Systematic *in silico* discovery of novel solute carrier-like proteins from proteomes

Gergely Gyimesi and Matthias A. Hediger

Membrane Transport Discovery Lab, Department of Nephrology and Hypertension and Department of Biomedical Research, Inselspital, University of Bern, Freiburgstrasse 15, CH-3010 Bern, Switzerland.

*For correspondence: Gergely Gyimesi (gergely.gyimesi@dbmr.unibe.ch); Matthias A. Hediger (matthias.hediger@ibmm.unibe.ch).

Abstract

Solute carrier (SLC) proteins represent the largest superfamily of transmembrane transporters, the systematic analysis of which is hampered by their functional and structural heterogeneity, despite their biological importance. Based on available nomenclature systems, we suspected that many as yet unidentified SLC transporters exist in the human genome. Here, we present criteria for defining "SLC-likeness" and apply them to curate a set of "SLC-like" protein families from the Transporter Classification Database (TCDB) and Protein families (Pfam) databases. Computational sequence similarity searches then surprisingly yielded ~130 more proteins in human with SLC-like properties, compared to previous annotations. Several of these novel putative SLC transporter proteins actually have documented transport activity in the scientific literature. We complete our overview of the SLC-ome by presenting an algorithm to classify SLC-like proteins into protein families, investigating their known functions and evolutionary relationships to similar proteins from 6 other clinically relevant experimental organisms, and pinpointing structural orphans. We envision that our work will serve as a stepping stone for future studies of the biological function and the identification of the natural substrates of the many under-explored SLC transporters, as well as the development of new therapeutic applications, including strategies for personalized medicine and drug delivery.

Introduction

The solute carrier (SLC) protein superfamily accounts for about 50% of all transport-related proteins and 10% of all membrane proteins encoded by the human genome. With more than 400 annotated members, it is the largest superfamily of membrane transporter proteins [1]. Membrane transporters and channels are the main entry routes for nutrients, ions, xenobiotics and serve as major exit routes for waste products and metabolites. The roles of SLC transporters as cellular gatekeepers, determinants of nutrient homeostasis and facilitators of drug metabolism and drug targeting has recently been revisited [2]. Around ~50% of currently annotated SLCs are predicted to be associated with human disease phenotypes and many SLCs are considered to represent promising drug targets or drug delivery systems or to affect drug ADMET (absorption, distribution, metabolism, extrusion, toxicity). It has recently become evident that the SLC superfamily offers an enormous unexplored therapeutic treasure. But while the list of approved drugs that target transporter proteins is increasing, many promising SLCs still remain unexplored, uncharacterized and underrepresented in the literature.

The SLC nomenclature system has traditionally been used to classify mammalian secondary active and facilitative transporters, including exchangers and antiporters, into families based on sequence identity per enquiry by the Human Gene Nomenclature Committee (HGNC) starting in the early

1990s [3]. Originally, SLCs assignments have been made for all membrane transport proteins that are not channels, ATP-driven pumps, aquaporins, porins of the outer mitochondrial membrane or ATP-binding cassette (ABC) transporters, having multiple transmembrane spanning segments usually and exhibiting transmembrane transport of a solute, or showing homology to membrane proteins having such features. Due to this construction, the SLC superfamily is structurally and functionally highly heterogeneous and thus most likely evolutionarily polyphyletic in origin. Because of these properties and the lack of common sequence patterns in the different SLC transporters, it has been difficult to assess how many SLC transporters exist in the human genome or which proteins could be predicted as SLC transporters. Proteins were typically added to the SLC system on a case-by-case basis. However, in view of recent requests to add new members into the SLC nomenclature, we suspected that the SLC system might be incomplete.

Despite their heterogeneity, SLC transporters seem to share common properties that probably were evolutionarily selected based on their suitability as facilitative transporters, secondary active transporters or exchangers. A remarkable property is that they have a symmetric inverted repeat architecture, which can be observed based on the available structures [4] and can sometimes also be detected at the sequence level [5,6]. Another important aspect is that the currently well-studied SLC transporters seem to follow an alternating access mechanism, meaning that the substrate-binding site is exposed on either one or the other side of the membrane, but not on both sides simultaneously [7,8]. Consequences of these properties are that SLC transporters typically contain many transmembrane helices (TMHs), and functionally exhibit saturable transport activity. In addition, most but not all SLC carriers transport water-soluble small molecules. These properties could be used as criteria to identify additional SLC transporter proteins.

In fact, there have been earlier attempts to gather additional SLCs from the human genome [9–11], as well as to classify them using automatic methods [12,13]. In this regard, one study [9] used BLAST searches to find SLC transporters that have local sequence similarities and found that 15 of the known SLC families fall into 4 phylogenetic clusters, which were termed α, β, γ, and δ groups. In addition, they have found 19 sequences that have previously not been described as SLCs. A later study [12] used a more sensitive HMM-HMM comparison-based method to identify locally similar regions in known SLC proteins. Visualization of the similarity network revealed visible protein clusters that correlated with existing SLC families. In addition, they identified two unannotated protein sequences that showed similarity to existing SLC proteins. A common limitation of these studies, however, is that they only searched for proteins that were similar to proteins already annotated as human SLC transporters. Nevertheless, these efforts using sequence similarity-based approaches have highlighted that there are additional as yet unannotated SLC transporters in human protein databases.

In our current study, we aimed to identify putative SLC transporters that might differ from those currently annotated in human. To do this, we turned to sequence databases and annotation systems that are phylogenetically broader and not limited to human proteins, and we developed criteria to define “SLC-like” proteins. Our method is thus more general than previous approaches and tackles the task of identifying SLC transporters on a semantic level. For this reason, we turned to the Transporter Classification DataBase (TCDB) to enable the extraction of missing SLC-like proteins.

The Transporter Classification DataBase (TCDB) is an alternative classification system that was created in the 1990s in parallel to the SLC nomenclature series [14,15]. It collects transport-related membrane proteins, including membrane receptors, transporters, ion channels, and membrane-anchored enzymes from all kingdoms of life, with a particular focus on proteins from lower organisms. The proteins in the TCDB are organized hierarchically into subfamilies, families and superfamilies based on phylogenetic and functional considerations, and each member in the database is given a five-segmented TC# similar to the EC# that is used for enzyme classification. In addition, a brief description is provided for each family that introduces identified members and contains links to the most important relevant papers.

However, the TCDB dataset is not directly applicable for creating an overview of the human SLC-ome. One of the reasons is that the TCDB is set up as a "representative database", which means that it only contains certain representative sequences from each family. In addition, there is no particular focus on human proteins, and in fact several annotated human SLC transporters are not present in the database.

It also seems problematic to consider all proteins in the TCDB that are annotated as part of the secondary transporter superfamily TC #2.A as being SLC-like. Indeed, many of the TCDB families annotated as part of TC #2.A exhibit structural or sequence features that do not match the characteristics of existing SLCs. Examples of this are the Trk K^+ transporters (#2.A.38), which display an ion-channel like structural fold [16], the GUP glycerol uptake proteins (#2.A.50), which exhibit based on follow-up studies enzymatic activity [17], and the Twin Arginine Targeting (Tat) family (#2.A.64), which are actually protein secretion complexes [18,19]. Since none of these families correspond to structural or functional characteristics of currently known SLC transporters, it is likely that not all proteins annotated under the TC #2.A superfamily are "SLC-like". Thus, while the TCDB could be a rich source of information for finding new transporters, it is clear that the perception of what a "transporter" should be according to TCDB does not always correspond to the typical properties of well-characterized SLC proteins. Additional filtering of TCDB data is therefore necessary.

As part of the TransportDB project, there were parallel efforts to collect secondary transporters from human and several other organisms [20–22]. To this end, the authors have built an automatic transporter annotation pipeline (TransAAP), which relies on BLAST searches, the Clusters of Orthologous Groups (COG) database [23], "selected HMMs for transporter protein families" [22] from the TIGRfam and Pfam databases [24,25] and hydropathy predictions of TMHMM [26]. This pipeline was used as a semiautomatic tool to annotate transporter-like proteins from the NCBI RefSeq database. However, based on the currently available TransportDB website, the resulting protein hits are neither linked to protein annotation databases, such as UniProt [27], nor are their official gene symbols or SLC names displayed. Therefore, no correlation with the existing SLC nomenclature is provided, and it is not trivial to say whether or not an existing SLC protein is included in the database.

Therefore, there is a clear need in the field to define the criteria of "SLC-likeness" and to identify and classify all proteins in humans and other species that exhibit "SLC-likeness". Thus, in our work, we have added semantic content to the term "SLC-likeness" by defining the essential criteria for it, and we have carried out an exhaustive search for proteins meet these criteria, both with manual curation of datasets and with automatic sequence similarity-based approaches.

Results

Elaboration of criteria for "SLC-likeness"

Since the TCDB takes a very inclusive approach to collecting membrane transport-related proteins from a broad range of biological organisms, we have selected it as the source database for our endeavors. However, as outlined in the introduction, selecting SLC-like proteins from the TCDB is non-trivial. Therefore, we have introduced a set of criteria based on current knowledge of SLC transporters in order to select SLC-like protein families from the TCDB. We believe that these criteria represent the most important properties of currently known SLC transporters. However, due to missing information, these criteria could not always be verified while filtering the TCDB for SLC-like protein families. Therefore, the criteria themselves were neither formulated very specifically nor strictly enforced, but merely served as a guideline during the selection process, and exceptions were made in several cases. The criteria used to infer "SLC-likeness" were as follows.

- Structure of the protein should be α-helical, with at least three transmembrane helices (TMHs). In exceptional cases, 2-TMH families were also included, but β-barrel proteins, mostly

β-structure proteins, membrane-anchored proteins, cyclic peptides and proteins consisting only of soluble domains (based on predictions or structural data) were excluded.
- The size of the transported substrate should fall within the small-molecule range (i.e. oligopeptides might be accepted as substrates but protein secretion systems are excluded). Also excluded are DNA-, RNA- and polysaccharide-transporting systems.
- Proteins where transport activity was used as a synonym for trafficking (i.e. protein or vesicle translocation within the cell) but otherwise seemingly having no small-molecule transmembrane transport activity were excluded. On this basis, chaperones and other proteins helping the insertion of nascent proteins into a cellular membrane were also excluded.
- Proteins with a channel-like mechanism were excluded, except in some rare cases. In particular, holins, toxins and other pore-forming proteins, and proteins bearing similarity to them, have been excluded.
- Proteins with enzymatic activity, or similarity to known enzymes were excluded. In some cases, where the protein was believed to contain both a transport domain and a soluble enzyme domain, the proteins have been included.
- Receptors that trigger endocytosis upon substrate binding were excluded. Only receptors were included where the receptor protein itself mediates the translocation of the substrate through the membrane, or the insertion of the substrate in the membrane, if that is its final location.
- ATP- and GTP-dependent transporters (e.g. ABC, ECF) were excluded.
- For some (mostly putative) transporter families, TCDB does not give an explanation why the proteins would be considered as transporters. Families with no resemblance to known transporters and no indication or argument as to why they would be transporters were excluded.

We believe that our criteria are sufficiently broad to allow the identification of all putative SLC-like transporters, while also being specific enough to distinguish them from other well-known, non-transport-related transmembrane protein families. In general, our criteria represent the first attempt to set up semantic guidelines defining "SLC-likeness", and proteins and protein families that meet the above criteria are referred to as "SLC-like" in our current study.

Search for novel "SLC-like" proteins

The above-mentioned criteria were applied to manually select protein families in the TCDB that fulfill these criteria based on the description of each third-level family from the TCDB database. Throughout this manuscript, we use the term "TCDB family" to refer to third-level groupings in the classification hierarchy (i.e. TC# x.y.z), while "subfamily" and "superfamily" refer to fourth-level (TC# x.y.z.w) and second-level (TC #x.y) classes, respectively. In the work presented herein, we have analyzed superfamilies TC #1.A, #2.A, #9.A and #9.B and the families, and in certain cases the subfamilies within them (see Table 1).

It was a significant curation effort to manually assess the 1534 subfamilies within 616 families in the above-mentioned superfamilies of the TCDB, which were expected to contain SLC transporter-like proteins (Table 1). In many cases, information from the scientific literature, transmembrane segment prediction (see Methods) or structural databases were taken into account in order to determine whether a particular family fulfills the "SLC-likeness" criteria. Strikingly, of the 1534 subfamilies examined, only 602 in 167 families were found to be SLC-like according to our criteria. In particular, of the 556 subfamilies within superfamily TC# 2.A ("porters"), only 501 appeared to meet our selection criteria. This underlines once more that the term "porter" in relation to solute transport ambiguous in this field and a clearer definition of the perception of a solute transporter protein is needed.

To streamline our curation efforts, we have expanded our analysis to include protein domains stored within the Pfam database [28], which aims to maintain a curated set of protein families, often

represented by functional domains. Notably, Pfam provides curated Hidden Markov-Models (HMMs) for each Pfam family to facilitate sequence similarity searches for the occurrence of those domains. At the same time, Pfam neither attaches importance to transport-related domains nor to membrane-spanning domains, but groups protein families into higher-order groups called clans. Pfam clans contain evolutionarily related families whose relationships are supported either by sequence similarity, structural similarity or other orthogonal biological evidence [29]. As part of our study, we asked which Pfam families describe functionally relevant membrane-spanning regions of SLC-like transporters. To this end, we used HMMER [30] to search for all Pfam families in the sequences of all TCDB members within the four superfamilies analyzed, as shown in Table 1. This was followed by a manual curation step in which the resulting Pfam families were filtered according to the same criteria that were applied to the selection of SLC-like families from the TCDB (see above), based on the descriptions of the families on the Pfam-website as well as the scientific literature. Special care has been taken to consistently include or exclude TCDB (sub)families and their corresponding Pfam family, if applicable. Our goal was to identify a set of Pfam families that describe the functional membrane-spanning regions of SLC-like transporters.

These curation efforts yielded 211 Pfam families bearing "SLC-like" properties, which likely contain the membrane-spanning regions of SLC-like proteins selected from the 336 Pfam families present in total in the TCDB sequences analyzed. Notably, many of those Pfam families that have been excluded represented soluble structural or regulatory domains. Following this initial round of selection, we took advantage of clan groupings in Pfam and extended our selection efforts to analyze Pfam families belonging to the same clan as the selected 211 SLC-like Pfam families. This was based on the observation that many SLC-like transporter families in Pfam appear grouped into clans, so that members of these clans might represent SLC-like transporters themselves. Such a "clan expansion" procedure resulted in 12 additional Pfam families that are evolutionarily related to SLC-like protein families, of which 8 Pfam families met our criteria as SLC-like. Interestingly, these 8 SLC-like Pfam families currently have no representatives with a modest score (bit score > 25) in the TCDB, while 4 of the 8 families are annotated in Pfam as domains of unknown function (DUF). Thus, a total of 219 SLC-like Pfam families were identified in our search (see Supplementary Table 1).

As a next step, we wanted to know whether the selected families either from the TCDB or from Pfam have representatives in the proteomes of human and other clinically relevant organisms. For this analysis, we selected 7 organisms due to their clinical relevance or scientific utility (*H. sapiens*, *R. norvegicus*, *M. musculus*, *G. gallus*, *D. rerio*, *D. melanogaster*, *C. elegans*), for which we downloaded all sequences from the UniProt database [27], including Swiss-Prot (curated) and TrEMBL (predicted) [31] entries. Sequences of proteins in the TCDB have been aligned within each SLC-like family and subfamily and the alignments converted to HMMs for sensitive sequence similarity searches (see Methods). In addition, HMMs of SLC-like Pfam families were downloaded from the Pfam database. HMM-based similarity searches were then performed on the sequences downloaded for the 7 organisms to find proteins similar to any of the SLC-like TCDB families or subfamilies, or SLC-like Pfam domains, followed by the clustering of sequence fragments to arrive at one representative protein sequence per gene (see Methods).

The results of our search for SLC-like proteins are summarized in numbers in Table 2. Briefly, 60-68 of the 167 TCDB families have representatives in the 7 organisms (67 in human), and the organisms seem to contain 435-676 SLC-like proteins (552 in human). In total, 3750 proteins have been found in the 7 organisms studied. Notably, the number of SLC-like transporters found in human in our search is ~130 higher than previously reported [1], indicating that the human SLC-ome may be significantly larger than previously thought.

After arriving at this initial set of SLC-like proteins, we proceeded by attempting to classify the proteins into protein families, as detailed below, followed by manual database and literature searches focused on outlier proteins (single protein in a family) and human proteins in order to exclude false positives, as detailed below. This process was repeated iteratively to finally arrive at a

consistent set of proteins and corresponding families. The numbers mentioned below correspond to this final set of proteins, and all subsequent analysis was carried out on this set.

Classification into families

As mentioned in the introduction, SLC carriers are likely of polyphyletic origin, and individual families can be so diverse that even sensitive sequence similarity-based methods may have difficulties grouping related SLCs [12]. In our experience, multiple sequence alignment-based methods were not able to cluster the identified SLC-like sequences and to recapitulate them in known SLC families, so we devised a custom method for clustering distantly related sequences into proteins families based on the introduction of "HMM fingerprints". An HMM fingerprint is a mathematical vector of numbers assigned to a protein sequence, where the numbers represent the similarity scores of that protein sequence against each of the TCDB families, subfamilies and Pfam families that we have selected to be SLC-like. Thus, two protein sequences that show a similar pattern in their HMM fingerprints indicate their similarity. The usefulness of an HMM fingerprint depends on a meaningful definition of HMMs used in the fingerprint, whereby we capitalize on the evolutionary principle in the construction of TCDB families, subfamilies as well as Pfam families. Nevertheless, our goal was not to reconstruct the evolutionary history of a set of proteins, but to group proteins that share similar sequence features. However, due to the transitivity of homology and since similarity to a group of proteins suggests homology, clusters derived using HMM fingerprinting are likely to contain homologous proteins.

The HMM fingerprint-based classification of the SLC-like proteins found in our search yielded 103 protein families in total, 94 of which had representatives in human (Figure 1). For existing SLC transporters, the generated families corresponded well to classical SLC families. Interestingly, outlier proteins were found in several families that did not cluster with their families at the threshold we used. Examples include SLC5A7, SLC10A7, SLC25A46, SLC30A9, SLC39A9, MPDU1/SLC66A5 as well as the SLC9B family and SLC35 subfamilies. This shows that the HMM fingerprints, and thus likely the sequences of these outlier proteins diverge from those of other members of their families, and the sequences of subfamilies seem to diverge in certain cases. Several classical SLC families also clustered together, such as SLC32-SLC36-SLC38, SLC2-SLC22, and SLC17-SLC18-SLC37 proteins, likely due to their high sequence similarity. In these cases, we have adjusted the threshold value to arrive at clusters that better correspond to existing SLC classification (see Methods).

Our analysis revealed 43 new proteins families in total in human, containing proteins that have not yet been annotated as SLCs. Interestingly, our search has also found new proteins that clustered into existing SLC families (Table 3).

Structural similarity analysis, fold assignment

For polytopic transmembrane proteins, structural similarity can provide support for evolutionary relatedness and at the same time a basis for homology-based model building efforts. On the other hand, the lack of predicted similarity to proteins of known structure could pinpoint interesting targets for structural biology efforts by highlighting proteins that are likely to belong to new fold families.

We performed HMM-based searches on the pdb70 database (sequences of the proteins in the Protein Data Bank clustered to 70% sequence identity, see Methods) to assess whether structural homologues are available for the proteins found. In total, for 80 of the 103 families, at least one similar protein was found with a corresponding structure in the Protein Data Bank. Importantly, 476 human SLC-like proteins have a homologue whose structure has been solved. On the other hand, 43 human SLC-like proteins belonging to 19 different families do not have homologues with a known structure and thus are likely to constitute novel fold families. For the classical SLC proteins, their best-scoring similar proteins from the pdb70 dataset and the corresponding structural fold families are summarized in Supplementary Table 3. Based on this, it appears that classical SLCs from families SLC34, SLC44, SLC48 and SLC51 are still "structural orphans".

Phylogenetic analysis

Model organisms can be useful to study the biological function of various proteins, including solute transporters. In order to relate the results to human, however, knowledge of orthologous gene pairs is necessary. To this end, we performed phylogenetic analysis on each family of SLC-like proteins corresponding to our clustering. In brief, unrooted phylogenetic trees for each SLC-like protein family from all organisms have been generated and reconciled with the species tree of the 7 organisms in our study to identify gene duplication and speciation events in their evolutionary history (see Methods). The resulting evolutionary trees are deposited in Supplementary File 1. Based on these trees, we carried out orthology analysis focused on human proteins and the human lineage. The resulting data is presented in Figure 2, showing relationships between human genes and their orthologs in the other 6 organisms in our study. In addition, gene clusters are presented that have arisen through gene duplication events in the evolutionary history of the human lineage, but the corresponding human genes have likely been lost.

Literature search on newly found SLC-like proteins

As alluded to above, our initial search was followed by thorough investigation of the available literature to check whether a description of transport activity for the newly found proteins is available. Surprisingly, our search has revealed human proteins for which transported substrates are known. These data have been included in Table 3.

We also found several sequences that we believe could be false positives for various reasons. These can be fragments of other proteins, non-human proteins, translated sequences of pseudogenes or proteins that have shown partial similarity to existing transporters or proteins in Table 3, but have well-known non-transporter functions. These proteins have been omitted from Table 3, and are instead summarized in Supplementary Table 2. After such filtering, 3676 SLC-like proteins have remained, out of which 519 were from human, and 130 of them were from non-classical SLC families, indicating putative novel SLC-like proteins.

Discussion

Our curation and search efforts have revealed a surprising 130 human proteins that are SLC-like but were not officially part of the SLC nomenclature. Interestingly, around 30 of them were addressed in an earlier study and referred to as “atypical SLC transporters” [10,11]. All of these atypical transporters have also been identified in our study, together with many others, including several that are not part of the major facilitator superfamily (MFS) or the amino acid-polyamine-cation (APC) transporter superfamily.

In the following sections, we discuss the resulting set of SLC-like proteins with special focus on the identities of the transported substrates and, if known, the structural and mechanistic aspects of transport. In certain cases, the proteins have already been officially included in the SLC nomenclature, following approval by the Human Gene Nomenclature Committee (HGNC). Where existing information about substrate and function is not available, we have speculated on these aspects using available information in the literature and considerations on sequence similarity. For several proteins, their classification will require further consideration. We also would like to articulate open questions about what further work is needed in order to identify further transporters. We believe that our approach has been useful to pinpoint proteins that have a high possibility of being novel transporters. Thus, specific biological, biochemical or structural efforts could focus on these specific targets highlighted in our work, all of which would contribute to a complete assessment of the SLC-ome in human cells.

Transporters with existing evidence for transport function

Importantly, our search highlighted several proteins for which our literature search uncovered previously reported evidence of transporter activity. Several of the proteins in these families have been assigned SLC family numbers and, in collaboration with the HGNC, included in the SLC

nomenclature. Among these hits are several mitochondrial transporters (MPC/SLC54, LETM/SLC55, Sideroflexins/SLC56), which have been reviewed before [32]. In terms of transported substrate, many of the proteins with documented transport activity appear to be ion transporters or exchangers (Table 3). In the next paragraph, we will highlight certain proteins and families that have particularly caught our attention.

Interestingly, the **SLC60** family contains two MFS-like proteins (MFSD4A, MFSD4B), of which MFSD4B has been shown to transport D-glucose and urea [33,34].

The **SLC61A1** protein (MFSD5) is the only protein in human that shows similarity to the #2.A.1.40 family of molybdate transporters and contains the "MFS_5" Pfam domain. It has been claimed to be the homolog of similar transporters from algae and plants, and complementation assays suggested its ability to transport molybdate [35]. While molybdenum is a biologically active trace element, not much is known about its transport and homeostasis in human [36].

Interestingly, the **TMEM163** protein clustered together with SLC30 zinc transporter (ZnT) proteins, since the "Cation_efflux" Pfam domain, representative of the SLC30 family, was present in its sequence, although at a low score and non-significant e-value (2.9e-3). In the TCDB, TMEM163 is also classified under subfamily #2.A.4.8, sharing a common family with SLC30 transporters (#2.A.4.2). Multiple sequence alignment as well as pairwise alignments with existing SLC30 members reveal very low sequence identity with SLC30 proteins (4.2-14.4%), albeit these numbers are similar to those of SLC30A9 (6.5-13.4%). Given the marginal similarity to the "Cation_efflux" domain, it is tempting to assume that SLC30 proteins and TMEM163 are distantly related. Indeed, TMEM163 has been shown to bind [37] and transport Zn^{2+} [38–40], and substitution of its proposed substrate-binding residues with alanine abolished Zn^{2+} efflux activity [40]. Transport has been demonstrated to be H^{+}-coupled, and the protein functioning as a dimer [38], while extruding Zn^{2+} from the cell [40]. Intracellularly, TMEM163 was originally shown to be expressed in synaptic vesicles [41]. In overexpression systems, it is localized to both the plasma membrane and intracellular membrane compartments [40]. TMEM163 has been linked to Parkinson's disease (PD) [42], even though the opposite conclusion has also been drawn [43]. TMEM163 has also been reported to be upregulated by olanzapine, a psychotropic drug prescribed for PD patients [44]. In addition, TMEM163 was also shown to be highly expressed in insulin secretory vesicles in human pancreas [45], and has been identified as a risk factor in type 2 diabetes [46,47]. Disruption of TMEM163 expression might impair insulin secretion at high glucose stimuli [45].

The **TMEM165** protein clustered into its own family and is the only protein in human containing the "UPF0016" Pfam domain and showing similarity to TCDB family #2.A.106.2. TMEM165 is a member of a highly conserved family of transmembrane proteins that is present in many species of eukaryotes and bacteria [48]. Initially, TMEM165 and its yeast homolog, Gdt1p, have been hypothesized to be Ca^{2+}/H^{+} exchangers [49,50]. However, recently, evidence has been mounting about its involvement in manganese (Mn^{2+}) homeostasis [50], and both Ca^{2+} and Mn^{2+} transport activity has directly been shown [51]. TMEM165 is localized to the trans-Golgi in human cells [48], and is proposed to play a crucial role in regulating Mn^{2+} uptake into the Golgi apparatus [48,50]. In line with this, its homologs in other organisms, also containing the UPF0016 domain, are also annotated as Mn^{2+} transporters [50]. Manganese plays an important role as a co-factor for enzymes involved in glycosylation, and impairment of TMEM165 function results in glycosylation defects. Indeed, mutations of TMEM165 found in patients with congenital disorder of glycosylation (CDG) type II hamper the transport function or localization of TMEM165 [51]. Due to the importance of TMEM165 in lactate biosynthesis [52], it has also been suggested that TMEM165 could be a transporter importing both Ca^{2+} [53] and Mn^{2+} into the Golgi in exchange for protons [50]. TMEM165 proteins contain two copies of the UPF0016 domain, and each domain contains a signature motif, E-φ-G-D-(K/R)-(T/S), where φ denotes a hydrophobic amino acid. The glutamic acid of the second motif, E248, has been shown to be crucial for affecting the glycosylation function of the Golgi but not the expression of the protein [54], and so can be speculated to form part of a binding site for

transporter function. However, in the absence of an experimentally determined structure, further investigation will be required to understand the transport mechanism of TMEM165.

Proteins with sequence similarity to existing transporters

Our search uncovered a large number of proteins that show sequence similarity and thus possible relationships to existing transporters in the SLC nomenclature. Since transport activity has not been demonstrated, these proteins are either orphan transporters or they could have transceptor functions. What follows is a comprehensive discussion of these proteins, as their similarity to transporters makes them ideal targets for further studies to elucidate their putative transporter activity.

Atypical transporters

A previous effort by Perland and coworkers has uncovered novel transporter-like proteins mostly from the MFS and APC superfamilies [11], which have also been recognized by our search. In general, the function of these atypical transporters is not well known, but some have been reported to be expressed in the brain, and their expression levels seem to be affected by nutrient availability [55–58]. For MFSD1, MFSD6 and UNC93A, the study of the *D. melanogaster* and *C. elegans* orthologs have provided some information on the loss-of-function phenotype [59–62]. MFSD8 and MFSD10 have been linked to the Wolf-Hirschhorn syndrome and to LINCL (late-infantile-onset neuronal ceroid lipofuscinoses), respectively [63,64], and MFSD8 seems to be localized in the lysosomes [65,66]. More studies about the biological function and transport activity of these proteins is required to fully understand their physiological roles.

Some of the “atypical” SLC-like transporters (e.g. MFSD8, MFSD9, MFSD10 and MFSD14 proteins) clustered together with members of the classical SLC18 family, which prompted us to examine the relationship of these and neighboring proteins in more detail. We constructed multiple alignments and a phylogenetic tree of the proteins one level above these proteins in our clustering dendrogram (i.e. members of the SLC17, SLC18, SLC37 families as well as MFSD8, MFSD9, MFSD10, MFSD14A-C and SLC22A18 proteins, Figure 3). As expected, the phylogenetic tree gives a better separation of these very similar proteins than the HMM fingerprint-based dendrogram, and the branch support values suggest a clear separation of the SLC17, SLC18 and SLC37 families. In addition, the phylogenetic tree highlights that the atypical SLC proteins MFSD8, MFSD9, MFSD10, MFSD14A-C as well as SLC22A18, while being more divergent, are likely to have evolved from a single common ancestor. The relationship between the MFSD9, MFSD10 and MFSD14A-B proteins also agrees with earlier studies [Perland2017a, Perland2017b]. Similarly, the evolutionary dendrogram created using all 7 organisms in our study for the SLC18 family (Supplementary File 1) suggests that MFSD9, MFSD10, MFSD14A-B and SLC22A18 likely share a common evolutionary origin and are thus more closely related to each other than to SLC18 proteins, while MFSD8 is more distantly related. Further studies may be required to elucidate the particular evolutionary relationship between these proteins.

Interestingly, **MFSD3** has clustered together with the SLC33A1 protein in our HMM fingerprint-based clustering analysis. Indeed, the “Acatn” (Acetyl-coenzyme A transporter 1) Pfam domain is present in MFSD3, albeit with a relatively low score, but with significant e-value (5.6e-12). Sequence alignment between SLC33A1 and MFSD3 gives 18.2% sequence identity. Even though the sequence identity between MFSD3 and SLC33A1 is relatively low, the “Acatn” domain was found only in these proteins. The relatedness of MFSD3 and SLC33A1 is also corroborated by previous results of other groups [57]. The biological function of MFSD3 is still unclear [57,67,68].

