

# Short Interspersed Nuclear Elements (SINEs)

Subjects: Genetics & Heredity

Contributor: HandWiki Huang

SINEs or Short Interspersed Nuclear Elements are sequences of non-coding DNA present at high frequencies in various eukaryotic genomes. They are a class of retrotransposons, DNA elements that amplify themselves throughout eukaryotic genomes, often through RNA intermediates. Short-interspersed nuclear elements are characterized by their size and method of retrotransposition. The literature differs on the length of the SINEs but there is a general consensus that they often range in length from about 100 to 700 base pairs (more or less, arbitrary cut-offs). Short-interspersed nuclear elements are transcribed by RNA polymerase III which is known to transcribe ribosomal RNA and tRNA, two types of RNA vital to ribosomal assembly and mRNA translation. SINEs, like tRNAs and many small-nuclear RNAs possess an internal promoter and thus are transcribed differently than most protein-coding genes. In other words, short-interspersed nuclear elements have their key promoter elements within the transcribed region itself. Though transcribed by RNA polymerase III, SINEs and other genes possessing internal promoters, recruit different transcriptional machinery and factors than genes possessing upstream promoters. The RNA coded by the short-interspersed nuclear element does not code for any protein product but is nonetheless reverse-transcribed and inserted back into an alternate region in the genome. For this reason, short interspersed nuclear elements are believed to have co-evolved with long interspersed nuclear element (LINES), as LINES do in fact encode protein products which enable them to be reverse-transcribed and integrated back into the genome. SINEs are believed to have co-opted the proteins coded by LINES which are contained in 2 reading frames. Open reading frame 1 (ORF 1) encodes a protein which binds to RNA and acts as a chaperone to facilitate and maintain the LINE protein-RNA complex structure. Open reading frame 2 (ORF 2) codes a protein which possesses both endonuclease and reverse transcriptase activities. This enables the LINE mRNA to be reverse-transcribed into DNA and integrated into the genome based on the sequence-motifs recognized by the protein's endonuclease domain. Furthermore, SINEs are known to share sequence homology with LINES which gives a basis by which the LINE machinery can reverse transcribe and integrate SINE transcripts. Alternately, some SINEs are believed to use a much more complex system of integrating back into the genome; this system involves the use random double-stranded DNA breaks (rather than the endonuclease coded by related long-interspersed nuclear elements creating an insertion-site). These DNA breaks are utilized to prime reverse transcriptase, ultimately integrating the SINE transcript back into the genome. SINEs nonetheless depend on enzymes coded by other DNA elements and are thus known as non-autonomous retrotransposons as they depend on the machinery of LINES, which are known as autonomous retrotransposons.

Keywords: sequence homology ; endonuclease ; transcriptional

---

## 1. Properties

Short-interspersed nuclear elements are believed to have parasitic origins in eukaryotic genomes. These SINEs have mutated and replicated themselves an incredible amount of times on an evolutionary time-scale and thus form many different lineages. Their early evolutionary origin has caused them to be ubiquitous in many eukaryotic lineages. In fact, the Alu element, a short-interspersed nuclear element of about 300 nucleotides is located at over a million copies throughout the human genome, which is over 10 percent of the total genome; this is not uncommon among other species.<sup>[1]</sup> The Alu family of sines are among an incredible amount of eukaryotic SINEs distinguished by different features including their length, sequence, and method of retrotransposition.

