Identification of a novel Nucleobase-Ascorbate Transporter family member in fish and amphibians

Diogo Oliveira*,1, André M. Machado*,1, Tiago Cardoso1, Mónica Lopes-Marques1, L. Filipe C. Castro1,2● and Raquel Ruivo1●

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

2Department 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 title: 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. Using phylogenetic and synteny analysis, we identify a fifth member of the vertebrate *slc23* complement (*slc23a5*), present 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, using available transcriptomic data, 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 discussed.

Keywords: Nucleobase-Ascorbate Transporters (NAT); Slc23; Slc23a5; Uric acid; Antioxidant

32 **1. Introduction**

33 Ascorbic acid (or its salt form ascorbate) is an essential enzyme cofactor, participating in
34 collagen and norepinephrine synthesis, as well as a potent antioxidant and free radical scavenger
35 [1,2]. Despite its importance, anthropoid primates, teleost fish, Guinea pigs, some bats and
36 Passeriformes bird species, lack a functional L-gulono-ϒ-lactone oxidase (GLO) gene and, thus,
37 the ability to catalyse the last step of ascorbic acid synthesis [3,4]. The scattered loss of the GLO
38 gene was suggested to evolve neutrally, with ascorbic acid requirements being met by dietary
39 intake [3]. Unlike ascorbate, uric acid is largely regarded as end-product of purine metabolism
40 and metabolic waste [5]. Yet, human-biased research has suggested alternative physiological
41 functions for uric acid or urate (uric acid salts), such as antioxidant and free radical scavenger,
42 similarly to ascorbic acid, counterbalancing disease-related oxidative stress [6-8].

43 Ascorbic and uric acid serve as substrates to active transporters that belong to the
44 Nucleobase-Ascorbate Transporter (NAT) family; NAT family members, which include Solute
45 carrier family 23 (Slc23) transporters, display a very broad evolutionary distribution, from
46 bacteria to metazoans [1]. Sodium-dependent Slc23 transporters were suggested to maintain
47 vertebrate ascorbic acid levels: with Slc23a1 responsible for renal and intestine ascorbic acid
48 absorption across brush border epithelia and Slc23a2, which exhibits a widespread expression,
49 responsible for ascorbic acid supply to target tissues and cells, including neurons [1,9,10]. In
50 vertebrates, two additional transporters were identified, the orphan transporter Slc23a3 and the
51 nucleobase transporter Slc23a4, described in rats but found pseudogenized in humans [1,11]. Rat
52 Slc23a4 is restricted to the small intestine and was shown to predominantly mediate sodium-
53 dependent uracil transport [11]. Slc23a3, on the other hand, is expressed in proximal renal tubules
54 but is unresponsive to both ascorbic acid and nucleobases [1].

55 NAT family members display a highly conserved signature motif: [Q/E/P]-N-x-G-x-x-x-x-
56 T-[R/K/G] [1,11-13]. Functional characterization of distinct member of the NAT family
57 highlighted the importance of the first amino acid position of the motif in the definition of

58 substrate specificity [1,11-13]. Thus, these transporters can be divided into three functional
59 groups. Slc23a1, Slc23a2 and Slc23a3 belong to the ascorbate group and harbour a conserved
60 proline (P) in the first amino acid position of the signature motif [1]. A glutamate (E) residue
61 defines the uracil group, including bacterial anion symporters and Slc23a4, while a glutamine (Q)
62 is present in the bacterial uric acid and/or xanthine anion symporters [1,11-13]. Topology
63 predictions, using proteins from the distinct groups, suggest a cytosolic location of the motif [1,
64 11, 12]. Here we revise the portfolio of NAT family members in vertebrates.

65 **2. Materials and Methods**

66 *2. 1Sequence mining*

67 Slc23 protein sequences were obtained from the GenBank database by BLASTp using
68 human Slc23a1 (NP_005838), Slc23a2 (NP_976072) and Slc23a3 (NP_001138362) sequences as
69 query. The sequences from *Lethenteron japonica* (Japanese lamprey) were obtained from the
70 Japanese Lamprey Genome Project (http://jlampreygenome.imcb.a-star.edu.sg/). The sequences
71 from *Leucoraja erinacea* (little skate) and *Scyliorhinus canicula* (small-spotted catshark) were
72 obtained as nucleotide sequences from SkateBase (http://skatebase.org) and translated. Accession
73 numbers for all the retrieved sequences are listed in Supplementary Table 1 (Table S1).

74 *2. 2 Synteny*

75 *Slc23* genes were mapped onto the respective species genomes, using the latest genome
76 assemblies available in GenBank database (https://www.ncbi.nlm.nih.gov/). Genes flanking
77 *Slc23a1*, *Slc23a2*, *Slc23a3*, *Slc23a4* and the novel *Slc23a*-like gene were identified and mapped using
78 the human loci as reference. If the target gene was not found, the syntenic genomic region was
79 retrieved using neighboring genes as reference.

80 *2. 3Phylogenetic Analysis*

81 The retrieved amino acid sequences were aligned using the MAFFT software web service
82 (http://mafft.cbrc.jp/alignment/software/) in default automatic settings, with L-INS-I refinement

83 method [14]. The alignment was stripped from columns with at least 20% of gaps resulting in an
84 alignment with 75 sequences and 526 positions. The output alignment was used to construct a
85 phylogenetic tree using PhyML 3.0 with Smart Model Selection web service (http://www.atgc-
86 montpellier.fr/phyml-sms/) in default settings, which selected the LG +G+I+F model [15].
87 Bayesian-like transformation of aLRT was selected to assess branch support. An additional
88 phylogenetic tree, with automatic bootstrapping (552 bootstraps) and using the default hill-
89 climbing algorithm, was constructed with RAxML, available in CIPRES Science Gateway.

90 *2.4Membrane topology prediction*

91 Slc23a5 membrane topology predictions were performed for selected species, using
92 distinct prediction web servers: TMHMM Server v. 2.0
93 (http://www.cbs.dtu.dk/services/TMHMM/) and TMPred (https://embnet.vital-
94 it.ch/software/TMPRED_form.html). Human Slc23a1 was used as control.

95 *2.5Filtering and quality control of Sequence Reading Archive (SRA)*
96 *datasets*

97 For RNA-seq analysis, we included 39 SRA datasets from brain, intestine, kidney, liver,
98 and testis tissues of *Homo sapiens* (human), *Mus musculus* (mouse), *Gallus gallus* (chicken), *Anolis*
99 *carolinensis* (American green anole), *Xenopus tropicalis* (western clawed frog), *Danio rerio*
100 (zebrafish), *Lepisosteus oculatus* (spotted gar) and *Oryzias latipes* (Japanese medaka), which were
101 obtained from the National Center for Biotechnology (NCBI) Sequence Read Archive (*SRA*)
102 (https://www.ncbi.nlm.nih.gov/sra/) (Table S2). The read quality assessment of each SRA dataset
103 was performed using FastQC software (https://www.bioinformatics.
104 Babraham.ac.uk/projects/fastqc/). Trimmomatic 0.36 [17] was used to trim and drop reads with
105 quality scores below 5, at leading and trailing ends, with an average quality score below 20 in a 4
106 bp sliding window and with less than 36 bases length.

107 *2.6Relative gene expression levels*

108 To assess the relative gene expression of *slc23* genes from the selected species, reference
109 sequences and the corresponding annotations were collected from NCBI and Ensembl [18]
110 databases. For human, mouse, lizard, zebrafish and spotted gar, the GTF and DNA files (genome)
111 were retrieved from Ensembl (Release 89) [18]. For chicken and western clawed frog, transcript
112 sequences (coding and non-coding) and mapping (gene/transcript) from NCBI Refseq database
113 were used after filtering [19] (for each species the rna.fasta file from
114 ftp://ftp.ncbi.nih.gov/genomes/refseq/ was used – See Table S3 and S4 for detail). The SRA reads
115 of each specimen were aligned to the reference using Bowtie2 (default parameters)[20], and gene
116 expression was quantified as transcript per million (TPM) using RSEM v.1.2.31 (RNA-seq by
117 Expectation Maximization) software package [21]. Isoform expression levels for each gene were
118 summed to derive the TPM values. For genes with low relative expression values, TPM values
119 were log2-transformed after adding a value of one; TPM values < 0.5 were considered unreliable
120 and substituted with zero.

