

## **Supplementary Material**

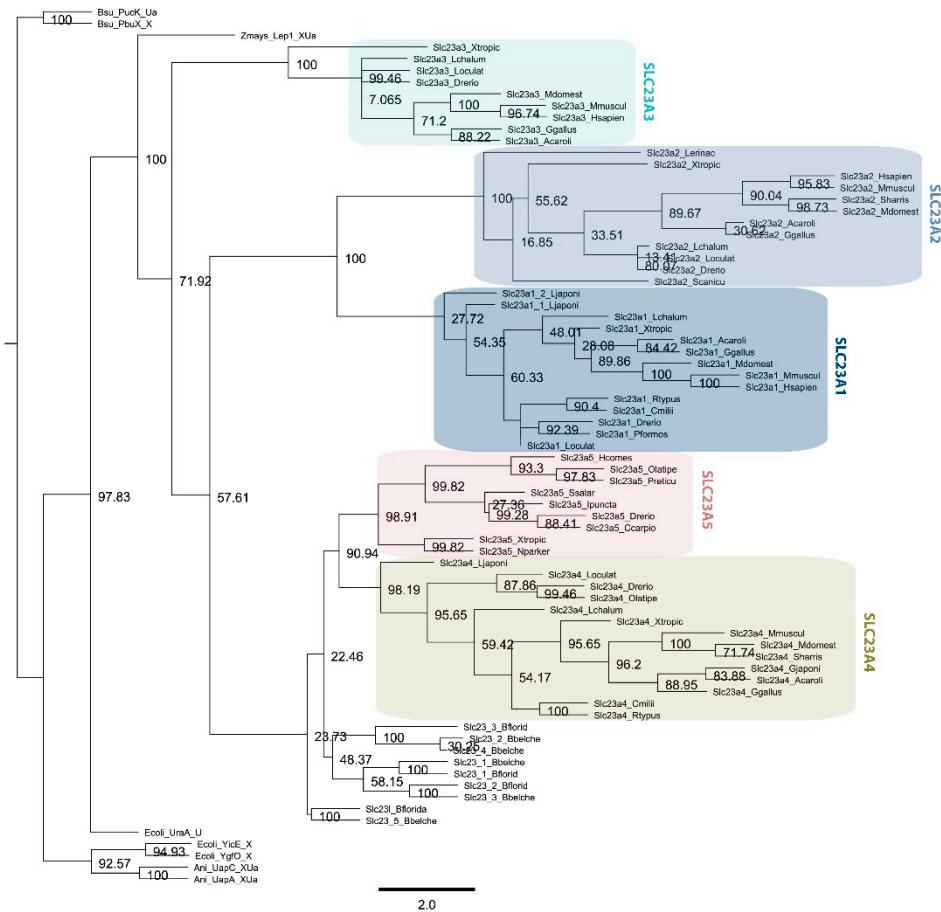
### **Identification of a Novel Nucleobase-Ascorbate Transporter Family Member in Fish and Amphibians**

#### **Supplementary Figures:**

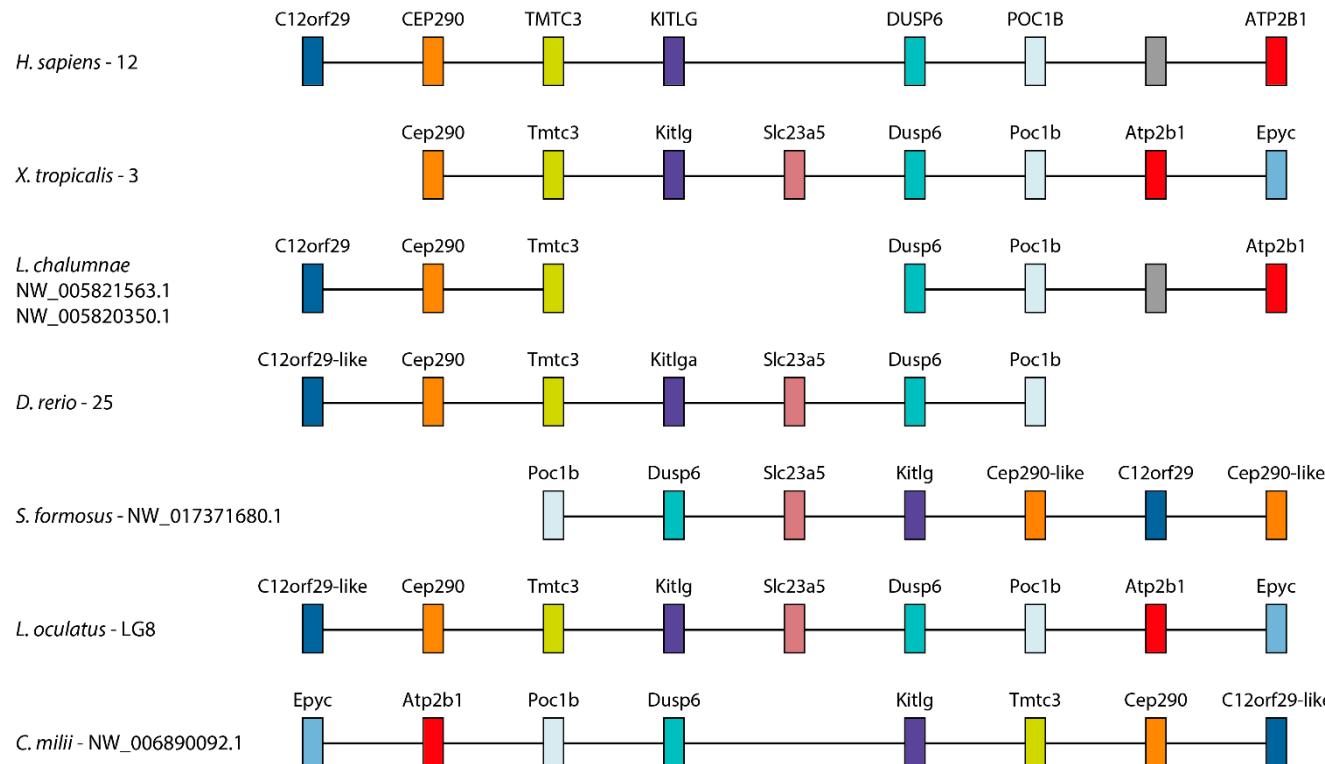
- Figure S1 – RaxML phylogenetic tree.
- Figure S2 - Synteny maps of *Slc23a5*.
- Figure S3 - Synteny maps of *Slc23a1*.
- Figure S4 - Synteny maps of *Slc23a2*.
- Figure S5 - Synteny maps of *Slc23a3*.
- Figure S6 - Synteny maps of *Slc23a4*.
- Figure S7 – S/T-X-Ø tripeptide conservation in *Slc23a1* and *Slc23a4* sequences.
- Figure S8 – Topology prediction of *Slc23a5*.

