Send Orders for Reprints to reprints@benthamscience.net

Current Proteomics, 2024, *21*, 140-161

REVIEW ARTICLE

The Relationship of Transposable Elements with Non-Coding RNAs in the Emergence of Human Proteins and Peptides

Rustam Nailevich Mustafin[1,*]

[1]*Bashkir State Medical University, 3, Lenin str., Ufa, 450008, Russia*

ARTICLE HISTORY

Received: March 27, 2024
Revised: June 27, 2024
Accepted: July 12, 2024

DOI:
10.2174/0115701646319572240805103747

Abstract: Transposable elements are the oldest structural and functional units that were formed during the emergence of life on Earth. The most ancient properties of transposable elements are the multifunctionality of their transcription and translation products and the formation of their many variants through processing, due to which transposable elements are key evolutionary sources of long non-coding RNAs, circular RNAs, microRNAs, proteins and peptides formation. Moreover, the same type of transposon can simultaneously serve as the source of the origin of all these molecules, providing the adaptive properties of living organisms, especially complex eukaryotes, including humans. The ancient ability of transposable elements for mutual integration due to their protein products interacting with DNA and RNA molecules, as well as for mutual regulation due to the functionality of their RNA, is the basis for the origin of many proteins and non-coding RNAs characterized by the same properties. This can explain the emergence of transcription factors from transposable elements, that is, proteins capable of interacting with the structures of DNA molecules due to the presence of specific amino acid sequences derived from transposable elements. This article presents facts about the origin during the evolution of many protein and non-coding RNA genes from transposable elements. Specific proteins and peptides translated from long non-coding RNAs, pri-microRNAs and circular RNAs are described, which reflect the origin of non-coding RNAs from transposable elements in evolution. These proteins and peptides are promising tools for the treatment of viral infections and drug-resistant tumors, since, together with non-coding RNAs, they are involved in antiviral and antitumor responses.

Keywords: Circular RNAs, long non-coding RNAs, microRNAs, origin, peptides, proteins, retroelements, transposable elements.

1. INTRODUCTION

According to a scientific article published in 2022 on deciphering the complete sequence of the human genome, the haploid set contains 3.055 billion nucleotide pairs. A total of 63,494 genes have been identified, of which only 19,969 are protein-coding genes (PCGs), responsible for the formation of 84,277 different transcripts [1]. A significant proportion consists of non-coding RNA (ncRNA) genes, the number of which in the GENECODE database is increasing due to modern molecular genetic research methods. Thus, if in 2012, 9277 long non-coding RNA genes (lncRNAs) were annotated in the human genome in the GENECODE database [2], in 2024 the statistics of this database (https://www.gencodegenes.org/human/stats.html) indicate the presence of 20,424 lncRNA genes responsible for the formation of 59,719 transcripts. In addition, 7565 small non-coding RNA genes and 14719 pseudogenes were annotated. It is interesting to note that 53.94% of the human genome is occupied by repeats, of which the main share is made up of transposable elements (TEs) [1]. There are class I TEs - retroelements (REs) and class II - DNA transposons. REs are transposed by reverse transcription of their own RNAs with the insertion of the formed complementary DNA (cDNA) into a new genomic locus using a "copy and paste" mechanism. REs are classified into LTR, DIRS, PLE, LINE and SINE orders. DNA transposons are classified into subclass 1, which includes TIR and Crypton, and subclass 2, which contains the Helitron and Maverick orders [3].

LINEs occupy 0.63 billion bp in the human genome, SINE – 0.39 billion bp, LTR – 0.27 billion bp, and DNA transposons – 0.108 billion bp. In total, TEs occupy about 1.4 billion bp, which is 46.7% of all DNA sequences [1]. Analysis of DNA sequences using specific oligonucleotides complementary to TEs, carried out back in 2011, showed that TEs sequences occupy at least 2/3 of the entire human genome [4]. This discrepancy in data is due to the key role of TEs in the emergence and evolution of non-coding RNA and PCGs, which suggests the multifunctionality of such genes as a specific property of TEs [5], since during the origin of life on Earth and the emergence of the DNA-RNA-proteins world from the RNA-DNA world, TEs served as universal sources of origin for a countless variety of peptide and protein molecules due to the processing of their transcripts [6]. Accordingly, TE transcripts are processed by

* Address correspondence to this author at the Bashkir State Medical University, 3, Lenin str., Ufa, 450008, Russia; E-mail: ruji79@mail.ru

1875-6247/24 $65.00+.00

various enzymatic systems to produce different variants of RNA molecules that can bind to ribosomes and form a huge variety of proteins and peptides. An analysis of the scientific literature shows the growing amount of information confirming this assumption. This article presents data on the ability of ncRNAs to translate into functional peptide and protein molecules, as well as the evolutionary role of ncRNAs in the emergence of new PCGs.

2. TRANSPOSONS ARE SOURCES OF NON-CODING RNA GENES

NcRNAs include long ncRNAs (lncRNAs) longer than 200 nucleotides and small ncRNAs (shorter than 200 nucleotides, including microRNAs 20–24 nucleotides long). MicroRNAs negatively regulate their target mRNAs by binding to their 3' untranslated region (UTR) and inhibiting translation or impairing the stability of the molecules. MicroRNA biogenesis consists of the processing of larger pri-miRNA molecules into shorter pre-miRNAs, followed by their further processing into mature miRNAs. A single microRNA can regulate multiple targets, participating in a variety of cellular functions. For example, the tumor suppressor miR-34a inhibits the expression of 700 individual PCGs [7]. Mature lncRNAs can perform protein-independent functions, that is, they are ribozymes. Many lncRNAs are named after the PCGs in which they are located. If they are complementary to genes, they are called antisense lncRNAs or NAT - natural antisense transcript. There are cis-NATs (overlap with complementary genes) and trans-NATs (transcribed from pseudogenes) [8].

The expression of lncRNAs is similar to that of PCGs, since lncRNAs are transcribed by RNA polymerase II and subject to splicing and capping [9]. LncRNAs can function alone or as part of ribonucleoproteins (RNPs). They are involved in transcription regulation by interacting with DNA-binding proteins, and changing histone modification by binding to histone-modifying complexes and RNA polymerase [10]. Functioning as ribozymes, lncRNAs promote tRNA processing, ribosome and telomere formation, and intron splicing. LncRNAs play an important role in the regulation of DNA synthesis, cell division, apoptosis, transcription, processing, splicing and translation of mRNA. LncRNAs such as ANRASSF1, ANRIL, BORDERLINE, Kcnq1ot1, NeST, and PINT are involved in histone modifications due to the recruitment of chromatin remodeling factors. LncRNAs Airn, ecCEBP, H19, Kcnq1ot1, PAPAS, pRNA, PTENpa1-AS, TARID and Xist serve as guides for specific DNA methylation. LncRNAs interacting with transcription factors to regulate PCGs include AK141205, AK028326, ES1,ES2, ES3, linc-RoR, Evx1as, and Hoxb5/6as. In addition, lncRNA ALU-RNA edits mRNA molecules, and lncRNAs Uchl1-as1, lincMD1, lincRNA-p21, and ½-sbsRNA are used for RNA interference in the regulation of mRNA translation [9]. Some lncRNAs (lincRNA-Cox2, lincRNA-p21) form RNPs that regulate specific sets of gene transcription. For example, lincRNA-p21 together with hnRNP-K form complexes that bind to specific genome regions and suppress gene transcription along the p53 pathway [8].