121 *2.7PCR-based validation of Slc23a5 expression*

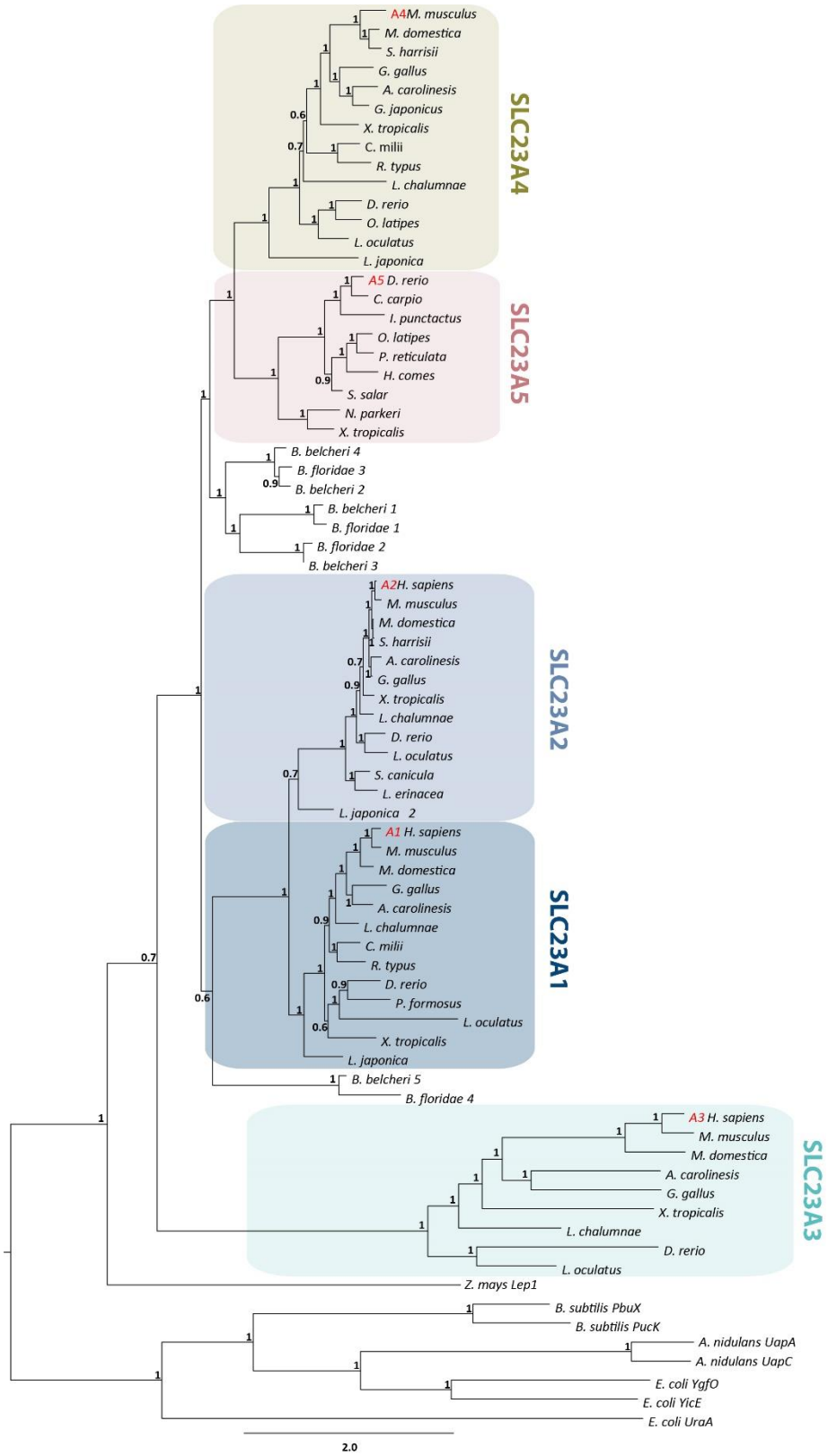
122 Total RNA extraction from *X. tropicalis* and *D. rerio* tissues was performed using the Illustra
123 RNAspin Mini RNA Isolation Kit animal tissues protocol with on-column DNAseI digestion (GE
124 Healthcare). RNA quality was visually assessed by electrophoresis and sample concentration was
125 determined using the Take3 micro-volume plate system (BioTek). First-strand cDNA was
126 synthesized from 250ng of total RNA using the iScriptTM cDNA Synthesis Kit (Bio-Rad),
127 according to the manufacturer's instructions. Gene expression of *Slc23a5* was analyzed by PCR
128 using the intro flanking expression primer pairs 5'-AGACGTTGAAGCCGTTACTGA-3' and 5'-
129 TGCTATTATGCCTCCAAAGGCT-3' for *X. tropicalis* and 5'-GGTGGAATGTTCCTCGTCAT-3'
130 and 5'-GATGAACATGTGGGTGGTGA-3' for *D. rerio*. PCR reactions were carried out on a tissue
131 panel containing heart, skin, stomach, intestine, liver, kidney, ovary and testis tissues for *X.*
132 *tropicalis* and brain, heart, skin, gut, liver, kidney, ovary and testis tissues for *D. rerio*. The PCR
133 was performed with Phusion® Flash high fidelity Master Mix (Thermofisher), 1μM of each

134 primer and 25ng/µl of template cDNA, with the following parameters: initial denaturation at
135 98°C for 1s, annealing at 58°C for *D. rerio* or 66ºC for *X. tropicalis* primers and elongation at 72°C
136 for 4s, with 30 cycles.

137 **3. Results**

138 *3.1Identification of a novel member of the NAT family in neopterygians*
139 *and amphibians*

140 Through database mining and phylogenetic analysis, we detected the presence of a novel
141 gene, *slc23a5*, found in neopterygians (Holostei and Teleostei) and amphibians (Figure 1 and
142 Figure S1). No *slc23a5* sequence was retrieved from chondrichthyans, chondrosteans (basal
143 actinopterygians) nor from the sarcopterygian *Latimeria chalumnae* (coelacanth). Subsequent loss
144 of *slc23a5* in chondrichthyans was further supported by the conservation of the genomic locus in
145 *Callorhynchus milii* (Australian ghostshark) (Figure S2). Regarding coelacanth, two scaffolds,
146 denoting conservation of the flanking genes, were found; while, no information was retrieved
147 regarding the chondrostean loci (Figure S2). The remaining genes of the *slc23* family were
148 retrieved and genomic loci were found generally conserved across vertebrate species, with
149 exception of the *slc23a4* loci in fish (Figures S3-S6). The branching pattern of our phylogenetic
150 analysis suggests that after the expansion of the *slc23a* gene family *slc23a5* was subsequently lost
151 in sharks, basal actinopterygians, coelacanth and amniotes (Figure 1 and Figure S2). In agreement,
152 we obtained three phylogenetic clusters possibly corresponding to three ancestral lineages
153 (Figure 1 and Figure S1): one composed of typical *slc23a1* and *slc23a2* genes and a second group
154 including the previously reported *slc23a4* and the novel *slc23a5* genes, both outgrouped by
155 invertebrate sequences. A third cluster, which outgroups the remaining *slc23* genes, included the
156 highly divergent *slc23a3* genes, mirrored by the long branch lengths (Figure 1).



157

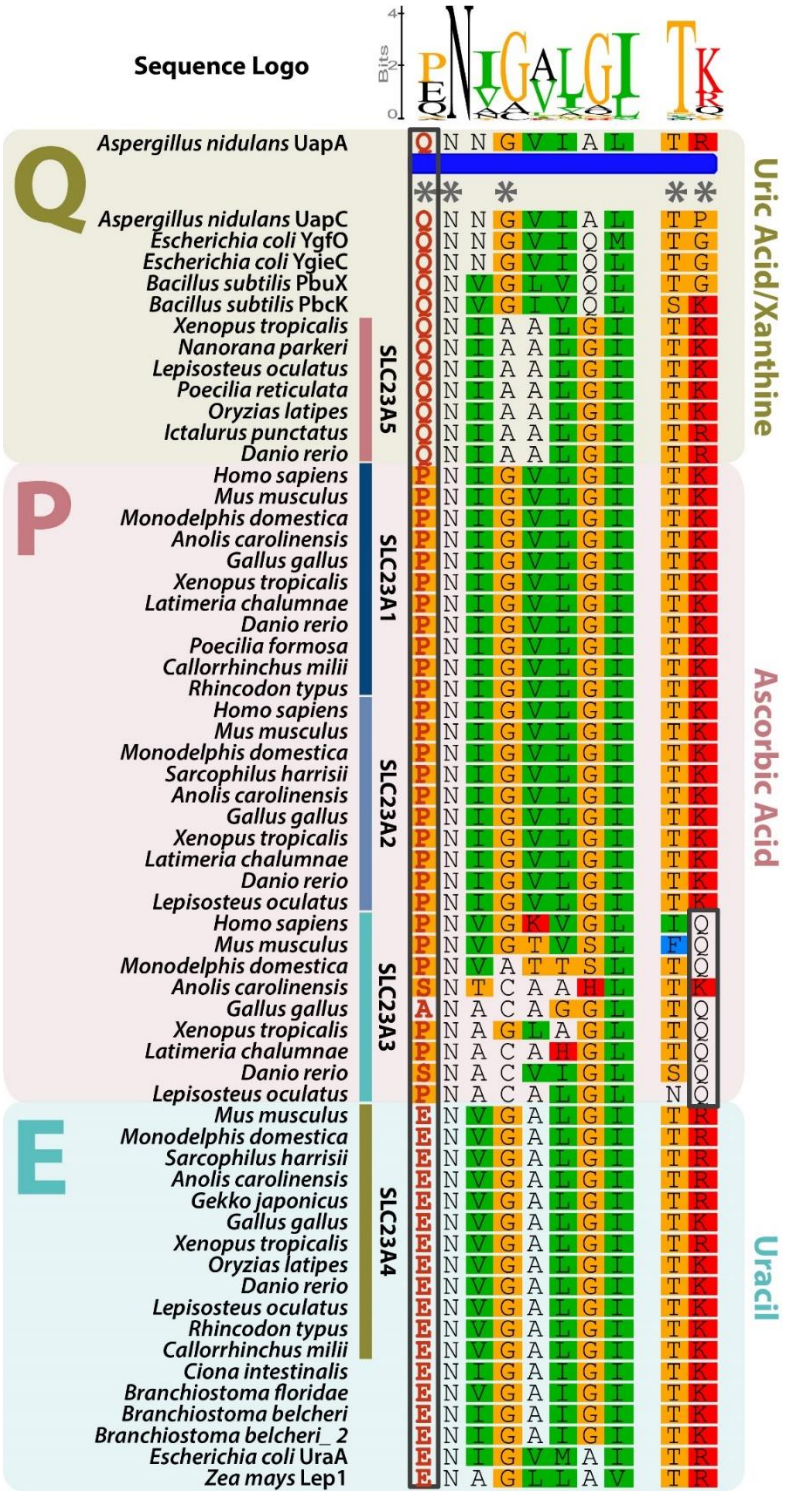
158 **Figure 1 – Maximum likelihood phylogenetic tree describing relationships among**