#### **Supplementary Tables:**

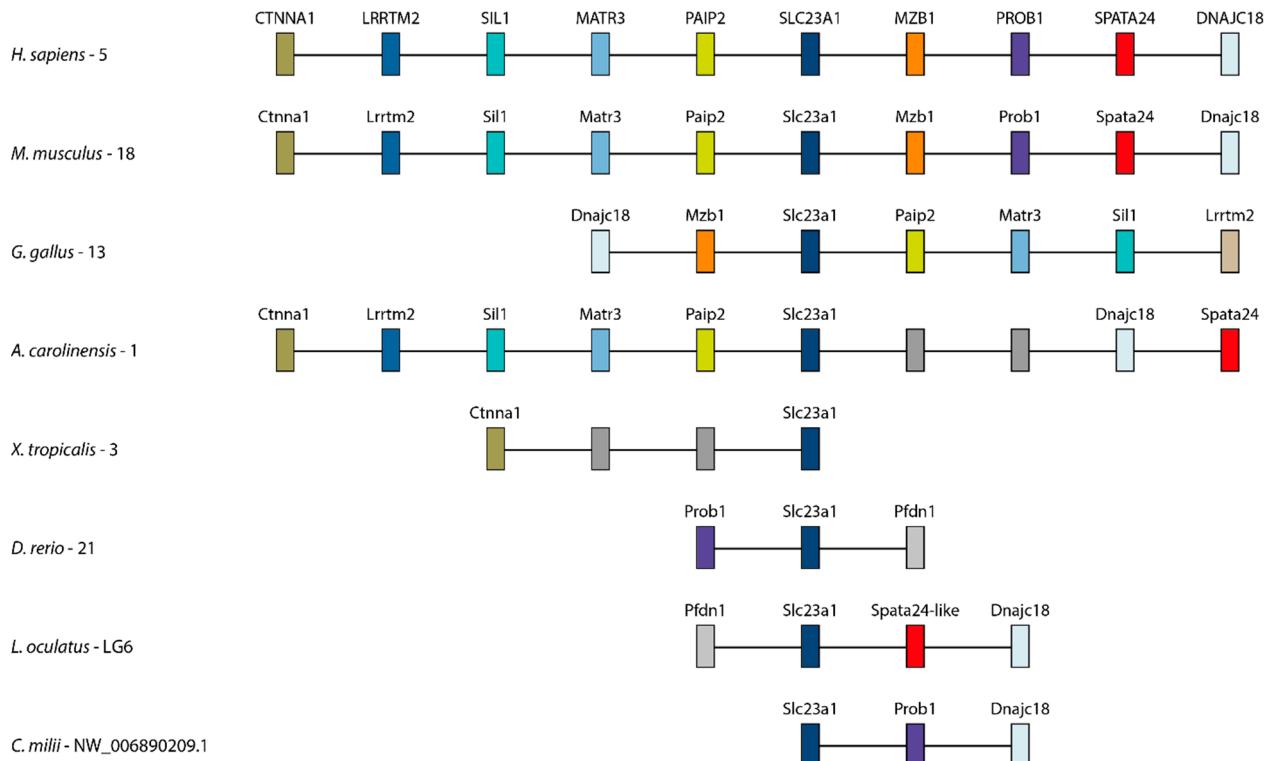
- Table S1 - Accession numbers.
- Table S2 - Accession numbers of the RNAseq files.
- Table S3 - Gene identifier for all genes included in gene expression analysis, collected from NCBI and ENSEMBLE Databases.
- Table S4 - Genome and GTF files retrieved from Ensemble database (Release 89) and Transcriptome files retrieved from NCBI used for this study.
- Table S5 - Relative gene expression levels of Slc23 family genes for eight species. The values are presented in log2 (TPM +1) ratios.
- Table S6 - Renal uric acid absorption and secretion: the urate transportome.



