J Mol Evol (2005) 61:481–490
DOI: 10.1007/s00239-004-0332-0

JOURNAL OF MOLECULAR EVOLUTION

A Small Family of Sushi-Class Retrotransposon-Derived Genes in Mammals and Their Relation to Genomic Imprinting

Neil A. Youngson, Sylvia Kocialkowski, Nina Peel, Anne C. Ferguson-Smith

Department of Anatomy, University of Cambridge, Downing St, Cambridge CB23DY, UK

Received: 19 November 2004 / Accepted: 29 March 2005 [*Reviewing Editor:* Dr. Juergen Brosius]

Abstract. Ty3/gypsy retrotransposons are rare in mammalian genomes despite their abundance in invertebrate and other vertebrate classes. Here we identify a family of nine conserved mammalian genes with homology to Ty3/gypsy retrotransposons but which have lost their ability to autonomously retrotranspose. Of these, five map to the X chromosome while the remaining four are autosomal. Comparative phylogenetic analyses show them to have strongest homology to the sushi-ichi element from *Fugu rubripes*. Two of the autosomal gene members, *Peg10* and *Rtl1*, are known to be imprinted, being expressed from the paternally inherited chromosome homologue. This suggests, consistent with the host-parasite response theory of the evolution of the imprinting mechanism, that parental-origin specific epigenetic control may be mediated by genomic ''parasitic'' elements such as these. Alternatively, these elements may preferentially integrate into regions that are differentially modified on the two homologous chromosomes such as imprinted domains and the X chromosome and acquire monoallelic expression. We assess the imprinting status of the remaining autosomal members of this family and show them to be biallelically expressed in embryo and placenta. Furthermore, the methylation status of *Rtl1* was assayed throughout development and was found to resemble that of actively, silenced repetitive elements rather than imprinted sequences. This indicates that the ability to undergo genomic imprinting is not an inherent property of all members of this family of retroelements. Nonetheless, the conservation but functional divergence between the different members suggests that they have undergone positive selection and acquired distinct endogenous functions within their mammalian hosts.

Key words: LTR retrotransposon — Genomic imprinting — Ty3-gypsy — Mammalian — Sushi-ichi — Bisulphite sequencing — CpG methylation

Introduction

More than 45% of the human genome is derived from repetitive parasitic elements such as transposons, retrotransposons, and retroviruses (Lander et al. 2001). These elements can be detrimental to the host and multiple strategies in different organisms have been utilized to combat this, many involving epigenetic mechanisms (Yoder et al. 1997; Birchler et al. 2000; Matzke et al. 2002). In mammalian genomes, these elements are generally silenced by mechanisms involving cytosine methylation. DNA methylation is also involved in genomic imprinting, a process that causes a subset of genes to be expressed in a parental-origin specific manner. Certain classes of a repetitive parasitic element appear to be methylated in a parental-origin specific manner in the gametes (Rubin et al. 1984; Sanford et al. 1987; Chesnokov and Schmid, 1995; Lane et al. 2003). Hence, a relationship between this and the evolution of the imprinting mechanism has been suggested (Barlow, 1993). The

The sequence data from this study have been submitted to Entrez/NCBI under accession nos. BN000778 to BN000783.

Correspondence to: A. C. Ferguson-Smith; *email:* afsmith@mole.bio.cam.ac.uk

finding of two closely related paternally expressed imprinted retrotransposon-like genes, *Rtl1* and *Peg10* (Seitz et al. 2003; Ono et al. 2003), prompted further investigation to determine whether there were other mammalian genes in this family, a detailed assessment of their phylogenetic origin, and an analysis of their imprinting status.

The mammalian genome is composed of several different types of transposable elements. Each type can be characterised by features including mode of replication, element structure, and relatedness to other families. LTR retrotransposons replicate through the reintegration of a reverse-transcribed copy of their transcripts, a mechanism they share with their closest relatives, the retroviruses. LTR retrotransposon sequences consist of *gag, pol*, and in some cases the *env* gene flanked by LTRs, and are divided into subgroups based on sequence similarity usually determined through alignment of conserved motifs within their respective *pol* genes. The Ty3/gypsy family is found in disparate eukaryotic taxa including fungi, plants, insects, tunicates, and echinoderms (Miller et al. 1999). Vertebrate copies are found in the genomes of fish, amphibia, reptiles, and mammals and are most closely related to elements from fungi suggesting that the vertebrate elements originated via horizontal transfer between these two phyla (Miller et al. 1999; Butler et al. 2001; Volff et al. 2001). Considering their abundance in other taxa, it is perhaps surprising that very few have been described in mammalian species. Sequence analysis of *Rtl1* and *Peg 10* (below) indicates that they are incapable of autonomous retrotransposition. However, the relative rarity of Ty3/gypsy elements in the mammalian genome, combined with the preservation of their open reading frames in mammals and their location within unusually epigenetically regulated domains, suggests that this class of retrotransposon may have evolved an endogenous function in their mammalian hosts. Furthermore, they could also be subject to imprinted expression or might even confer imprinting on their surrounding regions. These latter possibilities were investigated, firstly, by assaying novel sushi-type genes for imprinting and, secondly, by studying the methylation status of the *Rtl1* Ty3/gypsy-homologous region throughout development. The second experiments determined whether the *Rtl1* DNA methylation dynamics were comparable to those either of imprinted genes or actively silenced repetitive elements.

Methods

Sequence Analyses

Amino-acid sequences from the mouse sushi-type ORFs as were aligned to other members of the Ty3/gypsy retrotransposon family using ClustalW WWW Service at the European Bioinformatics Institute (http://www.ebi.ac.uk/clustalw) (default settings used) and adjusted manually with GeneDoc (Nicholas et al. 1997). The mouse sushi-type ORFs were also tested for non-retrotransposon-related protein motifs by searching the ExPASy (Expert Protein Analysis System) Prosite database (http://ca.expasy.org/prosite). Conservation between human and mouse was investigated using the BLASTZ (Schwartz et al., 2003) of the human genome (hg16-July 03) Tight Subset of Alignments (http://genome.ucsc.edu/).