LncRNAs can act as decoys for transcription factors by mimicking the sequence or structure of target DNA. These lncRNAs include gas5, PANDA, DHFR minor, and Lethe. Like mRNAs, long ncRNAs can also interact with microRNAs through hybridization due to partial sequence complementarity. Some lncRNAs that bind to miRNAs serve as molecular sponges that dampen the effects of miRNAs on their target mRNAs [8]. The possibility of interactions with transcription factors of lncRNAs may be due to their evolutionary origin from TEs, which were important evolutionary sources for the emergence of sequences of binding sites with transcription factors due to the specific sequences of their nucleotides [11]. Indeed, many mature lncRNAs are identical to TE sequences [12]. On average, 41% of lncRNA exons are proven to originate from TEs, and 83% of them contain at least one TE fragment. This provides the basis for the formation of functional domains of RNA molecules involved in a variety of biological reactions. In addition to participating in the structural configurations of mature transcripts, TEs form polyadenylation sites, promoters, splice donor and acceptor sites, that is, they are key evolutionary sources of lncRNA world [13]. Individual families of TEs are most frequently found in lncRNA genes. Interspecies differences in the composition of TEs in lncRNA genes are also characteristic, which is determined by the peculiarities of the distribution of TEs in these genomes [12].

One of the reasons for the domestication of TE genes by host genomes as sources of lncRNAs is the functionality of TE transcripts [5], which is due to their ancient universal properties as sources of life on Earth [6]. For example, Alu inserts into lncRNA genes and forms the structures required to interact with mRNA through short imperfect nucleotide pairings. LTR-REs also possess regulatory signals for ncRNAs. Thousands of lncRNA genes derived from LTR-REs have been identified in the human genome [14]. These lncRNAs control the pluripotency network by altering chromatin structure. They participate in blastocyst formation and further embryogenesis [15]. In addition, LTR-REs can directly serve as lncRNA genes [16]. LINE1 transcripts can also function as lncRNAs themselves, interacting with specific regions of chromatin and regulating gene expression in early embryogenesis. For example, by binding to Nucleolin and KAP1, LINE1s cause both the activation of ribosomal DNA genes and the suppression of many two-cell embryo genes through Dux silencing [17]. That is, in addition to transpositions using protein products, RE transcripts subjected to alternative processing exhibit the properties of ribozymes. This reflects their ancient conservative characteristic, which was key to the emergence of life on Earth. Accordingly, the relationship of TEs with lncRNAs and their role in the emergence of new PCGs may reflect similar processes at the early stages of the evolution of living things, when the continuity of ribozyme functions by protein molecules appeared.

The tissue specificity of lncRNAs exceeds that of proteins. Moreover, in the regulation of stem cell differentiation they interact with TEs [18]. LncRNAs, formed from intergenic regions of eukaryotic genomes, as well as from overlapping and antisense patterns relative to the adjacent PCGs

that they regulate, are characterized by tissue-specific transcription [19]. This allows lncRNAs to largely determine a variety of cellular phenotypes, especially in the brain [20]. Like proteins, the secondary structures of lncRNAs have a modular organization and form discrete domains consisting mainly of TE sequences [13]. The ability of TE transcripts to be processed to form mRNA by the spliceosome or functional RNA molecules by ribonucleases is the basis for the "search" for optimal ways of regulating gene networks in connection with the ability of lncRNA to be translated into peptides [21-23]. This can explain the emergence of PCGs in evolution from lncRNAs genes, which reflects the identical properties of TE genes in evolution [24-27]. LncRNAs are also capable of further processing to form microRNAs [28], the precursors of which, in turn, are also translated into peptides [29, 30]. In this case, lncRNA genes can simultaneously serve as microRNA genes that can be translated [31, 32]. In addition, TEs are also important sources of miRNA genes (Fig. **1**). Back in 2016, the MDTE DB database was published, which included 661 human microRNAs derived from TEs [33].

3. ORIGIN OF PROTEINS FROM TRANSPOSABLE ELEMENTS IN EVOLUTION

TEs in evolution were sources of new PCG formation through direct domestication of TEs genes, exonization of TEs sequences inserted into introns or non-coding regions, and the formation of pseudogenes using retroelement enzymes (Fig. **2**). The emergence of new PCGs by TEs plays an important role in evolution and speciation. Even the formation of pseudogenes contributes to changes in gene regulatory networks, since retrocopies are flanked by TEs sequences and, due to mutations, can acquire new functions and participation in new networks [34], controlled by complex epigenetic systems involving non-coding RNAs derived from TEs [5]. Spliceosomal introns also originated from TEs in evolution [35], with subsequent new insertions of TEs into these introns, which contributed to alternative splicing transcripts formation. If, during evolution, new protein isoforms formed contributed to better adaptation, these variants were fixed at the species level and exonization occurred. Various exons derived from TEs have been identified in the human genome. This occurs due to the evolutionary selection of more adaptive regulatory networks controlling alternative splicing of PCGs. As a result, organisms survive in a number of generations in which such a regulatory network promotes the formation of adaptive alternative splicing variants of PCGs, formed due to the exonization of TEs in the introns of these genes [36]. The highly abundant Alu retroelements in the introns of PCGs, constituting 11% of the human genome, represent a typical example of exonization of TEs. It has been shown that the use of the Alu sequences as protein-coding exons leads to the formation of new protein isoforms, promoting increased functional diversity. As a result, thousands of human PCGs express Alu-derived exons. Most exons are generated from the right arm, albeit left arm exonization is also frequent, and Alu elements are predominantly incorporated in the antisense to the gene. The Alu sequence contains multiple sites resembling consensus splice signals and other splicing regulatory elements. Alu exon activation occurs through mutations, usually at the 5' end of the Alu and in the surrounding introns, that create new and typically weak splice junctions and splicing regulatory elements. In most cases, Alu elements are incorporated as alternatively spliced exons, with the new Alu-containing transcript becoming a minor isoform. Using the eXAlu deep learning model, it was shown that the number of Alu elements in the human genome subject to exonization is about 110,000, which is 21 times more than represented in the GENOCODE database [37].

REs cause reverse transcription of mRNA to form cDNA, which they insert into the genome as a retrocopy (Fig. **3**). The terms "retrogenes" or "pseudogenes" are synonymous with retrocopies. They do not contain ancestral regulatory sequences, so neighboring sites are used to control them or new ones are formed due to mutations. In mammals, retrogenes generated by LINE reverse transcriptase have been shown to contain characteristic sequences in their flanking regions: a TTTT/AA endonuclease cleavage site, duplicated target sites, and a polyadenine tail priming reverse transcription. Pseudogenes in animals can also be formed by LTR-RE enzymes. For example, 15 polymorphic chimeric retrogenes flanked by LTR sequences have been identified in Drosophila [38]. Arabidopsis, rice and many other plant species also contain in their genomes many retrogenes, some of which are functional (tomato has the *Sun* gene, Arabidopsis has the *CYP98A8* and *CYP98A9* genes). Retrocopies in plants are flanked mainly by LTRs since LTR REs predominate in their genomes [39]. In rice, expression of more than 66% of retrocopies has been determined, although at a lower level than their parent genes, but with tissue-specific correlation of their levels [40].

