Available online at www.sciencedirect.com

SCIENCE DIRECT

Gene 345 (2005) 101–111

GENE
SECTION
EVOLUTIONARY GENOMICS

www.elsevier.com/locate/gene

Transposable elements as a source of genetic innovation: expression and evolution of a family of retrotransposon-derived neogenes in mammals

Jürgen Brandt, Sabrina Schrauth, Anne-Marie Veith, Alexander Froschauer, Torsten Haneke, Christina Schultheis, Manfred Gessler, Cornelia Leimeister, Jean-Nicolas Volff*

Biofuture Research Group, Physiologische Chemie I, Biozentrum, University of Würzburg, am Hubland, D-97074 Würzburg, Germany

Received 4 October 2004; received in revised form 27 October 2004; accepted 9 November 2004
Available online 25 December 2004
Received by M. Porto

Abstract

A family of functional neogenes called *Mart*, related to the *gag* gene of Sushi-like long terminal repeat retrotransposons from fish and amphibians, is present in the genome of human (11 genes) and other primates, as well as in mouse (11 genes), rat, dog (12 genes), cat, and cow. *Mart* genes have lost their capacity of retrotransposition through non-functionalizing rearrangements having principally affected long terminal repeats and *pol* open reading frame. Most *Mart* genes are located on the X chromosome in different mammals. Sequence database analysis suggested that *Mart* genes are present in opossum (marsupial), but absent from the genome of chicken. Hence, the *Mart* gene family might have been formed from Sushi-like retrotransposon(s) after the split of birds and mammals (310 myr ago), but before the divergence between placental mammals and marsupials (170 myr ago). RT-PCR analysis showed that at least six *Mart* genes are expressed during mouse embryonic development, with in situ hybridization analysis revealing rather ubiquitous expression patterns. *Mart* expression was also detected in adult mice, with some genes being expressed in all tissues tested, while others showed a much more restricted expression pattern. Although additional analysis will be required to establish the function of the retrotransposon-derived *Mart* neogenes, these observations support the evolutionary importance of retrotransposable elements as a source of genetic novelty.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Ty3/gypsy; LTR; Sushi; Gag; Molecular domestication

1. Introduction

Transposable elements can be generally considered as DNA parasites capable of hitchhiking genomes over very long periods of evolution (Burke et al., 1998; Curcio and Derbyshire, 2003). Within genomes, they can disrupt resident genomic sequences or induce other kinds of genomic rearrangements, either through ectopic recombination between non-allelic copies or by transposition itself (Deininger et al., 2003; Kazazian, 2004). Hence, transposable elements can occasionally have deleterious effects on host genes and organisms. On the other hand, as a source of genomic variation, they certainly play a major role in the evolution of genomes. Segments derived from transposable elements are frequently found in regulatory and coding regions of host genes. These sequences are in some cases required for proper expression and functionality of the proteins (Nekrutenko and Li, 2001; Ganko et al., 2003; Jordan et al., 2003; van de Lagemaat et al., 2003). Functional retrotransposable elements can directly play a beneficial role for the host organism, as observed for the telomeric retrotransposons of *Drosophila* (Pardue and DeBaryshe, 2003).

Several cases of neofunctionalization have been described, in which transposable element genes have been recruited as neogenes with new cellular functions and have concomitantly lost their ability to transpose (Miller et al., 1999a,b). Such a phenomenon of "molecular domestication", "co-option", or "exaptation" has been observed for both major classes of transposable elements, the

* Corresponding author. Tel.: +49 0 931 888 4167; fax:+49 0 931 888 4150.
E-mail address: volff@biozentrum.uni-wuerzburg.de (J.-N. Volff).

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.11.022

retrotransposable elements (transposing through a mechanism involving mRNA reverse transcription; Curcio and Derbyshire, 2003) and the DNA transposable elements (transposing without reverse transcription). Neogenes of unknown functions related to DNA transposons have been found in various organisms (Pinsker et al., 2001; Sarkar et al., 2003). There is also functional evidence that the RAG proteins, which catalyze the V(D)J recombination necessary for the assembly of immunoglobulin and T-cell-receptor genes in developing lymphocytes, might have evolved from a transposase (Agrawal et al., 1998; Hiom et al., 1998). The telomerase, a reverse transcriptase involved in the replication of the ends of eukaryotic linear chromosomes (telomeres), might be derived from the reverse transcriptase of a retrotransposable element (Eickbush, 1997; Nakamura and Cech, 1998). Domesticated endogenous retrovirus envelope genes are involved in placentation in primates (Mi et al., 2000; Blaise et al., 2003; Mallet et al., 2004). The *Fv1* gene, which inhibits murine leukemia virus infection in mice, is derived from the *gag* region of an endogenous retrovirus (Best et al., 1996). An intron-containing neogene of unknown function, which probably evolved from the integrase gene of an LTR retrotransposon, is present in mammalian genomes (Llorens and Marin, 2001).

Several genes related to Sushi, a family of Ty3/gypsy long terminal repeat (LTR) retrotransposons first described in teleost fish (Poulter and Butler, 1998), have been identified in human and mouse genomes (Steplewski et al., 1998; Nagasaki et al., 1999; Butler et al., 2001; Ono et al., 2001; Volff et al., 2001; Shigemoto et al., 2001; Lynch and Tristem, 2003; Gorinsek et al., 2004; Brandt et al., 2004). Functional Sushi retrotransposons themselves are apparently absent from mammalian genomes (Volff et al., 2003), and Sushi-related neogenes have lost their ability to retrotranspose through non-functionalizing mutations before the divergence between human and mouse (Lynch and Tristem, 2003; Brandt et al., 2004). However, intact open reading frames have been conserved over more than 75 myr of mammalian evolution, and these genes evolved under purifying (negative) selection (Lynch and Tristem, 2003; Brandt et al., 2004). Taken together, these observations strongly suggest that mammalian Sushi-related sequences correspond to domesticated retrotransposons that have adopted new cellular functions. Here we analyze available genomic sequence databases for the presence of Sushi-related neogenes, and compare their expression patterns during mouse embryogenesis and in adult tissues by RT-PCR and in situ hybridization.

2. Materials and methods

2.1. DNA/RNA isolation and analysis

Genomic DNA extraction and Southern blot analysis were performed according to standard protocols (Volff et al., 1999 and references therein). Southern blots were hybridized with cloned PCR products as probes at 42 °C in 45% formamide, 0.1% sodium pyrophosphate, 50 mM Tris–HCl pH 7.5, 5×SSC, 1% SDS, 5×Denhardt's reagent, 100 mg/l calf thymus DNA, and were washed in 1×SSC, 1% SDS at 55–60 °C. Primers used in (RT-) PCR reactions (35 cycles) are available on request. RNA was isolated using the TRIZOL reagent (Gibco BRL). Reverse transcription was primed using random hexanucleotides. Absence of genomic DNA in RNA preparations was demonstrated by the absence of PCR products without reverse transcriptase (not shown). In situ hybridization of CD1 outbred mice (Charles River laboratories) embryo sections using digoxigenin-labeled antisense and sense riboprobes was performed as described (Leimeister et al., 1998). E12.5 and E14.5 correspond to embryonic days; the morning of the vaginal plug was defined as embryonic day 0.5 (E0.5).

