universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Development of a KNIME workflow to link disease-associated alteration of metabolite levels with SLC transporter substrates“

verfasst von / submitted by

Michael Wildfellner, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magister pharmaciae (Mag. pharm.)

Wien, 2022 / Vienna, 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 605
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Pharmazie
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Gerhard Ecker
Mitbetreut von / Co-Supervisor:	Dipl.-Ing. (FH) Dr. Daniela Digles

DANKSAGUNG

An dieser Stelle möchte ich mich bei all jenen bedanken, die mich während der Anfertigung dieser Masterarbeit unterstützt haben.

Bedanken möchte ich mich in erster Linie bei der Mitbetreuerin meiner Arbeit Dr. Daniela Digles, für Ihre wertvolle Unterstützung sowohl in der Entwicklung als auch bei der Umsetzung der Idee dieser Masterarbeit. Danke auch dafür, dass ich mich mit meinen Fragen jederzeit an Dich wenden konnte.

Ich möchte mich auch bei Prof. Gerhard Ecker für die Möglichkeit bedanken, meine Masterarbeit in seiner Forschungsgruppe zu schreiben.

Weiters möchte ich mich auch bei allen Mitgliedern der Pharmacoinformatics Research Group für Ihre Hilfsbereitschaft und die gute Arbeitsatmosphäre bedanken.

Abschließend möchte ich meiner Familie, insbesondere meinen Eltern, für ihre Unterstützung danken. Danke für die Dinge, die Ihr mir auf meinen Lebensweg mitgegeben habt.

ABSTRACT

The heterogeneous superfamily of solute carriers represents the largest group of membrane transporters and enables the transport of a large number of substances across plasma membranes. The importance of SLCs stems from their involvement in a variety of physiological but also pathological processes, whereby also their pharmacological relevance, for example by influencing the pharmacokinetics of drugs, is noteworthy. However, although individual family members are well described today and there has been increased interest in researching this transport protein family on a broader basis, especially in recent years, the functions and properties of a large number of SLC transporters are still unclear.

On the one hand, the work presented here aims to uncover previously unknown associations between changes in metabolite levels, diseases and SLC transporters by combining data on metabolite signatures of diseases with data on SLC-disease relationships. On the other hand, an attempt is made to describe possible new SLC substrates using the same approach. For the implementation of this idea an appropriate workflow was created with the data analytics platform KNIME.

The data set resulting from this integration shows associations between diseases and SLCs, which have already been described in different studies. This indicates that the other results may also contain relevant, not yet described correlations. Also the second resulting data set describing possible new substrates of SLCs provides reasonable, if not really surprising, results.

ZUSAMMENFASSUNG

Die heterogene Superfamilie der Solute Carrier bildet die größte Gruppe unter den Membrantransportern und ermöglicht die Passage einer großen Anzahl von Substanzen durch Plasmamembranen. Bedeutend sind SLCs dabei aufgrund ihrer Beteiligung an einer Vielzahl physiologischer, aber auch pathologischer Prozesse, wobei auch ihre Relevanz in pharmakologischen Fragestellungen, beispielsweise durch eine mögliche Beeinflussung der Pharmakokinetik von Wirkstoffen, beachtenswert ist. Obwohl einzelne SLC-Transporter bereits gut beschrieben sind und das Interesse an einer umfangreichen Erforschung der ganzen Familie gerade in den letzten Jahren gestiegen ist, bleiben die Funktionen und Eigenschaften vieler SLCs weiter unklar.

Die hier vorgestellte Arbeit zielt zum einen darauf ab, bisher unbekannte Zusammenhänge zwischen Änderungen von Metabolitenspiegeln, Krankheiten und SLC-Transportern aufzudecken. Dafür werden Daten zu krankheitsbedingt veränderten Metabolitenkonzentrationen mit Daten über Beziehungen zwischen SLCs und Krankheiten kombiniert. Andererseits wird im selben Ansatz auch versucht potentielle unbekannte Substrate von SLC-Transportern zu beschreiben. Um diesen Plan umzusetzen, wurde mit der Datenanalyseplattform KNIME ein entsprechender Workflow geschaffen.

Der aus dieser Integration resultierende Datensatz zeigt bereits bekannte Assoziationen zwischen Krankheiten und SLCs, die bereits in unterschiedlichen Studien beschrieben wurden. Dies deutet darauf hin, dass auch in den anderen erhaltenen Ergebnissen relevante, noch nicht beschriebene Zusammenhänge enthalten sein könnten. Auch der zweite resultierende Datensatz, der potentielle neue Substrate von SLC-Transportern beschreibt, liefert vernünftige, wenn auch nicht unbedingt überraschende Ergebnisse.

CONTENTS

1 INTRODUCTION

The first chapter is intended to give an insight into the huge family of solute carriers and their fundamental importance in physiological and pathological processes. Section 1.1 first gives a short overview on the characteristics of solute carriers and their essential role in membrane transport. After that, section 1.2 briefly illustrates the relevance of this superfamily of transporters in the occurrence of diseases, their use as drug targets and their important role as parts of pharmacokinetic pathways, based on some significant examples. Following to that, section 1.3 highlights the importance of systematic research on solute carriers, before section 1.4 clarifies why metabolite-disease associations were used as the starting point for this work.

1.1 SOLUTE CARRIERS AND MEMBRANE TRANSPORT

A fundamental concept of life is that cells need to be separated from the extracellular space via biological membranes. These membranes function as physical barriers controlling the cellular uptake and release of compounds. Moreover, such barriers also exist in the intracellular space where they are required for the creation of intracellular compartments. Since the exchange of compounds through passive diffusion (a concept that has not yet been finally confirmed) is assumed to be restricted to rather small molecules with high lipophilicity, the transfer of most molecules relies on transmembrane proteins that mediate their transport. [1]

Such membrane transporters can be classified and grouped into different families on the basis of their function or sequence. The solute carrier (SLC) superfamily represents the largest family of membrane transporters as it currently comprises about 460 members classified into 65 families. [2] Other important groups of membrane transporters include the ATP-binding cassette (ABC) transporters, ion channels and ATPases. Transporters that are included

in an SLC family have a sequence similarity of over 20% in relation to at least one other transporter of the family, [3] while SLC families among each other often do not share a relevant homology in terms of their amino acid sequence. The membership of a transporter family in the SLC superfamily is based on the functional similarity of transporter proteins rather than their pylogenetic origin. [4] Because of this, the SLC superfamily represents a highly diverse class of membrane transporters, which means that their members differ in many of their features such as their structure, localisation, substrates, etc.

For the transport of molecules across biological membranes, SLCs use different so-called energy-coupling mechanisms, the most prominent of which, namely the facilitative transport and the secondary active transport, are discussed here. [5] In facilitative transport the electrochemical gradient of the compound itself is used to drive the transport, making it a passive transport. Carriers, in contrast to ion channels, which also use the electrochemical gradient of their substrates to facilitate their transportation, have specific binding sites and transport their substrates with a fixed stoichiometry in each transportation cycle. The transport capacity of channels is delimited by their open state probability but several dimensions higher than the capacity of carriers. [3] Besides that, SLCs also use secondary active transport, where substrates can be moved against their concentration gradient using the free energy provided by the simultaneous gradient-dependent transport of so-called coupled compounds. These are in most cases ions whose electrochemical gradient is generated by ATPases. [3] Secondary active transporters either carry their substrates in the same direction (symporters) or in the opposite direction (antiporters). [1] In contrast to transporters classified to the SLC superfamily, members of the other main group of membrane transporters, the ABC superfamily, mediate the efflux of compounds through primary active transport, which is directly powered by ATP hydrolysis. [6]

As mentioned before, membrane transporters do not only occur on the cell membrane but also on plasma membranes that create intracellular compartments. According to the paper "A substrate-based ontology for human solute carriers" by Meixner et al. where the authors describe the creation of a manually curated annotation of human SLCs, the most common localisation of annotated SLCs was the cell membrane with an occurrence of almost two thirds.

In regard to localisation on intracellular compartments mitochondria was found as most frequent. [7] Looking at the expression of SLCs in different tissues, there apparently is some kind of redundancy, meaning that a majority of SLCs can be found in a wide range of different tissues, whereas some other SLCs are highly limited to specific cell types. [8] The SLC6 and SLC18 families are a concrete example for this kind of tissue specificity since their occurrence is restricted to neurons where they control the levels of various neurotransmitters in the synaptic cleft. [2]

The spectrum of substrates transported by SLCs is quite diverse, including essential nutrients like amino acids, sugars and vitamins, as well as inorganic and organic ions, metabolites, drugs and more. [2] In their paper, Meixner et al. also depict the distribution of experimentally confirmed substrates of SLCs to different classes of compounds, showing that the transportation of ions and amino acids were the most common for annotated SLCs. [7] Another interesting point to consider is that SLC transporters obviously vary in their substrate specificity. While some SLCs mediate the transport of just a narrow range of substrates, such as for instance the amino acid transporters of the SLC7 family, or even are specific for only one specific substrate, others transport broad ranges of compounds from different chemical classes, like the transporters of the SLCO family (OATs). [6]

Despite their vital role in cellular functioning, due to their importance in regulating the transmembrane influx and efflux of compounds, the majority of SLCs remains chronically understudied. In fact, a substantial portion of 28% are considered as functional orphans given that they do not have any confirmed substrates. [7] Likewise to our in many cases limited knowledge of associated substrates, also other features of many SLCs are still unclear. An overview of the occurrences of unknown SLC characteristics can be found in Figure 1 which originates from the paper "A substrate-based ontology for human solute carriers" by Meixner et al. [7]

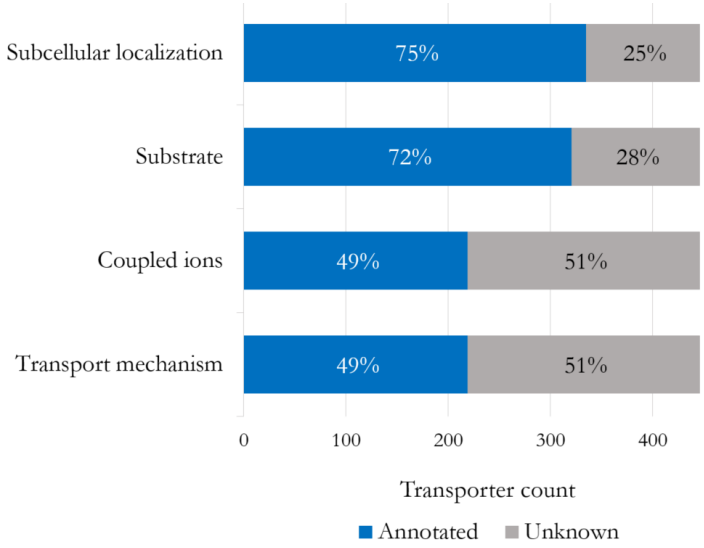


FIGURE 1: Frequencies of unknown annotations for 446 SLCs [7]

1.2 THE THERAPEUTIC RELEVANCE OF SLCs

Due to their crucial role in a large variety of physiological processes, it is not surprising that dysfunctions of solute carriers lead to imbalances in the disposition of their substrates and thereby contribute or even lead to a diverse spectrum of diseases. The relevance of this connection is clearly emphasized by the fact that mutations of around 190 SLC genes have been connected to human diseases in the Online Mendelian Inheritance in Man (OMIM) database up to now. [2]

1.2.1 SLCs in Mendelian and complex disorders

Mendelian diseases are defined as disorders that are caused by a single-gene mutation, with conspicuous familial inheritance patterns being typical of such monogenic diseases. In comparison to that, in the development of complex diseases, multiple genes and often also environmental factors are involved. Prominent examples of Mendelian diseases include Huntington's disease, sickle cell anemia and cystic fibrosis. Although monogenic diseases are numerous, most of them are comparatively rare. [9,10]

Interestingly, a considerable number of mutations in SLC genes that lead to dysfunctions or a deficient expression of the corresponding transporter have been verified to be causal for monogenic diseases. Such genotypic mutations are primarily induced by loss of function mutations of SLC transporters with very specific functions and a limited number of substrates. The resulting phenotypes of such mutations are highly diverse and can affect nearly every organic system. Some of them are harmless, whereas others have severe consequences for the individual. In their review "SLC Transporters as Therapeutic Targets: Emerging Opportunities" by Lin et al, the authors present a tabular overview of 84 SLCs whose mutations are associated with a total of 100 Mendelian diseases (see Supplementary Table S1 of the review), which was used in the validation of this work's results (see Discussion). [6]

Because the pathologic effects of Mendelian diseases are caused by specific alterations in the distribution of defined endogenous compounds these diseases provide distinct cause-effect relationships. By uncovering these connections between diseases and a changed handling of certain compounds, the generation of deeper insights into pathophysiological processes and also the development of new therapies could be promoted. [6]

As an example where a known SLC gene mutation has aided the development of a whole new class of drugs, the story of SGLT2-inhibitors can be mentioned. In familial renal glucosuria (aka diabetes renalis) glucose is excreted through the urine, although there are no increased blood levels of glucose. The cause for this rather benign trait, that usually does not lead to serious consequences, are mutations in the sodium-dependent glucose cotransporter (SGLT2 or SLC5A2) which is located in the proximal tubule and mediates the uptake of glucose from the primary urine. Evidence from the observation of patients with this disorder was an important argument to promote research efforts on the development of potent inhibitors of SGLT2. In their paper "Familial Renal Glucosuria and SGLT2: From a Mendelian Trait to a Therapeutic Target" [11] Santer and Calado refer to the phenotypic effects of the mutations that cause familial renal glucosuria as a "natural analogy" to the drug-induced inhibition of SGLT2. Both cases result in an increased excretion of glucose due to a reduced renal resorption, on the one hand caused by the reversible pharmacological effect of a drug that can be dosed, and on the other hand by a life-long reduced or even loss of function mutation of the transporter. Dapagliflozin was the first member of the family to be licensed for the treatment of diabetes in 2012 by the European Medicines Agency (EMA). Meanwhile SGLT2 inhibitors represent an important extension in the field of oral antidiabetics. Furthermore, several members of the family are now also applied in the treatment of heart failure and chronic kidney disease. [11,12]

As a matter of course, deficient SLC transporters, which are causal for monogenic diseases, would themselves be possible targets for the treatment of the disease. One possible approach could be the use of high-throughput screening methods to find compounds able to activate the transporter activity. Taking into account that most Mendelian diseases are rare (so-called orphan diseases), which is a major barrier in drug development, also drug repurposing

might be a reasonable approach to face the higher economical and logistical difficulties. [6]

Another important aspect to consider when talking about the variability of SLC-expressing genes are the population specific prevalences for different mutations. Schaller and Lauschke, who analyzed the genetic variability of SLC genes in different ethnic groups state in their paper "The genetic landscape of the human solute carrier (SLC) transporter superfamily", that they found an enormous share of 83% of all variants that were predicted to impact the SLC function had a certain amount of population specificity. This confirms that genetic variability is population-specific to a large extent, highlighting the relevance that ethnicity has in the predisposition of Mendelian diseases. Furthermore, also the efficacy and toxicity of drugs can be affected, due to the fact that an altered transporter function may have significant influence on the pharmacokinetics of a drug (see section 1.2.3 SLCs and pharmacokinetics). [13]

Besides their connection to a variety of monogenic diseases, numerous SLCs were linked to the appearance of a spectrum of common multicausal diseases. To identify associations between particular SLC variants and the development of such complex diseases, genome-wide association studies (GWAS) are an important tool. Techniques like this led to the discovery of a number of SLC transporter risk loci, whose polymorphisms may play a role as one among other multifactorial causes to contribute in the genesis of various diseases, including metabolic diseases like type 2 diabetes and gout, neurological diseases like depression and Alzheimer's disease, disorders connected with the immune system (e.g. asthma and inflammatory bowel disease) as well as cardiovascular diseases (e.g. high blood pressure) and cancer. [14] Table 1, adapted from the review "The SLC transporter in nutrient and metabolic sensing, regulation and drug development" by Zhang et al. presents links between SLC transporters and common diseases, which were used to validate the resulting associations from this work (see Discussion).

SLC	Human disease	Known substrates	Reference
SLC2A2	T2DM, Insulin resistance	Facilitated glucose transporter	[15]
SLC16A11	T2DM	Pyruvate	[16]
SLC30A8	T2DM, Insulin resistance	Zinc transporter 8	[15]
SLC6A1	Anxiety disorders	GABA transporter	[17]
SLC6A12	Schizophrenia	GABA transporter	[18]
SLC6A15	Depression	Branched-chain amino acids, particularly leucine, valine, isoleucine, and methionine	[19]
SLC30A10	Neurologic, hepatic, and hematologic disturbances	Manganese	[20]
SLC24A4	Alzheimer's disease	Calcium	[21]
SLC2A9	Gout	Urate	[22,23]
SLC16A9	Gout	Urate	[23,24]
SLC17A1	Gout	Sodium-dependent phosphate transporter 1 (uric acid)	[23,25,26]
SLC17A3	Gout	Urate	[22]
SLC22A11	Gout	Organic anion transporter 4	[24,27]
SLC22A12	Gout	Urate transporter 1	[23,26,27]
SLC4A7	Elevated blood pressure	Electroneutral Na^+/HCO_3^- cotransporter NBCn1	[28–31]
SLC6A13	Elevated blood pressure, chronic kidney disease (CKD)	GABA transporter	[32–34]
SLC8A1	Elevated blood pressure	Sodium-calcium exchanger 1	[35]
SLC12A1	Blood pressure variation	Kidney-specific sodium–potassium–chloride cotransporter	[36]
SLC12A3	Blood pressure variation	Renal thiazide-sensitive sodium-chloride cotransporter	[36]
SLC14A2	Elevated blood pressure	Urea transporter	[35]
SLC22A4/5	Elevated blood pressure	Gothioneine and carnitine	[37]
SLC24A3	Elevated blood pressure	K^+-dependent Na^+/Ca^{2+} exchanger 3	[35]
SLC35F1	Elevated blood pressure	-	[35]
SLC39A8	Elevated blood pressure	Zinc	[29]
SLC39A13	Elevated blood pressure	Zinc	[29]
SLC25A32	Blood pressure	-	[38]
SLC7A9	CKD	Cystine and neutral and dibasic amino acids	[33]
SLC34A1	CKD	Sodium–phosphate cotransporter	[33]
SLC22A2	CKD	Metformin, cisplatin, and lamivudine	[33,34]
SLC22A5	Asthma	Carnitine transporter	[39,40]
SLC30A8	Asthma	Zinc transporter 8	[41]
SLC22A23	Bronchodilator responsiveness in asthma	-	[42]
SLC25A15	Bronchodilator responsiveness in asthma	-	[43]

TABLE 1: Links of SLC transporters to common human diseases [14]

1.2.2 SLCs as drug targets

On account of the evident importance of SLC transporters in a variety of common diseases, especially in metabolic disorders, there are examples of approved drugs that exert their effects through an action on SLC transporters. According to the paper "A Call for Systematic Research on Solute Carriers" by César-Razquin et al. in 2015, there were 12 FDA-approved drug classes whose primary functional mechanism was an effect on SLC transporters. An overview of these approved drugs and further drugs in clinical development can be found in Table 1 of this paper. Furthermore, numerous other drugs interact with SLCs beyond their primary effect on a different target. Although, in many cases it is not entirely clear if these additional effects are of relevance for the pharmacological effect or potential side effects of the drug. [44]

Famous approved drug classes whose therapeutic effect is based on the inhibition of SLC transporters are important for various therapeutic areas. For example, the highly active class of loop diuretics works through an inhibition of $Na^+K^+2Cl^-$ symporters (aka SLC12A1) in the loop of Henle, whereas thiazides inhibit the Na^+Cl^- symporter (aka SLC12A3). For both classes this results in a decreased reabsorption of sodium and thus an elevated diuresis, making them important drugs in the management of high blood pressure and oedema caused by heart failure. The antagonism on members of the SLC6 transporter family represents another case for SLC transporters being targets of high importance for modern pharmaceutical treatment. NET (SLC6A2), DAT (SLC6A3) and SERT (SLC6A4) mediate the uptake of the monoamine neurotransmitters noradrenaline, dopamine and serotonin out of the synaptic cleft. These transport proteins serve as the primary targets of drugs like selective serotonin reuptake inhibitors (SSRI), serotonin-noradrenaline reuptake inhibitors (SNRI) and tricyclic antidepressants (TCA) that are used to treat depression and other neuropsychiatric disorders. SGLT2-Inhibitors used as oral antidiabetics, which were mentioned before, and inhibitors of urate transporter 1 (URAT1 aka SLC22A12), that are used for lowering uric acid levels due to their functionality to impede the reuptake of uric acid in the kidney, are further examples for SLC targeting drugs. [6]

1.2.3 SLCs and pharmacokinetics

Since a relevant number of SLCs has drugs among their substrates, they often play central roles in the absorption, distribution, and elimination of drugs by mediating the permeation across often poorly permeable cell barriers. Examples for such passages, where the transition through carriers is of importance are the transport across the blood-brain barrier, the hepatic uptake and biliary excretion as well as the tubular secretion in the kidneys. SLC-mediated transport is also important for the absorption of certain drugs in the intestine. Here it must be taken into account, however, that for many drugs passive membrane diffusion may also play a significant role for the permeation into the intestinal epithelium. [45]

Since the therapeutic effect of drugs highly depends on pharmacokinetic processes, changes in the functionality of SLC transporters can also influence the success of a therapy. These changes can be caused by polymorphisms or also by allosteric inhibitory or activatory effects. Thus, carrier-mediated transport can also entail drug-drug or nutrient-drug-interactions. A suitable example for an interaction of clinical relevance is the inhibition of OATP1B1 (aka SLCO1B1), which mediates the uptake of several commonly used drugs into hepatocytes. Some popular drugs, like for instance cyclosporin, act as inhibitors of this transporter and thereby induce elevated plasma concentrations of OATP1B1 substrates. Given that statins are important substrates of OATP1B1, their reduced hepatic uptake increases the risk of serious side effects like myopathy or even rhabdomyolysis. [46] As previously described, also polymorphisms may influence the pharmacokinetic functions of SLCs. Such polymorphisms have also been found for OATP1B1, leading to higher exposures of different drugs, including statins. Although these specific genomic features might be relatively rare, quite a number of patients could be affected, considering the high prevalence of a medication with statins. [47]

Another relevant example concerns metformin, the most frequently used substance to treat type 2 diabetes. The organic anion transporter 1 (OCT1 or SLC22A1) is known to take part in the absorption of metformin from the intestine, in its elimination through mediating the uptake of metformin into

hepatocytes, and for its renal excretion. A relevant proportion of patients medicated with metformin develops sometimes severe gastrointestinal side effects, which affect the adherence to the therapy. Several studies were able to confirm polymorphisms of SLC22A1, which lead to a reduced function of the transporter, as predisposing conditions for metformin intolerance. Moreover, some commonly used drugs (e.g. verapamil) operate as OCT1 inhibitors and increase the probability for an intolerance in addition. [48]

Because many therapeutic compounds have been identified as substrates or affect the activity of transporters in general, and polypharmacy is an increasingly important issue in aging societies, the possibility of drug-drug interactions must be considered. For this reason the FDA emphasises the high relevance of screening for transporter-mediated interactions during drug-development in their recent guidelines. [49]

1.3 SYSTEMATIC RESEARCH ON SLCs

As discussed in chapter 1.2, the family of solute carriers is of high prominence not only in physiological processes, but also in the formation and further development of diseases. Although some important drug classes that use SLCs as their targets have been developed, the therapeutic potential of this wide range of possible drug targets does not seem to be utilized adequately at present, especially when comparing SLCs with other membrane protein superfamilies like for instance the G-protein coupled receptors (GPCR), for which broader approaches in research have been realised. The thesis that the low number of SLCs to be intentionally targeted by approved drugs is above all caused by an underexploration of the superfamily is also supported by the fact that small-molecule inhibitors have been identified for most of the SLC transporters that have been investigated properly, making solute carriers a family with a good druggability in general. [44]

One principal problem when talking about the progress that has been achieved in SLC research, is that a majority of publications deals with a relatively small proportion of all SLC transporters. César-Razquin et al. took a closer look at this issue in their paper "A Call for Systematic Research on Solute Carriers" and found that SLCs bear by far the greatest asymmetry in the number of publications for single proteins compared to other gene families. Some members of the superfamily are explored to a high extent, while most others did not receive much research attention yet. [44]

Now why does a systematic approach in SLC research make sense and what findings can be expected? First of all, the similarities between SLCs that could be determined in expansive deorphanisation attempts may indicate relations between the characteristics of different SLCs, as for example in their transport mechanism or in their substrate specificity. [50] Another point is that through studying whole families of proteins, the acquired knowledge on one transporter and developed tools for its characterization could promote the research prospects on other members of the family. This approach has already been rewarding when it was applied to other gene families, leading to an elucidation of their structure and accessibility as potential drug targets. [44]

There seem to be different reasons why whole-scale research efforts on SLCs were rather restrained for a long time. First, investigation is often hindered due to a lack of applicable research tools like functional assays. A second problem: There are repositories on the basic properties of SLC transporters, but such collections for other data such as information on expression profiles, connections between gene variants and diseases or collections of reagents have long been missing. Although the SLC gene nomenclature system was introduced by Hediger in the 1990s, the use of multiple nomenclatures also still poses a problem. Furthermore, the holistic effects of SLC transporters in human physiology and a substrate specificity that is often low contributed to the supposed unsuitability of SLCs as drug targets. [50,51]

In order to enhance the systematic research on SLCs and to increase the utilisation of this transporter superfamily as drug targets, the RESOLUTE (Research Empowerment on Solute Carriers) consortium was founded. Different members from academia and pharmaceutical industry participate in the project that is funded by the European Innovative Medicines Initiative (IMI), the European Federation of Pharmaceutical Industries and Associations (EFPIA) and the EU. [52] The key intentions of this venture are to enable scientific investigation via providing research tools to deorphanise SLCs in a systematic approach, to establish a reliable source for assured information of SLC transporters and to generate functional assays for prioritized SLCs to further clarify their biological roles. In the 5 year term of this project that started in 2018, all the results and also methods are made openly accessible for the scientific community in order to boost the progress in SLC research. [50]

1.4 METABOLITE-DISEASE ASSOCIATIONS - A PROMISING STARTING POINT

One considerable problem with SLC transporter research is that the focus is often just placed on their activities in the transportation of drugs. Although it is understandable to address concerns over the safety of drugs that might arise from their interaction with SLC transporters, clarifying their endogenous roles must not be forgotten. A better understanding of the physiological function of SLCs might open new insights into the mechanisms of various diseases and also on adverse drug reactions, given that a relevant number of side effects is caused by interactions of drugs with SLC transported metabolites. [53]

Now what exactly is meant by "metabolites" in this context? The term refers to small molecules that are either transformed or generated in metabolic processes in order to produce energy from nutrients or to build up proteins, lipids, nucleic acids and carbohydrates. All of which are compounds needed to retain correct functioning of single cells as well as the whole organism. In the wide field of metabolic processes, the transport of substances inside and outside of cells has a crucial role to maintain an accurate functionality. The high relevance of SLCs in metabolism, owing to their function to transport a broad spectrum of metabolites, is further affirmed by their expression profiles, as SLC transporters can be found in metabolic organs like the liver, kidney, intestine and also the brain in particularly high levels. [54,55]

A fundamental condition for this work is that the presence of certain metabolites and their accumulation or reduction can be associated with the occurrence of specific diseases, respectively. Very quickly we associate diseases like cancer with an altered metabolism, but in fact it can be assumed that every disease is connected with dysregulations of metabolic pathways, resulting in an alteration of metabolite levels. Because these changes in the concentrations of metabolites can be seen as signatures of diseases, Cheng et al. collected the information on metabolite-disease associations and created "MetSigDis", which should provide a comprehensive resource for relationships between diseases and related metabolites (see section 2.1.1 for further information on "MetSigDis"). [56]

In this work the data from "MetSigDis", which depicts a variety of disease associated alterations of metabolites, is combined with the SLC transporters that are associated with these diseases and besides from that also exhibit transporting activity for the corresponding metabolites. Figure 2 gives a graphic overview of the described combination and the main data sources that are used. To realize the integration of data from different sources the KNIME analytics platform, an open source data analytics tool, was used to create a workflow capable of this task.

There were two basic objectives for this work. The first aim was to gain information on previously unknown associations between metabolites, diseases and SLC transporters from their connection, which was done via two different approaches. The second aim addressed in this work was whether an integration of data from metabolites whose alterations are associated with corresponding diseases and disease-SLC data could also be considered as a source for potential new substrates of the covered SLC transporters. The implementation of this idea via a similarity search is based on the assumption that molecules that are structurally similar to confirmed substrates may also be transported by the SLC transporter.

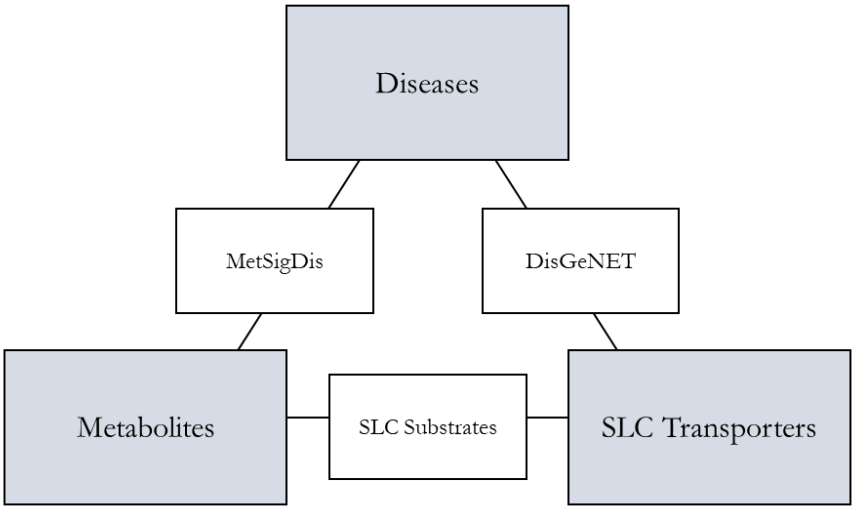


FIGURE 2: Combination of disease associated alterations of metabolites with corresponding SLC transporters

2 METHODS

In this chapter the methods used to achieve the objectives of this work, above all the goal of linking disease-related changes in metabolite levels with SLC transporters, are outlined. First, section 2.1 provides an overview and brief description of the data sources and tools used in this work. Then, in section 2.2, the construction and functionality of the workflow is explained in more detail.

