A taxonomically-restricted Uric Acid transporter provides insight into antioxidant equilibrium

Diogo Oliveira*[,1], André Machado*[,1], Tiago Cardoso[1], Mónica Lopes-Marques[1], L. Filipe C. Castro[1,2]• and Raquel Ruivo[1]•

[1]CIIMAR – Interdisciplinary Centre of Marine and Environmental Research, U. Porto – University of Porto, Porto, Portugal

[2]Department of Biology, Faculty of Sciences, U. Porto - University of Porto, Portugal

*Equal contribution

•Corresponding authors at: CIIMAR, Terminal de Cruzeiros do Porto de Leixões, Av. General Norton de Matos s/n, 4450-208 Matosinhos, Portugal. Tel.: +351 223 401 831

Running head: Novel uric acid transporter in fish and amphibians

Abstract

Nucleobase-Ascorbate Transporter (NAT) family includes ascorbic acid, nucleobases and uric acid transporters, with a broad evolutionary distribution. In vertebrates, four members have been previously recognized, the ascorbate transporters Slc23a1 and Slc3a2, the nucleobase transporter Slc23a4 and an orphan transporter SLC23A3. Here we identify a fifth member of the vertebrate *slc23* complement (*slc23a5*), expressed in neopterygians (gars and teleosts) and amphibians, and clarify the evolutionary relationships between the novel gene and known *slc23* genes. Further comparative analysis puts forward uric acid as the preferred substrate for Slc23a5. Gene expression quantification suggests kidney and testis as major expression sites in *Xenopus tropicalis* (western clawed frog) and *Danio rerio* (zebrafish). Additional expression in brain was detected in *D. rerio*, while in the Neoteleostei *Oryzias latipes* (medaka) *slc23a5* expression is restricted to brain. The biological relevance of the retention of an extra transporter in fish and amphibians is examined: with respect to the (1) antioxidant role of uric and ascorbic acid in seminal fluid and brain, (2) the ability to endogenously synthesize ascorbic acid and (3) the morphological adaptations of the male urogenital system.

Keywords: Nucleobase-Ascorbate Transporters, Slc23a5, Uric acid, Antioxidant, Gene duplication, Molecular evolution, Comparative studies

41 **Introduction**

42 Ascorbic acid (or its salt form ascorbate) is an essential enzyme cofactor,
43 participating in collagen and norepinephrine synthesis, as well as a potent antioxidant
44 and free radical scavenger (Burzle et al., 2013, Linster & Van Schaftingen, 2007). Despite
45 its importance, anthropoid primates, teleost fish, Guinea pigs, some bats and
46 Passeriformes bird species, lack a functional L-gulono-ϒ-lactone oxidase (*GLO*) gene and,
47 thus, the ability to catalyse the last step of ascorbic acid synthesis (Drouin et al., 2011,
48 Dutta Gupta et al., 1973). The scattered loss of the *GLO* gene was suggested to evolve
49 neutrally, with ascorbic acid requirements being met by dietary intake (Drouin et al.,
50 2011). Unlike ascorbate, uric acid is largely regarded as end-product of purine
51 metabolism and metabolic waste (Wright, 1995). Yet, human-biased research has
52 suggested alternative physiological functions for uric acid or urate (uric acid salts)
53 (Bobulescu & Moe, 2012). Due to loss of the *uricase* (UOX) enzyme in hominids, which
54 converts uric acid from hepatic nucleotide metabolism into the excretion product
55 allantoin, humans have the highest uric acid levels among animals (Wu et al., 1992,
56 Bobulescu & Moe, 2012, Anzai et al., 2012, Alvarez-Lario & Macarrón-Vicente, 2001, Oda
57 et al., 2002). In hominids, the fate of *UOX* followed *GLO* silencing, suggestive of an
58 increased selective pressure for an alternative anti-oxidant system (Bobulescu & Moe,
59 2012, Alvarez-Lario & Macarrón-Vicente, 2001). In humans, uric acid is thus proposed to
60 act as antioxidant and free radical scavenger, similarly to ascorbic acid, counterbalancing
61 disease-related oxidative stress (Bobulescu & Moe, 2012, Alvarez-Lario & Macarrón-
62 Vicente, 2001, Oda et al., 2002)

63 Ascorbate occurs in two forms. Uptake and cellular trafficking of the oxidized form
64 (dehydroascorbic acid) involves the facilitative hexose transporters comprising the

65 glucose transporters (GLUT1, GLUT3 and GLUT4), belonging to the Solute Carrier family
66 2 (SLC2) (Burzle et al., 2013, Corti et al., 2010). Ascorbic acid levels (reduced form), on
67 the other hand, are maintained by an active mechanism that requires sodium-
68 dependent transporters from the SLC23 family: Slc23a1 responsible for renal and
69 intestine ascorbic acid absorption across brush border epithelia and Slc23a2, which
70 exhibits a widespread expression and mediates ascorbic acid supply, from blood and
71 fluids, to target tissues and cells, including neurons (Burzle et al., 2013, Corti et al., 2010,
72 Savini et al., 2008). Slc23a1 and Slc23a2 belong to the Nucleobase-Ascorbate
73 Transporter (NAT) family, which has a very broad evolutionary distribution, from
74 bacteria to metazoans, and displays a highly conserved signature motif defining
75 substrate specificity: [Q/E/P]-N-x-G-x-x-x-x-T-[R/K/G] (Burzle et al., 2013, Koukaki et al.,
76 2005, Papakostas & Frillingos, 2012). In vertebrates, two additional transporters were
77 identified, the orphan *transporter* Slc23a3 and the nucleobase transporter Slc23a4,
78 described in rats but found pseudogenized in humans (Burzle et al., 2013, Yamamoto et
79 al., 2010). Rat Slc23a4 is restricted to the small intestine and was shown to
80 predominantly mediate sodium-dependent uracil transport (Yamamoto et al., 2010).
81 Human Slc23a3, on the other hand, is expressed in proximal renal tubules but is
82 unresponsive to both ascorbic acid and nucleobases (Burzle et al., 2013). Besides the
83 described animal transporters, the NAT family includes several transport proteins from
84 bacteria, fungi and plants, which can be divided into three functional groups (Figure 1).
85 Slc23a1, Slc23a2 and Slc23a3 belong to the ascorbate group and harbour a conserved
86 proline (P) in the first amino acid position of the signature motif (Burzle et al., 2013). A
87 glutamate (E) residue defines the uracil group, including bacterial anion symporters and
88 Slc23a4, while a glutamine (Q) is present in the bacterial uric acid and/or xanthine anion

89 symporters (Koukaki et al., 2005, Yamamoto et al., 2010, Burzle et al., 2013). Here we
90 identify a novel member of the NAT family, *slc23a5*, expressed in neopterygian fishes
91 (gars and teleosts) and amphibians. Combining relative gene expression patterns and
92 comparative analyses of protein motifs, we put forward the functional relevance of this
93 taxonomically restricted gene in the context of oxidative stress in the aquatic
94 environment.
95
96 **Materials and Methods**
97 *Sequence mining*
98 Slc23 protein sequences were obtained from the GenBank database by BLASTp
99 using human Slc23a1, Slc23a2 and Slc23a3 sequences as query. The sequences from
100 *Lethenteron japonica* (Japanese lamprey) were obtained from the Japanese Lamprey
101 Genome Project. The sequences from *Leucoraja erinacea* and *Scyliorhinus canicula* were
102 obtained as nucleotide sequences from SkateBase and translated. Accession numbers
103 for all the retrieved sequences are listed in Supplementary Table 1 (Table S1).
104
105 *Synteny and paralogy*
106 Synteny and paralogy maps were constructed based on information about the
107 genomic location of the *Slc23* members' flanking genes from the GenBank database.
108 Paralogs were identified according to the information available on Ensembl, and
109 ancestral chromosomal partitioning was established as per reference (Nakatani et al.,
110 2007, Putnam et al., 2008).
111
112 *Phylogenetic Analysis*

113 The retrieved sequences were aligned using the MAFFT software web service
114 (http://mafft.cbrc.jp/alignment/software/) in default automatic settings, with L-INS-I
115 refinement method (Katoh & Standley, 2013). The output alignment was used to
116 construct a phylogenetic tree using PhyML 3.0 with Smart Model Selection web service
117 (http://www.atgc-montpellier.fr/phyml-sms/) in default settings, which selected the LG
118 +G+I+F model (Guindon et al., 2010).
119
120 *Membrane topology prediction*
121 Slc23a1 and Slc23a5 membrane topology prediction was performed using the
122 CCTOP prediction web server (http://cctop.enzim.ttk.mta.hu/) (Dobson et al., 2015).
123
124 *Filtering and quality control of SRA datasets*
125 For Rna-seq analysis, we included 39 SRA datasets from brain, intestine, kidney,
126 liver, and testis tissues of Human (*Homo sapiens*), mouse (*Mus musculus*), chicken
127 (*Gallus gallus*), lizard (*Anolis carolinensis*), western clawed frog (*Xenopus tropicalis*),
128 zebrafish (*Danio rerio*), spotted gar (*Lepisosteus oculatus*) and Japanese medaka (*Oryzias*
129 *latipes*), which were obtained from the National Center for Biotechnology (NCBI)
130 Sequence Read Archive (*SRA)* (https://www.ncbi.nlm.nih.gov/sra/) (Table S2). The read
131 quality assessment of each SRA dataset was performed using FastQC software
132 (https://www.bioinformatics. Babraham.ac.uk/projects/fastqc/). Trimmomatic 0.36
133 (Bolger et al., 2014) was used to trim and drop reads with quality scores below 5, at
134 leading and trailing ends, with an average quality score below 20 in a 4 bp sliding window
135 and with less than 36 bases length.
136

137 *Relative gene expression levels*

138 To assess the relative gene expression of *slc23* genes from the selected species,
139 reference sequences and the corresponding annotations were collected from NCBI and
140 Ensemble (Yates et al., 2016) databases. For human, mouse, lizard, zebrafish and
141 spotted gar the GTF and DNA files (genome) were retrieved from Ensemble (Release 89)
142 (Yates et al., 2016). For chicken and western clawed frog transcript sequences (coding
143 and non-coding) and mapping (gene/transcript) from NCBI Refseq database were used
144 after filtering (O'Leary et al., 2016) (for each species the rna.fasta file from
145 ftp://ftp.ncbi.nih.gov/genomes/refseq/ was used – See Table S3 for detail). The SRA
146 reads of each specimen were aligned to the reference using Bowtie2 (default
147 parameters)(Langmead & Salzberg, 2012), and gene expression was quantified as
148 transcript per million (TPM) using RSEM v.1.2.31 (RNA-seq by Expectation Maximization)
149 software package (Li & Dewey, 2011). Isoform expression levels for each gene were
150 summed to derive the TPM values. For genes with low relative expression values TPM
151 values were $\log_2$-transformed after adding a value of one; TPM values < 0.5 were
152 considered unreliable and substituted with zero.

