Nucleic Acids Research, 1998, Vol. 26, No. 20 **4541–4550**

SURVEY AND SUMMARY

Does SINE evolution preclude Alu function?

Carl W. Schmid*

Section of Molecular and Cellular Biology and Department of Chemistry, University of California at Davis, Davis, CA 95616, USA

Received May 6, 1998; Revised and Accepted July 1, 1998

ABSTRACT

The evolution, mobility and deleterious genetic effects of human Alus are fairly well understood. The complexity of regulated transcriptional expression of Alus is becoming apparent and insight into the mechanism of retrotransposition is emerging. Unresolved questions concern why mobile, highly repetitive short interspersed elements (SINEs) have been tolerated throughout evolution and why and how families of such sequences are periodically replaced. Either certain SINEs are more successful genomic parasites or positive selection drives their relative success and genomic maintenance. A complete understanding of the evolutionary dynamics and significance of SINEs requires determining whether or not they have a function(s). Recent evidence suggests two possibilities, one concerning DNA and the other RNA. Dispersed Alus exhibit remarkable tissue-specific differences in the level of their 5-methylcytosine content. Differences in Alu methylation in the male and female germlines suggest that Alu DNA may be involved in either the unique chromatin organization of sperm or signaling events in the early embryo. Alu RNA is increased by cellular insults and stimulates protein synthesis by inhibiting PKR, the eIF2 kinase that is regulated by double-stranded RNA. PKR serves other roles potentially linking Alu RNA to a variety of vital cell functions. Since Alus have appeared only recently within the primate lineage, this proposal provokes the challenging question of how Alu RNA could have possibly assumed a significant role in cell physiology.

1. CONTINUOUS SUCCESSION OF DISTINCT SINES

Mammalian DNAs typically contain hundreds of thousands of copies of short interspersed repeated sequences called SINEs (1). The number of SINEs that are fixed in the mammalian genome is all the more remarkable when it is recognized that these sequences transposed into their genomic loci through RNA intermediates (retrotransposition). Thus SINEs must have been a tremendous source of insertional mutagenesis throughout mammalian evolution. Given their abundance and mobility, evolutionary considerations have naturally dominated research on SINEs. Results from those studies provide the starting point for considering other aspects of SINEs, including their possible functionality. Excellent reviews, including a recent monograph edited by Maraia, document generally accepted background information (Sections 1–5).

The most extensively studied mammalian SINE, human Alu, exemplifies most features of this unusual class of sequences. There are nearly 1 000 000 Alus per haploid genome (1), corresponding to an average genomic spacing of 3 kb. Individual Alus share a 282 nt consensus sequence which is typically followed by a 3′ A-rich region resembling a poly(A) tail (Fig. 1). The Alu consensus sequence is a divergent tandem dimer in which the two monomer units are separated by a short A-rich region, a vestige of what must have been a 3′ A-rich region that flanked the ancestral monomer (1,2; Fig. 1). Except for a 30 nt insertion in the right monomer, Alu monomers are homologous to SRP RNA, also known as 7SL RNA (Fig. 1). Most Alus are flanked by short direct repeats which are the duplicated insertion site (1; Fig. 1).

Except for rodents and primates, SINEs in all other animals examined are unrelated to SRP RNA but are instead homologous to tRNAs; even plants contain tRNA SINEs, indicating that the earliest eukaryotic SINEs must have been derived from tRNAs (1,3; Fig. 2). Moreover, all highly repetitive eukaryotic SINEs belong to either the SRP or a tRNA superfamily. (Different tRNA superfamilies are not distinguished here; 3.) Rodents contain both SRP RNA and tRNA related SINEs, usually called B1 and B2 repeats respectively (1,4; Fig. 2). Rodent B1 repeats essentially resemble the left human Alu monomer (Figs 1 and 2). Prosimian SINEs include full-length dimeric Alus, B1-like Alu monomers, B2-like/tRNA SINEs and composite elements consisting of both SINE superfamilies. This intermediate composition suggests a transition between the SINEs in rodents and higher primates (1; Fig. 2). Sequence analysis indicates that rodent B1 and primate Alu repeats are ultimately derived from a single founder (5). As discussed later (Section 4), human Alu subfamilies also result from individual founders. The question of whether the very earliest tRNA SINEs might have been rooted in a single primordial ancestor (Fig. 2) remains to be answered (Section 11).

The most recent common ancestor of human and rodent must have contained tRNA SINEs (Fig. 2). Decrepit, fossil tRNA

*Tel: +1 530 752 3003; Fax: +1 530 752 3085; Email: cwschmid@ucdavis.edu

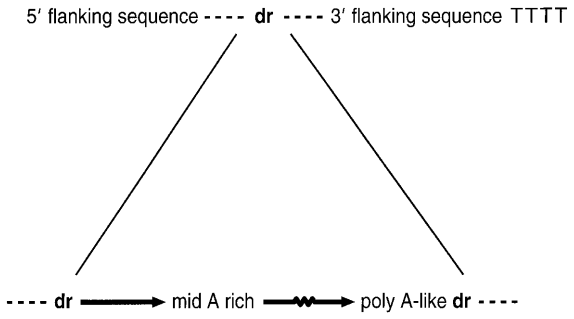

Figure 1. Consensus Alu structure (1). Direct repeats (dr) flanking an Alu result from duplication of its empty genomic insertion site. Similarly, 5′ and 3′ sequences flanking an Alu are contributed by the unique genomic locus in which it resides. Four or more T residues are sufficient to terminate pol III-directed transcription so that termination and length of the primary Alu transcript are determined by its unique 3′ flanking sequence. The consensus Alu sequence is usually followed by an A-rich region resembling a poly(A) tail. As depicted by the two solid arrows, the 282 nt Alu consensus sequence consists of an inexact duplication of two monomer units which are homologous to SRP RNA. These two monomers are separated by a mid A-rich region and, as depicted by the wavy line, the right monomer contains an additional sequence (~30 nt) that is absent in the left monomer. Rodent B1 repeats essentially resemble left Alu monomers.

SINEs, called MIRs, that pre-date rodent–human divergence are buried in human DNA (1,6,7; Fig. 2). Sequence database searches, hybridization analysis and library screening with rodent B2 probes have failed to identify other tRNA SINEs in human DNA (6–8; unpublished results). Thus all available evidence, albeit negative, indicates that the previously successful mammalian tRNA SINEs are now either extinct or severely reduced in copy number within the human genome.

The reasons why Alus flourished while tRNA SINEs died in the higher primate genome are unknown. Perhaps Alus are merely better genomic parasites than tRNA SINEs since ‘No cellular function...is required to explain...the behavior or persistence of middle repetitive sequences as a class’ (9–11). However, there is no *a priori* reason to dismiss the possibility that Alus provide a selective advantage to their host which drives their retrotranspositional success. While some evidence suggests that Alus may serve one or more functions (Sections 8, 10 and 11), the two explanations are not exclusive, since a successful parasite optimizes its requirements with those of the host. As extraordinarily successful genomic symbionts, Alus may have established a state of nearly complete neutrality, a ‘genomic peace’ (Section 3), or may instead compensate their host with selective advantages. In the extreme, Alus could serve a vital function that precedes their genomic proliferation. Competing themes throughout this review are how either the retrotransposition pathway or the host’s requirements might select for SINEs.

