Title	Ancestral Y-linked genes were maintained by translocation to the X and Y chromosomes fused to an autosomal pair in the Okinawa spiny rat Tokudaia muenninki
Author(s)	Murata, Chie; Kuroki, Yoko; Imoto, Issei; Kuroiwa, Asato
Citation	Chromosome research, 24(3), 407-419 https://doi.org/10.1007/s10577-016-9531-y
Issue Date	2016-09
Doc URL	http://hdl.handle.net/2115/67083
Rights	The final publication is available at link.springer.com
Type	article (author version)
File Information	MS_HUSCAP.pdf

[Instructions for use](#)

**Ancestral Y-linked genes were maintained by translocation to the X and Y chromosomes
fused to an autosomal pair in the Okinawa spiny rat *Tokudaia muenninki***

Chie Murata, Yoko Kuroki, Issei Imoto, Asato Kuroiwa

C. Murata

Department of Human Genetics, Institute of Biomedical Sciences, Tokushima University
Graduate School, 3-18-15 Kuramoto-cho, Tokushima 770-8503, Japan

Y. Kuroki

RIKEN, Center for Integrative Medical Sciences, 1-7-22 Suehiro-cho, Tsurumi, Yokohama,
Kanagawa 230-0045, Japan

Present address: Division of Pediatric Disease Genomics, Department of Genome
Medicine, National Research Institute for Child Health and Development, 2-10-1, Okura,
Setagaya-ku, Tokyo 157-8535, Japan

I. Imoto

Department of Human Genetics, Institute of Biomedical Sciences, Tokushima University
Graduate School, 3-18-15 Kuramoto-cho, Tokushima 770-8503, Japan

A. Kuroiwa

Laboratory of Animal Cytogenetics, Faculty of Science, Hokkaido University, Kita 10 Nishi 8,
Kita-ku, Sapporo, Hokkaido 060-0810, Japan

Tel: +81-11-706-2752; Fax: +81-11-706-2619; E-mail: asatok@sci.hokudai.ac.jp

Correspondence: A. Kuroiwa

Keywords: sex chromosome, evolution, translocation, gene amplification, RNA-seq, BAC

A short running title: Y-linked gene translocations in the Okinawa spiny rat.

Introduction

The mammalian Y chromosome is male-specific and triggers the differentiation of the developing fetus into a male. The Y is one of the smallest chromosomes with the fewest genes, despite previous homology to the X chromosome (Ross et al. 2005). Most present therians (placentas and marsupials) maintain the sex chromosomes formed in the therian ancestor; a core set of Y-linked genes are shared among many eutherian species, while many multiple-copy genes are species- or lineage-specific (Paria et al. 2011; Bellott et al. 2014; Cortez et al. 2014). Models of Y gene loss dynamics revealed that Y chromosomal gene decay proceeds rapidly upon differentiation and then markedly levels off when small, but apparently essential, gene repertoires are defined (Hughes et al. 2012; Bellott et al. 2014; Cortez et al. 2014). In particular, the sex-determining gene *SRY* has persisted over long evolutionary time periods in most therians, with rare examples of *SRY* loss in the *Tokudaia* and *Ellobius* lineages (Just et al. 1995; Soullier et al. 1998; Vogel et al. 1998; Suto et al. 2001). Several long-lived ancestral Y-linked genes were also lost in at least one mammalian lineage during recent evolution in the clade (Bellot et al. 2014). However, other genes have been conserved by compensatory mechanisms, e.g., the rescue of Y-linked gene loss via gene transposition to other chromosomes (Hughes et al. 2015).

This compensatory mechanism was first detected in *Tokudaia* species (Kuroiwa et al. 2010). The Amami and Tokunoshima spiny rats, *Tokudaia osimensis* and *Tokudaia tokunoshimensis*, respectively, have an XO/XO sex chromosome constitution; they lack the entire Y chromosome, except for a small region translocated to the X chromosome (Honda et al. 1977, 1978; Kuroiwa et al. 2010). The proto-Y region on the X chromosome contains at least four Y-linked genes, *EIF2S3Y*, *KDM5D*, *TSPY*, and *ZFY*, but functional *RBMY* and *SRY* have been absent from the genome since the secondary Y chromosome loss (Arakawa et al. 2002; Kuroiwa et al. 2010). Furthermore, *KDM5D* and *EIF2S3Y* are expressed in the gonads of both sexes in *T. osimensis* (Kuroiwa et al. 2010).

The Okinawa spiny rat *Tokudaia muenninki* has the XX/XY sex chromosome

constitution and multiple *SRY* copies on the large Y chromosome fused to an autosome (Tsuchiya et al. 1989; Murata et al. 2010, 2012). It has an ancestral Y chromosome region and many *SRY* sequences, but these *SRY* loci may have not retained sex-determining functions (Murata et al. 2010; Kimura et al. 2014). *SRY* proteins bind to the core region within a testis-specific enhancer of *SOX9* (TESCO) with SF1 to upregulate *SOX9* expression in the undifferentiated gonads of XY embryos of humans and mice (Sekido et al. 2008). TESCO of the genus *Tokudaia* lost enhancer activity due to mutations in its *SRY*- and SF1-binding sites, and *T. muenninki* *SRY* does not activate the mutant TESCO or mouse TESCO (Kimura et al. 2014).

In our previous study, we found that two Y-linked genes, *UTY* and *DDX3Y*, are single-copy genes on the pericentromeric region of Yp, which is the ancestral Y chromosome region in *T. muenninki* (Murata et al. 2012). By contrast, most other Y-linked genes are excessively duplicated, like the *SRY*s. We investigated six apparently multi-copy Y-linked genes, *ZFY*, *UBA1Y*, *EIF2S3Y*, *TSPY*, *USP9Y*, and *RBMV*, and two single-copy genes, *UTY* and *DDX3Y*, to examine whether the functions of multiple-copy genes were compensated by translocation or transposition. We performed RNA-seq using three tissue types to identify transcripts of ancestral Y-linked genes. We then conducted reverse transcription PCR (RT-PCR) and 3' rapid amplification of cDNA ends (RACE) using six tissue types from an adult male to examine the functional conservation of these Y-linked genes based on expression patterns and the complete coding sequences, and to determine whether mRNA sequence variants existed for each gene. Furthermore, we performed fluorescent *in situ* hybridization (FISH) using cDNA clones of Y-linked genes as probes to display their chromosomal localization. Using quantitative real-time PCR (qPCR) with genomic DNA, we estimated the copy number of each Y-linked gene in the male genome. The six genes were excessively duplicated on the Y chromosome, and several copies of *ZFY* and *UBA1Y* were located on the X chromosome. We then performed next-generation sequencing of five bacterial artificial chromosome (BAC) clones including *UTY* and/or *DDX3Y* to reveal the genomic structure of the two genes as well as adjacent genes in the

84 ancestral Y euchromatic region. Our results show that functional Y-linked genes are
85 compensated by translocations to other sex chromosomal regions and amplifications in *T.*
86 *muenninki*.

Materials and Methods

Sample preparation

T. muenninki is endangered (The IUCN Red List of Threatened Species; <http://www.iucnredlist.org/>, 3/8/2016) and has been protected by the Japanese government as a natural treasure since 1972 (Yamada et al. 2010). With permission from the Agency for Cultural Affairs and the Ministry of the Environment in Japan, all spiny rats were released at their capture sites after a small piece was cut from the tip of their tail. Using the tail tissues, cell cultures and DNA extraction were performed as previously reported (Murata et al. 2010). Total RNA was extracted from the brain, liver, testis, kidney, lung, and spleen of one adult male *T. muenninki* who died of accidental causes in March 2013 using an RNeasy Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. All experimental procedures conformed to the guidelines for animal experimentation of Tokushima University and all institutional and national guidelines for the care and use of laboratory animals were followed.

RNA-seq and data assembly

The cDNA libraries from the brain, liver, and testis were constructed using the TruSeq RNA Sample Preparation Kit v2 (Illumina, San Diego, CA, USA). Three indexed libraries from each tissue were pooled and sequenced using one lane to obtain 101 nucleotide single-end reads with an Illumina HiSeq 2500 RNA-seq data were deposited in the DDBJ Sequence Read Archive under the accession number DRA004624. *De novo* assembly and reference mapping were performed after removing polyA/T tails and filtering low-quality reads. *De novo* assembly was carried out using Trinity (Grabherr et al. 2011; Haas et al. 2013) with the default settings. BLAST searches using the assembled sequences were performed with Blast+ ver. 2.2.30 (Camacho et al. 2009; <ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>). The National Center for Biotechnology Information (NCBI) transcript reference sequences (refseq_rna.*.tar.gz) were downloaded via FTP from <ftp://ftp.ncbi.nlm.nih.gov/blast/db/>. The *de*

de novo assembled sequences matched to Y-linked genes, and these X homologs in the BLAST search were extracted using the 'grep' command. To fill in gaps between the extracted contigs, RNA-seq reads were mapped to RefSeq mRNA sequences from the mouse (refMrna.fa.gz downloaded from <http://hgdownload.cse.ucsc.edu/goldenPath/mm10/bigZips/>) using BWA 0.5.9 with a 15% uniform base error rate (Li and Durbin 2009). The Integrative Genomics Viewer was used to visualize reads mapped to the reference genome (Robinson et al. 2011). Accession numbers of the eight X-linked gene sequences determined in this analysis are listed in Table S1.

RT-PCR and 3' RACE

RT-PCR and 3' RACE were performed using testis tissue to determine the complete coding sequences of ancestral Y-linked genes by filling in gaps in the *de novo* assembled sequences and to confirm the sequence variation for the transcripts of each Y-linked gene. RT-PCR using multiple tissues from a male was also conducted to determine the expression patterns of the Y-linked genes. The cDNA was synthesized using the SuperScript III First-Strand System for RT-PCR Kit (Thermo Fisher Scientific, MA, USA). The primers for each RT-PCR amplification were designed using the partial transcript sequences determined by *de novo* assembly and the ancestral Y genome sequence determined by BAC sequencing. *ACTB* was used as a positive control, and its primer pair was designed from the common sequence of mouse and rat. The primer sequences are shown in Table S2. Amplification was carried out in a 20 µl mixture containing 0.5 µl of cDNA, 0.3 µM each primer, 0.2 mM each of the four deoxynucleotide triphosphates (dNTPs), 0.25 U of PrimeSTAR GXL DNA Polymerase, and a one-fifth volume of 5× PrimeSTAR GXL Buffer (TaKaRa, Shiga, Japan). The RT-PCR conditions were as follows: 98°C for 2 min; 30 cycles at 98°C for 10 s, 55–60°C for 15 s, and 68°C for 30 s to 4 min; and a final extension at 72°C for 7 min. For 3' RACE, cDNA synthesis was initiated at the poly(A) tail of mRNA using the adapter primer 5'-TGG AAG AAT TCG CGG CCG CAG TTT TTT TTT TTT TTT T-3' with the PrimeScript II 1st Strand cDNA Synthesis Kit (TaKaRa). The

unknown coding region of *USP9Y* was amplified using the Ex Taq HS System (TaKaRa) with the gene-specific primer 5'-CAC CAG CAC CAC AAG ATA GA-3' and the universal amplification primer 5'-TGG AAG AAT TCG CGG CCG CAG-3'. PCR products were cloned using the TArget Clone Kit (ToYoBo, Osaka, Japan) after overhanging dA was added to the 3' ends of the products using A-attachment mix (ToYoBo). Both the PCR products and cDNA clones were sequenced using an ABI 3500xL Genetic Analyzer. The identified Y gene and homologous X sequences were compared with orthologous sequences of the mouse (*Mus musculus*), rat (*Rattus norvegicus*), *T. osimensis*, and *Apodemus sylvaticus* using ClustalX 2.0 and MEGA5.2 (Larkin et al. 2007; Tamura et al. 2008). Furthermore, we performed phylogenetic analyses for each X/Y gene pair including *T. muenninki* and other rodent genes using MEGA5.2 to confirm whether X- and Y-linked genes formed separate clusters. Accession numbers of the Y-linked transcript sequences determined by RT-PCR and 3' RACE with *de novo* assembled sequences are listed in Table S1. Accession numbers of the X and Y gene sequences in other rodents are also listed in Table S1.

Chromosome preparation and FISH

The preparation of R-banded chromosomes was performed as described by Matsuda et al. (1992), Matsuda and Chapman (1995), and Kobayashi et al. (2008). The cDNA clones of *ZFY* (2,263 bp), *UBAIY* (1,778 bp), *EIF2S3Y* (1,458 bp), *TSPY* (691 bp), *USP9Y-2* (3,616 bp), and *RBMY* (1,123 bp) were used as probes. The cDNA clones *ZFX* (2,816 bp) and *UBAI* (2,114 bp) from the brain were also used, and the primer pairs used to detect the two X-linked genes by RT-PCR are shown in Table S2. The FISH analysis using the cDNA clones was performed as described by Kuroiwa et al. (2010). The FISH signals were observed under an Olympus fluorescence microscope BX53, and the images were captured with a DP73 digital camera (Olympus, Tokyo, Japan).

qPCR

The copy number of each amplified ancestral Y-linked gene was quantified using gene-specific primers designed using Primer3 ver. 0.4.0 (<http://bioinfo.ut.ee/primer3-0.4.0/>). For each qPCR, *DDX3Y* and *UTY*, single-copy genes in the *T. muenninki* male genome, were used as an internal control for copy number. The primer pairs used for this analysis are shown in Table S2. To achieve maximum efficiency of the qPCR, the product size was limited to 90–109 bp.

qPCR was conducted using TaKaRa SYBR *Premix Ex Taq* (Perfect Real Time) in accordance with the manufacturer's protocol, and amplification was detected using the Applied Biosystems 7500 Real-Time PCR System. The qPCR reactions were performed in triplicate in a 10 µl reaction volume in a 96-well plate under the following conditions: 95°C for 30 s, followed by 40 cycles of 95°C for 5 s and 60°C for 34 s. Genomic DNA samples from five *T. muenninki* males were used in the analysis. The presence of a single melting temperature peak, which represented a single amplicon, was used to validate primer specificity, and the slope and R^2 values for serial dilutions of genomic DNA were used to confirm the efficiency of the reaction. To detect potential contamination during plate preparation, nuclease-free water was included in each set of reactions as a negative control.

BAC clone sequencing, data assembly, and sequence comparison

BAC libraries of a single male *T. muenninki* were previously constructed (Murata et al. 2015). Five BAC clone sequences were determined, including *UTY* and/or *DDX3Y* genes (TMB1-066O12, 237P18, 265P11, 324H06, and 389K21), by multiplex sequencing on an Illumina HiSeq 1000 using a 100 bp paired-end protocol, as described in our previous study (Murata et al. 2015). The primers used to screen the BAC libraries were designed within an exon from the transcript sequences of the two genes determined by RT-PCR and sequencing. The primer pairs used for gene amplification were as follows: *UTY*, Primers no. 2 and 3 in Table S2; *DDX3Y*, 5'-GGA AGG ATT CAT TGA CTT TAG TA-3' and 5'-TCT GTG ATC TGT CTC

CAT GA-3'. Amplification was carried out in a 10 µl mixture containing 1.0 µl of library DNA, 0.5 µM each primer, 0.2 mM each of the four dNTPs, 0.5 U of BIOTaq HS DNA Polymerase, and a one-tenth volume of 10× ImmoBuffer. PCR was performed under the following conditions: denaturation for 10 min at 95°C, followed by 45 cycles of denaturation for 30 s at 95°C, annealing for 30 s at 55°C, and extension for 30 s at 72°C. BAC sequences were *de novo* assembled as previously reported (Murata et al. 2015). Gaps within *de novo* assembled sequences were bridged by PCR and sequencing. Finally 372,407 bp of combined sequence from five BAC clones was determined without gaps (LC066213).

To capture genome rearrangements, a dot-plot analysis was performed for the *T. muenninki* Y and mouse Y regions. The mouse genome sequence (NC_000087) was obtained from NCBI. A nucleotide-based alignment was generated using NUCmer, which is a mode in MUMmer 3.0, for the large-scale genome comparisons with default settings (minimum length of a maximal exact match: 20, and minimum cluster length: 65; Kurtz et al. 2004). The exon-intron boundaries were defined for the combined sequence of five BAC clones by comparison to the transcript sequences of *T. muenninki* and orthologous genes in the mouse genome. To verify whether the genes located on the ancestral Y genome were transcribed into mRNAs, the predicted mRNA sequences were compared with the actual mRNA sequences of *T. muenninki* using ClustalX 2.0. Furthermore, the sequences surrounding *SRY* with masked repetitive elements were compared between *T. muenninki*, mouse, and rat using ClustalX 2.0. Repetitive elements were identified using the RepeatMasker Web Server (<http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>).

Results

Identification of expressed Y-linked genes and their X homologs

Y-linked genes diverge more rapidly than other chromosomal genes among species. Therefore, the *de novo* assembly of RNA-seq data is a useful strategy to find Y-linked genes without a genome sequence, such as in non-laboratory animals. We performed RNA-seq using three tissue types to identify Y-linked genes in *T. muenninki*. We generated 115,946,604, 101,801,400, and 111,401,004 single-end reads from the testis, brain, and liver, respectively. A total of 91–92% of the sequenced reads passed the quality control check and were used for *de novo* Trinity assembly. After the *de novo* assembly of all available reads from the three tissues, we obtained 111,185 components with an N50 of 3,437 bp, defined as the maximum length such that 50% of the total assembled sequences were contained in contigs of at least this length. We extracted the partial transcript sequences that matched eight mouse and rat Y-linked genes (*ZFY*, *UBA1Y*, *EIF2S3Y*, *TSPY*, *UTY*, *DDX3Y*, *USP9Y*, and *RBMV*) by BLAST searches of the *de novo* assembled sequences. We also identified the complete coding sequences of six X-linked genes (*ZFX*, *UBA1X*, *KDM6A/UTX*, *DDX3X*, *USP9X*, and *RBMX*) from the assembled sequences, and we determined the sequences of two genes, *EIF2S3X* and *TSPYL2*, by filling in gaps using the read sequences mapped to the mouse RefSeq mRNA sequences. The similarities of the nucleotide and amino acid sequences were 91.8–97.6% and 93.4–99.4%, respectively, between X-linked genes of *T. muenninki* and mouse (Table S3).

Predicted amino acid sequences and expression patterns of Y-linked genes

Using RT-PCR and 3' RACE with an adult testis tissue sample, we determined the complete coding sequences of seven Y-linked genes, *ZFY*, *EIF2S3Y*, *TSPY*, *UTY*, *DDX3Y*, *USP9Y*, and *RBMV*, but not *UBA1Y* (Fig. S1). The identification of X and Y genes was performed by phylogenetic analysis with each X/Y gene sequences of *T. muenninki* and other rodents, revealing that their Y-linked genes were grouped into the same cluster and formed separate

cluster to the X homologs (Fig. S2). We observed several nucleotide substitution sites within the transcript sequences of *ZFY*, *EIF2S3Y*, *TSPY*, *USP9Y*, and *RBMY*, and no sequence variation within the *UTY* and *DDX3Y* transcripts (Fig. S1). The predicted amino acid sequences of the seven Y-linked genes exhibited 63.3–99.6% similarity between *T. muenninki* and other rodents (Table 1). Some of the *ZFY* transcripts had 1 bp and 7 bp deletions at positions 369 and 494–496, respectively, in *T. muenninki* (Fig. S1a). Therefore, the predicted amino acid sequences of the transcripts without either deletion were compared with the mouse and *T. osimensis* orthologs (Table 1, Fig. S1b). We did not detect sequence variation in *UBA1Y*, but we observed a partial coding sequence caused by an internal stop codon. The internal stop codon was generated by a frameshift 86 bp deletion, which resulted in the deletion of more than 230 amino acid residues (Fig. S1c, d). The predicted amino acid sequences of *EIF2S3Y*, *TSPY*, *UTY*, and *DDX3Y* were highly similar to the orthologs of other rodents without frameshift mutations (Table 1, Fig. S1e–l). In *USP9Y*, we identified a difference between *T. muenninki* and mouse orthologs at the last 22 amino acid residues due to a frameshift mutation (Fig. S1m–n). Based on a sequence comparison of *RBMY* among *T. muenninki*, mouse, and rat, we inferred that the frameshift mutation occurred in the mouse lineage, and sequence similarity was higher for the comparison between *T. muenninki* and the rat than between *T. muenninki* and the mouse (Table 1, Fig. S1o, p).

We examined the expression patterns of the eight Y-linked genes by RT-PCR using six tissues obtained from an adult male. We observed higher *ZFY* expression in the testis and brain than in other tissues, and we detected *UBA1Y* expression in all tissue types, except for the kidney. *EIF2S3Y*, *UTY*, and *DDX3Y* showed ubiquitous expression, *TSPY* and *RBMY* were specifically expressed in the testis, and *USP9Y* was expressed more highly in the testis than in other tissues in *T. muenninki* (Fig. 1).

Chromosomal localization of Y-linked genes

We mapped the *T. muenninki* cDNA clones of six Y-linked genes, *ZFY*, *UBA1Y*, *EIF2S3Y*, *TSPY*, *USP9Y*, and *RBMY*, to chromosomes of a male by direct R-banding FISH. All six genes were located on the long arm of the Y chromosome as multiple-copy genes (Fig. 2). *ZFY* was also distributed on the distal region of Xq as well as Yq, and *ZFX* was located on the middle region of Xq (Fig. 2a–c, Fig. S3a, b). Furthermore, we observed FISH signals of *UBA1Y* along the heterochromatin block of the X chromosome as well as Yq (Fig. 2d–f), and we detected *UBA1* signals in the pericentromeric region of Xq (Fig. S3c, d). *EIF2S3Y* and *USP9Y* copies were mainly located on the pericentromeric and middle region of Yq, respectively (Fig. 2g–l). The other four genes were amplified on the pericentromeric and middle region of Yq (Fig. 2a–f, m–r). We also observed *USP9Y* on the pericentromeric region of Yp (Fig. 2j–l), which was the predicted ancestral Y euchromatic region including two single-copy genes *UTY* and *DDX3Y* (Murata et al. 2012).

Copy numbers of Y-linked genes

We performed qPCR using genomic DNA of five *T. muenninki* males to estimate the copy numbers of the six Y-linked genes (*ZFY*, *UBA1Y*, *EIF2S3Y*, *TSPY*, *USP9Y*, and *RBMY*) that were identified as multiple-copy genes in the FISH analysis (Fig. 2). We also examined the copy number of *SRY*, which was previously demonstrated to have multiple copies (Murata et al. 2010). We used *UTY* and *DDX3Y* as internal control genes with a single copy per male diploid genome. Setting the copy number value of *UTY* (or *DDX3Y*) to 1.0, the relative copy numbers for *DDX3Y* (or *UTY*), *ZFY*, *UBA1Y*, *EIF2S3Y*, *TSPY*, *USP9Y*, *RBMY*, and *SRY* were approximately 1, 63, 66, 10, 61, 34, 50, and 61, respectively (Fig. 3, Table 1).

Genomic structure of ancestral Y genes

We screened five BAC clones including *UTY* and/or *DDX3Y*, which are single-copy genes on Yp, to identify the genomic structure of ancestral Y genes. We determined a 372.4 kb sequence

by next-generation sequencing using the BAC clones. Based on a comparison between *T. muenninki* and the mouse Y chromosome sequence, we identified a large homologous block without apparent inversions (Fig. 4). We obtained the sequences of *SRY*, *EIF2S3Y*, *TSPY*, *UTY*, *DDX3Y*, and *USP9Y* (Fig. 4a, b). We detected the reversed *SRY* sequence upstream of *EIF2S3Y* in *T. muenninki*, although *SRY* was located more than 1 Mb downstream of *USP9Y* in the mouse genome (Fig. 4c). Exons 1–4 and exon 12, including the complete 447 bp coding region, were absent from *EIF2S3Y*, although the upstream and downstream sequences of *EIF2S3Y* were found in the BAC clone sequence of *T. muenninki* (Fig. 4, Fig. S1e). We did not detect exons 1–28 in *USP9Y* because the terminal region of the BAC sequence was within intron 29 of *USP9Y* (Fig. 4a). Additionally, the ancestral *USP9Y* genomic sequence had an internal stop codon within exon 29 (Fig. S1m). The predicted transcript sequences inferred from the genomic sequence of three multiple-copy genes, *EIF2S3Y*, *TSPY*, and *USP9Y*, were different from the actual mRNA sequences (Table 1, Fig. S1e, g, m). For single-copy genes, the *UTY* and *DDX3Y* mRNA sequences were completely consistent with the predicted sequences from the genome, and the exon-intron boundaries for the two genes corresponded to those of mouse orthologs, except for a 3 bp slip of an acceptor splicing site in exon 16 of *DDX3Y* (Table 1, Fig. S1i, k and Fig. S4).

Function and structure of the duplicated *SRY*

The *SRY* copy located immediately upstream of *EIF2S3Y* (named *SRY25*) in the euchromatic region was compared with previously reported *SRY* sequences of *T. muenninki*, mouse, and rat in the coding region and the surrounding region, to reveal whether it functions as a sex-determining gene. *SRY25* matches the previously identified *SRY5*, which is a pseudogene with an internal stop codon (Fig. S1q). In the region upstream of *SRY/Sry*, an approximately 3.1 kb sequence in the rat showed more than 80% and 71% similarity with *T. muenninki* *SRY* sequences segmented by long interspersed element (LINE) insertions (Fig. 5a, Fig. S5a). We did

not detect conservation between mice and the other two species, except for an approximately 0.7 kb region just before the start codon (Fig. 5a, Fig. S5a). Downstream of *SRY/Sry*, we observed more than 6 kb of three conserved segments separated by insertions of several retrotransposable elements between mouse and *T. muenninki*, but the conservation was limited to an approximately 1.3 kb region between rat and the other two species (Fig. 5b, Fig. S5b). In the *SRY* surrounding region, sequence similarity was higher between *T. muenninki* and other rodents than between mouse and rat.

We then identified the distal end of a *SRY* duplication unit to identify factors related to *SRY* amplification in *T. muenninki* by comparing the genomic region surrounding *SRY* between *T. muenninki* and other rodents. We found a long-terminal repeat (LTR) at the distal end of a *SRY* duplication unit, which was adjacent to the region with high similarity to mouse upstream of *Eif2s3y* (Fig. 5b).

Discussion

We determined the complete coding sequences of *ZFY*, *EIF2S3Y*, *TSPY*, *UTY*, *DDX3Y*, *USP9Y*, and *RBMY*, but not *UBA1Y*, by RNA-seq, RT-PCR, and 3' RACE (Fig. S1). We observed higher sequence similarity between *T. muenninki* and mouse than between mouse and rat for the single-copy genes *UTY* and *DDX3Y*, but slightly lower similarity between *T. muenninki* and mouse/*Apodemus sylvaticus* (the Wood mouse) than between mouse/*A. sylvaticus* and rat for the multiple-copy genes *EIF2S3Y*, *TSPY*, *USP9Y*, and *RBMY* (Table 1). Based on molecular phylogenetic studies, the *Tokudaia* lineage (sister to *Apodemus*) is more closely related to *Mus* than *Rattus* (Rowe et al. 2008), suggesting that the rate of evolution of the latter four Y-linked genes accelerated in *T. muenninki* relative to other rodents owing to the lineage-specific gene amplification. In the mouse and rat, *ZFY*, *UBA1Y*, *USP9Y*, *TSPY*, and *RBMY* are predominantly expressed in testes, whereas *EIF2S3Y*, *UTY*, and *DDX3Y* are broadly expressed in male tissues (Bellott et al. 2014; GEO Profiles in National Center for Biotechnology Information). Similar expression patterns are observed in *T. muenninki* and other rodents for the Y-linked genes, except for *UBA1Y* (Fig. 1). Our results indicated that the *T. muenninki* genome retained the seven functional Y-linked genes. We also observed several nucleotide substitution sites within the transcript sequences of *ZFY*, *EIF2S3Y*, *TSPY*, *USP9Y*, and *RBMY* (Fig. S1), indicating that mRNAs of each gene are transcribed from multiple gene regions with different sequences. Additionally, some of the *ZFY* transcripts had deletions within the coding region in *T. muenninki*, suggesting that both functional and pseudogene sequences of *ZFY* are expressed in *T. muenninki* testes (Fig. S1a).

