THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Phylogenetic reclassification of vertebrate melatonin receptors to include Mel1d

OPEN ACCESS

Phylogenetic reclassification of vertebrate melatonin receptors to include Mel1d

Elsa Denker [*], Lars O. E. Ebbesson [*], David G. Hazlerigg [†], Daniel J. Macqueen [‡]

[*] *NORCE Environment, NORCE Norwegian Research Centre, 5008 Bergen, Norway*

[†] *Department of Arctic and Marine Biology, UiT: the Arctic University of Norway, 9037 Tromsø, Norway*

[‡] *The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, EH25 9RG Edinburgh, United Kingdom*

13

Running title: Mel1d - a new vertebrate melatonin receptor

15

Keywords: Melatonin receptor; Mel1d; vertebrate evolution; phylogenetic analysis;

comparative genomics; conserved synteny; teleost fish

18

Corresponding authors:

Daniel J. Macqueen, The Roslin Institute, University of Edinburgh, Easter Bush

Campus, Midlothian, EH25 9RG, United Kingdom. Phone: (+44) 131 651 9249.

Email: daniel.macqueen@roslin.ed.ac.uk

23

Elsa Denker, Integrative Fish Biology Group, NORCE Environment, NORCE

Norwegian Research Center, Thormøhlensgate 55, 5008 Bergen, Norway. Phone:

(+47) 55 58 44 16. Email: elsa.denker@uib.no

27

28

29   **ABSTRACT**

30

31   The circadian and seasonal actions of melatonin are mediated by high affinity G-

32   protein coupled receptors (melatonin receptors, MTRs), classified into

33   phylogenetically distinct subtypes based on sequence divergence and

34   pharmacological characteristics. Three vertebrate MTR subtypes are currently

35   described: MT1 (MTNR1A), MT2 (MTNR1B), and Mel1c (MTNR1C / GPR50), which

36   exhibit distinct affinities, tissue distributions and signaling properties. We present

37   phylogenetic and comparative genomic analyses supporting a revised classification

38   of the vertebrate MTR family. We demonstrate four ancestral vertebrate MTRs,

39   including a novel molecule hereafter named Mel1d. We reconstructed the evolution

40   of each vertebrate MTR, detailing genetic losses in addition to gains resulting from

41   whole genome duplication events in teleost fishes. We show that Mel1d was lost

42   separately in mammals and birds and has been previously mistaken for an MT1

43   paralogue. The genetic and functional diversity of vertebrate MTRs is more complex

44   than appreciated, with implications for our understanding of melatonin actions in

45   different taxa. The significance of our findings, including the existence of Mel1d, are

46   discussed in an evolutionary and functional context accommodating a robust

47   phylogenetic assignment of MTR gene family structure.

48

49

INTRODUCTION

51

Melatonin is an ancient eukaryotic signalling molecule that regulates diverse biological functions. While best known as a regulator of biological rhythms in humans, this hormone also regulates energy balance, temperature, behavior, blood pressure, and seasonal reproduction. Melatonin is secreted by the pineal gland and targets the brain as well as peripheral tissues (Hardeland *et al.* 2011, Slominski *et al.* 2012), but is also produced by several tissues, eliciting paracrine effects (Weaver and Reppert 1990). The actions of melatonin depend on the spatiotemporal expression of high-affinity melatonin receptors (MTR), representing a specific class of G protein-coupled receptor (GPCR).

61

Three paralogous MTR family members have been characterized in jawed vertebrates, namely MT1 (Mel1a / MTNR1A), MT2 (Mel1b / MTNR1B), and Mel1c (MTNR1C / GPR50 in mammals) (Reppert *et al.* 1994, 1995a, 1995b). Despite showing overlap in expression, these different MTRs have evolved unique functions. MT1 has a higher affinity for melatonin than MT2 (Dubocovich and Markowska 2005), and in mammals, Mel1c has lost the ability to bind melatonin (Dufourny *et al.* 2008), though it does modulate melatonin signaling via its association with MT1 (Levoye *et al.* 2006). While MT1 associates with a range of G proteins to activate several distinct signalling pathways, eliciting wide-ranging cellular effects (Witt-Enderby *et al.* 2003), MT2 associates with a single G protein (Jockers *et al.* 2008). Owing to such functional divergence, different MTRs may have very distinct biological effects, even when expressed in the same cell types (*e.g.* Dubocovich and Markowska 2005).

74

75   A past study demonstrated melatonin binding in the brain of jawed vertebrates and

76   lamprey, but not in hagfishes or amphioxus (Vernadakis *et al.* 1998). Thus, it is likely

77   that high-affinity MTRs were present in the vertebrate ancestor, and were secondarily

78   lost in some jawless fishes, as noted for several other traits (*e.g.* reduction of

79   vertebrae-like elements - Ota et al. 2011; *Dlx* genes - Sugahara *et al.* 2013; reviewed

80   in Kuraku 2013). MTR-like GPCR genes have also been discovered in urochordates,

81   cephalochordates, hemichordates and echinoderms (Kamesh *et al.* 2008, Nordstrom

82   *et al.* 2008, Krishnan *et al.* 2013), but their evolutionary affinity to the vertebrate

83   MTRs remains ambiguous. The distinct MTRs of jawed vertebrates potentially

84   originated during two rounds (2R) of whole genome duplication (WGD) at the stem of

85   vertebrate evolution (*e.g.* Dehal and Boore, 2005), though this is yet to be

86   established. Additional expansions in the MTR family of fishes (*e.g.* Shang &

87   Zhdanova 2007; Hong *et al.* 2014) may owe to a further round of teleost-specific

88   WGD ('Ts3R') in the common teleost ancestor, or additional lineage-specific WGD

89   events in some lineages, *e.g.* the salmonid-specific 4R ('Ss4R') (Macqueen and

90   Johnston, 2014; Lien *et al.* 2016), though, again, this has not been properly explored.

91

92   The overarching goal of this study was to re-examine the evolutionary history of

93   vertebrate MTRs, using data in publically-available sequence databases for robust

94   phylogenetic and comparative genomic reconstructions. Our findings concretely

95   demonstrate a fourth ancestral MTR ('Mel1d'), along with teleost-specific expansions

96   in MTR diversity, likely owing to Ts3R and Ss4R. With a new evolutionary framework

97   in place we reinterpret findings on vertebrate MTR sequence divergence and

98   expression from past studies. Overall, this study highlights substantial unexplored

99   diversity in MTR signalling within vertebrates, pointing to new lines of investigation.

100

MATERIALS AND METHODS

102

103 **Sequence and phylogenetic analyses**

104 Amino acid sequences encoded by MTR or FAT protocadherin family member genes

105 were collected from representative jawed vertebrate species with high-quality

106 genome assemblies. Details of these sequences are given in Table S1 (MTR) and

107 Table S2 (FAT), which include database accession numbers and nomenclature

108 matching the findings of our phylogenetic analyses. As a start point for the analysis,

109 MTR/FAT proteins of human (*i.e.* MT1/MT2/Mel1c/GPR50 or FAT1/2/3) were used as

110 queries in BLASTp (Altschul *et al.* 1997) searches to identify homologues within the

111 NCBI database (https://www.ncbi.nlm.nih.gov/). We also used the Ensembl genome

112 browser (https://www.ensembl.org/) to collect MTR family proteins from several

113 species, using the EnsemblCompara method (Vilella *et al.* 2009).

114

115 The sequences were aligned using MAFFT v.7 (Katoh and Standley, 2013) with

116 default settings and subjected to quality filtering using GBlocks with default settings

117 (Talavera and Castresana, 2007). Final alignments of 300 (MTR) and 2,540 (FAT)

118 amino acid positions (Additional Dataset 1) were used for tree building, done using

119 Bayesian (BY) and maximum likelihood (ML) (MTR) or just ML (FAT) methods. ML

120 trees were generated using IQ-TREE (Nguyen *et al.* 2015) via the IQ-TREE

121 webserver (Trifinopoulos *et al.* 2016), employing the best-fitting amino acid

122 substitution model selected with ModelFinder (Kalyaanamoorthy *et al.* 2017) under

123 the Bayesian information criterion. The best fit models were JTT+F+I+ G4 for MTR

124 and JTT+G4+I for FAT, where 'JTT' is the general matrix of Jones *et al.* 1992, '+I'

125 includes empirical estimation of the proportion of invariant sites, '+F' includes

126 empirical estimation of amino acid frequencies and '+G4' denotes estimation of the

gamma distribution parameter with 4 rate classes. The stability of branching in the ML trees was assessed using 1,000 ultrafast bootstrap replicates, (Hoang *et al.* 2018). The BY analysis (MTR dataset) was done in BEAST v1.8.3 (Drummond *et al.* 2012), employing an uncorrelated relaxed clock model (Drummond *et al.* 2005) and a Yule speciation prior (Gernhard, 2008), along with the best-fitting substitution model selected by IQ-TREE. A Markov chain Monte Carlo (MCMC) chain of 50 million generations was generated and sampled every 5,000 generations. Convergence of the MCMC chain was assessed using Tracer v1.7.1 http://beast.bio.ed.ac.uk/tracer). A maximum clade credibility tree was generated in TreeAnnotator (Drummond *et al.* 2012) after removal of the first 10% sampled trees.

Comparative genomic and sequence analyses

Synteny analyses were performed using Ensembl genome browser annotations via the Genomicus platform (Nguyen *et al.* 2018). These analyses were supplemented with data from NCBI GenBank for species not available in Ensembl. Gene prediction and annotation for *Lethenteron camtschaticum* was performed using FGENESH (Soloyvev *et al.* 2006). Comparative analyses of MTR family amino acid sequences was done using the final alignment described above (note: the Gblocks filtering step served to remove flanking regions outside the transmembrane/loop regions, which were unaltered). The sequence similarity of the proposed vestigial MTR-like pseudogenes identified in our synteny analyses was established using BLASTx within the NCBI database.