2. SINE FAMILIES GROW BY ACCUMULATING NEW MEMBERS

To understand the dynamics of SINE evolution, the fate of human Alus has been traced by comparisons with their orthologs in other primates (1). (Primate phylogeny is qualitatively depicted in Fig. 2 to follow this and subsequent discussions.) These comparisons, which emphasize Alus mapping within globin gene clusters, indicate that the great majority of human Alus post-date the

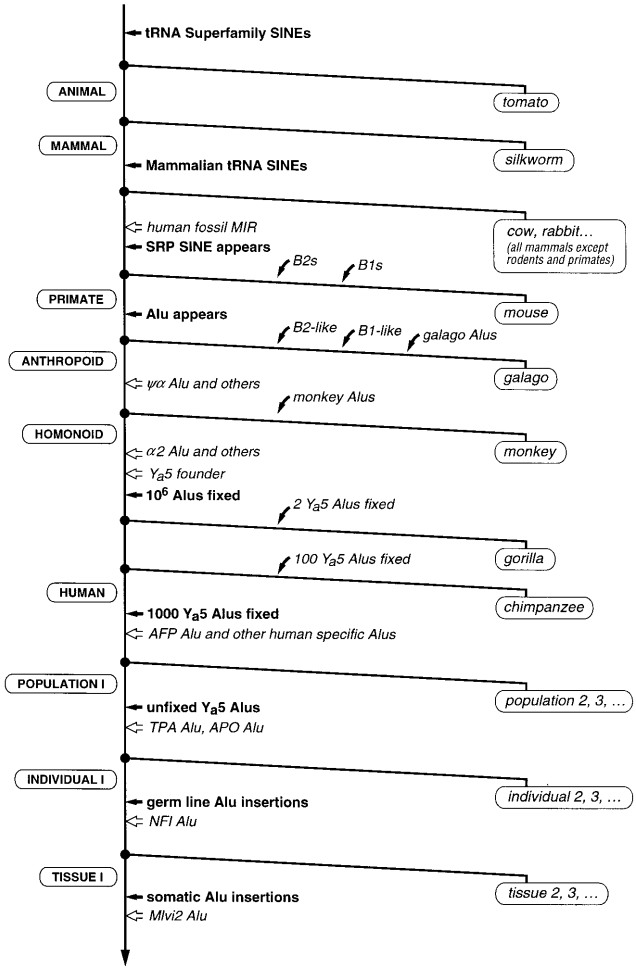

Figure 2. SINE phylogeny and accumulation within a human DNA (1). The straight time line on the left qualitatively traces the evolutionary history and sequence of events by which human Alus and SINEs accumulated in the DNA from a particular tissue from a particular person. Solid arrows with names in bold indicate the times at which SINE families appeared, and open arrows with names in italic indicate particular human SINEs as inferred from phylogenetic comparisons. Branching compares either definite species or hypothetical populations, individuals and tissues as outgroups (lower case) to trace the lineage of accumulated SINEs and human Alus within a particular tissue from a particular person (upper case). Although most of the branchings depicted have been established by comparing orthologous loci in outgroups, common ancestry has not been established for inheritance of either the tRNA superfamily in plants and animals or the SRP superfamily in rodent and primate. In principle either of these two events could have resulted from convergent evolution and not common inheritance, as assumed for simplicity in this tree. Somatic insertion of the Mlvi-2 Alu has not been proven.

divergence of the prosimian lineage but pre-date chimpanzee divergence (Fig. 2). Some older Alus even pre-date human–monkey divergence (Fig. 2). Alus are rarely subject to sequence conversion (12) and, except for the special case of CpG dinucleotides, accumulate point mutations at the rate expected for unselected DNA sequences (1,2). DNA methylation accelerates the mutation rate of Alu CpG dinucleotides (1,2,13,14; Sections 6 and 7). In summary, phylogenetic comparisons generally indicate that Alus are rather immobile, are stably inherited over relatively long evolutionary times and, like the fossil SINE example, are eventually obliterated by accumulated point mutations (Fig. 2).

As one exception to this otherwise stable inheritance, unequal Alu–Alu recombination occasionally prunes Alus from the genome (Section 3).

3. ALUS HAVE NOT ESTABLISHED GENOMIC PEACE

In three documented cases, *de novo* germline retrotransposition of Alus within genes has caused inherited disease phenotypes (1,15; Fig. 2). As an important side point, *de novo* Alu insertions are not restricted to the germline and cause somatic mutations (16; Fig. 2). Presumably, the 1 000 000 Alus that are fixed within the population represent the fraction of insertions that were neutral or at least tolerable. However, even fixed Alus have altered gene expression. Alus and other mammalian SINEs influence the transcription of neighboring genes, direct polyadenylation of the resulting mRNA and Alu fragments are spliced into mature mRNAs mutating the protein products (1,15,17,18). Positive adaptations involving individual Alus can be best viewed as a later consequence rather than as a cause of their presence (9–11).

Merely transmitting SINEs imposes a genetic load. Because of their ubiquitous genomic distributions, SINEs present a large target for homologous unequal recombination which, by duplicating and deleting sequences or by scrambling non-homologous chromosomes, also causes disease phenotypes. Currently, 30 examples of inherited genetic diseases result from ectopic Alu–Alu recombination (1,15). Unequal Alu–Alu pairing would erase flanking direct repeats in the recombinant (Fig. 1). Since the majority (~80%) of Alus have direct repeats, most Alus have not ectopically recombined, despite their very long residence within the genome (1; Fig. 2). Alu–Alu recombination might be suppressed by their sequence divergence (19) or negative selection might reduce fixation of the unequal recombinants (1,15). The divergence (~20%) of garden variety Alus (1) would be sufficient to suppress recombination (19).

The possibility that selfish DNAs might stimulate recombination should not be confused with function (9–11). Moreover, Alus evidently fail to satisfy the ideal postulated for the most successful selfish DNA sequences (9,10); they are not neutral parasites as their insertion, presence and maintenance cause disease phenotypes.

4. SINE FAMILIES RESULT FROM ANCESTRAL FOUNDERS

Human Alus also belong to distinct subfamilies of different evolutionary ages (1,2,20,21). For example, the consensus sequence of the young Ya5 subfamily differs from that of the next older Y subfamily by five concerted mutations. Almost all of the 1000 members of the human Ya5 subfamily integrated following the divergence of human from chimpanzee (1,20; Fig. 2). As expected for a recently active subfamily, some Ya5 Alus are not fixed in the human population but segregate according to pedigrees (22–24; Fig. 2). Furthermore, the Ya5 subfamily as well as at least two other subfamilies remain retrotranspositionally active within the contemporary human lineage (1,15; Fig. 2). The activity of several distinct subfamilies indicates the corresponding presence of multiple Alu source genes (25). As further evidence for their youth, very young Alus almost exactly match their subfamily consensus sequences (1,2; Section 7).

The appearance and expansion of young Alu subfamilies, such as Ya5, continues a consistent theme in SINE evolution, the constant succession of sequence families and subfamilies, raising the question of how new SINE families appear (Fig. 2). The following shows that the Ya5 subfamily stems from a single founder.

The presence of 100 Ya5 Alus in chimpanzee indicates that this subfamily's founder pre-dates human–chimpanzee divergence (1,25,26; Fig. 2). Fortuitously, there are only two Ya5 Alus in gorilla, providing a tight phylogenetic bottleneck to isolate the founder (27; Fig. 2). One of these is gorilla specific but the other corresponds to human and chimpanzee orthologs (27; Fig. 2). Further phylogenetic comparisons show that this presumptive Ya5 founder was encoded by a source gene for the older Y subfamily and, by drift, acquired the diagnostic base substitutions that identify its progeny (27; Fig. 2). Sequences flanking this Alu founder stimulate its transcription, suggesting a plausible explanation for its relative success and the appearance of a new SINE subfamily (28).

However, the Ya5 founder subsequently acquired additional mutations compared with the most recently inserted Ya5 subfamily members (1,27). Thus this Alu founder is not the immediate parent of its progeny but rather a grandparent. This interpretation is confirmed by the later appearance and simultaneous activity of the Ya8 subfamily, a descendent of the human Ya5 subfamily (1,2). Perhaps several of the founder's progeny inserted into even more fortuitous loci in human, making the human subfamily more successful than those in gorilla and chimpanzee. Since human, gorilla and chimpanzee all share the same Ya5 Alu founder, the great differences in this subfamily's expansion in these three species show that retrotranspositional success is not deterministic. Retrotransposition is a multistep process so that the relative success of different Alus stochastically depends upon the simultaneous occurrence of several events, as described in the following section.

