UNIVERSITY OF CALIFORNIA SAN DIEGO

Systems Biology Analysis Reveals Eight SLC22 Transporter Subgroups, Including OATs, OCTs, and OCTNs, and Drosophila SLC22 orthologs related to OATs, OCTs, and OCTNs regulate development and responsiveness to oxidative stress

A thesis submitted in partial satisfaction of the requirements of the degree Master of Science

in

Biology

by

Darcy Christine Engelhart

Committee in charge:

Professor Sanjay K. Nigam, Chair
Professor Milton Saier, Co-Chair
Professor Scott Rifkin

2020

The Thesis of Darcy Christine Engelhart is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

_____

_____

Co-Chair

_____

Chair

University of California San Diego

2020

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ABC	ATP-Binding Cassette
BLAST-P	Basic local alignment searching tool - Protein
BOCT1	Brain organic cation transporter 1
CCTOP	Constrained Consensus TOPology prediction server
CNS	Central nervous system
CT1/2	Carnitine transporter 1/2
ECD	Extracellular domain
ENCODE	Encyclopedia of DNA elements
FANTOM5	Functional annotation of the mammalian genome project
FASTA	Pairwise sequence alignment tool
FLIPT1/2	Fly-like putative transporter 1/2
GLK	Gut-liver-kidney axis
GTEx	Genotype-Tissue Expression project
GWAS	Genome wide association study
hOA/CTx	Human organic anion/cation transporter x
IBD	Irritable bowel disease
ICD	Intracellular domain
ICM	Internal Coordinate Mechanics
Lcn2-R	Lipocalin-2 receptor
MAFFT	Multiple alignment using fast Fourier transform
MEK	Mitogen-activated protein kinase kinase
MEME	Multiple expectation-maximum for Motif Elicitation
mOA/CTx	Mouse organic anion/cation transporter x
OAT	Organic anion transporter
OATP	Organic Anion Transporting Polypeptide
OATS1	Oat Subgroup 1
OATS2	Oat Subgroup 2
OATS3	Oat Subgroup 3
OATS4	Oat Subgroup 4
OCT	Organic cation transporter
OCTN	Organic carnitine transporter
PQ	Paraquat
RBC	Red Blood Cell
SLC22	Solute carrier family 22
SNP	Single nucleotide polymorphism
SSearch	Smith-Waterman Search

TMD	Transmembrane domain
TOPO2	Transmembrane protein display software
UCSC	University of California, Santa Cruz
UST	Unknown substrate transporter

LIST OF SUPPLEMENTAL FILES

engelhart_01_SLC22_specificity.xlsx

engelhart_02_SLC22_tissue.xlsx

engelhart_03_OATS4.xlsx

engelhart_04_ICM_homology.xlsx

engelhart_05_DM_SLC22_overview.xlsx

engelhart_06_supplementary_materials.zip

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Sanjay Nigam for his support and guidance during the two and a half years that I worked in the Nigam lab. He has been an invaluable resource and an extremely active and patient mentor. I would also like to thank Dr. Milton Saier and the members of the Saier Lab for their scientific input and Dr. Scott Rifkin and Dr. Michael Baker for their feedback on my project. I want to give special thanks to Dr. Kevin Bush who mentored me during my earliest days in the lab and continues to support me and my scientific career. He displayed endless patience teaching me cell culture and the ins and outs of working in a lab.

I am especially grateful to Dr. Priti Azad for her support and active participation in my investigation of SLC22. Without her and Ying Lu-Bo, I would not have been able to pursue fascinating work with *Drosophila melanogaster*. I would also like to thank Dr. Gabriel Haddad for allowing me to use his fly room and his lab's expertise to pursue this research. I'd also like to thank Da Shi for contributing homology studies, Jeffry Granados for being constantly available for scientific discussion and Suwayda Ali for dedicating most of her free time to data collection and maintenance of our lab's fly stocks. I have been so incredibly fortunate to have many collaborators and mentors throughout my time in the Nigam lab and I am grateful to every person who provided me with their expert opinions and support.

Chapter 1, in full, has been submitted for publication of the material as it may appear in International Journal of Molecular Sciences, 2020, Engelhart, Darcy C.; Granados, Jeffry C.; Shi, Da; Saier, Milton H.; Baker, Michael E.; Abagyan, Reuben; Nigam, Sanjay K., 2020. The thesis author was the primary investigator and author of this paper.

Chapter 2, in full, has been submitted for publication of the material as it may appear in International Journal of Molecular Sciences, 2020, Engelhart, Darcy C.; Azad, Priti; Ali,

Suwayda; Granados, Jeffry C.; Haddad, Gabriel G.; Nigam, Sanjay K., 2020. The thesis author was the primary investigator and author of this paper.

ABSTRACT OF THE THESIS

Systems Biology Analysis Reveals Eight SLC22 Transporter Subgroups, Including OATs, OCTs, and OCTNs, and Drosophila SLC22 orthologs related to OATs, OCTs, and OCTNs regulate development and responsiveness to oxidative stress

by

Darcy Christine Engelhart

Master of Science in Biology

University of California San Diego, 2020

Professor Sanjay K. Nigam, Chair
Professor Milton Saier, Co-Chair

The SLC22 family of transmembrane transport proteins consists of at least 31 known members that transport organic anions (OATs), organic cations (OCTs) and zwitterions (OCTNs). Despite their interaction with many endogenous metabolites and signaling molecules, many of these transporters are referred to as “drug” transporters since they facilitate the

movement of pharmaceutical drugs. However, their evolution and conservation as a group underscores key endogenous functions.

Chapter 1 suggests a re-designation of SLC22 into nine subclades with four new, functional subclades arising out of the previously defined Oat subclade - referred to as Oat subgroups. These new Oat subgroups are: OATS1 (SLC22A6, SLC22A8, and SLC22A20), OATS2 (SLC22A7), OATS3 (SLC22A11, SLC22A12, and slc22a22), and OATS4 (SLC22A9, SLC22A10, SLC22A24, SLC22A25 plus the rodent expansion of slc22a19, and slc22a26-30). We also propose the reclassification of SLC22A16 into the preexisting OCTN subclade (SLC22A4, SLC22A5, and slc22a21) and the renaming of the Oct-related subclade, which now contains only SLC22A15, to the Octn-related subclade.

Due to the functional similarities between the *Drosophila melanogaster* (fruit fly) hindgut and Malpighian tubules and the human intestines and kidneys, where most, if not all SLC22 family members can be localized, we are assessing the functionality of SLC22 transport proteins using flies as a model organism. Chapter 2 explores orthologous relationships between D. melanogaster transporters and their human counterparts. It also assesses fly transporters for their role in management of oxidative stress.

I.

CHAPTER 1. Systems Biology Analysis Reveals Eight SLC22 Transporter Subgroups, Including OATs, OCTs, and OCTNs, and Drosophila SLC22 orthologs related to OATs, OCTs, and OCTNs

1.1 ABSTRACT

The SLC22 family of OATs, OCTs, and OCTNs is emerging as a central hub of endogenous physiology. Despite often being referred to as "drug" transporters, they facilitate the movement of metabolites and key signaling molecules. An in-depth reanalysis supports a reassignment of these proteins into eight functional subgroups, with four new subgroups arising from the previously defined OAT subclade: OATS1 (SLC22A6, SLC22A8, SLC22A20), OATS2 (SLC22A7), OATS3 (SLC22A11, SLC22A12, Slc22a22), and OATS4 (SLC22A9, SLC22A10, SLC22A24, SLC22A25). We propose merging the OCTN (SLC22A4, SLC22A5, Slc22a21) and OCT-related (SLC22A15, SLC22A16) subclades into the OCTN/OCTN-related subgroup. Using data from GWAS, in vivo models, and in vitro assays, we developed an SLC22 transporter-metabolite network and similar subgroup networks, which suggest how multiple SLC22 transporters interact to regulate a particular metabolite. Subgroup associations include: OATS1 with signaling molecules, uremic toxins, and odorants, OATS2 with cyclic nucleotides, OATS3 with uric acid, OATS4 with conjugated sex hormones, particularly etiocholanolone glucuronide, OCT with neurotransmitters, and OCTN/OCTN-related with ergothioneine and carnitine derivatives. Our data suggest that the SLC22 family can work among itself, as well as with other ADME genes, to optimize levels of numerous metabolites and signaling molecules, as proposed by the Remote Sensing and Signaling Theory**.**

1.2 INTRODUCTION

The SLC (solute carrier) gene family includes 65 families with over 400 transporter genes. In humans, 52 of these families are expressed, encompassing more than 395 genes and it has been estimated that ~2000 (10% of the genome) human genes are transporter-related [1]. Various solute carrier 22 (SLC22) members are expressed on both the apical and basolateral surfaces of epithelial cells where they direct small molecule transport between body fluids and vital organs, such as the kidney, liver, heart, and brain [2]. SLC22 transporters are also found in circulating cell types such as erythrocytes (e.g. SLC22A7), monocytes, and macrophages (e.g. SLC22A3, SLC22A4, SLC22A15, and SLC22A16) [3], [4]. With recent calls for research on solute carriers, there has been a large influx of data over the past five years, including novel roles in remote sensing and signaling, leading to the need for a more comprehensive understanding of the functional importance of transporters [5].

The SLC22 family is comprised of at least 31 transporters and is found in species ranging from *Arabidopsis thaliana* of the plant kingdom to modern day humans [6], [7]. Knowledge surrounding this family of proteins has expanded greatly since its proposed formation in 1997, when SLC22A6 (OAT1, originally known as novel kidney transporter or NKT) was first cloned [8]. Its homology to SLC22A1 (OCT1) and SLC22A7 (OAT2/NLT) led to the establishment of a new family (SLC22, TC# 2.A.1.19) of transport proteins within the Major Facilitator Superfamily (TC# 2.A.1, MFS) as classified by the IUBMB-approved Transporter Classification (TC) system [8], [9]. These proteins all share 12 α-helical transmembrane domains (TMD), a large extracellular domain (ECD) between TMD1 and TMD2, and a large intracellular domain (ICD) between TMD6 and TMD7 [10]. Research has shown these transporters to be integral participants in the movement of drugs, toxins, and endogenous metabolites and signaling

molecules, such as prostaglandins, urate, α-ketoglutarate, carnitine, and cyclic nucleotides across the cell membrane [11].

As key players in small organic molecule transport, SLC22 members are hypothesized to play a role in the Remote Sensing and Signaling Theory [12]–[15]. The Remote Sensing and Signaling Theory posits that ADME genes – conventionally viewed as central to the absorption (A), distribution (D), metabolism (M), and elimination (E) of drugs, namely drug transporters and enzymes – aid in maintaining homeostasis through remote communication between organs via metabolites and signaling molecules in the blood that may in turn regulate gene expression [16]. This remote communication is supported by the example of serum uric acid levels. In the setting of compromised kidney function, the increase in serum uric acid seems to be partly mitigated through a compensatory increase in the expression and/or function of ABCG2 in the intestine, which allows the excretion of uric acid in the feces rather than the urine [17], [18]. Current research is focusing on determining the ways in which these transporters collaborate to regulate metabolite levels throughout the body [19].

Rather than maintaining a simple division of SLC22 into organic anion transporters (OATs), organic cation transporters (OCTs) and organic zwitterion/cation transporters (OCTNs), previous evolutionary studies have identified six phylogenetic "subclades" – OAT, OAT-like, OAT-related, OCT, OCTN-related, and OCTN – within the OAT and OCT "major clades" [10]. These subclades consist, on average, of three to four members with the exception of the OAT subclade which claims more than half of the 31 known members of SLC22 [10]. Although these subclades are phylogenetically sound, the endogenous functions of many SLC22 members within the six subclades remain ill-defined or unknown. With the emergence of new data, we performed a re-analysis of the SLC22 family to better characterize the functional, endogenous grouping of

these transporters. Our re-analysis shows eight apparent subgroups, with four of these subgroups arising out of the previously defined (but very large) OAT subclade. Because these groupings are more closely related to well-known OATs rather than OCTs, OCTNs or other subclades, we refer to these as OAT subgroups (OATS1, OATS2, OATS3, OATS4).

We considered many factors in our re-analysis of SLC22 and subsequent designation of functionally based subgroups. To better describe the subgroups while still highlighting the nuances of each individual transporter, we utilized data from genomic loci, tissue expression, sequence similarity searches, proteomic motif searches, and functional transporter-metabolite data from GWAS, in vitro assays, and in vivo models. In place of phylogenetic studies, we performed multiple sequence alignments (MSA) and generated guide-trees that are based on sequence similarity or homology and thus provide more insight into function than solely phylogenetic studies. While the SLC22 family is composed of putative transporters, some members, like Slc22a20 and Slc22a17, have proposed mechanisms that differ from those of classic transporters [20], [21]. To that effect, we explored the sequence similarities between SLC22 transporters and non-transport related proteins. We also used systems biology tools to develop an SLC22 transporter-metabolite networks as well as networks for each subgroup. This analysis elucidates the diversity of the endogenous functions of SLC22 transporters in various tissues and provides an updated functional framework for assigning each transporter to a subgroup. Considering the importance of SLC22 transporters, forming functional groups that incorporate endogenous substrates and tissue expression patterns can help better define their roles in intra-organ, inter-organ, and inter-organismal communication

1.3 MATERIALS AND METHODS

Data collection

SLC22 human and mouse sequences were collected from the National Center for Biotechnology Information (NCBI) protein database. Sequences were confirmed and genomic loci were recorded using the University of California Santa Cruz (UCSC) genome browser by searching within each available species on the online platform (https://genome.ucsc.edu/cgi-bin/hgGateway) [76]. The NCBI BLASTp web-based program was used to find similar sequences to those from the NCBI protein database. BLASTp was run with default parameters using query SLC22 sequences from human or mouse. The database chosen was "non-redundant protein sequence" (nr) and no organisms were excluded [77]. Tissue expression of all human SLC22 members was collected from the Human Protein Atlas, GTEx dataset, Illumina Body Map, ENCODE dataset, and RNA-seq datasets available on the EMBL-EBI Expression Atlas (https://www.ebi.ac.uk/gxa/home) [78]. Tissue expression data were also collected via extensive literature search.

Sequence Alignment and Guide-trees

Sequences for SLC22 were aligned using Clustal-Omega with default parameters via the online platform provided by the European Bioinformatics Institute (EMBL-EBI) (https://www.ebi.ac.uk/Tools/msa/clustalo/), as well as MAFFT (Multiple alignment using fast Fourier transform) and ICM-Pro v3.8-7c [79]–[82]. Clustal-Omega, MAFFT and ICM-Pro v3.8-7c produced similar topologies. These alignments were then visualized using The Interactive Tree of Life (http://itol.embl.de/) [83]. Topology was analyzed by branch length values, which are a result of the neighbor-joining method. This method calculates the number of amino acid changes between each organism and the common ancestor from which it branched. It then adopts

the minimum-evolution criteria (ME) by building a tree which minimizes the sum of all branch lengths to visually display relatedness [84]. SSearch36 was utilized to compare representative sequences of all members of the Drug:H+ Antiporter-1 (12 Spanner) (DHA1) Family (2.A.1.2) and the the SLC22A family (2.A.1.19) from the Transporter Classification Database (http://www.tcdb.org/) with the Cyanate Porter (CP) Family (2.A.1.17) as an outgroup to further investigate the belongingness of SLC22A18 in either the SLC22 or DHA1 family [9], [85]. SSearch36 is an exhaustive comparison method that uses the Smith-Waterman (SW) algorithm to compare FASTA files find sequence similarities [85], [86].

ICM-Pro v3.8-7c was used to align sequences in FASTA format as well as perform homology searches of all human and mouse SLC22 sequences against ICM-Pro's curated database of all known proteins [82]. Threshold for homology significance was determined by the probability of structural insignificance (pP), defined as the negative log of the probability value of a homology comparison. Alignments were discarded if the pP value was less than 5.0, indicating that the homology shared between two sequences is likely not due to random sequence similarities.

Motif Analysis

Motif comparisons were performed on the subgroups of the OAT Subclade using the Multiple Expectation-maximum for Motif Elicitation (MEME) (http://meme-suite.org/tools/meme) suite [87]. A threshold of 20 motifs containing a range of 6-20 amino acid length was set with the normal discovery mode. This detection method yielded a set of evolutionarily conserved motifs within all Oat Subclade sequences (n=57) as well as a set of evolutionarily conserved motifs for each of the four proposed Oat subgroups. These motifs were then mapped onto 2D topologies of one member from each of the newly proposed Oat subgroups

(SLC22A6 for OatS1, SLC22A7 for OatS2, SLC22A12 for OatS3, and SLC22A9 for OatS4). A separate motif analysis was also performed for the rodent expansion consisting of slc22a19, and slc22a26-30 and was mapped onto mouse slc22a27. Transmembrane domains (TMDs) of these transporters were predicted by the Constrained Consensus TOPology prediction server (CCTOP) (http://cctop.enzim.ttk.mta.hu/) [88]. TMD locations and the motif locations were entered into TOPO2 (http://www.sacs.ucsf.edu/TOPO2/) to visualize the 2D representation of the transporters with the OAT subclade's evolutionarily conserved motifs shown in blue and each subgroup's evolutionarily conserved motifs shown in red [89].

SNP, mutation, in-vitro, knockout, and drug transport data

To determine the diversity of substrate specificity, the number of drugs that list SLC22 members as a target on DrugBank were recorded [22]. The Metabolomics GWAS server was utilized to determine SNPs within all SLC22 members. The dataset produced by Shin et al. (2014) with the cohort KORA+TwinsUK (blood) and the association of single metabolites was chosen. This dataset was searched by gene symbol (e.g. SLC22A6) [35], [90]. The EMBL GWAS Catalog and Metabolomix's table of all published GWAS with metabolomics (http://www.metabolomix.com/list-of-all-published-gwas-with-metabolomics/) were also utilized in searching for SNPs and their effect on metabolite transport by SLC22 members [91]. Current literature available on the NCBI gene database under Gene References Into Functions (Gene RIFs) (https://www.ncbi.nlm.nih.gov/gene/about-generif) was used to search for non-synonymous mutations that did not affect protein expression yet affected transport of metabolites and/or drugs. These methods were accompanied by an extensive literature search for in-vitro transport and knockout data. Most in-vitro data comes from tissue culture assays from a variety of cell lines while most in-vivo data comes from genetic or chemical knockout mice. Metabolite data was

abstracted from the aforementioned databases and confirmed via scrutinous literature review. The import from table feature on Cytoscape 3.7.2 was used to generate functional networks for the entire SLC22 family and the subgroups [23]. A spring embedded layout was applied to the networks and the subgroups were color coded manually.

1.4 RESULTS

Emerging data continues to indicate the centrality of the SLC22 family (particularly OATs, OCTs, and OCTNs) in endogenous physiology [5], [16]. Our thorough reanalysis of the previously described phylogenetic subclades [10] revealed eight functional subgroups: OATS1, OATS2, OATS3, OATS4, OAT-like, OAT-related, OCT and OCTN/OCTN-related (Table 1). By thus grouping this large family of proteins, we highlighted differences in substrate selectivity, showing that each member has a unique profile of associated metabolites. Based on the number of different metabolites it interacts with, each SLC22 transporter can be classified as relatively mono-, oligo-, or multi-specific. In what follows, publicly available data from GWAS, in vitro, and in vivo datasets were used to build functional networks that support the subgroups (Figure 1). In addition to these functional data and systems biology analyses, subgroups were also supported by structural, genomic, and other analyses explained below. Because some SLC22 members remain understudied, we also investigated low level sequence identity with non-transport proteins to better characterize these "orphaned" transporters.

Table 1. Updated SLC22 Family Subgroups. The SLC22 family was previously separated into 6 phylogenetic subclades. We propose a reclassification into 8 subgroups based on functional data and supported by the methods described in the text.