Data Availability

Supplemental material described in the paper is available at Figshare: https://gsajournals.figshare.com/s/56f29e83cc0ec8748842. Fig. S1. ML phylogenetic

153   analysis of MTRs in vertebrates. This analysis was done using IQ-TREE with a high-

154   confidence alignment of eighty MTRs (300 amino acid positions; Additional Dataset

155   1) and the best-fitting amino acid substitution model (JTT+F+I+G4). Numbers on

156   branches are bootstrap support values. Other details as in the Fig. 1 legend (see

157   main text) Table S1 provides details of all protein sequences used for phylogenetic

158   analyses of the vertebrate MTR family. Table S2 provides details of all sequences

159   used for phylogenetic analyses of the vertebrate FAT protocadherin family

160   Additional Dataset 1 is the MTR sequence alignment used for phylogenetic analysis

161   and comparative sequence analysis. Additional Dataset 2 is the FAT alignment used

162   for phylogenetic analysis.

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

181   **RESULTS**

182

183   **Four MTRs are retained in jawed vertebrates**

184   We identified eighty unique MTR family member proteins in sequence databases

185   representing a standardized set of eighteen jawed vertebrate lineages (see

186   MATERIALS AND METHODS). A phylogenetic analysis was done using a BY

187   method (Fig. 1) incorporating a relaxed molecular clock model, which allows

188   estimation of the most plausible root location in the tree (Drummond *et al.* 2006; *e.g.*

189   Macqueen and Wilcox 2014; Redmond *et al.* 2018). Four distinct MTR clades (Fig. 1)

190   had strong statistical support (posterior probability, PP: >0.96), and each was

191   represented by cartilaginous fish, as well as ray-finned and lobe-finned fish lineages,

192   with branching patterns closely matching expected species phylogeny (Fig. 1). Three

193   of these clades correspond to known ancestral vertebrate MTR family members (*e.g.*

194   Dufourny *et al.* 2008). The fourth clade is hereafter named 'Mel1d'. The same four

195   clades were strongly supported in an unrooted ML phylogenetic analysis (bootstrap

196   support: >96%) congruent with the BY tree (Fig. S1).

197

198   These analyses indicate that four distinct MTRs existed in the jawed vertebrate

199   ancestor. However, the phylogenetic affinity of the four MTRs remains equivocal in

200   the BY analysis, with moderate support for Mel1d/MT1 (PP: 0.87) and MT2/Mel1C

201   (PP: 0.53) being paralogues, which can be explained parsimoniously by 2R (Fig. 1).

202

203   **Evolutionary history of individual vertebrate MTRs**

9

204   Expanding on the above findings, we reconstructed a more detailed evolutionary

205   history for each ancestral MTR in jawed vertebrates, accommodating gene losses, in

206   addition to gains resulting from WGD events in teleosts (summarized in Fig. 2).

207

208   Mel1d was encoded by a single gene in all represented species (Fig. 1, Fig. 2a)

209   including teleosts, consistent with the loss of any paralogues created during Ts3R

210   and Ss4R. In lobe-finned fish, Mel1d was identified in a coelacanth, an amphibian,

211   and two reptiles, but was not identified in the mammals and birds represented in our

212   trees (Fig. 2a). As only a small number of bird and mammals were included, we

213   decided to search more broadly for Mel1d orthologues. Hence, BLAST searches of

214   the complete set of proteins stored in NCBI for mammals (~4.6 million) and birds

215   (~2.8 million) were done using reptile Mel1d orthologues as the query. Though

216   hundreds of bird and mammal genomes are available in NCBI with protein-level

217   annotations (spanning the diversity of each lineage), the top mammal/bird hit for

218   reptile Mel1D was always MT1/MTNR1A (not shown). Considering our current

219   understanding of amniote phylogeny (*e.g.* Chiari *et al.* 2012), our data requires that

220   independent losses of Mel1d occurred in the ancestors to birds and mammals.

221

222   For all studied vertebrate species outside teleosts, we identified one copy of MT1,

223   barring spotted gar, where MT1 was not identified (Fig. 1, Fig. 2b); its trace was

224   retrieved in the genome after further analyses (see section below), representing a

225   sequence annotated as a pseudogene. Several teleost species retain two or more

226   ancestral MT1 copies (PP: 0.99, Fig. 1), which can be explained by Ts3R. These

227   duplicates have been annotated in zebrafish as "Mtnr1aa" and "Mtnr1ab" (ZFIN 2008

228   - ZNC nomenclature, cloned as "'ZMel1a1" and "ZMel1a3" by Shang & Zhdanova

229   2007). Consequently, we maintained the same 'a' and 'b' nomenclature in all species

230   according to inferences of orthology with zebrafish (Fig. 1). The two teleost-specific

231   MT1 paralogues were not present in all teleost lineages, with MT1b absent from the

232   studied acanthopterygians (tilapia and pufferfish). Salmonid-specific paralogues of

233   MT1a (MT1a1 and 1a2) were identified, ancestral to three salmonid species (PP: 1.0,

234   Fig. 2b), consistent with retention from Ss4R, though only a single copy of MT1b was

235   retained in the same three species, suggesting ancestral loss following Ss4R (Fig. 1,

236   Fig. 2b).

237

238   We identified one copy of MT2 in non-teleost vertebrate lineages, and evidence for

239   teleost-specific paralogues (Fig. 2c). Two MT2 paralogues were identified in

240   Ostariophysi members (zebrafish and Mexican cavefish) and northern pike

241   (Protacanthoptergii); however, only one MT2 copy was identified in Acanthopterygii

242   members (Nile tilapia and pufferfish) (Fig. 1, Fig. 2c). Branching patterns among

243   these duplicates were not well resolved when considering species phylogeny. An

244   ancestral teleost duplication event (*e.g.* Ts3R) predicts two paralogous MT2 teleost

245   clades, each containing teleosts branching after expected species relationships (as

246   seen for MT1a/b). However, a clade containing zebrafish "Mtnr1ba" (ZFIN 2008,

247   "ZMel1b2" in Shang & Zhdanova 2007) branched outside other fish (including the

248   non-teleost spotted gar) in both the BY and ML trees (Fig. 1 and S1). Internal to the

249   spotted gar, there were two teleost MT2 clades, one containing zebrafish "Mtnr1bb"

250   (ZFIN 2008, "ZMel1b1" in Shang & Zhdanova 2007) and other teleost lineages

251   (northern pike and Acanthopterygii members), while the other contained a separate

252   northern pike sequence and all MT2 sequences from salmonids. Given the strong

253   support for the clade containing zebrafish "Mtnr1bb" (PP:1.0, Bootstrap support:

254   100%), we considered all sequences therein to be orthologous, and named them

255   MT2b (to maintain the zebrafish "b" nomenclature) (Fig. 2c). We named the

256   remaining teleost MT2 sequences as MT2a (Fig. 2c), under the hypothesis that

257   orthology to zebrafish MT2a was obscured by a long-branch attraction artefact (note

258   the long-branch length leading to Ostariophysi members for MT2a; Fig. S1). This

259   scenario is parsimonious, as it allows for a single ancestral teleost duplication (*e.g.*

260   Ts3R) rather than several lineage-specific MT2 gains. Accordingly, we propose that

261   MT2a was lost in the ancestor to *Oreochromis* and *Takifugu*, while two salmonid

262   duplicates of MT2a (MT2a1 and 2a2) were retained from Ss4R (Fig. 1 and S1, Fig.

263   2c). No copies of MT2b were identified in salmonids, suggesting a loss in the

264   common salmonid ancestor (Fig. 2c).

265

266   As shown elsewhere (Dufourny *et al.* 2008), Mel1c and mammalian GPR50 proteins

267   grouped together in a well-supported clade (Fig. 1). A single Mel1c copy was

268   identified in all teleosts barring salmonids, which evidently lack Mel1c (Fig. 2d). This

269   is consistent with a scenario where one Mel1c paralogue was lost early following

270   Ts3R, and an additional loss occurred in the common salmonid ancestor (Fig. 2d).

271

272   **Synteny analysis supports phylogenetic assignment of vertebrate MTRs**

273   Next, to gain an independent line of evidence to support our phylogenetic

274   reconstructions, we compared the genomic regions harboring MTR-encoding genes

275   among a range of vertebrate lineages. The local gene neighborhood containing each

276   MTR family member was more or less conserved across jawed vertebrate evolution,

277   defining identifiable synteny groups specific to each ancestral MTR (Fig. 3), including

278   teleost and salmonid-specific paralogues (Fig. 4). The genomic neighborhood

279   containing the single *MTR* locus of lampreys did not conserve synteny with an

280   equivalent region containing any single MTR gene in gnathostomes. Instead, the

281   genes surrounding the single *MTR* locus of lampreys showed notable similarity to a

282 combination of genes located around the various gnathostome MTRs (Fig. 4e). This

283 lends support to an ancestral origin of MTRs in the vertebrate lineage, but does not

284 allow us to pinpoint the relationship of lamprey MTR to the four MTR family members

285 of jawed vertebrates. One possible interpretation is that the duplications generating

286 four gnathostome MTR genes occurred after the cyclostomes and gnathostomes

287 split, with the lamprey genomic neighborhood reflecting a derived representation of

288 the ancestral vertebrate state. However, the current consensus is that at least one

289 round of WGD is shared by cyclostomes and gnathostomes (*e.g.* Kuraku *et al.* 2009,

290 Stadler *et al.* 2004). In this scenario, conserved synteny between a single genomic

291 region in the former to multiple blocks in the latter may be explained by one or more

292 shared duplications followed by lineage-specific rediploidization, as proposed by

293 Robertson *et al.* 2017.

294

295 **Genetic linkage between *MTR* and *FAT* genes**

296 Tandem-linked MTR and *FAT* protocadherin gene family members are strongly

297 conserved in all vertebrates (Fig. 3, Fig. 4). Specifically, *MT1*, *Mel1d*, and *MT2* were

298 almost always in tandem with *FAT1*, *2*, and *3*, respectively (Fig. 3, Fig. 4). This

299 association was absent for *Mel1c*, in addition to MTR co-orthologues from a sea

300 squirt (Fig. 3f) and the Florida lancelet (not shown), defining this as a vertebrate-

301 specific feature. Past studies have noted genetic linkage between MTR and FAT

302 genes. For example, the FAT3-MT2 locus is involved in diabetes risk, with several

303 SNPs involved in disease located between the two genes, implying potential

304 functional links (*e.g.* Prokopenko *et al.* 2009, Dupuis *et al.* 2010). While, the reason

305 for co-evolution of these loci is yet to be determined, the tandem organization of FAT

306 and MTR genes indicates selective pressure to maintain an association that may be

307 underpinned by a conserved feature of vertebrate physiology.