2.2. Sequence analysis

Sequence analysis was performed using the GCG Wisconsin package (Version 10.3, Accelrys, San Diego, CA). Multiple sequence alignments were generated using PileUp from the GCG Wisconsin package and ClustalX (Thompson et al., 1997). Phylogenetic analysis was performed on a 75 amino-acid alignment of Gag-like sequences using the neighbour-joining method (Saitou and Nei, 1987; 1000 pseudosamples; the distance parameters used were mean character difference and among-site rate variation) as implemented in PAUP* (Swofford D.L., Sinauer Associates, Sunderland, Massachusetts). Iterative BLAST analysis was essentially performed using sequence databases accessible from the NCBI server (http://www.ncbi.nlm.nih.gov/BLAST/).

3. Results and discussion

3.1. Multiple gag-related Mart genes in placental mammals

Using sequences from the Sushi retrotransposon of the pufferfish *Takifugu rubripes* (Poulter and Butler, 1998) as queries in a BLAST analysis of available public sequence databases, nine different genes or groups of genes were detected in human, chimpanzee, mouse (Fig. 1), rat, and dog (Table 1; Figs. 2 and 3). These genes, called *Mar* in a previous publication (Brandt et al., 2004), have now been renamed *Mart* (*Ma*mmalian *r*etro*t*ransposon-derived), to avoid any confusion with mammalian Mariner DNA transposons (Robertson and Zumpano, 1997). Several *Mart* genes were also detected in other primates, as well as in pig, cat, sheep, and cow (Table 1 and Fig. 3). *Mart8* was found to be amplified at least in human (3 copies), mouse (3 copies), and dog (4 copies).

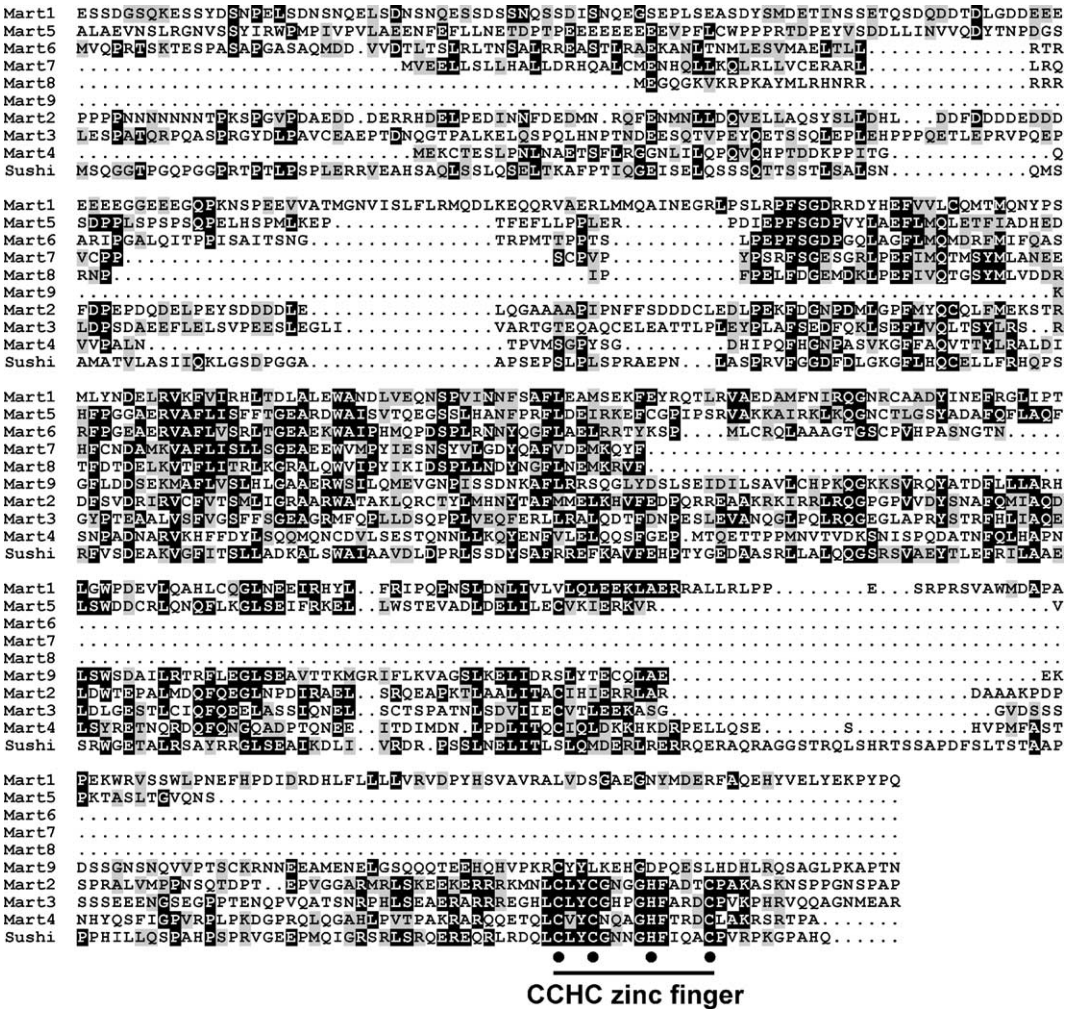

Fig. 1. Gag-like conceptual translation products of the mouse *Mart* genes. Conserved cysteine and histidine residues in the $CX_2CX_4HX_4C$ zinc finger motif are indicated. Only one Mart8 sequence is shown. For more clarity, low-complexity N- and/or C-terminal regions of some proteins have been removed. Identical residues are in black, conservative substitutions in gray. Accession numbers: Gag protein of Sushi: AAC33525; other accession numbers are given in Table 1.

All *Mart* genes encode putative proteins with similarity to the Gag protein of Sushi-like retrotransposons (Figs. 1 and 2). However, the putative nucleic acid-binding $CX_2CX_4HX_4C$ zinc finger is present only in Mart2, Mart3, and Mart4. In some *Mart* genes, the *gag*-like region within the open reading frame is associated with low complexity regions generally consisting of short tandem repeats. There is no evidence that these additional regions might have been acquired from other parts of the genome, or that they contain any known functional protein domain. Therefore they might have evolved from the original retrotransposon sequence, but we can not exclude that they might originate from other unrelated genomic sequences.