2.1 DATABASES AND TOOLS

2.1.1 MetSigDis

MetSigDis is a manually curated resource that covers metabolic signatures for various diseases and thereby forms the starting point for this work. According to the authors of the paper in which this library was introduced [56], MetSigDis contains 6849 manually curated relationships between 2420 metabolites that correlate with 129 diseases. Included information derives from reviewed PubMed literature about metabolomic assays which identified certain metabolites as signatures for specific diseases. Besides the information on metabolite-disease associations, every entry comprises the observed species (in a large part Homo sapiens), the analytical platform used in the metabolomic study (in most cases NMR, GC-MS or LC-MS), the tissue as well as the metabolite alteration (if increased or decreased) and a reference to the original paper. A great benefit for the utilization of this data set in the workflow was that besides the disease and metabolite names also the codes of standardised vocabularies are included (e.g. Disease Ontology IDs (DOIDs) and Human Metabolome Database (HMDB) codes). This eases both the analysis of the MetSigDis data and also the connection with other data sets substantially. [56]

The relationships contained in MetSigDis are freely available and can be downloaded from their webpage (http://www.bio-annotation.cn/MetSigDis/).

Furthermore, the web interface also provides interactive visualisations of the metabolite disease network (MDN) and the human disease network (HDN) that the authors created using the collected data. In addition, a search engine is also provided from which the detailed information on each covered metabolite-disease association can be retrieved. [57] For this work, the data set was downloaded on the 13.10.2021, but unfortunately it did not contain all the information that was promised in the paper. For further information on that issue see the data limitations section in the discussion part.

2.1.2 DisGeNET

The DisGeNET platform contains associations of genes and variants to human diseases. DisGeNET v7.0, which is the current version of the platform (accessed on 28.10.2021), covers over 1.1 million gene-disease associations between 21 671 genes and 30 170 diseases or other abnormalities. The contained gene disease associations (GDA) and variant disease associations (VDA) derive from different types of source databases and include data from expert curated repositories, animal models, as well as from GWAS catalogues. For detailed information on the original data sources see their webpage (https://www.disgenet.org/dbinfo/). [58] For the application in this work, only curated GDAs are used, which originate from UniProt [59], the Comparative Toxicogenomics Database (CTD) [60], Orphanet [61], the Clinical Genome Resource (ClinGen) [62], the Genomics England PanelApp [63], the Cancer Genome Interpreter (CGI) [64] and PsyGeNET [65].

DisGeNET was used to retrieve the associated genes for the diseases covered in the data table from MetSigDis. There are multiple ways to access the information of this database, for the purpose of this work the accession through the DisGeNET REST API was chosen (see section 2.2.2).

2.1.3 SLC substrates list

For the pivotal task to compare the metabolites with confirmed SLC substrates the SLC Substrates Structure Tool is used. The data set derives from the KNIME workflow SLC_substrates of Catrin Gabrail [66], which was edited by Daniela Digles (unpublished work). The workflow automatedly collects annotated SLC substrates from several databases and creates a table containing substrates for the respective SLC transporters. The sources from which the substrates information is obtained are the Metrabase [67], Bioparadigms [3,68], TCDB [69], CeMM [70], UCSF-FDA [71], ChEMBL [72] and the IUPHAR/BPS Guide to Pharmacology [73].

In order to make the workflow compatible with mine, it was necessary to make some adjustments to this workflow. One change that was made was to remove some unclear designations of substrates and on the other hand to separate rows that contained several substrates together in one cell into individual rows. These changes were made in the databases area of the workflow and affected for instance the sources TCDB and IUPHAR/BPS Guide to Pharmacology. Furthermore, I created a table with InChI codes that were missing for numerous substrates of the data set (see Supplementary Table 3 in the appendix) and then connected it to the resulting data set of the SLC_substrates workflow. Since the stereochemistry was relevant for the application in my workflow, I changed the settings of the Standardizer_RDKit_Component, created by Jennifer Hemmerich [74], so that the stereochemistry is not removed, but only cleaned.

2.1.4 Metabolite Structures data set

The Metabolite Structures data set used to add SDF-formatted chemical structures and identifiers (e.g. InChI codes) to the respective metabolites derives from the Human Metabolome Database (HMDB). The HMDB contains comprehensive metabolomic data about human metabolites and covers their biological roles, concentrations, associations with the metabolic pathways of

diseases, reference spectra et cetera. The data is available to the public and can also be searched on their web page. Data sets with different emphases are provided on the HMDB download page. For the purposes of this work the Metabolite Structures data set version 4.0 was used. [75]

2.1.5 RESOLUTE SLC list

This data set, which derives from the RESOLUTE project, contains a collection of 446 SLC transporters and provides several identifiers (e.g. HGNC ID, Entrez Gene ID, UniProt accession number etc.) as well as their protein sequence, known isoforms and some other information. In this work, the RESOLUTE SLC list is used to filter out all metabolite-disease associations that are not related to SLC transporters.

2.1.6 KNIME Analytics Platform

KNIME [76] is a freely available and open-source data analytics workflow system. One major advantage of the platform is its graphical user interface, which enables even users without programming knowledge to create their own workflows. Such a workflow consists of assembled nodes that perform specific functionalities and whose settings can be adapted to the respective use case. These single code units can be linked through a connection line in order to carry data or models from one node to the next. Instead of running the whole workflow, nodes can also be executed selectively and intermediate results can be checked for each step. For a better understanding of the structural organisation of complex workflows, the use of metanodes and components, which encapsulate separate workflow-steps, is extremely helpful. [77,78] Because of the visual representation and the intelligible way to set up workflows through knitting together preimplemented nodes, KNIME workflows are relatively easy

to understand and to construct. Furthermore, the open-source philosophy allows researchers and programmers to develop and provide their own extensions to the scientific community and due to that, a multitude of extensions for a variety of research interests is available. [78,79]

This work was developed using KNIME version 4.5.0. The requirements to run the workflow are KNIME Distance Matrix, KNIME Javasnippet, MOE Extensions for KNIME, KNIME REST Client Extension, KNIME JSON-Processing, RDKit Nodes Feature, KNIME Base Chemistry Types & Nodes, BioSolveIT Interfaces, KNIME Excel Support, KNIME-CDK, KNIME Quick Forms, Schrödinger Extensions for KNIME, Vernalis KNIME Nodes and KNIME Python Integration.

2.2 THE WORKFLOW

In order to create a better overview, the workflow can be divided into different steps (graphically depicted in Figure 3), which are described in more detail in the following sections. The first two steps are about obtaining the required data from MetSigDis and DisGeNET. After adding structural information and excluding irrelevant information in step 3, two different approaches are taken to connect the gained information on metabolite-disease associations and disease-gene associations with the SLC substrates list in order to accomplish the wanted interconnection of metabolites, diseases and SLC transporters. One way to realise this objective was to use similarity search, this approach is covered in step 4. The alternative approach, which is illustrated in step 5, uses joining to achieve the corresponding goal.

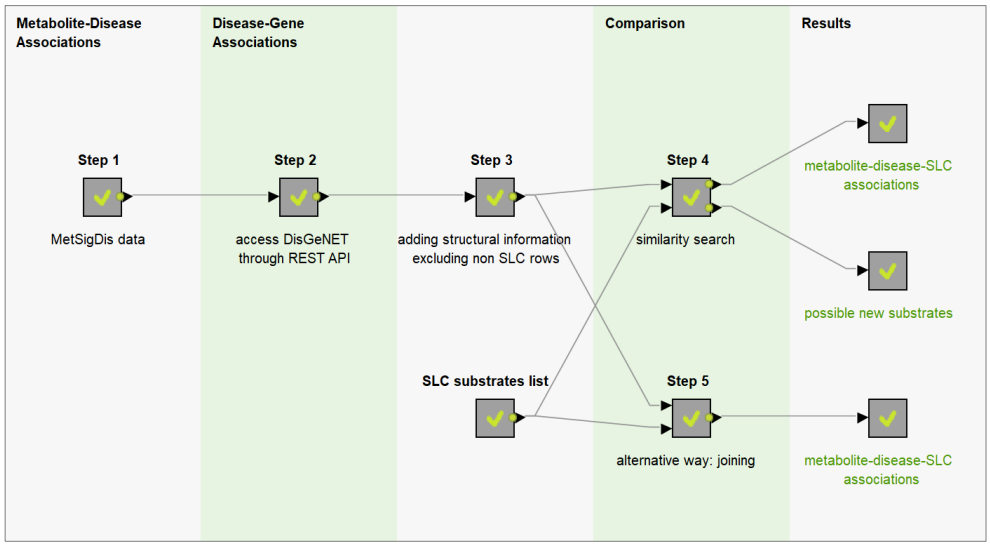


FIGURE 3: Schematic overview of the workflow and its subdivision into the individual steps

2.2.1 Step 1: Download and processing of MetSigDis data

The easiest way to get data into a KNIME workflow is to read the data from files that are located on a local file system. Different kinds of reader nodes are available for the respective file formats. As mentioned before, the data from MetSigDis can easily be downloaded from the webpage, which was done in this case. After that, the data is then imported into the workflow using a *CSV Reader node*. Since it is useful to be able to share the workflow together with the data sets included, the settings of the reader node under the "Read from" item are set to "Relative to Current workflow data area".

Now that the data is available in the workflow, some simple data processing steps are taken, including the exclusion of data rows that contain information on non-human organisms like drosophila or mouse, using the basic *Row Filter node*. In the next step, irrelevant columns of the data table are excluded by applying the *Column Filter node*. Because some cells of the DOID column and the disease name column contain multiple disease IDs and names, which would hinder a correct attribution in the following step 2 of the REST API call, these cells need to be split, which is done using *Cell Splitter nodes* and an *Ungroup node*. In a further step, the disease IDs, which all have the prefix "DOID:", are standardized with the *String Manipulation node* so that they only consist of the digits. Rows with DOID column cells containing the value "NA" are excluded. The nodes for executing all these data processing steps are shown in Figure 4, while the organization of the resulting table can be seen in Figure 5.

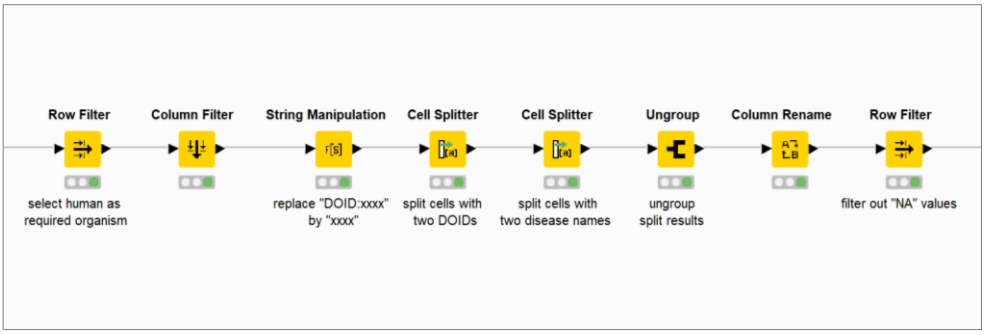


FIGURE 4: Data processing steps after reading the MetSigDis data into the workflow

Row ID	Tissue	Signature Name	Signature ID	Metabolite Alteration	DOID	Disease Name
Row437	blood	5-methoxytryptophan	HMDB02339	Increased	3910	lung adenocarcinoma
Row438	blood	linoleic acid	HMDB00673	Increased	3910	lung adenocarcinoma
Row439	blood	oleic acid	HMDB00207	Increased	3910	lung adenocarcinoma
Row440	prostate tissues	sorbitol	HMDB00247	Increased	1470	major depressive disorder
Row441	prostate tissues	uric acid	HMDB00289	Increased	1470	major depressive disorder
Row442	prostate tissues	azelaic acid	HMDB00784	Increased	1470	major depressive disorder
Row443	prostate tissues	quinolinic acid	HMDB00232	Increased	1470	major depressive disorder
Row444	prostate tissues	hippuric acid	HMDB00714	Increased	1470	major depressive disorder
Row445	prostate tissues	tyrosine	HMDB00158	Increased	1470	major depressive disorder
Row446	Blood	HR-	?	Decreased	1612	breast cancer
Row447	Blood	HR+	?	Increased	1612	breast cancer
Row448	bladder	Alanine	HMDB00161	Increased	11054	Bladder Cancer
Row449	bladder	Glutamate	HMDB03339	Increased	11054	Bladder Cancer
Row450	bladder	Glutamine	HMDB00641	Increased	11054	Bladder Cancer
Row451	bladder	Valine	HMDB00883	Increased	11054	Bladder Cancer
Row452	bladder	Isoleucine	HMDB33923	Increased	11054	Bladder Cancer
Row453	bladder	Leucine	HMDB00687	Increased	11054	Bladder Cancer
Row454	bladder	Phenylalanine	HMDB00159	Increased	11054	Bladder Cancer
Row455	bladder	Tyrosine	HMDB00158	Increased	11054	Bladder Cancer
Row456	Blood	myristic	?	Decreased	9352	type 2 diabetes mellitus
Row457	Blood	palmitic	?	Decreased	9352	type 2 diabetes mellitus
Row458	Blood	stearic acid	HMDB00827	Decreased	9352	type 2 diabetes mellitus
Row459	Blood	pyroglutamic acid	HMDB00267	Decreased	9352	type 2 diabetes mellitus

FIGURE 5: Extract from the resulting table after processing the MetSigDis data. From left to right: The tissue where the metabolite levels were measured, the altered metabolite with its HMDB code and the DOID plus the disease in which the metabolite level is changed

2.2.2 Step 2: Accessing DisGeNET through its REST API

To access information from DisGeNET, the library offers several ways. One way is that required data can be downloaded from the websites download-section, which was also done in an earlier attempt while developing this workflow. But since working with the downloaded data set has some disadvantages, such as the point that the data set consists mainly of information irrelevant for this use case, this approach was abandoned and the more accurate way of simply retrieving the desired data via REST API calls was chosen.

In a nutshell, what exactly is a REST API call? Basically, an Application Programming Interface (API) allows users to interact with a web service that provides programmatic access. By using API calls, specific data can be retrieved from a database without having to use the graphical interface or to download the whole data set. Many of these APIs conform to the constraints of the Representational State Transfer (REST) architecture style, which consists of rules that describe how web sources are defined and addressed. To perform API calls and to receive the required information, it is important to maintain a correct syntax in the HTTP request. KNIME offers the opportunity to perform data

retrievals through an API directly in the workflow, using the *REST Web Service nodes*. [80,81] In case of this workflow, the *GET Request node* was used.

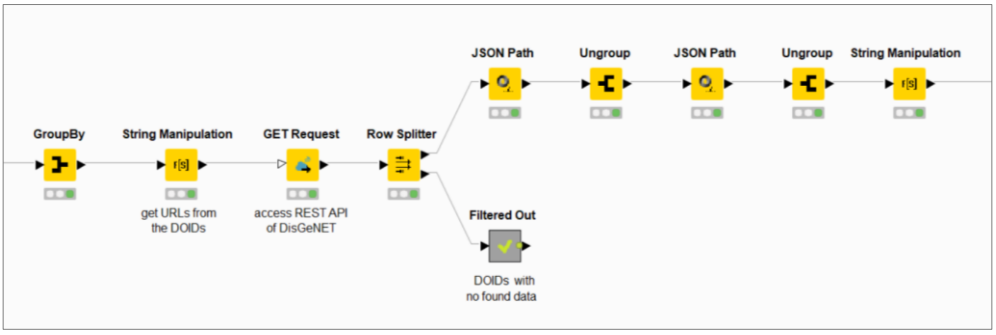


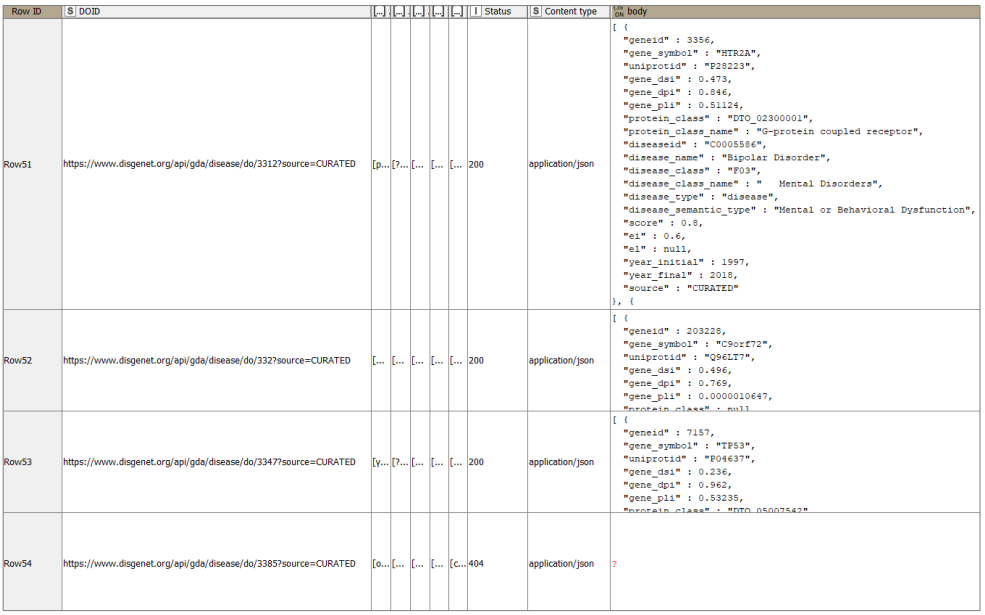
FIGURE 6: Accessing DisGeNET through its REST API and transforming the JSON formatted data into separated readable columns

The tasks described in the following paragraphs can also be viewed in Figure 6, which depicts the structural organisation of the workflow step 2.

In order to perform an API call only once, even if the disease (represented by its DOID) is mentioned several times in the table, the first step is to group the data table on the DOID column and keep all the other columns as lists. After that, the *String Manipulation node* is used to create the required URLs, consisting out of the base URL of the DisGeNET API server (https://www.disgenet.org/api/), the part for requesting the gene-disease associations through using the disease ontology identification (gda/disease/do/), the DOID itself (xxxx) and the requirement to obtain just information from curated sources (source=CURATED). All the information how to access the wanted data from DisGeNETs REST API with the correct URL syntax is provided in a clear manner on the databases platform. [82]

In the next step, the *GET Request node* is used to access the REST API of DisGeNET with the created URLs. Therefore, some settings of the node need to be adjusted in its configuration window. In the connections settings tab, the URL column is chosen and the timeout(s) are raised from the very low default setting, which increases the probability that a request will be successful. Furthermore, respective request headers are added to provide the necessary information for the API to tailor the response. After running the *GET Request node*, a new column including the HTTP response codes "200" for a successful

request and "404" for no found result appears in the output table. The retrieved information is presented in JSON format and contained in another column (see Figure 7).



Row ID	S DOID	[...]	[...]	[...]	[...]	[...]	[...]	I Status	S Content type	{} body

Row51	https://www.disgenet.org/api/gda/disease/do/3312?source=CURATED	[p...	[?...	[...	[...	[...	[...	200	application/json	[{
										"geneid" : 3356,
										"gene_symbol" : "HTR2A",
										"uniprotid" : "P28223",
										"gene_dsi" : 0.473,
										"gene_dpi" : 0.846,
										"gene_pli" : 0.51124,
										"protein_class" : "DTO_02300001",
										"protein_class_name" : "G-protein coupled receptor",
										"diseaseid" : "C0005586",
										"disease_name" : "Bipolar Disorder",
										"disease_class" : "F03",
										"disease_class_name" : " Mental Disorders",
										"disease_type" : "disease",
										"disease_semantic_type" : "Mental or Behavioral Dysfunction",
										"score" : 0.8,
										"ei" : 0.6,
										"el" : null,
										"year_initial" : 1997,
										"year_final" : 2018,
										"source" : "CURATED"
										}, {

Row52	https://www.disgenet.org/api/gda/disease/do/332?source=CURATED	[...	[...	[...	[...	[...	[...	200	application/json	[{ "geneid" : 203228, "gene_symbol" : "C9orf72", "uniprotid" : "Q96LT7", "gene_dsi" : 0.496, "gene_dpi" : 0.769, "gene_pli" : 0.0000010647, "protein_class" : null
Row53	https://www.disgenet.org/api/gda/disease/do/3347?source=CURATED	[y...	[?...	[...	[...	[...	[...	200	application/json	[{ "geneid" : 7157, "gene_symbol" : "TP53", "uniprotid" : "P04637", "gene_dsi" : 0.236, "gene_dpi" : 0.962, "gene_pli" : 0.53235, "protein_class" : "DTO_05007542"
Row54	https://www.disgenet.org/api/gda/disease/do/3385?source=CURATED	[o...	[...	[...	[...	[...	[c...	404	application/json	?

FIGURE 7: Output table after executing *GET Request node*. From left to right: The URL for each API call, the listed content columns, the response status and the retrieved data in JSON format can be seen.

In the subsequent step, all rows with DOIDs for which no results were found are filtered out, using the *Row Splitter node*. Following to that, the new data needs to be transformed into a readable table, therefore *JSON Path nodes* are used.

Because there are multiple genes that are associated with a single disease, the resulting output column contains the JSON-formatted information about multiple genes for each requested disease. Therefore, an intermediate step is necessary to separate the acquired gene-disease associations. After ungrouping all other rows that were previously grouped, another *String Manipulation node* is used to return from the URLs to the plain DOIDs. Afterwards, not needed columns are filtered out. The organization of the resulting table can be seen in Figure 8.

Row ID	DOID	Signature Name	Signature ID	Metabolit...	Tissue	...	GeneID	Gene_Symbol	UniprotID	DiseaseID	Disease_Name
Row...	14330	Significant Differ...	?	Increased	urine	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	14330	aminobutyric acid	?	Increased	urine	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	14330	uric acid	HMDB00289	Decreased	Plasma	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	14330	glutathione	HMDB00125	Increased	Plasma	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	14330	2-hydroxy-2-deo...	?	Increased	Plasma	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	14330	hypoxanthine	HMDB00157	Decreased	Plasma	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	14330	uric acid	HMDB00289	Decreased	Plasma	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	14330	xanthosine	HMDB00299	Decreased	Plasma	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	14330	xanthine	HMDB00292	Decreased	Plasma	...	9045	RPL14	P50914	C0030567	Parkinson Disease
Row...	1470	sorbitol	HMDB00247	Increased	prostate tissues	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	uric acid	HMDB00289	Increased	prostate tissues	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	azelaic acid	HMDB00784	Increased	prostate tissues	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	quinolinic acid	HMDB00232	Increased	prostate tissues	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	hippuric acid	HMDB00714	Increased	prostate tissues	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	tyrosine	HMDB00158	Increased	prostate tissues	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	5- hydroxyindole...	?	Decreased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	tryptophane	HMDB00929	Decreased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	kynurenine	HMDB00684	Decreased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	xanthine methion...	?	Decreased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	homovanillic acid	HMDB00118	Decreased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	5-hydroxyindolea...	HMDB00763	Increased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	tyrosine/ hydroxy...	?	Increased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	methionine	HMDB00696	Increased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	glutathione	HMDB00125	Increased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder
Row...	1470	xanthine	HMDB00292	Increased	cerebrospinal...	...	6532	SLC6A4	P31645	C0011581	Depressive disorder

FIGURE 8: Kept new columns from DisGeNET: GeneID, Gene Symbol, UniprotID, DiseaseID and Disease Name

2.2.3 Step 3: Adding metabolite structure and identifiers

With the intention to make the MetSigDis-derived metabolites comparable to the SLC substrates list, standardised structural information for these molecules has to be introduced to the data set. For this purpose, the data set is joined with the Metabolite Structures data set (see Figure 9), which was downloaded from the Human Metabolome Database (HMDB) and read into the workflow with a *SDF reader node*. After adjusting the HMDB codes of the two data sets to the same number of digits (using *String Manipulation node*), the two tables are then joined via the HMDB codes (Inner Join – to get all matching rows).

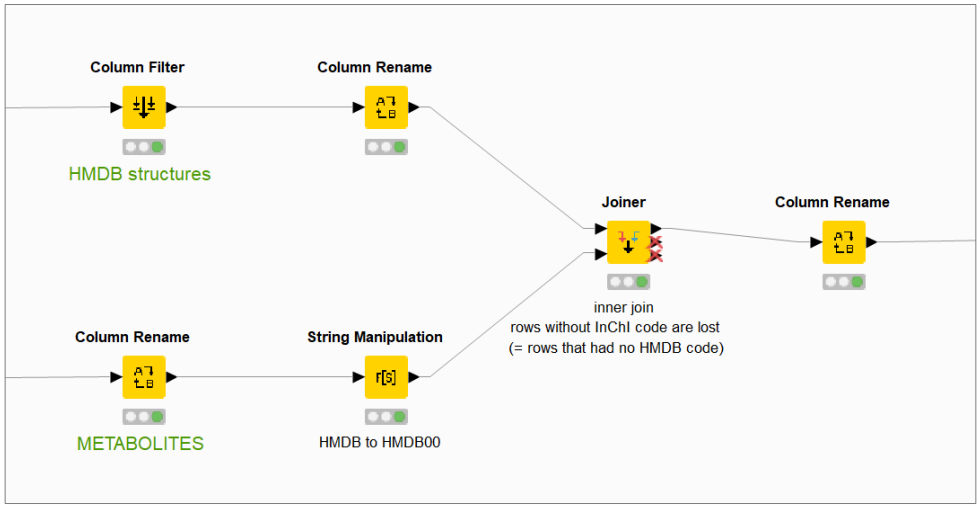


FIGURE 9: Adding metabolite structure and identifiers

After joining and filtering, the resulting table now includes several new columns covering molecular characteristics and identifiers for the metabolites, such as an SDF depiction of the molecule, its molecular formula, InChI code and InChI key. For a representation of the resulting table see Figure 10, which shows an extract from the output after step 3.

Subsequent to this step of adding structural information for the metabolites, the quite large data set should be reduced. Information which is not of interest for the objectives of this work should therefore be removed. For this purpose, the data is joined (via Entrez Gene ID) with the RESOLUTE SLC list, which generates a downsized data table that only consists of metabolite-disease rows that contain SLC transporters.

Row ID	Metabolite_Molecule	Metabolite	MetaboliteID	InChI code	Disease_MetSigDis	DOID	Metabolite...	Tissue	HGNC Symbol	Entrez GeneID
Row9547		Trigonelline	HMDB0000875	InChI=1S/C7...	lung cancer	1324	Decreased	Urine	SFXN1	94081
Row9548		Trigonelline	HMDB0000875	InChI=1S/C7...	lung cancer	1324	Decreased	Urine	SLC4A7	9497
Row9549		xanthurenic acid	HMDB0000881	InChI=1S/C1...	Parkinson's disease	14330	Decreased	urine	SLC18A2	6571
Row9550		xanthurenic acid	HMDB0000881	InChI=1S/C1...	Parkinson's disease	14330	Decreased	urine	SLC6A3	6531
Row9551		xanthurenic acid	HMDB0000881	InChI=1S/C1...	Parkinson's disease	14330	Decreased	urine	SLC2A14	144195
Row9552		xanthurenic acid	HMDB0000881	InChI=1S/C1...	Parkinson's disease	14330	Decreased	urine	SLC30A10	55532
Row9553		Valine	HMDB0000883	InChI=1S/C5...	prostate cancer	10283	Increased	prostate t...	SLC22A3	6581
Row9554		Valine	HMDB0000883	InChI=1S/C5...	prostate cancer	10283	Increased	prostate t...	SLC5A5	6528

FIGURE 10: Extract of the resulting table containing metabolite columns, disease columns and SLC columns

2.2.4 Step 4: Similarity Search

After the creation of this output table, which is now used in the following analysis steps, the actual core step of the workflow follows. In the process of the similarity search a set of molecules, the query table, is compared with a reference table of molecules. In fact, not the molecules themselves, but their molecular

fingerprints that comprise the structural information of each molecule encoded as a series of bits are used. The course of the process is that for each row of the query table, the reference table is searched for rows that meet the similarity criteria. As a result the user then gets the information if a similar molecule has been found and, if so, the number of the found row as well as the degree of similarity are shown in separate columns. The similarity value can range from 0 to 1, with a similarity value of 1 indicating that the two molecules are identical.

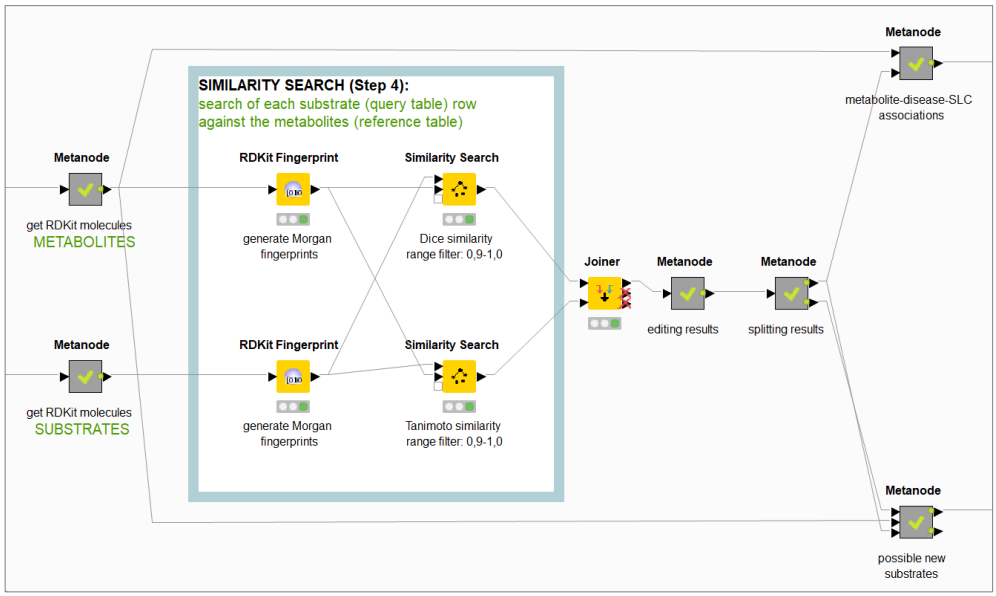


FIGURE 11: Organisation of step 4: Creation of RDKit molecules, translation into Morgan fingerprints, actual similarity search (Dice and Tanimoto similarity), postprocessing steps, splitting the results after manually analysing them.

In case of this workflow, the two data tables to be compared are on the one side the table that descends from the MetSigDis data set and on the other side the data table from the SLC substrates list. Figure 11 shows how the steps of preparation, similarity search and the following processing steps are organised. For the actual search the *Similarity Search node* takes each single row from the SLC substrates list with the verified substrates and searches the MetSigDis-descending table. But before this similarity search can be run, some preparation steps have to be completed (see Figure 12). First of all, both tables are grouped over the molecule columns, while all other columns are bundled up into lists. This step ensures that the subsequent search runs more efficiently and therefore faster, leading to a data table that does not get out of hand in terms of size.