Former Groupings		Updated Groupings	
Subclade	**Members**	**Subgroup**	**Members**
OAT	A6, A7, A8, A9, A10, A11, A12, A19, A20, A22, A24, A25, A26, A27, A28, A29, A30	OATS1	A6, A8, A20
		OATS2	A7
		OATS3	A11, A12, a22
		OATS4	A9, A10, A24, A25
OAT-like	A13, A14	OAT-like	A13, A14
OAT-related	A17, A18. A23, A31	OAT-related	A17, A18. A23, A31
OCTN-related	A15, A16	OCTN/OCTN related	A4, A5, A15, A16, a21
OCTN	A4, A5		
OCT	A1, A2, A3	OCT	A1, A2, A3

Analysis of Substrate Specificity and Selectivity Helps Categorize Mono-, Oligo-, and Multi-Specificity of SLC22 Members

The concept of multi-, oligo-, and mono-specific SLC22 transporters was supported in part based on the number of unique drugs that are known to interact with each SLC22 member (engelhart_01_SLC22_specificity.xlsx) [22]. Several SLC22 members (e.g., OAT1, OCT2) are best known as "drug" transporters and due to this association, many have been extensively tested as potential drug targets. While drugs are not the primary focus of this research, the number of drugs a transporter is linked to is indicative of how many structurally different substrates it can interact with. This may translate to endogenous compounds from different metabolic pathways. As interest in solute carriers has increased over the past decade, there has been a large influx of functional data (engelhart_06_supplementary_materials.zip). We used these data to validate our initial specificity assignments, and found that, for the most part, the metabolite data were in agreement with the drug data. A transporter linked to many unique drugs was often linked to many

unique metabolites. For example, OATS1 members SLC22A6 and SLC22A8 are linked to 100 or more drugs, respectively. This is reflected in the metabolite data, as each transporter was associated with at least 50 unique metabolites, confirming their multi-specific nature. OATS4 members SLC22A9, SLC22A10, SLC22A24, and SLC22A5 are understudied with respect to drugs. As a group, they are only associated with three drugs, making it difficult to predict their substrate selectivity. Endogenously, the group appears to have relatively mono-specific members that are dedicated to conjugated sex steroids, and oligo-specific members which are linked to conjugated sex hormones, short chain fatty acids, and bile acids.

Construction of Functional Networks from Metabolite-Transporter Interaction Data Support the Eight Subgroups

To visualize these transporter-metabolite interactions, which were acquired from a combination of GWAS, in vivo, and in vitro studies, we created networks using Cytoscape [23]. These networks allowed us to see the extent of unique and overlapping substrate specificity between transporters in the SLC22 family and within the proposed subgroups engelhart_06_supplementary_materials.zip). The networks also provide, for the first time, a systems biology lens into subgroup (as opposed to single transporter) function. In these networks, all edges are undirected and represent a statistically significant result linking an SLC22 member to a metabolite. To give an example, the OATS1 network uses the members (SLC22A6, SLC22A8, SLC22A20) as central nodes. Each associated metabolite is connected to the member, and the networks are then combined to represent the entire subgroup and demonstrate how a metabolite may be linked to multiple transporters (engelhart_06_supplementary_materials.zip). Functional data were available for 21 of 31 known SLC22 transporters. The trimmed SLC22 network is displayed in Figure 1, the individual

subgroup networks are in engelhart_06_supplementary_materials.zip, and the total SLC22 network is in engelhart_06_supplementary_materials.zip. The compiled data with transporter, metabolite, study, quantitative metric, and citation are present in engelhart_06_supplementary_materials.zip.

While there is no single metabolite that is associated with all SLC22 transporters, some are linked to multiple family members, and thus may be a hallmark of the subgroup or family as a whole. These metabolites are prostaglandin E2, prostaglandin F2, estrone sulfate, uric acid, carnitine and creatinine, which are each linked to at least five different SLC22 members, respectively (engelhart_06_supplementary_materials.zip). This result demonstrated that SLC22, as a group, is involved in regulating several metabolic processes, ranging from blood vessel dilation through prostaglandins to cellular energy production through carnitine [24], [25]. This also implies that the particular structural features of the SLC22 family in general (12 TMD, large ECD between TMD1 and TMD2, and large ICD between TMD6 and TMD7) lends itself well to interacting with these compounds. This is further supported by the subgroup-specific network analyses and motif analysis we performed (Figure 1).

In Figure 1, which has removed metabolites linked to only one transporter, 24 SLC22 proteins are linked to 79 unique metabolites, highlighting the physiological relevance of this family. This representation also brings attention to the number of shared substrates among SLC22, with 222 total edges present in the trimmed network. The multi-specific, oligo-specific, mono-specific nature of the family members suggests that these proteins may be able to compensate for the reduced function of another member. Furthermore, several of the metabolites interacting with the proteins (prostaglandins, carnitine derivatives, bile acids) belong to different

metabolic pathways, indicating that many processes, at both the systemic and cellular level, are dependent upon SLC22.

OATS1(SLC22A6, SLC22A8, SLC22A20) Handles A Wide Variety of Metabolites, Signaling Molecules, Uremic Toxins, and Odorants

Several metabolites have been identified as substrates of SLC22A6 (OAT1) and SLC22A8 (OAT3). While many are unique, there is notable overlap. Both OAT1 and OAT3 interact with uremic toxins (indoxyl sulfate, p-cresol sulfate, uric acid) and gut microbiome derived products (indolelactate, 4-hydroxyphenylacetate), as well as many of the more general SLC22 metabolites, like prostaglandin E2, prostaglandin F2, uric acid, and creatinine [26]–[30]. SLC22A20 (OAT6), while not as well-studied, has affinity for several odorants and short chain fatty acids that are also associated with OAT1 [31]. OAT1 and OAT3 are clearly multi-specific, and OAT6 appears to be oligo-specific, as it handles both odorants and some short chain fatty acids. With respect to remote signaling, the shared metabolites among these transporters (engelhart_06_supplementary_materials.zip) are noteworthy because of their tissue localization (engelhart_02_SLC22_tissue.xlsx). OAT1 and OAT3 are primarily expressed in the kidney proximal tubule, with some expression in other tissues, like the choroid plexus and retina (engelhart_02_SLC22_tissue.xlsx). OAT6, however, is expressed in the olfactory mucosa of mice, presumably reflecting its affinity for odorants [21], [31], [32]. In the kidney, OAT1 and OAT3, along with many other SLC22 transport proteins, help regulate the urine levels of many metabolites and signaling molecules which may potentially facilitate inter-organismal communication. For example, a volatile compound in one organism may be excreted into the urine through OAT1 and then somehow sensed by another individual of the same or different species through a mechanism involving OAT6 in the olfactory mucosa [12].

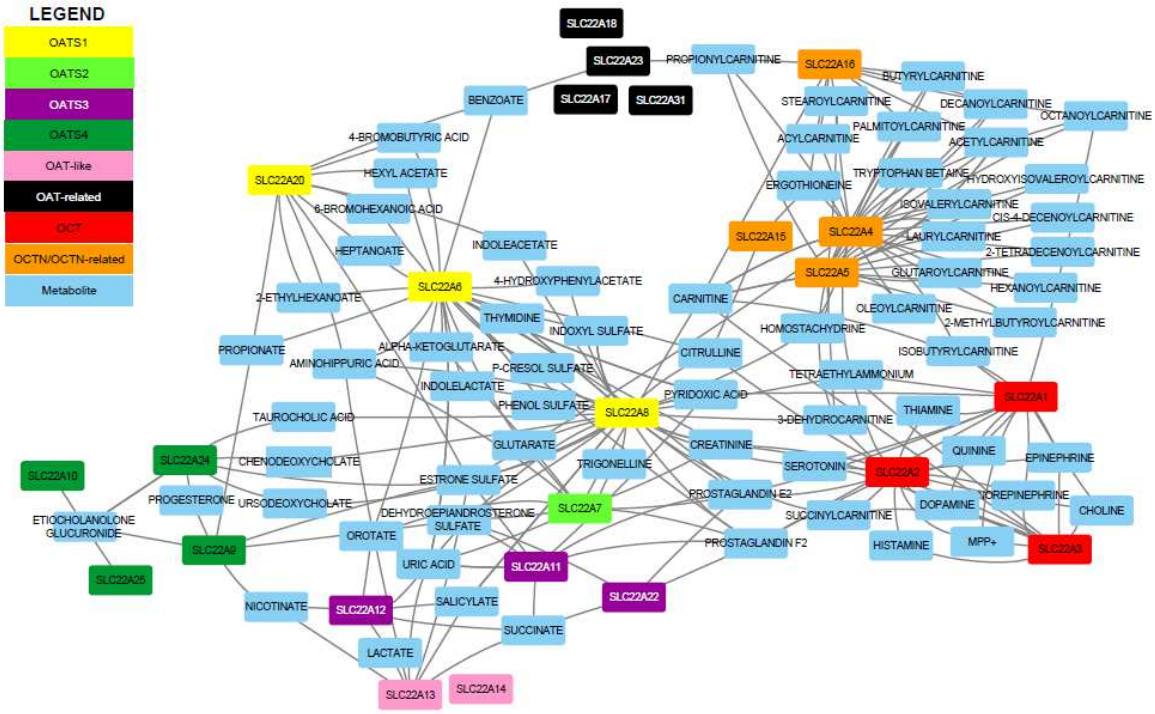


Figure 1. Pruned SLC22 Network. All SLC22 transporters with functional data were initially included. Metabolites associated with only one transporter were removed for improved visualization. SLC22 transporters and metabolites are colored nodes. Each edge represents a significant transporter-metabolite association. Multiple edges connecting one metabolite to a specific transporter were bundled (e.g., in vitro and GWAS support).

OATS2 (SLC22A7) is a Systemically-Expressed Transporter of Organic Anions and Cyclic Nucleotides

SLC22A7 (OAT2) is the only member of the OATS2 subgroup and is associated with prototypical SLC22 substrates, such as prostaglandins, carnitine, creatinine, and uric acid [29], [33]–[35]. Evolutionarily, OAT2 is an obvious single member subgroup with a distinct branching pattern and single common ancestor within our generated guide trees (Figure 2 engelhart_06_supplementary_materials.zip). OAT2 is also linked to cyclic nucleotides and dicarboxylic acids, which when taken with the previous metabolites, creates a unique profile worthy of its own subgroup (engelhart_06_supplementary_materials.zip) [36]. Another distinguishing feature of OAT2 is its tissue expression patterns (engelhart_02_SLC22_tissue.xlsx). While its expression in the liver and kidney are common to

many SLC22 members, it has been localized to circulating red blood cells, where it may function in cyclic nucleotide transport [3]. Its expression in a mobile cell type and transport of cyclic nucleotides raises the possibility that it may act as an avenue for signaling.

OATS3 (SLC22A11, SLC22A12, Slc22a22) Functions to Balance Uric Acid and Prostaglandins

In humans, SLC22A11 (OAT4) and SLC22A12 (URAT1) share only two substrates, uric acid and succinate (engelhart_06_supplementary_materials.zip) [37], [38]. Uric acid is a beneficial metabolite in the serum as it is thought to be responsible for more than half of human antioxidant activity in the blood [39]. However, high levels of uric acid can be harmful and are associated with gout [40]. URAT1 is associated with very few metabolites and is best understood for its role in uric acid reabsorption in the kidney proximal tubule, making it relatively mono-specific [37]. OAT4, on the other hand, has been shown to transport prostaglandins and conjugated sex hormones in addition to uric acid, making it oligo-specific [41]–[43]. URAT1 is almost exclusively expressed in the kidney, and OAT4 is expressed in the kidney, placenta, and epididymis (engelhart_02_SLC22_tissue.xlsx). The more diverse tissue expression of SLC22A11 seems consistent with its wider range of substrates. The subgroup differs in rodents because mice do not express Oat4. Instead, the rodent subgroup is composed of Slc22a12, known as renal-specific transporter (Rst) in mice, and Slc22a22, known as prostaglandin-specific organic anion transporter (Oat-pg). While Rst and Oat-pg do not share substrate specificity, together, they combine to play the role of URAT1 and OAT4 by handling uric acid and prostaglandins [44].

OATS4 (SLC22A9, SLC22A10, SLC2A24, SLC22A25) Members are Specifically Associated with Conjugated Sex Hormones

GWAS analyses support the association of all human members of this subgroup with one common metabolite, etiocholanolone glucuronide, a conjugated sex hormone, with a p-value of $4.12x10^{-27}$ or lower for all members (engelhart_02_SLC22_tissue.xlsx, engelhart_06_supplementary_materials.zip) [45]. While this group shares at least one conjugated sex hormone, SLC22A24 and SLC22A9 appear to be more oligo-specific transporters, with SLC22A9 linked to short chain fatty acids and SLC22A24 linked to bile acids [46], [47]. SLC22A10 and SLC22A25 are only linked to conjugated sex hormones, making them relatively mono-specific transporters (engelhart_06_supplementary_materials.zip) [45]. In terms of tissue expression, there is a distinct correlation between patterns and shared function amongst human OATS4 members (engelhart_03_OATS4.xlsx). We predict that all four members are conjugated sex steroid transporters with SLC22A9, A10, and A25 showing high expression in the liver where conjugation of glucuronides and sulfates to androgens and other gonadal steroids occurs [46]. SLC22A24 has low expression levels in the liver but is highly expressed in the proximal tubule, where it is predicted to reabsorb these conjugated steroids [46]. This subgroup also includes a large rodent-specific expansion, consisting of Slc22a19 and Slc22a26-30. Although the rodent-specific expansion is greatly understudied, transport data for rat Slc22a9/a24 shows shared substrate specificity for estrone sulfate with SLC22A24, but not for bile acids or glucuronidated steroids, which is consistent with the lack of glucuronides in rat urine and serum [46]. While sulfatases are extremely highly conserved amongst humans, rats, and mice, the separation of rodent- and nonrodent-specific OATS4 groups may be due to the species differences in expression and function of glucuronidases [48]. Despite their distinct differences

from human OATS4 members in sequence similarity studies and minimal functional data, the rodent-specific transporters are also highly expressed in both liver and kidney [49].

OAT-Like (SLC22A13, SLC22A14) has Potentially Physiologically Important Roles

Very little functional data is available for the OAT-like subgroup. SLC22A13 (OAT10/ORCTL3) has been well characterized as a transporter of both urate and nicotinate, but SLC22A14 has no available transport data [50]. However, N′-methyl nicotinate is increased in the plasma levels of self-reported smokers, and GWAS studies have implicated SNPs in the SLC22A14 gene to be associated with success in smoking cessation [51], [52]. Although this data does not directly relate SLC22A14 to nicotinate, it suggests a possible route of investigation into the functional role of this transporter, one that may, in some ways, overlap with that of OAT10. SLC22A13 is primarily expressed in the kidney, and although we found no human protein expression data for SLC22A14, transcripts for this gene are found at low levels in the kidney and notably high levels the testis (engelhart_02_SLC22_tissue.xlsx), which is in concordance with its critical role in sperm motility and fertility in male mice [53]. Future studies are required to determine the functional classification of this subgroup; however, our genomic localization and sequence-based analyses provide enough data to support the notion that these two belong in their own individual subgroup.

OAT-Related (SLC22A17, SLC22A18, SLC22A23, SLC22A31) is Anomalous Amongst SLC22 Members but has Interesting Functional Mechanisms and Disease Associations

The OAT-related subgroup is an outlier within the SLC22 family, consisting of the orphan transporters SLC22A17, SLC22A18, SLC22A23, and SLC22A31. SLC22A17 and SLC22A23 are strongly related, with greater than 30% shared amino acid identity. When these two transporters were initially identified together as BOCT1 (SLC22A17) and BOCT2

(SLC22A23), it was noted that they both show high expression levels in the brain, as well as a nonconserved amino terminus that may negate prototypical SLC22 function [54]. SLC22A17 is known as LCN2-R (Lipocalin receptor 2) and is reported to mediate iron homeostasis through binding and endocytosis of iron-bound lipocalin, as well as exhaustive protein clearance from the urine as shown by high affinities for proteins such as calbindin [20], [55]. SLC22A23 has no confirmed substrates, but SNPs and mutations within this gene have medically-relevant phenotypic associations such as QT elongation, inflammatory bowel disease, endometriosis-related infertility, and the clearance of antipsychotic drugs [56]–[58]. SLC22A31 is the most understudied transporter of the SLC22 family but has been associated with right-side colon cancer [59]. SLC22A18 remains an outlier and lacks the characteristic SLC22 large ECD. Its membership within the SLC22 family is arguable due to high sequence similarity with the DHA H^+-antiporter family (engelhart_06_supplementary_materials.zip) [10]. Further study is required to confirm if the OAT-related members share substrates or if their sequence diversity and deviations from classical physical SLC22 member characteristics are the reason for their phylogenetic association.

OCT (SLC22A1, SLC22A2, SLC22A3) Members are Characteristic Organic Cation Transporters with High Affinities for Monoamine Neurotransmitters and Other Biologically Important Metabolites and Signaling Molecules

The OCT subclade of SLC22A1 (OCT1), SLC22A2 (OCT2), and SLC22A3 (OCT3) has ample data to support its formation and has been widely accepted and utilized as the prototypical subgroup of organic cation transporters. All three members of this subgroup transport monoamine neurotransmitters, carnitine derivatives, creatinine and the characteristic OCT substrates, MPP+ and TEA (engelhart_06_supplementary_materials.zip) [35], [60]–[64]. All

three members of this subgroup are expressed in the liver, kidney, and brain (engelhart_02_SLC22_tissue.xlsx), When considered together with the transport of neurotransmitters, this subgroup serves as an example of inter-organ communication between the brain and the kidney-liver axis via transporters. The systemic levels of these neurotransmitters and thus, their availability to the brain can be regulated by the expression of OCT subgroup members in the liver, where the metabolites can be enzymatically modified, and expression in the kidney, which may serve as an excretory route [7].

OCTN/OCTN-Related (SLC22A4, SLC22A5, SLC22A15, SLC22A16) Subgroup Consists of Prototypical Carnitine and Ergothioneine Transporters

The OCTN/OCTN-Related subgroup is a combination of two previously established subclades, OCTN and OCTN-related [10]. Previous studies have mistakenly named SLC22A15 as CT1 (carnitine transporter 1), but this name actually belongs to SLC22A5 (OCTN2) [13]. GWAS data show that SLC22A4 (OCTN1), SLC22A5 (OCTN2/CT1), and SLC22A16 (FLIPT2/CT2) are heavily linked to carnitine and its derivatives [35]. This is consistent with in vitro data showing that OCTN2 and FLIPT2 are carnitine transporters [65], [66]. Although OCTN1 has lower affinity for carnitine than OCTN2 and FLIPT2, it has high affinity for the endogenous antioxidant ergothioneine, which GWAS data suggests may be a shared metabolite with both SLC22A15 (FLIPT1) and FLIPT2 (SLC22A16) (engelhart_06_supplementary_materials.zip) [35], [67]. SLC22A15 is associated with many complex lipids that are not characteristic of any other SLC22 transporter [45]. Although data is very limited, this anomalous SLC22 member so far appears to only share one potential substrate with this subgroup, but its inclusion is supported by multiple sequence alignments focusing on the intracellular loop and tissue expression patterns. Most other subgroups in this family are

limited to a few tissues, mainly the liver and kidney, but the members of the OCTN/OCTN-Related subgroup are all expressed in at least five tissues as well as circulating immune cells (engelhart_02_SLC22_tissue.xlsx) [4], [7]. This broad tissue expression pattern, in conjunction with our network analysis, supports the notion that these transporters' main task is transporting carnitine derivatives, as carnitine metabolism is an energy producing mechanism in nearly every cell. It may also play a role in regulating levels of the antioxidant ergothioneine, which appears to be a unique substrate of this subgroup [24], [68].

Multiple Sequence Alignment Further Supports the Classification of Subgroups

Our new subgroupings are primarily based on the endogenous function of the transporters, but they are also supported by additional analyses. These analyses are necessary, as structural and evolutionary similarities can predict functional traits that have yet to be discovered. Though the previously established phylogenetic subclades remain sound, our re-analysis includes new and updated amino acid sequences that support the proposed subgroups with more confidence, especially when investigating similarities within functional regions [10]. MSA programs were favored over phylogenetics because MSA searches are based upon structural similarities rather than evolutionary relatedness [69]. These structural similarities, especially in the large ECD (extracellular domain) and large ICD (intracellular domain) of SLC22 proteins, may indicate shared function.