We observed excessive amplification of the ancestral Y-linked genes on the Y chromosome in *T. muenninki* (Fig. 2, Fig. 3). Based on a FISH analysis, multiple copies of six genes (*ZFY*, *UBA1Y*, *EIF2S3Y*, *TSPY*, *USP9Y*, and *RBMY*) were mainly localized to the long arm of the Y chromosome. Our previous study demonstrates that the long arm of the Y chromosomes is entirely heterochromatic (Murata et al 2010). *EIF2S3Y* and *USP9Y* copies were located on the

pericentromeric and middle region of Yq, respectively (Fig. 2g–l), while the other four genes and *SRY* were on the pericentromeric and middle regions (Fig. 2a–f, m–r, Murata et al. 2010). Consistent with the FISH results, which showed that *EIF2S3Y* and *USP9Y* copies were not located on the middle and pericentromeric regions, respectively, we detected fewer copies of these two genes than the other multiple-copy genes located in both regions by qPCR (Fig. 3). These results indicated that the Y segment that includes the seven Y-linked genes translocated from Yp to Yq, and the segment was subsequently amplified in the middle and pericentromeric regions of Yq after the loss of *EIF2S3Y* and *USP9Y*, respectively. Y-linked gene amplification is common during mammalian evolution; for example, it has been reported for *ZFY*, *RBMY*, *TSPY*, and *SRY* (Bianchi et al. 1993; Chai et al. 1998; Verkaar et al. 2004; Turner et al. 2007). However, this is the first report of the massive amplification of these seven Y-linked genes in a mammalian species. We also found that the Y-linked genes were translocated to the X chromosome at least two times in *T. muenninki*. *ZFY* and *UBA1Y* were distributed on the distal region of Xq and the centromeric heterochromatin of the X chromosome, respectively (Fig. 2a–f). In the *T. muenninki* X chromosome, the long and short arms consist of the ancestral X chromosome and a fused autosome, respectively (Murata et al. 2012). Dobigny et al. (2004) suggested that constitutive heterochromatin between the X chromosome and the autosome fused to the X prevented the spread of facultative heterochromatinization of the inactive X to an autosome. Therefore, *UBA1Y* would be translocated to the centromeric heterochromatin of the X chromosome and would then be involved in this amplification related to the autosomal fusion in *T. muenninki*.

The order and direction of the five Y-linked genes determined by BAC clone sequencing of *T. muenninki* revealed that the genome structure of the ancestral Y euchromatic region was highly conserved between *T. muenninki* and mouse, except for *SRY* (Fig. 4), suggesting that we identified the ancestral genomic sequences of three multiple-copy genes (*EIF2S3Y*, *TSPY*, and *USP9Y*) in the Y euchromatic region of *T. muenninki*. We observed the

loss of five exons in *EIF2S3Y* and the presence of an internal stop codon in *USP9Y* (Fig. 4, Fig. S1). The deletion in *EIF2S3Y* and a missense mutation in *USP9Y* might occur in each original sequence after duplication. The predicted transcript sequences inferred from the genomic sequence of *TSPY* as well as *EIF2S3Y* and *USP9Y* were different from the actual transcript sequences (Fig. S1), suggesting that the amplified copies of these genes were functional, unlike the original sequences. Thus, our results provide evidence for compensatory effects of the amplified gene copies.

We showed that the *SRY* surrounding region was highly conserved between *T. muenninki* and other rodents, although *SRY25* was a nonfunctional pseudogene with the internal stop codon (Fig. S1q). There is a possibility that *SRY* could be duplicated on the Y chromosome to compensate its function like as other Y-linked genes. However, it is suggested that *SRY* has not retained sex-determining functions (Murata et al. 2010; Kimura et al. 2014).

The 5' and 3' flanking region of *SRY25* showed high similarity to mouse and rat (Fig. 5), suggesting that the gene was duplicated by not transposition but translocation. The 5' flanking region of *SRY25* was segmented by LINE insertion (Fig. 5a yellow colored), indicating that this LINE insertion occurred after *SRY* duplication. In contrast, a LTR were found at the distal end of a conserved *SRY* unit (Fig 5b, pink colored). Furthermore, the downstream of this LTR showed high homology to the 5' flanking region of mouse *Eif2S3y* (Fig. 5b, black colored). These observations indicated that the LTR was boundary for *SRY* duplication unit and the duplication occurred via these repetitive sequences. Several genetic mechanisms can determine the numbers of repeating units in tandem arrays, including gene amplification by processes such as replication slippage, rolling circle amplification, and unequal exchange (Charlesworth et al. 1994). Because replication slippage and rolling circle amplification involve only a short stretch of DNA at a time, these processes probably play only minor roles in the amplification of very long repeat units (Charlesworth et al. 1994). By contrast, the expansion of repeat units may be strongly associated with unequal exchange; this pattern has been inferred based on a

sequence analysis of satellites consisting of higher-order repeat units (Cabot et al. 1993; Warburton et al. 1993). Furthermore, extensive accumulation of retrotransposable elements has been observed in the Y chromosomes of *Mus* species (Eicher et al. 1989). Actually, unequal sister chromatid exchange mediated by retrotransposable elements could have been involved in *SRY* amplification along the heterochromatic block of the X chromosome in Cabrera's vole, *Microtus cabreræ* (Fernández et al. 2002; Marchal et al. 2008). Gene amplification by unequal exchange mediated by retrotransposable elements has also been described for human olfactory receptor genes (Glusman et al. 2000), the primate γ -globin gene (Fitch et al. 1991), and mouse zinc-finger protein genes (Chen et al. 2002). Therefore, the amplification process for the Y-linked genes of *T. muenninki* can be explained by the same mechanism mediated by retrotransposable elements. In support of this hypothesis, we did not detect an intronless sequence generated by retrotransposition in the six multiple Y-linked genes (*ZFY*, *UBA1Y*, *EIF2S3Y*, *TSPY*, *USP9Y*, and *RBM1Y*) by PCR using the genomic DNA of *T. muenninki* (data not shown).

Our results revealed that the seven Y-linked genes were translocated to the X and Y chromosomes fused to an autosomal pair and were subsequently amplified in *T. muenninki*. Additionally, we showed that the amplified copies of the Y-linked genes are functional, unlike the original genes in the ancestral Y euchromatic region. The rescue of Y-linked genes via transposition to other chromosomes is widespread among mammals (Hughes et al. 2015). In the common ancestor of the two XO/XO species, *T. osimensis* and *T. tokunoshimensis*, the Y chromosome was lost after a small euchromatic region containing essential Y-linked genes was translocated to the X chromosome and a few Y-linked gene transposed to autosomes (Kuroiwa et al. 2010; Fig. 6). In the lineage of the *T. muenninki*, the Y chromosome was likely rescued from the risk of loss by sex-autosome fusions (Murata et al. 2012; Fig. 6). Nonetheless, the Y-linked genes were specifically translocated to the other region of the Y chromosome in the *T. muenninki* lineage, suggesting that the translocation was necessary to rescue Y-linked genes

from progressive degeneration in the ancestral Y region. As the Y chromosome diverged from the X chromosome, the correction of replication errors by homologous recombination was not possible, leading to the accumulation of deleterious mutations in Y-linked genes (Bachtrog 2013). Therefore, it is hypothesized that becoming multiple copied genes is one strategy to survive on the non-recombining Y (Bellott et al. 2014). Also, gene duplication coupled with gene conversion between duplicates can potentially maintain gene function by counteracting evolutionary Y decay (Rozen et al. 2003; Charlesworth 2003; Noordam and Repping 2006). A combination of analytical and computer simulation methods also showed that gene conversion effectively maintains the functionality of Y-linked duplicates that have already become fixed, although it has little impact on the probability that duplicates become fixed within a population (Connallon and Clark 2010). Therefore, the functions of *T. muenninki* Y-linked genes may have been maintained not only by translocation, but also by amplification on the sex chromosomes fused to an autosomal pair.

Figure Legends

Figure 1. Expression patterns of eight Y-linked genes in six tissue types from an adult male of *Tokudaia muenninki*.

ACTB was used as a positive control. N.C.: negative control. Marker: AccuRuler 1 kb DNA RTU Ladder (Maestrogen).

Figure 2. FISH mapping of six Y-linked genes on *T. muenninki* male chromosomes.

All six genes, *ZFY* (a–c), *UBA1Y* (d–f), *EIF2S3Y* (g–i), *USP9Y* (j–l), *TSPY* (m–o), and *RBMY* (p–r), were mapped along the long arm of the Y chromosome. The hybridization signals are indicated by arrowheads. Propidium iodide-stained R- (a, b, d, e, g, h, j, k, m, n, p, q) and Hoechst G-banding patterns (c, f, i, l, o, r) are shown. The scale bars represent 10 μ m. (s) A summary of the Y-linked gene locations on the sex chromosomes. *: Murata et al. 2012; **: Murata et al. 2010.

Figure 3. Quantification of the copy numbers of ancestral Y-linked genes in *T. muenninki*.

The copy number of *UTY* (a) or *DDX3Y* (b) was set to 1.0, and the relative values for the other eight genes were estimated in five males by qPCR. The graph shows the means $\pm$ SD.

Figure 4. Comparison of the Y chromosome region of mouse (x-axis) to the ancestral Y chromosome region of *T. muenninki* (y-axis).

(a) Dot-plot alignments of a 391 kb region of mouse (NC_000087: 960,000–1,350,695) to a 372.4 kb region of *T. muenninki* (LC066213: 1–372,407). (b) Dot-plot alignments of an 8.8 kb region of mouse (NC_000087: 2,655,600–2,664,400) to a 15.9 kb region of *T. muenninki* (LC066213: 14,001–29,875). These alignments were generated using a mummerplot script and gnuplot. Aligned segments are represented as dots or lines in the NUCmer alignment. The positions of the genes on the mouse Y chromosome are shown on the dot-plot alignments. (c)

Comparison of Y gene order between the mouse and *T. muenninki*. *SRY* was located approximately 1.2 Mb downstream of *USP9Y* in the mouse genome.

Figure 5. Sequence conservation of the regions upstream (a) and downstream (b) of *SRY* among *T. muenninki*, mouse, and rat.

Percentages between bars indicate sequence similarity between the segments. Species-specific regions, except for retrotransposable elements, are colored white. TMU: *T. muenninki*.

Figure 6. Evolution of the X and Y chromosomes and genes in genus *Tokudaia*.

The lineage of *T. muenninki* was firstly diverged in genus *Tokudaia* (Murata et al. 2010). Sex-autosome fusion (Murata et al. 2012), Y-to-Y and Y-to-X gene translocations occurred in the lineage. In the common ancestor of *T. osimensis* and *T. tokunoshimensis*, a small euchromatic region of Y chromosome translocated to a distal end of X chromosome, and several Y genes transposed to autosome (Kuroiwa et al. 2010).

Supplementary Figure Legends

Figure S1. Alignment of the nucleotide and amino acid sequences of nine Y-linked genes between *T. muenninki* and other rodents.

The coding region from the mRNA and/or genomic sequences of *ZFY* (a, b), *UBAIY* (c, d), *EIF2S3Y* (e, f), *TSPY* (g, h), *UTY* (i, j), *DDX3Y* (k, l), *USP9Y* (m, n), *RBMV* (o, p), and *SRY* (q) in *T. muenninki* were compared with orthologous sequences in mouse, rat, *T. osimensis*, and/or *A. sylvaticus*. Identical residues are indicated by dots. TMU: *T. muenninki*. The deletion and the internal stop codon are colored in gray (a) and pink (m), respectively. Asterisks show the stop codon sites in the amino acid sequences (b, d, f, h, j, l, n, p). Single nucleotide and amino acid polymorphisms within *T. muenninki* transcripts of each gene are shown in red letters (e–h, m–p). Different nucleotide sites of genomic sequences from the mRNAs of *T. muenninki* are shown in blue letters (e, g, m, o). (q) The start codons and the stop codons are colored in blue. *SRY25* is the genomic sequence observed upstream of *EIF2S3Y* in this study. *SRY1–24* sequences of *T. muenninki* are reported by Murata et al. 2010, of which only three (*SRY1–3*) have the conserved coding sequences. The repeat sequence was partially skipped in the mouse (600 bp), and the skipped site is shown by double slashes.

Figure S2. Neighbor-joining trees constructed with each X/Y gene sequence of *T. muenninki* and other rodents.

Number above and below branches refer to bootstrap percentages. The values less than 70% are not shown. TMU: *T. muenninki*; MMU: *M. musculus*; RNO: *R. norvegicus*; ASY: *A. sylvaticus*

Figure S3. Chromosomal location of two X-linked genes in *T. muenninki*.

A fluorescent *in situ* hybridization (FISH) analysis showed that *ZFX* (a, b) and *UBA1* (c, d) are distributed on the middle and pericentromeric regions of Xq, respectively. The hybridization signals are indicated by arrowheads. R- (a, c) and G-banding (b, d) patterns are shown. The

511 scale bars represent 10 μ m.

512

513 **Figure S4. Nucleotide alignment of the acceptor splicing site within intron 15 of *DDX3Y***
514 **between *T. muenninki* and mouse.**

515 The intronic sites are colored in gray for each species. Arrows show the boundary between an
516 exon and intron. Asterisks show the identical sites between two sequences.

517

518 **Figure S5. Nucleotide sequence alignment in the conserved upstream (a) and downstream**
519 **(b) regions of *SRY* among *T. muenninki*, mouse, and rat.**

520 Identical residues are indicated by dots. The skipped site is shown by double slashes. TMU: *T.*
521 *muenninki*.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

The authors thank F. Yamada, T. Jogahara, C. Koshimoto, T. Mitani, and K. Nakata for their efforts in capturing the animals and the Fujii Memorial Institute of Medical Sciences, Tokushima University, for support with fluorescence microscopy. This work was supported by JSPS KAKENHI Grant Numbers 23870023 and 26870410.

Authors' Contributions

CM conceived of and designed the study, performed molecular and cytogenetic experiments and all data analysis, and drafted the manuscript. YK conducted RNA-seq. II participated in total RNA and BAC DNA extraction and commented on the manuscript. AK conceived of and designed the study, and participated in drafting the manuscript. All authors read and approved of the final manuscript.

537 **References**

- 538 Arakawa Y, Nishida-Umehara C, Matsuda Y, Sutou S, Suzuki H (2002) X-chromosomal
539 localization of mammalian Y-linked genes in two XO species of the Ryukyu spiny rat.
540 Cytogenet. Genome Res. 99:303-309
- 541 Bachtrog D (2013) Y-chromosome evolution: emerging insights into processes of
542 Y-chromosome degeneration. Nat Rev Genet 14:113-124
- 543 Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho TJ, Koutseva N, Zaghlul S,
544 Graves T, Rock S, followed by et al (2014) Mammalian Y chromosomes retain widely
545 expressed dosage-sensitive regulators. Nature 508:494-499
- 546 Bianchi NO, Bianchi MS, Bailliet G, de la Chapelle A. (1993) Characterization and sequencing
547 of the sex determining region Y gene (*Sry*) in *Akodon* (Cricetidae) species with sex
548 reversed females. Chromosoma 102:389-395
- 549 Cabot EL, Doshi P, Wu ML, Wu CI (1993) Population genetics of tandem repeats in centromeric
550 heterochromatin: unequal crossing over and chromosomal divergence at the *Responder*
551 locus of *Drosophila melanogaster*. Genetics 135:477-487
- 552 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009)
553 BLAST+: architecture and applications. BMC Bioinformatics. 2009 10:421
- 554 Chai NN, Zhou H, Hernandez J, Najmabadi H, Bhasin S, Yen PH (1998) Structure and

555 organization of the *RBMY* genes on the human Y chromosome: transposition and
 556 amplification of an ancestral autosomal *hnRNPG* gene. *Genomics* 49:283-289
 557 Charlesworth B (2003) The organization and evolution of the human Y chromosome. *Genome*
 558 *Biol* 4:226
 559 Charlesworth B, Sniegowski P, Stephan W (1994) The evolutionary dynamics of repetitive DNA
 560 in eukaryotes. *Nature* 371:215-220
 561 Chen HH, Liu TY, Huang CJ, Choo KB (2002) Generation of two homologous and intronless
 562 zinc-finger protein genes, *zfp352* and *zfp353*, with different expression patterns by
 563 retrotransposition. *Genomics* 79:18-23
 564 Connallon T, Clark AG (2010) Gene duplication, gene conversion and the evolution of the Y
 565 chromosome. *Genetics* 186:277-286
 566 Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A, Waters PD, Grutzner F,
 567 Kaessmann H (2014) Origins and functional evolution of Y chromosomes across
 568 mammals. *Nature* 508:488-493
 569 Dobigny G, Ozouf-Costaz C, Bonillo C, Volobouev V (2004) Viability of X-autosome
 570 translocations in mammals: an epigenomic hypothesis from a rodent case-study.
 571 *Chromosoma* 113:34-41
 572 Eicher EM, Hutchison KW, Phillips SJ, Tucker PK, Lee BK (1989) A repeated segment on the

573 mouse Y chromosome is composed of retroviral-related, Y-enriched and Y-specific
 574 sequences. Genetics 122:181-92
 575 Fernandez R, Barragan MJ, Bullejos M, Marchal JA, Martinez S, Diaz de la Guardia R, Sanchez
 576 A (2002) Mapping the *SRY* gene in *Microtus cabreræ*: a vole species with multiple *SRY*
 577 copies in males and females. Genome 45:600-603
 578 Fitch DH, Bailey WJ, Tagle DA, Goodman M, Sieu L, Slightom JL (1991) Duplication of the
 579 gamma-globin gene mediated by L1 long interspersed repetitive elements in an early
 580 ancestor of simian primates. Proc Natl Acad Sci U S A 88:7396-7400
 581 Glusman G, Sosinsky A, Ben-Asher E, Avidan N, Sonkin D, Bahar A, Rosenthal A, Clifton S,
 582 Roe B, Ferraz C, Demaille J, Lancet D (2000) Sequence, structure, and evolution of a
 583 complete human olfactory receptor gene cluster. Genomics 63:227-45
 584 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
 585 Raychowdhury R, Zeng Q followed by et al (2011) Full-length transcriptome assembly
 586 from RNA-Seq data without a reference genome. Nat Biotechnol 29:644-652
 587 Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D,
 588 Li B, Lieber M followed by et al (2013) *De novo* transcript sequence reconstruction from
 589 RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc
 590 8:1494-1512

591 Honda T, Suzuki H, Itoh M (1977) An unusual sex chromosome constitution found in the
 592 Amami spinous country-rat, *Tokudaia osimensis osimensis*. Jpn J Genet 52:247-249
 593 Honda T, Suzuki H, Itoh M, Hayashi K (1978) Karyotypical differences of the Amami spinous
 594 country-rats, *Tokudaia osimensis osimensis* obtained from two neighbouring islands. Jpn.
 595 J. Genet. 53:297-299
 596 Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Graves T, Fulton RS, Dugan S, Ding Y,
 597 Buhay CJ, Kremitzki C, followed by et al (2012) Strict evolutionary conservation
 598 followed rapid gene loss on human and rhesus Y chromosomes. Nature 483:82-86
 599 Hughes JF, Skaletsky H, Koutseva N, Pyntikova T, Page DC (2015) Sex
 600 chromosome-to-autosome transposition events counter Y-chromosome gene loss in
 601 mammals. Genome Biol 16:104.
 602 Just W, Rau W, Vogel W, Akhverdian M, Fredga K, Graves JA, Lyapunova E (1995) Absence of
 603 Sry in species of the vole *Ellobius*. Nat Genet 11:117-118
 604 Kimura R, Murata C, Kuroki Y, Kuroiwa A (2014) Mutations in the testis-specific enhancer of
 605 *SOX9* in the *SRY* independent sex-determining mechanism in the genus *Tokudaia*. PLoS
 606 One 9:e108779
 607 Kobayashi T, Yamada F, Hashimoto T, Abe S, Matsuda Y, Kuroiwa A (2008) Centromere
 608 repositioning in the X chromosome of XO/XO mammals, Ryukyu spiny rat. Chromosome

609 Res 16:587-593

610 Kuroiwa A, Ishiguchi Y, Yamada F, Abe S, Matsuda Y (2010) The process of a Y-loss event in
611 an XO/XO mammal, the Ryukyu spiny rat. *Chromosoma* 119:519-526

612 Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL (2004)
613 Versatile and open software for comparing large genomes. *Genome Biol* 5:R12

614 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F,
615 Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W
616 and Clustal X version 2.0. *Bioinformatics* 23:2947-2948

617 Li H and Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler
618 Transform. *Bioinformatics* 25:1754-60

619 Marchal JA, Acosta MJ, Bullejos M, Diaz de la Guardia R, Sanchez A (2008) Origin and spread
620 of the *SRY* gene on the X and Y chromosomes of the rodent *Microtus cabreræ*: Role of
621 L1 elements. *Genomics* 91:142-151

622 Matsuda Y, Chapman VM (1995) Application of fluorescence in situ hybridization in genome
623 analysis of the mouse. *Electrophoresis* 16:261-272

624 Matsuda Y, Harada YN, Natsuume-Sakai S, Lee K, Shiomi T, Chapman VM (1992) Location of
625 the mouse complement factor H gene (*cfh*) by FISH analysis and replication R-banding.
626 *Cell Genet* 61:282-285

627 Murata C, Kuroki Y, Imoto I, Tsukahara M, Ikejiri N, Kuroiwa A (2015) Initiation of
 628 recombination suppression and PAR formation during the early stages of neo-sex
 629 chromosome differentiation in the Okinawa spiny rat, *Tokudaia muenninki*. BMC Evol
 630 Biol 15:234

631 Murata C, Yamada F, Kawauchi N, Matsuda Y, Kuroiwa A (2010) Multiple copies of *SRY* on the
 632 large Y chromosome of the Okinawa spiny rat, *Tokudaia muenninki*. Chromosome Res
 633 18:623-634

634 Murata C, Yamada F, Kawauchi N, Matsuda Y, Kuroiwa A (2012) The Y chromosome of the
 635 Okinawa spiny rat, *Tokudaia muenninki*, was rescued through fusion with an autosome.
 636 Chromosome Res 20:111-125

637 Noordam MJ, Repping S (2006) The human Y chromosome: a masculine chromosome. Curr
 638 Opin Genet Dev 16:225-232

639 Paria N, Raudsepp T, Pearks Wilkerson AJ, O'Brien PC, Ferguson-Smith MA, Love CC, Arnold
 640 C, Rakestraw P, Murphy WJ, Chowdhary BP. (2011) A gene catalogue of the euchromatic
 641 male-specific region of the horse Y chromosome: comparison with human and other
 642 mammals. PLoS One 6:e21374

643 Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP
 644 (2011) Integrative Genomics Viewer. Nature Biotechnology 29:24-26

645 Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR,
 646 Burrows C, Bird CP, followed by et al (2005) The DNA sequence of the human X
 647 chromosome. *Nature* 434:325-337

648 Rowe KC, Reno ML, Richmond DM, Adkins RM, Steppan SJ (2008) Pliocene colonization and
 649 adaptive radiations in Australia and New Guinea (Sahul): multilocus systematics of the
 650 old endemic rodents (Muroidea: Murinae). *Mol Phylogenet Evol* 47:84-101

651 Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Page
 652 DC (2003) Abundant gene conversion between arms of palindromes in human and ape Y
 653 chromosomes. *Nature* 423:873-876

654 Sekido R, Lovell-Badge R (2008) Sex determination involves synergistic action of SRY and
 655 SF1 on a specific *Sox9* enhancer. *Nature* 453:930-934

656 Soullier S, Hanni C, Catzeflis F, Berta P, Laudet V (1998) Male sex determination in the spiny
 657 rat *Tokudaia osimensis* (Rodentia:Muridae) is not Sry dependent. *Mamm Genome*
 658 9:590-592

659 Sutou S, Mitsui Y, Tsuchiya K (2001) Sex determination without the Y chromosome in two
 660 Japanese rodents *Tokudaia osimensis osimensis* and *Tokudaia osimensis* spp. *Mamm*
 661 *Genome* 12:17-21

662 Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2008) MEGA5: Molecular

663 Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and
 664 Maximum Parsimony Methods. *Mol Biol Evol* 28:2731-2739
 665 Tsuchiya K, Wakana S, Suzuki H, Hattori S, Hayashi Y (1989) "Taxonomic study of *Tokudaia*
 666 (Rodentia: Muridae): I. Genetic differentiation. *Memoirs Nat Sci Museum, Tokyo*
 667 22:227-234
 668 Turner ME, Martin C, Martins AS, Dunmire J, Farkas J, Ely DL, Milsted A (2007) Genomic and
 669 expression analysis of multiple *Sry* loci from a single *Rattus norvegicus* Y chromosome.
 670 *BMC Genet* 8-11
 671 Verkaar EL, Zijlstra C, van't Veld EM, Boutaga K, van Boxtel DC, Lenstra JA (2004)
 672 Organization and concerted evolution of the ampliconic Y-chromosomal *TSPY* genes
 673 from cattle. *Genomics* 84:468-474
 674 Vogel W, Jainta S, Rau W, Geerkens C, Baumstark A, Correa-Cerro LS, Ebenhoch C, Just W
 675 (1998) Sex determination in *Ellobius lutescens*: the story of an enigma. *Cytogenet Cell*
 676 *Genet* 80:214-221
 677 Warburton PE, Wayne JS, Willard HF (1993) Nonrandom localization of recombination events in
 678 human alpha satellite repeat unit variants: implications for higher-order structural
 679 characteristics within centromeric heterochromatin. *Mol Cell Biol* 13:6520-6529
 680 Yamada F, Kawauchi N, Nakata K, Abe S, Kotaka N, Takashima A, Murata C, Kuroiwa A

681 (2010) Rediscovery after thirty years since the last capture of the critically endangered
682 Okinawa spiny rat *Tokudaia muenninki* in the northern part of Okinawa Island. Mammal
683 Study 35:243-255