Following to that, the *Molecule Type Cast node* is then used to change the InChI code column from type string to type InChI, which is necessary for the *RDKit From InChI node* to create RDKit molecules out of the InChI codes. In the settings window of this node the options of sanitizing the molecule and removing hydrogens are checked in order to achieve a certain level of standardisation before performing the similarity search.

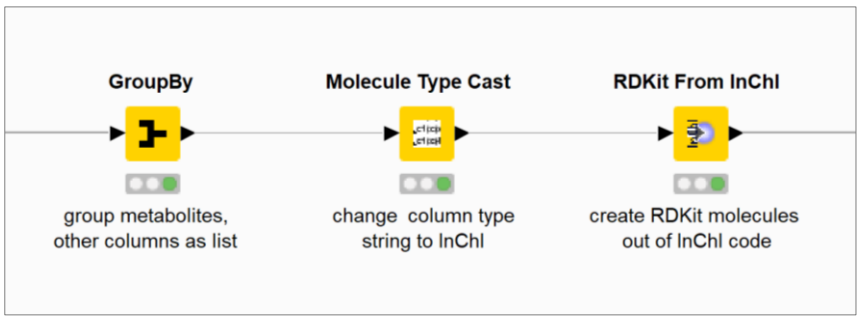


FIGURE 12: Preparation steps before the similarity search performed for both the metabolites and the substrates data set

Subsequently to that, the structural information of the RDKit molecules from both tables has to be converted into bit-based fingerprints so that the *Similarity Search nodes* can process the presented information. For this purpose, the *RDKit Fingerprint node* is the means of choice. The settings of this node offer the possibility to choose between several types of fingerprints, for this work the choice fell on Morgan fingerprints. [83] In general, fingerprints provide a computational representation of chemical structures that consists of single bits which encode chemical and molecular features of the molecule. Each bit stands for a molecular feature and holds the information if the defined feature is present "1" or absent "0". [84] Morgan fingerprints are circular fingerprints that use the Morgan algorithm to give each molecule a unique sequential numbering. The neighborhood of each atom is taken into account, whereby the selected radius is decisive, as this in turn determines the size of the encoded fragments. [85] Although the use of MACCS fingerprints was also tried out in earlier attempts, Morgan fingerprints obviously outperformed them in terms of delivering correct results.

For the actual step of the similarity search, the corresponding node offers some setting options (see Figure 13). On the one hand, in the distance selection, the user can choose between different algorithms for calculating the molecular similarity between the individual fingerprints. Both chosen methods for the calculation, the Dice similarity coefficient and the Tanimoto algorithm have proven to be very suitable for this kind of analysis. How the coefficients are calculated can be seen from the formulas in Figure 14. [86] In the search option settings it is selected for this use case that the similarity and not the distance should be calculated, in the neighbor selection the option nearest (most similar) is chosen. A range filter is used to reduce useless output and to get only relevant results with a similarity between 0.90 and 1.00. The neighbor count must be set higher than in the default setting (=1) to enable multiple results being obtained for single queries.

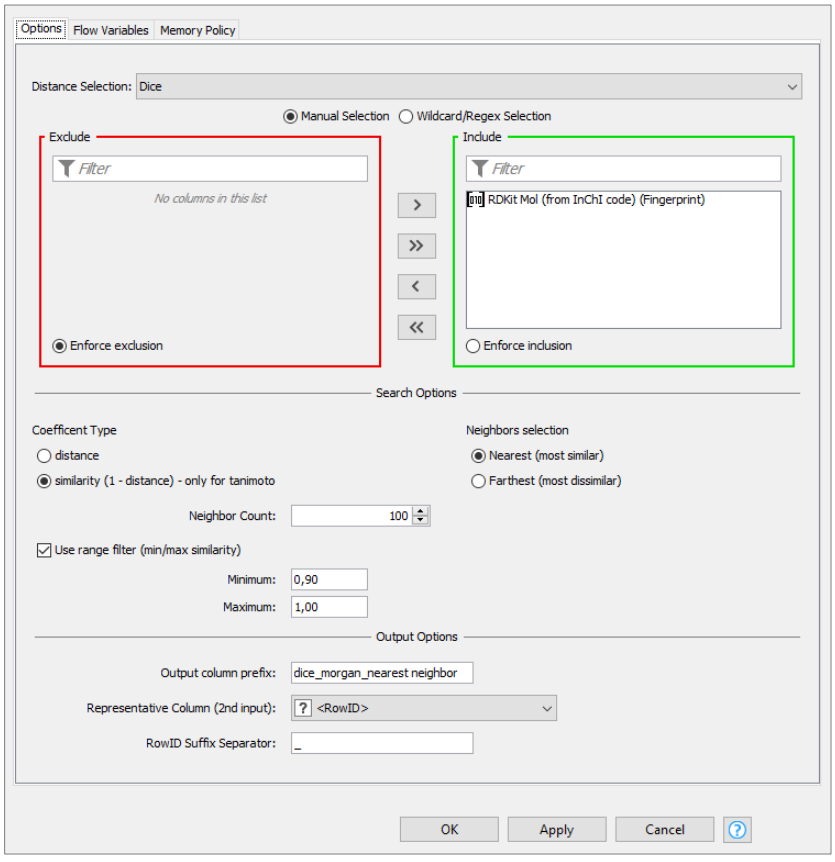


FIGURE 13: Settings of the *Similarity Search node*

$$T_c(A,B) = \frac{c}{a + b - c} \qquad D_c(A,B) = \frac{c}{\frac{1}{2}(a + b)}$$

a … count of features present in molecule A
b … count of features present in molecule B
c … count of features shared by molecules A and B

FIGURE 14: Formulas for the calculation of the Tanimoto and Dice coefficients [86]

Row ID	RDKit Mol (from InChI code)	Unique concatenate(substrate)	List(SLC)	dice_morgan...	dice_morgan...	tanimoto_morgan...	tanimoto_m...
Row4854		L-carnitine	[SLC22A16,SLC22A16,S...	Row634	1	Row632	1
Row4855		L-carnitine	[SLC22A16,SLC22A16,S...	Row634	1	Row633	1
Row4856		L-carnitine	[SLC22A16,SLC22A16,S...	Row634	1	Row634	1
Row4857		Methylarginine	[SLC25A29]	?	?	?	?
Row4858		Homoarginine, L-homoarginine	[SLC25A29,SLCO4C1]	Row511	1	Row511	1
Row4859		Homoarginine, L-homoarginine	[SLC25A29,SLCO4C1]	Row564	0.941	Row511	1
Row4860		Homoarginine, L-homoarginine	[SLC25A29,SLCO4C1]	Row565	0.941	Row511	1
Row4861		Homoarginine, L-homoarginine	[SLC25A29,SLCO4C1]	Row566	0.941	Row511	1
Row4862		acetylcholine, Acetylcholine	[SLC18A3,SLC18A3,SLC...	?	?	?	?

FIGURE 15: Joined results after similarity search with columns containing the row ID of rows from the MetSigDis-derived table with a similarity ranging from 0.90 to 1.00

The results from the Dice and the Tanimoto calculation are then combined through joining via the substrates InChI codes (see Figure 15). Since the results of the two searches differ in part, the *Column Aggregator node* is in the next step used to identify differences in the row IDs found, whereby the two row ID columns are displayed in the output in form of unique concatenate and unique count. After that, all cells containing two row IDs are split (*Cell Splitter node*) and then ungrouped into new rows. Query rows that do not match to any row of the reference table are excluded (see Figure 16).

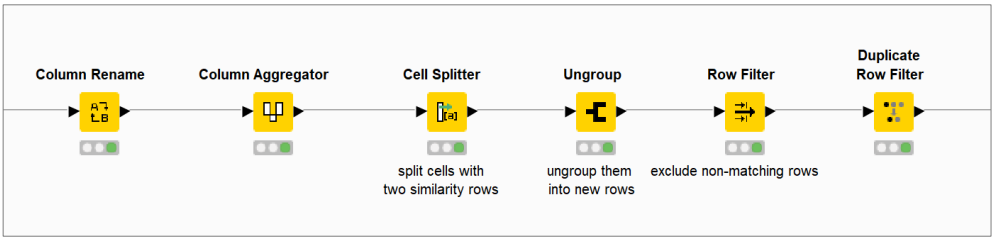


FIGURE 16: Editing the results from similarity search

It should be noted that the Morgan fingerprints used in the similarity search are sometimes not able to distinguish between different molecules. However, in order to realize the previously defined first goal of searching for metabolite-disease-SLC associations, it is necessary to find identical molecules between the two datasets. A manual check is therefore necessary to ensure this. Before performing this manual control step, the data table is divided into three different parts (see Figure 17). The first and largest part consists of all rows exhibiting the similarity value 1 for both Tanimoto and Dice similarity, most of the results are correct in this part. The second part includes all rows that have a Tanimoto similarity value of 1, but a Dice similarity value below 1. A reverse case does not appear in the results. The third part then contains all rows with a score of less than 1 for both similarity values. After manually controlling the results of all three tables, they are separated into two appropriate outputs. After that, these split results are then concatenated into one table containing all the rows that carry associations of identical molecules and another table containing all the results that were not identical but at least similar.

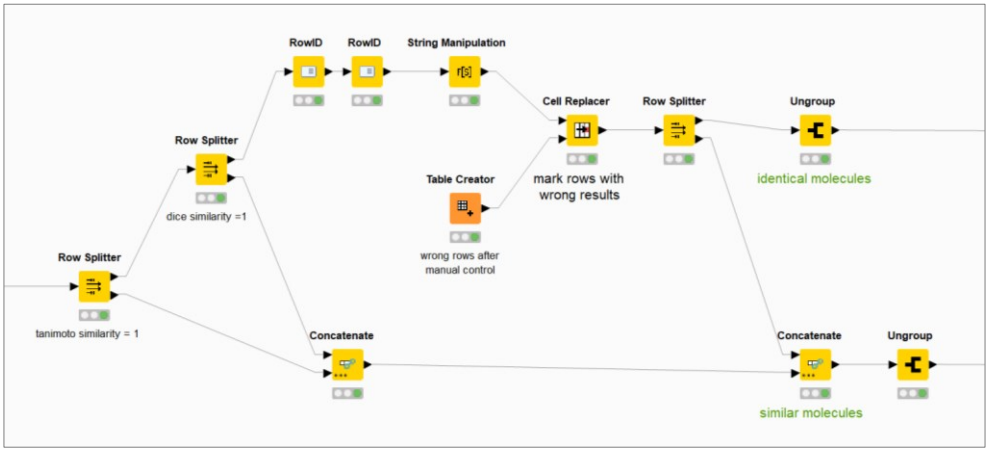


FIGURE 17: Separating rows with identical molecule matches from rows with non-identical matches

Matches of non-identical molecules occur more frequently, for example, when the properties of a molecule highly depends on its stereochemistry or when the molecule bears long carbon chains. For a more detailed explanation of the results that were filtered out, see the results and discussion chapters.

The idea was that the results with a similarity of less than 1 or incorrect associations of molecules with a similarity of 1, which are all combined in the lower output table, could serve as a source for possible new substrates for the corresponding SLC transporters. For this purpose, relevant columns from the metabolites data set, which were previously excluded in the similarity search, are added by joining via the corresponding row ID. Since a large part of the SLCs is represented in the SLC substrates list with several substrates and in the similarity search each row of the query table is compared to each row of the reference table, it follows that the present table also contains data rows in which the two compared molecules are indeed not identical, but the molecules are nevertheless substrates of the corresponding transporters. In the subsequent step, known SLC substrate combinations are therefore filtered out through joining it with the correct similarity search results via the SLCs and the InChI codes. After that, the table is grouped over the SLC and InChI columns to get a clearer arrangement of the data. Figure 18 shows all these processing steps to get to the possible new substrates output table. A different attempt that was made was to filter the resulting data set in such a way that only data rows remain whose columns refer to the same SLC. Therefore, the *Rule-based Row Splitter node* was used to exclude rows that do not contain the same SLC in the respective columns deriving from the metabolites and the substrates table. However, the output was greatly reduced by this additional step (see the results part).

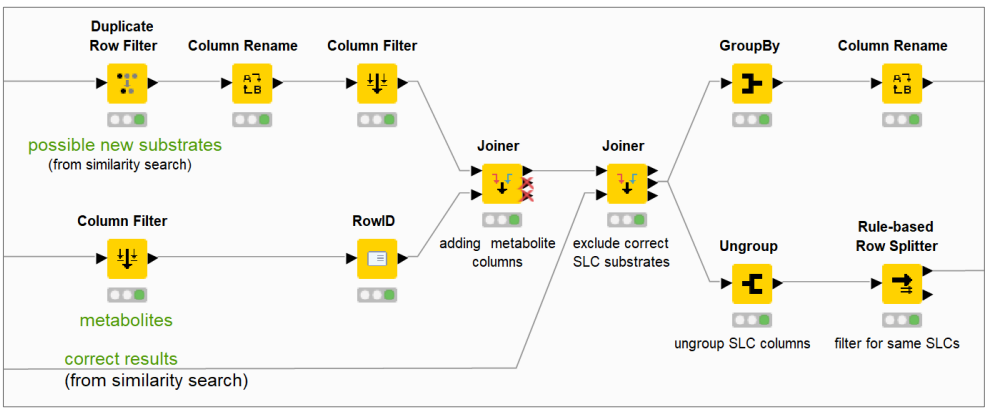


FIGURE 18: Processing of the possible new substrates data table

As far as the table with the correct results is concerned, the procedure of adding the columns from the metabolites data set is quite similar (see Figure 19). However, in addition to joining via the corresponding row ID, a *Rule-based Row Splitter node* (see Figure 20) is then used for filtering in such a way that the resulting data set only contains rows that have the same SLCs for the columns derived from the metabolites table and those that descend from the substrates table. Thus, the desired goal to interconnect metabolites, diseases and SLC transporters can be achieved. For specific information on the resulting data of these two outputs see the results and discussion part.

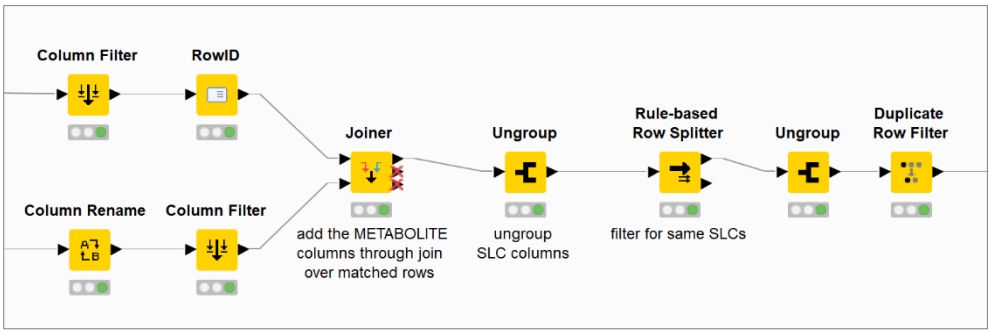


FIGURE 19: Processing of the metabolite-disease-SLC associations data table

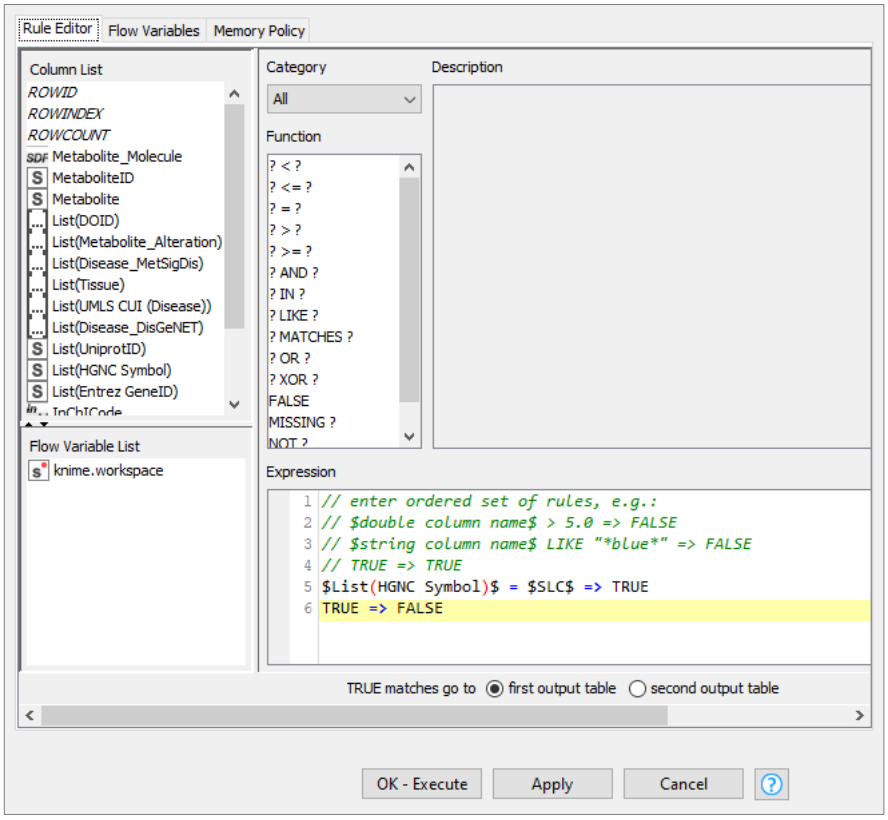


FIGURE 20: Settings of the *Rule-based Row Splitter*

2.2.5 Step 5: Alternative way: Joining

In this alternative approach to combine the metabolites data set with the substrates data set, the two tables are simply joined over their InChI codes, instead of performing the similarity search step. However, the great advantage of this method that only molecules which have the exact same InChI code can be joined also has the disadvantage that little differences in the structural representation, like for instance concerning hydrogens, charges or an inconsistent stereochemistry result in the structures not being recognized as equal. Due to that it is crucial for the outcome of this approach to use the exact same structure standardisation for both of the data sets.

To achieve this requirement, the metabolites data set is normalised using the *Standardiser_RDKit_Component* created by Jennifer Hemmerich [74], which was also used in the SLC_substrates workflow from which the SLC substrates list derives. All the settings of the component are also adopted from this workflow (for the settings of the component see Figure 22). After performing the structure standardisation step, both the metabolites table and the substrates table are grouped over the InChI and SLC columns, before they are then joined over the standardised InChI codes and the SLCs. Following to that, the table is then edited and columns that were previously aggregated as lists are ungrouped (see Figure 21).

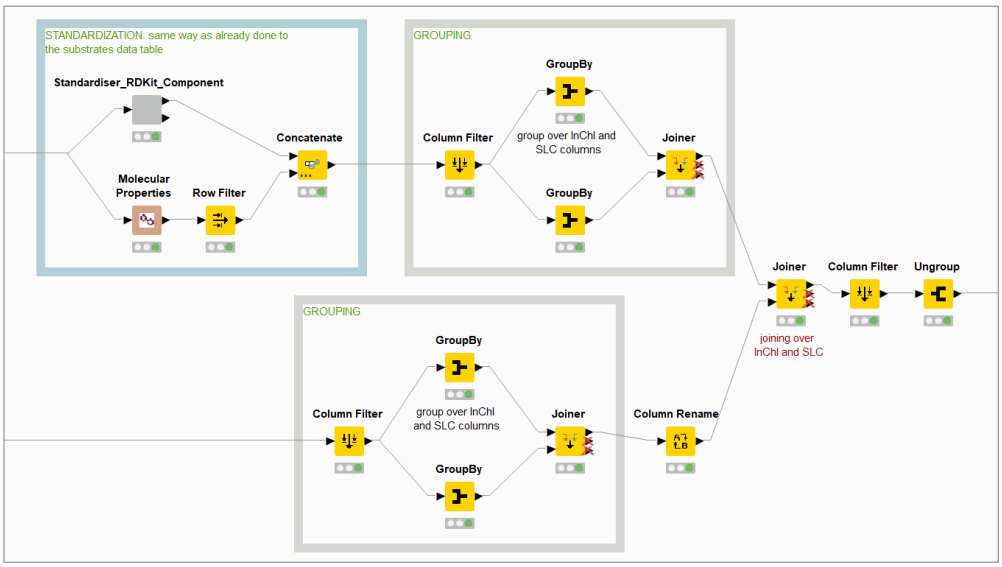


FIGURE 21: Overview of the joining approach: First the structure standardisation of the metabolite structures, then the same grouping step for both the metabolites table and the substrates table, followed by the joining via InChI codes and SLCs.

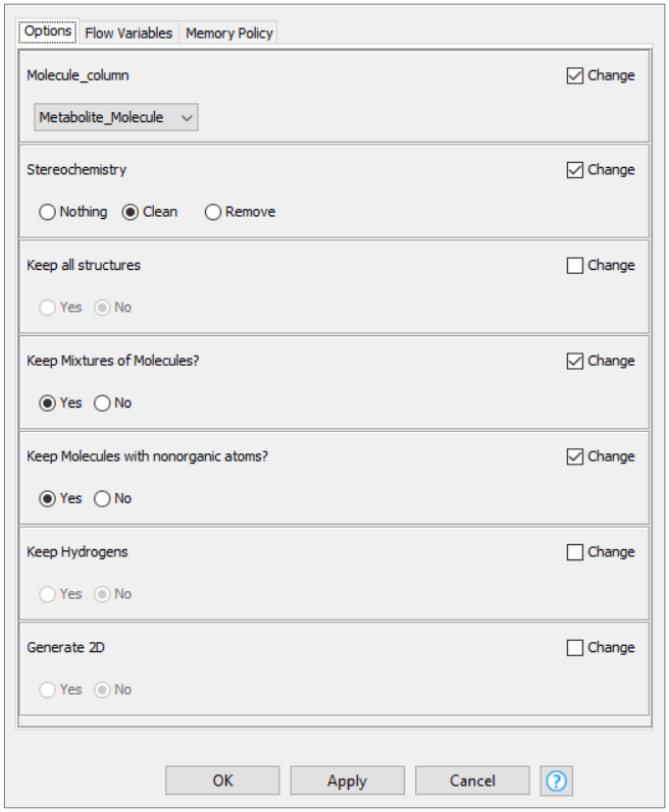


FIGURE 22: Settings of the *Standardiser_RDKit_Component*

3 RESULTS

The MetSigDis data set, which was used as the starting point for this work, contains 4651 metabolite-disease associations between 92 different diseases (counted as DOIDs) and 724 metabolites (counted as HMDB codes) after carrying out the processing steps. When the DisGeNET database was subsequently queried, 71 of the 92 diseases (as DOIDs) were found to have associations with a total of 5032 proteins. Accordingly, all data rows from the 21 queries for which there was no successful response are excluded from the further analysis. Furthermore, also in the step of adding the structure information data rows are lost, namely all those that do not contain a HMDB code. This joining step reduces the size of the data table from previously 911 292 rows to 646 027 rows (a reduction of almost 30 percent) containing 625 different metabolites. After data rows that do not relate to SLC transporters are then removed, which reduces the data set to a relatively compact table of 14 282 rows containing 132 SLCs, the workflow divides into two paths in order to link the received data table with the SLC Substrate list. The results of the workflow finally derive from the similarity search and from the alternative joining way.

3.1 METABOLITE-DISEASE-SLC ASSOCIATIONS: RESULTS FROM SIMILARITY SEARCH

After filtering all non-identical molecular matches, which are of interest as possible new substrates, but not for this approach, the similarity search path results in a data set of 1957 rows including 220 different SLCs and a total of 175 substrates. After that, it is now also necessary to connect the data tables via the SLCs. Through this step the desired goal of linking disease-associated alterations of metabolites with SLC transporters can be achieved. The end result is a table with 1404 rows containing associations between 47 different SLCs, 52 metabolites and 39 diseases. For more clarity, the resulting data set is given in the appendix (Supplementary Table 1) in such a way that the individual

metabolites are collected in the form unique concatenate. This reduces the data table to a total of 450 rows.

3.2 METABOLITE-DISEASE-SLC ASSOCIATIONS: RESULTS FROM JOINING PATH

In this alternative way, joining via InChI codes and SLCs results in a data set that includes 119 rows, which incorporate correlations between 37 SLCs, 38 substrates and 17 diseases. A comparison of the two approaches shows that the similarity search method delivers more results (manually controlled and with a similarity = 1) than this alternative approach, which is displayed in Figure 23.

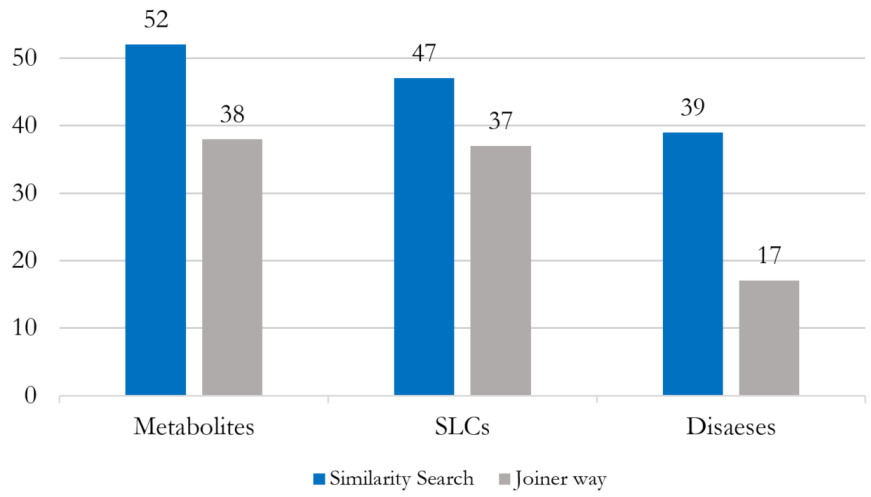


FIGURE 23: Results from the similarity search versus the results from the alternative joiner approach

An essential step to improve the performance of the joiner path was to update the SLC substrates list by manually adding missing InChI codes via a search in CHEBI and PubChem. For this purpose, an Excel table consisting of 479 rows with InChI codes that were missing for numerous substrates of the data set was created. After including this data table into the workflow from which the SLC substrates list derives from, the resulting new data set was saved and then

integrated into this work's workflow. The second decisive step, as described in more detail in the methods section, was to implement the same structure standardization for both, the substrates table and the metabolites table, which was done by using the *Standardiser_RDKit_Component*. Through this and furthermore through the enhancement of the substrates data set it was possible to significantly improve the output of the joining approach. The results increased from originally 18 SLCs, 12 substrates and 8 diseases to 37 SLCs, 38 substrates and 17 diseases. Of course the addition of missing information to the substrates list also led to an improvement of the similarity search results.

3.3 RESULTS: POSSIBLE NEW SUBSTRATES

The second output table derived from the similarity search with the name "possible new substrates" contains on the one hand all those data rows with similarity values below 1 (similarity: 0.90 <= x < 1.00) and on the other hand all data rows that, despite a calculated similarity value of 1, only provide similar and not identical molecules. The final output of this approach is a table of 267 rows containing 57 different molecules that might be possible new substrates for the 85 included SLC transporters. The output data set is given in the appendix as Supplementary Table 2. As for the attempt to additionally relate the resulting data set to the same SLCs, the output was greatly reduced to a table of 58 rows containing information on only 6 SLCs and 5 substances.

4 DISCUSSION

4.1 SIMILARITY SEARCH VERSUS JOINING PATH

As one could already see in Figure 23, the similarity search approach yielded more results (manually controlled and with similarity = 1) for the connection of metabolites, diseases and SLC transporters. But how can this difference be explained and what are the actual advantages and disadvantages of each approach? On the one hand, joining via InChI codes, is quite prone to overlook identical molecules. This can happen due to an inconsistent structure standardisation, which leads to the problem that the same molecule can provide different chemical identifiers, depending on whether the molecule is protonated, charged or represented with its stereochemistry, for example. This is also the reason why it can be difficult to merge data sets from different origins, since different chemical databases usually implement different structure standardisation processes and for this reason often supply different identifiers for the same molecule. [80] Apparently, in some cases these differences cannot be completely compensated by using the *Standardiser_RDKit_Component*, which would be a reason for the different extent of results from the two approaches. Here lies the main advantage of the similarity search approach, namely that the structure of a molecule itself is searched to meet the defined similarity or distance criteria of another molecule. The threshold value can be changed in such a way that molecules with different similarity values can be identified, which also offers the possibility to realise the approach of describing potential new substrates. However, there is of course the disadvantage here that a manual check of the results is necessary, which can be carried out relatively easy for this volume of data, but is associated with an ever-increasing effort in the case of extensive data sets. One limitation due to this manual check arises for the repeated implementation of the workflow with changed data sets. The difficulty lies in the step of filtering out incorrect matches from the similarity search, since this exclusion is done for manually selected rows. This is also the most obvious advantage of the joiner path, which immediately delivers correct results without the need for manual intervention.

In general, the similarity search delivers quite good results but certain chemical properties such as stereochemistry and long carbohydrate chains are not taken into account, or only to a limited extent, which is why the manual check is necessary. The problem that stereochemistry is not recognised arises not from the similarity search step itself, but from the use of fingerprints like Morgan or MACCS, since they are 2D molecular descriptors, which means that they cannot encode three-dimensional structures of molecules. [84]

In order to find out to what extent different stereoisomers like for instance sugars are distinguished, I did a try-out with a data set consisting of diverse monosaccharides. It turned out that the similarity search step cannot differentiate between particular stereoisomers like the members of the aldohexoses group. For example: Glucose is, besides others, recognized as mannose or galactose and vice versa. Furthermore, no distinction can be made between D- and L-isomers. However, the similarity search step is able to distinguish between pyranose and furanose forms. Of course, this also raises the question of the extent to which SLCs themselves distinguish between different sugars at all.

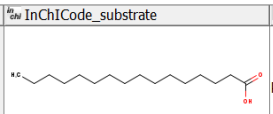
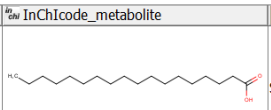
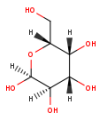
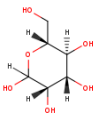
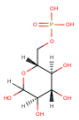
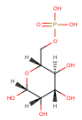
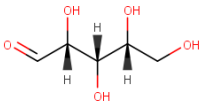
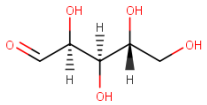
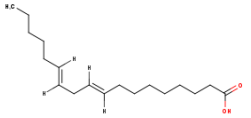
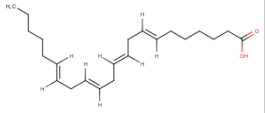
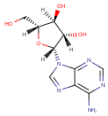
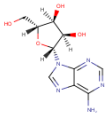
Row ID	InChICode_substrate	substrate	InChIcode_metabolite	metabolite
Row0		Palmitate		Stearate
Row1		D-galactose		D-Mannose
Row2		glucose 6-phosphate		mannose-6-phosphate
Row3		xylose, Xylose		arabinose
Row4		Linoleate		Adrenic acid
Row5		vidarabine		Adenosine

FIGURE 24: Selected examples of wrong matches with both Dice and Tanimoto similarity of 1

In reviewing the similarity search results manually, there were some aspects that appeared. On the one hand, that all those rows that erroneously show a Dice and Tanimoto similarity of 1 are errors due to not recognising stereochemistry or errors due to a lacking differentiation of long carbohydrate chains. It is therefore not surprising that the faulty molecules are mainly fatty acids and monosaccharides or substrates that contain sugars in their structure. Figure 24 provides representative examples of such wrong matches.