Initially, pseudogenes do not contain introns, as they are formed by reverse processing of mature mRNAs. However, during evolution, they can acquire introns due to intragenic insertions of TEs, the preservation of which also indicates the inclusion of retrocopies in the functioning of regulatory networks. If the new pseudogene protein product promotes better adaptation and survival, the pseudogene may functionally replace the parent gene or develop a new function [35]. On average, about 10% of all eukaryotic PCGs contain one or more retrocopies. The number of pseudogenes is significantly higher than those determined by standard identification methods, which is associated with their 5' truncation [41]. The generated pseudogenes can be used in complex regulatory gene networks involving TEs and epigenetic factors since retrocopies serve as sources of transcription of non-coding RNAs, which was proven by analyzing the distribution of microRNA clusters involving Alu and L1 in the human genome [42]. Due to this, new opportunities for mutual regulation with other genes and loci of the genome are formed, which potentially serves as a source of new adaptive functions. To determine the origin of PCGs from TEs, various approaches are used, one of which is the identification of conserved domains specific to TEs. As a result, it is possible to identify many genes that arose in evolution from TE sequences and are used for the needs of hosts. Some of these PCGs form tandem clusters of gene families [43].

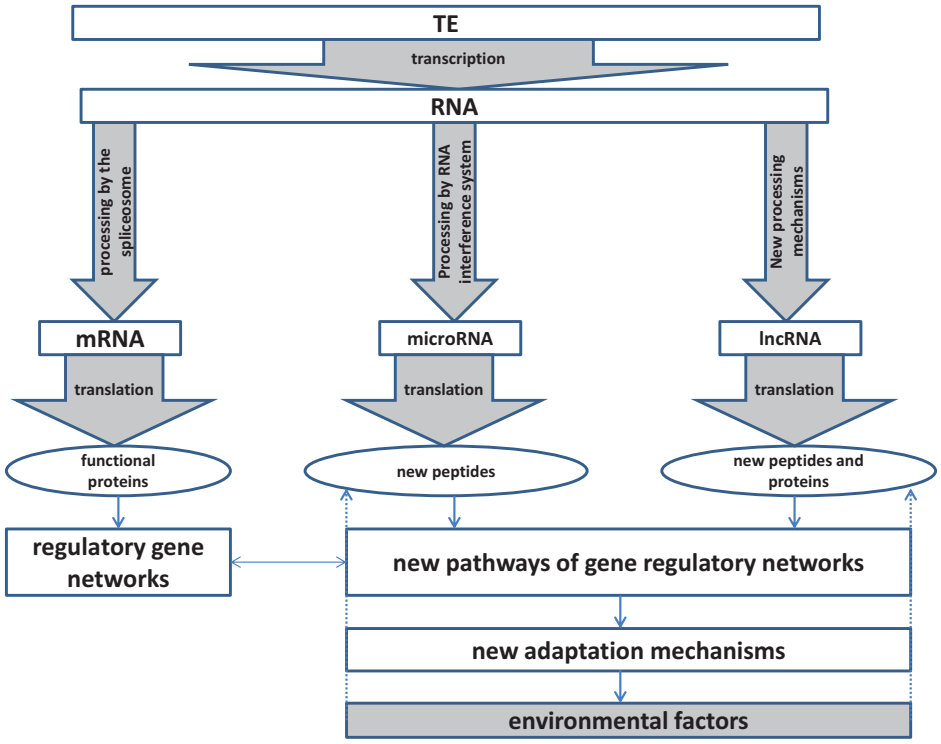

Fig. (1). Scheme of various proteins and peptides formation from transposable element transcripts due to their processing by various systems (splicing, RNA-induced silencing complex (RISC), alternative splicing), which allows transposable elements to participate in the regulation of new regulatory networks. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

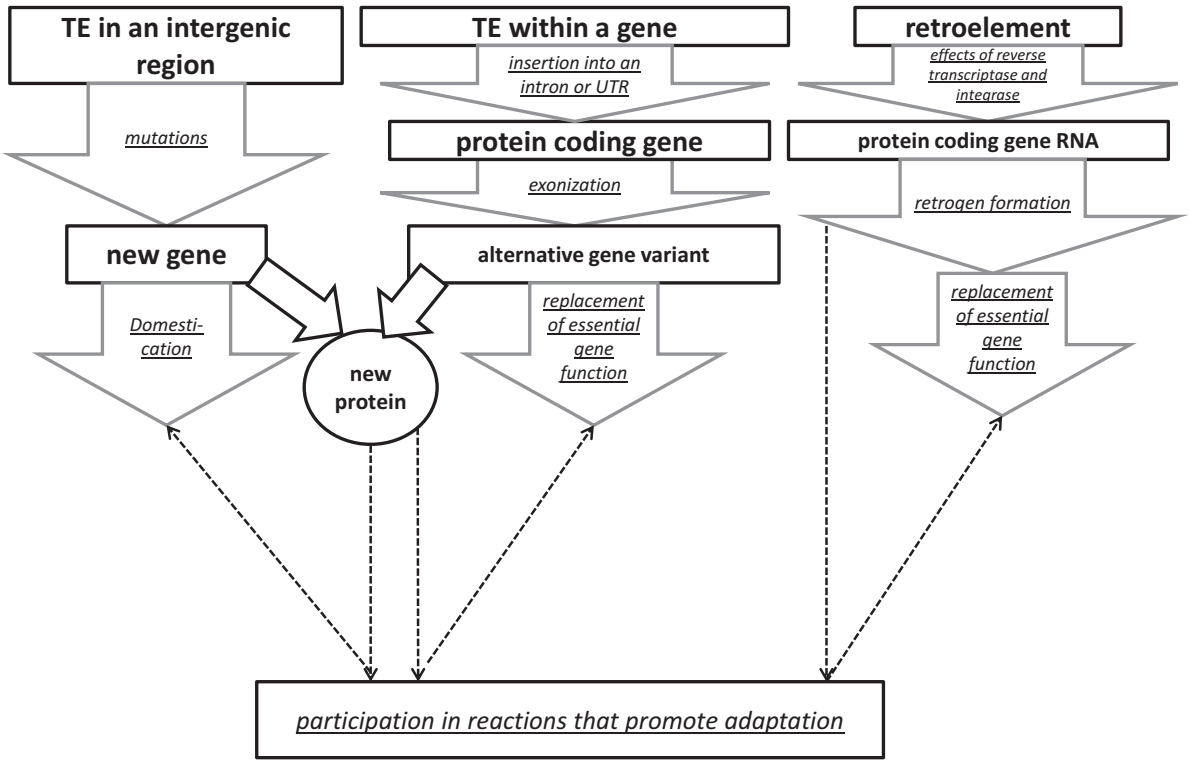

Fig. (2). Scheme of the formation of protein-coding genes from transposable elements (TEs) during evolution in three different ways: through domestication, through exonization, and through the formation of pseudogenes using retroelement enzymes. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