The **TMEM104** protein in our analysis clustered together with amino acid transporter families SLC32, SLC36 and SLC38. Based on multiple and pairwise sequence alignments and sequence identity, TMEM104 was most similar to SLC38A7 (13.3-15.1%), SLC38A8 (13.1-15.1%), and SLC36A1 (10.9-16.1%). Interestingly, TMEM104 also bears moderate similarity to the “Aa_trans” Pfam domain, which describes the transmembrane region of SLC38 proteins. In our SLC classification dendrogram

(Figure 1), TMEM104 clustered with SLC38 proteins, even though it seems to be an outlier from the family, similarly to SLC38A9. In addition, TMEM104, SLC38A7 and SLC38A8 all show low similarity to the "Trp_Tyr_perm" Pfam domain, which describes bacterial tyrosine and tryptophan permeases. Despite the low sequence similarity to SLC38 members, these data suggest that TMEM104 might be an amino acid transporter distantly related to the SLC38 family. To get a more detailed picture of the evolutionary relationship of TMEM104 and the SLC38, SLC36 and SLC32 families, we constructed a multiple alignment and a phylogenetic tree of these proteins (Figure 4). While the tree undoubtedly separates the SLC32 and SLC36 clades due to high branch support values, TMEM104 could not be clearly separated from the SLC38 family, and it likely has a similar relationship to the rest of the family as SLC38A9, playing a transceptor role in cells. However, there is currently no experimental evidence of this and the biological function of TMEM104 remains elusive.

Proteins similar to SLC35 transporters

Interestingly, our search revealed several proteins that show sequence similarity to transporters of the SLC35 family. SLC35 proteins are currently classified into subfamilies A-G, which have relatively low sequence identity among them (4.0-22.0%). SLC35 transporters belong to the family of "DMT" (drug-metabolite transporters), which is classified in TCDB family #2.A.7, and corresponds to a clan of Pfam families, called DMT. Currently known substrates of human SLC35 members include nucleotide-sugar conjugates [69]. However, the substrate range of this superfamily is substantially larger [70].

The **TMEM144** proteins harbors its dedicated Pfam domain called "TMEM144", which itself is a member of the DMT clan of transporters. Its relatedness to the DMT family is further corroborated by high-scoring similarity of the TCDB subfamily #2.A.7.8 to the sequence of TMEM144. Otherwise, functionally, the protein is uncharacterized, although it might be related to sterol metabolism/transport, because its function has been linked to bovine milk cholesterol levels [71], the hypothalamic-gonadal axis and testosterone response [72]. It is also highly expressed in the hypothalamus [72].

TMEM234 is classified in the TCDB #2.A.7.32 family and also contains a corresponding "TMEM234" Pfam domain, which is a member of the "DMT" clan of Pfam domains as well. The physiological role of TMEM234 is not known. However, in zebrafish, its homolog might play a role in the formation of the kidney filtration barrier, as its knockdown causes proteinurea [73].

In our clustering analysis, **TMEM241** clustered with the SLC35 family very closely. Its HMM fingerprint shows similarity to the TC #2.A.7.13 subfamily, and weak similarity to the "TPT" Pfam domain (which also belongs to the "DMT" Pfam clan) over the whole length of the protein. The proteins in the #2.A.7.13 family are Golgi GDP-mannose:GMP antiporters from plants, yeast and other organisms [74,75], but not from vertebrates. Nevertheless, the protein seems to be present in many higher organisms according to the Swiss-Prot database. However, these protein are not listed in the TCDB. The biological function of TMEM241 is still unknown, but it has been suggested to affect serum triglyceride levels [76].

Due to the sequence diversity of the SLC35 family, we were interested in the relationships between individual proteins. To this end, we have built a phylogenetic tree of human proteins that showed similarity to existing SLC35 transporters (Figure 5). In the tree, most SLC35 subfamilies could be resolved as a single clade, while TMEM241 and TMEM234 form clades with SLC35D and SLC35F3-5 proteins with a support value of 0.91 and 0.71, respectively. The relationship of TMEM241 with the SLC35D subfamily is also supported by our HMM fingerprint-based clustering results. TMEM241 shows 12.0-21.4% sequence identity with SLC35D proteins. In contrast, our phylogenetic tree with SLC35 proteins from all 7 organisms (Supplementary File 1) indicated that TMEM241 is most closely related to SLC35E4. On the other hand, TMEM234 only weakly associated with SLC35F proteins, with sequence identities 3.6-9.2%. TMEM144 appears to be only distantly related to SLC35 proteins. The

elucidation of evolutionary relationships between proteins in the SLC35 thus likely requires further investigation.

Others

GPR155 is an enigmatic protein that seems to be a concatenation of a membrane transporter domain (Pfam: "Mem_trans") and a G-protein coupled receptor (GPCR) domain, which might be the reason why it is annotated as a GPCR. The membrane transporter part seems to be most similar to TCDB #2.A.69.3 subfamily proteins, which are annotated as malate/malonate transporters in the Auxin Efflux Carrier (AEC) family (#2.A.69). Gene structure analysis suggests that the concatenation is real [77], and both the human, mouse and fruitfly proteins seem to contain 17 TMHs according to UniProt annotations.

The "Mem_trans" domain is only present in GPR155 from all human proteins analyzed, and matches the first 10 TMHs of the protein in a 5+5 arrangement. In our structural search, the second half of this transporter domain (TMHs 6-10) of GPR155 exhibits similarity to the N-terminal half of sodium/bile transporters of the AsbT fold (SLC10 family). On the other hand, the last 7 TMHs of GPR155 (TMHs 11-17) indeed show similarity to GPCR-fold (7-TM) proteins with known structure, with highest similarity to structures of the human Smoothened receptor homolog (PDB ID: 6OT0). Interestingly, in our search, the N-terminal half of the transporter domain of GPR155 did not show any structural homologues.

The precise function of GPR155 still remains elusive. However, because highest expression levels were found in the brain, especially in GABAergic neurons, it might play a role in GABAergic neurotransmission [77]. It also has been suggested that GPR155 might play a role in neurons involved in motor brain function as well as sensory information processing [77]. In *D. melanogaster*, knockdown of the homologous gene, "anchor", resulted in increased wing size and thickened veins [78]. This phenotype was similar to what appeared in bone morphogenetic protein (BMP) signaling gain-of-function experiments [78]. GPR155 has also been linked to a number of different cancers [79,80].

Somewhat surprisingly, **KDEL** receptor proteins also turned up in our search (Supplementary Table 2). The structure of KDEL receptors is known and interestingly reveals similarity to the structure of SWEET transporters, having 7 TMHs in a 3+1+3 arrangement [81]. Indeed KDEL receptors have a common evolutionary origin with SWEET as well as PQ-loop transporters [82], which are represented in human by the SLC50 and SLC66 families, respectively. Oddly, while KDEL receptors have been well characterized, they are not known to harbor transport activity.

The **RFT1** protein was originally thought to be a scramblase of lipid-linked origosaccharides [83]. However, these molecules have at least 12-14 sugar moieties, so given their size, it is unlikely that a single transporter could catalyze their flipping. Later studies refuted the scramblase concept and suggested instead that RFT1 could serve as an accessory protein to a flippase, but would not act as a flippase itself [84–86].

Nevertheless, the corresponding "Rft1" Pfam domain shows similarity to multidrug and toxic compound extrusion (MATE) transporters (SLC47 family) and belongs to a clan of Pfam domains ("MviN_MATE") that contain transporters as well. In line with this, human RFT1 showed significant similarity to MATE transporters in our search for structural homologs, indicating likely structural similarity. Some members of the corresponding TC #2.A.66.3 subfamily also contain weak hits of the "MatE" Pfam domain. Thus, while RFT1 shows similarity to existing transporters, its biological function is still unclear.

The C-terminal half of the **TMEM245** protein shows weak similarity to TC #2.A.86 proteins (Autoinducer-2 Exporter/AI-2E family), which contain both small-molecule exporters [87,88] as well as Na^+/H^+ antiporter proteins [89–91]. Accordingly, TMEM245 also has weak similarity to the corresponding Pfam domain ("AI-2E_transport").

The HMMs of TC #2.A.86 and #2.A.86.1 match from residues 444-866 of human TMEM245, which are the last 6 TMHs according to UniProt predictions. The last 5 TMHs are separated from the previous ones by a slightly larger loop. This architecture is similar to the 3+5 arrangement of the previously described bacterial $Na^+(Li^+)/H^+$ antiporter TC #2.A.86.1.14 according to UniProt predictions (accession code: NLHAP_HALAA). This bacterial protein also matched the full-length "AI-2E_transport" domain from Pfam, while only the last 5 TMHs of TMEM245 match with C-terminal region of "AI-2E_transport". The human TMEM245 protein contains 14 TMHs in total according to UniProt predictions. Thus, TMEM245 might have a transporter-like domain at the C-terminus. In terms of the structure, we have not found any similarities to proteins with known structure. Therefore both TMEM245 and bacterial exporters and antiporters in the TC #2.A.86 family are likely to have a yet uncharacterized tertiary structure. Functionally, the TMEM245 protein also remains elusive.

The **TMEM41A**, **TMEM41B**, and **TMEM64** proteins clustered to the same family in our results. These are the only proteins in human that show any similarity to the "SNARE_assoc" Pfam domain, as well as to the TCDB family #9.B.27. While no protein with this domain or from this family has direct evidence for transport activity, Pfam reports SCOOP-based similarity [92] of the "SNARE_assoc" domain with "Sm_multidrug_ex", which is a domain encoding transporter proteins. Some members of the family in the TCDB have been proposed to be "cation:proton importers" (#9.B.27.2.2) or "selenite transport proteins" (#9.B.27.2.3).

The most well-characterized member of the human protein family is TMEM41B. Interestingly, a recent study reported a putative structure generated *ab initio* using evolutionary covariance-derived information [93]. Strikingly, this structural model shows features reminiscent of secondary transporters, such as a tandem internal repeat with two-fold rotational symmetry, and the authors suggest a H^+ antiporter activity as a mechanism of transport [93].

While the exact function of TMEM41B is still unclear, it forms a complex with vacuole membrane protein 1 (VMP1), also harboring the "SNARE_assoc" domain, and both are required for autophagosome formation [94,95]. Tmem41b localized to mitochondria-associated ER membranes [96–98]. Interestingly, TMEM41B seems to be an absolutely required factor for SARS-CoV-2 [99], and probably also flaviviral [100] infection, possibly by facilitating a membrane curvature that is beneficial for viral replication [100].

The proteins **TMEM184A**, **TMEM184B**, **TMEM184C** clustered together with **SLC51A** (family of transporters of steroid-derived molecules) in our analysis. While human TMEM184B and SLC51A are included in the TCDB as members of family #2.A.82, TMEM184A and C are not. Independently, the "Solute_trans_a" Pfam domain was found in all four proteins with high scores and significance, but not in other human proteins. Therefore, it is likely that the four proteins, TMEM184A-C and SLC51A, are homologous. In spite of this, sequence identity between TMEM184 proteins and SLC51A is low (12.3%-13.6%), but moderate among TMEM184 proteins (26.5%-62.0%). All four proteins are predicted to harbor 7 TM helices according to UniProt, yet our search has found no similar proteins with known structures.

TMEM184A was identified as a heparin receptor in vascular cells [101], but no transport activity has been reported. Interestingly, while SLC51A is known to function as a bile acid transporter [102–104], TMEM184B has been proposed to be responsible for ibuprofen uptake [105]. This is interesting in view of the partial chemical similarity between steroid acids and ibuprofen, both harboring a hydrophobic hydrocarbon part and a carboxyl moiety. TMEM184C resides in a locus that has been suggested to be responsible for the pathogenicity of X-linked congenital hypertrichosis syndrome [106], but no transport activity has been suggested.

Putative transporters

Our search also identified proteins whose transport activity is either controversial or not characterized, and which do not show sequence similarity to transporters of known function. Thus, the proteins in these families require further investigation to uncover their function.

The **CNNM1-4** proteins (also called ACDP1-4) are distant homologs of the cyclins, but have no documented enzymatic activity. Instead, CNNM proteins belong to a highly conserved family of Mg^{2+} transport-related proteins [107], and CNNM2 and CNNM4 have been proposed to be the long sought after basolateral Na^{+}/Mg^{2+} exchangers in the kidney and intestine, respectively [108,109]. The function of these proteins is, however, controversial [110–113], and there are hypotheses that CNNM proteins *per se* are not Mg^{2+} transporters [114]. Most recently, however, the structure of a bacterial homolog, CorC, has been resolved, revealing its membrane topology, as well as a conserved Mg^{2+}-binding site [115]. Strikingly, the Mg^{2+} ions in the structure are fully dehydrated, in contrast to those in other known Mg^{2+} channel structures [115], which makes it unlikely that the proteins function via a channel-like transport mechanism. In line with this, the authors suggest an alternating-access exchange mechanism [115], however, further studies are required to understand how and whether CNNM proteins might be able to mediate the translocation of Mg^{2+} ions across the membrane.

A family of 4 lysosomal-associated transmembrane proteins (**LAPTM4A**, **LAPTM4B**, **LAPTM5** and sequence **B4E0C1**) turned up in our search, corresponding to the TCDB subfamily #2.A.74.1 and Pfam domain "Mtp" (mouse transporter protein). The family also includes an uncharacterized transcript with the UniProt accession "B4E0C1". Originally, the mouse transporter protein (Mtp, ortholog of LAPTM4A) was characterized as a transporter mediating the transport of nucleosides and nucleobases between the cytoplasm and intracellular compartments [116], and was later also associated with multidrug-resistance (MDR) in yeast, where its expression changed the subcellular compartmentalization of a heterogenous group of compounds [117,118]. LAPTM4A was shown to be involved in glycosylation and glycolipid regulation [119,120]. All three proteins seem to be lysosomal [116,121,122]. Nevertheless, these proteins appear to interact with other characterized transporters, such as SLC22A2 (hOCT2), SLC7A5/SLC3A2 (LAT1/4F2hc) and MDR-related ABC transporters [123–125]. But it has been claimed that they are not *per se* transporters, but rather regulatory factors, either assisting the localization and targeting or the function of other transporters [125]. Interestingly, the transcript "B4E0C1" appears to have 4 TMHs at its N-terminus, which is identical to human LAPTM5 apart from a ~40-amino acid insertion between TMH3 and TMH4. This region also shows significant similarity to both the "Mtp" Pfam domain and the TC #2.A.74.1 subfamily. However, the C-terminal region of the transcript is identical to the C-terminal segment of "actin filament-associated protein 1-like 1" protein (UniProt accession Q8TED9). We did not find any similar fusion sequences in the other organisms we analyzed. The "Mtp" Pfam domain, which is the hallmark of the family, belongs to the "Tetraspannin" Pfam clan, which has no other domains with annotated transporter function and no similarity to existing transporters. Structural information is also not available. Therefore, the transport function of these proteins requires further investigation.

Our search identified four proteins in human (**LMBR1**, **LMBR1L**, **LMBD1/LMBRD1**, **LMBRD2**) bearing the "LMBR1" Pfam domain, which clustered into two families in our analysis. These proteins correspond to TCDB family #9.A.54. The LMBD1 protein, encoded by the LMBRD1 gene, was suggested to function as a vitamin B12 (cobalamin) transporter, exporting vitamin B12 from the lysosomes into the cytoplasm [126]. However, it was later shown that LMBD1 actually interacts with ABCD4 and assists in its lysosomal trafficking [127], and that ABCD4 transports vitamin B12 even in the absence of LMBD1 [128]. Therefore, it is likely that LMBD1 itself is not a vitamin B12 transporter. LMBD1 was originally coined as having "significant homology" to lipocalin membrane receptors [126], and indeed the LIMR (lipocalin-1-interacting membrane receptor) protein, encoded by the LMBR1L gene, is responsible for binding lipocain 1 (LCN-1) with high affinity [129,130]. LMBRD2 was

proposed to be a regulator of β2-adrenoceptor signaling [131], while the first protein identified in the family, LMBR1, was associated with polydactyly and limb malformations [132,133]. However, its physiological role is still elusive. The proteins seem to contain 9 TMHs in a 5+4 arrangement according to UniProt predictions, but the tertiary structure of the proteins is still unknown, and no homologs with a known structure were found in our search.

Two MagT1-like proteins (**MAGT1** and **TUSC3**), as well as **OSTC** and **OSTCL** (oligosaccharyltransferase complex subunit) turned up in our search, showing similarity to TCDB family #1.A.76 members. MAGT1 and TUSC3 also have high-scoring hits for the Pfam domain "OST3_OST6", which is characteristic of members of the oligosaccharyltransferase (OST) complex. TUSC3 (also called N33) was first identified as a tumor suppressor gene [134], and its presence, together with that of MAGT1 in the OST complex has been attested later on [135–139]. Therefore, it was suggested that these proteins act as oxidoreductases [138]. Meanwhile, MAGT1 and TUSC3 were also proposed to act as Mg^{2+} transporters [140,141]. On the other hand, recent structural findings of human MAGT1 [139] indicated that this protein may not function as a transporter or channel due to the lack of substrate-binding site or pore. OSTC (also called DC2) has similarly been shown to be part of the OST complex and to have a structure similar to MAGT1 [139]. Whether MagT1-like proteins still have a transport function remains to be clarified.

The **TMEM14A**, **TMEM14B** and **TMEM14C** proteins are the only ones in human containing the "Tmemb_14" Pfam domain. While Pfam lists this domain as functionally uncharacterized, a plant protein (FAX1) containing this domain was suggested to be involved in fatty acid export from chloroplasts [142]. However, the physiological roles of TMEM14A and TMEM14B in human remain elusive. The third member of the family, TMEM14C, was identified as a putative mitochondrial protein whose transcript is consistently coexpressed with proteins from the core machinery of heme biosynthesis [143]. It was later shown that TMEM14C mediates the import of protoporphyrinogen IX (PPgenIX) into the mitochondrial matrix [144,145]. While the structure of TMEM14C was solved using nuclear magnetic resonance (NMR) [146], showing a bundle of three TM helices and an amphipathic helix, the transport mechanism remains elusive. Interestingly, despite their proposed function in the mitochondria, TargetP-2.0 [147] did not predict a mitochondrial targeting sequence in the amino acid sequence of any of the human TMEM14 proteins in our hands.

TMEM205 is a 4-TMH protein according to UniProt annotations, which was linked to cisplatin resistance [148]. The protein is expressed mostly in liver, pancreas and adrenal glands, and is present on the plasma membrane [148]. TMEM205-mediated resistance was shown to be selective towards platinum-based drugs, such as cisplatin and oxaliplatin, but not carboplatin [149]. While structural information about the protein is not available, mutagenesis studies of TMEM205 showed that mutating sulfur-containing residues, especially in TMH2 and TMH4, diminishes the effect of cisplatin resistance [149]. Nevertheless, neither the biological function nor the physiological substrates of TMEM205 are known.

Transporters with hydrophobic substrates

In addition to proteins that transport solutes or are similar to transporters that typically translocate water-soluble small molecule compounds, our search has uncovered numerous proteins that have been reported to take part in modulating the intracellular distribution of hydrophobic or amphipathic compounds, such as cholesterol, fat-soluble vitamins, lipids and fatty acids (Table 3). Although these proteins do not, strictly speaking, transport so-called "solutes", they translocate small hydrophobic molecules that have fundamental biological functions. Thus they belong to the SLC superfamily as well. Accordingly, they have been integrated in our search and some of them have already been included in the SLC nomenclature (SLC59, SLC63, SLC65). In view of the biological and pharmaceutical importance of the transport mechanisms of hydrophobic substrates, our results on this topic will be discussed in a separate paper.

Conclusions

Our study represents the first systematic correlation of the SLC and TCDB nomenclature schemes. Many of the transport proteins discovered in our search are underexplored and there is limited information about them, although they often have important physiological roles and/or potentially represent new therapeutic targets. Even with proteins that have been studied for their physiological involvement, it was often not taken into account that they could have a transport function. Numerous proteins uncovered in our search have similarity to proteins with transport function in other organisms, but their physiological substrates remain unknown. These proteins will be interesting targets for deorphanization studies, to reveal their natural substrates. In our work, we also highlight proteins for which transport activity has been controversial, and more specific analyses are required to clarify their biological function. In addition, our search reveals new SLC-like proteins that have no structural information. This hinders a deeper understanding of their transport mechanism. Future structure determinations would be of crucial importance to accelerate validation of the identified proteins. The combination of all these efforts would greatly facilitate the completion of the SLC-ome in human cells. Thus, our study points out important directions in which future studies could help resolve the lack of information about SLC transporters, which will help unlock their therapeutic potential.

Acknowledgements

We acknowledge support by the Swiss National Science Foundation for the grants # CRSII5_180326 ("The role of mitochondrial carriers in metabolic tuning and reprogramming by calcium flow across membrane contact sites") and # 310030_182272 ("Intestinal absorption of transition metals in human health and disease") as well as by the Swiss National Research Programme NRP 78, # 4078P0_198281 ("New insights into the COVID-19 pandemic"). Calculations were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern.

Methods

Transmembrane segment prediction

In certain cases, especially with less well-characterized putative transporters of the TC# 9.A and 9.B superfamilies, transmembrane helix prediction based on HMMTOP [150] was used to assess the number of transmembrane helices.

HMM building for TCDB families and subfamilies

Sequences of selected TCDB families and subfamilies were aligned using PSI-Coffee 11.00 [151] and NCBI BLAST+ 2.6.0 [152] using the "nr" BLAST database of 2018-02-12. The alignments were turned into HMMs using the "hmmbuild" command of HMMER 3.1b2 [30] using default settings.

Sequence similarity search

All protein belonging to each of the 7 organisms studied were downloaded from UniProt into a FASTA file. Sequence similarity search was performed using the HMMs downloaded for selected Pfam families and those generated for selected TCDB families and subfamilies using "hmmsearch" from HMMER 3.1b2. Hits with bit scores larger than 50 were used for further analysis.

Sequence clustering for fragment elimination

Since the downloaded protein set from UniProt contained fragments as well as predicted open reading frames and sequences from genomic screening methods, we strived to retain one sequence per gene for further analysis. In order to achieve this, hits yielded by the sequence similarity search were clustered using the following method. First, all-against-all similarity searches were performed using NCBI BLAST+ 2.4.0 [152]. Sequences were assigned the same cluster if they either share common gene annotations according to their UniProt records, or a high-scoring segment pair (HSP) with more than 95% sequence identity based on the all-against-all BLAST search. For gene

annotation, the fields Gene Symbol (GN), HGNC symbol, GeneID, UniGene, FlyBase, KEGG identifiers were used from UniProt records. Ambiguous or conflicting annotations, as well as annotations conflicting with BLAST-reported sequence similarity were detected and resolved manually. Afterwards, clusters were reduced to representative sequences. For clusters containing a single Swiss-Prot sequence, that sequence was taken as representative. For cluster with no Swiss-Prot sequence, the longest sequence of the cluster was taken as representative. Clusters with more than one Swiss-Prot sequence were manually analyzed and split if necessary.

HMM fingerprint-based sequence clustering

We introduce the concept of a "HMM fingerprint", which is a mathematical vector of numbers assigned to a protein sequence, corresponding to the bit scores of similarity to each of the set of HMMs used in our analysis, consisting of the HMMs of TCDB families and subfamilies, as well as Pfam HMMs. We have restricted the number of HMMs to those that gave a hit with bit score > 25, in total 513 HMMs. Two proteins sequences that are related are expected to show similarity to a similar subset of TCDB families, subfamilies, or Pfam families, and therefore a similar pattern in their HMM fingerprints. In turn, if two protein sequences show similarity to the same subset of TCDB families, subfamilies or Pfam families, as indicated by a similar HMM fingerprint, then they can be expected to be related. Once the HMM fingerprint has been assigned to each protein sequence found in our search, the unweighted pair group method with arithmetic mean (UPGMA) method [SokalMichener1958] was used using the cosine metric to arrive at a hierarchical clustering of the sequences. The tree representing the clustering was cut at 0.7 cosine distance to arrive at branches that formed the basis of protein families. For some families, the threshold was either elevated to join outlier sequences (SLC5A7, SLC10A7, SLC25A46, SLC30A9, SLC39A9, MPDU1/SLC66A5, SLC9B subfamily, SLC35 subfamilies) or reduced to separate highly similar families (SLC32-36-38, SLC2-22, SLC17-18-37).

Structural search

Sequences of SLC-like proteins were turned into HMMs using "hhblits" from the HH-suite3 package [153], using the UniRef30 database of 2020-06 [154,155], 3 iterations, an E-value threshold of 1e-3 for inclusion, and a probability threshold of 0.35 for MAC re-alignment. The HMMs were searched against the "pdb70" database of 2021-08-04 [154] with no MAC realignment, "predicted vs predicted" secondary structure scoring, and amino acid score of 1. The resulting hits were checked against PDB annotations of transmembrane helices, and were accepted if at least 3 TM segments were contained within the aligned region and the E-value of the hit was less than 1e-4.

Phylogenetic trees

Selected groups of SLC-like proteins were aligned using Clustal Omega 1.2.1 [156,157] with 5 iterations and default settings. Smart Model Selection 1.8.4 [158] and PhyML 3.3.20190909 [159] were used to generate the phylogenetic trees, with 10 random starting trees and using the approximate likelihood ratio test aLRT method [160]. Trees in main figures were visualized using TreeViewer 1.2.2 [https://treeviewer.org/]. Tree rooting, rearrangement (with threshold 0.9) and reconciliation with the species tree was done using NOTUNG 2.9.1.5 [161]. Reconciled phylogenetic trees were visualized using custom Python scripts in the style used by NOTUNG. Orthologs were identified using the reconciled phylogenetic trees and custom Python scripts.

References

1. Hediger MA, Clémençon B, Burrier RE, Bruford EA. The ABCs of membrane transporters in health and disease (SLC series): introduction. Mol Aspects Med. 2013;34: 95–107. doi:10.1016/j.mam.2012.12.009

2. César-Razquin A, Snijder B, Frappier-Brinton T, Isserlin R, Gyimesi G, Bai X, et al. A Call for Systematic Research on Solute Carriers. Cell. 2015;162: 478–487. doi:10.1016/j.cell.2015.07.022

3. Hediger MA, Romero MF, Peng J-B, Rolfs A, Takanaga H, Bruford EA. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. Pflugers Arch. 2004;447: 465–468. doi:10.1007/s00424-003-1192-y
4. Bai X, Moraes TF, Reithmeier RAF. Structural biology of solute carrier (SLC) membrane transport proteins. Mol Membr Biol. 2018; 1–32. doi:10.1080/09687688.2018.1448123
5. Choi S, Jeon J, Yang J-S, Kim S. Common occurrence of internal repeat symmetry in membrane proteins. Proteins. 2008;71: 68–80. doi:10.1002/prot.21656
6. Forrest LR. Structural Symmetry in Membrane Proteins. Annu Rev Biophys. 2015;44: 311–337. doi:10.1146/annurev-biophys-051013-023008
7. Jardetzky O. Simple allosteric model for membrane pumps. Nature. 1966;211: 969–970. doi:10.1038/211969a0
8. Forrest LR, Krämer R, Ziegler C. The structural basis of secondary active transport mechanisms. Biochim Biophys Acta. 2011;1807: 167–188. doi:10.1016/j.bbabio.2010.10.014
9. Fredriksson R, Nordström KJV, Stephansson O, Hägglund MGA, Schiöth HB. The solute carrier (SLC) complement of the human genome: phylogenetic classification reveals four major families. FEBS Lett. 2008;582: 3811–3816. doi:10.1016/j.febslet.2008.10.016
10. Perland E, Fredriksson R. Classification Systems of Secondary Active Transporters. Trends Pharmacol Sci. 2017;38: 305–315. doi:10.1016/j.tips.2016.11.008
11. Perland E, Bagchi S, Klaesson A, Fredriksson R. Characteristics of 29 novel atypical solute carriers of major facilitator superfamily type: evolutionary conservation, predicted structure and neuronal co-expression. Open Biol. 2017;7. doi:10.1098/rsob.170142
12. Schlessinger A, Matsson P, Shima JE, Pieper U, Yee SW, Kelly L, et al. Comparison of human solute carriers. Protein Science. 2010;19: 412–428. doi:10.1002/pro.320
13. Schlessinger A, Yee SW, Sali A, Giacomini KM. SLC classification: an update. Clin Pharmacol Ther. 2013;94: 19–23. doi:10.1038/clpt.2013.73
14. Saier MH. Molecular phylogeny as a basis for the classification of transport proteins from bacteria, archaea and eukarya. Adv Microb Physiol. 1998;40: 81–136.
15. Saier MH, Reddy VS, Moreno-Hagelsieb G, Hendargo KJ, Zhang Y, Iddamsetty V, et al. The Transporter Classification Database (TCDB): 2021 update. Nucleic Acids Res. 2021;49: D461–D467. doi:10.1093/nar/gkaa1004
16. Cao Y, Jin X, Huang H, Derebe MG, Levin EJ, Kabaleeswaran V, et al. Crystal structure of a potassium ion transporter, TrkH. Nature. 2011;471: 336–340. doi:10.1038/nature09731
17. Bosson R, Jaquenoud M, Conzelmann A. GUP1 of Saccharomyces cerevisiae encodes an O-acyltransferase involved in remodeling of the GPI anchor. Mol Biol Cell. 2006;17: 2636–2645. doi:10.1091/mbc.e06-02-0104
18. Berks BC, Sargent F, Palmer T. The Tat protein export pathway. Mol Microbiol. 2000;35: 260–274. doi:10.1046/j.1365-2958.2000.01719.x
19. Berks BC, Sargent F, De Leeuw E, Hinsley AP, Stanley NR, Jack RL, et al. A novel protein transport system involved in the biogenesis of bacterial electron transfer chains. Biochim Biophys Acta. 2000;1459: 325–330. doi:10.1016/s0005-2728(00)00168-7
20. Ren Q, Kang KH, Paulsen IT. TransportDB: a relational database of cellular membrane transport systems. Nucleic Acids Res. 2004;32: D284-288. doi:10.1093/nar/gkh016
21. Ren Q, Chen K, Paulsen IT. TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. Nucleic Acids Res. 2007;35: D274-279. doi:10.1093/nar/gkl925
22. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. Nucleic Acids Res. 2017;45: D320–D324. doi:10.1093/nar/gkw1068
23. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, et al. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 2001;29: 22–28. doi:10.1093/nar/29.1.22

24. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E. TIGRFAMs and Genome Properties in 2013. Nucleic Acids Res. 2013;41: D387-395. doi:10.1093/nar/gks1234
25. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44: D279-285. doi:10.1093/nar/gkv1344
26. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305: 567–580. doi:10.1006/jmbi.2000.4315
27. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47: D506–D515. doi:10.1093/nar/gky1049
28. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. Nucleic Acids Res. 2019;47: D427–D432. doi:10.1093/nar/gky995
29. Finn RD, Mistry J, Schuster-Böckler B, Griffiths-Jones S, Hollich V, Lassmann T, et al. Pfam: clans, web tools and services. Nucleic Acids Res. 2006;34: D247-251. doi:10.1093/nar/gkj149
30. Eddy SR. Accelerated Profile HMM Searches. PLoS Comput Biol. 2011;7: e1002195. doi:10.1371/journal.pcbi.1002195
31. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A. UniProtKB/Swiss-Prot. Methods Mol Biol. 2007;406: 89–112. doi:10.1007/978-1-59745-535-0_4
32. Gyimesi G, Hediger MA. Sequence Features of Mitochondrial Transporter Protein Families. Biomolecules. 2020;10. doi:10.3390/biom10121611
33. Horiba N, Masuda S, Takeuchi A, Takeuchi D, Okuda M, Inui K. Cloning and characterization of a novel Na+-dependent glucose transporter (NaGLT1) in rat kidney. J Biol Chem. 2003;278: 14669–14676. doi:10.1074/jbc.M212240200
34. Nawata CM, Dantzler WH, Pannabecker TL. Alternative channels for urea in the inner medulla of the rat kidney. Am J Physiol Renal Physiol. 2015;309: F916-924. doi:10.1152/ajprenal.00392.2015
35. Tejada-Jiménez M, Galván A, Fernández E. Algae and humans share a molybdate transporter. Proc Natl Acad Sci USA. 2011;108: 6420–6425. doi:10.1073/pnas.1100700108
36. Zhu W, Spiga L, Winter S. Transition metals and host-microbe interactions in the inflamed intestine. Biometals. 2019;32: 369–384. doi:10.1007/s10534-019-00182-8
37. Barth J, Zimmermann H, Volknandt W. SV31 is a Zn2+-binding synaptic vesicle protein. J Neurochem. 2011;118: 558–570. doi:10.1111/j.1471-4159.2011.07344.x
38. Waberer L, Henrich E, Peetz O, Morgner N, Dötsch V, Bernhard F, et al. The synaptic vesicle protein SV31 assembles into a dimer and transports Zn2. J Neurochem. 2017;140: 280–293. doi:10.1111/jnc.13886
39. Cuajungco MP, Kiselyov K. The mucolipin-1 (TRPML1) ion channel, transmembrane-163 (TMEM163) protein, and lysosomal zinc handling. Front Biosci (Landmark Ed). 2017;22: 1330–1343.
40. Sanchez VB, Ali S, Escobar A, Cuajungco MP. Transmembrane 163 (TMEM163) protein effluxes zinc. Arch Biochem Biophys. 2019;677: 108166. doi:10.1016/j.abb.2019.108166
41. Burré J, Zimmermann H, Volknandt W. Identification and characterization of SV31, a novel synaptic vesicle membrane protein and potential transporter. J Neurochem. 2007;103: 276–287. doi:10.1111/j.1471-4159.2007.04758.x
42. Wang L, Li N-N, Lu Z-J, Li J-Y, Peng J-X, Duan L-R, et al. Association of three candidate genetic variants in ACMSD/TMEM163, GPNMB and BCKDK /STX1B with sporadic Parkinson's disease in Han Chinese. Neurosci Lett. 2019;703: 45–48. doi:10.1016/j.neulet.2019.03.019
43. Chang K-H, Chen C-M, Chen Y-C, Fung H-C, Wu Y-R. Polymorphisms of ACMSD-TMEM163, MCCC1, and BCKDK-STX1B Are Not Associated with Parkinson's Disease in Taiwan. Parkinsons Dis. 2019;2019: 3489638. doi:10.1155/2019/3489638
44. Lauterbach EC. Psychotropic drug effects on gene transcriptomics relevant to Parkinson's disease. Prog Neuropsychopharmacol Biol Psychiatry. 2012;38: 107–115. doi:10.1016/j.pnpbp.2012.03.011

45. Chakraborty S, Vellarikkal SK, Sivasubbu S, Roy SS, Tandon N, Bharadwaj D. Role of Tmem163 in zinc-regulated insulin storage of MIN6 cells: Functional exploration of an Indian type 2 diabetes GWAS associated gene. Biochem Biophys Res Commun. 2019. doi:10.1016/j.bbrc.2019.11.117
46. Tabassum R, Chauhan G, Dwivedi OP, Mahajan A, Jaiswal A, Kaur I, et al. Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. Diabetes. 2013;62: 977–986. doi:10.2337/db12-0406
47. Bai H, Liu H, Suyalatu S, Guo X, Chu S, Chen Y, et al. Association Analysis of Genetic Variants with Type 2 Diabetes in a Mongolian Population in China. J Diabetes Res. 2015;2015: 613236. doi:10.1155/2015/613236
48. Foulquier F, Amyere M, Jaeken J, Zeevaert R, Schollen E, Race V, et al. TMEM165 deficiency causes a congenital disorder of glycosylation. Am J Hum Genet. 2012;91: 15–26. doi:10.1016/j.ajhg.2012.05.002
49. Demaegd D, Foulquier F, Colinet A-S, Gremillon L, Legrand D, Mariot P, et al. Newly characterized Golgi-localized family of proteins is involved in calcium and pH homeostasis in yeast and human cells. Proc Natl Acad Sci USA. 2013;110: 6859–6864. doi:10.1073/pnas.1219871110
50. Foulquier F, Legrand D. Biometals and glycosylation in humans: Congenital disorders of glycosylation shed lights into the crucial role of Golgi manganese homeostasis. Biochim Biophys Acta Gen Subj. 2020;1864: 129674. doi:10.1016/j.bbagen.2020.129674
51. Stribny J, Thines L, Deschamps A, Goffin P, Morsomme P. The human Golgi protein TMEM165 transports calcium and manganese in yeast and bacterial cells. J Biol Chem. 2020;295: 3865–3874. doi:10.1074/jbc.RA119.012249
52. Snyder NA, Palmer MV, Reinhardt TA, Cunningham KW. Milk biosynthesis requires the Golgi cation exchanger TMEM165. J Biol Chem. 2019;294: 3181–3191. doi:10.1074/jbc.RA118.006270
53. Reinhardt TA, Lippolis JD, Sacco RE. The Ca(2+)/H(+) antiporter TMEM165 expression, localization in the developing, lactating and involuting mammary gland parallels the secretory pathway Ca(2+) ATPase (SPCA1). Biochem Biophys Res Commun. 2014;445: 417–421. doi:10.1016/j.bbrc.2014.02.020
54. Lebredonchel E, Houdou M, Potelle S, de Bettignies G, Schulz C, Krzewinski Recchi M-A, et al. Dissection of TMEM165 function in Golgi glycosylation and its Mn2+ sensitivity. Biochimie. 2019;165: 123–130. doi:10.1016/j.biochi.2019.07.016
55. Perland E, Lekholm E, Eriksson MM, Bagchi S, Arapi V, Fredriksson R. The Putative SLC Transporters Mfsd5 and Mfsd11 Are Abundantly Expressed in the Mouse Brain and Have a Potential Role in Energy Homeostasis. PLoS ONE. 2016;11: e0156912. doi:10.1371/journal.pone.0156912
56. Lekholm E, Perland E, Eriksson MM, Hellsten SV, Lindberg FA, Rostami J, et al. Putative Membrane-Bound Transporters MFSD14A and MFSD14B Are Neuronal and Affected by Nutrient Availability. Front Mol Neurosci. 2017;10: 11. doi:10.3389/fnmol.2017.00011
57. Perland E, Hellsten SV, Lekholm E, Eriksson MM, Arapi V, Fredriksson R. The Novel Membrane-Bound Proteins MFSD1 and MFSD3 are Putative SLC Transporters Affected by Altered Nutrient Intake. J Mol Neurosci. 2017;61: 199–214. doi:10.1007/s12031-016-0867-8
58. Bagchi S, Perland E, Hosseini K, Lundgren J, Al-Walai N, Kheder S, et al. Probable role for major facilitator superfamily domain containing 6 (MFSD6) in the brain during variable energy consumption. Int J Neurosci. 2020;130: 476–489. doi:10.1080/00207454.2019.1694020
59. Valoskova K, Biebl J, Roblek M, Emtenani S, Gyoergy A, Misova M, et al. A conserved major facilitator superfamily member orchestrates a subset of O-glycosylation to aid macrophage tissue invasion. Elife. 2019;8. doi:10.7554/eLife.41801
60. Landis GN, Bhole D, Tower J. A search for doxycycline-dependent mutations that increase Drosophila melanogaster life span identifies the VhaSFD, Sugar baby, filamin, fwd and Cctl genes. Genome Biol. 2003;4: R8. doi:10.1186/gb-2003-4-2-r8
61. Kim KW, Tang NH, Piggott CA, Andrusiak MG, Park S, Zhu M, et al. Expanded genetic screening in Caenorhabditis elegans identifies new regulators and an inhibitory role for NAD+ in axon regeneration. Elife. 2018;7: e39756. doi:10.7554/eLife.39756

62. Ceder MM, Aggarwal T, Hosseini K, Maturi V, Patil S, Perland E, et al. CG4928 Is Vital for Renal Function in Fruit Flies and Membrane Potential in Cells: A First In-Depth Characterization of the Putative Solute Carrier UNC93A. Front Cell Dev Biol. 2020;8: 580291. doi:10.3389/fcell.2020.580291
63. Hannes F, Hammond P, Quarrell O, Fryns J-P, Devriendt K, Vermeesch JR. A microdeletion proximal of the critical deletion region is associated with mild Wolf-Hirschhorn syndrome. Am J Med Genet A. 2012;158A: 996–1004. doi:10.1002/ajmg.a.35299
64. Damme M, Brandenstein L, Fehr S, Jankowiak W, Bartsch U, Schweizer M, et al. Gene disruption of Mfsd8 in mice provides the first animal model for CLN7 disease. Neurobiol Dis. 2014;65: 12–24. doi:10.1016/j.nbd.2014.01.003
65. Siintola E, Topcu M, Aula N, Lohi H, Minassian BA, Paterson AD, et al. The novel neuronal ceroid lipofuscinosis gene MFSD8 encodes a putative lysosomal transporter. Am J Hum Genet. 2007;81: 136–146. doi:10.1086/518902
66. von Kleist L, Ariunbat K, Braren I, Stauber T, Storch S, Danyukova T. A newly generated neuronal cell model of CLN7 disease reveals aberrant lysosome motility and impaired cell survival. Mol Genet Metab. 2019;126: 196–205. doi:10.1016/j.ymgme.2018.09.009
67. Li Y, Yang X, Yang J, Wang H, Wei W. An 11-gene-based prognostic signature for uveal melanoma metastasis based on gene expression and DNA methylation profile. J Cell Biochem. 2018. doi:10.1002/jcb.28151
68. Nicoletti CF, Pinhel MS, Noronha NY, Jácome A, Crujeiras AB, Nonino CB. Association of MFSD3 promoter methylation level and weight regain after gastric bypass: Assessment for 3 y after surgery. Nutrition. 2020;70: 110499. doi:10.1016/j.nut.2019.04.010
69. Song Z. Roles of the nucleotide sugar transporters (SLC35 family) in health and disease. Mol Aspects Med. 2013;34: 590–600. doi:10.1016/j.mam.2012.12.004
70. Västermark Å, Almén MS, Simmen MW, Fredriksson R, Schiöth HB. Functional specialization in nucleotide sugar transporters occurred through differentiation of the gene cluster EamA (DUF6) before the radiation of Viridiplantae. BMC Evol Biol. 2011;11: 123. doi:10.1186/1471-2148-11-123
71. Do DN, Schenkel FS, Miglior F, Zhao X, Ibeagha-Awemu EM. Genome wide association study identifies novel potential candidate genes for bovine milk cholesterol content. Sci Rep. 2018;8: 13239. doi:10.1038/s41598-018-31427-0
72. Prentice LM, d'Anglemont de Tassigny X, McKinney S, Ruiz de Algara T, Yap D, Turashvili G, et al. The testosterone-dependent and independent transcriptional networks in the hypothalamus of Gpr54 and Kiss1 knockout male mice are not fully equivalent. BMC Genomics. 2011;12: 209. doi:10.1186/1471-2164-12-209
73. Rodriguez PQ, Oddsson A, Ebarasi L, He B, Hultenby K, Wernerson A, et al. Knockdown of Tmem234 in zebrafish results in proteinuria. Am J Physiol Renal Physiol. 2015;309: F955-966. doi:10.1152/ajprenal.00525.2014
74. Dean N, Zhang YB, Poster JB. The VRG4 gene is required for GDP-mannose transport into the lumen of the Golgi in the yeast, Saccharomyces cerevisiae. J Biol Chem. 1997;272: 31908–31914. doi:10.1074/jbc.272.50.31908
75. Baldwin TC, Handford MG, Yuseff MI, Orellana A, Dupree P. Identification and characterization of GONST1, a golgi-localized GDP-mannose transporter in Arabidopsis. Plant Cell. 2001;13: 2283–2295. doi:10.1105/tpc.010247
76. Rodríguez A, Gonzalez L, Ko A, Alvarez M, Miao Z, Bhagat Y, et al. Molecular Characterization of the Lipid Genome-Wide Association Study Signal on Chromosome 18q11.2 Implicates HNF4A-Mediated Regulation of the TMEM241 Gene. Arterioscler Thromb Vasc Biol. 2016;36: 1350–1355. doi:10.1161/ATVBAHA.116.307182
77. Trifonov S, Houtani T, Shimizu J-I, Hamada S, Kase M, Maruyama M, et al. GPR155: Gene organization, multiple mRNA splice variants and expression in mouse central nervous system. Biochem Biophys Res Commun. 2010;398: 19–25. doi:10.1016/j.bbrc.2010.05.162

78. Wang XC, Liu Z, Jin LH. Anchor negatively regulates BMP signalling to control Drosophila wing development. Eur J Cell Biol. 2018;97: 308–317. doi:10.1016/j.ejcb.2018.04.007
79. Shimizu D, Kanda M, Tanaka H, Kobayashi D, Tanaka C, Hayashi M, et al. GPR155 Serves as a Predictive Biomarker for Hematogenous Metastasis in Patients with Gastric Cancer. Sci Rep. 2017;7: 42089. doi:10.1038/srep42089
80. Umeda S, Kanda M, Sugimoto H, Tanaka H, Hayashi M, Yamada S, et al. Downregulation of GPR155 as a prognostic factor after curative resection of hepatocellular carcinoma. BMC Cancer. 2017;17: 610. doi:10.1186/s12885-017-3629-2
81. Bräuer P, Parker JL, Gerondopoulos A, Zimmermann I, Seeger MA, Barr FA, et al. Structural basis for pH-dependent retrieval of ER proteins from the Golgi by the KDEL receptor. Science. 2019;363: 1103–1107. doi:10.1126/science.aaw2859
82. Saudek V. Cystinosin, MPDU1, SWEETs and KDELR belong to a well-defined protein family with putative function of cargo receptors involved in vesicle trafficking. PLoS ONE. 2012;7: e30876. doi:10.1371/journal.pone.0030876
83. Helenius J, Ng DTW, Marolda CL, Walter P, Valvano MA, Aebi M. Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. Nature. 2002;415: 447–450. doi:10.1038/415447a
84. Frank CG, Sanyal S, Rush JS, Waechter CJ, Menon AK. Does Rft1 flip an N-glycan lipid precursor? Nature. 2008;454: E3-4; discussion E4-5. doi:10.1038/nature07165
85. Gottier P, Gonzalez-Salgado A, Menon AK, Liu Y-C, Acosta-Serrano A, Bütikofer P. RFT1 Protein Affects Glycosylphosphatidylinositol (GPI) Anchor Glycosylation. J Biol Chem. 2017;292: 1103–1111. doi:10.1074/jbc.M116.758367
86. Verchère A, Cowton A, Jenni A, Rauch M, Häner R, Graumann J, et al. Complexity of the eukaryotic dolichol-linked oligosaccharide scramblase suggested by activity correlation profiling mass spectrometry. Sci Rep. 2021;11: 1411. doi:10.1038/s41598-020-80956-0
87. Nobre LS, Al-Shahrour F, Dopazo J, Saraiva LM. Exploring the antimicrobial action of a carbon monoxide-releasing compound through whole-genome transcription profiling of Escherichia coli. Microbiology (Reading). 2009;155: 813–824. doi:10.1099/mic.0.023911-0
88. Herzberg M, Kaye IK, Peti W, Wood TK. YdgG (TqsA) controls biofilm formation in Escherichia coli K-12 through autoinducer 2 transport. J Bacteriol. 2006;188: 587–598. doi:10.1128/JB.188.2.587-598.2006
89. Dong P, Wang L, Song N, Yang L, Chen J, Yan M, et al. A UPF0118 family protein with uncharacterized function from the moderate halophile Halobacillus andaensis represents a novel class of Na+(Li+)/H+ antiporter. Sci Rep. 2017;7: 45936. doi:10.1038/srep45936
90. Wang L, Zou Q, Yan M, Wang Y, Guo S, Zhang R, et al. Polar or Charged Residues Located in Four Highly Conserved Motifs Play a Vital Role in the Function or pH Response of a UPF0118 Family Na+(Li+)/H+ Antiporter. Front Microbiol. 2020;11: 841. doi:10.3389/fmicb.2020.00841
91. Shao L, Xu T, Zheng X, Shao D, Zhang H, Chen H, et al. A novel three-TMH Na+/H+ antiporter and the functional role of its oligomerization. J Mol Biol. 2021;433: 166730. doi:10.1016/j.jmb.2020.166730
92. Bateman A, Finn RD. SCOOP: a simple method for identification of novel protein superfamily relationships. Bioinformatics. 2007;23: 809–814. doi:10.1093/bioinformatics/btm034
93. Mesdaghi S, Murphy DL, Sánchez Rodríguez F, Burgos-Mármol JJ, Rigden DJ. In silico prediction of structure and function for a large family of transmembrane proteins that includes human Tmem41b. F1000Res. 2020;9: 1395. doi:10.12688/f1000research.27676.2
94. Morita K, Hama Y, Izume T, Tamura N, Ueno T, Yamashita Y, et al. Genome-wide CRISPR screen identifies TMEM41B as a gene required for autophagosome formation. J Cell Biol. 2018;217: 3817–3828. doi:10.1083/jcb.201804132
95. Moretti F, Bergman P, Dodgson S, Marcellin D, Claerr I, Goodwin JM, et al. TMEM41B is a novel regulator of autophagy and lipid mobilization. EMBO Rep. 2018;19. doi:10.15252/embr.201845889

96. Van Alstyne M, Lotti F, Dal Mas A, Area-Gomez E, Pellizzoni L. Stasimon/Tmem41b localizes to mitochondria-associated ER membranes and is essential for mouse embryonic development. Biochem Biophys Res Commun. 2018;506: 463–470. doi:10.1016/j.bbrc.2018.10.073
97. Morita K, Hama Y, Mizushima N. TMEM41B functions with VMP1 in autophagosome formation. Autophagy. 2019;15: 922–923. doi:10.1080/15548627.2019.1582952
98. Shoemaker CJ, Huang TQ, Weir NR, Polyakov NJ, Schultz SW, Denic V. CRISPR screening using an expanded toolkit of autophagy reporters identifies TMEM41B as a novel autophagy factor. PLoS Biol. 2019;17: e2007044. doi:10.1371/journal.pbio.2007044
99. Schneider WM, Luna JM, Hoffmann H-H, Sánchez-Rivera FJ, Leal AA, Ashbrook AW, et al. Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. bioRxiv. 2020; 2020.10.07.326462. doi:10.1101/2020.10.07.326462
100. Hoffmann H-H, Schneider WM, Rozen-Gagnon K, Miles LA, Schuster F, Razooky B, et al. TMEM41B Is a Pan-flavivirus Host Factor. Cell. 2021;184: 133-148.e20. doi:10.1016/j.cell.2020.12.005
101. Farwell SLN, Kanyi D, Hamel M, Slee JB, Miller EA, Cipolle MD, et al. Heparin Decreases in Tumor Necrosis Factor α (TNFα)-induced Endothelial Stress Responses Require Transmembrane Protein 184A and Induction of Dual Specificity Phosphatase 1. J Biol Chem. 2016;291: 5342–5354. doi:10.1074/jbc.M115.681288
102. Dawson PA, Hubbert M, Haywood J, Craddock AL, Zerangue N, Christian WV, et al. The heteromeric organic solute transporter alpha-beta, Ostalpha-Ostbeta, is an ileal basolateral bile acid transporter. J Biol Chem. 2005;280: 6960–6968. doi:10.1074/jbc.M412752200
103. Seward DJ, Koh AS, Boyer JL, Ballatori N. Functional complementation between a novel mammalian polygenic transport complex and an evolutionarily ancient organic solute transporter, OSTalpha-OSTbeta. J Biol Chem. 2003;278: 27473–27482. doi:10.1074/jbc.M301106200
104. Wang W, Seward DJ, Li L, Boyer JL, Ballatori N. Expression cloning of two genes that together mediate organic solute and steroid transport in the liver of a marine vertebrate. Proc Natl Acad Sci USA. 2001;98: 9431–9436. doi:10.1073/pnas.161099898
105. Rasmussen RN, Christensen KV, Holm R, Nielsen CU. Nfat5 is involved in the hyperosmotic regulation of Tmem184b: a putative modulator of ibuprofen transport in renal MDCK I cells. FEBS Open Bio. 2019;9: 1071–1081. doi:10.1002/2211-5463.12630
106. Zhu H, Shang D, Sun M, Choi S, Liu Q, Hao J, et al. X-linked congenital hypertrichosis syndrome is associated with interchromosomal insertions mediated by a human-specific palindrome near SOX3. Am J Hum Genet. 2011;88: 819–826. doi:10.1016/j.ajhg.2011.05.004
107. Quamme GA. Molecular identification of ancient and modern mammalian magnesium transporters. Am J Physiol, Cell Physiol. 2010;298: C407-429. doi:10.1152/ajpcell.00124.2009
108. Stuiver M, Lainez S, Will C, Terryn S, Günzel D, Debaix H, et al. CNNM2, encoding a basolateral protein required for renal Mg2+ handling, is mutated in dominant hypomagnesemia. Am J Hum Genet. 2011;88: 333–343. doi:10.1016/j.ajhg.2011.02.005
109. Yamazaki D, Funato Y, Miura J, Sato S, Toyosawa S, Furutani K, et al. Basolateral Mg2+ extrusion via CNNM4 mediates transcellular Mg2+ transport across epithelia: a mouse model. PLoS Genet. 2013;9: e1003983. doi:10.1371/journal.pgen.1003983
110. Funato Y, Furutani K, Kurachi Y, Miki H. CrossTalk proposal: CNNM proteins are Na+ /Mg2+ exchangers playing a central role in transepithelial Mg2+ (re)absorption. J Physiol. 2018;596: 743–746. doi:10.1113/JP275248
111. Arjona FJ, de Baaij JHF. CrossTalk opposing view: CNNM proteins are not Na+ /Mg2+ exchangers but Mg2+ transport regulators playing a central role in transepithelial Mg2+ (re)absorption. J Physiol (Lond). 2018;596: 747–750. doi:10.1113/JP275249
112. Funato Y, Furutani K, Kurachi Y, Miki H. Rebuttal from Yosuke Funato, Kazuharu Furutani, Yoshihisa Kurachi and Hiroaki Miki. J Physiol. 2018;596: 751. doi:10.1113/JP275706
113. Arjona FJ, de Baaij JHF. Rebuttal from Francisco J. Arjona and Jeroen H. F. de Baaij. J Physiol. 2018;596: 753–754. doi:10.1113/JP275705

114. Sponder G, Mastrototaro L, Kurth K, Merolle L, Zhang Z, Abdulhanan N, et al. Human CNNM2 is not a Mg(2+) transporter per se. Pflugers Arch. 2016;468: 1223–1240. doi:10.1007/s00424-016-1816-7
115. Huang Y, Jin F, Funato Y, Xu Z, Zhu W, Wang J, et al. Structural basis for the Mg2+ recognition and regulation of the CorC Mg2+ transporter. Sci Adv. 2021;7: eabe6140. doi:10.1126/sciadv.abe6140
116. Hogue DL, Ellison MJ, Young JD, Cass CE. Identification of a novel membrane transporter associated with intracellular membranes by phenotypic complementation in the yeast Saccharomyces cerevisiae. J Biol Chem. 1996;271: 9801–9808. doi:10.1074/jbc.271.16.9801
117. Cabrita MA, Hobman TC, Hogue DL, King KM, Cass CE. Mouse transporter protein, a membrane protein that regulates cellular multidrug resistance, is localized to lysosomes. Cancer Res. 1999;59: 4890–4897.
118. Hogue DL, Kerby L, Ling V. A mammalian lysosomal membrane protein confers multidrug resistance upon expression in Saccharomyces cerevisiae. J Biol Chem. 1999;274: 12877–12882. doi:10.1074/jbc.274.18.12877
119. Tian S, Muneeruddin K, Choi MY, Tao L, Bhuiyan RH, Ohmi Y, et al. Genome-wide CRISPR screens for Shiga toxins and ricin reveal Golgi proteins critical for glycosylation. PLoS Biol. 2018;16: e2006951. doi:10.1371/journal.pbio.2006951
120. Yamaji T, Sekizuka T, Tachida Y, Sakuma C, Morimoto K, Kuroda M, et al. A CRISPR Screen Identifies LAPTM4A and TM9SF Proteins as Glycolipid-Regulating Factors. iScience. 2019;11: 409–424. doi:10.1016/j.isci.2018.12.039
121. Shao G-Z, Zhou R-L, Zhang Q-Y, Zhang Y, Liu J-J, Rui J-A, et al. Molecular cloning and characterization of LAPTM4B, a novel gene upregulated in hepatocellular carcinoma. Oncogene. 2003;22: 5060–5069. doi:10.1038/sj.onc.1206832
122. Adra CN, Zhu S, Ko JL, Guillemot JC, Cuervo AM, Kobayashi H, et al. LAPTM5: a novel lysosomal-associated multispanning membrane protein preferentially expressed in hematopoietic cells. Genomics. 1996;35: 328–337. doi:10.1006/geno.1996.0364
123. Grabner A, Brast S, Sucic S, Bierer S, Hirsch B, Pavenstädt H, et al. LAPTM4A interacts with hOCT2 and regulates its endocytotic recruitment. Cell Mol Life Sci. 2011;68: 4079–4090. doi:10.1007/s00018-011-0694-6
124. Milkereit R, Persaud A, Vanoaica L, Guetg A, Verrey F, Rotin D. LAPTM4b recruits the LAT1-4F2hc Leu transporter to lysosomes and promotes mTORC1 activation. Nat Commun. 2015;6: 7250. doi:10.1038/ncomms8250
125. Li L, Wei XH, Pan YP, Li HC, Yang H, He QH, et al. LAPTM4B: a novel cancer-associated gene motivates multidrug resistance through efflux and activating PI3K/AKT signaling. Oncogene. 2010;29: 5785–5795. doi:10.1038/onc.2010.303
126. Rutsch F, Gailus S, Miousse IR, Suormala T, Sagné C, Toliat MR, et al. Identification of a putative lysosomal cobalamin exporter altered in the cblF defect of vitamin B12 metabolism. Nat Genet. 2009;41: 234–239. doi:10.1038/ng.294
127. Kawaguchi K, Okamoto T, Morita M, Imanaka T. Translocation of the ABC transporter ABCD4 from the endoplasmic reticulum to lysosomes requires the escort protein LMBD1. Sci Rep. 2016;6: 30183. doi:10.1038/srep30183
128. Kitai K, Kawaguchi K, Tomohiro T, Morita M, So T, Imanaka T. The lysosomal protein ABCD4 can transport vitamin B12 across liposomal membranes in vitro. J Biol Chem. 2021;296: 100654. doi:10.1016/j.jbc.2021.100654
129. Wojnar P, Lechner M, Merschak P, Redl B. Molecular cloning of a novel lipocalin-1 interacting human cell membrane receptor using phage display. J Biol Chem. 2001;276: 20206–20212. doi:10.1074/jbc.M101762200
130. Hesselink RW, Findlay JBC. Expression, characterization and ligand specificity of lipocalin-1 interacting membrane receptor (LIMR). Mol Membr Biol. 2013;30: 327–337. doi:10.3109/09687688.2013.823018

131. Paek J, Kalocsay M, Staus DP, Wingler L, Pascolutti R, Paulo JA, et al. Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling. Cell. 2017;169: 338-349.e11. doi:10.1016/j.cell.2017.03.028
132. Lettice LA, Heaney SJH, Purdie LA, Li L, de Beer P, Oostra BA, et al. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet. 2003;12: 1725–1735. doi:10.1093/hmg/ddg180
133. Ianakiev P, van Baren MJ null, Daly MJ, Toledo SP, Cavalcanti MG, Neto JC, et al. Acheiropodia is caused by a genomic deletion in C7orf2, the human orthologue of the Lmbr1 gene. Am J Hum Genet. 2001;68: 38–45. doi:10.1086/316955
134. MacGrogan D, Levy A, Bova GS, Isaacs WB, Bookstein R. Structure and methylation-associated silencing of a gene within a homozygously deleted region of human chromosome band 8p22. Genomics. 1996;35: 55–65. doi:10.1006/geno.1996.0322
135. Knauer R, Lehle L. The oligosaccharyltransferase complex from Saccharomyces cerevisiae. Isolation of the OST6 gene, its synthetic interaction with OST3, and analysis of the native complex. J Biol Chem. 1999;274: 17249–17256. doi:10.1074/jbc.274.24.17249
136. Knauer R, Lehle L. The oligosaccharyltransferase complex from yeast. Biochim Biophys Acta. 1999;1426: 259–273. doi:10.1016/s0304-4165(98)00128-7
137. Kelleher DJ, Karaoglu D, Mandon EC, Gilmore R. Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. Mol Cell. 2003;12: 101–111. doi:10.1016/s1097-2765(03)00243-0
138. Cherepanova NA, Shrimal S, Gilmore R. Oxidoreductase activity is necessary for N-glycosylation of cysteine-proximal acceptor sites in glycoproteins. J Cell Biol. 2014;206: 525–539. doi:10.1083/jcb.201404083
139. Ramírez AS, Kowal J, Locher KP. Cryo-electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B. Science. 2019;366: 1372–1375. doi:10.1126/science.aaz3505
140. Goytain A, Quamme GA. Identification and characterization of a novel mammalian Mg2+ transporter with channel-like properties. BMC Genomics. 2005;6: 48. doi:10.1186/1471-2164-6-48
141. Zhou H, Clapham DE. Mammalian MagT1 and TUSC3 are required for cellular magnesium uptake and vertebrate embryonic development. Proc Natl Acad Sci USA. 2009;106: 15750–15755. doi:10.1073/pnas.0908332106
142. Li N, Gügel IL, Giavalisco P, Zeisler V, Schreiber L, Soll J, et al. FAX1, a novel membrane protein mediating plastid fatty acid export. PLoS Biol. 2015;13: e1002053. doi:10.1371/journal.pbio.1002053
143. Nilsson R, Schultz IJ, Pierce EL, Soltis KA, Naranuntarat A, Ward DM, et al. Discovery of genes essential for heme biosynthesis through large-scale gene expression analysis. Cell Metab. 2009;10: 119–130. doi:10.1016/j.cmet.2009.06.012
144. Yien YY, Robledo RF, Schultz IJ, Takahashi-Makise N, Gwynn B, Bauer DE, et al. TMEM14C is required for erythroid mitochondrial heme metabolism. J Clin Invest. 2014;124: 4294–4304. doi:10.1172/JCI76979
145. Yien YY, Ringel AR, Paw BH. Mitochondrial transport of protoporphyrinogen IX in erythroid cells. Oncotarget. 2015;6: 20742–20743. doi:10.18632/oncotarget.5124
146. Klammt C, Maslennikov I, Bayrhuber M, Eichmann C, Vajpai N, Chiu EJC, et al. Facile backbone structure determination of human membrane proteins by NMR spectroscopy. Nat Methods. 2012;9: 834–839. doi:10.1038/nmeth.2033
147. Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, et al. Detecting sequence signals in targeting peptides using deep learning. Life Sci Alliance. 2019;2. doi:10.26508/lsa.201900429
148. Shen D-W, Ma J, Okabe M, Zhang G, Xia D, Gottesman MM. Elevated expression of TMEM205, a hypothetical membrane protein, is associated with cisplatin resistance. J Cell Physiol. 2010;225: 822–828. doi:10.1002/jcp.22287

149. Gallenito MJ, Qasim TS, Tutol JN, Prakash V, Dodani SC, Meloni G. A recombinant platform to characterize the role of transmembrane protein hTMEM205 in Pt(II)-drug resistance and extrusion. Metallomics. 2020;12: 1542–1554. doi:10.1039/d0mt00114g
150. Tusnády GE, Simon I. The HMMTOP transmembrane topology prediction server. Bioinformatics. 2001;17: 849–850. doi:10.1093/bioinformatics/17.9.849
151. Chang J-M, Di Tommaso P, Taly J-F, Notredame C. Accurate multiple sequence alignment of transmembrane proteins with PSI-Coffee. BMC Bioinformatics. 2012;13 Suppl 4: S1. doi:10.1186/1471-2105-13-S4-S1
152. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10: 421. doi:10.1186/1471-2105-10-421
153. Steinegger M, Meier M, Mirdita M, Vöhringer H, Haunsberger SJ, Söding J. HH-suite3 for fast remote homology detection and deep protein annotation. BMC Bioinformatics. 2019;20: 473. doi:10.1186/s12859-019-3019-7
154. Remmert M, Biegert A, Hauser A, Söding J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods. 2011;9: 173–175. doi:10.1038/nmeth.1818
155. Mirdita M, von den Driesch L, Galiez C, Martin MJ, Söding J, Steinegger M. Uniclust databases of clustered and deeply annotated protein sequences and alignments. Nucleic Acids Res. 2017;45: D170–D176. doi:10.1093/nar/gkw1081
156. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7: 539. doi:10.1038/msb.2011.75
157. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. Protein Sci. 2018;27: 135–145. doi:10.1002/pro.3290
158. Lefort V, Longueville J-E, Gascuel O. SMS: Smart Model Selection in PhyML. Mol Biol Evol. 2017;34: 2422–2424. doi:10.1093/molbev/msx149
159. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59: 307–321. doi:10.1093/sysbio/syq010
160. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol. 2006;55: 539–552. doi:10.1080/10635150600755453
161. Chen K, Durand D, Farach-Colton M. NOTUNG: a program for dating gene duplications and optimizing gene family trees. J Comput Biol. 2000;7: 429–447. doi:10.1089/106652700750050871

Figures and Tables

Table 1

Families and subfamilies examined from the TCDB for SLC-like proteins. Superfamilies marked by arrows have been analyzed. The superfamily #1.A was examined since several existing SLC families are classified here, while superfamily #2.A was expected to contain most known SLC proteins. Superfamilies #9.A and #9.B were also examined. Numbers show the total number of level 3 families and level 4 subfamilies in each superfamily.