Strand-Specific RT-PCR

Four micrograms of total RNA was DNase treated with Promega RQ1 DNase according to the manufacturer's instructions. We reverse transcribed RNA using a strand-specific primer and Invitrogen Superscript III RNase H^- reverse transcriptase according to the manufacturer's instructions with the RT step 55°C for 1 h. PCR was done with Bioline Biotaq DNA polymerase and buffers. PCR cycles: M15f/M15r2 95°C 5 min, [95°C 30 s, 58°C 30 s, 72°C 30 s] × 35, 72°C 5 min; 14C1F2/14C1R2 94°C 2 min, 94°C 15 s, 60°C 30 s, 72°C 40 s]×35, 72°C 5 min. Primers: M15f GCTGGAGA AGCTGGGATTGC; M15r2 AGACATCCGCGTCCAATAGG; 14C1F2 CCGATGCGTTGGAGACAATC; 14C1R2 GCCTCAC CGTCTTATAGGTG.

Rtl1 *Methylation Analyses*

MatDi(12) and PatDi(12) mice and their normal littermates were collected as described by Takada et al. (2002). DNA isolation and Southern analysis is as described previously (Ferguson-Smith et al. 1993). Normal sperm was purified from a C3H/He male. Isolation of PGCs was conducted as described previously (Pesce and De Felici, 1995). The probe corresponding to the *Rtl1* internal CpG island was produced with the primers: AW060f CCGAACGATG CTCTCCAAGTG; AW060r TTGGAACCTGTGAGCGCCAG.

Oocytes and preimplantation embryos were collected from superovulated (CBA x C57BL/6)F1 females after mating with Fl males. DNA was isolated and subjected to sodium bisulphite mutagenesis treatment of genomic DNA. PCR from the mutagenised template and cloning of the amplified product was carried out with two different techniques. Some were subjected to the agarose-bead technique described by Takada et al. (2002) and others were processed using the simpler non-agarose bead treatment that is described in Arnaud et al. (2003). The results obtained for each tissue were the same whichever technique was used. Controls for PCR and cloning bias were conducted using test templates of known CpG and TpG proportions. No bias was found. The primer sequences for the nested PCR reactions were: CpG1F1:5′ TTTTGATTATTGTTTTTAGTTTGTTGAAGTTT; CpG1F2: 5′ GAAGTTTAGAGGGTTTTGATTTAGAAAGAT; CpG1R1: 5′ ATATACCTTCCACTCTCCCTACTATCTAC; CpG1R2; 5′ AT TACTTCACCCCACCTCCACCATA.

F1/R1 are the outer pair and F2/R2 are the inner pair.

Results

Sequence Identification, Classification, Conservation, and Expression

A TBLASTN search (using http://www.ensembl.org was conducted using human PEG10 (BAB43951) against the mouse genomic sequence and identified seven putative genes with high homology (Evalue

Table 1. Features of mouse sushi class retrotransposon sequences

Gene	Chromosomal location	*gag* homologous ORF length (bps)	BLASTP homology to Sushi-ichi *gag* (AAC33525)	*pol* homologous ORF length (bps)	BLASTP homology to Sushi-ichi *pol* (AAC33526)	Expression evidence	Human mouse conservation[a]
Rtl1	12F2	4077 (*gag* and *pol*)	74 E = 8e-12	*gag* and *pol* are same ORF	182 E = 5e-44	Seitz et al. (2003)	79.2% over 2526 bp (14q32)
Peg10	6A1	1128	113 E = 7e-24	1668	64 E = 6e-9 and 41 E = 0.045	Steplewski et al. (1998)	79.3% over 1255 bp (7q21)
Sushi-15E3	15E3	564	45 E = 8e-4	- - - - - - - - -	- - - - - - - - -	ESTs and an mRNA	80.3% over 748 bp (22q13)
Sushi-14C1	14C1	- - - - - - - - -	- - - - - - - - -	2904	68 E = 8e-10	ESTs	83.9% over 2821 bp (14q11)
Ldoc1	XA5	453	36 E = 0.19	- - - - - - - - -	- - - - - - - - -	BC052689	80.1% over 412 bp (Xq27)
Sushi-XC3	XC3	1596	55 E = 2e-6	- - - - - - - - -	- - - - - - - - -	AK032151	78.9% over 2660 bp (Xql3)
Sushi-XD	XD	1659	70 E = 2e-10	633	81 E = 1e-13	AK084983	80.9% over 3618 bp (Xq21)
Sushi-XF2	XF2	4101[b]	55 E = 9e-06	- - - - - - - - -	- - - - - - - - -	None in mouse	79.2% over 5788 bp (Xq23)
Sushi-XF2b	XF2	915	52 E = 1e-05	- - - - - - - - -	- - - - - - - - -	ESTs	73.8% over 940 bp (Xq23)

Note. The size (in bp) and percent nucleotide identity within each aligning block is given. The chromosomal location of the orthologous human region is in parentheses.
[a]Determined by result of Blastz of Human genome (July 2003), Tight Subset of Best Alignments. Only the largest aligning block within the ORF is described.
[b]ORF manually adjusted to start with methionine.

< 0.0018, score 139). Further BLASTP analysis (http://www.ncbi.nlm.nih.gov/BLAST/) confirmed the closest protein homology outside these genes to be with the sushi-ichi or sushi-san Gag proteins (Table 1, Fig. 1). The names of the novel elements refer to their homology to sushi-ichi and their mouse chromosomal location; *Sushi-XF2b* is the second element on band F2 of the X chromosome. On the basis of the Gag protein homology alone, phylogenetic classification showed all the novel sequences to be closely related to sushi-ichi itself or with the previously known mouse sushi-type elements, all of which emerge from the same branch (Fig. 2). It is unclear whether any phylogenetic divergence is due to them originating through retrotransposition of different elements, or is a consequence of sequence corruption/positive selection over time. We next used the sushi-ichi Pol protein in a TBLASTN ensemble search, which identified multiple retroviral sequences, but also identified a further mouse sushi-type sequence (*Sushi-14C1*) (Fig. 3). Three of the genes initially identified by *gag* homology also had associated *pol* homologous regions (Figs. 3, 4). As with the *gag* homologous sequences, BLASTP searches with all *pol* homologous regions confirmed their closest homologies to be with sushi-ichi (Table 1). The conservation of all the nine mouse sushi-type sequences with human was determined using BLASTZ comparative analyses of their sequences with the human genome (http://genome.ucsc.edu/). All elements exhibited strong conservation at the nucleotide level. With the exception of *Sushi-XF2b*, conservation between the two species was greater than 75%. Furthermore, all elements (including *Sushi-XF2b*), mapped to chromosomal regions of syntenic homology between mouse and human (Table 1).