153

154 *PCR-based Slc23a5 expression*

155 Gene expression of *Slc23a5* was further analyzed by PCR for *Xenopus tropicalis* and
156 *Danio rerio* using the primer pairs 5'-AGACGTTGAAGCCGTTACTGA-3' and 5'-
157 TGCTATTATGCCTCCAAAGGCT-3' in the former and 5'-GGTGGAATGTTCCTCGTCAT-3' and
158 5'-GATGAACATGTGGGTGGTGA-3' in the latter. These primers map to a small gene
159 region flanking an intron. PCR reactions were carried out on a tissue panel containing

160 heart, skin, stomach, intestine, liver, kidney, ovary and testis tissues for *X. tropicalis* and
161 heart, skin, gut, liver, kidney, ovary and testis tissues for *D. rerio*.
162
163 **Results and Discussion**
164 Through database mining and phylogenetic analysis, we detected the presence of
165 a novel gene, *slc23a5*, found in neopterygians (Holostei and Teleostei) and amphibians
166 (Fig 1). No *slc23a5* sequence was retrieved from chondrichthyans, chondrosteans (basal
167 actinopterygians) nor from the sarcopterygian *Latimeria chalumnae* (coelacanth).
168 Secondary loss of *slc23a5* in chondrichthyans was further supported by the conservation
169 of the genomic *locus* in *Callorhynchus milii* (Australian ghostshark) (Fig S1). Regarding
170 coelacanth, two scaffolds, denoting conservation of the flanking genes, were found;
171 while, no information was retrieved regarding the chondrostean *loci* (Fig S1). The
172 remaining genes of the *slc23* family were retrieved and genomic *loci* were found
173 generally conserved across species, with exception of the *slc23a4 loci* in fish (Fig S2-S5).
174 The branching pattern of our phylogenetic analysis suggests that this novel *slc23a5* gene
175 emerged, along with the *slc23a4* paralogue, from an ancestral duplication and that
176 *slc23a5* was secondarily lost in sharks, basal actinopterygians, coelacanth and amniotes
177 (Fig 1 and Fig S1). In agreement, we obtained three phylogenetic clusters possibly
178 corresponding to three ancestral lineages (Fig 1): one composed of typical *slc23a1* and
179 *slc23a2* genes and a second group including the previously reported s*lc23a4* and the
180 novel *slc23a5* genes, both outgrouped by invertebrate sequences. A third cluster, which
181 outgroups the remaining *slc23* genes, included the highly divergent *slc23a3* genes,
182 mirrored by the long branch lengths (Fig 1). To confirm ancestral duplication events we
183 conducted paralogy analysis of the human *loci* using two chromosomal partitioning

184 approaches (Putnam et al., 2008, Nakatani et al., 2007). While some reorganization is
185 observed, our combined analysis seems to support a 2R origin for the *slc23a1*/*slc23a2*
186 and *slc23a4*/*slc23a5* gene *loci*, while *slc23a3* might stand as an independent lineage (Fig
187 S6). Also, human paralogues of the genes flanking an invertebrate *slc23*-like sequence,
188 from *Branchiostoma belcheri*, were mapped back to the same chromosomal partition as
189 the human *slc23a3* sequence, corroborating the existence of a *slc23a3*-like invertebrate
190 gene (Fig S6). Further comparative protein sequence analysis highlighted the presence
191 of a NAT family signature motif with a conserved Q in substrate determining position,
192 placing Slc23a5 in the uric acid and/or xanthine group (Fig 2). Besides the typical
193 signature motif, a carboxyl terminus tripeptide motif was identified in some SLC23
194 family members: a S/T-X-Ø tripeptide shown to interact PDZ domains of scaffolding
195 proteins belonging to the Sodium/Hydrogen Exchanger Regulatory Factor, NHERF,
196 family (Anzai et al., 2012). NHERF family members, such as PDZ-Containing Kidney
197 Protein 1 (PDZK1), are mostly expressed in the apical side of absorptive epithelia of
198 kidney, intestine and liver, and establish protein-protein interaction networks with a
199 wide range of proteins, including receptors, transporters and channels (Dev, 2004,
200 Clapéron et al., 2011). Besides its scaffolding role, NHERF family members were also
201 shown to modulate the activity of protein assemblages (Clapéron et al., 2011, Raghuram
202 et al., 2001, Wang et al., 2000). Curiously, S/T-X-Ø are found highly conserved across
203 species in Slc23a1 and Slc23a4, in agreement with their renal and intestinal expression
204 and role in ascorbic acid and nucleobase absorption. Slc23a5, on the other hand, exhibits
205 a highly conserved tripeptide motif in Otocephala (Cypriniformes, Characiformes and
206 Siluriformes) but not Osteoglossomorpha (Fig 3). The Euteleostei Esociformes and
207 Salmoniformes exhibit an inverted motif while in the Neoteleostei no conservation was

208 found. Finally, topology predictions suggest a conserved cellular import topology, similar
209 to other family members (Fig S7).

210 Our findings indicate a restricted distribution of *slc23a5* to *taxa* that share an
211 aquatic nature of their habitat. Why would these retain a transporter for uric acid, or its
212 precursor, xanthine? Here we hypothesize that, beyond its metabolic waste nature, uric
213 acid may act as an antioxidant in fish and amphibians: notably in *GLO*-deficient teleost.
214 To gain further insight into the functional relevance of this additional gene, we collected
215 RNA-seq data and determined relative expression levels of *Slc23* genes. Gene expression
216 analysis suggested that, in *Danio rerio* (zebrafish) and *Xenopus tropicalis* (western
217 clawed frog), *slc23a5* is mostly expressed in kidney and testis (Fig 3 and Tables S4 and
218 S5). Expression in brain was also detected in *D. rerio* (zebrafish) and was found to be the
219 exclusive expression site in the Neoteleostei *Oryzias latipes* (medaka). The holostean
220 *Lepisosteus oculatus* (spotted gar), on the other hand, yielded no significant expression
221 in the tested tissues. Gene expression levels were further confirmed by PCR in *D. rerio*
222 and *X. tropicalis*. The gonadal expression provides a relevant clue for the retention of
223 the extra isoform and supports the antioxidant role of uric acid. In external fertilizers,
224 such as most fish and amphibians, reactive oxygen species were suggested to decrease
225 motility and thus fertilisation success (Lahnsteiner & Mansour, 2010). In agreement,
226 high concentrations of uric acid were detected, and suggested to act as the main non-
227 enzymatic antioxidant mechanism, in seminal fluid of several fish species: including
228 Cypriniformes, Esociformes, Salmoniformes and the Neoteleostei Gadiformes and
229 Perciformes (Ciereszko et al., 1999, Dzyuba et al., 2014b, Lahnsteiner & Mansour, 2010,
230 Lahnsteiner et al., 2010). Yet, regarding amphibians, no information on seminal fluid
231 composition was retrieved (Keogh et al., 2017). Nonetheless, *in vitro* assays suggested

232 that other antioxidant species, such as ascorbic acid or vitamin E, were unsuitable at
233 offsetting oxidative stress and improving sperm motility in *Litoria booroolongensis*
234 (Booroolong frog) and suggested uric acid as a pertinent candidate (Keogh et al., 2017).
235 In this context, the absence of *slc23a5* sequences in Chondrichthyes is possibly due to
236 their internal fertilization strategies derived from ovoviviparous or viviparous
237 reproductive modes (Goodwin et al., 2002). Accordingly, viviparity is also observed in
238 the sarcopterygian *L. chalumnae* (Goodwin et al., 2002).

239 Besides reproductive categories, other morphological and metabolic traits can
240 justify the seemingly scattered retention of *slc23a5* among fish; for instance, specific
241 anatomical features such as the presence of a urogenital connection: described in
242 Elasmobranchii and in most non-teleost Osteichthyes, yet absent in Cyclostomata and
243 Teleostei (Blün, 1986). By connecting testicular and renal ducts, antioxidant
244 requirements could be met by renal secretion into urine. In agreement, in the
245 Chondrostei *Acipenser ruthenus*, full sperm maturation and motility was shown to
246 require dilution by urine (Dzyuba et al., 2014a). An urogenital connection is also present
247 in amphibians (Blün, 1986). Additionally, other antioxidant mechanisms could exist.
248 Besides uric acid, ascorbic acid is also present in seminal fluid of *Oncorhynchus mykiss*
249 (rainbow trout) (Ciereszko & Dabrowski, 1995). Seminal ascorbate was reported to be
250 affected by seasonal variations and semen quality was shown to improve upon dietary
251 ascorbic acid supplementation (Ciereszko & Dabrowski, 2000, Ciereszko & Dabrowski,
252 1995), contrasting with the amphibian *L. booroolongensis* (Keogh et al., 2017). In the
253 case of non-teleost fish, such as Chondrostei (*A. ruthenus*) and Holostei (*L. oculatus*), a
254 functional *GLO* gene is retained, providing endogenously synthesised ascorbate
255 (Moreau & Dabrowski, 2000). Curiously, no gonadal expression was detected in the

256 Neoteleostei *O. latipes*. Neoteleosts exhibit varied reproduction modes ranging from
257 external fertilizers to internal fertilizers with viviparity (Forsman et al., 2015, Bone &
258 Moore, 2008, Goodwin et al., 2002, Scott & Fuller, 1976). Also, neoteleosts were shown
259 to exhibit a distinctive gonadal morphology, the lobular testis, whereas the remaining
260 bony fish exhibit tubular testis (Parenti & Grier, 2004). Although still obscure, the
261 absence of testicular *slc23a5* expression in the externally fertilizer *O. latipes* could be
262 due to specific morphological adaptations of neoteleosts.

263 Besides the gonadal expression, we detected significant expression of *slc23a5* in
264 teleost brain (Fig 3). Again, an antioxidant role can be postulated for neuronal uric acid
265 (Uttara et al., 2009). Active swimming behaviour requires a high metabolic rate,
266 mirrored by an increased oxygen consumption, subsequently leading to reactive oxygen
267 species production (Lopez-Cruz et al., 2011). In agreement, fish antioxidant levels were
268 shown to be correlated with their swimming activity (Lopez-Cruz et al., 2011). Neuronal
269 cell populations are particularly sensitive to oxidative stress, thus, uric acid could
270 participate in the neutralization of oxidative species to by-pass oxidative stress, notably
271 in *GLO*-deficient fish. Yet, information on neuronal uric acid levels in fish is currently
272 absent.