Full length sequence analysis via Clustal-Omega, MAFFT, and ICM-Pro v3.8-7 supports the division of SLC22 into eight subgroups (Figure 2A, engelhart_06_supplementary_materials.zip). While the OATS1, OATS2, OATS4, OAT-like, OAT-related, and OCT subgroups are supported by full-length sequence analyses, OATS3 and OCTN/OCTN-Related required a more rigorous investigation. To further clarify "borderline" subgroup assignments from the full-length

sequence analysis, sequence similarity between the ECDs and ICDs of all human and mouse SLC22 members was determined using ICM-Pro v3.8-7, and the results were visualized via guide trees (Figure 2B, Figure 2C). ECD alignment preserved all eight subgroups, with the exception of SLC22A15 in the OCTN/OCTN-Related subgroup. In contrast, ICD alignment preserved only the OATS4, OATS2, and OCT subgroups.

The branching pattern of OATS3 member Oat-pg (Slc22a22) differs between tree variations. These analyses consistently indicate a similar relationship between Oat-pg and OATS3, as well as OATS4. However, in an analysis of the SLC22 ECDs, it is most closely associated with OATS3 over any other subgroup. This, in conjunction with shared substrate specificity with both SLC22A12 and SLC22A11, and not OATS4 members, supports its membership within the OATS3 subgroup [29], [37], [38], [44].

In full-length sequence alignments, the grouping of SLC22A4, SLC22A5, and Slc22a21 is consistently conserved, while the topology of both SLC22A15 and SLC22A16 is irregular. Despite this, analysis of the large ECD shows similarity between all OCTN/OCTN-related members other than SLC22A15. Previous analyses have noted the large difference between the ECD of SLC22A15 and all other SLC22 members which is supported by our analysis in Figure 2B [10]. Interestingly, there appears to be some similarity between the large intracellular domains of SLC22A16 and SLC22A15. Although much of the support for the establishment of the OCTN/OCTN-related subgroup comes from functional data (engelhart_06_supplementary_materials.zip), the described MSA analyses highlight shared structural, and possibly functional, regions.

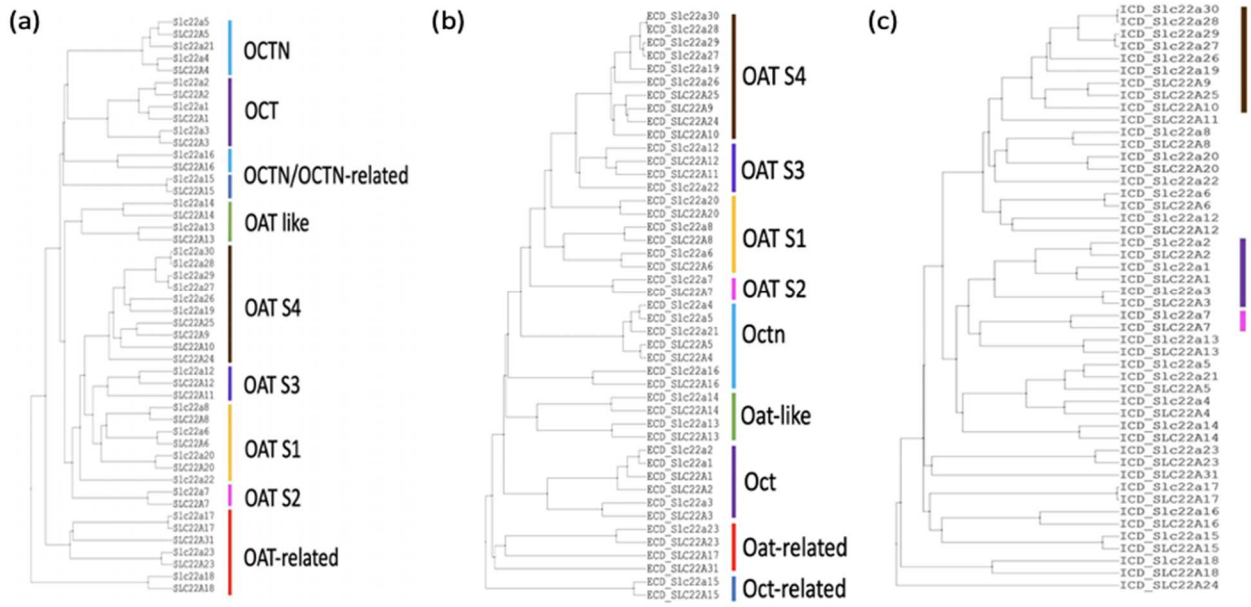


Figure 2. Multiple sequence alignment using ICM-Pro v3.8-7c tree of SLC22 members implies function. All known mouse and human SLC22 sequences, excluding Slc22a18, were aligned using ICM-Pro v3.8-7c sequence similarity-based alignment. A) Full sequence. B) Extracellular loop (not including Slc22a18, due to its lack of a characteristic large extracellular loop between TMD1 and TMD2). C) Intracellular loop.

Analysis of Genomic Localization Highlights Evolutionary Relatedness of Subgroup Members and Suggests Basis of Coregulation

Genomic clustering within the SLC22 family has been previously described [10]. Specifically, genes found in tandem on the chromosome, such as OAT-like members SLC22A13 and SLC22A14, are predicted to have arisen from duplication events, indicating a strong evolutionary relationship. Despite the majority of the OAT subclade being found on chromosome 11 in humans and chromosome 19 in mice, clustering within the chromosome supports the division of the OAT subclade into smaller subgroups. For example, OATS4 members SLC22A9, SLC22A10, SLC22A24, and SLC22A25 appear in tandem on human chromosome 11 within the UST (Unknown Substrate Transporter) region of the genome. This region is analogous to the UST region within the mouse genome on chromosome 19, where the mouse-specific OATS4 members Slc22a19, Slc22a26, Slc22a27, Slc22a28, Slc22a29, and

Slc22a30 reside as well as the rat UST region on chromosome 1 that contains Slc22a9/a24, Slc22a9/a25, Ust4r and Ust5r (engelhart_02_SLC22_tissue.xlsx) [70], [71]. It has been proposed that genes within clusters, to some degree, are coordinately regulated and thus are predicted to have similar overall tissue expression patterns [70], [72], [73]. Support for shared regulatory mechanisms of subgroup members within genomic clusters can be inferred from similar patterns of tissue expression or by expression of subgroup members along a common axis of metabolite transport such as the gut-kidney-liver axis. Genomic localization from the UCSC Genome browser and resultant tissue expression patterns for all SLC22 members are shown in engelhart_02_SLC22_tissue.xlsx.

Analysis of OAT Subgroup Specific Motifs Highlight Patterns Potentially Involved in Specificity

Motif analyses revealed subgroup specific motifs within functionally important regions, such as the large ICD, large ECD, and the region spanning TMD9 and TMD10, for all novel OAT subgroups [10], [74]. However, the number of unique residues appears to be correlated to the range of substrate specificity.

Of the newly proposed OAT subgroups, OATS2 claims the smallest number of subgroup-specific amino acid motifs and is the only subgroup without a specific motif in TMD9 (Figure 3B). The lack of multiple subgroup-specific regions is interesting not only because this subgroup consists of a single transporter but also because this may be indicative of a more promiscuous transporter with a wide range of substrates, which is substantiated by the functional data as described earlier in "*OATS2 (SLC22A7) is a Systemically-Expressed Transporter of Organic Anions and Cyclic Nucleotides*". This pattern is also seen in OATS1, which consists of multi- and oligo-specific transporters OAT1, OAT3, and OAT6. In addition to having few subgroup-specific motifs, the multi/oligo-specific nature of this subgroup is reflected by the shared

evolutionary conservation of the large extracellular domain with other OAT subclade members (Figure 3A).

To further clarify the membership of Oat-pg in OATS3, evolutionarily conserved motifs were determined between all three members, as well as just Slc22a11 and Slc22a12. This analysis revealed a total of ten evolutionarily conserved amino acid motifs between all three members, eight of which are present in the analysis of only OAT4 and URAT1 (Figure 3, engelhart_06_supplementary_materials.zip). Specifically, both analyses exhibited a notably large motif in the large intracellular loop found at D313-Q332 on URAT1 and Q312-G331 on OAT-PG (Figure 3C, Figure 3D). This larger number of conserved regions seems consistent with a more limited range of substrates (eg. uric acid and prostaglandins) [41].

Motif analysis was performed separately on the OATS4 rodent and non-rodent specific subgroups and the entirety of the OATS4 subgroup members. In all analyses, OATS4 claims the largest number of evolutionarily conserved and subgroup-specific amino acid residues amongst the OAT subgroups, supporting selective substrate specificity, possibly for conjugated sex steroids (Figure 3E, Figure 3F). In the case of non-rodent transporters, a unique motif spans the sixth extracellular domain and TMD12. This region is predicted to govern substrate specificity of transporters of the MFS, to which the SLC22 family belongs [74]. Recent publications defining the substrate specificity of SLC22A24 point to a more narrow range of substrates and conservation of this specific region amongst OATS4 members may explain the association of conjugated steroid hormones with SLC22A9, SLC22A10, SLC22A24, and SLC22A25 in GWAS studies [45], [46]. Although further analysis is required to fully understand the relationship between structure and substrate specificity in SLC22 transporters, we provided a basis for

investigation into specific regions that may determine functional patterns. The sequences and p-values for each motif are in engelhart_06_supplementary_materials.zip.

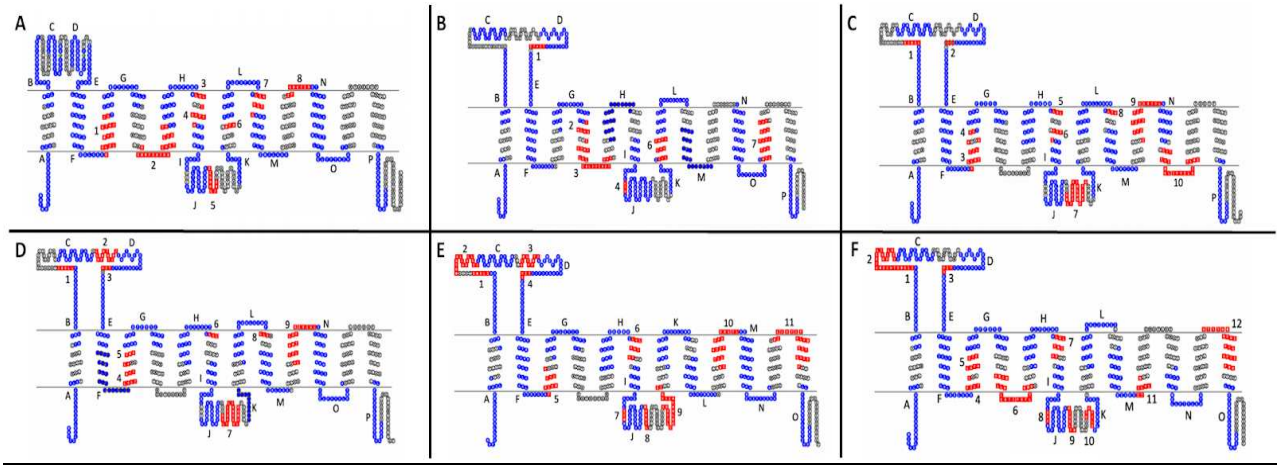


Figure 3. Evolutionarily conserved motifs for each subgroup within the OAT major subgroup mapped onto 2D topology of prototypical members. A) OATS1 mapped onto SLC22A6 (OAT1). B) OATS2 mapped onto hSLC22A7 (OAT2). C) OATS3 mapped onto hSLC22a12 (URAT1). D) OATS3 mapped onto mSlc22a22 (OAT-PG). E) OATS4 mapped onto hSLC22A9 (OAT7). F) OATS4 mapped onto mSlc22a7. In each panel, red sequences are subgroup specific motifs, blue sequences are OAT-major subgroup motifs, and green diamonds represent non-synonymous SNPs that affect serum metabolite levels. Conserved OAT-major subgroup motifs are assigned letters and specific, conserved OAT subgroup motifs are numbered. Data, including motif sequence identities, exact locations, and p-values can be found in engelhart_06_supplementary_materials.zip.

Sequence Similarity Study Suggests Novel Potential Functions To Explore and Possible Tertiary Structure of SLC22

Each SLC22 member is a putative transporter, but there is evidence that suggests some members may have alternative mechanisms of action [31], [55]. To further explore this possibility and to potentially find sequence similarity to other proteins, the specific amino acid sequences for the extracellular and intracellular loops of each SLC22 member were compared to all proteins in the ICM-Pro v3.8-7c database. The extracellular loop of mouse Slc22a16 shares 26% sequence identity (pP=5.4) with chicken beta-crystallin B3 (CRBB3). Beta-crystallin is a structural protein mainly comprised of beta sheets [75]. The similarity between the ECD of mouse Slc22a16 and CRBB3 could point to potential for a beta sheet like configuration. Because

none of the SLC22 family members have been crystallized, any insight into tertiary structure is of interest.

SLC22A31, a member of the divergent OAT-Related subclade, is the most ambiguous member of the SLC22 family with no functional data available. An investigation of the human SLC22A31 large ECD shows at least 30% shared sequence identity with RNA-binding protein 42 (RBM42) in mouse, rat, cow and human. This analysis also showed a 37% sequence identity (pP=5.5) shared between the ECD of hSLC22A31 and human heterochromatinization factor BAHD1. These and other interesting sequence similarities to proteins, including those involved in signaling, are noted in engelhart_04_ICM_homology.xlsx.

1.5 DISCUSSION

In the years following the establishment of the previous SLC22 subclades, there has been a notable increase in functional data, particularly with respect to endogenous substrates, concerning these transporters and their substrates [10]. With these data, we are now in a position to better characterize the biology of these transporters, which play important physiological roles and are implicated in certain diseases. However, our newly proposed subgroups are not entirely dependent on functional data, as we have considered multiple approaches including phylogenetics, multiple sequence alignments, evolutionarily conserved motifs, sequence homology, and both tissue and genomic localization. Each of these approaches has individual value in that they reveal unique characteristics of each transporter; yet it is the combination of multiple approaches that ensures the full variety of available data (though still incomplete) for these transporters is considered when forming functional subgroups. We support the subgroups with a thorough literature search of metabolites associated with SLC22 proteins.

Although the functional data is inherently biased due to the high level of interest in some SLC22 members, particularly the "drug" transporters OAT1, OAT3, OCT1, and OCT2, for the majority of the transporters, there is enough data to create functional subgroups that play distinct and overlapping roles in metabolism (Figure 1, engelhart_06_supplementary_materials.zip). Genomic localization reveals evolutionary information and provides insight on how genes may arise from duplication events. Phylogenetic analysis determines the evolutionary relatedness of these proteins, while MSA, motif analysis, and sequence homology focus on structural similarities, which can be indicative of function. We often see that members of a subgroup are expressed in the same tissues or along functional axes. For example, substrates transported from the liver via SLC22 transporters (e.g., SLC22A1, OCT1) can be either excreted into or retrieved

from the urine by other SLC22 members (SLC22A2, OCT2) of the same subgroup. Establishment of these functional subgroups may also inform future virtual screenings for metabolites of understudied transporters.

Protein families are established based on shared ancestry and structural similarity, which is commonly considered grounds for shared functionality. This is exemplified amongst SLC22 members with the generally shared structural characteristics of 12 TMDs, a large extracellular loop between TMD1 and TMD2, and a smaller intracellular loop between TMD6 and TMD7. Despite these shared features, we show here that there are many functional differences between these transporters. Although our analyses mostly align with previous evolutionary studies when considering ancestry, here, we show that phylogenetic grouping is not always reflective of similar structure and function. For example, although the previously established OCTN subclade of SLC22A4, SLC22A5 and Slc22a21 does not share common ancestry with Slc22a16, the newly proposed group shares functional similarity and ECD homology. Thus, by expanding our investigation beyond phylogenetic relationships, we can now more appropriately group proteins from the same family and better understand their roles in endogenous physiology.

An important concept in the Remote Sensing and Signaling Network is that of multi-specific, oligo-specific, and relatively mono-specific transporters working in a coordinated function [16]. Multi-specific transporters are able to interact with a wide variety of structurally different compounds, oligo-specific with a smaller variety, and relatively mono-specific transporters are thought to interact with only one or a few substrates. Existing functional data suggests that it is unlikely that any truly mono-specific transporters exist within the SLC22 family, yet the different subgroups we have formed imply that multi-specific, oligo-specific, and relatively mono-specific transporters are more likely to form subgroups with transporters that

share substrate specificity. Multi-specific transporters, like those in the OATS1 and OCT subgroups, handle a diverse set of drugs, toxins, endogenous metabolites, and signaling molecules [14], [60]. Conversely, the OATS4 subgroup appears to be a collection of relatively mono-specific transporters with an affinity for conjugated sex steroid hormones, specifically etiocholanolone glucuronide, which is also supported by a recent study focused on SLC22A24, a member of the OATS4 subgroup [45], [46]. Previous evolutionary studies have suggested that multi-specific transporters arose before the mono-specific transporters [10]. As evolution has progressed, more specific transporters have developed to handle the burden of changing metabolism. The multi-specific transporters have been more extensively characterized because of their importance in pharmaceuticals, but in the case of endogenous metabolic diseases, the oligo and mono-specific transporters may be more appropriate targets for drugs or therapies.

One of the best examples of multi-specific transporters working in concert with oligo, and mono-specific transporters is the regulation of uric acid [17], [18]. Handling of uric acid mainly occurs in the kidney, but when renal function is compromised, multi-specific transporters regulate their expression to compensate. Two proteins, SLC22A12 and SLC2A9, are expressed in the proximal tubule and are nearly exclusively associated with uric acid. The multi-specific transporters SLC22A6 and SLC22A8 are also present in the proximal tubule and are able to transport uric acid. When the kidney is damaged, one would expect serum uric acid levels to increase because most of the proteins involved in its elimination are in the kidney. However, this is partly mitigated due to the increased expression of ABCG2 and/or functional activity in the intestine [17], [18]. SLC2A9 is a relatively mono-specific transporter and ABCG2 (BCRP) is a multi-specific ABC transporter, and other uric acid transporters can be considered oligo-specific (eg. SLC22A11). The example of uric acid serves to illustrate how, when certain mono-, oligo-,

and multi-specific transporters are unable to perform their primary function, multi-specific transporters of the same or different function (even of the ABC superfamily) can use their shared substrate specificity to mitigate the consequences. It is generally assumed that all SLC22 family members are transporters. However, Slc22a17, a member of the outlier Oat-related subclade, functions as an endocytosed iron-bound lipocalin receptor and some SLC22 members have been suggested to function as "transceptors" due to homology with GPCR odorant receptors and shared odorant substrates [20], [21]. Thus, to better understand the SLC22 family members' individually unique functions and their placement into subgroups/subclades, we compared the full-length amino acid sequences, large ECDs, and large ICDs of all SLC22 family members to a database of known proteins.

When considering such a large number of proteins, the function on both local and systemic levels of metabolites is likely to be impacted. The SLC22 family is a central hub of coexpression for ADME (absorption, distribution, metabolism, excretion)-related genes in non-drug treated conditions, which underscores their importance in regulating endogenous metabolism through the transport of small molecules [16]. In the context of the Remote Sensing and Signaling Theory (RSST), it is essential to understand substrate specificity of different SLC22 members and the 8 subgroups.

