

From 'JUNK' to Just Unexplored Noncoding Knowledge: the case of transcribed Alus

Rajesh Pandey and Mitali Mukerji

Abstract

Non-coding RNAs (ncRNAs) are increasingly being implicated in diverse functional roles. Majority of these ncRNAs have their origin in the repetitive elements of genome. Significantly, increase in genomic complexity has been correlated with increase in repetitive content of the genome. Primate-specific Alu repeats, belonging to SINE class of repeats, is the most abundant repeat class inhabiting the human genome. Of the many possible functional roles of Alu repeats, they have been shown to modulate human transcriptome by virtue of harboring diverse array of functional RNA pol II TFBS, cryptic splice-site-mediated Alu exonization and as probable miRNA targets. Retro-transposition of Alu harboring TFBS has shaped up gene-specific regulatory networks. Alu exonized transcripts are raw material for dsRNA-mediated A–I editing leading to nuclear retention of transcripts and change in miRNA target. miRNA targets within Alu may titrate the effective miRNA or transcript concentration, thus acting as 'miRNA sponge'. Differential levels of Alu RNA during different conditions of stress also await clear functional understanding. These have contributed toward evolution of complex regulatory repertoire leading to the evolution of primate-specific functions. Recent reports of co-localization of pol II and pol III binding sites near the gene and elsewhere in the genome, increase the possibility of dynamic co-ordination between both pol II and pol III determining the ultimate transcriptional outcome. Dynamic and functional Alu repeats seem to be centrally placed to modulate the transcriptional landscape of human genome.

Keywords: non-coding RNA; Alu repeat; TFBS; NAT; miRNA; compartmentalization

Non-coding DNA and Genome Complexity

The proponents of molecular biology have always gone to great lengths to emphasize on what we know today as the Central Dogma. It precisely holds that for any biological expression to occur at the genotypic or phenotypic level, a gene must transcribe into an mRNA and this must then go on to make a protein. This has usually been interpreted to mean that genetic information flows from DNA to proteins via RNA, where RNA is a mere intermediate in the information flow cascade. However,

only 2% of our genes functionally code for proteins; ~50–75% of human genome is transcribed, and 98% of the transcripts do not get translated [1, 2]. Further, genomic complexity neither correlates with chromosome number (popularly known as *K* value paradox), genome size (*C* value paradox) or the number of genes (*N* value paradox) [3]. This was confirmed after human genome sequencing which revealed that humans have only approximately 21 500 protein-coding genes which are comparable to lesser complex genomes of *Drosophila* and *Arabidopsis* [4]. In contrast, non-coding component

Corresponding authors. Mitali Mukerji, Rajesh Pandey, Genomics and Molecular medicine, Institute of Genomics and Integrative Biology (CSIR-IGIB), Council of Scientific and Industrial Research (CSIR), Mall Road, Near Jubilee Hall, Delhi 110007, India. Tel: +91 11 27662738; Fax: +91 11 27667471; E-mail: mitali@igib.res.in, rajeshp@igib.in

Rajesh Pandey During my PhD from CSIR-IGIB, I explored the functional significance of regulatory elements within Alu in human transcriptome. My research interest is to understand the genome regulatory mechanism being modulated by noncoding RNAs in general and Alu in particular.

Mitali Mukerji My research is in the broad area of genome variation and its effect on human phenotypes and susceptibility to diseases. My group is studying the role of Alu elements in evolution of novel regulatory networks through functional genomics approaches.

of the genome has increased substantially with genome complexity. These regions have assumed importance in recent times as they seem to hold a large number of cues that might explain the observed complexity in humans [5, 6]. This makes us sit back and question some of our good old notions of biology.

ncRNAs have no apparent ORF for translation, and it include 'housekeeping' ncRNAs [such as transfer RNAs (tRNAs), ribosomal RNAs (rRNAs) and small nuclear RNAs (snRNAs)], small/short ncRNAs (miRNA and piRNA), repetitive element-associated ncRNAs (Alu RNA) and long ncRNAs (lncRNAs). ncRNAs are transcribed by different RNA polymerase complexes, e.g. Alu is transcribed by pol III whereas miRNAs, snoRNAs, some snRNAs and some scRNAs are mostly transcribed by pol II [7–9]. A large class of ncRNAs like 7SL [10], 7SK [11] and miRNAs [12] are highly conserved across species and are being implicated in many genome functions [13]. Although conserved *cis*-regulatory elements in human genome like proximal promoter elements, enhancers, silencers and insulators comprise

a larger fraction of the regulatory repertoire [14]. ncRNA can assume dynamic secondary structures and affect various cellular functions like splicing, editing, silencing through antisense, methylation and nuclear retention of transcripts [15–17]. They can also be a source of miRNA biogenesis and miRNA target. These RNAs govern diverse aspects of development, cellular differentiation and external/internal stimuli driven spatio-temporal expression [18–22]. They are also known to regulate transcription factors (TFs) as co-activators or co-repressors [23–26]. Different tiers of regulation across phyla from bacteria to humans can be regulated through variety of ncRNAs like *cis/trans* antisense RNA and transcribed pseudogene (Table 1). The ability of some RNAs to function as ncRNA or mRNA/protein depending on the cellular state has added a new dimension to the already rich repertoire of functions transacted by RNAs (Table 2). Different class of ncRNAs has also been correlated with many disease states like cancer, neurological and congenital diseases [27] as enumerated in Table 3.

Table 1: Expanding functional landscape of ncRNAs

Organism	ncRNA	Function
<i>Xenopus laevis</i>	<i>Xlsirt</i>	Anchors RNAs in vegetal cortex for normal embryonic development [28].
Bacteria	Antisense RNA	Affect bacterial virulence through its 3'-UTR, 5'-UTR, <i>cis</i> - and <i>trans</i> -antisense RNAs [29].
<i>Drosophila melanogaster</i>	<i>Hsr-ω</i> Sat III-induced ncRNAs	Involved in intranuclear trafficking and availability of hnRNPs [30]. Nuclear stress granules (nSGs), harbors Sat III repeat-induced ncRNAs, HSF1, splicing factors and hnRNPs which modulate chromatin structure [31].
Humans	<i>NRON</i> Transcribed pseudogenes	Represses <i>NFAT</i> , a TF essential for T-cell receptor-mediated immune response [25]. Act as decoy for functional mRNAs, involved in sequestering TFs and miRNAs [32].

ncRNAs have been part of the transcriptome regulatory network across diverse species right from bacteria to humans. Different type of regulatory RNAs like ncRNA, antisense RNA and transcribed pseudogenes affect different functions from development to decoy for TFs or miRNAs.

Table 2: RNA with dual functions

Organism	Function	
	ncRNA	Protein/mRNA
Bacteria	<i>SgrS</i> 3' region mediates mRNA degradation and translational repression, during sugar phosphate stress [33].	<i>SgrT</i> (43 aa polypeptide) 5' region inhibits glucose transport [34].
<i>Drosophila melanogaster</i>	<i>oskar</i> during oogenesis is required for development [35].	Functions as mRNA during embryonic stage [35].
Human	<i>SRA</i> co-activates nuclear receptors like estrogen, progesterone and androgen [36]. <i>p53</i> and <i>Mdm2</i> functional interplay regulates <i>p53</i> synthesis [38].	<i>SRAP</i> interacts with <i>SRA</i> to inhibit myogenic differentiation through <i>MyoD</i> [37]. <i>p53</i> and <i>Mdm2</i> also regulate <i>p53</i> ubiquitination by proteasome [38].

ncRNAs also display duality of function. In addition to its role as ncRNA, they also modulate cellular functions either as protein or mRNA.

Table 3: Dark side of ncRNA

Disease	ncRNAs
In cancer	
Breast cancer	SRA [39], H19 [40], BC200 [41], miR-155 [42] and miR-21 [43].
Prostate cancer	DD3 [44] and PCGEM1 [45].
Colon cancer	OCCI [46].
In neurological disorders	
Alzheimer's	BC200 [47].
Schizophrenia and bipolar disorder	DISC2 [48].
Spinocerebellar ataxia type 8	SCA8 (KLHL1 antisense) [49].
In congenital syndromes	
Autistic disorder	RAY1/ ST7 [50].
Prader–Willi syndrome	ZNF127AS [51] and IPW [52].
DiGeorge syndrome	DGCR5 [53].
Psoriasis	PRINS [54].

Differential levels of ncRNA have also been functionally correlated with disease states like cancer, neurological disorders, congenital syndromes and developmental disorders.

RNA Regulation: Economic and Efficient choice

Cells can benefit by using RNA molecules for regulation in diverse ways.

- (i) *Through its increased regulatory repertoire:* The repertoire of RNA is enriched through its multi-dimensional roles based on its location in 5'UTR, 3'UTR, intronic and intergenic regions. RNAs also have unique ability to fold in three-dimensional space and hybridize in a sequence-specific manner to other nucleic acids. dsRNAs act as substrates for A–I editing and can form sense–antisense pairs.
- (ii) *Through transient expression:* Regulation by RNA is more transient as they can be sequestered based on requirement or titrated through mechanisms such as 'miRNA sponges' and sense–antisense pairs.
- (iii) *At post-transcriptional level:* Regulatory RNA can also act at post-transcriptional level, through miRNAs and nonsense-mediated decay (NMD).
- (iv) *In different cellular compartments:* These RNAs can mediate differential transport of molecules thereby modulating downstream functions.

Some of these functions like increased regulatory repertoire by RNA mediated regulation are exclusive to RNA whereas other functions are also shared

by the proteins. Thus it is remarkable that RNA in a cell has more than one way to achieve a single objective and provides choice for more than one way of differential regulation [29]. RNAs are proposed to modulate networks that process discrete information packets to yield continuously varying responses, thus functioning like digital-to-analog converters, allowing the expansion of complexity in biological systems, well beyond the scope of purely protein-based regulatory networks [55]. One class of ncRNA that almost reflects all the characteristic features displayed by ncRNA community as a whole, are primate-specific Alu repeats. As highlighted in Figure 1, transposons constitute ~45% of the human sequences and are comprised of SINE elements, LINEs, long terminal repeats (LTRs), DNA transposons, MIR repeats, etc. Among the SINEs, Alus are the predominant fraction and occupy nearly 11% of the human genome. Moreover, compared to other repeats, Alu transcripts are also the most abundant. Alu density is enriched in intragenic regions (12.5%) compared to intergenic density of 9.6% and nearly 75% of the genes having at least one or more Alu insertion. Within genes, Alus occupy 12.8% intronic and 1.6% exonic regions. Alu repeats were also observed to be significantly enriched in genes belonging to specific functional categories like metabolism, transport and signaling whereas they are sparse in genes coding structural proteins and information pathways [56].

