleterious. Similarly to the study [26], we selected the missense SNPs that were predicted as deleterious by at least four bioinformatic tools due to increase in the precision of estimation. Those SNPs with deleterious effects are shown in Tables 1 and 2. The genotype and allele frequencies of the known high-risk missense SNPs in different populations are given in Table 5.

Prediction of structural stability of MATE1 and MATE2 proteins

The effects of 13 high-risk missense SNPs of *SLC47A1* and 11 high-risk missense SNPs of *SLC47A2* on protein stability were determined by MUpro and I-Mutant 3.0 algorithmic programs by comparing free energies. When the DDG values of the mutations obtained from MUpro and I-Mutant softwares <0, the mutations were predicted to destabilize protein. Furthermore, the mutations with the DDG value <-1 had significantly decreasing effect on protein stability. Reliability index (RI) was generated by I-Mutant 3.0 web tool. RI values are in the range of 0 and 10. 10 refers to the highest reliability value, whereas 0 indicates the lowest reliability. It was demonstrated that P186T had the highest reliability value for protein stability, whereas Y203C and G285V with minimal reliability value (see Table 3). In Table 4, it was shown that V118G and K106E had the highest and lowest reliability values for protein stability, respectively. The findings revealed that 12 out of 13 high-risk missense SNPs decreased the stability and 1 missense SNP was found to increase the stability of MATE1, whereas 11 deleterious missense SNPs decreased MATE2 protein stability (Tables 3, 4). The *SLC47A1* 2 variants P186T, L116P and the *SLC47A2* 3 variants I158N, L112P, V118G unanimously exhibited $\Delta\Delta G$ values less than -1 kcal/mol computed by two tools, and these variants are considered to disrupt the structure and function of MATE1 and MATE2 proteins.

Prediction of the effects of high-risk missense SNPs on ligand-binding regions of MATE1 and MATE2

The findings of COACH web server show that (2s)-2,3-dihydroxypropyl(7z)-pentadec-7-enoate binding the human MATE1 protein occupies the rank 3 position with a C-score of 0.04 with residue R118.

Table 1 Deleterious missense SNPs in the human *SLC47A1* gene determined by bioinformatic tools

SNP ID	Nucleotide substitution	Amino acid change	SIFT result	SIFT score	PolyPhen-2 result	PolyPhen-2 score	PROVEAN result (cut-off = −2.5)	PROVEAN score	PhD-SNP result	PhD-SNP RI	PMut result	PMut score
rs77138970	G/A	R118Q	DEL	0.003	Pro-damg	0.999	DEL	−3.847	Neutral	1	Diseased	0.65
rs77630697	G/A	G64D	DEL	0.002	Pro-damg	1.000	DEL	−6.576	Diseased	5	Diseased	0.83
rs143542564	C/T	S151F	DEL	0.001	Pro-damg	0.995	DEL	−4.770	Diseased	3	Diseased	0.65
rs145557304	A/G	Y203C	DEL	0.004	Pro-damg	1.000	DEL	−6.345	Diseased	4	Neutral	0.33
rs145720500	T/C	M269T	DEL	0.002	Pro-damg	1.000	DEL	−5.911	Diseased	3	Diseased	0.65
rs147768037	G/A	G424R	DEL	0.001	Pro-damg	1.000	DEL	−7.544	Diseased	8	Diseased	0.87
rs149729794	C/A	P186T	DEL	0.002	Pro-damg	1.000	DEL	−7.578	Diseased	3	Diseased	0.63
rs200191180	T/C	L116P	DEL	0.001	Pro-damg	1.000	DEL	−6.598	Diseased	6	Diseased	0.82
rs200450505	G/A	G113D	DEL	0	Pro-damg	1.000	DEL	−6.731	Diseased	6	Diseased	0.77
rs201666623	G/T	G285V	DEL	0	Pro-damg	1.000	DEL	−8.845	Diseased	6	Diseased	0.73
rs202140517	G/A	G67S	DEL	0.025	Pro-damg	0.965	DEL	−5.623	Diseased	0	Neutral	0.32
rs372035240	G/T	W274L	DEL	0.013	Pro-damg	0.998	DEL	−12.807	Diseased	3	Neutral	0.47
rs375863644	C/T	R261C	DEL	0.012	Pro-damg	0.994	DEL	−4.335	Diseased	2	Neutral	0.40

DEL deleterious, *pro-damg* probably damaging

Table 2 Deleterious missense SNPs in the human *SLC47A2* gene determined by bioinformatic tools