The RSST proposes that a network of ADME genes (drug transporters, drug metabolizing enzymes, and various regulation proteins) regulates the levels of hundreds if not thousands of small organic molecules with "high informational content" including key metabolites and signaling molecules involved in intra-organ, inter-organ, and inter-organismal remote communication. The RSST would seem to imply that organisms are constantly solving a multi-objective optimization problem, where balancing each particular compound's serum

concentration represents a single objective. Each compound present in the blood has a range of healthy concentrations, and when the concentration is outside of that range, the body must address it, in part through the regulation of transporters and enzymes. Due to their wide range of tissue expression and diverse functional roles at body fluid interfaces, the particular combination of transporters and enzymes are critical variables necessary for solving this multi-objective optimization problem. Transporters regulate the entry and exit of substrates to and from cells, but enzymes are responsible for the altering of these compounds. To use a simple hypothetical example, if a metabolite's serum concentration is too high, a transporter with high affinity for that metabolite can move it into the cell, where an enzyme with high affinity for the substrate can change it so that it may re-enter the circulation or be more readily cleared from the body. The existence of abundant multi-specific and oligo-specific transporters and enzymes, in addition to relatively mono-specific ones, expressed differentially in tissues and at body fluid interfaces, allows for a highly flexible and responsive complex adaptive system that not only maintains homeostasis in blood, tissue, and body fluid compartments (e.g. CSF), but also helps restore it after acute or chronic perturbations.

Thus, together, transporters and enzymes have tremendous potential to manage levels of metabolites, signaling molecules, and antioxidants in the circulation and in specific tissues. By developing functional groupings for the SLC22 family, we can better understand the metabolic networks in which they function and how their expression is utilized to regulate concentrations of metabolites, signaling molecules (e.g. cyclic nucleotides, prostaglandins, short chain fatty acids, and sex steroids), antioxidants (ergothioneine, uric acid), and other molecules affecting diverse aspects of homeostasis (e.g. lipocalin). Although this analysis focuses on the SLC22

family, a similar approach can be applied to develop a deeper understanding of other families of transporters and enzymes.

In the past, the majority of functional data has come from transport assays using cells overexpressing a specific SLC22 transporter and a single metabolite of interest. These assays lack uniformity and, as the OAT knockouts have shown, are not necessarily reflective of endogenous physiology [26], [28], [30]. Recently, GWAS studies have linked many metabolites to polymorphisms in SLC22 genes, and in vivo metabolomic studies using knockout models have also identified several metabolites that may be substrates of transporters [26], [28], [30], [45]. In upcoming years, the integration of multiple types of omics data related to SLC22 family members with functional studies of transporters and evolutionary analyses will likely produce a more fine-grained picture of the roles of these and other transporters in inter-organ and inter-organismal Remote Sensing and Signaling.

Chapter 1, in full, has been submitted for publication of the material as it may appear in International Journal of Molecular Sciences, 2020, Engelhart, Darcy C.; Granados, Jeffry C.; Shi, Da; Saier, Milton H.; Baker, Michael E.; Abagyan, Reuben; Nigam, Sanjay K., 2020. The thesis author was the primary investigator and author of this paper.

1.6 REFERENCES

[1] Hediger, M. A.; Clémençon, B.; Burrier, R. E.; Bruford, E. A. The ABCs of Membrane Transporters in Health and Disease [SLC Series]: Introduction. *Molecular Aspects of Medicine*. April 2013, pp 95–107. https://doi.org/10.1016/j.mam.2012.12.009.

[2] Lamhonwah, A. M.; Hawkins, C. E.; Tam, C.; Wong, J.; Mai, L.; Tein, I. Expression Patterns of the Organic Cation/Carnitine Transporter Family in Adult Murine Brain. *Brain Dev.* 2008, *30* [1], 31–42. https://doi.org/10.1016/j.braindev.2007.05.005.

[3] Sager, G.; Smaglyukova, N.; Fuskevaag, O. M. The Role of OAT2 [SLC22A7] in the Cyclic Nucleotide Biokinetics of Human Erythrocytes. *J. Cell. Physiol.* 2018, *233* [8], 5972–5980. https://doi.org/10.1002/jcp.26409.

[4] Minuesa, G.; Purcet, S.; Erkizia, I.; Molina-Arcas, M.; Bofill, M.; Izquierdo-Useros, N.; Casado, F. J.; Clotet, B.; Pastor-Anglada, M.; Martinez-Picado, J. Expression and Functionality of Anti-Human Immunodeficiency Virus and Anticancer Drug Uptake Transporters in Immune Cells. *J. Pharmacol. Exp. Ther.* 2008, *324* [2], 558–567. https://doi.org/10.1124/jpet.107.131482.

[5] César-Razquin, A.; Snijder, B.; Frappier-Brinton, T.; Isserlin, R.; Gyimesi, G.; Bai, X.; Reithmeier, R. A.; Hepworth, D.; Hediger, M. A.; Edwards, A. M.; A Call for Systematic Research on Solute Carriers. *Cell*. Elsevier: Austria July 30, 2015, pp 478–487. https://doi.org/10.1016/j.cell.2015.07.022.

[6] Lelandais-Brière, C.; Jovanovic, M.; Torres, G. A. M.; Perrin, Y.; Lemoine, R.; Corre-Menguy, F.; Hartmann, C. Disruption of AtOCT1, an Organic Cation Transporter Gene, Affects Root Development and Carnitine-Related Responses in Arabidopsis. *Plant J.* 2007, *51* [2], 154–164. https://doi.org/10.1111/j.1365-313X.2007.03131.x.

[7] Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; Tissue-Based Map of the Human Proteome. *Science [80-.]*. 2015, *347* [6220]. https://doi.org/10.1126/science.1260419.

[8] Lopez-Nieto, C. E.; You, G.; Bush, K. T.; Barros, E. J. G.; Beier, D. R.; Nigam, S. K. Molecular Cloning and Characterization of NKT, a Gene Product Related to the Organic Cation Transporter Family That Is Almost Exclusively Expressed in the Kidney. *J. Biol. Chem.* 1997, *272* [10], 6471–6478. https://doi.org/10.1074/jbc.272.10.6471.

[9] Saier, M. H.; Reddy, V. S.; Tamang, D. G.; Västermark, Å. The Transporter Classification Database. *Nucleic Acids Res.* 2014, *42* [D1]. https://doi.org/10.1093/nar/gkt1097.

[10] Zhu, C.; Nigam, K. B.; Date, R. C.; Bush, K. T.; Springer, S. A.; Saier, M. H.; Wu, W.; Nigam, S. K. Evolutionary Analysis and Classification of OATs, OCTs, OCTNs, and Other SLC22 Transporters: Structure-Function Implications and Analysis of Sequence Motifs. *PLoS One* 2015, *10* [11], 140569. https://doi.org/10.1371/journal.pone.0140569.

[11] Nigam, S. K. The SLC22 Transporter Family: A Paradigm for the Impact of Drug Transporters on Metabolic Pathways, Signaling, and Disease. *Annu. Rev. Pharmacol. Toxicol.* 2018, *58* [1], 663–687. https://doi.org/10.1146/annurev-pharmtox-010617-052713.

[12] Ahn, S. Y.; Nigam, S. K. Toward a Systems Level Understanding of Organic Anion and Other Multispecific Drug Transporters: A Remote Sensing and Signaling Hypothesis. *Molecular Pharmacology*. September 2009, pp 481–490. https://doi.org/10.1124/mol.109.056564.

[13] Wu, W.; Dnyanmote, A. V.; Nigam, S. K. Remote Communication through Solute Carriers and ATP Binding Cassette Drug Transporter Pathways: An Update on the Remote Sensing and Signaling Hypothesis. *Molecular Pharmacology*. May 2011, pp 795–805. https://doi.org/10.1124/mol.110.070607.

[14] Nigam, S. K.; Bush, K. T.; Martovetsky, G.; Ahn, S. Y.; Liu, H. C.; Richard, E.; Bhatnagar, V.; Wu, W. The Organic Anion Transporter [OAT] Family: A Systems Biology Perspective. *Physiol. Rev.* 2015, *95* [1], 83–123. https://doi.org/10.1152/physrev.00025.2013.

[15] Nigam, S. K. What Do Drug Transporters Really Do? *Nature Reviews Drug Discovery*. Nature Publishing Group January 1, 2014, pp 29–44. https://doi.org/10.1038/nrd4461.

[16] Rosenthal, S. B.; Bush, K. T.; Nigam, S. K. A Network of SLC and ABC Transporter and DME Genes Involved in Remote Sensing and Signaling in the Gut-Liver-Kidney Axis. *Sci. Rep.* 2019, *9* [1], 11879. https://doi.org/10.1038/s41598-019-47798-x.

[17] Bhatnagar, V.; Richard, E. L.; Wu, W.; Nievergelt, C. M.; Lipkowitz, M. S.; Jeff, J.; Maihofer, X.; Nigam, S. K. Analysis of ABCG2 and Other Urate Transporters in Uric Acid Homeostasis in Chronic Kidney Disease: Potential Role of Remote Sensing and Signaling. *Clin. Kidney J.* 2016, *9* [3], 444–453. https://doi.org/10.1093/ckj/sfw010.

[18] Yano, H.; Tamura, Y.; Kobayashi, K.; Tanemoto, M.; Uchida, S. Uric Acid Transporter ABCG2 Is Increased in the Intestine of the 5/6 Nephrectomy Rat Model of Chronic Kidney Disease. *Clin. Exp. Nephrol.* 2014, *18* [1], 50–55. https://doi.org/10.1007/s10157-013-0806-8.

[19] Nigam, S. K.; Bhatnagar, V. The Systems Biology of Uric Acid Transporters: The Role of Remote Sensing and Signaling. *Current Opinion in Nephrology and Hypertension*. Lippincott Williams and Wilkins July 1, 2018, pp 305–313. https://doi.org/10.1097/MNH.0000000000000427.

[20] Langelueddecke, C.; Roussa, E.; Fenton, R. a; Wolff, N. a; Lee, W. K.; Thévenod, F. Lipocalin-2 [24p3/Neutrophil Gelatinase-Associated Lipocalin [NGAL]] Receptor Is Expressed in Distal Nephron and Mediates Protein Endocytosis. *J. Biol. Chem.* 2012, *287* [1], 159–169. https://doi.org/10.1074/jbc.M111.308296.

[21] Kaler, G.; Truong, D. M.; Sweeney, D. E.; Logan, D. W.; Nagle, M.; Wu, W.; Eraly, S.

A.; Nigam, S. K. Olfactory Mucosa-Expressed Organic Anion Transporter, Oat6, Manifests High Affinity Interactions with Odorant Organic Anions. *Biochem. Biophys. Res. Commun.* 2006, *351* [4], 872–876. https://doi.org/10.1016/j.bbrc.2006.10.136.

[22] Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* 2018, *46* [D1], D1074–D1082. https://doi.org/10.1093/nar/gkx1037.

[23] Smoot, M. E.; Ono, K.; Ruscheinski, J.; Wang, P. L.; Ideker, T. Cytoscape 2.8: New Features for Data Integration and Network Visualization. *Bioinformatics* 2011, *27* [3], 431–432. https://doi.org/10.1093/bioinformatics/btq675.

[24] Bremer, J. Carnitine. Metabolism and Functions. *Physiological Reviews*. 1983, pp 1420–1480. https://doi.org/10.1152/physrev.1983.63.4.1420.

[25] Blumberg, A. L.; Denny, S. E.; Marshall, G. R.; Needleman, P. Blood Vessel Hormone Interactions: Angiotensin, Bradykinin, and Prostaglandins. *Am. J. Physiol.* 1977, *232* [3]. https://doi.org/10.1152/ajpheart.1977.232.3.h305.

[26] Wikoff, W. R.; Nagle, M. A.; Kouznetsova, V. L.; Tsigelny, I. F.; Nigam, S. K. Untargeted Metabolomics Identifies Enterobiome Metabolites and Putative Uremic Toxins as Substrates of Organic Anion Transporter 1 [Oat1]. *J. Proteome Res.* 2011, *10* [6], 2842–2851. https://doi.org/10.1021/pr200093w.

[27] Vallon, V.; Eraly, S. A.; Rao, S. R.; Gerasimova, M.; Rose, M.; Nagle, M.; Anzai, N.; Smith, T.; Sharma, K.; Nigam, S. K.; A Role for the Organic Anion Transporter OAT3 in Renal Creatinine Secretion in Mice. *Am. J. Physiol. - Ren. Physiol.* 2012, *302* [10]. https://doi.org/10.1152/ajprenal.00013.2012.

[28] Wu, W.; Bush, K. T.; Nigam, S. K. Key Role for the Organic Anion Transporters, OAT1 and OAT3, in the in Vivo Handling of Uremic Toxins and Solutes. *Sci. Rep.* 2017, *7* [1]. https://doi.org/10.1038/s41598-017-04949-2.

[29] Kimura, H.; Takeda, M.; Narikawa, S.; Enomoto, A.; Ichida, K.; Endou, H. Human Organic Anion Transporters and Human Organic Cation Transporters Mediate Renal Transport of Prostaglandins. *J. Pharmacol. Exp. Ther.* 2002, *301* [1], 293–298. https://doi.org/10.1124/jpet.301.1.293.

[30] Bush, K. T.; Wu, W.; Lun, C.; Nigam, S. K. The Drug Transporter OAT3 [SLC22A8] and Endogenous Metabolite Communication via the Gut–Liver– Kidney Axis. *J. Biol. Chem.* 2017, *292* [38], 15789–15803. https://doi.org/10.1074/jbc.M117.796516.

[31] Wu, W.; Bush, K. T.; Liu, H. C.; Zhu, C.; Abagyan, R.; Nigam, S. K. Shared Ligands between Organic Anion Transporters [OAT1 and OAT6] and Odorant Receptors. *Drug Metab. Dispos.* 2015, *43* [12], 1855–1863. https://doi.org/10.1124/dmd.115.065250.

[32] Monte, J. C.; Nagle, M. A.; Eraly, S. A.; Nigam, S. K. Identification of a Novel Murine

Organic Anion Transporter Family Member, OAT6, Expressed in Olfactory Mucosa. *Biochem. Biophys. Res. Commun.* 2004, *323* [2], 429–436. https://doi.org/10.1016/j.bbrc.2004.08.112.

[33] Shen, H.; Liu, T.; Morse, B. L.; Zhao, Y.; Zhang, Y.; Qiu, X.; Chen, C.; Lewin, A. C.; Wang, X. T.; Liu, G.; Characterization of Organic Anion Transporter 2 [SLC22A7]: A Highly Efficient Transporter for Creatinine and Species-Dependent Renal Tubular Expression. *Drug Metab. Dispos.* 2015, *43* [7], 984–993. https://doi.org/10.1124/dmd.114.062364.

[34] Köttgen, A.; Albrecht, E.; Teumer, A.; Vitart, V.; Krumsiek, J.; Hundertmark, C.; Pistis, G.; Ruggiero, D.; O'Seaghdha, C. M.; Haller, T.; Genome-Wide Association Analyses Identify 18 New Loci Associated with Serum Urate Concentrations. *Nat. Genet.* 2013, *45* [2], 145–154. https://doi.org/10.1038/ng.2500.

[35] Shin, S. Y.; Fauman, E. B.; Petersen, A. K.; Krumsiek, J.; Santos, R.; Huang, J.; Arnold, M.; Erte, I.; Forgetta, V.; Yang, T. P.; An Atlas of Genetic Influences on Human Blood Metabolites. *Nat. Genet.* 2014, *46* [6], 543–550. https://doi.org/10.1038/ng.2982.

[36] Sun, W.; Wu, R. R.; Van Poelje, P. D.; Erion, M. D. Isolation of a Family of Organic Anion Transporters from Human Liver and Kidney. *Biochem. Biophys. Res. Commun.* 2001, *283* [2], 417–422. https://doi.org/10.1006/bbrc.2001.4774.

[37] Enomoto, A.; Kimura, H.; Chairoungdua, A.; Shigeta, Y.; Jutabha, P.; Cha, S. H.; Hosoyamada, M.; Takeda, M.; Sekine, T.; Igarashi, T.; Molecular Identification of a Renal Urate-Anion Exchanger That Regulates Blood Urate Levels. *Nature* 2002, *417* [6887], 447–452. https://doi.org/10.1038/nature742.

[38] Hagos, Y.; Stein, D.; Ugele, B.; Burckhardt, G.; Bahn, A. Human Renal Organic Anion Transporter 4 Operates as an Asymmetric Urate Transporter. *J. Am. Soc. Nephrol.* 2007, *18* [2], 430–439. https://doi.org/10.1681/ASN.2006040415.

[39] Maxwell, S. R. J.; Thomason, H.; Sandler, D.; Leguen, C.; Baxter, M. A.; Thorpe, G. H. G.; Jones, A. F.; Barnett, A. H. Antioxidant Status in Patients with Uncomplicated Insulin-Dependent and Non-Insulin-Dependent Diabetes Mellitus. *Eur. J. Clin. Invest.* 1997, *27* [6], 484–490. https://doi.org/10.1046/j.1365-2362.1997.1390687.x.

[40] Sakiyama, M.; Matsuo, H.; Shimizu, S.; Nakashima, H.; Nakayama, A.; Chiba, T.; Naito, M.; Takada, T.; Suzuki, H.; Hamajima, N.; A Common Variant of Organic Anion Transporter 4 [OAT4/SLC22A11] Gene Is Associated with Renal Underexcretion Type Gout. *Drug Metab. Pharmacokinet.* 2014, *29* [2], 208–210.

[41] Cha, S. H.; Sekine, T.; Kusuhara, H.; Yu, E.; Kim, J. Y.; Kim, D. K.; Sugiyama, Y.; Kanai, Y.; Endou, H. Molecular Cloning and Characterization of Multispecific Organic Anion Transporter 4 Expressed in the Placenta. *J. Biol. Chem.* 2000, *275* [6], 4507–4512. https://doi.org/10.1074/jbc.275.6.4507.

[42] Tomi, M.; Eguchi, H.; Ozaki, M.; Tawara, T.; Nishimura, S.; Higuchi, K.; Maruyama, T.;

Nishimura, T.; Nakashima, E. Role of OAT4 in Uptake of Estriol Precursor 16α-Hydroxydehydroepiandrosterone Sulfate into Human Placental Syncytiotrophoblasts from Fetus. *Endocrinology* 2015, *156* [7], 2704–2712. https://doi.org/10.1210/en.2015-1130.

[43] Skwara, P.; Schömig, E.; Gründemann, D. A Novel Mode of Operation of SLC22A11: Membrane Insertion of Estrone Sulfate versus Translocation of Uric Acid and Glutamate. *Biochem. Pharmacol.* 2017, *128*, 74–82. https://doi.org/10.1016/j.bcp.2016.12.020.

[44] Shiraya, K.; Hirata, T.; Hatano, R.; Nagamori, S.; Wiriyasermkul, P.; Jutabha, P.; Matsubara, M.; Muto, S.; Tanaka, H.; Asano, S.; A Novel Transporter of SLC22 Family Specifically Transports Prostaglandins and Co-Localizes with 15-Hydroxyprostaglandin Dehydrogenase in Renal Proximal Tubules. *J. Biol. Chem.* 2010, *285* [29], 22141–22151. https://doi.org/10.1074/jbc.M109.084426.

[45] Long, T.; Hicks, M.; Yu, H. C.; Biggs, W. H.; Kirkness, E. F.; Menni, C.; Zierer, J.; Small, K. S.; Mangino, M.; Messier, H.; Whole-Genome Sequencing Identifies Common-to-Rare Variants Associated with Human Blood Metabolites. *Nat. Genet.* 2017, *49* [4], 568–578. https://doi.org/10.1038/ng.3809.

[46] Wah Yee, S.; Stecula, A.; Chien, H.-C.; Zou, L.; Feofanova, E. V.; Borselen, V.; Wun Kathy Cheung, K.; Yousri, N. A.; Suhre, K.; Kinchen, J. M.; Unraveling the Functional Role of the Orphan Solute Carrier, SLC22A24 in the Transport of 2 Steroid Conjugates through Metabolomic and Genome-Wide Association Studies 3 4. *PLOS Genet.* 2019, *15* [9], e1008208. https://doi.org/10.1101/648238.