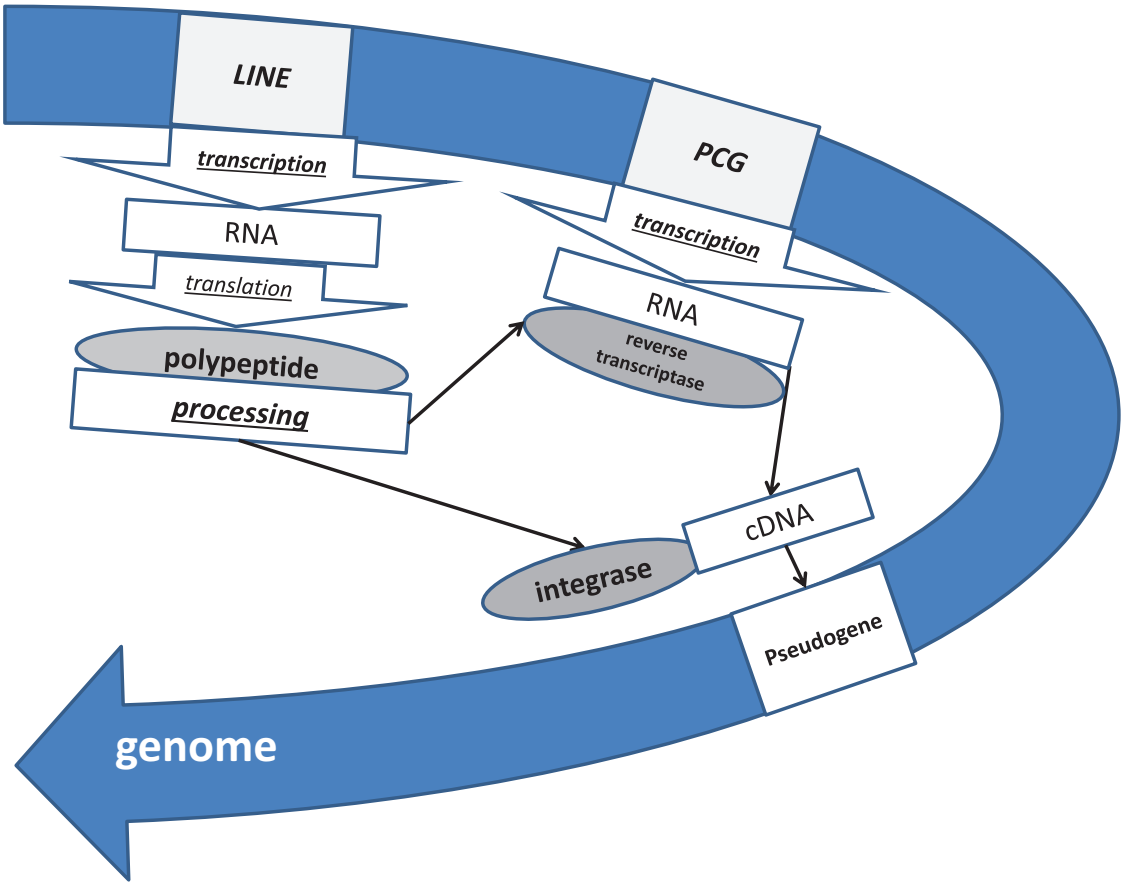

Fig. (3). Scheme of pseudogenes (retrocopies) formation using LINE retroelements. Reverse transcriptase of the LINE element produces complementary DNA (cDNA), which is then inserted into a new genomic locus using LINE integrase. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

The origin of new genes through the recruitment of TE domains is the reason for their spread due to the possibility of participation in adaptive reactions. Evidence for domestication of TEs in evolution has been obtained due to the presence of functional differences between autonomous and domesticated TEs [44]. The formation of new functional domains from sequences encoded in TE genes can explain the exonization of thousands of genes in the human genome [36, 37]. The use of TE genes for host needs can be explained by the mutual regulation of TE proteins with their own sequences, resulting in TEs becoming the evolutionary sources of both transcription factors and binding sites for these transcription factors. This provides the possibility of self- and mutual control of TEs, inactivated as a result of the accumulation of mutations in them and domesticated for the needs of the hosts [11]. Genes of both DNA transposons, such as transposase, and retroelements are domesticated due to the transformation of regulatory networks, which also evolve as a result of the dynamic influences of new TEs insertions into eukaryotic genomes [44].

TEs turned out to be sources of the emergence of conservative genes that play a key role in the evolutionary transformations of living organisms. For example, RE reverse transcriptase became the basis for telomerase formation [36, 45]. In *Drosophila* evolution, telomerase was replaced by HeT-A, TART (telomere-associated retrotransposon) and TAHRE (telomere-associated and HeT-related element) retroelements [46]. The mammalian centromeric protein CENP-B is homologous to the transposase (DNA transposon enzyme) pogo from which it originated [47]. Back in 2005, it was shown that in vertebrates, REs became the source of more than 1000 genes in evolution [48]. A large number of genes derived from TEs have been identified in animals, plants and yeast [49]. TE-derived PCGs can encode transcription factors through the use of DNA-binding domains of TEs, which is important for controlling genome function. Moreover, many binding sites for transcription factors also originate from TEs sequences [11]. This indicates the global role of TEs in the formation of regulatory gene networks [49, 50], which may also involve non-coding RNAs and their translation products (Fig. **4**). TE transcripts are selected for functional fitness in parallel with their proteins since ncRNAs are formed in evolution from transposons [33, 35] and are capable of being translated into peptides [21-23, 29, 51, 52].