273 Additionally, renal Slc23a5, detected in *D. rerio* and *X. tropicalis* could participate
274 in the maintenance of uric acid, or xanthine, levels for antioxidant purposes, osmotic
275 regulation and/or excretion. Kidneys play a crucial role in the maintenance of both
276 ascorbate and urate levels through the excretion and reabsorptions processes (Anzai et
277 al., 2012, Bobulescu & Moe, 2012). Regarding urate, several transporters, localized to
278 specific cellular regions (apical or basolateral), seem to cooperate for efficient urate
279 handling in the human proximal renal tubules: the urate transportome (Anzai et al.,

280 2012, Bobulescu & Moe, 2012). In kidney absorptive epithelium, transporters located on
281 the filtrate-epithelial interface (apical side) interact with the scaffolding protein PDZK1
282 (Anzai et al., 2012). A PDZK1 interacting motif was found in Slc23a5 highly conserved in
283 Cypriniformes, Characiformes and Siluriformes, but not in other fish groups or
284 amphibians (Fig 4). No relationship between salt and freshwater environments was
285 found. Thus, the presence/absence of this peptide signal might reflect distinct uric acid
286 flows in the kidney: an apically expressed Slc23a5 likely participates in absorptive uric
287 acid flows, while a basolateral expression would be related to uric acid secretion (Fig 4).
288 In agreement, *X. tropicalis*, *D. rerio* and *L. oculatus* appear to lack the full complement
289 of renal urate transporters (Table S6). A similar scenario is observed for birds and
290 reptiles; however, their excretory system represents a particular adaptation for uric acid
291 excretion (Braun, 1998, Singer, 2003). With this in mind, *X. tropicalis* and *D. rerio*
292 illustrate the two opposing mechanisms: while *D. rerio* captures, *X. tropicalis* secretes
293 uric acid into renal filtrate. Besides simple excretion, and given the renal-testicular
294 morphology of amphibians, uric acid secretion into urine could also promote antioxidant
295 build-up for reproductive purposes.

296 Curiously, despite its inability to mediate transport, Slc23a3 is also expressed in
297 proximal renal tubules, along Slc23a1 and Slc23a5 (Burzle et al., 2013). Comparative
298 sequence analysis places this orphan *transporter* in the ascorbic acid group, due to
299 conservation of the P in the first amino acid position of the NAT family signature motif,
300 crucial for substrate selection. Yet, significant sequence variation was observed in
301 Teleostei and Sauropsida (Fig 1). A conserved, yet apparently disruptive, Q was also
302 evident in the last amino acid position of the signature motif: as a mirror image of the
303 conserved P (Fig 1). An interesting, yet untested, hypothesis suggests Slc23a3 as a sensor

304 for downstream regulation of transportome expression. Slc23a1 expression is, indeed,
305 modulated by ascorbic acid levels (Burzle et al., 2013). Transporter to transreceptor
306 shifts, leading to nutrient sensing by transporter homologues, were previously
307 acknowledged in yeast (Kriel et al., 2011, Karhumaa et al., 2010, Poulsen et al., 2005,
308 Wu et al., 2006). Thus, renal ascorbic and uric acid handling could share common
309 mechanistic features and future studies should address the cooperative handling, if any,
310 of uric and ascorbic acid in kidney proximal tubules.

311 Taken together, these observations put forward the novel Slc23a5 as a uric acid
312 transporter and support the antioxidant role of uric acid, notably in GLO-deficient fish.
313 The retention of an additional SLC23 gene, with presumed uric acid transport function,
314 seems to parallel antioxidant requirements for (1) aquatic reproductive strategies, by
315 maintaining semen integrity in external fertilizers, and (2) to counterbalance increasing
316 metabolic rates, notably in the brain of active swimmers. These observations further
317 support the (1) role of both ascorbic and uric acids in the maintenance of an optimum
318 body redox potential, and the (2) need to equilibrate antioxidant scavenging
319 requirements. Additionally, our results emphasize the morphological and physiological
320 diversity observed in kidneys and gonads, as well as the distinct levels of elaboration and
321 interaction between these ontogenically related organs across evolutionary groups. In
322 aquatic animals, and developing amniotes, excretory functions are mainly orchestrated
323 by the mesonephric kidney (Blün, 1986). During amniote ontogeny, the mesonephros
324 relays its excretory function to the metanephros, with the mesonephros giving rise to a
325 separate seminal duct (Blün, 1986). Within aquatic animals, distinct levels of complexity
326 were described regarding urine and semen transportation: with contrasting urogenital
327 connections between teleosts and non-teleosts (Blün, 1986). Thus, although uric acid

does seem beneficial for sperm motility and survival in aquatic animals, the means by which this antioxidant, and other antioxidant mechanisms, are maintained and conveyed to the seminal fluid appears diverse and should be further explored.

Supplementary material: Supplementary tables and figures including synteny and paralogy analysis and relative gene expression levels, as well as reference tables.

Acknowledgments

This work was supported by the MarInfo – Integrated Platform for Marine Data Acquisition and Analysis (reference NORTE-01-0145-FEDER-000031), a project by the North Portugal Regional Operational Program (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF) and by the European Regional Development Fund (ERDF) through the COMPETE - Operational Competitiveness Programme and POPH - *Operational Human Potential Programme*.

345 **References**

346 Alvarez-Lario, B. & Macarrón-Vicente, J. (2001) Evolution of Uric Acid Metabolism in Humans.
347 In: *eLS.* pp. John Wiley & Sons, Ltd.
348 Anzai, N., Jutabha, P., Amonpatumrat-Takahashi, S. & Sakurai, H. 2012. Recent advances in renal
349 urate transport: characterization of candidate transporters indicated by genome-wide
350 association studies. *Clin. Exp. Nephrol.* **16**: 89-95.
351 Blün, V. (1986) Comparative Anatomy of the Urogenital System. In: *Vertebrate Reproduction, a*
352 *Textbook.* pp. 43-74. Springer-Verlag, Berlin Heidelberg New York Tokyo.
353 Bobulescu, I. A. & Moe, O. W. 2012. Renal transport of uric acid: evolving concepts and
354 uncertainties. *Adv. Chronic Kidney Dis.* **19**: 358-71.
355 Bolger, A. M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina
356 sequence data. *Bioinformatics* **30**: 2114-20.
357 Bone, Q. & Moore, R. H. 2008. *Biology of Fishes* Taylor & Francis Group, New York.
358 Braun, E. J. 1998. Comparative renal function in reptiles, birds, and mammals. *Semin. Avian*
359 *Exotic. Pet. Med.* **7**: 62-71.
360 Burzle, M., Suzuki, Y., Ackermann, D., Miyazaki, H., Maeda, N., Clemencon, B., et al. 2013. The
361 sodium-dependent ascorbic acid transporter family SLC23. *Mol. Aspects Med.* **34**: 436-
362 54.
363 Ciereszko, A. & Dabrowski, K. 1995. Sperm quality and ascorbic acid concentration in rainbow
364 trout semen are affected by dietary vitamin C: an across-season study. *Biol. Reprod.* **52**:
365 982-8.
366 Ciereszko, A. & Dabrowski, K. 2000. Effect of ascorbic acid supplement in vitro onrainbow trout
367 sperm viability. *Aquacult. Int.* **8**: 1-8.
368 Ciereszko, A., Dabrowski, K., Kucharczyk, D., Dobosz, S., Goryczko, K. & Glogowski, J. 1999. The
369 presence of uric acid, an antioxidantive substance, in fish seminal plasma. *Fish Physiol.*
370 *Biochem.* **21**: 313-315.
371 Clapéron, A., Mergey, M. & Fouassier, L. 2011. Roles of the scaffolding proteins NHERF in liver
372 biology. *Clin. Res. Hepatol. Gastroenterol.* **35**: 176-181.
373 Corti, A., Casini, A. F. & Pompella, A. 2010. Cellular pathways for transport and efflux of
374 ascorbate and dehydroascorbate. *Arch. Biochem. Biophys.* **500**: 107-15.
375 Dev, K. K. 2004. Making protein interactions druggable: targeting PDZ domains. *Nat. Rev. Drug*
376 *Discov.* **3**: 1047-1056.
377 Dobson, L., Remenyi, I. & Tusnady, G. E. 2015. CCTOP: a Consensus Constrained TOPology
378 prediction web server. *Nucleic Acids Res.* **43**: W408-12.
379 Drouin, G., Godin, J. R. & Page, B. 2011. The genetics of vitamin C loss in vertebrates. *Curr.*
380 *Genomics* **12**: 371-8.
381 Dutta Gupta, S., Choudhury, P. K. & Chatterjee, I. B. 1973. Synthesis of l-ascorbic acid from d-
382 glucurono-1,4-lactone conjugates by different species of animals. *Int. J. Biochem.* **4**: 309-
383 314.
384 Dzyuba, B., Cosson, J., Boryshpolets, S., Bondarenko, O., Dzyuba, V., Prokopchuk, G., et al. 2014a.
385 In vitro sperm maturation in sterlet, Acipenser ruthenus. *Reprod. Biol.* **14**: 160-3.
386 Dzyuba, V., Dzyuba, B., Cosson, J., Boryshpolets, S., Yamaner, G., Kholodniy, V., et al. 2014b. The
387 antioxidant system of sterlet seminal fluid in testes and Wolffian ducts. *Fish Physiol.*
388 *Biochem.* **40**: 1731-9.
389 Forsman, A., Tibblin, P., Berggren, H., Nordahl, O., Koch-Schmidt, P. & Larsson, P. 2015. Pike Esox
390 lucius as an emerging model organism for studies in ecology and evolutionary biology:
391 a review. *J. Fish Biol.* **87**: 472-479.
392 Goodwin, N. B., Dulvy, N. K. & Reynolds, J. D. 2002. Life-history correlates of the evolution of live
393 bearing in fishes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci* **357**: 259-267.