None of the mammalian *Mart* sequences apparently corresponds to a functional autonomous retrotransposon (Fig. 2). All sequences lack long terminal repeats and present various types of deletions having principally affected the reverse transcriptase/integrase region (see Lynch and Tristem, 2003; Brandt et al., 2004). Orthologous sequences in mouse and human display similar non-functionalizing rearrangements (Fig. 2), indicating that the capacity of retrotransposition has been lost before the divergence between both organisms, at least 75 myr ago. Accordingly, most *Mart* genes are single-copy. This was confirmed by Southern blot hybridization in the mouse (Fig. 4). In addition, introns are present in at least four Mart genes at identical orthologous positions in human and mouse (not shown, see Brandt et al., 2004). The duplicated copies of *Mart8* in human, mouse, and dog are located in a same genomic region and are flanked by large stretches of similar sequences (not shown). This indicated that *Mart8* duplicates have been probably formed by local events of segmental duplications rather than by retrotransposition. In the human genome, both *MART4* on Xq23 and the *MART8* gene cluster on Xq26 are close (61 kb and 232 kb, respectively) to a (pseudo)gene related to the *HMG4* gene encoding the high-mobility group protein 4 (LOC203510 on Xq23 and LOC392545 on Xq26; data not shown). This suggests that some *Mart* and *HMG4*-like genes have been co-duplicated on the X chromosome through segmental duplication during evolution.

Table 1
Retrotransposon-derived *gag*-related *Mart* genes

Gene	Organism	Aka	Chromosome	Accession number
Mart1	*Homo sapiens* (human)	*HUR1*	14q32	XM_370776
	Pan troglodytes (chimpanzee)			AACZ01151180
	Otolemur garnettii (galago)			AC148728
	Mus musculus (mouse)	*Mor1, Rtl1*	12F1	BK001261
	Rattus norvegicus (rat)	*Rar1*	6q32	XM_243381
	Bos taurus (cow)			AC148746
	Ovis aries (sheep)	*Peg11*	18q	AF354168
	Canis familiaris (dog)		8	AAEX01040213
Mart2	*Homo sapiens* (human)	*PEG10*	7q21	BC050659
	Pan troglodytes (chimpanzee)			AACZ01060433
	Papio anubis (baboon)			AC092529
	Callithrix jacchus (marmoset)			AC148261
	Mus musculus (mouse)	*Myef-3, Edr*	6A1	AB091827
	Rattus sp. (rat)			AI599367
	Sus scrofa (pig)			CF788043
	Felis catus (cat)			AC108191
	Canis familiaris (dog)		14	AAEX01048894
	Rhinolophus ferrumequinum (bat)			AC149024
	Sorex araneus (shrew)			AC148354
	Monodelphis domestica (opossum)			448448810 (WGS)
Mart3	*Homo sapiens* (human)	*ZCCHC5*	Xq13	NM_152694
	Pan troglodytes (chimpanzee)			AADA01082408
	Mus musculus (mouse)		XD	NM_199468
	Rattus norvegicus (rat)		Xq31	XM_228502
	Canis familiaris (dog)		X	AAEX01050855
Mart4	*Homo sapiens* (human)	*FLJ46608*	Xq23	AK128465
	Pan troglodytes (chimpanzee)			AACZ01078749
	Mus musculus (mouse)		XF2	AL807791
	Rattus norvegicus (rat)		Xq14	AC094904
	Canis familiaris (dog)		X	AAEX01051357
Mart5	*Homo sapiens* (human)	*KIAA2001*	Xq13	XM_291322
	Pan troglodytes (chimpanzee)			AADA01037057
	Pongo pygmaeus (orangutan)			CR542615
	Macaca fascicularis (macaque)			AB060816
	Mus musculus (mouse)		XC3	NM_183318
	Rattus norvegicus (rat)		Xq31	XM_228548
	Bos taurus (cow)			AW464288
	Canis familiaris (dog)		X	AACN010604025
Mart6	*Homo sapiens* (human)	*LDOC1L*	22q13	NM_032287
	Pan troglodytes (chimpanzee)			AACZ01405866
	Pongo pygmaeus (orangutan)			CR557003
	Mus musculus (mouse)		15E2	AL611986
	Rattus norvegicus (rat)		7q34	XM_235542
	Canis familiaris (dog)		10	AAEX01056350
Mart7	*Homo sapiens* (human)	*LDOC1*	Xq27	AB019527
	Pan paniscus (pygmy chimpanzee)		X	AY459388
	Gorilla gorilla (gorilla)		X	AY459387
	Mus musculus (mouse)		XA5	BC052689
	Rattus norvegicus (rat)		Xq36	XM_346368
	Bos taurus (cow)			CO880785
	Canis familiaris (dog)		X	AAEX01051796
Mart8	*Homo sapiens* (human; 3 genes)	*CXX1*	Xq26	AL136169
	Pan troglodytes (chimpanzee)			AADA01342505
	Macaca mulatta (macaque)			CO579834
	Mus musculus (mouse; 3 genes)		XA5	AL807398
	Rattus norvegicus (rat)			CF114446
	Bos taurus (cow)			CK948047
	Sus scrofa (pig)			BG609236
	Canis familiaris (dog; 4 genes)		X	AAEX01059888
Mart9	*Homo sapiens* (human)	*KIAA1318, RGAG1*	Xq23	AB037739
	Pan troglodytes (chimpanzee)			AACZ01265624
	Mus musculus (mouse)		XF2	XM_142197
	Rattus norvegicus (rat)		X	XM_235717
	Canis familiaris (dog)		X	AAEX01056279