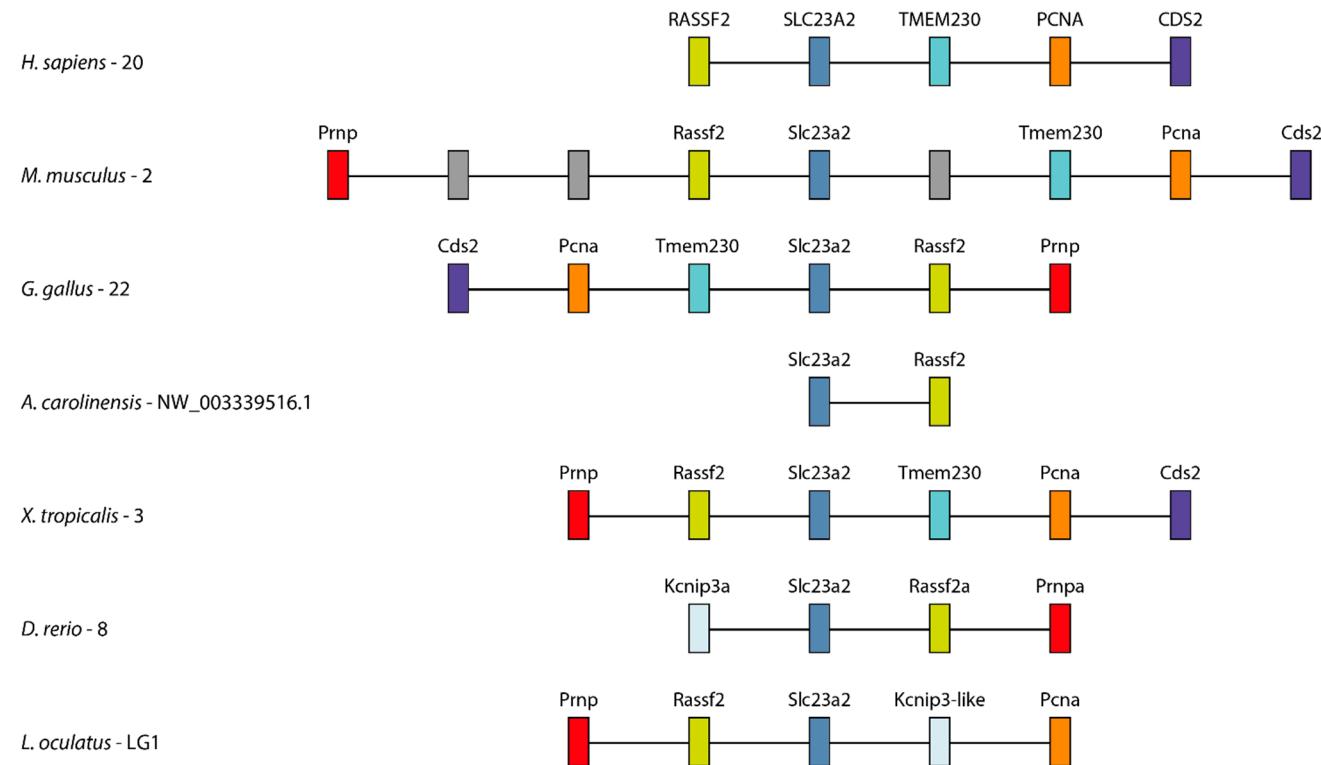
**Figure S1.** RaxML phylogenetic tree. Node values represent bootstrap values.



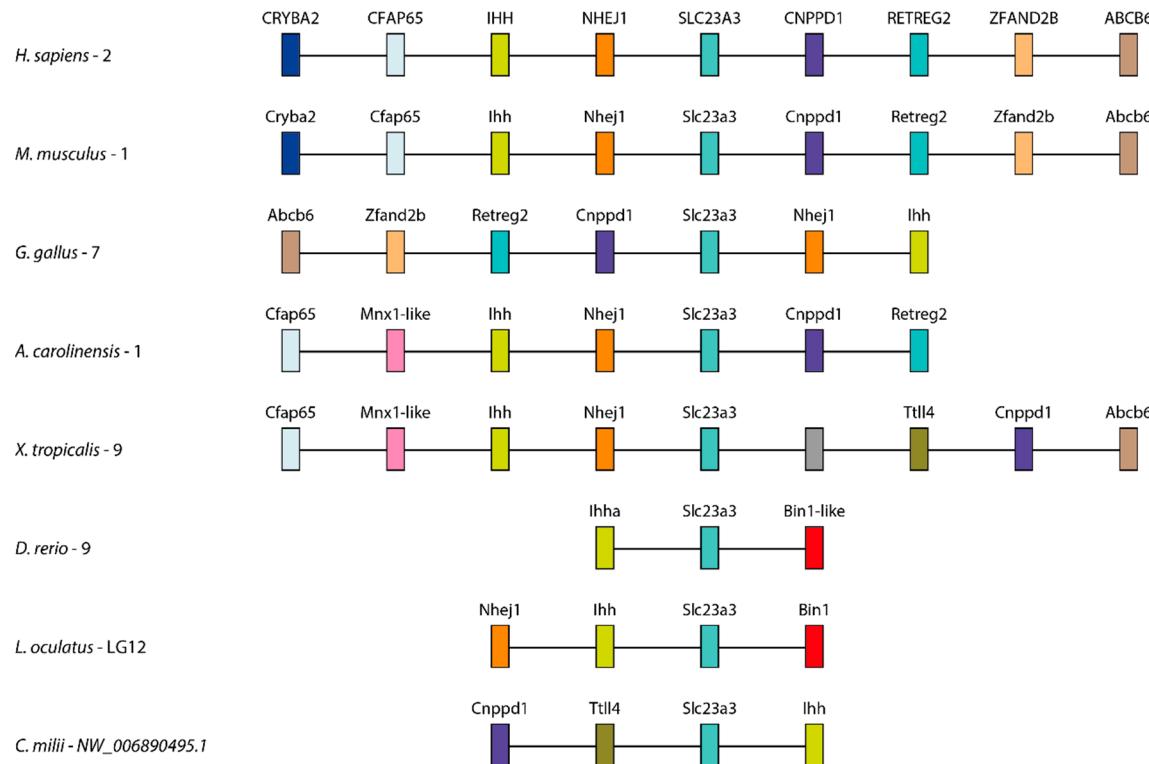
**Figure S2.** Synteny maps of *Slc23a5*. Conservation of the *Slc23a5* locus in major fish lineages, *Xenopus tropicalis* and *Homo sapiens*. Numbering after species corresponds to chromosome or scaffold. Information is presently absent for the Chondrostei *Acipenser ruthenus*. Grey boxes indicate non-conserved genes.