5. FACTORS DETERMINING RETROTRANSPOSITIONAL SUCCESS

Insertion of new Alus within the inherited genome requires: (i) their germline transcription; (ii) post-transcriptional availability of the resulting RNA; (iii) reverse transcriptase and other factors necessary for inserting the resulting cDNA. Retrotranspositional success depends upon satisfying each of these requirements (20).

The transcriptional activity of IAP retrotransposons highlights an essential concept. An IAP source gene is expressed only at a particular stage in spermatogenesis and, conversely, somatic transcription of other IAPs is irrelevant to retrotransposition (29). Unfortunately, our current understanding of Alu transcription depends upon studies using cultured somatic cells (Section 8).

Primary polymerase (pol) III-directed Alu transcripts, referred to as full-length (fl) Alu RNA, are short lived (~20 min) cytoplasmic transcripts but some are processed into a stable small cytoplasmic (sc) Alu RNA (1,30,31). Similarly, B1 RNA is processed into scB1 RNA. scAlu/B1 RNA formation probably competes with retrotransposition (31). This processing is modulated by the interaction of La protein with very subtle sequence features in the pol III terminator which results from the 3′ sequences flanking the Alu/B1 insertion site (32,33; Fig. 1). The stability of flAlu RNA is sensitive to the binding of SRP 9/14, which depends upon the Alu subfamily identity, and there is also a tight correlation between scAlu sequence conservation and the interaction with SRP 9/14 (34; Section 11). Alu sequences and their SRP binding proteins evidently co-evolved (34,35). Both the subfamily identity and divergence of an Alu further determine its RNA

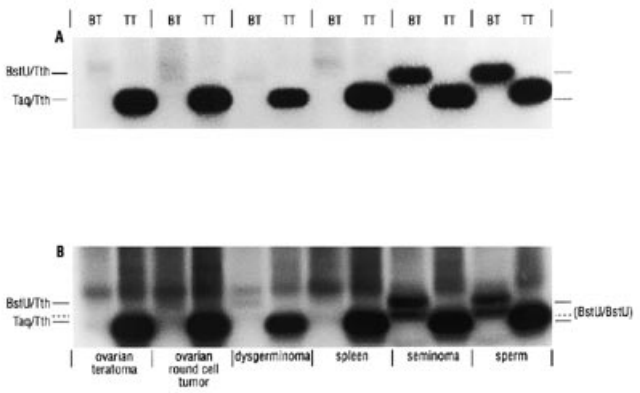

Figure 3. Tissue-specific Alu methylation (45). Human DNAs from indicated tissues were digested with either *Bst*UI/*Tth*111I (labeled BT) or *Taq*I/*Tth*111I (labeled TT) and hybridized under conditions that are specific for either the young Ya5 subfamily (**A**) or older Ya subfamily (**B**). *Bst*UI in the BT digest cleaves only its unmethylated site and, for comparison, the TT digest provides a measure of total DNA. Note that the *Bst*UI site is largely uncut (methylated) in a somatic tissue, spleen, and in a female germline tissue, dysgerminoma, but is mostly cut (unmethylated) in Ya5 Alus from male germline tissues, sperm and seminoma.

folding, which is likely to affect the transcript's lifetime as a potential intermediate in retrotransposition (36,37). In summary, a complex interplay of subfamily identity, divergence, accidents of its 3′ flanking sequence and RNA-binding proteins all probably contribute post-transcriptionally to the relative success of different Alus.

Repetitive LINE retrotransposons encode both a reverse transcriptase and an endonuclease that are required for their autonomous retrotransposition (38,39); SINEs do not. The following suggests that repetitive LINEs probably contribute these factors to SINEs: tRNA superfamily SINEs are a composite of a 5′ tRNA-related sequence and 3′ non-tRNA sequences (40,41). The homology of these 3′ SINE sequences to LINEs suggests that tRNA SINEs resulted from priming the reverse transcription of a LINE template (40,41). This model has been extended to the SRP superfamily of SINEs. Alus and LINEs share consensus flanking direct repeats (42). These flanking direct repeats are generated by the LINE endonuclease which cleaves the genomic insertion site to prime the RNA intermediate for reverse transcription (39,42). Thus, considering both the tRNA and SRP RNA SINE superfamilies together, the proposed association of SINE retrotransposition with LINEs neatly provides a source of reverse transcriptase, a mechanism for priming SINE cDNA synthesis and an activity for inserting the resulting cDNA product (39,42). The pairing of SINEs with LINEs further predicts that the loss of a LINE family might cause the parallel extinction of its cognate SINE family (41; Fig. 2).

Surprisingly, the LINE1 reverse transcriptase as well as the reverse transcriptases encoded by some other retrotransposable elements function best when acting *in cis* (43). This introduces a further requirement of delivering the SINE RNA intermediate to the site at which the reverse transcriptase is synthesized. The association of flAlu RNA with SRP 9/14 provides an attractive mechanism for presenting Alu RNA to freshly synthesized reverse transcriptase at the ribosome (43). Furthermore, the poly(A) tail of LINE1 evidently interacts with a LINE1-encoded factor, suggesting a role for this A-rich sequence in retrotansposition and, by implication, the A-rich region associated with Alu repeats (43). Thus this model also potentially explains the presence of a poly(A)-like tail on pol III-directed Alu transcripts. These considerations suggest that retrotranspositionally successful Alus will be selected for both the presence of a suitable A-rich region and for a folded RNA structure which favors binding of SRP proteins. One potential weakness of this proposed model is that tRNA SINEs are not expected to bind SRP proteins. However, this criticism is not insurmountable as other plausible mechanisms might deliver tRNA-related SINEs to a ribosome.

The possible involvement of SRP proteins and cognate LINE pairing in SINE retrotransposition are not exclusive models and might act in concert to select for successful elements (41,43). Moreover, the many conserved features that are associated with retrotranspositionally active SINEs are likely the result of compounding multiple levels of selection.

The proposed mechanism for SINE retrotransposition requires the simultaneous germline expression of SINEs and LINEs. SINE transcriptional expression is inhibited by DNA methylation (Section 8), so that demethylation of an Alu subset in the male germline (Section 6) probably derepresses expression of Alu source genes.

6. METHYLATION—HOW ALUS MIGHT FUNCTION AS DNA

Depending upon the tissue, slightly less than 1% of human DNA consists of 5-methylcytosine (5-meC) residues, which are overwhelmingly contained within CpG dinucleotides (reviewed in 1,44,45). The CpG content of young Alus (9%) is 9-fold greater than that of human DNA, so that Alus, which constitute nearly 10% of human DNA, account for a substantial fraction (~33%) of the genome's potential methylation sites (1,2,13,14). The rapid replacement of consensus Alu CpGs by TpGs (and equivalently its complement CpA) indicates their germline methylation (1,2,13,14) and Alu CpGs are almost totally methylated in somatic tissues such as spleen (44,45; Fig. 3).

The insertion of foreign DNA sequences, including mouse B1 elements, causes changes in the methylation of sequences flanking the insertion site (46,47). Presumably, young Alus might also act *in cis* to direct the methylation of their immediate flanking regions. If so, evolutionarily recent Alu insertions could lead to significant differences in the DNA methylation of related species, such as chimpanzee and human (26). Changes that occur in the germline methylation of Alus are discussed below. These observations coupled with the widespread belief that DNA methylation has important roles in embryonic development and that developmental differences are more important to speciation than mutations in genes raise the intriguing possibility that Alu dimorphisms might contribute to speciation (discussed in 26).

Most members of a young Alu subfamily are completely unmethylated in mature sperm and earlier stages of spermatogenesis (45,48,49; Fig. 3). Despite this subgroup's complete demethylation, some young Alus and a majority of older Alus are partially or completely methylated in sperm. The unknown rules governing sperm Alu methylation are probably enforced by an Alu-binding protein (45,48–50).