In terms of the proportion of false results, a total of 86 of the 512 rows that have a Dice and Tanimoto similarity of 1 give false positive matching molecules. As for rows with a Tanimoto similarity equal to 1 and a Dice similarity below 1, there are just non-identical molecule pairings, indicating that Dice similarity has more stringent requirements and hence fewer false results.

Another detail to consider is which fingerprints should be used to describe the molecular structures of the metabolites and substrates. In the development of this workflow Morgan and MACCS fingerprints were tried out. Here, the Morgan fingerprints clearly outperformed the MACCS fingerprints, producing fewer false results. In concrete terms, the similarity search using MACCS fingerprints resulted in 1200 rows with a similarity of 1 for the Tanimoto calculation method and 961 rows for the Dice similarity search. On the other hand, the search with Morgan fingerprints resulted in a number of 573 rows with a similarity of 1 for the Tanimoto similarity search and 538 rows for the Dice similarity. The resulting rows from the Morgan fingerprint searches overlapped the results from the searches with MACCS fingerprints. A manual inspection revealed that there were no correct results among the additional output rows from the MACCS fingerprint searches.

4.2 METABOLITE-DISEASE-SLC ASSOCIATIONS: INTERPRETATION

The first objective was to link the disease-altered metabolites to the SLC transporters responsible for their transport. This goal could be achieved with the methods described above and resulted in Supplementary Table 2 which contains the found associations between metabolites, diseases and SLC transporters.

A comparison with a table from the review "The SLC transporter in nutrient and metabolic sensing, regulation and drug development" by Zhang et al. (also see section 1.2.1), which provides a representative summary of SLC transporters relevant in human diseases [14], shows that at least 5 of the presented associations between SLC transporters and diseases also appear in the results of this work. An example for such a finding is the association between SLC2A2 and type 2 diabetes mellitus. This connection found is affirmed by an earlier GWAS study that found SLC2A2 as one of several gene loci for insulin resistance and type 2 diabetes mellitus. [15] Other such results for which there is supporting evidence from research include the association of SLC6A12 with schizophrenia [18], SLC6A15 as a gene locus associated with a risk of major depression [19] and a connection of the SLC5A8 and SLC2A1 expression with different kinds of cancer [87,88]. Results like these suggest that further relevant associations of metabolite transporting SLCs with human diseases may be included in the output of this work.

A comparison with the table of Mendelian diseases associated with SLC transporters from the paper "SLC Transporters as Therapeutic Targets: Emerging Opportunities" by Lin et al. [6] (also see section 1.2.1) did not reveal any overlaps with the results of this work. However, this is primarily due to the fact that the initial data set from MetSigDis contains only very few Mendelian diseases.

4.3 POSSIBLE NEW SUBSTRATES: INTERPRETATION

The second objective to describe possible new substrates for SLC transporters was realised by searching the metabolites table resulting from step 3 for molecules with a relative high similarity to already identified substrates. Supplementary Table 3, which results from this approach, contains the possible new substrates for the respective SLC transporters. Most of these resulting data rows describe molecules from the structural groups of monosaccharides (mainly aldohexoses), amino acids (proteinogenic and non-proteinogenic), fatty acids and other carboxylic acids. Moreover, the output table also contains oligopeptides, nucleotides or nucleosides and other structures that cannot be assigned to the mentioned groups. For an overview on the distribution of the group membership of the possible new substrates, see Figure 24.

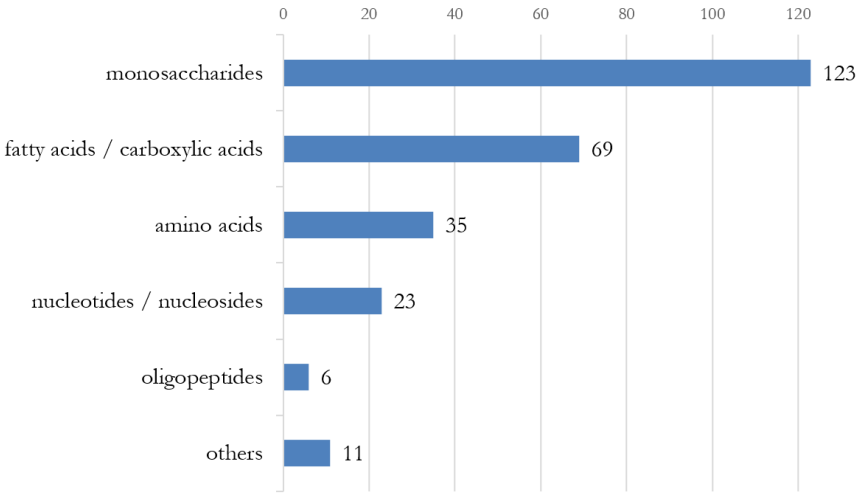


FIGURE 25: Distribution of the metabolites from the 267 resulting rows on different structural groups

The results found are quite reasonable, since SLCs often transport molecules which are very similar to each other. As an example, SLC2A3 (aka GLUT3) mediates the transport of different monosaccharides like for instance glucose, galactose or xylose. Other molecules that are similar to these already established ones (e.g. ribose and arabinose) are then found as possible new substrates for this transporter. At the same time, the fact that only similar molecules are identified is also a major disadvantage of this approach. As a matter of fact, there

won't appear any unexpected candidates as possible new substrates of SLCs, which of course would be more exciting for a further analytic elucidation, since such exceptional compounds may not have been tested yet. For further analyses, a possible adaptation would be to lower the similarity range (currently: 0.90 <= x <= 1.00), for example to a similarity of >= 0.80.

A promising prospective idea would be to link my workflow to metabolomics data from the RESOLUTE project. The data obtained from targeted and untargeted metabolomics studies, which presents relations between the expression of SLC transporters and changes in the concentration of metabolites in the intra- or extracellular space of specific cell lines, could serve as starting data for the search of new substrates. Similar to the data used in the current approach for finding potential substrates, the metabolomics data also does not provide distinct evidence whether the detected metabolite itself is transported by the SLC or whether the concentration change is caused by coupled processes. However, as recent examples such as the deorphanization of SLC16A9 and SLC22A24 using metabolomic GWAS [89,90] show, metabolomics data generally provides a promising source for the identification of possible new SLC substrates.

4.4 DATA LIMITATIONS

A central point for the creation of this work was the publical availability of data. However, some problems appeared in the use of some of the databases mentioned in the databases and tools section. One issue that already arose at the very beginning in the development of this work was that after downloading the data file from the MetSigDis webpage it turned out that the table did not contain the full MetSigDis data set. In fact, it only included 5567 rows instead of the 6849 metabolite-disease relationships as stated in the paper and on the webpage. An attempt to contact the authors was unsuccessful because the e-mail addresses given in the contact section no longer existed. It appears as if this resource was left unmanaged since it was created.

One further limitation and a point for possible adaptations is provided in step 3. After adding the molecular structures and identifiers, all rows are filtered out for which the data cannot be added due to missing HMDB codes. It is important to consider that through filtering out a relevant number of rows relevant information might be lost. In fact, nearly thirty percent of the metabolite rows are eliminated, although it has to be noted that the rows that are excluded mostly contain ambiguous metabolite notations or designate whole metabolite classes. Examples for such designations would be "NTP" (nucleoside triphosphate) or "total cholines", which would anyway deliver rather inconclusive results. But since the metabolites without HMDB codes are not only unassignable designations (in some cases there are also distinct names), the loss could be partially compensated by manually adding missing HMDB codes to the MetSigDis data set beforehand.

Another point for a potential adaptation lies in a further improvement of the substrates list. Although many InChI codes for respective substrates have already been added, which led to a significant enhancement of the results, the extensive SLC substrates list still offers a potential for improvements.

In conclusion, it can be said that the chosen approaches are suitable for achieving the objectives defined at the beginning of this work. The first aim, to gain insights on associations between metabolites, diseases and SLCs is mainly achieved through the similarity search approach. As it turned out, already known correlations appear in the results, which indicates that more relevant associations could be contained in the rest of the output. The second aim of finding new potential substrates for SLCs was also achieved using the similarity search approach. Since SLCs often transport molecules that are similar to each other, the results obtained are quite reasonable, although not surprising.

5 APPENDIX

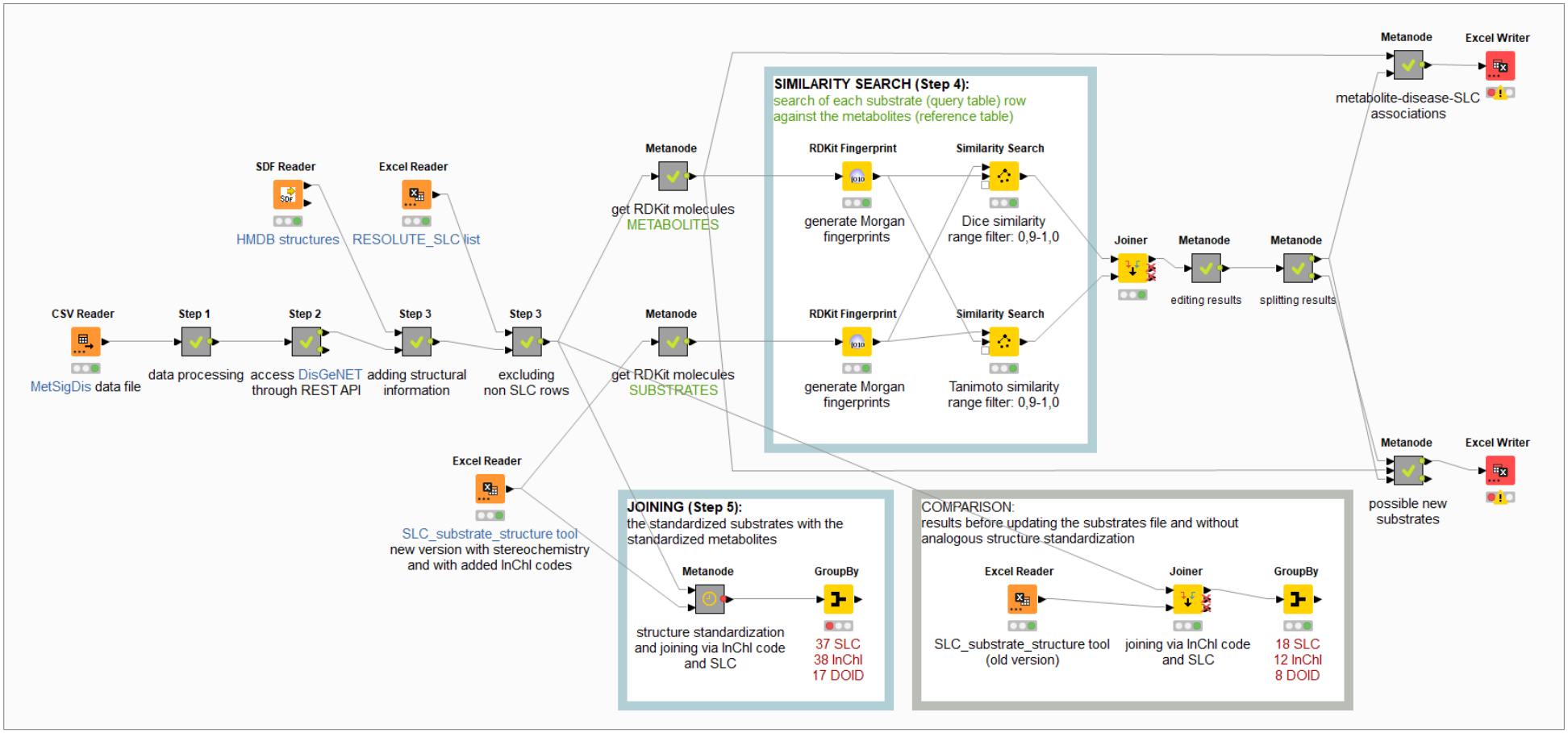


FIGURE 26: Workflow for the connection of disease associated alterations of metabolites with corresponding SLC transporters

SUPPLEMENTARY TABLE 1: Output table: Altered levels of SLC-transported metabolites in corresponding tissues as a result of given diseases.
For more clarity, the metabolites were collected in the form unique concatenate and some columns were filtered.

SLC	Metabolites	Metabolite Alteration	Tissue	Disease
FLVCR1	D-Glucose	Increased	urine	hepatocellular carcinoma
SFXN1	alanine	Decreased	breast cancer cell lines	breast cancer
SFXN1	Alanine, alanine	Decreased	serum	colorectal cancer
SFXN1	glycine	Decreased	Plasma	congestive heart failure
SFXN1	Alanine, Cysteine	Decreased	Serum	esophagus squamous cell carcinoma
SFXN1	alanine	Decreased	blood plasma	lung cancer
SFXN1	Glycine	Decreased	frozen tissue samples	lung cancer
SFXN1	Glycine	Decreased	urine	obesity
SFXN1	alanine	Decreased	Serum	primary biliary cirrhosis
SFXN1	Glycine	Decreased	cell lines	prostate cancer
SFXN1	alanine	Decreased	prostate cancer cell lines	prostate cancer
SFXN1	alanine	Decreased	cerebrospinal fluid	schizophrenia
SFXN1	Alanine	Difference	Serum	Alzheimer's disease
SFXN1	glycine, alanine	Difference	Serum	cancer
SFXN1	glycine, alanine	Difference	glioma	malignant glioma
SFXN1	Alanine	Increased	bladder	Bladder Cancer
SFXN1	glycine	Increased	urine	Parkinson's disease
SFXN1	alanine	Increased	Colon Cancer Initiating Cells	colon cancer
SFXN1	glycine	Increased	blood	colorectal cancer
SFXN1	Glycine, Cysteine	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SFXN1	Cysteine	Increased	Blood	esophageal cancer
SFXN1	Alanine	Increased	Plasma	esophageal cancer
SFXN1	Alanine	Increased	urine	obesity
SFXN1	Alanine	Increased	visceral and subcutaneous adipose tissue	obesity
SFXN1	alanine	Increased	Prostate cell lines	prostate cancer
SFXN1	Alanine	Increased	cell lines	prostate cancer
SFXN1	Glycine, glycine, Alanine	Increased	prostate tissues	prostate cancer
SFXN1	glycine, alanine	Increased	plasma	schizophrenia
SFXN1	Alanine	Increased	Lung tissue	squamous cell carcinoma
SFXN1	Alanine	Increased	serum	squamous cell carcinoma
SFXN1	Glycine	Increased	serum	type 1 diabetes mellitus
SFXN1	alanine	Increased	Blood	type 2 diabetes mellitus
SLC16A1	Pyruvate, Lactate	Decreased	urine	Celiac disease
SLC16A1	glycolate, acetoacetate	Decreased	Colon Cancer Initiating Cells	colon cancer

SLC16A1	Pyruvate, Lactate	Decreased	serum	colorectal cancer
SLC16A1	Pyruvate	Decreased	urine	esophageal cancer
SLC16A1	Lactate	Decreased	urine	obesity
SLC16A1	lactate	Decreased	Serum	primary biliary cirrhosis
SLC16A1	Lactate	Decreased	cell lines	prostate cancer
SLC16A1	lactate	Decreased	cerebrospinal fluid	schizophrenia
SLC16A1	acetoacetate, 3-hydroxybutyrate	Decreased	plasma	schizophrenia
SLC16A1	Pyruvate	Decreased	serum	type 1 diabetes mellitus
SLC16A1	lactate	Difference	1st trimester pregnant woman plasma	Down syndrome
SLC16A1	Pyruvate, Lactate	Difference	renal cell	kidney cancer
SLC16A1	3-hydroxybutyrate	Difference	plasma	mental depression
SLC16A1	lactate	Difference	Brain	schizophrenia
SLC16A1	lactate	Increased	Brain	bipolar disorder
SLC16A1	Lactate	Increased	breast cancer cell lines	breast cancer
SLC16A1	lactate	Increased	blood	colorectal cancer
SLC16A1	Lactate	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC16A1	3-hydroxybutyrate	Increased	serum	colorectal cancer
SLC16A1	glycolate, lactate	Increased	blood	congestive heart failure
SLC16A1	Lactate	Increased	Blood	esophageal cancer
SLC16A1	Lactate	Increased	Plasma	esophageal cancer
SLC16A1	Lactate	Increased	intestinal epithelial cells	inflammatory bowel disease
SLC16A1	lactate	Increased	Lung Cancer Tissue	lung cancer
SLC16A1	lactate	Increased	blood plasma	lung cancer
SLC16A1	Lactate	Increased	frozen tissue samples	lung cancer
SLC16A1	3-hydroxybutyrate	Increased	Serum	primary biliary cirrhosis
SLC16A1	Lactate	Increased	prostate cancer cell lines	prostate cancer
SLC16A1	lactate	Increased	plasma	schizophrenia
SLC16A1	Lactate	Increased	Lung tissue	squamous cell carcinoma
SLC16A3	Lactate	Decreased	urine	Celiac disease
SLC16A3	Lactate	Decreased	serum	colorectal cancer
SLC16A3	Lactate	Decreased	urine	obesity
SLC16A3	Lactate	Decreased	cell lines	prostate cancer
SLC16A3	Lactate	Difference	renal cell	kidney cancer
SLC16A3	Lactate	Increased	breast cancer cell lines	breast cancer
SLC16A3	Lactate	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC16A3	Lactate	Increased	Blood	esophageal cancer
SLC16A3	Lactate	Increased	Plasma	esophageal cancer
SLC16A3	Lactate	Increased	intestinal epithelial cells	inflammatory bowel disease
SLC16A3	Lactate	Increased	frozen tissue samples	lung cancer
SLC16A3	Lactate	Increased	prostate cancer cell lines	prostate cancer

SLC16A3	Lactate	Increased	Lung tissue	squamous cell carcinoma
SLC16A7	Lactate	Decreased	urine	Celiac disease
SLC16A7	Lactate	Decreased	serum	colorectal cancer
SLC16A7	Lactate	Decreased	urine	obesity
SLC16A7	Lactate	Decreased	cell lines	prostate cancer
SLC16A7	Lactate	Difference	renal cell	kidney cancer
SLC16A7	Lactate	Increased	breast cancer cell lines	breast cancer
SLC16A7	Lactate	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC16A7	Lactate	Increased	Blood	esophageal cancer
SLC16A7	Lactate	Increased	Plasma	esophageal cancer
SLC16A7	Lactate	Increased	intestinal epithelial cells	inflammatory bowel disease
SLC16A7	Lactate	Increased	frozen tissue samples	lung cancer
SLC16A7	Lactate	Increased	prostate cancer cell lines	prostate cancer
SLC16A7	Lactate	Increased	Lung tissue	squamous cell carcinoma
SLC17A6	glutamate	Decreased	breast and blood	breast cancer
SLC17A6	glutamate	Decreased	serum	colorectal cancer
SLC17A6	glutamate	Difference	Serum	cancer
SLC17A6	glutamate	Difference	Brain	schizophrenia
SLC17A6	glutamate	Increased	Brain	bipolar disorder
SLC17A6	glutamate	Increased	Plasma	congestive heart failure
SLC17A6	glutamate	Increased	cell	non-small cell lung carcinoma
SLC17A7	glutamate	Decreased	breast and blood	breast cancer
SLC17A7	glutamate	Decreased	serum	colorectal cancer
SLC17A7	glutamate	Difference	Serum	cancer
SLC17A7	glutamate	Difference	Brain	schizophrenia
SLC17A7	glutamate	Increased	Brain	bipolar disorder
SLC17A7	glutamate	Increased	Plasma	congestive heart failure
SLC17A7	glutamate	Increased	cell	non-small cell lung carcinoma
SLC18A1	5-HT	Decreased	Blood	schizophrenia
SLC18A2	5-HT	Decreased	Blood	schizophrenia
SLC18A2	Serotonin	Increased	Plasma	alcohol abuse
SLC19A1	folate	Decreased	Colon Cancer Initiating Cells	colon cancer
SLC1A1	Asparagine	Decreased	urine	Celiac disease
SLC1A1	Asparagine	Decreased	the plasma amino acids	autistic disorder
SLC1A1	glutamate	Decreased	breast and blood	breast cancer
SLC1A1	glutamate	Decreased	serum	colorectal cancer
SLC1A1	glutamate	Difference	Serum	cancer
SLC1A1	glutamate	Difference	Brain	schizophrenia
SLC1A1	Asparagine	Increased	bladder	Bladder Cancer
SLC1A1	glutamate	Increased	Brain	bipolar disorder

SLC1A1	Asparagine	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC1A1	glutamate	Increased	Plasma	congestive heart failure
SLC1A1	glutamate	Increased	cell	non-small cell lung carcinoma
SLC1A2	glutamate	Decreased	breast and blood	breast cancer
SLC1A2	glutamate	Decreased	serum	colorectal cancer
SLC1A2	glutamate	Difference	Serum	cancer
SLC1A2	glutamate	Difference	Brain	schizophrenia
SLC1A2	L-Asparagine	Increased	Plasma	alcohol abuse
SLC1A2	glutamate	Increased	Brain	bipolar disorder
SLC1A2	glutamate	Increased	Plasma	congestive heart failure
SLC1A2	glutamate	Increased	cell	non-small cell lung carcinoma
SLC1A3	Asparagine	Decreased	urine	Celiac disease
SLC1A3	Asparagine	Decreased	the plasma amino acids	autistic disorder
SLC1A3	glutamate	Decreased	breast and blood	breast cancer
SLC1A3	glutamate	Decreased	serum	colorectal cancer
SLC1A3	glutamate	Difference	Serum	cancer
SLC1A3	glutamate	Difference	Brain	schizophrenia
SLC1A3	Asparagine	Increased	bladder	Bladder Cancer
SLC1A3	glutamate	Increased	Brain	bipolar disorder
SLC1A3	Asparagine	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC1A3	glutamate	Increased	Plasma	congestive heart failure
SLC1A3	glutamate	Increased	cell	non-small cell lung carcinoma
SLC1A6	glutamate	Decreased	breast and blood	breast cancer
SLC1A6	glutamate	Decreased	serum	colorectal cancer
SLC1A6	glutamate	Difference	Serum	cancer
SLC1A6	glutamate	Difference	Brain	schizophrenia
SLC1A6	glutamate	Increased	Brain	bipolar disorder
SLC1A6	glutamate	Increased	Plasma	congestive heart failure
SLC1A6	glutamate	Increased	cell	non-small cell lung carcinoma
SLC1A7	Asparagine	Decreased	urine	Celiac disease
SLC1A7	Asparagine	Decreased	the plasma amino acids	autistic disorder
SLC1A7	glutamate	Decreased	breast and blood	breast cancer
SLC1A7	glutamate	Decreased	serum	colorectal cancer
SLC1A7	Glutamate	Decreased	urine	esophageal cancer
SLC1A7	Aspartate, Glutamate	Decreased	intestinal epithelial cells	inflammatory bowel disease
SLC1A7	Aspartate, Glutamate	Decreased	cell lines	prostate cancer
SLC1A7	glutamate	Difference	Serum	cancer
SLC1A7	glutamate	Difference	Brain	schizophrenia
SLC1A7	Asparagine, Glutamate	Increased	bladder	Bladder Cancer
SLC1A7	glutamate	Increased	Brain	bipolar disorder

SLC1A7	Aspartate, Asparagine, Glutamate	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC1A7	glutamate	Increased	Plasma	congestive heart failure
SLC1A7	Aspartate, Glutamate	Increased	Blood	esophageal cancer
SLC1A7	glutamate	Increased	cell	non-small cell lung carcinoma
SLC1A7	Glutamate	Increased	prostate tissues	prostate cancer
SLC1A7	Glutamate	Increased	Lung tissue	squamous cell carcinoma
SLC1A7	Glutamate	Increased	serum	type 1 diabetes mellitus
SLC22A1	Spermidine	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC22A2	L-carnitine	Increased	Serum	obesity
SLC22A3	norepinephrine	Decreased	cerebrospinal fluid	Alzheimer's disease
SLC22A3	norepinephrine	Decreased	Prostate cell lines	prostate cancer
SLC22A8	4-Aminohippuric acid	Difference	Serum	fatty liver disease
SLC22A8	4-Aminohippuric acid	Difference	Serum	type 2 diabetes mellitus
SLC28A1	cytidine	Decreased	breast and blood	breast cancer
SLC28A1	Uridine, Cytidine	Decreased	cell lines	breast cancer
SLC28A1	Adenosine	Decreased	Lung Cancer Tissue	lung cancer
SLC28A1	Uridine, Cytidine	Difference	Serum	fatty liver disease
SLC28A1	Uridine, Cytidine	Difference	Serum	hypertension
SLC28A1	Uridine, Cytidine	Difference	Serum	type 2 diabetes mellitus
SLC28A1	Adenosine	Increased	cell lines	breast cancer
SLC28A1	Adenosine	Increased	intestinal epithelial cells	inflammatory bowel disease
SLC29A1	Hypoxanthine	Decreased	Urine	cocaine dependence
SLC29A1	Hypoxanthine	Difference	Serum	hypertension
SLC29A1	Hypoxanthine	Difference	Serum	type 2 diabetes mellitus
SLC29A1	Hypoxanthine	Increased	bladder	Bladder Cancer
SLC29A1	Hypoxanthine	Increased	Plasma	alcohol abuse
SLC29A1	Uracil, Hypoxanthine	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC29A1	Hypoxanthine	Increased	urine	hepatocellular carcinoma
SLC29A1	Uracil	Increased	renal cell	renal cell carcinoma
SLC2A1	Glucose	Decreased	urine	Down syndrome
SLC2A1	galactose, glucose	Decreased	breast and blood	breast cancer
SLC2A1	Glucose	Decreased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC2A1	Glucose	Decreased	Serum	esophagus squamous cell carcinoma
SLC2A1	glucose	Decreased	Lung Cancer Tissue	lung cancer
SLC2A1	glucose	Decreased	blood plasma	lung cancer
SLC2A1	Glucose	Decreased	frozen tissue samples	lung cancer
SLC2A1	glucose	Decreased	prostate tissues	prostate cancer
SLC2A1	glucose	Decreased	plasma	schizophrenia
SLC2A1	Glucose	Decreased	Lung tissue	squamous cell carcinoma
SLC2A1	glucose	Difference	Plasma	cardiovascular system disease

SLC2A1	Galactose, Glucose	Difference	renal cell	kidney cancer
SLC2A1	glucose	Difference	cerebrospinal fluid	schizophrenia
SLC2A1	Glucose	Increased	urine	Celiac disease
SLC2A1	glucose	Increased	gray matter	amyotrophic lateral sclerosis
SLC2A1	Galactose	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC2A1	D-Galactose, D-Mannose, D-Glucose	Increased	urine	hepatocellular carcinoma
SLC2A1	Glucose	Increased	urine	obesity
SLC2A10	galactose, glucose	Decreased	breast and blood	breast cancer
SLC2A10	glucose	Decreased	Lung Cancer Tissue	lung cancer
SLC2A10	glucose	Decreased	blood plasma	lung cancer
SLC2A10	glucose	Decreased	prostate tissues	prostate cancer
SLC2A10	glucose	Decreased	plasma	schizophrenia
SLC2A10	glucose	Difference	Plasma	cardiovascular system disease
SLC2A10	glucose	Difference	cerebrospinal fluid	schizophrenia
SLC2A10	glucose	Increased	gray matter	amyotrophic lateral sclerosis
SLC2A2	galactose, glucose	Decreased	breast and blood	breast cancer
SLC2A2	glucose	Decreased	Lung Cancer Tissue	lung cancer
SLC2A2	glucose	Decreased	blood plasma	lung cancer
SLC2A2	glucose, fructose	Decreased	prostate tissues	prostate cancer
SLC2A2	glucose	Decreased	plasma	schizophrenia
SLC2A2	glucose, fructose	Difference	Plasma	cardiovascular system disease
SLC2A2	glucose	Difference	cerebrospinal fluid	schizophrenia
SLC2A2	glucose	Increased	gray matter	amyotrophic lateral sclerosis
SLC2A2	D-Galactose, D-Mannose, D-Glucose	Increased	urine	hepatocellular carcinoma
SLC2A2	fructose	Increased	Blood	type 2 diabetes mellitus
SLC2A5	glucose	Decreased	breast and blood	breast cancer
SLC2A5	glucose	Decreased	Lung Cancer Tissue	lung cancer
SLC2A5	glucose	Decreased	blood plasma	lung cancer
SLC2A5	glucose	Decreased	prostate tissues	prostate cancer
SLC2A5	glucose	Decreased	plasma	schizophrenia
SLC2A5	glucose	Difference	Plasma	cardiovascular system disease
SLC2A5	glucose	Difference	cerebrospinal fluid	schizophrenia
SLC2A5	glucose	Increased	gray matter	amyotrophic lateral sclerosis
SLC38A2	alanine	Decreased	breast cancer cell lines	breast cancer
SLC38A2	alanine	Decreased	serum	colorectal cancer
SLC38A2	glycine, glutamine	Decreased	Plasma	congestive heart failure
SLC38A2	alanine, glutamine	Decreased	blood plasma	lung cancer
SLC38A2	alanine	Decreased	Serum	primary biliary cirrhosis