■ : *UTY**, *DDX3Y**, *USP9Y*

■ : *ZFY*, *UBA1Y*, *EIF2S3Y*, *TSPY*, *SRY***, *RBMY*

■ : *ZFY*, *UBA1Y*, *TSPY*, *USP9Y*, *SRY***, *RBMY*

■ : *UBA1Y*

■ : *ZFY*

a

Copy number in a male genome

0 10 20 30 40 50 60 70 80 90

b

Copy number in a male genome

0 10 20 30 40 50 60 70 80 90

a

ZFY-TMU	W TGGATGAAG	ATGAAATTGA	ATTGACCCCA	CAAGAAGGAA	ACTCATT K TT	TGATGG A MTA	G-----	---TACAAAT	GGATGGTGAC	CAGAATATTG
Zfy1-mouse	A.....			G.....A..	.G.....C..	A..	.GAGCTGATG	CAG....C..	A.....	T.G...
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	TGGAAGTACA	AGAAACTGTT	TTTACAAC T A	ATT C AGATGT	AACTGTGCAT	AATTTT T W T C	S TGATGATCC	AGACCCACTT	ATAATTCAAG	ATGTTATTGA
Zfy1-mouse			...TT.G...			GT..	C....A....	..G.T..G..		
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	AGCTGTTCTT	ATTGAAGATG	TTCAC T TCTC	ACATATTTTA	GAGGAAACAG	ATATATCTGA	CAAT G Y C ATT	ATT C CTGA S C	AAGTGCTCGA	TTTAGATAGA
Zfy1-mouse	..A.....		GT..		..A.....		T....	G....	T...A..	G..C.
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	GCTGAAGAAG	TGTCTTTAGC	ACAGTTCTCA	ATTCCAGATA	TTTTAGCTTC	CAGTATTACA	TCAAC C W C AT	TGAGTATGCC	TGAATATGTC	TTGATGAGTG
Zfy1-mouse	..A.....		T..	C..	A.A..	TG.....G	T...T..	...C.....	C.....	
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	AAGCTATACA	GGTGTCTGAT	GTAGGACAAA	TTGA A S A AGT	GTTTCATGAT	AGCTTAGTAG	AAACAGAA G T	CATCACTGAT	CCTCTGACAG	CC R ACATTTC
Zfy1-mouse		T.....	TT	C....	.A.....	...C.T....			...A.A....	.TG...C...
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	AGAAATACTG	GTAGCAGATG	GTGCTTCTGA	AGCAGTCTTA	GATTCTAGTG	GGATGCCTCT	GGAGCAGCAA	GATGATGCCA	AAATCAACTG	TGAGGATTAT
Zfy1-mouse	...T.....	T	...T.....		C....			...A...A..		
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	CTTATGATTT	CTTTGGATGA	TGCTAGCAAA	ACAGAT Y ATG	AAGGTTCC T C	TGAAGTTACC	ATGAATGCAG	AGTCAGAAAC	TGATCCTTAT	AAATTGGATG
Zfy1-mouse	..G.....G.		AC.....	G....CT..				G...	T...C..	
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	GAACATCTCC	TGA A R T TCTC	AAG K TG C ACA	TTTTTAAAGC	TGGC C MTGAA	GAGGTTGATG	TAGGAGAAAC	TATA S ATGTG	GTAGAGAGTG	AGAATGAAAA
Zfy1-mouse	A.G.....	A...G..A..	...G..TG..	..C.....	..A.TCA...	.T..A...AT		...C.C.CA		...CCA...
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	CGGC A K G AA	GCTGAAGTCA	TTGATCAGAG	C AR C AGC R TT	CAT G K T CCCA	GAG---ACAT	TTATATTCCA	GCCAGTGATT	CTCAGAAGGA	AGAAGA---T
Zfy1-mouse	T....T...		CA.....	..CT...A..	.G..T.....	...TCA....	GT..		A.....	AGA.
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	ACTAAAGTAA	TTGTAGGAGA	TGAAGATGCT	GGTGGCACAG	CTGCAGATAC	TCCTGAGCAT	GGACAACAGA	TGGATGACAG	TGAAATAAAA	GCAGCTTTCC
Zfy1-mouse	...G.....						.AG.....	T...		
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	TGCCTATTGC	ATGGGCGGCA	ACTTATG---	-----	-----	-----	-----	-----	-----	-----
Zfy1-mouse	.A.....	...A.A...G.	G.....ATA	ATAATTCTGA	TGAAATTGAA	GACCAGAATG	TCACTGCCAG	TGCTCTCTTA	AACCAAGATG	AGTCTGGTGG
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	-----	-----	-----	-----	-----	-----	-----CA	GTAATTGTTG	CTCCTGATGG	ACAAACTTTA
Zfy1-mouse	CCTTGACAGA	GTACCAAAAC	AAAAATCAAA	AAAGAAAAAA	AGACCTGAAT	CCAAACAGTA	CCAGTCAG..	A..T.....		
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	ATAGTCTATC	CTTGCA T GTT	TTGT G S A AAA	ATATTTAAGA	CCAAGAGGTT	TT T R A AAAGA	M AC A TT A AAA	ACCATCCTGA	ATACCTTGCT	AAGAAAAAAT
Zfy1-mouse	CGT.....		GG...	.A.....	A.....	...G.....	C.T.CA...			...T.....
ZFY-TOS	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
ZFY-TMU	ATCACTGTAC	CGACTGTGAT	TACAGTACCA	A S A AGAAGAT	AAGCTTACAT	AATCACATGG	A R A GCCACAA	GCTAACCATT	R AGACAGAAA	AGACCACAGA
Zfy1-mouse		T..G.....		..C.....		T.	.G.....		A.....	T..
ZFY-TOS	-----	-----	-----	-----	T.		.G.....		A.....	
ZFY-TMU	AT K T R ATGAC	TGTGAGAAGC	ATCTTTCTCA	TGCTGGGACT	TTGTGTACTC	ACAA A CA R T	GCATAAGAA	AAAGGAGTCA	ACAAAACATA	TAAGTGTAAG
Zfy1-mouse	..G.G.....	...AG...A				A.	C....		G	
ZFY-TOS	..G.A.....					A.			A....	
ZFY-TMU	TTCTGTGACT	ATGAAACAGC	TGAACAG A Y A	TTATTGAATC	ACCAC M ATTT	GGTGGTCCAC	GGCAAGAAAT	TTCTCAAGT	TT R TGTAGAA	TGTGGCAAAG
Zfy1-mouse			C..		CT...		A.G.....	CA..	..G..G....	T....
ZFY-TOS	-----	C..	C..		C..C				..G.....	
ZFY-TMU	R TTTCC G K R	CCCATCAGAG	CTCAAAAAGC	ACATAAGAGT	T CAC A CTGGA	GAGAAGCCCT	ATGAATGTCA	GTATTGTGAG	TAC A M G TCTG	CAA A CTCTTC
Zfy1-mouse	G.....T.A	CA		C....	A..				...A.....	..G.....
ZFY-TOS	G.....T.A			..G....C.	CT.....				...A.....	
ZFY-TMU	TA A CTTGAAA	ACACATATAA	AA A CTAAGCA	TAATAAGGAG	ATACCACTGA	AGTGTGGCAT	CTGTCT Y CTG	ACTTT Y TCAG	A M A C Y AAAGA	GGCTCAGCAA
Zfy1-mouse	C.....	..T.....	..T.....	..G...A..			C...	C...	..T.AC...	
ZFY-TOS			G.....				...C..C...	C....	..C..C....	
ZFY-TMU	CATGCTGTTG	TACACCAAGA	AAG C MS A ACA	CAT C M G TGTT	CATATTGCAA	CC R TAAGAGT	TCAA A CTCA V	GTGATTTAA A	G Y G ACATATA	ATTTCAGTTC
Zfy1-mouse	C...C	.G.....	...AG....	...AA....	..C.....	..A.....	A		..C...C...	C....
ZFY-TOS			...AAG...	T...A....		..A.....	A	T.....	..C.....	
ZFY-TMU	ACACAAAGGA	CTATCCTCAC	AAGTG K GACA	TGT K CAGCAA	AGGGTTTCAT	AGG C Y T TCAG	AACTCA A R A A	GCATGTGGCT	AC Y CM A AAAA	GTA A AAAAAAT
Zfy1-mouse	C	G.....T	..A..T...	..G.....	..A.....	...C.....	G..		..C.AT...	
ZFY-TOS			T....	..G.....		...C..C	G..		..C.A.....	T.....
ZFY-TMU	GCAC C A R T R T	AGACACTGTG	ACTTTGAGAG	TCCAGAT C MA	TTTCTGCTGA	GT C G C MA T AT	TCTCTCAGTT	CACACAAAGA	ATGTTCCATT	TAAGTG Y AAG
Zfy1-mouse	A.G.		A.T..	C..	T.	..A.C....	C..			C....T...
ZFY-TOS	A.G.		CG	..C.....		...C....				T...
ZFY-TMU	AG R TGTAAAA	AGGGATTTAG	ACAACAGTGT	GAGCTTCAAA	ATCATATGAA	GACCCATAGT	GGCAGGAAA S	TGTATCATTG	TGAATACTGT	G A R W ATAGCA
Zfy1-mouse	..A.....	...A...CA			CG.....	C...	A..C.A...G	.C.....G..	...G.....	..AT.....
ZFY-TOS	..A.....					C...	G		...G.....	..AT.....

ZFY-TMU CCACAGATAC CTCAGGCTTT AAGC**RG**CAS**G** TTATCTCCAT TCATA**Y**GAAA GACTATCCT**M** ACCGCTGTGA GTAT**Y**TGCAGG AAAGGATTCT GGAGACCCTC
Zfy1-mouse ...A...G.T... ..GC..C.C... ..C... ..A... ..C.TC...A.C
ZFY-TOSTG..C.C... ..C... ..C... ..C... ..C.....

ZFY-TMU GGAAAAGAR**T** AAACACATAA TACGACATCA TAAAGAAGTT GSCGTG**Y**CCT AA
Zfy1-mouseA. C.....G..... ---... .G.C..C... ..
ZFY-TOST.G..... .G...C... ..

b

ZFY-TMU-mRNA1 MDEDEIELTP QEGNS**F**FD**G**L ----VQMDGD QNIVEVQETV FTNSDVTVH NF**Y**PDDPDPL IIQDVIEAVL IEDVHFHSHIL EETDISDN**A**I IPDQVLDLDR
ZFY-TMU-mRNA2**L**..**I** ----..... ..**FR**..... ..**V**..**E**.....
ZFY-TMU-mRNA3**L**..**I** ----..... ..**FR**..... ..**V**..**E**.....
Zfy1-mouse E.EK....I GADA.H..S. .IV..... .LA..... .V..N.GSVN..C....V. ..E...N.GT

ZFY-TMU-mRNA1 AEEVSLAQFS IPDILASSIT **ST****T**L**S**MP**E**YV LMSEAIQVSD VGQIEQVFHD SLVETEVIDT PLTAN**I**SEIL VADGASEAVL DSSGMPL**E**QQ DDAKINCEDY
ZFY-TMU-mRNA2**S**..... ..**D**..... ..**D**.....
ZFY-TMU-mRNA3**S**..... ..**D**.....
Zfy1-mouseLT.G.. .S.T...H.H... .HF...I..I..DT.D.. ...CV..... .ND.....

ZFY-TMU-mRNA1 LMISLDD**A**SK TD**Y**EGSSEVT MNAESETDPY KLDGTSPEVL KVHIFKAGPE EVDVGETI**H**V VESENENGKE AEVIDQSSSI **H**VPRD-IYIP ASDSQKEE-D
ZFY-TMU-mRNA2**N**..**H**..... ..**D**..... ..**D**..... ..**G**...-... ..**V**..**E**.....
ZFY-TMU-mRNA3**N**..**H**..... ..**D**..... ..**N**..... ..**G**...-... ..**V**..**E**.....
Zfy1-mouse ..M...EP.. A.L..... ..G..SS ..EA....I ..C.L..DS. VDEL....ATK..N. ...T...T.. R...VN..MSE.

ZFY-TMU-mRNA1 TKVIVGDEDA GGT**V**ADTPEH GQQMDDSEIK AAF**L**PIAWAA TY----- ----- ----- ----- -----A VIVAPDGQTL
ZFY-TMU-mRNA2**A**..... ..----- ----- ----- -----
ZFY-TMU-mRNA3**A**..... ..----- ----- ----- -----
Zfy1-mouse .E..... ..A..... E...V....T. A.DNNSDEIE DQNV**T**ASALL NQDESGGLDR VPKQKSKKKK RPESKQYQS. IF.....

ZFY-TMU-mRNA1 IVYPCMF**C**GK IFKTKRFLKR HIKNHPEYLA KKKYHCTDCD YSTNKKISLH NHMESHK**L**TI KTEK**T**TE**F**DD CEKHL**S**HAGT LCTHK**T**M**H**KE KGVNKTYKCK
ZFY-TMU-mRNA2**A**..... ..**CN**..... ..**V**..**E**.....
ZFY-TMU-mRNA3**A**..... ..**C**..... ..**V**..**E**.....
Zfy1-mouse R..... K..... .T..... N.....E..C.. .R.N..... .T.....C...

ZFY-TMU-mRNA1 FCDYETAEQ**I** LLNHH**H**LVVH GKKFPQVCVE CGK**G****F****R****H**PSE LKKHIRVHTG EKP**Y**ECQYCE YKSANSSNLK THIKTKHNKE IPLKCGICLL TFS**D**TKEAQQ
ZFY-TMU-mRNA2**T**..... ..**N**..... ..**D**..**R**.....
ZFY-TMU-mRNA3**T**..... ..**N**..... ..**D**..**R**.....
Zfy1-mouseTL.... R...HI.G.A.....D.....S..S... ..N.....

ZFY-TMU-mRNA1 HAVVHQES**R**T HQCSYCNHKS SNS**S**DLKRHI ISVHTKDYPH KCD**M**SKGFH RPSELKKHVA **T**HKSKKM**H**Q**C** RHCD**F**ESP**D**F FLLSRHILSV HTKNVPFKCK
ZFY-TMU-mRNA2**T**..... ..**R**..... ..**F**..... ..**P**.....**Y**.....**Q**.....
ZFY-TMU-mRNA3**T**..**P**..... ..**R**..... ..**F**..... ..**P**.....**Y**.....**Q**.....
Zfy1-mouse ...L..... ..H..... ..A.....N.....H...A

ZFY-TMU-mRNA1 RCKKGFRQQC ELQN**H**MK**T**HS **RR**KVYHCEYC EYST**T**D**I**SGF **KR****H**VISI**H**TK DYPHRCEYCR KGF**R**RPSE**K****N** KHIMRHHKE**A** **G**VP*
ZFY-TMU-mRNA2**G**..... ..**T**..... ..**W**..... ..**V**.....
ZFY-TMU-mRNA3**G**..... ..**T**..... ..**Q**..... ..**W**.....**S**.....**I**.....**V**..**A**..**S**..
Zfy1-mouse ...E.Q... ..T..... S...Q.... ..K.A... ..S.DF.KQ.....-V .L..

C

UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	ATGTCCAGCT	CAGTCTGTGC	CAAGAAACGC	AAAGTGTCTG	GACCTGACTC	TAGTCTGGAT	TCTAGTTGGT	CTCCTACCTA	TTCTGTGATG	TTTGGCGTGC	
UBA1Y-rat	ATGTCCAGCT	CGGTCTGTGC	CAAGAAACGC	AGAGTGTCTG	GACCTGACTC	TGAGGTGGAT	TCTAGTTGGC	CTTCTACCCA	TTCTGTGATG	TTTGGCGCAC	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	CCCCAGGGCC	AACCAACGAA	ATGTCAAAAA	ATAAGGAAAT	GGACATAGAT	GAAAGCCTTT	ACTCCCGCCA	GCTGTATGTA	CTCGGCCATG	AGGCAATGAA	
UBA1Y-rat	CCCCAGGGCC	AACACAGCGGA	ATGTCAAAAA	ACAAGGAAAT	GGATATAGAT	GAAAGCCTTT	ACTCCCGCCA	GCTGTATGTA	CTAGGTCATG	AGGCAATGAA	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	GCATCTCCAG	GCTTCCAGTG	TACTGATATC	AGGCCTGCAG	GGTCTGGGTG	TGGAAATTGC	CAAGAATATC	ATCCTTGGTG	GGGTCAAGGC	TGTCACTCTC	
UBA1Y-rat	ACATCTCCAG	ACGTCCAGTG	TACTGATATC	AGGCCTGCAG	GGTCTGGGTG	TGGAAATTGC	CAAGAATATC	ATCCTTGGTG	GGGTCAAGGC	TGTCACTCTC	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	CATGATCAGG	GCATTGCCCA	GTGGGCTGAT	CTGTCTTCCC	AGTTCTGCCT	ACGTGAGGAA	GATATTGGTA	AAAACCGAGC	TGAGATATCC	CAACCACGCC	
UBA1Y-rat	CATGACCAGG	GCACTGCCCA	GTGGGCTGAT	CTGTCTTCCC	AGTTCTACCT	GCATGAGGAA	GATATTGGCA	AAAATCGAGC	TGAGGTATCC	CAACCACGCC	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	TTGCTGAACT	CAACAGTTAC	GTTCTGTGT	TTGCCACAC	TGGACCTCTA	ATTGAGGAGT	TCCTTAGTGG	TTTTCAGGTA	GTGGTTCCTA	CCAACACTCC	
UBA1Y-rat	TTGCTGAACT	CAACAGTTAT	GTTCTGTGTC	ACACTTACAC	TGGACCTCTA	GTTGACGACT	TCCTTAGTGG	TTTTCAGGTG	GTGGTTCCTA	CCAACACTCC	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	TTTGGAAATAT	CAGCTGCAGG	TGGGTGAATT	CTGCCATAGC	CACGGAATCA	AGCTGGTAGT	AGCAGACACT	CGGGGCCCTAG	TTGGACAAC	CTTCTGTGAC	
UBA1Y-rat	TTTGGAAATAT	CAGCTGCAGG	TGGGTGAATT	CTGCCATAGC	CATGGAATCA	AGCTGGTAGT	AGCAGACACT	CGGGGCCCTAG	TTGGACAAC	CTTCTGTGAC	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	TTTGAGAGAG	AAATGATTCT	CACTGATTCA	AATGGAGAAC	AGCCACTCAG	TGCCATGGTT	TCAATGATCA	CTAAGGAGAA	TCCAGGGATT	GTCACTTGCC	
UBA1Y-rat	TTTGAGAGAG	AAATGATTCT	CACTGATGCA	AATGGAGAAC	AGCCACTCAG	TGCTATGGTT	TCAATGATCA	CTAAGGAGAA	TCCAGGGATT	GTCACTTGCT	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	TGGAGGACTC	CCGGCATGGA	TTTGAAAGTG	GTGACTTTAT	CTCTTTCACA	GAAGTTTCAGG	GCATGAGTGA	ACTCAATGGC	ATTGGTCCCA	TAGAGATCAA	
UBA1Y-rat	TGGAGGAAAC	CCGGCATGGA	TTTGAAAGTG	GTGACTTTGT	CTCTTTCACA	GAAGTTCAAG	GCATGAGTGA	ACTGAATGGC	ATTGGTCCCTA	TGGAGATCAA	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	AGTTCTGGGT	CCCTATACCT	TTAGTATCTG	TGATACCTCC	AGCTTCTCTG	AGTACATCCG	TGGAGGCATT	GTCAGTCAAG	TGAAAGTACC	TCGGAAGATC	
UBA1Y-rat	AGTTCTGGGT	CCCTATTCCT	TTAGTATCTG	TGATACCTCC	AGCTTCTCTG	AGTACACCCG	TGGAGGCATT	GTCAGTCAAG	TGAAAGTATC	TCAGAAGATC	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	AATTTTAAAC	CCCTACTTGC	CTCGCTGGCA	GAGCCAGAGT	TTGTGGTCAC	AGATTTTGCT	AAGTGTCTGT	ATCCTGCTCA	GCTCCACATT	GGCTTCCAGG	
UBA1Y-rat	AGTTTAAAT	CCCTAGTTGC	CTCGCTGGCA	GAGCCAGAGT	TTGTGATAAC	AGATTTTGCT	AAGTGTCTGT	GCCCTGTCTA	GCTCCACATT	GGCTTCCAGG	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	CCCTGCATCA	GTTCGTACT	CAGCACAGCA	GGCCTCCTCG	GCCCCATAAT	GAGGAGGATG	CTGAGGAAAT	GGTGACCTTA	GCCCAATCTG	TGAATGCTCA	
UBA1Y-rat	CCCTGCATCA	GTTCTGTACT	CAGCACAGCA	GGCCTCCTCG	GCCCCATAAT	GAGGAGGATG	CTGCAGAAAT	GGTGACCTTA	GCACAGGCTG	TGAATGCTCA	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	AGCTTTGCCA	GCAGTTCAGC	AGGATTGCCT	GGACATCGAC	CTTATCCGGA	AGTTGGCCTA	TGTAGCAGCT	GGGGACCTGG	CACCCATGAA	TGCCTTTTTT	
UBA1Y-rat	ATCTTTGCCA	GCAGTTCAGC	AGGATTGCCT	AGATATAGAC	CTCATCCGGA	AGTTGGCCTA	TGTAGCAGCT	GGGGATCTGG	CACCCATGAG	TGCTTTTCATT	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	GGTGGTTTGG	CTGCCCAGGA	GGTCATGAAG	GCTTGTTCTG	GAAAGTTTAT	GCCTATTAGG	CAGTGGCTGT	ACTTTGATGC	CCTTGAATGT	CTCCCAGAGC	
UBA1Y-rat	GGTGGTTTGG	CTGCCCAGGA	GGTCATGAAG	GCTTGTTCTG	GAAAGTTTAT	GCCTATTAGG	CAGTGGCTGT	ACTTTGATGC	CCTCGAATGT	CTCCCAGAGC	
UBA1Y-TMU	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
UbalY-mouse	ACAGAGTGGC	CTTCATGGAA	GATAAGTGCC	TGCCACACCA	GAACCGTTAT	GATGGGCAAG	TGGCTGTATT	TGGATCAGAC	CTACAAGAGA	AGCTTGGCAA	
UBA1Y-rat	ACAGAGTGGC	CTTCATGGAA	GATAAGTGCC	TGCCGCGCCA	GAACCGTTAC	GATGGGCAAG	TGGCGGTATT	TGGATCAGAC	CTACAAGAGA	AGCTTGGCAA	
UBA1Y-TMU	NNNNNNNNNN	NNNNNGGTAG	GTGCAGGTGC	TATTGGCTGT	GAGCTGCTCA	AGAACCTTGC	CATGATTGGT	CTTGTTGCGC	GGGAGGTTGG	AGAAATCAC	
UbalY-mouse	GCAGAAATAC	TTCT.....		C.....		T....	C	..C..C..T.	A..		
UbalY-rat	GCAGAAGTAC	TTCT.....		C.....T..		T....	C	C..T.	..A...G..		
UBA1Y-TMU	GTACAGACA	TGGACACCAC	TGAGAAGTCA	AACCTGAACA	GACAGTTTCT	CTTCCGTCCC	CGGGACGTCA	CAAAGTTAA	GTCTGAGACT	GATGCTGCTG	
UbalY-mouse		T			..G.....		T....A...	A....		..C....A.	
UbalY-rat		T				...T..C...	T....T....	A....		..C....A.	
UBA1Y-TMU	CAGTGCATGA	CATAAATCCA	CACATCAGGG	TATTCAGCCA	CCAGAATCGA	GTAGGCTCTG	AGACAGAACA	CATCTATGAT	GATGACTTCT	TCCAAAACCT	
UbalY-mouse	G...		A			C.A.		..G.....		G...	
UBA1Y-rat	A.G...			..G.G....				..G.....			
UBA1Y-TMU	GGATGGTGTG	GCCAATGCC	TAGACAATGT	GGATGCTCGA	TTGTACATGG	ACCGCCGTTG	TGTTTACTAT	TGTAAGCCCTC	TCCTGGAATC	AGGCACTCTG	
UbalY-mouse					..A...G...			C.....	..G.....	G...	
UBA1Y-rat		T.			..A.....			C.....	..G.....	C....AT..	
UBA1Y-TMU	GGCACCACAG	GCAACGTGCA	GGTGGTTGTT	CCCTTCTTGA	CAGAATCTTA	CAGCTCAAGC	CAGGACCCAC	CTGAGAAATC	CATCTCCATC	TGCACACTGA	
UbalY-mouse						T..			...C.....		
UBA1Y-rat			A.....						...C.....		
UBA1Y-TMU	AGAACTTCCC	TAATGCCATT	GAGCACACCC	TGCAGTGGGC	TCGGGATGAG	TTTGAAGGGC	TCTTCAAGCC	GTCAGCCGAA	AACATTAACC	AGTACCTCAC	
UbalY-mouse			G				A	T....	..TG.....	..A.....	
UBA1Y-rat		C.....	T....			A.	..G.....A		..TG.....	..A.....T	
UBA1Y-TMU	GGACCCCAAG	TTCATGGAGA	GGACACTGCA	GCTAGCTGGC	ACCCAGCCTT	TGGAAGTACT	GGAGGCTATA	CAGTACAGCC	TGGTCTCTGA	CAGGCCACAG	
UbalY-mouse		C		C...			C...	..C.G....		G.....	
UBA1Y-rat		C						...G.....		G.....	
UBA1Y-TMU	ACTTGGGCCG	ACTGTGTGAA	GTGGGCCTAC	CAGCACTGGC	ACACCCAGTA	TTCTCACAA	ATCCAGCAGT	TACTGCACAA	CTTCCCTCCA	GACCAGCTTA	
UbalY-mouse		C	T.....					..G.....		..CA.....	
UBA1Y-rat		C	T.....					..G.....			

UBA1Y-TMU	CAAGCTCTGG	AGCACTTTTT	TGGTCAGGAC	CAAAACGCTG	TCCACATCCC	CTCACCTTTG	ACACAAACAA	TCCCTACAT	CTGGATTATG	TGATGGCTGC
UbalY-mouse	.G.....				T		...T.....	G...		
UBA1Y-rat		G.T.....		.T.....	A					.T.....
UBA1Y-TMU	TGCCAACCTG	TTTGCTAAGA	CATATGGACT	AGGAGGGTCC	CAGGACTGTG	CTGCAGTGGC	CAAAGTCCTG	CAGTCTCTGC	CGGTCCCCAA	GTTTGCCCCC
UbalY-mouse		C...	C.....			...TG.....	C.....		.A.....	T...
UBA1Y-rat		C...	C.....	..A.....		...T...A.	..C.C.....		.A.C.....	T...
UBA1Y-TMU	AAGTCTGGCA	TCAGGATCCG	TGTTTCTGAG	CAGGAGCTGC	AGAGTACCAG	TGCCACTGCC	AATGTTG---	-----	-----	-----
UbalY-mouse		A				CA...	.T..A..ACA	GCCACCTGGA	GGAACTCAAG	ACTGCACTTC
UBA1Y-rat		A	A			A.T.	G...ACA---	GCCACCTAGA	GGAACTCAAG	ACTTCACTAC
UBA1Y-TMU	-----	-----	-----	-----	-----	---GATGACG	ACAGCAACTT	TGACATGGAT	TTCATTGTGG	CAGCATCCAA
UbalY-mouse	CTACTCCAGA	CAAGCTGCTT	GGATTCAAGA	TGTACCCCAT	TGACTTTGAG	AAG.....		.C.....		
UBA1Y-rat	CTACTCCAGA	CAAGATGCTT	GGATTCAAGA	TGCATCCCAT	TGATTTCGAG	AAG.....T.		.C.....		
UBA1Y-TMU	CCTCCGAGCT	GAAAACTATG	GCATTTCCCC	TGCAGACCGG	CACAAGAGCA	AACTGATTGC	AGGCAAGATC	ATCCCAGCCA	TTGCAACCAC	CACATCAGCT
UbalY-mouse	A			A.....	..T.....					T...
UBA1Y-rat	G..A		A...C.....		..T.....		...G.....	T.		T.....
UBA1Y-TMU	GTAGTTGGCC	TTGTGTGTCT	GGAGCTCTAC	AAGGTGGTTC	AGGGCCACCA	ACAGCTAGAG	TCTTATAAAA	ACAGTTTAT	CAACTTGGCT	CTGCCTTTCT
UbalY-mouse	A.....		G...			..A..G...				G.
UBA1Y-rat		.C.....	G..T		...T.....	..A..G..T	...T.....			
UBA1Y-TMU	TTAGCTTCTC	TGTACCTCTG	GCTCCAGCGG	GTCACCAGTA	CTATGATCAA	GAGTGGACAT	TGTGGGATCA	CTTTGATGTG	CAGGGACTGC	AACCTAGCGG
UbalY-mouse	T...	..C.....	A.T	...T.....		C	G		..A.....	
UBA1Y-rat	.C.....	..C.....A	G.T	A.....	A..	A.....	G	T.....		T..
UBA1Y-TMU	TGAGGAGATG	ACCCTCAAGC	AGTTTATAGA	CTACTTTTAA	GACAGAGCAC	AAACTGGAAG	TCACCATGCT	GTCCCAGGCT	GTGTCCATGC	TCTATTCCTT
UbalY-mouse		...T.G...				..G.....	...T.....			TG.
UBA1Y-rat	..A.....		C...			..G.....A	A..	A.....		
UBA1Y-TMU	CTTGATGCCA	GCCACCAAGC	TCAAGGAACG	GTTGGATCAG	CTGATGACAG	AGATCGTGAG	CCGTGTGTCA	AAGCGAAAAC	TGGGCCATCA	TGTGAAGTCC
UbalY-mouse	...T.....	...G.....	...A.....		.C.....	...T.....	.T.....	...A.....		...A
UBA1Y-rat	...T.....	A.	..C.....		.C.....	...T.....		...G.....	...G...	C.....
UBA1Y-TMU	GTTGTGTTTG	AGATGTGCTG	CAACAATGAG	AGTGGAGAAG	ACATCGAGGT	TCCTTATATA	TGATACATCA	TCCGCTGA		
UbalY-mouse	C.....	..C.....	G...C	C.	...A.....	G..	C.....			
UBA1Y-rat	C.....	..C.....		C.	...T..A..	G..	C.....	..G.....		

d

UBA1Y-TMU	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
UbalY-mouse	MSSSVLSKKR	KVSGPDSSLD	SSWSPTY SVM	FGVPPGPTNE	MSKNKEMDID	ESLYSRQLYV	LGHEAMKHLQ	ASSVLISGLQ	GLGVEIAKNI	ILGGVKAVTL
UBA1Y-rat	MSSSVLSKKR	RVSGPDSEVD	SSWPSTHSVM	FGAPPGPNSG	MSKNKEMDID	ESLYSRQLYV	LGHEAMKHLQ	TSSVLISGLQ	GLGVEIAKNI	ILGGVKAVTL
UBA1Y-TMU	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
UbalY-mouse	HDQGIAQWAD	LSSQFCLREE	DIGKNRAEIS	QPRLAELNSY	VPVFAYTGPL	IEEFLSGFQV	VVLTNTPLEY	QLQVGEFCHS	HGIKLVVADT	RGLVGQLFCD
UBA1Y-rat	HDQGTAWAD	LSSQFYLHEE	DIGKNRAEVS	QPRLAELNSY	VPVHTYTGPL	VDDFLSGFQV	VVLTNTPLEY	QLQVGEFCHS	HGIKLVVADT	RGLVGQLFCD
UBA1Y-TMU	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
UbalY-mouse	FGEEMILTDS	NGEQPLSAMV	SMITKENPGI	VTCLEDSRHG	FESGDFISFT	EVQGMSELNG	IGPIEIKVLG	PYTFISICDTS	SFSEYIRGGI	VSQVKVPRKI
UBA1Y-rat	FGEEMILTDA	NGEQPLSAMV	SMITKENPGI	VTCLEETRHG	FESGDFVSFT	EVQGMSELNG	IGPMEIKVLG	PYSFSICDTS	SFSEYTRGGI	VSQVKVSQKI
UBA1Y-TMU	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
UbalY-mouse	NFKPLLASLA	EPEFVVTDFE	KCCHPAQLHI	GFQALHQFCT	QHSRPPRPHN	EEDAEEVLTL	AQSVNAQALP	AVQQDCLDID	LIRKLAYVAA	GDLAPMNAFF
UBA1Y-rat	SFKSLVASLA	EPEFVITDFA	KCCRPALHI	GFQALHQFCT	QHSRPPRPHN	EEDAAEMVTL	AQAVNAQSLP	AVQQDCLDID	LIRKLAYVAA	GDLAPMSAFI
UBA1Y-TMU	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
UbalY-mouse	GGLAAQEVKM	ACSGKFMPIR	QWLYFDALC	LPEHRVAFME	DKCLPHQNRY	DGQVAVFGSD	LQEKLQKQKY	FLVGAGAIGC	ELLKNFAMIG	LGCGEDGEIT
UBA1Y-rat	GGLAAQEVKM	ACSGKFMPIR	QWLYFDALC	LPEHRVAFME	DKCLPRQNRY	DGQVAVFGSD	LQEKLQKQKY	FLVGAGAIGC	ELLKNFAMIG	LGCGEGETIT
UBA1Y-TMU	??????????S	NLNROQLFRP	RDVTKLKSET	DAAAVHDINP	HIRVFSHQNR	VGSETEHIYD	DDFFQNLGDV	ANALDNVDAR	LYMDRRRCVY	CKPLLESGTL
UbalY-mouse	VTDMDTIEK.		W.I.....	A...R....	...I.....	..P...V..	K....		..V.....	R.....
UBA1Y-rat	VTDMDTIEK.		W.....	A...R....	...C.....	..P...V..				R.....
UBA1Y-TMU	GTKGNVQVVV	PFLTESYSSS	QDPPEKSISI	CTLKNFPNAI	EHTLQWARDE	FEGLFKPSAE	NINQYLTPDK	FMERTLQLAG	TQPLEVLEAI	QYSLVLHRPQ
UbalY-mouse			...P.....		...V.....	Q...	..V.....			HC...Q...
UBA1Y-rat			...P.....			...Q...	..V...M...			.C...Q...
UBA1Y-TMU	TWADCVKWAY	QHWHTQYSHN	IQQLLHNFPF	DQLTSSGALF	WSGPKRCPPH	LTFDTNPNPLH	LDYVMAAANL	FAKTYGLGGS	QDCAAAVAKVL	QSLPVPKFPAP
UbalY-mouse	T...			A.....		...I.....		..Q.....	...V...L.	
UBA1Y-rat	T...			...V..				..Q...E..	...TTL.	...A....
UBA1Y-TMU	KSGIRIRVSE	QELQSTSATA	NVG*							
UbalY-mouse	H...	T	IDDSHLEELK	TALPTPDKLL	GFKMPIDFE	KDDDSNFHMD	FIVAASNLR	ENYGISPADR	HKSKLIAGKI	IPAIATTSTA
UBA1Y-rat	H...	V	DD-SHLEELK	TSLPTPDKML	GFKMHPIDFE	KDDDSNFHMD	FIVAASNLR	ENYDIPPADR	HKSKLIAGKI	IPAIATTSTA
UbalY-mouse	IVGLVCLLELY	KVVQGHQQLE	SYKNSFINLA	LPLFSFSAPL	APECHQYYDQ	EWTLWDRFDV	QGLQPSGEEM	TLKQFLDYFK	TEHKLEVIML	SQGVSMLYSV
UBA1Y-rat	VVGLVCLLELY	KVVQGHQQLD	SFKNSFINLA	LPPFSFSAPL	APGYHQYYDK	KWTLWDRFDV	QGLQPSGEEM	TLKQFLDYFK	TEHKLEITIL	SQGVSMLYSF
UbalY-mouse	FMPASKLKER	LDQP-MTEIV	SCVSKQKLGH	HVKSLVFELC	CNSDSGDDIE	VPYVRYIIR*				
UBA1Y-rat	FMPATKLQER	LDQP-MTEIV	SRVSKRKLQ	HVKSLVFELC	CNNESGDDIE	VPYVRYIIG*				