Despite the existence of this unmethylated subgroup in sperm, the rapid decay of Alu CpGs to TpGs still requires that the vast majority are methylated in the germline (1,2,13,14). [As a separate issue, CpG dinucleotides in recently inserted Alus are mostly intact (Section 7).] Satisfying this requirement, Alus in

ovarian dysgerminoma (Fig. 3) and oocytes are mostly methylated (45). But this observation also means that embryos inherit entirely different maternal and paternal Alu methylation patterns (45).

This difference in inherited Alu methylation patterns could be required either early in fertilization or late in sperm maturation. Genomic imprinting (differential expression of the parental genomes) requires that the parental genomes can be distinguished in early embryos, which might be accomplished by their differential methylation (51). Because Alus are ubiquitously distributed throughout human DNA, their differential methylation in sperm and oocyte has been previously recognized to make them ideally suited to signal imprinting (1,45). Alternatively, sperm chromatin is entirely unlike that of any other cell type. Approximately 85% of sperm DNA is organized as nucleoprotamine and 15% as nucleohistone (52,53). Even the nucleohistone composition in sperm differs markedly from that in somatic cells: sperm histones are hyperacetylated (53). Furthermore, the sperm genome is compartmentalized in a sequence-specific manner between these two very different chromatins (52,54). Histone deacetylation is mediated by DNA methylation (55), possibly linking Alu hypomethylation, sperm histone hyperacteylation and the unusual chromatin organization in sperm. Demethylated Alus might direct sperm histone hyperacetylation so that differences in Alu methylation might ultimately direct the sequence-specific packing of DNA within the two types of sperm chromatin. This suggestion implies a possible correlation between the compartmentalization of single copy sequences within these two types of sperm chromatin and the density and methylation status of their neighboring Alu elements. Alternatively, the structure of chromatin in the developing male germline may direct Alu retroposition. Sperm maturation apparently involves a series of modified histones that serve as transition proteins during chromatin repackaging (53). Conceivably, differences caused by these alterations in histones during sperm chromatin development could open certain regions for Alu insertion and, consequently, the germline inheritance of Alus within the affected chromatin subdomains. [As a note of caution before further considering these issues, the sequence-specific packing of DNA in sperm chromatin has only been examined in human (52,54) and the demethylation of a subset of sperm SINEs has only been reported in human and monkey (45,48,49). Conceivably, either of these two observations might be peculiar to primates. Determining the sequence-specific packing of sperm chromatin and the methylation status of SINEs in sperm DNA of other mammals would provide an immediate, evolutionary test of whether these phenomenon have more general significance.]

Each of the two suggested roles for hypomethylation of sperm Alus entertains a DNA function that requires the ubiquitous genomic distribution of a family of sequences. Except for its CpG content, the requirements for this sequence might be so lax that unrelated SINEs could be functionally equivalent. A major difficulty with this suggestion is the further implication that Alus must have assumed or inherited one or the other of these suggested roles from earlier tRNA SINEs. As discussed above, alterations in the structure of chromatin during male germline development might direct retrotransposition and, as a consequence, it is at least conceivable that Alus could have replaced and succeeded tRNA SINEs at whatever regions within the genome are critical to these putative functions. In agreement with this suggestion, the human genome has been subdivided into regions having distinct base compositions, CpG content and repetitive sequence content (56). There is evidence that these genomic subdivisions have been conserved throughout mammalian evolution, implying that Alus may have indeed succeeded mammalian tRNA SINEs within these genomic subdivisions. This idea leads to the testable prediction that Alu-rich subdivisions of the human genome would also be rich in fossil tRNA SINEs (6,7).

Whether or not either of the two possible functions considered for sperm Alu hypomethylation is correct, these observations on sequence-specific packing of DNA in sperm chromatin (52,54), the hypomethylation of an Alu subset in sperm DNA (45,48,49) and the conservation of distinct genomic (56) subdivisions together suggest the existence of an incompletely understood large scale hierarchical organization(s) of human DNA which might either result from or cause the insertion of Alu elements.

7. RETAINING CpG DINUCLEOTIDES IN RECENTLY INSERTED ALUS

Although methylation causes a rapid transition of CpG to TpG the most recently inserted Alus largely retain consensus CpGs (1,2,13,14). Three possibilities could account for this observation. Active source genes might either: (i) enjoy complete protection from germline methylation; (ii) be extremely young; (iii) be restored by an unknown mechanism, potentially creating Alu CpGs.

Oocyte Alus are mostly methylated (45) but a small subset of young Alus that are unmethylated in the oocyte might also belong to the larger subset of unmethylated sperm Alus. The possibility that select Alu source genes may be protected from methylation in both oocyte and sperm cannot be entirely ruled out (below).

Alternatively, the presently active Ya5 Alu source genes are certainly younger than the divergence time of human, chimpanzee and gorilla (Section 4). Consequently, retrotranspositional activity might be restricted to the extremely young Alus that have intact CpGs. Methylation inhibits Alu transcription (Section 8) and younger Alus are far less methylated than older Alus within the male germline (Section 6), potentially providing them with a transcriptional advantage, at least within the paternal lineage. CpG→TpG transitions within the internal promoter elements decreases Alu transcriptional activity (57), providing another possible transcriptional advantage for Alus having intact CpGs. Post-transcriptional events involving RNA folding might also select for intact CpG dinucleotides in source gene transcripts (Sections 5 and 8). While the case is far from proven, several fairly well-documented processes potentially favor the retrotranspositional activity of the youngest Alus as defined by their CpG content.

Biological selection might require a minimum CpG content within at least some Alu elements (for example Sections 6 and 11) and, interestingly, the Ya5 Alu subfamily has both gained and lost a CpG dinucleotide compared with the next older Ya Alu subfamily (2,13,14). Conceivably, CpGs are actively restored and occasionally generated *de novo* during retrotransposition. However, other residues within the most recently inserted Alus almost exactly match their subfamily consensus, re-emphasizing the conclusion that young Alus are apparently the products of young source genes (1,2,13,14) and not the products of older sources genes having rescued or newly created CpG dinucleotides. These same considerations also indirectly argue against the first possibility that source genes are completely protected from germline methylation, since the products of those older genes would be betrayed by their divergence at other sites. Thus I

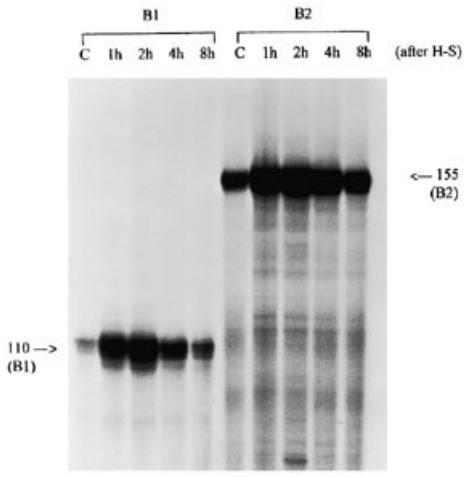

Figure 4. Heat shock increases accumulation SINE RNAs (71). Primer extension by reverse transcriptase is used to assay for mouse B1 and B2 RNAs in control cells (C) and heat-shocked cells that had recovered for the indicated time periods. Note the substantial increases in both B1 (SRP superfamily) and B2 (tRNA superfamily) RNAs during heat shock recovery.

presently favor the possibility that one or more retrotrans-positional advantages select for the activity of extremely young Alus and cause the retention of CpGs in recently inserted Alus.

8. ALU TRANSCRIPTION IS CONTROLLED AT MANY LEVELS

Considering their internal A box and B box promoters, Alus are conceptually a vast multigene family consisting of 1 000 000 dispersed members (1,4). While not systematically investigated, this internal promoter is sufficiently forgiving that most Alus seem to be transcribed *in vitro*. Yet, despite this transcriptional potential, flAlu RNA is usually scarce in cultured cells (1,58). Similarly, B1 RNA is also usually expressed at low levels (Fig. 4, lane 1).

How might a gene family consisting of 1 000 000 dispersed members be regulated? Results reviewed in this section suggest that flAlu RNA expression is determined at many levels, including DNA methylation, chromatin context, *cis*-acting elements, *trans*-acting factors and post-transcriptional processing (59; Fig. 5). Complicating this categorical analysis, the activity of factors for transcription and RNA processing depends upon the *cis* elements both within and flanking individual Alus, making expression of each Alu idiosyncratic and cell type dependent.