394 Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. 2010. New
395 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
396 performance of PhyML 3.0. *Syst. Biol.* **59**: 307-21.
397 Karhumaa, K., Wu, B. & Kielland-Brandt, M. C. 2010. Conditions with high intracellular glucose
398 inhibit sensing through glucose sensor Snf3 in Saccharomyces cerevisiae. *J. Cell.*
399 *Biochem.* **110**: 920-5.
400 Katoh, K. & Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7:
401 improvements in performance and usability. *Mol. Biol. Evol.* **30**: 772-80.
402 Keogh, L. M., Byrne, P. G. & Silla, A. J. 2017. The effect of antioxidants on sperm motility
403 activation in the Booroolong frog. *Anim. Reprod. Sci.* **183**: 126-131.
404 Koukaki, M., Vlanti, A., Goudela, S., Pantazopoulou, A., Gioule, H., Tournaviti, S., et al. 2005. The
405 nucleobase-ascorbate transporter (NAT) signature motif in UapA defines the function of
406 the purine translocation pathway. *J. Mol. Biol.* **350**: 499-513.
407 Kriel, J., Haesendonckx, S., Rubio-Texeira, M., Van Zeebroeck, G. & Thevelein, J. M. 2011. From
408 transporter to transceptor: signaling from transporters provokes re-evaluation of
409 complex trafficking and regulatory controls: endocytic internalization and intracellular
410 trafficking of nutrient transceptors may, at least in part, be governed by their signaling
411 function. *Bioessays* **33**: 870-9.
412 Lahnsteiner, F. & Mansour, N. 2010. A comparative study on antioxidant systems in semen of
413 species of the Percidae, Salmonidae, Cyprinidae, and Lotidae for improving semen
414 storage techniques. *Aquaculture* **307**: 130-140.
415 Lahnsteiner, F., Mansour, N. & Plaetzer, K. 2010. Antioxidant systems of brown trout (Salmo
416 trutta f. fario) semen. *Anim. Reprod. Sci.* **119**: 314-21.
417 Langmead, B. & Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*
418 **9**: 357-9.
419 Li, B. & Dewey, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or
420 without a reference genome. *BMC Bioinformatics* **12**: 323.
421 Linster, C. L. & Van Schaftingen, E. 2007. Vitamin C. *FEBS J.* **274**: 1-22.
422 Lopez-Cruz, R. I., Dafre, A. L. & Filho, D. W. (2011) Oxidative Stress in Sharks and Rays. In:
423 *Oxidative Stress in Aquatic Ecosystems*, (Abele, D., Vazquez-Medina, J. P. & Zenteno-
424 Savin, T., eds.). pp. 157-164. Wiley-Blackwell.
425 Moreau, R. & Dabrowski, K. 2000. Biosynthesis of ascorbic acid by extant actinopterygians. *J.*
426 *Fish Biol.* **57**: 733-745.
427 Nakatani, Y., Takeda, H., Kohara, Y. & Morishita, S. 2007. Reconstruction of the vertebrate
428 ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome*
429 *Res.* **17**: 1254-1265.
430 O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. 2016.
431 Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion,
432 and functional annotation. *Nucleic Acids Res.* **44**: D733-45.
433 Oda, M., Satta, Y., Takenaka, O. & Takahata, N. 2002. Loss of urate oxidase activity in hominoids
434 and its evolutionary implications. *Mol. Biol. Evol.* **19**: 640-53.
435 Papakostas, K. & Frillingos, S. 2012. Substrate selectivity of YgfU, a uric acid transporter from
436 Escherichia coli. *J. Biol. Chem.* **287**: 15684-95.
437 Parenti, L. R. & Grier, H. J. 2004. Evolution and phylogeny of gonad morphology in bony fishes.
438 *Integr. Comp. Biol.* **44**: 333-48.
439 Poulsen, P., Wu, B., Gaber, R. F., Ottow, K., Andersen, H. A. & Kielland-Brandt, M. C. 2005. Amino
440 acid sensing by Ssy1. *Biochem. Soc. Trans.* **33**: 261-4.
441 Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., et al. 2008.
442 The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453**: 1064-
443 1071.

Raghuram, V., Mak, D.-O. D. & Foskett, J. K. 2001. Regulation of cystic fibrosis transmembrane conductance regulator single-channel gating by bivalent PDZ-domain-mediated interaction. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 1300-1305.

Savini, I., Rossi, A., Pierro, C., Avigliano, L. & Catani, M. V. 2008. SVCT1 and SVCT2: key proteins for vitamin C uptake. *Amino Acids* **34**: 347-55.

Scott, D. B. C. & Fuller, J. D. 1976. The reproductive biology of Scleropages formosus (Müller & Schlegel) (Osteoglossomorpha, Osteoglossidae) in Malaya, and the morphology of its pituitary gland. *J. Fish Biol.* **8**: 45-53.

Singer, M. A. 2003. Do mammals, birds, reptiles and fish have similar nitrogen conserving systems? *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **134**: 543-58.

Uttara, B., Singh, A. V., Zamboni, P. & Mahajan, R. T. 2009. Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options. *Curr. Neuropharmacol.* **7**: 65-74.

Wang, S., Yue, H., Derin, R. B., Guggino, W. B. & Li, M. 2000. Accessory Protein Facilitated CFTR-CFTR Interaction, a Molecular Mechanism to Potentiate the Chloride Channel Activity. *Cell* **103**: 169-179.

Wright, P. A. 1995. Nitrogen excretion: three end products, many physiological roles. *J. Exp. Biol.* **198**: 273-81.

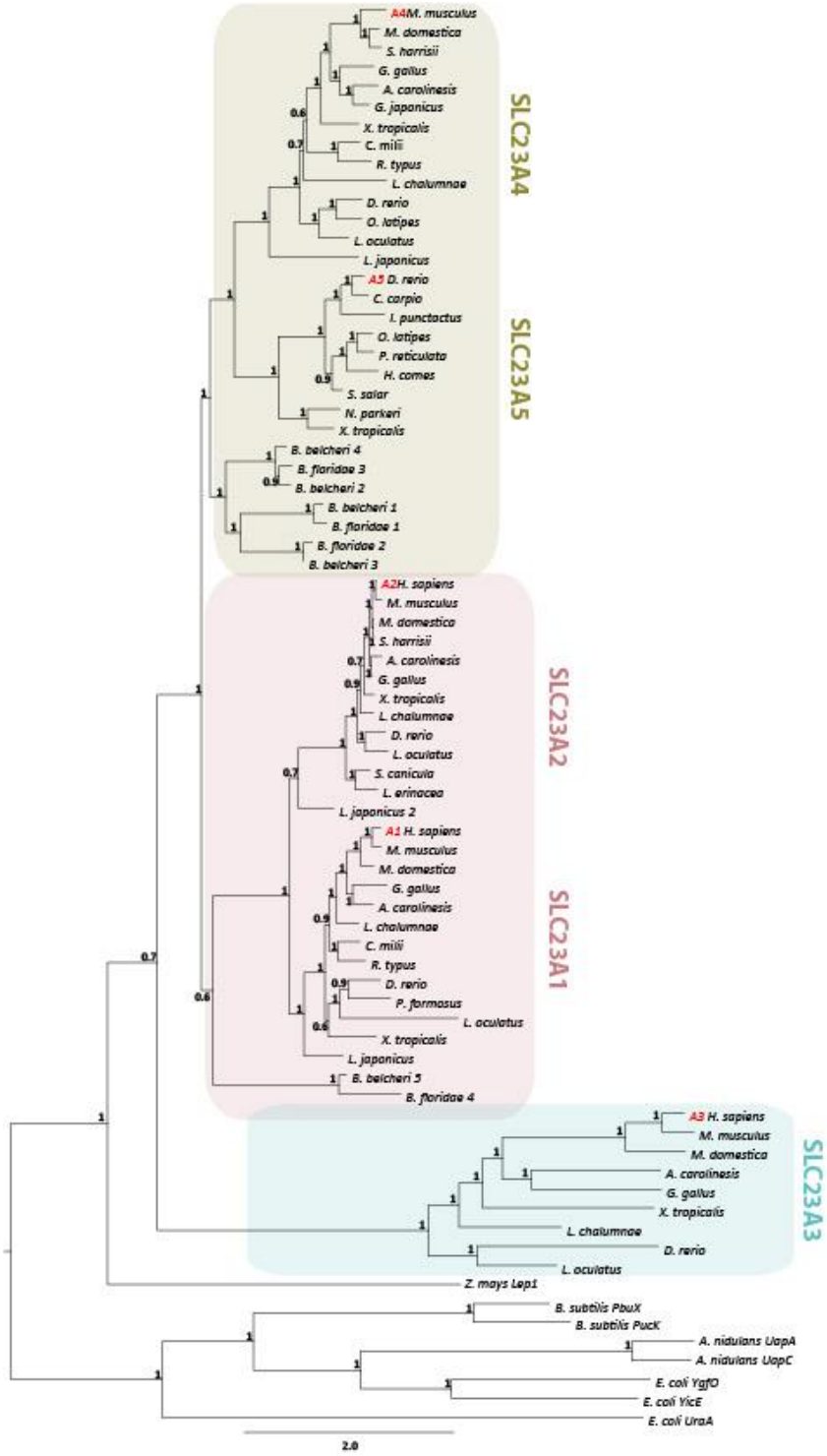
Wu, B., Ottow, K., Poulsen, P., Gaber, R. F., Albers, E. & Kielland-Brandt, M. C. 2006. Competitive intra- and extracellular nutrient sensing by the transporter homologue Ssy1p. *J Cell Biol* **173**: 327-31.

Wu, X. W., Muzny, D. M., Lee, C. C. & Caskey, C. T. 1992. Two independent mutational events in the loss of urate oxidase during hominoid evolution. *J. Mol. Evol.* **34**: 78-84.

Yamamoto, S., Inoue, K., Murata, T., Kamigaso, S., Yasujima, T., Maeda, J. Y., et al. 2010. Identification and functional characterization of the first nucleobase transporter in mammals: implication in the species difference in the intestinal absorption mechanism of nucleobases and their analogs between higher primates and other mammals. *J. Biol. Chem.* **285**: 6522-31.

Yates, A., Akanni, W., Amode, M. R., Barrell, D., Billis, K., Carvalho-Silva, D., et al. 2016. Ensembl 2016. *Nucleic Acids Res.* **44**: D710-6.

476 **Figures**



477

478 **Figure 1 – Maximum likelihood phylogenetic tree describing relationships among**

479 **Nucleobase-Ascorbate Transporter (NAT) family members.** Node values represent

480 branch support using the aBayes algorithm. Accession numbers for all sequences are

481 provided in Supplementary Materials.