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

160 support using the aBayes algorithm. Accession numbers for all sequences are provided in
161 Supplementary Materials.
162

163 *3.2Slc23a5 signature motifs*

164 Further comparative protein sequence analysis highlighted the presence of a NAT family
165 signature motif with a conserved Q in the substrate determining position. According to previous
166 functional characterization of NAT family members, this residue is found in uric acid and/or
167 xanthine transporters, thus placing Slc23a5 in the uric acid and/or xanthine group (Figure 2)[1,11-
168 13]. Besides the typical signature motif, a carboxyl terminus tripeptide motif was identified in
169 some SLC23 family members: a S/T-X-Ø tripeptide. S/T-X-Ø motifs were shown to interact PDZ
170 domains of scaffolding proteins belonging to the Sodium/Hydrogen Exchanger Regulatory
171 Factor, NHERF, family [22]. S/T-X-Ø are found highly conserved across species in Slc23a1 and
172 Slc23a4, in agreement with their renal and intestinal expression and role in ascorbic acid and
173 nucleobase absorption (Figure S7). Slc23a5, on the other hand, exhibits a highly conserved
174 tripeptide motif in Otocephala (Cypriniformes, Characiformes and Siluriformes) but not
175 Osteoglossomorpha (Figure 3A). The Euteleostei Esociformes and Salmoniformes exhibit an
176 inverted motif while in the Neoteleostei no conservation was found. Finally, topology predictions
177 yielded similar transmembrane profiles for Slc23a5 sequences from *D. rerio*, *O. latipes*, *Salmo salar*
178 (Atlantic salmon) and *X. tropicalis*, when compared to human Slc23a1 (Figure S8). The NAT motif
179 was predicted cytosolic in all analyzed sequences. This suggests a conserved cellular import
180 topology, as observed in other family members.



181

182 **Figure 2 – Alignment of the signature motif of the Nucleobase-Ascorbate Transporter (NAT)**

183 **family.** NAT family members display a highly conserved signature motif defining substrate

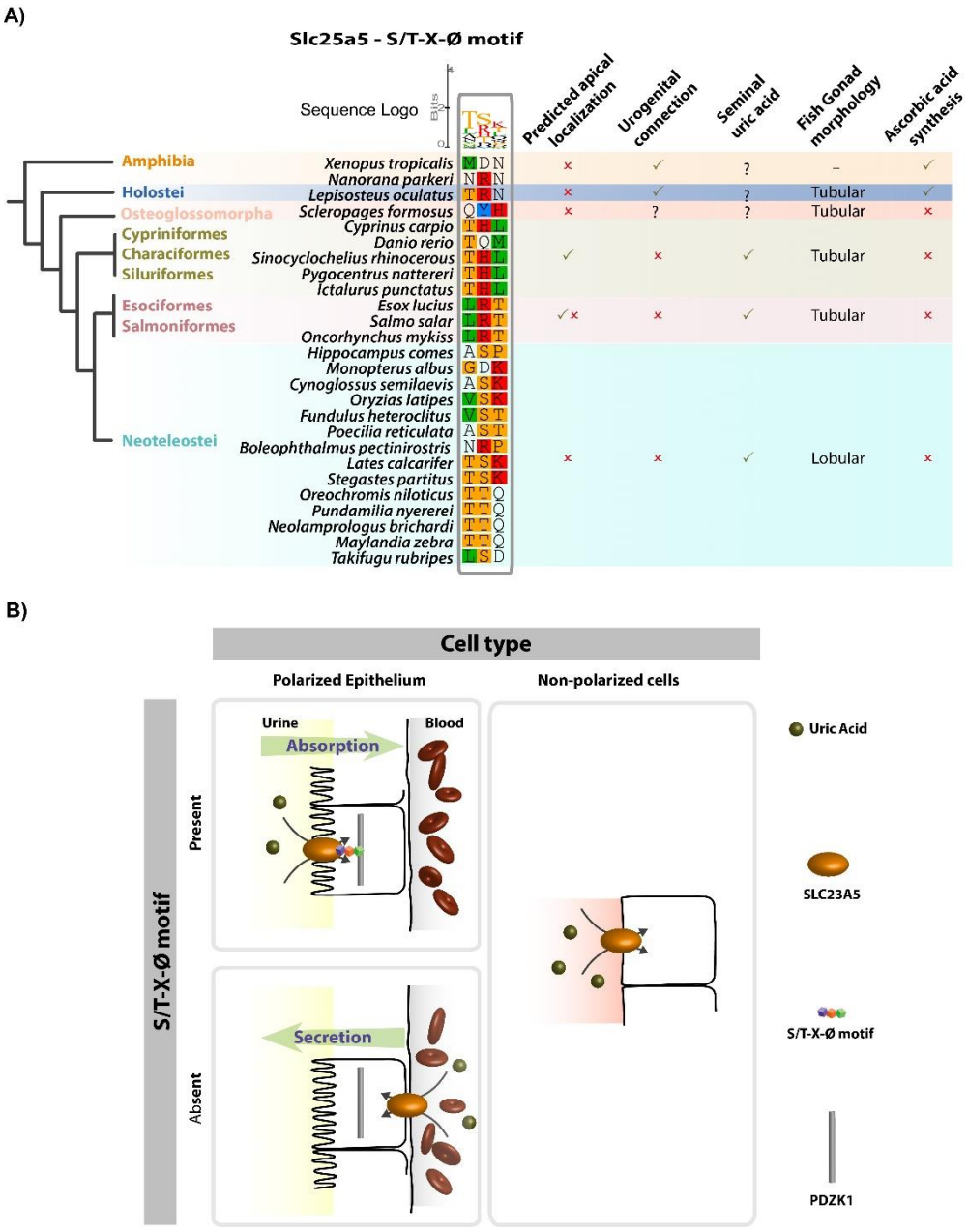
184 specificity: [Q/E/P]-N-x-G-x-x-x-x-T-[R/K/G]. Asterisks indicate conserved positions. The first

185 amino acid position, with residues highlighted in red within a black box, defines substrate

186 preference: Q – Uric acid/xanthine group; P – Ascorbate group and E – Uracil group. The mirror

187 Q in the Slc23a3 group is also highlighted by a black box.

188



189

190 **Figure 3 – The S/T-X-Ø tripeptide provides insight into the cellular function of Slc23a5**

191 **homologues from fish and amphibians.** (A) Alignment of the PDZ-binding tripeptide in Slc23a5

192 homologues from fish and amphibians. The S/T-X-Ø tripeptide is found in Cypriniformes,

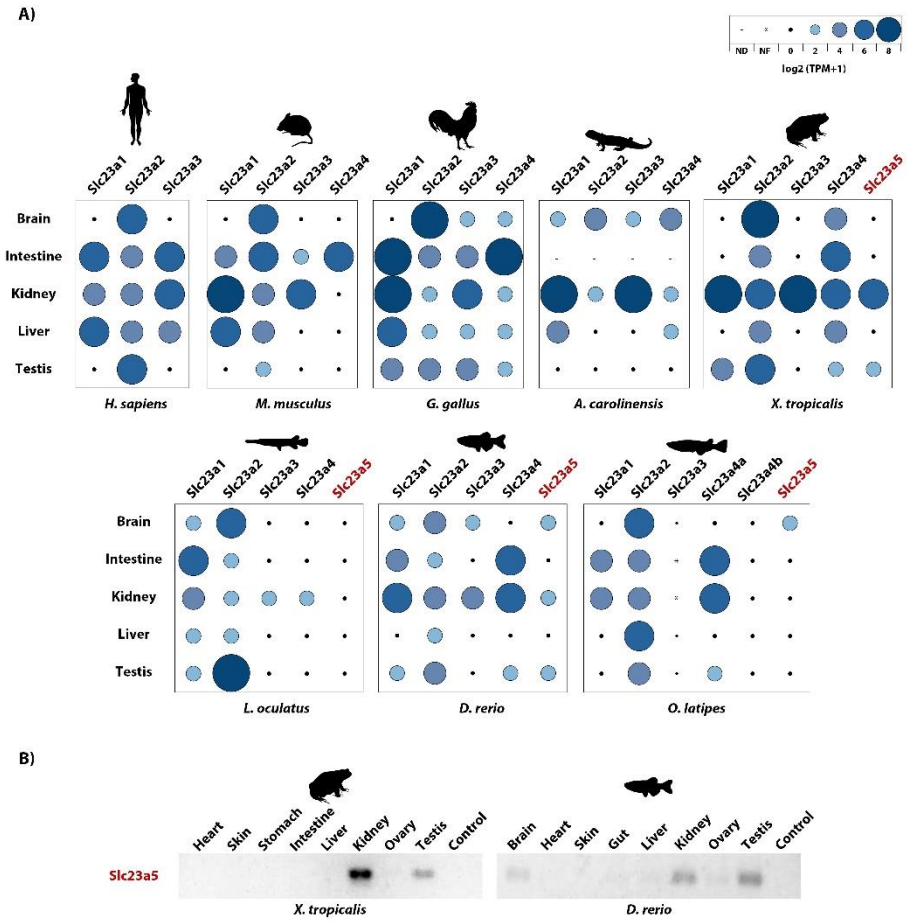
193 Characiformes and Siluriformes while the Euteleostei Esociformes and Salmoniformes exhibit an