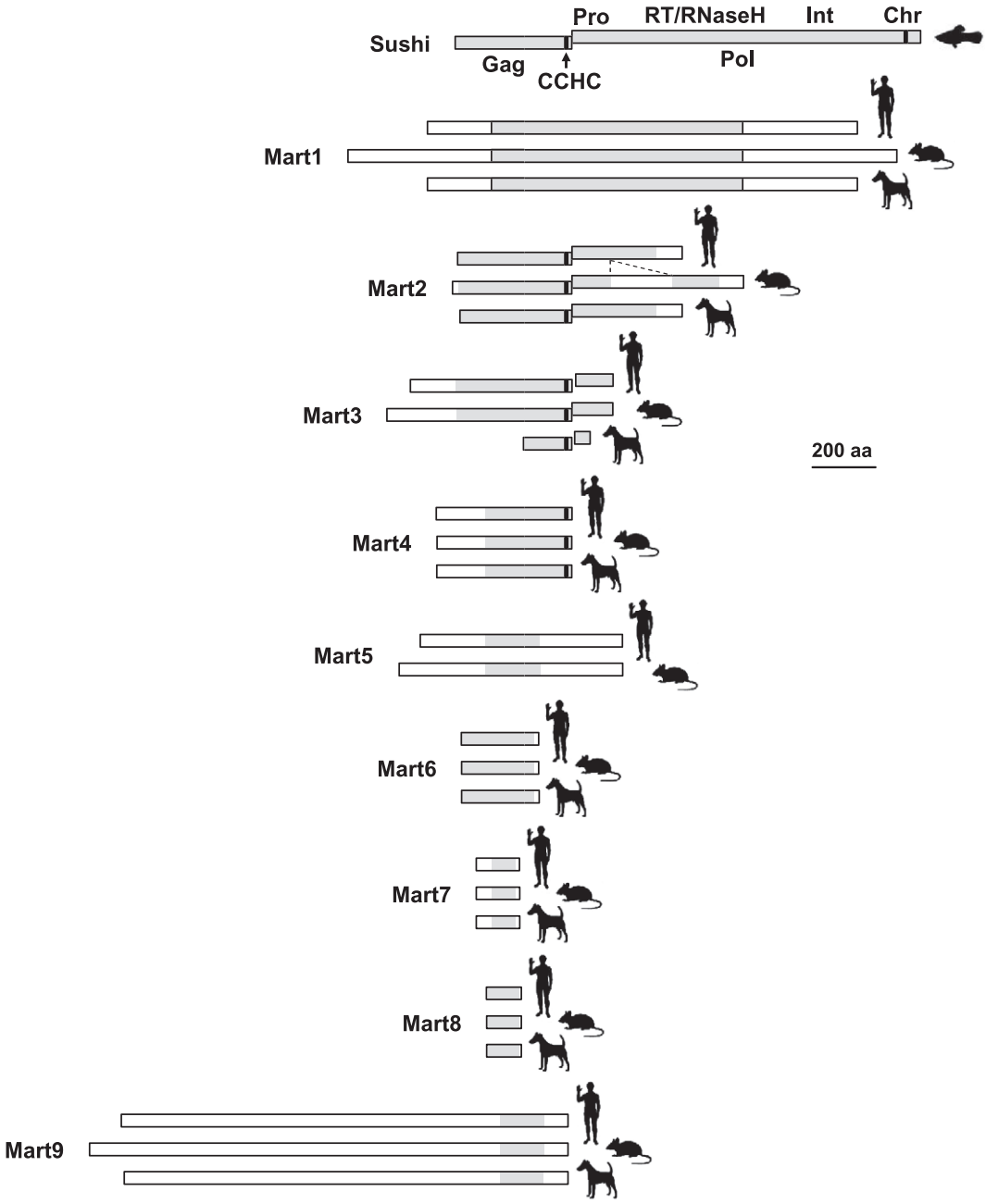

Fig. 2. Structure of Mart conceptual proteins in human, mouse, and dog. Gray boxes, region of similarity with Sushi proteins; black boxes, CCHC zinc finger; white boxes, regions of low complexity with no significant similarity with any known protein. The *gag/pol* frameshift is shown for Sushi, Mart2, and Mart3. Chr, chromodomain; Int, integrase; Pro, protease; RT, reverse transcriptase.

Strikingly, most *Mart* genes are located on the X, but not on the Y chromosome in human, mouse, rat, and dog. Exceptions are *Mart1*, *Mart2*, and *Mart6*, which reside on different autosomes in placental mammals (Table 1).

3.2. Are Mart genes mammal-specific?

The presence of *Mart* genes outside of the placental mammalian lineage was further investigated by sequence database analysis. Interestingly, *Mart2* could be identified in whole genome shotgun (WGS) trace sequences from a marsupial, the gray short-tailed opossum *Monodelphis domestica* (Table 1). Consistent with the phylogenetic position of marsupials in the mammalian lineage, the translation product from opossum was basal to the group of Mart2 sequences from placental mammals in the molecular phylogeny (Fig. 3).

The first assembly of the genome of the chicken *Gallus gallus* and WGS sequences traces was searched for the presence of *Mart* genes. Neither *Mart* sequences nor classical Sushi-like retrotransposons were identified. In amphibians, including the frogs *Xenopus laevis* and

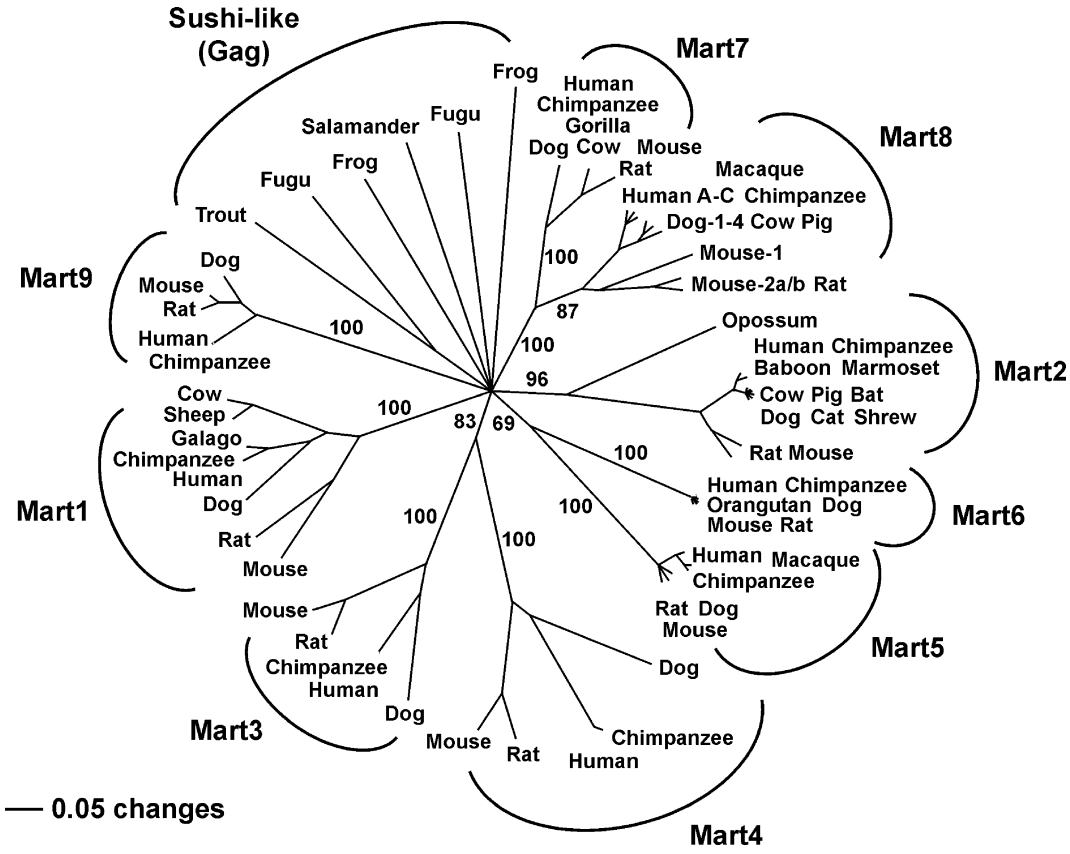

Fig. 3. Phylogenetic analysis of mammalian Mart proteins. The tree (neighbour-joining, 1000 pseudosamples) is unrooted. The bootstrap values supporting the different clades are given as percentages. Accession numbers: AAC33525 and AAG60684 (fugu *T. rubripes*), CA371598 (trout *Oncorhynchus mykiss*), CK797365 and AC149870 (frog *X. laevis*), CN055626 (salamander *A. tigrinum*). Other accession numbers are given in Table 1. Only examples of Sushi-like elements, which are multicopy in fish and amphibians, have been included in the analysis.

Xenopus tropicalis (WGS traces) and the salamander *Ambystoma tigrinum*, classical Sushi-like LTR retrotransposons were found but no *Mart* gene was detected. No *Mart* genes were identified in non-mammalian species. Taken together, these results suggest that *Mart* genes are mammal-specific neogenes derived from retrotransposons, which have been formed before the split between marsupials and placental mammals at least 170 myr ago (Kumar and Hedges, 1998). Classical Sushi-like LTR retrotransposons have been apparently eliminated from both the mammalian and chicken (bird) lineages.