Furthermore, SINEs are characterized by their different modules, which are essentially a sectioning of their sequence. SINEs can, but do not necessarily have to possess a head, a body, and a tail. The head, is at the 5' end of short-interspersed nuclear elements and is an evolutionarily derived from an RNA synthesized by RNA Polymerase III such as ribosomal RNAs and tRNAs; the 5' head is indicative of which endogenous element that SINE was derived from and was able to parasitically utilize its transcriptional machinery.<sup>[2]</sup> For example, the 5' of the Alu sine is derived from 7SL RNA, a sequence transcribed by RNA Polymerase III which codes for the RNA element of SRP, an abundant ribonucleoprotein.<sup>[3]</sup> The body of SINEs possess an unknown origin but often share much homology with a corresponding LINE which thus allows SINEs to parasitically co-opt endonucleases coded by LINES (which recognize certain sequence motifs). Lastly, the

3' tail of SINEs is composed of short simple repeats of varying lengths; these simple repeats are sites where two (or more) short-interspersed nuclear elements can combine to form a dimeric SINE.<sup>[4]</sup> Short-interspersed nuclear elements which do not only possess a head and tail are called simple SINEs whereas short-interspersed nuclear elements which also possess a body or are a combination of two or more SINEs are complex SINEs.<sup>[2]</sup>

In essence, short interspersed nuclear elements are genetic parasites which have evolved very early in the history of eukaryotes to utilize protein machinery within the organism as well as to co-opt the machinery from similarly parasitic genomic elements. The simplicity of these elements make them incredibly successful at persisting and amplifying (through retrotransposition) within the genomes of eukaryotes. These "parasites" which have become ubiquitous in genomes can be very deleterious to organisms as discussed below. However, eukaryotes have been able to integrate short-interspersed nuclear elements into different signaling, metabolic and regulatory pathways and have become a great source of genetic variability. They seem to play a particularly important role in the regulation of gene expression and the creation of RNA genes as discussed in Sines and Gene-Regulation. This regulation extends to chromatin re-organization and the regulation of genomic architecture; furthermore, the different lineages, mutations, and activity among eukaryotes make short-interspersed nuclear elements an incredible useful tool in phylogenetic analysis.

## **2. SINEs & Pseudogenes**

The activity of SINEs however has genetic vestiges which do not seem to play a significant role, positive or negative, and manifest themselves in the genome as pseudogenes. SINEs however should not be mistaken as RNA pseudogenes.<sup>[2]</sup> In general, pseudogenes are generated when processed mRNAs of protein-coding genes are reverse-transcribed and incorporated back into the genome (RNA pseudogenes are reverse transcribed RNA genes).<sup>[5]</sup> Pseudogenes are generally functionless as they descend from processed RNAs independent of their evolutionary-context which includes introns and different regulatory elements which enable transcription and processing. These pseudogenes, though non-functional may in some cases still possess promoters, CpG islands, and other features which enable transcription; they thus can still be transcribed and may possess a role in the regulation of gene expression (like SINEs and other non-coding elements).<sup>[6]</sup> Pseudogenes thus differ from SINEs in that they are derived from transcribed- functional RNA whereas SINEs are DNA elements which retrotranspose by co-opting RNA genes transcriptional machinery. However, there are studies which suggest that retro-transposable elements such as short-interspersed nuclear elements are not only capable of copying themselves in alternate regions in the genome but are also able to do so for random genes too.<sup>[6]</sup> Thus SINEs can be playing a vital role in the generation of pseudogenes, which themselves are known to be involved in regulatory networks. This is perhaps another means by which SINEs have been able to influence and contribute to gene-regulation.

## **3. Mechanism of Retrotransposition**

The theory that short-interspersed nuclear elements have evolved to utilize the retrotransposon machinery of long-interspersed nuclear elements is supported by studies which examine the presence and distribution of LINEs and SINEs in taxa of different species.<sup>[8]</sup> For example, LINEs and SINEs in rodents and primates show very strong homology at the insertion-site motif.<sup>[8]</sup> Such evidence is a basis for the proposed mechanism in which integration of the SINE transcript can be co-opted with LINE-coded protein products. This is specifically demonstrated by a detailed analysis of over 20 rodent species profiled LINEs and SINEs, mainly L1's and B1's respectively; these are families of LINEs and SINEs found at high frequencies in rodents along with other mammals.<sup>[8]</sup> The study sought to provide phylogenetic clarity within the context of LINE and SINE activity.