309 FAT family sequences also provide an independent source of phylogenetic

310 information that may help reconstruct the evolution of the genomic regions containing

311 linked MTR genes. In an ML analysis performed with FAT proteins from

312 representative vertebrate species, we observed three clades (FAT1, 2 and 3) that

313 branched according to expected species relationships (Fig. 5). When the ML tree

314 was midpoint rooted, FAT1 (linked to MT1) and FAT2 (linked to Mel1d) were sister

315 groups (Fig. 5), consistent with the sister relationship of MT1 and Mel1d recovered by

316 the MTR phylogeny. Further, the teleost duplications observed for MTR genes were

317 clearly identifiable in the respective tandem FAT genes (Fig. 5). Finally, the well-

318 supported branching of salmonid FAT3a sequences with zebrafish FAT3a (i.e. linked

319 to the *MT2a* gene, Fig 3c) adds weight to the hypothesis that salmonid/pike MT2

320 sequences are orthologous to zebrafish *MT2a* (Fig. 2c).

321

322 **Synteny analyses support MTR losses**

323 The conservation of synteny across vertebrate taxa in genomic regions containing

324 MTR genes provides useful information on MTR genes inferred to be absent in

325 sequence databases. In this respect, we observed that the genomic regions

326 containing *Mel1d* in reptiles, frogs and fishes have matched syntenic regions in the

327 human and chicken genomes (Fig. 3d). Consequently, the regions predicted to

328 contain *Mel1d* in human and chicken have been properly assembled and are

329 otherwise well annotated, consistent with *bone-fide* genetic losses of *Mel1d* in these

330 species. The same approach allowed us to detect a pseudogene likely to be a

331 vestige of *Mel1c* in Atlantic salmon (LOC106568030) (Fig. 3d), and a gene annotated

332 as 'non-coding' bearing similarity with *MT1* (according to BLAST) at the predicted

333 MT1 locus in spotted gar (LOC107077181) (Fig. 3a). Further, a second *FAT2*

14

334 paralogue was detected in Atlantic salmon, supporting our previous conclusion of an

335 ancestral loss of one Mel1d copy following Ss4R. Similarly, a second *FAT3*

336 paralogue was detected in *Oreochromis*, non-paired with an *MTR2* gene (Fig. 3c),

337 confirming the loss of *MT2a* in this species.

338

339 **Comparative sequence analysis of Mel1d with other MTRs**

340 Having established that Mel1d is an ancestral vertebrate MTR, we sought to compare

341 the primary amino acid sequence of this molecule to other MTR family members,

342 hoping to gain clues on its function considering existing literature (Fig. 6).

343

344 We first examined the MTR transmembrane domains and ligand-binding residues,

345 which have known functional importance. The characteristic seven transmembrane

346 domain structure (TMDs) of all MTRs, critical for GPCR structure and ligand binding

347 (Baldwin 1994), were conserved in Mel1d, MT1, MT2 and Mel1c (Fig. 6). Indeed,

348 most of the residues identified as key for melatonin binding are readily identifiable in

349 the Mel1d transmembrane domains (Fig. 6), in particular TM3, 6 and 7 (Gubitz &

350 Reppert 2000, Kokkola *et al.* 2003, Mazna *et al.* 2005, 2008, Chan & Wong 2013).

351 The only notable difference in the TMDs was that several Mel1d orthologues had

352 threonine replacements at position 254, specific to this MTR. This position is

353 important for melatonin binding in MT2 (valine-291 on human MT2), which was not

354 reported for MT1 (Mazna *et al.* 2005). Outside the TMDs, two additional melatonin

355 binding residues (asparagine-102 of the conserved NRY motif and alanine-238) were

356 conserved in Mel1d (Fig. 6). Interestingly, a mutation in the second extracellular loop

357 of GPR50 linked to the loss of melatonin binding function in mammals (Clément *et al.*

358 2017) was absent in Mel1d (Fig. 6).

359

360   Other key sites conserved in Mel1d included cysteine-78 and cysteine-155,

361   responsible for a conserved disulfide bridge essential to MTR structure (Fig. 6). In

362   addition, residues important for G protein activation and trafficking of MT1 (Kokkola

363   *et al.* 2005) were all conserved in Mel1d (green arrows on Fig. 6). Putative

364   palmitoylation site in MT1 and MT2 (cysteine-314 in MT1 and cysteine- 332 in MT2,

365   Sethi *et al.* 2008) required for G-protein interaction (light blue arrow on Fig. 6) were

366   either not identified (MT2 cysteine-332) in Mel1d or absent from most species (MT1

367   cysteine-314)*.* However a proximal conserved cysteine in position 294 of Mel1d (Fig.

368   6) may fulfil a similar function. Several phosphorylation sites have been suggested in

369   the C-terminal cytoplasmic tail of MT1 and MT2, which might be important for β-

370   arrestin-dependent receptor internalization (Ebisawa *et al.* 1994, Sethi *et al.* 2008,

371   yellow arrows on Fig. 6). One of these sites is present on Mel1d, at position 288,

372   however only in coelacanth and tetrapods. None of the other phosphorylation sites

373   are present because of the shorter length of Mel1d, and this could be linked to

374   differences in phosphorylation properties.

375

376   **Residue changes distinguishing Mel1d from other MTRs**

377

378   The above analyses confirm that Mel1d has most of the canonical residues for

379   melatonin binding and MTR structure/function. We next sought to identify conserved

380   differences between Mel1d and the other MTRs, as candidates to impart functional

381   properties unique to Mel1d.

382

383   Five extracellular or intracellular positions in Mel1d show substantial differences with

384   either one or all other MTRs (Fig. 6). In most Mel1d orthologues, the first extracellular

385   loop contains lysine (positive charged) at position 38, which is typically asparagine

16

386   (neutrally charged) in the other MTRs. At position 144, which is almost always fixed

387   as glycine in MT1, MT2 and Mel1c, Mel1d orthologues retain glutamic acid or

388   aspartic acid. This replacement is presumed functionally significant, as glycine

389   provides high conformational flexibility (Betts and Russell 2003), while glutamic and

390   aspartic acid are highly negatively charged. At position 246, MTRs usually conserve

391   proline (except for the two derived GPR50 from mammals), but Mel1d shows a high

392   diversity of residues with diverse functional properties, suggesting a distinct mode of

393   selective pressure. In the same loop (position 242), a gap is observed in all Mel1d

394   sequences at an amino acid position that is variable among the other MTRs. Finally,

395   a notable difference between Mel1d and MT1 is observed in position 119, in the

396   second intracellular loop. Most MT1 sequences have aspartic acid at this position,

397   while Mel1d conserves asparagine or serine, leading to a major difference in charge.

398

399

400

401

402

403

404

405

406

407

408

409

410

411

DISCUSSION

413

414   Our unequivocal demonstration of a new ancestral vertebrate MTR forces a revision

415   of current models for the origin and diversity of MTRs, and has biological implications

416   for vertebrate lineages conserving distinct MTR gene repertoires.

417

418   It seems important to ask why Mel1d has previously been missed as a unique MTR,

419   when the gene is readily detectable in sequence databases. This is likely partly due

420   to a historic assumption that the MTR gene family structure of birds and mammals

421   (i.e. MT1, MT2 and Mel1c) is representative for all vertebrates. Mel1d has high

422   similarity with MT1, and has tended to be named 'mtnr1a-like' in genome databases.

423   In addition, previous phylogenetic studies of MTRs have been based on small

424   datasets (*e.g.* Reppert *et al.* 1995a; Mazurais *et al.* 1999; Park *et al.* 2006, 2007a,b;

425   Shang & Zhdanova 2007; Hong *et al.* 2014), with biases in the taxa sampled, and

426   could not by design distinguish Mel1d and MT1. A single past study noted a *Xenopus*

427   MTR sequence that did not group with MT1, MT2 or Mel1c and concluded the

428   existence of a novel MTR (Mel1d) (Shiu *et al.* 1996); correctly according to our

429   findings. Our study benefits from a much broader survey of vertebrate MTR

430   sequences, allowing us to conclude that Mel1d is at least 450 million years old,

431   having been present in the jawed vertebrate ancestor.

432

433   Our phylogenetic reconstruction of MTRs will help the field going forwards, as

434   researchers can be certain of which family member (including teleost-specific

435   paralogues) they are studying, allowing more reliable conclusions in comparative

436   studies of function and gene expression. We show that teleost-specific paralogues of

437   MT1 are easily distinguished from Mel1d and provide a scheme to allow researchers

18

438   to match teleost MTRs formerly named under several nomenclature systems to a

439   single phylogenetically-assigned naming system accommodating orthologues and

440   paralogues (Table 1).

441

442   **Insights into Mel1d function: reinterpreting expression data in teleosts**

443   While not being previously recognized as a unique vertebrate MTR, Mel1d has

444   already been studied in various teleosts (Table 1). These past studies demonstrate

445   that the *Mel1d* transcript is abundantly expressed in a manner like other MTR family

446   members, but showing differences that may underlie unique functions. A pattern

447   seems conserved across multiple species, where Mel1d and MT1a expression is

448   higher in brain and retina, respectively (*e.g.* Park *et al.* 2006, 2007a,b; Ikegami *et al.*

449   2009: Confente *et al.* 2010; Hong *et al.* 2014). Mel1d tends to be more strongly

450   expressed in brain regions associated with visual perception (*e.g.* Mazurais *et al.*

451   1999; Gaildrat and Falcón, 2000; Shi *et al.* 2004; Confente *et al.* 2010; Hong *et al.*

452   2014). Many peripheral tissues were reported to express Mel1d with species-specific

453   differences and in a distinct manner to other MTRs (Park *et al.* 2006,

454   2007a,b; Ikegami *et al.* 2009; Confente *et al.* 2010; Hong *et al.* 2014). Such data

455   suggests involvement of Mel1d in photoreceptive processes, along with broader

456   regulatory roles in the physiological functions of peripheral organs.