3.3. Expression of Mart genes in mouse embryos and adult tissues

Expression of *Mart* genes was analyzed in tissues from E14.5-old mouse embryos (Fig. 5). *Mart1-6* were found to be expressed in different embryonic organs including brain,

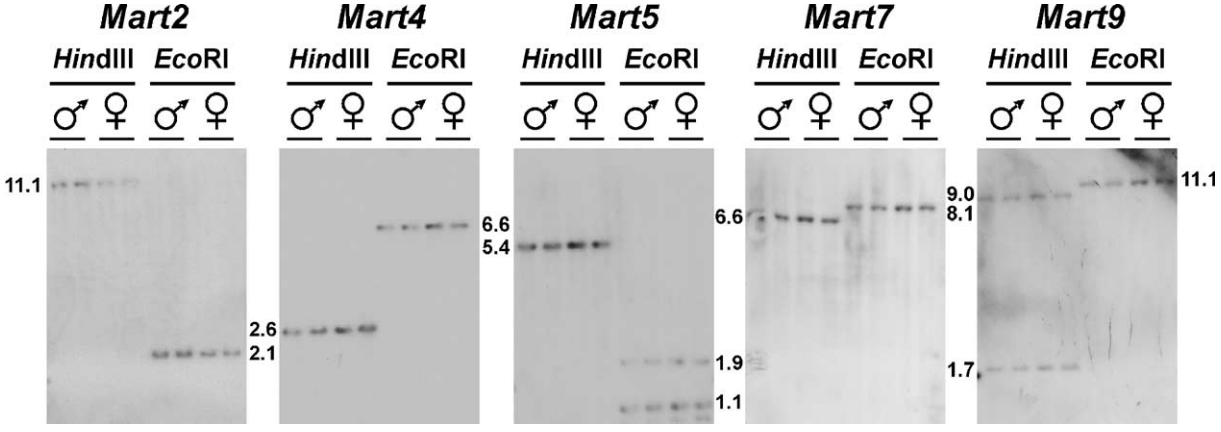

Fig. 4. Southern blot analysis of mouse *Mart* genes. The size of restriction fragments is given in kb.

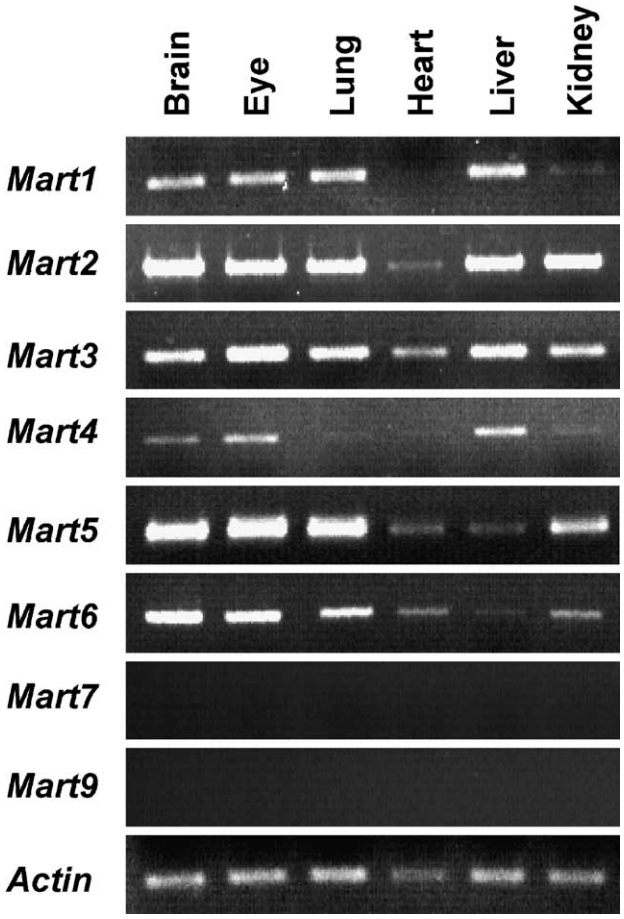

Fig. 5. RT-PCR analysis of the expression of *Mart* genes in mouse embryonic tissues at E14.5. *Mart8* has not been analyzed. Actin was used as a control.

eye, lung, liver, and kidney. Generally, the lowest level of expression was found in the heart. No expression of *Mart7* and *Mart9* was observed in the tissues tested (*Mart8* has not been tested).

In order to analyze the spatial distribution of *Mart* transcripts, in situ hybridizations were performed on sagittal sections of E12.5- and E14.5-old mouse embryos. Consistent with the RT-PCR analysis, no strict specificity of expression was observed for *Mart1* and *Mart2* (Fig. 6) as well as for *Mart3*, *Mart4*, *Mart5*, and *Mart6* (not shown). *Mart1* transcripts were detected in almost all tissues of E12.5 mouse embryos (Fig. 6A). A much weaker staining of E12.5 sections is observed with the sense probe, with some staining of the brain cortex in most sections (Fig. 6B). At E14.5, *Mart1* is again expressed almost ubiquitously (Fig. 6C–D). A stronger *Mart1* signal was observed in nervous and muscle tissues. Expression was detected in the entire brain, the spinal cord, and the dorsal root ganglia of the peripheral nervous system. Examples of muscle expression of *Mart1* are the tongue, the intercostal muscles, and the fibers of the crus of the diaphragm, while the heart is stained much more faintly. *Mart2* is likewise expressed ubiquitously in E12.5 and E14.5 embryo sections, with the same preferential staining of nervous and muscle tissues but a higher expression than *Mart1* in cartilage (Fig. 6E–J).

Expression of *Mart* genes was also detected in mouse adult tissues by RT-PCR (Fig. 7). Some *Mart* genes, including *Mart2*, *Mart5*, and *Mart6,* were found to be expressed in all tissues tested. Other *Mart* genes showed a more restricted expression pattern, like *Mart7*, which is only expressed in brain and ovary. A very faint expression of *Mart9* was detected in brain (not visible on Fig. 7). Two genes, *Mart3* and *Mart7*, are apparently expressed in ovary but not in testis.

3.4. Conclusions

A family of neogenes derived from *Sushi*-like retrotransposons is present in the genome of primates and other mammals. These genes have lost their ability to retrotranspose during evolution through non-functionalizing mutations. Subsequently, their *gag*-related open reading frame evolved under purifying (negative) selection (Lynch and Tristem, 2003; Brandt et al., 2004). This strongly supports the hypothesis that *Mart* genes have been "domesticated" during evolution and fulfil now some retrotransposition-independent functions useful for the host.

