Genetica **107:** 209–238, 1999.

Genomes were forged by massive bombardments with retroelements and retrosequences

Jürgen Brosius
Institute of Experimental Pathology/Molecular Neurobiology, ZMBE, University of Münster, Von-Esmarch-Str. 56, D-48149 Münster, Germany (Phone: +49 251 835 8511; Fax: +49 251 835 8512; E-mail: RNA.world@uni-muenster.de)

Accepted 11 February 2000

Key words: evolution, genomic symbionts, host defense, impact on genomes, retronuons, reverse transcription, template switching

Abstract

Retroposition is an efficient route to move coding regions around the genome 'in search' of novel regulatory elements and to shotgun regulatory elements into the genome 'in search' of new target genes. The templates for such retrogenes are mRNAs, and for regulatory retronuons (nuon = any definable nucleic acid sequence) usually small non-mRNAs (snmRNAs). An example in support of the 'master gene' model for SINEs (short interspersed repetive elements) is provided with neuronal BC1 RNA. Furthermore, an alternative explanation of LINE (long interspersed repetive elements) involvement in the generation of SINEs is given. I will also argue that the status of transposable elements with respect to the host resembles more symbiosis than parasitiasis and that host defense is often lenient as if even to 'tolerate or support' retronuons. Finally the paradox of evolution's lack of foresight and the future exaptive use of retronuons is being dealt with by referring to W.F. Doolittle's 'Hierarchical Approaches to Genome Evolution'.

Introduction

The first sequence of a large ribosomal RNA (Brosius et al., 1978) was not worth much for secondary structure predictions until additional rRNA sequences became available (Woese et al., 1983). Likewise, the scientific value of the entire human genome will multiply when not only another mammalian genome (mouse) but also additional primate sequences including that of our closest relative, the chimpanzee, will become available. I can hardly disagree with opinions voicing that the 3–5% protein coding regions of higher eukaryotic genomes will be boring and that the excitement lies in understanding the remainder of the total sequence. Consequently, it is highly significant to extend genomic sequencing of eukaryotic model organisms featuring compressed genomes (pufferfish or the plant *Arabidopsis thaliana*) to organisms with genome sizes that are in the 'normal' range for higher Eucarya. In vertebrates, it is evident already that retroelements and retrosequences – here termed retronuons where a nuon is any definable nucleic acid sequence (Brosius & Gould, 1992, 1993) – have contributed the majority of the extra genomic mass. Retronuon integration has a wide spectrum of effects on the host genome ranging from none to dramatic and from negative (Deininger & Batzer, 1999) to positive. The potential impact of these retronuons on genomic function (Georgiev, 1984; Brosius, 1991; Brosius & Gould, 1992; Shapiro, 1992; McDonald, 1993; Makalowski, Mitchell & Labuda, 1994; Nouvel, 1994; Flavell, 1995; Wessler, Bureau & White, 1995; Brosius & Tiedge, 1996; Britten, 1996; Kidwell & Lisch, 1997; Schmid, 1998; Jurka, 1998; Tomilin, 1999; Brosius, 1999a, b, c) is gradually being accepted. The dimension of the effect depends on whether the event took place in somatic cells or the germ line.

Importantly, the effect on the host depends largely on the location of the newly integrated retronuon with respect to resident genes or regulatory elements. In comparison to the last review on the subject (Brosius, 1999b), updated tables with examples of

Table 1. Vertebrate regulatory elements or parts of coding regions generated by retroelements

Retronuon	Gene under its influence	Species	Ancestor of source nuon	Serves as	References
LTR	cDNA 7, cDNA$_\gamma$	Human	THE-1	Polyadenylation signal	Paulson et al. (1987)
LTR	Sex-limited protein (slp)	Mouse	5′ LTR of C-type retrovirus (*imp1*)	Promoter	Stavenhagen and Robins (1988); Robins and Samuelson (1992); Ramakrishnan and Robins (1997)
LTR	Oncomodulin	Rat (but not mouse)	IAP	Promoter and first exon	Banville and Boie (1989)
LTR	MIPP	Mouse	IAP	Promoter	Chang-Yeh et al. (1991)
LTR	AF-3	Human	RTVL-H	Promoter	Feuchter et al. (1992)
LTR	AF-4 (CDC4L homology)	Human	RTVL-H	Promoter	Feuchter et al. (1992)
LTR	PLT	Human	RTVL-H	Polyadenylation signal	Goodchild et al. (1992)
LTR	cH-6	Human	RTVL-H	Polyadenylation signal	Mager (1989)
LTR	cH-7	Human	RTVL-H	Polyadenylation signal	Mager (1989)
LTR	cPB-3	Human	RTVL-H	Polyadenylation signal	Mager (1989)
LTR	PLA2L (phospholipase A2 homology)	Human	RTVL-H	Promoter	Feuchter-Murthy et al. (1993)
LTR	Calibindin D28K	Human	RTVL-H	Promoter	Liu and Abraham (1991)
LTR	ZNF80 zinc finger gene	Human	ERV9	Promoter	Di Cristofano et al. (1995)
LTR	Growth factor pleitropin (PTN)	Human	HERV-E (RTVL-1)	Trophoblast-specific promoter	Schulte et al. (1996, 1998)
LTR	Leptin receptor (OBRa)	Human	HERV-K	Alternative splicing, inclusion of 67 LTR-derived aa into C-terminus of OBRa protein	Kapitonov and Jurka (1999)
LTR-IS	A1	Mouse	MuRRS	Polyadenylation signal	Baumruker et al. (1988)
LTR-IS	A3	Mouse	MURRS	Polyadenylation signal	Baumruker et al. (1988)
LTR	Aromatase	Chicken	Retrovirus	Promoter and 5′ exon	Matsumine et al. (1991)
CR1	Lysozyme	Chicken	Retrovirus	Transcriptional silencer	Baniahmad et al. (1987)
L1	Thymidylate synthase	Mouse	LINE	Polyadenylation signal	Harendza and Johnson (1990)
L1	Insulin I gene	Rat	LINE	Transcriptional silencer	Laimins et al. (1986)
LINE	α_s1-casein E	Goat	LINE	mRNA stability	Pérez et al. (1994)
L1	Apolipoprotein(a)	Human	LINE	Transcriptional enhancer	Yang et al. (1998)
L1	Proteasome activator PA28β (PMSE2b)	Mouse	LINE	Promoter	Zaiβ and Kloetzel (1999)
HERV-E	Salivary amylase gene	Human	Retrovirus	Promoter	Samuelson et al. (1990); Emi et al. (1988); Ting et al. (1992)
HRES-1	Transaldolase	Human	Retrovirus	Part of the coding sequence	Banki et al. (1994)

Table 1. Continued

Retronuon	Gene under its influence	Species	Ancestor of source nuon	Serves as	References
The-1, IAP	Immunoglobulin heavy chain	Human Mouse	Retrotransposon	Protein sequence building blocks	Hakim et al. (1994)
LTR	Leptin	Human	MER11	Placental enhancer	Bi et al. (1997)
ALF	Annexin VI, interleukin-4, protein kinase C-β	Human	LINE-2	Potent T-cell-specific silencer	Donnelly et al. (1999)
Bov-B LINE	Bucentaur (b*bcnt*)	Ruminantia	LINE	Large part of protein coding sequence	Takahashi et al. (1998)

For a definition of the differences between retroelements (this table) and retrosequences (Tables 2–7) see text.
Not all examples are proven exaptations. Especially events that date back not much more than a few million years could only be potential exaptations (potaptations according to Brosius & Gould, 1992, 1993).
NGF-inducible cAMP-extinguishable retrovirus-like (NICER) elements have been described (Cho et al., 1990). No association with a gene under their control has been reported.

recruited/exapted (Gould & Vrba, 1982) retronuons in vertebrates are presented. Furthermore, tables with examples that cannot be defined as exaptations as of yet have been added. These may be termed potential exaptations or potaptations (Brosius & Gould, 1992, 1993). In addition, generation of SINEs (short interspersed repetitive elements) from 'master genes' and the involvement of LINEs (long interspersed repetitive elements) in their biogenesis is dealt with. Furthermore, measures of the host (and mobile element) are discussed that may mitigate negative effects of transposable elements on the host. Finally, I will expand on questions concerning the perseverance of reverse transcription in genomes, despite the lack of an obvious need in the cell, on parasitism versus symbiosis, the lack of foresight in evolution and on hierarchical approaches to genome evolution (Doolittle, 1989) in the light of retronuon biology.

Biogenesis of SINEs and effects of retronuons on neighboring genes

Transposable elements (TEs) include DNA transposons and elements that disperse via RNA intermediates, so-called retroposons. Retroposons can be further divided into retroelements and retrosequences. Retroelements comprise proviruses and LTR (long terminal repeat) retrotransposons as well as non-LTR retrotransposons such as LINEs. Their common denominator is the presence of an open reading frame (ORF) encoding reverse transcriptase at least in the (rarer) full-length versions. Retrosequences arise from virtually any cellular RNA that has the potential to be reverse transcribed in the cell: ribosomal, messenger and small stable RNAs (Brosius & Gould, 1992; Nouvel, 1994; Brosius, 1999b). Reverse transcriptases encoded by the aforementioned retroelements play an important part in their biogenesis (Singer, 1995; Jurka, 1998).

Parts of retroelements are efficient modulators of gene expression

Often the remnants of retroelements such as lone LTRs or other domains have been recruited as regulatory elements that act on targeted genes as promoters, transcriptional enhancers/silencers, splice sites, polyadenylation signals and even as protein coding domains (Table 1).

Small RNA derived retrosequences are frequent modulators of gene expression

Retrosequences derived from small RNA templates (mainly SINEs) whose generation apparently depends on retroelement encoded reverse transcriptases play a similar role as parts of the aforementioned retroelements in regulating targeted genes (Table 2).

Models of SINE amplification: most repeats are transposable or only certain master genes can give rise to new SINEs

There are still two conflicting models concerning the mode of SINE generation as summarized by (Brookfield, 1994): "One is that they are non-autonomous transposable elements. Here the Alu sequence transposes sequentially from site to site, retaining its capacity for further transposition. . . . This model sees

Table 2. Vertebrate regulatory elements or parts of coding regions generated by retrosequences

Retronuon	Gene/nuon that is under its influence	Species	Ancestor of source nuon	Serves as	References
B2	MHC class I genes	Mouse	$tRNA^{Lys}$	Polyadenylation signal	Kress et al. (1984)
B2	$B2^{+}$ $mRNA_x$	Mouse	$tRNA^{Lys}$	Polyadenylation signal	Ryskov et al. (1984)
B2	Glutathione S-transferase	Mouse	$tRNA^{Lys}$	Polyadenylation signal	Rothkopf et al. (1986)
B2	Various	Rodents	$rRNA^{Lys}$	mRNA stability	Clemens (1987)
B2	Fourth component of complement (C4) in $H\text{-}2^{k}$ haplotype	Mouse	$tRNA^{Lys}$	Located in intron 10; reduces expression rate to 1/10 of non-$H\text{-}2^{k}$ mice	Zheng et al. (1992)
B2	Muscle γ-phosphorylase kinase	Mouse	$tRNA^{Lys}$	Polyadenylation signal	Maichele et al. (1993)
B2	MOK-2 zinc-finger protein	Mouse	$tRNA^{Lys}$	Exerts a negative *cis*-acting effect on *MOK-2* promoter activity	Arranz et al. (1994)
B2	Leukemia inhibitory factor receptor (LIFR)	Mouse	$tRNA^{Lys}$	Generating new splice variant that leads to a soluble form of LIFR	Owczarek et al. (1996); Michel et al. (1997)
C repeats	MHC	Rabbit	$tRNA^{Gly}$	Polyadenylation signal	Rebiere et al. (1987); Krane and Harsdison (1990)
C repeats	Major apoprotein of pulmonary surfactant	Rabbit	$tRNA^{Gly}$	Polyadenylation signal	Boggaram et al. (1988); Krnae and Hardison (1990)
C repeats	Cytochrome P-450 isozyme 4	Rabbit	$tRNA^{Gly}$	Polyadenylation signal	Okino et al. (1985); Krane and Hardison (1990)
CHR-1 repeats	EP3B and EP3C prostaglandin E2 receptors	Bovine	$tRNA^{Glu}$	Protein domain	Shimamura et al. (1998)
ID		Rat	$tRNA^{Ala}$ via neural BC1 RNA	Enhancer	McKinnon et al. (1986)
ID	pIL2	Rat	$tRNA^{Ala}$ via neural BC1 RNA	mRNA stability	Glaichenhaus and Cuzin (1987)
B1	pIL2, pIL8	Mouse	SRP RNA	mRNA stability	Vidal et al. (1993)
B1	Immunoglobulin κ light chain	Mouse	SRP RNA	Negative regulation of transcription	Saksela and Baltimore (1993)
Alu	Haptoglobin related gene	Human	SRP RNA	Transcriptional enhancer	Oliviero and Monaci (1988)
Alu	θ1 globin	Higher primates	SRP RNA	CCAAT box of promoter	Kim et al. (1989)
Alu	ε-globin	Human	SRP RNA	Transcriptional modulation	Wu et al. (1990)
Alu	7.02 bidirectional promoter	Monkey	SRP RNA	Transcriptional reducer	Saffer and Thurston (1989)
Alu	c-myc	Human	SRP RNA	Transcriptional modulation	Tomilin et al. (1990)
Alu	Adenosine deaminase	Human	SRP RNA	Transcriptional enhancer	Aronow et al. (1992)
Alu	Proliferating cell nuclear antigen (PCNA)	Human	SRP RNA	Transcriptional silencer	Sell et al. (1992)
Alu	Mitochondrial hinge protein	Human	SRP RNA	Transcriptional enhancer	Liu and Bradner (1993)
Alu	SV40 origin	Human	SRP RNA	Transcriptional enhancer	Saegusa et al. (1993)
Alu	FcεRI-γ	Human	SRP RNA	Transcriptional regulation (positive and negative)	Brini et al. (1993)

Table 2. Continued

Retronuon	Gene/nuon that is under its influence	Species	Ancestor of source nuon	Serves as	References
Alu	Keratin 18 (human)	Mouse (trans-genic)	SRP RNA	Transcriptional insulation; Alus provide retinoic acid receptor binding sites	Thorey et al. (1993); Neznanov and Oshima (1993); Vansant and Reynolds (1995)
Alu	CD8α	Human	SRP RNA	Transcriptional enhancer (located in last intron)	Hambor et al. (1993)
Alu	α-3 acetylcholine receptor subunit	Human	SRP RNA	Alternative splicing	Mihovilovic et al. (1993)
Alu	In several protein coding regions	Primates	SRP RNA	Generating new splice variants, potentially contributing new protein domains	Reviewed in Makalowski et al. (1994)
Alu	Interferon receptor, $IFNR_1$ R-2	Human	SRP RNA	Alt. splicing, part of protein cod. region	Mullersman and Pfeffer (1995)
Alu-J	Double-stranded RNA-specific editase (RED1/ADAR2)	Human	SRP RNA	40 Alu-derived aa are added via alternative splicing; protein product has enzymatic activity	Gerber et al. (1997)
Alu-J	Cathepsin B	Human	SRP RNA	Alt. splicing of exon 2 in 5′ UTR	Berquin et al. (1997)
Alu	β1C-2 integrin subunit	Human	SRP RNA	Alt. splicing, part of protein cod. region	Svineng et al. (1998)
Alu	DNA (cytosine-5) methyltransferase (CpG MTase)	Higher Primates	SRP RNA	Alt. splicing, part of protein cod. region	Hsu et al. (1999)
Alu	7.8 kb RNA	Human	SRP RNA	Induction of expression of a ST receptor in *trans*	Almenoff et al. (1994)
Alu	Wilms' tumor gene (WT1)	Human	SRP RNA	Intronic transcriptional silencer	Hewitt et al. (1995)
Alu	BRCA-1 gene, ERF-3	Human	SRP RNA	Estrogen-dependent transcriptional enhancers	Norris et al. (1995)
Alu	Parathyroid hormone gene	Human	SRP RNA	Negative calcium response element	McHaffie and Ralston (1995)
Alu	Poly(ADP-ribosyl) transferase (ADPRT) gene	Human	SRP RNA	Transcription regulation	Schweiger et al. (1995); Oei et al. (1997)
Alu	Potentially many genes	Human	SRP RNA	Transcriptional modulation via binding of YY1 protein	Humphrey et al. (1996)
Alu	Myeloperoxidase gene promoter	Human	SRP RNA	Composite SP1-thyroid hormone-retinoic acid response element	Piedrafita et al. (1996)
Alu	α3 nicotinic receptor subunit	Human	SRP RNA	Transcription modulation	Fornasari et al. (1997)
RRE[a]	Erythropoietin recept. prom.	Mouse	?	Transcription inhibitor	Youssoufian and Lodish (1993)
Highly repet. element[a]	c-Ha-ras	Human	?	Blocks transcriptional readthrough	Lowndes et al. (1990)
MIR	Nicotinic acetylcholine receptor α subunit	Human	tRNA	Generating new splice variant, contributes to protein coding region	Murnane and Morales (1995)
MIR	β-tubulin	Human	tRNA	Polyadenylation signal	Murnane and Morales (1995)

Table 2. Continued

Retronuon	Gene/nuon that is under its influence	Species	Ancestor of source nuon	Serves as	References
MIR	Follitropin receptor	Sheep	tRNA	Polyadenylation signal	Murnane and Morales (1995)
MIR	Clone c-zrog02	Human	tRNA	Polyadenylation signal	Murnane and Morales (1995)
MIR	Clone NIB1273	Human	tRNA	Polyadenylation signal	Murnane and Morales (1995)
γ-actin	Salivary amylase gene	Human	mRNA	Promoter	Samuelson et al. (1990; 1996) Emi et al. (1988)

For a definition of the differences between retroelements (Table 1) and retrosequences see text.
Not all examples are proven exaptations. Especially events that date back not much more than a few million years could only be potential exaptations (potaptations according to Brosius and Gould, 1992, 1993).
[a]Resemblance to a known repetitive element not yet established.

Alus as selfish DNAs and does not require them ever to provide any useful function for the organism. An alternative model is the single source or 'master gene' model, envisaging Alu sequences as pseudogenes created by retroposition from a single functioning gene located at a specific locus . . . . This model hypothesizes that the source (or 'master') Alu is conserved by selection arising from its as-yet–unknown function." Much of the evidence favors the latter, master gene model (Willard, Nguyen & Schmid, 1987; Deininger & Slagel, 1988; Britten et al., 1988; Quentin, 1988; Labuda & Striker, 1989; Jurka & Milosavljevic, 1991; Deininger et al., 1992). The relationship of the first known master gene for a SINE, BC1 RNA, to ID repetitive elements (Kim et al., 1994; Deininger, et al., 1996) further supports this model. The findings on the interrelationship of BC1 RNA, its gene and related ID repetitive elements is further illuminating in that it regurgitates a frequently overlooked fact: black and white scenarios are relatively rare in biology; reality often lies between extremes (albeit sometimes much closer to one than to the other). BC1 RNA, generated by retroposition of a $tRNA^{Ala}$, has been exapted into a function in rodent nerve cells (Brosius & Tiedge, 1995). Depending on the rodent species, BC1 RNA is also the master gene for a few hundred up to 10,000 ID repetitive elements generated by retroposition. In some rodent species there is a large number of additional ID elements ($\geq$ 100,000 in rat) whose consensus sequences point to a few different master genes. In the rat, up to four additional master genes most likely belong to the few ID elements that are actively being transcribed. For a more detailed discussion see (Brosius, 1999b).