Of the nine sushi retrotransposon-related sequences identified, two sequences corresponded to the *Rtl1*/*Peg11*/*Mor1* and *Peg10*/*Edr*/*MyEF-3* genes. Further characterisation of the novel sequences was undertaken to determine their genomic properties and potential for expression. Table 1 summarises these results. All contain open reading frames (ORFs); two are short but the remainder range from 1.1 to 4.1 kb. The presence of expressed sequence tags (ESTs) and cDNAs within the EST and non-redundant (nr) genomic databases provides evidence for the expression of all but *Sushi-XF2*, although its human ortholog was identified in a brain cDNA library (Nagase et al. 2000). *Sushi-15E3, Sushi-XC3*, and *Sushi-XD* all have full-length cDNAs encompassing their ORFs. *Ldoc1* is associated with several ESTs, the longest terminating at each end of the EST cluster and chosen for further sequence analysis. The human gene *KIAA1305* is orthologous to *Sushi-14C1* and the expression of both is supported by abundant EST evidence though no one EST covers the whole mouse

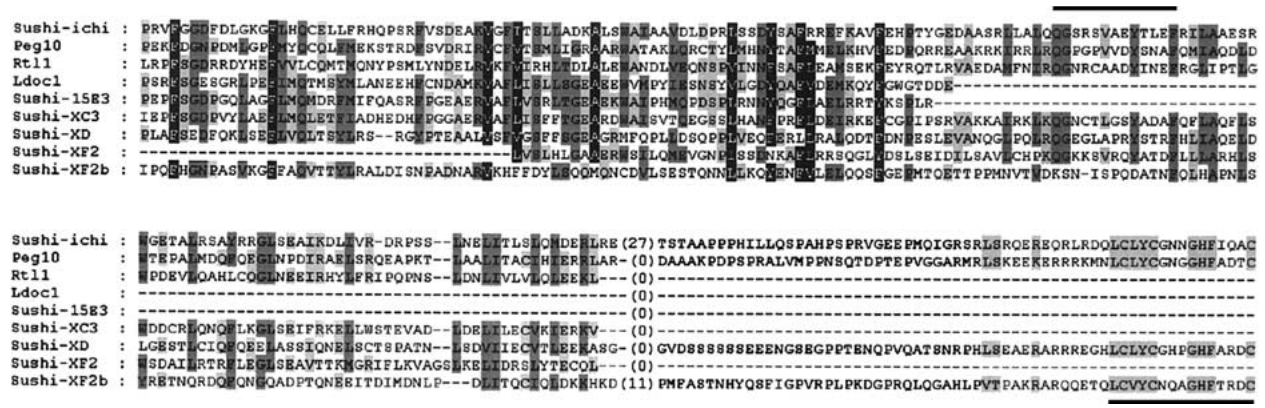

Fig. 1. Amino acid alignment of sushi-type retrotransposon gag sequences in mouse. The central gag conserved motif is overlined. Potential nucleocapsid (NC) nucleic acid binding domain is underlined.

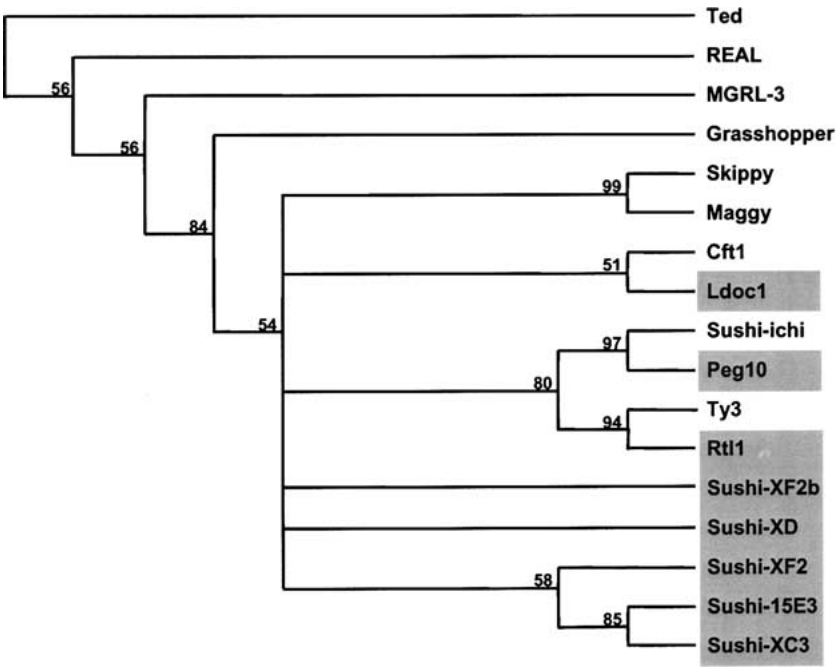

Fig. 2. Phylogenetic grouping of novel mammalian retrotransposon sequences among Ty3/gypsy retrotransposons on the basis of *gag* homology. Phylogeny was obtained using the *gag* sequence delineated in Figure 1. Trees were generated by the neighbor-joining distance algorithm (Saito and Nei 1987; bootstrap analysis 1000 replicates, Gap sites ignored). Nodes receiving less than 50% support were collapsed. The numbers indicate the percentage bootstrap support for the remaining nodes. The mammalian genes are in grey boxes.

ORF. Therefore, the whole ORF sequence was chosen for further analysis.