[47] Ho, J. S.; Anzai, N.; Enomoto, A.; He, X.; Do, K. K.; Endou, H.; Kanai, Y. Novel Liver-Specific Organic Anion Transporter OAT7 That Operates the Exchange of Sulfate Conjugates for Short Chain Fatty Acid Butyrate. *Hepatology* 2007, *45* [4], 1046–1055. https://doi.org/10.1002/hep.21596.

[48] Sardiello, M.; Annunziata, I.; Roma, G.; Ballabio, A. Sulfatases and Sulfatase Modifying Factors: An Exclusive and Promiscuous Relationship. *Hum. Mol. Genet.* 2005, *14* [21], 3203–3217. https://doi.org/10.1093/hmg/ddi351.

[49] Papatheodorou, I.; Fonseca, N. A.; Keays, M.; Tang, Y. A.; Barrera, E.; Bazant, W.; Burke, M.; Füllgrabe, A.; Fuentes, A. M. P.; George, N.; Expression Atlas: Gene and Protein Expression across Multiple Studies and Organisms. *Nucleic Acids Res.* 2018, *46* [D1], D246–D251. https://doi.org/10.1093/nar/gkx1158.

[50] Bahn, A.; Hagos, Y.; Reuter, S.; Balen, D.; Brzica, H.; Krick, W.; Burckhardt, B. C.; Sabolic, I.; Burckhardt, G. Identification of a New Urate and High Affinity Nicotinate Transporter - Human Organic Anion Transporter 10 [HOAT10, SLC22A131. *Journal of Biological Chemistry*. p .

[51] Uhl, G. R.; Liu, Q. R.; Drgon, T.; Johnson, C.; Walther, D.; Rose, J. E.; David, S. P.; Niaura, R.; Lerman, C. Molecular Genetics of Successful Smoking Cessation: Convergent Genome-Wide Association Study Results. *Arch. Gen. Psychiatry* 2008, *65* [6], 683–693. https://doi.org/10.1001/archpsyc.65.6.683.

[52] Uhl, G. R.; Drgon, T.; Li, C. Y.; Johnson, C.; Liu, Q. R. Smoking and Smoking Cessation in Disadvantaged Women: Assessing Genetic Contributions. *Drug and Alcohol Dependence*. October 1, 2009, pp S58-63. https://doi.org/10.1016/j.drugalcdep.2009.03.012.

[53] Maruyama, S. Y.; Ito, M.; Ikami, Y.; Okitsu, Y.; Ito, C.; Toshimori, K.; Fujii, W.; Yogo, K. A Critical Role of Solute Carrier 22a14 in Sperm Motility and Male Fertility in Mice. *Sci. Rep.* 2016, *6*. https://doi.org/10.1038/srep36468.

[54] Bennett, K. M.; Liu, J.; Hoelting, C.; Stoll, J. Expression and Analysis of Two Novel Rat Organic Cation Transporter Homologs, SLC22A17 and SLC22A23. *Mol. Cell. Biochem.* 2011, *352* [1–2], 143–154. https://doi.org/10.1007/s11010-011-0748-y.

[55] Wan-Jie, C.; Wei-Yi, O.; Gavin, D. Lipocalin 2 and Lipocalin 2 Receptor in Neuron-Glia Interactions Following Injury. *Front. Cell. Neurosci.* 2016, *10*. https://doi.org/10.3389/conf.fncel.2016.36.00027.

[56] Zhao, Z. Z.; Croft, L.; Nyholt, D. R.; Chapman, B.; Treloar, S. A.; Hull, M. L.; Montgomery, G. W. Evaluation of Polymorphisms in Predicted Target Sites for Micro RNAs Differentially Expressed in Endometriosis. *Mol. Hum. Reprod.* 2011, *17* [2], 92–103. https://doi.org/10.1093/molehr/gaq084.

[57] Serrano León, A.; Amir Shaghaghi, M.; Yurkova, N.; Bernstein, C. N.; El-Gabalawy, H.; Eck, P. Single-Nucleotide Polymorphisms in SLC22A23 Are Associated with Ulcerative Colitis in a Canadian White Cohort. *Am. J. Clin. Nutr.* 2014, *100* [1], 289–294. https://doi.org/10.3945/ajcn.113.080549.

[58] Åberg, K.; Adkins, D. E.; Liu, Y.; McClay, J. L.; Bukszár, J.; Jia, P.; Zhao, Z.; Perkins, D.; Stroup, T. S.; Lieberman, J. A.; Genome-Wide Association Study of Antipsychotic-Induced QTc Interval Prolongation. *Pharmacogenomics J.* 2012, *12* [2], 165–172. https://doi.org/10.1038/tpj.2010.76.

[59] Bien, S. A.; Su, Y. R.; Conti, D. V.; Harrison, T. A.; Qu, C.; Guo, X.; Lu, Y.; Albanes, D.; Auer, P. L.; Banbury, B. L.; Genetic Variant Predictors of Gene Expression Provide New Insight into Risk of Colorectal Cancer. *Hum. Genet.* 2019, *138* [4], 307–326. https://doi.org/10.1007/s00439-019-01989-8.

[60] Amphoux, A.; Vialou, V.; Drescher, E.; Brüss, M.; La Cour, C. M.; Rochat, C.; Millan, M. J.; Giros, B.; Bönisch, H.; Gautron, S. Differential Pharmacological in Vitro Properties of Organic Cation Transporters and Regional Distribution in Rat Brain. *Neuropharmacology* 2006, *50* [8], 941–952. https://doi.org/10.1016/j.neuropharm.2006.01.005.

[61] Busch, A. E.; Quester, S.; Ulzheimer, J. C.; Gorboulev, V.; Akhoundova, A.; Waldegger, S.; Lang, F.; Koepsell, H. Monoamine Neurotransmitter Transport Mediated by the Polyspecific Cation Transport ROCT1. *FEBS Lett.* 1996, *395* [2–3], 153–156. https://doi.org/10.1016/0014-5793[96]01030-7.

[62] Gründemann, D.; Liebich, G.; Kiefer, N.; Köster, S.; Schömig, E. Selective Substrates for Non-Neuronal Monoamine Transporters. *Mol. Pharmacol.* 1999, *56* [1], 1–10. https://doi.org/10.1124/mol.56.1.1.

[63] Urakami, Y.; Kimura, N.; Okuda, M.; Inui, K. I. Creatinine Transport by Basolateral Organic Cation Transporter HOCT2 in the Human Kidney. *Pharm. Res.* 2004, *21* [6], 976–981. https://doi.org/10.1023/B:PHAM.0000029286.45788.ad.

[64] Wu, X.; Kekuda, R.; Huang, W.; Fei, Y. J.; Leibach, F. H.; Chen, J.; Conway, S. J.; Ganapathy, V. Identity of the Organic Cation Transporter OCT3 as the Extraneuronal Monoamine Transporter [Uptake2] and Evidence for the Expression of the Transporter in the Brain. *J. Biol. Chem.* 1998, *273* [49], 32776–32786. https://doi.org/10.1074/jbc.273.49.32776.

[65] Enomoto, A.; Wempe, M. F.; Tsuchida, H.; Shin, H. J.; Cha, S. H.; Anzai, N.; Goto, A.; Sakamoto, A.; Niwa, T.; Kanai, Y.; Molecular Identification of a Novel Carnitine Transporter Specific to Human Testis: Insights into the Mechanism of Carnitine Recognition. *J. Biol. Chem.* 2002, *277* [39], 36262–36271. https://doi.org/10.1074/jbc.M203883200.

[66] Ohashi, R.; Tamai, I.; Yabuuchi, H.; Nezu, J. I.; Oku, A.; Sai, Y.; Shimane, M.; Tsuji, A. Na+-Dependent Carnitine Transport by Organic Cation Transporter [OCTN2]: Its Pharmacological and Toxicological Relevance. *J. Pharmacol. Exp. Ther.* 1999, *291* [2], 778–784.

[67] Gründemann, D.; Harlfinger, S.; Golz, S.; Geerts, A.; Lazar, A.; Berkels, R.; Jung, N.; Rubbert, A.; Schömig, E. Discovery of the Ergothioneine Transporter. *Proc. Natl. Acad. Sci. U. S. A.* 2005, *102* [14], 5256–5261. https://doi.org/10.1073/pnas.0408624102.

[68] Cheah, I. K.; Halliwell, B. Ergothioneine; Antioxidant Potential, Physiological Function and Role in Disease. *Biochimica et Biophysica Acta - Molecular Basis of Disease*. May 2012, pp 784–793. https://doi.org/10.1016/j.bbadis.2011.09.017.

[69] Ogden, T. H.; Rosenberg, M. S. Multiple Sequence Alignment Accuracy and Phylogenetic Inference. *Syst. Biol.* 2006, *55* [2], 314–328. https://doi.org/10.1080/10635150500541730.

[70] Wu, W.; Baker, M. E.; Eraly, S. A.; Bush, K. T.; Nigam, S. K. Analysis of a Large Cluster of SLC22 Transporter Genes, Including Novel USTs, Reveals Species-Specific Amplification of Subsets of Family Members. *Physiol. Genomics* 2009, *38* [2], 116–124. https://doi.org/10.1152/physiolgenomics.90309.2008.

[71] Yokoyama, H.; Anzai, N.; Ljubojevic, M.; Ohtsu, N.; Sakata, T.; Miyazaki, H.; Nonoguchi, H.; Islam, R.; Onozato, M.; Tojo, A.; Functional and Immunochemical Characterization of a Novel Organic Anion Transporter Oat8 [Slc22a9] in Rat Renal Collecting Duct. *Cell. Physiol. Biochem.* 2008, *21* [4], 269–278. https://doi.org/10.1159/000129385.

[72] Eraly, S. A.; Monte, J. C.; Nigam, S. K. Novel Slc22 Transporter Homologs in Fly,

Worm, and Human Clarify the Phylogeny of Organic Anion and Cation Transporters. *Physiol. Genomics* 2004, *18* [1], 12–24. https://doi.org/10.1152/physiolgenomics.00014.2004.

[73] Eraly, S. A.; Nigam, S. K. Novel Human CDNAs Homologous to Drosophila Orct and Mammalian Carnitine Transporters. *Biochem. Biophys. Res. Commun.* 2002, *297* [5], 1159–1166. https://doi.org/10.1016/S0006-291X[02]02343-4.

[74] Quistgaard, E. M.; Löw, C.; Guettou, F.; Nordlund, P. Understanding Transport by the Major Facilitator Superfamily [MFS]: Structures Pave the Way. *Nature Reviews Molecular Cell Biology*. 2016, pp 123–132. https://doi.org/10.1038/nrm.2015.25.

[75] Inana, G.; Piatigorsky, J.; Norman, B.; Slingsby, C.; Blundell, T. Gene and Protein Structure of a β-Crystallin Polypeptide in Murine Lens: Relationship of Exons and Structural Motifs. *Nature* 1983, *302* [5906], 310–315. https://doi.org/10.1038/302310a0.

[76] Casper, J.; Zweig, A. S.; Villarreal, C.; Tyner, C.; Speir, M. L.; Rosenbloom, K. R.; Raney, B. J.; Lee, C. M.; Lee, B. T.; Karolchik, D.; The UCSC Genome Browser Database: 2018 Update. *Nucleic Acids Res.* 2018, *46* [D1], D762–D769. https://doi.org/10.1093/nar/gkx1020.

[77] Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* 1990, *215* [3], 403–410. https://doi.org/10.1016/S0022-2836[05]80360-2.

[78] Madeira, F.; Park, Y. mi; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; The EMBL-EBI Search and Sequence Analysis Tools APIs in 2019. *Nucleic Acids Res.* 2019, *47* [W1], W636–W641. https://doi.org/10.1093/nar/gkz268.

[79] Sievers, F.; Higgins, D. G. Clustal Omega for Making Accurate Alignments of Many Protein Sequences. *Protein Sci.* 2018, *27* [1], 135–145. https://doi.org/10.1002/pro.3290.

[80] Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* 2011, *7*. https://doi.org/10.1038/msb.2011.75.

[81] Katoh, K.; Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* 2013, *30* [4], 772–780. https://doi.org/10.1093/molbev/mst010.

[82] Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM—A New Method for Protein Modeling and Design: Applications to Docking and Structure Prediction from the Distorted Native Conformation. *J. Comput. Chem.* 1994, *15* [5], 488–506. https://doi.org/10.1002/jcc.540150503.

[83] Letunic, I.; Bork, P. Interactive Tree Of Life [ITOL] v4: Recent Updates and New

Developments. *Nucleic Acids Res.* 2019, *47* [W1], W256–W259. https://doi.org/10.1093/nar/gkz239.

[84] The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* 1987. https://doi.org/10.1093/oxfordjournals.molbev.a040454.

[85] Reddy, V. S.; Saier, M. H. BioV Suite - A Collection of Programs for the Study of Transport Protein Evolution. *FEBS J.* 2012, *279* [11], 2036–2046. https://doi.org/10.1111/j.1742-4658.2012.08590.x.

[86] Smith, T. F.; Waterman, M. S. Identification of Common Molecular Subsequences. *J. Mol. Biol.* 1981, *147* [1], 195–197. https://doi.org/10.1016/0022-2836[81]90087-5.

[87] Bailey, T. L.; Elkan, C. Fitting a Mixture Model by Expectation Maximization to Discover Motifs in Biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 1994, *2*, 28–36.

[88] Dobson, L.; Reményi, I.; Tusnády, G. E. CCTOP: A Consensus Constrained TOPology Prediction Web Server. *Nucleic Acids Res.* 2015, *43* [W1], W408–W412. https://doi.org/10.1093/nar/gkv451.

[89] Johns, S. J. TOPO2, Transmembrane protein display software http://www.sacs.ucsf.edu/TOPO2/.

[90] Suhre, K.; Shin, S. Y.; Petersen, A. K.; Mohney, R. P.; Meredith, D.; Wägele, B.; Altmaier, E.; Deloukas, P.; Erdmann, J.; Grundberg, E.; Human Metabolic Individuality in Biomedical and Pharmaceutical Research. *Nature* 2011, *477* [7362], 54–62. https://doi.org/10.1038/nature10354.

[91] Kastenmüller, G.; Raffler, J.; Gieger, C.; Suhre, K. Genetics of Human Metabolism: An Update. *Human Molecular Genetics*. October 15, 2015, pp R93–R101. https://doi.org/10.1093/hmg/ddv263.

II.

CHAPTER 2. Drosophila SLC22 orthologs related to OATs, OCTs, and OCTNs regulate development and responsiveness to oxidative stress

2.1 ABSTRACT

The SLC22 family of transporters is widely expressed, evolutionarily conserved, and plays a major role in regulating homeostasis by transporting small organic molecules such as metabolites, signaling molecules, and antioxidants. Analysis of transporters in fruit flies provides a simple yet orthologous platform to study the endogenous function of drug transporters *in vivo*. Evolutionary analysis of *Drosophila melanogaster* SLC22 orthologs reveals that, while many of the 25 SLC22 fruit fly orthologs do not fall within previously established SLC22 subclades, at least four members appear orthologous to mammalian SLC22 members (Slc22a16:*CG6356*, Slc22a15:*CG7458, CG7442* and Slc22a18:*CG3168*). We functionally evaluated the role of SLC22 transporters in *Drosophila melanogaster* by knocking down 13 of these genes. Three SLC22 ortholog knockdowns – Slc22a18 ortholog *CG3168,* Slc22a16 ortholog *CG6356, and CG7084*/SLC22A – did not undergo eclosion and was lethal at the pupa stage, indicating the developmental importance of these genes. Additionally, knocking down four SLC22 members increased resistance to oxidative stress via paraquat testing (*CG4630* $p<0.05$*, CG6006* $p<0.05$*, CG6126* $p<0.01$ *and CG16727* $p<0.05$). Consistent with recent evidence that SLC22 is central to a Remote Sensing and Signaling Network (RSSN) involved in signaling and metabolism, these phenotypes support a key role for SLC22 in handling reactive oxygen species

2.2 INTRODUCTION

SLC (solute carrier) proteins are the second largest family of membrane proteins in the human genome after G protein-coupled receptors (GPCRs) and are relatively understudied given how much of the genome they represent. SLC22 has been identified as a central hub of coexpression with almost every other SLC subfamily and appears to be one of the major hubs of coexpression amongst SLCs, ATP-binding cassette proteins (ABCs), and drug-metabolizing enzymes (DMEs) as well as the predominant hub for coexpression with phase I and phase II DMEs [1], [2]. This central position within coexpression analyses of healthy, non-drug-treated tissues highlights the crucial role that these transporters likely play in endogenous physiology, as proposed in the Remote Sensing and Signaling Theory [3]. To better understand the systemic functionality of the SLC22 family using a highly conserved but simpler model organism than mice, we chose to disrupt this central metabolic hub in *Drosophila*. Our observation of both developmental and oxidative stress phenotypes further underscores the importance of these transporters as developmental regulators and mediators of exogenous stressors.

We utilized *Drosophila melanogaster* as a model system to gain insight into the potential physiological reasons for evolutionary conservation of SLC22 and to investigate their role in mediation of oxidative stress. Evolutionary studies suggest that, in addition to animals, the SLC22 family is conserved in members of the fungi kingdom such as the unicellular eukaryote *S. cerevisiae*, as well as *A. thaliana* of the plant kingdom. However, these species lack physiologically "parallel" systems that could provide insight into the function of human SLC22 transporters [4]. Due to the similarities between *Drosophila melanogaster* physiology and human systems, such as shared functions of the *Drosophila* hindgut and Malpighian tubules and the human intestines and kidneys, the fruit fly serves as a valuable model for human renal and

intestinal disease states. Approximately 65% of human disease-associated genes have putative orthologs in Drosophila and within functional regions, these fly genes can share up to 90% amino acid or DNA sequence identity with their human orthologs [5]. With identification of 25 SLC22 proteins in the *Drosophila* genome and availability of reliable RNAi lines for 17 of these proteins from the Bloomington Drosophila Stock Center, the fruit fly provides a feasible platform for our scientific inquiry [5]–[7].

Although there are no established SLC22 orthologs between fly and human, there is evidence that some SLC22 fly genes share substrates and possibly functionality with human SLC22 members. Two of these genes, *CarT*/carcinine transporter (*CG9317*) and *BalaT*/β-alanine transporter (*CG3790*), play major roles in histamine recycling. In *Drosophila* photoreceptor neurons, *CarT* mediates the uptake of carcinine, an inactive metabolite that results from the conjugation of β-alanine and histamine [8]. Carcinine has been detected in mammalian tissues such as the human intestine and is transported by human OCT2 in both in-vitro and in-vivo studies [8], [9]. *BalaT* mediates the recycling of β-alanine, which is necessary for histamine homeostasis in *Drosophila* photoreceptor synapses [10]. In addition to the imperative role of histamine in *Drosophila* neurotransmission, histamine and histamine receptors (HRs) have broad physiological and regulatory functionality in both the cardiovascular and central nervous systems in which Oct2 and Oct3, which share high homology *CarT* and *Balat,* with are expressed in higher order species such as mice and humans [8], [10]–[13]. This relationship is supported by shared substrate specificity for monoamines, such as carcinine and other neurotransmitters, as well as similar neuronal expression patterns of fly and human genes [9], [14], [15]. Despite *CarT* knockdowns resulting in blindness and complete loss of photoreceptor transmission and *BalaT* knockdowns severely disrupting vision and inhibiting photoreceptor synaptic transmission, both

Oct2 and Oct3 knockout mice show no phenotypic abnormalities [8], [10], [16], [17]. To better utilize *Drosophila* as a model of human SLC22 proteins and direct future studies, a homology-based analysis was performed with all known fruit fly SLC22 orthologs in the frame of well-established SLC22 members from human, mouse, and other common model organisms. This analysis finds at least 10 of the fruit fly orthologs within the previously-defined SLC22 subclades [18]. Four of these fly genes share common ancestry with single SLC22 members which is characteristic of a direct, functional ortholog.