194 inverted motif. Predicted cellular localization is indicated as well as morphological and

physiological adaptations among the different fish groups. (B) Schematic representation of tripeptide-dependent uric acid flows in polarized epithelia and non-polarized cells. In polarized epithelia (i.e. kidney) the presence of the tripeptide motif suggests interaction with the scaffolding protein PDZK1 which localizes apically, in the absorptive side.

3.3Slc23a5 expression patterns

To gain further insight into the functional relevance of this additional gene, we collected RNA-seq data and determined relative expression levels of *Slc23* genes. Gene expression analysis suggested that, in *D. rerio* and *X. tropicalis*, slc23a5 is mostly expressed in kidney and testis, further confirmed by PCR (Figure 4 and Table S5). Expression in brain was also detected in *D. rerio* and was found to be the exclusive expression site in the Neoteleostei *O. latipes*. The holostean *L. oculatus*, on the other hand, yielded no significant expression in the tested tissues.



208 **Figure 4 – Expression patterns of *Slc23* isoforms across vertebrates.** (A) RNAseq-based
209 quantification. Relative gene expression levels are provided as log_2-transformed transcript per
210 million (TPM) values, from 0 to 8 TPM's, the NF and ND means gene not found and no SRA data
211 available to a specific tissue. (B) PCR-based amplification of *Slc23a5* transcripts in *D. rerio* and *X.*
212 *tropicalis* tissues.
213

214 **4. Discussion**

215 Our findings indicate a restricted distribution of *slc23a5* to taxa that share an aquatic nature
216 of their habitat. In fact, *slc23a5* was found to be expressed in kidney and testis of both *X. tropicalis*
217 and *D. rerio*. In the latter, as well as in the Neoteleostei *O. latipes*, brain expression was also
218 detected. Comparative analysis using previously characterized NAT family transporters placed
219 this novel transporter in the uric acid and/or xanthine transporter group. In fact, similarly to uric
220 acid and/or xanthine transporters, Slc23a5 proteins harbor a Q in the first position of the NAT
221 family signature motif ([Q/E/P]-N-x-G-x-x-x-x-T-[R/K/G]). Yet, further functional studies are
222 required to fully clarify the molecular function of Slc23a5 proteins, notably its uric acid and/or
223 xanthine binding abilities.

224 Given the nature of NAT family transporter ligands, including ascorbic and uric acids,
225 xanthine or nucleobases, renal Slc23a5 could participate in the homeostatic maintenance of
226 excretion products, in osmotic regulation and/or in antioxidant processes. In fact, kidneys play a
227 crucial role in the maintenance of both ascorbic and uric acid levels through the excretion and
228 reabsorptions processes [6,22]. In kidney absorptive epithelium, transporters located on the
229 filtrate-epithelial interface (apical side) interact with the scaffolding protein PDZ-Containing
230 Kidney Protein 1 (PDZK1) [22]. PDZK1 belongs to the NHERF family of proteins which are
231 mostly expressed in the apical side of absorptive epithelia of kidney, intestine and liver, and
232 establish protein-protein interaction networks with receptors, transporters and channels [23,24].
233 Besides its scaffolding role, NHERF family members were also shown to modulate the activity of

234 protein assemblages [24-26]. Within Slc23 family members, S/T-X-Ø motifs are found in Slc23a1
235 and Slc23a4 proteins, which are present in renal and intestinal polarized epithelia and participate
236 in ascorbic acid and nucleobase absorption, and absent in Slc23a2. Curiously, a highly conserved
237 PDZK1 interacting motif (S/T-X-Ø) was also detected in Slc23a5 from Cypriniformes,
238 Characiformes and Siluriformes, but not in other fish groups or amphibians. No relationship
239 between salt and freshwater environments was found. Thus, the presence/absence of this peptide
240 signal might reflect distinct flows in the kidney: an apically expressed Slc23a5 likely participates
241 in absorptive flows, while a basolateral expression would be related to secretion (Figure 3B).

242 In the human proximal renal tubules, several transporters, localized to specific cellular
243 regions (apical or basolateral), seem to cooperate for efficient urate handling: named the urate
244 transportome [6,22]. Thus, if the predicted uric acid transport activity is confirmed, *X. tropicalis*
245 and *D. rerio* Slc23a5 could illustrate the two opposing mechanisms: while *D. rerio* captures, *X.*
246 *tropicalis* secretes uric acid into renal filtrate. In agreement, *X. tropicalis*, *D. rerio*, as well as *L.*
247 *oculatus*, appear to lack the full complement of renal urate transporters (Table S6). A similar
248 scenario is observed for birds and reptiles; however, their excretory system represents a particular
249 adaptation for uric acid excretion [27,28]. Besides simple excretion, and given that high
250 concentrations of uric acid were detected, and suggested to act as the main non-enzymatic
251 antioxidant mechanism, in seminal fluid of several fish species (including Cypriniformes,
252 Esociformes, Salmoniformes and the Neoteleostei Gadiformes and Perciformes [29-32]), this
253 putative uric acid transport activity could be maintained for reproductive purposes, in agreement
254 with the observed testicular expression. In fact, in external fertilizers, such as most fish and
255 amphibians, reactive oxygen species were suggested to decrease motility and thus fertilisation
256 success [29]. Yet, regarding amphibians, no information on seminal fluid composition was
257 retrieved [33]. Nonetheless, in vitro assays suggested that other antioxidant species, such as
258 ascorbic acid or vitamin E, were unsuitable at offsetting oxidative stress and improving sperm
259 motility in *Litoria booroolongensis* (Booroolong frog) and suggested uric acid as a pertinent

260 candidate [33]. In this context, the absence of *slc23a5* sequences in Chondrichthyes could be due
261 to their internal fertilization strategies derived from ovoviviparous or viviparous reproductive
262 modes [34]. Accordingly, viviparity is also observed in the sarcopterygian *L. chalumnae* [34].

263 Besides reproductive categories, other morphological and metabolic traits could justify the
264 seemingly scattered retention of *slc23a5* among fish in the context of the proposed uric acid
265 transport activity; for instance, specific anatomical features such as the presence of a urogenital
266 connection: described in Elasmobranchii and in most non-teleost Osteichthyes, yet absent in
267 Cyclostomata and Teleostei [35]. By connecting testicular and renal ducts, antioxidant
268 requirements could be met by renal secretion into urine. In agreement, in the Chondrostei
269 *Acipenser ruthenus*, full sperm maturation and motility was shown to require dilution by urine
270 [36]. An urogenital connection is also present in amphibians [35]. Additionally, other antioxidant
271 mechanisms could exist. Besides uric acid, ascorbic acid is also present in seminal fluid of
272 *Oncorhynchus mykiss* (rainbow trout) [37]. Seminal ascorbate was reported to be affected by
273 seasonal variations and semen quality was shown to improve upon dietary ascorbic acid
274 supplementation [37,38], contrasting with the amphibian *L. booroolongensis* [33]. In the case of non-
275 teleost fish, such as Chondrostei (*A. ruthenus*) and Holostei (*L. oculatus*), a functional GLO gene is
276 retained, providing endogenously synthesised ascorbate [39]. Curiously, no gonadal expression
277 was detected in the Neoteleostei *O. latipes*. Neoteleosts exhibit varied reproduction modes
278 ranging from external fertilizers to internal fertilizers with viviparity [34,40-42]. Also, neoteleosts
279 were shown to exhibit a distinctive gonadal morphology, the lobular testis, whereas the
280 remaining bony fish exhibit tubular testis [43]. Although still obscure, the absence of testicular
281 slc23a5 expression in the externally fertilizer *O. latipes* could be due to specific morphological
282 adaptations of neoteleosts.

