Tandem duplications in the C-terminal domain of the mesotocin receptor exclusively identified among East Eurasian thrushes.
Re-submitted to *Journal of Molecular Evolution*

**Research article**

Tandem duplications in the C-terminal domain of the mesotocin receptor exclusively identified among East Eurasian thrushes

Hideaki Abe¹, ³, Isao Nishiumi² and Miho Inoue-Murayama¹,*

¹Wildlife Research Center, Kyoto University, Kyoto, 606-8203, Japan

²Department of Zoology, National Museum of Nature and Science, Tsukuba, 305-0005, Japan

³ Current address:

Department of Anatomy, University of Otago, Dunedin, 9054, New Zealand

*Correspondence to:

Miho Inoue-Murayama
Wildlife Research Center of Kyoto University
2-24 Tanaka-Sekiden-cho, Sakyo-ku, Kyoto, 606-8203, Japan
Telephone: +81-75-751-4375  Fax:+81-75-771-4394
e-mail: mmurayama@wrc.kyoto-u.ac.jp
Mesotocin is a neurohypophyseal hormone found in some non-mammalian vertebrates, including birds, reptiles and amphibians. In this study, we identified and characterized 18-amino acid duplications in the C-terminal domain of the mesotocin receptor (MTR), specifically found in *Turdus* thrushes (Aves: Passeriforms: Turdidae). These duplicated elements are located in the distal part of the C-terminal tails of MTR and consist of amino acids that are highly conserved among major vertebrates. Intraspecific polymorphisms in a variable number of tandem duplications are commonly found in East Eurasian *Turdus*, but not in any other genus of Turdidae. Moreover, the genus *Turdus* can be further classified into 2 groups according to the presence or absence of a 3-amino acid deletion just adjacent to the putative palmitoylation site in the cytoplasmic C-terminal tail. The phylogeny presented here strongly supports the conspecific group of 4 East Eurasian thrushes (*Turdus pallidus*, *T. chrysolaus*, *T. obscurus*, and *T. celaenops*). Our findings, therefore, provide a new synapomorphy that can be used for phylogenetic assumptions and shed a light on the history of diversification within Eurasian *Turdus* clades.

**Keywords:** mesotocin receptor — C-terminal domain — tandem duplication — polymorphism — insertions and deletions — *Turdus*
Mesotocin (MT) is the non-mammalian homolog of mammalian oxytocin, and this neurohypophysial hormone exerts biological effects by binding to G protein-coupled receptors (GPCRs) on the cell surface. GPCRs are structurally characterized by 7 transmembrane domains, intracellular and extracellular loops, and amino- and C-terminal segments (Gimpl and Fahrenholz 2001; Lolait et al. 1992). We focused on length variations in the C-terminal domain of peptide hormone receptor as a key factor in influencing signal transduction pathways. Hoare et al. (1999) experimentally illustrated that artificial truncation in the C-terminal tail of the human oxytocin receptor has critical effects in receptor-G protein interactions. Furthermore, our recent study demonstrated that the arginine vasotocin receptor has inter- and intraspecific variations in the length of C-terminal domains, suggesting conformational differences in this region (Abe et al. 2012). Thus, we consider it worth investigating whether the mesotocin receptor (MTR) also has conserved sequence motifs and amino acid residues in the cytoplasmic C-tail.