SLC22 members such as OAT1 (Slc22a6), OCT1 (Slc22a1), OAT2 (Slc22a7) are transiently expressed throughout development in tissues that show minimal or no expression in adulthood [19]. Oct1-3, Octn1, Oat1, Oat3 and Urat1 knockout (KO) mice are fertile, viable and show no general phenotypic abnormalities except for the Oat3 KO's decreased blood pressure and the Octn1 KO's increased susceptibility to intestinal inflammation [6], [16], [17], [20]–[25]. The only SLC22 KO mouse line with a reported clear developmental phenotype is the Octn2 KO [24]. Octn2 (Slc22a5) is the main transporter of carnitine in the bodies of both humans and mice and mutations in this gene are associated with systemic carnitine deficiency [26], [27]. The Octn2 knockout line is also referred to as the JVS (juvenile visceral steatosis) line because of the defects in fatty acid oxidation due to carnitine deficiency that results in abnormal accumulation of lipids. Without carnitine supplementation, Octn2 KO mice develop dilated cardiomyopathy, fatty livers and steatosis of other organs, and expire in 3-4 weeks [24]. Although OAT KO's (including URAT1) and OCT KO's have abnormal levels of metabolites and signaling molecules, with only one clear developmental phenotype observed thus far in mice, determining the functional importance of these genes in *Drosophila* could provide insight for orthologous developmental roles in mice and humans given their interesting developmental expression

patterns [19], [28]–[31]. As an initial developmental screen, we created ubiquitous RNAi knockdowns driven by a ubiquitous *da-GAL4* driver of 13 of the SLC22 orthologs and observed their development. A ubiquitous driver was chosen due to the diverse expression patterns of human SLC22 members [32]. Fruit flies have distinct, easily observable developmental stages of which the egg and pupa stage are the most sensitive to environmental stressors and RNAi knockdowns [33], [34]. We show that, of the 13 RNAi knockdowns, three are lethal at the pupa stage, for the first time implicating *Slc22a15*, *Slc22a16*, and *Slc22a18* genes in development.

Paraquat (PQ) resistance tests were performed on ubiquitously expressing knock-down SLC22 lines that progressed to the adult stage. Paraquat is an herbicide and neurotoxicant that is known to cause Parkinson's disease [35]. Low levels of this herbicide can induce redox cycling that yields high levels of reactive oxygen species (ROS), causing systemic oxidative stress [36]. Because of this, it is used as a tool for investigation of acquired resistance to oxidative stress in *Drosophila melanogaster* [37]. As a major contributor to the pathogenesis of a multitude of human diseases, such as cardiovascular disease, metabolic syndrome, neurological disorders, and general cell and tissue degradation associated with aging, oxidative stress, and the mechanisms with which we manage free radicals, are of extreme interest [38]. Ubiquitous RNAi knockdowns of at least some SLC22 members might be predicted to affect resistance to oxidative stress because many SLC22 members transport or affect serum levels of antioxidants. Some examples observed in both mice and humans are OCTN1 (*Slc22a4*) and ergothioneine (EGT), URAT1 (*Slc22a12*) and uric acid, and OAT1/OAT3 (*Slc22a6/a8*) and uric acid as well as TCA (tricarboxylic acid) intermediates such as the oxoacid, α-ketoglutarate and dietary flavonoids [6], [22], [39], [40]. Additionally, carcinine, the characteristic substrate of the fly SLC22 member, *CarT*, is transported by hOCT2 (SLC22A2) and has antioxidant properties [41]. Strikingly, our

studies revealed that ubiquitous RNAi knockdown of four SLC22 genes, resulted in significantly increased oxidative stress resistance at one or more time points.

2.3 MATERIALS AND METHODS

Data collection

SLC22 human and mouse sequences were collected manually from the NCBI protein database. Sea Urchin and C. elegans sequences were collected manually from EchinoBase (http://www.echinobase.org/Echinobase/) and WormBase (https://www.wormbase.org/#012-34-5), respectively[71], [72]. Sequences were confirmed using the UCSC genome browser by searching within each available species on the online platform (https://genome.ucsc.edu/cgi-bin/hgGateway) [73]. The NCBI BLASTp web-based program was used to find sequences similar to those that were searched for manually [74]. BLASTp was run with default parameters using query SLC22 sequences from human or mouse. The database chosen was non-redundant protein sequence (nr), and no organisms were excluded. SLC22 fruit fly orthologs were determined from FlyBase (http://flybase.org/reports/FBgg0000667.html), and sequences were collected manually from the NCBI protein database [75]. Genomic locations of all transporters in question for fruit fly were determined from FlyBase. Drosophila tissue expression data was collected from FlyAtlas (http://flyatlas.org/atlas.cgi) [15].

Phylogenetic analysis

Sequences for SLC22 were aligned using Clustal-Omega (Clustal-W) and MAFFT (Multiple alignment using fast Fourier transform) with default parameters via the online platform provided by the European Bioinformatics Institute (EMBL-EBI) (https://www.ebi.ac.uk/Tools/msa/clustalo/) [76]–[78]. Clustal-W and MAFFT produced similar topologies. These alignments were then visualized using The Interactive Tree of Life (http://itol.embl.de/) [79]. Topology confidence was additionally confirmed by branch length values, which are a result of the neighbor-joining method which calculates the number of amino

acid changes between the organism at the end of the branch and the common ancestor from which it branched to visually display relatedness [80].

Drosophila strains and genetics

Drosophila stocks were fed on standard cornmeal-molasses-yeast diet and kept at room temperature [45]. *Gal4* and RNAi lines were obtained from the Bloomington Drosophila Stock Center (Indiana University, USA) [81]. Ubiquitous RNAi via the GAL4/UAS was used to downregulate the following putative SLC22 transporters: *Balat* (*CG3790*), *CG6231*, *CG4630*, *Orct* (*CG6331*), *Orct2* (EP1027, *CG13610*), *CarT (38E.10*, *CG9317*), *SLC22A* (*CG7442*), *CG14855*, *CG16727*, *CG7333*, *CG8654*, *CG6126*, *CG6006*, *CG7084*, *CG7458*, *CG3168*, *CG6356* [5], [51]. Male SLC22 RNAi stocks were crossed to *da-GAL*4 female virgins to produce an F1 generation with ubiquitous downregulation of the specific SLC22 transporters [82].

RNAi developmental screens and paraquat exposure

F1 offspring were observed from the egg stage through eclosion. Developmental phenotypes were defined as normal development of the F1 generation up until the failure to reach eclosion and surpass the pupa stage. Male F1 flies aged two to seven days after eclosion were tested for paraquat sensitivity as defined by survival. Both parent lines were tested in parallel as controls. Three replicates of 10 flies each were tested per strain. Flies were fed on a 3mm Whatmann paper soaked with 10mM paraquat *(N,N'-dimethyl-4,4'-bipyridinium dichloride*, Sigma) in 10% sucrose. Fresh paraquat was added daily. For the initial 60 hours, the number of dead flies were recorded every 12 hours. All tests were performed at room temperature. In order to avoid unnecessary stress, flies were not starved before adding paraquat. Significance of survival trends was assessed using one-way ANOVA followed by a post hoc Tukey's t-test.

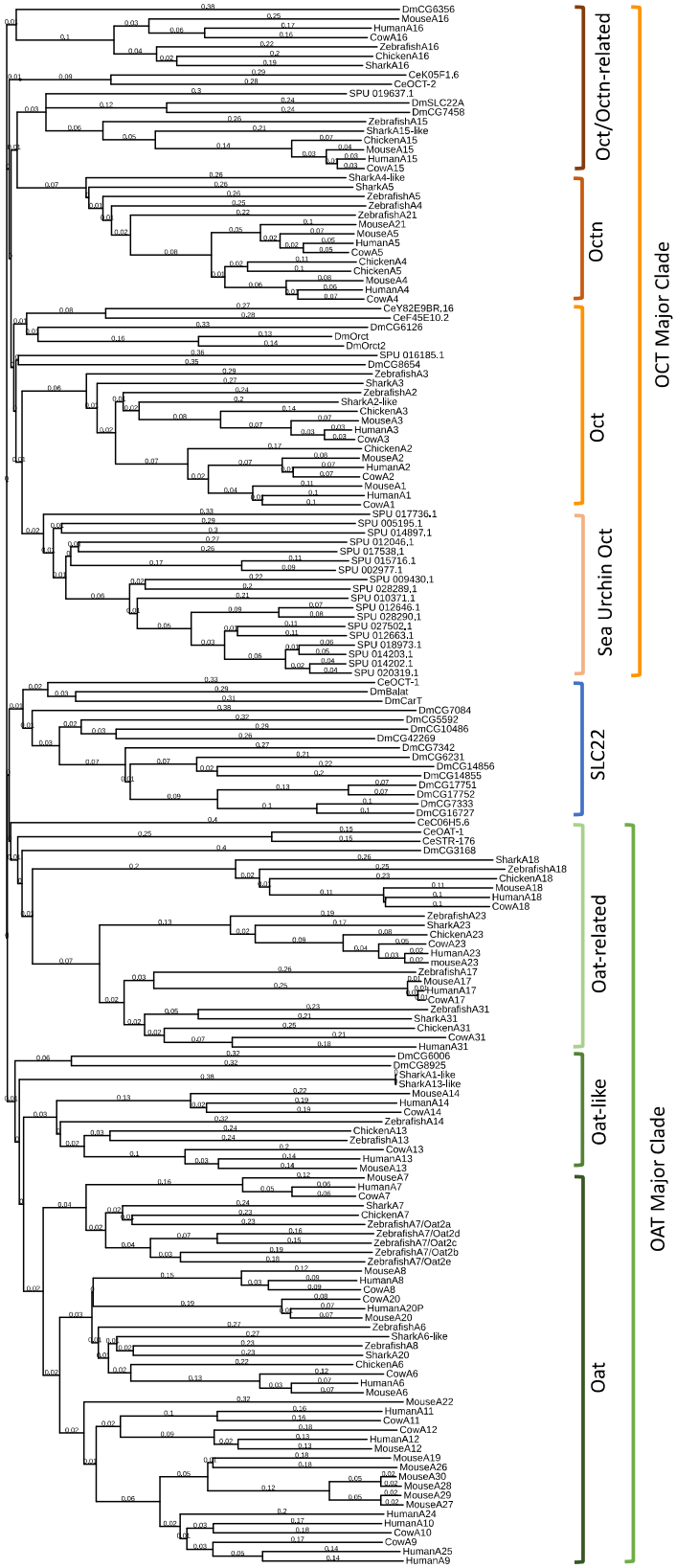
2.4 RESULTS

Drosophila melanogaster SLC22 phylogenetic and genomic analysis

As in mammalian genomes, fly SLC22 genes exist in clusters. The majority of SLC22 genes in Drosophila are found on chromosome 3R with many members found in tandem with other SLC22 orthologs. One notably large cluster consists of 6 SLC22 genes (CG733, CG734, CG17751, Cg17752, CG16727, and CG6231). Inclusion of D. melanogaster orthologous genes in a homology-based analysis of all SLC22 members across a multitude of species (engelhart_05_DM_SLC22_overview.xlsx; Fig 1) resulted in the observation of at least four members that appear orthologous to mammalian SLC22 members (CG6356: Slc22a16, CG7458 and SLC22A/CG7442:Slc22a15 and CG3168:Slc22a18) and an additional six members that can be preliminarily assigned to the individual subclades. The subclades of SLC22 are based on phylogenetic relatedness and functional characterization. They exist within two major clades – OAT and OCT. Although recently revised, [42] the original definitions still stand here: Oat, Oat-like, Oat-related, Oct, Octn, and Oct/Octn-related [17]. When a GUIDANCE 2.0 alignment was performed and all sequences with a GUIDANCE score of <0.6 were removed, the only topology change observed was the omission of all Slc22a18 sequences and the reassignment of CG3168 to the large fly SLC22 transporter group, indicating that it may have sequence homology with Slc22a18, but not the other members of the Oat-related subclade (Fig S1). Interestingly, CG3168 is the only SLC22 ortholog that is localized to the X chromosome in flies. CG6006 and CG8654 fall within the Oat-related subclade and CG6126, CG8654, Orct/CG6331, and Orct2/CG13610 appear to be part of the Oct subclade. The remaining 15 orthologs form their own group outside of the SLC22 subclades and are considered to be mostly organic cation transporters [5], [18]. In summary, in flies, SLC22 appears to have at least some orthologous genes that, based off of

sequence analysis, may prove to be useful models for their relatively understudied human counterparts.

Figure 4. Guide tree of the SLC22 Transporter Family Using 167 Sequences. Sequences from human, mouse, cow, chicken, shark, zebrafish, sea urchin (SPU), C. elegans (Ce), and fruit fly (Dm) were aligned and tree was generated using Clustal Omega (using default parameters). The tree was viewed using Interactive Tree of Life (iTOL). Branch length values are calculated via the Kimura method[44]. Large sea urchin expansion within the Oct Major clade is labeled "Sea Urchin Oct". Sequences that fall between the Oat Major Clade (green) and Oct Major Clade (orange) are denoted as SLC22 (blue).


DmCG6356
MouseA16
HumanA16
CowA16
ZebrafishA16
ChickenA16
SharkA16
CeK05F1.6
CeOCT-2
SPU 019637.1
DmSLC22A
DmCG7458
ZebrafishA15
SharkA15-like
ChickenA15
MouseA15
HumanA15
CowA15
SharkA4-like
SharkA5
ZebrafishA5
ZebrafishA4
ZebrafishA21
MouseA21
MouseA5
HumanA5
CowA5
ChickenA4
ChickenA5
MouseA4
HumanA4
CowA4
CeY82E9BR.16
CeF45E10.2
DmCG6126
DmOrct
DmOrct2
SPU 016185.1
DmCG8654
ZebrafishA3
SharkA3
ZebrafishA2
SharkA2-like
ChickenA3
MouseA3
HumanA3
CowA3
ChickenA2
MouseA2
HumanA2
CowA2
MouseA1
HumanA1
CowA1
SPU 017736.1
SPU 006195.1
SPU 014897.1
SPU 012046.1
SPU 017538.1
SPU 015716.1
SPU 002977.1
SPU 009430.1
SPU 028289.1
SPU 010371.1
SPU 012646.1
SPU 028290.1
SPU 027502.1
SPU 012663.1
SPU 018973.1
SPU 014203.1
SPU 014202.1
SPU 020319.1
CeOCT-1
DmBalat
DmCarT
DmCG7084
DmCG5592
DmCG10486
DmCG42269
DmCG7342
DmCG6231
DmCG14856
DmCG14855
DmCG17751
DmCG17752
DmCG7333
DmCG16727
CeC06H5.6
CeOAT-1
CeSTR-176
DmCG3168
SharkA18
ZebrafishA18
ChickenA18
MouseA18
HumanA18
CowA18
ZebrafishA23
SharkA23
ChickenA23
CowA23
HumanA23
mouseA23
ZebrafishA17
MouseA17
HumanA17
CowA17
ZebrafishA31
SharkA31
ChickenA31
CowA31
HumanA31
DmCG6006
DmCG8925
SharkA1-like
SharkA13-like
MouseA14
HumanA14
CowA14
ZebrafishA14
ChickenA13
ZebrafishA13
CowA13
HumanA13
MouseA13
MouseA7
HumanA7
CowA7
SharkA7
ChickenA7
ZebrafishA7/Oat2a
ZebrafishA7/Oat2d
ZebrafishA7/Oat2c
ZebrafishA7/Oat2b
ZebrafishA7/Oat2e
MouseA8
HumanA8
CowA8
CowA20
HumanA20P
MouseA20
ZebrafishA6
SharkA6-like
ZebrafishA8
SharkA20
ChickenA6
CowA6
HumanA6
MouseA6
MouseA22
HumanA11
CowA11
CowA12
HumanA12
MouseA12
MouseA19
MouseA26
MouseA30
MouseA28
MouseA29
MouseA27
HumanA24
HumanA10
CowA10
CowA9
HumanA25
HumanA9
Oct/Octn-related
Octn
Oct
Sea Urchin Oct
OCT Major Clade
SLC22
Oat-related
Oat-like
Oat
OAT Major Clade

Developmental phenotypes of D. melanogaster SLC22 knockdowns

One of the many advantages of using Drosophila melanogaster as a model organism is its distinct, easily visualized developmental stages. Additionally, RNAi knockdowns of any gene in Drosophila show pupal lethality at a rate of about 15% [34]. These developmental observations provide valuable information regarding the developmental function of orthologous genes that may have compensatory mechanisms in higher order species. Three SLC22 ubiquitous knockdowns (CG3168, CG6356 and CG7084) proved to be lethal at the pupa stage when crossed with the ubiquitous da-GAL4 driver line. Crosses were repeated three times to confirm phenotypes. CG6356 appears to be a direct ortholog of Slc22a16, a carnitine transporter related to OCTNs. In addition, CG3168, which also arrests at the pupa stage is an apparent ortholog of the poorly understood SLC22 member, Slc22a18. To our knowledge, the murine knockouts of these genes have not been reported.

PQ Resistance Test of D. melanogaster SLC22 knockdowns

Paraquat testing is commonly used in *Drosophila* to determine oxidative stress resistance in which increased survival is correlated to increased resistance to oxidative stress [36], [37]. Previous studies have established reliable dose-response curves for paraquat testing in *D. melanogaster* [45]. SLC22 transport proteins in the proximal tubule cells of the kidney take small molecules, such as the antioxidants and SLC22 characteristic substrates uric acid and ergothioneine, into cells to be later excreted [39], [40], [43], [46]. By blocking this route of excretion, levels of these small molecules, such as antioxidants (including dietary flavonoids), are expected to increase in the *Drosophila* hemolymph and increased hemolymph levels of antioxidants would confer resistance to oxidative stress. Through paraquat testing, we show that knocking down SLC22 members in *Drosophila* significantly increases resistance to oxidative

stress at different time points in at least four knock-down lines (*CG4630* $p<0.05$*, CG6006* $p<0.05$*, CG6126* $p<0.01$ *and CG16727* $p<0.05$) when compared to parent and *da-GAL4* control lines (Fig 2, Fig S2-4). The most apparent oxidative stress resistant phenotype is observed for the knock-down of *CG6126*, showing statistically significant increased survival at 36-, 48- and 60-hour time points with 100% survival of the RNAi knockdown flies for all three time points – and an average of about 40% at 36h, 20% at 48h, and 10% at 60h for the parent lines which were used as a control ($p<0.002$). In mice, it is known that SLC22 transporters like Oat1, Oat3, Rst (URAT1), and OCTN1 directly regulate key antioxidants such as uric acid, EGT, flavonoids, and TCA intermediates [6], [25], [47]. Whether or not these fly transporters directly or indirectly regulate redox states will be explored in future studies.