283 A similar role could be postulated for neuronal uric acid [44]. Active swimming behaviour
284 requires a high metabolic rate, mirrored by an increased oxygen consumption, subsequently
285 leading to reactive oxygen species production [45]. In agreement, fish antioxidant levels were

shown to be correlated with their swimming activity [45]. Neuronal cell populations are particularly sensitive to oxidative stress, thus, uric acid could participate in the neutralization of oxidative species to by-pass oxidative stress, notably in GLO-deficient fish. Yet, information on neuronal uric acid levels in fish is currently absent.

5. Conclusion

Taken together, these observations put forward the retention of an additional *Slc23* gene, in neopterygians and amphibians, with restricted expression patterns and presumed uric acid transport function. Yet, extensive comparative functional analysis should be conducted to clarify the physiological relevance of this novel receptor in aquatic organisms.

Supplementary material: Supplementary tables and figures including synteny, phylogeny and topology analysis, as well as reference tables.

Acknowledgments : This work was supported by Coral—Sustainable Ocean Exploitation (Norte-01-0145-FEDER-000036), 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.

References

1. Burzle, M.; Suzuki, Y.; Ackermann, D.; Miyazaki, H.; Maeda, N.; Clemencon, B.; Burrier, R.; Hediger, M.A. The sodium-dependent ascorbic acid transporter family SLC23. *Mol. Aspects Med.* **2013**, *34*, 436-454, doi:10.1016/j.mam.2012.12.002.
2. Linster, C.L.; Van Schaftingen, E. Vitamin C. *FEBS J.* **2007**, *274*, 1-22, doi:10.1111/j.1742-4658.2006.05607.x.
3. Drouin, G.; Godin, J.R.; Page, B. The genetics of vitamin C loss in vertebrates. *Curr. Genomics* **2011**, *12*, 371-378, doi:10.2174/138920211796429736.
4. Dutta Gupta, S.; Choudhury, P.K.; Chatterjee, I.B. Synthesis of l-ascorbic acid from d-glucurono-1,4-lactone conjugates by different species of animals. *Int. J. Biochem.* **1973**, *4*, 309-314, doi:http://dx.doi.org/10.1016/0020-711X(73)90053-0.

5. Wright, P.A. Nitrogen excretion: three end products, many physiological roles. *J. Exp. Biol.* **1995**, *198*, 273-281.
6. Bobulescu, I.A.; Moe, O.W. Renal transport of uric acid: evolving concepts and uncertainties. *Adv. Chronic Kidney Dis.* **2012**, *19*, 358-371, doi:10.1053/j.ackd.2012.07.009.
7. Alvarez-Lario, B.; Macarrón-Vicente, J. Evolution of Uric Acid Metabolism in Humans. In *eLS*, John Wiley & Sons, Ltd: 2001; 10.1002/9780470015902.a0024618.
8. Oda, M.; Satta, Y.; Takenaka, O.; Takahata, N. Loss of urate oxidase activity in hominoids and its evolutionary implications. *Mol. Biol. Evol.* **2002**, *19*, 640-653.
9. Corti, A.; Casini, A.F.; Pompella, A. Cellular pathways for transport and efflux of ascorbate and dehydroascorbate. *Arch. Biochem. Biophys.* **2010**, *500*, 107-115, doi:10.1016/j.abb.2010.05.014.
10. Savini, I.; Rossi, A.; Pierro, C.; Avigliano, L.; Catani, M.V. SVCT1 and SVCT2: key proteins for vitamin C uptake. *Amino Acids* **2008**, *34*, 347-355, doi:10.1007/s00726-007-0555-7.
11. Yamamoto, S.; Inoue, K.; Murata, T.; Kamigaso, S.; Yasujima, T.; Maeda, J.Y.; Yoshida, Y.; Ohta, K.Y.; Yuasa, H. 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.* **2010**, *285*, 6522-6531, doi:10.1074/jbc.M109.032961.
12. Koukaki, M.; Vlanti, A.; Goudela, S.; Pantazopoulou, A.; Gioule, H.; Tournaviti, S.; Diallinas, G. The nucleobase-ascorbate transporter (NAT) signature motif in UapA defines the function of the purine translocation pathway. *J. Mol. Biol.* **2005**, *350*, 499-513, doi:10.1016/j.jmb.2005.04.076.
13. Papakostas, K.; Frillingos, S. Substrate selectivity of YgfU, a uric acid transporter from Escherichia coli. *J. Biol. Chem.* **2012**, *287*, 15684-15695, doi:10.1074/jbc.M112.355818.
14. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772-780, doi:10.1093/molbev/mst010.
15. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307-321, doi:10.1093/sysbio/syq010.
16. Dobson, L.; Remenyi, I.; Tusnady, G.E. CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Res.* **2015**, *43*, W408-412, doi:10.1093/nar/gkv451.
17. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114-2120, doi:10.1093/bioinformatics/btu170.
18. Yates, A.; Akanni, W.; Amode, M.R.; Barrell, D.; Billis, K.; Carvalho-Silva, D.; Cummins, C.; Clapham, P.; Fitzgerald, S.; Gil, L., et al. Ensembl 2016. *Nucleic Acids Res.* **2016**, *44*, D710-716, doi:10.1093/nar/gkv1157.
19. O'Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufo, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D., et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733-745, doi:10.1093/nar/gkv1189.
20. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357-359, doi:10.1038/nmeth.1923.
21. Li, B.; Dewey, C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **2011**, *12*, 323, doi:10.1186/1471-2105-12-323.
22. Anzai, N.; Jutabha, P.; Amonpatumrat-Takahashi, S.; Sakurai, H. Recent advances in renal urate transport: characterization of candidate transporters indicated by genome-wide

association studies. *Clin. Exp. Nephrol.* **2012**, *16*, 89-95, doi:10.1007/s10157-011-0532-z.

23. Dev, K.K. Making protein interactions druggable: targeting PDZ domains. *Nat. Rev. Drug Discov.* **2004**, *3*, 1047-1056.
24. Clapéron, A.; Mergey, M.; Fouassier, L. Roles of the scaffolding proteins NHERF in liver biology. *Clin. Res. Hepatol. Gastroenterol.* **2011**, *35*, 176-181, doi:http://dx.doi.org/10.1016/j.clinre.2010.11.009.
25. Raghuram, V.; Mak, D.-O.D.; Foskett, J.K. Regulation of cystic fibrosis transmembrane conductance regulator single-channel gating by bivalent PDZ-domain-mediated interaction. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 1300-1305.
26. Wang, S.; Yue, H.; Derin, R.B.; Guggino, W.B.; Li, M. Accessory Protein Facilitated CFTR-CFTR Interaction, a Molecular Mechanism to Potentiate the Chloride Channel Activity. *Cell* **2000**, *103*, 169-179, doi:10.1016/S0092-8674(00)00096-9.
27. Braun, E.J. Comparative renal function in reptiles, birds, and mammals. *Semin. Avian Exotic. Pet. Med.* **1998**, *7*, 62-71, doi:http://dx.doi.org/10.1016/S1055-937X(98)80044-3.
28. Singer, M.A. Do mammals, birds, reptiles and fish have similar nitrogen conserving systems? *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2003**, *134*, 543-558.
29. Lahnsteiner, F.; Mansour, N. A comparative study on antioxidant systems in semen of species of the Percidae, Salmonidae, Cyprinidae, and Lotidae for improving semen storage techniques. *Aquaculture* **2010**, *307*, 130-140, doi:http://dx.doi.org/10.1016/j.aquaculture.2010.07.011.
30. Ciereszko, A.; Dabrowski, K.; Kucharczyk, D.; Dobosz, S.; Goryczko, K.; Glogowski, J. The presence of uric acid, an antioxidantive substance, in fish seminal plasma. *Fish Physiol. Biochem.* **1999**, *21*, 313-315, doi:10.1023/A:1007886121663.
31. Dzyuba, V.; Dzyuba, B.; Cosson, J.; Boryshpolets, S.; Yamaner, G.; Kholodniy, V.; Rodina, M. The antioxidant system of sterlet seminal fluid in testes and Wolffian ducts. *Fish Physiol. Biochem.* **2014**, *40*, 1731-1739, doi:10.1007/s10695-014-9963-2.
32. Lahnsteiner, F.; Mansour, N.; Plaetzer, K. Antioxidant systems of brown trout (Salmo trutta f. fario) semen. *Anim. Reprod. Sci.* **2010**, *119*, 314-321, doi:10.1016/j.anireprosci.2010.01.010.
33. Keogh, L.M.; Byrne, P.G.; Silla, A.J. The effect of antioxidants on sperm motility activation in the Booroolong frog. *Anim. Reprod. Sci.* **2017**, *183*, 126-131, doi:http://dx.doi.org/10.1016/j.anireprosci.2017.05.008.
34. Goodwin, N.B.; Dulvy, N.K.; Reynolds, J.D. Life-history correlates of the evolution of live bearing in fishes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci* **2002**, *357*, 259-267, doi:10.1098/rstb.2001.0958.
35. Blün, V. Comparative Anatomy of the Urogenital System. In *Vertebrate Reproduction, a Textbook*, Springer-Verlag: Berlin Heidelberg New York Tokyo, 1986; 10.1007/978-3-642-71074-2pp. 43-74.
36. Dzyuba, B.; Cosson, J.; Boryshpolets, S.; Bondarenko, O.; Dzyuba, V.; Prokopchuk, G.; Gazo, I.; Rodina, M.; Linhart, O. In vitro sperm maturation in sterlet, Acipenser ruthenus. *Reprod. Biol.* **2014**, *14*, 160-163, doi:10.1016/j.repbio.2014.01.003.
37. Ciereszko, A.; Dabrowski, K. Sperm quality and ascorbic acid concentration in rainbow trout semen are affected by dietary vitamin C: an across-season study. *Biol. Reprod.* **1995**, *52*, 982-988.
38. Ciereszko, A.; Dabrowski, K. Effect of ascorbic acid supplement in vitro onrainbow trout sperm viability. *Aquacult. Int.* **2000**, *8*, 1-8, doi:10.1023/A:1009253330451.
39. Moreau, R.; Dabrowski, K. Biosynthesis of ascorbic acid by extant actinopterygians. *J. Fish Biol.* **2000**, *57*, 733-745, doi:10.1111/j.1095-8649.2000.tb00271.x.