EIF2S3Y-TMU-mRNA	ATGGCGGGCG	GCGAGGCTGG	AGACTCTCTT	GGGCAGCCGC	ATCTTTCTCG	TCAGGATCTT	GCTA M TTTGG	ATGTTACCAA	GTTGAC K CCG	CTTTCACATG
EIF2S3Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
EIF2S3Y-TOS			A....C				...C.....		T...	
Eif2s3y-mouse		.A..A..C..	T.T.A....C				..C.C.....		T...	G..
EIF2S3Y-rat		.A..A..C..	..TGA....C		..C.....		..C.C.....		T...	..A.....
EIF2S3Y-TMU-mRNA	AAATTATCAG	CAGACAAGCC	ACAATTAATA	TAGGCACAAT	TGGTCATGTT	GCTCATGGGA	AATCTACAGT	TGTAAAAGCC	ATTTCTGGTG	TTCACACTGT
EIF2S3Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
EIF2S3Y-TOS										
Eif2s3y-mouse						A.				
EIF2S3Y-rat										...T.....
EIF2S3Y-TMU-mRNA	CCGATTCAAA	AATGAAC TAG	AAAGGAATAT	TACCATAAAA	CTTGGATATG	CTAATGCCAA	GATTTATAAG	CTTGATGACT	CAAGTTGTCC	TCGACCAGAA
EIF2S3Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
EIF2S3Y-TOS									A..	
Eif2s3y-mouse							A.....			
EIF2S3Y-rat		G.								
EIF2S3Y-TMU-mRNA	TGCTAC A KAT	CTTGTGGAAG	TAGTACACCT	GATGAGTTTC	CTTCAGATAT	TCCAGGGACC	AAAGGAAACT	TCAGACTAGT	CAGACATGTT	TCCTTTGTTG
EIF2S3Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
EIF2S3Y-TOS	G..									
Eif2s3y-mouse	..T....G..									
EIF2S3Y-rat	..T....G..	G..				C.....T.				
EIF2S3Y-TMU-mRNA	ATTGTCCTGG	TCATGATATT	TTGATGGCAA	CTATGCTGAA	TGGGGCAGCA	GTGATGGATG	CAGCTCTTCT	GTTGATAGCT	GGTAAT G RAT	CTTGTCTCTA
EIF2S3Y-TMUgenome									A..	
EIF2S3Y-TOS									A..	
Eif2s3y-mouse									A..	
EIF2S3Y-rat					..T.....				A..	
EIF2S3Y-TMU-mRNA	ACCTCAGACT	TCTGAACACC	TAGCAGCCAT	TGAAATTATG	AAGCTAAAC	ATATTTTGAT	TCTGCAAAAT	AAAATTGATT	TGGTGAAAGA	AAGTCAGGCT
EIF2S3Y-TMUgenome						A			A	
EIF2S3Y-TOS						A			A	
Eif2s3y-mouse			.G..T....			A				..C.....
EIF2S3Y-rat			.G..T....			A				
EIF2S3Y-TMU-mRNA	AAAGAACAGT	ATGAACAGAT	ACTTGCAATT	GTACAGGGTA	CAGTAGCTGA	AGGAGCTCCT	ATTATTCCAA	TTTCTGCTCA	GTTAAAATAC	AATATTGAAG
EIF2S3Y-TMUgenome										
EIF2S3Y-TOS										
Eif2s3y-mouse					C..					
EIF2S3Y-rat		G.								
EIF2S3Y-TMU-mRNA	TTGTATGTGA	ATATATA--G	TAAAGAAAAT	TCCAGTACCT	CTAAGAGACT	TTACTTCAGA	ACCCCGACTT	ATTGTTATTC	GGTCTTTTGA	TGTTAACAAA
EIF2S3Y-TMUgenome		 TA								
EIF2S3Y-TOS		--								
Eif2s3y-mouse		G.....--								
EIF2S3Y-rat	...T.....	G.....--		A.....					A	
EIF2S3Y-TMU-mRNA	CCTGGCTGTG	AAGTTGATGA	CCTTAAAGGG	GGGTAGCTG	GTGGTAGTAT	TTTAAAAGGT	GTATTAAAGG	TGGGCCAAGA	GATAGAAGTG	AGACCTGGGA
EIF2S3Y-TMUgenome				A						
EIF2S3Y-TOS							...C.....			
Eif2s3y-mouse							...C.....	A...		...T.....
EIF2S3Y-rat		T.					...A.....			...T.....
EIF2S3Y-TMU-mRNA	TTGTTTCTAA	AGATGGAGAA	GGGAAGCTTA	TGTGTAAACC	AATCTTTTCC	AAGATTGTAT	CGCTTTTTCG	AGAACACAAT	GATCTTCAGT	ATGCTGCTCC
EIF2S3Y-TMUgenome						A	A			
EIF2S3Y-TOS										
Eif2s3y-mouse		...C.....					...C.....			
EIF2S3Y-rat		...CA.T...				...G.....	A			...A.....
EIF2S3Y-TMU-mRNA	AGGTGGTCTT	ATTGGAGTTG	GAACAAAAAT	TGACCCAACA	TTATGCCGAG	CAGATAGAAT	GGTTGGGCAG	GTCTTGGTGT	CTGTTGGAGC	ATTACCTGAG
EIF2S3Y-TMUgenome										
EIF2S3Y-TOS				G						
Eif2s3y-mouse				G						
EIF2S3Y-rat										
EIF2S3Y-TMU-mRNA	ATATTCACAG	AATTAGAAAT	TTCTACTTTC	CTACTAAGAC	GGCTCCTAGG	TGTACGTACA	GAAGGAGACA	AAAAAGCAGC	AAAAGTTCAA	AAACTATCCA
EIF2S3Y-TMUgenome										
EIF2S3Y-TOS										
Eif2s3y-mouse		.G.....		...G....				.G.....		..G.....
EIF2S3Y-rat				T...G....	.A.....	..T.....		...T....	..G.....	
EIF2S3Y-TMU-mRNA	AGAATGAAGT	ACTCATGGTA	AACATAGGGT	CCTTGTCGAC	AGGAGGCAGA	GTTAGTGCAG	TCAAGGCAGA	TTTGGGTAAA	ATTGTTCTCA	CCAATCCTGT
EIF2S3Y-TMUgenome		G.A	T	A				A		
EIF2S3Y-TOS										...A.....
Eif2s3y-mouse		G		...T....					...A.....	...A.....
EIF2S3Y-rat			...A.....	...A..T...						
EIF2S3Y-TMU-mRNA	GTGCACAGAA	GTAGGAGAAA	AAATTGCTCT	TAGCCGACGA	GTTGAGAAGC	ACTGGCGGTTT	AATCGGTGGG	GGCCAGATAA	GAAGAGGAGT	GACTATCAAG
EIF2S3Y-TMUgenome	... T			A		A ---	-----			-----
EIF2S3Y-TOS	A.....									..C.....
Eif2s3y-mouse	A.....			A.....	...A.....		...T.....		...C.....	
EIF2S3Y-rat			..T.....		...A.....		...T.....			
EIF2S3Y-TMU-mRNA	CCAACAATAG	ATGATGAATG	A							
EIF2S3Y-TMUgenome	-----	-----	-							
EIF2S3Y-TOS			.							
Eif2s3y-mouse			.							
EIF2S3Y-rat			.							

[illegible]

g

TSPY-TMU-mRNA	ATGGAGAGTC	TGGAGGAGGG	GAGTGTGAGG	GTGGCTGCCC	AGGAATTCCA	GGATCTGGTC	CTCCTGGAAA	GCAGC Y TTCT	GGAGGGA S GG	ACTCCAATAC
TSPY-TMUgenome		CT	 G.C	.. C		 G.	A	C.....	G.....	
TSPY-ASY		CC..A.....	G.C	C.....	G	.. G		C.....	GA.	G.....C..
TSPY1-rat	A.T	CT.....A	G.C	C.....	G	.. A	A	C.....	GA.	C.CT..G...
TSPY-TMU-mRNA	AGGCACCAGA	GCAGAGCCCA	GGTGCTCCTG	CTGAGGACAA	TCGGATGGTG	AAGCCTCTGG	TAATAGGTGG	TCCTGTGGAG	GACTGGATAA	CCCTTCCAGA
TSPY-TMUgenome			----	-- G	.. A	 A		.. G	 C	
TSPY-ASY		T..		.. G	.. A				G..C	
TSPY1-rat				.. G	C.. A		T.....	A	..T...G..C	G....
TSPY-TMU-mRNA	GGTGCTACTT	CACCTGACA	CGTGTGAGGA	GGGGCACGTG	GCTGCTGTCG	GGGAGTTGAA	AGTCACAGAG	GTGGC R GATG	TCAAGAATGA	S GGTGAAGAA
TSPY-TMUgenome					 G			G.....		G.....
TSPY-ASY				G...A.	..T...AAG.			TG.....		GA..C...G
TSPY1-rat	A.		A..	...TG...A.	...A..A.G.	C.....		...TA.....		G.....G
TSPY-TMU-mRNA	ATAGAGCAGC	ACAAGCAGGA	TGGTGAAGAA	GATCAGCAGC	CTGAGCTGGA	AAAGGACCCA	GAGCAGGCCT	GTGAGTCTAA	AGACCAGCAT	GGGATACAAA
TSPY-TMUgenome	.. A	.. G	-	--.....	.. C			.. T		
TSPY-ASY	G..A..G...		A.....	A.	...AT.....			...T.....	...T.....	
TSPY1-rat	G..A..G.T.	.G..A.....	A.....C..			...A.....	...A.....	...T.....	...T.....A	
TSPY-TMU-mRNA	GGCTGCCACA	GCCAAGGGGA	GGAACGTCAG	AAAGCAGATC	TAAATGGAG	GAGCTGGAAC	TTCTTCAACT	GGAGCTCAGC	TTTGTGAATG	CCCGCTATAG
TSPY-TMUgenome		.T.C	T .. G							.T
TSPY-ASY		..T..G.T..T	..G.....	..C.....					T	G...
TSPY1-rat	...T...T	.T..G.T..C	..GC.....	..C.....						G....
TSPY-TMU-mRNA	TGGAGATCTT	GCACGGATTA	AAACCCAGGT	GGCCAAAATG	CGCAAACCTC	ACTTTGAGGG	CAGAAAGACC	ATAATCGAGG	GCATCCCAGG	CTTCTGGGCC
TSPY-TMUgenome	 C	.. A	.. AA			 C		.. T		T
TSPY-ASY	C.T..	..C.....	..G.AA...			C.....		C.....		T
TSPY1-rat	C...C.T..	..C.....	..G.AA.A..		...G.....	TC.		C.....	...T.....	T
TSPY-TMU-mRNA	AAAGCTATGA	TGAACCATCC	CCAGATGTCA	TCCATCATCA	GCAACCAAGA	TGAAGATTTA	TTAAGTTACA	TGTTGAGCTC	GGAGGTGGAG	GAGTACAACC
TSPY-TMUgenome			 G			 C	.. G .. C	 T	 CA	
TSPY-ASY				..T...T.	...T.....	C.....	..G..C.....	T		
TSPY1-rat			C			C.....	C.G..C.....	T		T.....
TSPY-TMU-mRNA	CTGGGCCAAG	GATGTGCAGA	ATGATGTTTT	CCTTTAGTGA	AAACCCATAT	TTCCGGAATG	ATGTTGTCAC	TAAAGACTAT	CAACTCAGTG	TCATTGGATA
TSPY-TMUgenome				T C	 C		.. CA T	.. G	 CA	
TSPY-ASY	TG..			T.....	C	...A.....	CA.....	..G.....	CA	
TSPY1-rat	...TTG..	G		T.....	C		CA.....	..G..T...	CA	
TSPY-TMU-mRNA	CAAAGAGTCA	GATTCCAGTA	CAATAGAGTG	GGTTGGACAG	GCTGAGAATG	GATATGCCAA	CTGCATGCAA	GAAACTACCA	GACTCACTTT	CTTCAACTGG
TSPY-TMUgenome		 T		.. A	.. T.C	.. G		.. T	.. T	
TSPY-ASY	A.....			..A.....	C.....	..T.....		..C.....		
TSPY1-rat		G.....		..A.....	C.....	..T.....		..C.....		
TSPY-TMU-mRNA	TTGTGTGCAC	ACAAATTCCC	TGGATCCAAC	AGGATTGCTG	AGATAATCAT	GGATGACCTG	TGGCCAAATC	CTCTTTACTA	TTACCCTAAG	GAAGACCACT
TSPY-TMUgenome				.. A				.. G		T
TSPY-ASY			..C.....							
TSPY1-rat	C.....	...N.....	..C.....			...T...			C.....	
TSPY-TMU-mRNA	CATGA									
TSPY-TMUgenome										
TSPY-ASY										
TSPY1-rat										

h

TSPY-TMU-mRNA1	MESLEEGSVR	VAAQEFQDLV	LLESS L LEG G	TPIQAPEQSP	GAPAEDNRMV	KPLVIGGPVE	DWITLPEVLL	HPDTCEEGHV	AAVGELKVTE	VADVKN E GEE
TSPY-TMU-mRNA2			... F .. R							... D ...
TSPY-ASY	...P.....G	..P...RE..	E	A.T.....S	...G..Q..		..VP.....	DE	V.K.....	..V...SQ.
TSPY1-rat	..NS..E..G	..P...RE..	E	PSV.....	...G..Q..	C...	..VP.A...H	DE	.TMA.....	..V.....
TSPY-TMU-mRNA1	IEQHKQDGEE	DQQPELEKDP	EQACESKDQH	GIQRLPQPRG	GTAESRSKME	ELELLQLELS	FVNARYSGDL	ARIKTQVAKM	RKPHFEGRKT	IIEGIPGFWA
TSPY-TMU-mRNA2										
TSPY-ASY	VKE...E...		...D.....	SG.	...R.....		..S.C..AF	...AK.....	...R.....	..Q.....
TSPY1-rat	VKDQ...E..Q	N.	...D...Q	...LSG.	..P..R...		...C..AF	...AK....	..R...DR...	..Q.....
TSPY-TMU-mRNA1	KAMMNH PQMS	SIISNQDEDL	LSYMLSSEVE	EYNPGPRMCR	MMFSFSENPY	FRNDVVTKDY	QLSVIGYKES	DSSTIEWVGQ	AENGYANCMQ	ETTRLTFFNW
TSPY-TMU-mRNA2										
TSPY-ASY			...L.....	...L.....	...F.....		...I.....	I...	..H.....	D.....
TSPY1-rat			...L.....	...L.....	...F.....	...I.....	...I.....	...I...	..H.....	D.....
TSPY-TMU-mRNA1	LCAHKFPGSN	RIAEIIMDDL	WPNPLYYPK	EDHS*						
TSPY-TMU-mRNA2										
TSPY-ASY										
TSPY1-rat	?.									

UTY-TMU	ATGAAATCAT	ACGGATTGCC	TTTCACTACT	GCTTTCCTAG	GTAATGAGGA	AGAGAAAAATG	GCGGCGAAAA	AGGCTAGAGG	CGAGGGCGAA	GAGGGGTCCT
UTY-TMUgenome	T.	AT.	.C.....	...GC.....		..A.....	G..			
Uty-mouse	...C.GATGC	CA.TGAA.TT	.GAT.T.CTA	AGAA..TCG.		..A.....	G..			
UTY-rat						..A.....	G..			
UTY-TMU	TCAGTCTCAC	AGTCGAAGAG	AAGAAGGCGC	TTTGTGGATT	AGATAGCAGC	TTCTTCGGGT	TCTTGAGCCG	ATGCAAAGAG	GGCGCCAAGA	TGAAGACGCT
UTY-TMUgenome							C..	T		
Uty-mouse							T	T	...T....	
UTY-rat					...C....		T	T	...T....	
UTY-TMU	GTTGAACAAG	GCAATTCGCT	GCTATGAATC	TTTAATAGTA	AAAGCTGAAG	GAAAAGTGGA	ATCAGAATTC	TTTTGTCAAT	TAGGTCACCT	CAATCTCTTA
UTY-TMUgenome										
Uty-mouse		A.	T.....G..				T..			
UTY-rat		AT.					T..			
UTY-TMU	CTGGAAGATT	ATTCAAAAGC	ATTGTCATCT	TACCAGAGAT	ACTACAGTTT	ACAGACTGAC	TACTGGAAGA	ATGCTGCCTT	TTTATATGGT	CTTGTTTTGG
UTY-TMUgenome										
Uty-mouse										
UTY-rat							A..			
UTY-TMU	TCTACTTCTA	CTACAATGCA	TTTCAGTGGG	CAATTAGAGC	ATTTCAAGAA	GTCCTTTATG	TTGACCCCTAA	CTTTTGCAAG	GCCAAGGAAA	TTCATTTACG
UTY-TMUgenome										
Uty-mouse	...T.....				T....G..			T.....	...A....	.C.....
UTY-rat			...A....					T.....		
UTY-TMU	ACTTGGTTTT	ATGTTCAAAG	TGAATACAGA	TTATGAGTCT	AGTTTAAAGC	ATTTTCAACT	GGCACTGATC	GACTGTAATG	TATGCACCTT	GTCCAGTGTC
UTY-TMUgenome										
Uty-mouse		A		...A..		...G..	..C.....		.C..T..T.	
UTY-rat	C...	..A.....				...G..	...T..A.GG	...C.....	.C..T....	T
UTY-TMU	GAAATTCAGT	TTCATATTGC	CCACTTATAT	GAAACCCAGA	GGAAATACCA	TTCTGCAAAG	GCAGCCTATG	AACAACCTTTT	ACAAATTGAA	AGCCTGCCAT
UTY-TMUgenome										
Uty-mouse			T.....						...G.....	..T.....
UTY-rat			T.....	...TT....		A				...T....
UTY-TMU	CACAAGTCAA	AGCAGCTGTA	TTGCAACAGT	TAGGTTGGAT	GCACCATAAT	ATGGATCTAG	TAGGAGACAA	GACCATCAAA	GAAAGATATG	CTATTCAGTA
UTY-TMUgenome										
Uty-mouse		...A....				A	.T.....	T...C....	.G.....	
UTY-rat		...A....					T...C....	T...C....		.C.....
UTY-TMU	TCTCCAAAAG	TCATTGGAGG	AAGATCCTAA	TTCTGGTCAA	TCTTGGTATT	TCCTTGGGAG	GTGTTACTCA	TGTGTTGGGA	AGATTCAAGA	TGCCTTTGTA
UTY-TMUgenome										
Uty-mouse						...A..		...A.....	.G....G..	
UTY-rat	...T.....			...C...		...A..	G	...A.....	.AG....G..	
UTY-TMU	TCTTATAGGC	AATCCATTGA	TAAATCAGAA	GCAAGTGCAG	ACACATGGTG	TTCAATAGGT	GTGCTCTATC	AACAGCAAAA	TCAACCTATG	GATGCTTTGC
UTY-TMUgenome										
Uty-mouse	...C....						...G....			...C....
UTY-rat							...T....			...C....
UTY-TMU	AAGCCTATAT	ATGTGCTGTA	CAGTTGGATC	ATGGACATGC	AGCAGCCTGG	ATGGACTTAG	GTATTCTTTA	TGAATCTTGC	AACCAACCCC	AGGATGCCAT
UTY-TMUgenome										
Uty-mouse										
UTY-rat	...T.....		...A....					...C...	...A....	...C....
UTY-TMU	TAAATGCTAC	TTAAATGCAA	CTCGGAGCAA	AAGCTGTAGT	AACACCTCTG	CACCTTATATC	AAGAATTAAG	TTTTTACAGG	CTCAGTTATA	TAACCTTCCA
UTY-TMUgenome										
Uty-mouse		G		.TCT...A.		...C...			G.G	
UTY-rat						TG....			G.G	
UTY-TMU	CAAAATAGTC	TACAGAATAA	AACTAAATTA	CTTCCTAGTA	TTGAGGAGGC	ATGGAGCCTA	CCAATCCCCG	CAGAGCTTAC	CTCCAGGCAG	GGTGCCATGA
UTY-TMUgenome							A...			
Uty-mouse	...G....									
UTY-rat	...G....									
UTY-TMU	ACACAGCACA	GCAGAATGTT	TCTGATAGCT	GGAATAGTGT	CCAGACAGCT	TCTCATTATC	CAGTACAGCA	GAAACTTTAT	GCACAGTGCT	TCACAACACA
UTY-TMUgenome										
Uty-mouse		...G....	...C....			..A...C..T		...G....	A.....	...G....
UTY-rat					...C....	..A...C..		.C..G....	A.....	
UTY-TMU	GAAATTACAG	AGTCTAGGTA	AAGATCAGAA	ACCTCCCCTT	CAAACCTGGC	CCACACGATA	TCTCCAAGCA	GCTAACACTA	ATAATCAGAA	TCAGAATGGA
UTY-TMUgenome										
Uty-mouse		...T.T...	...AC...	...A...	T			...G....	.G.....	
UTY-rat		G...T...	...AC...	.G...G...	T	.T...G...	.A...T...	...TG.GT.G		
UTY-TMU	ACTCATACTC	TGCTCAAAA	TTTAAAGGGG	GATGCTCAAA	ACTACTACTC	CTTACATATT	CCTACCTCAG	AAGAACAAAA	GATCATTAAAC	TTTACCAAAG
UTY-TMUgenome										
Uty-mouse	.A.....	...G....	.C...A...		.TC---.T	...G....				
UTY-rat	.A.....T		.C.....	...A....	.C...CT..				...T....	
UTY-TMU	AGAGAAAGGA	TTCAAGAAGT	AAACCCCTGA	GTTCTAAAAC	AAGTAGGCAA	GATATAGACA	CT---GATAT	CTGTGTTTCT	GCAAAGGAAC	ATTCTAATCA
UTY-TMUgenome							---GATAT			
Uty-mouse	...C....		...T....	C.....	...A...	...G....	..ATCTA...	AA.	...A.G...	
UTY-rat	...C....		...T....	.C.....		...G....	.ATCTA...	..A...AA.	AA.	
UTY-TMU	TATTCATCAG	ATCTCATCAG	TTCTATTTC	CACCCCTAAT	AATGAAGATT	CTGTGTCACC	AAATCTAACA	ATTGTAGACA	ATCCTCAGCT	TTCTGTCTTT
UTY-TMUgenome										
Uty-mouse	...T.....		...G....	.G.....	..A...A.		.G..T...T.			...A....
UTY-rat	G...	...A....		.GT.....			...T...T.	T.		...G....

UTY-TMU	GTTGAAGAAG	CTATTGATAA	TGTGGACCAT	GACACTGGAA	CTTGTGACAA	AGTCAATAAT	GTTTCATCCAG	CCATTACACAA	AAAGTCTGAT	AATCTTACTG
UTY-TMUgenome										
Uty-mouse	...G...A	...C..C..		...T.....			T..	C.....	T...	
UTY-rat	..A.G.....	...G..C.G		...T..A...			A	..C.....	...C.....	
UTY-TMU	CTTCTTCACC	CTCTTCACGC	AGTTCCACAG	AAACACTTTC	TCCAAAGCTC	ACTGAGCAG-	-----ACACA	CATTGTTACC	AGCTTTACCA	GTCCTCACAG
UTY-TMUgenome						-----	-----			
Uty-mouse			.T.....		..T.....	...A..-	-----		T..	
UTY-rat					..T....G..	A	TGCAC.....		T..	..C...A...
UTY-TMU	TGGATTACAC	ACAACATAACG	GAGAA---CA	CAAGAACTCA	GAGAGTTCTG	TCAGTGTATA	TGTTGGCCTC	AGACCTAGAC	CTCAAATCAT	ACCATCAATG
UTY-TMUgenome			-----							
Uty-mouse		...T..T..	...GGG..	TG.....T..		C.....		T..		
UTY-rat	G..	...T...T..	..G.---C	TG.....	..A.....A	G				G...
UTY-TMU	TCTGTGTCCA	TATACTCCAG	TTCAACAGAA	GTTCTGAAAG	CATGCAGGAA	TTTAGGTAAA	AATGGCTTAT	CTGATGGTCA	TATTCTCTTG	GATATATGCC
UTY-TMUgenome										
Uty-mouse					.G.....G	.C.....	...T..G..	..A.....		
UTY-rat		G..				.C.....	...T..G..	..AA.....		
UTY-TMU	CACCTCCAAG	ACCACCAACT	TCACCATATC	CTCCCTTACC	AAAGGAGAAG	TTGAATCCAC	CAACTCCTAG	TATTTATTTG	GAAAAAAGC	GTGATGCCTT
UTY-TMUgenome										
Uty-mouse			..G.....	...T.....			.C.....		...A.....	
UTY-rat			C..		..A.....		.C.....			
UTY-TMU	TTTCCCTCCA	TTACATCAAT	TTTGTATCAA	TCCAAAAAAC	CCTGTTACAG	TAATACGTGG	TCTTGCTGGA	GCTCTTAAGT	TAGATCTTGG	ACTTTTCTCT
UTY-TMUgenome										
Uty-mouse	..T.....		...C.....			...T..C..				
UTY-rat			...C.....		G..	...T.....	C.....			
UTY-TMU	ACCAAAACTT	TGGTAGAAGT	TAACAATGAG	CATATCGTAG	AAGTAAGGAC	CCAATTGTTG	CAGCCAGCTG	ATGAAAACTG	GGATCCCTCT	GGAACAAAGA
UTY-TMUgenome										
Uty-mouse		G.C	A		...G.....	A	A..			
UTY-rat		C			...G.....	...T...	A..		...T...	
UTY-TMU	AAATCTGGCG	ATATGAAAAC	AAATGTTCTC	ATACTACAAT	TGCTAAATAT	GCACAGTACC	AGGCCTGCTC	CTTCCAGGAA	TCATTGAGAG	AAGAAACTGC
UTY-TMUgenome										
Uty-mouse		T	..A.....							...A..A
UTY-rat		G.....T	..A.....							...A..A
UTY-TMU	GAGAAGAACA	CAAGTTAAAG	ACTATTGAGA	TAATGAATCA	ACATGTTTCA	ATACTTCTGG	AAGGAGACAG	AAAGCACCCCT	TTAAACCCTA	AAAATGTGGG
UTY-TMUgenome										
Uty-mouse	A.....					..A.....				
UTY-rat	...A....					..T...CA.		...T...T..		...C....
UTY-TMU	ATTAACATTG	ATCTATCTGA	TAACAAAAAG	TGGAAGTTGC	AGTTACATGA	GCTGACTAAG	CTTCCTGCTT	TTGTACGTGT	TGTATCAGCA	GGGAATCTTC
UTY-TMUgenome										
Uty-mouse		...G.....				..A....				
UTY-rat		..C..G....								
UTY-TMU	TCAGTCATGT	TGGCTATACT	ATACTGGGGA	TGAATTCAGT	TCAACTATGT	ATGAAAGTTC	CAGGGAGTAG	AATACCAGGT	CACCAAGAAA	ATAACAATTT
UTY-TMUgenome										
Uty-mouse			...T.....							
UTY-rat		C	...A....		...T...			G.....		...C...
UTY-TMU	CTGTTCTGTT	AACATAAATA	TTGGTCCAGG	TGATTGTGAA	TGGTTTGTTG	TACCCGAAGA	TTATTGGGGT	GTTCTGAATG	ACTTCTGTGA	AAAAAATAAC
UTY-TMUgenome										
Uty-mouse						...T.....				...T...
UTY-rat						...T.....				...T...
UTY-TMU	TTGAATTTTC	TGATGAGCTC	TTGGTGGCCG	AACCTTGAG	ATCTTTATGA	AGCAAAATGT	CCTGTTTATA	GGTTTATTCA	AAGACCTGGA	GATTTGGTCT
UTY-TMUgenome										
Uty-mouse	C.	T..							..C.....T	T..
UTY-rat	CT	T..								
UTY-TMU	GGATAAATGC	AGGCACGTGT	CACTGGGTTC	AAGCTATTGG	GTGGTGTAAC	AATATCACCT	GGAATGTTGG	TCCACTCACA	GCCTTCCAGT	ATAAATTGGC
UTY-TMUgenome										
Uty-mouse				.G.....		..C.....			..A.....	
UTY-rat				.G.....		..C.....				
UTY-TMU	AGTTGAACGA	TATGAATGGA	ACAAATTGCA	AAGTATCAAG	TCAGTAGTAC	CCATGGTTCA	TCTTTCATGG	AATATGGCAC	GAAATATCAA	AGTTTCAGAT
UTY-TMUgenome										
Uty-mouse			.T.....	...G.....						...C.....
UTY-rat	A..									...C.....
UTY-TMU	CCAAAGCTTT	TTGAAATGAT	TAAATATTGT	CTCCTCAAGA	TTCTCAAGCA	GTGTCAGACT	TTGAGAGAAG	CTCTTGTTGC	AGCAGGGAAA	GAGGTTTTGT
UTY-TMUgenome										
Uty-mouse					..A.....	T.....				
UTY-rat			...G..A--	-----	-----	-----	-----	-----	-----	-----
UTY-TMU	GGCATGGGAG	GATAAATGAT	GAACCAGCTC	CTTACTGTAG	CATCTGTGAG	GTGGAAGTTT	TTAATTTGCT	TTTTGTCACT	AACGAAAGCA	ATTCTGAAAA
UTY-TMUgenome										
Uty-mouse							...A..		..T.....	...C.....
UTY-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
UTY-TMU	GACATATATA	GTACATTGTC	AGAACTGTGC	ACGAAAAACA	AGTGGAAATT	TGGAATAATT	TGTGGTGCTA	GAACAGTACA	AAATGGACGA	CCTGATTCAA
UTY-TMUgenome										
Uty-mouse							..A.....	...A....	G..	
UTY-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

UTY-TMUGTCTATGACC AATTTACATT AGCTCCTTCA TTATCGTCTG CATCTTAA

UTY-TMUgenome.....