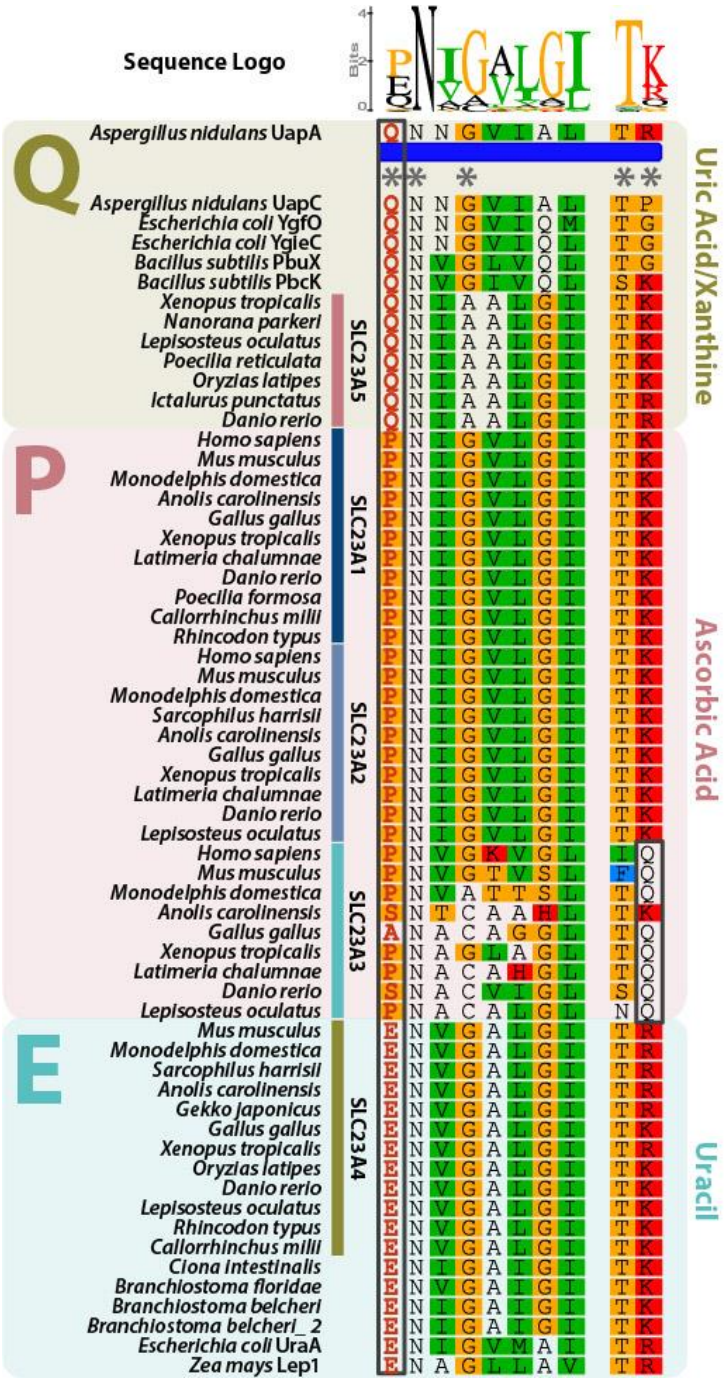


Figure 2 – Alignment of the signature motif of the Nucleobase-Ascorbate Transporter (NAT) family. NAT family members display a highly conserved signature motif defining substrate specificity: [Q/E/P]-N-x-G-x-x-x-x-T-[R/K/G]. Asterisks indicate conserved positions. The first amino acid position, with residues highlighted in red within a black box, defines substrate preference: Q – Uric acid/xanthine group; P – Ascorbate group and E – Uracil group. The mirror Q in the Slc23a3 group is also highlighted by a black box.

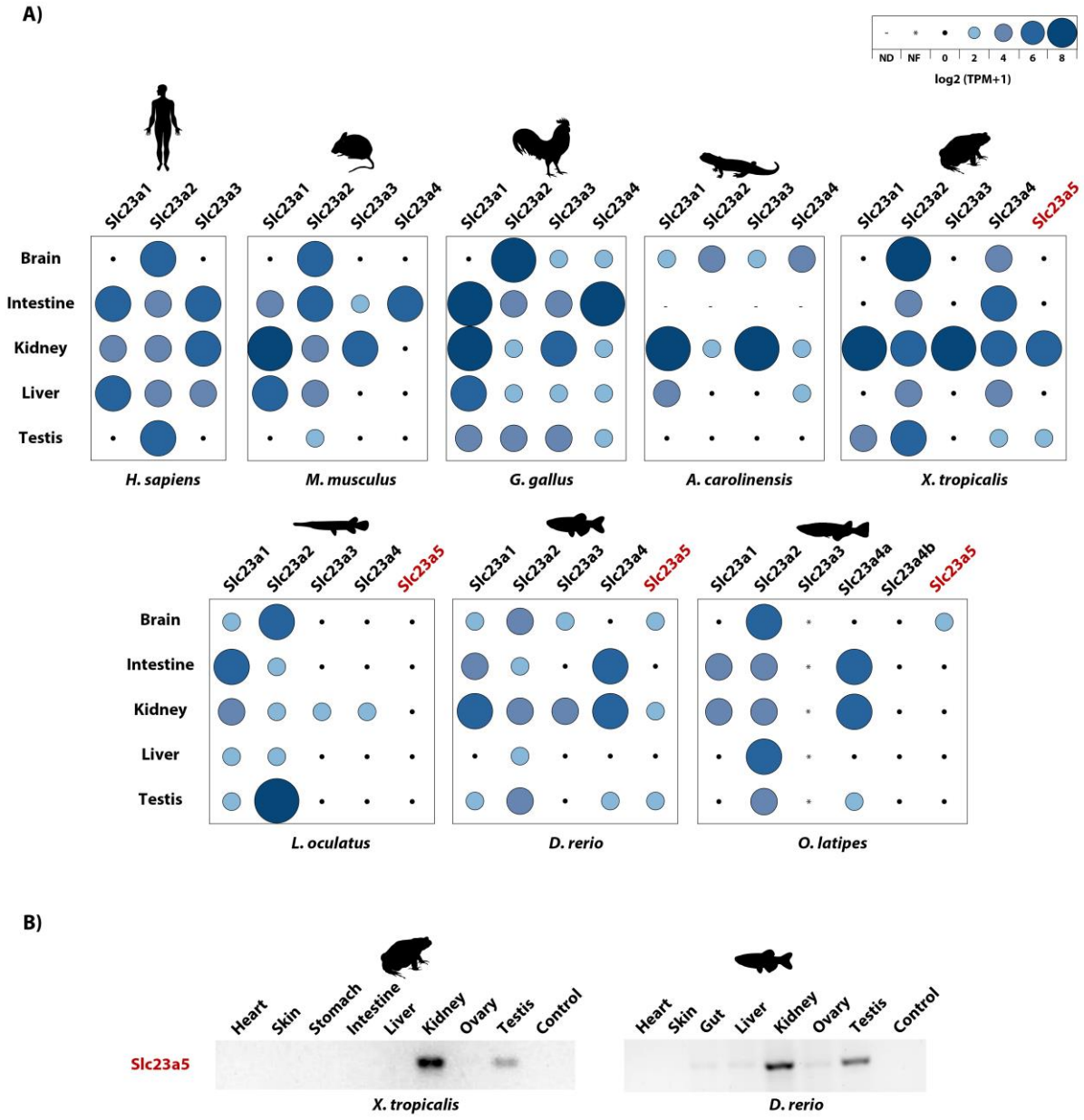


Figure 3 – Expression patterns of *Slc23* isoforms across vertebrates. (A) RNAseq-based quantification. Relative gene expression levels are provided as $\log_2$-transformed transcript per million (TPM) values, from 0 to 8 TPM's, the NF and ND means gene not found and no SRA data available to a specific tissue. (B) PCR-based amplification of *Slc23a5* transcripts in *D. rerio* and *X. tropicalis* tissues.

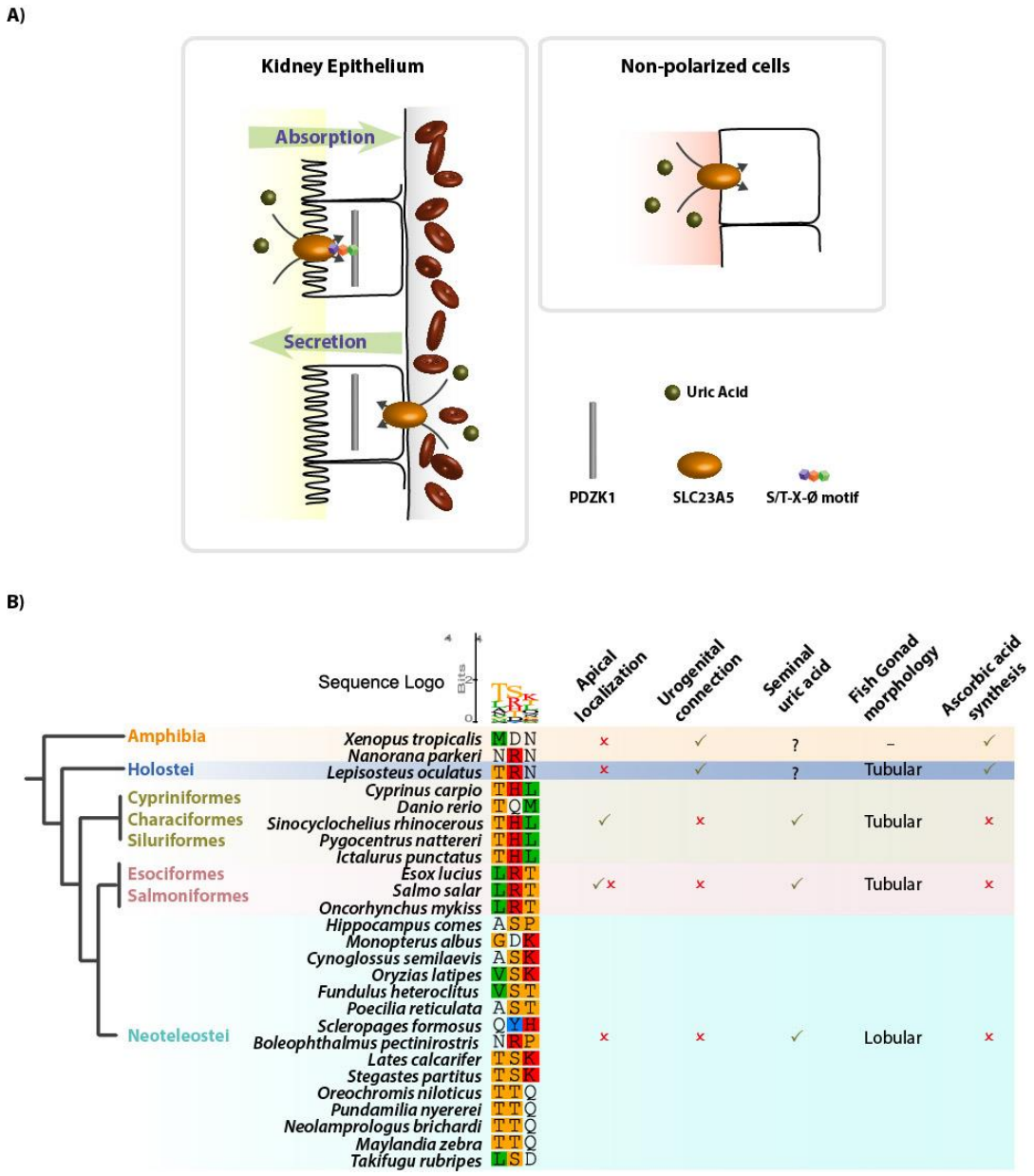


Figure 4 – The S/T-X-Ø tripeptide provides insight into the cellular function of Slc23a5 homologues from fish and amphibians. (A) Schematic representation of tripeptide-dependent uric acid flows in kidney epithelium and non-polarized cells. (B) Alignment of the PDZ-binding tripeptide in Slc23a5 homologues from fish and amphibians. The S/T-X-Ø tripeptide is found in Cypriniformes, Characiformes and Siluriformes while the Euteleostei Esociformes and Salmoniformes exhibit an inverted motif. Predicted cellular localization is indicated as well as morphological and physiological adaptations among the different fish groups.