SLC38A2	alanine	Decreased	prostate cancer cell lines	prostate cancer
SLC38A2	alanine, glutamine	Decreased	cerebrospinal fluid	schizophrenia
SLC38A2	glutamine	Decreased	gastric	stomach cancer
SLC38A2	glycine, alanine	Difference	Serum	cancer
SLC38A2	glycine, alanine	Difference	glioma	malignant glioma
SLC38A2	glutamine	Difference	Brain	schizophrenia
SLC38A2	glycine	Increased	urine	Parkinson's disease
SLC38A2	alanine	Increased	Colon Cancer Initiating Cells	colon cancer
SLC38A2	glycine	Increased	blood	colorectal cancer
SLC38A2	glutamine	Increased	Cerebrospinal Fluid	malignant glioma
SLC38A2	alanine	Increased	Prostate cell lines	prostate cancer
SLC38A2	glycine	Increased	prostate tissues	prostate cancer
SLC38A2	glycine, alanine	Increased	plasma	schizophrenia
SLC38A2	alanine, glutamine	Increased	Blood	type 2 diabetes mellitus
SLC38A9	Tyrosine	Decreased	serum	colorectal cancer
SLC38A9	Tyrosine	Decreased	Blood	esophageal cancer
SLC38A9	Tyrosine	Decreased	urine	esophageal cancer
SLC38A9	Tyrosine	Increased	bladder	Bladder Cancer
SLC38A9	Tyrosine	Increased	the plasma amino acids	autistic disorder
SLC38A9	Tyrosine	Increased	urine	hepatocellular carcinoma
SLC38A9	Tyrosine	Increased	urine	obesity
SLC45A2	Glucose	Decreased	urine	Down syndrome
SLC45A2	Glucose	Decreased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC45A2	Glucose	Decreased	Serum	esophagus squamous cell carcinoma
SLC45A2	Glucose	Decreased	frozen tissue samples	lung cancer
SLC45A2	Glucose	Decreased	Lung tissue	squamous cell carcinoma
SLC45A2	Glucose	Difference	renal cell	kidney cancer
SLC45A2	Glucose	Increased	urine	Celiac disease
SLC45A2	Glucose	Increased	urine	obesity
SLC51A	dehydroisoandrosterone sulfate	Decreased	Serum	primary biliary cirrhosis
SLC51A	taurocholate	Increased	Serum	primary biliary cirrhosis
SLC51B	dehydroisoandrosterone sulfate	Decreased	Serum	primary biliary cirrhosis
SLC51B	taurocholate	Increased	Serum	primary biliary cirrhosis
SLC5A3	myo-inositol	Decreased	frozen tissue samples	lung cancer
SLC5A3	myoinositol	Decreased	renal cell	renal cell carcinoma
SLC5A3	myo-inositol	Decreased	Brain	schizophrenia
SLC5A3	myo-inositol	Difference	glioma	malignant glioma
SLC5A3	myo-inositol	Increased	Brain	bipolar disorder
SLC5A3	myo-inositol	Increased	renal cell	kidney cancer
SLC5A3	myo-inositol	Increased	Cerebrospinal Fluid	malignant glioma

SLC5A8	Pyruvate, Lactate	Decreased	urine	Celiac disease
SLC5A8	acetoacetate, Pyroglutamic acid	Decreased	Colon Cancer Initiating Cells	colon cancer
SLC5A8	Pyruvate, Lactate	Decreased	serum	colorectal cancer
SLC5A8	Acetate	Decreased	Blood	esophageal cancer
SLC5A8	Acetate	Decreased	esophagus	esophageal cancer
SLC5A8	Pyruvate	Decreased	urine	esophageal cancer
SLC5A8	Acetate	Decreased	frozen tissue samples	lung cancer
SLC5A8	Lactate	Decreased	urine	obesity
SLC5A8	lactate	Decreased	Serum	primary biliary cirrhosis
SLC5A8	Lactate	Decreased	cell lines	prostate cancer
SLC5A8	lactate	Decreased	cerebrospinal fluid	schizophrenia
SLC5A8	acetoacetate, 3-hydroxybutyrate	Decreased	plasma	schizophrenia
SLC5A8	Pyruvate	Decreased	serum	type 1 diabetes mellitus
SLC5A8	lactate	Difference	1st trimester pregnant woman plasma	Down syndrome
SLC5A8	Pyruvate, Lactate	Difference	renal cell	kidney cancer
SLC5A8	3-hydroxybutyrate	Difference	plasma	mental depression
SLC5A8	lactate	Difference	Brain	schizophrenia
SLC5A8	lactate	Increased	Brain	bipolar disorder
SLC5A8	Lactate	Increased	breast cancer cell lines	breast cancer
SLC5A8	lactate	Increased	blood	colorectal cancer
SLC5A8	Lactate, 5-Oxoproline	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC5A8	Acetate, 3-hydroxybutyrate	Increased	serum	colorectal cancer
SLC5A8	lactate	Increased	blood	congestive heart failure
SLC5A8	Lactate	Increased	Blood	esophageal cancer
SLC5A8	Acetate, Lactate	Increased	Plasma	esophageal cancer
SLC5A8	Pyroglutamic acid	Increased	urine	hepatocellular carcinoma
SLC5A8	Lactate	Increased	intestinal epithelial cells	inflammatory bowel disease
SLC5A8	lactate	Increased	Lung Cancer Tissue	lung cancer
SLC5A8	lactate	Increased	blood plasma	lung cancer
SLC5A8	Lactate	Increased	frozen tissue samples	lung cancer
SLC5A8	3-hydroxybutyrate	Increased	Serum	primary biliary cirrhosis
SLC5A8	Lactate	Increased	prostate cancer cell lines	prostate cancer
SLC5A8	lactate	Increased	plasma	schizophrenia
SLC5A8	Lactate	Increased	Lung tissue	squamous cell carcinoma
SLC5A8	5-Oxoproline	Increased	serum	type 1 diabetes mellitus
SLC6A1	GABA	Decreased	plasma	mental depression
SLC6A1	gamma-aminobutyric acid	Decreased	Brain	schizophrenia
SLC6A11	gamma-aminobutyric acid	Decreased	Brain	schizophrenia
SLC6A12	glycine	Decreased	Plasma	congestive heart failure
SLC6A12	gamma-aminobutyric acid	Decreased	Brain	schizophrenia

SLC6A12	glycine	Difference	Serum	cancer
SLC6A12	glycine	Difference	glioma	malignant glioma
SLC6A12	glycine	Increased	urine	Parkinson's disease
SLC6A12	glycine	Increased	blood	colorectal cancer
SLC6A12	glycine	Increased	prostate tissues	prostate cancer
SLC6A12	glycine	Increased	plasma	schizophrenia
SLC6A13	GABA	Decreased	plasma	mental depression
SLC6A14	Asparagine, Valine, Methionine, Proline, Isoleucine, Leucine	Decreased	urine	Celiac disease
SLC6A14	Asparagine, Valine, Leucine	Decreased	the plasma amino acids	autistic disorder
SLC6A14	Valine, Leucine	Decreased	serum	colorectal cancer
SLC6A14	Isoleucine	Decreased	Plasma	congestive heart failure
SLC6A14	Methionine	Decreased	Blood	esophageal cancer
SLC6A14	Leucine	Decreased	esophagus	esophageal cancer
SLC6A14	Valine, Leucine	Decreased	Serum	esophagus squamous cell carcinoma
SLC6A14	Methionine	Decreased	intestinal epithelial cells	inflammatory bowel disease
SLC6A14	Methionine	Decreased	frozen tissue samples	lung cancer
SLC6A14	Valine, Proline, Leucine	Decreased	serum	squamous cell carcinoma
SLC6A14	Proline, Leucine	Difference	Serum	fatty liver disease
SLC6A14	Valine, Proline, Leucine	Difference	Serum	hypertension
SLC6A14	Valine, Proline, Leucine	Difference	Serum	type 2 diabetes mellitus
SLC6A14	Asparagine, Valine, Isoleucine, Leucine	Increased	bladder	Bladder Cancer
SLC6A14	Methionine, Isoleucine	Increased	the plasma amino acids	autistic disorder
SLC6A14	Asparagine	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLC6A14	Proline	Increased	serum	colorectal cancer
SLC6A14	Leucine	Increased	Blood	esophageal cancer
SLC6A14	Valine, Isoleucine, Leucine	Increased	Plasma	esophageal cancer
SLC6A14	Valine, Isoleucine, Leucine	Increased	urine	obesity
SLC6A14	Methionine, Leucine	Increased	visceral and subcutaneous adipose tissue	obesity
SLC6A14	Valine	Increased	prostate tissues	prostate cancer
SLC6A14	Isoleucine	Increased	Blood	type 2 diabetes mellitus
SLC6A15	methionine	Decreased	1st trimester pregnant woman plasma	Down syndrome
SLC6A15	methionine	Decreased	serum	colorectal cancer
SLC6A15	methionine	Difference	cerebrospinal fluid	Alzheimer's disease
SLC6A15	methionine	Difference	Serum	cancer
SLC6A15	methionine	Increased	cerebrospinal fluid	major depressive disorder
SLC6A15	methionine	Increased	Cerebrospinal Fluid	malignant glioma
SLC6A18	Valine, Leucine	Decreased	urine	Celiac disease
SLC6A18	Valine, Leucine	Decreased	the plasma amino acids	autistic disorder

SLC6A18	Valine, Leucine	Decreased	serum	colorectal cancer
SLC6A18	Leucine	Decreased	esophagus	esophageal cancer
SLC6A18	Phenylalanine	Decreased	urine	esophageal cancer
SLC6A18	Valine, Leucine	Decreased	Serum	esophagus squamous cell carcinoma
SLC6A18	Valine, Leucine, Phenylalanine	Decreased	serum	squamous cell carcinoma
SLC6A18	Alpha-alanine, Serine, Leucine, Phenylalanine	Difference	Serum	fatty liver disease
SLC6A18	Alpha-alanine, Serine, Valine, Leucine, Phenylalanine	Difference	Serum	hypertension
SLC6A18	Alpha-alanine, Serine, Valine, Leucine, Phenylalanine	Difference	Serum	type 2 diabetes mellitus
SLC6A18	Valine, Leucine, Phenylalanine	Increased	bladder	Bladder Cancer
SLC6A18	Phenylalanine	Increased	the plasma amino acids	autistic disorder
SLC6A18	Leucine, Phenylalanine	Increased	Blood	esophageal cancer
SLC6A18	Valine, Leucine	Increased	Plasma	esophageal cancer
SLC6A18	Valine, Leucine, Phenylalanine	Increased	urine	obesity
SLC6A18	Serine, Leucine	Increased	visceral and subcutaneous adipose tissue	obesity
SLC6A18	Valine	Increased	prostate tissues	prostate cancer
SLC6A18	Serine	Increased	serum	type 1 diabetes mellitus
SLC6A4	5-HT	Decreased	Blood	schizophrenia
SLC6A4	Serotonin	Increased	Plasma	alcohol abuse
SLC6A5	glycine	Decreased	Plasma	congestive heart failure
SLC6A5	glycine	Difference	Serum	cancer
SLC6A5	glycine	Difference	glioma	malignant glioma
SLC6A5	glycine	Increased	urine	Parkinson's disease
SLC6A5	glycine	Increased	blood	colorectal cancer
SLC6A5	glycine	Increased	prostate tissues	prostate cancer
SLC6A5	glycine	Increased	plasma	schizophrenia
SLC6A6	Taurine	Decreased	urine	Down syndrome
SLC6A6	Taurine	Decreased	intestinal epithelial cells	inflammatory bowel disease
SLC6A6	Taurine	Difference	Serum	fatty liver disease
SLC6A6	Taurine	Difference	Serum	hypertension
SLC6A6	Taurine	Difference	Serum	type 2 diabetes mellitus
SLC6A6	Taurine	Increased	frozen tissue samples	lung cancer
SLC6A9	glycine	Decreased	Plasma	congestive heart failure
SLC6A9	glycine	Difference	Serum	cancer
SLC6A9	glycine	Difference	glioma	malignant glioma
SLC6A9	glycine	Increased	urine	Parkinson's disease
SLC6A9	glycine	Increased	blood	colorectal cancer
SLC6A9	glycine	Increased	prostate tissues	prostate cancer

SLC6A9	glycine	Increased	plasma	schizophrenia
SLC7A1	lysine	Decreased	Serum	primary biliary cirrhosis
SLC7A1	lysine, arginine, Histidine	Decreased	Prostate cell lines	prostate cancer
SLC7A1	lysine	Difference	Plasma	Alzheimer's disease
SLC7A1	arginine	Difference	Serum	cancer
SLC7A1	Histidine	Difference	Serum	fatty liver disease
SLC7A1	Histidine	Difference	Serum	hypertension
SLC7A1	lysine	Difference	glioma	malignant glioma
SLC7A1	Histidine	Difference	Serum	type 2 diabetes mellitus
SLC7A1	Histidine	Increased	bladder	Bladder Cancer
SLC7A1	Histidine	Increased	the plasma amino acids	autistic disorder
SLC7A1	lysine	Increased	Cerebrospinal Fluid	malignant glioma
SLC7A1	Histidine	Increased	urine	obesity
SLC7A1	Histidine	Increased	prostate tissues	prostate cancer
SLC7A10	valine	Decreased	Plasma	Huntington's disease
SLC7A10	alanine	Decreased	breast cancer cell lines	breast cancer
SLC7A10	alanine	Decreased	serum	colorectal cancer
SLC7A10	glycine	Decreased	Plasma	congestive heart failure
SLC7A10	alanine, valine	Decreased	blood plasma	lung cancer
SLC7A10	alanine	Decreased	Serum	primary biliary cirrhosis
SLC7A10	alanine	Decreased	prostate cancer cell lines	prostate cancer
SLC7A10	alanine	Decreased	cerebrospinal fluid	schizophrenia
SLC7A10	glycine, alanine	Difference	Serum	cancer
SLC7A10	glycine, alanine	Difference	glioma	malignant glioma
SLC7A10	valine	Increased	CSF	Alzheimer's disease
SLC7A10	glycine	Increased	urine	Parkinson's disease
SLC7A10	alanine	Increased	Colon Cancer Initiating Cells	colon cancer
SLC7A10	glycine	Increased	blood	colorectal cancer
SLC7A10	alanine	Increased	Prostate cell lines	prostate cancer
SLC7A10	glycine	Increased	prostate tissues	prostate cancer
SLC7A10	glycine, alanine	Increased	plasma	schizophrenia
SLC7A10	valine	Increased	urine	schizophrenia
SLC7A10	alanine, valine	Increased	Blood	type 2 diabetes mellitus
SLCO2A1	Lactate	Decreased	urine	Celiac disease
SLCO2A1	Lactate	Decreased	serum	colorectal cancer
SLCO2A1	Lactate	Decreased	urine	obesity
SLCO2A1	lactate	Decreased	Serum	primary biliary cirrhosis
SLCO2A1	Lactate	Decreased	cell lines	prostate cancer
SLCO2A1	lactate	Decreased	cerebrospinal fluid	schizophrenia
SLCO2A1	lactate	Difference	1st trimester pregnant woman plasma	Down syndrome

SLCO2A1	Lactate	Difference	renal cell	kidney cancer
SLCO2A1	lactate	Difference	Brain	schizophrenia
SLCO2A1	lactate	Increased	Brain	bipolar disorder
SLCO2A1	Lactate	Increased	breast cancer cell lines	breast cancer
SLCO2A1	lactate	Increased	blood	colorectal cancer
SLCO2A1	Lactate	Increased	colorectal cancer and adjacent normal controls	colorectal cancer
SLCO2A1	lactate	Increased	blood	congestive heart failure
SLCO2A1	Lactate	Increased	Blood	esophageal cancer
SLCO2A1	Lactate	Increased	Plasma	esophageal cancer
SLCO2A1	Lactate	Increased	intestinal epithelial cells	inflammatory bowel disease
SLCO2A1	lactate	Increased	Lung Cancer Tissue	lung cancer
SLCO2A1	lactate	Increased	blood plasma	lung cancer
SLCO2A1	Lactate	Increased	frozen tissue samples	lung cancer
SLCO2A1	Lactate	Increased	prostate cancer cell lines	prostate cancer
SLCO2A1	lactate	Increased	plasma	schizophrenia
SLCO2A1	Lactate	Increased	Lung tissue	squamous cell carcinoma
SLCO2B1	glutamate	Decreased	breast and blood	breast cancer
SLCO2B1	glutamate	Decreased	serum	colorectal cancer
SLCO2B1	glutamate	Difference	Serum	cancer
SLCO2B1	glutamate	Difference	Brain	schizophrenia
SLCO2B1	glutamate	Increased	Brain	bipolar disorder
SLCO2B1	glutamate	Increased	Plasma	congestive heart failure
SLCO2B1	glutamate	Increased	cell	non-small cell lung carcinoma

SUPPLEMENTARY TABLE 2: Possible new substrates for respective SLC transporters and confirmed substrates of the corresponding SLC

SLC	UniProt ID	Possible new substrate	MetaboliteID (Possible new substrates)	Confirmed substrates
FLVCR1	Q9Y5Y0	D-Galactose	HMDB0000143	D-glucose, glucose, Glucose
FLVCR1	Q9Y5Y0	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
FLVCR1	Q9Y5Y0	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
FLVCR1	Q9Y5Y0	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
MFSD4B	Q5TF39	D-Galactose	HMDB0000143	D-glucose, glucose, Glucose
MFSD4B	Q5TF39	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
MFSD4B	Q5TF39	D-Ribose	HMDB0000283	alpha-Me-glucose
MFSD4B	Q5TF39	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC15A1	P46059	phenylalanyltryptophan	HMDB0029006	Trp-Trp-Trp
SLC15A1	P46059	tryptophylglutamate	HMDB0029082	Trp-Glu-Asp
SLC15A1	P46059	glutamyltyrosine	HMDB0028831	Glu-Phe-Tyr
SLC15A1	P46059	gamma-glutamylphenylalanine	HMDB0029156	Glu-Phe-Tyr
SLC15A1	P46059	homocarnosine	HMDB0000745	Carnosine, POLAPREZINC, carnosine
SLC15A4	Q8N697	homocarnosine	HMDB0000745	Carnosine, POLAPREZINC, carnosine
SLC17A5	Q9NRA2	Ribonic acid	HMDB0000867	gluconate, Gluconic acid
SLC17A5	Q9NRA2	D-(+)-Galacturonic acid 1	HMDB0002545	glucuronic acid, D-glucuronic acid
SLC17A9	Q9BYT1	ADP-ribose	HMDB0001178	ADP
SLC1A4	P43007	Pipecolate	HMDB0000070	Proline, L-proline, proline
SLC22A1	O15245	spermine	HMDB0001256	spermidine, Spermidine
SLC22A1	O15245	Cytidine	HMDB0000089	Cytarabine, cytarabine
SLC22A11	Q9NSA0	cholic acid	HMDB0000619	bile salts
SLC22A16	Q86VW1	spermine	HMDB0001256	spermidine, Spermidine
SLC22A4	Q9H015	Cytidine	HMDB0000089	Cytarabine, cytarabine
SLC25A15	Q9Y619	Homoarginine	HMDB0000670	Arginine, L-arginine, arginine
SLC25A15	Q9Y619	Ornithine	HMDB0003374	Lysine, L-lysine, lysine
SLC25A16	P16260	Heptadecanoic acid	HMDB0006497	CoA, CoA and congeners, Coenzyme A
SLC25A17	O43808	ADP-ribose	HMDB0001178	NAD+
SLC25A17	O43808	Riboflavin	HMDB0001520	ADP
SLC25A17	O43808	Heptadecanoic acid	HMDB0006497	FMN
SLC25A2	Q9BXI2	Homoarginine	HMDB0000670	CoA, CoA and congeners, Coenzyme A
SLC25A21	Q9BQT8	azelaic acid	HMDB0000784	Arginine, L-arginine, arginine
SLC25A21	Q9BQT8	tetradecanedioate	HMDB0000872	Pimelate
SLC25A23	Q9BV35	ADP-ribose	HMDB0001178	Pimelate
SLC25A24	Q6NUK1	ADP-ribose	HMDB0001178	ADP
SLC25A25	Q6KCM7	ADP-ribose	HMDB0001178	ADP
SLC25A25	Q6KCM7	ADP	HMDB0001341	ATP, ATP-Mg2+

SLC25A29	Q8N8R3	Ornithine	HMDB0003374	ATP, ATP-Mg2+
SLC25A31	Q9H0C2	ADP-ribose	HMDB0001178	Lysine, L-lysine, lysine
SLC25A33	Q9BSK2	UTP	HMDB0000285	ADP
SLC25A36	Q96CQ1	inosinic acid	HMDB0000175	UDP
SLC25A36	Q96CQ1	UTP	HMDB0000285	IDP
SLC25A4	P12235	ADP-ribose	HMDB0001178	UDP
SLC25A42	Q86VD7	ADP-ribose	HMDB0001178	ADP
SLC25A42	Q86VD7	Heptadecanoic acid	HMDB0006497	ADP
SLC25A5	P05141	ADP-ribose	HMDB0001178	CoA, CoA and congeners, Coenzyme A
SLC25A6	P12236	ADP-ribose	HMDB0001178	ADP
SLC27A2	O14975	Stearate	HMDB0000827	ADP
SLC27A2	O14975	Palmitelaidic acid	HMDB0012328	Palmitate
SLC27A2	O14975	Pelargonic acid	HMDB0000847	Oleate
SLC27A2	O14975	Myristoleate (14:1n5)	HMDB0002000	Palmitate
SLC27A2	O14975	n-Eicosanoic acid	HMDB0002212	Oleate
SLC27A2	O14975	margarate	HMDB0002259	Palmitate
SLC27A2	O14975	Palmitoleate	HMDB0003229	Palmitate
SLC27A2	O14975	mead acid	HMDB0010378	Oleate
SLC27A2	O14975	Docosadienoate (22:2n6)	HMDB0061714	Linoleate
SLC27A2	O14975	1-dihomo-linoleoylglycerophosphocholine (20:2n6)	HMDB0061864	Linoleate
SLC27A2	O14975	5-dodecenoate (12:1n7)	HMDB0000529	Oleate
SLC27A2	O14975	arachidonic acid	HMDB0001043	Oleate
SLC27A2	O14975	Capric acid	HMDB0000511	Linoleate
SLC27A2	O14975	caproate	HMDB0000535	Palmitate
SLC27A2	O14975	10-nonadecenoate (19:1n9)	HMDB0013622	Palmitate
SLC27A2	O14975	Adrenic acid	HMDB0002226	Oleate
SLC27A2	O14975	eicosenoate	HMDB0002231	Linoleate
SLC27A2	O14975	cis-vaccenate (18:1n7)	HMDB0003231	Oleate
SLC27A2	O14975	Laurate	HMDB0000638	Oleate
SLC27A2	O14975	docosapentaenoate (n6 DPA; 22:5n6)	HMDB0001976	Palmitate
SLC27A2	O14975	Myristate	HMDB0000806	Linoleate
SLC27A2	O14975	nonadecanoate	HMDB0000772	Palmitate
SLC27A2	O14975	Pentadecanoic acid	HMDB0000826	Palmitate
SLC27A4	Q6P1M0	Stearate	HMDB0000827	Palmitate
SLC27A4	Q6P1M0	Palmitelaidic acid	HMDB0012328	Palmitate
SLC27A4	Q6P1M0	Pelargonic acid	HMDB0000847	Oleate
SLC27A4	Q6P1M0	Myristoleate (14:1n5)	HMDB0002000	Palmitate
SLC27A4	Q6P1M0	n-Eicosanoic acid	HMDB0002212	Oleate
SLC27A4	Q6P1M0	margarate	HMDB0002259	Palmitate

SLC27A4	Q6P1M0	Palmitoleate	HMDB0003229	Palmitate
SLC27A4	Q6P1M0	Docosadienoate (22:2n6)	HMDB0061714	Oleate
SLC27A4	Q6P1M0	1-dihomo-linoleoylglycerophosphocholine (20:2n6)	HMDB0061864	Oleate
SLC27A4	Q6P1M0	5-dodecenoate (12:1n7)	HMDB0000529	Oleate
SLC27A4	Q6P1M0	Capric acid	HMDB0000511	Oleate
SLC27A4	Q6P1M0	caproate	HMDB0000535	Palmitate
SLC27A4	Q6P1M0	10-nonadecenoate (19:1n9)	HMDB0013622	Palmitate
SLC27A4	Q6P1M0	Adrenic acid	HMDB0002226	Oleate
SLC27A4	Q6P1M0	eicosenoate	HMDB0002231	Oleate
SLC27A4	Q6P1M0	cis-vaccenate (18:1n7)	HMDB0003231	Oleate
SLC27A4	Q6P1M0	Laurate	HMDB0000638	Oleate
SLC27A4	Q6P1M0	Linoleic acid	HMDB0000673	Palmitate
SLC27A4	Q6P1M0	Myristate	HMDB0000806	Oleate
SLC27A4	Q6P1M0	nonadecanoate	HMDB0000772	Palmitate
SLC27A4	Q6P1M0	Pentadecanoic acid	HMDB0000826	Palmitate
SLC27A6	Q9Y2P4	Stearate	HMDB0000827	Palmitate
SLC27A6	Q9Y2P4	Palmitelaidic acid	HMDB0012328	Palmitate
SLC27A6	Q9Y2P4	Pelargonic acid	HMDB0000847	Oleate
SLC27A6	Q9Y2P4	Myristoleate (14:1n5)	HMDB0002000	Palmitate
SLC27A6	Q9Y2P4	n-Eicosanoic acid	HMDB0002212	Oleate
SLC27A6	Q9Y2P4	margarate	HMDB0002259	Palmitate
SLC27A6	Q9Y2P4	Palmitoleate	HMDB0003229	Palmitate
SLC27A6	Q9Y2P4	mead acid	HMDB0010378	Oleate
SLC27A6	Q9Y2P4	Docosadienoate (22:2n6)	HMDB0061714	Linoleate
SLC27A6	Q9Y2P4	1-dihomo-linoleoylglycerophosphocholine (20:2n6)	HMDB0061864	Linoleate
SLC27A6	Q9Y2P4	5-dodecenoate (12:1n7)	HMDB0000529	Oleate
SLC27A6	Q9Y2P4	arachidonic acid	HMDB0001043	Oleate
SLC27A6	Q9Y2P4	Capric acid	HMDB0000511	Linoleate
SLC27A6	Q9Y2P4	caproate	HMDB0000535	Palmitate
SLC27A6	Q9Y2P4	10-nonadecenoate (19:1n9)	HMDB0013622	Palmitate
SLC27A6	Q9Y2P4	Adrenic acid	HMDB0002226	Oleate
SLC27A6	Q9Y2P4	eicosenoate	HMDB0002231	Linoleate
SLC27A6	Q9Y2P4	cis-vaccenate (18:1n7)	HMDB0003231	Oleate
SLC27A6	Q9Y2P4	Laurate	HMDB0000638	Oleate
SLC27A6	Q9Y2P4	docosapentaenoate (n6 DPA; 22:5n6)	HMDB0001976	Palmitate
SLC27A6	Q9Y2P4	Myristate	HMDB0000806	Linoleate
SLC27A6	Q9Y2P4	nonadecanoate	HMDB0000772	Palmitate
SLC27A6	Q9Y2P4	Pentadecanoic acid	HMDB0000826	Palmitate

SLC29A1	Q99808	Cytidine	HMDB0000089	Palmitate
SLC29A2	Q14542	Cytidine	HMDB0000089	Cytarabine, cytarabine
SLC29A3	Q9BZD2	ADP-ribose	HMDB0001178	Cytarabine, cytarabine
SLC29A3	Q9BZD2	ADP	HMDB0001341	ATP, ATP-Mg2+
SLC2A1	P11166	D-Galactose	HMDB0000143	ATP, ATP-Mg2+
SLC2A1	P11166	D-Ribose	HMDB0000283	mannose, D-mannose, Mannose
SLC2A1	P11166	b-glucose	HMDB0000516	galactose, Galactose
SLC2A10	O95528	D-Galactose	HMDB0000143	galactose, Galactose
SLC2A10	O95528	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A10	O95528	D-Ribose	HMDB0000283	galactose, Galactose
SLC2A10	O95528	b-glucose	HMDB0000516	galactose, Galactose
SLC2A11	Q9BYW1	D-Tagatose	HMDB0003418	galactose, Galactose
SLC2A11	Q9BYW1	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC2A11	Q9BYW1	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A11	Q9BYW1	arabinose	HMDB0029942	D-glucose, glucose, Glucose
SLC2A11	Q9BYW1	D-Ribose	HMDB0000283	Hexoses
SLC2A11	Q9BYW1	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC2A12	Q8TD20	D-Galactose	HMDB0000143	D-glucose, glucose, Glucose
SLC2A12	Q8TD20	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A12	Q8TD20	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC2A12	Q8TD20	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC2A14	Q8TDB8	D-Galactose	HMDB0000143	D-glucose, glucose, Glucose
SLC2A14	Q8TDB8	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A14	Q8TDB8	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC2A14	Q8TDB8	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC2A2	P11168	D-Tagatose	HMDB0003418	D-glucose, glucose, Glucose
SLC2A2	P11168	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC2A2	P11168	D-Ribose	HMDB0000283	mannose, D-mannose, Mannose
SLC2A2	P11168	b-glucose	HMDB0000516	galactose, Galactose
SLC2A3	P11169	D-Galactose	HMDB0000143	galactose, Galactose
SLC2A3	P11169	arabinose	HMDB0029942	mannose, D-mannose, Mannose
SLC2A3	P11169	D-Ribose	HMDB0000283	xylose, Xylose
SLC2A3	P11169	b-glucose	HMDB0000516	galactose, Galactose
SLC2A4	P14672	D-Galactose	HMDB0000143	galactose, Galactose
SLC2A4	P14672	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A4	P14672	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC2A4	P14672	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC2A5	P22732	D-Tagatose	HMDB0003418	D-glucose, glucose, Glucose
SLC2A5	P22732	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC2A5	P22732	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose

SLC2A5	P22732	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC2A5	P22732	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC2A6	Q9UGQ3	D-Galactose	HMDB0000143	D-glucose, glucose, Glucose
SLC2A6	Q9UGQ3	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A6	Q9UGQ3	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC2A6	Q9UGQ3	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC2A7	Q6PXP3	D-Tagatose	HMDB0003418	D-glucose, glucose, Glucose
SLC2A7	Q6PXP3	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC2A7	Q6PXP3	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A7	Q6PXP3	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC2A7	Q6PXP3	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC2A8	Q9NY64	D-Tagatose	HMDB0003418	D-glucose, glucose, Glucose
SLC2A8	Q9NY64	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC2A8	Q9NY64	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A8	Q9NY64	D-Ribose	HMDB0000283	galactose, Galactose
SLC2A8	Q9NY64	b-glucose	HMDB0000516	galactose, Galactose
SLC2A9	Q9NRM0	D-Tagatose	HMDB0003418	galactose, Galactose
SLC2A9	Q9NRM0	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC2A9	Q9NRM0	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC2A9	Q9NRM0	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC2A9	Q9NRM0	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC33A1	O00400	Heptadecanoic acid	HMDB0006497	D-glucose, glucose, Glucose
SLC35B1	P78383	ADP-ribose	HMDB0001178	CoA, CoA and congeners, Coenzyme A
SLC35B1	P78383	ADP	HMDB0001341	ATP, ATP-Mg2+
SLC35D1	Q9NTN3	UDP-N-acetyl-glucosamine	HMDB0000290	ATP, ATP-Mg2+
SLC36A1	Q7Z2H8	Pipecolate	HMDB0000070	UDP-N-acetyl-D-galactosamine, UDP-GlcNAc
SLC36A2	Q495M3	Pipecolate	HMDB0000070	Proline, L-proline, proline
SLC36A4	Q6YBV0	Pipecolate	HMDB0000070	Proline, L-proline, proline
SLC37A1	P57057	D-Galactose	HMDB0000143	Proline, L-proline, proline
SLC37A1	P57057	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC37A1	P57057	mannose-6-phosphate	HMDB0001078	D-glucose, glucose, Glucose
SLC37A1	P57057	D-Ribose	HMDB0000283	glucose 6-phosphate
SLC37A1	P57057	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC37A2	Q8TED4	mannose-6-phosphate	HMDB0001078	D-glucose, glucose, Glucose
SLC37A4	O43826	mannose-6-phosphate	HMDB0001078	glucose 6-phosphate
SLC38A2	Q96QD8	Pipecolate	HMDB0000070	glucose 6-phosphate
SLC38A4	Q969I6	Pipecolate	HMDB0000070	Proline, L-proline, proline
SLC38A4	Q969I6	Homoarginine	HMDB0000670	Proline, L-proline, proline
SLC38A4	Q969I6	Ornithine	HMDB0003374	Arginine, L-arginine, arginine
SLC38A9	Q8NBW4	Pipecolate	HMDB0000070	Lysine, L-lysine, lysine