Mouse Sushi-Element Incompetence for Autonomous Retrotransposition

The ability of each novel element to autonomously retrotranspose was considered by sequence comparison with closely related replication competent retrotransposons (Figs. 1, 3, 4). It was determined that the elements were not capable of retrotransposition. First, none of the identified elements have recognisable LTRs. *Ldoc1, Sushi-15E3, Sushi-XC3, Sushi-XF2*, and *Sushi-XF2b* have only *gag* homology and no surrounding sequences were found to be related to sushi-type *pol* proteins. Conversely, *Sushi-14C1* only has homology to the 3′ end of the RT domain, the RNase H domain, and the IN domain sequences of *pol* but no surrounding protease site, CHR, or *gag* homology. *Peg10* has *pol* homology in a second ORF generated through a-1 frameshift but is truncated within the RT domain and thus lacks the remainder of the 3′ domains. *Rtl1* (reported as *Mor1*) was previously considered to be transpositionally incompetent as it lacks LTRs and a CHR domain and has mutations within its putative RT, RNaseH, and IN domains (Lynch and Tristem, 2003). Of the novel elements. *Sushi-XD* has the largest number of recognisable retrotransposon domains. Its *gag* regions contain both the central *gag* and zinc-finger (NC) conserved domains and it contains an ORF with *pol* homology generated by a −1 frameshift. The *pol* homologous area contains a protease active site motif but it truncates short of the RT domain and, therefore, like *Peg10*, lacks the 3′ domains.

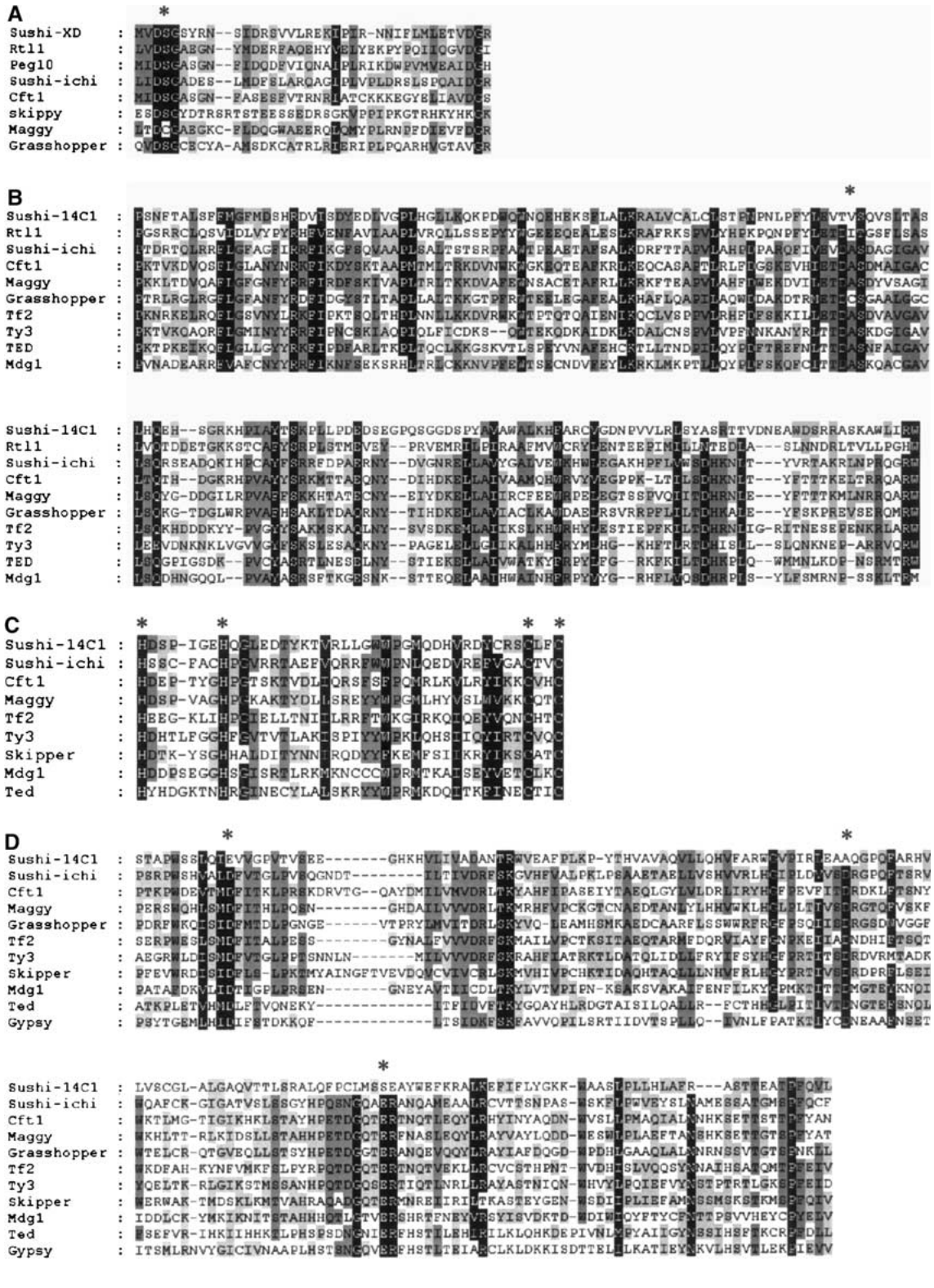

Fig. 3. Mouse Sushi-type element homology to Ty3/gypsy *pol* gene domains. **A:** Amino-acid alignment of Sushi-XD, *Rtl1*, and Peg10 protease domain homologous region to Ty3/gypsy retrotransposons. Highly conserved DSG protease active site asterisked. **B:** Amino acid alignment of Sushi-14C1 and *Rtl1* RNaseH homologous region to Ty3/gypsy retrotransposons. RNaseH highly conserved DAS motif asterisked. Mouse Sushi-type element homology to Ty3/gypsy pol gene domains. **C:** Amino acid alignment of Sushi-14C1 N-terminal integrase homologous region to Ty3/gypsy retrotransposons. Highly conserved H-H-C-C residues asterisked. **D:** Amino acid alignment of Sushi-14C1 central integrase homologous region to Ty3/gypsy retrotransposons. Highly conserved D-D-E residues asterisked.