The study arrived at a candidate taxa believed to be the first instance of L1 LINE extinction; it expectedly discovered that there was no evidence to suggest that B1 SINE activity occurred in species which did not have L1 LINE activity.<sup>[8]</sup> Also, the study suggested that B1 short-interspersed nuclear element silencing in fact occurred before L1 long-interspersed nuclear element extinction; this is due to the fact that B1 SINEs are silenced in the genus most-closely related to the genus which does not contain active L1 LINEs (though the genus with B1 SINE silencing still contains active L1 LINEs).<sup>[8]</sup> Another genus was also found which similarly contained active L1 long-interspersed nuclear elements but did not contain B1 short-interspersed nuclear elements; the opposite scenario, in which active B1 SINEs were present in a genus which did not possess active L1 LINEs was not found.<sup>[8]</sup> This result was expected and strongly supports the theory that SINEs have evolved to co-opt the RNA-binding proteins, endonucleases, and reverse-transcriptases coded by LINEs. In taxa which do not actively transcribe and translate long-interspersed nuclear elements protein-products, SINEs do not have the theoretical foundation by which to retrotranspose within the genome. The results obtained in Rinehart et al. are thus very supportive of the current model of SINE retrotransposition.

## 4. SINEs and Gene-Regulation

### 4.1. SINEs as Long Non-Coding RNA

Changes in chromosome structure influence gene expression primarily by affecting the accessibility of genes to transcriptional machinery. The chromosome has a very complex and hierarchical system of organizing the genome. This system of organization, which includes histones, methyl groups, acetyl groups, and a variety of proteins and RNAs allows different domains within a chromosome to be accessible to polymerases, transcription factors, and other associated proteins to different degrees.<sup>[9]</sup> Furthermore, the shape and density of certain areas of a chromosome can affect the shape and density of neighboring (or even distant regions) on the chromosome through interaction facilitated by different proteins and elements. Non-coding RNAs such as short-interspersed nuclear elements, which have been known to associate with and contribute to chromatin structure, can thus play huge role in regulating gene expression.<sup>[10]</sup> For example, long non-coding RNAs have been known to help initiate expression of Ubx by directing the Ash1 protein to regulatory elements within the Hox gene set; Ash1 modifies chromatin structure in such a way as to increase the expression of Ubx.<sup>[11]</sup> Short-interspersed-nuclear-elements similarly can be involved in gene regulation by modifying genomic architecture .

In fact Usmanova et al. 2008 suggested that short-interspersed nuclear elements can serve as direct signals in chromatin rearrangement and structure. The paper examined the global distribution of SINEs in mouse and human chromosomes and determined that this distribution was very similar to genomic distributions of genes and CpG motifs.<sup>[12]</sup> The distribution of SINEs to genes was significantly more similar than that of other non-coding genetic elements and even differed significantly from the distribution of long-interspersed nuclear elements.<sup>[12]</sup> This suggested that the SINE distribution was not a mere accident caused by LINE-mediate retrotransposition but rather that SINEs possessed a role in gene-regulation. Furthermore, SINEs frequently contain motifs for YY1 polycomb proteins.<sup>[12]</sup> YY1 is a zinc-finger protein that acts as a transcriptional repressor for a wide-variety of genes essential for development and signaling.<sup>[13]</sup> Polycomb protein YY1 is believed to mediate the activity of histone deacetylases and histone acetyltransferases to facilitate chromatin re-organization; this is often to facilitate the formation of heterochromatin (gene-silencing state).<sup>[14]</sup> Thus, the analysis suggests that short-interspersed nuclear elements can function as a 'signal-booster' in the polycomb-dependent silencing of gene-sets through chromatin re-organization.<sup>[12]</sup> In essence, it is the cumulative effect of many types of interactions that leads to the difference between euchromatin, which is not tightly packed and generally more accessible to transcriptional machinery, and heterochromatin, which is tightly packed and generally not accessible to transcriptional machinery; SINEs seem to play an evolutionary role in this process.