		Level 3 families (TC# x.y.z)			Level 4 subfamilies (TC #x.y.z.w)		
		Total	**Total examined**	**Found to be "SLC-like"**	**Total**	**Total examined**	**Found to be "SLC-like"**
1 – Channels/Pores		556			997		
→	1.A – α-Type Channels		112	7		321	17
2 – Electrochemical Potential-driven Transporters		178			561		
→	2.A – Porters (uniporters, symporters, antiporters)		129	114		556	501
3 – Primary Active Transporters		47			250		
4 – Group Translocators		18			36		
5 – Transmembrane Electron Carriers		17			42		
8 – Accessory Factors Involved in Transport		163			252		
9 – Incompletely Characterized Transport Systems		388			657		
→	9.A – Recognized Transporters of Unknown Biochemical Mechanism		69	20		122	37
→	9.B – Putative Transport Proteins		306	25		535	45
	Total:	**1367**			**2795**		
			616	**167**		**1534**	**602**

Table 2

Results of the initial search for "SLC-like" proteins. The search was performed using HMM-based sequence similarity analysis based on selected families and subfamilies from the TCDB, as well as selected Pfam models that likely encode transmembrane domains of "SLC-like" transporters (see text for details). The table shows the number of families, subfamilies from the TCDB and the number of Pfam models that had representative sequences in each organism. In addition, the initial number of found "SLC-like" proteins is shown.

Found:	**Level 3 families (TC #x.y.z)**	**Level 4 subfamilies (TC #x.y.z.w)**	**Pfam models**	**Proteins**
Human (*H. sapiens*)	67	310	82	552
Mouse (*M. musculus*)	68	316	82	535
Rat (*R. norvegicus*)	68	318	81	540
Zebrafish (D. *rerio*)	66	314	82	676
Chicken (G. *gallus*)	64	307	79	475
Fruitfly (*D. melanogaster*)	61	276	68	435
Roundworm (*C. elegans*)	60	285	71	537

Table 3

Novel SLC-like proteins. The table shows SLC-like human proteins from our search that are not in the "classical" group of SLCs (SLC1-52 families). The most similar (highest-scoring) TCDB family/subfamily and Pfam family are shown, as well as the most similar (highest-scoring) protein with structural information based on the pdb70 dataset. Substrate information, where available, was retrieved from available literature. Based on our search, SLC families SLC53-66 have recently been incorporated into the nomenclature in collaboration with the HUGO/HGNC.

Family name	UniProt ID	Gene symbol	TCDB family	Pfam family	Protein name	Closest PDB	Structural fold	Substrates	Remarks
SLC18	Q8NHS3	MFSD8	2.A.1.2	MFS_1	Major facilitator superfamily domain-containing protein 8	7mjsX	MFS	orphan	lysosomal [PMID:17564970], related to CLN7 (ceroid lipofuscinosis type 7) [PMID:17564970,PMID:19177532], might be related to mTOR signaling [PMID:29514215]
SLC18	Q8NBP5	MFSD9	2.A.1.2	MFS_1	Major facilitator superfamily domain-containing protein 9	4zw9A	MFS	orphan	
SLC18	Q14728	MFSD10	2.A.1.2	MFS_1	Major facilitator superfamily domain-containing protein 10	6s4mA	MFS	indomethacin [PubChem:3715,PMID:18638446]	
SLC18	Q96MC6	MFSD14A	2.A.1.2	MFS_1	Hippocampus abundant transcript 1 protein	7mjsX	MFS	orphan	affected by nutrient availability [PMID:28179877], loss causes globozoospermia in mice [PMID:27107036]
SLC18	Q5SR56	MFSD14B	2.A.1.2	MFS_1	Hippocampus abundant transcript-like protein 1	7mjsX	MFS	orphan	affected by nutrient availability [PMID:28179877]
SLC18	Q5VZR4	MFSD14C	2.A.1.2	MFS_1	Hippocampus abundant transcript-like protein 2	6kklA	MFS	orphan	
SLC18	Q96BI1	SLC22A18	2.A.1.2	MFS_1	Solute carrier family 22 member 18	6e9cA	MFS	orphan	also has an antisense gene (SLC22A18AS)
SLC22	Q7L0J3	SV2A (SLC22B1)	2.A.1.22	Sugar_tr	Synaptic vesicle glycoprotein 2A	4j05A	MFS	galactose [PubChem:6036,PMID:25326386]; also binds levetiracetam [PubChem:5284583,PMID:15210974], seletracetam [PubChem:9942725,PMID:18183537], brivaracetam [PubChem:9837243,PMID:14736235,PMID:17785672].	
SLC22	Q7L1I2	SV2B (SLC22B2)	2.A.1.22	MFS_1	Synaptic vesicle glycoprotein 2B	4j05A	MFS	orphan	
SLC22	Q496J9	SV2C (SLC22B3)	2.A.1.22	MFS_1	Synaptic vesicle glycoprotein 2C	4j05A	MFS	orphan	neuronal receptor of botulinum toxin A [PMID:31783045]
SLC22	Q8N4V2	SVOP (SLC22B4)	2.A.1.82	Sugar_tr	Synaptic vesicle 2-related protein	4zw9A	MFS	nicotinate [PubChem:937,PMID:21953179]; binds nucleotides in the TM region: 8-azido-ATP [PubChem:2735431,PMID:19390693]; ATP [PubChem:5957,PMID:19390693]; GTP [PubChem:135398633,PMID:19390693]; TTP [PubChem:64968,PMID:19390693]; CTP	

								[PubChem:6176,PMID:19390693]; NAD [PubChem:5892,PMID:19390693]	
SLC22	Q8N434	SVOPL (SLC22B5)	2.A.1.82	MFS_1	Putative transporter SVOPL	4j05A	MFS	orphan	
SLC33	Q96ES6	MFSD3	2.A.1.57	Acatn	Major facilitator superfamily domain-containing protein 3	7dl9A	MFS	orphan	
SLC35	Q24JQ0	TMEM241	2.A.7.13	TPT	Transmembrane protein 241	5ogeD	NST	orphan	
SLC51	Q6ZMB5	TMEM184A	2.A.82	Solute_trans_a	Transmembrane protein 184A			orphan	receptor for heparin [PMID:26769966]
SLC51	Q9Y519	TMEM184B	2.A.82	Solute_trans_a	Transmembrane protein 184B			orphan	might modulate the transport of ibuprofen [PMID:31066233]
SLC51	Q9NVA4	TMEM184C	2.A.82	Solute_trans_a	Transmembrane protein 184C			orphan	
SLC53	Q9UBH6	XPR1 (SLC53A1)	2.A.94	EXS	Xenotropic and polytropic retrovirus receptor 1			phosphate [PubChem:1061,PMID:23791524]	
SLC54	Q9Y5U8	MPC1 (SLC54A1)	2.A.105	MPC	Mitochondrial pyruvate carrier 1			pyruvate [PubChem:107735,PMID:22628554,PMID:22628558]	
SLC54	O95563	MPC2 (SLC54A2)	2.A.105	MPC	Mitochondrial pyruvate carrier 2			pyruvate [PubChem:107735,PMID:22628558]	
SLC54	P0DKB6	MPC1L (SLC54A3)	2.A.105	MPC	Mitochondrial pyruvate carrier 1-like protein			pyruvate [PubChem:107735,PMID:27317664]	
SLC55	O95202	LETM1 (SLC55A1)	2.A.97	LETM1	Mitochondrial proton/calcium exchanger protein			K+/H+ [PubChem:813,PubChem:1038,PMID:15138253,PMID:17925330,PMID:17541427,PMID:20197279]	Ca2+/H+ transport function [PMID:19797662,PMID:27669901] is debated [PMID:24616706].
SLC55	Q2VYF4	LETM2 (SLC55A2)	2.A.97	LETM1	LETM1 domain-containing protein LETM2, mitochondrial			orphan	
SLC55	Q6P1Q0	LETMD1 (SLC55A3)	2.A.97	LETM1	LETM1 domain-containing protein 1			orphan	
SLC56	Q9H9B4	SFXN1 (SLC56A1)	2.A.54	Mtc	Sideroflexin-1			L-serine [PubChem:5951,PMID:30442778]	
SLC56	Q96NB2	SFXN2 (SLC56A2)	2.A.54	Mtc	Sideroflexin-2			orphan	
SLC56	Q9BWM7	SFXN3 (SLC56A3)	2.A.54	Mtc	Sideroflexin-3			orphan	expressed in the mitochondria of neurons [PMID:31177362]
SLC56	Q6P4A7	SFXN4 (SLC56A4)	2.A.54	Mtc	Sideroflexin-4			orphan	knockout affects Fe-S biogenesis [PMID:31873120], and mutations can be causative of macrocytic anemia [PMID:24119684]
SLC56	Q8TD22	SFXN5 (SLC56A5)	2.A.54	Mtc	Sideroflexin-5			orphan	
SLC57	Q7RTP0	NIPA1 (SLC57A1)	2.A.7.25	Mg_trans_NIPA	Magnesium transporter NIPA1	5i20D	NST	Mg2+ [PubChem:888,PMID:17166836,PMID:18667602]; to a lesser extent Sr2+ [PubChem:104798,PMID:17166836,PMID:18667602],Fe2+ [PubChem:27284,PMID:17166836,PMID:1866	

								7602],Co2+ [PubChem:104729,PMID:17166836,PMID:18667602]	
SLC57	Q8N8Q9	NIPA2 (SLC57A2)	2.A.7.25	Mg_trans_NIPA	Magnesium transporter NIPA2	6ukjA	NST	very specific to Mg2+ [PubChem:888,PMID:18667602]	
SLC57	Q6NVV3	NIPAL1 (SLC57A3)	2.A.7.25	Mg_trans_NIPA	Magnesium transporter NIPA3	6ukjA	NST	Mg2+ [PubChem:888,PMID:18667602]; Sr2+ [PubChem:104798,PMID:18667602]; Ba2+ [PubChem:104810,PMID:18667602]; Fe2+ [PubChem:27284,PMID:18667602]; Cu2+ [PubChem:27099,PMID:18667602]	
SLC57	Q9H841	NIPAL2 (SLC57A4)	2.A.7.25	Mg_trans_NIPA	NIPA-like protein 2	6ukjA	NST	Mg2+ [PubChem:888,PMID:18667602]; Sr2+ [PubChem:104798,PMID:18667602]; Ba2+ [PubChem:104810,PMID:18667602]	NIPA4 referenced in [PMID:18667602] points to GenBank entry NM_024759, which corresponds to NIPAL2/SLC57A4, in contrast to the nomenclature in UniProt.
SLC57	Q6P499	NIPAL3 (SLC57A5)	2.A.7.25	Mg_trans_NIPA	NIPA-like protein 3	6ukjA	NST	orphan	knockout mice have a pleiotropic phenotype [PMID:19738379]
SLC57	Q0D2K0	NIPAL4 (SLC57A6)	2.A.7.25	Mg_trans_NIPA	Magnesium transporter NIPA4	6ukjA	NST	orphan	associated with autosomal recessive congenital ichthyosis [PMID:20301593]
SLC58	Q9H0U3	MAGT1 (SLC58A1)	1.A.76.1	OST3_OST6	Magnesium transporter protein 1	6s7tH	MagT	Mg2+ [PubChem:888,PMID:15804357]	Might have a channel-like mechanism [PMID:19940067].
SLC58	Q13454	TUSC3 (SLC58A2)	1.A.76.1	OST3_OST6	Tumor suppressor candidate 3	6s7tH	MagT	Mg2+ [PubChem:888,PMID:18667602,PMID:19940067], Fe2+ [PubChem:27284,PMID:18667602,PMID:19940067], Cu2+ [PubChem:27099,PMID:18667602,PMID:19940067], Mn2+ [PubChem:27854,PMID:19940067]	
SLC59	Q8NA29	MFSD2A (SLC59A1)	2.A.2.3	MFS_2	Sodium-dependent lysophosphatidylcholine symporter 1	7mjsX	MFS	LPC DHA (docosahexaenoic acid) [PubChem:10415542,PMID:24828044]; LPC palmitate [PubChem:460602,PMID:24828044]; TopFluor LPE [CAS:2260795-55-7,PMID:24828044]	
SLC59	A6NFX1	MFSD2B (SLC59A2)	2.A.2.3	MFS_2	Major facilitator superfamily domain-containing protein 2B	7mjsX	MFS	sphingosine-1-phosphate (S1P) [PubChem:5283560,PMID:29563527,PMID:29045386,PMID:33785361]	
SLC59	Q6NUT3	MFSD12	2.A.2.7	MFS_2	Major facilitator superfamily domain-containing protein 12	7mjsX	MFS	cysteine [PubChem:5862,PMID:33208952]	plays a role in skin pigmentation [PMID:29025994]
SLC59	Q14CX5	MFSD13A	2.A.2.3	MFS_2	Transmembrane protein 180	7mjsX	MFS	orphan	
SLC60	Q8N468	MFSD4A (SLC60A1)	2.A.1.7	MFS_1	Major facilitator superfamily domain-containing protein 4A	4j05A	MFS	orphan	
SLC60	Q5TF39	MFSD4B (SLC60A2)	2.A.1.7	MFS_1	Sodium-dependent glucose transporter 1	7ckoA	MFS	alpha-methyl-d-glucopyranoside [PubChem:64947,PMID:12590146]; D-glucose [PubChem:5793,PMID:12590146]; urea [PubChem:1176,PMID:26423860]	

SLC61	Q6N075	MFSD5 (SLC61A1)	2.A.1.40	MFS_5	Molybdate-anion transporter	6vbgA	MFS	molybdate [PubChem:24621,PMID:21464289]	
SLC62	Q9HCJ1	ANKH (SLC62A1)	2.A.66.9	ANKH	Progressive ankylosis protein homolog	4z3pA	MATE	pyrophosphate [PubChem:644102,PMID:10894769]	
SLC63	Q9H2V7	SPNS1 (SLC63A1)	2.A.1.49	MFS_1	Protein spinster homolog 1	6v4dA	MFS	orphan	
SLC63	Q8IVW8	SPNS2 (SLC63A2)	2.A.1.49	MFS_1	Protein spinster homolog 2	7mjsX	MFS	sphingosine-1-phosphate (S1P) [PubChem:5283560,PMID:19074308,PMID:21084291]; phosphorylated Fingolimod (FTY720-P) [PubChem:9908268,PMID:21084291]; dihydrosphingosine-1-phosphate (DH-S1P) [PubChem:644260,PMID:21084291]; phyto-S1P [PubChem:10883829,PMID:21084291]; C17-S1P [PubChem:5283559,PMID:21084291]	
SLC63	Q6ZMD2	SPNS3 (SLC63A3)	2.A.1.49	MFS_1	Protein spinster homolog 3	7mjsX	MFS	orphan	
SLC64	Q9HC07	TMEM165 (SLC64A1)	2.A.106.2	UPF0016	Transmembrane protein 165			Ca2+/H+ [PubChem:271,PubChem:1038,PMID:23569283,PMID:27008884]; Mn2+ [PubChem:27854,PMID:27008884,PMID:28270545]	
SLC65	O15118	NPC1 (SLC65A1)	2.A.6.6	Patched	NPC intracellular cholesterol transporter 1	6w5rA	RND	cholesterol [PubChem:5997,PMID:17989073,PMID:17989072,PMID:27410046]	
SLC65	Q9UHC9	NPC1L1 (SLC65A2)	2.A.6.6	Patched	NPC1-like intracellular cholesterol transporter 1	6v3hA	RND	cholesterol [PubChem:5997,PMID:14976318]	
SLC65	Q13635	PTCH1	2.A.6.6	Patched	Protein patched homolog 1	6n7kD	RND	cholesterol [PubChem:5997,PMID:33199907]	might be a multi-drug transporter [PMID:30110910]
SLC65	Q9Y6C5	PTCH2	2.A.6.6	Patched	Protein patched homolog 2	6n7kD	RND	likely cholesterol [PubChem:5997,PMID:33199907] based on functional similarity to PTCH1	
SLC65	Q96NR3	PTCHD1	2.A.6.6	Patched	Patched domain-containing protein 1	7dzpA	RND	orphan	associated with autism spectrum disorder and intellectual disability [PMID:20844286], might be related to the transport of a metabolite in the kynurenine pathway [PMID:31515500]
SLC65	Q3KNS1	PTCHD3	2.A.6.6	Patched	Patched domain-containing protein 3	6w5rA	RND	orphan	
SLC65	Q6ZW05	PTCHD4	2.A.6.6	Patched	Patched domain-containing protein 4	6w5rA	RND	orphan	
SLC65	Q12770	SCAP	2.A.6.6	Sterol-sensing	Sterol regulatory element-binding protein cleavage-activating protein	7etwB	SCAP_full	cholesterol [PubChem:5997,PMID:33446483], 25-hydroxycholesterol [PubChem:65094,PMID:33446483]	
SLC66	Q6ZP29	PQLC2 (SLC66A1)	2.A.43.2	PQ-loop	Lysosomal amino acid transporter 1 homolog	5xpdA	SWEET	L-Arg [PubChem:6322,PMID:22822152]; L-Lys [PubChem:5962,PMID:22822152]; L-His [PubChem:6274,PMID:22822152]; L-canavanine [PubChem:439202,PMID:23169667]; L-ornithine [PubChem:6262,PMID:23169667];	

								cysteamine-cysteine mixed disulfide [PubChem:424083,PMID:23169667]	
SLC66	Q8N2U9	PQLC1 (SLC66A2)	2.A.43	PQ-loop	PQ-loop repeat-containing protein 1	5xpdA	SWEET	orphan	
SLC66	Q8N755	PQLC3 (SLC66A3)	2.A.43.3	PQ-loop	PQ-loop repeat-containing protein 3	5xpdA	SWEET	orphan	
SLC66	O60931	CTNS (SLC66A4)	2.A.43.1	PQ-loop	Cystinosin	5xpdA	SWEET	L-cystine [PubChem:67678,PMID:11689434]; L-selenocystine [PubChem:207306]; L-cystathionine [PubChem:439258]; H+ [PubChem:1038,PMID:11689434,PMID:22232659]	
SLC66	O75352	MPDU1 (SLC66A5)	2.A.43.3	PQ-loop	Mannose-P-dolichol utilization defect 1 protein	5xpdA	SWEET	orphan	linked to congenital disorder of glycosylation [PMID:29721919,PMID:11733564]
SLC66	A1A4F0	PQLC2L	2.A.43.2	PQ-loop	Putative uncharacterized protein PQLC2L	5xpdA	SWEET	orphan	
pSLC.ARV1	Q9H2C2	ARV1	9.A.19	Arv1	Protein ARV1			orphan	might play a role in membrane sterol homeostasis [PMID:23668914] or the flipping of a GPI intermediate [PMID:32449190], also linked to epileptic encephalopathy [PMID:32462292]
pSLC.Battenin	Q13286	CLN3	2.A.57.5	CLN3	Battenin	6ob7A	MFS	orphan	linked to Batten/CLN3 disease [PMID:34274435]
pSLC.CLCN	P35523	CLCN1	2.A.49.2	Voltage_CLC	Chloride channel protein 1	6coyA	CLC	channel for chloride [PubChem:312,PMID:9565403]; thiocyanate [PubChem:9322,PMID:9565403]; perchlorate [PubChem:123351,PMID:9565403]; bromide [PubChem:259,PMID:9565403]; nitrate [PubChem:943,PMID:9565403]; chlorate [PubChem:104770,PMID:9565403]; iodide [PubChem:30165,PMID:9565403]	linked to myotonia [PMID:29845874]
pSLC.CLCN	P51788	CLCN2	2.A.49.2	Voltage_CLC	Chloride channel protein 2	6coyA	CLC	channel for chloride [PubChem:312,PMID:1311421]; bromide [PubChem:259,PMID:1311421]; iodide [PubChem:30165,PMID:1311421]	associated with a variety of diseases [PMID:28534947]
pSLC.CLCN	P51790	CLCN3	2.A.49.2	Voltage_CLC	H(+)/Cl(-) exchange transporter 3	7cq5D	CLC	antiport of chloride [PubChem:312,PMID:29917234]; and H+ [PubChem:1038,PMID:29917234]	present in endosomal compartments and synaptic vesicles [PMID:11182090], also in synaptic-like microvesicles (SLMV) [PMID:15073168]
pSLC.CLCN	P51793	CLCN4	2.A.49.2	Voltage_CLC	H(+)/Cl(-) exchange transporter 4	7cq5D	CLC	antiport of chloride [PubChem:312,PMID:16034421,PMID:16034422]; and H+ [PubChem:1038,PMID:16034421,PMID:16034422]; also nitrate [PubChem:943,PMID:9873029]; bromide	

								[PubChem:259,PMID:9873029]; iodide [PubChem:30165,PMID:9873029]	
pSLC.CLCN	P51795	CLCN5	2.A.49.2	Voltage_CLC	H(+)/Cl(-) exchange transporter 5	7cq5D	CLC	antiport of chloride [PubChem:312,PMID:16034421,PMID:16034422]; and H+ [PubChem:1038,PMID:16034421,PMID:16034422]; also nitrate [PubChem:943,PMID:9873029]; bromide [PubChem:259,PMID:9873029]; iodide [PubChem:30165,PMID:9873029]	linked to Dent's disease [PMID:7874126,PMID:8575751,PMID:8559248]
pSLC.CLCN	P51797	CLCN6	2.A.49.3	Voltage_CLC	Chloride transport protein 6	7cq5D	CLC	antiport of chloride [PubChem:312,PMID:20466723]; and H+ [PubChem:1038,PMID:20466723]; also nitrate [PubChem:943,PMID:20466723]; iodide [PubChem:30165,PMID:20466723]	expressed in late endosomes of the neurons [PMID:16950870]
pSLC.CLCN	P51798	CLCN7	2.A.49.3	Voltage_CLC	H(+)/Cl(-) exchange transporter 7	7cq5D	CLC	antiport of chloride [PubChem:312,PMID:18449189]; and H+ [PubChem:1038,PMID:18449189]	expressed in lysosomes [PMID:18449189], associated with osteopetrosis [PMID:11207362], obligate partner of Ostm1 [PMID:16525474]
pSLC.CLCN	P51800	CLCNKA	2.A.49.2	Voltage_CLC	Chloride channel protein ClC-Ka	6qvuA	CLC	channel for chloride [PubChem:312,PMID:11734858]; bromide [PubChem:259,PMID:11734858]; nitrate [PubChem: 943,PMID:11734858]; iodide [PubChem:30165,PMID:11734858]	expressed in the kidney [PMID:8041726] and cochlea [PMID:11734858], interacts with the beta-subunit barttin [PMID:9326936]
pSLC.CLCN	P51801	CLCNKB	2.A.49.2	Voltage_CLC	Chloride channel protein ClC-Kb	6qvuA	CLC	channel for chloride [PubChem:312,PMID:11734858]; bromide [PubChem:259,PMID:11734858]; nitrate [PubChem: 943,PMID:11734858]; iodide [PubChem:30165,PMID:11734858]	expressed in the kidney [PMID:8041726] and cochlea [PMID:11734858], associated with Bartter's syndrome [PMID:9326936], interacts with the beta-subunit barttin [PMID:9326936]
pSLC.CNNM	Q9NRU3	CNNM1	1.A.112	DUF21	Metal transporter CNNM1	7m1tA	CorC_large	orphan	expressed in brain and testis [PMID:22399287]
pSLC.CNNM	Q9H8M5	CNNM2	1.A.112	DUF21	Metal transporter CNNM2	7m1tA	CorC_large	Mg2+ [PubChem:888,PMID:15899945]; Co2+ [PubChem:104729,PMID:15899945]; Mn2+ [PubChem:27854,PMID:15899945]; Sr2+ [PubChem:104798,PMID:15899945]; Ba2+ [PubChem:104810,PMID:15899945]; Cu2+ [PubChem:27099,PMID:15899945]; Fe2+ [PubChem:27284,PMID:15899945]; Na+ [PubChem:923,PMID:21397062]	expressed in kidney, lung, spleen, testis [PMID:22399287]; linked to dominant hypomagnesemia [PMID:21397062]; Mg2+ transport activity has been debated [PMID:27068403,PMID:29383729]
pSLC.CNNM	Q8NE01	CNNM3	1.A.112	DUF21	Metal transporter CNNM3	7m1tA	CorC_large	orphan	expressed in kidney, brain, lung, spleen, heart [PMID:22399287]

pSLC.CNNM	Q6P4Q7	CNNM4	1.A.112	DUF21	Metal transporter CNNM4	7m1tA	CorC_large	Mg2+? [PubChem:888,PMID:24339795]; maybe antiport of Na+ [PubChem:923,PMID:24339795]	expressed in the basolateral membrane of intestinal epithelia [PMID:24339795]; linked to Jalili syndrome [PMID:19200525,PMID:19200527]
pSLC.CitMHS	Q04671	OCA2	2.A.45.2	CitMHS	P protein	7jskA	AbgT	maybe chloride [PubChem:312,PMID:25513726]	related to melanosome function and albinism [PMID:7761348,PMID:7991586]
pSLC.Dispatched	Q96F81	DISP1	2.A.6.9	Patched	Protein dispatched homolog 1	6td6A	RND	hedgehog protein [PMID:10619433,PMID:12372301]	related to hedgehog signaling
pSLC.Dispatched	A7MBM2	DISP2	2.A.6.9	Sterol-sensing	Protein dispatched homolog 2	6xe6A	RND	orphan	
pSLC.Dispatched	Q9P2K9	DISP3	2.A.6	Patched	Protein dispatched homolog 3	6td6A	RND	cholesterol? [PubChem:5997,PMID:19179482]	expressed in neural cell types [PMID:28134287, might be related to thyroid hormone levels [PMID:19179482]
pSLC.GPR155	Q7Z3F1	GPR155	2.A.69.3	Mem_trans	Integral membrane protein GPR155	6lgvA	NhaA	orphan	homologous to the Drosophila anchor protein [PMID:29735293]
pSLC.LAPTM	B4E0C1		2.A.74.1	Mtp	cDNA FLJ61683, moderately similar to Lysosomal-associated multitransmembrane protein			orphan	
pSLC.LAPTM	Q15012	LAPTM4A	2.A.74.1	Mtp	Lysosomal-associated transmembrane protein 4A	6wvgA	Tspn	thymidine [PubChem:5789,PMID:8621662]; uridine [PubChem:6029,PMID:8621662]; multi-drug resistance [PMID:10212276]	localized to late endosomes/lysosomes [PMID:10519401]
pSLC.LAPTM	Q86VI4	LAPTM4B	2.A.74.1	Mtp	Lysosomal-associated transmembrane protein 4B			maybe ceramide [PubChem:5702612,PMID:26280656]; multi-drug resistance [PMID:20711237]	
pSLC.LAPTM	Q13571	LAPTM5	2.A.74.1	Mtp	Lysosomal-associated transmembrane protein 5	6wvgA	Tspn	orphan	localized to lysosomes [PMID:8661146]
pSLC.LMBR-A	Q9NUN5	LMBRD1	9.A.54.1	LMBR1	Probable lysosomal cobalamin transporter			orphan	related to B12 vitamin transport [PMID:27456980]
pSLC.LMBR-B	Q8WVP7	LMBR1		LMBR1	Limb region 1 protein homolog			orphan	associated with preaxial polydactyly [PMID:10945466,PMID:11606546]
pSLC.LMBR-B	Q6UX01	LMBR1L	9.A.54.1	LMBR1	Protein LMBR1L			orphan	lipocalin-1 receptor [PMID:23964685]
pSLC.LMBR-B	Q68DH5	LMBRD2	9.A.54.3	LMBR1	LMBR1 domain-containing protein 2			orphan	may play a role in beta-2-AR and AT1R receptor internalization [PMID:28388415]
pSLC.MFSD1	Q9H3U5	MFSD1	2.A.1.53	MFS_1	Major facilitator superfamily domain-containing protein 1	7mjsX	MFS	orphan	affected by altered nutrient intake [PMID:27981419], physically interacts with GLMP [PMID:31661432,PMID:32959924]
pSLC.MFSD6	Q6ZSS7	MFSD6	2.A.1.65	MFS_1_like	Major facilitator superfamily domain-containing protein 6	7bp3A	MFS	sugars?	its homolog from D. melanogaster is Sugar baby, which is related to life span