It is quite rare that retroposition yields a novel transcribed gene encoding a small RNA, such as BC1 RNA, BC200 RNA (see below) or transcribed SINEs, such as the postulated additional ID element master gene(s). Much more frequent is recruitment or exaptation of a SINE as new regulatory element for a targeted gene. This includes enhancement or silencing of transcription, providing (alternative) splice sites, polyadenylation signals, modulation of mRNA stability and small protein coding domains (Table 2).

Why do many SINEs contain LINE sequences at their 3′ end? Template switching as alternative explanation

SINEs and other retrosequences depend on LINEs and/or other retroelements to provide for the key enzyme reverse transcriptase ((Feng et al., 1996); see also below). In addition, LINEs have been suggested to play a further role in the propagation of SINEs (Ohshima et al., 1996; Okada et al., 1997) as well as mRNA derived retrosequences (Jurka, 1997). It is thought that a strong-stop DNA (generated by reverse transcription) with a primer tRNA was integrated into the 3′ portion of a LINE, which gave rise to a primordial SINE. Here, I represent an alternative explanation with an equally tight association of a LINE to a tRNA – reverse transcription of a LINE is aborted (presumably again through a strong stop), and then the template switched (Pathak & Hu, 1997) to a nearby (primer) tRNA or other small RNA, possibly co-packaged in virus-like particles (Kolosha & Martin, 1995). Some of these SINEs with 5′ tRNA domains and 3′ LINE domains – if transcribed – could be efficient founder genes for additional SINEs, presumably by providing more efficient priming sites for reverse transcriptase. However, not all SINEs feature this composite structure, notable exceptions being Alu, B1, B2, ID and S1 repetitive elements. Hence, alternative features must exist that earmark small RNAs as efficient master genes.

Table 3. Vertebrate genes generated by retrosequences

Retrogene; Pattern of expression; chromosome	Species	Source gene; pattern of expression; (# of introns); chromosome	Template for reverse transcription	Hallmarks of retrosequences			References
				intron loss	A-stretch at 3′ end of founder RNA	direct repeats	
Insulin I; Langerhans islets; (1)	Murids	Insulin II; Langerhans islets; (2)	Part. proc. hnRNA	(+)	+	(+)	Soares et al. (1985)
Phosphoglycerate kinase (Pgk-2); testes	Mammals	Pgk-1; constitutive; (10); X-linked	Mature mRNA	+	(+)	(+)	McCarrey and Thomas (1987); Boer et al. (1987); Adra et al. (1988)
Zfa; testes; chr 10	Mouse	Zfx; ubiquitous; (≥ 10); X-linked	Mature mRNA	+	(+)	+	Ashworth et al. (1990)
Pyruvate dehydrogenase (Pdha2); testes	Human mouse	Pdha1; constitutive; (10); X-linked	Mature mRNA	+	(+)	(+)	Dahl et al. (1990) Fitzgerald et al. (1992)
N-myc2; brain and liver tumours[a]	Sciuridae rodents, e.g. woodchucks	N-myc1; in development and various adult tissues; (2)	Mature mRNA	+	(+)	+	Fourel et al. (1990, 1992); Sugiyama et al. (1989, 1999); Robertson et al. (1991); Quigon et al. (1996)
NB-1 or CLP; epithelial tissue; chr 10	Human	Calmodulin CaMIII; ubiquitous; (5)	Mature mRNA	+		(+)	Yaswen et al. (1992); Rhyner et al. (1992); Berchtold et al. (1992)
Carcinoma associated antigen, GA733-1	Human	GA733-2; placenta, carcinoma; (8)	mRNA	+			Linnenbach et al. (1989, 1993)
Glutamate dehydrogenase (GLUD2); retina, testes, brain; X-linked	Human	GLUD1; ubiquitous; (13); chr 10	Mature mRNA	+	+	+	Shashidharan et al. (1994)
S-adenosyl- methionine decarboxylase (AMD2); liver and other tiss.; chr 12	Mouse	(AMD1); ubiquitous; (9)	Mature mRNA	+	+	+	Persson et al. (1995, 1999); Nishimura et al. (1998)
Glucose-6-phosphate dehydrogenase (G6PD-2); testes	Mouse	G6PD-1; constitutive (10); X-linked	Mature mRNA	+	+	+	Hendriksen et al. (1997)
Hypoxoxanthine transferase, *HPRT-2*; liver	Kangaroo	*HPRT-1*; ubiquitous; (8); X chr	Mature mRNA	+		(+)	Noyce et al. (1997); Noyce and Piper (1994)

Table 3. Continued

Poly(A) binding protein 2 (Pabp2); spermatogenic cells	Mouse	Pabp1; spermatogenic and somatic cells (several)	Mature mRNA	+			Kleene et al. (1998)
C_γ catalytic subunit of cAMP-dependent protein kinase[b]; testes; chr 9	Catarrhini primates	Cα catalytic subunit of cAMP- dependent protein kinase; ubiquitous; (∼9); chrom. 19	Mature mRNA	+	(+)	(+)	Reinton et al. (1998)
H430 encoding a splicing factor; pancreas, spleen, prostate etc.; chr 11	Human	PR264/SC35; thymus, spleen, kidney, lung etc.; (2); chr 17	Mature nRNA	+	(+)	+	Soret et al. (1998)
CDY genes (at least one family member); testes; Y chr	Anthropoidea	CDYL; ubiquitous; (9); chr 13 (human)	Mature mRNA	+			Lahn and Page (1999)
Proteasome activator PA28, β- subunit (PMSE2b); constitutive; chr 14	Mouse	PMSE2, gamma interferon inducible; (10); chr 11	Mature mRNA	+	+	+	Zaiβ and Kloetzel (1999)
Centrin, *Cetn1*; testes; chr 18	Mammals	*Cetn2*; neonatal testes, oviduct; (4); X chr	Mature mRNA	+		+	Hart et al. (1999)
XAP-5-like (X5L)[c]; chr 6	Human Mouse	XAP-5; (12); X- linked	Mature mRNA				Sedlacek et al. (1999)
BC1 RNA; neurons; chr 7 (mouse)	Rodents	$tRNA^{Ala}$; ubiquitous	Non-mRNA	n.a.	+	(+)	DeChiara and Brosius (1987); Martignetti and Brosius (1993a)
BC200 RNA; neurons; chr 2 (human)	Anthropoidea	Free Alu monomer	Non-mRNA	n.a.	+	(+)	Martignetti and Brosius (1993b)

Not all examples are proven exaptations. Especially events that date back not much more than a few million years could only be potential exaptations (potaptations according to Brosius & Gould, 1992, 1993).
[a]See also s-myc, sm-myc in rodents and mycL2 in primates; role in apoptosis.
[b]Protein product not confirmed yet.
[c]Intron in 5′ UT.

The role of SINEs in the evolutionary roulette

It has been argued before that the shotgunning of mobile regulatory elements through the genome is a powerful way to impose plasticity to the genome (McClintock, 1948; Georgiev, 1984; Brosius & Gould, 1992). Kermekchiev et al. (1991) have shown that various combinations of *in vitro* shuffled promoters and enhancers function with high efficiency. The authors proposed, therefore, that "a generally permissive enhancer/promoter interaction is of evolutionary benefit for higher eukaryotes: by enhancer shuffling, genes could be easily brought under a new type of inducibility/cell-type specificity". Retronuons as shown in Tables 1 and 2 are precisely the mobile elements that could have achieved and still accomplish a constant shuffling of regulatory elements on a trial-and-error basis.

The essence of the human/chimp divergence at the genomic level

A comparison between genomes of closely related species will reveal the driving forces of speciation. Especially the full sequences of human and chimpanzee will disclose that there are few genes that arose *de novo* in one of the two species. Naturally, we expect some genes in one species versus the other to have arisen by duplication (via recombination or

Table 4. Vertebrate genes probably generated by retrosequences

Retrogene; pattern of expression; chromosome	Species	Presumable source gene; pattern of expression; (# of introns); chromosome	Template f. reverse transcript.	Age of retrogene	References
Replication-dependent histone genes[a]	Various	'Replacement' variant histone genes	mRNA	Metazoans and plants	Reviewed in: Kedes et al. (1979); Hentschel and Birnstiel (1981)
G protein α subunit, Gi class, Gnaz	Mammals	Gnai; (8)	Part. proc. hnRNA		Wilkie et al. (1992)
G-protein coupled receptors[b]	Various	G-protein coupled receptors	mRNA		Reviewed in: Gentles and Karlin (1999); Brosius (1999)
Potassium channels	Various	Potassium channel	mRNA		Reviewed in: Strong et al. (1993)
Class III POU domain protiens e.g. SCIP (or Tst-1, Oct-6); or: Brn-3b; brain; X chr; early development and brain	Various	POU domain transcription factor; (multiple) Brn-3a or 3c; brain; chr 14 and chr 18	mRNA		Kuhn et al. (1991); Hara et al. (1992); Theil et al. (1994); Alvarez-Bolado et al. (1995); Atanasoski et al. (1995); Levavasseur et al. (1998)
Forkhead transcription factors, e.g.: brain factor-2 (HBF2)[c]; fetal brain; chr 14; MFH-1	Various	Brain factor-1 (HBF1); fetal brain; (1) chr 14	mRNA		Wiese et al. (1995); Ernstsson et al. (1996); Frank and Zoll (1998); Miura et al. (1997)
Inducible heat shock genes[d]	Various	Constitutive heat shock genes	mRNA	Ancient	Hunt and Morimoto (1985); Mues et al. (1986); Zakeri et al. (1988); Milner and Campbell (1990); Lim and Brenner (1999)
Genes encoded by herpesviruses	Various	Var. cellular intron-containing genes	mRNA		Reviewed by Brunovskis and Kung (1996); Martin (1999)
Protamines	Vertebrates	Protamines			States et al. (1992); Jankowski et al. (1986); Moir and Dixon (1988); Oliva and Dixon (1989); Retief et al. (1993); Schlüter and Engel (1995)
Non-histone chromosomal protein HMG-1	Mammals	HMG-2; (4)	mRNA	$\geq$ mamm. radiation	Stros and Dixon (1993); Stros et al. (1995)
Glycerol kinase (GyK)[e]; testes; chr 4	Human Mouse	GyK;constitutive; (18); X-linked	mRNA		Sargent et al. (1994b) Pan et al. (1999)
Antioxidant protein 2 related seq. (Aop2-rs1 and Aop2-rs2[f]); var. tissues; chr 2 and chr 4, respectively	Mouse	Antioxidant protein 2 (Aop2); var. include. heart, liver, kidney; (4); chr 1	mRNA		Phelan et al. (1998)
DNA ligase IV (LIG4)	Mouse		mRNA		Barnes et al. (1998)
Actin-like-7A and actin-like-7B (*ACTL7A, ACTL7B*); chr. 9	Mammals	Actin or an actin-related protein (ARP)	mRNA		Chadwick et al. (1999)
Sterol 12α-hydroxylase (*CYP8B1*); 3p21.3 (Hsa), 9qF4 (Mmu)	Human Mouse	*CYP8A1* (or *CYP7A1/7B1*)	mRNA	$\geq$ mamm. radiation	Gåfvels et al. (1999)

Table 4. Continued

Metalloproteinase-distintegrins (ADAM20, ADAM21)	Human	Intron-containing family members	mRNA	Poindexter et al. (1999)
αCP-1 RNA binding protein	Mammals	αCP-2; (≥ 12)	mRNA	Makeyev et al. (1999)
Germ cell-specific actin capping protein α (Gsg3 clone); chr 6	Rodents	Somatic cell type actin capping protein α (ACPα)	mRNA	Yoshimura et al. (1999)
1-Cys peroxiredoxin, 1-Cys Prx (CP-2 and CP-5)	Mouse	CP-3; (4)		Lee et al. (1999)

Intron loss in comparison to an intron containing paralogue candidate is the only (remaining) hallmark. The decision, whether a sequence belongs in this table or Table 3 is arbitrary in some cases. Likewise, inclusion in Table 4 versus Tables 5 or 6 is somewhat arbitrary. In situations where only one intron is present in the putative founder gene, it may have been acquired in the founder. (e.g. protamine genes). Clearly, not all examples can be proven exaptations (potaptations according to Brosius & Gould, 1992, 1993).
[a]We cannot rule out that the ancestral histone gene was intronless and some histone genes acquired introns in the 'intron late' scenario.
[b]Many genes encoding G-protein coupled receptors that lack introns in their coding regions feature an intron in the 5′ UT; presumably generated by acquisition of splice sites (Brosius & Gould, 1992).
[c]Intronless HBF-2 is clustered with HBF-1 (one intron in coding region) on chromosome 14q 11-13 (Wiese et al., 1995).
[d]A 71 kDa heat shock protein has been described in the human genome that contains 8 introns (Dworniczak & Mirault, 1987).
[e]In addition to the split gene there are at least six additional loci in humans; two are pseudogenes (Xq and chr. 1) two are active retrogenes (both chr. 4) - protein protein product not confirmed yet; the status of the remaining two genes needs to be established.
[f]Aop2-rs2 potentially encodes only a truncated popypeptide of 114 aa.

retroposition) after the split. Even then, five million years do not provide enough time to generate much sequence divergence. I predict, that the essential difference between these closely related primate species is not the emergence of a certain number of novel genes. Instead, we will find that the main variance will be observed:

(i) in the recruitment of short domains of protein sequences from anonymous sequences (or retronuon-derived sequences) previously located in introns (Gilbert, 1978) or intergenic sequences and
(ii) most importantly, in the differential expression of shared genes.

These changes of when and where genes are being expressed with numerous consequences (including governing different developmental processes) will be found to be largely determined by migrant retronuons.

Tip of the iceberg as many regulatory elements could have been derived from ancient retronuons

Clearly, Tables 1 and 2 exhibit only the tip of the iceberg for a number of reasons – once the entire sequences of genomes of higher Eucarya will become available, many more examples will become evident. Although juxtaposition of retronuons with known and unknown genes will become apparent by sequence analysis, the question remains whether such retronuons do have an effect on the targeted gene. A global approach in the postgenomic era is needed to identify most retroelements that exert an effect on associated genes. There is yet another important reason why even then these tables would remain incomplete to a large degree: on an evolutionary scale, there is only a relatively small time window when retronuons as shown in Tables 1 and 2 are discernible. In other words, about 200×10^6 years from now the retronuon from Table 1 or 2 that carried an enhancer to an existing gene will have disintegrated. What remains, for example, are short conserved enhancer sequence motifs, perhaps after some modification of the enhancer core itself. It is clear that retroposition not only took place in the last 200×10^6 years but probably since the transition from the RNP to the DNA worlds (Darnell & Doolittle, 1986; Flavell, 1995; Brosius, 1999c; Poole Jeffares & Penny, 1999). Therefore, it would be of little surprise if it turned out that the vast majority of such small regulatory motifs that today are visible and active near targeted genes were derived from ancient retronuons whose identities disappeared over time (Brosius & Gould, 1992); W. Herr, cited in (Brosius & Tiedge, 1996), (Britten, 1996, 1997).

Messenger RNA derived retrosequences jump too. Did all intronless genes arise via an RNA intermediate?

Just as regulatory elements can insert next to resident genes, retroposition also allows for the opposite

Table 5. Intronless* vertebrate genes (no further evidence of retrosequence origin)

Gene	Species	Pattern of expression	References
Interferons	Vertebrates		Reviewed in: Nagata et al. (1980); Lawn et al. (1981); Watkins et al. (1991); Roberts et al. (1998)
Ribonucleases[a]	Various		Carsana et al. (1988); Hamann et al. (1990); Samuelson et al. (1991); Sasso et al. (1991); Tiffany et al. (1996)
Mos	Mammals		Watson et al. (1982); Newman and Dai (1996)
Insulinoma associated, IA-1 (zinc finger); chr 20	Human	Neuroendocrine tumours	Lan et al. (1994); Li et al. (1997)
Transcription elongation factor SII or A (TCEA)	Human		Park et al. (1994); DiMarco et al. (1996)
Modifier of Na^+-D-glucose co-transport (hRS1)	Human		Lambotte et al. (1996)
Profilaggrin	Mouse, rat		Haydock and Dale (1986)
Thrombomodulin	Mammals	RA inducible	Jackman et al. (1987); Niforas et al. (1993)
HS, HGT-C2, HGT-B2, BIIIB4, HGT-F, cKer1 keratins; keratin-associated protein, Krtap12-1; KAP6; B2E and B2F (high sulfur protein genes, in hair follicles); keratin-associated proteins pmg-1 and pmg-2	Vertebrates	Hair, skin	Powell and Rogers (1986); Kuczek and Rogers (1987); Frenkel et al. (1989); Whitbread et al. (1991); Fratini et al. (1993); Mitsui et al. (1998); Cole and Reeves (1998); Kuhn et al. (1999)
Blood platelet membrane glycoprotein lbα, glycoprotein V (GPV); glycoproteins lbβ, IX[b]	Mammals	Platelets	Wenger et al. (1988); Lanza et al. (1993); Ravanat et al. (1997); Yagi et al. (1995)
Olfactory marker protein (OMP)	Rat		Danciger et al. (1989)
Melanin-concentrating hormone	Fish		Takayama et al. (1989)
Cerebellar degeneration-related antigen, CDR34	Human		Chen et al. (1990)
Leukosialin CD43	Mammals		Cyster et al. (1990); Shelley et al. (1990)
Nuclear pore glycoprotein p62	Rat		D'Onofrio et al. (1991)
N-acetyltransferases Nat1 and Nat2[a]	Vertebrates		Grant et al. (1989); Blum et al. (1990a–c); Martell et al. (1991)
Centromere protein, CENP-B	Mammals		Sullivan and Glass (1991); Bejarano and Valdivia (1996)
JUN protooncogene	Vertebrates		Hattori et al. (1988); Hartl et al. (1991)
Factor VIII-associated gene (F8A)[d]	Mammals	Ubiquitous	Levinson et al. (1992)
A-kinase anchor protein, AKAP 75	Bovine		Hirsch et al. (1992)
LAP, C/EBPα, β_1, δ_1, CRP2 or NF-IL6β CAAT/enhancer-binding proteins; basic region-leucine zipper class (bZIP)	Human		Landschulz et al. (1988); Akira et al. (1990); Chang et al. (1990); Descombes et al. (1990); Cao et al. (1991); Williams et al. (1991); Kinoshita et al. (1992)
Cytochrome b5	Rabbit		Takematsu et al. (1992)
Gap junction genes connexin 31.1 and 30.3; chr 4	Mouse	Skin	Hennemann et al. (1992)
Na^+-MI contransporter (SMIT/SLC5A3)	Human	Kidney and other tissues	Berry et al. (1995); Porcellati et al. (1999)
Myeloid zinc finger gene (MZF-1)	Human	Bone marrow	Hui et al. (1995)
U2 auxiliary factor binding protein related sequence U2AFBPL[e] chr 5 / U2afbp-rs	Human Mouse		Pearsall et al. (1996)

Table 5. Continued

(imprinted in mouse); chr 11			
Acetyltransferases AT1 and AT2	Rat	Various	Land et al. (1996)
Choriolysin H (HCE)	Teleost fish		Yasumasu et al. (1996)
Pw1 zinc-finger protein			Relaix et al. (1996)
Defensin (HNP-1)	Human		Takemura et al. (1996)
Rho/Rac-like RhoG GTPase[f] (ARHG)	Human		Le Gallic and Fort (1997)
Ventral prostate protein C7orf1	Human		Peacock et al. (1997)
Antiproliferative proteins Tob, ANA	Mammals	Ubiquitous	Yoshida et al. (1997, 1998); Guéhenneux et al. (1987)
Serine-threonine kinase genes Tsk1, Tsk2	Mouse		Galili et al. (1997)
Glycosylphosphatidylinositol synthesis gene PIGC			Hong et al. (1997)
Growth arrest-specific C16orf3	Human		Whitmore et al. (1998)
Sex determining gene SRY and SOX-3	Mammals		Foster and Graves (1994); Tucker and Lundrigan (1995); O'Neill et al. (1998)
2′, 5′-oligoadenylate-dependent RNAse (interferon inducible)	Human		Tnani and Bayard (1998)
Necdin	Mammals	Neurons	Uetsuki et al. (1996); Nakada et al. (1998)
Chondroitin 6-sulfotransferase (C6ST); chr 11	Human		Mazany et al. (1998)
Testes-specific protein Y-encoded-like, (human TSPYL chr 6; rodent Tspyl chr 10)	Mammals	Ubiquitous	Vogel et al. (1998)
Citrate synthase (CS); chr 12	Human		Goldenthal et al. (1998)
CXorf1	Human	Hippocampus	Redolfi et al. (1998)
Cholesterol 25-hydroxylase	Human, mouse		Lund et al. (1998)
Prion protein[g]	Mammals, chicken		Lee et al. (1998)
Insulin receptor substrate 4 (IRS-4)[h]	Mouse		Fantin et al. (1999)
110 kDa high molecular wt. basic nuclear protein (HMrBNP)	Flounder	Sperm	Watson and Davies (1999)
cded/lior	Mouse		Mishra et al. (1999)
Rab-like protein (Rlp-2)	Human		Peng et al. (1999)
Slow-kinetics immediate early gene ler5	Mouse		Williams et al. (1999)
transport modifier RS1	Rabbit		Reinhardt et al. (1999)
Sperizin, RING zinc-finger protein	Mouse	Haploid sperm cells	Fujii et al. (1999)
Malaria-inducible gene	Mouse	Spleen	Krücken et al. (1999)
KRML (MAFB); chr 20	Human	Hemapoietic tissue	Wang et al. (1999)
α-endosulfine (ENSA); chr 14	Human		Heron et al. (1999)
ZNF127 RING zinc-finger protein[i]	Human		Jong et al. (1999)
MAGEL2[i]	Human, mouse	Brain, placenta	Boccaccio et al. (1999)

*May have intron(s) upstream from coding region; a single intron in the coding region may also have been generated subsequent to retroposition (e.g. the monocyte-specific Dif-2 gene (Pietzsch et al., 1998) or the acidic 80 kDa protein kinase C substrate/MARCKS (Erusalimsky et al., 1991; Blackshear et al., 1992).