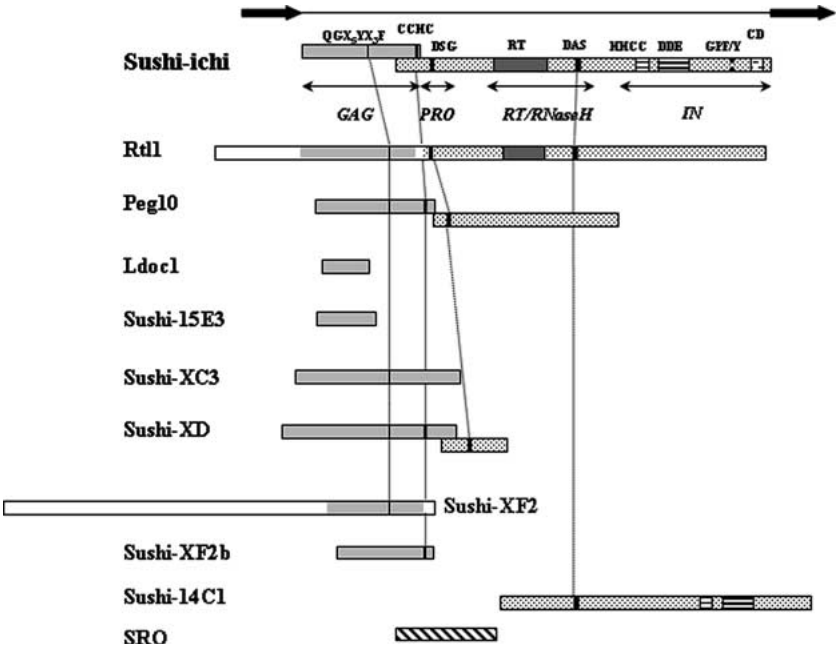

Fig. 4. Schematic representation (drawn to scale) of the structure of the sushi-ichi element including domain and subdomain organisation. Underneath is represented the nine mouse genes and their domains of homology. Gag domains are shaded grey and Pol domains, speckled. The bold arrows of the sushi sequence are the LTRs. Other regions are as follows: *CD* chromodomain; *PRO* protease domain; *IN* integrase domain, QGX_6YX_3F central gag Ty3/gypsy motif; *CCHC* RNA binding zinc-finger motif; *DSG* protease active site, *RT* reverse transcriptase domain, *DAS* RNaseH motif, *HHCC* & *DDE* & *GPY/F* integrase motifs. The hatched box represents the smallest region of overlap (SRO) common to the imprinted autosomal sequences (*Rtl1* and *Peg10*) but absent from the non-imprinted members (*Sushi-15E3* and *Sushi-14C1*).

Imprinted Expression

Five of the newly identified genes map to the X chromosome and four are autosomal. The X chromosome is subject to genomic imprinting with the paternally inherited X being preferentially inactivated in preimplantation mouse conceptuses (Takagi and Sasaki, 1975). Of the remaining two autosomal loci, neither map to known imprinted domains. However, unclustered imprinted genes have been documented on both mouse chromosomes 14 and 15 (Kato et al. 1998; Smith et al. 2003), hence the remaining two untested autosomal sequences. *Sushi-14C1* and *Sushi-15E3*, were tested for expression during development and for imprinting.

Sushi-15E3 encodes a 5-kb transcript detected in embryos by Northern blot analysis (data not shown). To investigate parental allele-specific expression, a strain-specific *HpaII* restriction fragment site polymorphism was identified between DBA2 and C57BL6/J, and 129/Sv and SWR laboratory strains of mice. Because *Rtl1* is known to have an imprinted sense transcript expressed from the paternally inherited chromosome and a reciprocally imprinted antisense transcript expressed from the maternal chromosome, strand-specific RT-PCR was used to identify and determine the imprinting status of transcripts from either orientation. No antisense transcripts were detected in any of the parental strains or crosses used in these experiments (data not shown). *HpaII* digestion of sense-strand specific RT-PCR products from RNA isolated from e14 embryos and placentas from DBA/2 x 129/Sv and C57BL6/J x SWR reciprocal crosses revealed *Sushi-15E3* to be expressed biallelically (Fig. 5A). Similarly, a strain specific *MseI* restriction fragment site polymorphism was identified between C57BL6/J and *Mus musculus molossinus* strains for *Sushi-14C1*. Strand-specific RT-PCR again found no antisense transcript in the ORF and digestion of product amplified from reciprocal crosses showed that expression of this gene was from both parental alleles in e16 embryos and placenta (Fig. 5B). Tissue-specific imprinting of these transcripts cannot be ruled out. However, our data show that, unlike *Rtl1* and *Peg10, Sushi-15E3* and *Sushi 14C1* are not subject to genomic imprinting in whole embryo and placenta.

Rtl1 *Methylation Analyses*

A CpG island within the *Rtl1* ORF was found to be heavily methylated throughout development (Fig. 6). Both the maternal and paternal chromosomes were hypermethylated in eggs and sperm and in pre- and post-implantation embryos. Reduced methylation was observed in e11.5 primordial germ cells at a time when genome-wide epigenetic reprogramming is in progress and erasure of methylation marks has begun. This methylation symmetry on the two parental chromosomes indicates that the two parental germlines do not methylate this retrotransposon-derived area differently and that the presence of a Ty3-gypsy homologous region in this domain is unlikely to confer differential epigenetic marks on the two parental chromosomes at this imprinted domain.