457

458   Rhythmical oscillations in the expression of Mel1d have also been reported, with

459   variations depending on species, organ and season. In zebrafish, a day/night

460   oscillation of MTR brain gene expression (peaking at night) was noted for all six MTR

461   paralogues, including Mel1d, with further expression upregulation in response to

462   melatonin administration (Shang & Zhdanova 2007). In golden rabbitfish, MT1a,

463   Mel1d and Mel1c expression was higher at night for brain and retina, with Mel1d

19

464 levels peaking at different times (Park *et al.* 2006, 2007a,b, 2014). In goldfish, Mel1d

465 was the only MTR showing rhythmical oscillations in optic tectum expression, while

466 the same was true for MT1a in retina, both peaking at the night-day transition

467 (Ikegami *et al.* 2008). In a marine pufferfish, Mel1d, MT1a and Mel1c showed

468 synchronous daily cycling of expression in the pineal gland with a nocturnal peak

469 (Ikegami *et al.* 2015). Conversely, in golden rabbitfish pineal gland, oscillations were

470 desynchronized for the same three MTRs (Park *et al.* 2006, 2007a,b). Daily

471 rhythmicity in Mel1d expression has also been observed in peripheral tissues (liver

472 and kidney) of golden rabbitfish, with higher expression during the day, opposite to

473 the brain/retina (Park *et al.* 2006, 2007b). In addition to daily variation in regulation,

474 Mel1d expression is regulated by other cycles, for example showing semilunar

475 oscillation in the diencephalon of mudskipper (Hong *et al.* 2014) and ultradiurnal

476 oscillation in a marine pufferfish, which may be circatidal (Ikegami *et al.* 2015). Mel1d

477 expression in the Senegalese sole exhibited stronger day-to-night and seasonal

478 variation than other MTR family members, with reciprocal differences recorded

479 between retina and optic tectum (Confente *et al.* 2010). Therefore, past work shows

480 that Mel1d is regulated during multiple biological cycles in teleosts, showing

481 variations distinct from other MTRs, implying functional distinctiveness.

482

483 **Functional divergence between Mel1d and MT1?**

484 High protein-level similarity between Mel1d and MT1, taken with the conservation of

485 all key residues in the MTR transmembrane domains, strongly implies that Mel1d

486 binds melatonin. Notably, residues showing conserved replacements between Mel1d

487 and MT1 are all located in extracellular or cytoplasmic loops, which is predicted to

488 impact interactions with other proteins, in particular signalling partners, rather than

489 melatonin. Strikingly, one of these sites corresponds to a documented human MT1

490   mutation studied *in vitro* (Chaste *et al.* 2010). The replacement of glycine-144 (MT1)

491   with glutamic acid or aspartic acid corresponds to a G166E mutation in human MT1,

492   associated with impaired activation of cAMP signalling, despite retention of strong

493   melatonin binding (Chaste *et al.* 2010). The elephant shark retains glutamic acid at

494   this position in both MT1 and Mel1d, suggesting this represents the ancestral state,

495   with functional divergence arising in the common ancestor to lobe and ray-finned

496   fishes. It is also intriguing to observe that Mel1d of two tetrapods have apparently

497   reverted to glycine in this position, indicating selection towards the ancestral residue.

498

499   **Why was Mel1d lost in mammals and birds?**

500   Further work will be needed to establish the extent of conservation in Mel1d function

501   and regulation across different vertebrate lineages. This should focus on reptiles and

502   amphibians, where the function of this gene has not been studied experimentally.

503   Such studies may help explain the specific biological requirements for Mel1d, and

504   reveal why the gene was lost independently in mammals and birds. It is notable that

505   mammals and birds stand out from other vertebrates when considering their

506   melatonin-dependent light detection and clock systems. Mammals have lost

507   extraocular light perception and relocated control of their biological clock away from

508   melatonin-producing pinealocytes to the suprachiasmatic nucleus (Falcón *et al.*

509   2009). Birds have both the ancestral pineal clock and melatonin production system,

510   but also independently developed a clock system in the homologue of the

511   suprachiasmatic nucleus and use retinal detection (Cassone 1991, Falcón *et al.*

512   2009). Another distinguishing feature specific to both groups is homeothermy, with

513   modulatory effects of melatonin on body temperature regulation reported in humans

514   (Cagnacci *et al.* 1992; Viswanathan *et al.* 1990) and Japanese quail (Underwood and

515   Edmonds 1995). Extrinsic temperature variation appears a less important zeitgeber

516 for the circadian clock of homeotherms relative to poikilotherms (Rensing and Ruoff,

517 2002), which are known to use melatonin to regulate behavioral thermoregulation

518 (Lutterschmidt *et al.* 2003). In addition, birds and mammals are the only vertebrates

519 that have evolved (through convergent mechanisms) stereotypical slow wave and

520 rapid eye movement sleep phases, linked to melatonin regulation in mammals (Lesku

521 *et al.* 2011). Such changes in the physiological role of melatonin and consequent re-

522 organization of melatonin response pathways, may have been the ultimate driver for

523 Mel1d redundancy and gene loss through relaxation of purifying selection.

524

525 Another melatonin-associated function that is present in vertebrate lineages retaining

526 Mel1d (in addition to lamprey), but lost in both mammals and birds, is the negative

527 regulation of pigmentation development in the dark, known as the "body-blanching

528 response" (Hamasaki and Eder 1977, Norris and Carr 2013). In fishes, melatonin is

529 thought to regulate chromatosome aggregation in different kinds of chromatophores

530 (Fujii 2000); Mel1d is expressed in the skin of mudskipper (together with MT1 - Hong

531 *et al.* 2014), the goldfish (together with MT2 and Mel1c - Ikegami *et al.* 2008) and the

532 sole (together with MT2 - Confente *et al.* 2010). In addition, in sole skin, Mel1d is the

533 only MTR to be up-regulated at night. It is therefore possible that Mel1d is involved in

534 skin physiology and pigment regulation in fish chromatophores.

535

536 **Expansion of the MTR repertoire of teleosts**

537 Contrary to mammals/birds, there has been a trend towards evolutionary expansion

538 in the MTR repertoire of teleosts, as observed in many gene families with paralogues

539 retained from Ts3R (Glasauer and Neuhauss, 2014) and Ss4R (Houston and

540 Macqueen, 2019). Interestingly, not all MTR family members were affected equally.

541 While we identified multiple paralogous copies of MT1 and MT2 - presumed to have

542 been retained from Ts3R and Ss4R - Mel1c and Mel1d were always single copy,

543 requiring repeated losses of paralogues generated during gene duplication or WGD

544 events. This is compatible with a hypothesis where the functions or expression-level

545 regulation of MT1 and MT2 can be divided among paralogous copies, following the

546 well-established subfunctionalization model, or potentially reflects fixation of new

547 adaptive functions among MT1/MT2 paralogues (Stoltzfus 1999 and Force *et al.*

548 1999). In this respect, we observed several amino acid substitutions between MT1a

549 vs. MT1b and MT2a vs. MT2b (Fig. 6), consistent with protein-level functional

550 divergence. Conversely, selection has operated in a distinct manner for Mel1c and

551 Mel1d, with any duplicates generated being quickly purged by selection for reasons

552 that remain to be established, but potentially linked to dosage constraints, or a

553 mechanism of regulation that cannot be divided across distinct loci.

554

555   CONCLUSIONS

556

557 Mel1d is one of four ancestral vertebrate MTRs that shows a wide phylogenetic

558 distribution but has been lost in mammals and birds. Compared to MT1, MT2 and

559 Mel1c, Mel1d has many conserved, but also divergent characteristics, both in terms

560 of protein sequence and spatio-temporal expression patterns of relevance to

561 chronobiological traits. Additional work is needed to characterize the functional

562 distinctiveness of Mel1d compared to other MTRs and to explain why unique MTR

563 repertoires have been conserved in different vertebrate lineages.

564

REFERENCES

576

577

578 Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman,

579 D.J., 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database

580 search programs. Nucleic Acids Res. 25(17): 3389-3402.

581

582 Baldwin, J.M., 1994 Structure and function of receptors coupled to G proteins. Curr.

583 Opin. Cell Biol. 6: 180-190.

584

585 Betts, M.J., Russell, R.B., 2003 Amino acid properties and consequences of

586 subsitutions, in *Bioinformatics for Geneticists*, M.R. Barnes, I.C. Gray eds, Wiley.

587

588 Cagnacci, A., Elliott, J.A., Yen, S.S., 1992 Melatonin: a major regulator of the

589 circadian rhythm of core temperature in humans. J Clin Endocrinol Metab. 75: 447-

590 452.

591

592 Cassone, V.M., 1991 Melatonin and suprachiasmatic nucleus function, pp. 309–323

593 in *Suprachiasmatic Nucleus: The Mind's Clock*, D.C. Klein, R.Y. Moore, S.M.

594 Reppert, Oxford Univ. Press, New York.

595

596 Chan, K.H., Wong, Y.H., 2013 A molecular and chemical perspective in defining MTR

597 subtype selectivity. Int J Mol Sci. 14: 18385-18406.

598

599 Chai, K., Liu, X., Zhang, Y., Lin. H., 2013 Day-night and reproductive cycle profiles of

600 melatonin receptor, kiss, and gnrh expression in orange-spotted grouper

601 (*Epinephelus coioides*). Mol Reprod Dev. 80: 535-548.

602

603 Chaste, P., Clément, N., Mercati, O., Guillaume, J.L., Delorme, R., Botros, H.G., *et*

604 *al.*, 2010 Identification of pathway-biased and deleterious MTR mutants in autism

605 spectrum disorders and in the general population. PLoS One. 5: e11495.

606

607 Chiari, Y., Cahais, V., Galtier, N., Delsuc, F., 2012 Phylogenomic analyses support

608 the position of turtles as the sister group of birds and crocodiles (Archosauria). BMC

609 Biol. 10: 65.

610

611 Choi, Y.J., Habibi, H.R., Choi, C.Y., 2016 Profiles of gonadotropin-inhibitory hormone

612 and melatonin during the sex change and maturation of cinnamon clownfish,

613 *Amphiprion melanopus*. Biochem Biophys Res Commun. 475: 189-93.

614

615 Clément, N., Renault, N., Guillaume, J.L., Cecon, E., Journé, A.S., Laurent, X., *et al.*

616 2018 Importance of the second extracellular loop for melatonin MT1 receptor function

617 and absence of melatonin binding in GPR50. Br J Pharmacol. 175: 3281-3297.