Uty-mouse.....G.....T.....

UTY-rat-----

j

UTY-TMU	MKSYGLPFTT	AFLGNEEEKM	AAKKARGEGE	EGSFSLTVEE	KKALCGLDSS	FFGFLSRCKE	GAKMKTLLNK	AIRCYESLIV	KAEGKVESEF	FCQLGHFNLL
Uty-mouse	SL..	.A.....K..	..E.....			T...D		..HF.....	D.	
UTY-rat	.QMPVKFDIL	RIS....K..	..E.....			L..D	.V.....	..H.....	D.	
UTY-TMU	LEDYSKALSS	YQRYISLQTD	YWKNAAFLYG	LGLVYFYNA	FQWAIRAFQE	VLYVDPNFCR	AKEIHLRLGF	MFKVNTDYES	SLKHFLQALI	DCNVCTLSSV
Uty-mouse								..M.....		
UTY-rat									R	.S.....
UTY-TMU	EIQFHIAHLY	ETQRKYHSAK	AAYEQLLQIE	SLPSQVKAHV	LQQLGWMHHN	MDLVGDKTIK	ERYAIQYLQK	SLEEDPNSGQ	SWYFLGRCYS	CVGKIQDAFV
Uty-mouse			T.	T.		I..N.T.				.I..V....
UTY-rat		.I.....		T.		N.T.				.I..V....
UTY-TMU	SYRQSIDKSE	ASADTWCSIG	VLYQQQNQPM	DALQAYICAV	QLDHGHAHAW	MDLGILYESC	NQPQDAIKCY	LNATRSKSCS	NTSALISRIK	FLQAQLYNLP
Uty-mouse								..A.....N	T....	C...
UTY-rat							..T.....		..V.....	C...
UTY-TMU	QNSLQNKTKL	LPSIEEAWSL	PIPAELTSRQ	GAMNTAQQNV	SDSWNSVQTA	SHYPVQQKLY	AQCFTTQKLQ	SLGKDQKPPF	QTGPTRYLQA	ANTNNQNQNG
Uty-mouse	.S.....			S.	..T.....	..HS....V.	T....A....	.F....Q...	...S.....	.S..D....
UTY-rat	.S.....				H.	..H....QV.	T.....	G....QR..	...S...IH.	.CVD.....
UTY-TMU	THTLPQNLKG	DAQNYYSLHI	PTSEEQKIIN	FTKERKDSRS	KPLSSKTSRQ	DIDT-DICVS	AKEHSNHIHQ	ISSVPISTLN	NEDSVSPNLT	IVDNPQLSVF
Uty-mouse	N.....S..	...H-F.R.		...S.....	.S.T....K	.R..SN...N	..K....Y.	S..	.KE....D.I	L
UTY-rat	N.....S..	.T..HS...	F.	...S.....	.S..P.....	.R..SN.Y.N	KN..E	.T.....S..	I	L
UTY-TMU	VEEAIDNVDH	DTGTCDKVNN	VHPAIHKKSD	NLTASSPSSA	SSTETLSPKL	TEQ--THIVT	SFTSPHSGLH	TTNGE-HKNS	ESSVSVNVGL	RPRPQIIPSM
Uty-mouse	.G.T.....	.I.....	..L....P.	..S.....	I.....L.	..--.....	..I.....	.I...G.E.L	..A.....	...S.....
UTY-rat	.G..M.S...	.SR.....	...TL...P.		L.V	..MH.....	..I..N...D	.I..G-PE..	...I..S...	
UTY-TMU	SVSIYSSSTE	VLKACRNLGK	NGLSDGHILL	DICPPPRPPT	SPYPPLPKKE	LNPPTPSIYL	ENKRDAFFPP	LHQFCINPKN	PVTVIRGLAG	ALKLDLGLFS
Uty-mouse		S...	...N.....							
UTY-rat			...N.....							
UTY-TMU	TKTLVEVNNE	HIVEVRTQLL	QPADENWDPS	GTKKIWRYEN	KCSHTTIAKY	AQYQACSFQE	SLREETARRT	QVKDYSDNES	TCSDTSGRRQ	KAPFKTIKCG
Uty-mouse	A...				.S.....		NE...		...N.....	
UTY-rat	A...	F.			.S.....		NE.K.		...I..R...	.V.....
UTY-TMU	INIDLSDNKK	WKLQLHELTK	LPAFVRVVSA	GNLLSHVGYT	ILGMNSVQLC	MKVPGSRIPI	HQENNNFCSV	NINIGPGDCE	WFVVPEDYWG	VLNDFCEKNN
Uty-mouse										
UTY-rat										
UTY-TMU	LNFLMSSWWP	NLEDLYEANV	PVYRFIQRPG	DLVWINAGTV	HWVQAIGWCN	NITWNVGELT	AFQYKLAVER	YEWNKLQSIK	SVVPMVHLSW	NMARNIKVSD
Uty-mouse								V.		
UTY-rat										
UTY-TMU	PKLFEMIKYC	LLKILKQCQT	LREALVAAGK	EVLWHGRIND	EPAPYCSICE	VEVENLLFVT	NESNSEKTYI	VHCQNCARKT	SGNLENFVVL	EQYKMDDLIO
Uty-mouse		H...					Q....			E....
UTY-rat	*	-----	-----	-----	-----	-----	-----	-----	-----	-----
UTY-TMU	VYDQFTLAPS	LSSAS*								
Uty-mouse										
UTY-rat	-----	-----								

DDX3Y-TMUmRNA	ATGAGTCAAG	TGGAAGCTGA	AAATACTGTC	GGCCTAGACC	AGCAATTCGC	AGGTCTTGAC	CTGAACTCAT	CAGATAATCC	AAATGGAGGA	GGAAATACAG
DDX3Y-TMUGenome										
Ddx3y_mouse	...C...G..	..G...G.C.	G....	T.T	T.....	A....	A	G.....		
DDX3_rat	.C.....A	..CG..G..	G..T	T....	T..	T.....	A.....-	--.....A	G.....	.TT.G....
DDX3Y-TMUmRNA	AGAGCAAAGG	GCGCTATATA	CCTCCTCACT	TAAGGAACAG	AGAAACATCT	AAAGGTGTCT	GTGATAAGGA	CAGTTCAGGA	TGGAGTTGTA	GTAAGATAAA
DDX3Y-TMUGenome										
Ddx3y_mouse						G....			C....	
DDX3_rat	.A.....							..A.....G	----	-----
DDX3Y-TMUmRNA	AGATGCCTAT	AGTAGTTTGG	GATCTCGTGA	TTCCAGAGGA	AAGCCCAATT	ATTTCAGTGA	TCCTGGAAGT	GGATCCAGGG	GAAGATTGTA	TGATCTTGGT
DDX3Y-TMUGenome										
Ddx3y_mouse	C	..C.....		G			..G.....		G..C..	A....
DDX3_rat	C	..C.....	G..		..A.....		..G.....	A	G....	C....
DDX3Y-TMUmRNA	CGAAGTGACT	ATGATGGTAT	TGGCGGTCGT	GACCGAACTG	GTTTTGGCAA	ATTT-----	-----	--CGTTGGGG	TGACAGATCA	GATGAAGATG
DDX3Y-TMUGenome						-----	-----			
Ddx3y_mouse	...A.....		..T.....		..A.....	...GAACGG	AGTGGTCATA	GT.....A.		
DDX3_rat	T....	..C.A....	G..	..A.....		...GAACGA	AGTGGTTATA	GT.....	A....	
DDX3Y-TMUmRNA	ATTGGTCAAA	ACCACTTCCA	CCAAGTGAAC	GCTTAGAGCA	AGAACTTTTT	TCTGGAGGAA	ACACAGGAAT	TAACTTTGAG	AAATATGATG	ATATACCACT
DDX3Y-TMUGenome										
Ddx3y_mouse	.C.....						G..			
DDX3_rat					C	T.	.TG...G..			
DDX3Y-TMUmRNA	AGAGGCAACT	GGAAATAATT	GTCCCTCCTCA	CATCGAAAAT	TTCAGTGTATG	TTGAGATGGG	AGAAATTATC	ATGGGGAACA	TTGAACTTAC	TCGTTATACT
DDX3Y-TMUGenome										
Ddx3y_mouse	..A.....			..T.....	A					C...
DDX3_rat				..T.....				A....		C...
DDX3Y-TMUmRNA	CGTCCTACTC	CAGTGCAAAA	ACATGCTATT	CCAATTATTA	AAGAAAAAAG	GGACTTAATG	GCTTGTGCTC	AAACAGGGTC	TGGAAAAACT	GCTGCATTTC
DDX3Y-TMUGenome										
Ddx3y_mouse										
DDX3_rat		...T..G..							...TGT...	
DDX3Y-TMUmRNA	TACTACCCAT	TCTCAGTCAG	ATTTATACAG	ATGGTCCAGG	AGAGGCTTTG	AAGGCTATGA	AGGAAAATGG	AAGATATGGT	CGCCGTAAAC	AGTACCCAAT
DDX3Y-TMUGenome										
Ddx3y_mouse										..A.....
DDX3_rat		C....					C..			..A.....
DDX3Y-TMUmRNA	CTCCTTGGTG	TTAGCCCCAA	CAAGAGAATT	GGCTGTGCAG	ATCTATGAGG	AAGCCAGAAA	ATTTTCATAC	AGGTCTAGAG	TCCGTCCTTG	TGTAGTGTAT
DDX3Y-TMUGenome										
Ddx3y_mouse								C.A.....		
DDX3_rat		...T....					G..	C.....	...A....	
DDX3Y-TMUmRNA	GGCGGGGCTG	ATACTGTTCA	GCAGATTCGG	GACTTAGAAC	GCGGATGCCA	CTTGTTAGTT	GCCACACCCG	GACGTCTAGT	GGACATGATG	GAAAGAGGAA
DDX3Y-TMUGenome										
Ddx3y_mouse	..T..T....				..T....		A.			
DDX3_rat	..T..T....	A.			..A....		A.			
DDX3Y-TMUmRNA	AGATTGGATT	AGATTTCTGC	AAATACTTAG	TGTTGGATGA	AGCCGATAGG	ATGTTGGATA	TGGGATTGTA	ACCTCAGATA	CGTCGTATAG	TTGAGCAGGA
DDX3Y-TMUGenome										
Ddx3y_mouse				..A.....	..T.....	..C.....		A..		
DDX3_rat	C.				..T..C..	..C.....				...A....
DDX3Y-TMUmRNA	CACAATGCCA	CCAAAGGGGG	TTCGTCATAC	TATGATGTTT	AGTGCTACTT	TCCCTAAGGA	GATACAGATG	CTTGCTCGTG	ACTTTTTAGA	TGAATATATC
DDX3Y-TMUGenome										
Ddx3y_mouse			C..			..T.....				
DDX3_rat	...T.....	A.	C..	C.....	..T.....					C...
DDX3Y-TMUmRNA	TTTCTGGCTG	TAGGCAGAGT	AGGCTCTACT	TCTGAGAACA	TCACACAGAA	AGTAGTATGG	GTTGAAGAAC	TAGATAAACG	GTCATTTTTA	CTTGATCTCT
DDX3Y-TMUGenome										
Ddx3y_mouse		A..						C..		
DDX3_rat							GT	..C....		
DDX3Y-TMUmRNA	TAAATGCAAC	AGGGAAGGAT	TCATTGACTT	TAGTATTTGT	GGAGACTAAA	AAGGGTGCAG	ACTCCCTGGA	GAACCTCTTA	TTTCAAGAAA	GATATTCGTG
DDX3Y-TMUGenome										
Ddx3y_mouse							..T..G....			G...
DDX3_rat		G....					..T..A....			G.A..
DDX3Y-TMUmRNA	TACTAGTATT	CATGGAGACA	GATCACAGAA	AGATAGAGAA	GAGGCCCTTC	ATCAGTTCCG	CTCAGGGAGA	AAGCCAATTT	TGGTGGCTAC	AGCTGTGGCA
DDX3Y-TMUGenome										
Ddx3y_mouse				..C...G		..C...T..	C..		..A....	
DDX3_rat			..G.....	...C...G		..C..A....	G		..A....	
DDX3Y-TMUmRNA	GCAAGAGGTC	TAGACATTTC	AAATGTGAAA	CATGTTATCA	ATTTTGATTT	GCCAAGTGAT	ATTGAAGAAT	ACGTGCATCG	TATTGGCCGC	ACAGGACGTG
DDX3Y-TMUGenome						C...				
Ddx3y_mouse								..T.....	A.....	
DDX3_rat								..T.....		C.
DDX3Y-TMUmRNA	TAGGAAACCT	TGGACTTGCC	ACCTCTTTTT	TCAATGAGAG	GAATTTGAAC	ATTACAAAAG	ATTTGTTAGA	TCTCCTTGTT	GAGGCTAAAC	AAGAAGTTCC
DDX3Y-TMUGenome										
Ddx3y_mouse						..C.....				G..
DDX3_rat		...T.....	...C....			G..		C..		A..
DDX3Y-TMUmRNA	ATCTTGGTTG	GAAAGTATGG	CCTACGAACA	CCACTACAAG	GGAAGCAGCC	GGGGACGTTT	AAAAAG---A	TTCAGTGGAG	GATTTGGAGC	CAGAGACTAT
DDX3Y-TMUGenome							-----			
Ddx3y_mouse	T.....		..T..T....		..T....	..T.....	..T....CAG.			
DDX3_rat			..T..T....		..A....		..T....CAG.		T..	

DDX3Y-TMUmRNA	CGACAGAGCA GTGGTTCTGC AAATTCTGGG TTTAATACTA ATCGTGCAAA CAGCAGCCGA AGTAGTGGTA GTAGCCACAA CAGAGGATTT GGTGGAGGTG
DDX3Y-TMUgenome	
Ddx3y_mouse	G....AG.. .C....C.. T.....
DDX3_rat	A....A..C..C.....
DDX3Y-TMUmRNA	GCTATGGAAG TTCTACAAT AATGATGGAT ATGGAGGAAA TTATAACTCC CAGGCAGTTG ACTGGTGGGG CAATTGA
DDX3Y-TMUgenome	
Ddx3y_mouse	G.T.
DDX3_rat	G.GCA....

DDX3Y-TMUmRNA	MSQVEAENTV GLDQQFAGLD LNSSDNPNNG GNTESKGRIY PPHLRNRETS KGVCDKDSSG WSCSKDKDAY SSFGSRDSRG KPNYFSDPGS GSRGRFDDL
Ddx3y_mouse	...A..S.AV... .K....Q...
DDX3_rat	MA..... M.-..Q... VS.....N.. ..-.....R..P.
DDX3Y-TMUmRNA	RSDDYDGIGGR DRTGFGKF-- ---RWGDRS DEDDWSKPLP PSERLEQEELF SGGNTGINFE KYDDIPVEAT GNNCPPHIEN FSDVEMGEII MGNIELTRYT
Ddx3y_mouse	.N.....ER SGHS..S...
DDX3_rat	C.AS. ..S.....ER SGYS....K.A.....
DDX3Y-TMUmRNA	RPTPVQKHAI PIIEKEKRLM ACAQTGSGKT AAFLLPILSQ IYTDGPGEAL KAMKENGRIY RRKQYPISLV LAPTRELA
Ddx3y_mouse	
DDX3_rat	C.
DDX3Y-TMUmRNA	GGADTVQQIR DLERGCHLLV ATPGRLVDM ERGKIGLDFC KYLVLEADR MLDMGFEPQI RRIVEQDTMP PKGVRHTMMF SATFPKEIQM LARDFLDEYI
Ddx3y_mouse	
DDX3_rat	
DDX3Y-TMUmRNA	FLAVGRVGST SENITQKVWV VEELDKRSFL LDLLNATGKD SLTLVFVETK KGADSLNENL FQERYSC
Ddx3y_mouse	A....
DDX3_rat	R.A.....
DDX3Y-TMUmRNA	ARGLDISNVK HVINFDLPSD IEEYVHRIGR TGRVGNLGLA TSFFNERNLN ITKDLDLLV EAKQEVPSWL ESMAYEHYK GSSRGRSK-- FSGGFGARDY
Ddx3y_mouse	SR
DDX3_rat	N.....SR
DDX3Y-TMUmRNA	RQSSGSANS GNTNRANSSR SSGSSHNRGF GGGGYGSFYNDGYGGNYNS QAVDWWGN*
Ddx3y_mouse	A. ..S.....G....
DDX3_rat	Y.... ..N..... T...T.....G..S .E.....

USP9Y-TMU-mRNA	ATGACAATCA	CAACTCGTGA	CTCTCCAGTA	GGAGAGAATG	AGAGCCAGGG	ACAGGTTCCA	GATGGACAAC	CTCAGCCTTC	TTTCCAACAG	AACCAG----
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		G				T...AC.T.T				ATCT
USP9Y-rat	GG...	G	G			C...T	G...T			ATCT
USP9Y-TMU-mRNA	-----	-----	---AATTCCC	CAGCAACTCC	TCCAGATGAA	CAAGGCCAAG	GTGATGCTCC	ACCGCAGCAC	GAAGAAGAAG	ATCCTTCATT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	CATCATCTGA	TTCTTCCAAT	GAG.C...T.	..A.....	...T.....			...A....T	G.	
USP9Y-rat	CATCATCTGA	TTCTTCCAAT	GAG.....T.		...T.....			...A....T		..C...A...
USP9Y-TMU-mRNA	TCCTCATGCT	GACCTGGCAA	AGTTAGAAGA	CATGATCAAC	AGGCCCCAAT	GGGTGGTTCC	TGTTTGGCCA	ATGGAGGAAT	TAGAAGTACT	TTTAGAAGCT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	A..				...T...G..			..A..GA...		
USP9Y-rat					..A...G..			..A..G....	..G.....	...G.....
USP9Y-TMU-mRNA	TCTATGGATC	TTACTAAAA	AGGCCTTGAT	GCTAAAAGTG	AAGCATGCCA	GCGGTTTTTT	CGAGATGTTT	TAAGTGTGTC	TTTTACCAAA	ATCCTCATGG
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	T....		...T..G...	.T.....		...T.....	..C.....		...T....	T....
USP9Y-rat	T....		...A.....	..A.....		C			...T....	T.C..
USP9Y-TMU-mRNA	ATGAGGCAGT	GACTAGATGG	AAGTCTGAAA	TTCATAGGTG	TATTATTAAC	AATATGCATC	GCCTAGTGGA	GCTTTGTGTG	GCCAAATTGT	CTCAGGACTG
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	T..	..G.G....	...T....			...CC....	...G....			
USP9Y-rat		..G.G....	...T....		C....A...	..C.CC...	...A....			
USP9Y-TMU-mRNA	GTTTCCATTT	CTAGAACTTC	TAGCAATGGC	ATTAAATCCT	CATTGCAAAT	TTCATGTCTA	TAATGGTGCA	CGTCCATGTG	AATCTGTGTC	CTCAAGTGTT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse			T.						...A....	
USP9Y-rat			..G....A..				...C....		...A....	
USP9Y-TMU-mRNA	CAGTTTCCTG	AAGGTGATCT	CTTTGCCTGT	TCTCCAGACC	TACATTCACC	TAAGGGTTGG	TTAGTGGATC	TCATTAACAC	ATTTGGCACA	CTAAATGGTT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		...A...A..				A..A.....	C.			
USP9Y-rat		...A...A..			...G.....	A..A.....	C.....		...T....	
USP9Y-TMU-mRNA	TCCAGATTTT	GCATGATCGT	TTCACTAGTG	GATCAGCATT	AAATGTTCAA	ACAATTGCAG	CTCTTATTAA	GCCATTTGGA	CAGTGCATATG	AATTTCTTAC
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse								...G.....		..G.....C..
USP9Y-rat			...T....		...C....			G..		..G.A.....
USP9Y-TMU-mRNA	TCAACATACA	TTAAAAAGT	ACTTCGTTCC	GGTTATAGAA	GTGGTTCCAC	AATTACTACA	AAAATTAAGT	AATGAAGAAC	TGAAAAAGA	AGCAAAGAGT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		...G.....	...A....	A.....	..A.....	..A..T....	C.			..A.....C.
USP9Y-rat			...A....	A.....	..T.....		...CC			C.
USP9Y-TMU-mRNA	GAAGTGAAAA	ATGATGCCAT	TTCAGTAATT	ATTAAATTTC	TAAAAAATTT	AGCTTCAAGA	ATTCCAGGAC	AAGAGGAAAC	TGTTAAAGAT	TTAGAAATTT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		...A....	...A.G..						...A...	C..
USP9Y-rat		T..	...A.G..			...C....		...A....	...A...	
USP9Y-TMU-mRNA	TTAGGCTAAA	AATGATACTC	AGATGGCTAC	AGATTTCTTC	TTTTAGTGGA	AAGATGAATG	CACTGAATGA	AATTAACAAA	GTGCTGTCAA	GTGTGTCATA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	T....		...T....		...A....				...T....	
USP9Y-rat	T....		...T....		...A...G		..G.....	G	...A....	
USP9Y-TMU-mRNA	T---ACACAT	CAACATGGTA	ATTTTGAAGA	AGAGTGGCTA	ACCGCTGAGC	AAATGGCAGA	GTGGATACAA	CAGAACAAATA	TCTTATCCAT	TGTTTTGCAA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	.TAT..G...	GG.....	...C.....		...T....	G...A....	G			...C.....
USP9Y-rat	.TAT.....	.GG.....	...C.....			G.....	A.....	..A..T....		A..C.....
USP9Y-TMU-mRNA	GATAGCCTAC	ATCAACCACA	GTATGTGGAA	AAGCTAGAGA	AGATTCTTCG	TTTATAAATT	AAAGAAAAGG	CTCTTACTTT	ACAGGATCTT	AATAATATCT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse				..AA.....		...G.....				
USP9Y-rat	T.					...G.....				
USP9Y-TMU-mRNA	GGGCAGCACA	GGCTGGAAAA	CATGAAGCTA	TTGTGAAGAA	TGTACATGAT	TTACTAGCAA	AATGGCTTG	GAATTTTCT	TCTGAACAAC	TTGATCATCT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse					...T.....				C.....	
USP9Y-rat					C			..G.....	C.....	
USP9Y-TMU-mRNA	TTTTGATTGC	TTTAAGGCAA	GCTGGATGAA	TGCAAGTAAA	AAACAACGTG	AAAAACTCCT	TGAATTGATA	CGTCGCCTTG	CAGAAGATGA	TAAAGATGGT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	C.....		CA.			A..		T...		
USP9Y-rat		T.	CA.					...T....	..C.....	
USP9Y-TMU-mRNA	GTAATGGTAC	ACAAAGTGTT	GAATCTTCTT	TGGAATCTAG	CTCACAGTGA	TGATGTACCT	GTAGATATCA	TCGATATGGC	TCTCAATGCC	CACTTAAAAA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	C..					...C..G...		..G...C...	...G....	...A.....
USP9Y-rat	C..	..G.....	T.			...G....		..G...C...	...G....	...A.....
USP9Y-TMU-mRNA	TACTCAATTA	TAGTTGTT Y C	CAGGATCGAG	ATACACAAAA	AATCCAGTGG	ATAGACTGCT	TTATTGAAGA	ATTTCGAACA	AATAATAAAT	GGGTATATCC
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	...GG....	C.					...A....			...C.....
USP9Y-rat	...TG....	C.....C.			G.....	...T....	...A....			
USP9Y-TMU-mRNA	TGCATTAAAA	CAAATTAAAG	AAATTGTGAG	CTTGTTTGGT	GAAGCACCTC	AAAATTTAAG	TCAAACATCAT	CAAAGTCCTC	GTGTATTTTA	TCGTCATGAT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse			...T....					C.		C
USP9Y-rat			..G.....	T.....			...G.....	..G.....	A.....	...C.....C