SNP ID	Nucleotide substitution	Amino acid change	SIFT result	SIFT score	PolyPhen-2 result	PolyPhen-2 score	PROVEAN result (cut-off = −2.5)	PROVEAN score	PhD-SNP result	PhD-SNP RI	PMut result	PMut score
rs111838634	T/C	K106E	DEL	0.005	Pro-damg	0.999	DEL	−3.667	Diseased	5	Neutral	0.26
rs112362579	A/T	I158N	DEL	0.001	Pro-damg	0.973	DEL	−6.254	Diseased	6	Diseased	0.83
rs139778919	A/G	L112P	DEL	0.001	Pro-damg	1.000	DEL	−6.831	Diseased	6	Diseased	0.83
rs141717616	G/A	P125S	DEL	0	Pro-damg	1.000	DEL	−7.960	N	4	Diseased	0.64
rs146901447	G/A	P162L	DEL	0.002	Pro-damg	1.000	DEL	−9.850	Diseased	8	Neutral	0.28
rs147661457	G/T	H61Q	DEL	0	Pro-damg	1.000	DEL	−7.782	Diseased	3	Neutral	0.49
rs148039766	G/T	A128E	DEL	0.001	Pro-damg	1.000	DEL	−4.975	N	3	Diseased	0.80
rs148775490	C/T	G81R	DEL	0.002	Pro-damg	1.000	DEL	−7.167	Diseased	5	Diseased	0.83
rs199674976	G/T	A298D	DEL	0.004	Pro-damg	0.991	DEL	−4.380	Diseased	7	Neutral	0.28
rs373244724	T/C	Y273C	DEL	0	Pro-damg	1.000	DEL	−8.066	Diseased	4	Neutral	0.40
rs374762995	A/C	V118G	DEL	0.001	Pro-damg	0.973	DEL	−6.865	Diseased	4	Diseased	0.80

DEL deleterious, *pro-damg* probably damaging

Table 3 Effects of missense SNPs on MATE1 protein stability by MUpro and I-Mutant 3.0

SNP ID	Amino acid change	MUpro result[a]	MUpro DDG	I-Mutant result*	I-Mutant DDG	I-Mutant RI
rs77138970	R118Q	Decrease	− 0.5256	Decrease	− 0.52	7
rs77630697	G64D	Decrease	− 0.9342	Decrease	− 1.03	6
rs143542564	S151F	Increase	0.4154	Increase	0.35	4
rs145557304	Y203C	Decrease	− 1.4685	Decrease	− 0.72	2
rs145720500	M269T	Decrease	− 1.5494	Decrease	− 0.69	7
rs147768037	G424R	Decrease	− 0.6050	Decrease	− 0.44	4
rs149729794	P186T	Decrease	− 1.2193	Decrease	− 1.30	9
rs200191180	L116P	Decrease	− 1.5929	Decrease	− 1.22	3
rs200450505	G113D	Decrease	− 0.2109	Decrease	− 0.95	5
rs201666623	G285V	Decrease	− 0.4049	Decrease	− 0.24	2
rs202140517	G67S	Decrease	− 0.9497	Decrease	− 1.18	7
rs372035240	W274L	Decrease	− 0.2482	Decrease	− 0.47	6
rs375863644	R261C	Decrease	− 1.0377	Decrease	− 0.67	3

[a] Protein stability: Increase or decrease; DDG values are calculated inkcal/mol

9-[2-(ethoxycarbonyl)phenyl]-3,6-bis(ethylamino)-2,7-dimethylxanthylium chloride binding the human MATE2 occupies the rank 7 position with a C-score of 0.03 at Y273. The rank 5 site binds (2s)-2,3-dihydroxypropyl(7z)-pentadec-7-enoate at V118 with a C-score of 0.04. A pocket multiplicity value greater than 40 determines precise estimation based on the RaptorX Binding tool. However, the greatest pocket multiplicity was 13 for MATE1 and MATE2.

Evolutionary conservation analysis of high-risk missense SNPs

ConSurf web server predicted the evolutionary conservancy of amino acids of the native MATE1 and MATE2 proteins. ConSurf tool parameters were set as: Homolog search algorithm: HMMER, E-value cut-off: 0.0001, maximum %ID between sequences: 95, calculation method: Bayesian. It determines the structural or functional amino acid residues of deleterious missense SNPs of MATE1 and MATE2 proteins based on solvent accessibility and phylogenetic conservation. The amino acids positioned in conserved sites were highly deleterious compared to the amino acids in nonconserved sites. The amino acid residues are estimated as structural if they are highly conserved and buried and as functional if they are highly conserved and exposed. As shown in Fig. 1, R118 and G67 are found to be highly conserved and exposed and functional, whereas G64, M269, P186, G113, G285 residues are highly conserved and buried and structural. As shown in Fig. 2, P125 and H61 residues were predicted as functional, whereas V118 was determined as structural (Table 5).