40. Forsman, A.; Tibblin, P.; Berggren, H.; Nordahl, O.; Koch‐Schmidt, P.; Larsson, P. Pike Esox lucius as an emerging model organism for studies in ecology and evolutionary biology: a review. *J. Fish Biol.* **2015**, *87*, 472-479, doi:10.1111/jfb.12712.
41. Bone, Q.; Moore, R.H. *Biology of Fishes* Taylor & Francis Group: New York, 2008.
42. Scott, D.B.C.; Fuller, J.D. The reproductive biology of Scleropages formosus (Müller & Schlegel) (Osteoglossomorpha, Osteoglossidae) in Malaya, and the morphology of its pituitary gland. *J. Fish Biol.* **1976**, *8*, 45-53, doi:10.1111/j.1095-8649.1976.tb03906.x.
43. Parenti, L.R.; Grier, H.J. Evolution and phylogeny of gonad morphology in bony fishes. *Integr. Comp. Biol.* **2004**, *44*, 333-348, doi:10.1093/icb/44.5.333.
44. Uttara, B.; Singh, A.V.; Zamboni, P.; Mahajan, R.T. Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options. *Curr. Neuropharmacol.* **2009**, *7*, 65-74, doi:10.2174/157015909787602823.
45. Lopez-Cruz, R.I.; Dafre, A.L.; Filho, D.W. Oxidative Stress in Sharks and Rays. In *Oxidative Stress in Aquatic Ecosystems*, Abele, D., Vazquez-Medina, J.P., Zenteno-Savin, T., Eds. Wiley-Blackwell: 2011; pp. 157-164.

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

436 **Supplementary Figures:**

437 Figure S1 – RaxML phylogenetic tree.

438 Figure S2 - Synteny maps of *Slc23a5*.

439 Figure S3 - Synteny maps of *Slc23a1*.

440 Figure S4 - Synteny maps of *Slc23a2*.

441 Figure S5 - Synteny maps of *Slc23a3*.

442 Figure S6 - Synteny maps of *Slc23a4.*

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

444 Figure S8 – Topology prediction of Slc23a5.

445 **Supplementary Tables:**

446 Table S1 - Accession numbers.

447 Table S2 - Accession numbers of the RNAseq files.

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

449 Table S4 - Genome and GTF files retrieved from Ensemble database (Release 89) and Transcriptome files retrieved from NCBI used for

450 this study.

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

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

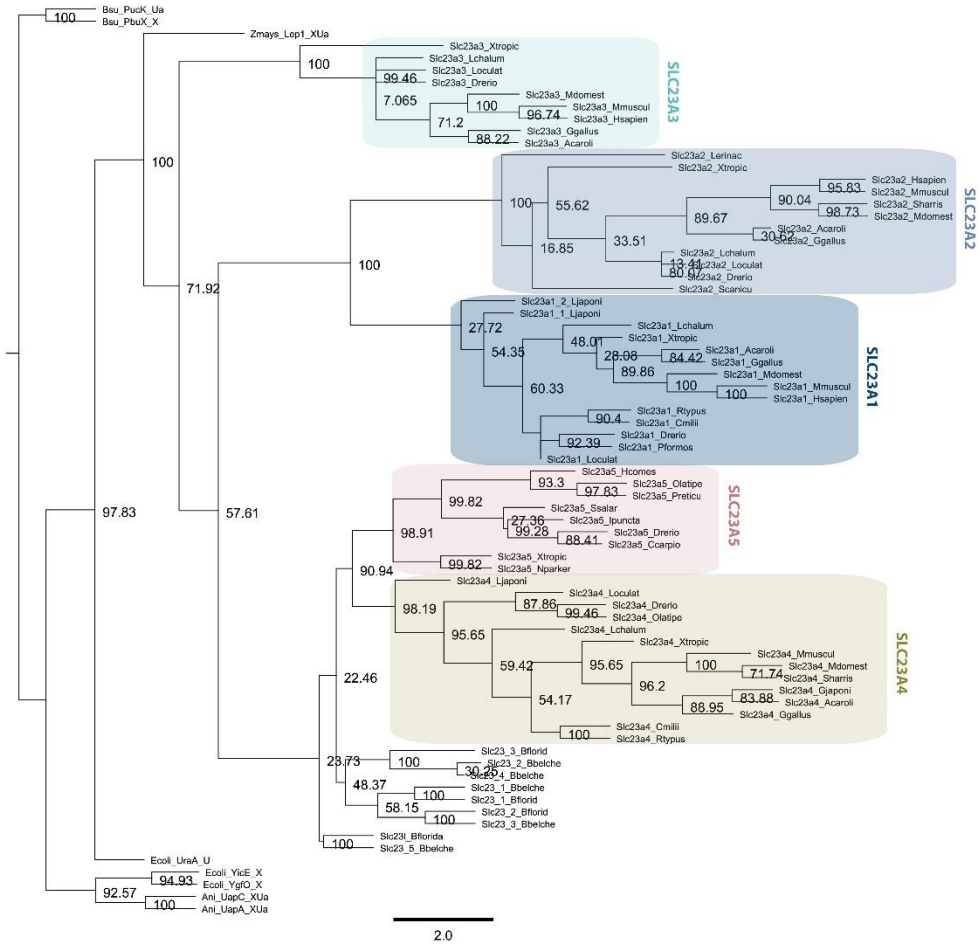
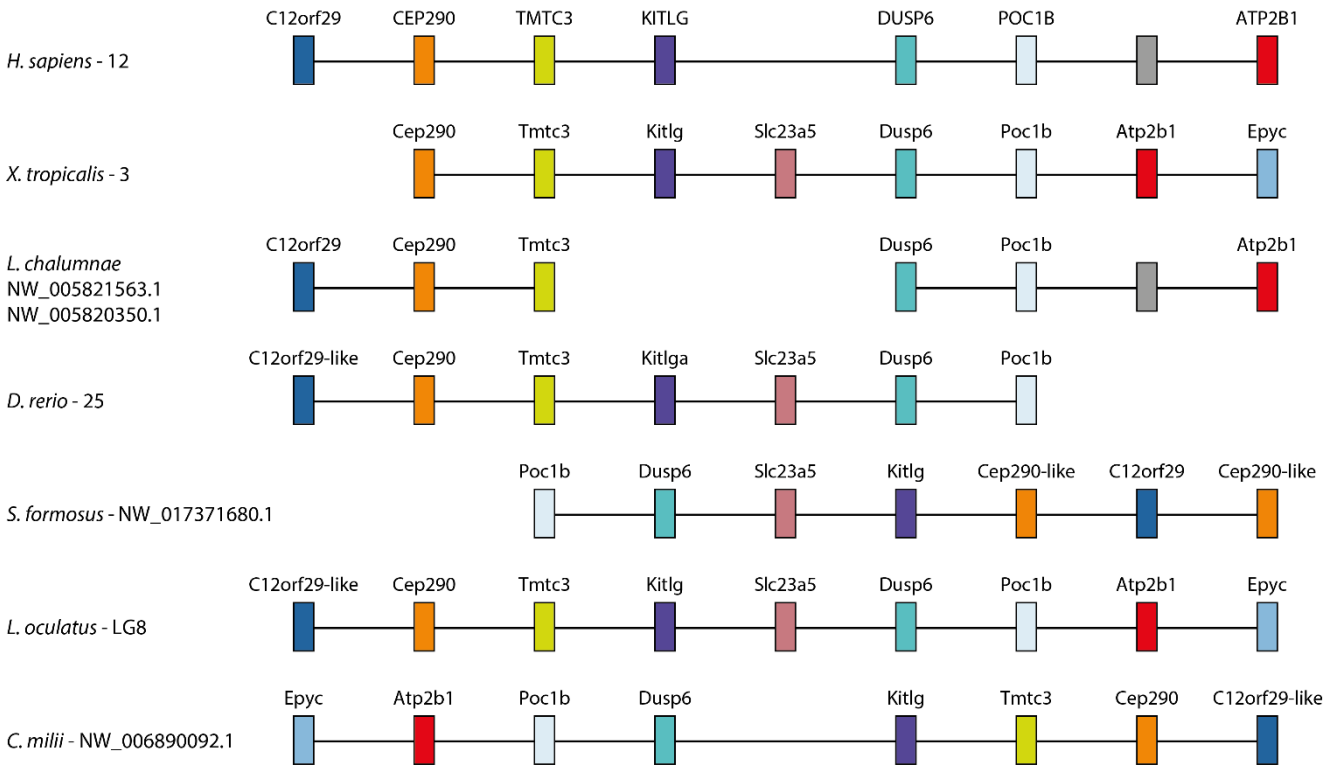


Figure S1 – RaxML phylogenetic tree. Node values represent bootstrap values.