									[PMID:12620118], related to energy consumption in mice [PMID:31906755]
pSLC.MFSD6	Q8IWD5	MFSD6L	2.A.1.65	MFS_1_like	Major facilitator superfamily domain-containing protein 6-like	4m64A	MFS	orphan	
pSLC.OSTC	B4DH36		1.A.76.2	OST3_OST6	cDNA FLJ52625	6s7oH	MagT	orphan	
pSLC.OSTC	Q9NRP0	OSTC	1.A.76.2	OST3_OST6	Oligosaccharyltransferase complex subunit OSTC	6s7oH	MagT	orphan	related to congenital disorders of glycosylation [PMID:32267060]
pSLC.OSTC	Q8TBU1	OSTCL	1.A.76.2	OST3_OST6	Oligosaccharyltransferase complex subunit-like	6ftg3	OSTC_part	orphan	
pSLC.RFT1	Q96AA3	RFT1	2.A.66.3	Rft-1	Protein RFT1 homolog	6cc4A	MATE	orphan	might be a scramblase of lipid-linked origosaccharides [PMID:11807558], however, this has been disputed [PMID:33446867]
pSLC.SIDT	Q9NXL6	SIDT1	1.A.79.1	SID-1_RNA_chan	SID1 transmembrane family member 1			cholesterol [PubChem:5997,PMID:28785058]	contradicts with originally proposed function in RNA uptake [PMID:21474576]
pSLC.SIDT	Q8NBJ9	SIDT2	1.A.79.1	SID-1_RNA_chan	SID1 transmembrane family member 2			cholesterol [PubChem:5997,PMID:28785058]	contradicts with proposed function in RNA uptake [PMID:27046251]
pSLC.STAR	Q14849	STARD3	9.B.64	MENTAL	StAR-related lipid transfer protein 3			probably cholesterol [PubChem:5997,PMID:15976441,PMID:15718238,PMID:12070139]	localizes to late endosomes [PMID:11053434]
pSLC.STAR	O95772	STARD3NL	9.B.64	MENTAL	STARD3 N-terminal-like protein			binds cholesterol [PubChem:5997,PMID:15718238]	
pSLC.STRA	Q9BX79	STRA6	2.A.90.1	RBP_receptor	Receptor for retinol uptake STRA6	5sy1A	STRA6	retinol [PubChem:445354,PMID:22815070]	expressed in the basolateral membrane of the retinal pigment epithelium and various other tissues [PMID:9203140,PMID:17255476], linked to Matthew-Wood syndrome [PMID:19309693]
pSLC.TMCO3	Q6UWJ1	TMCO3	2.A.37.1	Na_H_Exchanger	Transmembrane and coiled-coil domain-containing protein 3	6z3yA	NhaA	orphan	putative Na+/H+ exchanger, based on similarity
pSLC.TMEM14	Q9Y6G1	TMEM14A	2.A.126.1	Tmemb_14	Transmembrane protein 14A	2lopA	TMEM14_alt	orphan	may be involved in the regulation of the mitochondrial membrane potential [PMID:21723035]
pSLC.TMEM14	Q9NUH8	TMEM14B	2.A.126.1	Tmemb_14	Transmembrane protein 14B	2losA	TSPO	orphan	
pSLC.TMEM14	Q9P0S9	TMEM14C	2.A.126.1	Tmemb_14	Transmembrane protein 14C	2losA	TSPO	protoporphyrinogen IX [PubChem:121893,PMID:25157825]	localizes to the inner mitochondrial membrane [PMID:25157825], there are further links to its involvement in heme synthesis [PMID:32968076]
pSLC.TMEM14	A8MWL7	TMEM14DP	2.A.126.1	Tmemb_14	Transmembrane protein 14DP	2losA	TSPO	orphan	might be a pseudogene
pSLC.TMEM41-64	Q96HV5	TMEM41A	9.B.27.1	SNARE_assoc	Transmembrane protein 41A			orphan	
pSLC.TMEM41-64	Q5BJD5	TMEM41B	9.B.27.1	SNARE_assoc	Transmembrane protein 41B			scramblase of cholesterol [PubChem:5997,PMID:34074213];	localized to the ER [PMID:30093494], ab initio

								phosphatidylserine [PubChem:9547096,PMID:34074213]	structure prediction shows transporter-like features [PMID:33520197,PMID:33771928], required factor for SARS-CoV-2 and flavivirus infection [PMID:33052332,PMID:33338421]
pSLC.TMEM41-64	Q6YI46	TMEM64	9.B.27.5	SNARE_assoc	Transmembrane protein 64			orphan	associates with SERCA and modulates its function [PMID:23395171]
pSLC.TMEM104	Q8NE00	TMEM104	2.A.18.10	Aa_trans	Transmembrane protein 104	7kgvB	APC	orphan	might be a housekeeping gene [PMID:28386932]
pSLC.TMEM144	Q7Z5S9	TMEM144	2.A.7.8	TMEM144	Transmembrane protein 144	5i20D	NST	orphan	linked to bovine milk cholesterol content [PMID:30185830]
pSLC.TMEM163	Q8TC26	TMEM163	2.A.4.8	Cation_efflux	Transmembrane protein 163	6xpeA	CDF	Zn2+ [PubChem:32051,PMID:27917477,PMID:31697912]	expressed in synaptic vesicles [PMID:17623043], linked to Parkinson's disease [PMID:25064009]
pSLC.TMEM205	Q6UW68	TMEM205	9.A.55.1	DUF4149	Transmembrane protein 205			orphan	might be involved in cisplatin resistance [PMID:20589834,PMID:32789331]
pSLC.TMEM234	Q8WY98	TMEM234	2.A.7.32	TMEM234	Transmembrane protein 234	6ukjA	NST	orphan	knockdown causes proteinurea in zebrafish [PMID:26377798]
pSLC.TMEM245	Q9H330	TMEM245	2.A.86	AI-2E_transport	Transmembrane protein 245			orphan	
pSLC.TSPO	P30536	TSPO	9.A.24.1	TspO_MBR	Translocator protein	2mgyA	TSPO	protoporphyrin IX [PubChem:4971,PMID:3031675,PMID:25635101], cholesterol [PubChem:5997,PMID:9925760]	might be involved in cholesterol import into mitochondria [PMID:9925760,PMID:17692008], however, diverse other function have been suggested [PMID:34191248]
pSLC.TSPO	Q5TGU0	TSPO2	9.A.24.1	TspO_MBR	Translocator protein 2	4ryoA	TSPO	5-aminolevulinic acid [PubChem:137,PMID:31989647]	expressed in erythrocytes [PMID:21795748]
pSLC.UNC93	O43934	MFSD11	2.A.1.58	UNC-93	UNC93-like protein MFSD11	6vbgA	MFS	orphan	expressed in neurons, might play a role in energy homeostasis [PMID:27272503]
pSLC.UNC93	Q86WB7	UNC93A	2.A.1.58	UNC-93	Protein unc-93 homolog A	7c77B	MFS	orphan	could play an ion transporter role in the kidney [PMID:33163493]
pSLC.UNC93	Q9H1C4	UNC93B1	2.A.1.58	UNC-93	Protein unc-93 homolog B1	7cynC	MFS	orphan	helps with the trafficking of toll-like receptors to the endolysosomes [PMID:18305481
pSLC.XK	P51811	XK	2.A.112.1	XK-related	Membrane transport protein XK	7p14A	XK	orphan	part of the Kell blood group complex [PMID:16431037], linked to McLeod syndrome [PMID:11761473]

pSLC.XK	Q5GH77	XKR3	2.A.112.1	XK-related	XK-related protein 3	7p14A	XK	orphan	expressed in the testes [PMID:16431037]
pSLC.XK	Q5GH76	XKR4	2.A.112.1	XK-related	XK-related protein 4	7p14A	XK	phosphatidylserine [PubChem:9547096,PMID:25231987]	probably a scramblase [PMID:25231987] expressed in the brain [PMID:25231987]
pSLC.XK	Q6UX68	XKR5	2.A.112.1	XK-related	XK-related protein 5	7p14A	XK	orphan	
pSLC.XK	Q5GH73	XKR6	2.A.112.1	XK-related	XK-related protein 6	7p14A	XK	orphan	associated with depression and neuroticism [PMID:32109668,PMID:29559929,PMID:28552732], epilepsy [PMID:28533195], lupus erythematosus [PMID:29967481,PMID:20947557], and serum cholesterol and lipid levels [PMID:32024373,PMID:20160193]
pSLC.XK	Q5GH72	XKR7	2.A.112.1	XK-related	XK-related protein 7	7p14A	XK	orphan	
pSLC.XK	Q9H6D3	XKR8	2.A.112.1	XK-related	XK-related protein 8	7p14A	XK	phosphatidylserine [PubChem:9547096,PMID:23845944]	functions as a scramblase [PMID:23845944]
pSLC.XK	Q5GH70	XKR9	2.A.112.1	XK-related	XK-related protein 9	7p14A	XK	phosphatidylserine [PubChem:9547096,PMID:25231987]	probably a scramblase [PMID:25231987] expressed in the intestine [PMID:25231987]
pSLC.XK	Q6PP77	XKRX	2.A.112.1	XK-related	XK-related protein 2	7p14A	XK	orphan	expressed mostly in placenta and adrenal gland [PMID:16431037]

Figure 1

Dendrogram of human SLC-like protein families based on HMM fingerprint-based clustering. Classical and newly incorporated SLC families are shown with colors. Since SLC proteins are polyphyletic, which manifests itself in mathematically orthogonal HMM fingerprints that cannot be further clustered, the dendrogram does not join into a single branch.

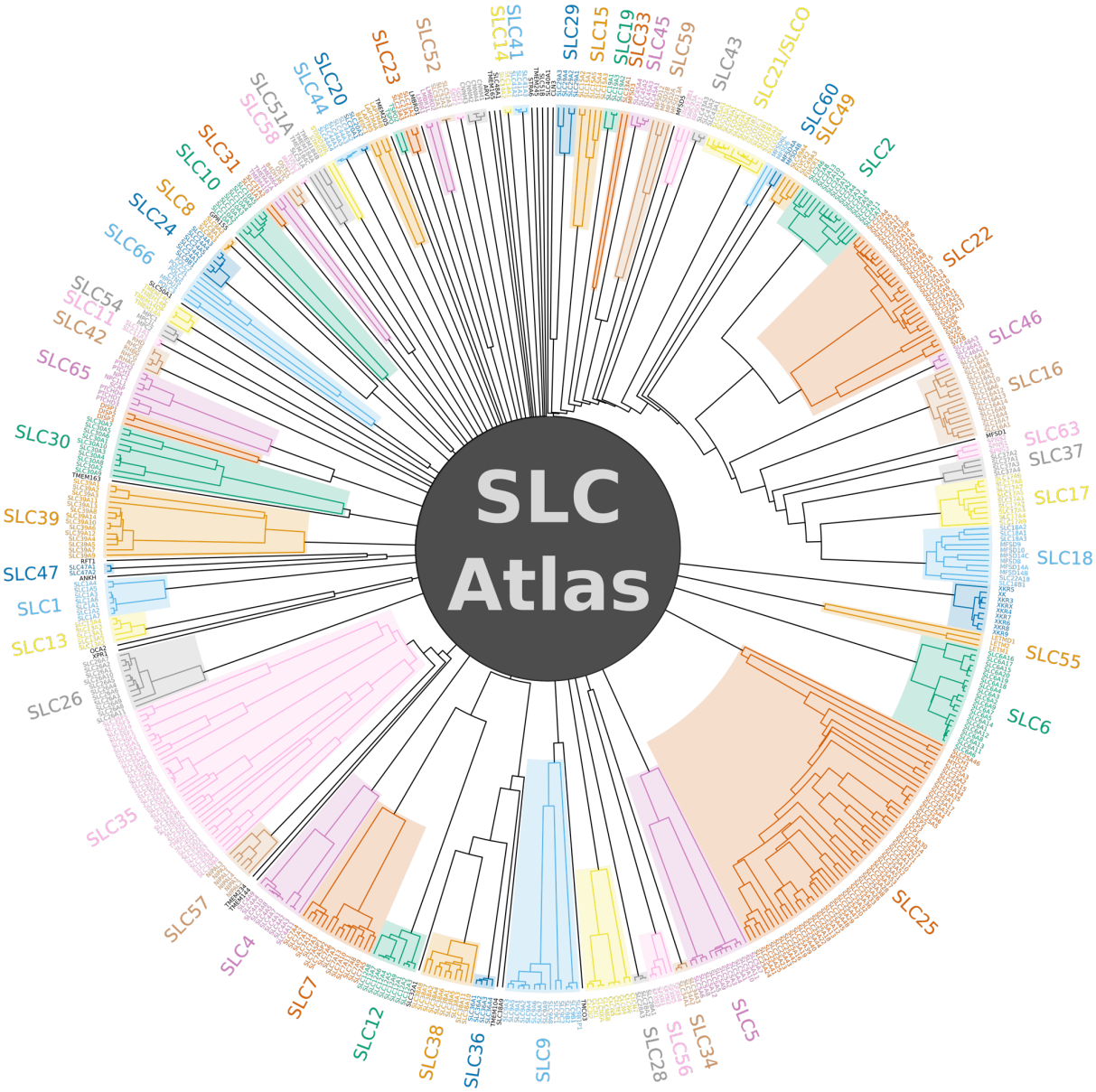


Figure 2

Orthology analysis of SLC-like proteins based on the human lineage. Each line shows a gene cluster that has evolved through gene duplication events along the human lineage. The human gene in the cluster, if present, is noted in the text label. Normal grey boxes in each column indicate a 1:1 orthology relationship between the human gene and a corresponding gene in the organism specified. Dark grey boxes indicate that the human gene has several orthologs, and light grey boxes indicate that several human genes share a common ortholog in that organism. White boxes denote no ortholog in the organism specified, while lines without a human gene name correspond to genes that have been lost in human.

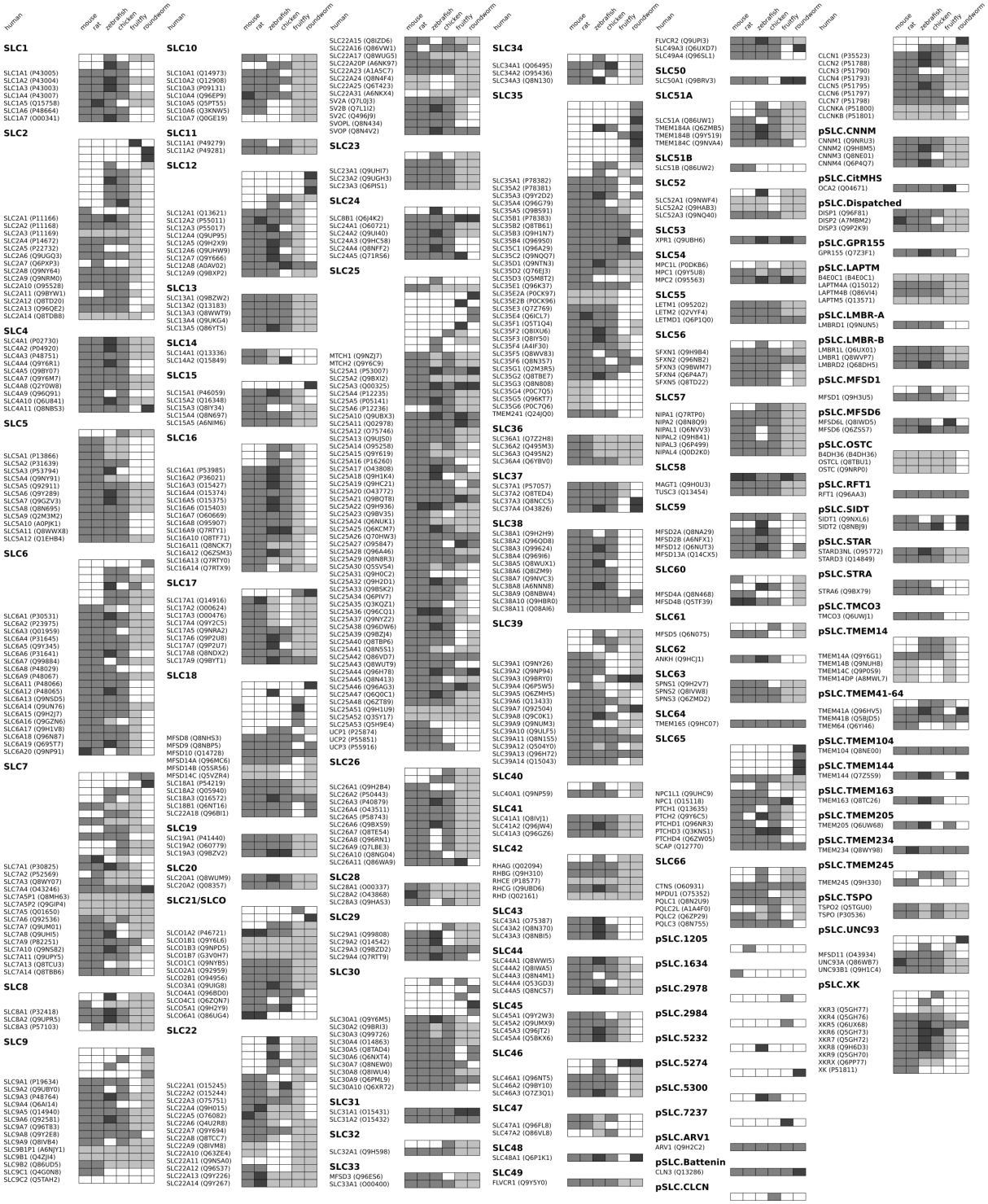


Figure 3

Phylogenetic tree of the SLC17, SLC18 and SLC37 families and proteins clustering in their neighborhood. Branch support values ≥ 0.7 are shown.

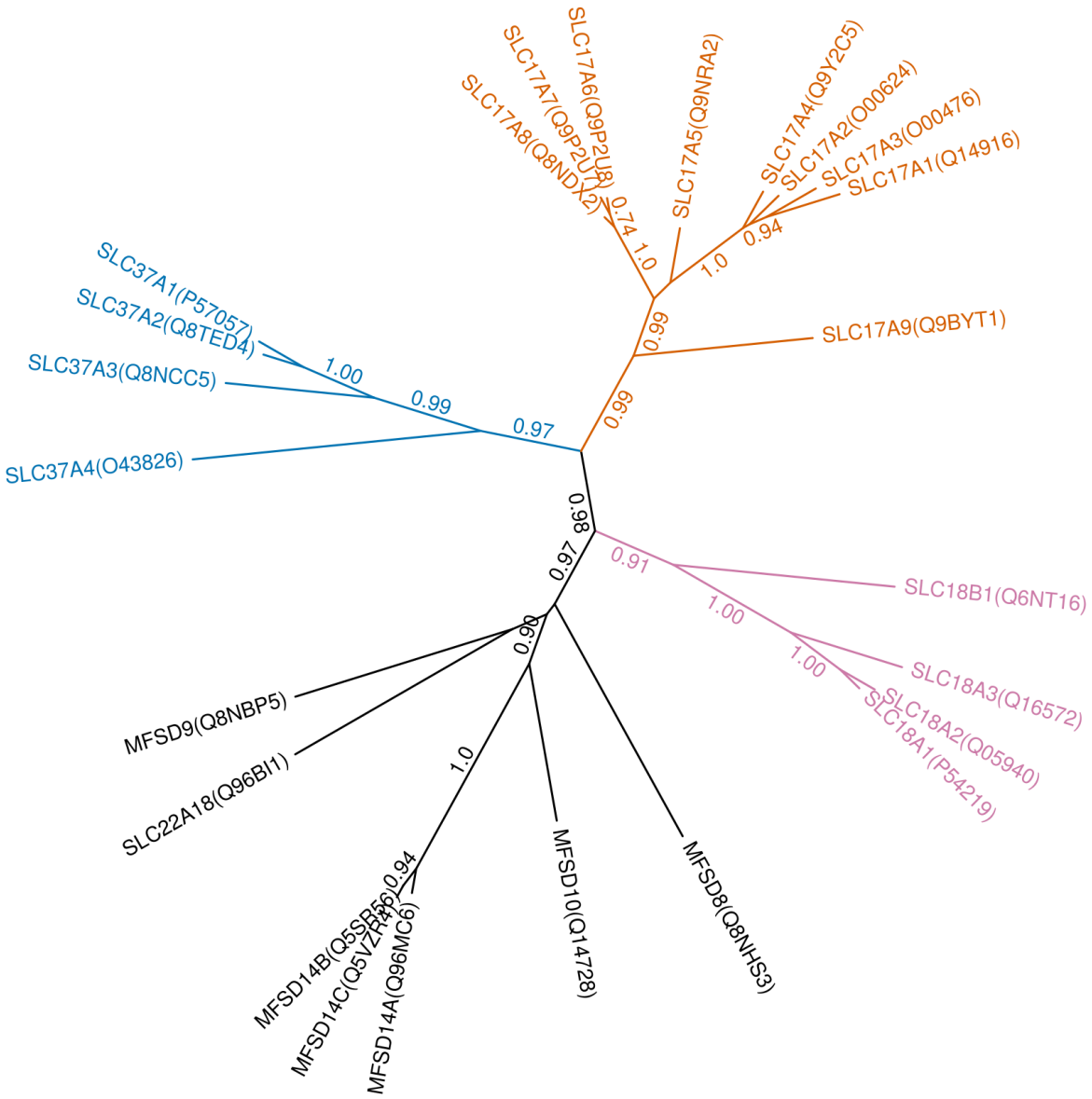


Figure 4

Phylogenetic tree of proteins clustering in the SLC38-36-32 families and TMEM104. Branch support values ≥ 0.5 are shown. SLC32 and SLC36 family members are colored gold and blue, respectively.

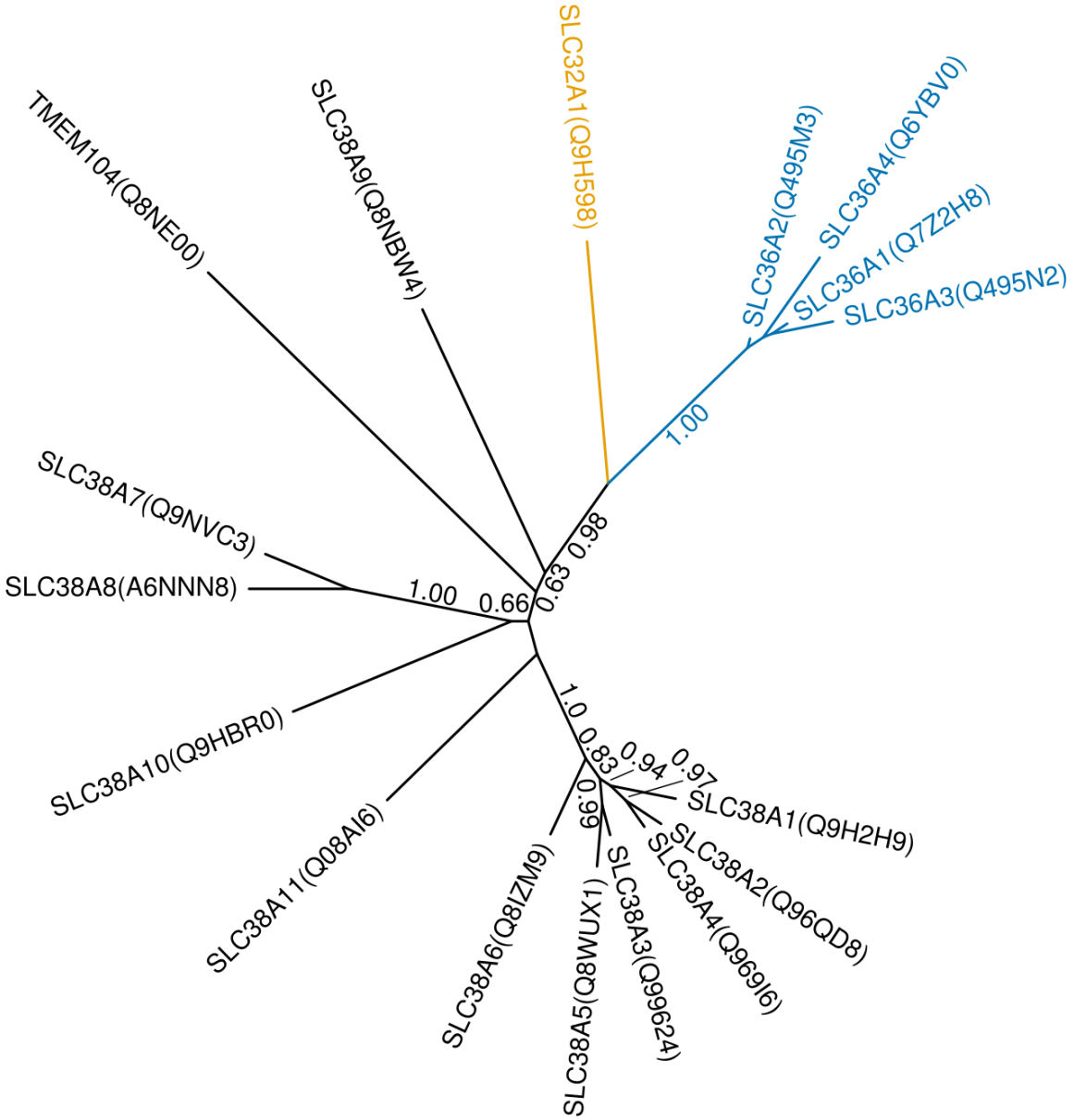
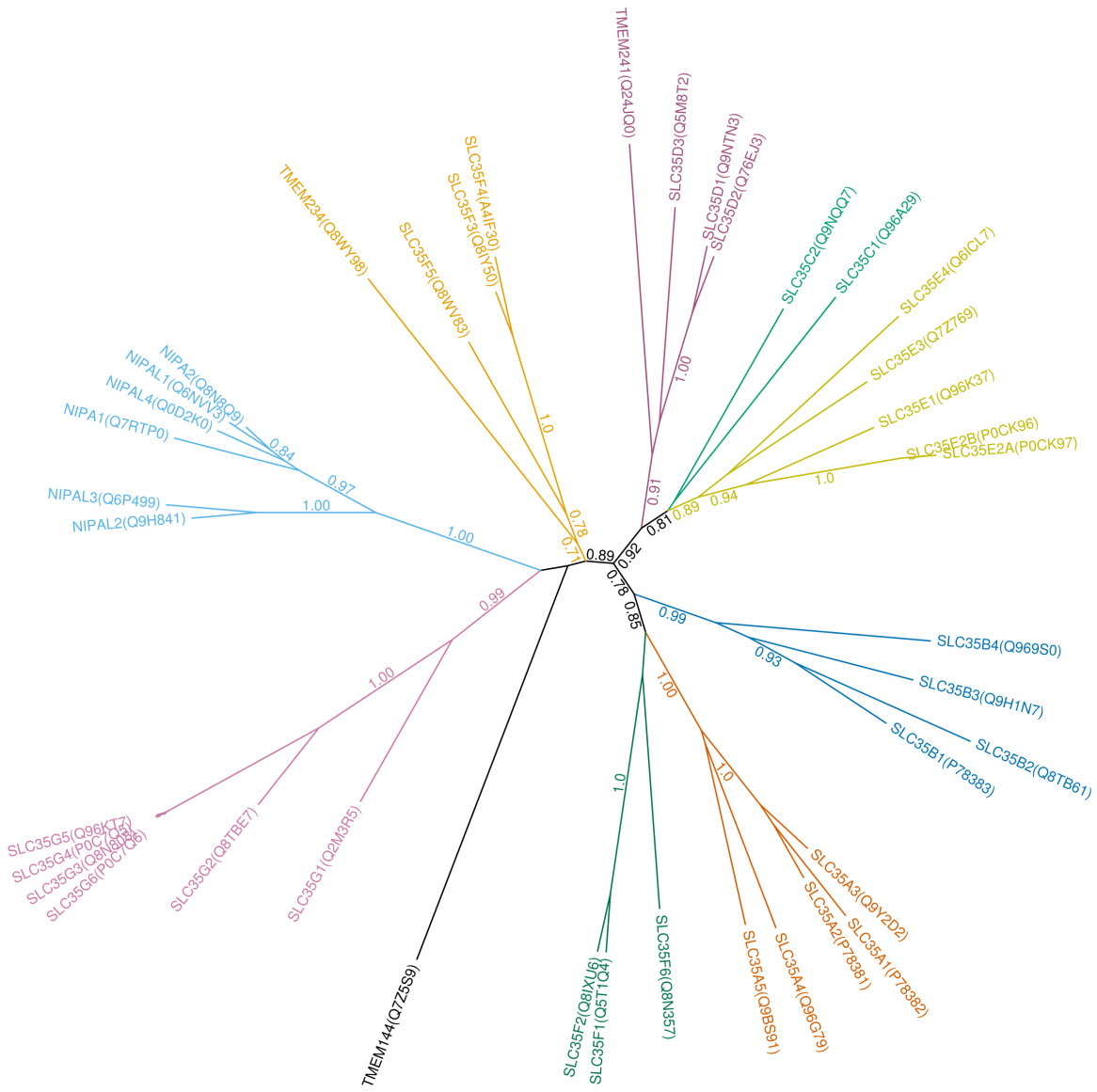


Figure 5

Phylogenetic tree of proteins with HMM fingerprints overlapping with the SLC35 family. Branch support values ≥ 0.7 are shown. The different subfamilies are colored in various colors.