In addition to characterization of novel polymorphisms, in this study we also present a hypothesis of phylogenetic relationships among East Eurasian thrushes by taking advantage of synapomorphic feature of insertions and deletions (indels) that are specifically found in Turdus clades. A synapomorphic character is one of the most reliable piece of information on phylogenetic
reconstitution, as long as parallel changes and/or reversions are unlikely between different taxa. Not only morphological characteristics but also genomic changes can be used as a landmark for identifying a taxonomic clade. Rare genomic changes (RGCs) occasionally provide unambiguous evidences for resolving the puzzle of phylogenetic problems (reviewed in Boore 2006; Rokas and Holland 2000). RGCs can be caused by various molecular events—retroposon integrations, gene order changes, and large-scale indels, etc. Some retroposons known as long and short interspersed repetitive elements (LINEs and SINEs, respectively) have attracted special attention of evolutionary biologists, because copies of these molecules shared at the same locus in 2 or more different taxa are derived from the same element originally inserted into the germline of a common ancestor (Shedlock and Okada 2000). Indeed, the presence or absence of SINEs insertions have been used for phylogenetic analyses in a variety of organisms such as salmons (Kido et al. 1991), cetaceans (Nikaido et al. 1999), and avian species (St. John et al. 2005; Watanabe et al. 2006). Moreover, some amino acid indels and substitutions that had occurred in functional proteins proved to be synapomorphic characters for avian clades (Ericson et al. 2000; Groth and Barrowclough 1999; Stapel et al. 1984).

The genus *Turdus* is the largest group of Turdidae and perhaps the most successful of any passerine genus in the world in terms of natural geographical penetration and degree of speciation. *Turdus* is the only thrush genus that is widely distributed throughout Eurasia, South and North America, and
Africa. Of its nearly 70 species, 8 are confined to Africa, 25 are essentially Asian, and 34 are
restricted to the America (Collar 2005). A previous study on phylogenetic relationships among true
thrushes indicated that Turdus radiations occurred near the end of the Miocene (7–8 million years
ago) and rapidly diversified into each continental lineage (Klicka et al. 2005). Although the
subsequent statistical analysis based on mitochondrial genes indicated a monophyly of each
continental clade, relatively weak nodal support was found within the Eurasian Turdus clade
(Voelker et al. 2007). Such a rapid radiation and speciation, which are general characteristics of
avian evolution, always hamper phylogenetic approaches using stochastic methods (see Poe and
Chubb 2004). Another striking feature of the genus Turdus is its apparent plumage
homogeneity—similar plumage patterns appear repeatedly across its wide range. This makes it more
complicated to define species boundaries and unravel the evolutionary history. Therefore, it is
important to gather information on RGCs, because one can obtain invaluable data sets for cladistic
assumptions if synapomorphic characters could be found elsewhere within problematic taxa. This
approach is expected to provide a better resolution of phylogenetic relationships and interesting
insights into speciation processes and timing at multiple taxonomic levels.

Material and methods
Avian specimens and DNA extraction

A wide range of avian species were used in this study to explore sequence and structural features of the C-terminal tails of the MTR. In addition, more than 200 Turdus specimens were particularly used to further investigate intraspecific polymorphisms in this domain of the receptor. These specimens were divided into 3 categories, depending on their sources: (1) live resources derived from zoos and aviaries, (2) carcasses originated either from wildlife conservation acts or from bird-strike events in Japanese airports, and (3) specimens obtained from avian research projects conducted by the National Museum of Nature and Science. Supplementary S-Table 1 summarizes the avian samples used for sequence alignment and their GenBank IDs. When we collected specimens from live animals, we strictly observed the domestic laws for animal rights and protection in Japan. DNA was extracted from various materials such as feathers, blood, and tissues (e.g., muscle or liver) using either of the 3 methods: DNeasy Blood and Tissue Kit (Qiagen, Tokyo, Japan), Puregene DNA Purification Kit (Gentra Systems, Minneapolis, MN, USA), or standard phenol-chloroform extraction (Sambrook et al. 1989).