Identification of *Mart2* in a marsupial species suggests that the domestication event(s) at the origin of the *Mart* gene family arose at least 170 myr ago (Fig. 8). This event might have taken place on the X chromosome, which has generated and recruited a disproportionately high number of functional retroposed gene during evolution (Emerson et al., 2004). Subsequently, the original *Mart* gene has probably been duplicated several times. The location of most *Mart* genes on the X chromosome in different mammals points to the involvement of segmental intrachromosomal duplications. The fact that all *Mart8* duplicates are located within a same restricted region of the X chromosome in human, mouse, and dog and are flanked by large stretches of similar sequences supports this hypothesis. Some *Mart* genes might have been co-duplicated together with neighbouring genes. Alternatively, multiple events of retrotransposon domestication might have occurred, since our phylogenetic analysis failed to demonstrate the monophyly of the Mart family.

The apparent absence of *Mart* genes from chicken, amphibians, and fish might indicate that these genes are specific of mammals. The related Sushi-like retrotransposons were detected in fish and amphibians, but not in chicken. Hence, Sushi-like retrotransposons might have been inactivated/domesticated after the divergence between reptiles/birds and mammals, but before the split between marsupials and placental mammals, with a second retrotransposon loss in the chicken lineage (Fig. 7). Alternatively, domestication and retrotransposon inactivation arose

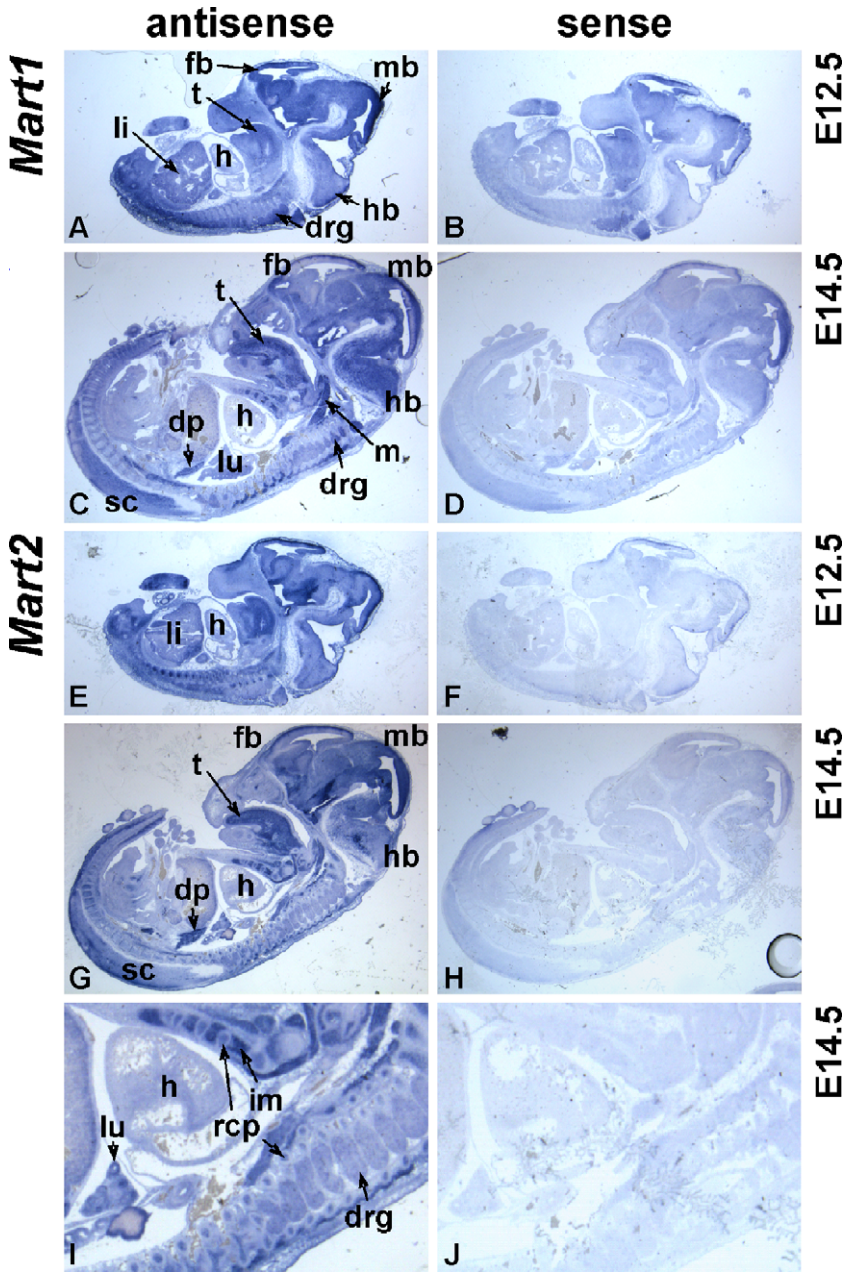

Fig. 6. Expression of *Mart1* and *Mart2* in sagittal sections of E12.5- and E14.5-old mouse embryos. (A–D) Expression of *Mart1* at E12.5 (A) and E14.5 (C); (E–J) expression of *Mart2* at E12.5 (E) and E14.5 (G). (I)–(J) are magnifications of (G)–(H) showing the expression of *Mart2* in rib cartilage primordia and intercostal muscles. (B), (D), (F), (H), (J) are hybridization controls with sense probes. Dp, diaphragm; drg, dorsal root ganglia; fb, forebrain; h, heart; hb, hindbrain; im, intercostal muscles; li, liver; lu, lung; mb, midbrain; m, muscle; rcp, rib cartilage primordium; sc, spinal cord; t, tongue.

before the split between mammals and birds, with subsequent loss of the *Mart* gene(s) in the chicken lineage. Clearly, additional data, particularly from birds and reptiles, will be necessary to retrace the evolutionary history of *Mart* genes and Sushi-like retrotransposons in vertebrates. Related retrotransposons have been identified in reptiles, but they are apparently paralogous to the Sushi/*Mart* lineage (Miller et al., 1999a,b; Gorinsek et al., 2004).

At least six *Mart* genes are expressed during mouse development. *Mart* expression is not restricted to particular regions of the embryo but appears rather ubiquitous. *Mart* genes present more or less specific expression patterns in adult tissues, with *Mart3* and *Mart7* apparently expressed in female but not in male gonads. Interestingly, some *Mart* genes are differentially expressed in some tumours and cancer cell lines (Okabe et al., 2003; Nagasaki et al., 2003).