618

619 Confente, F., Rendón, M., Besseau, L., Falcón, J., Muñoz-Cueto, J.A., 2010 MTRs in

620 a pleuronectiform species, *Solea senegalensis*: Cloning, tissue expression, day-night

621 and seasonal variations. Gen Comp Endocrinol. 167: 202-214.

622

623 Dehal, P., Boore, J.L., 2005 Two rounds of whole genome duplication in the

624 ancestral vertebrate. PLoS Biol. 3: e314.

625

626 Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006 Relaxed

627 phylogenetics and dating with confidence. PLoS Biol. 4: e88.

628

629 Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012 Bayesian phylogenetics

630 with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29: 1969-1973.

631

632 Dubocovich, M.L., Markowska, M., 2005 Functional MT1 and MT2 MTRs in

633 Mammals. Endocrine 27: 101-110.

634

635 Dufourny, L., Levasseur, A., Migaud, M., Callebaut, I., Pontarotti, P., Malpaux, B.,

636 Monget, P., 2008 GPR50 is the mammalian ortholog of Mel1c: Evidence of rapid

637 evolution in mammals. BMC Evol. Biol. 8: 105.

638

639 Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A.U.

640 *et al.*, 2010 New genetic loci implicated in fasting glucose homeostasis and their

641 impact on type 2 diabetes risk. Nat Genet. 42: 105-116.

642

643 Ebisawa, T., Karne, S., Lerner, M.R., Reppert, S.M., 1994 Expression cloning of a

644 high-affinity MTR from *Xenopus* dermal melanophores. Proc Natl Acad Sci USA. 91:

645 6133-6137.

646

647 Falcón, J., Besseau, L., Fuentès, M., Sauzet, S., Magnanou, E., Boeuf, G., 2009

648 Structural and functional evolution of the pineal melatonin system in vertebrates. Ann

649 N Y Acad Sci. 1163: 101-111.

650

651 Feng, N.Y., Bass, A.H., 2016 "Singing" Fish Rely on Circadian Rhythm and

652 Melatonin for the Timing of Nocturnal Courtship Vocalization. Curr Biol. 26: 2681-

653 2689.

654

655 Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., Postlethwait, J., 1999
656 Preservation of duplicate genes by complementary, degenerative mutations.
657 Genetics. 151: 1531-1545.

658

659 Fujii, R., 2000 The regulation of motile activity in fish chromatophores. Pigment Cell
660 Res. 13: 300-319.

661

662 Gaildrat, P., Falcón, J., 2000 Melatonin receptors in the pituitary of a teleost fish:
663 mRNA expression, 2-[(125)I]iodomelatonin binding and cyclic AMP response.
664 Neuroendocrinology. 72: 57-66.

665

666 Gauer, F., Masson-Pévet, M., Stehle, J., Pevet, P., 1994 Daily variations in MTR
667 density of rat pars tuberalis and suprachiasmatic nuclei are distinctly regulated. Brain
668 Res 641: 92-98.

669

670 Gernhard, T., 2008 The conditioned reconstructed process. J Theor Biol. 253: 769-
671 778.

672

673 Glasauer, S.M., Neuhauss, S.C., 2014 Whole-genome duplication in teleost fishes
674 and its evolutionary consequences. Mol Genet Genomics. 289: 1045-1060.

675

676 Gubitz, A.K., Reppert, S.M., 2000 Chimeric and point-mutated receptors reveal that a
677 single glycine residue in transmembrane domain 6 is critical for high affinity
678 melatonin binding. Endocrinology 141: 1236-1244.

679

680 Hardeland, R., Cardinali, D.P., Srinivasan, V., Spence, D.W., Brown, G.M., Pandi-

681 Perumal, S.R., 2011 Melatonin--a pleiotropic, orchestrating regulator molecule. Prog

682 Neurobiol. 93: 350-384.

683

684 Hamasaki, D.I., Eder, D.J., 1977 Adaptive Radiation of the Pineal System. pp.

685 7(5)497–548 in *Handbook of sensory physiology - The Visual System in Vertebrates*

686 (F. Crescitelli, ed.), Springer-Verlag, New York.

687

688 Herrera-Pérez, P., Del Carmen Rendón, M., Besseau, L., Sauzet, S., Falcón, J.,

689 Muñoz-Cueto, J.A., 2010 MTRs in the brain of the European sea bass: An *in situ*

690 hybridization and autoradiographic study. J Comp Neurol. 518: 3495-3511.

691

692 Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018

693 UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35: 518-

694 522.

695

696 Hong, L.Y., Hong, W.S., Zhu, W.B., Shi, Q., You, X.X., Chen, S.X., 2014 Cloning and

697 expression of MTRs in the mudskipper *Boleophthalmus pectinirostris*: their role in

698 synchronizing its semilunar spawning rhythm. Gen Comp Endocrinol. 195: 138-150.

699

700 Houston, R.D., Macqueen, D.J., 2019 Atlantic salmon (*Salmo salar L.*) genetics in the

701 21st century: taking leaps forward in aquaculture and biological understanding. Anim

702 Genet. 50: 3-14.

703

704 Ikegami, T., Azuma, K., Nakamura, M., Suzuki, N., Hattori, A., Ando, H., 2009 Diurnal

705 expressions of four subtypes of MTR genes in the optic tectum and retina of goldfish.

706 Comp Biochem Physiol A Mol Integr Physiol. 152: 219-224.

707

708 Ikegami, T., Maruyama, Y., Doi, H., Hattori, A., Ando, H., 2015 Ultradian oscillation in

709 expression of four MTR subtype genes in the pineal gland of the grass puffer, a

710 semilunar-synchronized spawner, under constant darkness. Front Neurosci. 9: 9.

711

712 Jin, Y.H., Park, J.W., Kim, J.H., Kwon, J.Y., 2013 The Expression Pattern of MTR 1a

713 Gene during Early Life Stages in the Nile tilapia (*Oreochromis niloticus*). Dev Reprod.

714 17: 45-53.

715

716 Jockers, R., Maurice, P., Boutin, J.A., Delagrange, P., 2008 MTRs,

717 heterodimerization, signal transduction and binding sites: what's new? Br J

718 Pharmacol. 154: 1182-1195.

719

720 Jones, D.T., Taylor, W.R., Thornton, J.M., 1992 The rapid generation of mutation

721 data matrices from protein sequences. Bioinformatics 8: 275-282.

722

723 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S.,

724 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat

725 Methods. 14: 587-589.

726

727 Kamesh, N., Aradhyam, G.K., Manoj, N., 2008 The repertoire of G protein-coupled

728 receptors in the sea squirt *Ciona intestinalis*. BMC Evol. Biol. 8: 129.

729

730   Katoh, K., Standley, D.M., 2013 MAFFT multiple sequence alignment software
731   version 7: improvements in performance and usability. Mol Biol Evol. 30: 772-780.

732

733   Kokkola, T., Foord, S.M., Watson, M.A., Vakkuri, O., Laitinen, J.T., 2003 Important
734   amino acids for the function of the human MT1 MTR. Biochem Pharmacol. 65: 1463-
735   1471.

736

737   Kokkola, T., Salo, O.M., Poso, A., Laitinen, J.T., 2005 The functional role of cysteines
738   adjacent to the NRY motif of the human MT1 MTR. J Pineal Res. 39: 1-11.

739

740   Krishnan, A., Almén, M.S., Fredriksson, R., Schiöth, H.B., 2013 Remarkable
741   similarities between the hemichordate (*Saccoglossus kowalevskii*) and vertebrate
742   GPCR repertoire. Gene. 526:122-133.

743

744   Kuraku, S., 2013 Impact of asymmetric gene repertoire between cyclostomes and
745   gnathostomes. Semin. Cell Dev. Biol. 24: 119-127.

746

747   Kuraku, S., Meyer, A., Kuratani, S., 2009 Timing of Genome Duplications Relative to
748   the Origin of the Vertebrates: Did Cyclostomes Diverge before or after? Mol Biol Evol
749   26: 47-59.

750

751   Lesku, J.A., Vyssotski, A.L., Martinez-Gonzalez, D., Wilzeck, C., Rattenborg, N.C.,
752   2011 Local sleep homeostasis in the avian brain: convergence of sleep function in
753   mammals and birds? Proc. R. Soc. B 278: 2419-2428.

754

755 Levoye, A., Dam, J., Ayoub, M.A., Guillaume, J.L., Couturier, C., Delagrange, P.,

756 Jockers, R., 2006 The orphan GPR50 receptor specifically inhibits MT1 MTR function

757 through heterodimerization. Embo J. 25: 3012-3023.

758

759 Lien, S., Koop, B.F., Sandve, S.R., Miller, J.R., Kent, M.P., Nome, T., *et al.*, 2016

760 The Atlantic salmon genome provides insights into rediploidization. Nature. 533: 200-

761 205.

762

763 Lutterschmidt, D.I., Lutterschmidt, W.I., Hutchison, V.H., 2003 Melatonin and

764 thermoregulation in ectothermic vertebrates: a review. Can. J. Zool. 81: 1-13.

765

766 Macqueen, D.J., Johnston, I.A., 2014 A well-constrained estimate for the timing of

767 the salmonid whole genome duplication reveals major decoupling from species

768 diversification. Proc Biol Sci. 281: 20132881.

769

770 Macqueen, D.J., Wilcox, A.H., 2014 Characterization of the definitive classical

771 calpain family of vertebrates using phylogenetic, evolutionary and expression

772 analyses. Open Biol. 4: 130219.

773

774 Mazna, P., Grycova, L., Balik, A., Zemkova, H., Friedlova, E., Obsilova, V., *et al.*,

775 2008 The role of proline residues in the structure and function of human MT2 MTR. J

776 Pineal Res. 45: 361-372.

777

778 Mazna, P., Berka, K., Jelinkova, I., Balik, A., Svoboda, P., Obsilova, V., *et al.*, 2005

779 Ligand binding to the human MT2 MTR: the role of residues in transmembrane

780 domains 3, 6, and 7. Biochem Biophys Res Commun. 332: 726-734.