Supplementary information

Supplementary Table 1

"SLC-like" Pfam domains and their clan memberships. The dataset is a result of our manual curation based on our "SLC-like" criteria (see text) and Pfam families present in protein sequences listed in the TCDB. Family names and data are from the Pfam database.

Family ID	Family name	Clan ID	Clan name	Family description
PF03390	2HCT			2-hydroxycarboxylate transporter family
PF00324	AA_permease	CL0062	APC	Amino acid permease
PF13520	AA_permease_2	CL0062	APC	Amino acid permease
PF01490	Aa_trans	CL0062	APC	Transmembrane amino acid transporter protein
PF06541	ABC_trans_CmpB			Putative ABC-transporter type IV
PF03806	ABG_transport	CL0182	IT	AbgT putative transporter family
PF13000	Acatn	CL0015	MFS	Acetyl-coenzyme A transporter 1
PF00873	ACR_tran	CL0322	RND_permease	AcrB/AcrD/AcrF family
PF01594	AI-2E_transport			AI-2E family transporter
PF06610	AlaE			L-alanine exporter
PF11744	ALMT	CL0307	FUSC	Aluminium activated malate transporter
PF00909	Ammonium_transp			Ammonium Transporter Family
PF07260	ANKH	CL0222	MviN_MATE	Progressive ankylosis protein (ANKH)
PF06081	ArAE_1	CL0307	FUSC	Aromatic acid exporter family member 1
PF10334	ArAE_2	CL0307	FUSC	Aromatic acid exporter family member 2
PF02040	ArsB	CL0182	IT	Arsenical pump membrane protein
PF03773	ArsP_1			Predicted permease
PF11449	ArsP_2			Putative, 10TM heavy-metal exporter
PF04161	Arv1			Arv1-like family
PF06826	Asp-Al_Ex	CL0064	CPA_AT	Predicted Permease Membrane Region
PF11700	ATG22	CL0015	MFS	Vacuole effluxer Atg22 like
PF03591	AzlC			AzlC protein
PF05437	AzlD			Branched-chain amino acid transport protein (AzlD)
PF02028	BCCT	CL0062	APC	BCCT, betaine/carnitine/choline family transporter
PF03594	BenE	CL0062	APC	Benzoate membrane transport protein
PF05525	Branch_AA_trans	CL0062	APC	Branched-chain amino acid transport protein
PF03092	BT1	CL0015	MFS	BT1 family
PF05232	BTP			Chlorhexidine efflux transporter
PF03596	Cad	CL0292	LysE	Cadmium resistance transporter
PF01545	Cation_efflux			Cation efflux family
PF09490	CbtA			Probable cobalt transporter subunit (CbtA)
PF03040	CemA			CemA family
PF04515	Choline_transpo			Plasma-membrane choline transporter
PF02417	Chromate_transp			Chromate transporter
PF03600	CitMHS	CL0182	IT	Citrate transporter
PF16980	CitMHS_2	CL0182	IT	Putative citrate transport
PF02487	CLN3			CLN3 protein
PF03601	Cons_hypoth698	CL0064	CPA_AT	Conserved hypothetical protein 698
PF05425	CopD	CL0430	CopD_like	Copper resistance protein D
PF08627	CRT-like	CL0184	DMT	CRT-like, chloroquine-resistance transporter-like
PF16965	CSG2	CL0184	DMT	Ceramide synthase regulator

PF02554	CstA	CL0062	APC	Carbon starvation protein CstA
PF13722	CstA_5TM			5TM C-terminal transporter carbon starvation CstA
PF04145	Ctr			Ctr copper transporter family
PF06808	DctM	CL0182	IT	Tripartite ATP-independent periplasmic transporter, DctM component
PF04290	DctQ			Tripartite ATP-independent periplasmic transporters, DctQ component
PF03605	DcuA_DcuB	CL0182	IT	Anaerobic c4-dicarboxylate membrane transporter
PF03606	DcuC	CL0182	IT	C4-dicarboxylate anaerobic carrier
PF04342	DMT_6	CL0184	DMT	Putative member of DMT superfamily (DUF486)
PF04657	DMT_YdcZ	CL0184	DMT	Putative inner membrane exporter, YdcZ
PF02683	DsbD	CL0292	LysE	Cytochrome C biogenesis protein transmembrane region
PF13386	DsbD_2	CL0292	LysE	Cytochrome C biogenesis protein transmembrane region
PF06237	DUF1011			Protein of unknown function (DUF1011)
PF07854	DUF1646	CL0182	IT	Protein of unknown function (DUF1646)
PF09843	DUF2070			Predicted membrane protein (DUF2070)
PF01595	DUF21			Cyclin M transmembrane N-terminal domain
PF09930	DUF2162			Predicted transporter (DUF2162)
PF10027	DUF2269	CL0430	CopD_like	Predicted integral membrane protein (DUF2269)
PF10101	DUF2339			Predicted membrane protein (DUF2339)
PF10821	DUF2567			Protein of unknown function (DUF2567)
PF11168	DUF2955	CL0307	FUSC	Protein of unknown function (DUF2955)
PF11377	DUF3180			Protein of unknown function (DUF3180)
PF11893	DUF3413			Domain of unknown function (DUF3413)
PF04165	DUF401	CL0182	IT	Protein of unknown function (DUF401)
PF13664	DUF4149	CL0430	CopD_like	Domain of unknown function (DUF4149)
PF15860	DUF4728	CL0347	Tetraspannin	Domain of unknown function (DUF4728)
PF04332	DUF475	CL0292	LysE	Protein of unknown function (DUF475)
PF15962	DUF4765			Domain of unknown function (DUF4765)
PF05684	DUF819	CL0064	CPA_AT	Protein of unknown function (DUF819)
PF06166	DUF979			Protein of unknown function (DUF979)
PF00892	EamA	CL0184	DMT	EamA-like transporter family
PF13536	EmrE	CL0184	DMT	Putative multidrug resistance efflux transporter
PF00810	ER_lumen_recept	CL0141	MtN3-like	ER lumen protein retaining receptor
PF04346	EutH			Ethanolamine utilisation protein, EutH
PF03124	EXS	CL0182	IT	EXS family
PF01770	Folate_carrier	CL0015	MFS	Reduced folate carrier
PF06963	FPN1	CL0015	MFS	Ferroportin1 (FPN1)
PF03239	FTR1	CL0292	LysE	Iron permease FTR1 family
PF04632	FUSC	CL0307	FUSC	Fusaric acid resistance protein family
PF12805	FUSC-like	CL0307	FUSC	FUSC-like inner membrane protein yccS
PF13515	FUSC_2	CL0307	FUSC	Fusaric acid resistance protein-like
PF07670	Gate			Nucleoside recognition
PF03323	GerA			Bacillus/Clostridium GerA spore germination protein
PF03616	Glt_symporter	CL0064	CPA_AT	Sodium/glutamate symporter
PF02447	GntP_permease	CL0182	IT	GntP family permease
PF04138	GtrA			GtrA-like protein
PF00955	HCO3_cotransp	CL0062	APC	HCO3- transporter family
PF04982	HPP	CL0307	FUSC	HPP family
PF16954	HRG			Haem-transporter, endosomal/lysosomal, haem-responsive gene
PF04955	HupE_UreJ	CL0292	LysE	HupE / UreJ protein
PF13795	HupE_UreJ_2	CL0292	LysE	HupE / UreJ protein

PF10688	Imp-YgjV			Bacterial inner membrane protein
PF04120	Iron_permease			Low affinity iron permease
PF03812	KdgT	CL0064	CPA_AT	2-keto-3-deoxygluconate permease
PF02705	K_trans	CL0062	APC	K+ potassium transporter
PF02652	Lactate_perm	CL0182	IT	L-lactate permease
PF01306	LacY_symp	CL0015	MFS	LacY proton/sugar symporter
PF07766	LETM1			LETM1-like protein
PF04791	LMBR1			LMBR1-like membrane protein
PF04172	LrgB			LrgB-like family
PF01810	LysE	CL0292	LysE	LysE type translocator
PF03956	Lys_export	CL0064	CPA_AT	Lysine exporter LysO
PF03817	MadL			Malonate transporter MadL subunit
PF03818	MadM			Malonate/sodium symporter MadM subunit
PF01914	MarC	CL0292	LysE	MarC family integral membrane protein
PF01554	MatE	CL0222	MviN_MATE	MatE
PF03547	Mem_trans	CL0064	CPA_AT	Membrane transport protein
PF10457	MENTAL			Cholesterol-capturing domain
PF07690	MFS_1	CL0015	MFS	Major Facilitator Superfamily
PF12832	MFS_1_like	CL0015	MFS	MFS_1 like family
PF13347	MFS_2	CL0015	MFS	MFS/sugar transport protein
PF05977	MFS_3	CL0015	MFS	Transmembrane secretion effector
PF06779	MFS_4	CL0015	MFS	Uncharacterised MFS-type transporter YbfB
PF05631	MFS_5	CL0015	MFS	Sugar-tranasporters, 12 TM
PF16983	MFS_MOT1	CL0062	APC	Molybdate transporter of MFS superfamily
PF07672	MFS_Mycoplasma	CL0015	MFS	Mycoplasma MFS transporter
PF01769	MgtE			Divalent cation transporter
PF05653	Mg_trans_NIPA	CL0184	DMT	Magnesium transporter NIPA
PF00153	Mito_carr			Mitochondrial carrier protein
PF03094	Mlo			Mlo family
PF10296	MMM1			Maintenance of mitochondrial morphology protein 1
PF03176	MMPL	CL0322	RND_permease	MMPL family
PF02659	Mntp	CL0292	LysE	Putative manganese efflux pump
PF03650	MPC	CL0141	MtN3-like	Mitochondrial pyruvate carriers
PF03820	Mtc			Tricarboxylate carrier
PF03083	MtN3_slv	CL0141	MtN3-like	Sugar efflux transporter for intercellular exchange
PF03821	Mtp	CL0347	Tetraspannin	Golgi 4-transmembrane spanning transporter
PF00893	Multi_Drug_Res	CL0184	DMT	Small Multidrug Resistance protein
PF03023	MurJ	CL0222	MviN_MATE	Lipid II flippase MurJ
PF01235	Na_Ala_symp	CL0062	APC	Sodium:alanine symporter family
PF01699	Na_Ca_ex			Sodium/calcium exchanger protein
PF03553	Na_H_antiporter	CL0182	IT	Na+/H+ antiporter family
PF06965	Na_H_antiport_1	CL0064	CPA_AT	Na+/H+ antiporter 1
PF13726	Na_H_antiport_2	CL0182	IT	Na+-H+ antiporter family
PF07399	Na_H_antiport_3	CL0182	IT	Putative Na+/H+ antiporter
PF00999	Na_H_Exchanger	CL0064	CPA_AT	Sodium/hydrogen exchanger family
PF02690	Na_Pi_cotrans			Na+/Pi-cotransporter
PF00939	Na_sulph_symp	CL0182	IT	Sodium:sulfate symporter transmembrane region
PF06450	NhaB	CL0182	IT	Bacterial Na+/H+ antiporter B (NhaB)
PF03824	NicO	CL0292	LysE	High-affinity nickel-transport protein
PF04973	NMN_transporter			Nicotinamide mononucleotide transporter

PF06813	Nodulin-like	CL0015	MFS	Nodulin-like
PF01566	Nramp	CL0062	APC	Natural resistance-associated macrophage protein
PF01733	Nucleoside_tran	CL0015	MFS	Nucleoside transporter
PF03825	Nuc_H_symport	CL0015	MFS	Nucleoside H+ symporter
PF04142	Nuc_sug_transp	CL0184	DMT	Nucleotide-sugar transporter
PF03137	OATP	CL0015	MFS	Organic Anion Transporter Polypeptide (OATP) family
PF16955	OFeT_1	CL0292	LysE	Ferrous iron uptake permease, iron-lead transporter
PF03169	OPT			OPT oligopeptide transporter protein
PF04756	OST3_OST6	CL0172	Thioredoxin	OST3 / OST6 family, transporter family
PF15048	OSTbeta			Organic solute transporter subunit beta protein
PF02460	Patched	CL0322	RND_permease	Patched family
PF16933	PelG			Putative exopolysaccharide Exporter (EPS-E)
PF01384	PHO4			Phosphate transporter family
PF04193	PQ-loop	CL0141	MtN3-like	PQ loop repeat
PF00854	PTR2	CL0015	MFS	POT family
PF02378	PTS_EIIC	CL0493	PTS_EIIC	Phosphotransferase system, EIIC
PF13303	PTS_EIIC_2	CL0493	PTS_EIIC	Phosphotransferase system, EIIC
PF03209	PUCC	CL0015	MFS	PUCC protein
PF16913	PUNUT	CL0184	DMT	Purine nucleobase transmembrane transport
PF14752	RBP_receptor			Retinol binding protein receptor
PF04506	Rft-1	CL0222	MviN_MATE	Rft protein
PF06379	RhaT	CL0184	DMT	L-rhamnose-proton symport protein (RhaT)
PF04479	RTA1			RTA1 like protein
PF01758	SBF	CL0064	CPA_AT	Sodium Bile acid symporter family
PF13593	SBF_like	CL0064	CPA_AT	SBF-like CPA transporter family (DUF4137)
PF05982	Sbt_1	CL0064	CPA_AT	Na+-dependent bicarbonate transporter superfamily
PF02667	SCFA_trans	CL0182	IT	Short chain fatty acid transporter
PF00375	SDF			Sodium:dicarboxylate symporter family
PF13630	SdpI			SdpI/YhfL protein family
PF02355	SecD_SecF	CL0322	RND_permease	Protein export membrane protein
PF11139	SfLAP	CL0292	LysE	Sap, sulfolipid-1-addressing protein
PF13965	SID-1_RNA_chan	CL0192	GPCR_A	dsRNA-gated channel SID-1
PF03842	Silic_transp			Silicon transporter
PF03522	SLC12			Solute carrier family 12
PF06027	SLC35F	CL0184	DMT	Solute carrier family 35
PF09335	SNARE_assoc			SNARE associated Golgi protein
PF00209	SNF	CL0062	APC	Sodium:neurotransmitter symporter family
PF03619	Solute_trans_a			Organic solute transporter Ostalpha
PF09546	Spore_III_AE			Stage III sporulation protein AE (spore_III_AE)
PF03845	Spore_permease	CL0062	APC	Spore germination protein
PF13782	SpoVAB			Stage V sporulation protein AB
PF00474	SSF	CL0062	APC	Sodium:solute symporter family
PF12349	Sterol-sensing	CL0322	RND_permease	Sterol-sensing domain of SREBP cleavage-activation
PF00083	Sugar_tr	CL0015	MFS	Sugar (and other) transporter
PF06800	Sugar_transport	CL0184	DMT	Sugar transport protein
PF00916	Sulfate_transp	CL0062	APC	Sulfate permease family
PF04143	Sulf_transp			Sulphur transport
PF01925	TauE	CL0292	LysE	Sulfite exporter TauE/SafE
PF01970	TctA			Tripartite tricarboxylate transporter TctA family
PF03741	TerC	CL0292	LysE	Integral membrane protein TerC family

PF06738	ThrE	CL0470	UMP_1	Putative threonine/serine exporter
PF12821	ThrE_2	CL0470	UMP_1	Threonine/Serine exporter, ThrE
PF03219	TLC	CL0015	MFS	TLC ATP/ADP transporter
PF07857	TMEM144	CL0184	DMT	Transmembrane family, TMEM144 of transporters
PF10639	TMEM234	CL0184	DMT	Putative transmembrane family 234
PF03647	Tmemb_14			Transmembrane proteins 14C
PF03151	TPT	CL0184	DMT	Triose-phosphate Transporter family
PF02133	Transp_cyt_pur	CL0062	APC	Permease for cytosine/purines, uracil, thiamine, allantoin
PF06609	TRI12	CL0015	MFS	Fungal trichothecene efflux pump (TRI12)
PF03222	Trp_Tyr_perm	CL0062	APC	Tryptophan/tyrosine permease family
PF03073	TspO_MBR			TspO/MBR family
PF08449	UAA	CL0184	DMT	UAA transporter family
PF05978	UNC-93	CL0015	MFS	Ion channel regulatory protein UNC-93
PF01169	UPF0016	CL0292	LysE	Uncharacterized protein family UPF0016
PF02694	UPF0060	CL0184	DMT	Uncharacterised BCR, YnfA/UPF0060 family
PF03653	UPF0093	CL0430	CopD_like	Uncharacterised protein family (UPF0093)
PF07168	Ureide_permease	CL0184	DMT	Ureide permease
PF03253	UT			Urea transporter
PF04892	VanZ			VanZ like family
PF03631	Virul_fac_BrkB			Virulence factor BrkB
PF01988	VIT1			VIT family
PF00654	Voltage_CLC			Voltage gated chloride channel
PF00860	Xan_ur_permease	CL0062	APC	Permease family
PF09815	XK-related			XK-related protein
PF03733	YccF			Inner membrane component domain
PF02325	YGGT			YGGT family
PF02588	YitT_membrane			Uncharacterised 5xTM membrane BCR, YitT family COG1284
PF02535	Zip	CL0184	DMT	ZIP Zinc transporter

Supplementary Table 2

False positive hits or incorrectly annotated human sequences found in our search. Gene symbol and protein name shown based on UniProt data. Most similar (highest scoring) TCDB families/subfamilies and Pfam families are shown.

UniProt ID	Gene symbol	TCDB family	TCDB score	Pfam family	Pfam score	Protein name	Remarks
B7Z8S2		1.A.11.4	176.1	Ammonium_transp	69.7	cDNA FLJ52981, moderately similar to Homo sapiens Rhesus blood group, B glycoprotein (RHBG), mRNA	Seems like a splice variant of RHBG_HUMAN, with the middle of the protein missing.
P18577	RHCE	1.A.11.4	506.9	Ammonium_transp	203	Blood group Rh(CE) polypeptide	Defines the Rh blood group, likely has a structural role in assembling the gas metabolon [PMID:18832935]
Q02161	RHD	1.A.11.4	509.3	Ammonium_transp	201.3	Blood group Rh(D) polypeptide	Defines the Rh blood group, likely has a structural role in assembling the gas metabolon [PMID:18832935]
Q8IXB1	DNAJC10	1.A.76	36.5	OST3_OST6	53.3	DnaJ homolog subfamily C member 10	Characterized enzyme. Part of this protein shows similarity to part of the OST3_OST6 Pfam domain, probably because the OST3_OST6 model also contains a (partial) thioredoxin domain, because the hits overlaps with the Thioredoxin domain hit annotated on the Pfam web site.
P13667	PDIA4	1.A.76.1	56.7	OST3_OST6	57.3	Protein disulfide-isomerase A4	Characterized enzyme. Similarly to protein Q8IXB1, this protein also gives a hit for the OST3_OST6 domain that overlaps with the hit for a Thioredoxin domain annotated as shown on the website for this sequence in Pfam.
Q14554	PDIA5	1.A.76.1	53.3	OST3_OST6	57.4	Protein disulfide-isomerase A5	Characterized enzyme. The same case as for proteins Q8IXB1 and P13667.
Q6ZUX1		1.A.112	72.5	DUF21	17.6	cDNA FLJ43252 fis, clone HEART2006909	Only similar proteins in UniProt are bacterial proteins with >50% identity.
B4DIU4		2.A.1.1	47.7	Sugar_tr	62	cDNA FLJ50540, highly similar to Solute carrier family 2, facilitated glucose transporter member 5	First 98 amino acids of 113 are identical to human SLC2A5, so probably a fragment or splice variant.
A0A024R4M7		2.A.3.3	90.5		0	HCG1658540, isoform CRA_a	The first 195 aa of this 205-aa sequence shows 53.3% identity to a segment of human SLC7A3. Contains 5 predicted TM helices according to UniProt, so likely is not a complete transporter sequence. BLAST search returns mammalian sequences with the highest identity of 78.7%.
Q9NWI3		2.A.3.8	194.7	AA_permease_2	39	cDNA FLJ20839 fis, clone ADKA02346	The C-terminal 185 aa of this 370-aa sequence is identical to the C-terminus of human SLC7A10.
Q75T53		2.A.5.4	359	Zip	226.4	Uncharacterized protein	BLAST search showed hits with 50-60% sequence identity from mites/ticks, and these were the most similar hits. Possibly not a human sequence.
P04035	HMGCR	2.A.6.6	530.9	Sterol-sensing	63.1	3-hydroxy-3-methylglutaryl-coenzyme A reductase	A well-characterized enzyme showing similarity to TC #2.A.6.6, Patched proteins, and also containing the Sterol-sensing domain. Also, the rate-limiting enzyme in cholesterol biosynthesis highly expressed in liver [25650254].
O43600		2.A.17.1	58.9	PTR2	28.8	Erythroid differentiation-related factor 2 (Fragment)	Shows highest similarity (96.2% identity) to inner segments of bacterial proteins, so likely not a human protein.
Q86YK3	STK4/SLC36AL	2.A.18.8	61.2		0	Kinase/transmembrane domain fusion protein (Fragment)	Proteins in this family without the Aa_trans Pfam domain are likely not transmembrane. Missing the canonical transporter domain of the family, and contains only 2 TMHs according to UniProt annotations.
Q8WXG9	ADGRV1	2.A.19.3	638.8		0	Adhesion G-protein coupled receptor V1	SLC8-similar proteins not containing the Pfam domain Na_Ca_ex. This protein is a false positive, which has been selected because it contain Calx-beta Pfam domains similarly to SLC8A1-3 proteins. However, it is a known GPCR.
Q86XX4	FRAS1	2.A.19.3	134.9		0	Extracellular matrix protein FRAS1	SLC8-similar proteins not containing the Pfam domain Na_Ca_ex. This protein is a false positive, which has been selected because it contain Calx-beta Pfam domains similarly to SLC8A1-3 proteins. However, it is a single-TMH protein with no known transporter function.
Q5SZK8	FREM2	2.A.19.3	145		0	FRAS1-related extracellular matrix protein 2	SLC8-similar proteins not containing the Pfam domain Na_Ca_ex. This protein is a false positive, which has been selected because it contain Calx-beta Pfam domains similarly to SLC8A1-3 proteins. However, it is a single-TMH protein with no known transporter function.

P0C091	FREM3	2.A.19.3	92.6		0	FRAS1-related extracellular matrix protein 3	SLC8-similar proteins not containing the Pfam domain Na_Ca_ex. This protein is a false positive, which has been selected because it contain Calx-beta Pfam domains similarly to SLC8A1-3 proteins. However, it is a single-TMH protein with no known transporter function.
B4E2E9		2.A.22.2	154.3	SNF	113.8	cDNA FLJ51458, highly similar to Sodium- and chloride-dependent neutral andbasic amino acid transporter B(0+)	This 149-aa sequence shows 92% identity to a segment of human SLC6A14, so is likely a fragment.
Q8N7R6		2.A.29.4	108.8		0	cDNA FLJ40434 fis, clone TESTI2039353, moderately similar to MITOCHONDRIAL PHOSPHATE CARRIER PROTEIN	Seems to contain only a single MCF repeat (UniProt). BLAST search returns vertebrate sequences, highest similarity was 81.9% identity with sequence F7I611.
B4DSH6		2.A.29.14	225.1	Mito_carr	118.4	cDNA FLJ53093, moderately similar to Mus musculus solute carrier family 25, member 23 (Slc25a23), mRNA	Seems to contain only two MCF repeats (UniProt). An alignment with human SLC25A44, SLC25A18 and mouse Slc25a23 and human SLC25A23 showed that this sequence is basically identical to the C-terminal half of Q96H78 (human SLC25A44), with a larger stretch of ~100 residues missing. So probably an alternatively spliced variant.
Q13717		2.A.31.1	689.5	HCO3_cotransp	438.5	Anion exchange protein 3 (Fragment)	Similar (92.8% identity) to the C-terminal part of SLC4A3, also does not start with a methionine.
B3KS22		2.A.36.1	279.9	Na_H_Exchanger	130	HCG2015407	96.6% identical to the gorilla protein A0A2I2YMF4, however, the gorilla protein is much larger, and this human sequence possibly contains a half transporter.
Q9BXG1		2.A.40.7	380.5	Xan_ur_permease	118.8	p41	BLAST search gives ~90% identical hits from Leuconostoc lactis (Guanine permease), so probably not a human protein.
Q59FC0		2.A.48	115.7	Folate_carrier	69.8	Solute carrier family 19, member 2 variant (Fragment)	The last 89 aa of this 125-aa sequence is identical to human SLC19A2, might be a fragment of a splice variant.
B4DE56		2.A.49.2	214	Voltage_CLC	75.5	cDNA FLJ59180, moderately similar to Chloride channel protein ClC-Ka	High similarity to the inner segment of human ClC-Ka (P51800, 76.8% identity, but C-terminal region almost identical), therefore probably a fragment.
B4E0R8		2.A.90	74.4	RBP_receptor	170.2	cDNA FLJ60210	Seems to match 3 TMHs of the region encoded by the RBP_receptor Pfam family. The full domain contains 9 TMHs, so this sequence is likely to be a fragment. BLAST searches returned no significantly similar hits in human.
B7ZMG3	KIAA1529	2.A.90.2	84.9	RBP_receptor	128.9	KIAA1529 protein	Seems to match a non-TM region of the RBP_receptor Pfam domain. Identical to human CCDC180 with an extra 139 aa at the N-terminus, which is a secreted protein according to UniProt annotations.
A4D146	LOC402642	2.A.100	61.5	FPN1	60.4	Solute carrier family 40 protein	Short protein (167 aa), might be a fragment. Contains a segment of 59 aa 69.5% identical to zebrafish SLC40A1-like sequence (A0A2R8Q2S9), and a segment 82 aa 53.7% identical to human SLC40A1. BLAST search returns SLC40A1-like sequences from mammals and vertebrates.
Q6J4J2		9.B.27.1	69	SNARE_assoc	26.5	Inflammatory related protein	Most similar proteins are bacteria (95.0% identity with N-terminal half of Q8XDY4), so likely not a human protein.
P49675	STAR	9.B.64	144.8		0	Steroidogenic acute regulatory protein, mitochondrial	Proteins in this family without the MENTAL Pfam domain are not transmembrane.
P59095	STARD6	9.B.64	76.1		0	StAR-related lipid transfer protein 6	Proteins in this family without the MENTAL Pfam domain are not transmembrane.
P24390	KDELR1	9.B.191.1	285.2	ER_lumen_recept	193.5	ER lumen protein-retaining receptor 1	Belongs to the family of KDEL receptor proteins, which are well-characterized receptors. However, KDEL receptors show sequence similarity to SWEET transporters.
P33947	KDELR2	9.B.191.1	278.3	ER_lumen_recept	188.8	ER lumen protein-retaining receptor 2	Belongs to the family of KDEL receptor proteins, which are well-characterized receptors. However, KDEL receptors show sequence similarity to SWEET transporters.
O43731	KDELR3	9.B.191.1	258.8	ER_lumen_recept	180.1	ER lumen protein-retaining receptor 3	Belongs to the family of KDEL receptor proteins, which are well-characterized receptors. However, KDEL receptors show sequence similarity to SWEET transporters.

Supplementary Table 3

Table of existing ("original"/classical) SLC transporters, showing correlation with TCDB families/subfamilies, Pfam families and PDB structures and structural folds. Gene symbol and protein name based on UniProt information are shown. Most similar (highest-scoring) TCDB families and subfamilies, as well as Pfam families are shown for each protein. The PDB structure from the pdb70 dataset that is most similar (highest-scoring) to each protein is shown, along with its fold family assignment by us, partly based on TCDB family names.

AbgT: p-Aminobenzoyl-glutamate Transporter family; Amt: Ammonium Transporter family; APC: Amino acid-Polyamine-Cation family; CDF: Cation Diffusion Facilitator family; CNT: Concentrative Nucleoside Transporter family; Ctr: Copper Transporter family; DAACS: Dicarboxylate/Amino Acid:Cation (Na+ or H+) Symporter family; MATE: Multidrug And Toxic compound Extrusion family; MCF: Mitochondrial Carrier Family; MFS: Major Facilitator Superfamily; MgtE: Mg2+ transporter-E family; NAT: Nucleobase/Ascorbate Transporter or Nucleobase:Cation Symporter-2 (NCS2) family; NCX: Sodium/Calcium exchanger family; NhaA: Sodium/proton antiporter family; NST: Nucleoside-Sugar Transporter family; PiT: Type III Sodium/phosphate cotransporter family; SWEET: Sugar Will Eventually be Transported family.; ZIP: Zrt/Irt-like Transporter family.