Interestingly, two autosomal *Mart* genes are subject to imprinting. *Mart1* is paternally expressed and located within imprinted orthologous loci in human, mouse, and sheep (Charlier et al., 2001; Seitz et al., 2003). A non-coding antisense transcript expressed from the maternal chromo-

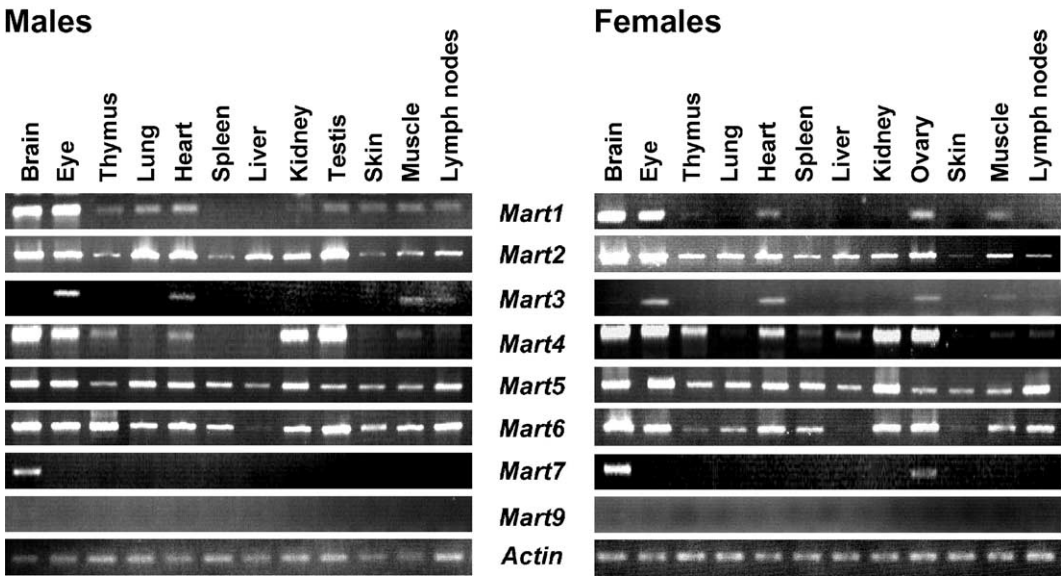

Fig. 7. RT-PCR analysis of the expression of *Mart* genes in mouse adult tissues. *Mart8* has not been analyzed. Actin was used as a control.

some contains two microRNAs with full complementarity to *Mart1* (Charlier et al., 2001; Seitz et al., 2003). This situation is reminiscent of the RNA interference-mediated silencing of transposable elements (Plasterk, 2002), and the microRNAs might function as small interfering RNAs silencing *Mart1* (Seitz et al., 2003). *Mart2* is also imprinted and is paternally expressed in human and mouse (Ono et al., 2001, 2003; Okita et al., 2003).

Additional functional data will be required to determine the function of *Mart* genes in mammals. It has been speculated that *Mart1* can restrict colonization of mammalian genomes by Ty3/gypsy retrotransposons (Lynch and Tristem, 2003). The presence of a zinc finger domain in Mart2, Mart3, and Mart4 suggests a function involving nucleic acid binding. Accordingly, Mart2 has been identified as a putative transcription factor able to

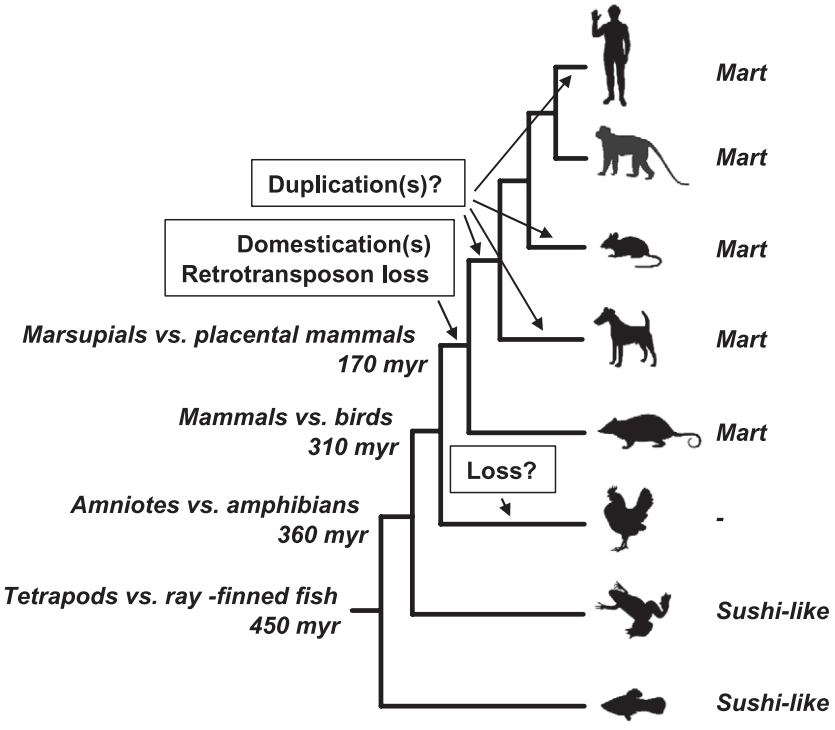

Fig. 8. Distribution and hypothetic evolution of Sushi-like retrotransposons and *Mart* neogenes in the vertebrate lineage.

specifically interact with regulatory sequences of the gene encoding the myelin basic protein (MBP), the major component of the myelin sheath of the central nervous system (Steplewski et al., 1998). In addition, MART2 exogenous expression is able to promote growth of human hepatoma cells. Human MART2 can interact with SIAH1, a mediator of apoptosis, and overexpression of MART2 decreases the cell death mediated by SIAH1 (Okabe et al., 2003). Finally, MART7 has been proposed to correspond to a regulator of the transcriptional response mediated by the nuclear factor kappa B (Nagasaki et al., 2003). Hence, some Mart genes might be involved in the control of cell proliferation and apoptosis, and serve as new targets for the treatment of cancer (Okabe et al., 2003; Nagasaki et al., 2003).

To conclude, the mammalian *Mart* gene family provides further evidence that transposable elements can not only contribute regulatory sequences and novel exons to resident genes, but also can evolve as neogenes with cellular functions. Maybe these functions were already performed by the ancestral Sushi-like retrotransposons and were uncovered in the *Mart* genes by the loss of the ability of retrotransposition. This might explain why some functional retrotransposable elements have been maintained in their hosts over very long periods of evolution.

Acknowledgements

This work was supported by the BioFuture program of the German Bundesministerium für Bildung und Forschung (to J.N.V). We are grateful to Manfred Schartl and to the members of the laboratory for encouragements and helpful discussions.

References

Agrawal, A., Eastman, Q.M., Schatz, D.G., 1998. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 394, 744–751.

Best, S., Le Tissier, P., Towers, G., Stoye, J.P., 1996. Positional cloning of the mouse retrovirus restriction gene *Fv1*. Nature 382, 826–829.

Blaise, S., de Parseval, N., Benit, L., Heidmann, T., 2003. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies *syncytin 2*, a gene conserved on primate evolution. Proc. Natl. Acad. Sci. U. S. A. 100, 13013–13018.

Brandt, J., Veith, A.-M., Volff, J.-N., 2004. A family of neofunctionalized Ty3/gypsy retrotransposon genes in mammalian genomes. Cytogenet. Genome Res. (in press).

Burke, W.D., Malik, H.S., Lathe III, W.C., Eickbush, T.H., 1998. Are retrotransposons long-term hitchhikers? Nature 392, 141–142.

Butler, M., Goodwin, T., Simpson, M., Singh, M., Poulter, R., 2001. Vertebrate LTR retrotransposons of the Tf1/sushi group. J. Mol. Evol. 52, 260–274.