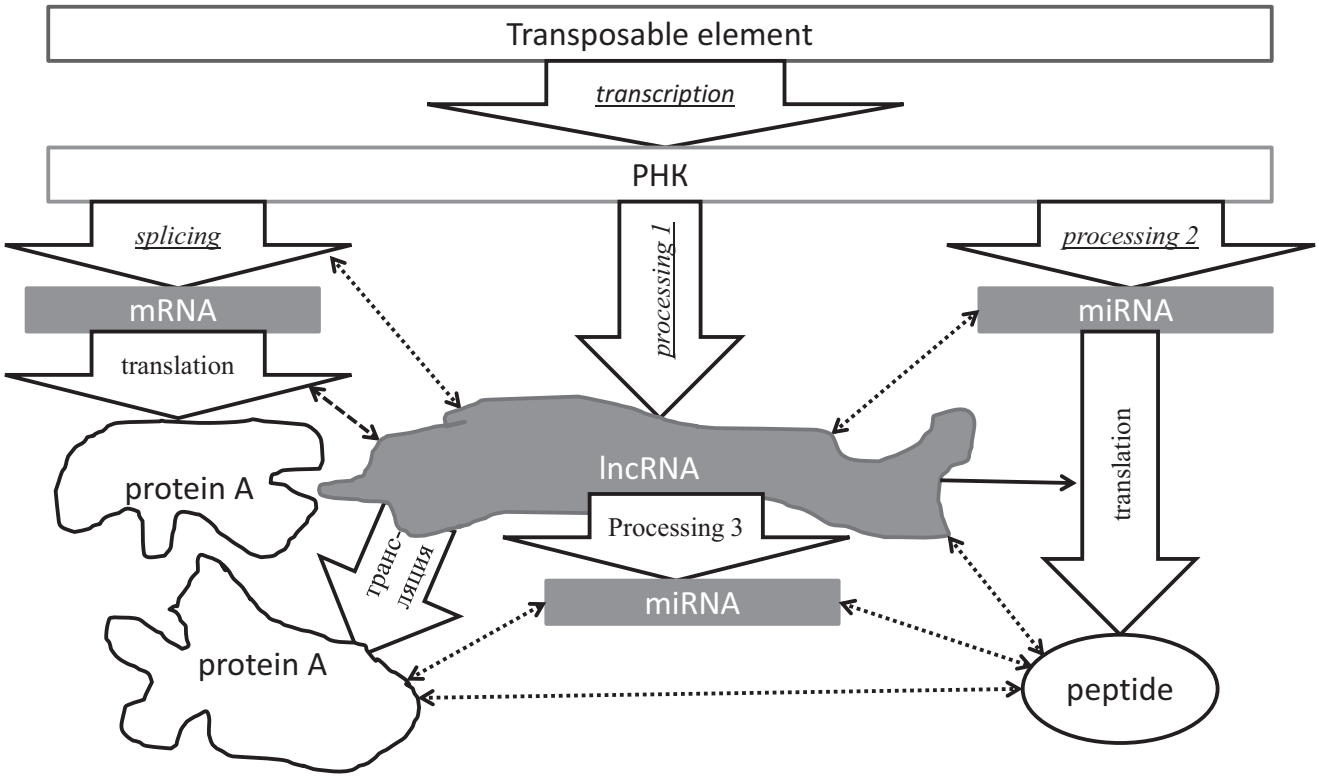

Fig. (4). Scheme of the polyfunctionality of TE transcripts as participants in gene regulatory networks. The scheme shows that three different processing products can be formed from the primary TE transcripts: 1) Messenger RNA (mRNA), which is translated to form the enzymes necessary for transposition and pseudogene formation. 2) Long non-coding RNAs (lncRNAs), which can be translated to form new proteins, and also processed to form mature microRNAs. 3) Pri-microRNA, which is then translated to form a peptide that can interact with various proteins. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

A large amount of evidence has accumulated for the origin of PCGs through the domestication of specific genes of various TEs. The source of the most important PCGs in animals, plants and fungi turned out to be DNA transposon transposase (Table **1**). Such conservative genes as *CENP-B* of all eukaryotes, *RAG* gene in vertebrates, *Daysleeper* gene in drosophila, *BUSTER1-3, ZBED1, ZBED4, ZBED5, P52rlPK* gene in arabidopsis originated from transposase [44, 53]. In humans, rats, mice, cows, pigs, chickens, frogs, and various species of fish, Harbinger transposases gave rise to genes encoding HARBI1 proteins, which serve as nucleases necessary for DNA rearrangement [54]. Transposases have become the basis for the formation of eukaryotic DNA-binding domains such as sprial-turn-sprhal (STH), zinc fingers (ZF), HD (homeodomain), KRAB (Kruppel-linked box), BTB (Broad-Complex, Tramtrack, and Bric-a-brac), SET (Su(-var), E(z) and Trithorax), SWIM (SWI2/SNF2 and MuDR), hATC (hAT C-terminal dimerization), LZ (leucine zipper) [49]. The protein products of DNA-transposons are characterized by lower contact order compared to randomly selected control proteins. This allows them to fold quickly and avoid aggregation [55]. The role of DNA-transposons as evolutionary sources of PCGs involved in epigenetic regulation should be noted. Thus, the *HDP1* (from transposase) and *HDP2* (from DNA-binding protein) genes, which interact with the components of the acetyltransferase complex IDM1, IDM2, IDM3, MBD7, involved in DNA demethylation, originated from Harbinger in Arabidopsis [56]. The hAT transposase gave rise to the chromatin modification factors BEAF-32 and HIM-17 [57], the centromeric protein Abp1 (in fungi) [49], and insulators involved in chromatin modification [58]. Genes derived from DNA-transposons are expressed into proteins THAP0, THAP1, E93 (*Eip93F* gene), involved in apoptosis, as well as THAP, LIN-36, LIN-15B, which control the cell cycle [57]. In yeast, DNA- transposon integrase gene became the source of the *Fob1p* gene, the product of which controls ribosomal RNA recombination [44, 53]. All this indicates the key role of TEs in the origin and evolution of proteins in eukaryotes, including humans.

Table 1. Human and other eukaryotes protein-coding genes originated from transposase.

DNA-Transposon, Encoding a Transposase	Transposon-Derived Gene	Gene-encoded Protein Function	Eukaryote Name
Tc1/mariner	*PAX6*	Transcription factor regulating organ development	*Homo sapiens*
	haT-like	Participation in reproduction	*Homo sapiens* and other mammals
	SETMAR	Non-homologous repair of double DNA breaks	
	METNASE	DNA integration and repair	
Transib, Tc1	*RAG1, RAG2*	Immune reactions (V(D)J recombination)	*Homo sapiens* and other jawed vertebrates
PiggyBac-like	*PGBD1*	Brain functioning	*Homo sapiens* and other vertebrates
Pogo	*CENP-B*	Centromere functioning	*Homo sapiens*, other animals and fungi
	JRK, JRKL	DNA and RNA binding activity in neurons	*Homo sapiens* and other mammals
P-element	*THAP0, THAP1*	Interferon-induced apoptosis	*Homo sapiens* and other mammals
	THAP2, 3, 4, 5, 6, 10, 11	DNA binding activity and metal ion binding activity	
	THAP7	Recruits HDAC3, NCoR deacetylase to specific DNA sites	
	E2F6	Repressor of E2F-dependent transcription in the S-phase	
hAT	*GTF2IRD2*	Transcription factor	*Homo sapiens* and other mammals
	P52rIPK	Inhibition of PKR protein kinase	
	ZBED1, 4, 5	Transcription regulation	
Harbinger	*HARBI1*	DNA rearrangement	
	NAIF1	Nuclear apoptosis factor	*Homo sapiens* and other vertebrates
Pogo	*Bab1, Bab2*	Morphogenesis of the ovaries and body	*Drosophila*
	Eip93F	Regulation of steroid-mediated programmed cell death during metamorphosis	Fungi
	Abp1	Segregation of chromosomes, formation of centromere heterochromatin	
P-element	*CDC14B*	Cell cycle inhibitor	Roundworms
	CTB-1	Regulation of transcription	
	HIM-17	Chromatin modification	
	LIN-36	Cell cycle inhibitor	
	phsa/pgga	Repressor-like protein with a DNA-binding motif	*Drosophila*
Mutator	*Aft1*	Transcription factor	*Saccharomyces cerevisiae, Lodderomyces elongisporus*
	Rcs1		
	MUG1	Regulation of transcription	*Arabidopsis*
Mutator-like	*FHY3, FAR1*	Phytochrome A pathways	
Harbinger	*HDP1, HDP2*	DNA demethylation	
hAT	*Daysleeper*	Development regulation	
hAT	*BEAF-32*	Insulator activity for chromatin regulation	*Drosophila*
hAT	*DREF*	DNA replication, cell differentiation	
hAT	*GON-14*	Regulation of development	*Caenorhabditis elegans*
hAT	*LIN-15B*	Cell cycle inhibition	