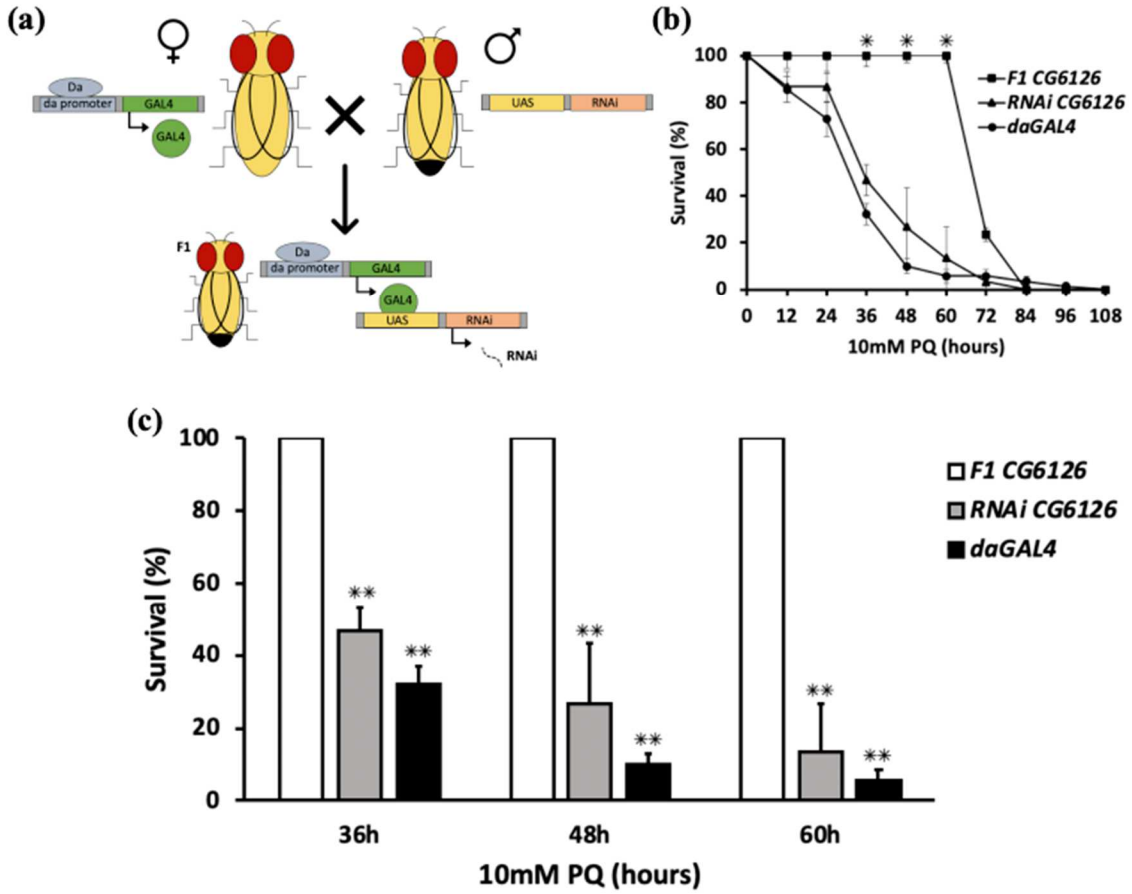


Figure 5. **CG6126 RNAi knockdown shows resistance to oxidative stress.** Parent lines: BDSC Stock 56038 (UAS/RNAi for CG6126), *da-GAL4*. Error bars represent standard error (S.E.). F1 v RNAi p=0.0012051 and F1 v daGAL4 p=0.0010053 at 36h. F1 v RNAi p=0.0010053 and F1 v daGAL4 p=0.0010053 at 48h. F1 v RNAi p=0.0010053 and F1 v daGAL4 p=0.0010053 at 60h. A) Schematic of GAL4/UAS system used to generate RNAi knockdown lines. B) Survival of CG6126 compared to parent lines over 108h. * indicates time point with statistically significant increased percent survival of F1 when compared to both parent lines. C) Survival at points of statistically significant increased survival of CG6126 RNAi knockdown flies compared to parent lines. * $p<0.05$, ** $p<0.01$.

2.5 DISCUSSION

Out of the seven transporters chosen by the International Transporter Consortium and the FDA for evaluation during drug development, three (Oat1, Oat3, and Oct2) are members of the SLC22 family [48]. Additionally, at least 17 SLC22 members are identified as drug transporters by the VARIDT database [49]. In addition to their pharmacological importance, many of these transporters transport metabolites that play a role in the response to endogenous stressors such as oxidative stress induced by reactive oxygen species (ROS). Using *Drosophila melanogaster* as a model organism, we sought to better understand the role of SLC22 in response to oxidative stress and development. Due to the lack of information regarding SLC22 fruit fly orthologs, we attempted to characterize and classify them utilizing multiple sequence alignments and RNAi knockdowns. Because there are minimal developmental phenotypes (apart from SLC22A5) for single SLC22 knockouts in humans or mice despite developmentally interesting and highly dynamic expression patterns, developmental phenotypes observed in *Drosophila* could help further our understanding of how SLC22 contributes to development in other organisms as well [19], [28]–[31]. Prior to our analysis, only three (*Balat, CarT* and *SLC22A*) out of the 25 established fruit fly SLC22 orthologs were functionally investigated beyond global tissue expression screens and homology studies [5], [10], [15], [50]–[52].

Alignment of fly orthologs with the SLC22 family shows at least ten members that fall within the established Oct, Octn, Oct/Octn-related, or Oat-related subclades. Four orthologs appear orthologous to individual SLC22 members, three of which proved to be lethal in ubiquitous RNAi knockdowns. The *Drosophila* gene *CG6356* shares distinct homology with Slc22a16 and RNAi knockdown of this gene resulted in arrest at the pupa stage. Based on what is known about Slc22a16 transport function and its membership in the Oct/Octn-related

subclade, it is possible that this arrest is due to a systemic imbalance of both carnitine and choline. Previous Drosophila developmental studies have found that proper levels of either carnitine or choline are necessary for flies to reach eclosion [53]. Although Slc22a16 (FLIPT2/OCT6/CT2) has not yet been evaluated for the ability to transport choline, it is an established carnitine transporter [54]. FLIPT2/OCT6/CT2 is homologous to two carnitine transporters of the SLC22 family, OCTN1 (Slc22a4) and OCTN2 (Slc22a5), which have been shown to transport acetylcholine and choline, respectively, in addition to acetylcarnitine and carnitine [55]–[58].

We observe that ubiquitous knocked down *SLC22A/CG7442* causes arrest at the pupa stage, confirming observations made in previous studies [50]. *SLC22A/CG7442*, which shares homology with Slc22a15, has been confirmed as a transporter of characteristic Oct and Octn metabolites MPP+, dopamine, serotonin, carnitine, TEA, choline, and acetylcholine [50]. The fly gene *CG7458* also groups with Slc22a15 but lacks any phenotypic data to infer function. With further analysis, these associations could provide a basis for investigation of the endogenous function of the orphan transporter Slc22a15. Developmental tissue expression studies show transiently high expression of SLC22A15 in vital organs such as the heart, liver, and kidneys [59]. This transporter is also known to be highly expressed in white blood cells in humans, which are present at the highest concentration at birth and decrease to normal, adult levels by two years of age [55]. In combination with the observed *SLC22A/CG7442* developmental phenotype, it appears likely that *CG7442* orthologs (such as SLC22A15) in other species may play a developmental role.

CG3168 groups with the Oat-related subclade, appearing to share direct ancestry with the orphan transporter, Slc22a18. Previous studies have observed high levels of *CG3168* expression

in glial cells during embryogenesis [60]. Slc22a18 has been shown to be expressed in low levels in the adult brain in the Human Protein Atlas, GTEx, and FANTOM5 RNA-seq studies [59]. It also has low expression levels in the human fetal brain [61]. Between adult and fetal brain, there is a pattern of consistent expression of Slc22a18 in the cerebral cortex. The cerebral cortex consists of ~75% glial cells which could represent partly orthologous expression patterns between *CG3168* and *Slc22a18* in different species. Further investigation of *CG3168* and its relationship to the orphan SLC22 member, *Slc22a18*, could build an understanding of how both of these genes are implicated in development.

In addition to phylogenetic and developmental functional screens, RNAi knockdown fly lines that progressed to adulthood were examined for resistance to oxidative stress via paraquat resistance testing. Four knock-down lines (*CG4630* $p<0.05$, *CG6006* $p<0.05$, *CG6126* $p<0.01$ and *CG16727* $p<0.05$) showed significantly greater resistance to oxidative stress. *CG16727* has no phenotypic or phylogenetic associations other than increased paraquat survival for crosses with two separate *da-GAL*4 driver lines. However, it is specifically expressed in the Malpighian tubules, which are often considered somewhat analogous to the mammalian kidneys where excretion of the antioxidant-acting oxoacids of the TCA cycle, uric acid, and flavonoids normally occurs via OATs [15], [62]–[64]. TCA intermediates pyruvate, oxaloacetate, and α-ketoglutarate are known to mediate oxidative stress responses, due to direct interaction of their α-ketoacid structure with reactive oxygen species such as H_2O_2 [47], [65], [66]. Due to the conservation of the TCA and its metabolites between *Drosophila* and humans, we raise the possibility that RNAi knockdowns of potential OAT orthologs would be more resistant to PQ due to increased systemic levels of oxoacids with antioxidant properties and the removal of the excretory route for these intermediates would result in increased serum levels of these

metabolites which would protect against oxidative stress. Further, investigation of a Malpighian tubule-specific knockdown of this gene would be necessary to assess this phenotype and hypothetical functionality. *CG4630*, *CG6006* and *CG6126* showed similar oxidative stress resistant phenotypes when crossed with one da-GAL4 driver line. All three of these transporters are expressed within the *Drosophila* excretory system but have a wider range of tissue expression than *CG16727*. Resistance to oxidative stress exhibited by these RNAi knockdown lines must be further examined by hemolymph analysis for classical SLC22 antioxidants such as urate, EGT, and the oxoacids of the TCA. Oxidative stress resistance is of particular interest in the search for SLC22 organic anion transporters (OATs) in fruit flies. Our homology-based analyses show no unambiguous OAT orthologs in fruit flies. However, it is possible that some SLC22 fly genes transport organic anions but do not share enough sequence similarity for multiple sequence alignment programs to determine their orthologous functions.

Although some SLC22 RNAi knockdown lines may not show a significant phenotype, it has been shown that knocking down specific OA transporters in *D. melanogaster* can affect the expression patterns of other transporters with similar functionality, indicating a mechanism of sensing and signaling tied to organic anion, cation, and zwitterion transporters (OATs, OCTs, and OCTNs) [67], [68]. Changes of expression levels of functionally similar transporters could provide further support for the Remote Sensing and Signaling Theory, in which drug-related proteins (eg. drug transporters and drug metabolizing enzymes) and signaling molecules mediate inter-organ communication to maintain physiological balance [69], [70]. For mammalian organs, a transporter and DME gene remote sensing and signaling network (RSSN) has recently been proposed [46].

Our findings show that the fruit fly is a useful model system to investigate understudied transporters, specifically Slc22a15, Slc22a16, and Slc22a18, as well as gain functional insight into the SLC22 gene family as a whole. Additionally, confirmation of apparently strong phylogenetic relationships could result in viable models to better understand the functionality and developmental role of SLC22A16 and SLC22A18 through *CG6356* and *CG3168,* respectively. While further study is necessary to understand the mechanism of oxidative stress resistance in certain RNAi knockdown lines, it will also be interesting to determine if there are increased levels of antioxidants in these lines and what those antioxidants might be. Given the substantial genetic and physiological conservation between mammals and *Drosophila*, these findings may support, in certain contexts, the use of fruit flies as a pre-clinical model organism for select SLC22 transporters, for instance, in elucidating their role in handling oxidative stress.

This thesis is coauthored with Jeffry Granados, Da Shi, Milton H. Saier, Michael E. Baker, Reuben Abagyan, Priti Azad, Suwayda Ali, Gabriel G. Haddad, and Sanjay Nigam. The thesis author was the primary author of both "Systems Biology Analysis Reveals Eight SLC22 Transporter Subgroups, Including OATs, OCTs, and OCTNs" and "Drosophila SLC22 orthologs related to OATs, OCTs, and OCTNs regulate development and responsiveness to oxidative stress."

Chapter 2, in full, has been submitted for publication of the material as it may appear in International Journal of Molecular Sciences, 2020, Engelhart, Darcy C.; Azad, Priti; Ali, Suwayda; Granados, Jeffry C.; Haddad, Gabriel G.; Nigam, Sanjay K., 2020. The thesis author was the primary investigator and author of this paper.

2.6 REFERENCES

[1] César-Razquin, A.; Snijder, B.; Frappier-Brinton, T.; Isserlin, R.; Gyimesi, G.; Bai, X.; Reithmeier, R. A.; Hepworth, D.; Hediger, M. A.; Edwards, A. M.; . A Call for Systematic Research on Solute Carriers. *Cell*. Elsevier: Austria July 30, 2015, pp 478–487. https://doi.org/10.1016/j.cell.2015.07.022.

[2] Rosenthal, S. B.; Bush, K. T.; Nigam, S. K. A Network of SLC and ABC Transporter and DME Genes Involved in Remote Sensing and Signaling in the Gut-Liver-Kidney Axis. *Sci. Rep.* 2019, *9* [1]. https://doi.org/10.1038/s41598-019-47798-x.

[3] Nigam, S. K. What Do Drug Transporters Really Do? *Nature Reviews Drug Discovery*. Nature Publishing Group January 1, 2014, pp 29–44. https://doi.org/10.1038/nrd4461.

[4] Höglund, P. J.; Nordström, K. J. V; Schiöth, H. B.; Fredriksson, R. The Solute Carrier Families Have a Remarkably Long Evolutionary History with the Majority of the Human Families Present before Divergence of Bilaterian Species. *Mol. Biol. Evol.* 2011, *28* [4], 1531–1541. https://doi.org/10.1093/molbev/msq350.

[5] Limmer, S.; Weiler, A.; Volkenhoff, A.; Babatz, F.; Klämbt, C. The Drosophila Blood-Brain Barrier: Development and Function of a Glial Endothelium. *Front. Neurosci.* 2014, *8*, 365. https://doi.org/10.3389/fnins.2014.00365.

[6] Eraly, S. A.; Vallon, V.; Rieg, T.; Gangoiti, J. A.; Wikoff, W. R.; Siuzdak, G.; Barshop, B. A.; Nigam, S. K. Multiple Organic Anion Transporters Contribute to Net Renal Excretion of Uric Acid. *Physiol. Genomics* 2008, *33* [2], 180–192. https://doi.org/10.1152/physiolgenomics.00207.2007.

[7] Pandey, U. B.; Nichols, C. D. Human Disease Models in Drosophila Melanogaster and the Role of the Fly in Therapeutic Drug Discovery. *Pharmacol. Rev.* 2011, *63* [2], 411–436. https://doi.org/10.1124/pr.110.003293.

[8] Xu, Y.; An, F.; Borycz, J. A.; Borycz, J.; Meinertzhagen, I. A.; Wang, T. Histamine Recycling Is Mediated by CarT, a Carcinine Transporter in Drosophila Photoreceptors. *PLoS Genet.* 2015, *11* [12]. https://doi.org/10.1371/journal.pgen.1005764.

[9] Ogasawara, M.; Yamauchi, K.; Satoh, Y. I.; Yamaji, R.; Inui, K.; Jonker, J. W.; Schinkel, A. H.; Maeyama, K. Recent Advances in Molecular Pharmacology of the Histamine Systems: Organic Cation Transporters as a Histamine Transporter and Histamine Metabolism. *Journal of Pharmacological Sciences*. May 2006, pp 24–30. https://doi.org/10.1254/jphs.FMJ06001X6.

[10] Han, Y.; Xiong, L.; Xu, Y.; Tian, T.; Wang, T. The β-Alanine Transporter BalaT Is Required for Visual Neurotransmission in Drosophila. *Elife* 2017, *6*. https://doi.org/10.7554/eLife.29146.

[11] Hattori, Y.; Hattori, K.; Matsuda, N. Regulation of the Cardiovascular System by

Histamine. In *Handbook of Experimental Pharmacology*; Springer New York LLC, 2017; Vol. 241, pp 1–20. https://doi.org/10.1007/164_2016_15.

[12] Hu, W.; Chen, Z. The Roles of Histamine and Its Receptor Ligands in Central Nervous System Disorders: An Update. *Pharmacology and Therapeutics*. Elsevier Inc. July 1, 2017, pp 116–132. https://doi.org/10.1016/j.pharmthera.2017.02.039.

[13] Wang, Y.; Moussian, B.; Schaeffeler, E.; Schwab, M.; Nies, A. T. The Fruit Fly Drosophila Melanogaster as an Innovative Preclinical ADME Model for Solute Carrier Membrane Transporters, with Consequences for Pharmacology and Drug Therapy. *Drug Discovery Today*. Elsevier Ltd October 1, 2018, pp 1746–1760. https://doi.org/10.1016/j.drudis.2018.06.002.

[14] Mayer, F. P.; Schmid, D.; Owens, W. A.; Gould, G. G.; Apuschkin, M.; Kudlacek, O.; Salzer, I.; Boehm, S.; Chiba, P.; Williams, P. H.; . An Unsuspected Role for Organic Cation Transporter 3 in the Actions of Amphetamine. *Neuropsychopharmacology* 2018, *43* [12], 2408–2417. https://doi.org/10.1038/s41386-018-0053-5.

[15] Chintapalli, V. R.; Wang, J.; Dow, J. A. T. Using FlyAtlas to Identify Better Drosophila Melanogaster Models of Human Disease. *Nature Genetics*. June 2007, pp 715–720. https://doi.org/10.1038/ng2049.

[16] Zwart, R.; Verhaagh, S.; Buitelaar, M.; Popp-Snijders, C.; Barlow, D. P. Impaired Activity of the Extraneuronal Monoamine Transporter System Known as Uptake-2 in Orct3/Slc22a3-Deficient Mice. *Mol. Cell. Biol.* 2001, *21* [13], 4188–4196. https://doi.org/10.1128/mcb.21.13.4188-4196.2001.

[17] Jonker, J. W.; Wagenaar, E.; van Eijl, S.; Schinkel, A. H. Deficiency in the Organic Cation Transporters 1 and 2 [Oct1/Oct2 [Slc22a1/Slc22a2]] in Mice Abolishes Renal Secretion of Organic Cations. *Mol. Cell. Biol.* 2003, *23* [21], 7902–7908. https://doi.org/10.1128/mcb.23.21.7902-7908.2003.

[18] Zhu, C.; Nigam, K. B.; Date, R. C.; Bush, K. T.; Springer, S. A.; Saier, M. H.; Wu, W.; Nigam, S. K. Evolutionary Analysis and Classification of OATs, OCTs, OCTNs, and Other SLC22 Transporters: Structure-Function Implications and Analysis of Sequence Motifs. *PLoS One* 2015, *10* [11], 140569. https://doi.org/10.1371/journal.pone.0140569.

[19] Pavlova, A.; Sakurai, H.; Leclercq, B.; Beier, D. R.; Yu, A. S. L.; Nigam, S. K. Developmentally Regulated Expression of Organic Ion Transporters NKT[OAT1], OCT1, NLT[OAT2], and Roct. *Am. J. Physiol. - Ren. Physiol.* 2000, *278* [4 47-4], F635–F643. https://doi.org/10.1152/ajprenal.2000.278.4.F635.

[20] Vallon, V.; Eraly, S. A.; Wikoff, W. R.; Rieg, T.; Kaler, G.; Truong, D. M.; Ahn, S. Y.; Mahapatra, N. R.; Mahata, S. K.; Gangoiti, J. A.; . Organic Anion Transporter 3 Contributes to the Regulation of Blood Pressure. *J. Am. Soc. Nephrol.* 2008, *19* [9], 1732–1740. https://doi.org/10.1681/ASN.2008020180.

[21] Jonker, J. W.; Wagenaar, E.; Mol, C. A. A. M.; Buitelaar, M.; Koepsell, H.; Smit, J. W.;

Schinkel, A. H. Reduced Hepatic Uptake and Intestinal Excretion of Organic Cations in Mice with a Targeted Disruption of the Organic Cation Transporter 1 [Oct1 [Slc22a1]] Gene. *Mol. Cell. Biol.* 2001, *21* [16], 5471–5477. https://doi.org/10.1128/mcb.21.16.5471-5477.2001.

[22] Eraly, S. A.; Vallon, V.; Vaughn, D. A.; Gangoiti, J. A.; Richter, K.; Nagle, M.; Monte, J. C.; Rieg, T.; Truong, D. M.; Long, J. M.; . Decreased Renal Organic Anion Secretion and Plasma Accumulation of Endogenous Organic Anions in OAT1 Knock-out Mice. *J. Biol. Chem.* 2006, *281* [8], 5072–5083. https://doi.org/10.1074/jbc.M508050200.

[23] Sweet, D. H.; Miller, D. S.; Pritchard, J. B.; Fujiwara, Y.; Beier, D. R.; Nigam, S. K. Impaired Organic Anion Transport in Kidney and Choroid Plexus of Organic Anion Transporter 3 [Oat3 [Slc22a8]] Knockout Mice. *J. Biol. Chem.* 2002, *277* [30], 26934–26943. https://doi.org/10.1074/jbc.M203803200.