Genomic organization of Alus

The consensus sequence of Alu is comprised of ~280 bp, formed by head to tail dimerisation of two similar but non-identical monomers joined by variable length of functional poly-A and a polymorphic poly-A tail at the 3' end [57]. The two monomers are called as FRAM and FLAM, indicate the fossil right arm monomer and fossil left arm monomer, respectively. These monomers, in turn, have arisen from a 7SL RNA progenitor through a deletion of 141 nt [57]. The FLAM contains functional bipartite RNA pol III promoter comprising of Box A and B elements [58]. It also contains binding sites for many RNA pol II TFs. A schematic representation of consensus Alu sequence in Figure 2 summarizes the structure of Alu. Alu harbors a total of 23 cryptic splice sites in both 5' and 3' orientation. Unlike other pol III promoters, B-box promoter element is capable of initiating weak

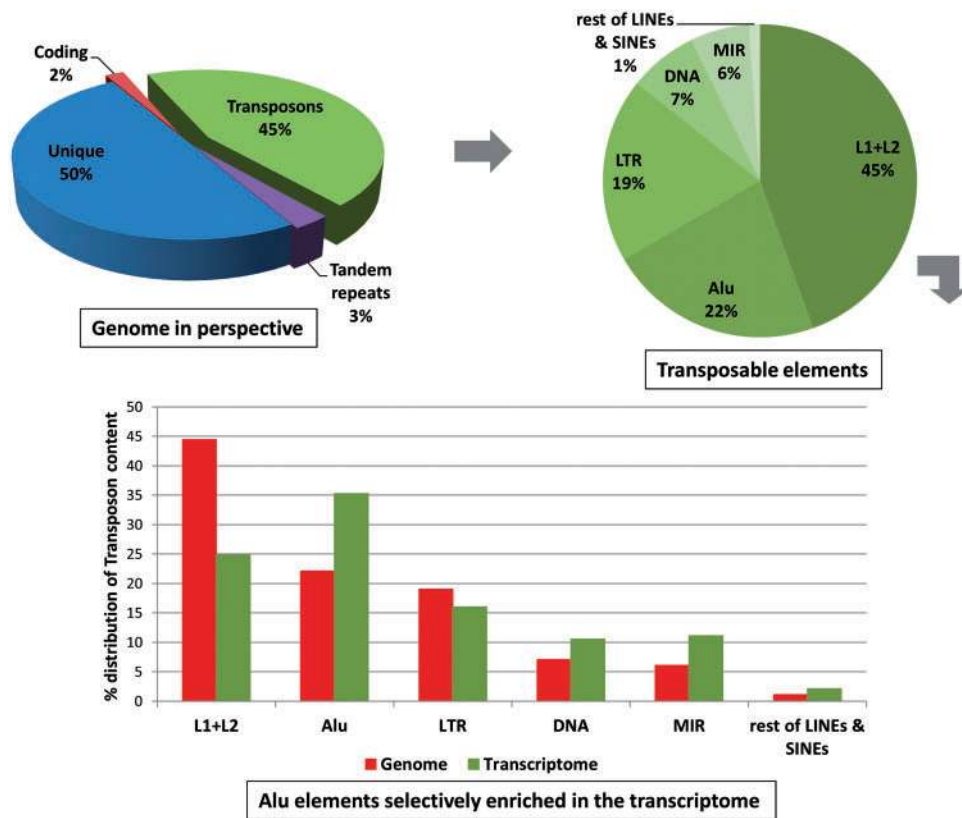


Figure 1: Preferential transcription of Alu contrary to its genomic coverage. Broadly, genome is made up of ~45% TEs and ~55% non-repetitive elements. Transposons are constituted mainly of Alus, LINES, LTR, DNA transposons and MIR repeats. Approximate percentage contribution of each class has been mentioned in the schematic diagram. Interestingly, of the total repeat content, Alus occupy ~22% of genomic space but contribute ~35% of the transcriptomic pool of repetitive elements.

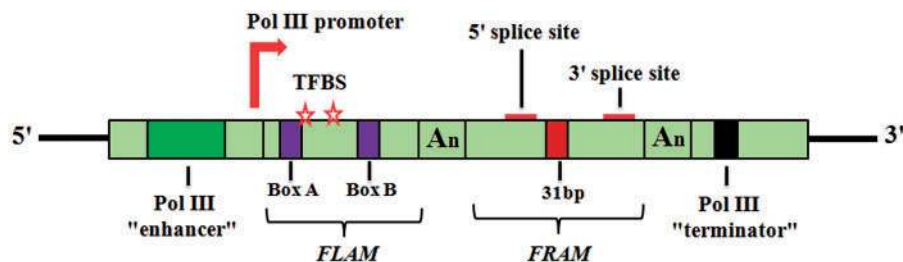


Figure 2: Basic structure of an Alu repeat showing different sequence features. Alu repeat is characterized by the FLAM and FRAM being joined by poly-A and having a poly-A tail at the 3' end. In addition to Alu being transcribed by nearby pol III promoter of a LINE element, it also harbors RNA pol III binding site in its B-box and a terminator sequence at variable distance from FRAM. The functional repertoire is enhanced by the presence of RNA pol II binding sites within Alu like that for RARE, Vit. D, p53, etc. Asterisks show Transcription Factor Binding Sites (TFBS).

transcription on its own and also acts as a transcriptional enhancer [59].

The fusion of two 7SL RNAs happened ~65 million years ago through retro-transposition via a single-stranded RNA intermediate transcribed by

internal RNA pol III promoter [60]. Importantly, Alus are the only active SINE in the human genome capable of retro-transposition via younger Alu subfamilies. According to recent estimates, Alus retrotranspose with a rate of approximately 1

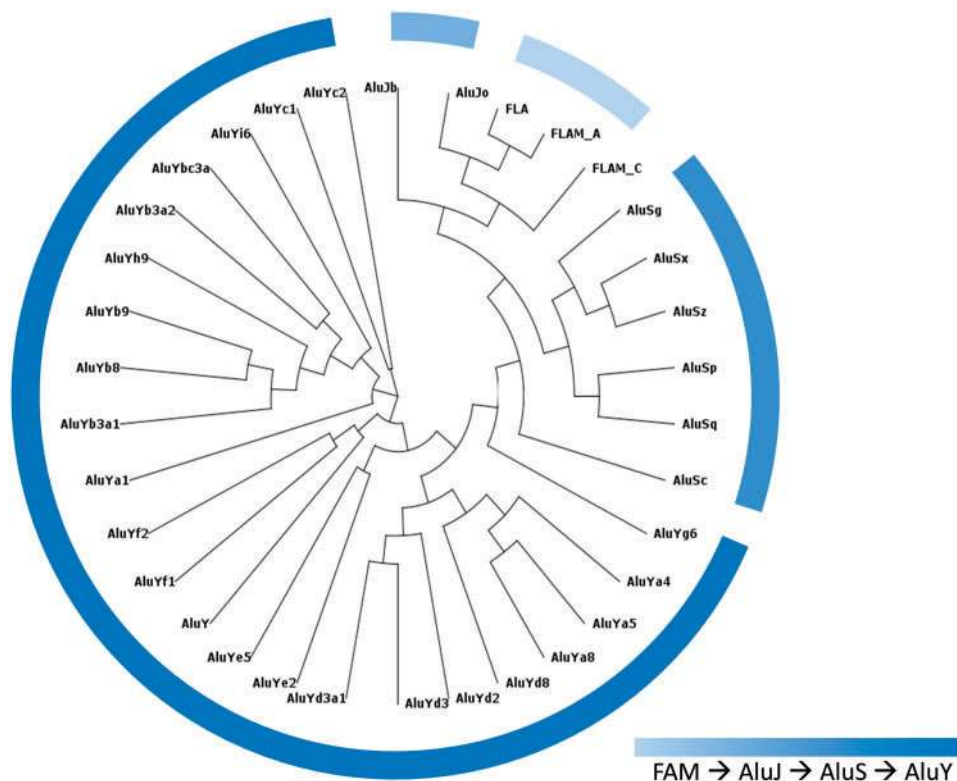


Figure 3: The phylogenetic tree of various subfamilies of Alu. Briefly, the 34 Alu subfamilies consensus sequence were obtained from Repbase (15.07 release) and were aligned in multiple-sequence alignment mode using MUltiple Sequence Comparison by Log-Expectation (MUSCLE). It was analyzed using RepeatMasker (version 3.2.7, January 2009). As per count of Alu subfamilies from hg18, FAM is ~4.4%, Alu J is 26.2%, Alu S is 57.5% and Alu Y is ~12% of the total number of Alus. The intensity of the outer color band is indicative of the Alu subfamily age, newer the subfamily, darker is the color.

integration every 20 births [61]. During expansion of Alu, it is speculated that the master gene acquired mutations which gave rise to Alu subfamilies, Alu J (oldest), Alu S (intermediate) and Alu Y (young), which share ~80–85% similarity among them but can be identified independently by nucleotides at diagnostic positions [62]. Maximum amplification rate was observed about 30 million years ago, the time when Alu Sx subfamily appeared, leading to their highest representation in the genome [63]. A phylogenetic tree of Alu subfamilies is depicted in Figure 3 based on repeats catalogued in Repbase (15.07 release). Of the 1.18×10^6 copies of Alu (hg18) in the human genome, the ancestral form, i.e. fossil arm monomer (FAM) has 52 097 copies, Alu J has 310 562 copies, Alu S has 682 137 copies and the youngest retro-transpositionally active Alu Y has 142 089 copies (to calculate the exact number of copies, alternate genome assemblies and unplaced contigs have been excluded from the analysis). The percentage contribution of Alu in human genome is higher among closely related primates

(Table 4). Depending on the location of Alu in the context of gene-like intergenic, exon, intron, promoter enhancer region or untranslated regions, it can affect functions at various hierarchies. When present in the:

- (i) Intergenic regions, Alus can mediate non-homologous recombination leading to genome shuffling, affect nucleosome positioning/exclusion and chromatin remodeling [64–66], and alter methylation status/imprinting [67–69]. Alu nucleosomes are proposed to serve as ‘anchors’ in organizing the chromatin in human cells. Alus are generally hypermethylated whereas during disease states like cancer and Alzheimer, they are hypomethylated. Methylation status of Alu has the potential to be used as prognostic marker to evaluate progression of cancer.
- (ii) Promoter region, may serve as TFBS or enhancers/repressor and thus regulating gene expression [70–73]. Alus have been the hub of host

Table 4: Genomic coverage of Alu in other primates

Primates	Genome size (bps)	Number of Alu (copies)	Genomic coverage of Alu (bps)	Percentage of genome
Human (hgl8)	3.08×10^9	11 93 407	3.10×10^8	10.08
Chimpanzee (panTro2)	3.35×10^9	11 96 587	3.05×10^8	9.09
Rhesus (rheMac2)	3.01×10^9	11 38 432	2.89×10^8	9.6

A brief summary of the percentage contribution of Alu repeats in closely related primate genomes. The percentage of Alu is higher in humans and importantly younger Alu subfamilies are still undergoing retro-transposition. The version used for calculations are mentioned in the bracket.

of TFs (discussed in detail in ‘Alus and TFBS’ section). Depending upon the position of Alu in the promoter region, they either act as enhancer or repressor.