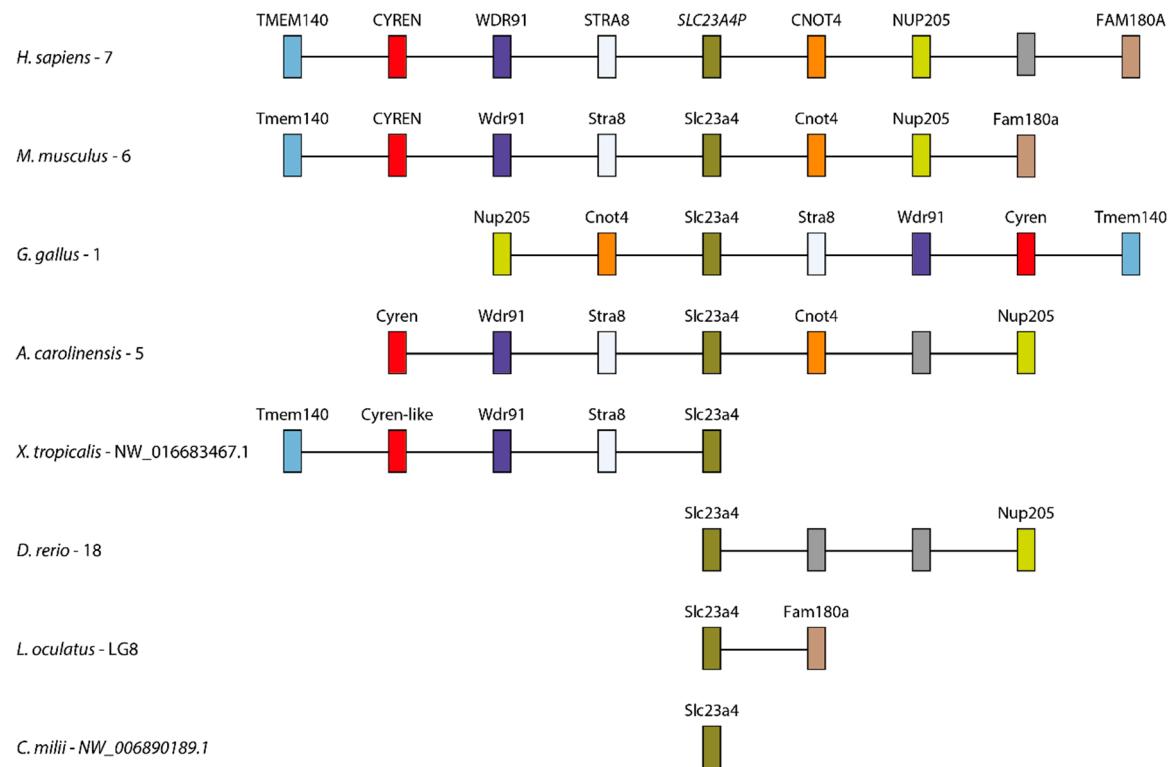
**Figure S3.** Synteny maps of *Slc23a1*. Conservation of the *Slc23a1* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes.



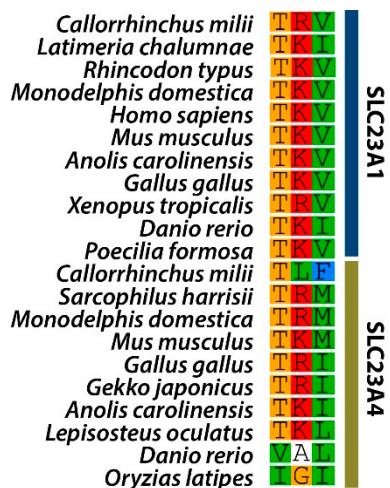
**Figure S4.** Synteny maps of *Slc23a2*. Conservation of the *Slc23a2* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes..