USP9Y-TMU-mRNA	TTAATCAACC	AAC TTCAGCG	TAATCATGCT	TTAGTTAC Y T	TGGTAGCAGA	AAACCTTGCA	ACTTACATGA	ACAGCATAAG	ATTGTATGCT	AAAGACCATG
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	G..	A	T.	T.			G.....			.G...T....
USP9Y-rat	G..	A	T.	T.				C.		.G...T....
USP9Y-TMU-mRNA	AAGATTATGA	CCCACAAACT	ATAAGACTAG	GAAGTAGATA	CAGTCATGTT	CAAGAAGTCC	AAGAACGACT	TAATTTCTCT	AGATTCTTAT	TGAAGGAAGG
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse			G.....				G..		C	T..
USP9Y-rat							G..			T..
USP9Y-TMU-mRNA	TAAACTCCGG	CTCTGTGCTT	CCCAAGCAAA	ACAAATATGG	AACTGTTTAG	CAGAGAACGC	AGTTTACTTT	ACTGATCGAG	AAGCCTGTTT	TATGTGGTAT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	.C.....T..	T..	G..			T..	C.....	.G.....		
USP9Y-rat	.C.....T..	T..	G..		C.....	T..		.G.....		
USP9Y-TMU-mRNA	TCCAAATTAA	TGGGCGATGA	ACCTGAATTA	GAGCCTGATC	TTAATAAGGA	GTTTTTTGAA	AGTAATGTGC	TTCAGCTTGA	TCCTTCTCTA	TTAACTGAAA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		...A.....	C..	C.T.....A		C..				
USP9Y-rat		...A.....	C..	.T.....A	.C.....					
USP9Y-TMU-mRNA	ATGGAATAAA	ATGTTTTGAA	AGATTTTTC	AAACTGTTAA	TTGGCGAGAA	GGGAAACTAA	TGATAAAAAG	AAAAGTCTAT	ATGATGAATG	ATTTGGACTT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	G..G..				..T.....			...A.....	G..	
USP9Y-rat					..T.....			C	G..	
USP9Y-TMU-mRNA	AATAGGATTA	GACTACCTTT	GGAAGGTTGT	GATTCAGAGT	AATGACAATA	TTTCAAGCAG	AGCAATAGAT	CTCCTTAAAG	AGATACACAC	AAGCCTTGGT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse				TG...				..T.....	T...	
USP9Y-rat		T...		TG...		...A...	..T...		T.T..	
USP9Y-TMU-mRNA	CCCAGATTAC	AAGCTAATCA	GGTAGTAATT	CATGAAGACT	TTATTCAATC	TTGCTTTGAT	CGCTTGAATG	CCTCATATGA	TACACTGAAT	GTTTGGGATA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	...A...G.						A.		...T..TG.	
USP9Y-rat	..T.....		T..				..TA...AT	.G.....	...GT..C..	
USP9Y-TMU-mRNA	GTGATAAAGA	CAATATTTTT	AGCTGTGCAA	GACAAGAAGC	CATCCAGATG	GTTCGCATAT	TAACGGTTTT	AAGGGAATAC	ATAAATGAAT	GTGACAGTGG
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	...A.....				...GA...		C..		...G.....	A.....A
USP9Y-rat					...GA..A		..T.....		...G.....	A
USP9Y-TMU-mRNA	TTATCATGTG	GAAAGGATGA	TTCTACCTAT	GTCAAGAGCA	TTTCGTGGGA	ATCACTTGTC	TTTTACAGTT	CGATTTCCAA	ACCAGGGCAA	AGAAGTTGAT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	A.				A.	.A.....	A			...G....G
USP9Y-rat	...C...A.		...G.....		..C.....	.A.....	.C.....		...A.....	
USP9Y-TMU-mRNA	GATTTGGAACA	TAATGTCTCA	TACAAATGCT	ACCATTGGTT	CTGTGCGACG	CTGCATTCTC	AATTGTATGA	ATGCCAATGT	AGCCCATACA	GAAATTGAAC
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		..T.....	C..				...C.....	..T.....		A.....
USP9Y-rat		..C.....	...G.....				...C...T.		.A.....	A.....
USP9Y-TMU-mRNA	TGTTTATAGA	TGGAGAGTTG	ATCGATTCTG	AAGATGACAG	AAAGTTAGTT	GAACAGTTAA	ATTTGAAAGA	TAAATCACTA	ATTACTGCTA	AATATATACA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	G	A	G.A.CG...		...C.....					...T.....
USP9Y-rat	G	A	..A.....		...C.....					...T.T....
USP9Y-TMU-mRNA	AATAAATTCC	ATTATGCCTT	CAAGTCCTGA	TAGCTCATCT	GATTCGTCAG	CTGAATCTCC	TGGAAACCAC	AGTCATTA--	-CTACAGAGA	TGTTCCCAAT
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		.AC.....				...G.C...	T	...A.TA	A.....	...T.....
USP9Y-rat		.AC.....			...A...	.C.G.....		T....A.TA	A...T.....	
USP9Y-TMU-mRNA	CCAGAAATGG	AAAAATGTTT	ACCTGGGGTG	ATAATGTCAC	TGCAGCCCAG	ATACATATCT	TTCTTTGGGC	AAGTTGCAGA	CTTAGGTAGC	ATGTTGACTA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		...G.....								G
USP9Y-rat	T.....	...T.....	...A..C..						...A...	
USP9Y-TMU-mRNA	TGCCAAGTCT	TAGAGATGGG	GCAAGAATAC	TTATGAAACT	TATGCCACCA	GATAGCACAA	TCATAGAAAA	ATTAAGAGCT	GTATGTTTCA	ACGATTTAAA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	C..	A					CTC....C.		C.....	..C..G...
USP9Y-rat	C..	A			A.....	..A.....	CT.....	C.....		..C.....
USP9Y-TMU-mRNA	TCTTGGAGAA	AGTCAACTTG	GTCAGTCTCT	TCACTCACTT	TTCTTTGGTT	CTTCTGCCTC	TGAAGCACTA	TATTTAACAG	AGGTAGTTTA	TATCTTGTTA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		..A.G.....	...A.....				CC...TG...			C.C.....
USP9Y-rat		..A.G.....	...A.....				CA...T...			C...C.....
USP9Y-TMU-mRNA	ATGCCTGCTG	GTGGACCTCT	GGCTGATATT	TCCTATGACT	TTCAGTATCA	CTTCTTGAAA	AGTGGTGGCT	TACCCTTGGT	ACTTAATATG	CTGATTCAAA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse		...C.....		...C.....				...TC.T..	...A.G....	...A....
USP9Y-rat	C.	...C..A...	A.....	...C.....		...A...	.A.....	.G...C.T.	...A.G....	...A....
USP9Y-TMU-mRNA	ATAACTTCTT	ACCAAATACA	GATGTTGAAA	CTCGTAGAGA	TGCTTACGTC	TGTGCCCTTA	AAATAGCCAA	ACTGTTGTTA	ACTACCATTG	GCTATGGCCA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse					T..	A.....			...T.G...	
USP9Y-rat					T..	A...T...		...C...		
USP9Y-TMU-mRNA	TGTCCAGGCC	ATAGCAGAAG	CTTGTCAGCC	TGTTGCAGAT	GATACAGATC	CTAAAACACC	GATTAACCAG	GTTACTCATG	ATCAAGCTGT	GGTGCTACAA
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
USP9Y_mouse	...T.....				.GC.....					
USP9Y-rat					.G.....		...A...	.C.....		...AT.....

USP9Y-TMU-mRNA	AATGCCCTTC	AGAGCATTC	TAATCCTTCT	TCTGAGTGTA	TACTTAGAAA	CGTATCAATG	CGCCTTGCTC	AACAGATATC	AGGGTTGGCC	TCAAGATATA	
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
USP9Y_mouse		.A.....		C.	.G.....	T.....G.T	.AT.....	...A.....			
USP9Y-rat				C.	.G....G..	T.....G..T	.T.....	.C..A.....		C.	
USP9Y-TMU-mRNA	TACCTGATAT	TTGTGTTATT	ARAGCTATAC	AAAAAATTAT	CTGGGCAACA	GGATGTGGAT	CATTAGAACT	AGTATTTAGT	CCAAATGAAG	ATATCACAGA	
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
USP9Y_mouse			.G.....	C..	G..		G..	...T.....			
USP9Y-rat			.G.....	C..	T...A..T..		G..		...C.C....	A.	
USP9Y-TMU-mRNA	AACCTTACAAG	ATGACCACCA	GTACAAGGGG	TAATCCAGAA	GTGAAAGATG	AGCAGGTTTG	CTGTGAAGCC	CTGGAAGTGA	TGACATTATG	TTTTGCTTTA	
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
USP9Y_mouse			A.	T....		.A..A.....					
USP9Y-rat			...T...A.			...A.....		..A....A.			
USP9Y-TMU-mRNA	ATTCCAACCTG	CGTTGGATTCT	ACTCAGTAAA	GAAAAGGCCT	GGCAGTCATT	TGTCATTGAC	TTACTATTGT	ACTGTCCTAG	CAAACTATA	CGGCAGTTGG	
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
USP9Y_mouse		.TA.....	A....						G....		
USP9Y-rat		.T.....				T					
USP9Y-TMU-mRNA	CACAGGAGCA	GTTCCTTTTA	ATATGTACCA	GATGTTGCAT	GGGACATAGA	CCTTTGCTGT	TCTTCATTAC	TCTACTCTTT	ACCATTTTGG	GGAGTGCAAC	
USP9Y-TMUgenome	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
USP9Y_mouse									A.T.		
USP9Y-rat										..G....G.	
USP9Y-TMU-mRNA	AGCAGAGAAG	GGTAAACATT	CCGATGTCTA	CTTCCCTCTT	TTAAGGCGTC	TTCTCACTTA	TGCTTACAAC	AGGAACATTC	ATATACCCAA	TGTTGACGTT	
USP9Y-TMUgenome	.A.....--	-----	-----	A....			A....				
USP9Y_mouse	TAAC.....		.A.....	A.A...				..C.....	.AG.T....	T...	
USP9Y-rat	.A.G.....		.A.....	...T..A...				..C.....	.C.....	T...	
USP9Y-TMU-mRNA	CTTCTCAATG	AAGAAATTGA	TTGGCTTAAA	AGGGTTAGGG	ATTACATTAA	AAATACAGGA	GAAACAAATA	TAGAAGACCC	AGTCTTAGAA	GGCCATCTTG	
USP9Y-TMUgenome	-----	-----	-----	...T....					..A.....		
USP9Y_mouse		.C.....					G		.A.....	G.	
USP9Y-rat		.C....A.				C	C.G.G		.A....G...	G.	
USP9Y-TMU-mRNA	GAGTAACAAA	AGAGTTACTG	TCTTTTCAGA	CTCCTGAGGA	AAAATATCAT	ATTGGTTGTA	AAACAGGAGG	TGCTAATCTT	GTTAAAGAAT	TAATTGATGA	
USP9Y-TMUgenome	...C.....	-----	-----	A....				..G.....			
USP9Y_mouse			T	AA...	..G.....					T.	
USP9Y-rat			..C....A.	...C..TA.	..G.....		..G.....	A			
USP9Y-TMU-mRNA	TTTCATCTTT	CCTGTATCCA	AAGCTTACCT	ACAATATATG	AGAAGTGGGG	AAC TGCCAAT	CAAGCAAGCC	GTTCCAGTTT	GTGGTTCACC	AGCTACAGTT	
USP9Y-TMUgenome	-----	...AC....	-----	-----	-----	-----	-----	-----	-----	-----	
USP9Y_mouse		..C.....		G.....		..T.....	T	A.....C.		A..	
USP9Y-rat		..C.....		G.....TCA		..A.....A	..A..A..T	A.....C.		...C.....	
USP9Y-TMU-mRNA	AATGCTGGTT	TTGAACTCCT	TGTAGCTTTG	GCTTTTGGTT	GTGTAAGGAA	TCTTGAACAG	ATAGTAAACT	GTTTGACTGA	ACTGTTTTAC	ATAGGCACAT	
USP9Y-TMUgenome	-----	-----	-----	-----	-----	..A.....	-----	-----	-----	C	
USP9Y_mouse		A..				..A.....				C	
USP9Y-rat		...G..A..			C....	..A.....		A.		C	
USP9Y-TMU-mRNA	CAGTAACTAC	TTGTGAAGCC	CTTGCTGAAT	GGGAATATTT	ACCACCTGTT	GAATCCCGCC	CACCAAAAGG	ATTTGTGGGA	CTGAAAAATG	CTGGTGCTAC	
USP9Y-TMUgenome	A..	.G.....	-----	-----	-----	.G.....	-----	-----	-----		
USP9Y_mouse			G...G....		..C.....	.GGC....					
USP9Y-rat	.T...A...	A-----				.G.C.....			..C.....		
USP9Y-TMU-mRNA	TTGTTATATG	AACTCTGTGA	TCCAACAGCT	ATACATGATT	CCTTCAATCA	GGAACAGTAT	TCTTGCAATT	GAAAGCATAT	GCAATGATAC	AGATGGTGAT	
USP9Y-TMUgenome	C....	-----	-----	-----	-----	-----	-----	-----	..G.....		
USP9Y_mouse								.T.....	.G.G.....	...A....	
USP9Y-rat		..T.....	.T.....	G.....				...A....	.G.....	...A....	
USP9Y-TMU-mRNA	ATATTCAAAG	ATGGAAAGCA	CGACAGTGAG	AATAATGCTG	ATCCCAGAGA	TGATGTATTCT	GGATATCCTC	ATCAATTTGA	AGATAAACCA	ACATTAAGTA	
USP9Y-TMUgenome	-----	...A....	G.....	T....	-----		A.....	-----	-----	-----	
USP9Y_mouse		G..A.....	G.....	G...T..			A.....				
USP9Y-rat	.C.....	..A..G...	G.....A	AT..		T				A..	
USP9Y-TMU-mRNA	AAGTAGAAGA	TCGAAAAGAG	TACAACATTG	GTGTTTTAAA	GCACCTTCAG	ATAACCTTTG	GACATTTAGC	TGCTTCACAA	CTACAATACT	ATGTGCCAAG	
USP9Y-TMUgenome	-----	-----	-----		A.....	-----	-----	T....	T....		
USP9Y_mouse		..G.....	C.....	C.....	A.....			T...		...A....	
USP9Y-rat	...C.....	C			A.....		G...	T...	G....		
USP9Y-TMU-mRNA	AGGATTTTGG	CAGCAATTCA	GGCTTTGGGG	TGAACCTGTT	AATCTTCGTG	AGCAACATGA	TGCATTAGAG	TTTTTTAATT	CTTTGGTGGA	TAGTTTAGAT	
USP9Y-TMUgenome	-----	-----	-----	-----	A..	.A.....	.A.....A.	-----	-----	-----	
USP9Y_mouse		.A.....		C..							
USP9Y-rat		.A.....		..G.....	C...						
USP9Y-TMU-mRNA	GAAGCTTTTA	AAGCATTAGG	ACATCCAACCT	GTGCTAAGTA	AAGTTTTTGG	AGGATCATTT	GCTGATCAGA	AGATTTGCCA	AGGCTGCCCA	CATCGGTATG	
USP9Y-TMUgenome	-----	-----	-----	-----	CC...	C...	A.	-----	-----	...A....	
USP9Y_mouse		...T.....	.T.....	...T.....						...A....	
USP9Y-rat		...T.....	C.....		C...	C...	A.			...A....	
USP9Y-TMU-mRNA	AATGTGAAGA	ATCTTTTACA	ACATTAAATA	TAGACATTAG	AAATCACCAA	AATCTTCTTG	ATTCTTTGGA	ACAGTATATT	AAAGGAGATT	TATTAGAGGG	
USP9Y-TMUgenome	A..	-----	-----	-----	-----	-----	-----	C		T..	
USP9Y_mouse			G					G.....G..			
USP9Y-rat			..T....CG				C....				
USP9Y-TMU-mRNA	TGCAAAATGCA	TATCACTGTG	AAAAATGTGA	TAAAAAGGTT	GATACAGTAA	AACGCTTACT	GATTAATAAA	CTACCTTCTG	TTCTTGCTAT	CCAACTGAAA	
USP9Y-TMUgenome	-----	..T.....	A....	-----	T....	-----	-----	-----	C.....	-----	
USP9Y_mouse						T.			A....		
USP9Y-rat						.G..C....					

Internal
stop codon

USP9Y-TMU-mRNA	CGATTTGACT	ATGATTGGGA	AAGAGAATGT	GCAATTAAAT	TCAATGATTA	CTTTGAATTT	CCAAGAGAAT	TGGATATGGA	GCCTTATACA	R TAACAGGTG
USP9Y-TMUgenome							A			G.....
USP9Y_mouse				C.....					A.....C.....	G..G.....
USP9Y-rat	..T.....			C.....				..A.....	A.....	G.GG.....
USP9Y-TMU-mRNA	CCACAAAGCT	GGAAGGTGAC	AGTATAAGTC	TACAGACTCA	ATTGATTAAG	CAGAATGAGC	AGTCTGAAAG	TGTGATACCA	GGAAGCACAA	AATATAGACT
USP9Y-TMUgenome										
USP9Y_mouse			...G...A..	C.....	A			...C.....		...C.....
USP9Y-rat			G..G..CA..	C.....		..A.....	A	...G.....		...C.....
USP9Y-TMU-mRNA	TGTTGGTGTA	CTTGTTTACA	GTGGTCAAGC	AAATGGTGGA	CATTATTATT	CTTATATCAT	TCAAAGAAAT	GGTAAAGACA	GTAAGAGAGG	TCGCTGGTTC
USP9Y-TMUgenome					 C					
USP9Y_mouse	G	... A				...C.....			 A ..	.. A
USP9Y-rat	G	... A	...C.....				G..		A...T.....	
USP9Y-TMU-mRNA	AAATTTGATG	ATGGAGATGT	AACGGAATGT	AAAATGGATG	ATGATGAAGA	AATGAAAAAT	CAATGCTTCG	GTGGAGAGTA	TATGGGAAAA	GTTTTTGATC
USP9Y-TMUgenome			 A ..	 A ..	G				 T	
USP9Y_mouse			...A.....				..G....T.		G..	
USP9Y-rat		C..	...A.....				..G....T.	...G.....	G..	
USP9Y-TMU-mRNA	ATATGATGAA	GCGGATGTCA	TACAGGCGTC	AGAAAAGATG	GTGGAATGCT	TATATTCTTT	TTTATGAACG	GATTGATATA	ACTGATGAAA	ATGATGAGAT
USP9Y-TMUgenome			 T ..							
USP9Y_mouse		...T.....				..C.....		...G.....	G	
USP9Y-rat		...T.....						...G.....	G	..A.....
USP9Y-TMU-mRNA	GATAACATAT	ATATCAGAGC	TAACAATCAC	CAGACCTCAT	CAGATTATGC	CACCTGCCAT	TGAGAGAATT	GTGTGGAAAC	AGAATGTGCA	GTTCTGTCAT
USP9Y-TMUgenome				... G			 A .. G ..			A
USP9Y_mouse	A.....		...T.T..		T	...A.....	G..	..A.....		A.....
USP9Y-rat		T.G.....	...T..		T	...A..T..	G..	..A.....		A.....
USP9Y-TMU-mRNA	AACCAAAATG	AGTTTAGCTT	AGAATATTTT	CAGTTTATTA	AAAAACTGCT	TACATGTAAT	GCTGTGTAAT	TAAACCCGTC	TCCAGGACAA	GATCACTTGC
USP9Y-TMUgenome						A				... A ...
USP9Y_mouse								...G.....		
USP9Y-rat	..T.....	..A.....					...A.....			A.
USP9Y-TMU-mRNA	TTCCAGAAGC	AGAAGATATT	ACAATGATCA	GCATTCAGCT	TGCTGCTAGA	TTCTCTTTTA	CTACTGGATT	TCATACTAAG	GAAATAGTCC	GTGGACCTGC
USP9Y-TMUgenome									A	A T .. T
USP9Y_mouse	..A.....	C			...T.....				A...A...	
USP9Y-rat	CA.....	A					..A.....		A...T...	..C.....
USP9Y-TMU-mRNA	CAGTGACTGG	TATGATGTGC	TGTGTATTCT	CCTCCGTCAC	AGCAAGAATG	TACGTTTTTG	GTTTGCTCAT	AATGTCCTTT	TTAATGTATC	AAATCGCTTC
USP9Y-TMUgenome			... G							
USP9Y_mouse	..A.....	...C.CA..	..T.....				...T.....	C.		...T...
USP9Y-rat	..A.....	...C.C..	...C.....			..G.....		C.		...C...T...
USP9Y-TMU-mRNA	TCTGAGTACC	TTCTGGAATG	TCCAATTGCA	GAAGTCAGAG	ATACATTTGC	AAAGCTTATA	GTATTTATTG	CACATTTTTC	TTTGCAAGAT	GGATCTTCCC
USP9Y-TMUgenome		... A								
USP9Y_mouse			...G.....	...A.....	G.....					...T...
USP9Y-rat			...G.....		G.....		...T.....			--
USP9Y-TMU-mRNA	CTTCACCTTT	TACATCACCT	TTTGCAACTC	CAGGACCTTA	TAGTCAGATC	TATAACAATT	TAAGCTTGAG	CGATCATTTA	CTAAAAACAG	TATTAAGTCT
USP9Y-TMUgenome						C .. G				
USP9Y_mouse			...A.....			...G.....		T.....	...G...	
USP9Y-rat	-...T....	..G.....				..G.....	..G..T....	G	...G...	
USP9Y-TMU-mRNA	CTTGAGAAGG	GAAGTTTCTG	AACATGGGCG	ACATTTACAA	CAATATTTC	ATTTATTTAT	AATGTATGCC	AGTTTAGGTT	TGGCAGAAAA	GACCCAGCTT
USP9Y-TMUgenome	 A		... T							
USP9Y_mouse								..C.....		
USP9Y-rat										
USP9Y-TMU-mRNA	CTAAAACTGA	ATGTACCTGC	AACCTTTATG	CTTGTGCTTT	TGAATGAAGG	ACCAGGCCCC	CCAATTCAAT	ATCAATATGC	TGAATTAAGC	AAGTTACATT
USP9Y-TMUgenome						... T			... A	
USP9Y_mouse		...T.....			..G..G...	... T .. T	...G..A...		...G..G...	
USP9Y-rat		A.			..G.....	C....T...T	...A...		...G..G...	...C.....
USP9Y-TMU-mRNA	CAATAGTGTC	TCAGCTGATC	CGATGTTGCA	GTGTCATC	AAGAATGCAG	TCTTCTATTA	ATGGTAACCC	CCCTCTTCCA	AATCCATTTG	GTGACCCAAA
USP9Y-TMUgenome			... A		... C		... A .. A		... A	
USP9Y_mouse	..G.....		A.....			...C.....				
USP9Y-rat	..G...A...					..C...A..	...T...			
USP9Y-TMU-mRNA	TTTGTACAAA	CCTATAATGC	CAATTCAGCA	GCATGTGGCT	GACATTTTAT	TTATGAGAAC	CACTTATATG	AAGAAAGTTA	TTGAAGACTG	CAGTAACTCA
USP9Y-TMUgenome										
USP9Y_mouse	..A.....	..G.....		AA.....			...C...			
USP9Y-rat	..A.....G			..A.....			...G...	..A..A...		
USP9Y-TMU-mRNA	GAAGAACCAG	TCAAATTACT	TCGATTCTGC	TGCTGGGAGA	ATCCCCAGTT	TTCATGTTCT	GTTCTCAGTG	AACTGCTGTG	GCAGGTGCA	TATTCATATG
USP9Y-TMUgenome	... G									
USP9Y_mouse	..G..TA...		..TC.....				...T..C..	...A...		C...C...
USP9Y-rat	..G...A...		...C.....				...C.....	...A...		
USP9Y-TMU-mRNA	AATATGAACT	TCAACCATAT	TTGGATCTAC	TTTTGCAAAT	TATTCTGCTT	GAAGACTCCT	GGCAAGCTCA	CAGAATTCAT	AATGCACTTA	AA-GGAATTC
USP9Y-TMUgenome						... A ...			... C	... A
USP9Y_mouse	C.....				...T...	..G.....				...-
USP9Y-rat	C.....	...G.....C								...-...A.
USP9Y-TMU-mRNA	GAAATGATCG	AGACGGACTG	TTTGATACAA	TCCAGCGCTC	CAAAAATCAC	CATCAAAAAA	AAGCCTATCA	ATGTATAAAA	TGGATGGTAA	CTCTCTTTAA
USP9Y-TMUgenome	 T								... T	
USP9Y_mouse	C.....A	...T.....		...A...	...T...		G.....			
USP9Y-rat	C..G.....	...T.....		...A..T	T.....T	...G...	G.....			

USP9Y-TMU-mRNA	CAGTTGTCCT	GTTGCCTACC	AGATCTTACA	GGGTAATGAA	GACCTTAAAA	AAAAGTGGAC	TTGGGCTGTG	GAATGGCTTG	GAGATGAACT	TGAAAGAAAA
USP9Y-TMUgenome										
USP9Y_mouse			G..	G..		.T.....	A..			
USP9Y-rat	...C.....			...G...G.		C.....			...C.....	G
USP9Y-TMU-mRNA	CCATATTCTG	GCAATCCTCA	GTACACTTAC	AGCAACTGGT	CTCCTCCAGT	ACAAAGCAGC	GAAACAGCAG	ATGGTTATTT	TTTAGAGAGA	TCTCAGAGTG
USP9Y-TMUgenome	A.....							...C.....		
USP9Y_mouse			A.....			AT	A		A.	T...
USP9Y-rat			C.....		A.	AT	A			T...
USP9Y-TMU-mRNA	CAAGGATGAC	ACTTGCAAAA	GCTTGTGACC	TGTGTCTGGA	AGAGGACCCT	GATGACCAAG	ATGCCCCAGA	TGAGCATGAG	TCACCAGCAC	CACAAGATAG
USP9Y-TMUgenome	...A.....			T...	...A...CA..			TC..	...A...T.	AT...
USP9Y_mouse	...AA.....A	...A.....	T.	...A...CA..			TC..	...A...T.	AT...	
USP9Y-rat	...A.....		T.	...A...CA..	A..				T...	A..
USP9Y-TMU-mRNA	AACATTTTAT	CTTTATTTCG	ACAGATCTCA	GTATCAACAG	AATTATGTAC	CTGAACAGCC	ATATTCAGGT	CCAGCAGCAC	ATTATTTGAA	TAAGCCTCAG
USP9Y-TMUgenome										
USP9Y_mouse		A.		C.....			G.T.....	T..	..C.C.....	..C.....
USP9Y-rat		T.	.T.....		G.			TT.A...A..	..C.C.....	.G.A.....
USP9Y-TMU-mRNA	AAAAAAAAAGA	TGAACCACAA	GAAGCTCAGG	AAAGCAGTGA	AGAAATATCT	TCATCTCTAA	TAAAGGATCA	GTAA		
USP9Y-TMUgenome										
USP9Y_mouse	T--..	.A.....	..A...T.	A..		...G.....	...A.....			
USP9Y-rat	T--..		..A...T.		GA...		...A.....	..G.		

Internal stop codon

n

USP9Y-TMU-mRNA1	MTITTRDSPV	GENESQGVVP	DGQPQPSFQQ	NQ-----	-NSPATPPDE	QGQGDAPPQH	EEEDPSFPHA	DLAKLEDMIN	RPQWVVPVLP	MEELEVLLLEA
USP9Y-TMU-mRNA2				-----	-.....					
USP9Y_mouse	G...	TS		..ISSDSSN	ET..T...Y.		T		..SR.....	KG.....
USP9Y-rat	..V...G...	A.	...S.....	..ISSDSSN	E.....		T...		..R.....	KG.....
USP9Y-TMU-mRNA1	SMDLTKKGLD	AKSEACQRF	RDVLTVSFTK	ILMDEAVTRW	KSEIHRCIIN	NMHLVELCV	AKLSQDWFPF	LELLAMALNP	HCKPHVYNGA	RPCESVSSSV
USP9Y-TMU-mRNA2										
USP9Y_mouse	.I.....	V.....	S.	SG.	.F.....	.T.....		I.		
USP9Y-rat	.I.....I.			..T...SG.	.F.....	.T...M...		I.		
USP9Y-TMU-mRNA1	QFPEGDLFAC	SPDLHSPKGW	LVDLINTFGT	LNGFQILHDR	FTSGSALNVQ	TIAALIKPFG	QCYEFLTQHT	LKKYFVPVIE	VVPQLLQKLS	NEELKKEAKS
USP9Y-TMU-mRNA2										
USP9Y_mouse	...DE.....							..R...I...	...I...T	T.T
USP9Y-rat	...DE....	...R....			..I.....		..C.Y....	...I....	T	T
USP9Y-TMU-mRNA1	EVKND AISVI	IKFLKNLASR	IPQEETVKD	LEIFRLKMIL	RWLQISSFSG	KMNALNEINK	VLSSVS Y-TH	QHGNFEEEWL	TAEQMAEWIQ	QNNILSIVLQ
USP9Y-TMU-mRNA2										
USP9Y_mouse	T..M.		N	..T.....	.L...N.		Y.	G...S....	.V.R.T...	
USP9Y-rat	M.		N	..L...N.			Y..	R...S....	..R.....	
USP9Y-TMU-mRNA1	DSLHQPYQVE	KLEKILRFII	KEKALTQLDL	NNIWAAQAGK	HEAIVKNVHD	LLAKLAWNFS	SEQLDHLFDC	FKASWMNASK	KQREKLELEI	RRLAEDDKDG
USP9Y-TMU-mRNA2										
USP9Y_mouse		.I.....V.					P.....	T...		
USP9Y-rat		V.				D.	P.....	T...		
USP9Y-TMU-mRNA1	VMVHKVLNLL	WNLASDSDVP	VDIIDMALNA	HLKILNYS	CS	QDRDTQKIQW	IDCFIEEFRT	NNKWVIPALK	QIKEICSLFG	EAPQNL SQTH
USP9Y-TMU-mRNA2				HLKILNYS	F					
USP9Y_mouse	..A.....		..M.L..S.	.I...D.						
USP9Y-rat	..AQ.....		..M.L..S.	.I...D.						R..H.....
USP9Y-TMU-mRNA1	LINQLQRNHA	LVTLV AENLA	TYMNSIRLYA	KDHEDYDPQT	IRLGSRYSHV	QEVQERLNFL	RFL LKEGKLR	LCASQAKQIW	NCLAENAVYF	TDREACFMWY
USP9Y-TMU-mRNA2										
USP9Y_mouse	..S...H...		A.....	R.....	V.....		D.Q.W	..V.....		S.....
USP9Y-rat	..S...H...			R.....			D.Q.W		H.....	S.....
USP9Y-TMU-mRNA1	SKLMGDEPEL	EPDLNKEFFE	SNVLQLDPSL	LTENG I KCFE	RFFKTVNWRE	GKLMIKRVY	MMNDLDLIGL	DYLWKVVIQS	NDNISSRAID	LLKEIHTSLG
USP9Y-TMU-mRNA2										
USP9Y_mouse	D.	H..I.....		..M.....	C.	I.	..D.....		..D.....	Y....
USP9Y-rat	D.	D..I.....		C.			..D.....		..D..N...	Y....
USP9Y-TMU-mRNA1	PRLQANQVVI	HEDFIQSCFD	RLNASYDTLN	VLDSDKDNIF	SCARQEAIQM	VRILT V LREY	INECDSGYHV	ERMILPMSRA	FRGNHLSFTV	RFPNQGKEVD
USP9Y-TMU-mRNA2										
USP9Y_mouse	.K.....		..K.....C	...E.....	R.		.S.Y..D.E		..K.....	E
USP9Y-rat			.MKS.....H		RI		.S...D..E		..K...L..	S....
USP9Y-TMU-mRNA1	LDLIMSH TNA	TIGSVRRCIL	NCMNAVVAHT	EIELFIDGEL	IDSEDDRKL V	EQLNLKDKSL	ITAKYIQINS	IMPSSPDSSS	DSSAESPGNH	SH--YRDVPN
USP9Y-TMU-mRNA2										
USP9Y_mouse	...L.....		.R..V....	K....G...	VA.....		...F.....	N.....	...GP....	..NN...S.
USP9Y-rat	...L.....		.RI...T.	K....G...			...FL....	N.....	...G....	C.NN.....
USP9Y-TMU-mRNA1	PEMEKCLPGV	IMSLQPRYIS	FLWQVADLGS	MLTMPSLRDG	ARILMKLMPP	DSTIIEKLRA	VCSDDLNLGE	SQLGQSLHSL	FFGSSASEAL	YLTEVVYIIL
USP9Y-TMU-mRNA2										
USP9Y_mouse				..V.T...		...TL.Q...	L...HV...	RR.....	QV.	T..
USP9Y-rat	S...N.....			..T...		.N.T..N..	...H.....	RR.....	KV.	
USP9Y-TMU-mRNA1	MPAGGLADI	SYDFQYHFLK	SGGLPLVLNM	LIQNNFLPNT	DVETRDAYV	CALKIAKLLL	TTIGYGHVQA	IAEACQPVAD	DTDPKTPINQ	VTHDQAVVLQ
USP9Y-TMU-mRNA2										
USP9Y_mouse	...A.....	.S.....	S.		F	S.....	.IV.....		G.....	
USP9Y-rat	...AH.....	.S.....	N.....S.		F	S.....			G.....	A.....