Polymerase chain reaction amplification
To amplify the C-terminal region of the MTR, primers were designed on the basis of the nucleotide sequences of the zebra finch deposited in GenBank (Accession No. XM_00218826). The following oligo primers were designed for amplification of avian MTR: MTf2 (5′–GCCTCCCCCTTCATCATCG–3′), MTr3 (5′–AAGCTCCTGTGGCTCGTG–3′), MTr4 (5′–CGGCAGCTGAGCACGAAG–3′), and MTr7 (5′–TCGCGGGCGCTGCACGAA–3′). We mainly used the MTf2/MTr3 primer pair, but MTf2/MTr4 was employed for amplification of Phasianidae (Galliformes) and Anatidae (Anseriformes). MTf2/MTr7 was used in the amplifications of Turdus thrushes, because the MTr3 annealing site overlapped with lineage-specific tandem duplicated elements. In each case, polymerase chain reaction (PCR) was carried out in a PCR cocktail containing approximately 10–50 ng of genomic DNA, 0.25 µM each primer, 0.2 mM dNTPs, 5 × tuning buffer, and 0.375 U of G-Taq DNA polymerase (Hokkaido System Science, Hokkaido, Japan). PCR conditions comprised an initial incubation at 95°C for 2 min, followed by 40 cycles of 95°C for 15 s and 60°C for 30 s.

Sequencing and multiple alignments

The nucleotide sequences of the C-terminal domain of MTR were determined in a total of 44 avian
species belonging to 31 families (17 orders). Sequencing was performed on an ABI 3130xl automated sequencer (Applied Biosystems, Foster City, CA, USA) using BigDye Terminator Cycle Sequencing reagents v3.1, according to the manufacturer’s instructions (Applied Biosystems). Nucleotide sequences were translated into amino acids and then aligned by using CLUSTAL W (Thompson et al. 1994) implemented in BioEdit ver. 7.9.9.0 (Hall 1999). Each band of the amplified alleles was purified from the gel and subjected to DNA sequencing. Information on amino acid sequences of the C-terminal tail of vertebrate oxytocin and MT receptors were collected either from previous papers (Akhundova et al. 1996; Hoare et al. 1999) or from GenBank sequences.

Genotyping and sex determination

Multiple sequence alignment revealed the existence of tandem duplicated elements in the distal C-terminal tails of MTR specifically in Turdus thrushes. Thus, the pattern of length polymorphisms was further investigated by genotype analysis using capillary electrophoresis. MTf2 was fluorescently labeled with 6-carboxyfluorescein (6-FAM), and the HD500 Rox size standard (Applied Biosystems) was used as an internal standard. PCR products were electrophoresed on an ABI 3130xl sequencer, followed by peak detection with the Genemapper software (Applied Biosystems). When we could not judge the gender according to the external features of the cloaca,
sexes were determined using the DNA-based identification method described by Griffiths et al. (1998). The exact fragment size of partial CHD-W and CHD-Z on automated capillary electrophoresis was taken from the empirical data of Lee et al. (2010). We evaluated the extent of statistical heterogeneity in allelic distributions between sampling locations (west Honsyu \(n = 39\) vs. Kyusyu \(n = 14\)) and sexes (male \(n = 55\) vs. female \(n = 29\)) of the pale thrush (*Turdus pallidus*) using Fisher’s exact test.

**Results**

*Structural characteristics in the C-terminal domain of the avian mesotocin receptor*

In this study we determined the nucleotide sequences of the C-terminal domain of MTR in 44 avian species belonging to 17 orders. These sequences were deposited in the GenBank database (Accession Nos. AB634778–AB634824, and AB743549). We did not find any stop codon or frame-shift mutation in the sequence alignment; hence, we excluded the possibility that these sequences could be derived from MTR pseudogenes. The schematic secondary structure of the target region is shown in Figure 1. Multiple alignments showed that the distal part of TM VII (position...
318–335) and its juxtamembrane region (position 336–343) are completely conserved among all avian species (Fig. 2). The conserved NPxxY motif, which is characteristic of rhodopsin GPCRs (Fritze et al. 2003), could be identified in TM VII (position 328–332). A double cysteine motif in the C-terminus (positions 349, 350) was conserved in all of the avian and mammalian species; these residues have been suggested by the CSS-Palm software (Ren et al. 2008) to be putative sites for palmitoylation. Passerine birds could be distinguished from other avian species because of the common deletion of 2 amino acids in the middle of the C-terminal domain (positions 364, 365).