456



457

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

461

462

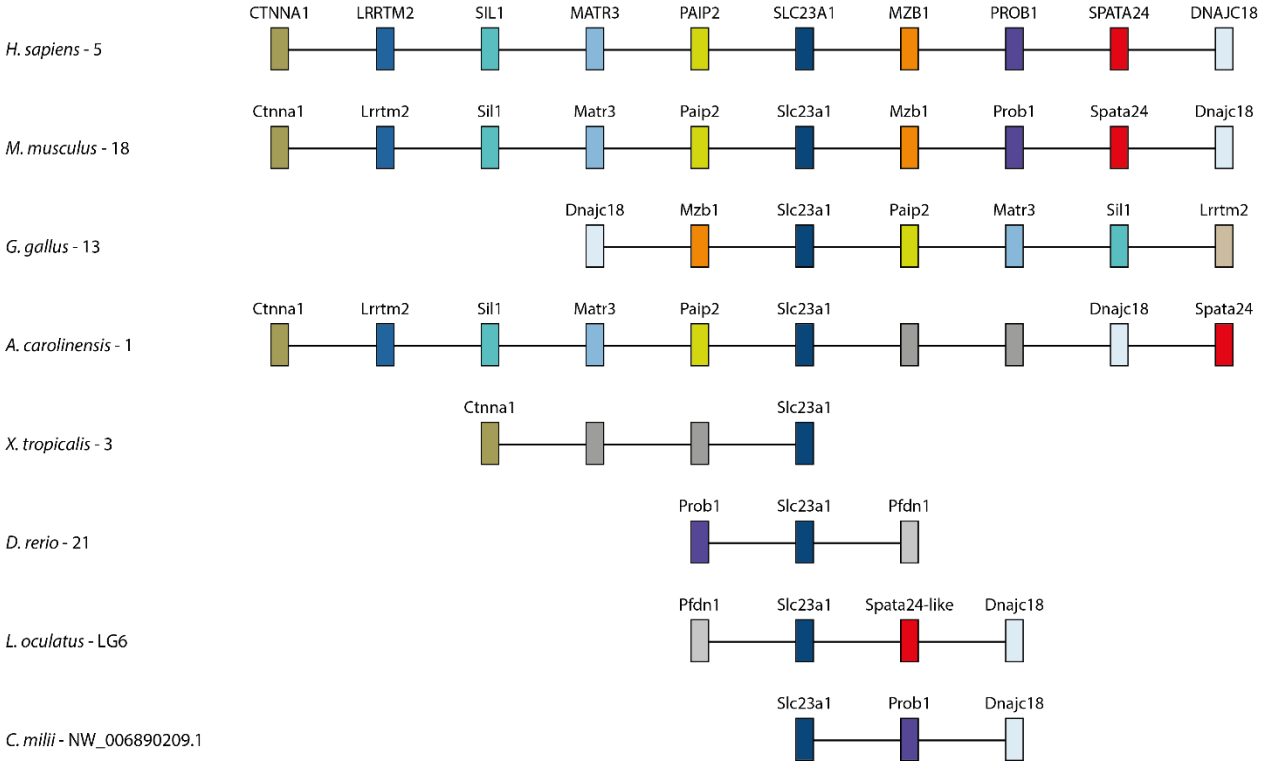


Figure S3 - Synteny maps of *Slc23a1*. Conservation of the *Slc23a1* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes.

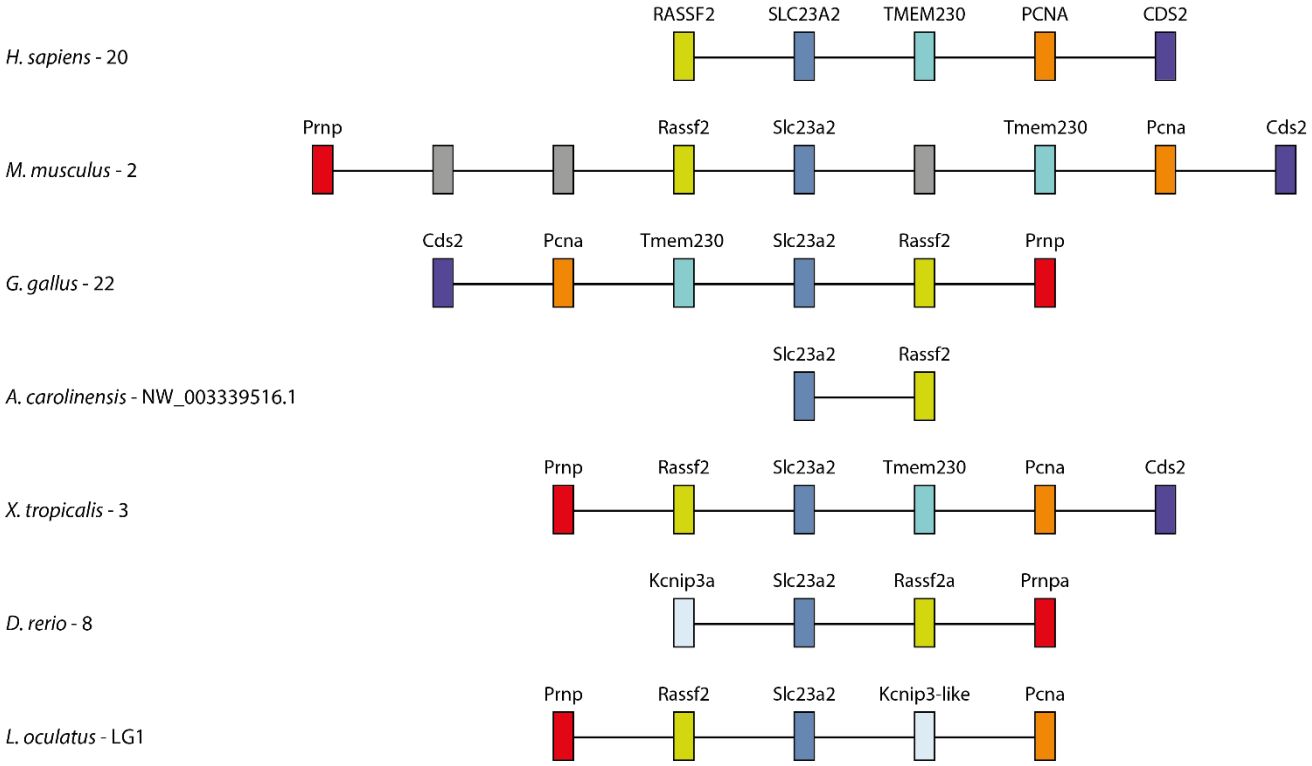


Figure S4 - Synteny maps of *Slc23a2*. Conservation of the *Slc23a2* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes.

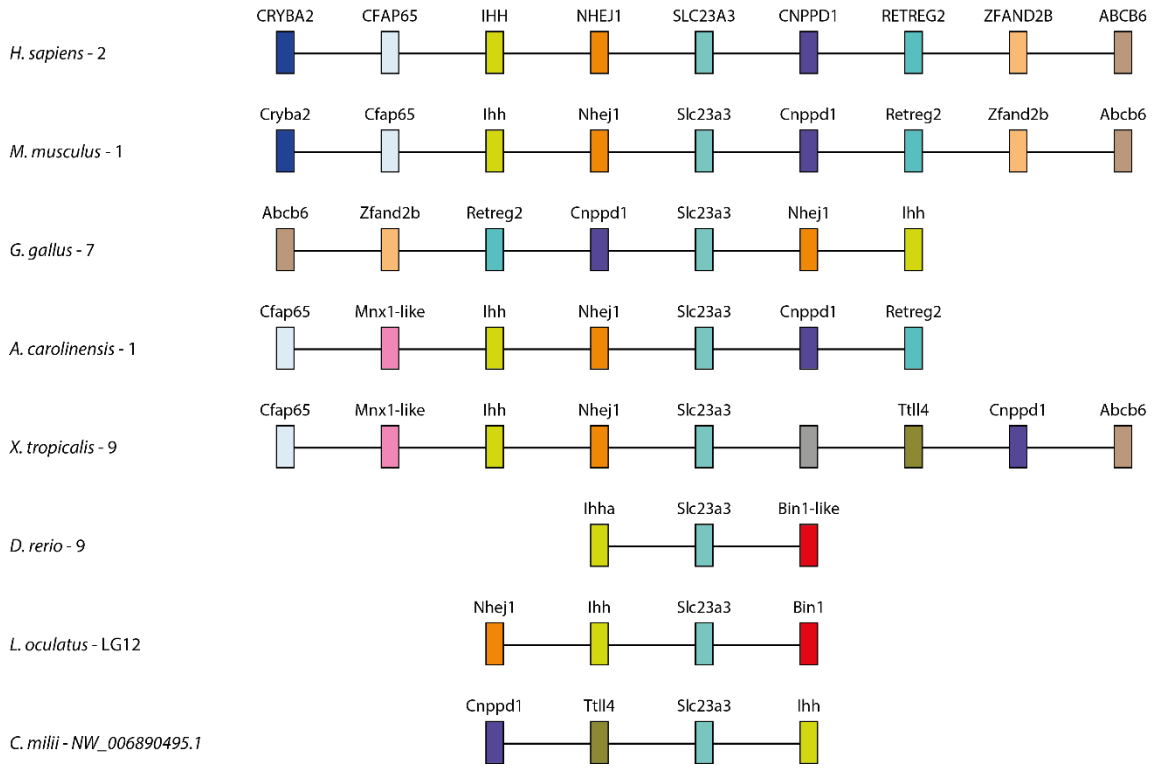


Figure S5 - Synteny maps of *Slc23a3*. Conservation of the *Slc23a3* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. Grey boxes indicate non-conserved genes.