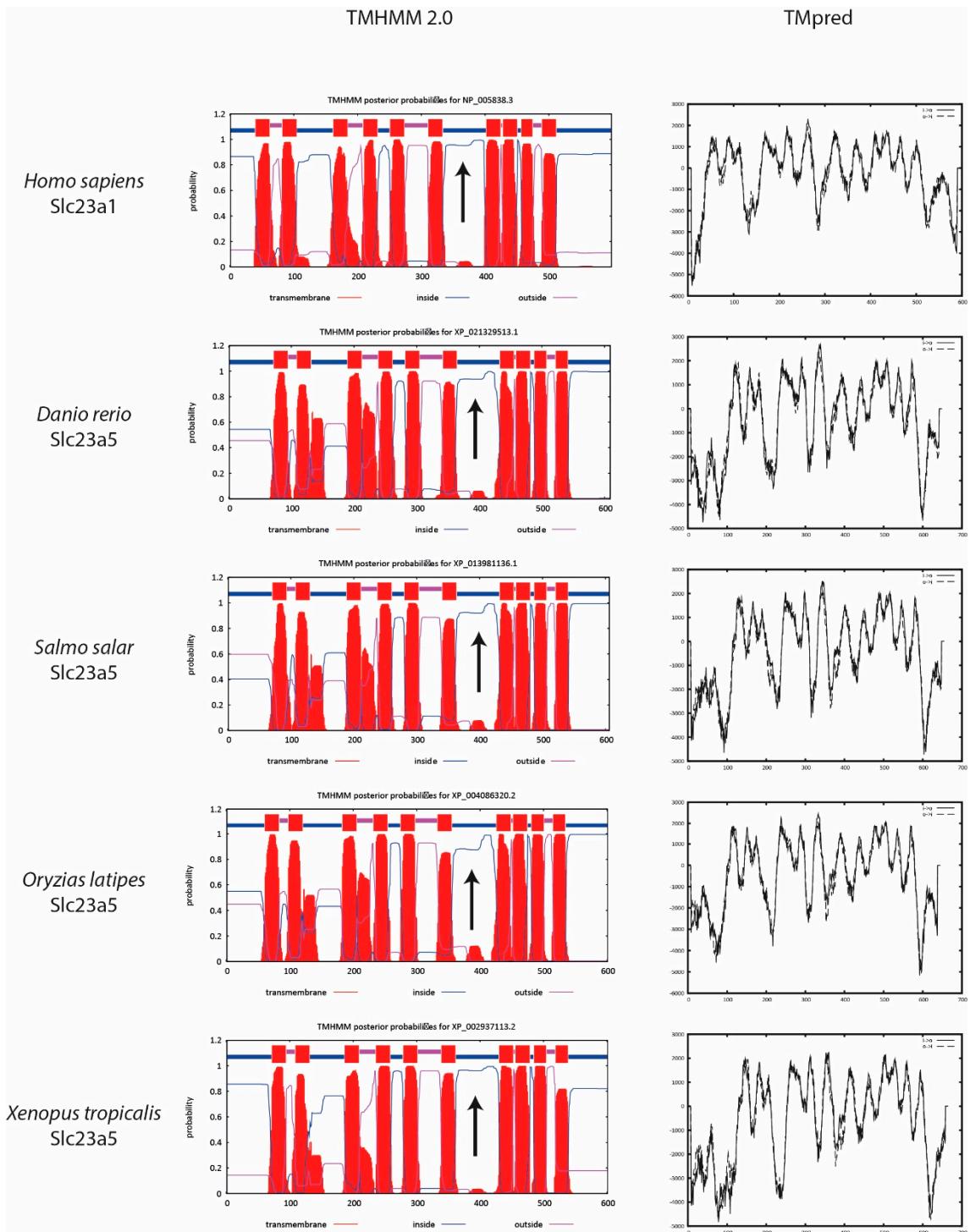
**Figure S5.** Synteny maps of *Slc23a3*. Conservation of the *Slc23a3* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes.



**Figure S6.** Synteny maps of *Slc23a4*. Conservation of the *Slc23a4* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. The human homologue is indicated as pseudogenized (*SLC23A4P*). Grey boxes indicate non-conserved genes.



**Figure S7.** S/T-X-Ø tripeptide conservation in Slc23a1 and Slc23a4 sequences.



**Figure S8.** Topology prediction of Slc23a5 sequences: from *D. rerio*, *O. latipes*, *Salmo salar* (Atlantic salmon) and *X. tropicalis*, and human Slc23a1 used as control. TMHMM Server v. 2.0 (left) and hydropathy plots (right) are provided. Arrows denote the position of NAT family motif.

**Supplementary Tables:**

**Table S1.** Accession numbers.

Species	Gene	Accession Number	Common name
<i>Aspergillus nidulans</i>	UapA	CAA50681	-
<i>Aspergillus nidulans</i>	UapC	CAA56190	-
<i>Bacillus subtilis</i>	PbuX	KIX81273	-
<i>Bacillus subtilis</i>	PucK	KIX81727	-
<i>Escherichia coli</i>	YgfO	CDU41213.1	-
<i>Escherichia coli O25b:H4</i>	UraA	ANK02789	-
<i>Escherichia coli O25b:H4</i>	YicE	ANK04183	-
<i>Zea mays</i>	Lep1	AAB17501	Maize
<i>Branchiostoma belcheri</i>	Slc23-like	XP_019630437	Amphioxus
<i>Branchiostoma belcheri</i>	Slc23-like	XP_019636191	Amphioxus
<i>Branchiostoma belcheri</i>	Slc23-like	XP_019633776	Amphioxus
<i>Branchiostoma belcheri</i>	Slc23-like	XP_019616280	Amphioxus
<i>Branchiostoma belcheri</i>	Slc23-like	XP_019646793	Amphioxus
<i>Branchiostoma floridae</i>	Slc23-like	XP_002600710	Florida lancelet
<i>Branchiostoma floridae</i>	Slc23-like	XP_002597306	Florida lancelet
<i>Branchiostoma floridae</i>	Slc23-like	XP_002595086	Florida lancelet
<i>Branchiostoma floridae</i>	Slc23-like	XP_002611161	Florida lancelet
<i>Anolis carolinensis</i>	Slc23a1	XP_008113637	Green anole
<i>Callorhinus milii</i>	Slc23a1	XP_007904441	Ghost shark
<i>Cyprinus carpio</i>	Slc23a1	XP_018975086	Common carp
<i>Danio rerio</i>	Slc23a1	NP_001166970	Zebrafish
<i>Gallus gallus</i>	Slc23a1	XP_004944825	Chicken
<i>Homo sapiens</i>	Slc23a1	NP_005838	Human
<i>Ictalurus punctatus</i>	Slc23a1	XP_017348904	Channel catfish
<i>Latimeria chalumnae</i>	Slc23a1	XP_005989382	Coelacanth
<i>Lepisosteus oculatus</i>	Slc23a1	XP_015205203	Spotted gar
<i>Lethenteron japonicum</i>	Slc23a1	JL2986	Japanese lamprey
<i>Monodelphis domestica</i>	Slc23a1	XP_016278852	Gray short-tailed opossum
<i>Mus musculus</i>	Slc23a1	NP_035527	House mouse
<i>Oryzias latipes</i>	Slc23a1	XP_004076777	Japanese rice fish
<i>Poecilia formosa</i>	Slc23a1	XP_007556555	Amazon molly
<i>Rhincodon typus</i>	Slc23a1	XP_020385431	Whale shark
<i>Xenopus tropicalis</i>	Slc23a1	XP_002932349	Tropical clawed frog
<i>Anolis carolinensis</i>	Slc23a2	XP_003229574	Green anole
<i>Danio rerio</i>	Slc23a2	XP_021327808	Zebrafish
<i>Gallus gallus</i>	Slc23a2	XP_015152797	Chicken
<i>Homo sapiens</i>	Slc23a2	NP_976072	Human
<i>Latimeria chalumnae</i>	Slc23a2	XP_014352307	Coelacanth
<i>Lepisosteus oculatus</i>	Slc23a2	XP_015197789	Spotted gar
<i>Lethenteron japonicum</i>	Slc23a2	JL15308	Japanese lamprey
<i>Leucoraja erinacea</i>	Slc23a2	ctg10612	Little Skate
<i>Monodelphis domestica</i>	Slc23a2	XP_007476346	Gray short-tailed opossum
<i>Mus musculus</i>	Slc23a2	NP_061294	House mouse
<i>Sarcophilus harrisii</i>	Slc23a2	XP_003757995	Tasmanian devil