Discussion

The sushi-ichi family of retrotransposons described here is unusual in that its members have been con-

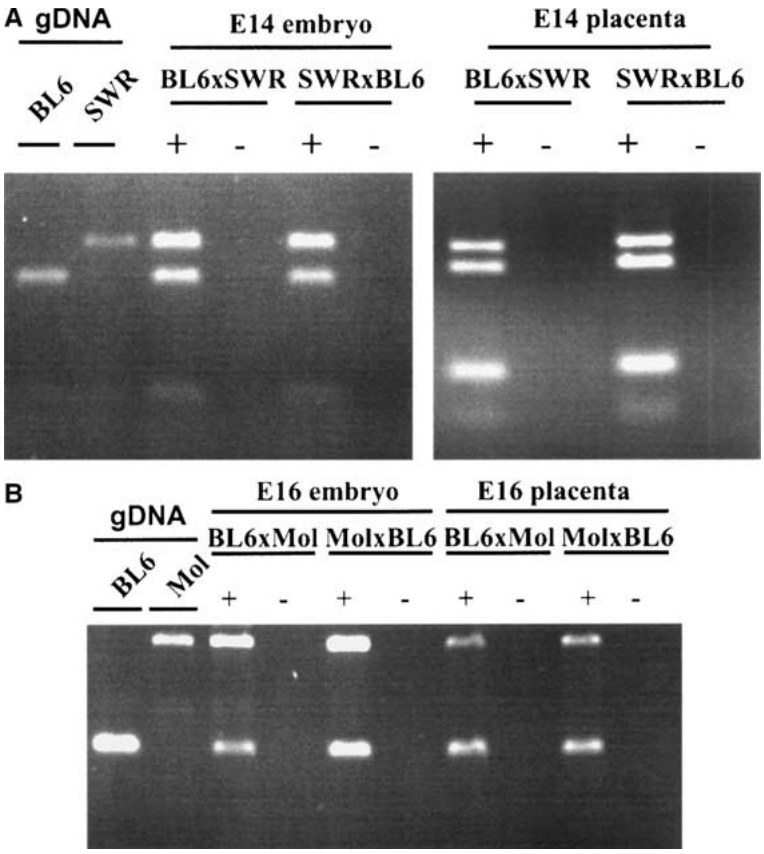

Fig. 5. *Sushi-15E3* and *Sushi-14C1* are not imprinted in embryo or placenta. **A:** Strand-specific RT-PCR followed by allele-specific digestion of *Sushi-15E3*. Both parental alleles are present in F1 hybrid cDNA in reciprocal crosses indicating biallelic expression in embryo and placenta. **B:** Strand-specific RT-PCR followed by allele-specific digestion of *Sushi-14C1*. Both parental alleles are present in the F1 hybrid cDNA in reciprocal crosses indicating biallelic expression in embryo and placenta. + with reverse transcriptase; – in the absence of reverse transcriptase.

served in the genomes of mammals. Importantly, they are conserved without the ability to autonomously retrotranspose and all thus far tested are expressed. The absence of LTRs at any of the mouse sushi-type genes is surprising and most likely further emphasises the importance of the protein coding regions in the conservation of the elements. To date, human *LDOC1* and mouse and human *PEG10* are the only members whose proteins have been attributed endogenous function. LDOC1 has an anti-proliferative effect functioning as a tumour suppressor through interaction between NF-KB and the LDOC1 leucine zipper domain (Nagasaki et al. 2003). This domain is also present in mouse *Ldoc1, Sushi-XC3*, PEG10, and *Sushi-XD* but not any other sushi-type sequences. Conversely, human and mouse *Peg10* have oncogenic properties (Shigemoto et al. 2001; Okabe et al. 2003). Other work has suggested that, like Syncytin-1 a retrovirally-derived gene located close by, human PEG10 has a placental function (Smallwood et al. 2003). Mouse *Peg10*, however, has been postulated to play a role in myelination as it binds at the promoter of myelin basic protein (*MBP*) (Steplewski et al. 1998). This function for the mouse gene is a notable departure from the proposed protein–protein interactions attributed to human PEG10 and LDOC1. This difference in function between the mouse and human PEG 10 may reflect the absence of a leucine-zipper domain in the mouse gene. The variety of functions attributed to LDOC1, PEG10, and Peg10 implies that the sushi-type genes are a highly adaptive class of proteins predominantly involved in gene regulation through utilisation of the DNA and protein-binding properties of the retrotransposon-derived NC and CA domains.

The data reported here show that all members of the sushi-ichi family of retrotransposons are not inherently subject to genomic imprinting. This suggests that the ability of sushi-ichi type retrotransposons to be imprinted lies either within sequences specific to the element or is a feature of the genomic environment in which the element is located. Comparative sequence analysis of the four autosomal members of this family identifies a small region of overlap present in the imprinted members and absent from the non-imprinted members (Fig. 4). This area extends from the central gag motif to just beyond the protease domain but does not include the NC domain as this is absent in *Rtl1*. Alternatively. *Rtl1, Peg10*, and the X-linked members of this family may have inserted into asymmetrically epigenetically-regulated regions such as the inactive X chromosome and imprinted domains and been influenced by the regional control elements conferring long-range allele-specific gene activity and repression. This is consistent with the biallelic expression of *Sushi-15E3* and *Sushi-*