Prediction of protein–protein interactions

STRING online database elucidates the functional protein–protein interactions in a cell. Protein–protein interaction was predicted in order to identify the functional partners of MATE1 and MATE2. Moreover, it was aimed to determine whether the high-risk missense SNPs that disrupt the structure of MATE1 may affect the interaction with MATE2 or the high-risk missense SNPs that disrupt the structure of MATE2 may affect the interaction with MATE1. Through analysis of MATE1 protein with STRING, functional interactants with high confidence (score 0.7) were indicated as follows: SLC22A2, SLC22A1, SLC22A8, SLC22A4, SLC29A4, SLC22A16, SLC29A3, SLC6A19, SLC22A5, SLCO1B1. Functional partners of MATE2 with high confidence (score 0.7) were indicated as follows: SLC22A2, SLC22A1, SLC22A8, SLC22A4, SLC22A5, SLCO1B3, ENSG00000257046, SLC29A4, SLC22A3, SLCO1B1. The interactions of these proteins are given in Figs. 3 and 4. It was determined that MATE1 protein does not interact with MATE2 protein as shown in Figs. 3 and 4 in this study.

Discussion

Metformin is the most common prescribed drug that is used for the treatment of T2DM. The metformin response is variable due to the interindividual variation of pharmacokinetics which is based on strong genetic background. Multi-drug and toxin extrusion proteins (MATEs), i.e. MATE1 and MATE2, are determinants of the pharmacokinetics of metformin and promote the efflux of metformin into the urine [27].

Table 4 Effects of missense SNPs on MATE2 protein stability by MUpro and I-Mutant 3.0

SNP ID	Amino acid change	MUpro result*	Mupro DDG	I-Mutant result[a]	I-Mutant DDG	I-Mutant RI
rs111838634	K106E	Decrease	−0.9204	Decrease	−0.33	0
rs112362579	I158N	Decrease	−1.4481	Decrease	−1.98	5
rs139778919	L112P	Decrease	−1.786	Decrease	−1.40	5
rs141717616	P125S	Decrease	−0.7955	Decrease	−1.01	7
rs146901447	P162L	Decrease	−0.1120	Decrease	−0.80	8
rs147661457	H61Q	Decrease	−0.4299	Decrease	−0.34	6
rs148039766	A128E	Decrease	−0.3744	Decrease	−0.02	4
rs148775490	G81R	Decrease	−0.6689	Decrease	−0.70	4
rs199674976	A298D	Decrease	−0.6131	Decrease	−0.64	5
rs373244724	Y273C	Decrease	−0.7329	Decrease	−0.91	4
rs374762995	V118G	Decrease	−2.619	Decrease	−1.84	10

[a] Protein stability: Increase or decrease; DDG values are calculated in kcal/mol

Structural conformation and protein stability are significant for the biological activity and function of protein and modulation of biomolecules. Therefore, it is necessary to elucidate the effects of high-risk missense SNPs of MATE1 and MATE2. In the current study, we attempted in silico analysis to reveal the most damaging missense SNPs and how they affect the structure and/or function of MATE1 and MATE2 proteins.

Among 428 missense SNPs in the human *SLC47A1* gene, finally 13 damaging missense SNPs were screened by the use of five bioinformatics prediction tools: SIFT, PolyPhen-2, PROVEAN, PhD-SNP, PMut, whereas 11 deleterious missense SNPs among 446 missense SNPs in the human *SLC47A2* gene were analyzed.

Protein stability is crucial for the structure, function and activity of a protein [28] and decreased protein stability and misfolding are the main outcomes of pathogenic nonsynonymous mutations [29]. The deleterious missense SNPs that might affect the stability of the human MATE1 and MATE2 proteins were obtained using MUpro and I-Mutant. Among 13 missense SNPs, 12 missense SNPs decreased the stability of MATE1 protein, whereas 11 missense SNPs had a decreasing effect on the MATE2 stability. Only 2 missense SNPs and 3 missense SNPs had significantly decreasing effect on MATE1 and MATE2 proteins, respectively. The conformational structure of a protein is modulated by protein stability, and therefore, function of a protein is determined by stability. Changes in protein stability might lead to degradation, misfolding or aggregation of proteins [30]. Furthermore, evolutionary conservancy in the protein sequence is necessary to identify whether a mutation may affect the host negatively. We obtained that highly damaging missense SNPs (R118Q, P186T) with great conservation scores were located in the highly conserved sites and hence affect pharmacokinetics of metformin by inactivating MATE1 and MATE2.