<i>Scyliorhinus canicula</i>	Slc23a2	ctg15352	Small-spotted catshark
<i>Xenopus tropicalis</i>	Slc23a2	NP_001120161	Tropical clawed frog
<i>Anolis carolinensis</i>	Slc23a3	XP_008108299	Green anole
<i>Danio rerio</i>	Slc23a3	XP_017213243	Zebrafish
<i>Gallus gallus</i>	Slc23a3	XP_015145565	Chicken
<i>Homo sapiens</i>	Slc23a3	NP_001138362	Human
<i>Latimeria chalumnae</i>	Slc23a3	XP_014343771	Coelacanth
<i>Lepisosteus oculatus</i>	Slc23a3	XP_015214926	Spotted gar
<i>Monodelphis domestica</i>	Slc23a3	XP_007501758	Gray short-tailed opossum
<i>Mus musculus</i>	Slc23a3	NP_919314	House mouse
<i>Xenopus tropicalis</i>	Slc23a3	XP_012826889	Tropical clawed frog
<i>Anolis carolinensis</i>	Slc23a4	XP_016849222	Green anole
<i>Callorhinus milii</i>	Slc23a4	XP_007903282	Elephant shark
<i>Danio rerio</i>	Slc23a4	NP_001013353	Zebrafish
<i>Gallus gallus</i>	Slc23a4	XP_015145041	Chicken
<i>Gekko japonicus</i>	Slc23a4	XP_015276774	Schlegel's Japanese gecko
<i>Latimeria chalumnae</i>	Slc23a4	XP_005991386	Coelacanth
<i>Lepisosteus oculatus</i>	Slc23a4	XP_006633863	Spotted gar
<i>Lethenteron japonicum</i>	Slc23a4	JL7088	Japanese lamprey
<i>Monodelphis domestica</i>	Slc23a4	XP_007504433	Gray short-tailed opossum
<i>Mus musculus</i>	Slc23a4	XP_006506197	House mouse
<i>Oryzias latipes</i>	Slc23a4	XP_004069901	Japanese medaka
<i>Rhincodon typus</i>	Slc23a4	XP_020391140	Whale shark
<i>Sarcophilus harrisii</i>	Slc23a4	XP_012406506	Tasmanian devil
<i>Xenopus tropicalis</i>	Slc23a4	XP_012815547	Tropical clawed frog
<i>Boleophthalmus pectinirostris</i>	Slc23a5	XP_020774843	Great blue-spotted mudskipper
<i>Cynoglossus semilaevis</i>	Slc23a5	XP_016888339	Tongue sole
<i>Cyprinus carpio</i>	Slc23a5	XP_018944039	Common carp
<i>Danio rerio</i>	Slc23a5	XP_009296238	Zebrafish
<i>Fundulus heteroclitus</i>	Slc23a5	XP_021172742	Mummichog
<i>Hippocampus comes</i>	Slc23a5	XP_019750396	Tiger tail seahorse
<i>Ictalurus punctatus</i>	Slc23a5	XP_017328967	Channel catfish
<i>Lates calcarifer</i>	Slc23a5	XP_018548144	Barramundi perch
<i>Lepisosteus oculatus</i>	Slc23a5	XP_015208244	Spotted gar
<i>Maylandia zebra</i>	Slc23a5	XP_014264577	Zebra mbuna
<i>Monopterus albus</i>	Slc23a5	XP_020448997	Swamp eel
<i>Nanorana parkeri</i>	Slc23a5	XP_018412235	-
<i>Neolamprologus brichardi</i>	Slc23a5	XP_006782389	Fairy cichlid
<i>Nothobranchius furzeri</i>	Slc23a5	XP_015819289	Turquoise killifish
<i>Oncorhynchus mykiss</i>	Slc23a5	XP_021432371	Rainbow trout
<i>Oreochromis niloticus</i>	Slc23a5	XP_005454237	Nile tilapia
<i>Oryzias latipes</i>	Slc23a5	XP_004086320	Japanese medaka
<i>Poecilia reticulata</i>	Slc23a5	XP_008408845	Guppy
<i>Pundamilia nyererei</i>	Slc23a5	XP_005731628	-
<i>Pygocentrus nattereri</i>	Slc23a5	XP_017573200	Red-bellied piranha
<i>Salmo salar</i>	Slc23a5	XP_013981136	Atlantic salmon
<i>Scleropages formosus</i>	Slc23a5	XP_018602707	Asian bonytongue