- (iii) Intronic region, affect transcriptional activity, produce alternate splice variants through its exonization and A–I edited transcripts [74–79]. Alu exons derived from intronic splicing of Alu are proposed to be one of the basic ingredients for evolution as they are known to regulate the human genes and vary in expression across different tissues.
- (iv) Exons, can disrupt reading frame leading to truncated protein or loss of gene activity (pseudogene formation), genetic disorders or evolution of new functions [80–82]. Exons derived from Alu elements have also been associated with disease states in 40 cases. Alu insertion leads to human-specific transcript variants as well as disease like ‘Menkes’ due to insertion induced mutation.
- (v) UTR can affect alternate transcript isoforms in a tissue-specific manner and provide binding site for miRNAs [56, 78, 83, 84]. Alternative transcripts mediated by Alu exonization also show lineage-specific evolutionary events during primate evolution.

Additionally, intra- and inter-chromosomal recombination events mediated by Alu retro-transposition or Alu–Alu non-homologous recombination leads to chromosomal alteration, deletions and duplications (Table 5) and in some cases has resulted in genetic diseases [95, 96]. A summary of the regulatory role of Alu by virtue of its presence at different levels of genomic regulation namely, DNA, RNA and protein level has been summarized in Figure 4. The dynamic hub of pol II TFBS and pol III promoter within Alus make it an interesting player which modulates both pol II and pol III mediated transcriptional responses. Through

Table 5: Alu shapes genomic landscape

Alu	Function
Alu insertion polymorphism	Human population diversity marker [85–87]. Population identification [88]. DNA fingerprinting [89].
Alu recombination	DNA methylation status [69]. Canonical polyadenylation signal [90]. Positioning of nucleosomes in chromatin [64, 91]. Segmental duplication and CNV [92]. Sequence-specific interaction with nuclear proteins [93].
Alu methylation	Tissue-specific pattern of methylation. Especially, it seems to have shaped methylome of human cerebellum [94].

Alu elements drive genomic rearrangements by virtue of insertion and recombination events. For example recombination-mediated DNA methylation imparts tissue-specific pattern of methylation which may shape the transcriptome in particular tissues.

recombination events, Alus are dispersed in a non-random manner in the human genome. By virtue of these characteristics, Alus harboring regulatory sites could have shaped novel regulatory networks [56, 72]. The possible non-coding regulatory networks in which Alus could be involved extensively are discussed below.

Alus and TFBS

Older Alu subfamilies, with passage of time, have accumulated mutations that have led to creation of TFBS for many nuclear receptors. This has resulted in widespread distribution of RNA pol II regulatory sites from these RNA pol III dependant elements [72]. The plasticity or specificity has also been contributed by different Alu subfamilies, with an important role of variable midA-stretch joining the two monomers. Evidences have demonstrated that Alu elements have evolved to harbor hormone response elements (HRE) that overlap with the

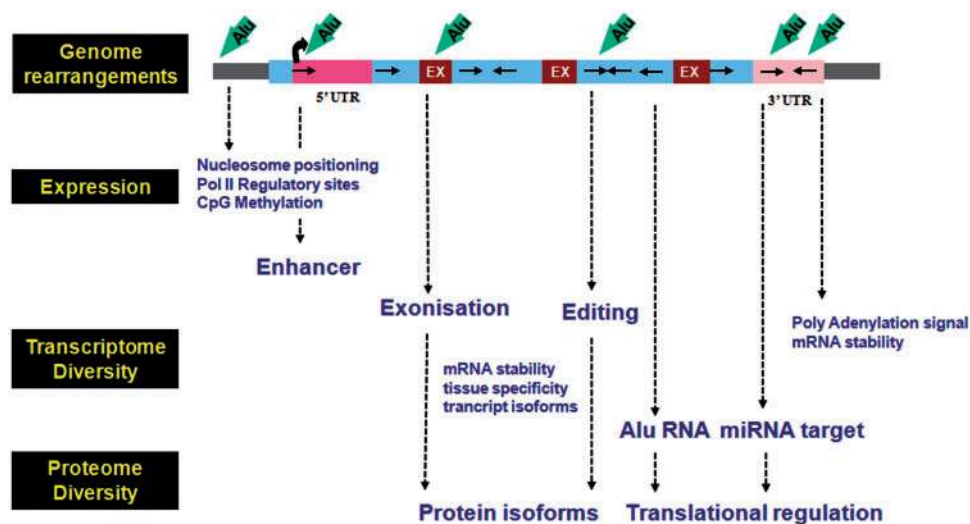


Figure 4: Regulatory repertoire of Alu shaped by its preferential and non-random genomic distribution. All the four-tiers of regulation, namely, genome rearrangement/shuffling by Alu mediated non-homologous recombination, expression by chromatin modulation through Alu nucleosome, transcriptome through Alu-mediated transcript diversity and finally at the proteome level by Alu exons. All of this contributes and co-ordinate with each other to modulate human genomic diversity.

internal promoter of the Alu. These several hexameric sequences are related to a consensus binding site, AGGTCA, recognizable by members of the nuclear receptor super-family of ligand-activated TFs, including receptors for retinoic acid (RAR) [97], thyroid hormone (T3) (THR) [98], vitamin D (VDR) [99], progesterone [71], estrogen [100], gluco-corticoid [101], growth hormone [102], steroids and a number of orphan receptors. It has been recently shown that ~90% of retinoic acid response element (RARE) sites reside within Alu and 95.5% of these are present within Alu S subfamily [103]. Many of the putative binding sites within Alu correspond to TFs associated with early markers of development processes [104]. Alus also harbor many orphan receptors like liver X receptors (LXRs) or *PPAR α* , which are members of nuclear receptor super-family that play important role in functions like lipid homeostasis and regulation of pro-inflammatory genes, respectively [105, 106].

In recent times, there have been interesting reports highlighting the role of TFBS within Alu in modulating immune response. Human cathelicidin antimicrobial peptide (*CAMP*) gene is the direct target of Vitamin D receptor and is strongly up-regulated in response to 1, 25-dihydroxyvitamin D₃. The binding site (BS) for vitamin D receptor is present within the Alu repeat present upstream of the gene [99]. This BS shows primate-specific purifying

selection of the exapted Alu exclusively in human and primate lineages. Evolutionary selection to place the *CAMP* gene under regulation of vitamin D pathway in primates potentiates the innate immune response and may counter the anti-inflammatory properties of vitamin D [107]. It also helps to emphasize the point that presence of TFBS within Alu is a non-random phenomenon as it provides functional advantage to the host and is under positive selection. Exaptation of Alu as an additional promoter for human neuronal apoptosis inhibitory protein (*NAIP*) gene has led to a novel protein isoform of *NAIP* [108]. Recent studies have also highlighted that *p53* binding site within Alu are generated by methylation and de-amination of CpGs on the genomic scale [70, 109]. It has been recently shown that an ankyrin-repeat containing protein p200 from an endoparasite *Ehrlichia chaffeensis* translocates to host nucleus and binds to Alu Sx elements in promoter and intronic regions of genes. These are enriched in biological processes such as transcription, apoptosis, ATPase activity and structural proteins associated with the nucleus and membrane-bound organelles [110]. Similarly, origin of *NF- κ B* BS from Alu families in the proximal promoter region of the *IFN-1* regulatory region also emphasize the role of Alu in innate immune response, in response to bacterial lipo poly-saccharide (LPS) membrane [111].

Transcribed Alu: Alu RNA

During normal cellular state, Alu RNAs accumulate at very low levels 10^3 – 10^4 molecules/cell but there is transient increase of nearly 20-fold in the levels of Alu RNA during stress, viral infection and cancer [112, 113]. Such pol III transcribed Alu RNA are referred to as ‘free Alu RNAs’. Alu is also transcribed as part of pol II transcription units of protein- and non-protein coding RNAs, termed as ‘embedded Alu RNAs’. This is a by-product of integration of Alu into various locations in the human genome [114]. As mentioned earlier, transcribed Alus contribute nearly 35% (Figure 1) of transcribed repetitive elements but an estimation of exact proportion of Alus has not been determined yet. It is compounded by the fact that due to the repetitive nature of Alu, it is not possible to map the transcribed free Alu RNA although embedded Alu RNAs can be mapped back to their genomic location. Free Alu RNA and embedded Alu RNA are regulated by different set of transcription rules governing pol III and pol II promoters. Pol III promoter element in older subfamilies have accumulated mutations and hence have reduced capability of driving transcription but in the process have gained pol II regulatory signals. Whereas, though newer subfamilies have less of pol II regulatory sites, their pol III promoters are in functional state [72]. Thus, older subfamilies would be transcribed more as embedded Alu RNA and newer subfamily as free Alu RNA. Most of the embedded Alus are preferentially present in the 3'UTR region of transcripts. Both free and embedded Alus play an important role in transcription and translational modulation. Alu RNA modulates functions like protein translation, alternative splicing and A–I editing. During heat shock stress, elevated Alu RNA acts as transcriptional co-repressor through direct interaction with RNA pol II in the promoter regions of heat shock responsive genes [115]. Conversely, the level of Alu RNA is down-regulated in human macrophages infected with *Leishmania*, a parasitic protozoan, during its initial stage of infection. The down-regulation is achieved through degradation of TFIIC110, which is an essential component of RNA pol III mediated transcription of Alu via B-box promoter element [116]. It is possible that one of the purposes of knocking down of ncRNA gene expression like Alu RNA and 7SL RNA in macrophages would be to alleviate ncRNA mediated transcriptional arrest in these cells so that the parasites can establish infection as well as proliferate. Alu

RNAs assemble with cellular proteins into ribonucleoprotein complexes and can be processed into the smaller scAlu RNAs [117]. The potential of Alu RNA as translational modulator through PKR (protein Kinase activated by RNA)-dependent mechanism has also been shown [118].

It is also equally plausible that ncRNA transcription in general and Alu transcription in particular could be non-functional. The skepticism about the functionality of pervasive transcription, mainly contributed by ncRNA has led to studies trying to understand the veracity of this phenomenon. These studies conclude that most ‘dark matter’ transcripts have low abundance than the known exons and that the genome is not as pervasively transcribed as widely perceived [119]. There has also been vigorous debate about what constitutes evidence for functional role of thousands of ncRNA loci which are located outside the protein-coding genes [120]. Possibly, this can only be answered through cellular phenotypes arising out of disruption of non-coding loci.