USP9Y-TMU-mRNA1	NALQSI PNPS	SECILRN VSM	RLAQQISGLA	SRYIPDICVI	RAIQKIIWAT	CGGSLELVFS	PNEDITETYK	MTTSTRGNPE	VKDEQVCCEA	LEVMTLCFAL
USP9Y-TMU-mRNA2					K.....					
USP9Y_mouse		...M....V	H.....		A			S.L.		
USP9Y-rat		...M....I	...H.....		TS		..H....K..	S...		
USP9Y-TMU-mRNA1	IPTALDSLK	EKAQSFVID	LLLYCPSKTI	RQLAQEQFFL	ICTRCCMGHR	PLLFFITLLF	TILGSATAEK	GKHSDEVYFPL	LRLLTYAYN	RNIHIPNVDV
USP9Y-TMU-mRNA2										
USP9Y_mouse	...M...N.		V				...G.AN.	T.		S..QV....
USP9Y-rat							AT.			S.....
USP9Y-TMU-mRNA1	LLNEEIDWLK	RVRDYIKNTG	ETNIEDPVLE	GHLGVTKELL	SFQTPEEKYH	IGCKTGGANL	VKELIDDFIF	PVSKAYLQYM	RSGELPIKQA	VPVCGSPATV
USP9Y-TMU-mRNA2										
USP9Y_mouse	...D.....		...V...I.		...S..K..		Y...	..A.....		I.....I
USP9Y-rat	...D..N...		..SV...I.		...DK...	...A.....		..A.....S	...M.N..T	I.....
USP9Y-TMU-mRNA1	NAGFELLVAL	AFGCVRNLEQ	IVNCLTELFY	IGTSVTTCEA	LAWEYLPPV	ESRPPKGFVG	LKNAGATCYM	NSVIQQLYMI	PSIRNSILAI	ESICNDTDGD
USP9Y-TMU-mRNA2										
USP9Y_mouse		K.		...P.....	VG.....	GP.....				D..WS...D.
USP9Y-rat		K.	M.	...P.N.---		GP.....				..N.W....D.
USP9Y-TMU-mRNA1	IFKDGKHDSE	NNADPRDDVF	GYPHQFEDKP	TLSKVEDRKE	YNIGVLKHLQ	ITFGHLAASQ	LQYYVPRGFW	QQFRLWGEVP	NLREQHDALE	FFNSLVDSLD
USP9Y-TMU-mRNA2										
USP9Y_mouse	...GE.Q...	...V.....	R.....		...A.....		K...			
USP9Y-rat	T...ERQ...	...I.....		...N.A...D						
USP9Y-TMU-mRNA1	EAFKALGHPT	VLSKVFGGSF	ADQKICQGCP	HRYECEESFT	TLNIDIRNHQ	NLLDSLEQYI	KGDLLEGANA	YHCEKCDKKV	DTVKRLLIKK	LPSVLAIQLK
USP9Y-TMU-mRNA2										
USP9Y_mouse	Y...	...L.....			...V.....	V				...T...
USP9Y-rat		L...L...			...V.....					
USP9Y-TMU-mRNA1	RFDYDWEREC	AIKFNDYFEF	PRELDMEPYT	VTGATKLEGD	SISLQTQLIK	QNEQSESVIP	GSTKYRLVGV	LVHSGQANGG	HYYSYIIQRN	GKDSKRGRWF
USP9Y-TMU-mRNA2				I.....						
USP9Y_mouse				..A.....	..VNP...					SH..
USP9Y-rat				..A.....	GVHP.....	N.V.				...NN....
USP9Y-TMU-mRNA1	KFDDGDVTEC	KMDDDEEMKN	QCFGGEYMGK	VFDHMMKRMS	YRRQKRWWNA	YILFYERIDI	TDENDEMITY	ISELTITRPH	QIMPPAIERI	VWKQNVQFLH
USP9Y-TMU-mRNA2										
USP9Y_mouse			E			M...	...D..I...	...F....	...S.....S	
USP9Y-rat			E			M...	...DN....	L.....	...S.....S	
USP9Y-TMU-mRNA1	NQMQSLEYF	QFIKKLLTCN	AVYLNPA PQG	DHLLPEAEDI	TMISIQLAAR	FLFTTG FH TK	EIVRG PASDW	YDVL CILLRH	SKNVRFWFAH	NVLFNVS NRF
USP9Y-TMU-mRNA2										
USP9Y_mouse			...S.....		S.		K.I...N..	..A.....	V.	
USP9Y-rat			...P.....				K.....N..	..A.....		
USP9Y-TMU-mRNA1	SEYLLECPIA	EV RDTFAKLI	VFIAHFSLQD	GSSPSPFTSP	FATPGPYSQI	YNNLSLS DHL	LKTVLSLLRR	EVSEHGRHLQ	QYFNLFIMYA	SIGLAEKTQL
USP9Y-TMU-mRNA2										
USP9Y_mouse	S.	..I.G.....		S.	..N.....	..D.....	..A.....			
USP9Y-rat	S.	...G.....		...-S.A.		..D.....	..A.....			
USP9Y-TMU-mRNA1	LKLNVPATFM	LVSLNEGPGP	PIQYQYAE LS	KLHSIVS QLI	RCCSVSSRMQ	SSINGNPPLP	NPF G DPNLSQ	PIMPIQQHVA	DILFMRTTYM	KKVIEDCSNS
USP9Y-TMU-mRNA2										
USP9Y_mouse		...D.....	..VK.....	...V.....				...N.....		
USP9Y-rat	T...	...D.....	..K.....G	...V.....				...N...	V	...I.....
USP9Y-TMU-mRNA1	EEPVKLLRFC	CWENPQFSCS	VLSELLWQVA	YSYEYELQPY	LDLLQIILL	EDSWQAHR IH	NALKGIRNDR	DGLFDTIQRS	KNHHQKKAYQ	CIKWMVTLFN
USP9Y-TMU-mRNA2										
USP9Y_mouse	..DT...L.			H.HA.....	F		...P..Q	H.	...R...	
USP9Y-rat	..T.....			..A.....			...PS..	...HF	...R...	
USP9Y-TMU-mRNA1	SCP VAYQILQ	GNEDLKKKWT	WAVEWLGDEL	ERKPYSGNPQ	YTYSNWSPPV	QSSETADGYF	LERSQSARMT	LAKACDLCLE	EDPDDQDAPD	EHESPAPQDR
USP9Y-TMU-mRNA2										
USP9Y_mouse		..G...N...	..M.....			..N...N..	..K.H..K.K	..T....YP.	L.	..V.H....
USP9Y-rat		..G.....		...H	I	..N...N..	...H.....	P.	..T.....	...E.
USP9Y-TMU-mRNA1	TFYLYSHRSQ	YQQNYVPEQP	YSGPAAHYLN	KPQK KKMNHK	KLRKAVKKYL	HL*				
USP9Y-TMU-mRNA2										
USP9Y_mouse	H		F...S.H..	N...NDKPQE	THESNEEISS	C.IKDQ*				
USP9Y-rat		...S.....	...LTE.H..	E...NDEPQE	THESSEEESS	S.IKDQ*				

O

RBMY-TMU	ATGGCAGAAA	CTCACCTGCC	TGGGAAAATT	TTTGTAGGAG	GCCTCAACAT	AGAGACCAGA	CAAAAGAGAC	TTCAAGTAAT	ATTTGAGAGA	TTTGGACCAG
Rbmy-mouse		..A.T.A...		...A.....		..A.....	CT.	A...	G....	
RBMY-rat		G.G.G.A...		...A.....	T.		CT.		G....	A
RBMY-TMU	TATTAAATGT	TATTTTGTATG	AGAGATCGTG	AAACTAAGAA	GTCAAGAGGC	TTTGCTTT SA	TTACTTTTAA	ACATGCTGCA	GATGCTAAGA	ATGCTGTCAA
Rbmy-mouse	..GC.CG...			...T.....		CC	CCG	..GCCT...		
RBMY-rat	..AC.C.....			...T.....		CG	C.	..TG.C...G.		
RBMY-TMU	AAAAATGAAT	GGAGTGTCTT	TGGATGGAAA	AAGAATTAAA	GTAGAACAAG	CCA R GAGACC	ATCATCACTT	GAACGTGGTA	GAAAGCAGAG	ACCACAATCC
Rbmy-mouse	.G.....	AT..			...A.....	..G.....		..A.....	..C...A....	C....
RBMY-rat	.G.....C					..GA.....		..A.GA...	..C...AG...	...TT...
RBMY-TMU	TTTTTAAGA-	-----	-----	-----	-----	-----	GCAAGAGGTT	GTCCTTCACG	TGGAGGGAAT	TTGGATGGTG
Rbmy-mouse	C....A	CCAGAGGTGC	TTCAAGAATT	CTGAAATGTG	GAAGAGGAGG	AAGAAGCAGA	A..G	T.	..A.....C	...G....
RBMY-rat	C....A	CCAGAGGTGC	TTCAAGAATT	CTGAAATGTG	GAAGAGGAGG	AAGTAGCAGA	.G...A..C		..A.A.A...	A..
RBMY-TMU	ATAGATATAC	T-----	GACATGAGTT	CTTCTGGGAG	ACACTTTTACA	GTTAAAAGAA	GTTCATCTTC	AAGAAGGAAT	GGTGGTCCTT	CTAAAAGACC
Rbmy-mouse		..CCTAATTTC	A.TG.....		G..		A.C.....	..A...G..	..A.CCC....	T.
RBMY-rat	G.....	..CCTAATTTC	A...A....		G..		..A.....	..A...G.A	..CC.....	T.
RBMY-TMU	TGCAACTTCT	GCTCTGACAA	G---CAATTC	TGGACTGAGA	GGACAAGAAC	CACGTGGAAG	ACAGATTTCT	AGAAATATGC	CAAGAAGAGA	ACCAGTGGCT
Rbmy-mouse		...A.....	..AAG...A.	A...	...G.....	...A.C....	..G.....		...G.....	G..G.C.T..
RBMY-rat			..AGC...C.		...G.....	...A.....	..G..G.G...			G....AT..
RBMY-TMU	TCCAGAACAG	ATGAGTATCC	ATTACCAAGG	GATGATGACC	AGTCTTCTAA	TGATAGTTAT	TCAAGGAGAT	ACGCATCAAC	ATCCAGAGGA	TACAGTGACT
Rbmy-mouse	G..			...T...G..		G---	-----A..	..T.A.....		...T.....
RBMY-rat	G..	...A.....		GT	..T.....			..T..G....	...T.....	...C.....
RBMY-TMU	ATGGCAATTA	CAGCTCTCAG	GAGGAACATA	CATCTGAAGT	CTTCAGTGAT	CATGCTGGCT	ATCGTGGGGG	ACATAATAGA	GATTTCCTCG	ACTATTTAAG
Rbmy-mouse		..CA.....G.	..A...AG.G	A...			...T...A..	..G.G.....		..A.....
RBMY-rat		G.	..T.....G	...A....			..C.A.....	..G.G.....		..A..C.....
RBMY-TMU	TGGAAGTTCA	TACAGAGATA	CAGATGGAAG	TTACGGAAAC	TTCCGTGAGG	CACCATCTGC	AAGATGGACT	TACACTGGAA	ACAACCACTA	TGATGATTAT
Rbmy-mouse	A....		..T.CA....	...T...GA	A....		...G.T---	-----	...G...	C...
RBMY-rat			..T...A...	...T...G.	...A....	A.....	...G.C...		...G...	C...
RBMY-TMU	AGTAGCTCAC	GAGATGGATA	TGGTGGAAGA	CGAGAAATAT	ATTCAAGCGG	CCAAAGTGAC	ATATATTCAA	GTGGCTATGA	ACATAGTGGT	AGA---GAAG
Rbmy-mouse	...A.....	A.....		G....CC..	..CAT...AA	..G...A..	C....	..A.....	..G.....	...CAA....
RBMY-rat		...G.....	G..	CT.	..C...G..A.	..G.....	----	-----	..G.....	...C....--
RBMY-TMU	TGCTTCCACC	T---GTTGAT	AGGGAATACT	TTGAGCATGG	AGGCAGACAA	GGAAGAGGGC	ATTCACCTAT	GGATAGGCTG	TATATTGCAT	CTCGTGAAAC
Rbmy-mouse		ACCTA....		...C.G..A	G.....G	..A.....	A	...G.....	...G.....	T.
RBMY-rat		---A.....		C....G...	G.....G..	..A.....	..C.....		...GG....	...A...T.
RBMY-TMU	ATATAATAGT	AGT-TCAAGA	TTTGGGGGAT	CCCATGGAGA	TCATGCAGAA	AGCAGATATG	AAAGTGGAGG	CAGGTACTAA		
Rbmy-mouse	...C.G....	..A.AC....		G	..G..G....	-				
RBMY-rat	...C.C..T.	...-.....G		A.....G	A....	..C.....		CGT...		

p

RBMY-TMU-mRNA1	MAETHLPGKI	FVGGLNIETR	QKRLQVIFER	FGPVLNVILM	RDRETKKSRG	FA F ITFKHAA	DAKNAVKKMN	GVSLDGKRIK	VEQA R RPSSL	ERGRKRPQS
RBMY-TMU-mRNA2						... L			... K	
Rbmy-mouse	...NQ....	..I.....K..	..T...E..G.	...AR....		...L..RRL.	E..	...I.....	..K.....	..S.S.K..P.
RBMY-rat	...SEQ....	..I....L...	..T....G.	...ITH....		...FV..QCPG	E..		...R.....	..SS.R..L.
RBMY-TMU-mRNA1	FLR-----	-----	ARGCPSRGGN	LDGDRYT---	DMSSSGRHFT	VKRSSSSSRN	GGPSKRPATS	ALT--NSGLR	GQEPGRGRQIS	RNMPREPVA
RBMY-TMU-mRNA2	...-----	-----		---				-----		
Rbmy-mouse	..S..TRGASRI	LKCGRGGRSR	..SG..CE..	..G.....PNF	NV.....A	..NP..K.D	DP....S...	..Q.RS.T...	..R..HR.E..	...G..AS
RBMY-rat	..S..TRGASRI	LKCGRGSSSR	G.SR...ER.	..DG...PNF	NI.....A	...T..K.E	..P....S...	...RA.P...	..R..H..EV.	S
RBMY-TMU-mRNA1	SRTDEYPLPR	DDDQSSNSDY	SRRYASTSRG	YSDYGNYSQ	EEHTSEVFS	HAGYRGGHN	DFSDYLSGSS	YRDTDGSYGN	FREAPSARWT	YTGNNHYDDY
RBMY-TMU-mRNA2										
Rbmy-mouse	..R.....	..YG.....R-	--K.E.....	..C.....H.R	..SA.K....	...L..RD.	...E....N.	...YR...R	..H...G-	--...R....
RBMY-rat	..R.....	..GY.....		..R.....R	D..A.K....	...H..RD.	...E.....	...YE...S	..H.E....G.	...R....
RBMY-TMU-mRNA1	SSSRDGYGGR	REIYSSGQSD	IYSSGYEHSG	R-EVLPP-VD	REYFEHGGRQ	GRGHSPMDRL	YIASRETYNS	SSRFGGSHGD	HAESRYESGG	RY*
RBMY-TMU-mRNA2				-.....-						...
Rbmy-mouse	..N.Q.....	G.P.I.NR.N	...D..R..	..Q.....PI.	...DRE...	E.....K.G.	..S....S.S.	NTKIW.IPWR	SWRKQI*	
RBMY-rat	...G....G	..T..GSR..	---...R..	..Q--...-I.	...S.R...E	E.....	..R..H.S.TI	...Y..G	..E.T.....	..R.

SRY25	AGTGTTCAGC	CCTACAGCCA	CAGGATATCT	TAAGCTCCGG	AGAAGAGACA	AACTTTGAGA	CTGGTGACAA	CTGTCTTGAG	AGCATGGAGA	GCCATGTCAA
SRY5										
SRY6	G			.C		...A				C
SRY15	G			...T						C
SRY20	G						...C		.A	C
SRY9	AA	...A					.A	C	.T	A
SRY12	AA	...A					.A	C	.T	A
SRY13	A	...A					.A	C	.T	A
SRY23	A	...A					.A	C	.T	A
SRY24	A	...A					.A	C	GT	A
SRY11	A	.A	...A				.A	C	.T	A
SRY18	A	...A					.A	C	.T	A
SRY17	A	...A					.A	C	.T	A
SRY21	A	...A		...T			.A	C	.T	A
SRY14	A	...A		C			.A	C	.T	A
SRY4	A	...A					.A	C	.T	A
SRY19	A	...A			...A		.A	C	.T	A
SRY16	..A	...A					.A	C	.T	A
SRY22	A	...A					.A	C	.T	A
SRY7	A	...A					.A	C	.T	A
SRY1	...C	...A	...A				.A	T	...C	.A
SRY2	...C	...A	...A				.A	T	...C	.A
SRY3		...A		...AA			.A	T	...C	.T
SRY8		...A			...T		.A		...C	.T
SRY10		...A			...T		GA		...C	.T
Sry-mouse			...T	...A	...T	...A	...GT	...G	...T	...A
SRY-rat	..A	T	G	...C	.A	...T	...T	...G	...G	...G

SRY25	GTGCCCTTG	AATGCATTTA	TGGTGTGGTT	CCATGGTGAG	ATGAGCAAGT	TGTCCCAGAA	GAAT-CCAGC	ATGCAAAATA	CAGATATCAG	CAAGCAACAG
SRY5							...-			
SRY6				.C			...-			...AT
SRY15				.C		...A	...-T			...T
SRY20				.C		...A	...-			...T
SRY9	.CA	...A		.C	TG	...G	.CA	...C	...G	...C
SRY12	.CA	...A		.C	TG	...G	.C	...C	...G	...C
SRY13	.CA	...A		.C	TG	...G	.C	...C	...G	...C
SRY23	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY24	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY11	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY18	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY17	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY21	.CA	...A		.C	TG	...G	.C	...C	...G	...A
SRY14	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY4	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY19	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY16	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY22	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY7	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY1	...A			.C	TG	...G	.C	...C	...G	...T
SRY2	.C	...A		.C	TG	...G	.C	...C	...G	...T
SRY3	.CA	...A		.C	TG	...G	.C	...C	...G	...T
SRY8	.C	...A		.C	TG	...G	.C	...C	...G	...T
SRY10	.C	...A		.C	TG	...G	.C	...C	...G	...T
Sry-mouse	.C	...A		.C	G	...CA	...G	...C	...G	...G
SRY-rat	.C	...A		.C	G	...A	...G	...T	...G	...T

SRY25	GGATACAGGT	GGAAAAGCCT	TACAGAAGCT	GAAAAAAAG	GCCCTTTTTC	CAGGAGGCAT	AGAGACTGAA	GGCCCTTCAC	AGAGAGAAAT	ATCCAAACTA
SRY5										
SRY6										
SRY15	...T			...T						
SRY20										
SRY9				...C		...A	...C	...T	...A	...C
SRY12				...C		...A	...C	...T	...A	...C
SRY13				...C		...A	...C	...T	...A	...C
SRY23				...C		...A	...C	...T	...A	...C
SRY24				...C		...A	...C	...T	...A	...C
SRY11				...C		...A	...C	...T	...A	...C
SRY18				...C		...A	...C	...T	...A	...C
SRY17				...C		...A	...C	...T	...A	...C
SRY21				...C		...A	...C	...T	...A	...C
SRY14				...C		...A	...C	...T	...A	...C
SRY4				...C	A	...A	...C	...T	...A	...C
SRY19				...C		...A	...C	...T	...A	...C
SRY16				...C		...A	...C	...T	...A	...C
SRY22				...C		...A	...C	...T	...A	...C
SRY7				...C		...A	...C	...T	...A	...C
SRY1				...C		...A	...C	...T	...A	...C
SRY2				...C		...A	...C	...T	...A	...C
SRY3				...C		...A	...C	...T	...A	...C
SRY8		...T		...G	...C		...A	...T	...A	...C
SRY10	-			...G	...C		...A	...T	...A	...C
Sry-mouse	...G			...C	...-		...T	...AT	...A	...C
SRY-rat	...TCA			...C	...-		...GC	...A	...A	...C

Internal stop codon

SRY25	TAAATATCAG	CCTCGTCAAA	GGGCTAAAGT	GCCACAGAGG	AGTGGCACTT	TGTAGCCTGC	AGTTGCCTCA	ACAAAAGTGT	ACAACATGCT	GCAGTGGGAC
SRY5										
SRY6						T...				
SRY15						T...				-
SRY20							G...			
SRY9		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY12		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY13		...A..GG.	.A.....	...T...	...A.....	...C.....	...A.....	...CG..		
SRY23		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY24		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY11		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY18		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY17		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY21		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY14		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY4	A	...A..GG.	.A.....	...T...		...C.....	...A.....	...GCG..		
SRY19	A	...A..G.	.A.....	...T...		...C.....	...A.....	...A...GCG..		
SRY16		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY22		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY7		...A..GG.	.A.....	...T...		...C.....	...A.....	...CG..		
SRY1		...A..GG.	.A.....			...C.....A		...CG..		
SRY2		...A..GG.	.A.....			...C.....		...TG..		
SRY3		...A..G.	.A.....			...C.....		...TG..		
SRY8	A.....	...A..GG.	.A.....			...C.....		...CA..		
SRY10		...A..GG.	.A.....T.			...C.....		...CG..		
Sry-mouse		...A..GG.	T.		...T...	AC.....		...C.....	...C.T...	
SRY-rat		...A..G.	...T...		...TAT...	...C...G...A		...C.....	...C.....	...A.....

SRY25	AGGAACCCAC	ATGCCATCAC	ATACAGACAA	GACTGGGCTA	GAAGTGTACA	CCTGTCTTCC	AAGAACCAGC	AAAGCTTTTA	TTTGCAGCCT	GTGGACATCA
SRY5										
SRY6					C.....	C...	...C.....			
SRY15					C.....	C...	...C.....			
SRY20					C.....	C...	...C.....			
SRY9					...G...G...	C...		CCAGCAGCAG		ACCAGCAGCA
SRY12					...G...G...	C...				
SRY13	T...				...G...G...	C...				
SRY23					...G...G...	C...				
SRY24					...G...G...	C...				
SRY11					...G...G...	C...				
SRY18			T.		...G...G...	...TC...				
SRY17					...G...G...	C...				
SRY21					...G...G...	C...				
SRY14					...G...G...	C...				
SRY4					...G...G...	C...				
SRY19					...G...G...	C...				
SRY16					...G...G...	C...				
SRY22					...G...G...	C...				
SRY7					...G...G...	C...		A		
SRY1			T.		...G...G...	C...		...C...		
SRY2			T.		...G...G...	C...		...C...		
SRY3					...G...G...	C...		...C...		
SRY8					...G...G...	...CA...				
SRY10					...G...G...	C...				...G...
Sry-mouse			...G...	...AG.	...G...C...	...AC...	...A.....		...G.....	...T...T...C
SRY-rat	AC...T...	CA.T...T	...G...G		...G...C...	...A...C...	A	...A...	...A.....	...C

SRY25	CCACTGG---	-----G	CAGAAGCAGC	AGCAGCAGCA	GCA-----	-----	-----ACA	CCAGCAGCAG	CACAAGTTCC	ACCAGCAGCA
SRY5	---	-----			-----	-----	-----			
SRY6	---	-----			-----	-----	-----			
SRY15	---	-----			-----	-----	-----			
SRY20	---	-----	A.		-----	-----	-----			
SRY9	---	-----	...A.....	...GCAGCAG	CTGCA----	-ACAGCA...		...T...		
SRY12	---	-----	...A...A.	...GCAGCAG	CTGCA----	-ACAGCA...				
SRY13	---	-----	...A.....	...GCAGCAG	CAGCA----	-ACAGCA...				
SRY23	---	-----	...A.....	...GCAGCAG	CTGCA----	-ACAGCA...				
SRY24	---	-----	...A.....	...GCAGCAG	CTGCA----	-ACAGCA...				
SRY11	---	-----	...A.....	...GCAGCAG	CAGCA----	-ACAGCA...				
SRY18	---	-----	...A.....	...GCAGCAG	CAGCAGC---	AACAGCA...				
SRY17	---	-----	...A.....	...GCAGCAA	CAGCA----	-----				
SRY21	---	-----	...A.....	...GCAGCAG	CAACA----	----GCA...				
SRY14	---	-----	...A.....	...GCAGCAG	CAGCA----	-ACAGCA...				
SRY4	---	-----	...A...A.	...GCAGCAG	CAGCA----	-ACAGCA...				
SRY19	---	-----	...A...A.	...GCAGCAG	CAGCA----	-ACAGCA...				
SRY16	---	-----	...A.....	...GCAGCAG	CAGCA----	-ACAGCA...				
SRY22	---	-----	...T.....	...GCAGCAG	CA-----	-ACAGCA...				
SRY7	---	-----	...A.....	...GCAGCAA	CAGCA----	-----				
SRY1	---	-----	...A.....	...GCAGCAG	CAGCA----	-ACAGCA...				...C
SRY2	---	-----	...A.....	...GCAGCAG	CAGCA----	-ACAGCA...				
SRY3	---	-----	...A.....	-----	-----	-----				
SRY8	---	-----			-----	-----	-----G.			
SRY10	---	-----			-----	-----	-----G.			
Sry-mouse	---	-----	...CCT...	...GCAGCAG	CAGCAGC---	-AG//CAG.	G.....	...GC.....	...G.C.....	
SRY-rat	ATA	CCCACTACA.	...A.....	...C.....	...GCAGCAC	GTGCACCTGC	AGCAGCAG..	G.....	...C.....	...T...

SRY25	CCTCACCTAC	TTACTGACAG	CTGACATCAC	TGGTGAGCA-	---TACACCA	TAGCAGGAGC	ACCTCAGCAA	AGCCCTGTGG	TAGGCAGTCT	CATAACACTG
SRY5					---					
SRY6					---					..G.....
SRY15			...T....		---	T				..G.....
SRY20			...T....	..C....	---	..G..		.T...A...		..G.....
SRY9					---			.T.....		..G.....
SRY12					---			.T...A...		..G.....
SRY13					---			.T.....		..G.....
SRY23					---			.T.....		..G.....
SRY24					---			.T.....		..G...T..
SRY11	A				---			.T.....		..G.....
SRY18	A				---			.T.....		..G.....
SRY17	A				---	..T...		.T.....		..G.....
SRY21	A				---			.T.....		..G.....
SRY14					---			.T.....		..G.....
SRY4					---			.T.....		..G.....
SRY19					---			.T.....		..G.....
SRY16					---			.T.....		..G.....
SRY22				..T....	---			.T.....		..G.....
SRY7					---	T.		.T.....		..GG.....
SRY1					---			.T.....		..G.....
SRY2		...T....			---			.T.....		..G.....
SRY3					---		A...	.T.....		..G.....
SRY8					---					..G.....
SRY10					---					..G.....
Sry-mouse	G..T.....	A...			---	..C....	C		.T.....	..G.....
SRY-rat	G..A.....	T.CA	...T....	C.GA	CCC.G....			..A.....A	...T.A...	..G.T...