[24] Shekhawat, P. S.; Yang, H. S.; Bennett, M. J.; Carter, A. L.; Matern, D.; Tamai, I.; Ganapathy, V. Carnitine Content and Expression of Mitochondrial β-Oxidation Enzymes in Placentas of Wild-Type [OCTN2+/+] and OCTN2 Null [OCTN2 -/-] Mice. *Pediatr. Res.* 2004, *56* [3], 323–328. https://doi.org/10.1203/01.PDR.0000134252.02876.55.

[25] Kato, Y.; Kubo, Y.; Iwata, D.; Kato, S.; Sudo, T.; Sugiura, T.; Kagaya, T.; Wakayama, T.; Hirayama, A.; Sugimoto, M.; . Gene Knockout and Metabolome Analysis of Carnitine/Organic Cation Transporter OCTN1. *Pharm. Res.* 2010, *27* [5], 832–840. https://doi.org/10.1007/s11095-010-0076-z.

[26] Nezu, J. I.; Tamai, I.; Oku, A.; Ohashi, R.; Yabuuchi, H.; Hashimoto, N.; Nikaido, H.; Sai, Y.; Koizumi, A.; Shoji, Y.; . Primary Systemic Carnitine Deficiency Is Caused by Mutations in a Gene Encoding Sodium Ion-Dependent Carnitine Transporter. *Nat. Genet.* 1999, *21* [1], 91–94. https://doi.org/10.1038/5030.

[27] Urban, T. J.; Gallagher, R. C.; Brown, C.; Castro, R. A.; Lagpacan, L. L.; Brett, C. M.; Taylor, T. R.; Carlson, E. J.; Ferrin, T. E.; Burchard, E. G.; . Functional Genetic Diversity in the High-Affinity Carnitine Transporter OCTN2 [SLC22A5]. *Mol. Pharmacol.* 2006, *70* [5], 1602–1611. https://doi.org/10.1124/mol.106.028126.

[28] Momper, J. D.; Nigam, S. K. Developmental Regulation of Kidney and Liver Solute Carrier and ATP-Binding Cassette Drug Transporters and Drug Metabolizing Enzymes: The Role of Remote Organ Communication. *Expert Opinion on Drug Metabolism and Toxicology*. Taylor and Francis Ltd June 3, 2018, pp 561–570. https://doi.org/10.1080/17425255.2018.1473376.

[29] Momper, J. D.; Yang, J.; Gockenbach, M.; Vaida, F.; Nigam, S. K. Dynamics of Organic Anion Transporter-Mediated Tubular Secretion during Postnatal Human Kidney Development and Maturation. *Clin. J. Am. Soc. Nephrol.* 2019, *14* [4], 540–548. https://doi.org/10.2215/CJN.10350818.

[30] Sweeney, D. E.; Vallon, V.; Rieg, T.; Wu, W.; Gallegos, T. F.; Nigam, S. K. Functional Maturation of Drug Transporters in the Developing, Neonatal, and Postnatal Kidney. *Mol.*

Pharmacol. 2011, *80* [1], 147–154. https://doi.org/10.1124/mol.110.070680.

[31] Brouwer, K. L. R.; Aleksunes, L. M.; Brandys, B.; Giacoia, G. P.; Knipp, G.; Lukacova, V.; Meibohm, B.; Nigam, S.; M Rieder, M.; Wildt, S. De. Human Ontogeny of Drug Transporters: Review and Recommendations of the Pediatric Transporter Working Group. *Clinical Pharmacology and Therapeutics*. Nature Publishing Group September 1, 2015, pp 266–287. https://doi.org/10.1002/cpt.176.

[32] Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; . Tissue-Based Map of the Human Proteome. *Science [80-.]*. 2015, *347* [6220]. https://doi.org/10.1126/science.1260419.

[33] Bainbridge, S. P.; Bownes, M. Staging the Metamorphosis of Drosophila Melanogaster. *J. Embryol. Exp. Morphol.* 1981, *Vol.66*, 57–80.

[34] Chen, Y. W.; Song, S.; Weng, R.; Verma, P.; Kugler, J. M.; Buescher, M.; Rouam, S.; Cohen, S. M. Systematic Study of Drosophila MicroRNA Functions Using a Collection of Targeted Knockout Mutations. *Dev. Cell* 2014, *31* [6], 784–800. https://doi.org/10.1016/j.devcel.2014.11.029.

[35] McCormack, A. L.; Thiruchelvam, M.; Manning-Bog, A. B.; Thiffault, C.; Langston, J. W.; Cory-Slechta, D. A.; Di Monte, D. A. Environmental Risk Factors and Parkinson's Disease: Selective Degeneration of Nigral Dopaminergic Neurons Caused by the Herbicide Paraquat. *Neurobiol. Dis.* 2002, *10* [2], 119–127.

[36] Cohen, G. M.; D'Arcy Doherty, M. Free Radical Mediated Cell Toxicity by Redox Cycling Chemicals. *Br. J. Cancer* 1987, *55* [SUPPL. 8], 46–52.

[37] Bus, J. S.; Gibson, J. E. Paraquat: Model for Oxidant-Initiated Toxicity. *Environ. Health Perspect.* 1984, *55*, 37–46. https://doi.org/10.1289/ehp.845537.

[38] Roberts, C. K.; Sindhu, K. K. Oxidative Stress and Metabolic Syndrome. *Life Sciences*. May 22, 2009, pp 705–712. https://doi.org/10.1016/j.lfs.2009.02.026.

[39] Ames, B. N.; Cathcart, R.; Schwiers, E.; Hochstein, P. Uric Acid Provides an Antioxidant Defense in Humans against Oxidant- and Radical-Caused Aging and Cancer: A Hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* 1981, *78* [11 II], 6858–6862. https://doi.org/10.1073/pnas.78.11.6858.

[40] Cheah, I. K.; Halliwell, B. Ergothioneine; Antioxidant Potential, Physiological Function and Role in Disease. *Biochimica et Biophysica Acta - Molecular Basis of Disease*. May 2012, pp 784–793. https://doi.org/10.1016/j.bbadis.2011.09.017.

[41] Marchette, L. D.; Wang, H.; Li, F.; Babizhayev, M. A.; Kasus-Jacobi, A. Carcinine Has 4-Hydroxynonenal Scavenging Property and Neuroprotective Effect in Mouse Retina. *Investig. Ophthalmol. Vis. Sci.* 2012, *53* [7], 3572–3583. https://doi.org/10.1167/iovs.11-9042.

[42] Engelhart, D. C.; Granados, J.; Nigam, S. K.; Shi, D.; Saier, M. H.; Baker, M. E. Reclassification of SLC22 Transporters: Analysis of OAT, OCT, OCTN, and Other Family Members Reveals 8 Functional Subgroups. *bioRxiv* 2019, 2019.12.23.887299. https://doi.org/10.1101/2019.12.23.887299.

[43] Nigam, S. K. The SLC22 Transporter Family: A Paradigm for the Impact of Drug Transporters on Metabolic Pathways, Signaling, and Disease. *Annu. Rev. Pharmacol. Toxicol.* 2018, *58* [1], 663–687. https://doi.org/10.1146/annurev-pharmtox-010617-052713.

[44] Kimura, M. A Simple Method for Estimating Evolutionary Rates of Base Substitutions through Comparative Studies of Nucleotide Sequences. *J. Mol. Evol.* 1980, *16* [2], 111–120. https://doi.org/10.1007/BF01731581.

[45] Huang, H.; Haddad, G. G. Drosophila DMRP4 Regulates Responsiveness to O2 Deprivation and Development under Hypoxia. *Physiol. Genomics* 2007, *29* [3], 260–266. https://doi.org/10.1152/physiolgenomics.00166.2006.

[46] Nigam, A. K.; Li, J. G.; Lall, K.; Shi, D.; Bush, K. T.; Bhatnagar, V.; Abagyan, R.; Nigam, S. K. Unique Metabolite Preferences of the Drug Transporters OAT1 and OAT3 Analyzed by Machine Learning. *J. Biol. Chem.* 2020, jbc.RA119.010729. https://doi.org/10.1074/jbc.ra119.010729.

[47] Andrae, U.; Singh, J.; Ziegler-Skylakakis, K. Pyruvate and Related α-Ketoacids Protect Mammalian Cells in Culture against Hydrogen Peroxide-Induced Cytotoxicity. *Toxicol. Lett.* 1985, *28* [2–3], 93–98. https://doi.org/10.1016/0378-4274[85]90015-3.

[48] Huang, S. M.; Zhang, L.; Giacomini, K. M. The International Transporter Consortium: A Collaborative Group of Scientists from Academia, Industry, and the FDA. In *Clinical Pharmacology and Therapeutics*; 2010; Vol. 87, pp 32–36. https://doi.org/10.1038/clpt.2009.236.

[49] Yin, J.; Sun, W.; Li, F.; Hong, J.; Li, X.; Zhou, Y.; Lu, Y.; Liu, M.; Zhang, X.; Chen, N.; . VARIDT 1.0: Variability of Drug Transporter Database. *Nucleic Acids Res.* 2019. https://doi.org/10.1093/nar/gkz779.

[50] Gai, Y.; Liu, Z.; Cervantes-Sandoval, I.; Davis, R. L. Drosophila SLC22A Transporter Is a Memory Suppressor Gene That Influences Cholinergic Neurotransmission to the Mushroom Bodies. *Neuron* 2016, *90* [3], 581–595. https://doi.org/10.1016/j.neuron.2016.03.017.

[51] Eraly, S. A.; Monte, J. C.; Nigam, S. K. Novel Slc22 Transporter Homologs in Fly, Worm, and Human Clarify the Phylogeny of Organic Anion and Cation Transporters. *Physiol. Genomics* 2004, *18* [1], 12–24. https://doi.org/10.1152/physiolgenomics.00014.2004.

[52] Stenesen, D.; Moehlman, A. T.; Krämer, H. The Carcinine Transporter CarT Is Required in Drosophila Photoreceptor Neurons to Sustain Histamine Recycling. *Elife* 2015, *4*

[DECEMBER2015]. https://doi.org/10.7554/eLife.10972.

[53] Geer, B. W.; Vovis, G. F. The Effects of Choline and Related Compounds on the Growth and Development of Drosophila Melanogaster. *J. Exp. Zool.* 1965, *158* [2], 223–236. https://doi.org/10.1002/jez.1401580209.

[54] Enomoto, A.; Wempe, M. F.; Tsuchida, H.; Shin, H. J.; Cha, S. H.; Anzai, N.; Goto, A.; Sakamoto, A.; Niwa, T.; Kanai, Y.; . Molecular Identification of a Novel Carnitine Transporter Specific to Human Testis: Insights into the Mechanism of Carnitine Recognition. *J. Biol. Chem.* 2002, *277* [39], 36262–36271. https://doi.org/10.1074/jbc.M203883200.

[55] Eraly, S. A.; Nigam, S. K. Novel Human CDNAs Homologous to Drosophila Orct and Mammalian Carnitine Transporters. *Biochem. Biophys. Res. Commun.* 2002, *297* [5], 1159–1166. https://doi.org/10.1016/S0006-291X[02]02343-4.

[56] Pochini, L.; Scalise, M.; Di Silvestre, S.; Belviso, S.; Pandolfi, A.; Arduini, A.; Bonomini, M.; Indiveri, C. Acetylcholine and Acetylcarnitine Transport in Peritoneum: Role of the SLC22A4 [OCTN1] Transporter. *Biochim. Biophys. Acta - Biomembr.* 2016, *1858* [4], 653–660. https://doi.org/10.1016/j.bbamem.2015.12.026.

[57] Wu, X.; Huang, W.; Prasad, P. D.; Seth, P.; Rajan, D. P.; Leibach, F. H.; Chen, J.; Conway, S. J.; Ganapathy, V. Functional Characteristics and Tissue Distribution Pattern of Organic Cation Transporter 2 [OCTN2], an Organic Cation/Carnitine Transporter. *J. Pharmacol. Exp. Ther.* 1999, *290* [3], 1482–1492.

[58] Tamai, I.; Ohashi, R.; Nezu, J. I.; Sai, Y.; Kobayashi, D.; Oku, A.; Shimane, M.; Tsuji, A. Molecular and Functional Characterization of Organic Cation/Carnitine Transporter Family in Mice. *J. Biol. Chem.* 2000, *275* [51], 40064–40072. https://doi.org/10.1074/jbc.M005340200.

[59] Papatheodorou, I.; Fonseca, N. A.; Keays, M.; Tang, Y. A.; Barrera, E.; Bazant, W.; Burke, M.; Füllgrabe, A.; Fuentes, A. M. P.; George, N.; Expression Atlas: Gene and Protein Expression across Multiple Studies and Organisms. *Nucleic Acids Res.* 2018, *46* [D1], D246–D251. https://doi.org/10.1093/nar/gkx1158.

[60] Altenhein, B.; Becker, A.; Busold, C.; Beckmann, B.; Hoheisel, J. D.; Technau, G. M. Expression Profiling of Glial Genes during Drosophila Embryogenesis. *Dev. Biol.* 2006, *296* [2], 545–560. https://doi.org/10.1016/j.ydbio.2006.04.460.

[61] Lindsay, S. J.; Xu, Y.; Lisgo, S. N.; Harkin, L. F.; Copp, A. J.; Gerrelli, D.; Clowry, G. J.; Talbot, A.; Keogh, M. J.; Coxhead, J.; HDBR Expression: A Unique Resource for Global and Individual Gene Expression Studies during Early Human Brain Development. *Front. Neuroanat.* 2016, *10*. https://doi.org/10.3389/fnana.2016.00086.

[62] Berridge, M. J.; Oschman, J. L. A Structural Basis for Fluid Secretion by Malpighian Tubules. *Tissue Cell* 1969, *1* [2], 247–272. https://doi.org/10.1016/S0040-8166[69]80025-X.

[63] Wah Yee, S.; Stecula, A.; Chien, H.-C.; Zou, L.; Feofanova, E. V.; Borselen, V.; Wun Kathy Cheung, K.; Yousri, N. A.; Suhre, K.; Kinchen, J. M.; Unraveling the Functional Role of the Orphan Solute Carrier, SLC22A24 in the Transport of 2 Steroid Conjugates through Metabolomic and Genome-Wide Association Studies 3 4. *PLOS Genet.* 2019, *15* [9], e1008208. https://doi.org/10.1101/648238.

[64] Nigam, S. K.; Bush, K. T.; Martovetsky, G.; Ahn, S. Y.; Liu, H. C.; Richard, E.; Bhatnagar, V.; Wu, W. The Organic Anion Transporter [OAT] Family: A Systems Biology Perspective. *Physiol. Rev.* 2015, *95* [1], 83–123. https://doi.org/10.1152/physrev.00025.2013.

[65] Fork, C.; Bauer, T.; Golz, S.; Geerts, A.; Weiland, J.; Del Turco, D.; Schömig, E.; Gründemann, D. OAT2 Catalyses Efflux of Glutamate and Uptake of Orotic Acid. *Biochem. J.* 2011, *436* [2], 305–312. https://doi.org/10.1042/BJ20101904.

[66] Sawa, K.; Uematsu, T.; Korenaga, Y.; Hirasawa, R.; Kikuchi, M.; Murata, K.; Zhang, J.; Gai, X.; Sakamoto, K.; Koyama, T.; Krebs Cycle Intermediates Protective against Oxidative Stress by Modulating the Level of Reactive Oxygen Species in Neuronal HT22 Cells. *Antioxidants* 2017, *6* [1]. https://doi.org/10.3390/antiox6010021.

[67] Chahine, S.; Campos, A.; O'Donnell, M. J. Genetic Knockdown of a Single Organic Anion Transporter Alters the Expression of Functionally Related Genes in Malpighian Tubules of Drosophila Melanogaster. *J. Exp. Biol.* 2012, *215* [15], 2601–2610. https://doi.org/10.1242/jeb.071100.

[68] Chahine, S.; Seabrooke, S.; O'Donnell, M. J. Effects of Genetic Knock-down of Organic Anion Transporter Genes on Secretion of Fluorescent Organic Ions by Malpighian Tubules of Drosophila Melanogaster. *Arch. Insect Biochem. Physiol.* 2012, *81* [4], 228–240. https://doi.org/10.1002/arch.21066.

[69] Ahn, S. Y.; Nigam, S. K. Toward a Systems Level Understanding of Organic Anion and Other Multispecific Drug Transporters: A Remote Sensing and Signaling Hypothesis. *Molecular Pharmacology*. September 2009, pp 481–490. https://doi.org/10.1124/mol.109.056564.

[70] Wu, W.; Dnyanmote, A. V.; Nigam, S. K. Remote Communication through Solute Carriers and ATP Binding Cassette Drug Transporter Pathways: An Update on the Remote Sensing and Signaling Hypothesis. *Molecular Pharmacology*. May 2011, pp 795–805. https://doi.org/10.1124/mol.110.070607.

[71] Cary, G. A.; Cameron, R. A.; Hinman, V. F. EchinoBase: Tools for Echinoderm Genome Analyses. In *Methods in Molecular Biology*; Humana Press Inc., 2018; Vol. 1757, pp 349–369. https://doi.org/10.1007/978-1-4939-7737-6_12.

[72] Lee, R. Y. N.; Howe, K. L.; Harris, T. W.; Arnaboldi, V.; Cain, S.; Chan, J.; Chen, W. J.; Davis, P.; Gao, S.; Grove, C.; WormBase 2017: Molting into a New Stage. *Nucleic Acids Res.* 2018, *46* [D1], D869–D874. https://doi.org/10.1093/nar/gkx998.

[73] Casper, J.; Zweig, A. S.; Villarreal, C.; Tyner, C.; Speir, M. L.; Rosenbloom, K. R.; Raney, B. J.; Lee, C. M.; Lee, B. T.; Karolchik, D.; The UCSC Genome Browser Database: 2018 Update. *Nucleic Acids Res.* 2018, *46* [D1], D762–D769. https://doi.org/10.1093/nar/gkx1020.

[74] Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* 1990, *215* [3], 403–410. https://doi.org/10.1016/S0022-2836[05]80360-2.

[75] Marygold, S. J.; Crosby, M. A.; Goodman, J. L. Using Flybase, a Database of Drosophila Genes and Genomes. In *Methods in Molecular Biology*; 2016; Vol. 1478, pp 1–31. https://doi.org/10.1007/978-1-4939-6371-3_1.

[76] Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* 2011, *7*. https://doi.org/10.1038/msb.2011.75.

[77] Katoh, K.; Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* 2013, *30* [4], 772–780. https://doi.org/10.1093/molbev/mst010.

[78] Madeira, F.; Park, Y. mi; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; The EMBL-EBI Search and Sequence Analysis Tools APIs in 2019. *Nucleic Acids Res.* 2019, *47* [W1], W636–W641. https://doi.org/10.1093/nar/gkz268.

[79] Letunic, I.; Bork, P. Interactive Tree Of Life [ITOL] v4: Recent Updates and New Developments. *Nucleic Acids Res.* 2019, *47* [W1], W256–W259. https://doi.org/10.1093/nar/gkz239.

[80] The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* 1987. https://doi.org/10.1093/oxfordjournals.molbev.a040454.

[81] Perkins, L. A.; Holderbaum, L.; Tao, R.; Hu, Y.; Sopko, R.; McCall, K.; Yang-Zhou, D.; Flockhart, I.; Binari, R.; Shim, H. S.; The Transgenic RNAi Project at Harvard Medical School: Resources and Validation. *Genetics* 2015, *201* [3], 843–852. https://doi.org/10.1534/genetics.115.180208.

[82] Piccin, A. Efficient and Heritable Functional Knock-out of an Adult Phenotype in Drosophila Using a GAL4-Driven Hairpin RNA Incorporating a Heterologous Spacer. *Nucleic Acids Res.* 2001, *29* [12], 55e – 55. https://doi.org/10.1093/nar/29.12.e55.