[a]Several genes contain single introns in 5′ UT. [b]No intron in entire ORF, but 5′ UT. [c]Coding region contains a tandem hexapeptide repetitive structure; could have been exapted from a non-coding repeat region. [d]One of the human genes may be located in intron of factor VIII gene. [e]One of the human genes located on the X chromosome contains introns (Kitagawa et al., 1995). [f]Contains a large exon in the 5′ UT. [g]Mouse, sheep have two and humans one exon(s) in 5′ UT. [h]Related gene IRS-3 contains one intron in the coding region. IRS-1 and IRS-2 are thought to be intronless as well. However, Vassen et al. (1999) described an intron at the C-terminus of the IRS-2 ORF. [i]Located on chr 15q11-13 (Prader-Willi Syndrome, PWS, region); paternally expressed.

scenario: retrogenes, usually generated by reverse transcription of mRNAs, can insert next to resident promoter/enhancer elements (Brosius, 1991) and thus escape transcriptional silencing (the fate of most retrosequences). The retrogene is often expressed in a cell type and/or at a stage different from the founder gene. In many cases a retrogene, when different from its founder gene by modifications, can be recruited or exapted into a novel function (Table 3). In fact, probably the majority of intronless genes may have arisen by this mechanism although most of their hallmarks may not be recognizable anymore (Tables 4, 5). This would be certainly true if the 'intron-early' hypothesis was correct. Should the 'intron-late' camp be correct, at least some of the intronless genes may be ancient and have remained intronless from 1.7 to 1.0×10^9 years ago, the presumed period when spliceosomal introns arose (Cavalier-Smith, 1991), until now. Apart from the scenario that a retrogene is immediately active upon integration, there are scenarios where the retrogene is initially inactive and is being activated at a later time, for example, upon activation of flanking regulatory elements by mutation or acquisition of a retronuon.

Inactive retrogenes:future genes in the making?

It is conceivable that many retrogenes were reactivated after a considerable amount of transcriptional and/or translational latency. This time of inactivity facilitated more drastic changes in gene structure without the possible 'danger' of expressing intermediate molecules that had a large degree of similarity to the original ones but exerted detrimental effects as is the case in dominant negative mutations. As early as 1972, Arthur Koch (1972) had recognized the importance of untranslatable intermediates in enzyme evolution (for a more detailed discussion, see Brosius, 1999a, b).

Virtually any retrogene listed in Table 6 may be in such a 'stand-by' mode. One may say that even the majority of retrogenes, so-called inactive retropseudogenes, their ORFs annihilated with indels and mutations that introduce many stop codons in the coding region may be 'on call' (potaptation = potential exaptation) for future utility (exaptation) (Gould & Vrba, 1982). In summary, the relentless activity of retronuons in most vertebrate genomes had and will continue to have important consequences for the evolution of populations and species (Rose & Doolittle, 1983, 1989; McDonald, 1990, 1995).

Keeping retroposition in check

Countermeasures of the host

Despite potential advantages of retroposition on future evolvability, if a host 'permits' unrestricted retroposition, 'runaway' activity may result. Left without restraint, this may lead to disadvantages such as frequent inactivation or misregulation of essential genes and overinflation of genome size. It has been suggested, therefore, that the host is devising countermeasures to keep retroposition in check. For example, Patience wrote: '. . . animals harboring infectious ERVs have evolved mechanisms to prevent re-infection with their own potentially infectious ERV genomes' (Patience et al., 1997). This is perhaps an analogy to the arms race between host defense and external parasites (microorganisms/viruses). This analogy also includes gradual reduction of the parasite's virulence in order to ensure its own survival. It has been proposed that a widespread defensive measure of the host against retronuons is methylation of retroelements in order to shut down their transcription (Walsh et al., 1998; Wolffe & Matzke, 1999). While such measures may be effective against retroelements such as proviruses or intact LINEs, the methylation of truncated LINEs or SINEs such as Alu elements seems futile, since, for example, very few of the 10^6 Alu repeats are transcriptionally and hence retropositionally active. Likewise, promotion of mutations at methylated CpG sites is of little consequence for the propagation of Alu elements. In the case of SINEs, the only effective countermeasure would be silencing of the few active master genes. This, in turn, may be impossible if the RNA product of the master gene is under selective pressure. As a more effective measure, the activity of reverse transcription could be blocked provided there are no negative consequences for the host (see below). As a result, production of retroelements and retrosequences alike would be stopped, the only way 'out' for a parasitic retroelement being a change of host by horizontal transfer (Jordan et al., 1999).

Parasites or symbionts

Despite the realization that occasionally a TE is advantageous to the host, many authors consider transposable elements (TEs) including retronuons as genomic parasites (Doolittle & Sapienza, 1980; Orgel & Crick, 1980; Hickey, 1982; Kidwell & Lisch, 1997; Walsh, Chaillet & Bestor, 1998; Jordan, Matyunina

Table 6. Intronless vertebrate genes likely of retroposition origin – no proven activity of gene product (this does not exclude transcriptional or even translational activity)

Retrogene; pattern of expression; chromosome	Species	Source gene; pattern of expression; (# of introns); chromosome	Age of retrogene	References
Non-muscle tropomyosin (hTM_{NM}-1)	Human			MacLeod et al. (1983)
Metallothionein (MT-1Ψb)[a]	Rat			Andersen et al. (1986)
Sarcomeric actin α2	Frog	Actin		Stutz and Spohr (1987)
Glutamine synthetase (GSr)	Mouse	Glutamine synthetase (GSi)		Bhandari et al. (1991)
Glutamine synthetase (ΨGS)[b]	Human	Glutamine synthetase		Chakrabarti et al. (1995)
Heat stable antigen (2 ORFs)	Mouse	Heat stable antigen		Wenger et al. (1991)
Adenylate kinase 3 (AK3)[c]; chr 17	Human	AK3; chr 9		Xu et al. (1992)
Ferritin L subunit Lg	Mouse	ferritin L subunit; (3)		Renaudie et al. (1992)
Processed CD-MPR gene[d]; chr 3	Mouse	Cation-dependent mannose 6-phosphate receptor (CD-MPR); (7); chr 6		Ludwig et al. (1992)
Id2B[e]	Human	Helix-loop-helix protein Id2		Kurabayashi et al. (1993)
Casein kinase IIα	Human	Casein kinase IIα		Devilat and Carvallo (1993)
$\Psi EF1_A\#1$[f]	Bovine	CCAAT transcription factor subunit $EF1_A$		Ozer et al. (1993)
Ψ 5HT1D[g]	Human	serotonin receptor 5HT1D		Bard et a. (1995)
Protein kinase C ($\Psi PKC\zeta$[h])	Rat	PKCζ		Andrea and Walsh (1995)
FAU1P[i]; chr 18	Human	FAU1		Kas et al. (1995)
dbpB pseudogene[j]	Human	DNA binding protein dbpB		Kudo et al. (1995)
mif rp-1	Mouse	macrophage migration inhibitory factor (MIF)		Bozza et al. (1995)
Ferritin H subunit pseudogene	Human	Ferritin H subunit		Zheng et al. (1995, 1997)
Prothymosin α intronless	Mammals	prothymosin α		Varghese and Kronenberg (1991); Manrow et al. (1992); Rubtsov and Vartapetyan (1995)
Laminin receptor (37LRP/p40), intronless	Human	laminin receptor (37LRP/p40)		Jackers et al. (1996)
LAMRL5[k]	Human	67-kDA laminin receptor (LAMR1)		Richardson et al. (1998)
Ubiquitin-conjugating enzyme UBE2L1; chr 14	Human	ubiquitin-conjugating enzyme UBE2L3; (1); chr 22		Moynihan et al. (1996)
ΨAdh-2[l]	Mouse	Class III alcohol dehydrogenase (Adh-2); (8)		Foglio and Duester (1996)
MSSP-1 (transcriptional enhancer of c-myc)	Human	MSSP-2; (15)		Haigermoser et al. (1996)
Olfactory receptor pseudogenes	Human	Olfactory receptor genes		Crowe et al. (1996)
Hp53int1[m]	Human			Reisman et al. (1996)
Phosphoglycerate mutase brain isoform pseudogene ($\Psi PGAM1$[n])	Human	Phosphoglycerate mutase brain isoform (*PGAM1*)		Dierick et al. (1997)
α tubulin-related sequence; chr 11[o]	Human	Keratinocyte α tubulin		Devon et al. (1997)
r.pem2 homeobox gene[p]; epipydimis; X- linked	Rat	r.Pem, orphan homeobox gene; testes, ovary, placenta, epididymis; (5); chr 4		Nhim et al. (1997)
Leukocyte antigen C1pg-26[q]	Dog	Leukocyte antigen DLA class I; (7)		Burnett et al. (1997)

Table 6. Continued

Tdgf1-ps1; chr 16	Mouse	Teratocarcinoma-derived growth factor- 1; Tdgf1; (5); chr 9		Liguori et al. (1996, 1997)
TDGF3; X chr	Human	TDGF1; chr 3		Dono et al. (1991)
(FABP3-ps); chr 13	Human	Fatty acid binding protein FABP3, chr 1		Prinsen et al. (1997)
Serotonin-7 receptor (5-HT7Ψ)[r]	Human	5-HT7		Quian et al. (1998)
Mannose-binding protein-A; chr 10	Human			Guo et al. (1998)
ΨFGFR-3 (partial, antisense); fetal development; chr 1	Mouse	Fibroblast growth factor receptor (FGFR-3); chr 5		Weil et al. (1997)
Ψ ribosomal protein L7 (antisense)	Human	Ribosomal prot. L7		Hohlbaum et al. (1998)
Supt4h2; chr 10	Mouse	Supt4h; (4); chr 11		Chiang et al. (1998)
Ubiquitin conjugating E2 enzyme ubc9-psi1 and ubc-9-psi2	Mouse	ubc-9		Tsytsykova et al. (1998)
SMT3A and 3 SMT3B proc. pseudogenes	Mouse	Ubiquitin-like proteins		Chen et al. (1998)
ΨPTEN[s], chr 9	Human	PTEN/MMAC1/TEP1 phosphatase; chr 10		Dahia et al. (1998)
EIF4E2 translational initiation factor	Human	EIF4E1; (6)	Recent	Gao et al. (1998)
$EIF2_{\gamma}A$; testes; chr 12	Human	Euk. Translation initiation factor elF-2_{γ} ($EIF2_{\gamma}X$); X-linked		Ehrmann et al. (1998)
Ψ$hGABP_{\alpha t}$	Human	ets related $GAPB_{\alpha}$		Luo et al. (1999)
Proto-oncogene hPTTG2	Human	hPTTG1		Prezant et al. (1999)
CDC42-like; chr 4	Human	CDC42; chr 1		Nicole et al. (1999)
Spondyloepiphyseal dysplasia tarda gene (SEDLP)5; many tissues; chr 19	Human	Spondyloepiphyseal dysplasia tarda gene (SEDL or GPM6B); many tissues; (3); X-linked	Recent	Gedeon et al. (1999)
CK2α; chr 11	Human	CK2α; (12); chr 11		Wirkner and Pyerin (1999)

There are probably numerous additional retrogenes whose ORFs are not severely compromised or could yield a truncated polypeptide, partially in a different reading frame (e.g. Chen et al., 1982; Varshney & Gedamu, 1984; Dudov & Perry, 1984; Nojima et al., 1987; Kuzumaki et al., 1987; Srikantha et al., 1987; Seelan & Padmanaban, 1988; Nielsen & Trachsel, 1988; Kawaichi et al., 1992; Jun et al., 1997; Palmer et al., 1998). However, transcription and/or translation are not documented.

[a]This retrosequence is transcribed. Due to an insertion after codon 28 the ORF is shifted to with a different hypothetical C-terminus of an additional 35 amino acids instead of 33 aa in the correct MT-1 frame.

[b]This retrosequence is transcribed. The ORF is truncated but retains $\sim 2/3$ of the coding sequence; probably no protein product.

[c]Embedded in intron 10 of NF1 gene (located on human chr 17).

[d]This retrosequence is transcribed. The ORF is truncated after 141 codons (out of 278 possible) in murine CD-MPR. A soluble truncated form of CD-MPR encoding only the *n*-terminal extracytoplasmic region including codon 154 was functional in ligand binding and acid-dependent dissociation (Marron-Terada et al., 1998).

[e]This retrosequence is transcribed. Stop codon at aa 37, however.

[f]This retrosequence is transcribed. The ORF is truncated; probably no protein product.

[g]This retrosequence hypothetically encodes a 140 aa polypeptide most of which (aa 31–140) are similar to bovine EFI_A (324 aa total).

[h]This retrosequence is transcribed specifically in the brain. The ORF is truncated and no protein product could be identified by Western blots.

[i]This retrosequence is not transcribed but contains an intact ORF.

[j]One of 16 pseudogenes contains an intact ORF.

[k]Potentially active retrogenes may also exist in the mouse (Bignon et al., 1991).

[l]25 point mutations relative to Adh-2 cDNA, nevertheless ORF is intact, but no evidence for transcription, thus far.

[m]Located in the 10 kb first intron of p53 tumour supressor gene; no or short ORF.

[n]Located in intron 1 of Menckes disease gene (*ATP7A*, *MNK*).

[o]The ORF is truncated but retains 80% of the coding sequence.

[p]Although a processed mRNA was the founder of this retrogene, it acquired new splice sites that remove three premature stop codons yielding again an open reading frame – protein product not confirmed yet.

[q]Contains single ORF of 332 codons but no potential start codon in the *N*-terminal 2/3 of ORF; not likely to be functional.

[r]Transcribed but not translatable.

[s]ORF intact, hypothetical polypeptide somewhat smaller due to loss of first start codon; no evidence for transcription as of yet.

[t]This retrogene is transcribed in human myeloid cells, but a mutation at the site that corresponds to the ATG start methionine codon may prevent its translation.

Table 7. Intron containing vertebrate genes featuring large exons (probably of retrosequence origin)

Gene	Species	Pattern of expression	References
Developmentally regulated type X collagen	Chicken		Ninomiya et al. (1986)
C1r and *C1s* complement	Human		Tosi et al. (1989)
Follicle-stimulating hormone receptor (FSHR); LH, TSH receptors	Human		Gromoll et al. (1996); Misrahi et al. (1996)
Islet homeobox gene (isl 1)[a]	Mammals		Bozzi et al. (1996)

[a]Features intronless homeobox domain.

& McDonald, 1999). Personally, I prefer the term genomic symbionts for a number of reasons: The translation of 'symbiosis' from Greek (living together) is at least as close to TE reality as is 'parasitiasis' (living at one's cost). G.P. Georgiev (1998) who was among the first to realize the biological significance of mobile genomic elements considered them not as 'pure genomic parasites' but as 'genomic symbionts.' He continues, 'Their importance is in spreading signal sequences of different type throughout the genome to supply different genetic elements with such signal sequences. . . . One may speculate that mobile elements are, at least sometimes, involved in regulation of certain gene expression, RNA splicing, replication initiation etc.... the existence of mobile elements sharply increases the variability and, in this way, accelerates the process of evolution. Mobile elements can make organisms fit more readily to changing environmental conditions. Therefore, the advantages gained from their presence may outweigh the expenses spent for their replication and expression' (Georgiev, 1984). It is remarkable that Georgiev had this insight one and a half decades prior to gradual acceptance of such concepts within the scientific community.

As mentioned earlier, the truth hovers most of the time in between extremes. Who would place smallpox, measles, influenza, typhus, bubonic plague and other infectious disease causing microbes not into the category of human parasites? Yet, these infectious agents that had 'plagued' European populations for centuries (implementation of host-infective agent strategies reduced the virulence of some over time) gave initially a handful of 'conquistadores' a selective advantage over the natives of the Americas by killing as much as 95% of the population (Diamond, 1997). Therefore, it is not so straightforward anymore whether germs belong to the parasitic corner, when the evolution of populations is considered. Likewise, it is conceivable that host interactions have blunted the negative impacts of TEs including retronuons (Kidwell & Lisch, 1997). The impact of retronuons on individuals, populations or species is to be considered separately (Doolittle, 1989). What may be a parasite for an individual, may be a symbiont for a population or species.

Genome expansion of retronuons lessens their negative impacts

Apart from the aforementioned reduction of transpositional activity by methylation and other measures (Kidwell & Lisch, 1997) genome expansion is an effective way to reduce negative impacts of retronuon insertion. By blowing up the size of introns and intergenic regions, the chances that a retronuon hits a 'vulnerable' target (exon, promoter, or other regulatory element) is greatly reduced. This means that yeast and *Drosophila*, for example, are more assailable to TE activities with negative consequences as are plants (such as maize) or mammals. It will be interesting, for example, to compare retroposition activity in species with tightly packed genomes such as the pufferfish (*Fugu rubripes*) with similar, more bloated, sequences. The *Fugu* genome comprising ∼400 Megabases is about four-fold compressed in comparison to fish with larger introns and intergenic regions (zebrafish, *Danio rerio*, ∼1700 Mb) or even about eight-fold in comparison to mammals (∼3000 Mb). Interestingly, retronuons themselves generate this genome inflation and thus blunt their own 'virulence' in analogy to strategies of germs that reduce their virulence as mentioned above.