Alu Exonization and A–I Editing

In addition to the transcriptional landscape shaped by TFBS within Alu, cellular milieu is also replete with other important genomic variables shaped by Alu which could contribute to transcriptome diversity. Alu consensus sequence harbor 9 potential 5' and 14 potential 3' splice site, with most of them being present in the minus strand [56, 114, 121, 122]. This could potentiate Alu exonization leading to creation of novel splice isoforms. This phenomenon termed as ‘Alu alternative splicing’, exhibits tissue-specific expression patterns [78, 123]. An estimate says that ~5.2% of all identified alternative exons is derived from Alu elements. Alu exonization also generates new exons with altered mRNA translational efficiency. This could have global effects in primates through altered regulation of protein production of molecules, such as master transcriptional regulators in specific lineages [124]. Two Alu RNAs when present in a head to tail orientation in exonized transcripts are also substrate for A–I editing mediated by adenosine deaminase that specifically target double-stranded RNAs (*ADAR*) [125]. This process of RNA editing involves co-transcriptional or post-transcriptional modification of RNA, most prevalent being the hydrolytic deamination of adenosine (A) to inosine (I). >90% of all A–I editing occurs within Alus with ‘A’ at positions 27, 28, 136 and 162 being more susceptible to such editing,

compared to others [126]. Primates have high level of such editing occurring in their genomes. This opens up the avenue for speculating that appearance of Alu in the primate lineage led to such conspicuous editing, paving the path for primate evolution [127]. Inosine pairs with cytosine, not uracil, thereby creating novel secondary structure of RNA. Enhanced editing levels of Alus have been observed in human brain. This may add to the known significant enrichment of edited transcripts in the brain and their involvement in regulation of brain-specific transcripts [128]. Association of edited Alus with neuronal functions hints at the possible contribution of A-I edited transcripts in the development of higher brain function. Alu editing may serve as an alternate information mechanism based on the binary A/I code [129]. Knockdown experiments of *ADAR1* have shown high editing levels in non-coding sites of Alu in hESCs. This reverses to global decrease in editing levels during differentiation, particularly into the neural lineage. This highlights the role of A-I editing in development [130].

Alus and NATs

NATs are endogenous RNA molecules containing sequences that are complementary to other transcripts. NATs primarily belong to two subcategories, *cis*-NATs, which are transcribed from opposing DNA strands at the same genomic locus, and *trans*-NATs, which are transcribed from separate loci. NATs reflect pervasive nature of transcription from bacteria to human, which suggests that they are a fundamental component of gene regulation [131, 132]. In human, ~22% of transcripts form *cis* sense-antisense pairs [133] and *trans*-NATs comprise ~4% of the transcriptional units [134]. The size of NATs is variable ranging from a few bases to several kilobases [135]. Sense-antisense transcripts tend to be co-expressed or inversely expressed more frequently than would be expected by chance [136–138]. The humongous levels of NAT are also facilitated by (i) short introns in evolutionarily conserved antisense genes [139] and (ii) transcription of pseudogenes [140].

Genome-wide computation study shows that transcriptional start sites (TSSs) of NATs are overtly present within transposable elements (TEs) with as many as 48 718 human gene antisense TSS within TEs [141]. Interestingly, such TSSs are mostly present within 3' ends of the genes which are also most favored sites for Alu integration. Although it lacks

experimental evidence yet it is tempting to speculate that relative excess of antisense transcripts initiated from TEs and their enrichment closer to the 3' ends of TUs might be contributed, in part by Alus. This may yield *cis*-NATs with biologically significant regulatory activities as Alus harbor TFBS which may initiate transcription. Such circumstantial evidence awaits experimental validation at the genomic scale but evidence of human-specific antisense transcripts like *RNF144A*, *SYNE2* and *CAMCK4* induced by insertion of TEs is an indication in the right direction. Such NATs may have played a role in acquisition of human-specific traits [142]. Abundance of Alu in the human transcriptome especially in the 3'UTR of genes and Alus ability to promote transcription potentiates its involvement in NAT mediated regulation [141]. Besides, Alus as part of antisense transcriptome has already been reported. Antisense Alus belong to older subfamily and show significantly greater sequence divergence from their consensus sequence than Alus that do not co-locate with TSS of antisense transcripts [141].

NAT may mediate regulation through transcriptional interference, RNA masking, dsRNA-dependent mechanisms and RNA interference, and antisense-mediated methylation and mono-allelic expression. One way by which *cis*-NAT can mediate regulation could be through transcriptional interference due to collision between two bulky RNA pol II complexes transcribing from opposite strands [143]. The other mechanisms could involve RNA masking which may hinder processes that require protein-RNA interactions such as splicing, mRNA transport, polyadenylation, translation and degradation. For example, antisense transcript-mediated masking of splice-site in thyroid hormone receptor gene *erbA α* results in shift of balance between the two splice variants [144]; *Wrap53*, a NAT to *p53*, mediate *p53* induction in response to DNA damage by targeting the 5'UTR of *p53* mRNA [145]; regulation of *XIST* by an antisense lncRNA *Tsix* leading to X chromosome activation [146]. Alu element harboring TFBS can activate transcription from pol II promoter as well as downstream regions. Besides, Alus can also potentiate transcription of Alu RNA when present in antisense orientation as well as provide TFBS for antisense transcription. This could not only result in transcriptional interference but also sense-antisense transcript pairing in *cis*. Due to their ability to be exonized in antisense orientation in a gene, Alu exonized transcripts could be

substrates for a transcribing sense Alu from opposite strand or in *trans*. Alu-mediated genomic rearrangements leading to formation of pseudogenes may also help to enhance its NAT-mediated regulation.

Alus and miRNAs

Alus seem to share many features with miRNAs, the most characterized of the ncRNAs. These ~22 nt short ncRNAs that now seem to be essential for rapid and transient response to external or internal changes are involved in wide variety of functions like stress response, cell fate determination, morphogenesis regulation, synaptic plasticity, apoptosis, mRNA splicing, DNA methylation, circadian rhythms, angiogenesis control, cell cycle control, endocrinological regulation, immune-modulation and neuro-protection [147–161]. The regulatory potential and specificity are also enhanced by editing of miRNA sequence as well as in the target site of the miRNA [162–165]. miRNA majorly regulate at post-transcriptional level by affecting mRNA stability after binding to 3'UTR region of the transcripts or at the translational level [166–168]. Each miRNA on average targets nearly 200 mRNA sequences and overall ~40% of human genes [169].

In recent times, there have been bioinformatics predictions for origin of mammalian miRNAs from genomic repeats as well as miRNAs targeting the repeat elements. miRNAs have been mostly characterized to have target sites in the 3'UTR region of the transcripts. Alu exonization and especially its enrichment in the 3'UTR regions (84%) and the similarity of secondary structure of Alu RNA with the miRNA precursor potentiate Alu RNA in biogenesis of miRNA. Besides, presence of potential miRNA target sites within Alu also increase the possibility of Alu involvement in miRNA-based regulatory networks [170]. A–I editing of exonized transcripts adds another dimension as editing within miRNA or the target region of miRNA have been shown to increase the functional repertoire of miRNA. An *in silico* analysis revealed that around 30 human miRNAs which include paralogous miRNAs like miR-17-5p, miR-20a, miR-20b, miR-93, miR-106a and miR-106b, are implicated in cancer and show 5' seed complementarity against a specific site in the Alu sequences, which is highly conserved [83, 171]. Interestingly, several reports indicate that mRNAs containing Alu in their 3'UTRs are as a class associated with growth and differentiation and are subject to translational regulation. It has

also been reported that there has been co-evolution of Alu and miRNAs in the humans with probable feedback mechanism helping to maintain homeostatic state [172]. This dual relationship is postulated on the premise that duplication events involving Alu facilitated growth of miRNA cluster and the expression of miRNAs. On the other hand, the expressed miRNAs target free Alu RNAs to sequester the level of Alu RNA. The probability of cross-talk was increased further with the discovery that in addition to pol II, miRNAs can also be transcribed by pol III as in case of C19MC where miRNAs are transcribed as part of the poly-cistronic transcript along with Alu by RNA pol III [173, 174]. In the absence of rigorous experimental evidence, it is amply clear that most of the Alu–miRNA links are circumstantial and it may be by chance. Thus, one has to have more than usual stringent controls to understand the functionality of the probable cross-talk.

Compartmentalization of transcripts: Alu involvement?

Compartmentalization of transcripts, active or passive, allows cellular functions to be separated in specialized organelles and provides an additional level of gene regulation. The compartmentalization is closely related to the export kinetics of mRNA from the nucleus to the cytosol and the degradation rate of mRNA in the cytosol [175]. Defects in mRNA export can result in human diseases [176]. Recent study by Barthelson *et al.*, wherein nuclear and cytoplasmic transcripts were studied separately observed that both the compartments contain different, yet overlapping populations of transcripts with nuclear RNA especially enriched in non-coding sequences comprising of 41% intergenic and 25% intronic sequences [177, 178]. Differential compartmentalization of alternate transcript isoforms across cell types has been observed with cytosolic fraction having more and longer introns in their pre-mRNAs, more functional RNA folds and unique exons in the 3'-regions whereas nucleus enriched isoforms are more significantly associated with NMD [179]. Sub-cellular distribution of small RNAs has revealed that majority of miRNAs are imported to the nucleus with almost equal distribution of miRNAs between nuclear and cytosolic fraction [180].

The segregation/targeting are facilitated by particular signature sequences associated with individual transcripts. These signatures could be residing in Alus. Although, it has not been studied but it

would be worthwhile to explore the importance of 3'-UTR Alu integration, its potential to form secondary structures as embedded Alu RNA, subsequent A-I editing and retention of A-I edited transcripts. It has been reported that ~75% of A-I editing in transcriptome involves Alu repeats [181]. A-I hyper-edited transcripts are preferentially retained in the nucleus in spatio-temporal manner in membrane-less, transient but discrete ribonucleoprotein structure called 'paraspeckles'. *NEAT1* RNA is essential, along with *PSP1*, *p54* and other factors regulating A-I editing to initiate the de-novo assembly of paraspeckles. Interestingly, *NEAT1* RNA is not A-I edited which is consistent with its function of sequestering A-I edited transcripts [182, 183]. Retention of A-I was thought to be a norm rather than exception, with variable retention pattern in different tissues reflecting the editing levels [184] until altered retention pattern in hESCs were observed in comparison to the differentiated cells [185].