Note: *[44-49, 53-58]

The sources of many PCGs in evolution were also REs (Table **2**). The *Gin-1* gene in mammals, which is involved in the regulation of embryogenesis, originated from LTR-RE *integrase* gene. The following genes originated from the *GAG* retroelement gene in mammals: *Mart* (participates in the control of cell proliferation and apoptosis), *Ma* (involved in autoimmune response), *Fv1* (participates in limiting the replication of murine leukemia virus), *PEG10* (participates in parthenogenetic development), *Rtl1* (regulates the expression of paternal genes), *MyEF-3* (myelin basic protein transcription factor) [44, 53]. *SCAN* (transcription factors) gene family in vertebrates originated from the *GAG* gene of the Ty3/gypsy retroelements [34]. The *Syncytin-1, -2, -A, -B* genes originated from the *Env* genes of ERVs [44, 53]. Proteins derived from the *Env* gene of LTR retroelements function in antiviral host defense in several vertebrate species.

Table 2. Human and other eukaryotes protein-coding genes originated from retroelement genes.

Retroelement/ Source Gene	Retroelement-Derived Gene	Gene-encoded Protein Function	Eukaryote Name [Reference]
LTR-RE/ *integrase*	*Gin-1*	Embryogenesis regulation	*Homo sapiens* and other mammals [44, 53]
LTR-RE/ *GAG*	*Mart*	Control of cell proliferation and apoptosis	
	Ma	Immune response	
	MyEF-3	Transcription factor	
	Fv1	Antiviral response	
	PEG10	Regulation of cell proliferation, differentiation and apoptosis	
	Rtl1	Regulation of paternal genes expression	
LTR-RE/ *GAG*	*SCAN*	Transcription factor	*Homo sapiens* and other vertebrates [34]
ERV/ *Env*	*Syncytin*	Regulation of placental development	*Homo sapiens* and other vertebrates [44, 53]
LINE/*ORF1*	*L1TD1*	RNA-binding protein that functions in undifferentiated cells	*Homo sapiens* and other mammals [65]
ERV/protease	*SASPase*	Cell protease	*Homo sapiens* [66]
ERV/ *Env*	*Iris*	Antiviral response	*Drosophila* [60]

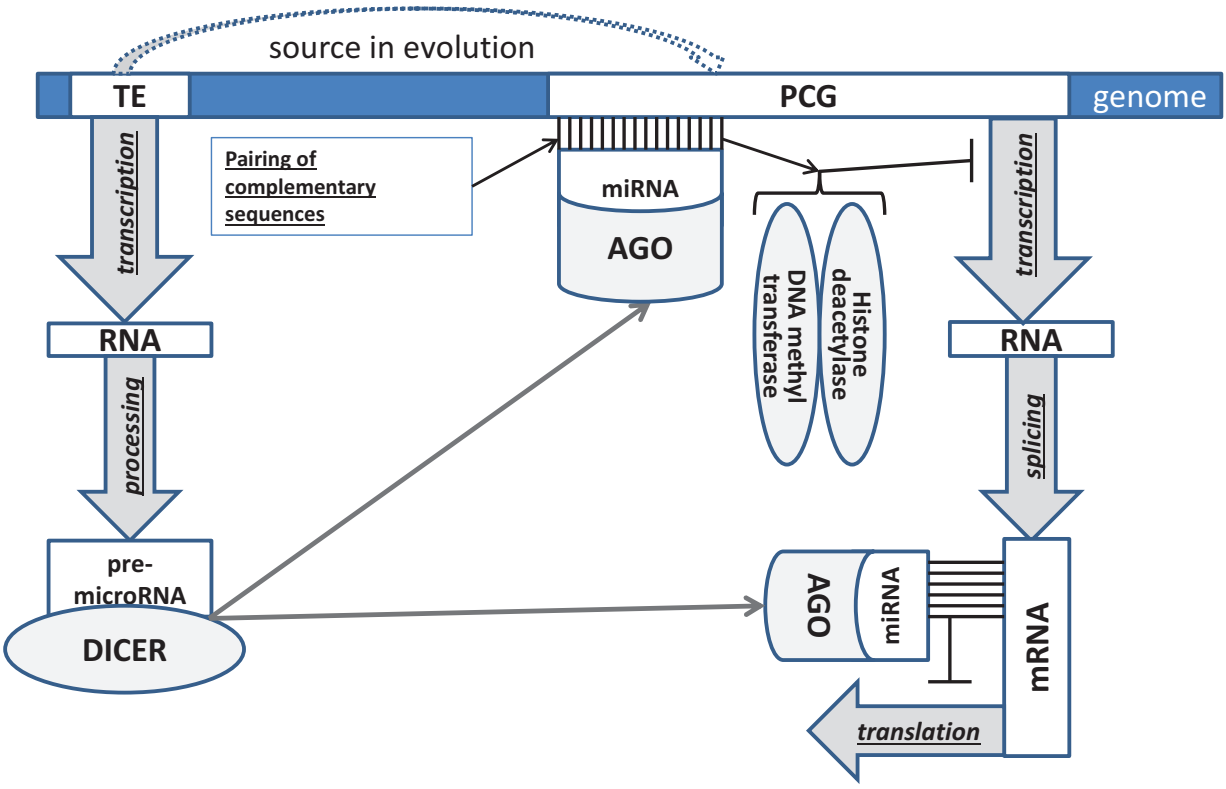

Fig. (5). Scheme of transposable elements participation in epigenetic control of genes encoding proteins derived from them in evolution (indicated by an arrow). MicroRNAs formed by processing transcripts of transposable elements can inhibit protein-coding gene expression at the transcriptional level due to complementary base pairing, as well as at the post-transcriptional level. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

They form surface (determines cell specificity, tropism for the host and cell type) and transmembrane (required for the fusion of the envelopes of the virus and the target cell) subunits. In all cases, endogenous *env* genes derived from different ERVs act as restriction factors for related exogenous retroviruses [59]. The homologue of the *env* gene of ERV in

drosophila is the *Iris* gene [60]. *Syncytin* genes in placental mammals are derived from the *Env* genes of LTR REs. These genes are involved in the formation of the placenta and include the primate-specific *Syncytin-1* (descended from HERV-W 25 million years ago) and *Syncytin-2* (descended from HERV-FRD 40 million years ago) [61]. Moreover, it has been shown that in rodents [62], carnivores [63], lagomorphs [64], and primates [61], *Syncytin* genes originated from REs as a result of independent evolutionary events, which confirms the key role of REs in the regulation of ontogenesis. The *L1TD1* gene, which encodes an RNA-binding protein that functions in undifferentiated cells, evolved from ORF1 LINE retroelements in humans and other mammals [65]. The human gene that encodes the protein named Skin Aspartic Protease (SASPase) evolved from ERV [66]. Domestication of TE genes for the needs of hosts leads to the fact that new PCGs contain transposon sequences not only in their introns and regulatory regions but also in conserved exons. This provides a basis for epigenetic control of PCGs at the transcriptional and posttranscriptional levels, with transposon-derived ncRNAs serving as universal RNA interference tools (Fig. **5**).