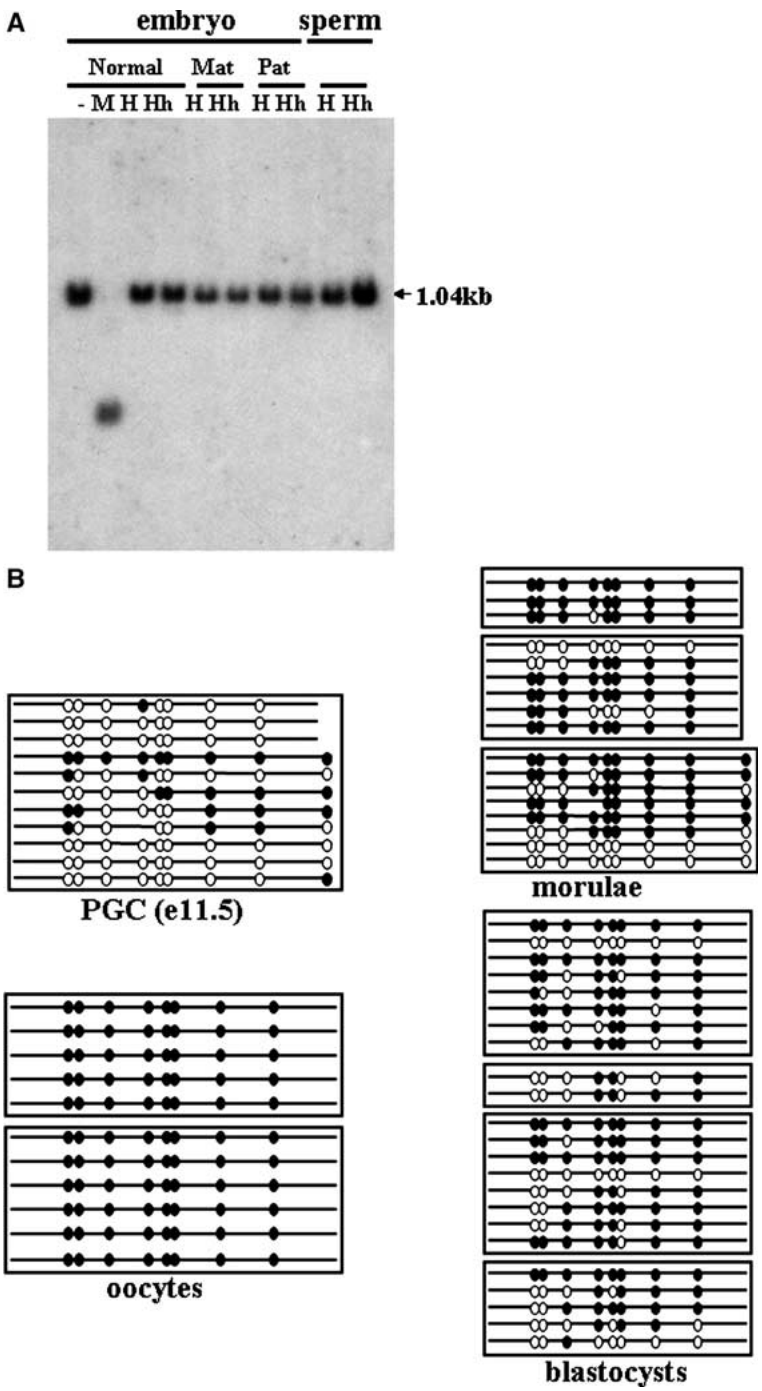

Fig. 6. A CpG island within *Rtl1* is not differentially methylated on the two parental chromosomes. **A:** Southern hybridisation of DNA isolated from embryos with uniparental duplications of mouse chromosome 12. Mat = Maternal duplication/paternal deficiency for chromosome 12. Pat = Paternal duplication/maternal deficiency for chromosome 12. Normal = littermate with normal maternal and paternal chromosome 12 contribution. All samples show lack of digestion with the methylation sensitive restriction enzymes *HpaII* and *HhaI* indicating that both parental chromosomes are fully methylated. **B:** Bisulphite sequencing analysis of DNA isolated from normal oocytes, morulae, blastocysts, and primordial germ cells. Methylated sites are represented by closed circles and unmethylated by open circles. Each line represents a different clone and bisulphite treatments were conducted on multiple different batches of material (indicated by boxing of clones). – = *BamH1* single digest; M = *BamH*1 + *Msp*1; H = *BamH*1 + *Hpa*II; Hh = *BamH*1 + *Hha*I double digests.

14C1, which are located 11.8 and 18.5 Mb, respectively, from the closest known imprinted genes, *Peg13* and *Htr2a* (Kato et al. 1998; Smith et al. 2003). Furthermore, there are several examples of unrelated imprinted genes that arose recently through non-autonomous retrotransposition of processed pseudogenes into regions that were, or subsequently became, imprinted (Chai et al. 2001; Hirotsune et al. 2003; Walter and Paulsen, 2003). Interestingly, no orthologues of *Rtl1, Peg10* or any of the other sushi-type sequences are found in the non-imprinted genomes of *Fugu* and chicken as determined by BLAST searches of genome databases (data not shown). However, despite this, one still cannot determine whether *Rtl1* and *Peg10* imprinting is due either to their insertion into an existing imprinted domain, or to a particular sequence they bear. Comparative analysis of the absence/presence and imprinting status of these and neighbouring genes in more closely related mammalian organisms that imprint with those that do not (such as marsupial and monotreme), might distinguish between these possibilities.

The methylation status of *Rtl1* suggests that the host genome recognises this element as "foreign" and, like for other "parasitic" elements such as IAPs, Alu elements, and LINES (Walsh et al. 1998; Hata and Sakaki, 1997; Lui et al. 1994), attempted to suppress the element by methylating CpG-rich sequences. It should be noted that the first exon of *Peg10* is associated with a tandem repeat-containing differentially methylated region (DMR) that is methylated on the maternally inherited allele. However, this region has no homology to Ty3/gypsy retrotransposons including *Rtl1* (Ono et al. 2003). In *Peg10*, the region bearing the Ty3/gypsy homology is 6.5 kb from exon 1, within exon 2. Exon 2 contains a CpG island; however, to our knowledge the methylation status of that exon has not been reported, so whether it is biallelically methylated like *Rtl1* remains to be determined.

In addition to investigating if *Rtl1* harboured a DMR, the possibility that it could be an insulator element was also considered. One of the eponymous members of Ty3/gypsy group, the *Drosophila* gypsy element, has been shown to act as a chromatin insulator functionally separating chromatin upstream of the element from that downstream (Gdula et al. 1996). Insulator function is involved in the mechanism of imprinting in at least one imprinted domain, the *Igf2-H19* locus (Hark and Tilghman, 1998; Bell and Felsenfeld, 2000). However, a simple sequence search revealed that neither *Rtl1* nor *Peg10* sequences are likely to be implicated in insulator function similar to the gypsy element as they do not contain conserved clusters of the TGCATA binding site for the insulator protein, Suppressor of Hairy-wing (Su[Hw]) (Spana et al. 1988).

More detailed analysis of the *in vivo* roles of the members of this family is likely to provide insight into the selective pressures that have given rise to this family of novel mammalian genes derived from mobile elements.

Acknowledgments. This study was supported by a PhD studentship from the BBSRC (N.A.Y.) and Wellcome Trust (N.P.) and by a grant from CR-UK. We thank Shuji Takada and Philippe Arnaud for help with the methylation analyses, Sarah Bray for assistance with the phylogenetic analysis, and Takashi Sado and Simão Rocha for technical assistance. We are grateful to Joanna Maldonado-Saldivia and Azim Surani for e11.5 primordial germ cells.