One may draw the following analogy: In a long-drawn trench war enemy artillerists (retroelements) infiltrate the opponent's field fortifications (host genome) and shoot cannons straight into the air (retroposition). Initially, the carnage inflicted by the projectiles is immense. Let's assume, however, that the remaining resident soldiers (genes, gene segments, regulatory

elements) are still viable and may move away from the craters (introns/intergenic regions) generated by the projectiles and thus geographically expand the fortifications (genome inflation). The next rounds of artillery will be less damaging until the soldiers are so widely separated that future rounds of artillery inflict very little damage. With time, for example, the craters fill with rainwater and serve as reservoirs for the maintenance of the resident forces (exaptation). What happens to the infiltrators? Some will also be hit by their own projectiles; they may multiply, but also may be killed captured, tied, by resident forces (host defense) and thus reduced in their effectiveness to operate the cannons. The resident forces may even tolerate them since they do not inflict much damage anymore, but their activities turn out to be useful in one way or the other (recruitment/exaptation). In such a constellation, the resident forces may have an advantage over the opposing forces and win the war (selective advantage) although nobody would have thought and planned so, at the initial stages of battle.

Of course, an opposite strategy that may be equally effective in banning retronuons from taking a foothold in genomes is tightening genomes. Organisms with very few and small introns and very little space between genes/operons can select efficiently against the spread of retronuons, and a potential parasite or symbiont can only persist when a certain 'consideration' develops towards the host. It is interesting to note that the expansive strategy is mostly associated with complex multicellular organisms with long generation times, while the contractive strategy is common in relatively simple organisms with very short generation times.

Does reverse transcriptase have a functional role in a modern cell, or do retroelements keep the gene alive?

A prerequisite for a full comprehension of retroposition is understanding the evolution and function of the enzyme reverse transcriptase in the cell and during biogenesis of retroelements (Becker, 1996; Doolittle & Feng, 1992). A fundamental question remains as to why reverse transcriptase still is present in so many lineages? Reverse transcriptase, the key enzyme for the transition from RNA to DNA as genetic material may have arisen from an RNA replicase during the final stages of the RNP world (Darnell & Doolittle, 1986; Flavell, 1995; Brosius, 1999c). It is interesting to note that a process that had its heyday several billion years ago during transition from the RNP to the DNA/RNA/protein world is still very active today doing the same as before – converting RNA into genomic DNA – as if the process had a beginning but in certain lineages no end (Jurka, 1998). Rather than a noise from bygone eras, reverse transcription has a dramatic effect on genomes of certain lineages even in quite recent times. If the situation of the almost completed sequence of human chromosome 22 is representative for the entire human genome, almost 40% of the DNA corresponds to retronuons (Dunham et al., 1999). These 40% are merely the retronuons that we can detect with our current biomathematical analyses of genomes. It is quite conceivable, that almost all of the remaining 55–57% of the genome that does not encode exons is derived from TEs, mainly retronuons. Apart from its historical role in converting RNA-based to DNA-based genomes and playing an instrumental role in gigantic genome expansions in certain lineages, an additional function of reverse transcriptase in modern cells remains obscure. In 1985, David Baltimore summarized it as follows: "... there has been no experimental demonstration that reverse transcription has an obligatory role in normal cellular physiology" (Baltimore, 1985). This is still true one and a half decades later, with the noted exception of a specialized reverse transcriptase activity in the maintenance of telomeres (Blackburn, 1992). It is unlikely, however, that the enzymes involved play a role during generation of other retronuons (but see (Pardue et al., 1996) where several retroelements were exapted as telomeres in *Drosophila*). Solving the question whether a cellular reverse transcriptase is involved in some vital cellular activity will be extremely challenging. Numerous copies of incapacitated and especially intact and active genes encoding reverse transcriptase provided by retroelements such as proviruses and LINE elements virtually exclude an experimental approach at this juncture. It is accepted that retroelements are very ancient; some authors state that "it seems plausible that retroelements were present at the genesis of living systems" (Flavell, 1995). Others "propose that the occasion of the juxtaposition (of a reverse transcriptase and a prokaryotic transposase) coincided with the invasion of primitive eukaryotes by endosymbiotic prokaryotes 1.5–2.0 billion years ago" (Doolittle & Feng, 1992) and "... the LINEs family of transposable elements is phylogenetically more primitive than the LTR-containing elements as reflected in their

biological distribution, the latter not yet having been identified in any protists. . . we can place the origin of the LINEs family after the introduction of organelles into eukaryotes and anticipate that they will not be found in those primitive eukaryotes like giardia. . . " (Doolittle & Feng, 1992). No matter when exactly retroelements arose, it is even plausible that retroelements alone kept reverse transcriptase alive by evading genomic purges (Jordan, Matyunina & McDonald, 1999) and even re-entering a lineage when countermeasures subside. The alternative is still a currently evasive intrinsic function in the cell, where reverse transcriptase is likely to be encoded by a cellular non-retroelement-derived gene. In the former case, retroposition is forced upon the cell by persistent albeit migrant retroelements, nevertheless conveying long-term benefits to a population or species. In the latter case, it is a by-product of an intrinsic function of the cell. Completely sequenced genomes of model organisms may help identify endogenous reverse transcriptase not derived from retroelements by biomathematical analysis. Experimentally, the feat of removing all retroelement-derived functional copies of the enzyme and then testing for remaining retropositional activity and/or loss of function will remain unattainable for some time, owing to their profusion in mammalian genomes.

Why is there so much extra DNA in many genomes even though its future use is not foreseeable?

Evolution cannot plan for the future: "Evolution has no foresight, and a genetic element cannot be selected because it might someday be of help. Once it is there, however, whatever the reason for its presence, such a structure might prove 'useful' and then become the target of some selective pressure on the host phenotype" (Jacob, 1982). Hence, it is difficult to understand why features that may have a potential use in the future convey a selective advantage to the host – especially if one considers that these elements may have serious negative impacts. An insightful and timely treatment of these issues can be read in an article by Doolittle (1989). The author concludes as follows: "Why do so many more species have repetitive sequences, or introns, or excess DNA in general, than we would otherwise, on the basis of our adaptationist understanding of molecular biology, have expected? I submit that an answer lies in the statement that species that have these things more frequently speciate than those which do not, regardless whether one regards this as a statement about selection as a process or selection as a force. Equally, one might argue that clades whose species retain relatively many of these genomic components less frequently go extinct for failure to adapt to environmental change than do clades whose species do not, even though there is no meaningful sense in which individual organisms within species of the first sort of clade are 'better adapted' at the organismal level than are individual organisms within species of clades of the second sort." In other words, TEs including retronuons are found in many extant species because many species that purged themselves of TEs for some short- to mid-term benefit (not being able to 'know' about their future exaptive use) are now extinct.

While the dream of solving many questions concerning the history and functional significance of retroposition will largely come true due to the advent of comparative (total) genomics, some questions may remain elusive forever, as intact DNA from long extinct lineages will likely be lost for all time.

Acknowledgements

I would like to thank Henri Tiedge and Jean-Marc Deragon for comments on an earlier version of the manuscript. For updated versions of tables please refer to: http://www-ifi.uni-muenster.de/exapted-retrogenes/tables.html

Appendix

Table references

Adra, C.N., N.A. Ellis & M.W. McBurney, 1988. The family of mouse phosphoglycerate kinase genes and pseudogenes. Somat. Cell. Mol. Genet. 14: 69–81.

Akira, S., H. Isshiki, T. Sugita, O. Tanabe, S. Kinoshita, Y. Nishio, T. Nakajima, T. Hirano & T. Kishimoto, 1990. A nuclear factor for *IL-6* expression (NF-IL6) is a member of a C/EBP family. Embo J. 9: 1897–1906.

Almenoff, J.S., J. Jurka & G.K. Schoolnik, 1994. Induction of heat-stable enterotoxin receptor activity by a human Alu repeat. J. Biol. Chem. 269: 16610–16617.

Alvarez-Bolado, G., M.G. Rosenfeld & L.W. Swanson, 1995. Model of forebrain regionalization based on spatiotemporal patterns of POU-III homeobox gene expression, birthdates, and morphological features. J. Comp. Neurol. 355: 237–295.

Andersen, R.D., B.W. Birren, S.J. Taplitz & H.R. Herschman, 1986. Rat metallothionein-1 structural gene and three pseudogenes, one of which contains 5′-regulatory sequences. Mol. Cell. Biol. 6: 302–314.

Andrea, J.E. & M.P. Walsh, 1995. Identification of a brain-specific protein kinase C zeta pseudogene (psi PKC zeta) transcript. Biochem. J. 310: 835–843.

Aronow, B.J., R.N. Silbiger, M.R. Dusing, J.L. Stock, K.L. Yager, S.S. Potter, J.J. Hutton & D.A. Wiginton, 1992. Functional analysis of the human adenosine deaminase gene thymic regulatory region and its ability to generate position-independent transgene expression. Mol. Cell. Biol. 12: 4170–4185.

Arranz, V., M. Kress & M. Ernoult-Lange, 1994. The gene encoding the MOK-2 zinc-finger protein: characterization of its promoter and negative regulation by mouse *Alu* type-2 repetitive elements. Gene 149: 293–298.

Ashworth, A., B. Skene, S. Swift & B.R. Lovell, 1990. Zfa is an expressed retroposon derived from an alternative transcript of the Zfx gene. Embo J. 9: 1529–1534.

Atanasoski, S., S.S. Toldo, U. Malipiero, E. Schreiber, R. Fries & A. Fontana, 1995. Isolation of the human genomic brain-2/N-Oct 3 gene (POUF3) and assignment to chromosome 6q16. Genomics 26: 272–280.

Baniahmad, A., M. Muller, C. Steiner & R. Renkawitz, 1987. Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elements. Embo J. 6: 2297–2303.

Banki, K., D. Halladay & A. Perl, 1994. Cloning and expression of the human gene for transaldolase. A novel highly repetitive element constitutes an integral part of the coding sequence. J. Biol. Chem. 269: 2847–2851.

Banville, D. & Y. Boie, 1989. Retroviral long terminal repeat is the promoter of the gene encoding the tumor-associated calcium-binding protein oncomodulin in the rat. J. Mol. Biol. 207: 481–490.

Bard, J.A., S.P. Nawoschik, B.F. O'Dowd, S.R. George, T.A. Branchek & R.L. Weinshank, 1995. The human serotonin 5-hydroxytryptamine 1D receptor pseudogene is transcribed. Gene 153: 295–296.

Barnes, D.E., G. Stamp, I. Rosewell, A. Denzel & T. Lindahl, 1998. Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. Curr Biol. 8: 1395–1398.

Baumruker, T., C. Gehe & I. Horak, 1988. Insertion of a retrotransposon within the 3′ end of a mouse gene provides a new functional polyadenylation signal. Nucleic Acids Res. 16: 7241–7251.

Bejarano, L.A. & M.M. Valdivia, 1996. Molecular cloning of an intronless gene for the hamster centromere antigen CENP-B. Biochim Biophys Acta. 1307: 21–25.

Berchtold, M.W., M. Koller, R. Egli, J.A. Rhyner, H. Hameister & E.E. Strehler, 1993. Localization of the intronless gene coding for calmodulin-like protein CLP to human chromosome 10p 13-ter. Hum Genet. 90: 496–500.

Berquin, I.M., M. Ahram & B.F. Sloane, 1997. Exon 2 of human cathepsin B derives from an *Alu* element. FEBS Lett. 419: 121–123.

Berry, G.T., J.J. Mallee, H.M. Kwon, J.S. Rim, W.R. Mulla, M. Muenke & N.B. Spinner, 1995. The human osmoregulatory Na+/myo-inositol cotransporter gene (SLC5A3): molecular cloning and localization to chromosome 21. Genomics 25: 507–513.

Bhandari, B., W.J. Roesler, K.D. DeLisio, D.J. Klemm, N.S. Ross & R.E. Miller, 1991. A functional promoter flanks an intronless glutamine synthetase gene. J. Biol. Chem. 266: 7784–7792.

Bi, S., O. Gavrilova, D.W. Gong, M.M. Mason & M. Reitman, 1997. Identification of a placental enhancer for the human leptin gene. J. Biol. Chem. 272: 30583–30588.

Bignon, C., M. Roux-Dosseto, M.E. Zeigler, M.S. Wicha & P.M. Martin, 1992. cDNA cloning and genomic analysis of a new multigene family sharing common phylogenetic and expression profiles with the laminin receptor gene. Biochem. Biophys. Res. Commun. 184: 1165–1172.

Blackshear, P.J., J.S. Tuttle, R.J. Oakey, M.F. Seldin, M. Chery, C. Philippe & D.J. Stumpo, 1992. Chromosomal mapping of the human (MACS) and mouse (Macs) genes encoding the MARCKS protein. Genomics 14: 168–174.

Blum, M., D.M. Grant, W. McBride, M. Heim & U.A. Meyer, 1990a. Human arylamine N-acetyltransferase genes: isolation, chromosomal localization, and functional expression. DNA Cell. Biol. 9: 193–203.

Blum, M., M. Heim & U.A. Meyer, 1990b. Nucleotide sequence of rabbit NAT1 encoding monomorphic arylamine N-acetyltransferase. Nucleic Acids Res. 18: 5287.

Blum, M., M. Heim, & U.A. Meyer, 1990c. Nucleotide sequence of rabbit NAT2 encoding polymorphic liver arylamine N-acetyltransferase (NAT). Nucleic Acids Res. 18: 5295.

Boccaccio, I., H. Glatt-Deeley, F. Watrin, N. Ro# ckel, M. Lalande & F. Muscatelli, 1999. The human *MAGEL2* gene and its mouse homologue are paternally expressed and mapped to the Prader-Willi region. Hum. Mol. Genet. 8: 2497–2505.

Boer, P.H., C.N. Adra, Y.F. Lau & M.W. McBurney, 1987. The testis-specific phosphoglycerate kinase gene pgk-2 is a recruited retroposon. Mol. Cell. Biol. 7: 3107–3112.

Boggaram, V., K. Qing & C.R. Mendelson, 1988. The major apoprotein of rabbit pulmonary surfactant. Elucidation of primary sequence and cyclic AMP and developmental regulation. J. Biol. Chem. 263: 2939–2947.

Bozza, M., L.F. Kolakowski, Jr., N.A. Jenkins, D.J. Gilbert, N.G. Copeland, J.R. David & C. Gerard, 1995. Structural characterization and chromosomal location of the mouse macrophage migration inhibitory factor gene and pseudogenes. Genomics 27: 412–419.

Brini, A.T., G.M. Lee & J.P. Kinet, 1993. Involvement of Alu sequences in the cell-specific regulation of transcription of the gamma chain of Fc and T cell receptors. J. Biol. Chem. 268: 1355–1361.

Brosius, J., 1999. Many G-protein-coupled receptors are encoded by retrogenes. Trends Genet. 15: 304–305.

Brunovskis, P. & H.J. Kung, 1996. Retrotransposition and herpesvirus evolution. Virus Genes. 11: 259–270.

Cao, Z., R.M. Umek & S.L. McKnight, 1991. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev. 5: 1538–1552.

Carsana, A., E. Confalone, M. Palmieri, M. Libonati & A. Furia, 1988. Structure of the bovine pancreatic ribonuclease gene: the unique intervening sequence in the 5′ untranslated region contains a promoter- like element. Nucleic Acids Res. 16: 5491–5502.

Chadwick, B.P., J. Mull, L.A. Helbling, S. Gill, M. Leyne, C.M. Robbins, H.W. Pinkett, I. Makalowska, C. Maayan, A. Blumenfeld, F.B. Axelrod, M. Brownstein, J.F. Gusella & S.A. Slaugenhaupt, 1999. Cloning, mapping, and expression of two novel actin genes, actin-like- 7A (ACTL7A) and actin-like-7B (ACTL7B), from the familial dysautonomia candidate region on 9q31. Genomics 58: 302–309.

Chakrabarti, R., R.B. McCracken, Jr., D. Chakrabarti & W.W. Souba, 1995. Detection of a functional promoter/enhancer in an intron-less human gene encoding a glutamine synthetase-like enzyme. Gene 153: 163–199.

Chang, C.J., T.T. Chen, H.Y. Lei, D.S. Chen & S.C. Lee, 1990. Molecular cloning of a transcription factor, AGP/EBP, that be-