781

782 Mazurais, D., Brierley, I., Anglade, I., Drew, J., Randall, C., Bromage, N., *et al.,* 1999

783 Central MTRs in the rainbow trout: comparative distribution of ligand binding and

784 gene expression. J. Comp. Neurol. 409: 313-324.

785

786 Norris, D.O., Carr, J.A., 2013 *Vertebrate endocrinology*. Academic Press.

787

788 Nordstrom, K.J., Fredriksson, R., Schioth, H.B., 2008 The amphioxus *Branchiostoma*

789 *floridae* genome contains a highly diversified set of G protein-coupled receptors.

790 BMC Evol. Biol. 8: 9.

791

792 Nguyen, L-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015 IQ-TREE: A fast

793 and effective stochastic algorithm for estimating maximum likelihood

794 phylogenies. Mol. Biol. Evol. 32: 268-274.

795

796 Nguyen, N.T.T., Vincens, P., Roest Crollius, H., Louis, A., 2018 Genomicus 2018:

797 karyotype evolutionary trees and on-the-fly synteny computing. Nucleic Acids Res.

798 46: D816-D822.

799

800 Ota, K.G., Fujimoto, S., Oisi, Y., Kuratani, S., 2011 Identification of vertebra-like

801 elements and their possible differentiation from sclerotomes in the hagfish. Nat

802 Commun. 2: 373.

803

804 Park, Y.J., Park, J.G., Takeuchi, Y., Hur, S.P., Lee, Y.D., Kim, S.J., Takemura, A.,

805 2014 Influence of moonlight on mRNA expression patterns of MTR subtypes in the

806 pineal organ of a tropical fish. Mar Genomics. 14: 67-70.

807

808 Park, Y.J., Park, J.G., Jeong, H.B., Takeuchi, Y., Kim, S.J., Lee, Y.D., Takemura, A.,

809 2007(a) Expression of the MTR Mel(1c) in neural tissues of the reef fish *Siganus*

810 *guttatus.* Comp Biochem Physiol A Mol Integr Physiol. 147: 103-111.

811

812 Park, Y.J., Park, J.G., Hiyakawa, N., Lee, Y.D., Kim, S.J., Takemura, A., 2007(b)

813 Diurnal and circadian regulation of a MTR, MT1, in the golden rabbitfish, *Siganus*

814 *guttatus.* Gen Comp Endocrinol. 150: 253-262.

815

816 Park, Y.J., Park, J.G., Kim, S.J., Lee, Y.D., Saydur Rahman, M., Takemura, A., 2006

817 MTR of a reef fish with lunar-related rhythmicity: cloning and daily variations. J Pineal

818 Res. 41: 166-174.

819

820 Prokopenko, I., Langenberg, C., Florez, J.C., Saxena, R., Soranzo, N., Thorleifsson,

821 G., *et al.*, 2009 Variants in MTNR1B influence fasting glucose levels. Nat. Genet. 41:

822 77-81.

823

824 Redmond, A.K., Macqueen, D.J., Dooley, H., 2018 Phylotranscriptomics suggests

825 the jawed vertebrate ancestor could generate diverse helper and regulatory T cell

826 subsets. BMC Evol Biol. 18: 169.

827

828 Rensing, L., Ruoff, P., 2002 Temperature effect on entrainment, phase shifting, and

829 amplitude of circadian clocks and its molecular bases. Chronobiol Int. 19: 807-864.

830

831 Reppert, S.M., Weaver, D.R., Ebisawa, T., 1994 Cloning and characterization of a

832 mammalian MTR that mediates reproductive and circadian responses. Neuron 13:

833 1177-1185.

834

835 Reppert, S.M., Godson, C., Mahle, C.D., Weaver, D.R., Slaugenhaupt, S.A., Gusella,

836 J.F., 1995(a) Molecular characterization of a second MTR expressed in human retina

837 and brain: the Mel1b-MTR. Proc Natl Acad Sci USA 92: 8734-8738.

838

839 Reppert, S.M., Weaver, D.R., Cassone, V.M., Godson, C., Kolakowski, L.F. Jr., 1995

840 (b) Melatonin receptors are for the birds: molecular analysis of two receptor subtypes

841 differentially expressed in chick brain. Neuron 15: 1003-1015.

842

843 Robertson, F.M., Gundappa, M.K., Grammes, F., Hvidsten, T.R., Redmond, A.K.,

844 Lien, S., *et al.*, 2017 Lineage-specific rediploidization is a mechanism to explain time-

845 lags between genome duplication and evolutionary diversification. Genome Biol. 18:

846 111.

847

848 Sauzet, S., Besseau, L., Herrera Perez, P., Covès, D., Chatain, B., Peyric, E., *et al.,*

849 *2008* Cloning and retinal expression of MTRs in the European sea bass,

850 *Dicentrarchus labrax*. Gen Comp Endocrinol. 157: 186-195.

851

852 Sethi, S., Adams, W., Pollock, J., Witt-Enderby, P.A., 2008, C-terminal domains

853 within human MT1 and MT2 MTRs are involved in internalization processes. J Pineal

854 Res. 45: 212-218.

855

856   Shang, E.H., Zhdanova, I.V., 2007 The circadian system is a target and modulator of

857   prenatal cocaine effects. PLoS One. 2: e587.

858

859   Shi, Q., Ando, H., Coon, S.L., Sato, S., Ban, M., Urano, A., 2004 Embryonic and

860   post-embryonic expression of arylalkylamine N-acetyltransferase and MTR genes in

861   the eye and brain of chum salmon (*Oncorhynchus keta*). Gen Comp Endocrinol. 136:

862   311-321.

863

864   Shiu, S.Y., Ng, N., Pang, S.F., 1996 A molecular perspective of the genetic

865   relationships of G-protein coupled MTR subtypes. J Pineal Res. 20: 198-204.

866

867   Slominski, R.M., Reiter, R.J., Schlabritz-Loutsevitch, N., Ostrom, R.S., Slominski,

868   A.T., 2012 Melatonin membrane receptors in peripheral tissues: distribution and

869   functions. Mol Cell Endocrinol. 351: 152-166.

870

871   Solovyev, V., Kosarev, P., Seledsov, I., Vorobyev, D., 2006 Automatic annotation of

872   eukaryotic genes, pseudogenes and promoters. Genome Biol. 7: 10.1-10.12.

873

874   Stadler, P.F., Fried, C., Prohaska, S.J., Bailey, W.J., Misof, B.Y., Ruddle, F.H.,

875   Wagner, G.P., 2004 Evidence for independent Hox gene duplications in the hagfish

876   lineage: a PCR-based gene inventory of *Eptatretus stoutii*. Mol Phylogenet Evol.

877   32:686-694.

878

879   Stoltzfus, A., 1999 On the possibility of constructive neutral evolution. J Mol Evol. 49:

880   169-181.

881

882 Sugahara, F., Murakami, Y., Adachi, N., Kuratani, S., 2013 Evolution of the

883 regionalization and patterning of the vertebrate telencephalon: what can we learn

884 from cyclostomes? Curr. Opin. Genet. Dev. 23: 475-483.

885

886 Talavera, G., Castresana, J., 2007 Improvement of phylogenies after removing

887 divergent and ambiguously aligned blocks from protein sequence alignments. Syst

888 Biol. 56: 564-577.

889

890 Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., Minh, B.Q., 2016 W-IQ-TREE: a

891 fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44:

892 W232-235.

893

894 Underwood, H., Edmonds, K., 1995 The circadian rhythm of thermoregulation in

895 Japanese quail: III. Effects of melatonin administration. J Biol Rhythms. 10: 284-298.

896

897 Vernadakis, A.J., Bemis, W.E., Bittman, E.L., 1998 Localization and partial

898 characterization of MTRs in amphioxus, hagfish, lamprey, and skate. Gen Comp

899 Endocrinol. 110: 67-78.

900

901 Vilella, A.J., Severin, J., Ureta-Vidal, A., Heng, L., Durbin, R., Birney, E., 2009

902 EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in

903 vertebrates. Genome Res. 19: 327-335.

904

905 Viswanathan, M., Laitinen, J.T., Saavedra, J.M., 1990 Expression of MTRs in arteries

906 involved in thermoregulation. Proc Natl Acad Sci USA 87: 6200-6203.

907

908 Weaver, D.R., Reppert, S.M., 1990 Melatonin receptors are present in the ferret pars

909 tuberalis and pars distalis, but not in brain. Endocrinology. 127: 2607-2609.

910

911 Witt-Enderby, P.A., Bennett, J., Jarzynka, M.J., Firestine, S., Melan, M.A., 2003

912 MTRs and their regulation: biochemical and structural mechanisms. Life Sciences 72:

913 2183-2198.

914

915 ZFIN, https://zfin.org

FIGURE LEGENDS

917

918   Fig. 1. Bayesian phylogenetic tree of MTR family evolution in jawed vertebrates. The

919   analysis was done using BEAST with a high-confidence alignment of eighty MTRs

920   (300 amino acid positions; Additional Dataset 1), an uncorrelated relaxed molecular

921   clock model and the best-fitting amino acid substitution model (JTT+F+I+G4).

922   Numbers on branches are posterior probability support. Three WGD events in

923   vertebrate evolution are shown (2R - ancestral to vertebrates; Ts3R - ancestral to

924   teleosts; Ss4R - ancestral to salmonids). A ML tree was performed using the same

925   data and is provided in Fig. S1.

926

927   Fig. 2. Proposed evolutionary history of each MTR family member, considering (a)

928   Mel1d, (b) MT1, (c) MT2m, and (d) Mel1c. Species inferred to have lost all copies of

929   a MTR gene are highlighted in dark red. Teleost species inferred to have lost

930   paralogues of MTR genes arising from the Ts3R and Ss4R events are highlighted in

931   light red.

932

933   Fig. 3. Conserved synteny between the genomic neighbourhood containing MTR

934   orthologues of different lineages, shown for (a) jawed vertebrate MT1, (b) jawed

935   vertebrate Mel1d, (c) jawed vertebrate MT2, (d) jawed vertebrate Mel1c, (e)

936   comparing MTR from two lamprey species with jawed vertebrates, and (f) comparing

937   a urochordate with vertebrates.

938

939   Fig. 4. Conserved synteny between the genomic neighbourhood containing MTR

940   paralogues retained from Ts3R and Ss4R, shown for (a) MT1a, (b) MT1b, (c) Mel1d,

941   (d) MT2a, (e) MT2b, and (f) Mel1c.