Family name	UniProt ID	Gene symbol	TCDB family	TCDB score	Pfam family	Pfam score	Protein name	Closest PDB	Structural fold
SLC1	P43005	SLC1A1	2.A.23.2	779.7	SDF	438.5	Excitatory amino acid transporter 3	6rvxB	DAACS
SLC1	P43004	SLC1A2	2.A.23.2	819.1	SDF	442	Excitatory amino acid transporter 2	6rvxB	DAACS
SLC1	P43003	SLC1A3	2.A.23.2	838.8	SDF	445.2	Excitatory amino acid transporter 1	6rvxB	DAACS
SLC1	P43007	SLC1A4	2.A.23.3	896	SDF	400.3	Neutral amino acid transporter A	6rvxB	DAACS
SLC1	Q15758	SLC1A5	2.A.23.3	935.5	SDF	399	Neutral amino acid transporter B(0)	6rvxB	DAACS
SLC1	P48664	SLC1A6	2.A.23.2	793.7	SDF	434.3	Excitatory amino acid transporter 4	6rvxB	DAACS
SLC1	O00341	SLC1A7	2.A.23.2	790.4	SDF	427.1	Excitatory amino acid transporter 5	6rvxB	DAACS
SLC2	P11166	SLC2A1	2.A.1.1	470.6	Sugar_tr	538.4	Solute carrier family 2, facilitated glucose transporter member 1	6thaA	MFS
SLC2	P11168	SLC2A2	2.A.1.1	496.4	Sugar_tr	534	Solute carrier family 2, facilitated glucose transporter member 2	5eqhA	MFS
SLC2	P11169	SLC2A3	2.A.1.1	469.2	Sugar_tr	522.6	Solute carrier family 2, facilitated glucose transporter member 3	4zw9A	MFS
SLC2	P14672	SLC2A4	2.A.1.1	456.4	Sugar_tr	508.3	Solute carrier family 2, facilitated glucose transporter member 4	4zw9A	MFS
SLC2	P22732	SLC2A5	2.A.1.1	426	Sugar_tr	465.8	Solute carrier family 2, facilitated glucose transporter member 5	4yb9D	MFS
SLC2	Q9UGQ3	SLC2A6	2.A.1.1	401.3	Sugar_tr	273.7	Solute carrier family 2, facilitated glucose transporter member 6	4j05A	MFS
SLC2	Q6PXP3	SLC2A7	2.A.1.1	414.9	Sugar_tr	401.8	Solute carrier family 2, facilitated glucose transporter member 7	4zw9A	MFS
SLC2	Q9NY64	SLC2A8	2.A.1.1	441.7	Sugar_tr	338.6	Solute carrier family 2, facilitated glucose transporter member 8	4zw9A	MFS
SLC2	Q9NRM0	SLC2A9	2.A.1.1	385.5	Sugar_tr	354.4	Solute carrier family 2, facilitated glucose transporter member 9	4zw9A	MFS
SLC2	O95528	SLC2A10	2.A.1.1	393.3	Sugar_tr	240.2	Solute carrier family 2, facilitated glucose transporter member 10	4zw9A	MFS
SLC2	Q9BYW1	SLC2A11	2.A.1.1	316.1	Sugar_tr	313.5	Solute carrier family 2, facilitated glucose transporter member 11	5eqhA	MFS
SLC2	Q8TD20	SLC2A12	2.A.1.1	456.8	Sugar_tr	332.1	Solute carrier family 2, facilitated glucose transporter member 12	4zw9A	MFS
SLC2	Q96QE2	SLC2A13	2.A.1.1	532.4	Sugar_tr	378.5	Proton myo-inositol cotransporter	4zw9A	MFS
SLC2	Q8TDB8	SLC2A14	2.A.1.1	458.9	Sugar_tr	513.9	Solute carrier family 2, facilitated glucose transporter member 14	4zw9A	MFS
SLC4	P02730	SLC4A1	2.A.31.1	1661.6	HCO3_cotransp	655.2	Band 3 anion transport protein	4yzfA	NAT
SLC4	P04920	SLC4A2	2.A.31.1	2333.9	HCO3_cotransp	693.2	Anion exchange protein 2	6caaB	NAT
SLC4	P48751	SLC4A3	2.A.31.1	2263.2	HCO3_cotransp	681.4	Anion exchange protein 3	6caaB	NAT
SLC4	Q9Y6R1	SLC4A4	2.A.31.2	1745.2	HCO3_cotransp	784.8	Electrogenic sodium bicarbonate cotransporter 1	6caaB	NAT
SLC4	Q9BY07	SLC4A5	2.A.31.2	1620.8	HCO3_cotransp	752.5	Electrogenic sodium bicarbonate cotransporter 4	6caaB	NAT
SLC4	Q9Y6M7	SLC4A7	2.A.31.2	1870.8	HCO3_cotransp	811.3	Sodium bicarbonate cotransporter 3	6caaB	NAT
SLC4	Q2Y0W8	SLC4A8	2.A.31.2	1807.3	HCO3_cotransp	820	Electroneutral sodium bicarbonate exchanger 1	6caaB	NAT
SLC4	Q96Q91	SLC4A9	2.A.31.2	1379.3	HCO3_cotransp	707.3	Anion exchange protein 4	6caaB	NAT
SLC4	Q6U841	SLC4A10	2.A.31.2	1834.1	HCO3_cotransp	822.1	Sodium-driven chloride bicarbonate exchanger	6caaB	NAT
SLC4	Q8NBS3	SLC4A11	2.A.31.4	2155.1	HCO3_cotransp	518	Sodium bicarbonate transporter-like protein 11	6caaB	NAT
SLC5	P13866	SLC5A1	2.A.21.3	884.6	SSF	586.7	Sodium/glucose cotransporter 1	2xq2A	APC
SLC5	P31639	SLC5A2	2.A.21.3	859.4	SSF	564.8	Sodium/glucose cotransporter 2	2xq2A	APC
SLC5	P53794	SLC5A3	2.A.21.3	867.7	SSF	538.8	Sodium/myo-inositol cotransporter	2xq2A	APC
SLC5	Q9NY91	SLC5A4	2.A.21.3	843.1	SSF	525.8	Solute carrier family 5 member 4	2xq2A	APC
SLC5	Q92911	SLC5A5	2.A.21.5	935.7	SSF	145.7	Sodium/iodide cotransporter	2xq2A	APC

SLC5	Q9Y289	SLC5A6	2.A.21.5	955.3	SSF	131.7	Sodium-dependent multivitamin transporter	2xq2A	APC
SLC5	Q9GZV3	SLC5A7	2.A.21.8	1189.2	SSF	120.6	High affinity choline transporter 1	2xq2A	APC
SLC5	Q8N695	SLC5A8	2.A.21.5	964	SSF	138.3	Sodium-coupled monocarboxylate transporter 1	2xq2A	APC
SLC5	Q2M3M2	SLC5A9	2.A.21.3	847.9	SSF	488.7	Sodium/glucose cotransporter 4	2xq2A	APC
SLC5	A0PJK1	SLC5A10	2.A.21.3	792	SSF	462.8	Sodium/glucose cotransporter 5	2xq2A	APC
SLC5	Q8WWX8	SLC5A11	2.A.21.3	821	SSF	471.9	Sodium/myo-inositol cotransporter 2	2xq2A	APC
SLC5	Q1EHB4	SLC5A12	2.A.21.5	988.3	SSF	153.3	Sodium-coupled monocarboxylate transporter 2	2xq2A	APC
SLC6	P30531	SLC6A1	2.A.22.3	937.7	SNF	775.7	Sodium- and chloride-dependent GABA transporter 1	6vrlA	APC
SLC6	P23975	SLC6A2	2.A.22.1	953.3	SNF	797	Sodium-dependent noradrenaline transporter	6vrlA	APC
SLC6	Q01959	SLC6A3	2.A.22.1	936.2	SNF	781.2	Sodium-dependent dopamine transporter	6vrlA	APC
SLC6	P31645	SLC6A4	2.A.22.1	949.9	SNF	755.7	Sodium-dependent serotonin transporter	6vrlA	APC
SLC6	Q9Y345	SLC6A5	2.A.22.2	1062.5	SNF	720.4	Sodium- and chloride-dependent glycine transporter 2	6vrlA	APC
SLC6	P31641	SLC6A6	2.A.22.3	989.4	SNF	798.9	Sodium- and chloride-dependent taurine transporter	6vrlA	APC
SLC6	Q99884	SLC6A7	2.A.22.2	915.5	SNF	762.7	Sodium-dependent proline transporter	6vrlA	APC
SLC6	P48029	SLC6A8	2.A.22.3	1022.1	SNF	802.9	Sodium- and chloride-dependent creatine transporter 1	6vrlA	APC
SLC6	P48067	SLC6A9	2.A.22.2	977.3	SNF	769.5	Sodium- and chloride-dependent glycine transporter 1	6vrlA	APC
SLC6	P48066	SLC6A11	2.A.22.3	1041.3	SNF	818	Sodium- and chloride-dependent GABA transporter 3	6vrlA	APC
SLC6	P48065	SLC6A12	2.A.22.3	1010.4	SNF	806.3	Sodium- and chloride-dependent betaine transporter	6vrlA	APC
SLC6	Q9NSD5	SLC6A13	2.A.22.3	1033.6	SNF	815.4	Sodium- and chloride-dependent GABA transporter 2	6vrlA	APC
SLC6	Q9UN76	SLC6A14	2.A.22.2	917.2	SNF	647	Sodium- and chloride-dependent neutral and basic amino acid transporter B(0+)	6vrlA	APC
SLC6	Q9H2J7	SLC6A15	2.A.22.6	948.3	SNF	739.4	Sodium-dependent neutral amino acid transporter B(0)AT2	6vrlA	APC
SLC6	Q9GZN6	SLC6A16	2.A.22.6	841.1	SNF	403.7	Orphan sodium- and chloride-dependent neurotransmitter transporter NTT5	6vrlA	APC
SLC6	Q9H1V8	SLC6A17	2.A.22.6	969.9	SNF	744.9	Sodium-dependent neutral amino acid transporter SLC6A17	6vrlA	APC
SLC6	Q96N87	SLC6A18	2.A.22.6	833.8	SNF	597.1	Inactive sodium-dependent neutral amino acid transporter B(0)AT3	6vrlA	APC
SLC6	Q695T7	SLC6A19	2.A.22.6	887	SNF	593.3	Sodium-dependent neutral amino acid transporter B(0)AT1	6m17C	APC
SLC6	Q9NP91	SLC6A20	2.A.22.6	840.1	SNF	583.1	Sodium- and chloride-dependent transporter XTRP3	6vrlA	APC
SLC7	P30825	SLC7A1	2.A.3.3	704.6	AA_permease_2	170.1	High affinity cationic amino acid transporter 1	7dsnB	APC
SLC7	P52569	SLC7A2	2.A.3.3	708.6	AA_permease_2	172.3	Cationic amino acid transporter 2	7dsnB	APC
SLC7	Q8WY07	SLC7A3	2.A.3.3	679.3	AA_permease_2	165	Cationic amino acid transporter 3	7dsnB	APC
SLC7	O43246	SLC7A4	2.A.3.3	608	AA_permease_2	164.3	Cationic amino acid transporter 4	7dsnB	APC
SLC7	Q01650	SLC7A5	2.A.3.8	653.4	AA_permease_2	213.9	Large neutral amino acids transporter small subunit 1	7dsnB	APC
SLC7	Q8MH63	SLC7A5P1	2.A.3.8	194.1	AA_permease_2	64	Putative L-type amino acid transporter 1-like protein MLAS	7dsnB	APC
SLC7	Q9GIP4	SLC7A5P2	2.A.3.8	179.6	AA_permease_2	56.3	Putative L-type amino acid transporter 1-like protein IMAA	7dsnB	APC
SLC7	Q92536	SLC7A6	2.A.3.8	657.6	AA_permease_2	220.4	Y+L amino acid transporter 2	7dsnB	APC
SLC7	Q9UM01	SLC7A7	2.A.3.8	642.3	AA_permease_2	203.6	Y+L amino acid transporter 1	7dsnB	APC
SLC7	Q9UHI5	SLC7A8	2.A.3.8	662.4	AA_permease_2	232.7	Large neutral amino acids transporter small subunit 2	7dsnB	APC
SLC7	P82251	SLC7A9	2.A.3.8	599	AA_permease_2	212	b(0,+)-type amino acid transporter 1	7dslB	APC
SLC7	Q9NS82	SLC7A10	2.A.3.8	636.8	AA_permease_2	205.4	Asc-type amino acid transporter 1	7dsnB	APC
SLC7	Q9UPY5	SLC7A11	2.A.3.8	607	AA_permease_2	197.1	Cystine/glutamate transporter	7dsnB	APC
SLC7	Q8TCU3	SLC7A13	2.A.3.8	456.3	AA_permease_2	135.9	Solute carrier family 7 member 13	7dsnB	APC
SLC7	Q8TBB6	SLC7A14	2.A.3.3	719.7	AA_permease_2	152.8	Probable cationic amino acid transporter	7dslB	APC
SLC8	P32418	SLC8A1	2.A.19.3	1869.3	Na_Ca_ex	210.9	Sodium/calcium exchanger 1	4k1cA	NCX
SLC8	Q9UPR5	SLC8A2	2.A.19.3	1730.6	Na_Ca_ex	192.2	Sodium/calcium exchanger 2	4kjrA	NCX
SLC8	P57103	SLC8A3	2.A.19.3	1812.3	Na_Ca_ex	199.4	Sodium/calcium exchanger 3	5hyaA	NCX
SLC9	P19634	SLC9A1	2.A.36.1	845.9	Na_H_Exchanger	296.7	Sodium/hydrogen exchanger 1	7dsxA	NhaA
SLC9	Q9UBY0	SLC9A2	2.A.36.1	869.6	Na_H_Exchanger	311.3	Sodium/hydrogen exchanger 2	7dsxA	NhaA
SLC9	P48764	SLC9A3	2.A.36.1	864.9	Na_H_Exchanger	287.6	Sodium/hydrogen exchanger 3	7dsxA	NhaA
SLC9	Q6AI14	SLC9A4	2.A.36.1	834.3	Na_H_Exchanger	299.6	Sodium/hydrogen exchanger 4	7dsxA	NhaA

SLC9	Q14940	SLC9A5	2.A.36.1	856.4	Na_H_Exchanger	291.9	Sodium/hydrogen exchanger 5	7dsxA	NhaA
SLC9	Q92581	SLC9A6	2.A.36.1	639.1	Na_H_Exchanger	319.5	Sodium/hydrogen exchanger 6	7dsxA	NhaA
SLC9	Q96T83	SLC9A7	2.A.36.1	660	Na_H_Exchanger	304	Sodium/hydrogen exchanger 7	7dsxA	NhaA
SLC9	Q9Y2E8	SLC9A8	2.A.36.1	577.5	Na_H_Exchanger	226.2	Sodium/hydrogen exchanger 8	7dsxA	NhaA
SLC9	Q8IVB4	SLC9A9	2.A.36.1	654.3	Na_H_Exchanger	291.1	Sodium/hydrogen exchanger 9	6z3yA	NhaA
SLC9	Q4ZJI4	SLC9B1	2.A.36.2	936.7	Na_H_Exchanger	103.8	Sodium/hydrogen exchanger 9B1	4d0aB	NhaA
SLC9	A6NJY1	SLC9B1P1	2.A.36.2	461.9	Na_H_Exchanger	61.9	Putative SLC9B1-like protein SLC9B1P1	4d0aB	NhaA
SLC9	Q86UD5	SLC9B2	2.A.36.2	944.3	Na_H_Exchanger	104.8	Sodium/hydrogen exchanger 9B2	4d0aB	NhaA
SLC9	Q4G0N8	SLC9C1	2.A.36.7	885.5	Na_H_Exchanger	96.8	Sodium/hydrogen exchanger 10	7dsxA	NhaA
SLC9	Q5TAH2	SLC9C2	2.A.36.7	788.2	Na_H_Exchanger	72.5	Sodium/hydrogen exchanger 11	7dsxA	NhaA
SLC10	Q14973	SLC10A1	2.A.28.1	442.8	SBF	165.9	Sodium/bile acid cotransporter	4n7wA	NhaA
SLC10	Q12908	SLC10A2	2.A.28.1	454.7	SBF	159.4	Ileal sodium/bile acid cotransporter	4n7wA	NhaA
SLC10	P09131	SLC10A3	2.A.28.1	449.2	SBF	149	P3 protein	4n7wA	NhaA
SLC10	Q96EP9	SLC10A4	2.A.28.1	494.5	SBF	117.8	Sodium/bile acid cotransporter 4	3zuxA	NhaA
SLC10	Q5PT55	SLC10A5	2.A.28.1	426.1	SBF	123.6	Sodium/bile acid cotransporter 5	4n7wA	NhaA
SLC10	Q3KNW5	SLC10A6	2.A.28.1	480.6	SBF	148.7	Solute carrier family 10 member 6	4n7wA	NhaA
SLC10	Q0GE19	SLC10A7	2.A.28.3	329.6	SBF_like	265	Sodium/bile acid cotransporter 7	3zuxA	NhaA
SLC11	P49279	SLC11A1	2.A.55.2	648.3	Nramp	390.2	Natural resistance-associated macrophage protein 1	5m8jA	APC
SLC11	P49281	SLC11A2	2.A.55.2	689.7	Nramp	399.7	Natural resistance-associated macrophage protein 2	5m8jA	APC
SLC12	Q13621	SLC12A1	2.A.30	1379	AA_permease	504	Solute carrier family 12 member 1	7d10A	APC
SLC12	P55011	SLC12A2	2.A.30	1423.3	SLC12	543.5	Solute carrier family 12 member 2	7d10A	APC
SLC12	P55017	SLC12A3	2.A.30	1259.8	AA_permease	462.9	Solute carrier family 12 member 3	7d10A	APC
SLC12	Q9UP95	SLC12A4	2.A.30	1311.7	AA_permease	217.9	Solute carrier family 12 member 4	6m22B	APC
SLC12	Q9H2X9	SLC12A5	2.A.30	1343.5	AA_permease	202.1	Solute carrier family 12 member 5	6m22B	APC
SLC12	Q9UHW9	SLC12A6	2.A.30	1316.9	AA_permease	213	Solute carrier family 12 member 6	6m22B	APC
SLC12	Q9Y666	SLC12A7	2.A.30	1318.3	AA_permease	201.3	Solute carrier family 12 member 7	6m22B	APC
SLC12	A0AV02	SLC12A8	2.A.30	722.4	AA_permease	149.7	Solute carrier family 12 member 8	6m1yB	APC
SLC12	Q9BXP2	SLC12A9	2.A.30	884.4	AA_permease	395.1	Solute carrier family 12 member 9	6m1yB	APC
SLC13	Q9BZW2	SLC13A1	2.A.47.1	746.2	Na_sulph_symp	323.4	Solute carrier family 13 member 1	7jskA	AbgT
SLC13	Q13183	SLC13A2	2.A.47.1	767.7	Na_sulph_symp	619.3	Solute carrier family 13 member 2	7jskA	AbgT
SLC13	Q8WWT9	SLC13A3	2.A.47.1	783.7	Na_sulph_symp	351	Solute carrier family 13 member 3	7jskA	AbgT
SLC13	Q9UKG4	SLC13A4	2.A.47.1	772.2	Na_sulph_symp	350.7	Solute carrier family 13 member 4	7jskA	AbgT
SLC13	Q86YT5	SLC13A5	2.A.47.1	772.4	Na_sulph_symp	431.9	Solute carrier family 13 member 5	7jskA	AbgT
SLC14	Q13336	SLC14A1	1.A.28.1	693.5	UT	370.4	Urea transporter 1	4ezcC	Amt
SLC14	Q15849	SLC14A2	1.A.28.1	1405.6	UT	712.9	Urea transporter 2	4ezcC	Amt
SLC15	P46059	SLC15A1	2.A.17.4	1012.4	PTR2	488.2	Solute carrier family 15 member 1	7nqkA	MFS
SLC15	Q16348	SLC15A2	2.A.17.4	1017.2	PTR2	461.8	Solute carrier family 15 member 2	7nqkA	MFS
SLC15	Q8IY34	SLC15A3	2.A.17.3	608.9	PTR2	276.3	Solute carrier family 15 member 3	4oh3A	MFS
SLC15	Q8N697	SLC15A4	2.A.17.3	609.7	PTR2	398.6	Solute carrier family 15 member 4	4oh3A	MFS
SLC15	A6NIM6	SLC15A5	2.A.17.3	526.8	PTR2	118.6	Solute carrier family 15 member 5	4oh3A	MFS
SLC16	P53985	SLC16A1	2.A.1.13	482	MFS_1	144.4	Monocarboxylate transporter 1	7ckoA	MFS
SLC16	P36021	SLC16A2	2.A.1.13	571.2	MFS_1	99.7	Monocarboxylate transporter 8	7ckoA	MFS
SLC16	O15427	SLC16A3	2.A.1.13	494.7	MFS_1	105.1	Monocarboxylate transporter 4	7ckoA	MFS
SLC16	O15374	SLC16A4	2.A.1.13	499.7	MFS_1	124.5	Monocarboxylate transporter 5	7ckoA	MFS
SLC16	O15375	SLC16A5	2.A.1.13	493.1	MFS_1	122.2	Monocarboxylate transporter 6	7ckoA	MFS
SLC16	O15403	SLC16A6	2.A.1.13	453.9	MFS_1	122.1	Monocarboxylate transporter 7	7ckoA	MFS
SLC16	O60669	SLC16A7	2.A.1.13	504.2	MFS_1	141.5	Monocarboxylate transporter 2	7bp3A	MFS
SLC16	O95907	SLC16A8	2.A.1.13	497.9	MFS_1	96.9	Monocarboxylate transporter 3	7bp3A	MFS
SLC16	Q7RTY1	SLC16A9	2.A.1.13	484.2	MFS_1	116.6	Monocarboxylate transporter 9	7ckoA	MFS
SLC16	Q8TF71	SLC16A10	2.A.1.13	561.5	MFS_1	117.1	Monocarboxylate transporter 10	7ckoA	MFS

SLC16	Q8NCK7	SLC16A11	2.A.1.13	398.9	MFS_1	108.8	Monocarboxylate transporter 11	7ckoA	MFS
SLC16	Q6ZSM3	SLC16A12	2.A.1.13	489.8	MFS_1	125.6	Monocarboxylate transporter 12	7ckoA	MFS
SLC16	Q7RTY0	SLC16A13	2.A.1.13	397.3	MFS_1	111.1	Monocarboxylate transporter 13	7bp3A	MFS
SLC16	Q7RTX9	SLC16A14	2.A.1.13	518.3	MFS_1	166.9	Monocarboxylate transporter 14	7ckoA	MFS
SLC17	Q14916	SLC17A1	2.A.1.14	418.3	MFS_1	164.1	Sodium-dependent phosphate transport protein 1	6v4dA	MFS
SLC17	O00624	SLC17A2	2.A.1.14	407.5	MFS_1	148.1	Sodium-dependent phosphate transport protein 3	6v4dA	MFS
SLC17	O00476	SLC17A3	2.A.1.14	334.5	MFS_1	119.5	Sodium-dependent phosphate transport protein 4	6v4dA	MFS
SLC17	Q9Y2C5	SLC17A4	2.A.1.14	447.9	MFS_1	184.4	Probable small intestine urate exporter	6v4dA	MFS
SLC17	Q9NRA2	SLC17A5	2.A.1.14	516.9	MFS_1	210.7	Sialin	6v4dA	MFS
SLC17	Q9P2U8	SLC17A6	2.A.1.14	657.2	MFS_1	169	Vesicular glutamate transporter 2	6v4dA	MFS
SLC17	Q9P2U7	SLC17A7	2.A.1.14	649.6	MFS_1	161.6	Vesicular glutamate transporter 1	6v4dA	MFS
SLC17	Q8NDX2	SLC17A8	2.A.1.14	650.6	MFS_1	158.1	Vesicular glutamate transporter 3	6v4dA	MFS
SLC17	Q9BYT1	SLC17A9	2.A.1.14	316.9	MFS_1	191.4	Solute carrier family 17 member 9	6e9nA	MFS
SLC18	P54219	SLC18A1	2.A.1.2	323.7	MFS_1	138.8	Chromaffin granule amine transporter	6v4dA	MFS
SLC18	Q05940	SLC18A2	2.A.1.2	331.7	MFS_1	152.5	Synaptic vesicular amine transporter	7mjsX	MFS
SLC18	Q16572	SLC18A3	2.A.1.2	303	MFS_1	125.6	Vesicular acetylcholine transporter	7mjsX	MFS
SLC18	Q6NT16	SLC18B1	2.A.1.2	201.4	MFS_1	159.6	MFS-type transporter SLC18B1	6kklA	MFS
SLC19	P41440	SLC19A1	2.A.48	736.2	Folate_carrier	564.9	Folate transporter 1	7bp3A	MFS
SLC19	O60779	SLC19A2	2.A.48	684.7	Folate_carrier	581.7	Thiamine transporter 1	4j05A	MFS
SLC19	Q9BZV2	SLC19A3	2.A.48	697.8	Folate_carrier	603.9	Thiamine transporter 2	4j05A	MFS
SLC20	Q8WUM9	SLC20A1	2.A.20.2	654.1	PHO4	470.2	Sodium-dependent phosphate transporter 1	6I85B	PiT
SLC20	Q08357	SLC20A2	2.A.20.2	639.7	PHO4	464.7	Sodium-dependent phosphate transporter 2	6I85B	PiT
SLC21/SLCO	P46721	SLCO1A2	2.A.60.1	799	OATP	561.9	Solute carrier organic anion transporter family member 1A2	7bp3A	MFS
SLC21/SLCO	Q9Y6L6	SLCO1B1	2.A.60.1	814.7	OATP	568.5	Solute carrier organic anion transporter family member 1B1	7ckoA	MFS
SLC21/SLCO	Q9NPD5	SLCO1B3	2.A.60.1	838.5	OATP	572	Solute carrier organic anion transporter family member 1B3	6e8jA	MFS
SLC21/SLCO	G3V0H7	SLCO1B7	2.A.60.1	755.9	OATP	522	Putative solute carrier organic anion transporter family member 1B7	7ckoA	MFS
SLC21/SLCO	Q9NYB5	SLCO1C1	2.A.60.1	851.4	OATP	599.6	Solute carrier organic anion transporter family member 1C1	7mjsX	MFS
SLC21/SLCO	Q92959	SLCO2A1	2.A.60	737.8	OATP	588.9	Solute carrier organic anion transporter family member 2A1	6e8jA	MFS
SLC21/SLCO	O94956	SLCO2B1	2.A.60.1	757.5	OATP	649	Solute carrier organic anion transporter family member 2B1	6lyyA	MFS
SLC21/SLCO	Q9UIG8	SLCO3A1	2.A.60.1	764.3	OATP	599.7	Solute carrier organic anion transporter family member 3A1	7mjsX	MFS
SLC21/SLCO	Q96BD0	SLCO4A1	2.A.60.1	723.5	OATP	600.2	Solute carrier organic anion transporter family member 4A1	7mjsX	MFS
SLC21/SLCO	Q6ZQN7	SLCO4C1	2.A.60	759.7	OATP	596.4	Solute carrier organic anion transporter family member 4C1	7mjsX	MFS
SLC21/SLCO	Q9H2Y9	SLCO5A1	2.A.60	848	OATP	618.5	Solute carrier organic anion transporter family member 5A1	6e8jA	MFS
SLC21/SLCO	Q86UG4	SLCO6A1	2.A.60	666.3	OATP	436.2	Solute carrier organic anion transporter family member 6A1	6vyhA	MFS
SLC22	O15245	SLC22A1	2.A.1.19	606	Sugar_tr	122.2	Solute carrier family 22 member 1	4zw9A	MFS
SLC22	O15244	SLC22A2	2.A.1.19	623.7	Sugar_tr	130.3	Solute carrier family 22 member 2	4zw9A	MFS
SLC22	O75751	SLC22A3	2.A.1.19	617.6	Sugar_tr	92	Solute carrier family 22 member 3	4j05A	MFS
SLC22	Q9H015	SLC22A4	2.A.1.19	562.2	MFS_1	89.1	Solute carrier family 22 member 4	5eqhA	MFS
SLC22	O76082	SLC22A5	2.A.1.19	593.9	Sugar_tr	92.9	Solute carrier family 22 member 5	4zw9A	MFS
SLC22	Q4U2R8	SLC22A6	2.A.1.19	655.1	Sugar_tr	114.1	Solute carrier family 22 member 6	7ckoA	MFS
SLC22	Q9Y694	SLC22A7	2.A.1.19	572.1	MFS_1	108.9	Solute carrier family 22 member 7	4zw9A	MFS
SLC22	Q8TCC7	SLC22A8	2.A.1.19	623.2	Sugar_tr	121.2	Solute carrier family 22 member 8	4j05A	MFS
SLC22	Q8IVM8	SLC22A9	2.A.1.19	544.7	MFS_1	102.3	Solute carrier family 22 member 9	5eqhA	MFS
SLC22	Q63ZE4	SLC22A10	2.A.1.19	570.8	Sugar_tr	89.6	Solute carrier family 22 member 10	5eqhA	MFS
SLC22	Q9NSA0	SLC22A11	2.A.1.19	568.8	Sugar_tr	95	Solute carrier family 22 member 11	5eqhA	MFS
SLC22	Q96S37	SLC22A12	2.A.1.19	566.5	MFS_1	81.3	Solute carrier family 22 member 12	5eqhA	MFS
SLC22	Q9Y226	SLC22A13	2.A.1.19	582.8	Sugar_tr	119.8	Solute carrier family 22 member 13	4j05A	MFS
SLC22	Q9Y267	SLC22A14	2.A.1.19	502.7	Sugar_tr	75.3	Solute carrier family 22 member 14	5eqhA	MFS
SLC22	Q8IZD6	SLC22A15	2.A.1.19	525.6	MFS_1	94.4	Solute carrier family 22 member 15	4zw9A	MFS
SLC22	Q86VW1	SLC22A16	2.A.1.19	558.9	Sugar_tr	93.8	Solute carrier family 22 member 16	5eqhA	MFS