506 SUPPLEMENTARY MATERIAL

507 **A taxonomically-restricted Uric Acid transporter provides insight into antioxidant**
508 **equilibrium**

509

510

511 **Supplementary Figures:**

512 **Figure S1 - Synteny maps of *Slc23a5*.**

513 **Figure S2 - Synteny maps of *Slc23a1*.**

514 **Figure S3 - Synteny maps of *Slc23a2*.**

515 **Figure S4 - Synteny maps of *Slc23a3*.**

516 **Figure S5 - Synteny maps of *Slc23a4.***

517 **Figure S6 – Paralogy analysis of the human *loci*.**

518 **Figure S7 – Topology prediction of Slc23a5.**

519 **Supplementary Tables:**

520 **Table S1 - Accession numbers.**

521 **Table S2 - Accession numbers of the RNAseq files.**

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

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

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

528 **Table S6 - Renal uric acid absorption and secretion: the urate transportome.**

529

530

531

532

533

534

Supplementary Figures:

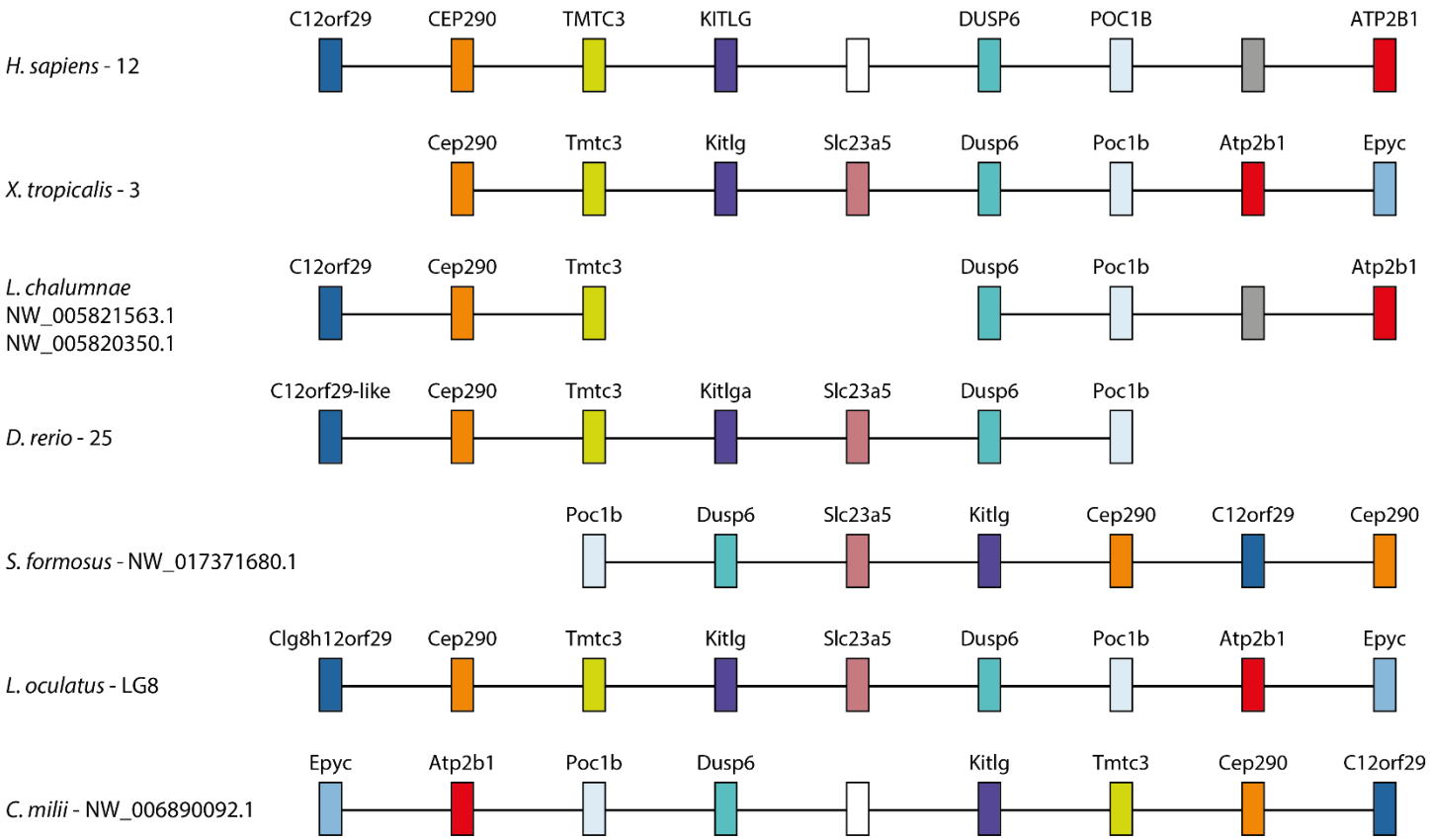


Figure S1 - Synteny maps of *Slc23a5*. Conservation of the *Slc23a5* locus in major fish lineages, *Xenopus tropicalis* and *Homo sapiens*. Numbering after species corresponds to chromosome or scaffold. Information is presently absent for the Chondrostei *Acipenser ruthenus*. Grey boxes indicate non-conserved genes.

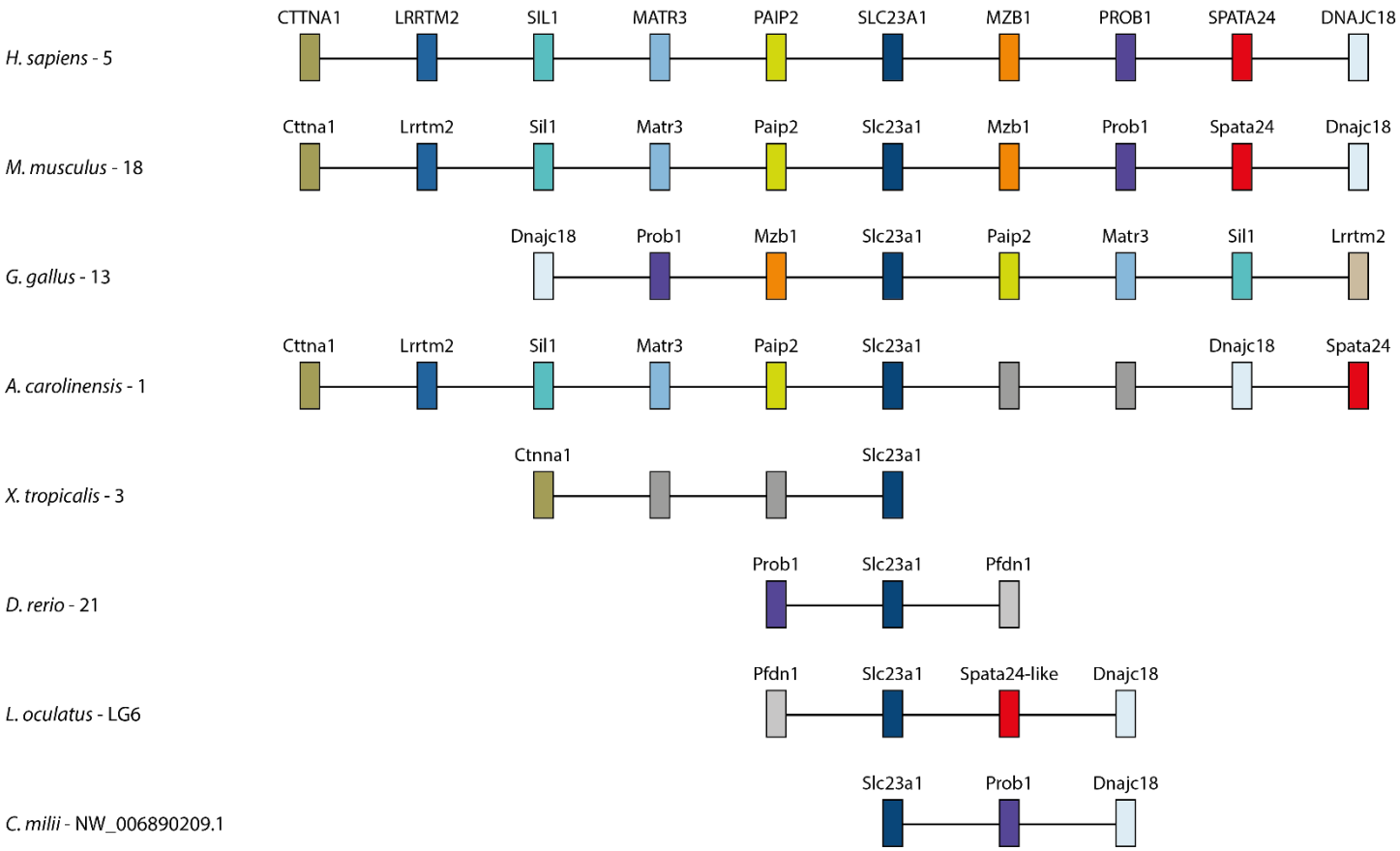


Figure S2 - Synteny maps of *Slc23a1*. Conservation of the *Slc23a1* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes.