SLC38A9	Q8NBW4	Homoarginine	HMDB0000670	Proline, L-proline, proline
SLC38A9	Q8NBW4	Ornithine	HMDB0003374	Arginine, L-arginine, arginine
SLC45A1	Q9Y2W3	D-Galactose	HMDB0000143	Lysine, L-lysine, lysine
SLC45A1	Q9Y2W3	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC45A1	Q9Y2W3	arabinose	HMDB0029942	galactose, Galactose
SLC45A1	Q9Y2W3	D-Ribose	HMDB0000283	L-glucose
SLC45A1	Q9Y2W3	b-glucose	HMDB0000516	galactose, Galactose
SLC45A2	Q9UMX9	D-Tagatose	HMDB0003418	galactose, Galactose
SLC45A2	Q9UMX9	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC45A2	Q9UMX9	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC45A2	Q9UMX9	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC45A2	Q9UMX9	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC45A3	Q96JT2	D-Tagatose	HMDB0003418	D-glucose, glucose, Glucose
SLC45A3	Q96JT2	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC45A3	Q96JT2	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC45A3	Q96JT2	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC45A3	Q96JT2	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC45A4	Q5BKX6	D-Tagatose	HMDB0003418	D-glucose, glucose, Glucose
SLC45A4	Q5BKX6	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC45A4	Q5BKX6	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC45A4	Q5BKX6	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC45A4	Q5BKX6	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC50A1	Q9BRV3	D-Galactose	HMDB0000143	D-glucose, glucose, Glucose
SLC50A1	Q9BRV3	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC50A1	Q9BRV3	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC50A1	Q9BRV3	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC5A1	P13866	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC5A1	P13866	D-Ribose	HMDB0000283	D-galactose
SLC5A1	P13866	Rhamnose	HMDB0000849	D-galactose
SLC5A1	P13866	b-glucose	HMDB0000516	Fucose
SLC5A10	A0PJK1	D-Tagatose	HMDB0003418	D-galactose
SLC5A10	A0PJK1	D-Ribose	HMDB0000283	D-fructose, fructose, Fructose
SLC5A10	A0PJK1	b-glucose	HMDB0000516	D-galactose
SLC5A2	P31639	D-Galactose	HMDB0000143	D-galactose
SLC5A2	P31639	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC5A2	P31639	D-Ribose	HMDB0000283	galactose, Galactose
SLC5A2	P31639	b-glucose	HMDB0000516	galactose, Galactose
SLC5A3	P53794	D-Galactose	HMDB0000143	galactose, Galactose
SLC5A3	P53794	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC5A3	P53794	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose

SLC5A3	P53794	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC5A4	Q9NY91	D-Galactose	HMDB0000143	D-glucose, glucose, Glucose
SLC5A4	Q9NY91	D-Mannose	HMDB0000169	D-glucose, glucose, Glucose
SLC5A4	Q9NY91	D-Ribose	HMDB0000283	D-glucose, glucose, Glucose
SLC5A4	Q9NY91	b-glucose	HMDB0000516	D-glucose, glucose, Glucose
SLC5A9	Q2M3M2	D-Tagatose	HMDB0003418	D-glucose, glucose, Glucose
SLC5A9	Q2M3M2	D-Galactose	HMDB0000143	D-fructose, fructose, Fructose
SLC5A9	Q2M3M2	D-Ribose	HMDB0000283	mannose, D-mannose, Mannose
SLC5A9	Q2M3M2	b-glucose	HMDB0000516	mannose, D-mannose, Mannose
SLC66A1		Homoarginine	HMDB0000670	mannose, D-mannose, Mannose
SLC66A1		Ornithine	HMDB0003374	Arginine, L-arginine, arginine
SLC6A12	P48065	Pipecolate	HMDB0000070	Lysine, L-lysine, lysine
SLC6A14	Q9UN76	Pipecolate	HMDB0000070	Proline, L-proline, proline
SLC6A15	Q9H2J7	Pipecolate	HMDB0000070	Proline, L-proline, proline
SLC6A18	Q96N87	Homoarginine	HMDB0000670	Proline, L-proline, proline
SLC6A7	Q99884	Pipecolate	HMDB0000070	Arginine, L-arginine, arginine
SLC7A1	P30825	Homoarginine	HMDB0000670	Proline, L-proline, proline
SLC7A1	P30825	Ornithine	HMDB0003374	Arginine, L-arginine, arginine
SLC7A2	P52569	Homoarginine	HMDB0000670	Lysine, L-lysine, lysine
SLC7A2	P52569	Ornithine	HMDB0003374	Arginine, L-arginine, arginine
SLC7A3	Q8WY07	Homoarginine	HMDB0000670	Lysine, L-lysine, lysine
SLC7A3	Q8WY07	Ornithine	HMDB0003374	Arginine, L-arginine, arginine
SLC7A6	Q92536	Homoarginine	HMDB0000670	Lysine, L-lysine, lysine
SLC7A7	Q9UM01	Homoarginine	HMDB0000670	Arginine, L-arginine, arginine
SLC7A8	Q9UHI5	Pipecolate	HMDB0000070	Arginine, L-arginine, arginine
SLC7A8	Q9UHI5	Homoarginine	HMDB0000670	Proline, L-proline, proline
SLC7A9	P82251	Homoarginine	HMDB0000670	Arginine, L-arginine, arginine
SLC7A9	P82251	Ornithine	HMDB0003374	Arginine, L-arginine, arginine
SLCO4A1	Q96BD0	cholic acid	HMDB0000619	Lysine, L-lysine, lysine
SLCO4C1	Q6ZQN7	Arginine	HMDB0000517	bile salts
SV2A	Q7L0J3	D-Mannose	HMDB0000169	Homoarginine, L-homoarginine
SV2A	Q7L0J3	D-Ribose	HMDB0000283	galactose, Galactose
SV2A	Q7L0J3	D-Glucose	HMDB0000122	galactose, Galactose
SV2A	Q7L0J3	b-glucose	HMDB0000516	galactose, Galactose

SUPPLEMENTARY TABLE 3: Table for updating the SLC substrates list with InChI codes that were missing

SLC	Substrate	InChI code	Source
ANKH	Pyrophosphate	InChI=1S/H4O7P2/c1-8(2,3)7-9(4,5)6/h(H2,1,2,3)(H2,4,5,6)/p-4	Bioparadigms
ANKH	Pyrophosphate	InChI=1S/H4O7P2/c1-8(2,3)7-9(4,5)6/h(H2,1,2,3)(H2,4,5,6)/p-4	GtoPdb
ANKH	pyrophosphate	InChI=1S/H4O7P2/c1-8(2,3)7-9(4,5)6/h(H2,1,2,3)(H2,4,5,6)/p-4	TCDB
FLVCR1	arsenite	InChI=1S/AsH3O3/c2-1(3)4/h2-4H	TCDB
LETM1	Ca2+	InChI=1S/Ca/q+2	GtoPdb
LETM1	Gd3+	InChI=1S/Gd/q+3	TCDB
LETM1	K+	InChI=1S/K/q+1	GtoPdb
LETM1	H+	InChI=1S/p+1	GtoPdb
MFSD4B	alpha-Me-glucose	InChI=1S/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2	GtoPdb
MFSD4B	α-Me-glucose	InChI=1S/C7H14O6/c1-12-7-6(11)5(10)4(9)3(2-8)13-7/h3-11H,2H2,1H3/t3-,4-,5+,6-,7+/m1/s1	Bioparadigms
MFSD4B	D-glucose	InChI=1S/C7H14O6/c1-12-7-6(11)5(10)4(9)3(2-8)13-7/h3-11H,2H2,1H3/t3-,4-,5+,6-,7+/m1/s1	GtoPdb
MFSD5	Molybdate	InChI=1S/Mo.4O/q;;;2*-1	Bioparadigms
MFSD5	molybdate	InChI=1S/Mo.4O/q;;;2*-1	GtoPdb
NIPA1	Co2+	InChI=1S/Co/q+2	GtoPdb
NIPA1	Fe2+	InChI=1S/Fe/q+2	GtoPdb
NIPA1	Sr2+	InChI=1S/Sr/q+2	GtoPdb
NIPAL1	Ba2+	InChI=1S/Ba/q+2	GtoPdb
NIPAL1	Cu2+	InChI=1S/Cu/q+2	GtoPdb
NIPAL1	Fe2+	InChI=1S/Fe/q+2	GtoPdb
NIPAL1	Mg2+	InChI=1S/Mg/q+2	GtoPdb
NIPAL1	Sr2+	InChI=1S/Sr/q+2	GtoPdb
NIPAL2	Ba2+	InChI=1S/Ba/q+2	GtoPdb
NIPAL2	Mg2+	InChI=1S/Mg/q+2	GtoPdb
NIPAL2	Sr2+	InChI=1S/Sr/q+2	GtoPdb
SLC10A1	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLC11A2	V3+	InChI=1S/V/q+3	CeMM
SLC12A1	Cl-	InChI=1S/ClH/h1H/p-1	TCDB
SLC12A1	Cl-	InChI=1S/ClH/h1H/p-1	TCDB
SLC12A1	K+	InChI=1S/K/q+1	TCDB
SLC12A1	Na+	InChI=1S/Na/q+1	TCDB
SLC13A1	thiosulfate	InChI=1S/H2O3S2/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC13A1	S2O32-	InChI=1S/H2O3S2/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC13A1	Thiosulfate	InChI=1S/H2O3S2/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC13A1	sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC13A1	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC13A1	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC13A1	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC13A1	selenate	InChI=1S/H2O4Se/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC13A1	SeO42-	InChI=1S/H2O4Se/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC13A3	Fumarate	InChI=1S/C4H4O4/c5-3(6)1-2-4(7)8/h1-2H,(H,5,6)(H,7,8)/b2-1+	CeMM

SLC13A4	Thiosulfate	InChI=1S/H2O3S2/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC13A4	sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC13A4	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC13A4	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC13A4	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC13A4	Selenate	InChI=1S/H2O4Se/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC15A1	protons	InChI=1S/p+1	Bioparadigms
SLC15A1	protons	InChI=1S/p+1	GtoPdb
SLC15A2	protons	InChI=1S/p+1	Bioparadigms
SLC15A2	protons	InChI=1S/p+1	GtoPdb
SLC15A3	Tri-DAP	InChI=1S/C15H26N4O8/c1-7(16)12(21)19-10(15(26)27)5-6-11(20)18-9(14(24)25)4-2-3-8(17)13(22)23/h7-10H,2-6,16-17H2,1H3,(H,18,20)(H,19,21)(H,22,23)(H,24,25)(H,26,27)/t7-,8+,9-,10+/m0/s1	GtoPdb
SLC15A3	protons	InChI=1S/p+1	Bioparadigms
SLC15A3	protons	InChI=1S/p+1	GtoPdb
SLC15A4	Tri-DAP	InChI=1S/C15H26N4O8/c1-7(16)12(21)19-10(15(26)27)5-6-11(20)18-9(14(24)25)4-2-3-8(17)13(22)23/h7-10H,2-6,16-17H2,1H3,(H,18,20)(H,19,21)(H,22,23)(H,24,25)(H,26,27)/t7-,8+,9-,10+/m0/s1	CeMM
SLC15A4	Tri-DAP	InChI=1S/C15H26N4O8/c1-7(16)12(21)19-10(15(26)27)5-6-11(20)18-9(14(24)25)4-2-3-8(17)13(22)23/h7-10H,2-6,16-17H2,1H3,(H,18,20)(H,19,21)(H,22,23)(H,24,25)(H,26,27)/t7-,8+,9-,10+/m0/s1	GtoPdb
SLC15A4	protons	InChI=1S/p+1	Bioparadigms
SLC16A10	Thyroxine (T4)	InChI=1S/C15H11I4NO4/c16-8-4-7(5-9(17)13(8)21)24-14-10(18)1-6(2-11(14)19)3-12(20)15(22)23/h1-2,4-5,12,21H,3,20H2,(H,22,23)	CeMM
SLC16A2	Thyroxine (T4)	InChI=1S/C15H11I4NO4/c16-8-4-7(5-9(17)13(8)21)24-14-10(18)1-6(2-11(14)19)3-12(20)15(22)23/h1-2,4-5,12,21H,3,20H2,(H,22,23)	CeMM
SLC16A2	3,3',5'-triiodothyronine (rT3)	InChI=1S/C15H12I3NO4/c16-9-3-7(4-12(19)15(21)22)1-2-13(9)23-8-5-10(17)14(20)11(18)6-8/h1-3,5-6,12,20H,4,19H2,(H,21,22)	CeMM
SLC16A2	Diiodothyronine (T2)	InChI=1S/C15H13I2NO4/c16-10-7-9(2-3-13(10)19)22-14-4-1-8(5-11(14)17)6-12(18)15(20)21/h1-5,7,12,19H,6,18H2,(H,20,21)	CeMM
SLC16A3	Acetate	InChI=1S/C2H4O2/c1-2(3)4/h1H3,(H,3,4)	CeMM
SLC17A1	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	Bioparadigms
SLC17A1	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	GtoPdb
SLC17A1	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC17A3	Estradiol-17-beta-glucuronide	InChI=1S/C24H32O8/c1-24-9-8-14-13-5-3-12(25)10-11(13)2-4-15(14)16(24)6-7-17(24)31-23-20(28)18(26)19(27)21(32-23)22(29)30/h3,5,10,14-21,23,25-28H,2,4,6-9H2,1H3,(H,29,30)/t14-,15-,16+,17+,18+,19+,20-,21+,23-,24+/m1/s1	CeMM
SLC17A3	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC17A4	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC17A5	NO3-	InChI=1S/NO3/c2-1(3)4/q-1	CeMM
SLC17A6	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC17A6	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC17A7	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC17A7	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC18B1	CDg16	InChI=1S/C27H23N5O3/c1-17-2-4-18(5-3-17)15-32(27(34)25-10-11-29-35-25)16-26(33)30-22-9-7-20-12-19-6-8-21(28)13-23(19)31-24(20)14-22/h2-14H,15-16,28H2,1H3,(H,30,33)	CeMM
SLC1A1	D/L-Asp	InChI=1S/C4H8N2O3/c5-2(4(8)9)1-3(6)7/h2H,1,5H2,(H2,6,7)(H,8,9)	Bioparadigms
SLC1A2	D/L-Asp	InChI=1S/C4H8N2O3/c5-2(4(8)9)1-3(6)7/h2H,1,5H2,(H2,6,7)(H,8,9)	Bioparadigms

SLC1A3	D/L-Asp	InChI=1S/C4H8N2O3/c5-2(4(8)9)1-3(6)7/h2H,1,5H2,(H2,6,7)(H,8,9)	Bioparadigms
SLC1A6	D/L-Asp	InChI=1S/C4H8N2O3/c5-2(4(8)9)1-3(6)7/h2H,1,5H2,(H2,6,7)(H,8,9)	Bioparadigms
SLC1A7	D/L-Asp	InChI=1S/C4H8N2O3/c5-2(4(8)9)1-3(6)7/h2H,1,5H2,(H2,6,7)(H,8,9)	Bioparadigms
SLC20A1	AsO43-	InChI=1S/AsH3O4/c2-1(3,4)5/h(H3,2,3,4,5)	GtoPdb
SLC20A1	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC20A1	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	GtoPdb
SLC20A1	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC20A2	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC20A2	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	GtoPdb
SLC20A2	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC22A1	Cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	Metrabase
SLC22A1	1-methyl-4-phenylpyridinium (MPP+)	InChI=1S/C12H12N/c1-13-9-7-12(8-10-13)11-5-3-2-4-6-11/h2-10H,1H3/q+1	CeMM
SLC22A1	4-(4-dimethylamino)styryl-N-methylpyridinium	InChI=1S/C16H19N2/c1-17(2)16-8-6-14(7-9-16)4-5-15-10-12-18(3)13-11-15/h4-13H,1-3H3/q+1	UCSF-FDA
SLC22A1	Tetraethylammonium (TEA)	InChI=1S/C8H20N/c1-5-9(6-2,7-3)8-4/h5-8H2,1-4H3/q+1	CeMM
SLC22A11	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLC22A2	cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	GtoPdb
SLC22A2	1-methyl-4-phenylpyridinium (MPP+)	InChI=1S/C12H12N/c1-13-9-7-12(8-10-13)11-5-3-2-4-6-11/h2-10H,1H3/q+1	CeMM
SLC22A2	4-(4-dimethylaminostyryl-N-methylpyridinium) (ASP+)	InChI=1S/C16H19N2/c1-17(2)16-8-6-14(7-9-16)4-5-15-10-12-18(3)13-11-15/h4-13H,1-3H3/q+1	CeMM
SLC22A2	4-(4-dimethylamino)styryl-N-methylpyridinium	InChI=1S/C16H19N2/c1-17(2)16-8-6-14(7-9-16)4-5-15-10-12-18(3)13-11-15/h4-13H,1-3H3/q+1	UCSF-FDA
SLC22A2	Tetraethylammonium (TEA)	InChI=1S/C8H20N/c1-5-9(6-2,7-3)8-4/h5-8H2,1-4H3/q+1	CeMM
SLC22A24	Estradiol-17-beta-glucuronide	InChI=1S/C24H32O8/c1-24-9-8-14-13-5-3-12(25)10-11(13)2-4-15(14)16(24)6-7-17(24)31-23-20(28)18(26)19(27)21(32-23)22(29)30/h3,5,10,14-21,23,25-28H,2,4,6-9H2,1H3,(H,29,30)/t14-,15-,16+,17+,18+,19+,20-,21+,23-,24+/m1/s1	CeMM
SLC22A24	Androstanediol glucuronide	InChI=1S/C25H40O8/c1-24-9-7-13(32-23-20(29)18(27)19(28)21(33-23)22(30)31)11-12(24)3-4-14-15-5-6-17(26)25(15,2)10-8-16(14)24/h12-21,23,26-29H,3-11H2,1-2H3,(H,30,31)/t12-,13+,14-,15-,16-,17-,18-,19-,20+,21-,23+,24-,25-/m0/s1	CeMM
SLC22A3	Cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	CeMM
SLC22A3	1-methyl-4-phenylpyridinium (MPP+)	InChI=1S/C12H12N/c1-13-9-7-12(8-10-13)11-5-3-2-4-6-11/h2-10H,1H3/q+1	CeMM
SLC22A4	Tetraethylammonium (TEA)	InChI=1S/C8H20N/c1-5-9(6-2,7-3)8-4/h5-8H2,1-4H3/q+1	CeMM
SLC22A5	Gamma-butryo-betaine	InChI=1S/C7H15NO2/c1-8(2,3)6-4-5-7(9)10/h4-6H2,1-3H3	TCDB
SLC22A5	Tetraethylammonium (TEA)	InChI=1S/C8H20N/c1-5-9(6-2,7-3)8-4/h5-8H2,1-4H3/q+1	CeMM
SLC22A6	3'-azido-3'-deoxythymidine (AZT)	InChI=1S/C10H13N5O4/c1-5-3-15(10(18)12-9(5)17)8-2-6(13-14-11)7(4-16)19-8/h3,6-8,16H,2,4H2,1H3,(H,12,17,18)/t6-,7+,8+/m0/s1	CeMM
SLC22A6	N-acetyl glutamate (NAG)	InChI=1S/C7H11NO5/c1-4(9)8-5(7(12)13)2-3-6(10)11/h5H,2-3H2,1H3,(H,8,9)(H,10,11)(H,12,13)/p-2/t5-/m0/s1	CeMM
SLC22A6	9-(2-Phosphonylmethoxyethyl) diaminopurine (PMEDAP)	InChI=1S/C8H12N5O4P/c9-7-6-8(11-3-10-7)13(4-12-6)1-2-17-5-18(14,15)16/h3-4H,1-2,5H2,(H2,9,10,11)(H2,14,15,16)	CeMM
SLC22A6	9-(2-Phosphonylmethoxy-oxyethyl)guanidine (PMEG)	InChI=1S/C8H12N5O5P/c9-8-11-6-5(7(14)12-8)10-3-13(6)1-2-18-4-19(15,16)17/h3H,1-2,4H2,(H2,15,16,17)(H3,9,11,12,14)	CeMM
SLC22A6	Tetrahydrobiopterin (BH4)	InChI=1S/C9H15N5O3/c1-3(15)6(16)4-2-11-7-5(12-4)8(17)14-9(10)13-7/h3-4,6,12,15-16H,2H2,1H3,(H4,10,11,13,14,17)	CeMM

SLC22A8	Edavarone sulfate	InChI=1S/C10H10N2O4S/c1-8-7-10(16-17(13,14)15)12(11-8)9-5-3-2-4-6-9/h2-7H,1H3,(H,13,14,15)	CeMM
SLC22A8	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLC22A9	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC25A10	thiosulphate	InChI=1S/H2O3S2/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC25A10	S2O32-	InChI=1S/H2O3S2/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC25A10	Thiosulfate	InChI=1S/H2O3S2/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC25A10	sulphate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC25A10	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC25A10	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC25A10	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	Bioparadigms
SLC25A10	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	GtoPdb
SLC25A10	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC25A16	CoA and congeners	InChI=1S/C21H36N7O16P3S/c1-21(2,16(31)19(32)24-4-3-12(29)23-5-6-48)8-41-47(38,39)44-46(36,37)40-7-11-15(43-45(33,34)35)14(30)20(42-11)28-10-27-13-17(22)25-9-26-18(13)28/h9-11,14-16,20,30-31,48H,3-8H2,1-2H3,(H,23,29)(H,24,32)(H,36,37)(H,38,39)(H2,22,25,26)(H2,33,34,35)/t11-,14-,15-,16+,20-/m1/s1	GtoPdb
SLC25A17	dPCoA	InChI=1S/C21H35N7O13P2S/c1-21(2,16(32)19(33)24-4-3-12(29)23-5-6-44)8-39-43(36,37)41-42(34,35)38-7-11-14(30)15(31)20(40-11)28-10-27-13-17(22)25-9-26-18(13)28/h9-11,14-16,20,30-32,44H,3-8H2,1-2H3,(H,23,29)(H,24,33)(H,34,35)(H,36,37)(H2,22,25,26)/t11-,14-,15-,16+,20-/m1/s1	Bioparadigms
SLC25A21	2-amino adipate	InChI=1S/C6H11NO4/c7-4(6(10)11)2-1-3-5(8)9/h4H,1-3,7H2,(H,8,9)(H,10,11)/t4-/m0/s1	TCDB
SLC25A23	ATP-Mg2+	InChI=1S/C10H16N5O13P3.Mg/c11-8-5-9(13-2-12-8)15(3-14-5)10-7(17)6(16)4(26-10)1-25-30(21,22)28-31(23,24)27-29(18,19)20;/h2-4,6-7,10,16-17H,1H2,(H,21,22)(H,23,24)(H2,11,12,13)(H2,18,19,20);/q;+2/p-4/t4-,6-,7-,10-;/m1./s1	Bioparadigms
SLC25A23	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC25A24	ATP-Mg2+	InChI=1S/C10H16N5O13P3.Mg/c11-8-5-9(13-2-12-8)15(3-14-5)10-7(17)6(16)4(26-10)1-25-30(21,22)28-31(23,24)27-29(18,19)20;/h2-4,6-7,10,16-17H,1H2,(H,21,22)(H,23,24)(H2,11,12,13)(H2,18,19,20);/q;+2/p-4/t4-,6-,7-,10-;/m1./s1	Bioparadigms
SLC25A24	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC25A24	Mg 2+	InChI=1S/Mg/q+2	TCDB
SLC25A25	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC25A25	Mg 2+	InChI=1S/Mg/q+2	TCDB
SLC25A3	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	Bioparadigms
SLC25A3	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC25A3	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC25A30	SO32-	InChI=1S/H2O3S/c1-4(2)3/h(H2,1,2,3)/p-2	CeMM
SLC25A30	S2O32-	InChI=1S/H2O3S2/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC25A30	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC25A30	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC25A42	dPCoA	InChI=1S/C21H35N7O13P2S/c1-21(2,16(32)19(33)24-4-3-12(29)23-5-6-44)8-39-43(36,37)41-42(34,35)38-7-11-14(30)15(31)20(40-11)28-10-27-13-17(22)25-9-26-18(13)28/h9-11,14-16,20,30-32,44H,3-8H2,1-2H3,(H,23,29)(H,24,33)(H,34,35)(H,36,37)(H2,22,25,26)/t11-,14-,15-,16+,20-/m1/s1	Bioparadigms
SLC26A1	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	Bioparadigms
SLC26A1	Oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	CeMM
SLC26A1	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	GtoPdb
SLC26A1	Oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	TCDB

SLC26A1	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC26A1	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC26A1	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC26A1	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC26A11	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	Bioparadigms
SLC26A11	HSO4-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-1	GtoPdb
SLC26A11	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC26A11	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC26A2	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	Bioparadigms
SLC26A2	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC26A2	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC26A2	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC26A2	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC26A3	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	Bioparadigms
SLC26A5	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	Bioparadigms
SLC26A5	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC26A6	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	Bioparadigms
SLC26A6	Oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	CeMM
SLC26A6	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	GtoPdb
SLC26A6	Oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	TCDB
SLC26A6	Hydrogen sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-1	TCDB
SLC26A6	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC26A6	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC26A6	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC26A7	Oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	TCDB
SLC26A7	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	Bioparadigms
SLC26A7	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC26A7	nitrate	InChI=1S/NO3/c2-1(3)4/q-1	TCDB
SLC26A8	Oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	CeMM
SLC26A8	oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	GtoPdb
SLC26A8	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC26A8	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	GtoPdb
SLC26A9	Oxalate	InChI=1S/C2H2O4/c3-1(4)2(5)6/h(H,3,4)(H,5,6)	CeMM
SLC26A9	SO42-	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	CeMM
SLC27A4	Bodipy-palmitate	InChI=1S/C28H43BF2N2O2/c1-21-19-23(3)32-27(21)25(28-22(2)20-24(4)33(28)29(32,30)31)17-15-13-11-9-7-5-6-8-10-12-14-16-18-26(34)35/h19-20H,5-18H2,1-4H3,(H,34,35)	CeMM
SLC28A1	3'-azido-3'-deoxythymidine (AZT)	InChI=1S/C10H13N5O4/c1-5-3-15(10(18)12-9(5)17)8-2-6(13-14-11)7(4-16)19-8/h3,6-8,16H,2,4H2,1H3,(H,12,17,18)/t6-,7+,8+/m0/s1	CeMM
SLC28A2	cladribrine	InChI=1S/C10H12ClN5O3/c11-10-14-8(12)7-9(15-10)16(3-13-7)6-1-4(18)5(2-17)19-6/h3-6,17-18H,1-2H2,(H2,12,14,15)/t4-,5+,6+/m0/s1	TCDB
SLC28A2	2'3'dideoxyinosine	InChI=1S/C10H12N4O3/c15-3-6-1-2-7(17-6)14-5-13-8-9(14)11-4-12-10(8)16/h4-7,15H,1-3H2,(H,11,12,16)/t6-,7+/m0/s1	TCDB
SLC29A4	Adenosine analogs	InChI=1S/C10H13N5O4/c11-8-5-9(13-2-12-8)15(3-14-5)10-7(18)6(17)4(1-16)19-10/h2-4,6-7,10,16-18H,1H2,(H2,11,12,13)/t4-,6-,7-,10-/m1/s1	CeMM
SLC29A4	1-methyl-4-phenylpyridinium (MPP+)	InChI=1S/C12H12N/c1-13-9-7-12(8-10-13)11-5-3-2-4-6-11/h2-10H,1H3/q+1	CeMM