4. ORIGIN OF PROTEIN-CODING AND PEPTIDE-CODING GENES FROM NON-CODING RNAS

The emergence of lncRNAs and microRNAs from TEs in evolution could be the reason for the multifunctionality of these ncRNAs. It means that despite the name "non-coding", they have the potential for translation into functional proteins and peptides. The reason is the universal property of TEs to contain nucleotide sequences that bind to ribosomes even in the presence of a non-canonical short open reading frame (ORF), which is due to the role of TEs as universal evolutionary sources of life on Earth [6]. Several phylogenetic studies have revealed the origin of evolutionarily novel PCGs from lncRNA genes [24-27]. A new gene, *BSC4*, has been identified in *Saccharomyces cerevisiae*, containing an ORF (which is absent in other closely related fungal species) encoding a protein 132 amino acids long. The *BSC4* gene product is involved in DNA repair pathways during the stationary phase of the fungus and contributes to its persistence when transferred to low-nutrient environments. In fungi of the species *S. paradoxus, S. mikatae, S. bayanus*, identical non-coding sequences are also transcribed, so they can be classified as the lncRNA gene, which in Saccharomyces cerevisiae in evolution became the source of a new PCG, domesticated in connection with adaptive functions [25].

A comparative analysis of the *Drosophila melanogaster* genome with other species of the genus *Drosophila* revealed 5 genes, 4 of which are located on the X chromosome and translated into peptides of 58, 79, 97 and 158 amino acids in length, 1 - on chromosome 2 with the encoding of a peptide of 127 amino acids. These genes are derived from non-coding DNA sequences transcribed in other species (lncRNA genes) [24]. In 2012, 24 evolutionarily young PCGs were identified, which evolved from lncRNAs genes. Of these, 11 genes encode proteins specific only to humans, and 13 others are conserved between humans and chimpanzees. Genes contain from 1 to 7 exons, and the length of gene translation products varies from 72 to 423 amino acids [26]. Such genes are called "orphan", that is, characteristic of a specific animal species and not found in other species, which is associated with their recent emergence in evolution from lncRNAs and transposon genes and participation in species-specific adaptive reactions. About 53% of "orphan" primate genes contain TE sequences, which indicates their use as sources of origin of these genes [67].

Confirmation of the role of lncRNAs as sources of "orphan" genes were obtained in studies in 2014, which found that ribosome-associated lncRNAs are characterized by low evolutionary conservation and contain homologs in other species from 0 to 15.6%, while for PCGs this figure more than 95% for vertebrates and 70 – 73% for plants and fungi. Analysis of the expression of lncRNAs in cells of 6 different eukaryotic species (*A. thaliana, S. cerevisiae, Danio rerio, D. melanogaster, Homo sapiens, Mus musculus*) showed similar coding potential and sequence restrictions with evolutionarily young proteins. The results showed that a significant proportion of lncRNAs are associated with ribosomes (from 28.6% in S. cerevisiae to 81.9% in mice), which indicated their expression. It was also revealed that evolutionarily young PCGs encoding experimentally tested proteins were characterized by properties common to lncRNAs genes. These PCGs contained short ORFs occupying a small portion of the transcript and also had a low coding score similar to that of lncRNAs. It indicates the origin of such evolutionarily young PCGs from lncRNAs genes due to the selection of peptides involved in properties adaptive for the survival of organisms [27]. The emergence of new PCGs from lncRNA genes in various animals and plants is due to their translation into peptides that participate in a variety of adaptive biological processes. Data on the role of lncRNAs as sources of PCGs indicate the possibility of the formation of larger polypeptide molecules from their gene transcripts [24-27]. This is possible due to the evolution of the resulting genes with an increase in their size due to insertions of TEs with their subsequent exonization [36, 37].

5. TRANSLATION OF HUMAN PROTEINS AND PEPTIDES FROM LONG NON-CODING RNAS AND PRIMIRNAS

In addition to the evolutionary role in the emergence of PCGs according to phylogenetic studies, more data are emerging on the direct translation of lncRNA in living organisms, including humans. In this case, the mechanisms of larger PCG formation from the entire or most of the lncRNA gene sequence may differ from the mechanisms of formation of a translated RNA molecule from lncRNA, the length

of which is significantly less than the mature lncRNA. For example, the skeletal muscle-specific peptide myoregulin (MLN) was discovered in mice and humans, formed from the ORF of an evolutionarily conserved lncRNA, which is designated LINC00948 in humans and AK009351 in mice. In both organisms, lncRNA genes consist of 3 exons and are more than 15,000 bp in length. At the same time, the OFR from which the MLN peptide is translated is only 138 bp long and is located in exon 3 of the gene. MLN shows similarities to phosphalambane and sarcolipin, which control muscle relaxation by regulating the uptake of calcium ions by the endoplasmic reticulum [21]. Another peptide, DWORF (dwarf open reading frame), produced during lncRNA translation, also interacts with the endoplasmic reticulum. It stimulates the activity of calcium ATPase SERCA (sarco/endoplasmic reticulum Ca2+-ATPase) by displacing its inhibitors (phospholamban, sarcolipin and MLN) in cardiomyocytes and skeletal muscles. The DWORF peptide is not expressed prenatally, but its production gradually increases in cardiac muscle in the postnatal period [22]. Various functional peptides produced during the translation of lncRNAs have also been identified in plants. These include COLDAIR and COOLAIR (control flowering time), IPS1 (regulates phosphate uptake), ENOD40 (participates in symbiosis with bacteria) [23].

Studies of human lncRNA translation are of greatest interest since the peptides produced during this process are specifically expressed or suppressed in malignant neoplasms and can serve as tools for cancer-targeted therapy. Full-length sequencing of translated RNAs and ribosome profiling published in 2019 indicate that 3,330 human lncRNAs bind to ribosomes with active translation elongation. The proteomic analysis allowed the authors to discover 308 new proteins and peptides formed as a result of the translation of lncRNAs [68]. Similar studies were carried out by other research groups, in which 128 peptides encoded by ncRNAs were identified [69]. It should be noted that some peptides are formed during the translation of lncRNAs not in normal cells of the body, but only in cancer cells, which indicates another mechanism for the evolution of cancer - due to the formation of new PCGs from lncRNAs. This reflects the general evolutionary mechanisms of tumors and the development of the living world, when possible variants of ncRNA translation are selected to form new adaptive properties of living organisms. For example, lncRNA LINC00675 is translated to produce a small, conserved 79 amino acid FORCP protein that is expressed by colorectal cancer cells, whereas it is not expressed by all normal human cells [70]. For the same type of cancer, the formation of the ASAP peptide from lncRNA LINC00467 was detected, which is not only expressed in normal cells but is also conserved for higher mammals, regulating mitochondrial function [71]. This indicates the development of different pathways of cancer evolution using both newly formed and conserved peptides involved in carcinogenesis. It was also shown that the expression of lncRNAs is influenced by changes in the internal environment of the body and the external environment, which indicates the potential of research in this direction to identify risk factors for cancer and the possibility of their correction. Thus, the expression of the YY1BM peptide is influenced not only by androgens (and therefore it is specifically translated in men), but also by cigarette smoke [72]. Several lncRNA genes that encoded MHC class I bound peptides have recently been described. Such peptides drove a potent antigen-specific CD8 T lymphocyte response, which translated into a significant delay in tumour growth and can be deployed as a cancer vaccine. Increased lncRNA Gm17173 (encoding peptide sequence RLAQLQTTI) and Gm37283 (encoding peptide sequence HIFSLHHF) expression were detected in tumor tissues. These peptides that were tested in mice could drive an antigen-specific T cell response when delivered as a vaccine [73]. The metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) is broadly expressed across cell types, almost exclusively as a nuclear long noncoding RNA. Within a 5' region of Malat1 containing short open reading frames. In differentiating neurons, a portion of Malat1 RNA redistributes to the cytoplasm, where it is translated to form the M1 peptide that modulates synaptic function [74].