<i>Sinocyclocheilus rhinocerous</i>	Slc23a5	XP_016366936	-
<i>Stegastes partitus</i>	Slc23a5	XP_008289900	Bicolor damselfish
<i>Takifugu rubripes</i>	Slc23a5	XP_003967776	Fugu rubripes
<i>Xenopus tropicalis</i>	Slc23a5	XP_002937113	Tropical clawed frog

**Table S2.** Accession numbers of the RNAseq files. PE – portable executable format.

<b>Spotted gar</b>	<b>Library</b>	<b>SRA Run</b>	<b>Read length</b>
Brain	PE	SRR1524250	100
Intestine	PE	SRR1524257	100
Kidney	PE	SRR1524255	100
Liver	PE	SRR1524254	100
Testis	PE	SRR1524260	100
<b>Zebrafish</b>	<b>Library</b>	<b>SRA Run</b>	<b>Read length</b>
Brain	PE	SRR1524238	100
Intestine	PE	SRR1524245	100
Kidney	PE	SRR1524243	100
Liver	PE	SRR1524242	100
Testis	PE	SRR1524249	100
<b>Chicken</b>	<b>Library</b>	<b>SRA Run</b>	<b>Read length</b>
Brain	PE	SRR924542	100
Intestine	PE	SRR3194289	100
Kidney	PE	SRR924553	100
Liver	PE	SRR924555	100
Testis	PE	SRR924547	100
<b>Human</b>	<b>Library</b>	<b>SRA Run</b>	<b>Read length</b>
Brain	PE	ERR315432	100
Intestine	PE	ERR315419	100
Kidney	PE	ERR315468	100
Liver	PE	ERR315451	100
Testis	PE	ERR315456	100
<b>Mouse</b>	<b>Library</b>	<b>SRA Run</b>	<b>Read length</b>
Brain	PE	SRR5171101	100
Intestine	PE	SRR5171080	100
Kidney	PE	SRR5171094	100
Liver	PE	SRR5171078	100
Testis	PE	SRR5171084	100
<b>Western clawed frog</b>	<b>Library</b>	<b>SRA Run</b>	<b>Read length</b>
Brain	PE	SRR579560	75
Intestine	SE	SRR1186999	50
Kidney	PE	SRR579562	75
Liver	PE	SRR579561	75
Testis	SE	SRR943353	100
<b>Anole lizard</b>	<b>Library</b>	<b>SRA Run</b>	<b>Read length</b>
Brain	PE	SRR579556	75
Kidney	PE	SRR579557	75
Liver	PE	SRR391651	100
Testis	PE	SRR5420127	50
<b>Japanese medaka</b>	<b>Library</b>	<b>SRA Run</b>	<b>Read length</b>
Brain	PE	SRR1524271	100
Intestine	PE	SRR1524278	100
Kidney	PE	SRR1524276	100
Liver	PE	SRR1524275	100
Testis	PE	SRR1628839	90

**Table S3.** Gene identifier for all genes included in gene expression analysis, collected from NCBI and ENSEMBLE Databases.

Organism	Gene Symbol	Gene ID NCBI	Gene ID Ensemble
Human	Slc23a1	9962	ENSG00000170482
	Slc23a2	9963	ENSG00000089057
	Slc23a3	151295	ENSG00000213901
Mouse	Slc23a1	20522	ENSMUSG00000024354
	Slc23a2	54338	ENSMUSG00000027340
	Slc23a3	22626	ENSMUSG00000026205
	Slc23a4	243753	ENSMUSG00000029847
Lizard	Slc23a1	100556489	ENSACAG00000016325
	Slc23a2	100566000	ENSACAG00000005986
	Slc23a3	100567833	ENSACAG00000003143
	Slc23a4	100563862	ENSACAG00000011027
Chicken	Slc23a1	416188	ENSGALG000000037642
	Slc23a2	419520	ENSGALG00000000195
	Slc23a3	429039	Not Found
	Slc23a4	417937	ENSGALG00000011717
Western clawed frog	Slc23a1	100493636	ENSXETG00000008402
	Slc23a2	100145200	ENSXETG00000016853
	Slc23a3	100486175	ENSXETG00000013878
	Slc23a4	100490216	ENSXETG00000020228
	Slc23a5	100486632	Not Found
Zebrafish	Slc23a1	559435	ENSDARG00000015033
	Slc23a2	100321249	ENSDARG00000017365
	Slc23a3	100333020	ENSDARG00000088891
	Slc23a4	503757	ENSDARG00000052497
	Slc23a5	798958	ENSDARG00000045792
Japanese medaka	Slc23a1	101155711	ENSORLG00000012701
	Slc23a2	101156743	ENSORLG00000002889
	Slc23a3	Not Found	Not Found
	Slc23a4 <sup>a</sup>	101173186	ENSORLG00000009915
	Slc23a4b	101157211	ENSORLG00000015991
	Slc23a5	101175057	ENSORLG00000019882
Spotted gar	Slc23a1	102691341	ENSLOC00000012533
	Slc23a2	102691255	ENSLOC00000015630
	Slc23a3	102699040	ENSLOC00000010877
	Slc23a4	102686145	ENSLOC00000016950
	Slc23a5	2682847	ENSLOC00000016369

**Table S4.** Genome and GTF files retrieved from Ensemble database (Release 89) and Transcriptome files retrieved from NCBI used for this study.