Future perspectives: Alu elements as 'ornaments' of human genome

With the paucity of experimental platforms to query for the functionality of repetitive elements in the human genome, systematic genome-wide studies for actively looking into the role of Alu at the global scale has not been carried out. Most of the observations have been serendipitous. Genome-wide study platforms like microarray and tiling array are devoid of probes to quantify ncRNA transcripts because of the technical limitations related to specificity and conservation. But with the availability of next-generation sequencing (NGS) platforms, it has been possible to undertake genome-wide studies involving RNA-seq and ChIP-seq to understand the role of non-coding part of the genome at the RNA and DNA level, respectively. NGS would also be of great help to explore RNA-protein interactions through sequencing of RNA-immunoprecipitated (RIP) samples.

From its initial days, when it was summarily rejected as 'genomic scrap', polluting and congesting the human genome; it has now been accepted and explored as 'genomic gems'. It is evident that Alus as non-coding components of the human genome could modulate transcriptome in diverse ways. This would range from providing TFBS, their specific regulation, exonization, editing and involvement in antisense to miRNA regulatory networks.

Additionally, through recombination they could further distribute these regulatory sites. Do these networks cross-talk and where could they assume importance? Parallel and peripheral information point toward the fact that ncRNA may be part of functional redundancy repertoire as a built-in safeguard for maintaining accurate regulation of the genome. Most of the characterized functional role of Alu repeats has been reported under different condition of stress, thus it is likely that Alus offer an alternative 'safe exit' route for survival during stress. It has also been reported that heat shock response in mammalian cells is RNA mediated [186]. Alus may be the unifying thread during different stress conditions as exemplified by the instances enumerated below. As Alus seem to be a partner in miRNA mediated regulatory networks, it would be worthwhile to explore Alu-miRNA cross-talk. During stress a concomitant transient increase in the level of Alu RNA along with heat shock proteins (HSPs) takes place till 12h and then declines. Identification of diverse array of functional TFBS within Alu raises the possibility of Alu also harboring heat shock factor (HSF) binding site which could induce expression of Alu RNA and also heat shock responsive genes directly. Alu RNA may mediate complementary binding and sequester the expression levels of sense transcripts. Though Alu antisense transcripts have been reported to be abundant in the human genome, there are very few reports probing functional relevance of these antisense transcripts. It is possible that these transcripts are transcribed in response to rapidly changing environmental stimuli. The other possibility could be through a direct cross talk between pol II and pol III machinery in stress as Alu RNA levels are elevated and they are also known to act as transcriptional co-repressor. It would be important to mention that Alu is not itself recognized by pol III machinery instead 'B-Box' promoter is recognized. The transcriptional machinery thus transcribes both 'free' and 'embedded' Alu RNAs. Recent observations show close association of pol III and many TFs, like c-Fos, c-Jun and c-Myc that are otherwise tightly associated with pol II genes [187]. The preferential association of human pol III complexes near functional pol II promoters indicates that TFIIC mediated recruitment of TFIIB is regulated in a locus-specific manner. As Alu 'B-Box' is recognized by the pol III machinery, specifically the TFIIC complex to initiate its transcription, these elements might be

involved in such cross-talk. On the other hand, pol II modulating pol III activity is exemplified by the presence of pol II epigenetic marks, like histone acetylation with pol III transcribed ncRNA genes. It has been observed that pol III bound and expressed ncRNA are marked by H3K4me1/2/3 modifications. The widespread co-occurrence of pol II and pol III perhaps reflect much larger role of this phenomenon in organizing the human genome into discrete functional domains and TE like Alu could play such 'insulator roles' [188, 189]. Alu has already been shown to confer such position-independent expression in transgenic mice.

Since Alus are predicted to be targets for miRNAs, they can affect the function of miRNA through titration of miRNA by providing alternate target sites. miRNA mediated post-transcriptional degradation of non-canonical transcript isoforms made in stress, may be one of the possible mechanisms by which such transcripts can be cleared from the cellular milieu when the gene product is not of immediate use [190]. Such transcripts can be earmarked by the presence of exonized Alu, which may act like 'miRNA sponges' to protect the genome against the increased levels of miRNA in any particular condition/stimuli. With evidence of pseudogene, like *PTENP1*, regulating the corresponding protein-coding mRNA by acting as a decoy for miRNA that bind to the common site in the 3'UTR region, such a function mediated by Alu is also possible [32, 191]. Localization of miRNA associated Argonaute to the stress granules (SG), which are temporarily formed foci for undecided transcripts during condition of stress, also highlights the close association between miRNA and stress [147, 192].

Stress response is also mediated through the activation of signaling pathways which lead to phosphorylation of target splicing factors and change their sub-cellular distribution, activity and/or association with multi-protein complexes. In response to stress, *hmRNP A1*, a nucleo-cytoplasmic shuttling protein that antagonizes serine/arginine-rich (SR) proteins during alternative splicing, accumulates in the cytoplasm localizing to the SGs. This accumulation causes an altered ratio of the antagonistic alternative splicing factors SF2/ASF and *hmRNP A1* in the nucleus and consequently affects alternative splicing regulation. Similarly, cell stress induces phosphorylation of the splicing factor *RBM4* (RNA-binding motif protein 4) and drives its cytoplasmic

accumulation and targeting to SGs, via the MKK3/6-p38 signaling pathway, where it inhibits translation. Overall it seems that *hmRNP A1* and *RBM4* are not the only exception but stress stimuli also influence the sub-cellular localization of several other RNA binding proteins, including the second step splicing factor *hSlu7*, required for correct 3' splice-site choice [193, 194] and *HuR* (human antigen R) [195]. These splicing factors may act on cryptic splice sites within Alu and add it as part of the coding exon. This may determine the fate of the transcripts, e.g. its retention in nucleus and compartmentalization or differential sensitivity to stress response. A summary of the validated and proposed regulatory functions mediated by Alu is present in Figure 5.

Finally, the increased levels of Alu RNA leading to toxicity of the cell in case of age-related macular degeneration has opened a totally new domain of research because till date it was only known that Alu RNA levels were either elevated or down-regulated. It has been shown that *staufen-1 (STAU1)*-mediated mRNA decay requires binding of *STAU1* to the 3'UTR of translationally active mRNA. This binding is facilitated by the presence of Alu in 3'UTR of mRNA and another Alu element in a cytoplasmic, polyadenylated lncRNA [196]. Although this is not a universal phenomena but it reflects the genome plasticity contributed by Alu which may vary in time and space, different tissues, stimuli and interacting partners. The functional role of transcribed Alus is just being realized whereby new quanta of information are being strung together to understand the Alu RNA mediated regulation. The variable levels of Alu RNA during infection further substantiate its functional involvement in physiological homeostasis. The role of Alu in governing the dynamics of host-pathogen interaction would be an interesting avenue to explore. Human complexity and plasticity is characterized by its ability to adjust to different conditions. To achieve this, there must be stand-by scenario available which provides flexibility. The abundance of TE like Alu may capacitate the human genome with such flexibility.

Key points

- ncRNA may be part of functional redundancy repertoire as a built-in safeguard for maintaining accurate regulation of the genome.
- Differential levels of Alu RNA during stress conditions like heat shock, cancer and infection potentiates it for functional role in host-pathogen interaction.

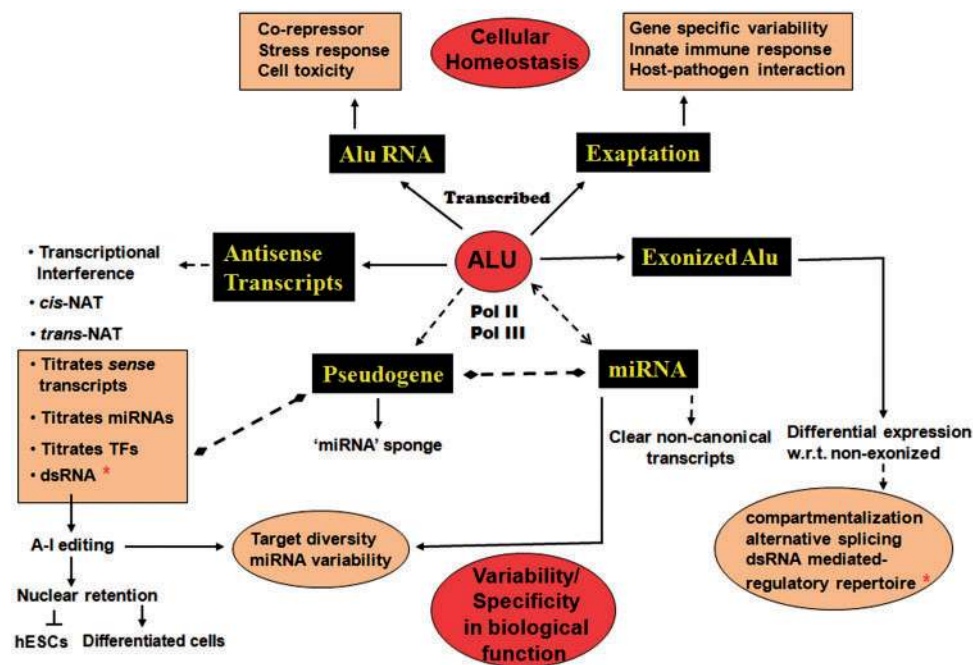


Figure 5: Transcriptomic diversity leading to genome plasticity and complexity being modulated and mediated by transcribed Alus. Arrows with continuous line represent known experimental/computational facts whereas dotted lines are plausible functions based on peripheral and parallel evidences. Asterisk on dsRNA represent shared functions. Alu repeats seem to be an integral component of genome regulatory repertoire particularly for maintaining cellular homeostasis during stress and also for variability/specificity in cellular functions.

- Possible direct cross-talk between pol II and pol III machinery during condition of stress which may be mediated by the presence of binding sites for both within Alu repeats.
- Probable miRNA targets within Alu potentiate the role of Alu as 'miRNA sponge' in response to external/internal stimuli.
- Alu exonised transcripts may determine the compartmentalization of transcripts across nucleo-cytosolic fractions.

Acknowledgement

Financial support to M.M. and R.P. from Council of Scientific and Industrial Research (CSIR) grant (NWP-0036) and Council of Scientific and Industrial Research-Senior Research Fellowship (CSIR-SRF to R.G.) is acknowledged. The authors would like to acknowledge Amit K. Mandal for valuable discussions.

FUNDING

Financial support to M.M. and R.P. from Council of Scientific and Industrial Research (CSIR) grant (NWP-0036) and Council of Scientific and Industrial Research-Senior Research Fellowship (CSIR-SRF to R.P.) is duly acknowledged.