Ligand-binding regions were determined by COACH and RaptorX Binding web servers. One potential binding region for MATE1 and two regions for MATE2 were identified. In the current study, it has been discovered that R118Q (MATE1), Y273C (MATE2) and V118G (MATE2) substitutions are implicated in the ligand-binding region and constitute the sphere of catalytic coordination, which may affect the binding affinities of MATE1 and MATE2 proteins. Therefore, mutations that are at these binding regions (R118Q-MATE1; Y273C, V118G-MATE2) may subsequently disturb protein function and transporting activities. In accordance with our findings, it has been reported that Y273C is associated with reduced transport activity [31].

The protein interaction network of MATE1 was determined by STRING, and the findings showed strong interaction of MATE1 with SLC22A2, SLC22A1, SLC22A8, SLC22A4, SLC29A4, SLC22A16, SLC29A3, SLC6A19, SLC22A5, SLCO1B1, implicated in the transport of some organic cations and drugs. Moreover, the results exhibited that MATE2 has 10 associated partners SLC22A2, SLC22A1, SLC22A8, SLC22A4, SLC22A5, SLCO1B3, ENSG00000257046, SLC29A4, SLC22A3 and SLCO1B1, involved in the excretion of organic anions and therapeutic agents. These proteins might interact to mediate the transport of molecules. Due to high-risk missense SNPs in MATE1 and MATE2, the transport mechanism may be disturbed that might eventually cause to alterations in pharmacokinetics.

The Pharmacogenetics Knowledge Base (PharmGKB) determines how variations in human genes impact drug response. It provides valuable information about human genetic variations, drug responses and associated