longs to members of the C/EBP family. Mol. Cell. Biol. 10: 6642–6653.
Chang- Yeh, A., D.E. Mold & R.C. Huang, 1991. Identification of a novel murine IAP-promoted placenta-expressed gene. Nucl. Acids Res. 19: 3667–3672.
Chen, A., H. Mannen & S.S. Li, 1998. Characterization of mouse ubiquitin-like SMT3A and SMT3B cDNAs and gene/pseudogenes. Biochem. Mol. Biol. Int. 46: 1161–1174.
Chen, M.J., T. Shimada, A.D. Moulton, M. Harrison & A.W. Nienhuis, 1982. Intronless human dihydrofolate reductase genes are derived from processed RNA molecules. Proc. Natl. Acad. Sci. USA 79: 7435–7439.
Chen, Y.T., W.J. Rettig, A.K. Yenamandra, C.A. Kozak, R.S. Chaganti, J.B. Posner & L.J. Old, 1990. Cerebellar degeneration-related antigen: a highly conserved neuroectodermal marker mapped to chromosomes X in human and mouse. Proc. Natl. Acad. Sci. USA 87: 3077–3081.
Chiang, P.W., R. Zhang, L. Stubbs, L. Zhang, L. Zhu & D.M. Kurnit, 1998. Comparison of murine Supt4h and a nearly identical expressed, processed gene: evidence of sequence conservation through gene conversion extending into the untranslated regions. Nucleic Acids Res. 26: 4960–4964.
Cho, K.O., B. Minsk & J.A. Wagner, 1990. NICER elements: a family of nerve growth factor-inducible cAMP- extinguishable retrovirus-like elements. Proc. Natl. Acad. Sci. USA 87: 3778–3782.
Clemens, M.J. 1987. A potential role for RNA transcribed from B2 repeats in the regulation of mRNA stability. Cell 49: 157–158.
Cole, S.E. & R.H. Reeves, 1998. A cluster of keratin-associated proteins on mouse chromosome 10 in the region of conserved linkage with human chromosome 21. Genomics 54: 437–442.
Coy, J.F., S. Dubel, P. Kioschis, K. Thomas, G. Micklem, H. Delius & A. Poustka, 1996. Molecular cloning of tissue-specific transcripts of a transketolase-related gene: implications for the evolution of new vertebrate genes. Genomics 32: 309–316.
Crowe, M.L., B.N. Perry & I.F. Connerton, 1996. Olfactory receptor-encoding genes and pseudogenes are expressed in humans. Gene 169: 247–249.
Csoka, A.B., S.W. Scherer & R. Stern, 1999. Expression analysis of six paralogous human hyaluronidase genes clustered on chromosomes 3p21 and 7q31. Genomics 60: 356–361.
Cyster, J., C. Somoza, N. Killeen & A.F. Williams, 1990. Protein sequence and gene structure for mouse leukosialin (CD43), a T lymphocyte mucin without introns in the coding sequence. Eur. J. Immunol. 20: 875–881.
D'Onofrio, M., M.D. Lee, C.M. Starr, M. Miller & J.A. Hanover, 1991. The gene encoding rat nuclear pore glycoprotein p62 is intronless. J. Biol. Chem. 266: 11980–11985.
Dahia, P.L., M.G. FitzGerald, X. Zhang, D.J. Marsh, Z. Zheng, T. Pietsch, A. von Deimling, F.G. Haluska, D.A. Haber & C. Eng, 1998. A highly conserved processed PTEN pseudogene is located on chromosome band 9p21. Oncogene 16: 2403–2406.
Dahl, H.H., R.M. Brown, W.M. Hutchison, C. Maragos & G.K. Brown, 1990. A testis-specific form of the human pyruvate dehydrogenase E1 alpha subunit is coded for by an intronless gene on chromosome 4. Genomics 8: 225–232.
Danciger, E., C. Mettling, M. Vidal, R. Morris & F. Margolis, 1989. Olfactory marker protein gene: its structure and olfactory neuron- specific expression in transgenic mice. Proc. Natl. Acad. Sci. USA 86: 8565–8569.
DeChiara, T.M. & J. Brosius, 1987. Neural BC1 RNA: cDNA clones reveal nonrepetitive sequence content. Proc. Natl. Acad. Sci. USA 84: 2624–2628.
Descombes, P., M. Chojkier, S. Lichtsteiner, E. Falvey & U. Schibler, 1990. LAP, a novel member of the *C/EBP* gene family, encodes a liver-enriched transcriptional activator protein. Genes Dev. 4: 1541–1551.
Devilat, I. & P. Carvallo, 1993. Structure and sequence of an intronless gene for human casein kinase II- alpha subunit. FEBS Lett. 316: 114–118.
Devon, R.S., K.L. Evans, J.C. Maule, S. Christie, S. Anderson, J. Brown, Y. Shibasaki, D.J. Porteous & A.J. Brookes, 1997. Novel transcribed sequences neighbouring a translocation breakpoint associated with schizophrenia. Am. J. Med. Genet. 74: 82–90.
Di Cristofano, A., M. Strazullo, L. Longo & G. La Mantia, 1995. Characterization and genomic mapping of the ZNF80 locus: expression of this zinc-finger gene is driven by a solitary LTR of ERV9 endogenous retroviral family. Nucleic Acids Res. 23: 2823–2830.
Dierick, H.A., J.F. Mercer & T.W. Glover, 1997. A phosphoglycerate mutase brain isoform (PGAM 1) pseudogene is localized within the human Menkes disease gene (ATP7 A). Gene 198: 37–41.
DiMarco, S.P., T.W. Glover, D.E. Miller, D. Reines & S.T. Warren, 1996. Transcription elongation factor SII (TCEA) maps to human chromosome 3p22 $--> p21.3. Genomics 36: 185–188.
Donnelly, S.R., T.E. Hawkins & S.E. Moss, 1999. A conserved nuclear element with a role in mammalian gene regulation. Hum. Mol. Genet. 8: 1723–1728.
Dono, R., N. Montuori, M. Rocchi, L. De Ponti-Zilli, A. Ciccodicola & M.G. Persico, 1991. Isolation and characterization of the CRIPTO autosomal gene and its X- linked related sequence. Am. J. Hum. Genet. 49: 555–565.
Dudov, K.P. & R.P. Perry, 1984. The gene family encoding the mouse ribosomal protein L32 contains a uniquely expressed intron-containing gene and an unmutated processed gene. Cell 37: 457–468.
Ehrmann, I.E., P.S. Ellis, S. Mazeyrat, S. Duthie, N. Brockdorff, M.G. Mattei, M.A. Gavin, N.A. Affara, G.M. Brown, E. Simpson, M.J. Mitchell & D.M. Scott, 1998. Characterization of genes encoding translation initiation factor elF-2gamma in mouse and human: sex chromosome localization, escape from X- inactivation and evolution. Hum. Mol. Genet. 7: 1725–1737.
Emi, M., A. Horii, N. Tomita, T. Nishide, M. Ogawa, T. Mori & K. Matsubara, 1988. Overlapping two genes in human DNA: a salivary *amylase* gene overlaps with a gamma-actin pseudogene that carries an integrated human endogenous retroviral DNA. Gene 62: 229–235.
Ernstsson, S., S. Pierrou, M. Hulander, A. Cederberg, M. Hellqvist, P. Carlsson & S. Enerback, 1996. Characterization of the human forkhead gene FREAC-4. Evidence for regulation by Wilms' tumor suppressor gene (WT-1) and p53. J. Biol. Chem. 271: 21094–21099.
Erusalimsky, J.D., S.F. Brooks, T. Herget, C. Morris & E. Rozengurt, 1991. Molecular cloning and characterization of the acidic 80-kDa protein kinase C substrate from rat brain. Identification as a glycoprotein. J. Biol. Chem. 266: 7073–7080.
Fantin, V.R., B.E. Lavan, Q. Wang, N.A. Jenkins, D.J. Gilbert, N.G. Copeland, S.R. Keller & G.E. Lienhard, 1999. Cloning, tissue expression, and chromosomal location of the mouse insulin receptor substrate 4 gene. Endocrinology 140: 1329–1337.
Feuchter, A.E., J.D. Freeman & D.L. Mager, 1992. Strategy for detecting cellular transcripts promoted by human endogenous long terminal repeats: identification of a novel gene (CDC4L) with homology to yeast CDC4. Genomics 13: 1237–1246.

Feuchter-Murthy, A.E., F.J.D. & M.D.L. 1993. Splicing of a human retrovirus to a novel phospolipase A2 related gene. Nucleic Acids Res. 21: 135–143.

Fitzgerald, J., W.M. Hutchison, and H.H. Dahl. 1992. Isolation and characterisation of the mouse pyruvate dehydrogenase *E1 alpha* genes. Biochim. Biophys. Acta 1131: 83–90.

Foglio, M.H. & G. Duester, 1996. Characterization of the functional gene encoding mouse class III alcohol dehydrogenase (glutathione-dependent formaldehyde dehydrogenase) and an unexpressed processed pseudogene with an intact open reading frame. Eur. J. Biochem. 237: 496–504.

Fornasari, D., E. Bataglioli, A. Flora, S. Terzano & F. Clementi, 1997. Structural and functional characterization of the human alpha 3 nicotinic subunit gene promoter. Molecular Pharmacol. 51: 250–261.

Foster, J.W. & J.A. Graves, 1994. An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. Proc. Natl. Acad. Sci. USA 91: 1927–1931.

Fourel, G., C. Transy, B.C. Tennant & M.A. Buendia, 1992. Expression of the woodchuck N-myc2 retroposon in brain and in liver tumors is driven by a cryptic N-myc promoter. Mol. Cell. Biol. 12: 5336–5344.

Fourel, G., C. Trepo, L. Bougueleret, B. Henglein, A. Ponzetto, P. Tiollais & M.A. Buendia, 1990. Frequent activation of N-myc genes by hepadnavirus insertion in woodchuck liver tumours. Nature 347: 294–298.

Frank, S. & B. Zoll, 1998. Mouse HNF-3/fork head homolog-1-like gene: structure, chromosomal location, and expression in adult and embryonic kidney. DNA Cell. Biol. 17: 679–688.

Fratini, A., B.C. Powell & G.E. Rogers, 1993. Sequence, expression, and evolutionary conservation of a gene encoding a glycine/tyrosine-rich keratin-associated protein of hair. J. Biol. Chem. 268: 4511–4518.

Frenkel, M.J., B.C. Powell, K.A. Ward, M.J. Sleigh & G.E. Rogers, 1989. The keratin BIIIB gene family: isolation of cDNA clones and structure of a gene and a related pseudogene. Genomics 4: 182–191.

Fujii, T., K. Tamura, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, K. Yomogida, H. Tanaka, Y. Nishimune, H. Nojima & Y. Abiko, 1999. Sperizin is a murine RING zinc-finger protein specifically expressed in Haploid germ cells. Genomics 57: 94–101.

Gafvels, M., M. Olin, B.P. Chowdhary, T. Raudsepp, U. Andersson, B. Persson, M. Jansson, I. Bjorkhem & G. Eggertsen, 1999. Structure and chromosomal assignment of the sterol *12alpha-hydroxylase* gene (CYP8B1) in human and mouse: eukaryotic cytochrome P-450 gene devoid of introns. Genomics 56: 184–196.

Galili, N., H.S. Baldwin, J. Lund, R. Reeves, W. Gong, Z. Wang, B.A. Roe, B.S. Emanuel, S. Nayak, C. Mickanin, M.I. Budarf & C.A. Buck, 1997. A region of mouse chromosome 16 is syntenic to the DiGeorge, velocardiofacial syndrome minimal critical region. Genome Res. 7: 399.

Gao, M., W. Rychlik & R.E. Rhoads, 1998. Cloning and characterization of human eIF4E genes. J. Biol. Chem. 273: 4622–4628.

Gedeon, A.K., A. Colley, R. Jamieson, E.M. Thompson, J. Rogers, D. Sillence, G.E. Tiller, J.C. Mulley & J. Gecz, 1999. Identification of the gene (SEDL) causing X-linked spondyloepiphyseal dysplasia tarda. Nat Genet. 22: 400–404.

Gentles, A.J. & S. Karlin, 1999. Why are human G-protein-coupled receptors predominantly intronless? Trends Genet. 15: 47–49.

Gerber, A.M., M.A. O'Connell & W. Keller, 1997. Two forms of human double-stranded RNA-specific editase 1 (hRED1) generated by the insertion of an Alu cassette. Rna 3: 453–463.

Glaichenhaus, N. & F. Cuzin, 1987. A role for ID repetitive sequences in growth- and transformation-dependent regulation of gene expression in rat fibroblasts. Cell. 50: 1081–1089.

Goldenthal, M.J., J. Marin-Garcia & R. Ananthakrishnan, 1998. Cloning and molecular analysis of the human citrate synthase gene. Genome 41: 733–738.

Goodchild, N.L., D.A. Wilkinson & D.L. Mager, 1992. A human endogenous long terminal repeat provides a polyadenylation signal to novel, alternatively spliced transcript in normal placenta. Gene 121: 287–294.

Grant, D.M. M. Blum, A. Demierre & U.A. Meyer, 1989. Nucleotide sequence of an intronless gene for a human arylamine N- acetyltransferase related to polymorphic drug acetylation. Nucleic Acids Res. 17: 3978.

Gromoll, J., E. Pekel & E. Nieschlag, 1996. The structure and organization of the human follicle-stimulating hormone receptor (FSHR) gene. Genomics 35: 308–311.

Guehenneux, F., L. Duret, M.B. Callanan, R. Bouhas, S. Hayette, C. Berthet, C. Samarut, R. Rimokh, A.M. Birot, Q. Wang, J.P. Magaud & J.P. Rouault, 1997. Cloning of the mouse BTG3 gene and definition of a new gene family (the BTG family) involved in the negative control of the cell cycle. Leukemia. 11: 370–375.

Guo, N., T. Mogues, S. Weremowicz, C.C. Morton & K.N. Sastry, 1998. The human ortholog of rhesus mannose-binding protein-A gene is an expressed pseudogene that localizes to chromosome 10. Mamm. Genome 9: 246–249.

Hakim, I., N. Amariglio, Z. Grossman, F. Simoni Brok, S. Ohno & G. Rechavi, 1994. The genome of the THE I human transposable repetitive elements is composed of a basic motif homologous to an ancestral immunoglobulin gene sequence. Proc. Natl. Acad. Sci. USA 91: 7967–7969.

Hamann, K.J., R.M. Ten, D.A. Loegering, R.B. Jenkins, M.T. Heise, C.R. Schad, L.R. Pease, G.J. Gleich & R.L. Barker, 1990. Structure and chromosome localization of the human eosinophil-derived neurotoxin and eosinophil cationic protein genes: evidence for intronless coding sequences in the ribonuclease gene superfamily. Genomics 7: 535–546.

Hambor, J.E., J. Mennone, M.E. Coon, J.H. Hanke & P. Kavathas, 1993. Identification and characterization of an *Alu*-containing, T-cell-specific enhancer located in the last intron of the human CD8a gene. Mol. Cell Biol. 13: 7056–7070.

Hara, Y., A.C. Rovescalli, Y. Kim & M. Nirenberg, 1992. Structure and evolution of four POU domain genes expressed in mouse brain. Proc. Natl. Acad. Sci. USA 89: 3280–3284.

Harendza, C.J. & L.F. Johnson, 1990. Polyadenylylation signal of the mouse thymidylate synthase gene was created by insertion of an L1 repetitive element downstream of the open reading frame. Proc. Natl. Acad. Sci. USA 87: 2531–2535.

Hart, P.E., J.N. Glantz, J.D. Orth, G.M. Poynter & J.L. Salisbury, 1999. Testis-specific murine centrin, Cetn 1: genomic characterization and evidence for retroposition of a gene encoding a centrosome protein. Genomics 60: 111–120.

Hartl, M., J.T. Hutchins & P.K. Vogt, 1991. The chicken junD gene and its product. Oncogene 6: 1623–1631.

Hattori, K., P. Angel, M.M. Le Beau & M. Karin, 1988. Structure and chromosomal localization of the functional intronless human JUN protooncogene. Proc. Natl. Acad. Sci USA 85:9148–9152.

Haydock, P.V. & B.A. Dale, 1986. The repetitive structure of the profilaggrin gene as demonstrated using epidermal profilaggrin cDNA. J. Biol. Chem. 261: 12520–12525.

Hendriksen, P.J., J.W. Hoogerbrugge, W.M. Baarends, P. de Boer, J.T. Vreeburg, E.A. Vos, T van der Lende & J.A. Grootegoed, 1997. Testis-specific expression of a functional retroposon

encoding glucose- 6-phosphate dehydrogenase in the mouse. Genomics 41: 350–359.

Hennemann, H., E. Dahl, J.B. White, H.J. Schwarz, P.A. Lalley, S. Chang, B.J. Nicholson & K. Willecke, 1992. Two gap junction genes, connexin 31.1 and 30.3, are closely linked on mouse chromosome 4 and preferentially expressed in skin. J. Biol. Chem. 267: 17225–17233.

Hentschel, C.C. & M.L. Birnstiel, 1981. The organization and expression of histone gene families. Cell. 25: 301–313.

Heron, L., A. Virsolvy, F. Apiou, A. Le Cam & D. Bataille, 1999. Isolation, characterization, and chromosomal localization of the human *ENSA* gene that encodes alpha-endosulfine, a regulator of beta-cell K(ATP) channels. Diabetes 48: 1873–1876.

Hewitt, S.M., G.C. Fraizer & G.F. Saunders, 1995. Transcriptional silencer of the Wilms' tumor gene WT1 contains an Alu repeat. J. Biol. Chem. 270: 17908–17912.

Hirsch, A.H., S.B. Glantz, Y. Li, Y. You & C.S. Rubin, 1992. Cloning and expression of an intron-less gene for AKAP 75, an anchor protein for the regulatory subunit of cAMP-dependent protein kinase II beta. J. Biol. Chem. 267: 2131–2134.

Hohlbaum, A.M. U. Krawinkel & P. Hemmerich, 1998. Structure and expression of a non-polyadenylated ribosome bound transcript carrying the sequence of human ribosomal protein L7 mRNA in antisense orientation. Biol. Chem. 379: 1193–1200.

Hong, Y., K. Ohishi, N. Inoue, Y. Endo, T. Fujita, J. Takeda & T. Kinoshita, 1997. Structures and chromosomal localizations of the glycosylphosphatidylinositol synthesis gene PIGC and its pseudogene PIGCP1. Genomics. 44: 347–349.

Hsu, D.W., M.J. Lin, T.L. Lee, S.C. Wen, X. Chen & C.K. Shen, 1999. Two major forms of DNA (cytosine-5) methyltransferase in human somatic tissues. Proc. Natl. Acad. Sci. USA 96: 9751–9756.

Hui, P., X. Guo & P.G. Bradford, 1995. Isolation and functional characterization of the human gene encoding the myeloid zinc finger protein MZF-1. Biochemistry 34: 16493–16502.

Humphrey, G.W., E.W. Englander & B.H. Howard, 1996. Specific binding sites for a pol III transcriptional repressor and pol II transcription factor YY1 within the internucleosomal spacer region in primate Alu repetitive elements. Gene. Expr. 6: 151–168.

Hunt, C. & R.I. Morimoto, 1985. Conserved features of eukaryotic hsp70 genes revealed by comparison with the nucleotide sequence of human hsp70. Proc. Natl. Acad. Sci. USA 82: 6455–6459.

Jackers, P., N. Clausse, M. Fernandez, A. Berti, F. Princen, U. Wewer, M.E. Sobel & V. Castronovo, 1996. Seventeen copies of the human 37 kDa laminin receptor precursor/p40 ribosome-associated protein gene are processed pseudogenes arisen from retropositional events. Biochim. Biophys. Acta. 1305: 98–104.

Jackman, R.W., D.L. Beeler, L. Fritze, G. Soff & R.D. Rosenberg, 1987. Human thrombomodulin gene is intron depleted: nucleic acid sequences of the cDNA and gene predict protein structure and suggest sites of regulatory control. Proc. Natl. Acad. Sci. USA 84: 6425–6429.

Jankowski, J.M., J.C. States & G.H. Dixon, 1986. Evidence of sequences resembling avian retrovirus long terminal repeats flanking the trout protamine gene. J. Mol. Evol. 23: 1–10.

Jong, M.T., T.A. Gray, Y. Ji, C.C. Glenn, S. Saitoh, D.J. Driscoll & R.D. Nicholls, 1999. A novel imprinted gene, encoding a RING zinc-finger protein, and overlapping antisense transcript in the Prader-Willi syndrome critical region. Hum. Mol. Genet. 8: 783–793.

Jun, D.Y., H.S. Kang, J.H. Seu & Y.H. Kim, 1997. Isolation and characterization of a processed pseudogene for murine cyclin D3. Mol. Cells 7: 278–283.

Kapitonov, V.V. & J. Jurka, 1999. The long terminal repeat of an endogenous retrovirus induces alterantive splicing and encodes an additional carboxy-terminal sequence in the human leptin receptor. J. Mol. Evol. 48: 248–251.

Kas, K., D. Stickens & J. Merregaert, 1995. Characterization of a processed pseudogene of human FAU1 on chromosome 18. Gene 160: 273–276.

Kawaichi, M., C. Oka, S. Shibayama, A.E. Koromilas, N. Matsunami, Y. Hamaguchi & T. Honjo, 1992. Genomic organization of mouse J kappa recombination signal binding protein (RBP-J kappa) gene. J. Biol. Chem. 267: 4016–4022.

Kedes, L.H., 1979. Histone genes and histone messengers. Annu. Rev. Biochem. 48: 837–870.

Kim, J.H., C.Y. Yu, A. Bailey, R. Hardison & C.K. Shen, 1989. Unique sequence organization and erythroid cell-specific nuclear factor-binding of mammalian theta 1 globin promoters. Nucleic Acids Res. 17: 5687–5700.

Kinoshita, S., S. Akira & T. Kishimoto, 1992. A member of the C/EBP family, NF-IL6 beta, forms a heterodimer and transcriptionally synergizes with NF-IL6. Proc. Natl. Acad. Sci. USA 89: 1473–1476.

Kitagawa, K., X. Wang, I. Hatada, T. Yamaoka, H. Nojima, J. Inazawa, T. Abe, K. Mitsuya, M. Oshimura, A. Murata, et al. 1995. Isolation and mapping of human homologues of an imprinted mouse gene U2af1-rs1. Genomics 30: 257–263.