References

Arnaud P, Monk D, Hitchins M, Gordon E, Dean W, Beechey CV, Peters J, Craigen W, Preece M, Stanier P, Moore GE, Kelsey G (2003) Conserved methylation imprints in the human and mouse GRB10 genes with divergent allelic expression suggests differential reading of the same mark. Hum Mol Genet 12:1005–1019

Barlow DP (1993) Methylation and imprinting: from host defense to gene regulation? Science 260:309–310

Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 405:482–485

Birchler JA, Bhadra MP, Bhadra U (2000) Making noise about silence: repression of repeated genes in animals. Curr Opin Genet Dev 10:211–216

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

Chai JH, Locke DP, Ohta T, Greally JM, Nicholfs RD (2001) Retrotransposed genes such as Frat3 in the mouse Chromosome 7C Prader-Willi syndrome region acquire the imprinted status of their insertion site. Mamm Genome 12:813–821

Chesnokov IN, Schmid CW (1995) Specific Alu binding protein from human sperm chromatin prevents DNA methylation. J Biol Chem 270:18539–18542

Ferguson-Smith AC, Sasaki H, Cattanach BM, Surani MA (1993) Parental-origin-specific epigenetic modification of the mouse H19 gene. Nature 362:751–755

Gdula DA, Gerasimova Tl. Corces VG (1996) Genetic and molecular analysis of the gypsy chromatin insulator of Drosophila. Proc Natl Acad Sci USA 93:9378–9383

Hark AT, Tilghman SM (1998) Chromatin conformation of the H19 epigenetic mark. Hum Mol Genet 7:1979–1985

Hata K, Sakaki Y (1997) Identification of critical CpG sites for repression of L1 transcription by DNA methylation. Gene 189:227–234

Hirotsune S, Yoshida N, Chen A, Garrett L, Sugiyama F, Takahashi S, Yagami K, Wynshaw-Boris A, Yoshiki A (2003) An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature 423:91–96

Kato MV, Ikawa Y, Hayashizaki Y, Shibata H (1998) paternal imprinting of mouse serotonin receptor 2A gene Htr2 in embryonic eye: a conserved imprinting regulation on the RB/Rb locus. Genomics 47:146–148

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M et al. (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W (2003) Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis 35:88–93

Liu WM, Maria RJ, Rubin CM, Schmid CW (1994) Alu transcripts: cytoplasmic localisation and regulation by DMA methylation. Nucleic Acids Res 22:1087–1095

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

Matzke MA, Aufsatz W, Kanno T, Mette MF, Matzke AJ (2002) Homology-dependent gene silencing and host defense in plants. Adv Genet 46:235–275

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

Nagasaki K, Schem C, von Kaisenberg C , Bialiek M, Rosel F, Jonat W, Maass N (2003) Leucine-zipper protein. LDOC1, inhibits NF-kappaB activation and sensitizes pancreatic cancer cells to apoptosis. Int J Cancer 105:454–458

Nagase T, Kikuno R, Ishikawa Kl. Hirosawa M, Ohara O (2000) Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro. DNA Res 7:65–73

Nicholas KB, Nicholas HB, Deerfield DW (1997) Genedoc: analysis and visualisation of genetic variation. EMBNEW.NEWS 4:14

Okabe H, Satoh S, Furukawa Y, Kato T, Hasegawa S, Nakajima Y, Yamaoka Y, Nakamura Y (2003) Involvement of PEG10 in human hepatocellular carcinogenesis through interaction with SIAH1. Cancer Res 63:3043–3048

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

Pesce M, De Felici M (1995) Purification of mouse primordial germ cells by MiniMACS magnetic separation system. Dev Biol 170:722–725

Rubin CM, VandeVoort CA, Teplitz RL, Schmid CW (1994) Alu repeated DNAs are differentially methylated in primate germ cells. Nucleic Acids Res 22:5121–5127

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

Sanford JP, Clark HJ, Chapman VM, Rossant J (1987) Differences in DNA methylation during oogenesis and spermatogeneis and their persistence during early embryogenesis in the mouse. Genes Dev 1:1039–1046

Schwartz S, Kent WJ, Smit A, Zhang Z, Baertsch R, Hardison RC, Haussler D, Miller W (2003) Human-mouse alignments with BLASTZ. Genome Res 13:103–107

Seitz H, Youngson N, Lin SP, Dalbert S, Paulsen M, Bachellerie JP, Ferguson-Smith AC, Cavaille J (2003) Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. Nat Genet 34:261–262

Shigemoto K, Brennan J, Walls E, Watson CJ, Stott D, Rigby PW, Reith AD (2001) Identification and characterisation of a developmental regulated mammalian gene that utilises-1 programmed ribosomal frameshifting. Nucleic Acids Res 29:4079–4088

Smallwood A, Papageorghiou A, Nicolaides K, Alley MK, Jim A, Nargund G, Ojha K, Campbell S, Banerjee S (2003) Temporal regulation of the expression of syncytin (HERV-W), maternally imprinted PEG10, and SGCE in human placenta. Biol Reprod 69:286–293

Smith RJ, Dean W, Konfortova G, Kelsey G (2003) Identification of novel imprinted genes in a genome-wide screen for maternal methylation. Genome Res 13:558–569

Spana C, Harrison DA, Corces VG (1988) The Drosophila melanogaster suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon. Genes Dev 2:1414–1423

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

Takada S, Paulsen M, Tevendale M, Tsai CE, Kelsey G, Cattanach BM, Ferguson-Smith AC (2002) Epigenetic analysis of the Dlk1-Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with Igf2-H19. Hum Mol Genet 11:77–86

Takagi N, Sasaki M (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature 256:640–642

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

Walsh CP, Chaillet JR, Bestor TH (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 20:116–117

Walter J, Paulsen M (2003) The potential role of gene duplications in the evolution of imprinting mechanisms. Hum Mol Genet 12 Spec No 2:R215–220

Yoder JA, Walsh CP, Bestor TH (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13:335–340