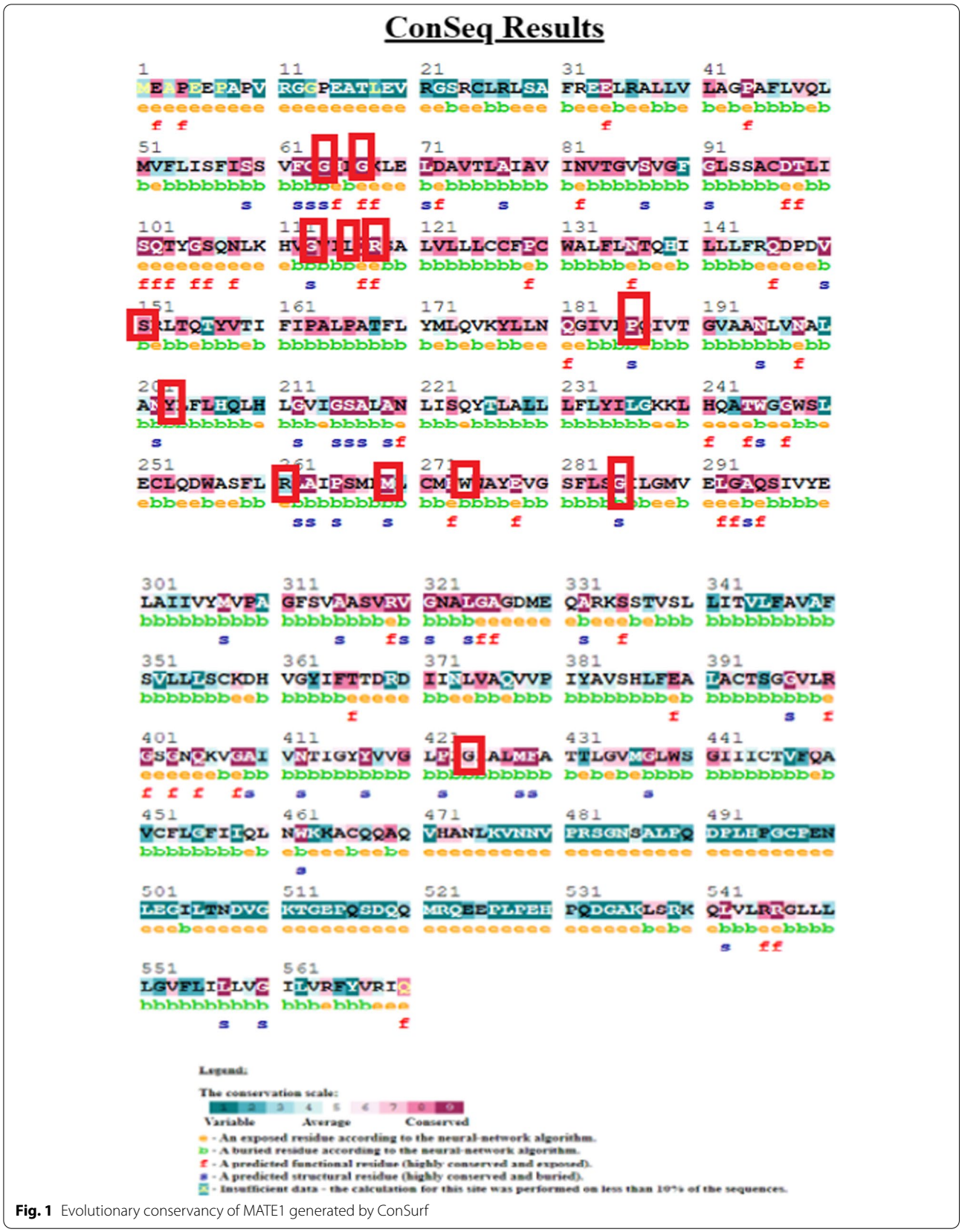

Fig. 1 Evolutionary conservancy of MATE1 generated by ConSurf

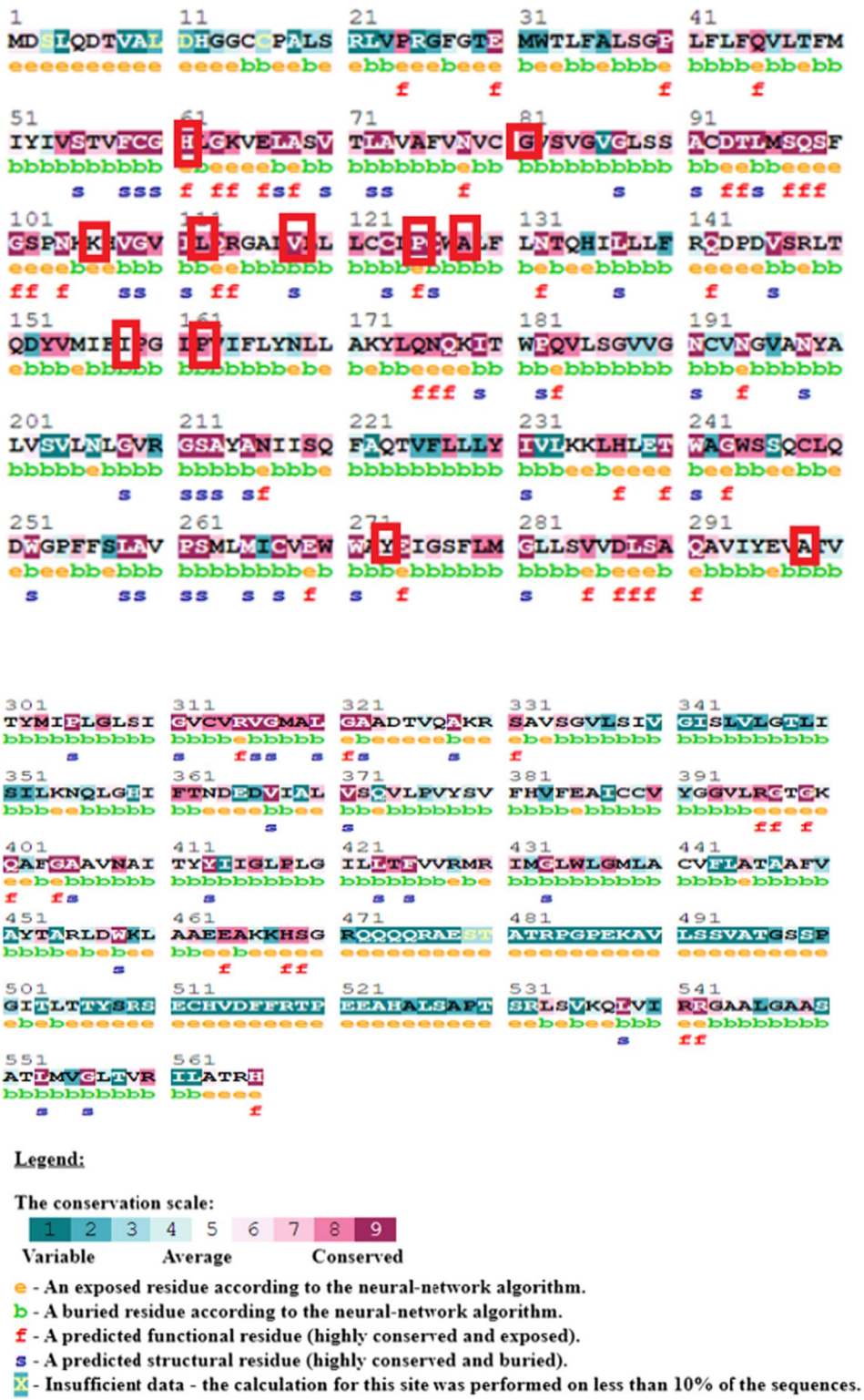

Fig. 2 Evolutionary conservancy of MATE2 generated by ConSurf

Table 5 Allele and genotype frequencies of the known high-risk missense SNPs in different populations