Kleene, K.C., E. Mulligan, D. Steiger, K. Donohue & M.A. Mastrangelo, 1998. The mouse gene encoding the testis-specific isoform of Poly(A) binding protein (Pabp2) is an expressed retroposon: intimations that gene expression in spermatogenic cells facilitates the creation of new genes. J. Mol. Evol. 47: 275–281.

Krane, D.E. & R.C. Hardison, 1990. Short interspersed repeats in rabbit DNA can provide functional polyadenylation signals. Mol. Biol. Evol. 7: 1–8.

Kress, M., Y. Barra, J.G. Seidman, G. Khoury & G. Jay, 1984. Functional insertion of an Alu type 2 (B2 SINE) repetitive sequence in murine class I genes. Science 226: 974–977.

Krücken, J., O. Stamm, H.P. Schmitt-Wrede, A Mincheva, P. Lichter & F. Wunderlich, 1999. Spleen-specific expression of the malaria-inducible intronless mouse gene imap38. J. Biol. Chem. 274: 24383–24391.

Kuczek, E.S. & G.E. Rogers, 1987. Sheep wool (glycine + tyrosine)-rich keratin genes. A family of low sequence homology. Eur. J. Biochem. 166: 79–85.

Kudo, S., M.G. Mattei & M. Fukuda, 1995. Characterization of the gene for dbpA, a family member of the nucleic- acid-binding proteins containing a cold-shock domain. Eur. J. Biochem. 231: 72–82.

Kuhn, F., C. Lassing, A. Range, M. Mueller, T. Hunziker, A. Ziemiecki & A.C. Andres, 1999. Pmg-1 and pmg-2 constitute a novel family of *KAP* genes differentially expressed during skin and mammary gland development. Mech. Dev. 86: 193–196.

Kuhn, R., E.S. Monuki & G. Lemke, 1991. The gene encoding the transcription factor SCIP has features of an expressed retroposon. Mol. Cell Biol. 11: 4642–4650.

Kuzumaki, T., T. Tanaka, K. Ishikawa & K. Ogata, 1987. Rat ribosomal protein L35a multigene family: molecular structure and characterization of three L35a-related pseudogenes. Biochim. Biophys. Acta. 909: 99–106.

Lahn, B.T. & D.C. Page, 1999. Retroposition of autosomal mRNA yielded testis-specific gene family on human Y chromosome. Nat. Genet. 21: 429–433.

Laimins, L., K.M. Holmgren & G. Khoury, 1986. Transcriptional 'silencer' element in rat repetitive sequences associated with the

rat insulin 1 gene locus. Proc. Natl. Acad. Sci. USA 83: 3151–3155.

Lambotte, S., M. Veyhl, M. Kohler, A.I. Morrison-Shetlar, R.K. Kinne, M. Schmid & H. Koepsell, 1996. The human gene of a protein that modifies Na(+)-D-glucose co-transport. DNA Cell. Biol. 15: 769–777.

Lan, M.S. Q. Li, J. Lu, W.S. Modi & A.L. Notkins, 1994. Genomic organization, 5′-upstream sequence, and chromosomal localization of an insulinoma-associated intronless gene, IA-1. J. Biol. Chem. 269: 14170–14174.

Land, S.J., R.F. Jones & C.M. King, 1996. Genetic analysis of two rat acetyltransferases. Carcinogenesis 17: 1121–1126.

Landschulz, W.H., P.F. Johnson, E.Y. Adashi, B.J. Graves & S.L. McKnight, 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes. Dev. 2: 786–800.

Lanza, F., M. Morales, C. de La Salle, J.P. Cazenave, K.J. Clemetson, T. Shimomura & D.R. Phillips, 1993. Cloning and characterization of the gene encoding the human platelet glycoprotein V. A member of the leucine-rich glycoprotein family cleaved during thrombin-induced platelet activation. J. Biol. Chem. 268: 20801–20807.

Lawn, R.M., J. Adelman, A.E. Franke, C.M. Houck, M. Gross, R. Najarian & D.V. Goeddel, 1981. Human fibroblast interferon gene lacks introns. Nucleic Acids Res. 9: 1045–1052.

Le Gallic, L. & P. Fort, 1997. Structure of the human ARHG locus encoding the Rho/Rac-like RhoG GTPase. Genomics 42: 157–160.

Lee, I.Y., D. Westaway, A.F. Smit, K. Wang, J. Seto, L. Chen, C. Acharya, M. Ankener, D. Baskin, C. Cooper, H. Yao, S.B. Prusiner & L.E. Hood, 1998. Complete genomic sequence and analysis of the prion protein gene region from three mammalian species. Genome Res. 8: 1022–1037.

Lee, T.H., S.L. Yu, S.U. Kim, Y.M. Kim, I. Choi, S.W. Kang, S.G. Rhee & D.Y. Yu. 1999. Characterization of the murine gene encoding 1-Cys peroxiredoxin and identification of highly homologous genes. Gene 234: 337–344.

Levavasseur, F., W. Mandemakers, P. Visser, L. Broos, F. Grosveld, D. Zivkovic & D. Meijer, 1998. Comparison of sequence and function of the Oct-6 genes in zebrafish, chicken and mouse. Mech. Dev. 74: 89–98.

Levinson, B., J.R. Bermingham, Jr., A. Metzenberg, S. Kenwrick, V. Chapman & J. Gitschier, 1992. Sequence of the human factor VIII-associated gene is conserved in mouse. Genomics 13: 862–865.

Li, Q., A.L. Notkins & M.S. Lan, 1997. Molecular characterization of the promoter region of a neuroendocrine tumor marker, IA-1. Biochem. Biophys. Res. Commun. 236: 776–781.

Liguori, G., L. De Gregorio, M. Tucci, C.T. Lago, A. Barra, T.A. Dragani & M. Persico, 1997. Mapping of the mouse *Tdgf1* gene and *Tdgf* pseudogenes. Mamm. Genome. 8: 502–505.

Liguori, G., M. Tucci, N. Montuori, R. Dono, C.T. Lago, F. Pacifico, F. Armenante & M.G. Persico, 1996. Characterization of the mouse *Tdgf1* gene and *Tdgf* pseudogenes. Mamm. Genome. 7: 344–348.

Lim, E.H. & S. Brenner, 1999. Short-range linkage relationships, genomic organisation and sequence comparisons of a cluster of the HSP70 genes in Fugu rubrips. Cell. Mol. Life Sci. 55: 668–678.

Linnenbach, A.J., B.A. Seng, S. Wu, S. Robbins, M. Scollon, J.J. Pyrc, T. Druck & K. Huebner, 1993. Retroposition in a family of carcinoma-associated antigen genes. Mol. Cell. Biol. 13: 1507–1515.

Linnenbach, A.J., J. Wojcierowski, S.A. Wu, J.J. Pyrc, A.H. Ross, B. Dietzschold, D. Speicher & H. Koprowski, 1989. Sequence investigation of the major gastrointestinal tumor-associated antigen gene family, GA733. Proc. Natl. Acad. Sci. USA 86: 27–31.

Liu, A.Y. & B.A. Abraham, 1991. Subtractive cloning of a hybrid human endogenous retrovirus and calbindin gene in the prostate cell line PC3. Cancer Res. 51: 4107–4110.

Liu, A.Y. & R.C. Bradner, 1993. Elevated expression of the human mitochondrial hinge protein gene in cancer. Cancer Res. 53: 2460–2465.

Lowndes, N.F., P. Bushel, L. Mendelsohn J. Wu, M.Y. Yen & M. Allan, 1990. A short, highly repetitive element in intron -1 of the human c-Ha-ras gene acts as a block to transcriptional readthrough by a viral promoter. Mol. Cell. Biol. 10: 4990–4995.

Ludwig, T., U. Ruther, R. Metzger, N.G. Copeland, N.A. Jenkins, P. Lobel & B. Hoflack, 1992. Gene and pseudogene of the mouse cation-dependent mannose 6-phosphate recetor. Genomic organization, expression, and chromosomal localization. J. Biol. Chem. 267: 12211–12219.

Lum, R. & M.L. Linial, 1997. Tail-to-head arrangement of a partial chicken glyceraldehyde-3- phosphate dehydrogenase processed pseudogene. J. Mol. Evol. 45: 564–570.

Lum, R. & M.L. Linial, 1998. Retrotransposition of nonviral RNAs in an avian packaging cell line. J. Virol. 72: 4057–4064.

Lund, E.G., T.A. Kerr, J. Sakai, W.P. Li & D.W. Russell, 1998. cDNA cloning of mouse and human cholesterol 25-hydroxylases, polytopic membrane proteins that synthesize a potent oxysterol regulator of lipid metabolism. J. Biol. Chem. 273: 34316–34327.

Luo, M., J. Shang, Z. Yang, C.P. Simkevich, C.L. Jackson, T.C. King & A.G. Rosmarin, 1999. Characterization and localization to chromosome 7 of psihGABPalpha, a human processed pseudogene related to the ets transcription factor, hGABPalpha. Gene 234: 119–126.

MacLeod, A.R. & K. Talbot, 1983. A processed gene defining a gene family encoding a human non-muscle tropomyosin. J. Mol. Biol. 167: 523–537.

Mager, D.L., 1989. Polyadenylation function and sequence variability of the long terminal repeats of the human endogenous retrovirus-like family RTVL-H. Virology 173: 591–599.

Maichele, A.J., N.J. Farwell & J.S. Chamberlain, 1993. A B2 repeat insertion generates alternate structures of the mouse muscle gamma-phosphorylase kinase gene. Genomics 16: 139–149.

Makalowski, W., G.A. Mitchell & D. Labuda, 1994. Alu sequences in the coding regions of mRNA: a source of protein variability. Trends Genet. 10: 188–193.

Makeyev, A.V., A.N. Chkheidze & S.A. Liebhaber, 1999. A set of highly conserved RNA-binding proteins, alphaCP-1 and alphaCP-2, implicated in mRNA stabilization, are coexpressed from an intronless gene and its intron-containing paralog. J. Biol. Chem. 274: 24849–24857.

Manrow, R.E., A. Leone, M.S. Krug, W.H. Eschenfeldt & S.L. Berger, 1992. The human prothymosin alpha gene family contains several processed pseudogenes lacking deleterious lesions. Genomics 13: 319–331.

Marron-Terada, P.G., K.E. Bollinger & N.M. Dahms, 1998. Characterization of truncated and glycosylation-deficient forms of the cation-dependent mannose 6-phosphate receptor expressed in baculovirus-infected insect cells. Biochemistry 37: 17223–17229.

Martell, K.J., K.P. Vatsis & W.W. Weber, 1991. Molecular genetic basis of rapid and slow acetylation in mice. Mol. Pharmacol. 40: 218–227.

Martignetti, J.A. & J. Brosius, 1993a. BC200 RNA: a neural RNA polymerase III product encoded by a monomeric *Alu* element. Proc. Natl. Acad. Sci. USA 90: 11563–11567.

Martignetti, J.A. & J. Brosius, 1993b. Neural BC1 RNA as an evolutionary marker: guinea pig remains a rodent. Proc. Natl. Acad. Sci. USA 90: 9698–9702.

Martin, W.J., 1999. Melanoma growth stimulatory activity (MGSA-GRO-alpha) chemokine genes incorporated into an African green monkey simian cytomegalovirus- derived stealth virus. Exp. Mol. Pathol. 66: 15–18.

Matsumine, H., M.A. Herbst, S.H. Ou, J.D. Wilson & M.J. McPhaul, 1991. Aromatase mRNA in the extragonadal tissues of chickens with the henny-feathering trait is derived from a distinctive promoter structure that contains a segment of a retroviral long terminal repeat. Functional organization of the Sebright, Leghorn, and Campine aromatase genes. J. Biol. Chem. 266: 19900–19907.

Mazany, K.D., T. Peng, C.E. Watson, I. Tabas & K.J. Williams, 1998. Human chondroitin 6-sulfotransferase: cloning, gene structure, and chromosomal localization. Biochim. Biophys. Acta. 1407: 92–97.

McCarrey, J.R. & K. Thomas, 1987. Human testis-specific *PGK* gene lacks introns and possesses characteristics of a processed gene. Nature 326: 501–505.

McHaffie, G.S. & S.H. Ralston, 1995. Origin of a negative calcium response element in an ALU-repeat: implications for regulation of gene expression by extracellular calcium. Bone 17: 11–14.

McKinnon, R.D., T.M. Shinnick & J.G. Sutcliffe, 1986. The neuronal identifier element is a cis-acting positive regulator of gene expression. Proc. Natl. Acad. Sci. USA 83: 3751–3755.

Michel, D., G. Chatelain, C. Mauduit, M. Benahmed & G. Brun, 1997. Recent evolutionary acquisition of alternative pre-mRNA splicing and 3′ processing regulations induced by intronic B2 SINE insertion. Nucleic Acids Res. 25: 3228–3234.

Mihovilovic, M., Y. Mai, M. Herbstreith, F. Rubboli, P. Tarroni, F. Clementi & A.D. Roses, 1993. Splicing of an anti-sense Alu sequence generates a coding sequence variant for the alpha-3 subunit of a neuronal acetylcholine receptor. Biochem. Biophys. Res. Commun. 197: 137–144.

Milner, C.M. & R.D. Campbell, 1990. Structure and expression of the three MHC-linked *HSP70* genes. Immunogenetics 32: 242–251.

Mishra, L., P. Yu, T. Cai, S.P. Monga & B. Mishra, 1999. Genomic structure, chromosomal mapping, and muscle-specific expression of a PH domain-associated intronless gene, cded/lior. Mamm. Genome. 10: 62–67.

Misrahi, M., N. Ghinea, M.T. Vu Hai, H. Loosfelt, G. Meduri, M. Atger, B. Gross, A. Jolivet & E. Milgrom, 1995. [Pituitary glycoprotein hormone receptors]. Ann. Endocrinol. (Paris) 56: 487–493.

Mitsui, S., A. Ohuchi, T. Adachi-Yamada, M. Hotta, R. Tsuboi & H. Ogawa, 1998. Structure and hair follicle-specific expression of genes encoding the rat high sulfur protein *B2* family. Gene 208: 123–129.

Miura, N., K. Iida, H. Kakinuma, X.L. Yang & T. Sugiyama, 1997. Isolation of the mouse (MFH-1) and human (FKHL 14) mesenchyme fork head-1 genes reveals conservation of their gene and protein structures. Genomics 41: 489–492.

Moir, R.D. & G.H. Dixon, 1988. Characterization of a protamine gene from the chum salmon (Oncorhynchus keta). J. Mol. Evol. 27: 8–16.

Moynihan, T.P., H.C. Ardley, J.P. Leek, J. Thompson, N.S. Brindle, A.F. Markham & P.A. Robinson, 1996. Characterization of a human ubiquitin-conjugating enzyme gene *UBEL2L3*. Mamm Genome. 7: 520–525.

Mues, G.I., T.Z. Munn & J.D. Raese, 1986. A human gene family with sequence homology to *Drosophila melanogaster* Hsp70 heat shock genes. J. Biol. Chem. 261: 874–877.

Mullersman, J.E. & L.M. Pfeffer, 1995. An Alu cassette in the cytoplasmic domain of an interferon receptor subunit. J. Interferon Cytokine Res. 15: 815–817.

Murnane, J.P. & J.F. Morales, 1995. Use of a mammalian interspersed repetitive (MIR) element in the coding and processing sequences of mammalian genes. Nucleic Acids Res. 23: 2837–2839.

Nagata, S., N. Mantei & C. Weissmann, 1980. The structure of one of the eight or more distinct chromosomal genes for human interferon-alpha. Nature 287: 401–408.

Nakada, Y., H. Taniura, T. Uetsuki, J. Inazawa & K. Yoshikawa, 1998. The human chromosomal gene for necdin, a neuronal growth suppressor, in the Prader-Willi syndrome deletion region. Gene 213: 65–72.

Newman, B. & Y. Dai, 1996. Transcription of c-mas protooncogene in the pig involves both tissue-specific promoters and alternative polyadenylation sites. Mol. Reprod. Dev. 44: 275–288.

Neznanov, N.S. & R.G. Oshima, 1993. cis regulation of the keratin 18 gene in transgenic mice. Mol. Cell. Biol. 13: 1815–1823.

Nhim, R.P., J.S. Lindsey & M.F. Wilkinson, 1997. A processed homeobox gene expressed in a stage, tissue and region specific manner in epididymis. Gene 185: 271–276.

Nicole, S., P.S. White, H. Topaloglu, P. Beigthon, M. Salih, F. Hentati & B. Fontaine, 1999. The human CDC42 gene: genomic organization evidence for the existence of a putative pseudogene and exclusion as a *SJS1* candidate gene. Hum. Genet. 105: 98–103.

Niforas, P., G.M. Sanderson, C.H. Bird & P. Bird, 1993. Characterization of the mouse thrombomodulin gene and functional analysis of the 5′ flanking region in F9 teratocarcinoma cells. Biochim. Biophys. Acta. 1173: 179–187.

Ninomiya, Y., M. Gordon, M. van der Rest, T. Schmid, T. Linsenmayer & B.R. Olsen, 1986. The developmentally regulated type X collagen gene contains a long open reading frame without introns. J. Biol. Chem. 261: 5041–5050.

Nishimura, K., M. Liisanantti, Y. Muta, K. Kashiwagi, A. Shirahata, M. Janne, K. Kankare, O.A. Janne & K. Igarashi, 1998. Structure and activity of mouse S-adenosylmethionine decarboxylase gene promoters and properties of the encoded proteins. Bichem. J. 332: 651–659.

Nojima, H., K. Kishi & H. Sokabe, 1987. Multiple calmodulin mRNA species are derived from two distinct genes. Mol. Cell. Biol. 7: 1873–1880.

Norris, J., D. Fan, C. Aleman, J.R. Marks, P.A. Futreal, R.W. Wiseman, J.D. Iglehart, P.L. Deininger & D.P. McDonnell, 1995. Identification of a new subclass of Alu DNA repeats which can function as estrogen receptor-dependent transcriptional enhancers. J. Biol. Chem. 270: 22777–22782.

Noyce, L., J. Conaty & A.A. Piper, 1997. Identification of a novel tissue-specific processed *HPRT* gene and comparison with X-linked gene transcription in the Australian marsupial Macropus robustus. Gene 186: 87–95.

Noyce, L. & A.A. Piper, 1994. Isolation of a potentially functional hprt processed pseudogene from the hill kangaroo macropus robustus. Gene 150: 361–365.

O'Neill, R.J., F.E. Brennan, M.L. Delbridge, R.H. Crozier & J.A. Graves, 1998. De novo insertion of an intron into the mammalian sex determining gene, SRY. Proc. Natl. Acad. Sci. USA 95: 1653–1657.

Oei, S.L., J. Griesenbeck, M. Schweiger, V. Babich, A. Kropotov & N. Tomilin. 1997. Interaction of the transcription factor YY1 with human poly(ADP- ribosyl) transferase. Biochem. Biophys. Res. Commun. 240: 108–111.

Okino, S.T., L.C. Quattrochi, H.J. Barnes, S. Osanto, K.J. Griffin, E.F. Johnson & R.H. Tukey, 1985. Cloning and characterization of cDNAs encoding 2,3,7,8-tetrachlorodibenzo-p-dioxin-inducible rabbit mRNAs for cytochrome P-450 isozymes 4 and 6. Proc. Natl. Acad. Sci. USA 82: 5310–5314.

Oliva, R. & G.H. Dixon, 1989. Chicken protamine genes are intronless. The complete genomic sequence and organization of the two loci. J. Biol. Chem. 264: 12472–12481.