Methylation inhibits pol III-directed Alu transcription *in vitro* and *in vivo* (57; unpublished results; Fig. 5) and Alus in somatic cells are usually highly methylated (Section 6). By globally repressing transcription (60), methylation might largely cause the low level of Alu expression. Alu hypomethylation in spermatocytes should derepress its source genes (Section 6) and, in the only case investigated, a *de novo* Alu insertion occurred in the paternal germline (61). Alus are also hypomethylated in placental and hydatidiform mole DNA and, by implication, other tissues associated with fetal development (49). Thus, Alu source genes are potentially derepressed in other tissues that might contribute to embryonic inheritance.

Alus are not the only repetitive sequence family to be extensively methylated (reviewed in 62), leading to the suggestion that the primary role of DNA methylation is to suppress parasitic elements (62). In evaluating this proposal, germ cells and somatic cells should be distinguished.

(i) The programmed demethylation and derepression of Alu retrotransposition in the inherited male germline (above) might be viewed as either contradicting this proposal or a collapse of the proposed protection that DNA methylation provides. Like Alus, the methylation of IAP elements is relaxed in the developing male germline and expression of an IAP source gene is restricted to a remarkably precise stage of spermatogenesis (29). Additional evidence strongly implies that demethylation is required for expression of IAP during this developmental stage. Also, expression of mouse LINE1 is tightly regulated for developmentally specific expression during embryogenesis such that it can contribute to primordial germ cells (63). The methylation of LINEs at these same stages of development remains to be investigated but their demethylation probably contributes to their derepression (63). In these three independent examples, mobile elements are subject to transcriptional repression due to methylation, their transcriptional repression is relaxed in exactly those tissues in which they might alter inheritance and, in each case, this relaxation is evidently a normally programmed event rather than the result of a cellular aberration or insult. These three mobile elements are not thought to direct their own demethylation; Alus in particular are entirely subject to regulation by factors supplied *in trans* by other genes. Thus each of these three elements is successfully exploiting a niche in the developmentally programmed pattern of DNA demethylation. Accordingly, DNA methylation fails to protect against retrotransposition because normal changes that occur during germ cell development make DNA methylation inherently unsuited for this function. Presumably, this developmental programming and the DNA methylation upon which it acts serve another role(s).

(ii) As previously discussed, methylation represses Alu transcription in somatic cells and probably does so through remodeling chromatin structure (below). Differences in Alu methylation in somatic tissues (49) suggest that Alu repression can be regulated in a tissue-specific manner. Other evidence supports a central role for Alu RNA in regulating protein homeostasis (Section 10), advancing a proposed Alu RNA function that certainly requires regulated Alu expression. Since such a large fraction of somatic DNA methylation resides within Alus, the proposal that this methylation is devoted to regulating transcription of this ‘multigene’ family (62) is entirely reasonable and can be regarded as another example of methylation regulating gene expression (60)

Chromatin structure inhibits Alu transcription nearly 100-fold, providing another level of global repression (64,65; Fig. 5). However, hypermethylated DNA organizes chromatin having hypoacetylated histones so that chromatin may merely mediate the transcriptional inhibition caused by Alu hypermethylation (55). As further evidence that Alu transcription is regulated by chromatin, the increase in Alu RNA caused by viral infection (64–67) has been attributed to an opening of Alu chromatin structure (64,65).

Accumulated mutations within their internal promoters eventually cripple older Alus, transcriptionally disadvantaging them as potential source genes (57; Fig. 5); a ‘multigene’ family consisting of 1 000 000 members certainly includes many

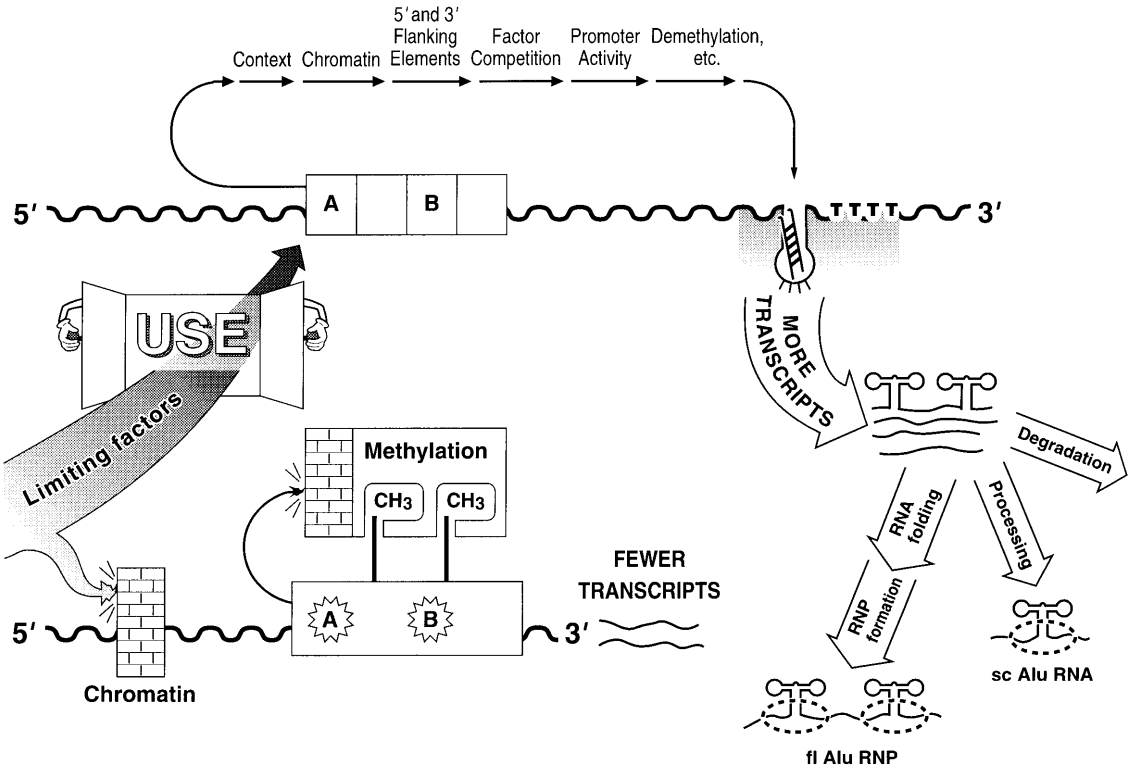

Figure 5. How 1 000 000 genes might be regulated. Many levels of control significantly (e.g. ≥10-fold) influence the steady-state abundance of different flAlu RNA transcripts (text; modified from 59). These include: (i) the fidelity of the A box–B box promoter elements; (ii) chromatin structure; (iii) chromosome context; (iv) methylation; (v) identity of upstream sequence elements (USE); (vi) downstream sequence elements including the terminator; (vii) limiting factors that specifically act upon these elements; (viii) Alu RNA folding and processing; (ix) association of flAlu RNA with binding proteins. Compounding nine or more significant levels of control can impose an extremely high degree of regulation on both the overall accumulation of flAlu RNA and the relative expression of particular Alus.

psuedogenes. Even in young Alus, the inherently weak internal A box–B box promoter is not sufficient for efficient pol III-directed transcription *in vivo* (30). Expression of the SRP RNA gene requires elements in its 5′ flanking sequence and these same elements stimulate expression of chimeric Alu constructs both *in vivo* and *in vitro* (30,68).

Sequences flanking Alus result accidentally from their insertion sites (Fig. 1), making the expression of each Alu potentially unique. In agreement with this suggestion, sequences flanking the Ya5 Alu founder gene (Section 4) stimulate its expression (28). Interestingly, p53 represses pol III-directed transcription in a promoter-dependent manner and, in particular, represses transcription from a basal A box–B box Alu promoter (69). However, flanking sequences derived from either the Ya5 Alu founder or the SRP RNA gene overcome p53-mediated repression (69). Taken together these data show how each Alu might be uniquely transcribed in a cell-dependent manner (Fig. 5).