In addition to directly affecting chromatin structure, there are a number of ways in which SINEs can potentially regulate gene expression. For example, long non-coding RNA can directly interact with transcriptional repressors and activators, attenuating or modifying their function.<sup>[15]</sup> This type of regulation can occur in different ways: the RNA transcript can directly bind to the transcription factor as a co-regulator; also, the RNA can regulate and modify the ability of co-regulators to associate with the transcription factor.<sup>[15]</sup> For example, Evf-2, a certain long non-coding RNA, has been known to function as a co-activator for certain homeobox transcription factors which are critical to nervous system development and organization.<sup>[16]</sup> Furthermore, RNA transcripts can interfere with the functionality of the transcriptional complex by interacting or associating with RNA polymerases during the transcription or loading processes.<sup>[15]</sup> Moreover, non-coding RNAs like SINEs can bind or interact directly with the DNA duplex coding the gene and thus prevent its transcription.<sup>[15]</sup>

Also, many non-coding RNAs are distributed near protein-coding genes, often in the reverse direction. This is especially true for short-interspersed nuclear elements as seen in Usmanova et al. These non-coding RNAs, which lie adjacent to or overlap gene-sets provide a mechanism by which transcription factors and machinery can be recruited to increase or repress the transcription of local genes. The particular example of SINEs potentially recruit the YY1 polycomb transcriptional repressor is discussed above.<sup>[12]</sup> Alternatively, it also provides a mechanism by which local gene expression can be curtailed and regulated because the transcriptional complexes can hinder or prevent nearby genes from being transcribed. There is research to suggest that this phenomenon is particularly seen in the gene-regulation of pluripotent cells.<sup>[17]</sup>

In conclusion, non-coding RNAs such as SINEs are capable of effecting gene expression on a multitude of different levels and in different ways. Short-interspersed nuclear elements are believed to be deeply integrated into a complex regulatory network capable of fine-tuning gene expression across the eukaryotic genome.

## 4.2. SINEs, microRNAs, & Disease

The role of short-interspersed nuclear elements in gene regulation within cells has been supported by multiple studies. One such study examined the correlation between a certain family of SINEs with microRNAs (in zebrafish).<sup>[18]</sup> The specific family of SINEs being examined was the Anamnia V-SINEs; this family of short interspersed nuclear elements is often found in the untranslated region of the 3' end of many genes and is present in vertebrate genomes.<sup>[18]</sup> The study involved a computational analysis in which the genomic distribution and activity of the Anamnia V-SINEs in *Danio rerio* zebrafish was examined; furthermore, these V-SINEs potential to generate novel microRNA loci was analyzed.<sup>[18]</sup> It was found that genes which were predicted to possess V-SINEs were targeted by microRNAs with significantly higher hybridization E-values (relative to other areas in the genome).<sup>[18]</sup> The genes that had high hybridization E-values were genes particularly involved in metabolic and signaling pathways.<sup>[18]</sup> Almost all miRNAs identified to have a strong ability to hybridize to putative V-SINE sequence motifs in genes have been identified (in mammals) to have regulatory roles.<sup>[18]</sup> These results which establish a correlation between short-interspersed nuclear elements and different regulatory microRNAs strongly suggest that V-SINEs have a significant role in attenuating responses to different signals and stimuli related to metabolism, proliferation and differentiation. Many other studies must be undertaken to establish the validity and extent of short-interspersed nuclear element retrotransposons' role in regulatory gene-expression networks. In conclusion, though not much is known about the role and mechanism by which SINEs generate miRNA gene loci it is generally understood that SINEs have played a significant evolutionary role in the creation of "RNA-genes", this is also touched upon above in SINEs & Pseudogenes.