SLC29A4	1-methyl-4-phenylpyridinium (MPP+) analogs	InChI=1S/C12H12N/c1-13-9-7-12(8-10-13)11-5-3-2-4-6-11/h2-10H,1H3/q+1	CeMM
SLC29A4	1-methyl-4-phenylpyridinium (MPP)	InChI=1S/C12H12N/c1-13-9-7-12(8-10-13)11-5-3-2-4-6-11/h2-10H,1H3/q+1	TCDB
SLC2A1	arsenite	InChI=1S/AsH3O3/c2-1(3)4/h2-4H	TCDB
SLC31A1	cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	Bioparadigms
SLC31A1	Cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	CeMM
SLC31A1	cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	GtoPdb
SLC31A2	cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	Bioparadigms
SLC31A2	Cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	CeMM
SLC31A2	cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	GtoPdb
SLC32A1	Glycine	InChI=1S/C2H5NO2/c3-1-2(4)5/h1,3H2,(H,4,5)	Bioparadigms
SLC32A1	GABA	InChI=1S/C4H9NO2/c5-3-1-2-4(6)7/h1-3,5H2,(H,6,7)	Bioparadigms
SLC34A1	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC34A1	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC34A2	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC34A2	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC34A3	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
SLC34A3	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC35B2	Adenosine 3'-phosphate 5'-phosphate	InChI=1S/C10H15N5O10P2/c11-8-5-9(13-2-12-8)15(3-14-5)10-6(16)7(25-27(20,21)22)4(24-10)1-23-26(17,18)19/h2-4,6-7,10,16H,1H2,(H2,11,12,13)(H2,17,18,19)(H2,20,21,22)/t4-,6-,7-,10-/m1/s1	TCDB
SLC35B2	3'-Phosphoadenosine-5'-phosphosulfate (PAPS)	InChI=1S/C10H15N5O13P2S/c11-8-5-9(13-2-12-8)15(3-14-5)10-6(16)7(27-29(17,18)19)4(26-10)1-25-30(20,21)28-31(22,23)24/h2-4,6-7,10,16H,1H2,(H,20,21)(H2,11,12,13)(H2,17,18,19)(H,22,23,24)/t4-,6-,7-,10-/m1/s1	CeMM
SLC35B3	3'-Phosphoadenosine-5'-phosphosulfate (PAPS)	InChI=1S/C10H15N5O13P2S/c11-8-5-9(13-2-12-8)15(3-14-5)10-6(16)7(27-29(17,18)19)4(26-10)1-25-30(20,21)28-31(22,23)24/h2-4,6-7,10,16H,1H2,(H,20,21)(H2,11,12,13)(H2,17,18,19)(H,22,23,24)/t4-,6-,7-,10-/m1/s1	CeMM
SLC37A1	phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	TCDB
SLC39A8	Se2+	InChI=1S/Se/q+2	CeMM
SLC46A1	N5-methyltetrafolate	InChI=1S/C20H25N7O6/c1-27-12(9-23-16-15(27)18(31)26-20(21)25-16)8-22-11-4-2-10(3-5-11)17(30)24-13(19(32)33)6-7-14(28)29/h2-5,12-13,22H,6-9H2,1H3,(H,24,30)(H,28,29)(H,32,33)(H4,21,23,25,26,31)/t12-,13-/m0/s1	GtoPdb
SLC47A1	Cisplatin	InChI=1S/2ClH.2H3N.Pt/h2*1H;2*1H3;/q;;;;+2/p-2	TCDB
SLC47A1	1-methyl-4-phenylpyridinium (MPP+)	InChI=1S/C12H12N/c1-13-9-7-12(8-10-13)11-5-3-2-4-6-11/h2-10H,1H3/q+1	CeMM
SLC47A1	N-dimethyl-4-4'-bipiridinium	InChI=1S/C12H14N2/c1-13-7-3-11(4-8-13)12-5-9-14(2)10-6-12/h3-10H,1-2H3/q+2	TCDB
SLC47A1	Tetraethylammonium (TEA)	InChI=1S/C8H20N/c1-5-9(6-2,7-3)8-4/h5-8H2,1-4H3/q+1	CeMM
SLC47A2	1-methyl-4-phenylpyridinium (MPP+)	InChI=1S/C12H12N/c1-13-9-7-12(8-10-13)11-5-3-2-4-6-11/h2-10H,1H3/q+1	CeMM
SLC47A2	Tetraethylammonium (TEA)	InChI=1S/C8H20N/c1-5-9(6-2,7-3)8-4/h5-8H2,1-4H3/q+1	CeMM
SLC4A10	NaHCO3-	InChI=1S/CH2O3.Na/c2-1(3)4;/h(H2,2,3,4);/q;+1/p-1	GtoPdb
SLC4A11	Boride	InChI=1S/B/q-3	TCDB
SLC4A11	B(OH)4-	InChI=1S/BH4O4/c2-1(3,4)5/h2-5H/q-1	CeMM
SLC4A11	borate	InChI=1S/BO3/c2-1(3)4/q-3	Bioparadigms
SLC4A11	NaHCO3-	InChI=1S/CH2O3.Na/c2-1(3)4;/h(H2,2,3,4);/q;+1/p-1	GtoPdb
SLC4A4	NaHCO3-	InChI=1S/CH2O3.Na/c2-1(3)4;/h(H2,2,3,4);/q;+1/p-1	GtoPdb
SLC4A5	NaHCO3-	InChI=1S/CH2O3.Na/c2-1(3)4;/h(H2,2,3,4);/q;+1/p-1	GtoPdb
SLC4A7	NaHCO3-	InChI=1S/CH2O3.Na/c2-1(3)4;/h(H2,2,3,4);/q;+1/p-1	TCDB

SLC4A8	NaHCO3-	InChI=1S/CH2O3.Na/c2-1(3)4;/h(H2,2,3,4);/q;+1/p-1	GtoPdb
SLC51A	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLC51A	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC51A	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC51B	Sulfate	InChI=1S/H2O4S/c1-5(2,3)4/h(H2,1,2,3,4)/p-2	TCDB
SLC5A12	Pyrazionate	InChI=1S/C5H4N2O2/c8-5(9)4-3-6-1-2-7-4/h1-3H,(H,8,9)	CeMM
SLC5A5	Tetrafluoroborate	InChI=1S/BF4/c2-1(3,4)5/q-1	TCDB
SLC5A5	Bromate	InChI=1S/BrHO3/c2-1(3)4/h(H,2,3,4)/p-1	TCDB
SLC5A5	SeCN-	InChI=1S/CHNSe/c2-1-3/h3H/p-1	CeMM
SLC5A5	ClO3-	InChI=1S/ClHO3/c2-1(3)4/h(H,2,3,4)/p-1	CeMM
SLC5A5	Chlorate	InChI=1S/ClHO3/c2-1(3)4/h(H,2,3,4)/p-1	TCDB
SLC5A5	ClO4-	InChI=1S/ClHO4/c2-1(3,4)5/h(H,2,3,4,5)/p-1	Bioparadigms
SLC5A5	ClO4-	InChI=1S/ClHO4/c2-1(3,4)5/h(H,2,3,4,5)/p-1	GtoPdb
SLC5A5	pertechnetate	InChI=1S/H2O.3O.Tc/h1H2;;;;/q;;;;+1/p-1	GtoPdb
SLC5A5	I-	InChI=1S/HI/h1H/p-1	TCDB
SLC5A5	Periodate	InChI=1S/HIO4/c2-1(3,4)5/h(H,2,3,4,5)/p-1	TCDB
SLC5A5	Na+	InChI=1S/Na/q+1	TCDB
SLC5A5	NO3-	InChI=1S/NO3/c2-1(3)4/q-1	Bioparadigms
SLC5A5	NO3-	InChI=1S/NO3/c2-1(3)4/q-1	CeMM
SLC5A5	NO3-	InChI=1S/NO3/c2-1(3)4/q-1	GtoPdb
SLC5A5	nitrate	InChI=1S/NO3/c2-1(3)4/q-1	TCDB
SLC5A6	lipoate	InChI=1S/C8H14O2S2/c9-8(10)4-2-1-3-7-5-6-11-12-7/h7H,1-6H2,(H,9,10)	Bioparadigms
SLC5A6	pantothenate	InChI=1S/C9H17NO5/c1-9(2,5-11)7(14)8(15)10-4-3-6(12)13/h7,11,14H,3-5H2,1-2H3,(H,10,15)(H,12,13)	Bioparadigms
SLC5A6	Panthothenate	InChI=1S/C9H17NO5/c1-9(2,5-11)7(14)8(15)10-4-3-6(12)13/h7,11,14H,3-5H2,1-2H3,(H,10,15)(H,12,13)	CeMM
SLC5A8	Acetate	InChI=1S/C2H4O2/c1-2(3)4/h1H3,(H,3,4)	CeMM
SLC5A8	acetic acid	InChI=1S/C2H4O2/c1-2(3)4/h1H3,(H,3,4)	GtoPdb
SLC5A8	Acetate	InChI=1S/C2H4O2/c1-2(3)4/h1H3,(H,3,4)	TCDB
SLC5A8	Proprionate	InChI=1S/C3H6O2/c1-2-3(4)5/h2H2,1H3,(H,4,5)	CeMM
SLC6A1	gamma-aminobutyric acid (GABA)	InChI=1S/C4H9NO2/c5-3-1-2-4(6)7/h1-3,5H2,(H,6,7)	TCDB
SLC7A11	cystine anionic form	InChI=1S/C6H12N2O4S2/c7-3(5(9)10)1-13-14-2-4(8)6(11)12/h3-4H,1-2,7-8H2,(H,9,10)(H,11,12)/p-2/t3-,4-/m0/s1	Bioparadigms
SLCO1A2	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLCO1B1	Arsenite	InChI=1S/AsH3O3/c2-1(3)4/h2-4H	Metrabase
SLCO1B1	arsenate	InChI=1S/AsH3O4/c2-1(3,4)5/h(H3,2,3,4,5)	Metrabase
SLCO1B1	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLCO1B1	TROGLITAZONE SULFATE	InChI=1S/C24H27NO8S2.H3N/c1-13-14(2)21-18(15(3)20(13)33-35(28,29)30)9-10-24(4,32-21)12-31-17-7-5-16(6-8-17)11-19-22(26)25-23(27)34-19;/h5-8,19H,9-12H2,1-4H3,(H,25,26,27)(H,28,29,30);1H3	ChEMBL
SLCO1B1	Estradiol-17 beta-D-glucuronide	InChI=1S/C24H32O8/c1-24-9-8-14-13-5-3-12(25)10-11(13)2-4-15(14)16(24)6-7-17(24)31-23-20(28)18(26)19(27)21(32-23)22(29)30/h3,5,10,14-21,23,25-28H,2,4,6-9H2,1H3,(H,29,30)/t14-,15-,16+,17+,18+,19+,20-,21+,23-,24+/m1/s1	TCDB
SLCO1B1	Monoglucuronosyl bilirubin	InChI=1S/C39H44N4O12/c1-7-20-19(6)36(50)43-27(20)14-25-18(5)23(10-12-31(46)54-39-34(49)32(47)33(48)35(55-39)38(52)53)29(41-25)15-28-22(9-11-30(44)45)17(4)24(40-28)13-26-16(3)21(8-	TCDB

		2)37(51)42-26/h7-8,13-14,32-35,39-41,47-49H,1-2,9-12,15H2,3-6H3,(H,42,51)(H,43,50)(H,44,45)(H,52,53)/b26-13+,27-14+/t32-,33-,34+,35-,39+/m0/s1	
SLCO1B1	Rifampicin	InChI=1S/C43H58N4O12/c1-21-12-11-13-22(2)42(55)45-33-28(20-44-47-17-15-46(9)16-18-47)37(52)30-31(38(33)53)36(51)26(6)40-32(30)41(54)43(8,59-40)57-19-14-29(56-10)23(3)39(58-27(7)48)25(5)35(50)24(4)34(21)49/h11-14,19-21,23-25,29,34-35,39,49-53H,15-18H2,1-10H3,(H,45,55)/b12-11+,19-14+,22-13-,44-20+/t21-,23+,24+,25+,29-,34-,35+,39+,43-/m0/s1	UCSF-FDA
SLCO1B1	Bisglucuronosyl bilirubin	InChI=1S/C45H52N4O18/c1-7-20-19(6)40(58)49-27(20)14-25-18(5)23(10-12-31(51)65-45-37(57)33(53)35(55)39(67-45)43(62)63)29(47-25)15-28-22(17(4)24(46-28)13-26-16(3)21(8-2)41(59)48-26)9-11-30(50)64-44-36(56)32(52)34(54)38(66-44)42(60)61/h7-8,13-14,32-39,44-47,52-57H,1-2,9-12,15H2,3-6H3,(H,48,59)(H,49,58)(H,60,61)(H,62,63)/b26-13+,27-14+/t32-,33-,34-,35-,36+,37+,38-,39-,44+,45+/m0/s1	TCDB
SLCO1B3	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLCO1B3	17-beta-glucuronosyl estradiol	InChI=1S/C24H32O8/c1-24-9-8-14-13-5-3-12(25)10-11(13)2-4-15(14)16(24)6-7-17(24)31-23-20(28)18(26)19(27)21(32-23)22(29)30/h3,5,10,14-21,23,25-28H,2,4,6-9H2,1H3,(H,29,30)/t14-,15-,16+,17+,18+,19+,20-,21+,23-,24+/m1/s1	CeMM
SLCO1B3	Rifampicin	InChI=1S/C43H58N4O12/c1-21-12-11-13-22(2)42(55)45-33-28(20-44-47-17-15-46(9)16-18-47)37(52)30-31(38(33)53)36(51)26(6)40-32(30)41(54)43(8,59-40)57-19-14-29(56-10)23(3)39(58-27(7)48)25(5)35(50)24(4)34(21)49/h11-14,19-21,23-25,29,34-35,39,49-53H,15-18H2,1-10H3,(H,45,55)/b12-11+,19-14+,22-13-,44-20+/t21-,23+,24+,25+,29-,34-,35+,39+,43-/m0/s1	UCSF-FDA
SLCO1B7	Estradiol-17-beta-glucuronide	InChI=1S/C24H32O8/c1-24-9-8-14-13-5-3-12(25)10-11(13)2-4-15(14)16(24)6-7-17(24)31-23-20(28)18(26)19(27)21(32-23)22(29)30/h3,5,10,14-21,23,25-28H,2,4,6-9H2,1H3,(H,29,30)/t14-,15-,16+,17+,18+,19+,20-,21+,23-,24+/m1/s1	CeMM
SLCO1C1	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLCO1C1	Estradiol-17-beta-glucuronide	InChI=1S/C24H32O8/c1-24-9-8-14-13-5-3-12(25)10-11(13)2-4-15(14)16(24)6-7-17(24)31-23-20(28)18(26)19(27)21(32-23)22(29)30/h3,5,10,14-21,23,25-28H,2,4,6-9H2,1H3,(H,29,30)/t14-,15-,16+,17+,18+,19+,20-,21+,23-,24+/m1/s1	CeMM
SLCO2B1	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLCO2B1	Thyroxine 4-O-beta-D-glucuronide	InChI=1S/C21H19I4NO10/c22-8-1-6(3-12(26)19(30)31)2-9(23)16(8)34-7-4-10(24)17(11(25)5-7)35-21-15(29)13(27)14(28)18(36-21)20(32)33/h1-2,4-5,12-15,18,21,27-29H,3,26H2,(H,30,31)(H,32,33)/t12-,13-,14-,15+,18-,21?/m0/s1	CeMM
SLCO3A1	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLCO4A1	ESTRONE SULFURIC ACID	InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,1H3,(H,20,21,22)/t14-,15-,16+,18+/m1/s1	ChEMBL
SLCO4C1	Asymmetrical dimethylarginine	InChI=1S/C8H18N4O2/c1-12(2)8(10)11-5-3-4-6(9)7(13)14/h6H,3-5,9H2,1-2H3,(H2,10,11)(H,13,14)/t6-/m0/s1	CeMM
SPNS2	C17-S1P	InChI=1S/C17H36NO5P/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-17(19)16(18)15-23-24(20,21)22/h13-14,16-17,19H,2-12,15,18H2,1H3,(H2,20,21,22)/b14-13+/t16-,17+/m0/s1	Bioparadigms
SPNS2	C17-S1P	InChI=1S/C17H36NO5P/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-17(19)16(18)15-23-24(20,21)22/h13-14,16-17,19H,2-12,15,18H2,1H3,(H2,20,21,22)/b14-13+/t16-,17+/m0/s1	GtoPdb
SPNS2	Sphingosine 1-phosphate (S1P)	InChI=1S/C18H38NO5P/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-18(20)17(19)16-24-25(21,22)23/h14-15,17-18,20H,2-13,16,19H2,1H3,(H2,21,22,23)/b15-14+/t17-,18+/m0/s1	CeMM

SPNS2	Sphingosine-1-phosphate (S1P)	InChI=1S/C18H38NO5P/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-18(20)17(19)16-24-25(21,22)23/h14-15,17-18,20H,2-13,16,19H2,1H3,(H2,21,22,23)/b15-14+/t17-,18+/m0/s1	GtoPdb
SPNS2	dihydrosphingosine-1-phosphate (DH-S1P)	InChI=1S/C18H40NO5P/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-18(20)17(19)16-24-25(21,22)23/h17-18,20H,2-16,19H2,1H3,(H2,21,22,23)/t17-,18+/m0/s1/i14T,15T/t14?,15?,17-,18+	GtoPdb
SPNS2	phyto-S1P	InChI=1S/C18H40NO6P/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-17(20)18(21)16(19)15-25-26(22,23)24/h16-18,20-21H,2-15,19H2,1H3,(H2,22,23,24)/t16-,17+,18-/m0/s1	Bioparadigms
SPNS2	phyto-S1P	InChI=1S/C18H40NO6P/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-17(20)18(21)16(19)15-25-26(22,23)24/h16-18,20-21H,2-15,19H2,1H3,(H2,22,23,24)/t16-,17+,18-/m0/s1	GtoPdb
SPNS2	phosphorylated Fingolimod	InChI=1S/C19H34NO5P/c1-2-3-4-5-6-7-8-17-9-11-18(12-10-17)13-14-19(20,15-21)16-25-26(22,23)24/h9-12,21H,2-8,13-16,20H2,1H3,(H2,22,23,24)	Bioparadigms
SPNS2	Phosphorylated fingolimod (FTY720-P)	InChI=1S/C19H34NO5P/c1-2-3-4-5-6-7-8-17-9-11-18(12-10-17)13-14-19(20,15-21)16-25-26(22,23)24/h9-12,21H,2-8,13-16,20H2,1H3,(H2,22,23,24)	GtoPdb
SV2A	selectracetam	InChI=1S/C10H14F2N2O2/c1-2-7(10(13)16)14-5-6(3-8(11)12)4-9(14)15/h3,6-7H,2,4-5H2,1H3,(H2,13,16)/t6-,7+/m1/s1	Bioparadigms
TMEM165	Ca2+	InChI=1S/Ca/q+2	GtoPdb
TMEM165	H+	InChI=1S/p+1	GtoPdb
TMEM165	protons	InChI=1S/p+1	TCDB
TUSC3	Cu2+	InChI=1S/Cu/q+2	GtoPdb
TUSC3	Fe2+	InChI=1S/Fe/q+2	GtoPdb
TUSC3	Mg2+	InChI=1S/Mg/q+2	GtoPdb
TUSC3	Mn2+	InChI=1S/Mn/q+2	GtoPdb
XPR1	Phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	Bioparadigms
XPR1	PO43-	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	CeMM
XPR1	Phosphate	InChI=1S/H3O4P/c1-5(2,3)4/h(H3,1,2,3,4)/p-3	GtoPdb

LIST OF FIGURES

LIST OF TABLES

REFERENCES

1. Müller-Esterl W. Biochemie. Buch. 2. Aufl. Heidelberg: Spektrum Akad. Verl; 2011. 738 p.

2. Pizzagalli MD, Bensimon A, Superti-Furga G. A guide to plasma membrane solute carrier proteins. The FEBS Journal. 2021;288(9):2784–835.

3. Hediger MA, Clémençon B, Burrier RE, Bruford EA. The ABCs of membrane transporters in health and disease (SLC series): Introduction. Molecular Aspects of Medicine. 2013 Apr 1;34(2):95–107.

4. Höglund PJ, Nordström KJV, Schiöth HB, Fredriksson R. The Solute Carrier Families Have a Remarkably Long Evolutionary History with the Majority of the Human Families Present before Divergence of Bilaterian Species. Mol Biol Evol. 2011 Apr;28(4):1531–41.

5. Colas C, Ung PMU, Schlessinger A. SLC Transporters: Structure, Function, and Drug Discovery. Medchemcomm. 2016 Jun 1;7(6):1069–81.

6. Lin L, Yee SW, Kim RB, Giacomini KM. SLC Transporters as Therapeutic Targets: Emerging Opportunities. Nat Rev Drug Discov. 2015 Aug;14(8):543–60.

7. Meixner E, Goldmann U, Sedlyarov V, Scorzoni S, Rebsamen M, Girardi E, Superti-Furga G. A substrate-based ontology for human solute carriers. Molecular Systems Biology. 2020 Jul 1;16(7):e9652.

8. O'Hagan S, Wright Muelas M, Day PJ, Lundberg E, Kell DB. GeneGini: Assessment via the Gini Coefficient of Reference "Housekeeping" Genes and Diverse Human Transporter Expression Profiles. Cell Syst. 2018 Feb 28;6(2):230-244.e1.

9. Mendelian disease [Internet]. TheFreeDictionary.com. [cited 2022 Mar 4]. Available from: https://medical-dictionary.thefreedictionary.com/Mendelian+disease

10. Raby BA. Inheritance patterns of monogenic disorders (Mendelian and non-Mendelian). Available from: https://www.uptodate.com/contents/inheritance-patterns-of-monogenic-disorders-mendelian-and-non-mendelian#H1177464532

11. Santer R, Calado J. Familial Renal Glucosuria and SGLT2: From a Mendelian Trait to a Therapeutic Target. CJASN. 2010 Jan 1;5(1):133–41.

12. EMA. Forxiga [Internet]. European Medicines Agency. 2018 [cited 2022 Mar 7]. Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/forxiga

13. Schaller L, Lauschke VM. The genetic landscape of the human solute carrier (SLC) transporter superfamily. Hum Genet. 2019;138(11):1359–77.

14. Zhang Y, Zhang Y, Sun K, Meng Z, Chen L. The SLC transporter in nutrient and metabolic sensing, regulation, and drug development. J Mol Cell Biol. 2018 Sep 18;11(1):1–13.

15. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Mägi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JRB, Egan JM, Lajunen T, Grarup N, Sparsø T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proença C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O'Connell J, Luan J, Elliott A, McCarroll SA, Payne F, Roccasecca RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Böttcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen YDI, Chines P, Clarke R, Coin LJM, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day INM, de Geus EJC, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllensten U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen AL, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PRV, Jørgensen T, Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley R, Mangino M, Manning AK, Martínez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orrù M, Pakyz R, Palmer CNA, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AFH, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S, Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurethsson G, Sijbrands EJG, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvänen AC, Tanaka T, Thorand B, Tichet J, Tönjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemsen G, Witteman JCM, Yarnell JWG, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, DIAGRAM Consortium, GIANT Consortium, Global BPgen Consortium, Borecki IB, Loos RJF, Meneton P, Magnusson PKE, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis GV, Serrano-Ríos M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WHL, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF, Anders Hamsten on behalf of Procardis Consortium, MAGIC investigators, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J, Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BWJH, Boomsma DI, Deloukas P, Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruokonen A, Jarvelin MR, Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis

GR, Wareham NJ, Sladek R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC, Barroso I. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 2010 Feb;42(2):105–16.

16. Rusu V, Hoch E, Mercader JM, Tenen DE, Gymrek M, Hartigan CR, DeRan M, von Grotthuss M, Fontanillas P, Spooner A, Guzman G, Deik AA, Pierce KA, Dennis C, Clish CB, Carr SA, Wagner BK, Schenone M, Ng MCY, Chen BH, MEDIA Consortium, SIGMA T2D Consortium, Centeno-Cruz F, Zerrweck C, Orozco L, Altshuler DM, Schreiber SL, Florez JC, Jacobs SBR, Lander ES. Type 2 Diabetes Variants Disrupt Function of SLC16A11 through Two Distinct Mechanisms. Cell. 2017 Jun 29;170(1):199-212.e20.

17. Thoeringer CK, Ripke S, Unschuld PG, Lucae S, Ising M, Bettecken T, Uhr M, Keck ME, Mueller-Myhsok B, Holsboer F, Binder EB, Erhardt A. The GABA transporter 1 (SLC6A1): a novel candidate gene for anxiety disorders. J Neural Transm. 2009 Jun;116(6):649–57.

18. Park HJ, Kim JW, Lee SK, Kim SK, Park JK, Cho AR, Chung JH, Song JY. Association between the SLC6A12 gene and negative symptoms of schizophrenia in a Korean population. Psychiatry Res. 2011 Oct 30;189(3):478–9.

19. Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, Tung JY, Hinds DA, Perlis RH, Winslow AR. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. Nat Genet. 2016 Sep;48(9):1031–6.

20. Quadri M, Federico A, Zhao T, Breedveld GJ, Battisti C, Delnooz C, Severijnen LA, Di Toro Mammarella L, Mignarri A, Monti L, Sanna A, Lu P, Punzo F, Cossu G, Willemsen R, Rasi F, Oostra BA, van de Warrenburg BP, Bonifati V. Mutations in SLC30A10 cause parkinsonism and dystonia with hypermanganesemia, polycythemia, and chronic liver disease. Am J Hum Genet. 2012 Mar 9;90(3):467–77.

21. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thorton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Morón FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fiévet N, Huentelman MW, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuiness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossù P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Deniz Naranjo MC, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, European Alzheimer's Disease Initiative (EADI), Genetic and Environmental Risk in Alzheimer's Disease, Alzheimer's Disease Genetic Consortium, Cohorts for Heart and Aging Research in Genomic Epidemiology, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannefelt L, Hakonarson H, Pichler S,

Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RF, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JS, Boerwinkle E, Riemenschneider M, Boada M, Hiltuenen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nöthen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013 Dec;45(12):1452–8.

22. Dehghan A, Köttgen A, Yang Q, Hwang SJ, Kao WL, Rivadeneira F, Boerwinkle E, Levy D, Hofman A, Astor BC, Benjamin EJ, van Duijn CM, Witteman JC, Coresh J, Fox CS. Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. Lancet. 2008 Dec 6;372(9654):1953–61.

23. Kolz M, Johnson T, Sanna S, Teumer A, Vitart V, Perola M, Mangino M, Albrecht E, Wallace C, Farrall M, Johansson A, Nyholt DR, Aulchenko Y, Beckmann JS, Bergmann S, Bochud M, Brown M, Campbell H, EUROSPAN Consortium, Connell J, Dominiczak A, Homuth G, Lamina C, McCarthy MI, ENGAGE Consortium, Meitinger T, Mooser V, Munroe P, Nauck M, Peden J, Prokisch H, Salo P, Salomaa V, Samani NJ, Schlessinger D, Uda M, Völker U, Waeber G, Waterworth D, Wang-Sattler R, Wright AF, Adamski J, Whitfield JB, Gyllensten U, Wilson JF, Rudan I, Pramstaller P, Watkins H, PROCARDIS Consortium, Doering A, Wichmann HE, KORA Study, Spector TD, Peltonen L, Völzke H, Nagaraja R, Vollenweider P, Caulfield M, WTCCC, Illig T, Gieger C. Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. PLoS Genet. 2009 Jun;5(6):e1000504.

24. Yang Q, Köttgen A, Dehghan A, Smith AV, Glazer NL, Chen MH, Chasman DI, Aspelund T, Eiriksdottir G, Harris TB, Launer L, Nalls M, Hernandez D, Arking DE, Boerwinkle E, Grove ML, Li M, Linda Kao WH, Chonchol M, Haritunians T, Li G, Lumley T, Psaty BM, Shlipak M, Hwang SJ, Larson MG, O'Donnell CJ, Upadhyay A, van Duijn CM, Hofman A, Rivadeneira F, Stricker B, Uitterlinden AG, Paré G, Parker AN, Ridker PM, Siscovick DS, Gudnason V, Witteman JC, Fox CS, Coresh J. Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. Circ Cardiovasc Genet. 2010 Dec;3(6):523–30.

25. Hollis-Moffatt JE, Phipps-Green AJ, Chapman B, Jones GT, van Rij A, Gow PJ, Harrison AA, Highton J, Jones PB, Montgomery GW, Stamp LK, Dalbeth N, Merriman TR. The renal urate transporter SLC17A1 locus: confirmation of association with gout. Arthritis Res Ther. 2012;14(2):R92.

26. Nakayama A, Nakaoka H, Yamamoto K, Sakiyama M, Shaukat A, Toyoda Y, Okada Y, Kamatani Y, Nakamura T, Takada T, Inoue K, Yasujima T, Yuasa H, Shirahama Y, Nakashima H, Shimizu S, Higashino T, Kawamura Y, Ogata H,

Kawaguchi M, Ohkawa Y, Danjoh I, Tokumasu A, Ooyama K, Ito T, Kondo T, Wakai K, Stiburkova B, Pavelka K, Stamp LK, Dalbeth N, Eurogout Consortium, Sakurai Y, Suzuki H, Hosoyamada M, Fujimori S, Yokoo T, Hosoya T, Inoue I, Takahashi A, Kubo M, Ooyama H, Shimizu T, Ichida K, Shinomiya N, Merriman TR, Matsuo H, Eurogout Consortium. GWAS of clinically defined gout and subtypes identifies multiple susceptibility loci that include urate transporter genes. Ann Rheum Dis. 2017 May;76(5):869–77.

27. Flynn TJ, Phipps-Green A, Hollis-Moffatt JE, Merriman ME, Topless R, Montgomery G, Chapman B, Stamp LK, Dalbeth N, Merriman TR. Association analysis of the SLC22A11 (organic anion transporter 4) and SLC22A12 (urate transporter 1) urate transporter locus with gout in New Zealand case-control sample sets reveals multiple ancestral-specific effects. Arthritis Res Ther. 2013;15(6):R220.

28. Lu X, Wang L, Lin X, Huang J, Charles Gu C, He M, Shen H, He J, Zhu J, Li H, Hixson JE, Wu T, Dai J, Lu L, Shen C, Chen S, He L, Mo Z, Hao Y, Mo X, Yang X, Li J, Cao J, Chen J, Fan Z, Li Y, Zhao L, Li H, Lu F, Yao C, Yu L, Xu L, Mu J, Wu X, Deng Y, Hu D, Zhang W, Ji X, Guo D, Guo Z, Zhou Z, Yang Z, Wang R, Yang J, Zhou X, Yan W, Sun N, Gao P, Gu D. Genome-wide association study in Chinese identifies novel loci for blood pressure and hypertension. Hum Mol Genet. 2015 Feb 1;24(3):865–74.