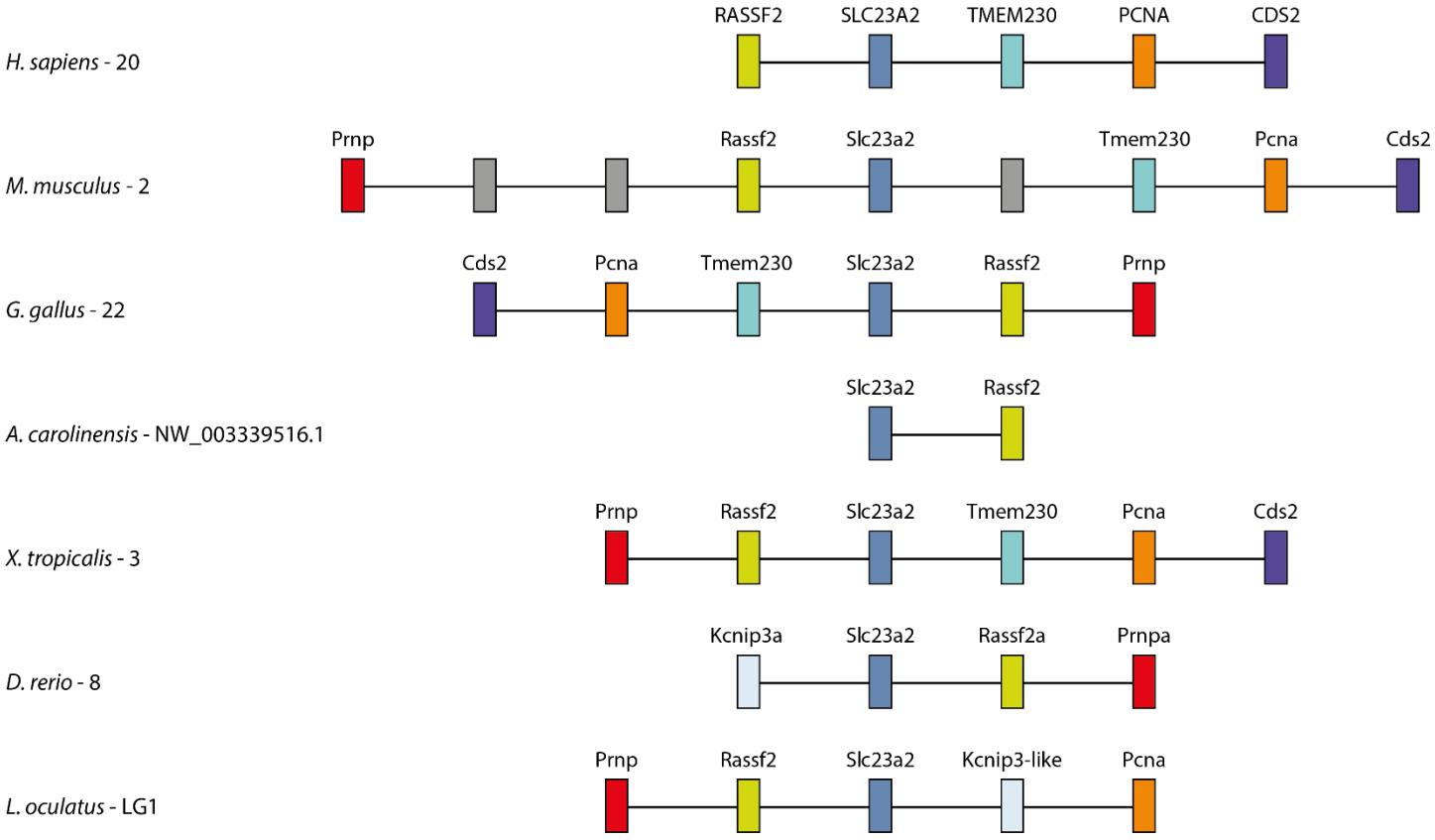


Figure S3 - Synteny maps of *Slc23a2*. Conservation of the *Slc23a2* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes.

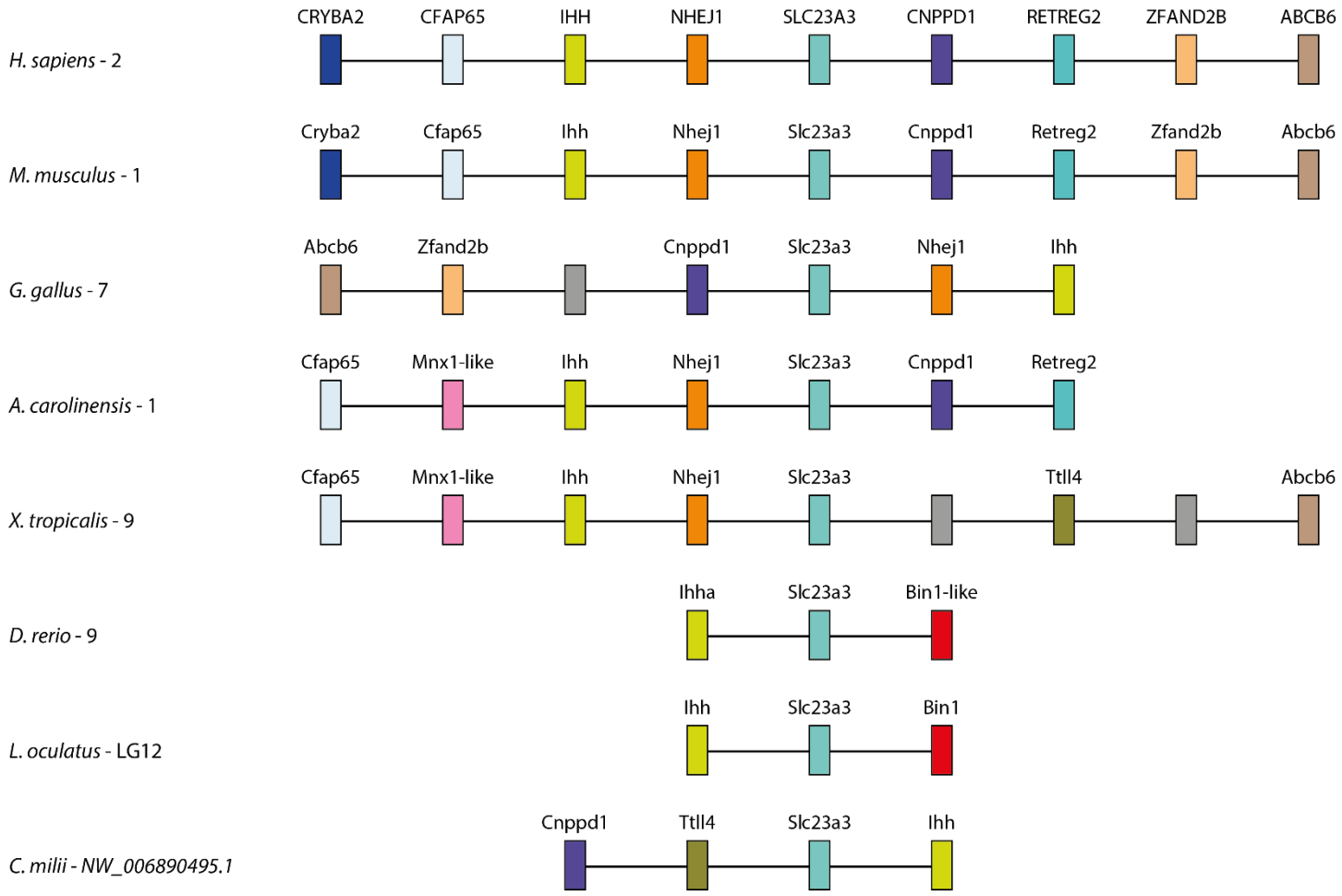


Figure S4 - Synteny maps of *Slc23a3*. Conservation of the *Slc23a3* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes.

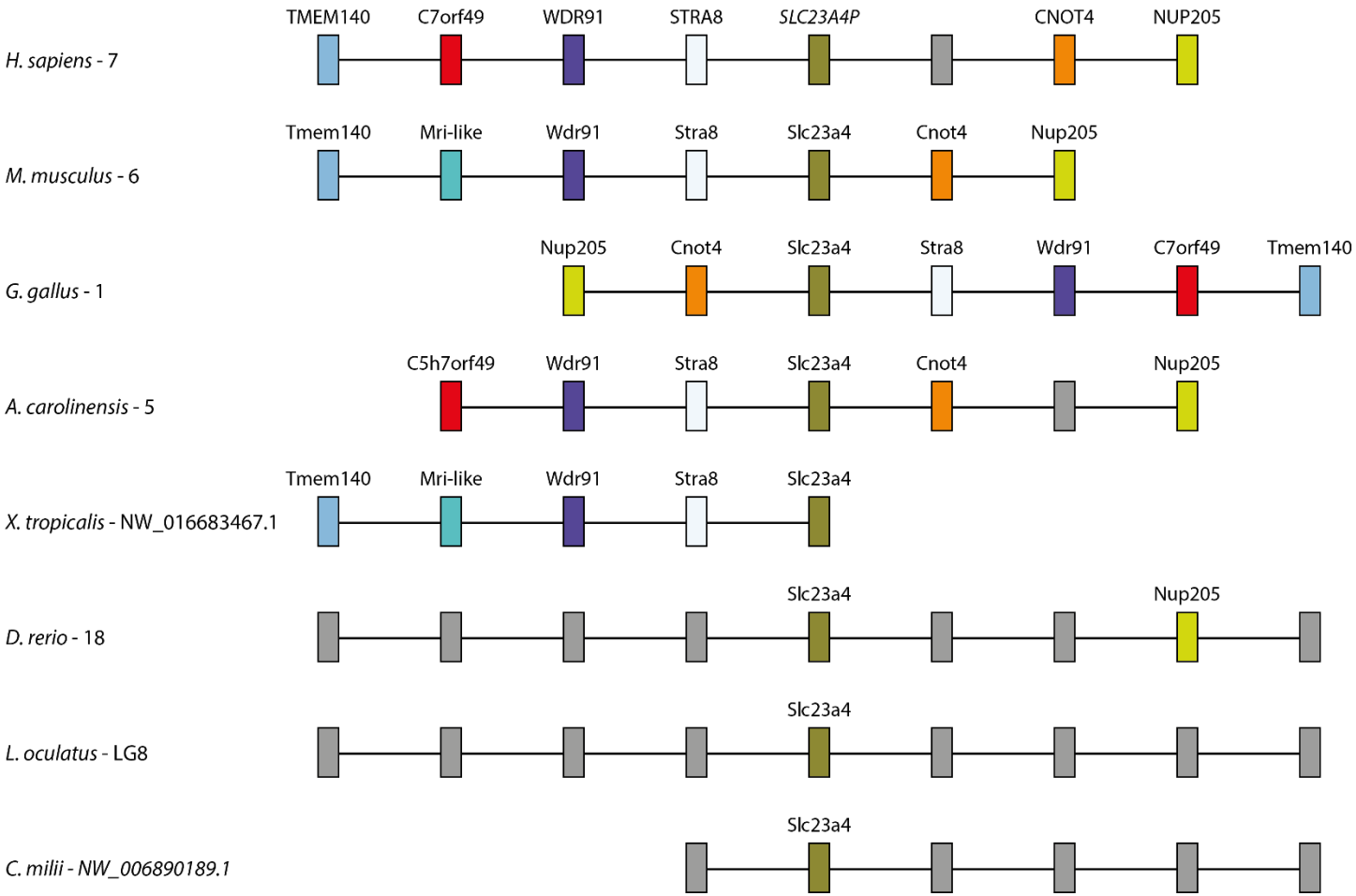


Figure S5 - Synteny maps of *Slc23a4.* Conservation of the *Slc23a4* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. The human homologue is indicated as pseudogenized (*SLC23A4P*). Grey boxes indicate non-conserved genes.