Furthermore, we found 54-bp duplicated elements (18 amino acids; KSNSSSFVLSCRSPSHRS) in Turdus thrushes. There were 4 types of alleles with the tandem repeat number varying from 1 to 4 [NR (wild-type), RT2, RT3, and RT4]. Four alleles were different by exactly 60-bp intervals. Global alignment with other vertebrates uncovered that the duplicated region was highly conserved not only among avian species but also among the mammalian species (Fig. 3). In particular, the KSXSSXFVLS (X: arbitrary amino acid) motif was completely conserved in all vertebrates except teleost fish. Although the functional significance of repeated motifs remains speculative, it should be mentioned that tandem duplicated motifs contain consensus sites for phosphorylation by the multifunctional protein kinase II (RxxS; Kemp and Pearson 1990).

Distribution of tandem duplicated alleles within Turdus thrushes
We detected intraspecific variations in the number of tandem duplications among *Turdus* thrushes, whereas the White’s thrush (*Zoothera dauma*) was monomorphic with the wild-type (i.e., non-duplicated) allele (Table 1). The other following avian species tested for length polymorphisms did not show any variation in this locus (zebra finch [*Taeniopygia guttata*; *n* = 32], Bengalese finch [*Lonchura striata var. domestica*; *n* = 15], red-throated parrot-finch [*Erythrura psittacea*; *n* = 8], canary [*Serinus canaria*; *n* = 10], rock eagle-owl [*Bubo bengalensis*; *n* = 8], and mallard [*Anas platyrhynchos*; *n* = 8]). These data indicate that duplication events occurred in the common ancestor of *Turdus* thrushes. There was little difference in the allelic frequency among Eurasian *Turdus* species. *RT2* was the predominant allele with a frequency ranging from 0.62 to 0.74; the frequency of the other alleles was similar among species.

**Phylogenetic assumption based on rare genomic changes**

In addition to tandem duplications in the C-terminal tails, we found a 3-amino acid deletion (Ser–Thr–Arg) at residues 351 to 353, just adjacent to a putative palmitoylation site (see Figs. 1 and 2). The non-duplicated and duplicated alleles with this deletion were expressed as *NR*–, *RT2*–, *RT3*–, and *RT4*–, and the frequency of each allele is shown in Table 1. The deletion was conserved in all
individuals belonging to 4 *Turdus* species (*T. pallidus*, brown-headed thrush *T. chrysolaus*,
eyebrowed thrush *T. obscurus*, and Izu thrush *T. celaenops*), but no deletion was detected in the
other *Turdus* (Naumann’s thrush *T. naumanni* and Japanese thrush *T.cardis*) and outgroup genera
\textit{Zoothera} (*Z. dauma*). Therefore, the East Eurasian thrushes can be divided into 3 groups according
to the presence or absence of these synapomorphic indels. Eurasian *Turdus* Group 1 (ETG1) has
characteristics of both tandem duplications and deletions, and includes *T. pallidus* and its conspecies
*T. chrysolaus, T. obscurus, and T. celaenops*. Eurasian *Turdus* Group 2 (ETG2) consists of the other
major *Turdus* thrushes (*T. naumanni* and *T.cardis*) whose only tandem duplications are detected in
the C-terminal domain. The thrush group who has neither tandem duplications nor the 3-amino acid
deletion is considered to form an outer clade of ETG1 and ETG2 (Fig. 4) and named OT (other
species of Turdidae).