Oliviero, S. & P. Monaci, 1988. RNA polymerase III promoter elements enhance transcription of RNA polymerase II genes. Nucleic Acids Res. 16: 1285–1293.

Owczarek, C.M., M.J. Layton, L.G. Robb, N.A. Nicola & C.G. Begley, 1996. Molecular basis of the soluble and membrane-bound forms of the murine leukemia inhibitory factor receptor alpha-chain. Expression in normal, gestating, and leukemia inhibitory factor nullizygous mice. J. Biol. Chem. 271: 5495–5504.

Ozer, J., R. Chalkley & L. Sealy, 1993. Isolation of the CCAAT transcription factor subunit EFIA cDNA and a potentially functional EFIA processed pseudogene from Bos taurus: insights into the evolution of the *EFIA/dbpB/YB-1* gene family. Gene 124: 223–230.

Palmer, D.B., J.H. McVey, R. Purohit, J. Picard & P.J. Dyson, 1998. Characterization of a recent retroposon inseretion on mouse chromosome 2 and localization of the cognate parental gene to chromosome 11. Mamm. Genome 9: 103–106.

Pan, Y., W.K. Decker, A.H. Huq & W.J. Craigen, 1999. Retrotransposition of glycerol kinase-related genes from the X chromosome to autosomes: functional and evolutionary aspects. Genomics 59: 282–290.

Park, H., K. Baek, C. Jeon, K. Agarwal & O. Yoo, 1994. Characterization of the gene encoding the human transcriptional elongation factor TFIIS. Gene 139: 263–267.

Paulson, K.E., A.G. Matera, and N. Deka & C.W. Schmid, 1987. Transcription of a human transposon-like sequence is usually directed by other promoters. Nucleic Acids Res. 15: 5199–5215.

Peacock, R.E., T.J. Keen & C.F. Inglehearn, 1997. Analysis of a human gene homologous to rat ventral prostate. 1 protein. Genomics 46: 443–449.

Pearsall, R.S., H. Shibata, A. Brozowska, K. Yoshino, K. Okuda, P.J. deJong, C. Plass, V.M. Chapman, Y. Hayashizaki & W.A. Held, 1996. Absence of imprinting in U2AFBPL, a human homologue of the imprinted mouse gene U2afbp-rs. Biochem. Biophys. Res. Commun. 222: 171–177.

Peng, H.L., S.D. Cheng, J.H. Lee & H.Y. Chang, 1999. Identification of a novel family of human Rab-like small GTP-binding proteins. Proc. Natl. Sci. Counc. Repub. China B. 23: 38–44.

Perez Jurado, L.A., Y.K. Wang, R. Peoples, A. Coloma, J. Cruces & U. Francke, 1998. A duplicate gene in the breakpoint regions of the 7q11.23 Williams–Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of BTK. Hum. Mol. Genet. 7: 325–334.

Pérez, M.J., C. Leroux, A.S. Bonastre & P. Martin, 1994. Occurrence of a LINE sequence in the 3′ UTR of the goat alpha s1-casein E-encoding allele associated with reduced protein synthesis level. Gene 147: 179–187.

Persson, K., O. Heby & F.G. Berger, 1999. The functional intronless S-adenosylmethionine decarboxylase gene of the mouse (Amd-2) is linked to the ornithine decarboxylase gene (Odc) on chromosome 12 and is present in distantly related species of the genus Mus. Mamm. Genome. 10: 784–788.

Persson, K., I. Holm & O. Heby, 1995. Cloning and sequencing of an intronless mouse S-adenosylmethionine decarboxylase gene coding for a functional enzyme strongly expressed in the liver. J. Biol. Chem. 270: 5642–5648.

Phelan, S.A., K.A. Johnson, D.R. Beier & B. Paigen, 1998. Characterization of the murine gene encoding Aop2 (antioxidant protein 2) and identification of two highly related genes. Genomics 54: 132–139.

Piedrafita, F.J., R.B. Molander, G. Vansant, E.A. Orlova, M. Pfahl & W.F. Reynolds, 1996. An Alu element in the myeloperoxidase promoter contains a composite SP1-thyroid hormone-retinoic acid response element. J. Biol. Chem. 271: 14412–14420.

Pietzsch, A., C. Buchler & G. Schmitz, 1998. Genomic organization, promoter cloning, and chromosomal localization of the Dif-2 gene. Biochem. Biophys. Res. Commun. 245: 651–657.

Poindexter, K., N. Nelson, R.F. DuBose, R.A. Black & D.P. Cerretti, 1999. The identification of seven metalloproteinase-disintegrin (ADAM) genes from genomic libraries. Gene 237: 61–70.

Porcellati, F., Y. Hosaka, T. Hlaing, M. Togawa, D.D. Larkin, A. Karihaloo, M.J. Stevens, P.D. Killen & D.A. Greene, 1999. Alternate splicing in human Na+-MI cotransporter gene yields differentially regulated transport isoforms. Am. J. Physiol. 276: C1325–C1337.

Powell, B.C., G.R. Cam, M.J. Fietz & G.E. Rogers, 1986. Clustered arrangement of keratin intermediate filament genes. Proc. Natl. Acad. Sci. USA 83: 5048–5052.

Powell, B.C., M.J. Sleigh, K.A. Ward & G.E. Rogers, 1983. Mammalian keratin gene families: organisation of genes coding for the B2 high-sulphur proteins os sheep wool. Nucleic Acids Res. 11: 5327–5346.

Prezant, T.R., P. Kadioglu & S. Melmed, 1999. An intronless homolog of human proto-oncogene *hPTTG* is expressed in pituitary tumors: evidence for *hPTTG* family. J. Clin. Endocrinol. Metab. 84: 1149–1152.

Prinsen, C.F., D.O. Weghuis, A.G. Kessel & J.H. Veerkamp, 1997. Identification of a human heart *FABP* pseudogene located on chromosome 13. Gene 193: 245–251.

Quignon, F., C.A. Renard, P. Tiollais, M.A. Buendia & C. Transy, 1996. A functional N-myc2 retroposon in ground squirrels: implications for hepadnavirus-associated carcinogenesis. Oncogene 12: 2011–2017.

Ravanat, C., M. Morales, D.O. Azorsa, S. Moog, S. Schuhler, P. Grunert, D. Loew, A. Van Dorsselaer, J.P. Cazenave & F. Lanza, 1997. Gene cloning of rat and mouse platelet glycoprotein V: identification of megakaryocyte-specific promoters and demonstration of functional thrombin cleavage. Blood 89: 3253–3262.

Rebiere, M.C., P.N. Marche & T.J. Kindt, 1987. A rabbit class I major histocompatibility complex gene with a T cell-specific expression pattern. J. Immunol. 139: 2066–2074.

Redolfi, E., C. Montagna, S. Mumm, M. Affer, L. Susani, R. Reinbold, F. Hol, P. Vezzoni, M. Cimino & I. Zucchi, 1998. Identification of CXorf1, a novel intronless gene in Xq27.3, expressed in human hippocampus. DNA Cell Biol. 17: 1009–1016.

Reinhardt, J., M. Veyhl, K. Wagner, S. Gambaryan, C. Dekel, A. Akhoundova, T. Korn & H. Koepsell, 1999. Cloning and characterization of the transport modifier RS1 from rabbit which was previously assumed to be specific for Na+-D-glucose contrasport. Biochim. Biophys. Acta. 1417: 131–143.

Reinton, N., T.B. Haugen, S. Orstavik, B.S. Skalhegg, V. Hansson, T. Jahnsen & K. Tasken, 1998. The gene encoding the C gamma catalytic subunit of cAMP-dependent protein kinase is a transcribed retroposon. Genomics 49: 290–297.

Relaix, F., X. Weng, G. Marazzi, E. Yang, N. Jenkins, S.E. Spence & D. Sassoon, 1996. *Pw1*, a novel zinc finger gene implicated in the myogenic and neuronal lineages. Dev. Biol. 177: 383–396.

Renaudie, F., L. Boulanger, B. Grandchamp & C. Beaumont, 1995. (Cloning, characterization and expression of mouse ferritin L subunit gene). C.R. Acad. Sci. III. 318: 431–437.

Renaudie, F., A.K. Yachou, B. Grandchamp, R. Jones & C. Beaumont, 1992. A second ferritin L subunit is encoded by an intronless gene in the mouse. Mamm. Genome 2: 143–149.

Retief, J.D., R.J. Winkfein & G.H. Dixon, 1993. Evolution of the monotremes. The sequences of the protamine P1 genes of platypus and echidna. Eur. J. Biochem. 218: 457-461.

Rhyner, J.A., M. Koller, I. Durussel Gerber, J.A. Cox & E.E. Strehler, 1992. Characterization of the human calmodulin-like protein expressed in *Escherichia coli*. Biochemistry 31: 12826–12832.

Roberts, R.M., L. Liu, Q. Guo, D. Leaman & J. Bixby, 1998. The evolution of the type I interferons. J. Interferon Cytokine Res. 18: 805–816.

Robertson, N.G., R.J. Pomponio, G.L. Mutter & C.C. Morton, 1991. Testis-specific expression of the human MYCL2 gene. Nucleic Acids Res. 19: 3129–3137.

Robins, D.M. & L.C. Samuelson, 1992. Retrotransposons and the evolution of mammalian gene expression. Genetica 86: 191–201.

Rothkopf, G.S., H.C. Telakowski, R.L. Stotish & C.B. Pickett, 1986. Multiplicity of glutathione S-transferase genes in the rat and association with a type 2 Alu repetitive element. Biochemistry 25: 993–1002.

Rubtsov lu, P. & A.B. Vartapetian, 1995. (New intronless members of human prothymosin alpha genes). Mol. Biol. (Mosk) 29: 1320–1325.

Ryskov, A.P., P.L. Ivanov, D.A. Kramerov & G.P. Georgiev, 1984. Universal orientation and 3′-terminal localization of repeated sequences in the *B2* family of mRNA. Mol. Biol. Mosk. 18: 92–103.

Saegusa, Y., M. Sato, I. Galli, T. Nakagawa, N. Ono, S.M. Iguchi-Ariga & H. Ariga, 1993. Stimulation of SV40 DNA replication and transcription by *Alu* family sequence. Biochim. Biophys. Acta 1172: 274–282.

Saffer, J.D. & S.J. Thurston, 1989. A negative regulatory element with properties similar to those of enhancers is contained within an *Alu* sequence. Mol. Cell Biol. 9: 355–364.

Saksela, K. & D. Baltimore, 1993. Negative regulation of immunoglobulin kappa light-chain gene transcription by a short sequence homologous to the murine *B1* repetitive element. Mol. Cell. Biol. 13: 3698–3705.

Samuelson, L.C, R.S. Phillips & L.J. Swanberg, 1996. *Amylase* gene structures in primates: retroposon insertions and promoter evolution. Mol. Biol. Evol. 13: 767–779.

Samuelson, L.C., K. Wiebauer, G. Howard, R.M. Schmid, D. Koeplin & M.H. Meisler, 1991. Isolation of the murine ribonuclease gene *Rib-1*: structure and tissue specific expression in pancreas and partoid gland. Nucleic Acids Res. 19: 6935–6941.

Samuelson, L.C., K. Wiebauer, C.M. Snow & M.H. Meisler, 1990 Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic *amylase* genes from a single gene during primate evolution. Mol. Cell. Biol. 10: 2513–2520.

Sargent, C.A., I.J. Chalmers, M. Leversha & N.A. Affara, 1994a. A rearrangement on chromosome 5 of an expressed human beta-glucuronidase pseudogene. Mamm. Genome 5: 791–796.

Sargent, C.A., C. Young, S. Marsh, S.M. Ferguson & N.A. Affara, 1994b. The glycerol kinase gene family: structure of the *Xp* gene, and related intronless retroposons. Hum. Mol. Genet. 3: 1317–1324.

Sasso, M.P., A. Carsana, E. Confalone, C. Cosi, S. Sorrentino, M. Viola, M. Palmieri, E. Russo & A. Furia, 1991. Molecular cloning of the gene encoding the bovine brain ribonuclease and its expression in different regions of the brain. Nucleic Acids Res. 19: 6469–6474.

Schlüter, G. & W. Engel, 1995. The rat Prm3 gene is an intronless member of the protamine gene cluster and is expressed in haploid male germ cells. Cytogenet. Cell Genet. 71: 352–355.

Schulte, A.M., S. Lai, A. Kurtz, F. Czubayko, A.T. Riegel & A. Wellstein, 1996. Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germline insertion of an endogenous retrovirus. Proc. Natl. Acad. Sci. USA 93: 14759–14764.

Schulte, A.M. & A. Wellstein, 1998. Structure and phylogenetic analysis of an endogenous retrovirus inserted into the human growth factor gene pleiotrophin. J Virol. 72: 6065–6072.

Schweiger, M., S.L. Oei, H. Herzog, C. Menardi, R. Schneider, B. Auer & M. Hirsch-Kauffmann, 1995. Regulation of the human poly(ADR-ribosyl) transferase promoter via alternative DNA racket structures. Biochimie 77: 480–485.

Sedlacek, Z., M.E.S. Dhorne-Pollet, C. Otto, D. Bock, G. Schütz & A. Poustka, 1999. Human and mouse XAP-5 and XAP-5-like (X5L) Genes: identification of an ancient functional retroposon differentially expressed in testis. Genomics 61: 125–132.

Selig, S., S. Bruno, J.M. Scharf, C.H. Wang, E. Vitale, T.C. Gilliam & L.M. Kunkel, 1995. Expressed cadherin pseudogenes are localized to the critical region of the spinal muscular atrophy gene. Proc. Natl Acad. Sci. USA 92: 3702–3706.

Sell, C., C.D. Chang, J. Koniecki, H.M. Chen & R. Baserga, 1992. A cryptopromoter is activated in the proliferating cell nuclear antigen gene of growth arrested cells. J. Cell. Physiol. 152: 177–184.

Shashidharan, P., T.M. Michaelidis, N.K. Robakis, A. Kresovali, J. Papamatheakis & A. Plaitakis, 1994. Novel human glutamate dehydrogenase expressed in neural and testicular tissues encoded by an X-linked intronless gene. J. Biol. Chem. 269: 16971–16976.

Shelley, C.S., E. Remold-O'Donnell, F.S. Rosen & A.S. Whitehead, 1990. Structure of the human sialophorin (CD43) gene. Identification of features atypical of genes encoding integral membrane proteins. Biochem. J. 270: 569–576.

Shimamura, M., M. Nikaido, K. Ohshima & N. Okada, 1998. A SINE that acquired a role in signal transduction during evolution. Mol. Biol. Evol. 15: 923–925.

Soares, M.B., E. Schon, A. Henderson, S.K. Karathanasis, R. Cate, S. Zeitlin, J. Chirgwin & A. Efstratiadis, 1985. RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon. Mol. Cell. Biol. 5: 2090–2103.

Soret, J., R. Gattoni, C. Guyon, A. Sureau, M. Popielarz, E. Le Rouzic, S. Dumon, F. Apiou, B. Dutrillaux, H. Voss, W. Ansorge, J. Stevenin & B. Perbal, 1998. Characterization of SRp46, a novel human SR splicing factor encoded by a PR264/SC35 retropseudogene. Mol. Cell Biol. 18: 4924–4934.

Srikantha, T., D. Landsman & M. Bustin, 1987. Retropseudogenes for human chromosomal protein HMG-17. J. Mol. Biol. 197: 405–413.

States, J.C., W. Connor, M.A. Wosnick, J.M. Aiken, L. Gedamu & G.H. Dixon, 1982. Nucleotide sequence of a protamine component *CII* gene of Salmo gairdnerii. Nucleic Acids Res. 10: 4551–4563.

Stavenhagen, J.B. & D.M. Robins, 1988. An ancient provirus has imposed androgen regulation on the adjacent mouse sex-limited protein gene. Cell 55: 247–254.

Strong, M., K.G. Chandy & G.A. Gutman, 1993. Molecular evolution of voltage-sensitive ion channel genes: on the origins of electrical excitability. Mol. Biol. Evol. 10: 221–242.

Stros, M. & G.H. Dixon, 1993. A retropseudogene for non-histone chromosomal protein HMG-1. Biochim. Biophys. Acta. 1172: 231–235.

Stros, M., J.D. Retief & G.H. Dixon, 1995. cDNA sequence and structure of a trout HMG-2 gene. Evidence for a trout- specific 3′-untranslated region. Gene 158: 181–187.

Stutz, F. & G. Spohr, 1987. A processed gene coding for a sarcomeric actin in Xenopus laevis and Xenopus tropicalis. Embo J. 6: 1989–1995.

Sugiyama, A., A. Kume, K. Nemoto, S.Y. Lee, Y. Asami, F. Nemoto, S. Nishimura & Y. Kuchino, 1989. Isolation and characterization of s-myc, a member of the rat myc gene family. Proc Natl. Acad. Sci. USA 86: 9144–9148.

Sugiyama, A., K. Noguchi, C. Kitanaka, N. Katou, F. Tashiro, T. Ono, M.C. Yoshida & Y. Kuchino, 1999. Molecular cloning and chromosomal mapping of mouse intronless myc gene acting as a potent apoptosis inducer. Gene 226: 273–283.

Sullivan, K.F. & C.A. Glass, 1991. CENP-B is a highly conserved mammalian centromere protein with homology to the helix-loop-helix family of proteins. Chromosoma 100: 360–370.

Svineng, G., R. Fassler & S. Johansson, 1998. Identification of beta1C-2, a novel variant of the integrin beta1 subunit generated by utilization of an alternative splice acceptor site in exon C. Biochem. J. 330: 1255–1263.

Takahashi, I., T. Nobukuni, H. Ohmori, M. Kobayashi, S. Tanaka, K. Ohshima, N. Okada, T. Masui, K. Hashimoto & S. Iwashita, 1998. Existence of bovine LINE repetitive insert that appears in the cDNA of bovine protein BCNT in ruminant, but not in human, genomes. Gene 211: 387–394.

Takayama, Y., C. Wada, H. Kawauchi & M. Ono, 1989. Structures of two genes coding for melanin-concentrating hormone of chum salmon. Gene 80: 65–73.

Takematsu, H., Y. Kozutsumi, A. Suzuki & T. Kawasaki, 1992. Molecular cloning of rabbit cytochrome b5 genes: evidence for the occurrence of two separate genes encoding the soluble and microsomal forms. Biochem. Biophys. Res. Commun. 185: 845–851.

Takemura, H., M. Kaku, S. Kohno, H. Tanaka, R. Yoshida, K. Ishida, R. Mizukane, H. Koga, K. Hara, T. Usui & T. Ezaki, 1996. Cloning and expression of human defensin HNP-1 genomic DNA in *Escherichia coli*. J. Microbiol. Methods 25: 287–293.

Theil, T., U. Zechner, C. Klett, S. Adolph & T. Moroy, 1994. Chromosomal localization and sequences of the murine *Brn-3* family of developmental control genes. Cytogenet. Cell Genet. 66: 267–271.

Thorey, I.S., G. Ceceña, W. Reynolds & R.G. Oshima, 1993. *Alu* sequence involvement in transcriptional insulation of the keratin 18 gene in transgenic mice. Mol. Cell. Biol. 13: 6742–6751.

Tiffany, H.L., J.S. Handen & H.F. Rosenberg, 1996. Enhanced expression of the eosinophil-derived neurotoxin ribonuclease (RNS2) gene requires interaction between the promoter and intron. J. Biol. Chem. 271: 12387–12393.

Ting, C.N., M.P. Rosenberg, C.M. Snow, L.C. Samuelson & M.H. Meisler, 1992. Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. Genes Dev. 6: 1457–1465.