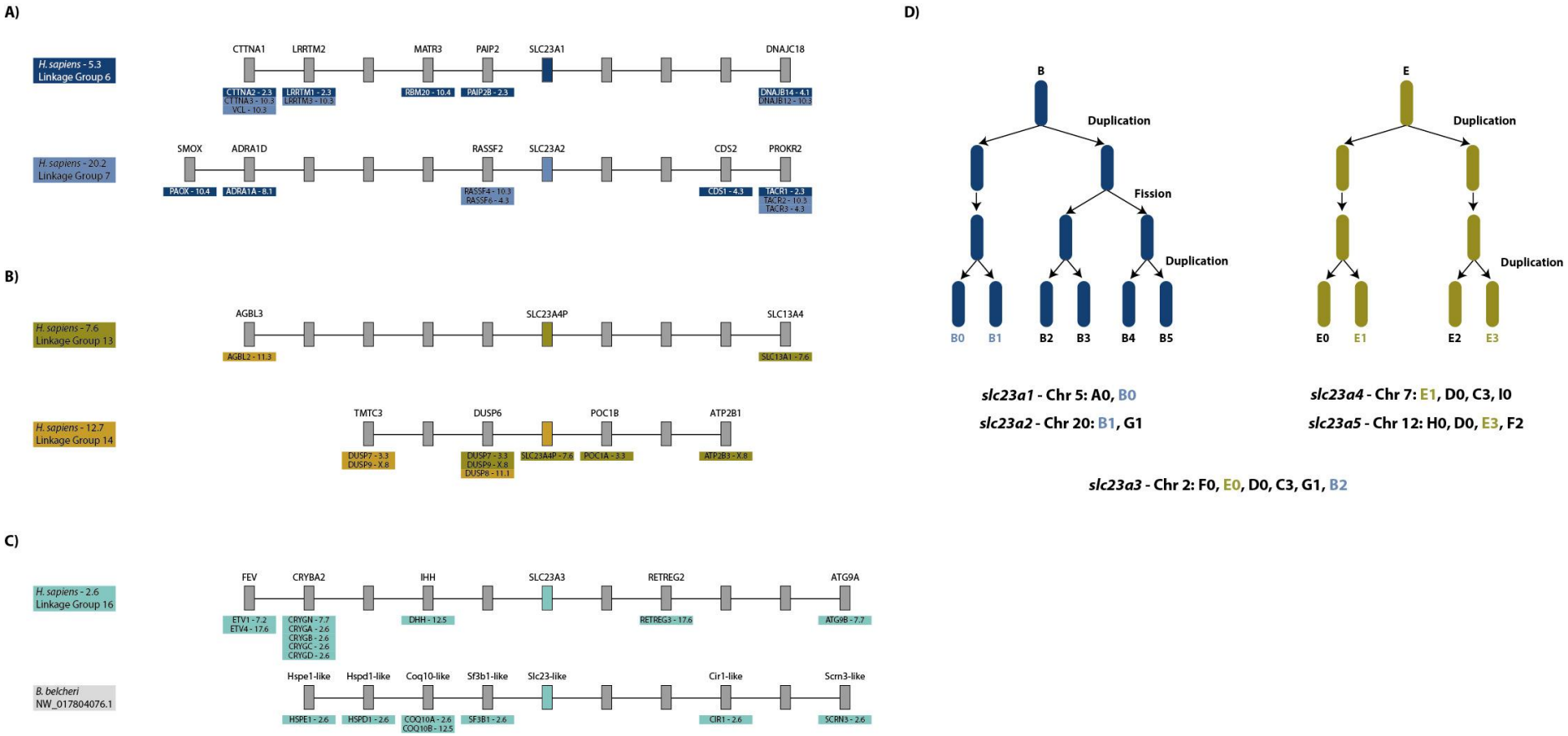


Figure S6 – Paralogy analysis of the human *loci*. (A-C) Reconstruction of ancestral genome partitions according to Putnam and colleagues (Putnam, et al. 2008) and (D) according to Nakatani and colleagues (Nakatani, et al. 2007). (A-C) Numbers following gene symbols correspond to predicted chromosomal partitions grouped within linkange groups, that is, groups of genes with a predicted common chromosomal origin in the vertebrates' ancestor; corresponding paralogues of neighbouring genes are indicated below. (C) An invertebrate Slc23-like locus, from

Branchiostoma belcheri is also shown with corresponding human homologues indicated below each gene. (D) Reconstruction of duplication and fission events in the predicted ancestral lineages including the Slc23a1/Slc23a2 pair (lineage B) and the Slc23a4 and Slc23a5 pair (lineage E).

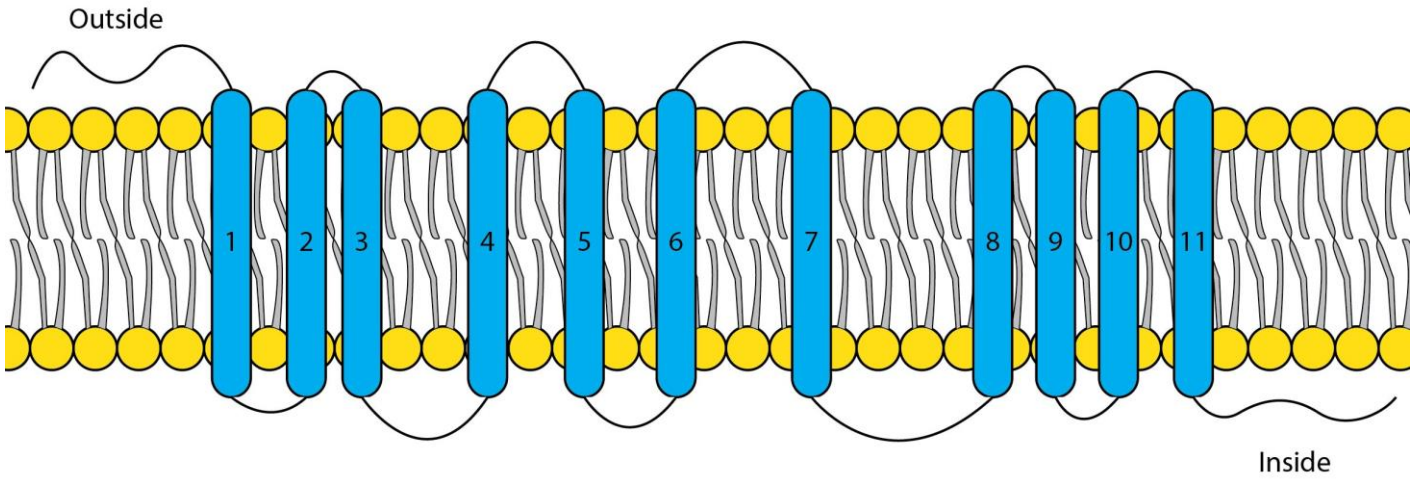


Figure S7 – Topology prediction of Slc23a5.

Supplementary Tables:

Table S1 - Accession numbers.

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

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

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

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

Spotted gar	Library	SRA Run	Read length
Brain	PE	SRR1524250	100
Intestine	PE	SRR1524257	100
Kidney	PE	SRR1524255	100
Liver	PE	SRR1524254	100
Testis	PE	SRR1524260	100
Zebrafish	**Library**	**SRA Run**	**Read length**
Brain	PE	SRR1524238	100
Intestine	PE	SRR1524245	100
Kidney	PE	SRR1524243	100
Liver	PE	SRR1524242	100
Testis	PE	SRR1524249	100
Chicken	**Library**	**SRA Run**	**Read length**
Brain	PE	SRR924542	100
Intestine	PE	SRR3194289	100
Kidney	PE	SRR924553	100
Liver	PE	SRR924555	100
Testis	PE	SRR924547	100
Human	**Library**	**SRA Run**	**Read length**
Brain	PE	ERR315432	100
Intestine	PE	ERR315419	100
Kidney	PE	ERR315468	100
Liver	PE	ERR315451	100
Testis	PE	ERR315456	100
Mouse	**Library**	**SRA Run**	**Read length**
Brain	PE	SRR5171101	100
Intestine	PE	SRR5171080	100
Kidney	PE	SRR5171094	100
Liver	PE	SRR5171078	100
Testis	PE	SRR5171084	100
Western clawed frog	**Library**	**SRA Run**	**Read length**
Brain	PE	SRR579560	75
Intestine	SE	SRR1186999	50
Kidney	PE	SRR579562	75
Liver	PE	SRR579561	75
Testis	SE	SRR943353	100
Anole lizard	**Library**	**SRA Run**	**Read length**
Brain	PE	SRR579556	75
Kidney	PE	SRR579557	75
Liver	PE	SRR391651	100
Testis	PE	SRR5420127	50
Japanese medaka	**Library**	**SRA Run**	**Read length**
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	ENSGALG00000037642
	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[a]	101173186	ENSORLG00000009915
	Slc23a4b	101157211	ENSORLG00000015991
	Slc23a5	101175057	ENSORLG00000019882
Spotted gar	Slc23a1	102691341	ENSLOCG00000012533
	Slc23a2	102691255	ENSLOCG00000015630
	Slc23a3	102699040	ENSLOCG00000010877
	Slc23a4	102686145	ENSLOCG00000016950
	Slc23a5	2682847	ENSLOCG00000016369

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

6 **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
	INSTESTINE	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

7

8

9

10

11

12

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

	Transporter	*H. sapiens*	*M. musculus*	*G. gallus*	*A. carolinensis*	*X. tropicalis*	*D. rerio*	*L. oculatus*
Basolateral Absorption	SLC2A9/GLUT9	✓	✓	✓	✓	✓	✓	✓
Apical Absorption	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	✓	✓	✓	✓	✓	✓	✓
Basolateral Secretion	SLC22A6/OAT1	✓	✓	✓	✓	✓	n.f.	n.f.
	SLC22A8/OAT3	✓	✓				n.f.	n.f.
Apical Secretion	SLC17A1/NPT1	✓	✓	n.f.	n.f.	n.f.	n.f.	n.f.
	SLC17A3/NPT4	✓	✓	n.f.	n.f.	n.f.	n.f.	n.f.

14

15

16

17