Organisms	Genome / Transcriptome Files	GTF FILE	Download date	Source
Mouse	Mus_musculus.GRCm38.dna.toplevel.fa	Mus_musculus.GRCm38.89.gtf	09/08/2017	Ensemble
Spotted gar	Lepisosteus_oculatus.LepOcu1.dna.toplevel.fa	Lepisosteus_oculatus.LepOcu1.89.gtf	09/08/2017	Ensemble
Anole lizard	Anolis_carolinensis.AnoCar2.0.dna.toplevel.fa	Anolis_carolinensis.AnoCar2.0.89.gtf	09/08/2017	Ensemble
Zebrafish	Danio rerio.GRCz10.dna.toplevel.fa	Danio rerio.GRCz10.89.gtf	09/08/2017	Ensemble
Human	Homo_sapiens.GRCh38.dna.toplevel.fa	Homo_sapiens.GRCh38.89.gtf	09/08/2017	Ensemble
Japanese medaka	Oryzias_latipes.MEDAKA1.dna.toplevel.fa	Oryzias_latipes.MEDAKA1.89.gtf	09/08/2017	Ensemble
Western clawed frog	GCF_000004195.3_Xenopus_tropicalis_v9.1_rna.fna	Xenopus_tropicalis_Gene_Transcript_Map.txt*	09/08/2017	Ncbi
Chicken	GCF_000002315.4_Gallus_gallus-5.0_rna.fna	Gallus_gallus_Gene_Transcript_Map.txt*	09/08/2017	Ncbi

**Table S5.** Relative gene expression levels of Slc23 family genes for eight species. The values are presented in log2 (TPM +1) ratios.

Organism	Tissue	Slc23a1	Slc23a2	Slc23a3	Slc23a4a	Slc23a4b	Slc23a5
Human	BRAIN	0,00	5,82	0,00	-	-	-
	INTESTINE	5,85	2,64	4,61	-	-	-
	KIDNEY	3,71	3,70	4,28	-	-	-
	LIVER	4,03	3,58	2,36	-	-	-
	TESTIS	0,00	4,59	0,00	-	-	-
Mouse	BRAIN	0,00	4,91	0,00	0,00	-	-
	INTESTINE	3,73	5,24	1,02	5,89	-	-
	KIDNEY	6,91	2,82	4,08	0,00	-	-
	LIVER	5,06	2,85	0,00	0,00	-	-
	TESTIS	0,00	1,63	0,00	0,00	-	-
Lizard	BRAIN	0,91	2,82	0,72	2,49	-	-
	KIDNEY	6,76	1,39	6,10	1,19	-	-
	LIVER	3,02	0,00	0,00	0,65	-	-
	TESTIS	0,00	0,00	0,00	0,00	-	-
Chicken	BRAIN	0,00	6,27	0,96	1,82	-	-
	INTESTINE	6,44	3,25	2,96	6,09	-	-
	KIDNEY	7,78	1,49	4,17	1,60	-	-
	LIVER	5,20	1,98	1,76	0,82	-	-
	TESTIS	2,35	2,08	2,03	1,85	-	-
Western clawed frog	BRAIN	0,00	6,45	0,00	2,54	-	0,00
	INTESTINE	0,00	2,23	0,00	5,65	-	0,00
	KIDNEY	7,63	4,08	6,53	5,00	-	5,40
	LIVER	0,00	3,41	0,00	2,14	-	0,00
	TESTIS	2,69	4,01	0,00	1,49	-	1,64
Zebrafish	BRAIN	0,85	3,73	0,71	0,00	-	1,01
	INTESTINE	2,69	1,21	0,00	4,00	-	0,00
	KIDNEY	5,35	2,27	3,42	4,78	-	0,93
	LIVER	0,00	1,26	0,00	0,00	-	0,00
	TESTIS	0,93	2,74	0,00	0,63	-	1,97
Japanese medaka	BRAIN	0,00	4,70	-	0,00	0,00	0,84
	INTESTINE	2,76	2,91	-	4,95	0,00	0,00
	KIDNEY	3,89	3,02	-	4,43	0,00	0,00
	LIVER	0,00	4,06	-	0,00	0,00	0,00
	TESTIS	0,00	3,20	-	0,60	0,00	0,00
Spotted gar	BRAIN	0,58	4,94	0,00	0,00	-	0,00
	INTESTINE	5,67	0,96	0,00	0,00	-	0,00
	KIDNEY	3,18	0,90	1,78	0,86	-	0,00
	LIVER	0,74	0,78	0,00	0,00	-	0,00
	TESTIS	1,92	7,21	0,00	0,00	-	0,00

**Table S6.** Renal uric acid absorption and secretion: the urate transportome. n.f. – not found.

	Transporter	<i>H. sapiens</i>	<i>M. musculus</i>	<i>G. gallus</i>	<i>A. carolinensis</i>	<i>X. tropicalis</i>	<i>D. rerio</i>	<i>L. oculatus</i>
<b>Basolateral Absorption</b>	SLC2A9/GLUT9	✓	✓	✓	✓	✓	✓	✓
<b>Apical Absorption</b>	SLC22A11/OAT4	✓	n.f.	n.f.	n.f.	n.f.	n.f.	n.f.
	SLC22A12/URAT1	✓	✓	n.f.	n.f.	n.f.	n.f.	n.f.
	SLC22A13/OAT10	✓	✓	✓	✓	✓	✓	✓
<b>Basolateral Secretion</b>	SLC22A6/OAT1	✓	✓	✓	✓	✓	n.f.	n.f.
	SLC22A8/OAT3	✓	✓	✓	✓	✓	n.f.	n.f.
<b>Apical Secretion</b>	SLC17A1/NPT1	✓	✓	n.f.	n.f.	n.f.	n.f.	n.f.
	SLC17A3/NPT4	✓	✓	n.f.	n.f.	n.f.	n.f.	n.f.