With such evidence suggesting that short-interspersed nuclear elements have been evolutionary sources for microRNA loci generation it is important to further discuss the potential relationships between the two as well as the mechanism by which the microRNA regulates RNA degradation and more broadly, gene expression. A microRNA is a non-coding RNA generally 22 nucleotides in length.<sup>[19]</sup> This non-protein coding oligonucleotide is itself coded by longer nuclear DNA sequence usually transcribed by RNA polymerase II which is also responsible for the transcription of most mRNAs and snRNAs in eukaryotes.<sup>[20]</sup> However, some research suggests that some microRNAs that possess upstream short-interspersed nuclear elements are transcribed by RNA polymerase III which is widely implicated in ribosomal RNA and tRNA, two transcripts vital to mRNA translation.<sup>[21]</sup> This provides an alternate mechanism by which short-interspersed nuclear elements could be interacting with or mediating gene-regulatory networks involving microRNAs.

The regions coding miRNA can be independent RNA-genes often being anti-sense to neighboring protein-coding genes, or can be found within the introns of protein-coding genes.<sup>[22]</sup> The co-localization of microRNA and protein-coding genes provides a mechanistic foundation by which microRNA regulates gene-expression. Furthermore, Scarpato et al. reveals (as discussed above) that genes predicted to possess short-interspersed nuclear elements (SINEs) through sequence analysis were targeted and hybridized by microRNAs significantly greater than other genes.<sup>[18]</sup> This provides an evolutionarily path by which the parasitic SINEs were co-opted and utilized to form RNA-genes (such as microRNAs) which have evolved to play a role in complex gene-regulatory networks.

The microRNAs are transcribed as part of longer RNA strands of generally about 80 nucleotides which through complementary base-pairing are able to form hairpin loop structures.<sup>[23]</sup> These structures are recognized and processed in the nucleus by the nuclear protein DiGeorge Syndrome Critical Region 8 (DGCR8) which recruits and associates with the Drosha protein.<sup>[24]</sup> This complex is responsible for cleaving some of the hair-pin structures from the pre-microRNA which is transported to the cytoplasm. The pre-miRNA is processed by the protein DICER into a double stranded 22 nucleotide.<sup>[25]</sup> Thereafter, one of the strands is incorporated into a multi-protein RNA-inducing silencing complex (RISC).<sup>[26]</sup> Among these proteins are proteins from the Argonaute family which are critical to the complex's ability to interact with and repress the translation of the target mRNA.<sup>[27]</sup>

Understanding the different ways in which microRNA regulates gene-expression, including mRNA-translation and degradation is key to understanding the potential evolutionary role of SINEs in gene-regulation and in the generation of microRNA loci. This, in addition to SINEs' direct role in regulatory networks (as discussed in SINEs as long non-coding RNAs) is crucial to beginning to understand the relationship between SINEs and certain diseases. Multiple studies have suggested that increased SINE activity is correlated with certain gene-expression profiles and post-transcription regulation of certain genes.<sup>[28][29][30]</sup> In fact, Peterson et al. 2013 demonstrated that high SINE RNA expression correlates with post-transcriptional downregulation of BRCA1, a tumor suppressor implicated in multiple forms of cancer, namely breast cancer.<sup>[30]</sup> Furthermore, studies have established a strong correlation between transcriptional mobilization of SINEs and certain cancers and conditions such as hypoxia; this can be due to the genomic instability caused by SINE activity as well as more direct-downstream effects.<sup>[29]</sup> SINEs have also been implicated in countless other diseases. In essence, short-

interspersed nuclear elements have become deeply integrated in countless regulatory, metabolic and signaling pathways and thus play an inevitable role in causing disease. Much is still to be known about these genomic parasites but it is clear they play a significant role within eukaryotic organisms.