The 3′ sequences flanking an Alu, which also fortuitously result from the insertion site, contain the pol III terminator (30,70; Fig. 5). Since four adjacent T residues in the coding strand suffice for termination, primary flAlu transcripts will on average extend ~200 nt into the 3′ flanking sequence. The steady-state expression of an Alu-like template is very sensitive to the structure of its terminator, including the number of T residues, the base sequence surrounding these T reidues and secondary structure adjacent to the transcript 3′-terminus (30,32,33,70). RNA lifetime experiments indicate that the terminator's effect on flAlu RNA accumulation results primarily from its influence on transcription rather than RNA stability (30). These effects are evidently mediated by the availability of La protein (31,33,70). Again, flanking sequences could uniquely influence each Alu's expression in a cell type-dependent manner.

Furthermore, as previously reviewed, processing of flAlu RNA into scAlu-like RNAs depends upon a complex interplay of various structural features of the particular Alu and its interaction with several proteins (31–33; Section 5). These post-transcriptional events would modulate the relative expression of different flAlu transcripts (Fig. 5).

Answering the question posed at the beginning of this section as to how 1 000 000 'genes' might be regulated, there is not one but rather multiple levels of Alu transcriptional control (Fig. 5). Methylation, chromatin structure, promoter mutations, unique 5′ and 3′ flanks and the action of various factors have each been observed to cause significant differences (~10-fold or more) in Alu transcriptional expression. Compounding six or more 10-fold effects would collectively control Alu expression over a range exceeding 1 000 000 (Fig. 5).

9. SINE RNA INDUCTION SUGGESTS THE POSSIBILITY OF FUNCTION

Cell stress dramatically increases the abundance of human flAlu RNA and other mammalian SINE RNAs (71,72; Fig. 4). For example, heat shock causes a nearly 100-fold increase in mouse

B1 RNA making this sparse transcript abundant (71; Fig. 4). Mouse B2 and rabbit C RNAs show similar increases, indicating that the heat shock response is conserved by the SRP RNA and tRNA SINE superfamilies (Fig. 4). In addition to heat shock, other classic cell stress treatments increase flAlu RNA (71). Cell stress does not change the lifetime of flAlu RNA but probably increases Alu transcription (71).

Viral infection or administering cycloheximide to cells also significantly increases the abundance of SINE RNA (66,67,71,73–75). The induction of Alu RNA by either cycloheximide or heat shock occurs <20 min after subjecting cells to these stresses (71). The rapidity of these responses suggests that they result from the modification of existing factors and do not involve either DNA demethylation or synthesis of new factors. Do these increases in SINE RNA merely reflect an aberrant breakdown in regulation or are they a controlled response?

10. PKR REGULATION UNIFIES THE EFFECTS OF CELLULAR INSULTS ON TRANSLATION, eIF2 PHOSPHORYLATION AND SINE RNA

Viral infection, inhibiting translational elongation and cell stress each cause complex, pleiotropic changes in cell physiology. But each increases SINE RNA abundance, alters protein synthesis through pathways involving eIF2α phosphorylation and causes changes in the activity of PKR, the eIF2 kinase that is regulated by double-stranded (ds)RNA (reviewed in 76). The known regulation of PKR by RNAs potentially unifies these observations. The binding of two or more PKRs to a long dsRNA increases its autophosphorylation activity (77–79). Phosphorylation activates PKR as an eIF2 kinase, which by phosphorylating eIF2α ultimately inhibits translational initiation (77–79). Small highly structured RNAs that sequester PKR as bound monomers inhibit its autoactivation, thereby potentially increasing the rate of protein synthesis (77–79).

Pursuing this clue, we find that overexpressed flAlu RNA: (i) increases protein synthesis; (ii) binds PKR; (iii) inhibits PKR activation (76). Alus are not amenable to the usual genetic tests of function, raising the possible question of whether these effects are an overexpression artifact. However, these effects are caused by levels of flAlu RNA overexpression that equal the levels resulting from cell stress, viral infection and so forth (76). Indeed, under stress conditions very scarce SINE RNAs accumulate to extremely high levels (66,71; Fig. 4). RNA gel shift assays indicate that flAlu RNA forms especially tight PKR complexes, making it a particularly potent PKR inhibitor (76; unpublished results). These observations support the proposals that flAlu RNA is a highly specialized PKR regulator and that increases in flAlu RNA caused by cell insults and viral infection are controlled responses to regulate PKR activation.

Biology implies that this proposed function for flAlu RNA cannot be peculiar to higher primates. Therefore, a corollary of this proposal is that cell stress-induced increases in the level of other SINE RNAs (Fig. 4) serve a similar function (Section 11).

As an eIF2 kinase, PKR maintains translational homeostasis (77–79). PKR also regulates transcription, possibly through NFκB (80–82), is required for TNFα-induced apoptosis and serves as a tumor suppresser (83–87). flAlu RNA could potentially regulate PKR's other activities (unpublished results) so that Alu and, more generally, SINE RNAs might serve a vital role in cell physiology.

11. HOW SINE RNAS COULD HAVE A COMMON FUNCTION

The processing of SRP SINE RNAs in intimate association with SRP proteins 9/14 (Section 5) strongly suggests the possibility that scAlu/B1 RNA participates in SRP-related activities (31; below). However, this intriguing possibility should be considered in the broader context of the central question raised in Section 1, whether a pre-existing function or the retrotranspositional pathway is ultimately responsible for maintaining SINEs and driving their evolution (Section 1). Since there is no reason to suspect that tRNA SINEs participate in an SRP-related activity, this possible function would be peculiar to rodents and primates (Fig. 2). Selfish DNA models are entirely compatible with such elements adapting to new functions (9–11) and there are numerous examples of Alu-like elements being involved in a variety of activities (18). Similarly, a pre-existing function that is common to all SINEs does not necessarily preclude the more recent SRP SINE superfamily from participating in other functions.

The absence of Alus in most mammals and tRNA SINEs in human (Fig. 2) indicates that either SINEs lack function or their sequence *per se* is not essential for function. PKR binding primarily requires only a minimum number of base pairs within an RNA secondary structure so that entirely unrelated RNA sequences can functionally substitute as PKR inhibitors (77,88). As the classic example, the adenovirus VAI RNA gene inhibits virally induced PKR activation; both protein synthesis and viral infectivity are impaired for VAI mutants. However, the gene for an entirely unrelated RNA, EBER1, rescues both infectivity and protein synthesis for VAI mutants (88). According to this suggestion, cell stress-induced transcripts from the tRNA SINE superfamily might serve the same PKR regulatory role as Alu RNA (71; Fig. 4).

Yet a minimal RNA secondary structure alone cannot suffice to inhibit PKR *in vivo*. Otherwise, cellular RNAs, e.g. rRNA alone, would present an extraordinary number of PKR binding sites overwhelming any possibility of signaling PKR with structured RNAs. Fragments of Alu that have been recruited into mRNAs present a similar problem (Section 3). Presumably, rRNA and other functional RNAs are unavailable for PKR binding because of their subcellular location or organization into RNP structures. Also, RNAs having other vital functions are poorly suited for signaling PKR since significantly increasing their abundance would disrupt normal function.

Because they apparently lack any other essential cellular function (certainly, none is presently known), SINE RNAs could be ideally suited for signaling PKR. SINE RNA abundance can be rapidly and dramatically increased without inadvertently disrupting other necessary functions (71; Fig. 4). Similarly, short-lived flAlu RNA (30) could be rapidly reduced to basal levels when no longer required for PKR inhibition. As otherwise functionless RNAs, the RNP structure and subcellular location of SINE RNAs could promote their PKR accessibility. In support of the notion that Alu RNPs are accessible for PKR binding, the only other proteins known to form Alu RNPs are two small SRP 9/14 proteins and La, which transiently binds to the 3′-ends of nascent pol III transcripts (31,33–35; Section 8). SRP 9/14 has a significantly lower binding affinity to right Alu monomers and the lowest affinity for the Ya5 subfamily right monomer (34). Ya5 flAlu RNA appears to be a far more potent PKR inhibitor than scAlu RNA (71), raising the possibility that the accessibility of its

right monomer determines its PKR activity. Perhaps an interplay between PKR and SRP 9/14 for flAlu RNA binding regulates its PKR accessibility. The PKR activity of other Alu subfamilies remains to be tested.