942

943   Fig. 5. ML phylogenetic analysis of FAT atypical protocadherins in jawed vertebrates.

944   The analysis was done using IQ-TREE with a high-confidence alignment of thirty-five

945   FAT proteins (2,540 amino acid positions; Additional Dataset 2) and the best-fitting

946   amino acid substitution model (JTT+G+I). Numbers on branches are bootstrap

947   support values. Other details are as in the Fig. 1 legend.

948

949   Fig. 6. Alignment used to compare amino acid positions among vertebrate MTR

950   proteins (matching to the alignment used for phylogenetic analysis; Additional

951   Dataset S1). Species abbreviations: Ac = *Anolis carolensis* (green anole lizard); Am

952   = *Astyanax mexicanus* (Mexican cavefish); Bt = *Bos taurus* (cattle); Cm =

953   *Callorhinchus milli* (elephant shark); Dr = *Danio rerio* (zebrafish); El = *Esox lucius*

954   (northern pike); Gg = *Gallus gallus* (chicken); Hs = *Homo sapiens* (human); Lc =

955   *Latimeria chalumnae* (coelacanth)*;* Lo = *Lepisosteus oculatus* (spotted gar); Oa =

956   *Ornithorhynchus anatinus* (platypus); On = *Oreochromis niloticus* (Nile tilapia); Ps =

957   *Pelodiscus sinensis* (Chinese softshell turtle); Tr = *Takifugu rubripes* (tiger pufferfish);

958   Xt = *Xenopus tropicalis* (western clawed frog). Detailed annotation of sequences

959   flagged up in the main text are provided within the figure.

960

961   Table 1. Phylogenetic assignment of teleost MTRs to a standardized nomenclature

962   system.

963

964

Mel1d

Ss4R ▸ 1
1
0.93
Ts3R ▸ 1
1
0.63
0.53
1
0.93
1

Oncorhynchus mykiss Mel1d
Oncorhynchus kisutch Mel1d
Salmo salar Mel1d
Esox lucius Mel1d
Oreochromis niloticus Mel1d
Takifugu rubripes Mel1d
Danio rerio Mel1d
Astyanax mexicanus Mel1d
Lepisosteus oculatus (holostei) Mel1d
Pelodiscus sinensis Mel1d
Anolis carolinensis Mel1d
Xenopus tropicalis Mel1d
Latimeria chalumnae (sarcopterygii) Mel1d
Callorhinchus milli (chondrichtyes) Mel1d

Teleosts
Tetrapods

MT1/MTNR1A

Ss4R ▸ 0.99
1
0.49 0.48
0.52
0.57
1
0.8
Ts3R ▸ 0.99 0.34
1
1
1
0.97
1
0.92
0.59
0.53
0.83
0.7
1
0.95

Oncorhynchus mykiss MT1a1
Salmo salar MT1a1
Oncorhynchus kisutch MT1a1
Oncorhynchus kisutch MT1a2
Oncorhynchus mykiss MT1a2
Salmo salar MT1a2
Astyanax mexicanus MT1a
Danio rerio MT1a
Esox lucius MT1a
Takifugu rubripes MT1a
Oreochromis niloticus MT1a
Oncorhynchus kisutch MT1b
Oncorhynchus mykiss MT1b
Salmo salar MT1b
Esox lucius MT1b
Danio rerio MT1b
Astyanax mexicanus MT1b
Lepisosteus oculatus (holostei) MT1
Callorhinchus milli (chondrichtyes) MT1
Latimeria chalumnae (sarcopterygii) MT1
Xenopus tropicalis MT1
Anolis carolinensis MT1
Pelodiscus sinensis MT1
Gallus gallus MT1
Bos taurus MT1
Homo sapiens MT1
Ornithorhynchus anatinus MT1

Teleosts
Tetrapods

MT2/MTNR1B

0.5
1
1 0.63
1
0.74
Ss4R ▸
Ts3R ▸ 0.9
1
1
0.86
1
0.52
1
0.91
0.99
0.77
0.51
0.99
1
1
0.8

Oncorhynchus kisutch MT2a2
Oncorhynchus mykiss MT2a2
Salmo salar MT2a2
Salmo salar MT2a1
Oncorhynchus mykiss MT2a1
Oncorhynchus kisutch MT2a1
Esox lucius MT2a
Oreochromis niloticus MT2b
Takifugu rubripes MT2b
Esox lucius MT2b
Danio rerio MT2b
Astyanax mexicanus MT2b
Lepisosteus oculatus (holostei) MT2
Danio rerio MT2a
Astyanax mexicanus MT2a
Callorhinchus milli (chondrichtyes) MT2
Gallus gallus MT2
Anolis carolinensis MT2
Pelodiscus sinensis MT2
Xenopus tropicalis MT2
Homo sapiens MT2
Bos taurus MT2
Ornithorhynchus anatinus MT2
Latimeria chalumnae (sarcopterygii) MT2

Teleosts
Tetrapods

2R ▸
0.87
0.53

Mel1c

Ts3R ▸ 0.73
1
0.99
1
1
0.54
1
0.93
0.77
0.88
0.94
0.89
0.96

Oreochromis niloticus Mel1c
Esox lucius Mel1c
Takifugu rubripes Mel1c
Danio rerio Mel1c
Astyanax mexicanus Mel1c
Lepisosteus oculatus (holostei) Mel1c
Anolis carolinensis Mel1c
Gallus gallus Mel1c
Pelodiscus sinensis Mel1c
Homo sapiens Mel1c
Bos taurus Mel1c
Ornithorhynchus anatinus Mel1c
Xenopus tropicalis Mel1c
Latimeria chalumnae (sarcopterygii) Mel1c
Callorhinchus milli (chondrichtyes) Mel1c

Teleosts
Tetrapods

0.06

a. Mel1d

Number of copies

Callorhinchus milli — 1
Latimeria chalumnae — 1
Xenopus tropicalis — 1
Ornithorhynchus anatinus — 0 — total loss in mammals
Bos taurus — 0
Homo sapiens — 0
Pelodiscus sinensis — 1
Anolis carolinensis — 1
Gallus gallus — 0 — total loss in birds
Lepisosteus oculatus — 1
Danio rerio — 1 — loss of one Mel1d duplicate
Astyanax mexicanus — 1
Oreochromis niloticus — 1
Takifugu rubripes — 1
Esox lucius — 1 — loss of 1 copy of the retained duplicate
Salmo salar — 1

Ts3R
Ss4R

b. MT1 (MTNR1A)

Number of copies

Callorhinchus milli — 1
Latimeria chalumnae — 1
Xenopus tropicalis — 1
Ornithorhynchus anatinus — 1
Bos taurus — 1
Homo sapiens — 1
Pelodiscus sinensis — 1
Anolis carolinensis — 1
Gallus gallus — 1
Lepisosteus oculatus — 0 — total loss in Holostei
Danio rerio — 2
Astyanax mexicanus — 2
Oreochromis niloticus — 1 — loss of MT1b
Takifugu rubripes — 1
Esox lucius — 2 — loss of one MT1b duplicate
Salmo salar — 3

Ts3R
Ss4R

c. MT2 (MTNR1B)

Number of copies

Callorhinchus milli — 1
Latimeria chalumnae — 1
Xenopus tropicalis — 1
Ornithorhynchus anatinus — 1
Bos taurus — 1
Homo sapiens — 1
Pelodiscus sinensis — 1
Anolis carolinensis — 1
Gallus gallus — 1
Lepisosteus oculatus — 1
Danio rerio — 2
Astyanax mexicanus — 2
Oreochromis niloticus — 1 — loss of MT2a
Takifugu rubripes — 1
Esox lucius — 2 — loss of all MT2b copies
Salmo salar — 2

Ts3R
Ss4R

d. Mel1c

Number of copies

Callorhinchus milli — 1
Latimeria chalumnae — 1
Xenopus tropicalis — 1
Ornithorhynchus anatinus — 1
Bos taurus — 1
Homo sapiens — 1
Pelodiscus sinensis — 1
Anolis carolinensis — 1
Gallus gallus — 1
Lepisosteus oculatus — 1
Danio rerio — 1 — loss of one Mel1c duplicate
Astyanax mexicanus — 1
Oreochromis niloticus — 1
Takifugu rubripes — 1
Esox lucius — 1
Salmo salar — 0 — total loss in salmonids