---

## References

1. Batzer M. and Cordaux R. (October 2009). The impact of retrotransposons on human gene evolution. *Nature Reviews Genetics* 10(10), 691-703.
2. Kramerov D. and Vassetzky N. (2012). SINEBase: a database and tool for SINE analysis. *Nucleic Acids Research* 41, D83-D89.
3. Kriegs J et al. (April 2007). Evolutionary history of 7SL RNA-derived SINEs in Supraprimates. *Trends Genet* 23(4), 158-161.
4. Okada N. et al. (1997). SINEs and Lines share common 3' sequences: a review. *Gene* 205: 229-243.
5. Vanin E. (1985). Process pseudogenes: characteristics and evolution. *Annu. Rev. Genet.* 19, 253-272.
6. Dewannieux M. et al. (September 2003). LINE-mediate retrotransposition of marked Alu sequences. *Nat. Genet.* 35(1), 41-48.
7. Jurka J. (December 2004). Evolutionary impact of human Alu repetitive elements. *Current Opinion in Genetics & Development* 14(6), 603-608.
8. Rinehart T. et al. (2005). SINE extinction preceded LINE extinction in sigmodontine rodents: implications for retrotranspositional dynamics and mechanism. *Cytogenet Genome Res* 110(1-4), 416-425.
9. Kiefer JC (April 2007). "Epigenetics in development". *Developmental Dynamics* 236 (4): 1144–1156.
10. Rodríguez-Campos A, and Azorín F. (2007). RNA is an integral component of chromatin that contributes to its structural organization. *PLOS ONE* 2(11), e1182.
11. Sanchez-Elsner et al. (February 2006). Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrathorax. *Science* 311(5764), 1118–1123.
12. Usmanova N et al. (2008). SINEs in mammalian genomes can serve as additional signals in formation of facultative heterochromatin. *Tsitologiya* 50(3), 256-260.
13. Shi Y et al. (October 1991). Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and a relief of repression by adenovirus E1A. *Cell* 67(2), 377-388.
14. Yao Y et al. (September 2001). Regulation of transcription factor YY1 by acetylation and deacetylation. *Molecular and Cellular Biology* 21(17), 5979-5991.
15. Goodrich J. and Kugel J. (August 2006). Non-coding-RNA regulators of RNA Polymerase II transcription. *Nature Reviews Molecular Cell Biology* 7(8), 612-616.
16. Feng J et al. (June 2006). The Evi-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & Development* 20(11), 1470-1484.
17. Sai Luo et al. (2016). Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. *Cell Stem Cell*. On-line.
18. Scarpato M et al. (September 2015). Short interspersed DNA elements and miRNAs: a novel hidden gene regulation layer in zebrafish? *Chromosome Res.* 23(3), 533-544.
19. Ambros V, (September 2004). The functions of animal microRNAs. *Nature* 431(7006), 350–5.
20. Lee Y et al. (October 2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBOJ* 23(20), 4051-4060.
21. Faller M and Guo F. (November 2008). MicroRNA biogenesis: there's more than one way to skin a cat. *Biochim. Biophys. Acta* 1779(11), 663-667.
22. Lau N et al. (October 2001). An abundant class of tiny RNAs with probable regulatory roles in *C. elegans*. *Science* 294(5543), 858-862.
23. Cai X et al. (December 2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10(12), 1957-1966.
24. Lee Y et al. (September 2003). The nuclear RNase III Drosha Initiates microRNA processing. *Nature* 425(6956), 415-419.

25. Bartel, DP (Jan 23, 2004). "MicroRNAs: genomics, biogenesis, mechanism, and function". *Cell* 116 (2): 281–97.
26. Schwarz D et al. (May 2002). Why do miRNAs live in the miRNP? *Genes Dev.* 16(9), 1025-1031.
27. Pratt A et al. (July 2009). The RNA-induced silencing complex: a versatile gene-silencing machine. *J. Biol. Chem.* 284(27), 17897-17901.
28. Natt D et al. (2015). High cortisol in 5 year-old children cause loss of DNA methylation in SINE retrotransposons: a possible role for ZNF263 in stress-related diseases. *Clin Epigenetics* 7(1), 91.
29. Pal A et al. (2010). Aberrant methylation and associated transcriptional mobilization of Alu elements contributes to genomic instability in hypoxia. *J Cell Mol Med* 14(11), 2646-2654.
30. Peterson M et al. (2013). High SINE RNA expression correlates with post-transcriptional down-regulation of BRCA1.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/80334>