The notion of an RNA regulating a kinase is odd and the idea that SINE RNAs might have been recruited for a PKR regulatory function during recent evolution approaches implausibility. The homology of yeast GCN2 to both tRNA synthetases and mammalian PKR suggests how SINE RNA might have acquired this function (77,79,80). GCN2 senses the cell's metabolic state by binding uncharged tRNAs which activate this eIF2 kinase to inhibit protein synthesis (77,79,89). GCN2 effectively combines aspects of tRNA synthetase and eIF2 kinase activities. This mechanism suggests a very ancient association between a primordial PKR's RNA-binding properties and the very deepest evolutionary roots of SINEs within the tRNA superfamiliy (Fig. 2). The original tRNA SINE transcripts would have been pre-endowed with the secondary structure necessary for signaling PKR. Because of its internal promoter, this original retrotransposed tRNA SINE would have retained the transcriptional potential of its source gene but, freed from the tRNA function of its antecedent, regulated expression of the resulting tRNA psuedogene(s) would have adapted to the requirements of a dedicated PKR inhibitor.

The proposal that SINE RNAs regulate protein synthesis by signaling PKR in response to cell stress implies a selective advantage for their maintenance within the genome. Consequently, any functional relationship between SINE RNA and PKR would select for successful SINE sequences. For example, binding studies indicate that flAlu RNA forms unusually tight PKR complexes, making it a particularly effective PKR inhibitor (76; unpublished results). While raising other questions, this proposed role for Alu RNA unifies our current knowledge of SINEs and presents testable hypotheses for their origins and functions.

Two entirely different speculations have been advanced for possible Alu functions, one involving DNA and the other RNA. At the risk of conflating two speculations, the demethylation of a major Alu subset in sperm (Section 6) almost certainly derepresses their transcription (Section 8) following fertilization. The resulting flAlu transcripts should then signal PKR and its attendant pathways, including especially protein synthesis (Section 10), in the early embryo.

ACKNOWLEDGEMENTS

I thank my colleagues Mel Green, Scott Hawley, Rich Maraia, Greg Matera and Nori Okada for valuable advice in preparing this manuscript. I especially thank my students and all other members of the extended Alu family who have taught us so much about ourselves. This research has been supported by USPHS grant GM 21346 and the Agricultural Experiment Station of the University of California.

REFERENCES

1 Schmid,C.W. (1996) *Prog. Nucleic Acid Res. Mol. Biol.*, **53**, 283–319.
2 Jurka,J. (1995) In Maraia,R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. Springer, New York, NY, pp. 25–42.
3 Okada,N. and Ohshima,K. (1995) In Maraia,R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. Springer, New York, NY, pp. 61–80.
4 Weiner,A.M., Deininger,P.L. and Efstratiadis,A. (1986) *Annu. Rev. Biochem.*, **55**, 631–661.
5 Quentin,Y. (1994) *Genetica*, **93**,203–215.
6 Jurka,J., Zietkiewicz,E. and Labuda,D. (1995) *Nucleic Acids Res.*, **23**, 170–175.
7 Smit,A.F. and Riggs,A.D. (1995) *Nucleic Acids Res.*, **23**, 98–102.
8 Rubin,C.M., Leeflang,E.P., Rinehart,F.P. and Schmid,C.W. (1992) *Genomics*, **18**, 322–328.
9 Doolittle,W.F. and Sapienza,C. (1980) *Nature*, **284**, 601–603.
10 Orgel,L.E. and Crick,F.H. (1980) *Nature*, **284**, 604–607.
11 Charlesworth,B., Sniegowski,P. and Stephan,W. (1994) *Nature*, **371**, 215–220.
12 Kass,D.H., Batzer,M.A. and Deininger,P.L. (1995) *Mol. Cell. Biol.*, **15**, 19–25.
13 Jurka,J. and Smith,T. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 4775–4778.
14 Britten,R.J., Baron,W.F., Stout,D.B. and Davidson,E.H. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 4770–4774.
15 Labuda,D., Zietkiewicz,E. and Mitchell,G.A. (1995) In Maraia,R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. Springer, New York, NY, pp. 106–123.
16 Economou-Pachnis,A. and Tsichlis,P.N. (1985) *Nucleic Acids Res.*, **13**, 8379–8387.
17 Makalowski,W., Mitchell,G.A. and Labuda,D. (1994) *Trends Genet.*, **10**, 188–193.
18 Makalowski,W. (1995) In Maraia,R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. Springer, New York, NY, pp. 81–104.
19 Waldman,A.S. and Liskay,R.M. (1988) *Mol. Cell. Biol.*, **8**, 5350–5357.
20 Schmid,C.W. and Maraia,R. (1992) *Curr. Opin. Genet. Dev.*, **2**, 874–882.
21 Batzer,M.A., Stoneking,M., Alegria-Hartman,M., Bazan,H., Kass,D.H., Shaikh,T.H., Novick,G.E., Ioannou,P.A., Scheer,W.D., Herrera,R.J. and Deininger,P.L. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 2288–2292.
22 Batzer,M.A., Rubin,C.M., Hellmann-Blumberg,U., Alegria-Hartman,M., Stern,J., Shaikh,T., Leeflang,E.P., Bazan,H., Deininger,P. and Schmid,C.W. (1995) *J. Mol. Biol.*, **247**, 418–427.
23 Batzer,M.A., Deininger,P.L., Hellmann-Blumberg,U., Jurka,J., Labuda,D., Rubin,C.M., Schmid,C.W., Zietkiewicz,E. and Zuckerkandl,E. (1996) *J. Mol. Evol.*, **42**, 3–6.
24 Batzer,M.A., Arcot,S.S., Phinney,J.W., Alegria-Hartman,M., Kass,D.H., Milligan,S.M., Kimpton,C., Gill,P., Hochmeister,M., Ioannou,P.A. *et al.* (1996) *J. Mol. Evol.*, **42**, 22–29.
25 Leeflang,E.P., Liu,W.-M., Hashimoto,C., Choudary,P.V. and Schmid,C.W. (1992) *J. Mol. Evol.*, **35**, 7–16.
26 Leeflang,E.P., Chesnokov,I. and Schmid,C.W. (1993) *J. Mol. Evol.*, **37**, 566–572.
27 Leeflang,E.P., Liu,W.-M., Chesnokov,I. and Schmid,C.W. (1993) *J. Mol. Evol.*, **37**, 559–565.
28 Chesnokov,I. and Schmid,C.W. (1996) *J. Mol. Evol.*, **42**, 30–36.
29 Dupressoir,A. and Heidmann,T. (1996) *Mol. Cell. Biol.*, **16**, 4495–4503.
30 Chu,W.-M., Liu,W.-M. and Schmid,C.W. (1995) *Nucleic Acids Res.*, **23**, 1750–1757.
31 Maraia,R.J. and Sarrowa,J. (1995) In Maraia,R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. Springer, New York, NY, pp. 163–196.
32 Maria,R.J., Chang,D.-Y., Wolffe,A.P., Vorce,R.L. and Hsu,K. (1992) *Mol. Cell. Biol.*, **12**, 1500–1506.
33 Maraia,R.J., Kenan,D.J. and Keene,J.D. (1994) *Mol. Cell. Biol.*, **14**, 2143–2158.
34 Sarrowa,J., Chang,D.Y. and Maraia,R.J. (1997) *Mol. Cell. Biol.*, **17**, 1144–1151.
35 Chang,D.Y., Sasaki-Tozawa,N., Green,L.K. and Maraia,R.J. (1995) *Mol. Cell. Biol.*, **15**, 2109–2116.
36 Sinnett,D., Richer,C., Deragon,J.M. and Labuda,D. (1992) *J. Mol. Biol.*, **226**, 689–706.
37 Sinnett,D., Richer,C., Deragon,J.M. and Labuda,D. (1991) *J. Biol. Chem.*, **266**, 8675–8678.
38 Moran,J.V., Holmes,S.E., Naas,T.P., DeBerardinis,R.J., Boeke,J.D. and Kazazian,H.H.,Jr (1996) *Cell*, **87**, 917–927.
39 Feng,Q., Moran,J.V., Kazazian,H.H.,Jr and Boeke,J.D. (1996) *Cell*, **87**, 905–196.
40 Ohshima,K., Hamada,M., Terai,Y. and Okada,N. (1996) *Mol. Cell. Biol.*, **16**, 3756–3764.
41 Okada,N., Hamada,M., Ogiwara,I. and Ohshima,K. (1997) *Gene*, **205**, 229–243.
42 Jurka,J. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 1872–1877.
43 Boeke,J.D. (1997) *Nature Genet.*, **16**, 6–7.
44 Schmid,C.W. (1991) *Nucleic Acids Res.*, **19**, 5613–5617.
45 Rubin,C.M., Vandevoort,C.A., Teplitz,R.L. and Schmid,C.W. (1994) *Nucleic Acids Res.*, **22**, 5121–5127.