Charlier, C., et al., 2001. Human-ovine comparative sequencing of a 250-kb imprinted domain encompassing the callipyge (clpg) locus and identification of six imprinted transcripts: *DLK1*, *DAT*, *GTL2*, *PEG11*, *antiPEG11*, and *MEG8*. Genome Res. 11, 850–862.

Curcio, M.J., Derbyshire, K.M., 2003. The outs and ins of transposition: from mu to kangaroo. Nat. Rev., Mol. Cell Biol. 4, 865–877.

Deininger, P.L., Moran, J.V., Batzer, M.A, Kazazian Jr., H.H., 2003. Mobile elements and mammalian genome evolution. Curr. Opin. Genet. Dev. 13, 651–658.

Eickbush, T.H., 1997. Telomerase and retrotransposons: which came first? Science 277, 911–912.

Emerson, J.J., Kaessmann, H., Betran, E., Long, M., 2004. Extensive gene traffic on the mammalian X chromosome. Science 303, 537–540.

Ganko, E.W., Bhattacharjee, V., Schliekelman, P., McDonald, J.F., 2003. Evidence for the contribution of LTR retrotransposons to *C. elegans* gene evolution. Mol. Biol. Evol. 20, 1925–1931.

Gorinsek, B., Gubensek, F., Kordis, D., 2004. Evolutionary genomics of chromoviruses in eukaryotes. Mol. Biol. Evol. 21, 781–798.

Hiom, K., Melek, M., Gellert, M., 1998. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. Cell 94, 463–470.

Jordan, I.K., Rogozin, I.B., Glazko, G.V., Koonin, E.V., 2003. Origin of a substantial fraction of human regulatory sequences from transposable elements. Trends Genet. 19, 68–72.

Kazazian Jr., H.H., 2004. Mobile elements: drivers of genome evolution. Science 303, 1626–1632.

Kumar, S., Hedges, S.B., 1998. A molecular timescale for vertebrate evolution. Nature 392, 917–920.

Leimeister, C., Bach, A., Gessler, M., 1998. Developmental expression patterns of mouse *sFRP* genes encoding members of the secreted frizzled related protein family. Mech. Dev. 75, 29–42.

Llorens, C., Marin, I., 2001. A mammalian gene evolved from the integrase domain of an LTR retrotransposon. Mol. Biol. Evol. 18, 1597–1600.

Lynch, C., Tristem, M., 2003. A co-opted gypsy-type LTR-retrotransposon is conserved in the genomes of humans, sheep, mice, and rats. Curr. Biol. 13, 1518–1523.

Mallet, F., et al., 2004. The endogenous retroviral locus *ERVWE1* is a bona fide gene involved in hominoid placental physiology. Proc. Natl. Acad. Sci. U. S. A. 101, 1731–1736.

Mi, S., et al., 2000. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403, 785–799.

Miller, K., Lynch, C., Martin, J., Herniou, E., Tristem, M., 1999a. Identification of multiple Gypsy LTR-retrotransposon lineages in vertebrate genomes. J. Mol. Evol. 49, 358–366.

Miller, W.J., McDonald, J.F., Nouaud, D., Anxolabéhère, D., 1999b. Molecular domestication—more than a sporadic episode in evolution. Genetica 107, 197–207.

Nagasaki, K., Manabe, T., Hanzawa, H., Maass, N., Tsukada, T., Yamaguchi, K., 1999. Identification of a novel gene, *LDOC1*, down-regulated in cancer cell lines. Cancer Lett. 140, 227–234.

Nagasaki, K., et al., 2003. Leucine-zipper protein, LDOC1, inhibits NF-kappaB activation and sensitizes pancreatic cancer cells to apoptosis. Int. J. Cancer 105, 454–458.

Nakamura, T.M., Cech, T.R., 1998. Reversing time: origin of telomerase. Cell 92, 587–590.

Nekrutenko, A., Li, W.H., 2001. Transposable elements are found in a large number of human protein-coding genes. Trends Genet. 17, 619–621.

Okabe, H., et al., 2003. Involvement of PEG10 in human hepatocellular carcinogenesis through interaction with SIAH1. Cancer Res. 63, 3043–3048.

Okita, C., Meguro, M., Hoshiya, H., Haruta, M., Sakamoto, Y.K., 2003. A new imprinted cluster on the human chromosome 7q21-q31, identified by human–mouse monochromosomal hybrids. Genomics 81, 556–559.

Ono, R., et al., 2001. A retrotransposon-derived gene, *PEG10*, is a novel imprinted gene located on human chromosome 7q21. Genomics 73, 232–237.

Ono, R., Shiura, H., Aburatani, H., Kohda, T., Kaneko-Ishino, T., Ishino, F., 2003. Identification of a large novel imprinted gene cluster on mouse proximal chromosome 6. Genome Res. 13, 1696–1705.

Pardue, M.L., DeBaryshe, P.G., 2003. Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. Annu. Rev. Genet. 37, 485–511.

Pinsker, W., Haring, E., Hagemann, S., Miller, W.J., 2001. The evolutionary life history of P transposons: from horizontal invaders to domesticated neogenes. Chromosoma 110, 148–158.

Plasterk, R.H., 2002. RNA silencing: the genome's immune system. Science 296, 1263–1265.

Poulter, R., Butler, M., 1998. A retrotransposon family from the pufferfish (fugu) *Fugu rubripes*. Gene 215, 241–249.

Robertson, H.M., Zumpano, K.L., 1997. Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. Gene 205, 203–217.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sarkar, A., et al., 2003. Molecular evolutionary analysis of the widespread piggyBac transposon family and related "domesticated" sequences. Mol. Genet. Genomics 270, 173–180.

Seitz, H., et al., 2003. Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. Nat. Genet. 34, 261–262.

Shigemoto, K., et al., 2001. Identification and characterisation of a developmentally regulated mammalian gene that utilises-1 programmed ribosomal frameshifting. Nucleic Acids Res. 29, 4079–4088.

Steplewski, A., Krynska, B., Tretiakova, A., Haas, S., Khalili, K., Amini, S., 1998. MyEF-3, a developmentally controlled brain-derived nuclear protein which specifically interacts with myelin basic protein proximal regulatory sequences. Biochem. Biophys. Res. Commun. 243, 295–301.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

van de Lagemaat, L.N., Landry, J.R., Mager, D.L., Medstrand, P., 2003. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. Trends Genet. 19, 530–536.

Volff, J.-N., Körting, C., Sweeney, K., Schartl, M., 1999. The non-LTR retrotransposon Rex3 from the fish Xiphophorus is widespread among teleosts. Mol. Biol. Evol. 16, 1427–1438.

Volff, J.-N., Körting, C., Schartl, M., 2001. Ty3/Gypsy retrotransposon fossils in mammalian genomes: did they evolve into new cellular functions? Mol. Biol. Evol. 18, 266–270.

Volff, J.-N., Bouneau, L., Ozouf-Costaz, C., Fischer, C., 2003. Diversity of retrotransposable elements in compact pufferfish genomes. Trends Genet. 19, 674–678.