29. Ehret GB, Ferreira T, Chasman DI, Jackson AU, Schmidt EM, Johnson T, Thorleifsson G, Luan J, Donnelly LA, Kanoni S, Petersen AK, Pihur V, Strawbridge RJ, Shungin D, Hughes MF, Meirelles O, Kaakinen M, Bouatia-Naji N, Kristiansson K, Shah S, Kleber ME, Guo X, Lyytikäinen LP, Fava C, Eriksson N, Nolte IM, Magnusson PK, Salfati EL, Rallidis LS, Theusch E, Smith AJP, Folkersen L, Witkowska K, Pers TH, Joehanes R, Kim SK, Lataniotis L, Jansen R, Johnson AD, Warren H, Kim YJ, Zhao W, Wu Y, Tayo BO, Bochud M, CHARGE-EchoGen consortium, CHARGE-HF consortium, Wellcome Trust Case Control Consortium, Absher D, Adair LS, Amin N, Arking DE, Axelsson T, Baldassarre D, Balkau B, Bandinelli S, Barnes MR, Barroso I, Bevan S, Bis JC, Bjornsdottir G, Boehnke M, Boerwinkle E, Bonnycastle LL, Boomsma DI, Bornstein SR, Brown MJ, Burnier M, Cabrera CP, Chambers JC, Chang IS, Cheng CY, Chines PS, Chung RH, Collins FS, Connell JM, Döring A, Dallongeville J, Danesh J, de Faire U, Delgado G, Dominiczak AF, Doney ASF, Drenos F, Edkins S, Eicher JD, Elosua R, Enroth S, Erdmann J, Eriksson P, Esko T, Evangelou E, Evans A, Fall T, Farrall M, Felix JF, Ferrières J, Ferrucci L, Fornage M, Forrester T, Franceschini N, Duran OHF, Franco-Cereceda A, Fraser RM, Ganesh SK, Gao H, Gertow K, Gianfagna F, Gigante B, Giulianini F, Goel A, Goodall AH, Goodarzi MO, Gorski M, Gräßler J, Groves C, Gudnason V, Gyllensten U, Hallmans G, Hartikainen AL, Hassinen M, Havulinna AS, Hayward C, Hercberg S, Herzig KH, Hicks AA, Hingorani AD, Hirschhorn JN, Hofman A, Holmen J, Holmen OL, Hottenga JJ, Howard P, Hsiung CA, Hunt SC, Ikram MA, Illig T, Iribarren C, Jensen RA, Kähönen M, Kang H, Kathiresan S, Keating BJ, Khaw KT, Kim YK, Kim E, Kivimaki M, Klopp N, Kolovou G, Komulainen P, Kooner JS, Kosova G, Krauss RM, Kuh D, Kutalik Z, Kuusisto J, Kvaløy K, Lakka TA, Lee NR, Lee IT, Lee WJ, Levy D, Li X, Liang KW, Lin H, Lin L, Lindström J, Lobbens S, Männistö S, Müller G, Müller-Nurasyid M, Mach F, Markus HS, Marouli E, McCarthy MI, McKenzie CA, Meneton P, Menni C, Metspalu A, Mijatovic V, Moilanen L, Montasser ME, Morris AD, Morrison AC, Mulas A,

Nagaraja R, Narisu N, Nikus K, O'Donnell CJ, O'Reilly PF, Ong KK, Paccaud F, Palmer CD, Parsa A, Pedersen NL, Penninx BW, Perola M, Peters A, Poulter N, Pramstaller PP, Psaty BM, Quertermous T, Rao DC, Rasheed A, Rayner NWNWR, Renström F, Rettig R, Rice KM, Roberts R, Rose LM, Rossouw J, Samani NJ, Sanna S, Saramies J, Schunkert H, Sebert S, Sheu WHH, Shin YA, Sim X, Smit JH, Smith AV, Sosa MX, Spector TD, Stančáková A, Stanton A, Stirrups KE, Stringham HM, Sundstrom J, Swift AJ, Syvänen AC, Tai ES, Tanaka T, Tarasov KV, Teumer A, Thorsteinsdottir U, Tobin MD, Tremoli E, Uitterlinden AG, Uusitupa M, Vaez A, Vaidya D, van Duijn CM, van Iperen EPA, Vasan RS, Verwoert GC, Virtamo J, Vitart V, Voight BF, Vollenweider P, Wagner A, Wain LV, Wareham NJ, Watkins H, Weder AB, Westra HJ, Wilks R, Wilsgaard T, Wilson JF, Wong TY, Yang TP, Yao J, Yengo L, Zhang W, Zhao JH, Zhu X, Bovet P, Cooper RS, Mohlke KL, Saleheen D, Lee JY, Elliott P, Gierman HJ, Willer CJ, Franke L, Hovingh GK, Taylor KD, Dedoussis G, Sever P, Wong A, Lind L, Assimes TL, Njølstad I, Schwarz PE, Langenberg C, Snieder H, Caulfield MJ, Melander O, Laakso M, Saltevo J, Rauramaa R, Tuomilehto J, Ingelsson E, Lehtimäki T, Hveem K, Palmas W, März W, Kumari M, Salomaa V, Chen YDI, Rotter JI, Froguel P, Jarvelin MR, Lakatta EG, Kuulasmaa K, Franks PW, Hamsten A, Wichmann HE, Palmer CNA, Stefansson K, Ridker PM, Loos RJF, Chakravarti A, Deloukas P, Morris AP, Newton-Cheh C, Munroe PB. The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. Nat Genet. 2016 Oct;48(10):1171–84.

30. Wang L, Li H, Yang B, Guo L, Han X, Li L, Li M, Huang J, Gu D. The Hypertension Risk Variant Rs820430 Functions as an Enhancer of SLC4A7. Am J Hypertens. 2017 Feb;30(2):202–8.

31. Ng FL, Boedtkjer E, Witkowska K, Ren M, Zhang R, Tucker A, Aalkjær C, Caulfield MJ, Ye S. Increased NBCn1 expression, Na+/HCO3- co-transport and intracellular pH in human vascular smooth muscle cells with a risk allele for hypertension. Hum Mol Genet. 2017 Mar 1;26(5):989–1002.

32. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T, Aulchenko Y, Lumley T, Köttgen A, Vasan RS, Rivadeneira F, Eiriksdottir G, Guo X, Arking DE, Mitchell GF, Mattace-Raso FUS, Smith AV, Taylor K, Scharpf RB, Hwang SJ, Sijbrands EJG, Bis J, Harris TB, Ganesh SK, O'Donnell CJ, Hofman A, Rotter JI, Coresh J, Benjamin EJ, Uitterlinden AG, Heiss G, Fox CS, Witteman JCM, Boerwinkle E, Wang TJ, Gudnason V, Larson MG, Chakravarti A, Psaty BM, van Duijn CM. Genome-wide association study of blood pressure and hypertension. Nat Genet. 2009 Jun;41(6):677–87.

33. A K, C P, Ca B, C F, M O, Nl G, A P, X G, Q Y, Av S, Jr O, M L, H S, T T, A I, S K, Sj H, Ad J, A D, A T, G P, Ej A, T Z, K L, Mc C, Nm PH, F K, A T, C H, T A, G E, Lj L, Tb H, E R, Bd M, De A, E B, M S, M C, A S, F G, J M, Ih de B, T H, T L, D S, Bm P, Mc Z, Ba O, M F, M P, M de A, St T, A S, A Z, Ps W, Rb S, S W, Tf M, Ts L, T I, N K, C M, He W, W K, L Z, T Z, I K, C M, Fb H, A J, W I, G Z, Sh W, Af W, H C, D E, S S, Ys A, Jf F, F R, Ag U, A H, M I, D N, A B, B K, L K, R M, M S, P K, M B, S C, K E, H V, Hk K, M N, U V, O P, V V, S B, An P, Pm R, Sl K, S B, Y L, Gc C, A F, T R, B P, I P, W W, V G, Ar S, J C, R S, L F, Mg S, Cm van D, I B, Bk K, I R, U G, Jf W, Jc W, Pp P, R R, N H, Di C, Wh K, Im H, Cs F. New loci associated with kidney function and chronic kidney

disease. Nature genetics [Internet]. 2010 May [cited 2022 Jun 8];42(5). Available from: https://pubmed.ncbi.nlm.nih.gov/20383146/

34. Liu CT, Garnaas MK, Tin A, Kottgen A, Franceschini N, Peralta CA, de Boer IH, Lu X, Atkinson E, Ding J, Nalls M, Shriner D, Coresh J, Kutlar A, Bibbins-Domingo K, Siscovick D, Akylbekova E, Wyatt S, Astor B, Mychaleckjy J, Li M, Reilly MP, Townsend RR, Adeyemo A, Zonderman AB, de Andrade M, Turner ST, Mosley TH, Harris TB, Rotimi CN, Liu Y, Kardia SLR, Evans MK, Shlipak MG, Kramer H, Flessner MF, Dreisbach AW, Goessling W, Cupples LA, Kao WL, Fox CS. Genetic Association for Renal Traits among Participants of African Ancestry Reveals New Loci for Renal Function. PLoS Genet. 2011 Sep 8;7(9):e1002264.

35. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, Ntalla I, Surendran P, Liu C, Cook JP, Kraja AT, Drenos F, Loh M, Verweij N, Marten J, Karaman I, Lepe MPS, O'Reilly PF, Knight J, Snieder H, Kato N, He J, Tai ES, Said MA, Porteous D, Alver M, Poulter N, Farrall M, Gansevoort RT, Padmanabhan S, Mägi R, Stanton A, Connell J, Bakker SJL, Metspalu A, Shields DC, Thom S, Brown M, Sever P, Esko T, Hayward C, van der Harst P, Saleheen D, Chowdhury R, Chambers JC, Chasman DI, Chakravarti A, Newton-Cheh C, Lindgren CM, Levy D, Kooner JS, Keavney B, Tomaszewski M, Samani NJ, Howson JMM, Tobin MD, Munroe PB, Ehret GB, Wain LV, International Consortium of Blood Pressure (ICBP) 1000G Analyses, BIOS Consortium, Lifelines Cohort Study, Understanding Society Scientific group, CHD Exome+ Consortium, ExomeBP Consortium, T2D-GENES Consortium, GoT2DGenes Consortium, Cohorts for Heart and Ageing Research in Genome Epidemiology (CHARGE) BP Exome Consortium, International Genomics of Blood Pressure (iGEN-BP) Consortium, UK Biobank CardioMetabolic Consortium BP working group. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat Genet. 2017 Mar;49(3):403–15.

36. W J, Jn F, Bj O, H Z, Mg L, Db S, C NC, Mw S, D L, Rp L. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nature genetics [Internet]. 2008 May [cited 2022 Jun 8];40(5). Available from: https://pubmed.ncbi.nlm.nih.gov/18391953/

37. Wain LV, Vaez A, Jansen R, Joehanes R, van der Most PJ, Erzurumluoglu AM, O'Reilly PF, Cabrera CP, Warren HR, Rose LM, Verwoert GC, Hottenga JJ, Strawbridge RJ, Esko T, Arking DE, Hwang SJ, Guo X, Kutalik Z, Trompet S, Shrine N, Teumer A, Ried JS, Bis JC, Smith AV, Amin N, Nolte IM, Lyytikäinen LP, Mahajan A, Wareham NJ, Hofer E, Joshi PK, Kristiansson K, Traglia M, Havulinna AS, Goel A, Nalls MA, Sõber S, Vuckovic D, Luan J, Del Greco M F, Ayers KL, Marrugat J, Ruggiero D, Lopez LM, Niiranen T, Enroth S, Jackson AU, Nelson CP, Huffman JE, Zhang W, Marten J, Gandin I, Harris SE, Zemunik T, Lu Y, Evangelou E, Shah N, de Borst MH, Mangino M, Prins BP, Campbell A, Li-Gao R, Chauhan G, Oldmeadow C, Abecasis G, Abedi M, Barbieri CM, Barnes MR, Batini C, Beilby J, Blake T, Boehnke M, Bottinger EP, Braund PS, Brown M, Brumat M, Campbell H, Chambers JC, Cocca M, Collins F, Connell J, Cordell HJ, Damman JJ, Davies G, de Geus EJ, de Mutsert R, Deelen J, Demirkale Y, Doney ASF, Dörr M, Farrall M, Ferreira T, Frånberg M, Gao H, Giedraitis V, Gieger C, Giulianini F, Gow AJ, Hamsten A, Harris TB, Hofman A, Holliday EG, Hui J,

Jarvelin MR, Johansson Å, Johnson AD, Jousilahti P, Jula A, Kähönen M, Kathiresan S, Khaw KT, Kolcic I, Koskinen S, Langenberg C, Larson M, Launer LJ, Lehne B, Liewald DCM, Lin L, Lind L, Mach F, Mamasoula C, Menni C, Mifsud B, Milaneschi Y, Morgan A, Morris AD, Morrison AC, Munson PJ, Nandakumar P, Nguyen QT, Nutile T, Oldehinkel AJ, Oostra BA, Org E, Padmanabhan S, Palotie A, Paré G, Pattie A, Penninx BWJH, Poulter N, Pramstaller PP, Raitakari OT, Ren M, Rice K, Ridker PM, Riese H, Ripatti S, Robino A, Rotter JI, Rudan I, Saba Y, Saint Pierre A, Sala CF, Sarin AP, Schmidt R, Scott R, Seelen MA, Shields DC, Siscovick D, Sorice R, Stanton A, Stott DJ, Sundström J, Swertz M, Taylor KD, Thom S, Tzoulaki I, Tzourio C, Uitterlinden AG, Völker U, Vollenweider P, Wild S, Willemsen G, Wright AF, Yao J, Thériault S, Conen D, Attia J, Sever P, Debette S, Mook-Kanamori DO, Zeggini E, Spector TD, van der Harst P, Palmer CNA, Vergnaud AC, Loos RJF, Polasek O, Starr JM, Girotto G, Hayward C, Kooner JS, Lindgren CM, Vitart V, Samani NJ, Tuomilehto J, Gyllensten U, Knekt P, Deary IJ, Ciullo M, Elosua R, Keavney BD, Hicks AA, Scott RA, Gasparini P, Laan M, Liu Y, Watkins H, Hartman CA, Salomaa V, Toniolo D, Perola M, Wilson JF, Schmidt H, Zhao JH, Lehtimäki T, van Duijn CM, Gudnason V, Psaty BM, Peters A, Rettig R, James A, Jukema JW, Strachan DP, Palmas W, Metspalu A, Ingelsson E, Boomsma DI, Franco OH, Bochud M, Newton-Cheh C, Munroe PB, Elliott P, Chasman DI, Chakravarti A, Knight J, Morris AP, Levy D, Tobin MD, Snieder H, Caulfield MJ, Ehret GB. Novel Blood Pressure Locus and Gene Discovery Using Genome-Wide Association Study and Expression Data Sets From Blood and the Kidney. Hypertension. 2017 Jul 24;HYPERTENSIONAHA.117.09438.

38. Fowdar JY, Grealy R, Lu Y, Griffiths LR. A genome-wide association study of essential hypertension in an Australian population using a DNA pooling approach. Mol Genet Genomics. 2017 Apr;292(2):307–24.

39. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, Himes BE, Levin AM, Mathias RA, Hancock DB, Baurley JW, Eng C, Stern DA, Celedón JC, Rafaels N, Capurso D, Conti DV, Roth LA, Soto-Quiros M, Togias A, Li X, Myers RA, Romieu I, Van Den Berg DJ, Hu D, Hansel NN, Hernandez RD, Israel E, Salam MT, Galanter J, Avila PC, Avila L, Rodriquez-Santana JR, Chapela R, Rodriguez-Cintron W, Diette GB, Adkinson NF, Abel RA, Ross KD, Shi M, Faruque MU, Dunston GM, Watson HR, Mantese VJ, Ezurum SC, Liang L, Ruczinski I, Ford JG, Huntsman S, Chung KF, Vora H, Li X, Calhoun WJ, Castro M, Sienra-Monge JJ, del Rio-Navarro B, Deichmann KA, Heinzmann A, Wenzel SE, Busse WW, Gern JE, Lemanske RF, Beaty TH, Bleecker ER, Raby BA, Meyers DA, London SJ, Mexico City Childhood Asthma Study (MCAAS), Gilliland FD, Children's Health Study (CHS) and HARBORS study, Burchard EG, Genetics of Asthma in Latino Americans (GALA) Study, Study of Genes-Environment and Admixture in Latino Americans (GALA2) and Study of African Americans, Asthma, Genes & Environments (SAGE), Martinez FD, Childhood Asthma Research and Education (CARE) Network, Weiss ST, Childhood Asthma Management Program (CAMP), Williams LK, Study of Asthma Phenotypes and Pharmacogenomic Interactions by Race-Ethnicity (SAPPHIRE), Barnes KC, Genetic Research on Asthma in African Diaspora (GRAAD) Study, Ober C, Nicolae DL. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. Nat Genet. 2011 Jul 31;43(9):887–92.

40. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M, Cookson WOCM, GABRIEL Consortium. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med. 2010 Sep 23;363(13):1211–21.

41. Himes BE, Jiang X, Hu R, Wu AC, Lasky-Su JA, Klanderman BJ, Ziniti J, Senter-Sylvia J, Lima JJ, Irvin CG, Peters SP, Meyers DA, Bleecker ER, Kubo M, Tamari M, Nakamura Y, Szefler SJ, Lemanske RF, Zeiger RS, Strunk RC, Martinez FD, Hanrahan JP, Koppelman GH, Postma DS, Nieuwenhuis MAE, Vonk JM, Panettieri RA, Markezich A, Israel E, Carey VJ, Tantisira KG, Litonjua AA, Lu Q, Weiss ST. Genome-Wide Association Analysis in Asthma Subjects Identifies SPATS2L as a Novel Bronchodilator Response Gene. PLoS Genet. 2012 Jul 5;8(7):e1002824.

42. Noguchi E, Sakamoto H, Hirota T, Ochiai K, Imoto Y, Sakashita M, Kurosaka F, Akasawa A, Yoshihara S, Kanno N, Yamada Y, Shimojo N, Kohno Y, Suzuki Y, Kang MJ, Kwon JW, Hong SJ, Inoue K, Goto Y ichi, Yamashita F, Asada T, Hirose H, Saito I, Fujieda S, Hizawa N, Sakamoto T, Masuko H, Nakamura Y, Nomura I, Tamari M, Arinami T, Yoshida T, Saito H, Matsumoto K. Genome-Wide Association Study Identifies HLA-DP as a Susceptibility Gene for Pediatric Asthma in Asian Populations. PLoS Genet. 2011 Jul 21;7(7):e1002170.

43. Ka D, Dg T, Cr G, Jm G, La R, S H, C E, Ss O, Sw Y, L L, Cd B, A ME, K S, A D, Ln B, Hj F, R K, Pc A, E BB, R C, Jg F, Ma L, F L, K M, D S, S T, W RC, S S, Jr RS, Rd H, Km G, Eg B. A genome-wide association study of bronchodilator response in Latinos implicates rare variants. The Journal of allergy and clinical immunology [Internet]. 2014 Feb [cited 2022 Jun 8];133(2). Available from: https://pubmed.ncbi.nlm.nih.gov/23992748/

44. César-Razquin A, Snijder B, Frappier-Brinton T, Isserlin R, Gyimesi G, Bai X, Reithmeier RA, Hepworth D, Hediger MA, Edwards AM, Superti-Furga G. A Call for Systematic Research on Solute Carriers. Cell. 2015 Jul 30;162(3):478–87.

45. Sugano K, Kansy M, Artursson P, Avdeef A, Bendels S, Di L, Ecker GF, Faller B, Fischer H, Gerebtzoff G, Lennernaes H, Senner F. Coexistence of passive and carrier-mediated processes in drug transport. Nat Rev Drug Discov. 2010 Aug;9(8):597–614.

46. Shitara Y. Clinical Importance of OATP1B1 and OATP1B3 in Drug–Drug Interactions. Drug Metabolism and Pharmacokinetics. 2011;26(3):220–7.

47. Giacomini KM, Huang SM, Tweedie DJ, Benet LZ, Brouwer KLR, Chu X, Dahlin A, Evers R, Fischer V, Hillgren KM, Hoffmaster KA, Ishikawa T, Keppler D, Kim RB, Lee CA, Niemi M, Polli JW, Sugiyama Y, Swaan PW, Ware JA, Wright SH, Yee SW, Zamek-Gliszczynski MJ, Zhang L. Membrane transporters in drug development. Nat Rev Drug Discov. 2010 Mar;9(3):215–36.

48. Dujic T, Zhou K, Donnelly LA, Tavendale R, Palmer CNA, Pearson ER. Association of Organic Cation Transporter 1 With Intolerance to Metformin in Type 2 Diabetes: A GoDARTS Study. Diabetes. 2014 Dec 15;64(5):1786–93.

49. Center for Drug Evaluation and Research (CDER). Clinical Drug Interaction Studies — Study Design, Data Analysis, and Clinical Implications Guidance for Industry. 2017;

50. RESOLUTE_booklet.pdf [Internet]. [cited 2022 Feb 8]. Available from: https://re-solute.eu/RESOLUTE_booklet.pdf

51. Dvorak V, Wiedmer T, Ingles-Prieto A, Altermatt P, Batoulis H, Bärenz F, Bender E, Digles D, Dürrenberger F, Heitman LH, IJzerman AP, Kell DB, Kickinger S, Körzö D, Leippe P, Licher T, Manolova V, Rizzetto R, Sassone F, Scarabottolo L, Schlessinger A, Schneider V, Sijben HJ, Steck AL, Sundström H, Tremolada S, Wilhelm M, Wright Muelas M, Zindel D, Steppan CM, Superti-Furga G. An Overview of Cell-Based Assay Platforms for the Solute Carrier Family of Transporters. Frontiers in Pharmacology [Internet]. 2021 [cited 2022 Feb 8];12. Available from: https://www.frontiersin.org/article/10.3389/fphar.2021.722889

52. Superti-Furga G, Lackner D, Wiedmer T, Ingles-Prieto A, Barbosa B, Girardi E, Goldmann U, Gürtl B, Klavins K, Klimek C, Lindinger S, Liñeiro-Retes E, Müller AC, Onstein S, Redinger G, Reil D, Sedlyarov V, Wolf G, Crawford M, Everley R, Hepworth D, Liu S, Noell S, Piotrowski M, Stanton R, Zhang H, Corallino S, Faedo A, Insidioso M, Maresca G, Redaelli L, Sassone F, Scarabottolo L, Stucchi M, Tarroni P, Tremolada S, Batoulis H, Becker A, Bender E, Chang YN, Ehrmann A, Müller-Fahrnow A, Pütter V, Zindel D, Hamilton B, Lenter M, Santacruz D, Viollet C, Whitehurst C, Johnsson K, Leippe P, Baumgarten B, Chang L, Ibig Y, Pfeifer M, Reinhardt J, Schönbett J, Selzer P, Seuwen K, Bettembourg C, Biton B, Czech J, de Foucauld H, Didier M, Licher T, Mikol V, Pommereau A, Puech F, Yaligara V, Edwards A, Bongers BJ, Heitman LH, IJzerman AP, Sijben HJ, van Westen GJP, Grixti J, Kell DB, Mughal F, Swainston N, Wright-Muelas M, Bohstedt T, Burgess-Brown N, Carpenter L, Dürr K, Hansen J, Scacioc A, Banci G, Colas C, Digles D, Ecker G, Füzi B, Gamsjäger V, Grandits M, Martini R, Troger F, Altermatt P, Doucerain C, Dürrenberger F, Manolova V, Steck AL, Sundström H, Wilhelm M, Steppan CM. The RESOLUTE consortium: unlocking SLC transporters for drug discovery. Nat Rev Drug Discov. 2020 Jul;19(7):429–30.

53. Nigam SK. What do drug transporters really do? Nat Rev Drug Discov. 2015 Jan;14(1):29–44.

54. Definition of metabolite - NCI Dictionary of Cancer Terms - National Cancer Institute [Internet]. 2011 [cited 2022 Apr 5]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/metabolite

55. Schumann T, König J, Henke C, Willmes DM, Bornstein SR, Jordan J, Fromm MF, Birkenfeld AL. Solute Carrier Transporters as Potential Targets for the Treatment of Metabolic Disease. Michel MC, editor. Pharmacol Rev. 2020 Jan 1;72(1):343–79.

56. Cheng L, Yang H, Zhao H, Pei X, Shi H, Sun J, Zhang Y, Wang Z, Zhou M. MetSigDis: a manually curated resource for the metabolic signatures of diseases. Briefings in Bioinformatics. 2019 Jan 18;20(1):203–9.

57. MetSigDis - HOME [Internet]. [cited 2022 Apr 13]. Available from: http://www.bio-annotation.cn/MetSigDis/

58. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, Furlong LI. The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Research. 2020 Jan 8;48(D1):D845–55.

59. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research. 2021 Jan 8;49(D1):D480–9.

60. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, Wiegers J, Wiegers TC, Mattingly CJ. Comparative Toxicogenomics Database (CTD): update 2021. Nucleic Acids Research. 2021 Jan 8;49(D1):D1138–43.

61. Orphanet: an online database of rare diseases and orphan drugs. Copyright, INSERM 1997. Available at http://www.orpha.net Accessed (20.06.2022).

62. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, Ledbetter DH, Maglott DR, Martin CL, Nussbaum RL, Plon SE, Ramos EM, Sherry ST, Watson MS, ClinGen. ClinGen--the Clinical Genome Resource. N Engl J Med. 2015 Jun 4;372(23):2235–42.

63. Martin AR, Williams E, Foulger RE, Leigh S, Daugherty LC, Niblock O, Leong IUS, Smith KR, Gerasimenko O, Haraldsdottir E, Thomas E, Scott RH, Baple E, Tucci A, Brittain H, de Burca A, Ibañez K, Kasperaviciute D, Smedley D, Caulfield M, Rendon A, McDonagh EM. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. Nat Genet. 2019 Nov;51(11):1560–5.

64. Tamborero D, Rubio-Perez C, Deu-Pons J, Schroeder MP, Vivancos A, Rovira A, Tusquets I, Albanell J, Rodon J, Tabernero J, de Torres C, Dienstmann R, Gonzalez-Perez A, Lopez-Bigas N. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. Genome Medicine. 2018 Mar 28;10(1):25.

65. Gutiérrez-Sacristán A, Grosdidier S, Valverde O, Torrens M, Bravo À, Piñero J, Sanz F, Furlong LI. PsyGeNET: a knowledge platform on psychiatric disorders and their genes. Bioinformatics. 2015 Sep 15;31(18):3075–7.

66. Gabrail C. Diplomarbeit: „Analysis of SLC families according to substrate similarity". :116.

67. Mak L, Marcus D, Howlett A, Yarova G, Duchateau G, Klaffke W, Bender A, Glen RC. Metrabase: a cheminformatics and bioinformatics database for small molecule transporter data analysis and (Q)SAR modeling. Journal of Cheminformatics. 2015 Jun 23;7(1):31.

68. BioParadigms [Internet]. [cited 2022 Jun 13]. Available from: https://www.bioparadigms.org/

69. Saier MH, Tran CV, Barabote RD. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. Nucleic Acids Res. 2006 Jan 1;34(Database issue):D181-186.

70. Home - CeMM [Internet]. [cited 2022 Jun 13]. Available from: https://cemm.at/

71. Morrissey KM, Wen CC, Johns SJ, Zhang L, Huang SM, Giacomini KM. The UCSF-FDA TransPortal: A Public Drug Transporter Database. Clinical Pharmacology & Therapeutics. 2012;92(5):545–6.

72. Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, Magariños MP, Mosquera JF, Mutowo P, Nowotka M, Gordillo-Marañón M, Hunter F, Junco L, Mugumbate G, Rodriguez-Lopez M, Atkinson F, Bosc N, Radoux CJ, Segura-Cabrera A, Hersey A, Leach AR. ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Research. 2019 Jan 8;47(D1):D930–40.

73. Harding SD, Armstrong JF, Faccenda E, Southan C, Alexander SPH, Davenport AP, Pawson AJ, Spedding M, Davies JA, NC-IUPHAR. The IUPHAR/BPS guide to PHARMACOLOGY in 2022: curating pharmacology for COVID-19, malaria and antibacterials. Nucleic Acids Res. 2022 Jan 7;50(D1):D1282–94.

74. Hemmerich J. KNIME Structure Standardisation Workflow [Internet]. Department of Pharmaceutical Chemistry, University of Vienna; 2020. Available from: https://github.com/PharminfoVienna/Chemical-Structure-Standardisation

75. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, Sayeeda Z, Lo E, Assempour N, Berjanskii M, Singhal S, Arndt D, Liang Y, Badran H, Grant J, Serra-Cayuela A, Liu Y, Mandal R, Neveu V, Pon A, Knox C, Wilson M, Manach C, Scalbert A. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res. 2018 Jan 4;46(D1):D608–17.

76. Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, Ohl P, Sieb C, Thiel K, Wiswedel B. KNIME: The Konstanz Information Miner. In: Preisach C, Burkhardt H, Schmidt-Thieme L, Decker R, editors. Data Analysis, Machine Learning and Applications. Berlin, Heidelberg: Springer; 2008. p. 319–26. (Studies in Classification, Data Analysis, and Knowledge Organization).

77. Samuel ELG, Holmes SL, Young DW. Processing binding data using an open-source workflow. Journal of Cheminformatics. 2021 Dec 11;13(1):99.

78. Sydow D, Wichmann M, Rodríguez-Guerra J, Goldmann D, Landrum G, Volkamer A. TeachOpenCADD-KNIME: A Teaching Platform for Computer-Aided Drug Design Using KNIME Workflows. J Chem Inf Model. 2019 Oct 28;59(10):4083–6.

79. Fillbrunn A, Dietz C, Pfeuffer J, Rahn R, Landrum GA, Berthold MR. KNIME for reproducible cross-domain analysis of life science data. Journal of Biotechnology. 2017 Nov 10;261:149–56.

80. Hemmerich J, Gurinova J, Digles D. Accessing Public Compound Databases with KNIME. 2020.

81. Fitzgerald A. API Calls: What They Are & How to Make Them in 5 Easy Steps [Internet]. [cited 2022 Apr 29]. Available from: https://blog.hubspot.com/website/api-calls

82. DisGeNET - a database of gene-disease associations [Internet]. [cited 2022 Apr 29]. Available from: https://www.disgenet.org/api/#/GDA/gdaByDiseaseMapping

83. Morgan HL. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. J Chem Doc. 1965 May 1;5(2):107–13.

84. 6.1: Molecular Descriptors [Internet]. Chemistry LibreTexts. 2019 [cited 2022 May 26]. Available from: https://chem.libretexts.org/Courses/Intercollegiate_Courses/Cheminformatics_OLCC_(2019)/6%3A_Molecular_Similarity/6.1%3A_Molecular_Descriptors

85. Apodaca RL. Computing Extended Connectivity Fingerprints [Internet]. [cited 2022 Jun 20]. Available from: http://depth-first.com/articles/2019/01/11/extended-connectivity-fingerprints/

86. Maggiora G, Vogt M, Stumpfe D, Bajorath J. Molecular Similarity in Medicinal Chemistry. J Med Chem. 2014 Apr 24;57(8):3186–204.

87. Gupta N, Martin PM, Prasad PD, Ganapathy V. SLC5A8 (SMCT1)-mediated transport of butyrate forms the basis for the tumor suppressive function of the transporter. Life Sci. 2006 Apr 18;78(21):2419–25.

88. Shulman RG, Rothman DL. The Glycogen Shunt Maintains Glycolytic Homeostasis and the Warburg Effect in Cancer. Trends Cancer. 2017 Nov;3(11):761–7.

89. Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wägele B, Altmaier E, Deloukas P, Erdmann J, Grundberg E, Hammond CJ, de Angelis MH, Kastenmüller G, Köttgen A, Kronenberg F, Mangino M, Meisinger C, Meitinger T, Mewes HW, Milburn MV, Prehn C, Raffler J, Ried JS, Römisch-Margl W, Samani NJ, Small KS, -Erich Wichmann H, Zhai G, Illig T, Spector TD, Adamski J, Soranzo N, Gieger C. Human metabolic individuality in biomedical and pharmaceutical research. Nature. 2011 Sep;477(7362):54–60.

90. Yee SW, Stecula A, Chien HC, Zou L, Feofanova EV, van Borselen M, Cheung KWK, Yousri NA, Suhre K, Kinchen JM, Boerwinkle E, Irannejad R, Yu B, Giacomini KM. Unraveling the functional role of the orphan solute carrier, SLC22A24 in the transport of steroid conjugates through metabolomic and genome-wide association studies. Evans WE, editor. PLoS Genet. 2019 Sep 25;15(9):e1008208.