SNP ID	Population	Allele: frequency (count)	Genotype: frequency (count)
rs77138970	African	G: 1.000 (1322)	G|G: 1.000 (661)
	American	G: 1.000 (694)	G|G: 1.000 (347)
	East Asian	G: 1.000 (1008)	G|G: 1.000 (504)
	European	G: 0.999 (1005)	G|G: 0.998 (502)
	South Asian	G: 1.000 (978)	G|G: 1.000 (489)
rs149729794	African-American	C: 0.9998 (4405)	C|C: 0.9995 (2202)
	European American	C: 1.000 (8600)	C|C: 1.000 (4300)
rs373244724	African-American	T: 0.9998 (4405)	T|T: 0.9995 (2202)
	European American	T: 1.000 (8600)	T|T: 1.000 (4300)
rs139778919	African	A: 0.997 (1318)	A|A: 0.994 (657)
	American	A: 1.000 (694)	A|A: 1.000 (347)
	East Asian	A: 1.000 (1008)	A|A: 1.000 (504)
	European	A: 1.000 (1006)	A|A: 1.000 (503)
	South Asian	A: 1.000 (978)	A|A: 1.000 (489)

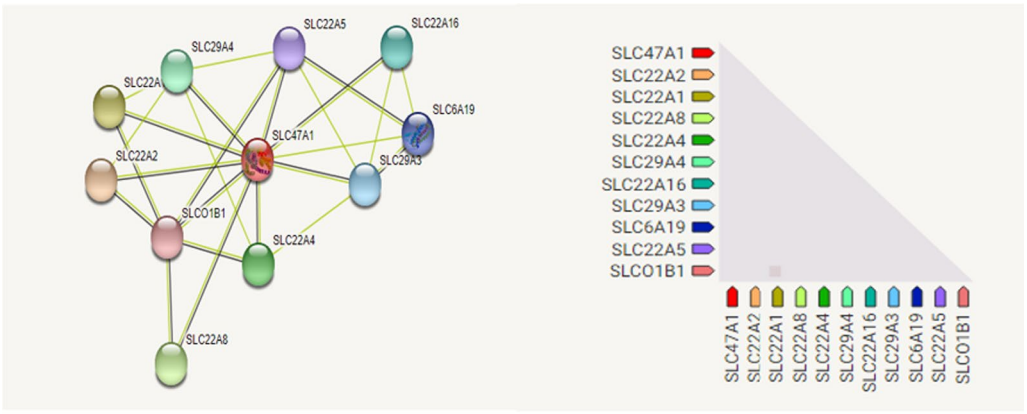

Fig. 3 **a** MATE1 protein–protein interactions with its partners **b** Coexpression pattern of the genes which correlates with the gene encodes MATE1

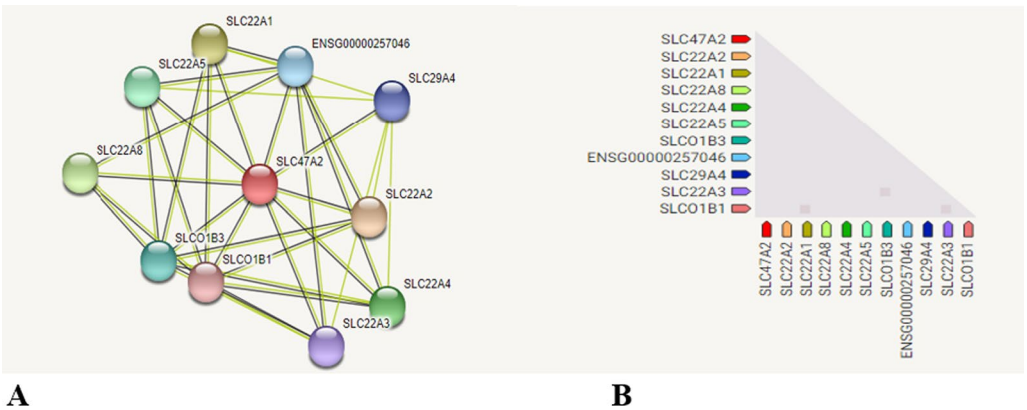

Fig. 4 **a** MATE2 protein–protein interactions with its partners **b** Coexpression pattern of the genes which correlates with the gene encodes MATE2

pathways [32]. According to the database, no known effects of the missense SNPs investigated in the present study on metformin pharmacokinetics were found. In a study conducted with Pakistani individuals with T2DM, it was reported that SLC47A1 rs77630697 (G64D) polymorphism was correlated with metformin response and metformin efficacy was reduced in patients that carry the mutant alleles [6]. It has been demonstrated that in vitro decreased metformin transport activity was associated with various nsSNPs in the human SLC47A2 gene [33]. SLC47A2 rs146901447 (P162L) and rs373244724 (Y273C) polymorphisms were reported to be resulted in decreased transport function and associated with metformin response and pharmacokinetics [11, 34].