46 Doeffler,W. (1993) In Jost,J.P. and Saluz,H. (eds), *DNA Methylation: Molecular Biology and Biological Significance*. Brinkhauser Verlag, Basel, Switzerland, pp. 263–299.
47 Hasse,A. and Schulz (1994) *J. Biol. Chem.*, **269**, 1821–1826.
48 Kochanek,S., Renz,D. and Doerfler,W. (1993) *EMBO J.*, **12**, 1141–1151.
49 Hellmann-Blumberg,U., Hintz,M.F., Gatewood,J. and Schmid,C.W. (1993) *Mol. Cell. Biol.*, **13**, 4523–4530.
50 Chesnokov,I. and Schmid,C.W. (1995) *J. Biol. Chem.*, **270**, 18539–18542.
51 Razin,A. and Cedar,H. (1994) *Cell*, **77**, 473–476.
52 Gatewood,J.G., Cook,G.R., Balhorn,R., Bradbury,E.M. and Schmid,C.W. (1987) *Science*, **236**, 962–964.
53 Gatewood,J.M., Cook,G.R., Balhorn,R., Schmid,C.W. and Bradbury,E.M. (1990) *J. Biol. Chem.*, **265**, 20662–20666.
54 Gardiner-Garden,M., Balleseros,M., Gordon,M. and Tam,P.P.L. (1998) *Mol. Cell. Biol.*, **18**, 3350–3356.
55 Kass,S.U., Pruss,D. and Wolffe,A.P. (1997) *Trends Genet.*, **13**, 444–449.
56 Bernardi,G. (1989) *Annu. Rev. Genet.*, **23**, 627–661.
57 Liu,W.M. and Schmid,C.W. (1993) *Nucleic Acids Res.*, **21**, 1351–1359.
58 Liu,W.-M., Maraia,R.J., Rubin,C.M. and Schmid,C.W. (1994) *Nucleic Acids Res.*, **22**, 1087–1095.
59 Schmid,C.W. and Rubin,C.M. (1995) In Maraia,R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. Springer, New York, NY, pp. 106–123.
60 Bird,A.P. (1995) *Trends Genet.*, **11**, 94–100.
61 Wallace,M.R., Anderson,L.B., Saulino,A.M., Gregory,P.E., Glover,T.W. and Collins,F.S. (1991) *Nature*, **353**, 864–866.
62 Yoder,J.A., Walsh,C.P. and Bestor,T.H. (1997) *Trends Genet.*, **13**, 335–340.
63 Trelogan,S.A. and Martin,S.L. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 1520–1524.
64 Russanova,V.R., Driscoll,C.T. and Howard,B.H. (1995) *Mol. Cell. Biol.*, **15**, 4282–4290.
65 Howard,B.H., Russanova,V.R. and Englander,E.W. (19995) In Maraia,R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. Springer, New York, NY, pp. 133–142.
66 Panning,B. and Smiley,J.R. (1993) *Mol. Cell. Biol.*, **13**, 3231–3244.
67 Panning,B. and Smiley,J.R. (1995) In Maraia,R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. Springer, New York, NY, pp. 143–161.
68 Bredow,S., Surig,D., Miller,J., Kleinert,H. and Benecke,B.J. (1990) *Nucleic Acids Res.*, **18**, 6779–6784.
69 Chesnokov,I., Chu,W.-M., Botchan,M.R. and Schmid,C.W. (1996) *Mol. Cell. Biol.*, **16**, 7084–7088.
70 Chu,W.-M., Ballard,R.E. and Schmid,C.W. (1997) *Nucleic Acids Res.*, **25**, 2077–2082.
71 Liu,W.-M., Chu,W.-M., Choudary,P.V. and Schmid,C.W. (1995) *Nucleic Acids Res.*, **23**, 1758–1765.
72 Fornace,A.J.,Jr and Mitchell,J.B. (1986) *Nucleic Acids Res.*, **14**, 5793–5811.
73 Jang,K.L. and Latchman,D.S. (1992) *Biochem. J.*, **284**, 667–673.
74 Jang,K.L., Collins,M.K. and Latchman,D.S. (1992) *J. Acquir. Immune Defic. Syndr.*, **5**, 1142–1147.
75 Singh,K., Carey,M., Saragosti,S. and Botchan,M. (1985) *Nature*, **314**, 553–556.
76 Chu,W.-M., Ballard,R., Carpick,B.W., Williams,B.R.G. and Schmid,C.W. (1998) *Mol. Cell. Biol.*, **18**, 58–68.
77 Clemens,M.J. (1996) In Hershey,J.W.B., Mathews,M.B. and Sonenburg,N. (eds), *Translational Control*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 139–172.
78 McMillan,N.A.J. and Williams,B.R.G. (1996) In Clemens,M.J. (ed.), *Structure and Function of the Interferon-Induced Protein Kinase, PKR and Related Enzymes in Protein Phosphorylation in Cell Growth Regulation*. Harwood, Amsterdam, The Netherlands, pp. 225–253.
79 Proud,C.G. (1995) *Trends Biochem. Sci.*, **20**, 241–246.
80 Kumar,A., Hague,J., Lacaste,J., Hiscott,J. and Williams,B.R.G. (1994) *Proc. Natl Acad. Sci. USA*, **9**, 6288–6292.
81 Kumar,A., Yang,Y.-L., Flati,V., Der,S., Kadereit,S., Hague,J., Deb,A., Reis,L., Weismann,C. and Williams,B.R.G. (1997) *EMBO J.*, **16**, 406–416.
82 Koromilas,A.E., Cantin,C., Craig,A.W., Jagus,R., Hiscott,J. and Sonenberg,N. (1995) *J. Biol. Chem.*, **270**, 25426–25434.
83 Koromilas,A.E., Roy,S., Barber,G.N., Katze,M.G. and Sonenberg,N. (1992) *Science*, **257**, 1685–1688.
84 Donzé,O., Jagus,R., Koromilas,A.E., Hershey,J.W. and Sonenberg,N. (1995) *EMBO J.*, **14**, 3828–3834.
85 Meurs,E.F., Galabru,J., Barber,G.N., Katze,M.G. and Hovanessian,A.G. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 232–236.
86 Yeung,M.C., Liu,J. and Lau,A.S. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 12451–12455.
87 Der,S.D., Yang,Y.L., Weissmann,C. and Williams,B.R. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 3279–3283.
88 Bhat,R.A. and Thimmappaya,B. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 4789–4793.
89 Hinnebusch,A.G. (1994) *Trends Biochem. Sci.*, **19**, 409–414.