In the next step, we tried to identify to which group the other *Turdus* species belong to, according to
the presence or absence of these synapomorphic indels. Two samples of the grey-backed thrush (*T.
hortulorum*) collected in central Japan had the genotype of RT2/RT2 and RT2/RT3, respectively, thus
belonging to ETG2 (Table 2). The other 2 Eurasian *Turdus* species (island thrush *T. poliocephalus
niveiceps* and rufous-throated thrush *T. ruficollis*) also had duplicated alleles that were
characteristics of ETG2 thrushes, whereas the Siberian thrush (*T. sibiricus*) could be clearly
distinguished from other *Turdus* species by the absence of a duplicated allele.
We detected intraspecific polymorphisms in a variable number of tandem duplications in all *Turdus* species (with large sample sizes), even though we did not detect a significant difference in the allelic distribution between sexes ($P = 0.383$) nor between sampling locations (Honsyu vs. Kyusyu; $P = 0.257$) in *T. pallidus* populations. This simply suggests that tandem duplications in the C-terminal tail would have no impact on the receptor functions, and that the current level of allelic distributions is determined by a random genetic drift under the neutral mutation model of evolution. One of the most plausible explanations for the retention of polymorphism within each *Turdus* species is gene conversion events associated with genetic drift. Recent study suggests that gene conversion can generate new haplotypes by transferring sections of DNA within and across duplicated loci in wild bird populations (Spurgin et al. 2011). Similar mechanism may contribute to generate genetic instability in the C-terminal domain of MTR. Gene conversion can also homogenize the level of variation among populations by increasing or decreasing of the tandem duplication unit. At present, however, we do not have an answer to the following fundamental questions: Why can only the genomes of *Turdus* thrushes accommodate structural changes in the C-terminal domain of MTR?
Why are duplicated elements so strictly conserved among vertebrates? To address these questions, we need to conduct site-directed mutagenesis and/or other biochemical experiments targeting C-terminal domain of MTR.

The phylogenetic relationship inferred from synapomorphic indels strongly supports the conspecific relationships among *T. pallidus*, *T. chrysolaus*, *T. obscurus*, and *T. celaenops*. This is consistent with molecular phylogenies based on the cytochrome c oxidase subunit I (*COI*) sequences (Saitoh et al. manuscript in preparation) and datasets from 3 mitochondrial genes (Voelker et al. 2007). Moreover, local ornithologists have reported that some of Eurasian Turdus belonging to EGT1 are temporally intermingled with each other for wintering in Japan, and that *T. chrysolaus* sometimes can be seen in the same flock with *T. pallidus*. All these findings suggest a close relationship and recent diversification of ETG1 thrushes. In fact, captive and natural hybridization was reported between several combinations of closely related Turdus species such as *T. chrysolaus* × *T. pallidus* (male) and *T. naumanni* × *T. ruficollis* (McCarthy 2006, p. 239). Natural hybrid zones between Eurasian Turdus species may help to maintain the allele frequency constant among each group of Turdus thrushes, even though our data presented here clearly reject the possibility of hybridization between ETG1 and ETG2 thrushes. *T. celaenops* is an exceptional species, because it is considered a sedentary bird that breeds and reproduces on small islands, in contrast to the other migratory thrushes. Therefore, *T. celaenops* is considered to have lost the rare alleles (*RT3*– and *RT4*–) recently due to local bottleneck
effects in some island populations.

Not only hybridization with closely related species but also other biological factors influence the magnitude of polymorphisms in MTR. *T. pallidus* is a migratory bird of eastern Asia, and southern and central parts of Japan are their primary wintering grounds (Brazil 1991). According to the first breeding record in the Hiroshima prefecture (located in the western part of Honsyu) in 1991, some populations of *T. pallidus* stay in Japan throughout the year (Ueno et al. 1993). Such a social and behavioral dimorphism would promote diversification and speciation in local populations, resulting in stochastic transitions of genetic variation. In our future research, the levels and patterns of polymorphisms in other neurotransmitter-related genes will be examined between resident and migrating populations of *T. pallidus*.
The authors wish to thank the following persons and institutions for providing avian specimens:

Kayoko Kameda (Lake Biwa Museum), Kimiyuki Tsuchiya (Applied Biology Co. Ltd.), Hitoshi Suzuki (Hokkaido University), Masatsugu Suzuki and Kumi Haraguchi (Gifu University), Masahiko Sakanashi (Matsubase Store House), Mizuki Mizutani (Fukui Nature Conservation Center), Yusaku Watanabe (Iwate Wildlife Rescue Center), Hideyuki Ito (Kyoto City Zoo), Tomoko Kawakami (Kujukushima Zoological and Botanical Garden), Taku Mizuta (Amami Wildlife Center), and Yuji Miyake (Yambaru Wildlife Center). We also thank Azusa Hayano (Kyoto University) for great assistance with laboratory work. This study was supported financially by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) with a Grant-in-aid for Science Research (#21310150 to MI-M) and Global Center of Excellence (GCOE) Program “Formation of a Strategic Base for Biodiversity and Evolutionary Research: from Genome to Ecosystem”
The authors declare that they have no conflict of interest.
References


Voelker G, Rohwer S, Bowie RCK, Outlaw DC (2007) Molecular systematics of a speciose,

<table>
<thead>
<tr>
<th>Species/allele*</th>
<th>2n</th>
<th>NR</th>
<th>NR⁻</th>
<th>RT2</th>
<th>RT2⁻</th>
<th>RT3</th>
<th>RT3⁻</th>
<th>RT4</th>
<th>RT4⁻</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. pallidus</em>*</td>
<td>170</td>
<td>25 (0.15)</td>
<td>120 (0.71)</td>
<td>22 (0.13)</td>
<td>3 (0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>T. chrysolaus</em></td>
<td>22</td>
<td>6 (0.27)</td>
<td>15 (0.68)</td>
<td>1 (0.05)</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>T. obscurus</em></td>
<td>10</td>
<td>1 (0.1)</td>
<td>7 (0.7)</td>
<td>1 (0.1)</td>
<td>1 (0.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>T. celaenops</em></td>
<td>86</td>
<td>22 (0.26)</td>
<td>64 (0.74)</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>T. cardis</em></td>
<td>52</td>
<td>9 (0.17)</td>
<td>32 (0.62)</td>
<td>9 (0.17)</td>
<td>2 (0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>T. naumanni</em></td>
<td>24</td>
<td>3 (0.13)</td>
<td>16 (0.67)</td>
<td>3 (0.03)</td>
<td>2 (0.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Zoothera dauma</em></td>
<td>18</td>
<td>18 (1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* NR (wild-type, no duplication), RT2 (2 tandem repeats), RT3 (3 tandem repeats), and RT4 (4 tandem repeats). A dash (—) beside allele name indicates the presence of the 3-amino acid deletion.

---

Table 1
Distributions of wildtype and duplicated alleles among *Turdus* and *Zoothera* thrushes
### Table 2
Genotyping in *Turdus* species with less than 3 specimens

<table>
<thead>
<tr>
<th>Species</th>
<th>n</th>
<th>Locations</th>
<th>Genotype(s)*</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. poliocephalus niveiceps</em></td>
<td>1</td>
<td>Taiwan</td>
<td>RT2/RT2</td>
<td>ETG2</td>
</tr>
<tr>
<td><em>T. ruficollis</em></td>
<td>1</td>
<td>Mongolia</td>
<td>NR/RT2</td>
<td>ETG2</td>
</tr>
<tr>
<td><em>T. hortulorum</em></td>
<td>2</td>
<td>central Japan</td>
<td>RT2/RT2, RT2/RT3</td>
<td>ETG2</td>
</tr>
<tr>
<td><em>T. sibiricus</em></td>
<td>2</td>
<td>northern Japan</td>
<td>NR/NR, NR/NR</td>
<td>OT</td>
</tr>
</tbody>
</table>

* NR (wild-type, no duplication), RT2 (2 tandem repeats), and RT3 (3 tandem repeats)
Fig. 1

Schematic structure of the C-terminal tail of the mesotocin receptor (MTR). The double cysteine residues for palmitoylation are highlighted and the amino acid residues highly conserved among vertebrates are shown in capital letters. The position of the 3-amino acid deletion (Ser–Thr–Arg) is boxed, and the tandem duplicated domain in Turdus thrushes is highlighted with gray background.