MATE1 and MATE2 are some of the determinants of metformin pharmacokinetics and important for its pharmacological action. These transporters provide metformin secretion into urine [35]. It has been demonstrated that MATE1 and MATE2 that play a role in renal clearance of metformin may modulate transcellular metformin transport and affect metformin response [27]. High-risk missense SNPs determined by the present study may affect elimination and/or accumulation of metformin in human tissues. These missense SNPs may lead to decreased transport level, metformin efflux and elevated tissue level. Disrupted MATE1 and MATE2 functions may lead to increased concentration of metformin and subsequently may cause to adverse effects [35]. In a study conducted with nonsynonymous SNPs, some nsSNPs were shown to be associated with decreased transport activity and protein expression of MATE2 [31]. High-risk missense SNPs may alter transport activity, protein expression and secretion of MATE1 and MATE2. Moreover, missense SNPs can disturb the interaction with metformin and consequently may affect its pharmacokinetics.

We suggest that high-risk deleterious missense SNPs may mediate the pharmacokinetics of metformin and can be associated with altered tissue distribution, renal clearance and metformin toxicity.

Conclusion

MATE1 and MATE2 proteins are significantly implicated in the pharmacokinetics of metformin. Stability and structural conformation of these proteins are vital to execute their functions as transporters. In the current study, various bioinformatic algorithms were utilized to predict the deleterious effects of missense SNPs on the structure and function of MATE1 and MATE2. This is the first study that determines the effects of missense SNPs on the structure and function of MATE1 and MATE2 transporter proteins. Our results demonstrate that R118Q for MATE1 and Y273C, V118G for MATE2 affect the binding of ligands. Furthermore, the SLC47A1 variants P186T, L116P and the SLC47A2 variants I158N, L112P, V118G are considered to significantly disrupt the structure and function of MATE1 and MATE2 proteins. As these transporter proteins are known to play crucial roles in metformin pharmacokinetics, we suppose that our results may direct the studies composed of the development of potential diagnostic and therapeutic strategies which need experimental evaluation and clinical trials based on the relationship between mutations and metformin response.

Abbreviations

HGMD: Human Gene Mutation Database; MAF: Minör allele frequency; MATE: Multidrug and toxin extrusion; NCBI: National Center for Biotechnology Information; PharmGKB: The Pharmacogenetics Knowledge Base; PhD-SNP: Predictor of human Deleterious Single Nucleotide Polymorphisms; PMAT: Plasma membrane monoamine transporter; PolyPhen-2: Polymorphism Phenotyping v2; PROVEAN: Protein Variation Effect Analyzer; SIFT: The Sorting Intolerant from Tolerant; SNP: Single nucleotide polymorphism; T2DM: Type 2 diabetes mellitus.

Acknowledgements

Not applicable.

Author contributions

OA performed all in silico analysis and wrote the manuscript. The author read and approved the final version of the manuscript.

Funding

This study did not receive any financial support.

Availability of data and materials

Not applicable.

Declarations

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The author declared that he has no conflict of interest.

Received: 1 July 2021 Accepted: 5 May 2022
Published online: 10 May 2022

References

1. Becker ML, Visser LE, van Schaik LHN (2009) Genetic variation in the organic cation transporter 1 is associated with metformin response in patients with diabetes mellitus. Pharmacogenom J 9:242–247
2. Li G, Goswami S, Giacomini KM et al (2012) Metformin pathways: pharmacokinetics and pharmacodynamics. Pharmacogenet Genom 22:820–827
3. Todd JN, Florez JC (2014) An update on the pharmacogenomics of metformin: progress, problems and potential. Pharmacogenomics 15:529–539
4. Pawlyk AC, Giacomini KM, McKeon C et al (2014) Metformin pharmacogenomics: current status and future directions. Diabetes 63:2590–2599