References

1. Huttenhofer A, Schattner P, Polacek N. Non-coding RNAs: hope or hype? *Trends Genet* 2005;**21**:289–97.

2. Prasanth KV, Spector DL. Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev* 2007;**21**:11–42.
3. Eddy SR. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet* 2001;**2**:919–29.
4. Lander ES, Linton LM, Birren B, *et al.* Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921.
5. Taft RJ, Pheasant M, Mattick JS. The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays* 2007;**29**:288–99.
6. Mattick JS. The functional genomics of noncoding RNA. *Science* 2005;**309**:1527–8.
7. Hannon GJ. RNA interference. *Nature* 2002;**418**:244–51.
8. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 2004;**5**:522–31.
9. Lipovich L, Johnson R, Lin CY. MacroRNA underdogs in a microRNA world: evolutionary, regulatory, and biomedical significance of mammalian long non-protein-coding RNA. *Biochim Biophys Acta* 2010;**1799**:597–615.
10. Blencowe BJ. Transcription: surprising role for an elusive small nuclear RNA. *Curr Biol* 2002;**12**:R147–9.
11. Yang Z, Zhu Q, Luo K, *et al.* The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* 2001;**414**:317–22.
12. Friedman RC, Farh KK, Burge CB, *et al.* Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009;**19**:92–105.

13. Chodroff RA, Goodstadt L, Sirey TM, *et al.* Long noncoding RNA genes: conservation of sequence and brain expression among diverse amniotes. *Genome Biol* 2010;**11**:R72.
14. Riethoven JJ. Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. *Methods Mol Biol* 2010;**674**: 33–42.
15. Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 2002;**415**:810–3.
16. Leeb M, Steffen PA, Wutz A. X chromosome inactivation sparked by non-coding RNAs. *RNA Biol* 2009;**6**:94–9.
17. Goodrich JA, Kugel JF. From bacteria to humans, chromatin to elongation, and activation to repression: the expanding roles of noncoding RNAs in regulating transcription. *Crit Rev Biochem Mol Biol* 2009;**44**:3–15.
18. Okamoto I, Arnaud D, Le BP, *et al.* Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* 2005;**438**:369–73.
19. Rinn JL, Kertesz M, Wang JK, *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 2007;**129**:1311–23.
20. Duret L, Chureau C, Samain S, *et al.* The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science* 2006;**312**:1653–5.
21. Johnson R, Teh CH, Kumarso G, *et al.* REST regulates distinct transcriptional networks in embryonic and neural stem cells. *PLoS Biol* 2008;**6**:e256.
22. Sheik MJ, Gaughwin PM, Lim B, *et al.* Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA* 2010;**16**:324–37.
23. Zhou Y, Zhong Y, Wang Y, *et al.* Activation of p53 by MEG3 non-coding RNA. *J Biol Chem* 2007;**282**:24731–42.
24. Martianov I, Ramadass A, Serra BA, *et al.* Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* 2007;**445**:666–70.
25. Willingham AT, Orth AP, Batalov S, *et al.* A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* 2005;**309**:1570–3.
26. Goodrich JA, Kugel JF. Dampening DNA binding: a common mechanism of transcriptional repression for both ncRNAs and protein domains. *RNA Biol* 2010;**7**: 305–9.
27. Tufarelli C, Stanley JA, Garrick D, *et al.* Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet* 2003;**34**: 157–65.
28. Allen L, Kloc M, Etkin LD. Identification and characterization of the Xsirt cis-acting RNA localization element. *Differentiation* 2003;**71**:311–21.
29. Gripenland J, Netterling S, Loh E, *et al.* RNAs: regulators of bacterial virulence. *Nat Rev Microbiol* 2010;**8**:857–66.
30. Prasanth KV, Rajendra TK, Lal AK, *et al.* Omega speckles – a novel class of nuclear speckles containing hnRNPs associated with noncoding hsr-omega RNA in *Drosophila*. *J Cell Sci* 2000;**113**(Pt 19):3485–97.
31. Sandqvist A, Sistonen L. Nuclear stress granules: the awakening of a sleeping beauty? *J Cell Biol* 2004;**164**:15–7.
32. Poliseno L, Salmena L, Zhang J, *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 2010;**465**:1033–38.
33. Maki K, Morita T, Otaka H, *et al.* A minimal base-pairing region of a bacterial small RNA SgrS required for translational repression of ptsG mRNA. *Mol Microbiol* 2010;**76**: 782–92.
34. Wadler CS, Vanderpool CK. A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc Natl Acad Sci USA* 2007;**104**:20454–9.
35. Jenny A, Hachet O, Zavorszky P, *et al.* A translation-independent role of oskar RNA in early *Drosophila* oogenesis. *Development* 2006;**133**:2827–33.
36. Colley SM, Leedman PJ. SRA and its binding partners: an expanding role for RNA-binding coregulators in nuclear receptor-mediated gene regulation. *Crit Rev Biochem Mol Biol* 2009;**44**:25–33.
37. Hube F, Velasco G, Rollin J, *et al.* Steroid receptor RNA activator protein binds to and counteracts SRA RNA-mediated activation of MyoD and muscle differentiation. *Nucleic Acids Res* 2011;**39**:513–25.
38. Candeias MM, Malbert-Colas L, Powell DJ, *et al.* P53 mRNA controls p53 activity by managing Mdm2 functions. *Nat Cell Biol* 2008;**10**:1098–105.
39. Hube F, Guo J, Chooniedass-Kothari S, *et al.* Alternative splicing of the first intron of the steroid receptor RNA activator (SRA) participates in the generation of coding and noncoding RNA isoforms in breast cancer cell lines. *DNA Cell Biol* 2006;**25**:418–28.
40. Lottin S, Adriaenssens E, Dupressoir T, *et al.* Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis* 2002;**23**: 1885–95.
41. Iacoangeli A, Lin Y, Morley EJ, *et al.* BC200 RNA in invasive and preinvasive breast cancer. *Carcinogenesis* 2004;**25**: 2125–33.
42. Tam W, Dahlberg JE. miR-155/BIC as an oncogenic microRNA. *Genes Chromosomes Cancer* 2006;**45**:211–2.
43. Iorio MV, Ferracin M, Liu CG, *et al.* MicroRNA gene expression deregulation in human breast cancer. *Cancer Res* 2005;**65**:7065–70.
44. Bussemakers MJ, van BA, Verhaegh GW, *et al.* DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res* 1999;**59**:5975–9.
45. Srikantan V, Zou Z, Petrovics G, *et al.* PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. *Proc Natl Acad Sci USA* 2000;**97**:12216–21.
46. Pibouin L, Villaudy J, Ferbus D, *et al.* Cloning of the mRNA of overexpression in colon carcinoma-1: a sequence overexpressed in a subset of colon carcinomas. *Cancer Genet Cytogenet* 2002;**133**:55–60.
47. Lukiw WJ, Handley P, Wong L, *et al.* BC200 RNA in normal human neocortex, non-Alzheimer dementia (NAD), and senile dementia of the Alzheimer type (AD). *Neurochem Res* 1992;**17**:591–7.
48. Millar JK, James R, Brandon NJ, *et al.* DISC1 and DISC2: discovering and dissecting molecular mechanisms underlying psychiatric illness. *Ann Med* 2004;**36**:367–78.
49. Mutsuddi M, Marshall CM, Benzow KA, *et al.* The spinocerebellar ataxia 8 noncoding RNA causes neurodegeneration and associates with staufen in *Drosophila*. *Curr Biol* 2004;**14**:302–8.
50. Vincent JB, Petek E, Thevarkunnel S, *et al.* The RAY1/ST7 tumor-suppressor locus on chromosome 7q31

- represents a complex multi-transcript system. *Genomics* 2002;**80**:283–94.
51. Jong MT, Gray TA, Ji Y, *et al.* A novel imprinted gene, encoding a RING zinc-finger protein, and overlapping antisense transcript in the Prader-Willi syndrome critical region. *Hum Mol Genet* 1999;**8**:783–93.
 52. Wevrick R, Kerns JA, Francke U. Identification of a novel paternally expressed gene in the Prader-Willi syndrome region. *Hum Mol Genet* 1994;**3**:1877–1882.
 53. Sutherland HF, Wadey R, McKie JM, *et al.* Identification of a novel transcript disrupted by a balanced translocation associated with DiGeorge syndrome. *Am J Hum Genet* 1996;**59**:23–31.
 54. Sonkoly E, Bata-Csorgo Z, Pivarcsi A, *et al.* Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, PRINS. *J Biol Chem* 2005;**280**:24159–67.
 55. Mattick JS, Gagen MJ. Mathematics/computation. Accelerating networks. *Science* 2005;**307**:856–8.
 56. Grover D, Kannan K, Brahmachari SK, *et al.* ALU-ring elements in the primate genomes. *Genetica* 2005;**124**:273–89.
 57. Ullu E, Tschudi C. Alu sequences are processed 7SL RNA genes. *Nature* 1984;**312**:171–2.
 58. Fuhrman SA, Deininger PL, LaPorte P, *et al.* Analysis of transcription of the human Alu family ubiquitous repeating element by eukaryotic RNA polymerase III. *Nucleic Acids Res* 1981;**9**:6439–56.
 59. Perez-Stable C, Shen CK. Competitive and cooperative functioning of the anterior and posterior promoter elements of an Alu family repeat. *Mol Cell Biol* 1986;**6**:2041–52.
 60. Weiner AM, Deininger PL, Efstratiadis A. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu Rev Biochem* 1986;**55**:631–61.
 61. Cordaux R, Hedges DJ, Herke SW, *et al.* Estimating the retrotransposition rate of human Alu elements. *Gene* 2006;**373**:134–7.
 62. Shen MR, Batzer MA, Deininger PL. Evolution of the master Alu gene(s). *J Mol Evol* 1991;**33**:311–20.
 63. Britten RJ. Evidence that most human Alu sequences were inserted in a process that ceased about 30 million years ago. *Proc Natl Acad Sci USA* 1994;**91**:6148–50.
 64. Englander EW, Howard BH. Nucleosome positioning by human Alu elements in chromatin. *J Biol Chem* 1995;**270**:10091–96.
 65. Tanaka Y, Yamashita R, Suzuki Y, *et al.* Effects of Alu elements on global nucleosome positioning in the human genome. *BMC Genomics* 2010;**11**:309.
 66. Bettecken T, Frenkel ZM, Trifonov EN. Human nucleosomes: special role of CG dinucleotides and Alu-nucleosomes. *BMC Genomics* 2011;**12**:273.
 67. Bollati V, Galimberti D, Pergoli L, *et al.* DNA methylation in repetitive elements and Alzheimer disease. *Brain Behav Immun* 2011;**25**:1078–83.
 68. Fabris S, Bollati V, Agnelli L, *et al.* Biological and clinical relevance of quantitative global methylation of repetitive DNA sequences in chronic lymphocytic leukemia. *Epigenetics* 2011;**6**:188–94.
 69. Witherspoon DJ, Watkins WS, Zhang Y, *et al.* Alu repeats increase local recombination rates. *BMC Genomics* 2009;**10**:530.
 70. Cui F, Sirotin MV, Zhurkin VB. Impact of Alu repeats on the evolution of human p53 binding sites. *Biol Direct* 2011;**6**:2.
 71. Jacobsen BM, Jambal P, Schittone SA, *et al.* ALU repeats in promoters are position-dependent co-response elements (coRE) that enhance or repress transcription by dimeric and monomeric progesterone receptors. *Mol Endocrinol* 2009;**23**:989–1000.
 72. Shankar R, Grover D, Brahmachari SK, *et al.* Evolution and distribution of RNA polymerase II regulatory sites from RNA polymerase III dependant mobile Alu elements. *BMC Evol Biol* 2004;**4**:37.
 73. Oei SL, Babich VS, Kazakov VI, *et al.* Clusters of regulatory signals for RNA polymerase II transcription associated with Alu family repeats and CpG islands in human promoters. *Genomics* 2004;**83**:873–82.
 74. Pastor T, Talotti G, Lewandowska MA, *et al.* An Alu-derived intronic splicing enhancer facilitates intronic processing and modulates aberrant splicing in ATM. *Nucleic Acids Res* 2009;**37**:7258–67.
 75. Schwartz S, Gal-Mark N, Kfir N, *et al.* Alu exonization events reveal features required for precise recognition of exons by the splicing machinery. *PLoS Comput Biol* 2009;**5**:e1000300.
 76. Schmitz J, Brosius J. Exonization of transposed elements: a challenge and opportunity for evolution. *Biochimie* 2011.
 77. Sakurai M, Yano T, Kawabata H, *et al.* Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. *Nat Chem Biol* 2010;**6**:733–40.
 78. Lin L, Shen S, Tye A, *et al.* Diverse splicing patterns of exonized Alu elements in human tissues. *PLoS Genet* 2008;**4**:e1000225.
 79. Athanasiadis A, Rich A, Maas S. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol* 2004;**2**:e391.
 80. Vorechovsky I. Transposable elements in disease-associated cryptic exons. *Hum Genet* 2010;**127**:135–54.
 81. Gu Y, Kodama H, Watanabe S, *et al.* The first reported case of Menkes disease caused by an Alu insertion mutation. *Brain Dev* 2007;**29**:105–8.
 82. Kim DS, Hahn Y. Identification of human-specific transcript variants induced by DNA insertions in the human genome. *Bioinformatics* 2011;**27**:14–21.
 83. Smalheiser NR, Torvik VI. Alu elements within human mRNAs are probable microRNA targets. *Trends Genet* 2006;**22**:532–6.
 84. Lee JR, Huh JW, Kim DS, *et al.* Lineage specific evolutionary events on SFTP gene: Alu recombination-mediated deletion (ARMD), exonization, and alternative splicing events. *Gene* 2009;**435**:29–35.
 85. Roy-Engel AM, Carroll ML, Vogel E, *et al.* Alu insertion polymorphisms for the study of human genomic diversity. *Genetics* 2001;**159**:279–90.
 86. Arcot SS, Wang Z, Weber JL, *et al.* Alu repeats: a source for the genesis of primate microsatellites. *Genomics* 1995;**29**:136–44.
 87. Batzer MA, egria-Hartman M, Deininger PL. A consensus Alu repeat probe for physical mapping. *Genet Anal Tech Appl* 1994;**11**:34–8.
 88. Novick GE, Novick CC, Yunis J, *et al.* Polymorphic human specific Alu insertions as markers for human identification. *Electrophoresis* 1995;**16**:1596–601.