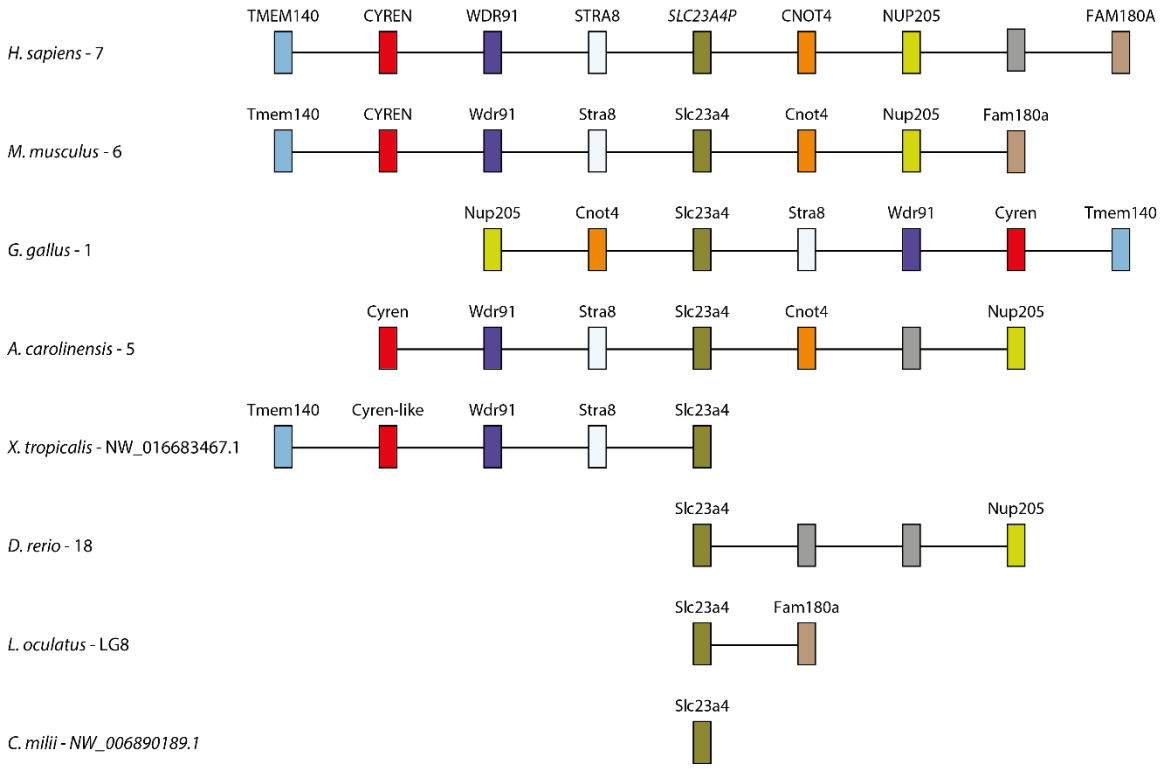


Figure S6 - Synteny maps of *Slc23a4*. Conservation of the *Slc23a4* locus in major vertebrate lineages. Numbering after species corresponds to chromosome or scaffold. The human homologue is indicated as pseudogenized (*SLC23A4P*). Grey boxes indicate non-conserved genes.

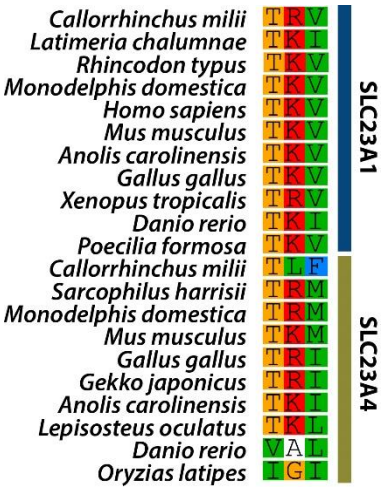
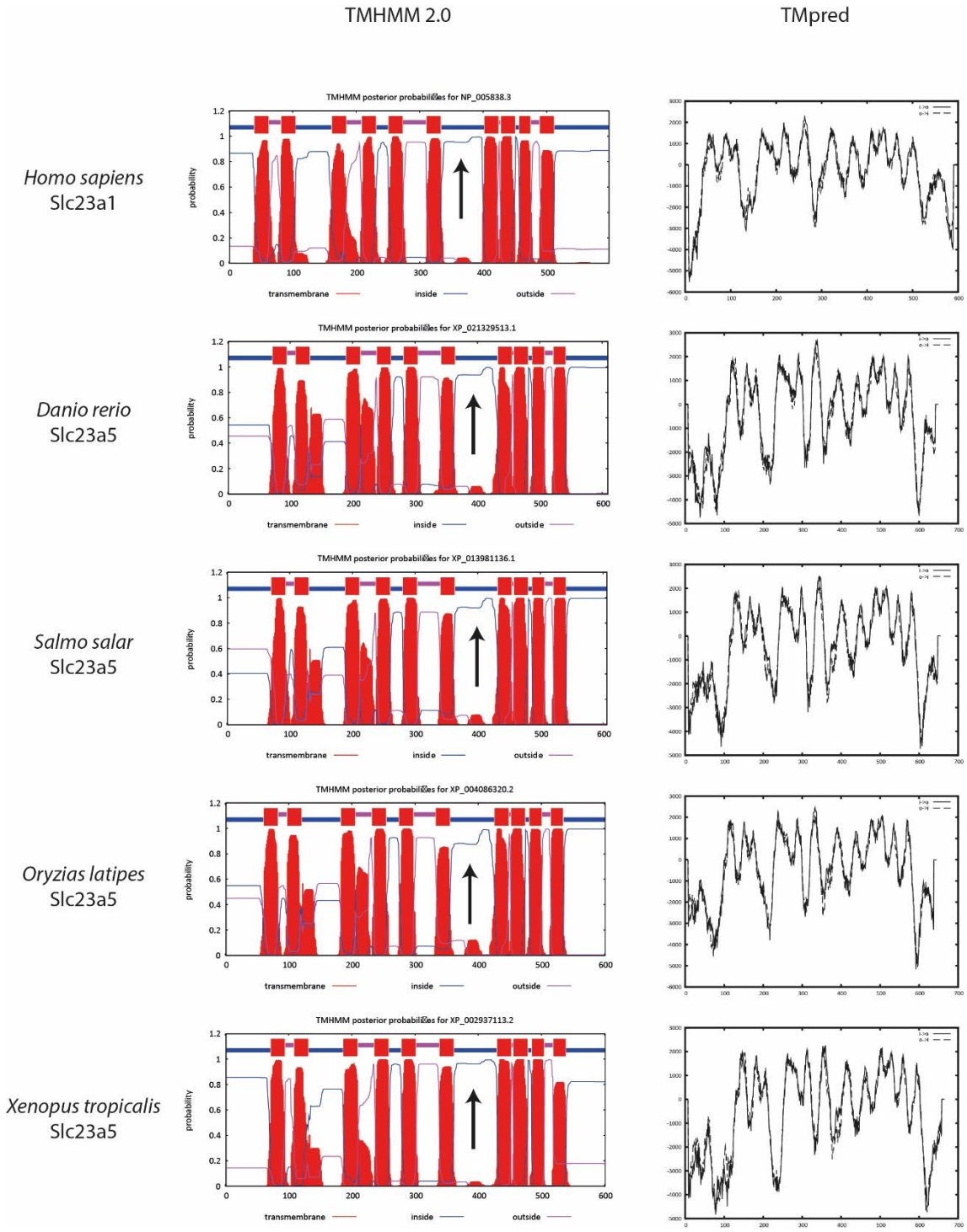


Figure S7 – S/T-X-Ø tripeptide conservation in Slc23a1 and Slc23a4 sequences.



481

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

486

487 **Supplementary Tables:**

488 **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

489

490 **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

491

492

493

494 **Table S3 - Gene identifier for all genes included in gene expression analysis, collected**
495 **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

496

497

498

499 **Table S4 - Genome and GTF files retrieved from Ensemble database (Release 89) and Transcriptome files retrieved from NCBI used for this**
500 **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

501

502

503

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

511 **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.

512