SLC22	Q8WUG5	SLC22A17	2.A.1.19	431.8	Sugar_tr	65.2	Solute carrier family 22 member 17	5eqhA	MFS
SLC22	A6NK97	SLC22A20P	2.A.1.19	603.8	Sugar_tr	83.4	Solute carrier family 22 member 20	5eqhA	MFS
SLC22	A1A5C7	SLC22A23	2.A.1.19	588.3	Sugar_tr	85.7	Solute carrier family 22 member 23	5eqhA	MFS
SLC22	Q8N4F4	SLC22A24	2.A.1.19	281.5	MFS_1	54.4	Solute carrier family 22 member 24	4ldsB	MFS
SLC22	Q6T423	SLC22A25	2.A.1.19	538.1	MFS_1	72.6	Solute carrier family 22 member 25	5eqhA	MFS
SLC22	A6NKX4	SLC22A31	2.A.1.19	422	Sugar_tr	55.4	Putative solute carrier family 22 member 31	5eqhA	MFS
SLC23	Q9UHI7	SLC23A1	2.A.40.6	978.9	Xan_ur_permease	306.1	Solute carrier family 23 member 1	5i6cA	NAT
SLC23	Q9UGH3	SLC23A2	2.A.40.6	955.9	Xan_ur_permease	301.4	Solute carrier family 23 member 2	5i6cA	NAT
SLC23	Q6PIS1	SLC23A3	2.A.40.6	771.3	Xan_ur_permease	292.6	Solute carrier family 23 member 3	5i6cA	NAT
SLC24	Q6J4K2	SLC8B1	2.A.19.4	380.5	Na_Ca_ex	161.7	Mitochondrial sodium/calcium exchanger protein	4kjrA	NCX
SLC24	O60721	SLC24A1	2.A.19.4	793	Na_Ca_ex	194	Sodium/potassium/calcium exchanger 1	4kjrA	NCX
SLC24	Q9UI40	SLC24A2	2.A.19.4	717.2	Na_Ca_ex	209.9	Sodium/potassium/calcium exchanger 2	4kjrA	NCX
SLC24	Q9HC58	SLC24A3	2.A.19.4	681.1	Na_Ca_ex	192.7	Sodium/potassium/calcium exchanger 3	5hyaA	NCX
SLC24	Q8NFF2	SLC24A4	2.A.19.4	647.8	Na_Ca_ex	193.8	Sodium/potassium/calcium exchanger 4	4kjrA	NCX
SLC24	Q71RS6	SLC24A5	2.A.19.4	517.4	Na_Ca_ex	170.7	Sodium/potassium/calcium exchanger 5	5hyaA	NCX
SLC25	P53007	SLC25A1	2.A.29.7	610.3	Mito_carr	207.5	Tricarboxylate transport protein, mitochondrial	4c9hB	MCF
SLC25	Q9BXI2	SLC25A2	2.A.29.19	691.4	Mito_carr	198.1	Mitochondrial ornithine transporter 2	4c9hB	MCF
SLC25	Q00325	SLC25A3	2.A.29.4	656.8	Mito_carr	163.2	Phosphate carrier protein, mitochondrial	4c9hB	MCF
SLC25	P12235	SLC25A4	2.A.29.1	522.3	Mito_carr	229.2	ADP/ATP translocase 1	1okcA	MCF
SLC25	P05141	SLC25A5	2.A.29.1	520.1	Mito_carr	225.1	ADP/ATP translocase 2	1okcA	MCF
SLC25	P12236	SLC25A6	2.A.29.1	524.1	Mito_carr	228.1	ADP/ATP translocase 3	1okcA	MCF
SLC25	P25874	SLC25A7 (UCP1)	2.A.29.3	544.7	Mito_carr	203.3	Mitochondrial brown fat uncoupling protein 1	4c9hB	MCF
SLC25	P55851	SLC25A8 (UCP2)	2.A.29.3	546.3	Mito_carr	223.6	Mitochondrial uncoupling protein 2	2lckA	MCF
SLC25	P55916	SLC25A9 (UCP3)	2.A.29.3	560.8	Mito_carr	217.2	Mitochondrial uncoupling protein 3	2lckA	MCF
SLC25	Q9UBX3	SLC25A10	2.A.29.2	376.2	Mito_carr	186.7	Mitochondrial dicarboxylate carrier	4c9hB	MCF
SLC25	Q02978	SLC25A11	2.A.29.2	434.9	Mito_carr	199.7	Mitochondrial 2-oxoglutarate/malate carrier protein	4c9hB	MCF
SLC25	O75746	SLC25A12	2.A.29.14	453.5	Mito_carr	237.9	Calcium-binding mitochondrial carrier protein Aralar1	4c9hB	MCF
SLC25	Q9UJS0	SLC25A13	2.A.29.14	455.9	Mito_carr	239.1	Calcium-binding mitochondrial carrier protein Aralar2	4c9hB	MCF
SLC25	O95258	SLC25A14	2.A.29.24	595.3	Mito_carr	220.5	Brain mitochondrial carrier protein 1	4c9hB	MCF
SLC25	Q9Y619	SLC25A15	2.A.29.19	694.4	Mito_carr	204	Mitochondrial ornithine transporter 1	4c9hB	MCF
SLC25	P16260	SLC25A16	2.A.29.12	481.6	Mito_carr	237.7	Graves disease carrier protein	4c9hB	MCF
SLC25	O43808	SLC25A17	2.A.29.20	504.1	Mito_carr	186.9	Peroxisomal membrane protein PMP34	4c9hB	MCF
SLC25	Q9H1K4	SLC25A18	2.A.29.14	363.6	Mito_carr	205.8	Mitochondrial glutamate carrier 2	4c9hB	MCF
SLC25	Q9HC21	SLC25A19	2.A.29.16	557.8	Mito_carr	222.3	Mitochondrial thiamine pyrophosphate carrier	4c9hB	MCF
SLC25	O43772	SLC25A20	2.A.29.8	394	Mito_carr	235.5	Mitochondrial carnitine/acylcarnitine carrier protein	4c9hB	MCF
SLC25	Q9BQT8	SLC25A21	2.A.29.2	334.9	Mito_carr	205.2	Mitochondrial 2-oxodicarboxylate carrier	4c9hB	MCF
SLC25	Q9H936	SLC25A22	2.A.29.14	367.9	Mito_carr	210.7	Mitochondrial glutamate carrier 1	4c9hB	MCF
SLC25	Q9BV35	SLC25A23	2.A.29.23	553.1	Mito_carr	238.5	Calcium-binding mitochondrial carrier protein SCaMC-3	4c9hB	MCF
SLC25	Q6NUK1	SLC25A24	2.A.29.23	551.2	Mito_carr	248.9	Calcium-binding mitochondrial carrier protein SCaMC-1	4c9hB	MCF
SLC25	Q6KCM7	SLC25A25	2.A.29.23	540.6	Mito_carr	241.1	Calcium-binding mitochondrial carrier protein SCaMC-2	4c9hB	MCF
SLC25	Q70HW3	SLC25A26	2.A.29.18	384.6	Mito_carr	166	S-adenosylmethionine mitochondrial carrier protein	4c9hB	MCF
SLC25	O95847	SLC25A27	2.A.29.24	542	Mito_carr	217.3	Mitochondrial uncoupling protein 4	4c9hB	MCF
SLC25	Q96A46	SLC25A28	2.A.29.5	595.5	Mito_carr	207	Mitoferrin-2	4c9hB	MCF
SLC25	Q8N8R3	SLC25A29	2.A.29.8	387.5	Mito_carr	218.5	Mitochondrial basic amino acids transporter	4c9hB	MCF
SLC25	Q5SVS4	SLC25A30	2.A.29.24	596.1	Mito_carr	228	Kidney mitochondrial carrier protein 1	4c9hB	MCF
SLC25	Q9H0C2	SLC25A31	2.A.29.1	509.1	Mito_carr	239.2	ADP/ATP translocase 4	1okcA	MCF
SLC25	Q9H2D1	SLC25A32	2.A.29.10	375.3	Mito_carr	232.2	Mitochondrial folate transporter/carrier	4c9hB	MCF
SLC25	Q9BSK2	SLC25A33	2.A.29.10	378.3	Mito_carr	217	Solute carrier family 25 member 33	4c9hB	MCF
SLC25	Q6PIV7	SLC25A34	2.A.29.15	550.4	Mito_carr	165.9	Solute carrier family 25 member 34	4c9hB	MCF
SLC25	Q3KQZ1	SLC25A35	2.A.29.15	555.6	Mito_carr	163.2	Solute carrier family 25 member 35	4c9hB	MCF

SLC25	Q96CQ1	SLC25A36	2.A.29.10	382.7	Mito_carr	227.4	Solute carrier family 25 member 36	4c9hB	MCF
SLC25	Q9NYZ2	SLC25A37	2.A.29.5	577.2	Mito_carr	212.4	Mitoferrin-1	4c9hB	MCF
SLC25	Q96DW6	SLC25A38	2.A.29.5	394.9	Mito_carr	197.9	Mitochondrial glycine transporter	4c9hB	MCF
SLC25	Q9BZJ4	SLC25A39	2.A.29.14	361.2	Mito_carr	202	Solute carrier family 25 member 39	4c9hB	MCF
SLC25	Q8TBP6	SLC25A40	2.A.29.14	361.8	Mito_carr	210	Solute carrier family 25 member 40	4c9hB	MCF
SLC25	Q8N5S1	SLC25A41	2.A.29.23	440	Mito_carr	212.9	Solute carrier family 25 member 41	4c9hB	MCF
SLC25	Q86VD7	SLC25A42	2.A.29.12	448	Mito_carr	239.9	Mitochondrial coenzyme A transporter SLC25A42	4c9hB	MCF
SLC25	Q8WUT9	SLC25A43	2.A.29.23	368.6	Mito_carr	203.9	Solute carrier family 25 member 43	4c9hB	MCF
SLC25	Q96H78	SLC25A44	2.A.29.14	313.6	Mito_carr	170.6	Solute carrier family 25 member 44	4c9hB	MCF
SLC25	Q8N413	SLC25A45	2.A.29.8	366.8	Mito_carr	215.4	Solute carrier family 25 member 45	4c9hB	MCF
SLC25	Q96AG3	SLC25A46	2.A.29.30	1010.6	Mito_carr	65.4	Solute carrier family 25 member 46	4c9hB	MCF
SLC25	Q6Q0C1	SLC25A47	2.A.29.8	382.2	Mito_carr	200.5	Solute carrier family 25 member 47	4c9hB	MCF
SLC25	Q6ZT89	SLC25A48	2.A.29.8	374.4	Mito_carr	202.2	Solute carrier family 25 member 48	4c9hB	MCF
SLC25	Q9NZJ7	SLC25A49 (MTCH1)	2.A.29.25	845.6	Mito_carr	49.7	Mitochondrial carrier homolog 1	4c9hB	MCF
SLC25	Q9Y6C9	SLC25A50 (MTCH2)	2.A.29.25	584.8	Mito_carr	59.7	Mitochondrial carrier homolog 2	4c9hB	MCF
SLC25	Q9H1U9	SLC25A51	2.A.29.2	319.3	Mito_carr	113.3	Solute carrier family 25 member 51	4c9hB	MCF
SLC25	Q3SY17	SLC25A52	2.A.29.2	318.7	Mito_carr	116.2	Solute carrier family 25 member 52	4c9hB	MCF
SLC25	Q5H9E4	SLC25A53	2.A.29.2	127.5	Mito_carr	85	Solute carrier family 25 member 53	4c9hB	MCF
SLC26	Q9H2B4	SLC26A1	2.A.53.2	849.4	Sulfate_transp	424.8	Sulfate anion transporter 1	7ch1B	NAT
SLC26	P50443	SLC26A2	2.A.53.2	924.8	Sulfate_transp	445.1	Sulfate transporter	7ch1B	NAT
SLC26	P40879	SLC26A3	2.A.53.2	998.1	Sulfate_transp	391.1	Chloride anion exchanger	7ch1B	NAT
SLC26	O43511	SLC26A4	2.A.53.2	1007.5	Sulfate_transp	353.3	Pendrin	7ch1B	NAT
SLC26	P58743	SLC26A5	2.A.53.2	990.7	Sulfate_transp	358.9	Prestin	7ch1B	NAT
SLC26	Q9BXS9	SLC26A6	2.A.53.2	1000.1	Sulfate_transp	369.7	Solute carrier family 26 member 6	7ch1B	NAT
SLC26	Q8TE54	SLC26A7	2.A.53.2	754.9	Sulfate_transp	312	Anion exchange transporter	7ch1B	NAT
SLC26	Q96RN1	SLC26A8	2.A.53.2	1058.5	Sulfate_transp	247.3	Testis anion transporter 1	7ch1B	NAT
SLC26	Q7LBE3	SLC26A9	2.A.53.2	967.6	Sulfate_transp	312	Solute carrier family 26 member 9	7ch1B	NAT
SLC26	Q8NG04	SLC26A10	2.A.53.2	604.8	Sulfate_transp	302.2	Solute carrier family 26 member 10	7ch1B	NAT
SLC26	Q86WA9	SLC26A11	2.A.53.1	650.9	Sulfate_transp	318.8	Sodium-independent sulfate anion transporter	7ch1B	NAT
SLC28	O00337	SLC28A1	2.A.41.2	1016.5	Gate	28.9	Sodium/nucleoside cotransporter 1	6kswA	CNT
SLC28	O43868	SLC28A2	2.A.41.2	1020.1	Gate	35.1	Sodium/nucleoside cotransporter 2	6kswA	CNT
SLC28	Q9HAS3	SLC28A3	2.A.41.2	1039.5	Gate	33.3	Solute carrier family 28 member 3	6kswA	CNT
SLC29	Q99808	SLC29A1	2.A.57.1	550.2	Nucleoside_tran	439.6	Equilibrative nucleoside transporter 1	6ob7A	MFS
SLC29	Q14542	SLC29A2	2.A.57.1	551.3	Nucleoside_tran	420.7	Equilibrative nucleoside transporter 2	6ob7A	MFS
SLC29	Q9BZD2	SLC29A3	2.A.57.1	462.4	Nucleoside_tran	214.4	Equilibrative nucleoside transporter 3	6ob7A	MFS
SLC29	Q7RTT9	SLC29A4	2.A.57.1	551.5	Nucleoside_tran	160.3	Equilibrative nucleoside transporter 4	6ob7A	MFS
SLC30	Q9Y6M5	SLC30A1	2.A.4.2	603.8	Cation_efflux	155.5	Zinc transporter 1	6xpdA	CDF
SLC30	Q9BRI3	SLC30A2	2.A.4.3	421.6	Cation_efflux	84.4	Zinc transporter 2	6xpeA	CDF
SLC30	Q99726	SLC30A3	2.A.4.3	497.7	Cation_efflux	153.7	Zinc transporter 3	6xpdA	CDF
SLC30	O14863	SLC30A4	2.A.4.3	570.6	Cation_efflux	185.4	Zinc transporter 4	6xpeA	CDF
SLC30	Q8TAD4	SLC30A5	2.A.4.4	522.6	Cation_efflux	163.7	Zinc transporter 5	6xpdA	CDF
SLC30	Q6NXT4	SLC30A6	2.A.4.4	409.1	Cation_efflux	113.1	Zinc transporter 6	6xpdA	CDF
SLC30	Q8NEW0	SLC30A7	2.A.4.4	437.6	Cation_efflux	151	Zinc transporter 7	6xpdA	CDF
SLC30	Q8IWU4	SLC30A8	2.A.4.3	493.3	Cation_efflux	148.1	Zinc transporter 8	6xpdA	CDF
SLC30	Q6PML9	SLC30A9	2.A.4.6	1373.1	Cation_efflux	107.2	Zinc transporter 9	3j1zQ	CDF
SLC30	Q6XR72	SLC30A10	2.A.4.2	523.5	Cation_efflux	152.8	Zinc transporter 10	6xpdA	CDF
SLC31	O15431	SLC31A1	1.A.56.1	173	Ctr	109.3	High affinity copper uptake protein 1	6m97A	Ctr
SLC31	O15432	SLC31A2	1.A.56.1	119.1	Ctr	104.2	Probable low affinity copper uptake protein 2	6m97A	Ctr
SLC32	Q9H598	SLC32A1	2.A.18.5	715.7	Aa_trans	331.5	Vesicular inhibitory amino acid transporter	7kgvB	APC
SLC33	O00400	SLC33A1	2.A.1.25	554.7	Acatn	411.7	Acetyl-coenzyme A transporter 1	7mjsX	MFS

SLC34	Q06495	SLC34A1	2.A.58.1	703	Na_Pi_cotrans	138.8	Sodium-dependent phosphate transport protein 2A		
SLC34	O95436	SLC34A2	2.A.58.1	726.3	Na_Pi_cotrans	129.3	Sodium-dependent phosphate transport protein 2B		
SLC34	Q8N130	SLC34A3	2.A.58.1	639.8	Na_Pi_cotrans	127.5	Sodium-dependent phosphate transport protein 2C		
SLC35	P78382	SLC35A1	2.A.7.12	412.1	Nuc_sug_transp	418.5	CMP-sialic acid transporter	6xboA	NST
SLC35	P78381	SLC35A2	2.A.7.12	469.6	Nuc_sug_transp	446.6	UDP-galactose translocator	6ukjA	NST
SLC35	Q9Y2D2	SLC35A3	2.A.7.12	437.6	Nuc_sug_transp	457.2	UDP-N-acetylglucosamine transporter	6xboA	NST
SLC35	Q96G79	SLC35A4	2.A.7.12	293.1	Nuc_sug_transp	103.2	Probable UDP-sugar transporter protein SLC35A4	6xboA	NST
SLC35	Q9BS91	SLC35A5	2.A.7.12	432.2	Nuc_sug_transp	176.1	Probable UDP-sugar transporter protein SLC35A5	6xboA	NST
SLC35	P78383	SLC35B1	2.A.7.11	398.2	UAA	189.4	Solute carrier family 35 member B1	5ogeD	NST
SLC35	Q8TB61	SLC35B2	2.A.7.11	487.8	UAA	252.1	Adenosine 3'-phospho 5'-phosphosulfate transporter 1	5ogeD	NST
SLC35	Q9H1N7	SLC35B3	2.A.7.11	399.1	UAA	313.7	Adenosine 3'-phospho 5'-phosphosulfate transporter 2	6ukjA	NST
SLC35	Q969S0	SLC35B4	2.A.7.10	518.6	UAA	390.3	UDP-xylose and UDP-N-acetylglucosamine transporter	5ogeG	NST
SLC35	Q96A29	SLC35C1	2.A.7.16	633.9	TPT	75.3	GDP-fucose transporter 1	5ogeD	NST
SLC35	Q9NQQ7	SLC35C2	2.A.7.9	259.2	TPT	92	Solute carrier family 35 member C2	5y78A	NST
SLC35	Q9NTN3	SLC35D1	2.A.7.15	530.3	TPT	58.6	UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter	5ogeD	NST
SLC35	Q76EJ3	SLC35D2	2.A.7.15	500.7	TPT	73.8	UDP-N-acetylglucosamine/UDP-glucose/GDP-mannose transporter	5ogeD	NST
SLC35	Q5M8T2	SLC35D3	2.A.7.15	507.2	EamA	27.9	Solute carrier family 35 member D3	6ukjA	NST
SLC35	Q96K37	SLC35E1	2.A.7.9	374.6	TPT	310.1	Solute carrier family 35 member E1	6ukjA	NST
SLC35	P0CK97	SLC35E2A	2.A.7.9	104.4	TPT	107.5	Solute carrier family 35 member E2A	6ukjA	NST
SLC35	P0CK96	SLC35E2B	2.A.7.9	297	TPT	312.3	Solute carrier family 35 member E2B	6ukjA	NST
SLC35	Q7Z769	SLC35E3	2.A.7.9	234.5	TPT	72.1	Solute carrier family 35 member E3	5ogeD	NST
SLC35	Q6ICL7	SLC35E4	2.A.7.9	233.7	TPT	82.1	Solute carrier family 35 member E4	5ogeD	NST
SLC35	Q5T1Q4	SLC35F1	2.A.7.24	354.6	SLC35F	495.8	Solute carrier family 35 member F1	6ukjA	NST
SLC35	Q8IXU6	SLC35F2	2.A.7.24	344.1	SLC35F	492.9	Solute carrier family 35 member F2	5ogeD	NST
SLC35	Q8IY50	SLC35F3	2.A.7.24	405.4	EamA	32.3	Putative thiamine transporter SLC35F3	6ukjA	NST
SLC35	A4IF30	SLC35F4	2.A.7.24	431.6	EamA	22.9	Solute carrier family 35 member F4	6ukjA	NST
SLC35	Q8WV83	SLC35F5	2.A.7.24	428	EamA	45.4	Solute carrier family 35 member F5	6ukjA	NST
SLC35	Q8N357	SLC35F6	2.A.7.12	321.6	EamA	51.9	Solute carrier family 35 member F6	5y78A	NST
SLC35	Q2M3R5	SLC35G1	2.A.7.28	369.1	EamA	96.5	Solute carrier family 35 member G1	5i20D	NST
SLC35	Q8TBE7	SLC35G2	2.A.7.28	421.1	EamA	70.6	Solute carrier family 35 member G2	6ukjA	NST
SLC35	Q8N808	SLC35G3	2.A.7.28	347.3	EamA	40.8	Solute carrier family 35 member G3	5ogeG	NST
SLC35	P0C7Q5	SLC35G4	2.A.7.28	354.6	EamA	41.3	Putative solute carrier family 35 member G4	5i20D	NST
SLC35	Q96KT7	SLC35G5	2.A.7.28	352.8	EamA	37.5	Solute carrier family 35 member G5	5ogeG	NST
SLC35	P0C7Q6	SLC35G6	2.A.7.28	353.9	EamA	42.3	Solute carrier family 35 member G6	5ogeG	NST
SLC36	Q7Z2H8	SLC36A1	2.A.18.8	816.6	Aa_trans	257.1	Proton-coupled amino acid transporter 1	7kgvB	APC
SLC36	Q495M3	SLC36A2	2.A.18.8	803.5	Aa_trans	248.3	Proton-coupled amino acid transporter 2	7kgvB	APC
SLC36	Q495N2	SLC36A3	2.A.18.8	737.4	Aa_trans	230.9	Proton-coupled amino acid transporter 3	7kgvB	APC
SLC36	Q6YBV0	SLC36A4	2.A.18.8	767.4	Aa_trans	258.7	Proton-coupled amino acid transporter 4	7kgvB	APC
SLC37	P57057	SLC37A1	2.A.1.4	575.6	MFS_1	123.4	Glucose-6-phosphate exchanger SLC37A1	1pw4A	MFS
SLC37	Q8TED4	SLC37A2	2.A.1.4	568.1	MFS_1	145	Glucose-6-phosphate exchanger SLC37A2	4j05A	MFS
SLC37	Q8NCC5	SLC37A3	2.A.1.4	515.3	MFS_1	129.5	Sugar phosphate exchanger 3	4j05A	MFS
SLC37	O43826	SLC37A4	2.A.1.4	474.2	MFS_1	171.1	Glucose-6-phosphate exchanger SLC37A4	6fmrA	MFS
SLC38	Q9H2H9	SLC38A1	2.A.18.6	440.8	Aa_trans	266.4	Sodium-coupled neutral amino acid transporter 1	7kgvB	APC
SLC38	Q96QD8	SLC38A2	2.A.18.6	478	Aa_trans	263.2	Sodium-coupled neutral amino acid transporter 2	7kgvB	APC
SLC38	Q99624	SLC38A3	2.A.18.6	479.3	Aa_trans	373.3	Sodium-coupled neutral amino acid transporter 3	7kgvB	APC
SLC38	Q969I6	SLC38A4	2.A.18.6	485.3	Aa_trans	274.8	Sodium-coupled neutral amino acid transporter 4	7kgvB	APC
SLC38	Q8WUX1	SLC38A5	2.A.18.6	450.3	Aa_trans	295.9	Sodium-coupled neutral amino acid transporter 5	7kgvB	APC
SLC38	Q8IZM9	SLC38A6	2.A.18.6	393.7	Aa_trans	261.1	Probable sodium-coupled neutral amino acid transporter 6	7kgvB	APC
SLC38	Q9NVC3	SLC38A7	2.A.18.6	340.7	Aa_trans	187.8	Putative sodium-coupled neutral amino acid transporter 7	7d10A	APC
SLC38	A6NNN8	SLC38A8	2.A.18.6	317	Aa_trans	180.4	Putative sodium-coupled neutral amino acid transporter 8	7kgvB	APC

SLC38	Q8NBW4	SLC38A9	2.A.18.9	1225	Aa_trans	98.2	Sodium-coupled neutral amino acid transporter 9	7kgvB	APC
SLC38	Q9HBR0	SLC38A10	2.A.18.6	680	Aa_trans	179.5	Putative sodium-coupled neutral amino acid transporter 10	7kgvB	APC
SLC38	Q08AI6	SLC38A11	2.A.18.6	340	Aa_trans	179	Putative sodium-coupled neutral amino acid transporter 11	7kgvB	APC
SLC39	Q9NY26	SLC39A1	2.A.5.3	313.6	Zip	199.1	Zinc transporter ZIP1	5tsaA	ZIP
SLC39	Q9NP94	SLC39A2	2.A.5.3	314.3	Zip	216.9	Zinc transporter ZIP2	5tsaA	ZIP
SLC39	Q9BRY0	SLC39A3	2.A.5.3	310.5	Zip	152.1	Zinc transporter ZIP3	5tsaA	ZIP
SLC39	Q6P5W5	SLC39A4	2.A.5.4	530.6	Zip	202.9	Zinc transporter ZIP4	5tsaA	ZIP
SLC39	Q6ZMH5	SLC39A5	2.A.5.4	469.3	Zip	220.5	Zinc transporter ZIP5	5tsaA	ZIP
SLC39	Q13433	SLC39A6	2.A.5.4	601.8	Zip	348.4	Zinc transporter ZIP6	5tsaA	ZIP
SLC39	Q92504	SLC39A7	2.A.5.4	354.4	Zip	268.3	Zinc transporter SLC39A7	5tsaA	ZIP
SLC39	Q9C0K1	SLC39A8	2.A.5.4	436.9	Zip	242.1	Zinc transporter ZIP8	5tsaA	ZIP
SLC39	Q9NUM3	SLC39A9	2.A.5.6	698.5	Zip	132.1	Zinc transporter ZIP9	5tsaA	ZIP
SLC39	Q9ULF5	SLC39A10	2.A.5.4	598	Zip	347.3	Zinc transporter ZIP10	5tsaA	ZIP
SLC39	Q8N1S5	SLC39A11	2.A.5.5	271.8	Zip	86.8	Zinc transporter ZIP11	5tsaA	ZIP
SLC39	Q504Y0	SLC39A12	2.A.5.4	532.3	Zip	217.4	Zinc transporter ZIP12	5tsaA	ZIP
SLC39	Q96H72	SLC39A13	2.A.5.4	298.4	Zip	181.8	Zinc transporter ZIP13	5tsaA	ZIP
SLC39	Q15043	SLC39A14	2.A.5.4	473.7	Zip	221.7	Zinc transporter ZIP14	5tsaA	ZIP
SLC40	Q9NP59	SLC40A1	2.A.100.1	734	FPN1	642	Solute carrier family 40 member 1	6vyhA	MFS
SLC41	Q8IVJ1	SLC41A1	1.A.26.2	858.3	MgtE	183.1	Solute carrier family 41 member 1	4u9nB	MgtE
SLC41	Q96JW4	SLC41A2	1.A.26.2	866.9	MgtE	174.8	Solute carrier family 41 member 2	4u9nB	MgtE
SLC41	Q96GZ6	SLC41A3	1.A.26.2	810.5	MgtE	163	Solute carrier family 41 member 3	4u9nB	MgtE
SLC42	Q02094	RHAG	1.A.11.4	625.4	Ammonium_transp	309.3	Ammonium transporter Rh type A	3hd6A	Amt
SLC42	Q9H310	RHBG	1.A.11.4	622.5	Ammonium_transp	244.2	Ammonium transporter Rh type B	3hd6A	Amt
SLC42	Q9UBD6	RHCG	1.A.11.4	696.7	Ammonium_transp	265.5	Ammonium transporter Rh type C	3hd6A	Amt
SLC43	O75387	SLC43A1	2.A.1.44	1014	MFS_1	55.5	Large neutral amino acids transporter small subunit 3	7ckoA	MFS
SLC43	Q8N370	SLC43A2	2.A.1.44	1027.6	MFS_1	56.6	Large neutral amino acids transporter small subunit 4	7ckoA	MFS
SLC43	Q8NBI5	SLC43A3	2.A.1.44	794.4	MFS_1	44.1	Solute carrier family 43 member 3	7ckoA	MFS
SLC44	Q8WWI5	SLC44A1	2.A.92	902.2	Choline_transpo	354.7	Choline transporter-like protein 1		
SLC44	Q8IWA5	SLC44A2	2.A.92	1006.1	Choline_transpo	369.2	Choline transporter-like protein 2		
SLC44	Q8N4M1	SLC44A3	2.A.92	824	Choline_transpo	303.6	Choline transporter-like protein 3		
SLC44	Q53GD3	SLC44A4	2.A.92	945.4	Choline_transpo	366.3	Choline transporter-like protein 4		
SLC44	Q8NCS7	SLC44A5	2.A.92	970	Choline_transpo	364.1	Choline transporter-like protein 5		
SLC45	Q9Y2W3	SLC45A1	2.A.2.4	772.4	MFS_1	64.5	Proton-associated sugar transporter A	7mjsX	MFS
SLC45	Q9UMX9	SLC45A2	2.A.2.4	656.2	MFS_1	47.5	Membrane-associated transporter protein	7mjsX	MFS
SLC45	Q96JT2	SLC45A3	2.A.2.4	579.4	MFS_1	48.5	Solute carrier family 45 member 3	7mjsX	MFS
SLC45	Q5BKX6	SLC45A4	2.A.2.4	795.8	MFS_2	32.9	Solute carrier family 45 member 4	7mjsX	MFS
SLC46	Q96NT5	SLC46A1	2.A.1.50	663.5	MFS_1	59.1	Proton-coupled folate transporter	7bc6A	MFS
SLC46	Q9BY10	SLC46A2	2.A.1.50	658.3	MFS_1	52.1	Thymic stromal cotransporter homolog	7bc6A	MFS
SLC46	Q7Z3Q1	SLC46A3	2.A.1.50	681.4	MFS_1	66	Solute carrier family 46 member 3	7bc6A	MFS
SLC47	Q96FL8	SLC47A1	2.A.66.1	553.5	MatE	210	Multidrug and toxin extrusion protein 1	7dqkB	MATE
SLC47	Q86VL8	SLC47A2	2.A.66.1	561.8	MatE	196.3	Multidrug and toxin extrusion protein 2	7dqkB	MATE
SLC48	Q6P1K1	SLC48A1	2.A.110.1	167.8	HRG	92.2	Heme transporter HRG1		
SLC49	Q9Y5Y0	SLC49A1 (FLVCR1)	2.A.1.28	763.5	MFS_1	94.4	Feline leukemia virus subgroup C receptor-related protein 1	6v4dA	MFS
SLC49	Q9UPI3	SLC49A2 (FLVCR2)	2.A.1.28	719.7	MFS_1	95.5	Feline leukemia virus subgroup C receptor-related protein 2	6v4dA	MFS
SLC49	Q6UXD7	SLC49A3	2.A.1.28	489.2	MFS_1	74.2	Solute carrier family 49 member A3	7ckoA	MFS
SLC49	Q96SL1	SLC49A4	2.A.1.28	584.3	MFS_1	29.8	Solute carrier family 49 member 4	7c77B	MFS
SLC50	Q9BRV3	SLC50A1	2.A.123.1	176.5	MtN3_slv	200.4	Sugar transporter SWEET1	5xpdA	SWEET
SLC51	Q86UW1	SLC51A	2.A.82-Solute_trans_a	404.7	Solute_trans_a	125	Organic solute transporter subunit alpha		
SLC51	Q86UW2	SLC51B	2.A.82-OSTbeta	162	OSTbeta	187.5	Organic solute transporter subunit beta		
SLC52	Q9NWF4	SLC52A1	2.A.125	535	DUF1011	122	Solute carrier family 52, riboflavin transporter, member 1	6ob7A	MFS

SLC52	Q9HAB3	SLC52A2	2.A.125	535.8	DUF1011	121.1	Solute carrier family 52, riboflavin transporter, member 2	6ob7A	MFS
SLC52	Q9NQ40	SLC52A3	2.A.125	562.6	DUF1011	125.7	Solute carrier family 52, riboflavin transporter, member 3	6ob7A	MFS

Supplementary File 1

Phylogenetic trees of SLC-like protein families with more than two members. Trees were generated using multiple alignment by ClustalO, maximum likelihood tree generation by PhyML, followed by tree reconciliation with the species tree using NOTUNG (see Methods). The species tree with internal names of putative ancestor taxa is shown on each page on the upper-right hand corner. The trees are shown as dendrograms and branch lengths are not indicative of evolutionary distance. Each tree leaf corresponds to an SLC-like protein sequence denoting a gene, labels show the gene symbol, UniProt accession and taxon name. Leaves with labels ending with "*LOST" denote putative genes lost in the indicated ancestral species. Red "D" denote gene duplication nodes, normal nodes correspond to speciation nodes. Light green numbers denote branch support values as calculated by NOTUNG.

<Link: slc-all-trees.pdf>