SRY25	GCCTCTG---
SRY5	---
SRY6	---
SRY15	---
SRY20	---
SRY9	---
SRY12	---
SRY13	---
SRY23	---
SRY24	---
SRY11	---
SRY18	---
SRY17	---
SRY21	---
SRY14	---
SRY4	---
SRY19	---
SRY16	---
SRY22	---
SRY7	---
SRY1	---
SRY2	...G...---
SRY3	---
SRY8	---
SRY10	---
Sry-mouse	...T.T---
SRY-rat	..TCTGCTCC

intron 15

TMU-DDX3Y	AGTTTGAGCATGAATTAATATTTTAATTGGATATCATACTAGATAATGTTTTTTCTTGC
mouse-DDX3Y	AGTT--AACA-GAATTGTTATTTTAATTGGATATCATTGTT-----TTTTTCTTGC
	**** * ** ***** ***** ** *****

mouse → TMU

exon 16

TMU-DDX3Y	AA-AGCAGATTCAGTGGAGGATTTGGAGCCAGAGACTATCGACAGAGCAGTGGTTCTGCA
mouse-DDX3Y	AATAGCAGATTCAGTGGAGGATTTGGAGCCAGAGACTATCGACAGAGCAGTGGTTCTGCA
	** *****

TMU-DDX3Y	AATTCTGGGTTTAATACTAATCGTGCAAACAGCAGCCGAAGTAGTGGTAGTAGCCACAAC
mouse-DDX3Y	AATGCTGGATTTAATAGTAACCGTGCCAATAGCAGCCGAAGTAGTGGTAGTAGCCACAAC
	*** ***** ** *****

TMU-DDX3Y	AGAGGATTTGGTGGAG
mouse-DDX3Y	AGAGGATTTGGTGGAG

upstream-TMU	AAGCTCCCA	AAGACGGGT	TCTCTTAAGT	CATTATCACA	TTCTTTTGGG	ATGGTCCTGT	GAGGAGAATT	ATAAAGGCAG	AGGCTATCCA	TGGAGTAGAC
upstream-rat	...TA.T..	.G...CT...	-..C	T.	..G.....	G.A....C..				...A..A.T
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	AACAGGCAAA	ATAGAAATCA	TGGCCGCAGT	GCACCCATA	CACGTTACCA	TAGGTCTCTA	TTAGGAAAGT	TAGAAAGATG	TCTCAGGTAG	GGAGTCTTCT
upstream-rat	.G.....	.-...TC..	A-...	...G..TG..	...A....T	G	C.....	G.CA.		C...
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	GATACCCTCA	GTGGGAGAAG	GAGAACTCTG	GGTGCTATGA	TCTAGATGAC	TGGCATTGTG	TGGATGGAGT	AAAGAGGCAG	GACTGGGACT	GACATGCCAT
upstream-rat	...G...TG	A.....A	AG.G.....	.A..T....	...G....G.	GC...	..CT....A	..T...A...		
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	TCAGTGACTC	TACTCCCTCA	GTAAGTAGTA	GTCAATTAGT	GTTATTTCAG	TGAGGGAAGT	AGGGGTGCAG	CAAACTGTGA	CTCAGGGGCA	CAAACTGGCA
upstream-rat	A.		.C....G...	G....	C..C.	A.-.	AG..	A.GGG.....	..T...AT.	G.....
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	AGACGCCAAA	ACATGCCCTT	ACCAAAGGAT	AGAAAGCCAA	AAGCCAAGGG	TGTTGGCTCA	ACAACA----	---GTGCCCT	TTCTTGAGCT	GAGTGAGGAA
upstream-rat	..T.....	T..	G...GT....	G..G.....	-.-----	CAG....TG	G..GA.TGGA	CGG.G.T...	G.....	
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	GTCAGAGAGA	GAAAGGCATT	CCCTCCCCCT	AATAGGCAAC	AAGCACTCTC	TTAGGCAGAC	ACCCAGAGAG	GCATCAGTGC	CTCCATCAGG	AGGGATCAGG
upstream-rat	-	T...	T.....T...	TT..	.T.....	A....	.G.A..C...	T....A...	A.	..CA.C....
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	TAATAACAGG	CTACCAGGGA	AGGCACTTGG	GGCACTTTGA	CTCCTGCCTC	AGAGTCCTGG	AGGAC-----	-----CCT	GCCTCAAATG	GAAAGCCCA
upstream-rat	C.....	AC.	T....	A.			ACACT	AATTGCT...	TG..	T.A..
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	GAATATTGGC	CTTGACAGAA	TTCCCTGTGA	AGCCCTCACA	ACCAAGGCCC	CAGTGTAGTT	CTCAGTGTTG	GAACATGTCC	TCACTAGCCA	CCCTGACTCA
upstream-rat	G....		C	CA.....	-----	C....	A....	.G.TG..A..	A..	.A.....
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	GGCACAGGTT	TCAAGTGTC	AGTTTGGGGT	TCTGAGAGTG	ACCCAGATAT	CCCAGCCCCA	CAGTGTGTTG	AAAGGTGCAG	GTTCACTGTC	AGCAC-----
upstream-rat	.A.....A.	A..A.	...C....T.	A	A.G	T...G	TAG.		C....T....	GGTTA
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	-----TCCC	CAACTTCCAT	TCCACCAGTA	CCTTCAAAGC	CCACCCTTTC	CCCACAAGAA	TCACCTTTTC	CTTTTATTGT	GTCCAAAACA	CAGGAAAAC
upstream-rat	TGGAAC....		C...T...G	.C.....	-.TGA....	T.....T.	...G.....	..G.....	C....	
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	G-//	LINE insertion		//--TGCCCC	TCC-TCTAAT	CTTGATTGTA	CATGGCAGCT	TTCAGAGCAG	AGACTGATGA	GCCCTGTCTG
upstream-rat	.CTAGGCCTC	ACCTGGGTGT	CTGG.....	..AC..TGA	..A.....	..CA.....	G-.....	G..T.....	-----	..G...T..
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	GGGCTGCTCC	AGACCTTCCC	TTGCTGATTA	TCAACAAGGT	GTGCTAGCTT	CCTACT----	---CTCCTTC	TTCAAGGATG	GACAAAGACC	AAGCAGGCAA
upstream-rat	T....	...T.....	C	G...A.	...GCC...	TCAA	GGT..A...	TCT	.G.C.....	T.T.C....T
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	GAATTAACAC	ACAATACAGA	ATGCTGACAA	ACAGCTGAGG	CAGGCTGACT	GGAGATTACT	CACAATG---	-----CT	TACTTGACTT	GAGTTGACAC
upstream-rat	...AT....	.G...TGA.	CAG..C	..AT.....	A....G..	..GC.....	G..TTT	CCTTTACT..	A..	...C..G..T
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	TCAGATCTCT	CTGCCTGAAG	TGACCTTGAC	TCAGCATCTG	CTGAACCTCC	-----	-----TGATA	GTCTTC----	-----	-----
upstream-rat		A	G...T...C.	T.C	.C....A..C	TTCTTGCCCA	AGAGC...C.	AA.C.TTCTG	TGCCTTCTTG	GGCCAAGCAG
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	-----GTGC	TCCTCCATTA	TTGAGCCAAG	-----	-----	-----AGCC	AAGAAACCTT	AGCATG----	-CCTTCTTAG	GCAAAGCAGA
upstream-rat	ACAAAT...G	.AGGT.C..C	...G.....	CAGACATATG	TGGTAGGTCC	ATCTTGG...	...C.GA.A.	.TGTG.TAGG	T....T.G.	..C.....
upstream-mouse	-----	-----	-----							

upstream-TMU	AACAACAATC	CTAATAAAAA	ATAATTTTAA	TATTACTATG	TAAGAAATTT	GTACAAATGA	CTTTTATGAT	GTATTGATAT	GACAAAGAAA	GGAAGAAAAG
upstream-rat	TGA..T....	C..	-.T.....G	CA	..G.....	C..	GC.....	.C..GT....	G.	.A.....C
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	TTC-TTGAAA	TCTTTGTTGA	TATTTATTTT	TTGTAGCTTG	AAACTCCAAA	GCTCAGCTTA	TGCATTAGTA	GTGATTTTGT	TTTGTTTTGT	TTTGTGTGTT
upstream-rat	...C....G.	.A.A.....	GT....GG.T	T	C..T..T...	AT....----	--.....C..	C..T.....	...TGG..--	-----
upstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
upstream-TMU	TGTTTTTCTC	TTTCATGGTG	TAGATATTTA	TACTGGGCTA	TTGAGCCAGG	ACAGACTATT	TCTTGGTGTT	TTCCATGCCG	CCTGCCTTAG	TGGACTAGGA
upstream-rat	----.C.T.G	.CC...T...	C....	..G.....	...C-..T..	GA.	--T..	C.T.....A	T..CG....C	
upstream-mouse	-----T..	...T.....			.A.....A		...A..T..	T.	G...	.A....T..
upstream-TMU	TACAAATGAT	AGGTTTCATG	CTTATCA---	TTAACT----	--ACTACTTT	TGCATGCTTT	TTTATTTCTC	ATGTTGGTCC	ATCCTAAATA	TCATGTACAC
upstream-rat	T....	T.T....C.A	T--	..G..CCTCT	CT..C....G	.A.....G		-.A....	..A.....	TG.
upstream-mouse	...T.....	C....	TTT	CCTTT	CC.....	...A...CG	.A.....	.C...A....	C.	.T.G....GT
upstream-TMU	TACAGGGCAA	AA-GAACTCA	TTAGTATCCT	TGTTGTTGTA	TCTTAATCCT	TGTTGTGACA	GAATGATCAA	ATAGAAAGAG	ATATTTTTTT	GAGGAGTGTA
upstream-rat		..A....GGC	G..	CA.....	G..	C..AG	A....C...	.G.....TT	G...G....	A.....
upstream-mouse		..-.....	.C.....GT.	CT.....--	A..			.G.A....C.	...A.....	A.....G
upstream-TMU	AAAAAAAAAG	TCTCTAACAT	TCACAAACCA	GAGGGGTTTT	GGTTGACAAT	GGGTAGGTGT	GTTGTTGTTG	TCTTAAAGTA	GATTAGTTGG	GTTGGACTAT
upstream-rat	...GTGC.-.	-...	.T....T...	C.....		AT.....T.	..-.G..G.	CT.C.....		.C....GTG
upstream-mouse	...GTGC.-.		TT		..C.....G.	ATC....TC	A-.AGA.G.	CT.C.....		.C.....G
upstream-TMU	GAAGGTCCTG	AAAGATGAGG	GACAAATATT	TTCTTGACACA	CTTTAAAAAC	TAAATCTTTT	ATTACAAACA	CATTTTGGAT	AGTAGCTTTT	AGCTCTTGCA
upstream-rat	.G.....C.	.GG..G....	..TG.....	A----	T..		G.....	..G..C....	.C.....	.C.C...A..
upstream-mouse	.G.....	..G..G....	..T.....	A....	.G.....T.T		G.....G..	TC	..G.....	A..
upstream-TMU	CTTTAAGTTT	TGACTTCTGT	CTCTGTCTCT	CTGTCTCTGT	CTCTCTCTGT	CT--CAATCT	CTGTCTCTCT	-----	----CATCT	CATCTCTGTC
upstream-rat	TA.C.	...T.....	G.	..C.G....C.	..G.T.....	...-T....		-----	-----C	A.
upstream-mouse			A.....G.	..C.G....	...G.....	..GT.TG...	T....G...	GTCTGTTTCT	CTTTG....A	
upstream-TMU	TTCTTGTTTG	CAATCATAAT	TCTTCTACTA	TGTTCAGCAT	TTCAGATCTT	GCTTTTTAGT	GTTCAGCCCT	ACAGCCACAG	GATATCTTAA	GCTCCGAGAG
upstream-rat	...C.....	C....	C....	CA.....		CAC...C...	A.....	TG..	.C..A....	T...T....
upstream-mouse	..T.....		C.T..			.A.....		T		A...T.A...
upstream-TMU	AGAGACAAAC	TTTGAGACTG	GTGACAACTG	TCTTGAGAGC						
upstream-rat	..G..-G.G.	G..GCA	GT..	...A..T...						
upstream-mouse	GT	G.....	T..	...A.....						

downstream-TMU	-----	-----	-----	-----	-----	-----CA	GGAGCACCTC	AGCAAAGCCC	TGTGGTAGGC	AGTCTCATAA
downstream-mouse	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-rat	CAGCAGCTAA	CCTACTTACT	GTCCACTGAT	ATCACTGGTG	ACCGACCCCTG	CACCATAG..		A.	...A...T	.A.....G.
downstream-TMU	CACTGGCCTC	TGCTCCTACC	TATGTCAACA	CTCCCCTTTC	CTATAGTTTT	ACAGCCAGAG	ATTCTTAGTT	GATTTGGATA	CTGATTTCCT	TCCTGTTGTA
downstream-mouse	T	.T.....	...C....	C..G	...GA....	TA..T.T...		.G...A...	...C...TT	CA.G
downstream-rat	T.....-		...C....	-..G	..G..A..-	TT..	C.	.G.....	...G.....	..TG.....
downstream-TMU	ATAAATGTCC	CGTTCCTGTT	CCTTCCTT--	TAGGGTT--	-----	CTGAAATGTG	TGTAAAGATG	CATAGTGTGT	AAAGAAGGAG	AATGCTTAGA
downstream-mouse	.A....-	-.G....C.	-	.T.....	-----	.A.....		TG.....	.G.....	..A.....
downstream-rat	G...	T.C....CC.	TT..TT..GG	.G...G.GTG	GGGGAGGAGA	.CA.....	G...	TG.....	..GA.....	
downstream-TMU	AATGCTA-CT	ACATATTTTT	TGTTTGACAC	AGTGCTAGTT	GCATTGTTGT	GGCATTTGAT	TCCACAGAAG	AAACCAGAGG	CTTAGACTGA	GATCCAATTT
downstream-mouse	T..-	G....	A..T	...A.....		A.		A	T...C...C.	
downstream-rat	.GG...TA..	C....	A...G	C..A...C.C	C..A.	A.	-.G.TG...	...T.-..	T.....C.	T.ATG....
downstream-TMU	AGATGAAATG	CATTATTCT	TACCTTAGCA	CAAAGAGAGC	AAACAGCTTT	AGAGCCAATA	CAGCATGCCA	TGCTAAAGTG	GGGGGTAGGG	GTGGGTAAGG
downstream-mouse			..GG.....	C...		GTG.T....			...C.....	...A....C
downstream-rat	.TC..T.C..	-----	----C...T	GC..			...T..A..		..A...G..A	.A...G....
downstream-TMU	TGGAGTTTTA	AAGGAAACAT	CAGAAGGG--	-----ATG	-CAGATATCC	AGCTGGGCAA	GAAAATTCTT	TTCATTCCTG	TTATGTAGTA	ATAAAGACC
downstream-mouse			...T.ACA	CATTGTA...	A..A.....					
downstream-rat			TA	CAC---CT	A..A.....		.G.....		..TGA...G.	G...T
downstream-TMU	ATTCATCCCC	CTTTCCCGA	GTTTAGCAGG	TTTTATAGGA	ACAAAGGCAG	CTGTTAATTC	CTCAGAGCCT	ATTATCCTAT	TCTTATCTCA	CTCTATGGCC
downstream-mouse		...T...C.	.A.A.T...	C....	A...	...G.....		...G.....	T...	...T.....
downstream-rat	AT..	C.	.A.A..A...	.G..C....	...G.....		T..C....C	...G.T...	...T...T.	AA..CCA..A.
downstream-TMU	AATAGGATCT	CCAAGGCATT	CCAAGGCAAA	GGTCAAGGGC	CCTTACTGGA	CGCCTGGCTG	CATATTCAAT	TTCTAGTGCC	ACCACAGGGA	ATATATATAA
downstream-mouse		..C....C		A.	T.....A.	T.....	...G..A..	...TTA...	C..AG	T
downstream-rat	.G.....	TGC.....	...T.....	A	.T.....	T.....A.	...T...A..G	...TTA...	AG...	.C....G...
downstream-TMU	ATT-----	-----	-----	-----	-----ATCA	CTCAGGTTGA	GCCTTCTAAT	GACTGGGATG	CAGCAATTCA	TGTCTGAGCA
downstream-mouse	..ATATATAT	GTGTATATAT	ATATATAATA	TGTAGGTATG	TATATA..	T.....	...A.....	...A.....	...T.G....	.T.....
downstream-rat	..-----	-----	-----	-----	-----	.T.....	.TA.....	...A...A	...T.G....	.T.....
downstream-TMU	GACCAAAACA	CAGTGTCA TG	AAATCAGGGC	TGAT---GT	GTCCCTTATAG	CATCACCATG	GAAACCATT C	TTTTTCTGTT	AGTCTTCAAT	ACTCATCTCT
downstream-mouse	.G.A.....	C...C	GGG.....	..GCTGTCA.	T.....C..		C.....-	-----A		T.
downstream-rat	.G.A.....	C.C.C	.GG..T...	..GCTGTCC.	TC..		G	.C.A....A	..GG...G.	
downstream-TMU	GAAAAACATA	ACACAACCTA	ATGCTGTGTA	ACTACAGGCT	GTGCCTCTTT	AGCATTCAAA	ACACTCCTTT	GTAGATCAAA	GGCAGGTCTC	TGTTTAT-GG
downstream-mouse		C.-.....	..A....A.	.AC.....	..T.....	.T..G....		.T.....	T...TC.T..	CA.
downstream-rat		-----A..	AA.	.AC..T...	C....	G....	GT.....	G.	CA..TC.A..	CA.
downstream-TMU	AGTCTCTCTC	T--AGTACAC	AAAAC TACAT	CATGAGAATG	TAAAGTCT-G	T-AGTGTGTG	TGGGGGGGGG	GATATTT--	-----	-----
downstream-mouse	G..T.....	-.-.C....		T...	T.	.G.....	.T.T.T.T.	TG.G.G.GTG	TGTGTGTGTG	TGTGTATGAC
downstream-rat	G.....	.CC..C....		...ATATGAC	.TCT...A-C	.GT.....CT	.C.T.T.CA	CG.GCG.GCC	AGTGTGCATG	GGTTTGCTCT
downstream-TMU	--TTCATTTC	TCCTCCACAC	TGTGCAAAGT	CCAGTTGTTT	CATGAAATAT	GCTATGAGCT	CTCTATAGTA	TCTTTGGGTA	AAATAAACCT	TTATCTTGTG
downstream-mouse	AT..T.G...	A	C		T.....C..		-	...T...	G...	...-----
downstream-rat	-T...C...	...T.....	.A..AT..AC		TG.A..C.C.	-----T..	...G..A..	.G..T.A..	...G..A..	...ACAA-
downstream-TMU	TTTATGCAGA	AGAATTCACA	CT--GCTTCT	CTTTTATTAT	CTATCTCTAA	AGAATAAATA	GTTTTATTAA	CATAAAAAAT	TTTATGTAAA	AGCACCACAG
downstream-mouse	-----T.	..G.G.....	T.TT.....		T...	G.CT	G.	...C.GC-	-----	-----
downstream-rat	..C....C.	G...G....	T.CT.....	-C	T...	C.	C.	...-G...	...G.A...	...GT....
downstream-TMU	AAAAAGCTAT	GTTTTTGA--	-CTAGGTAAG	AATCAGAAAC	AAAATGAAGA	GAGCA				

[illegible]

downstream-TMU	AATCTGAAAG	GCATAACAAC	TAAGTGAAAG	GAGAAATGAA	AGCAAGAGTA	TGCAGTCTTG	TGTTGGTAAT	AAGGTAAAGA	AGCACCGAGG	GGAAATAAAG
downstream-mouse		GG	...C.....	A..		.TA.....			...T...A...	..G.A.....
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	GACTATCCTA	CTCCATGTAG	GAAGAGCAGG	CCTTTGAGAC	ACAGTAGAAA	CAGCAGACAT	TAGGACTCAG	TGTGGAATTC	ATCTGCAGTT	TCCCTCCCAG
downstream-mouse		T.T..G....			...A.....	G...	C.....		A..	
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	CTCTGTTAAG	ATCAGGCTCT	GGGCCTGCTT	ATAATCTTAC	AGCTTGAGAC	TGAGTGGCCA	TGTTGAAAAG	GCTTAAATG	AACTG----	GGAGGAGAGG
downstream-mouse		C.TG..T...	A...		.A.....		G.....G...		GGGTT	..G..G.G..
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	TGTATGA--A	AAGGGAGGGG	AGTGATTAGT	GGTTAGG---	-GAGGGAATG	AGTGCCTCAG	GTAGATGGAT	CAGGGAGTGG	GCCCAGGCCC	TTTGCCATTT
downstream-mouse	.AGG..GAT.	..A.ATA...	..G.G.....	TGGG	T.....	.T..T....	..G.G.....	A.....		T.....
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	CCTTTGAGAA	AGTTATCACA	GGTTGTCAGG	AGCCATAA--	//	LTR insertion	//--AA	CTGAAAGTTT	AAAAGTGTAG	CCAGAGGTGG
downstream-mouse	C....	.C.....T.	...T.G.A.	...T...--	-----	-----	-----	T.....	A...--	-----
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	CCATTATCCA	AGAATATGGC	CACATTTGAG	TGGAAAAC TG	AGTTTGGTGG	GCTGTAAC TA	ACTCTAAAGA	CAGAGCTTTA	GAGTCTACTT	TTAAATTTGA
downstream-mouse	TG	G....	T.....C...	...T.....	T.....	A..-----	-----	T....	C.TC.	G.C..G
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	CTTAAATTTG	ACATGTCAGG	CAGGATTGGT	AGGGTGTTTT	ATCCCACTGC	CACCCAGGTT	TT-GGCAAAG	ATGAAAATGA	GCCCAGTATG	GAATCCCCAA
downstream-mouse		...A.....	C....			.T..T.A...	.T.....		GA.G..	.C.C...A.T
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	ACTCATAAGA	CTGTGGTGAG	A-TTTGTCAT	GTCAGTCAA-	ACCTGGAAGG	TGCCTAGAAC	CTCCAGGACG	TTGTGATCCA	GAGAAGGAAT	GAGAAGCT--
downstream-mouse		.GA.....G.	.C.....	G	.A...G...		A			CC
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	-----TCT	GGGGTCAGGG	CAGGGCACCT	CTCAGTCAGA	CCAGGATGGA	TGGTTACCTT	GCCATGCAGA	ATTTGTTTGT	GTGTGATGAT	AAGAATCC--
downstream-mouse	CACAGCT...	...A..T..T				CA.....	.T.C....C.	G-...		.GC.....AA
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	-----AAAGC	TCTCTTATCC	TGAGGATCAG	AGACTCTAGG	AAGAAGAGGT	CAGAGAAAGA	GTCTTAGGAA	GATGGATGGA	GAGGACTCAA	AGGATAGCTG
downstream-mouse	TGGAA.....			T.-		.C...G....		G		C.....
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	AAAGGACTGT	GATGACACTG	GGCTGACAGA	GAGTACAAAT	---GTCCCAG	GAAGTAAAT	TGCCCTCACA	GGAAGGCAGA	GGCCACAAGA	GGACAGAGGT
downstream-mouse	.GCAA.....	AG...		T...	GAA.....	..G.....	...T...T.	A...A.....	..T.....	..C....A.
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	CTCAGGTTAC	TCAGCTAACC	ATTAGCCT--	-----GTGTT	CAAAGAGTGT	GTGTGCTTCC	ATTTGAATGT	CTGCAGAGGA	GCATAAGAAA	AATGAGGCAG
downstream-mouse	..-....C..	.T.AG.....	AT	GTAAT...C.	..G.....		A	T.....	A.....	
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	AGAGGCAGAC	AGACTCAATT	TTATCAGGCT	GACAATGGGT	TATTGGAACA	TCAGTGAGAG	GCCTCAATAT	TTTGG---A	CAAAGGTTAA	GTGCACTTAA
downstream-mouse		A.	.C.....	...G...T.		G.....	.TG.....	CAGC.	TG...C....	...T.....
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	ATGTCAAGAG	ACAGGCTGAC	TCCATGACAA	GCTTCAAATA	GGCAGTCCAG	GGGATTATGC	CTAAAAACCC	AGGGCTAAAG	CACCTCTTCA	GAAACTGCCC
downstream-mouse	G.T...		G		T...			...C..T...	..A..AA...	C...A..TT.
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	CAACACCTGG	CCTGAAGTAT	GGACAATCGA	TCAGCAACAG	TTTCCAGCCC	TCATACCCAG	GTAAACAGTTC	AGAATGTCCC	TAAGAGAGAT	AGCCACCTAC
downstream-mouse	.TC...T...	...A.....A	G.G.G			...C.....		-----	-----	-----
downstream-rat	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
downstream-TMU	CCCTGAAGTA	TCCAACAACA								

Table. S1 Accession number list of the nucleotide sequences used in this study.

Genes	Accession No.				
	<i>T. muenninki</i>	<i>M. musculus</i>	<i>R. norvegicus</i>	<i>T. osimensis</i>	<i>A. sylvaticus</i>
<i>ZFY</i>	LC066211	NM_009570	-	E14275	-
<i>UBA1Y</i>	LC066205	NM_011667	XM_008773685	-	-
<i>EIF2S3Y</i>	LC066199	NM_012011	NM_001167666	AB441819	-
<i>TSPY</i>	LC066203	NR_027507	NM_022923	-	AF205112
<i>UTY</i>	LC066209	NM_009484	XM_008773696	-	-
<i>DDX3Y</i>	LC066197	NM_012008	NM_001167665	-	-
<i>USP9Y</i>	LC066207	NM_148943	XM_008773701	-	-
<i>RBMV</i>	LC066201	NM_001166384	GATN01000006	-	-
<i>ZFX</i>	LC066212	NM_001044386	-	-	-
<i>UBA1</i>	LC066206	NM_001136085	NM_001014080	-	-
<i>EIF2S3X</i>	LC066200	NM_012010	NM_001100542	-	-
<i>TSPYL2</i>	LC066204	NM_029836	NM_001191618	-	-
<i>KDM6A/UTX</i>	LC066210	NM_009483	XM_008773047	-	-
<i>DDX3X</i>	LC066198	NM_010028	NM_001108246	-	-
<i>USP9X</i>	LC066208	NM_009481	NM_001135893	-	-
<i>RBMX</i>	LC066202	NM_011252	NM_001025663	-	-

Table. S2 Primer pairs for RT-PCR and qPCR.

Genes	No.	Primers for RT-PCR	Primers for qPCR
<i>ZFY</i>	* ** ,	1 GTCAGGAGCTGACTTAGTAAGAA 2 GCTTCAAAGTCAACATCCCTA	AAGTCATCACTGATCCTCTGACA CCAGAGGCATCCCCTAGAA
<i>UBA1Y</i>	* ** ,	1 GGAGGTTGGAGAAATCACAG 2 AGGTCTTAGGAACAAGGAGGT	TGGAGAGGACACTGCAGCTA ACAGTCGGCCCAAGTCTGT
<i>EIF2S3Y</i>	* ** ,	1 CTTTCGTGAGCGAGATGGC 2 CCTGGTTCCAGCCAAATGT	TCTGAACACCTAGCAGCCATT TGTTCACTGTTCTTTAGCCTG/
<i>TSPY</i>	* ** ,	1 GGAAAAGGACCCAGAGCAG 2 GCTCTTTATCCTGGCAAGGTC	GGAGGAACTGCAGAAAGCAG GGGTTTTAATCCGTGCAAGA
<i>UTY</i>	*	1 ACAGCTGCTGCCGTAAAGTT 2 CAGAACTTCTGTTGAACTGGAG 3 TACCAGTCCTCACAGTGGAT 4 CAGGCTGAACAAGCAGGAAT	ATGTGGACCATGACACTGGA AACTGGCTGAAGAGGGTGAA — —
<i>DDX3Y</i>	*	1 CAAAGCTGTGTCCGTAAGCA 2 AACCAAGCAGGAAGGGAGTT	GTCGTGACCGAACTGGTTTT CTAAGCGTTCACTTGGTGGA
<i>USP9Y</i>		1 GGTGGTCACTTTGTGGTGTG 2 AGCATAAGTGAGAAGACGCCT 3 CCAGATGTTGCATGGGACA 4 TTCTTCACTGCTTTCCTGAGC	AGAGAGGTCGCTGGTTCAAA ACTCTCCACCGAAGCATTGA — —
<i>RBMY</i>	* ** ,	1 GACAAGAAAAATGGCAGAAAC 2 CCTTTAGTACCTGCCTCCAC	ATGGAGGCAGACAAGGAAGA GGATCCCCCAAATCTTGAAC
<i>SRY</i>		— —	AACCCACATGCCATCACATA TGTCACAGGCTGCAAATAA
<i>ZFX</i>	**	1 GCGTATAACTGTGCCCAGC 2 GAGTACCATCAAGGTAAGCAAGG	— —
<i>UBA1X</i>	**	1 GAGGACAAAGAGGCTCTGACA 2 ACTCCTTCCTTCCCTGGAGA	— —
<i>ACTB</i>	*	1 GGCATCCTGACCCTGAAGTA 2 TCTCAGCTGTGGTGGTGAAG	— —

*The primer pairs used for RT-PCR using multiple tissues to reveal the expression profiles.

**The primer pairs used for FISH analysis.

Table. S3 Sequence homology of *T. muenninki* X-linked genes with the orthologues of mouse

Gene name	Homology of nucleotide with mouse (%)	Homology of amino acid with mouse (%)
<i>ZFX</i>	94.6	95.4
<i>UBA1</i>	96.2	97.6
<i>EIF2S3X</i>	96.2	99.4
<i>TSPLY2</i>	94.9	93.4
<i>KDM6A/UTX</i>	97.6	96.2
<i>DDX3X</i>	97.2	99.4
<i>USP9X</i>	97.6	99.4
<i>RBMX</i>	91.8	96.9