89. Novick GE, Gonzalez T, Garrison J, *et al.* The use of polymorphic Alu insertions in human DNA fingerprinting. *EXS* 1993;**67**:283–91.
90. Roy-Engel AM, El-Sawy M, Farooq L, *et al.* Human retroelements may introduce intragenic polyadenylation signals. *Cytogenet Genome Res* 2005;**110**:365–71.
91. Tanaka Y, Yamashita R, Suzuki Y, *et al.* Effects of Alu elements on global nucleosome positioning in the human genome. *BMC Genomics* 2010;**11**:309.
92. Kim PM, Lam HY, Urban AE, *et al.* Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history. *Genome Res* 2008;**18**:1865–74.
93. Tomilin NV, Bozhkov VM, Bradbury EM, *et al.* Differential binding of human nuclear proteins to Alu subfamilies. *Nucleic Acids Res* 1992;**20**:2941–5.
94. Xie H, Wang M, Bonaldo MF, *et al.* High-throughput sequence-based epigenomic analysis of Alu repeats in human cerebellum. *Nucleic Acids Res* 2009;**37**:4331–40.
95. Deininger PL, Batzer MA. Alu repeats and human disease. *Mol Genet Metab* 1999;**67**:183–93.
96. Callinan PA, Batzer MA. Retrotransposable elements and human disease. *Genome Dyn* 2006;**1**:104–15.
97. Vansant G, Reynolds WF. The consensus sequence of a major Alu subfamily contains a functional retinoic acid response element. *Proc Natl Acad Sci USA* 1995;**92**:8229–33.
98. Piedrafita FJ, Molander RB, Vansant G, *et al.* An Alu element in the myeloperoxidase promoter contains a composite SP1-thyroid hormone-retinoic acid response element. *J Biol Chem* 1996;**271**:14412–20.
99. Gombart AF, Borregaard N, Koeffler HP. Human cathelicidin antimicrobial peptide (CAMP) gene is a direct target of the vitamin D receptor and is strongly up-regulated in myeloid cells by 1,25-dihydroxyvitamin D₃. *FASEB J* 2005;**19**:1067–77.
100. Norris J, Fan D, Aleman C, *et al.* Identification of a new subclass of Alu DNA repeats which can function as estrogen receptor-dependent transcriptional enhancers. *J Biol Chem* 1995;**270**:22777–82.
101. Babich V, Aksenov N, Alexeenko V, *et al.* Association of some potential hormone response elements in human genes with the Alu family repeats. *Gene* 1999;**239**:341–49.
102. Goodyer CG, Zheng H, Hendy GN. Alu elements in human growth hormone receptor gene 5' untranslated region exons. *J Mol Endocrinol* 2001;**27**:357–66.
103. Laperriere D, Wang TT, White JH, *et al.* Widespread Alu repeat-driven expansion of consensus DR2 retinoic acid response elements during primate evolution. *BMC Genomics* 2007;**8**:23.
104. Polak P, Domany E. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. *BMC Genomics* 2006;**7**:133.
105. Reynolds WF, Kumar AP, Piedrafita FJ. The human myeloperoxidase gene is regulated by LXR and PPARalpha ligands. *Biochem Biophys Res Commun* 2006;**349**:846–54.
106. Li Y, Bolten C, Bhat BG, *et al.* Induction of human liver X receptor alpha gene expression via an autoregulatory loop mechanism. *Mol Endocrinol* 2002;**16**:506–14.
107. Gombart AF, Saito T, Koeffler HP. Exaptation of an ancient Alu short interspersed element provides a highly conserved vitamin D-mediated innate immune response in humans and primates. *BMC Genomics* 2009;**10**:321.
108. Romanish MT, Nakamura H, Lai CB, *et al.* A novel protein isoform of the multicopy human NAIP gene derives from intragenic Alu SINE promoters. *PLoS One* 2009;**4**:e5761.
109. Zemojtel T, Kielbasa SM, Arndt PF, *et al.* Methylation and deamination of CpGs generate p53-binding sites on a genomic scale. *Trends Genet* 2009;**25**:63–6.
110. Zhu B, Nethery KA, Kuriakose JA, *et al.* Nuclear translocated Ehrlichia chaffeensis ankyrin protein interacts with a specific adenine-rich motif of host promoter and intronic Alu elements. *Infect Immun* 2009;**77**:4243–55.
111. Thomson SJ, Goh FG, Banks H, *et al.* The role of transposable elements in the regulation of IFN-lambda1 gene expression. *Proc Natl Acad Sci USA* 2009;**106**:11564–9.
112. Panning B, Smiley JR. Activation of RNA polymerase III transcription of human Alu elements by herpes simplex virus. *Virology* 1994;**202**:408–17.
113. Tang RB, Wang HY, Lu HY, *et al.* Increased level of polymerase III transcribed Alu RNA in hepatocellular carcinoma tissue. *Mol Carcinog* 2005;**42**:93–6.
114. Hasler J, Samuelsson T, Strub K. Useful 'junk': Alu RNAs in the human transcriptome. *Cell Mol Life Sci* 2007;**64**:1793–800.
115. Mariner PD, Walters RD, Espinoza CA, *et al.* Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Mol Cell* 2008;**29**:499–509.
116. Rana T, Misra S, Mittal MK, *et al.* Mechanism of down-regulation of RNA polymerase III-transcribed ncRNA genes in macrophages by Leishmania. *J Biol Chem* 2011;**286**:6614–26.
117. Berger A, Strub K. Multiple roles of Alu-related noncoding RNAs. *Prog Mol Subcell Biol* 2011;**51**:119–46.
118. Rubin CM, Kimura RH, Schmid CW. Selective stimulation of translational expression by Alu RNA. *Nucleic Acids Res* 2002;**30**:3253–61.
119. van BH, Nislow C, Blencowe BJ, *et al.* Most "dark matter" transcripts are associated with known genes. *PLoS Biol* 2010;**8**:e1000371.
120. Ponting CP, Belgard TG. Transcribed dark matter: meaning or myth? *Hum Mol Genet* 2010;**19**:R162–8.
121. Makalowski W, Mitchell GA, Labuda D. Alu sequences in the coding regions of mRNA: a source of protein variability. *Trends Genet* 1994;**10**:188–93.
122. Sorek R, Ast G, Graur D. Alu-containing exons are alternatively spliced. *Genome Res* 2002;**12**:1060–7.
123. Kneeling J, Graveley BR. The origins and implications of Alu alternative splicing. *Trends Genet* 2004;**20**:1–4.
124. Shen S, Lin L, Cai JJ, *et al.* Widespread establishment and regulatory impact of Alu exons in human genes. *Proc Natl Acad Sci USA* 2011;**108**:2837–42.
125. Kim DD, Kim TT, Walsh T, *et al.* Widespread RNA editing of embedded alu elements in the human transcriptome. *Genome Res* 2004;**14**:1719–25.
126. Levanon EY, Eisenberg E, Yelin R, *et al.* Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol* 2004;**22**:1001–5.
127. Eisenberg E, Nemzer S, Kinar Y, *et al.* Is abundant A-to-I RNA editing primate-specific? *Trends Genet* 2005;**21**:77–81.