Fig. 2

Comparison of the amino acid sequences at the C-terminal tail of mammalian oxytocin receptors (OXTRs), MTRs (amphibian, marsupial, and avian), and fish isotocin receptor. GenBank sequences are derived from rat (rOXTR; NP037003.2), human (hOXTR; NP000907.2), cow (cOXTR; NP776559.1), porcine (pOXTR; NP999192.1), canine (dOXTR; NP001185588.1), western clawed frog (fMTR; XM002936297.1), gray short-tailed opossum (oMTR; XP001375059.1), green anole (aMTR; XP003224939.1), non-passerine Neoaves (npMTR; AB634795—AB634816), zebra finch (zfMTR; XM002188266.1), chicken (cMTR; NP001026740.1), and zebrafish (fISR; XP001341507.1). RT2, RT3, and RT4 are duplicated elements found in Turdus thrushes. Highly conserved motifs are shaded, and the symbols ¶ and § represent S/G/N/A/T and S/G, respectively.
The sequence of npMTR is truncated because of internal primer design.

**Fig. 3**

Alignment of amino acid sequences in the distal part of transmembrane VII and the C-terminal tail of the avian mesotocin receptor. A dot or a dash denotes identity with the consensus sequence or a deletion, respectively. The GenBank sequence of zebra finch (XM002188266) is also used for alignment, and conserved residues in all avian taxa are highlighted. X in the sequence of Pthrush [RT3–] indicates alanine or glycine due to heterozygous nucleotides in this position. A deletion that could be identified only in Eurasian *Turdus* Group 1 (ETG1) is highlighted with gray background. The square brackets indicate the number of repetitive units in the C-terminal tail, and duplicated elements are shown at the bottom of the alignment (each repeat unit is boxed). Vertical bars represent the following categories: P, passerine birds; N, non-passerine Neoaves; G, Galliformes; and A, Anseriformes. Details of avian specimens are shown in S-Table 1.

**Fig. 4**

Phylogenetic relationships of East Eurasian thrushes inferred from synapomorphic characters at the C-terminal domain of the mesotocin receptor. The timings of 3-amino acid deletion and 18-amino
acid duplications are indicated as red circle and boxes, respectively. Eurasian Turdus thrushes can be divided into 2 groups, Eurasian Turdus Groups 1 and 2 (ETG1 and ETG2) according to these synapomorphic characters.

Supplementary material

Table S1 Details of avian specimens used for sequencing
| Consensus | S | V | S | K | K | S | N | S | S | T | F | V | L | S | C | R | S | S | S | Q | R | S |
| rOXTR     |   |   |   |   |   |   |   |   |   |   |   |   |   |   | R |   |   |   |   |   |   |   |
| hOXTR     | A |   |   |   |   |   |   |   |   |   |   |   |   | S | H |   |   |   |   |   |   |   |
| cOXTR     |   |   |   |   |   |   |   |   |   |   |   |   |   | Q | Y | R |   |   |   |   |   |   |
| pOXTR     |   |   |   |   |   |   |   |   |   |   |   |   |   | Q | H | K |   |   |   |   |   |   |
| dOXTR     |   |   |   |   |   |   |   |   |   |   |   |   |   | H | H |   |   |   |   |   |   |   |
| fMTR      | T | R |   |   |   |   |   |   |   |   |   |   |   | R | K |   |   |   |   |   |   |   |
| aMTR      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | L | K |   |   |   |   |   |
| oMTR      | A |   |   |   |   |   |   |   |   |   |   |   |   |   |   | R | K |   |   |   |   |   |
| npMTR     | T | $ |   |   |   |   |   |   |   |   |   |   |   | S |   | K | [reverse primer] |   |   |   |   |   |   |
| zfMTR     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | G | T | H |   |   |   |   |
| RT2/3/4   |   |   |   |   |   | F | Q |   |   |   |   |   |   |   |   | P | H |   |   |   |   |   |
| cMTR      |   | G | R |   |   | H |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| fISR      | Q | D | R |   |   |   |   | T | Y |   | I | K | S | T |   |   |   |   |   |   |   |   |

Fig. 3