Ts3R
Ss4R

▲ Whole genome duplication   ❋ Reduction in copy number   ✳ Loss of all copies
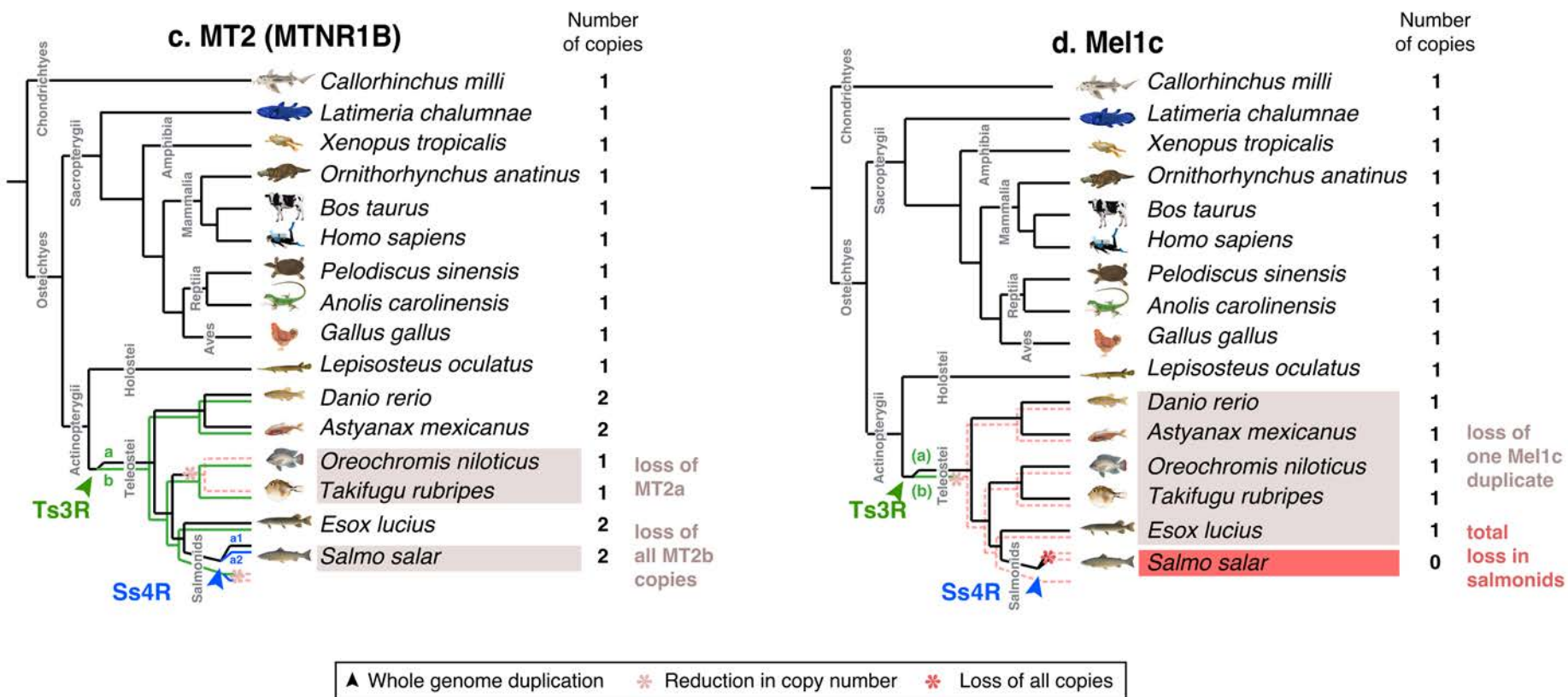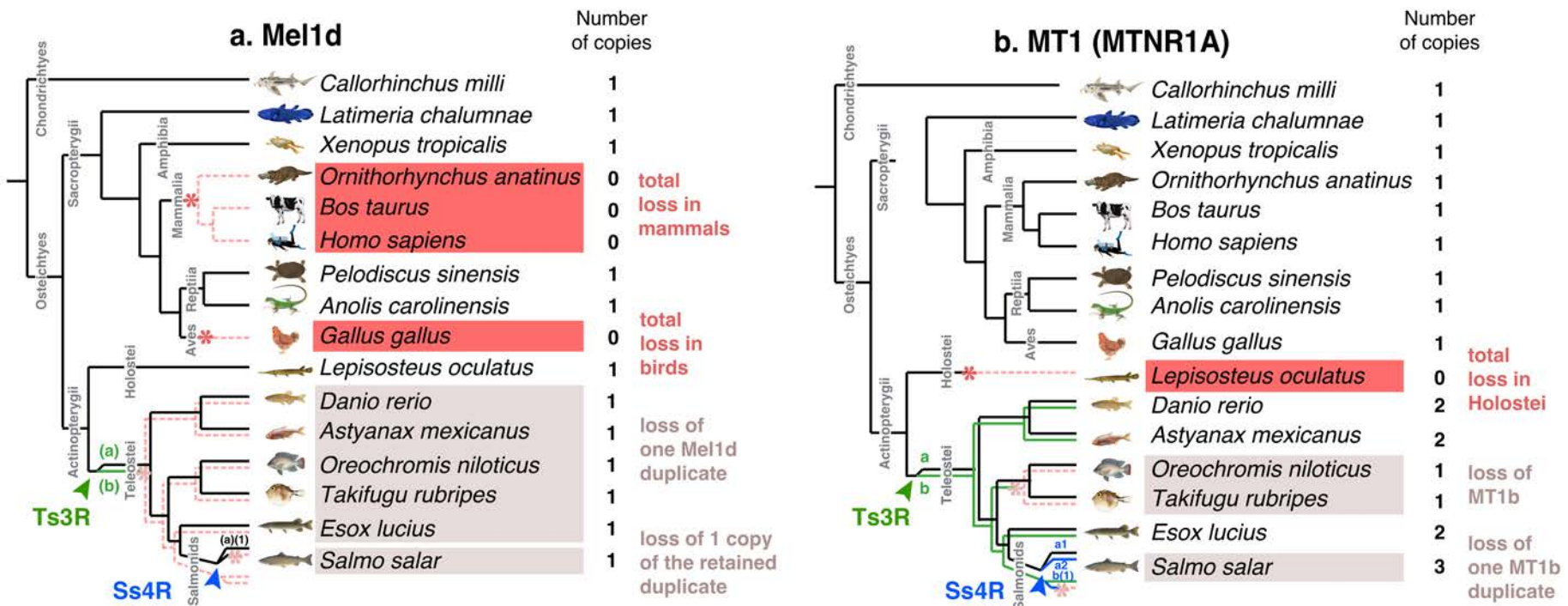
a. MT1

b. Mel1d

c. MT2

d. Mel1c

e. MTR

f. MTR-like

Table 1. Phylogenetic assignment of teleost MTRs to a standardized nomenclature system.

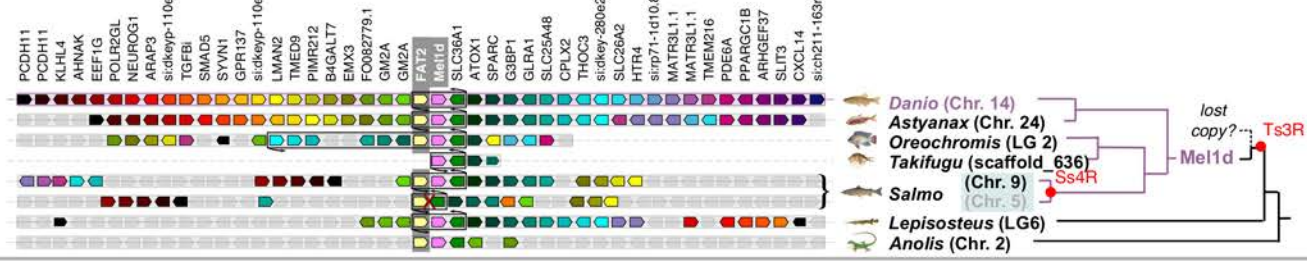| Species		Receptors attributed from literature vs orthology group assignment from this study						References
		MT1		MT2		Mel1c	Mel1d	
		MT1a	MT1b	MT2a	MT2b			
Dr	*Danio rerio* (zebrafish)	**Z1.7** (U31822.1)			**Z2.6** (U31824.1)		**Z1.4** (U31823.1)	**Reppert *et al.* 1995(b)**
		zMel1a1, Z1.7-4, mtnr1aa (NM_131393.1)	**zMel1a3** (XM_6889 89.6)	**zMel1b2, Z6.2, Mel1b-19, mtnr1ba** (NM_131395.1)	**zMel1b1, Z2.6-4, mtnr1bb** (NM_131394.1)	**zMel1c, Z2.3, mtnr1c** (NM_001161484.1)	**zMel1a2, Z1.4, mtnr1al** (NM_001159909.1)	**Shang & Zhdanova 2007**
Om	*Oncorhynchus mykiss* (rainbow trout)	**R1.7** (AF156262.1) = MT1a2					**R1.4** (AF178538.1)	**Mazurais *et al.* 1999**
El	*Esox lucius* (northern pike)			**P2.6** (AF188871.1)			**P1.4** (XM_010903666.1)	**Gaildrat and Falcón, 2000**
Oke	*Oncorhynchus keta* (chum salmon)	**mel1a** (AY356364.1) = MT1a2					**mel1b** (AY356365.1)	**Shi *et al.* 2004**
Sg	*Siganus guttatus* (golden rabbitfish)	**Mel1a** (DQ768087.1)				**Mel1c** (DQ768088.1)	**Mel1b** (DQ522314.1)	**Park *et al.*, 2006, 2007a,b, 2014**
Ca	*Carassius auratus* (goldfish)	**Mel1a1.7** (AB378058.1)			**Mel1b** (AB378059.1)	**Mel1c** (AB378060.1)	**Mel1a1.4** (AB378057.1)	**Ikegami *et al.* 2009**
		G1.7 (AB481372.1)		**G6.2** (AB481374.1)	**G2.6 type1** (AB481373.1)	**Mel1c** (AB481375.1)	**G1.4** (AB481371.1)	**Saito, unpublished**
Dl	*Dicentrarchus labrax* (sea bass)	**dlMT1** (EU378918.1)			**dlMT2** (EU378919.1)	**dlMel1c** (EU378920.1)		**Sauzet *et al.*, 2008 Herrera-Pérez P *et al.*, 2010**
Sse	*Solea senegalensis* (Senegal sole)				**ssMT2** (FM213464.1)	**ssMel1c** (FM213465.1)	**ssMT1** (FM213463.1)	**Confente *et al.* 2010**
On	*Oreochromis niloticus* (Nile tilapia)	**mel1a** (AY569971.1)						**Jin *et al.*, 2013**
Ec	*Epinephelus coioides* (orange-spotted grouper)	**MT1** (JX524508.1)			**MT2** (JX524509.1)			**Chai *et al.*, 2013**
Bp	*Boleophthalmus pectinirostris* (mudskipper)	**Mtnr1a1.7** (KC622030.1)			**Mtrn1b** (KC622031.1)	**Mtnr1c** (KC622032.1)	**Mtnr1a1.4** (KC622029.1)	**Hong *et al.* 2014**
Tn	*Takifugu niphobles* (grass puffer)	**mel1a1.7** (AB492764.1)			**mel1b** (AB492765.1)	**mel1c** (AB492766.1)	**mel1a1.4** (AB492763.1)	**Ikegami *et al.* 2015**
Pn	*Porichthys notatus* (plainfin midshipman - "singing" fish)	**mtnr1A1.7** (HQ007044)		**Mel1b** (KT878765.1)			**mtnr1a1.4** (HQ007045)	**Feng & Bass, 2016, Feng unpublished**
Amel	*Amphiprion melanopus* (cinnamon clownfish)	**MT-R1** (HM107821.1)						**Choi *et al.*, 2016**

Phylogenetic assignment according to findings of this study; previous publications using distinct nomenclature systems are provided. Sequences in red signal a significant change in assignment.