128. Jepson JE, Reenan RA. RNA editing in regulating gene expression in the brain. *Biochim Biophys Acta* 2008; **1779**:459–70.
129. Paz-Yaacov N, Levanon EY, Nevo E, *et al.* Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. *Proc Natl Acad Sci USA* 2010; **107**:12174–9.
130. Osenberg S, Paz YN, Safran M, *et al.* Alu sequences in undifferentiated human embryonic stem cells display high levels of A-to-I RNA editing. *PLoS One* 2010; **5**: e11173.
131. Sharma CM, Hoffmann S, Darfeuille F, *et al.* The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 2010; **464**:250–5.
132. Yelin R, Dahary D, Sorek R, *et al.* Widespread occurrence of antisense transcription in the human genome. *Nat Biotechnol* 2003; **21**:379–86.
133. Chen J, Sun M, Kent WJ, *et al.* Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res* 2004; **32**:4812–20.
134. Li JT, Zhang Y, Kong L, *et al.* Trans-natural antisense transcripts including noncoding RNAs in 10 species: implications for expression regulation. *Nucleic Acids Res* 2008; **36**: 4833–44.
135. He Y, Vogelstein B, Velculescu VE, *et al.* The antisense transcriptomes of human cells. *Science* 2008; **322**:1855–7.
136. Impey S, McCorkle SR, Cha-Molstad H, *et al.* Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. *Cell* 2004; **119**:1041–54.
137. Lercher MJ, Blumenthal T, Hurst LD. Coexpression of neighboring genes in *Caenorhabditis elegans* is mostly due to operons and duplicate genes. *Genome Res* 2003; **13**: 238–43.
138. Wang P, Yin S, Zhang Z, *et al.* Evidence for common short natural trans sense-antisense pairing between transcripts from protein coding genes. *Genome Biol* 2008; **9**:R169.
139. Lapidot M, Pilpel Y. Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO Rep* 2006; **7**:1216–22.
140. Muro EM, ndrade-Navarro MA. Pseudogenes as an alternative source of natural antisense transcripts. *BMC Evol Biol* 2010; **10**:338.
141. Conley AB, Miller WJ, Jordan IK. Human cis natural antisense transcripts initiated by transposable elements. *Trends Genet* 2008; **24**:53–56.
142. Kim DS, Hahn Y. Human-specific antisense transcripts induced by the insertion of transposable element. *Int J Mol Med* 2010; **26**:151–7.
143. Prescott EM, Proudfoot NJ. Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci USA* 2002; **99**:8796–801.
144. Hastings ML, Milcarek C, Martincic K, *et al.* Expression of the thyroid hormone receptor gene, *erbAalpha*, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. *Nucleic Acids Res* 1997; **25**:4296–300.
145. Mahmoudi S, Henriksson S, Corcoran M, *et al.* Wrap53, a natural p53 antisense transcript required for p53 induction upon DNA damage. *Mol Cell* 2009; **33**:462–71.
146. Kiyosawa H, Abe K. Speculations on the role of natural antisense transcripts in mammalian X chromosome evolution. *Cytogenet Genome Res* 2002; **99**:151–6.
147. Marsit CJ, Eddy K, Kelsey KT. MicroRNA responses to cellular stress. *Cancer Res* 2006; **66**:10843–8.
148. Bilen J, Liu N, Bonini NM. A new role for microRNA pathways: modulation of degeneration induced by pathogenic human disease proteins. *Cell Cycle* 2006; **5**: 2835–8.
149. Nelson PT, Wang WX, Rajeev BW. MicroRNAs (miRNAs) in neurodegenerative diseases. *Brain Pathol* 2008; **18**:130–8.
150. Reinhart BJ, Slack FJ, Basson M, *et al.* The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 2000; **403**:901–6.
151. Harfe BD, McManus MT, Mansfield JH, *et al.* The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. *Proc Natl Acad Sci USA* 2005; **102**:10898–903.
152. Nelson PT, Keller JN. RNA in brain disease: no longer just "the messenger in the middle". *J Neuropathol Exp Neurol* 2007; **66**:461–8.
153. Bothwell M, Giniger E. Alzheimer's disease: neurodevelopment converges with neurodegeneration. *Cell* 2000; **102**: 271–3.
154. Makeyev EV, Zhang J, Carrasco MA, *et al.* The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell* 2007; **27**:435–48.
155. Garzon R, Liu S, Fabbri M, *et al.* MicroRNA-29b induces global DNA hypomethylation and tumor suppressor gene reexpression in acute myeloid leukemia by targeting directly DNMT3A and 3B and indirectly DNMT1. *Blood* 2009; **113**: 6411–8.
156. Cheng HY, Papp JW, Varlamova O, *et al.* microRNA modulation of circadian-clock period and entrainment. *Neuron* 2007; **54**:813–29.
157. Yang WJ, Yang DD, Na S, *et al.* Dicer is required for embryonic angiogenesis during mouse development. *J Biol Chem* 2005; **280**:9330–5.
158. Vasudevan S, Tong Y, Steitz JA. Cell-cycle control of microRNA-mediated translation regulation. *Cell Cycle* 2008; **7**:1545–9.
159. Linsley PS, Schelter J, Burchard J, *et al.* Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression. *Mol Cell Biol* 2007; **27**:2240–52.
160. Cuellar TL, McManus MT. MicroRNAs and endocrine biology. *J Endocrinol* 2005; **187**:327–32.
161. Rodriguez A, Vigorito E, Clare S, *et al.* Requirement of bic/microRNA-155 for normal immune function. *Science* 2007; **316**:608–11.
162. Kawahara Y, Megraw M, Kreider E, *et al.* Frequency and fate of microRNA editing in human brain. *Nucleic Acids Res* 2008; **36**:5270–80.
163. Blow MJ, Grocock RJ, van DS, *et al.* RNA editing of human microRNAs. *Genome Biol* 2006; **7**:R27.
164. Kawahara Y, Zinshteyn B, Sethupathy P, *et al.* Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* 2007; **315**:1137–40.
165. Liang H, Landweber LF. Hypothesis. RNA editing of microRNA target sites in humans? *RNA* 2007; **13**:463–7.
166. Guo H, Ingolia NT, Weissman JS, *et al.* Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 2010; **466**:835–40.

167. Selbach M, Schwanhauser B, Thierfelder N, *et al.* Widespread changes in protein synthesis induced by microRNAs. *Nature* 2008;**455**:58–63.
168. Baek D, Villen J, Shin C, *et al.* The impact of microRNAs on protein output. *Nature* 2008;**455**:64–71.
169. Krek A, Grun D, Poy MN, *et al.* Combinatorial microRNA target predictions. *Nat Genet* 2005;**37**:495–500.
170. Daskalova E, Baev V, Rusinov V, *et al.* 3'UTR-located ALU elements: donors of potential miRNA target sites and mediators of network miRNA-based regulatory interactions. *Evol Bioinform Online* 2006;**2**:103–20.
171. Smalheiser NR, Torvik VI. Mammalian microRNAs derived from genomic repeats. *Trends Genet* 2005;**21**:322–6.
172. Lehnert S, Van LP, Thilakarathne PJ, *et al.* Evidence for co-evolution between human microRNAs and Alu-repeats. *PLoS One* 2009;**4**:e4456.
173. Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* 2006;**13**:1097–101.
174. Bortolin-Cavaille ML, Dance M, Weber M, *et al.* C19MC microRNAs are processed from introns of large Pol-II, non-protein-coding transcripts. *Nucleic Acids Res* 2009;**37**:3464–73.
175. Martin KC, Ephrussi A. mRNA localization: gene expression in the spatial dimension. *Cell* 2009;**136**:719–30.
176. Hurt JA, Silver PA. mRNA nuclear export and human disease. *Dis Model Mech* 2008;**1**:103–8.
177. Barthelson RA, Lambert GM, Vanier C, *et al.* Comparison of the contributions of the nuclear and cytoplasmic compartments to global gene expression in human cells. *BMC Genomics* 2007;**8**:340.
178. Trask HW, Cowper-Sal-lari R, Sartor MA, *et al.* Microarray analysis of cytoplasmic versus whole cell RNA reveals a considerable number of missed and false positive mRNAs. *RNA* 2009;**15**:1917–28.
179. Chen L. A global comparison between nuclear and cytosolic transcriptomes reveals differential compartmentalization of alternative transcript isoforms. *Nucleic Acids Res* 2010;**38**:1086–97.
180. Liao JY, Ma LM, Guo YH, *et al.* Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3' trailers. *PLoS One* 2010;**5**:e10563.
181. Barak M, Levanon EY, Eisenberg E, *et al.* Evidence for large diversity in the human transcriptome created by Alu RNA editing. *Nucleic Acids Res* 2009;**37**:6905–15.
182. Clemson CM, Hutchinson JN, Sara SA, *et al.* An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* 2009;**33**:717–26.
183. Mao YS, Sunwoo H, Zhang B, *et al.* Direct visualization of the co-transcriptional assembly of a nuclear body by non-coding RNAs. *Nat Cell Biol* 2011;**13**:95–101.
184. Levanon K, Eisenberg E, Rechavi G, *et al.* Letter from the editor: Adenosine-to-inosine RNA editing in Alu repeats in the human genome. *EMBO Rep* 2005;**6**:831–5.
185. Chen LL, Carmichael GG. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell* 2009;**35**:467–78.
186. Shamovsky I, Ivannikov M, Kandel ES, *et al.* RNA-mediated response to heat shock in mammalian cells. *Nature* 2006;**440**:556–60.
187. Barski A, Chepelev I, Liko D, *et al.* Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. *Nat Struct Mol Biol* 2010;**17**:629–34.
188. Raha D, Wang Z, Moqtaderi Z, *et al.* Close association of RNA polymerase II and many transcription factors with Pol III genes. *Proc Natl Acad Sci USA* 2010;**107**:3639–44.
189. Moqtaderi Z, Wang J, Raha D, *et al.* Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat Struct Mol Biol* 2010;**17**:635–40.
190. Yoon OK, Brem RB. Noncanonical transcript forms in yeast and their regulation during environmental stress. *RNA* 2010;**16**:1256–67.
191. Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods* 2007;**4**:721–6.
192. Pothof J, Verkaik NS, Hoeijmakers JH, *et al.* MicroRNA responses and stress granule formation modulate the DNA damage response. *Cell Cycle* 2009;**8**:3462–8.
193. Shomron N, Alberstein M, Reznik M, *et al.* Stress alters the subcellular distribution of hSlu7 and thus modulates alternative splicing. *J Cell Sci* 2005;**118**:1151–9.
194. Chua K, Reed R. The RNA splicing factor hSlu7 is required for correct 3' splice-site choice. *Nature* 1999;**402**:207–10.
195. Gallouzi IE, Brennan CM, Stenberg MG, *et al.* HuR binding to cytoplasmic mRNA is perturbed by heat shock. *Proc Natl Acad Sci USA* 2000;**97**:3073–8.
196. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* 2011;**470**:284–8.