Tnani, M. & B.A. Bayard, 1998. Lack of 2′,5′-oligoadenylate-dependent RNase expression in the human hepatoma cell line HepG2. Biochim. Biophys. Acta. 1402: 139–150.

Tomilin, N.V., A.S. Iguchi & H. Ariga, 1990. Transcription and replication silencer element is present within conserved region of human Alu repeats interacting with nuclear protein. Febs Lett. 263: 69–72.

Tosi, M., C. Duponchel, T. Meo & E. Couture-Tosi, 1989. Complement genes C1r and C1s feature an intronless serine protease domain closely related to haptoglobin. J. Mol. Biol. 208: 709–714.

Tsytsykova, A.V., E.N. Tsitsikov, D.A. Wright, B. Futcher & R.S. Geha, 1998. The mouse genome contains two expressed intronless retroposed pseudogenes for the sentrin/sumo-1/PIC1 conjugating enzyme Ubc9. Mol. Immunol. 35: 1057–1067.

Tucker, P.K. & B.L. Lundrigan, 1995. The nature of gene evolution of the mammalian Y chromosome: lessons from Sry. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 350: 221–227.

Uetsuki, T., K. Takagi, H. Sugiura & K. Yoshikawa, 1996. Structure and expression of the mouse necdin gene. Identification of a postmitotic neuron-restrictive core promoter. J. Biol. Chem. 271: 918–924.

Vansant, G. & W.F. Reynolds, 1995. The consensus sequence of a major Alu subfamily contains a functional retinoic acid response element. Proc. Natl. Acad. Sci. USA 92: 8229–8233.

Varghese, S. & H.M. Kronenberg, 1991. Rat thymosin beta 4 gene. Intron-containing gene and multiple retroposons. J. Biol. Chem. 266: 14256–14261.

Vidal, F., E. Mougneau, N. Glaichenhaus, P. Vaigot, M. Darmon & F. Cuzin, 1993. Coordinated posttranscriptional control of gene expression by modular elements including Alu-like repetitive sequences. Proc. Natl. Acad. Sci. USA 90: 208–212.

Vogel, T., O. Dittrich, Y. Mehraein, F. Dechend, F. Schnieders & J. Schmidtke, 1998. Murine and human TSPYL genes: novel members of the TSPY-SET-NAP1L1 family. Cytogenet. Cell Genet. 81: 265–270.

Wang, P.W., J.D. Eisenbart, S.P. Cordes, G.S. Barsh, M. Stoffel & M.M. Le Beau, 1999. Human KRML (MAFB): cDNA cloning genomic structure, and evaluation as a candidate tumor suppressor gene in myeloid leukemias. Genomics 59: 275–281.

Watkins, S.P., M.K. Jeacock, D. Savva & D.A. Shepherd, 1991. Ovine trophoblast protein-one: evidence for possible glycosylation. Int. J. Biochem. 23: 1013–1018.

Watson, C.E. & P.L. Davies, 1999. Recent and rapid amplification of the sperm basic nuclear protein genes in winter flounder. Biochim. Biophys. Acta. 1444: 337–345.

Watson, R., M. Oskarsson & G.F. Vande Woude, 1982. Human DNA sequence homologous to the transforming gene (mos) of Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA 79: 4078–4082.

Weil, D., M.A. Power, G.C. Webb & C.L. Li, 1997. Antisense transcription of a murine *FGFR-3* psuedogene during fetal development. Gene 187: 115–122.

Weller, P.A., R. Critcher, P.N. Goodfellow, J. German & N.A. Ellis, 1995. The human Y chromosome homologue of XG: transcription of a naturally truncated gene. Hum. Mol. Genet. 4: 859–868.

Wenger, R.H., M. Ayane, R. Bose, G. Kohler & P.J. Nielsen, 1991. The genes for a mouse hematopoietic differentiation marker called the heat-stable antigen. Eur. J. Immunol. 21: 1039–1046.

Wenger, R.H., N. Kieffer, A.N. Wicki & K.J. Clemetson, 1988. Structure of the human blood platelet membrane glycoprotein Ib alpha gene. Biochem. Biophys. Res. Commun. 156: 389–395.

Whitbread, L.A., K. Gregg & G.E. Rogers, 1991. The structure and expression of a gene encoding chick claw keratin. Gene 101: 223–229.

Whitmore, S.A., C. Settasatian, J. Crawford, K.M. Lower, B. McCallum, R. Seshadri, C.J. Cornelisse, E.W. Moerland, A.M. Cleton-Jansen, A.J. Tipping, C.G. Mathew, M. Savnio, A. Savoia, P. Verlander, A.D. Auerbach, C. Van Berkel, J.C. Pronk, N.A. Doggett & D.F. Callen, 1998. Characterization and screening for mutations of the growth arrest-specific *11 (GAS11)* and *C16orf3* genes at 16q24.3 in breast cancer. Genomics 52: 325–331.

Wiese, S., D.B. Murphy, A. Schlung, P. Burfeind, D. Schmundt, V. Schnulle, M.G. Mattei & U. Thies, 1995. The genes for human brain factor 1 and 2, members of the fork head gene family, are clustered on chromosome 14q. Biochim. Biophys. Acta. 1262: 105–112.

Wilkie, T.M., D.J. Gilbert, A.S. Olsen, X.N. Chen, T.T. Amatruda, J.R. Korenberg, B.J. Trask, P. de Jong, R.R. Reed, M.I. Simon et al. 1992. Evolution of the mammalian G protein alpha subunit multigene family. Nat. Genet. 1: 85–91.

Williams, M., M.S. Lyu, Y.L. Yang, E.P. Lin, R. Dunbrack, B. Birren, J. Cunningham & K. Hunter, 1999. ler5, a novel member of the slow-kinetics immediate-early genes. Genomics 55: 327–334.

Williams, S.C., C.A. Cantwell & P.F. Johnson, 1991. A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers *in vitro*. Genes Dev. 5: 1553–1567.

Wirkner, U. & W. Pyerin, 1999. CK2alpha loci in the human genome: structure and transcriptional activity. Mol. Cell Biochem. 191: 59–64.

Wu, J., G.J. Grindlay, P. Bushel, L. Mendelsohn & M. Allan, 1990. Negative regulation of the human epsilon-globin gene by transcriptional interference: role of an *Alu* repetitive element. Mol. Cell Biol. 10: 1209–1216.

Xu, G., P. O'Connell, J. Stevens & R. White, 1992. Characterization of human adenylate kinase 3 (AK3) cDNA and mapping of the *AK3* pseudogene to an intron of the *NF1* gene. Genomics 13: 537–542.

Yagi, M., S. Edelhoff, C.M. Disteche & G.J. Roth, 1995. Human platelet glycoproteins V and IX: mapping of two leucine-rich glycoprotein genes to chromosome 3 and analysis of structures. Biochemistry 34: 16132–16137.

Yang, G.C., N. Kunze, B. Baumgartner, Z.Y. Jiang, M. Sapp, R. Knippers & A. Richter, 1990. Molecular structures of two human DNA topoisomerase I retrosequences. Gene 91:247–253.

Yang, Z., D. Boffelli, N. Boonmark, K. Schwartz & R. Lawn, 1998. Apolipoprotein(a) gene enhancer resides within a *LINE* element. J. Biol. Chem. 273: 891–897.

Yasumasu, S., H. Shimada, K. Inohaya, K. Yamazaki, I. Iuchi, I. Yasumasu & K. Yamagami, 1996. Different exon-intron organizations of the genes for two astacin-like proteases, high choriolytic enzyme (choriolysin H) and low choriolytic enzyme (choriolysin L), the constituents of the fish hatching enzyme. Eur. J. Biochem. 237: 752–758.

Yaswen, P., A. Smoll, J. Hosoda, G. Parry & M.R. Stampfer, 1992. Protein product of a human intronless calmodulin-like gene shows tissue-specific expression and reduced abundance in transformed cells. Cell Growth Differ. 3: 335–345.

Yoshida, Y., S. Matsuda & T. Yamamoto, 1997. Cloning and characterization of the mouse tob gene. Gene 191: 109–113.

Yoshimura, Y., H. Tanaka, M. Nozaki, K. Yomogida, K. Shimamura, T. Yasunaga & Y. Nishimune, 1999. Genomic analysis of male germ cell-specific actin capping protein alpha. Gene 237: 193–199.

Youssouffian, H. & H. Lodish, 1993. Transcriptional inhibition of the murine erythroprotien receptor gene by an upstream repetitive element. Mol. Cell Biol. 13: 98–104.

Zaiβ, D.M., & P.M. Kloetzel, 1999. A second gene encoding the mouse proteasome activator PA28beta subunit is part of a *LINE1* element and is driven by a *LINE1* promoter. J. Mol. Biol. 287: 829–835.

Zakeri, Z.F., D.J. Wolgemuth & C.R. Hunt, 1988. Identification and sequence analysis of a new member of the mouse *HSP70* gene family and characterization of its unique cellular and developmental pattern of expression in the male germ line. Mol. Cell Biol. 8: 2925–2932.

Zhend, H., D. Bhavsar, I. Dugast, E. Zappone & J. Drysdale, 1997. Conserved mutations in human ferritin *H* pseudogenes: a second functional sequence or an evolutionary quirk? Biochim. Biophys. Acta 1351: 150–156.

Zheng, H.D., D. Bhavsar & J. Drysdale, 1995. An unusual human ferritin H sequence from chromosome 4. DNA Seq. 5: 173–175.

Zheng, J.H., N.-S. Shunnosuke, M. Takahashi & M. Nonaka, 1992. Insertion of the B2 sequence into intron 13 is the only defect of the *H-2^k C4* gene which causes low C4 production. Nucleic Acids Res. 20: 4975–4979.

References

Baltimore, D., 1985. Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. Cell 40: 481–482.

Becker, Y., 1996. A short introduction to the origin and molecular evolution of viruses. Virus Genes. 11: 73–77.

Blackburn, E.H., 1992. Telomerases. Annu. Rev. Biochem. 61: 113–129.

Britten, R.J., 1996. DNA sequence insertion and evolutionary variation in gene regulation. Proc. Natl. Acad. Sci. USA 93: 9374-9377.

Britten, R.J., 1997. Mobile elements inserted in the distant past have taken on important functions. Gene 205: 177–182.

Britten, R.J., W.F. Baron, D.B. Stout & E.H. Davidson, 1988. Sources and evolution of human Alu repeated sequences. Proc. Natl. Acad. Sci. USA 85: 4770–4774.

Brookfield, J.F., 1994. The human Alu SINE sequences – is there a role for selection in their evolution? Bioessays 16: 793–795.

Brosius, J., 1991. Retroposons – seeds of evolution. Science 251: 753.

Brosius, J., 1999a. Many G-protein-coupled receptors are encoded by retrogenes. Trends Genet. 15: 304–305.

Brosius, J., 1999b. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. Gene 238: 115–134.

Brosius, J., 1999c. Transmutation of tRNA over time. Nat. Genet. 22: 8–9.

Brosius, J. & S.J. Gould., 1992. On 'genomenclature': a comprehensive (and respectful) taxonomy for pseudogenes and other 'junk DNA'. Proc. Natl. Acad. Sci. USA 89: 10706–10710.

Brosius, J. & S.J. Gould, 1993. Molecular constructivity. Nature 365: 102.

Brosius, J., M.L. Palmer, P.J. Kennedy & H.F. Noller, 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. USA 75: 4801–4805.

Brosius, J. & H. Tiedge, 1995. Neural BC1 RNA: Dendritic localization and transport, pp. 289–330, in Localized RNAs, edited by H.D. Lipshitz, R:G: Landes, Austin, TX.

Brosius, J. & H. Tiedge, 1996. Reverse transcriptase – mediator of genomic plasticity. Virus Genes 11: 163–179.

Cavalier-Smith, T., 1991. Intron phylogeny: a new hypothesis. Trends Genet. 7: 145–148.

Darnell, J.E., & W.F. Doolittle, 1986. Speculations on the early course of evolution. Proc. Natl. Acad. Sci. USA 83: 1271–1275.

Deininger, P.L. & M.A. Batzer, 1999. Alu repeats and human disease. Mol. Genet. Metab. 67: 183–193.

Deininger, P.L., M.A. Batzer, C.A. Hutchison, III & M.H. Edgell, 1992. Master genes in mammalian repetitive DNA amplification. Trends Genet. 8: 307–311.

Deininger, P.L. & V.K. Slagel, 1988. Recently amplified Alu family members share a common parental Alu sequence. Mol. Cell. Biol. 8: 4566–4569.

Deininger, P.L., H. Tiedge, J. Kim & J. Brosius. 1996. Evolution, expression, and possible function of a master gene for amplification of an interspersed repeated DNA family in rodents, edited by W.E. Cohn and K. Moldave. Progr. Nucleic. Acids Res. 52: 67–88.

Diamond, J., 1997. Guns, Germs, and Steel: The Fates of Human Societies. W.W. Norton, New York.

Doolittle, R.F. & D.F. Feng, 1992. Tracing the origin of retroviruses. Curr. Top. Microbiol. Immunol. 176: 195–211.

Doolittle, W.F. & C. Sapienza, 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284: 601–603.

Doolittle, W.F., 1989. Hierarchical approaches to genome evolution. Can. J. Phil. 14 Suppl.: 101–133.

Dunham, I., N. Shimizu, B.A. Roe, S. Chissoe, A.R. Hunt, J.E. Collins, R. Bruskiewich, D.M. Beare, M. Clamp, L.J. Smink, R. Ainscough, J.P. Almeida, A. Babbage, C. Bagguley, J. Bailey, K. Barlow, K.N. Bates, O. Beasley, C.P. Bird, S. Blakey, A.M. Bridgeman, D. Buck, J. Burgess, W.D. Burrill & K.P. O'Brien, 1999. The DNA sequence of human chromosome 22. Nature 402: 489–495.

Feng, Q., J.V. Moran, H.H. Kazazian, Jr. & J.D. Boeke, 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87: 905-916.

Flavell, A.J., 1995. Retroelements, reverse transcriptase and evolution. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 110: 3–15.

Georgiev, G.P., 1984. Mobile genetic elements in animal cells and their biological significance. Eur. J. Biochem. 145: 203–220.

Gilbert, W., 1978. Why genes in pieces? Nature 271: 501.

Gould, S.J. & E.S. Vrba, 1982. Exaptation – a missing term in the science of form. Paleobiology 8: 4–15.

Hickey, D.A., 1982. Selfish DNA: a sexually-transmitted nuclear parasite. Genetics 101: 519–531.

Jacob, F., 1982. The Possible and the Actual. Pantheon Books, New York.

Jordan, I.K., L.V. Matyunina & J.F. McDonald, 1999. Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. Proc. Natl. Acad. Sci. USA 96: 12621–12625.

Jurka, J., 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc. Natl. Acad. Sci. USA 94: 1872–1877.

Jurka, J., 1998. Repeats in genomic DNA mining and meaning. Curr. Opin. Struct. Biol. 8: 333–337.

Jurka, J. & A. Milosavljevic, 1991. Reconstruction and analysis of human Alu genes. J. Mol. Evol. 32: 105–121.

Kermekchiev, M., M. Pettersson, P. Matthias & W. Schaffner, 1991. Every enhancer works with every promoter for all the combinations tested: could new regulatory pathways evolve by enhancer shuffling? Gene Expr. 1: 71–81.

Kidwell, M.G. & D. Lisch, 1997. Transposable elements as sources of variation in animals and plants. Proc. Natl. Acad. Sci. USA 94: 7704–7711.

Kim, J., J.A. Martignetti, M.R. Shen, J. Brosius & P. Deininger, 1994. Rodent BC1 RNA gene as a master gene for ID element amplification. Proc. Natl. Acad. Sci. USA 91: 3607–3611.

Koch, A.L., 1972. Enzyme evolution: I. The importance of untranslatable intermediates. Genetics 72: 297–316.

Kolosha, V.O. & S.L. Martin, 1995. Polymorphic sequences encoding the first open reading frame protein from LINE-1 ribonucleoprotein particles. J. Biol. Chem. 270: 2868–2873.

Labuda, D. & G. Striker, 1989. Sequence conservation in Alu evolution. Nucleic Acids Res. 17: 2477–2491.

Makalowski, W., G.A. Mitchell & D. Labuda, 1994. Alu sequences in the coding regions of mRNA: a source of protein variability. Trends Genet. 10: 188–193.

McClintock, B., 1948. Mutable loci in maize. Carnegie Inst. Wash. Yearbook 47: 155–169.

McDonald, J.F., 1990. Macroevolution and retroviral elements. Bioscience 40: 183–191.

McDonald, J.F., 1993. Evolution and consequences of transposable elements. Curr. Opin. Genet. Dev. 3: 855–864.

McDonald, J.F., 1995. Transposable elements: possible catalysts of organismic evolution. Trends Ecol. Evol. 10: 123–126.

Nouvel, P., 1994. The mammalian genome shaping activity of reverse transcriptase. Genetica 93: 191–201.

Ohshima, K., M. Hamada, Y. Terai & N. Okada, 1996. The 3′ ends of tRNA-derived short interspersed repetitive elements are derived from the 3′ ends of long interspersed repetitive elements. Mol. Cell. Biol. 16: 3756–3764.

Okada, N., M. Hamada, I. Ogiwara & K. Ohshima, 1997. SINEs and LINEs share common 3′ sequences: a review. Gene 205: 229–243.

Orgel, L.E. & Crick, F.H.C., 1980. Selfish DNA: the ultimate parasite. Nature 284: 604–607.

Pardue, M.L., O.N. Danilevskaya, K. Lowenhaupt, F. Slot & K.L. Traverse, 1996. Drosophila telomeres: new views on chromosome evolution. Trends Genet. 12: 48–52.

Pathak, V.K., and W.S. Hu, 1997. 'Might as well jump!' – Template switching by retroviral transcriptase, defective genome information, and recombination. Semin. Virol. 8: 141–150.

Patience, C., D.A. Wilkinson & R.A. Weiss, 1997. Our retroviral heritage. Trends Genet. 13: 116–120.

Poole, A., D. Jeffares & D. Penny, 1999. Early evolution: prokaryotes, the new kids on the block. Bioessays 21: 880–889.

Quentin, Y., 1988. The Alu family developed through successive waves of fixation closely connected with primate lineage history. J. Mol. Evol. 27: 194–202.

Rose, M.R. & W.F. Doolittle, 1983. Molecular biological mechanisms of speciation. Science 220: 157–162.

Schmid, C.W., 1998. Does SINE evolution preclude Alu function? Nucleic Acids Res. 26: 4541–4550.

Shapiro, J.A., 1992. Natural genetic engineering in evolution. Genetica 86: 99–111.

Singer, M.F., 1995. Unusual reverse transcriptases. J. Biol. Chem. 270: 24623–24626.

Tomilin, N.V., 1999. Control of genes by mammalian retroposons. Int. Rev. Cytol. 186: 1–48.

Walsh, C.P., J.R. Chaillet & T.H. Bestor, 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat. Genet. 20: 116–117.

Wessler, S.R., T.E. Bureau & S.E. White, 1995. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. Curr. Opin. Genet. Dev. 5: 814–821.

Willard, C., H.T. Nguyen & C.W. Schmid, 1987. Existence of at least three distinct Alu subfamilies. J. Mol. Evol. 26: 180–186.

Woese, C.R., R. Gutell, R. Gupta & H.F. Noller, 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol Rev. 47: 621–669.

Wolffe, A.P. & M.A. Matzke, 1999. Epigenetics: regulation through repression. Science 286: 481–486.