5. Kinaan M, Ding H, Triggle CR (2015) Metformin: an old drug for the treatment of diabetes but a new drug for the protection of the endothelium. Med Princ Pract 24:401–415
6. Moeez S, Khalid M, Khalid Z et al (2019) Genotypic and allelic distribution of polymorphic variants of gene SLC47A1 Leu125Phe (rs77474263) and Gly64Asp (rs77630697) and their association to the clinical response to metformin in adult Pakistani T2DM patients. Int J Med Health Res 13:294–305
7. Cook MN, Girman CJ, Stein PP et al (2007) Initial monotherapy with either metformin or sulphonylureas often fails to achieve or maintain current glycaemic goals in patients with type 2 diabetes in UK primary care. Diabet Med 24:350–358
8. Mousavi S, Kohan L, Yavarian M et al (2017) Pharmacogenetic variation of SLC47A1 gene and metformin response in type2 diabetes patients. Mol Biol Res Commun 6:91–94
9. Moeez S, Khalid Z, Jalil F et al (2019) Effects of SLC22A2 (rs201919874) and SLC47A2 (rs138244461) genetic variants on Metformin Pharmacokinetics in Pakistani T2DM patients. J Pak Med Assoc 69:155–163
10. Manolopoulos VG, Ragia G (2014) Chapter 30: pharmacogenomics of oral antidiabetic drugs. In: Padmanabhan S (ed) Handbook of pharmacogenomics and stratified medicine. Elsevier, New York, pp 683–713
11. Nies AT, Damme K, Kruck S et al (2016) Structure and function of multidrug and toxin extrusion proteins (MATEs) and their relevance to drug therapy and personalized medicine. Arch Toxicol 90:1555–1584
12. Sherry ST, Ward MH, Kholodov M et al (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29:308–311
13. Landrum MJ, Lee JM, Benson M et al (2018) ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res 46:D1062–D1067
14. Stenson PD, Mort M, Ball EV et al (2014) The human gene mutation database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. Hum Genet 133:1–9
15. Sim NL, Kumar P, Hu J et al (2012) SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res 40:W452–W457
16. Capriotti E, Fariselli P (2017) PhD-SNPg: a webserver and lightweight tool for scoring single nucleotide variants. Nucleic Acids Res 45:W247–W252
17. Adzhubei I, Jordan DM, Sunyaev SR (2013) Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 7:Unit7.20
18. Choi Y, Chan AP (2015) PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 31:2745–2747
19. Lopez-Ferrando V, Gazzo A, de la Cruz X et al (2017) PMut: a web-based tool for the annotation of pathological variants on proteins, 2017 update. Nucleic Acids Res 45:W222–W228
20. Cheng J, Randall AZ, Baldi P (2005) Prediction of protein stability changes for single-site mutations using support vector machines. Proteins Struct Funct Bioinform 62:1125–1132
21. Capriotti E, Fariselli P, Casadio R (2005) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 33:W306–W310
22. Yang J, Roy A, Zhang Y (2013) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformtics 29:2588–2595
23. Källberg M, Wang H, Wang S et al (2012) Template-based protein structure modeling using the RaptorX web server. Nat Protoc 7:1511–1522
24. Ashkenazy H, Abadi S, Martz E et al (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 44:W344–W350
25. Szklarczyk D, Gable AL, Nastou KC et al (2020) The STRING database in 2021: customizable protein–protein networks, and functional characterization of user uploaded gene/measurement sets. Nucleic Acids Res. 49(D1):D605–D612.
26. Arshad M, Bhatti A, John P (2018) Identification and in silico analysis of functional SNPs of human TAGAP protein: a comprehensive study. PLoS ONE 13:e0188143
27. Stocker SL, Morrissey KM, Yee SW et al (2013) The effect of novel promoter variants in MATE1 and MATE2 on the pharmacokinetics and pharmacodynamics of metformin. Clin Pharmacol Ther 93:186–194
28. Hossain S, Roy AS, Islam S (2020) In silico analysis predicting efects of deleterious SNPs of human RASSF5 gene on its structure and functions. Sci Rep 10:14542
29. Zhang M, Huang C, Wang Z et al (2020) In silico analysis of non-synonymous single nucleotide polymorphisms (nsSNPs) in the human GJA3 gene associated with congenital cataract. BMC Mol Cell Biol 21:12
30. Witham S, Takano K, Schwartz C et al (2011) A missense mutation in CLIC2 associated with intellectual disability is predicted by in silico modeling to afect protein stability and dynamics. Proteins Struct Funct Bioinform 79:2444–2454
31. Nishimura K, Ide R, Hirota T et al (2014) Identification and functional characterization of novel nonsynonymous variants in the human multidrug and toxin extrusion 2-K. Drug Metab Dispos 42:1432–1437
32. Thorn CF, Klein TE, Altman RB (2013) PharmGKB: the pharmacogenomics knowledge base. Methods Mol Biol 1015:311–320
33. Mannino GC, Andreozzi F, Sesti G (2019) Pharmacogenetics of type 2 diabetes mellitus, the route toward tailored medicine. Diabetes Metab Res Rev 35:e3109
34. Choi JH, Yee SW, Ramirez AH et al (2011) A common 5′-UTR variant in MATE2-K is associated with poor response to metformin. Clin Pharmacol Ther 90:674–684
35. Liang X, Giacomini KM (2017) Transporters involved in metformin pharmacokinetics and treatment response. J Pharm Sci 106:2245–2250

Publisher's Note