Viruses are also sources of lncRNAs that affect the proteome of the human genome. For example, Human Immunodeficiency Virus 1 lncRNA named ASP promotes viral latency by recruiting the Polycomb Repressor Complex 2 and promoting nucleosome assembly [75]. Viral RNA-host protein interactions are critical for the replication of flaviviruses. It was found that three conserved host RNA-binding proteins (RBPs) G3BP1, G3BP2 and CAPRIN1 in dengue virus (DENV-2) infection are regulators of the interferon response against DENV-2. Their antiviral activity was antagonized by the abundant DENV-2 non-coding subgenomic flaviviral RNA, which bound to G3BP1, G3BP2 and CAPRIN1, inhibited their activity and led to profound inhibition of interferon-stimulated genes mRNA translation [76]. This indicates the role of viral lncRNAs as adaptive transcripts aimed at combating host antiviral systems. Since exogenous viruses are evolutionarily associated with TEs [77], it can be assumed that the emergence of lncRNAs in evolution was also aimed at confronting host systems aimed at eliminating TEs from genomes. Adenovirus кодирует lncRNA named VA RNA relieves the cellular anti-viral blockade of protein synthesis through inhibition of the double-stranded RNA-activated protein kinase [78]. Since exogenous viruses probably evolved from TEs [77], it is possible that viral lncRNA can be translated to form peptides that have an inhibitory effect on the viruses themselves. This is likely due to host defense systems when translation of foreign RNAs is used against pathogens. Discovery of such peptides may lead to new directions in the treatment of viral infection. The basis for research in this direction is the results of studies that make it possible to identify peptides encoded by lncRNAs (Table **3**). Indeed Epstein-Barr virus encodes lncRNAs: oriPtLs and oriPtRs. While the oriPtLs are noncoding, some of the oriPtRs contain the BCRF1/vIL10 ORF and at the same time may serve a noncoding function [79].

Table 3. Human lncRNAs and peptides translated from their open reading frame.

lncRNA Name	Tissue where the Peptide is Expressed/ Reference	Peptide Name/ Length	Peptide Function
MALAT1	Wild-type neurons/ [74]	M1/35	Modulates synaptic function in the brain
CRNDE	Highly proliferating tissues/ [80]	CRNDEP/84	Regulates epithelial cells cell proliferation
LINC01420	Breast cancer and leukemia cells/ [81]	NoBody/68	Interact with mRNA decapitating proteins that remove the 5' cap of mRNA, promoting 5' to 3' decay, cancer stimulator
LINC00665	Breast cancer/ [82]	CIP2A-BP/52	Interacts with CIP2A oncogene, activating PP2A, which inhibits the PI3K/AKT/NFκB pathway, which leads to a decrease in the expression of MMP-2, MMP-9 and Snail, cancer inhibitor
LINC00908	Triple-negative breast cancer/ [83]	ASRPS/60	Inhibits STAT3 phosphorylation, which leads to decreased VEGF expression, cancer inhibitor
MAGI2-AS3	Breast cancer/ [84]	MAGI2-AS3-ORF5/ polypeptide	Affect the progression of breast cancer cells by binding to extracellular matrix (ECM)-related proteins
NR029453	Breast cancer/ [85]	CASIMO1/80	Interacts with squalene epoxidase to regulate lipid droplet clustering, cancer inhibitor
HOXB-AS3	Colon cancer/ [86]	HOXB-AS3/53	Suppresses hnRNP A1-mediated splicing of pyruvate kinase M by blocking arginine residues in the RGG motif of hnRNP A1, cancer inhibitor
Gm17173	Human HTC116 p53-/- colon cancer cells/ [73]	Peptide sequence RLAQLQTTI/9	Drive an antigen-specific T cell response, cancer inhibitor
Gm37283		Peptide sequence HIFSLHHF/8	
LOC90024	Colorectal cancer/ [87]	SRSP/130	Interacts with splicing regulators SRSF3, stimulating tumor splicing of RNA transcription factor Sp4, cancer stimulator
LINC00675	Colorectal cancer/ [70]	FORCP/79	Inhibits apoptosis of cancer cells, cancer stimulator
LINC00467	Colorectal cancer/ [71]	ASAP/94	Stimulates ATP synthetase activity by interacting with ATP5A and ATP5C, cancer stimulator
LINC-00266-1	Colorectal cancer/ [88]	RBRP/71	Interacts with RNA-binding proteins such as m6A reader IGF2BP1 to increase c-Myc stability and expression, cancer stimulator
BVES-AS1	Colorectal cancer/ [89]	BVES-AS1-201-50aa/50	Targeted and activated the Src/mTOR signaling pathway in cancer, cancer stimulator
LINC00278	Esophageal cancer/ [72]	YY1BM/21	Inhibits the interaction of YY1 with the androgen receptor, which reduces eEF2K expression, cancer inhibitor
LINC00998	Hepatocellular carcinoma/ [90]	SMIM30/59	Promotes the attachment of non-receptor tyrosine kinases SRC/YES to the membrane, activating the MAPK signaling pathway, cancer stimulator
NCBP2-AS2	Hepatocellular carcinoma / [91]	KRASIM/90	Interacts with the KRAS protein, inhibiting ERK signaling pathways, cancer inhibitor
LINC-PINT	Glioblastoma/ [92]	PINT87aa/87	Interacts with polymerase-associated factors PAF1c complex, cancer inhibitor
AF127577.4	Glioblastoma/ [93]	AF127577.4-ORF/29	Modulation of ERK2/METTL3 interaction, cancer inhibitor
HNF4A-AS1	Neuroblastoma/ [94]	HNF4A-AS1-51aa/ 51	eEF1A1-repressed SMAD4 transactivation, cancer stimulator
UBAP1-AST6	Lung cancer/ [68]	UBAP1-AST6	Stimulates cell proliferation, cancer stimulator
ASH1L-AS1	Acute myeloid leukemia/ [95]	APPLE/ 90	Promotes PABPC1-eIF4G interaction and facilitates mRNA circularization and eIF4F initiation complex assembly to support a specific pro--cancer translation program, cancer stimulator

It should be noted that drug resistance is an important characteristic of cancer cells, therefore, when considering the role of TEs and ncRNAs in the formation of proteins and peptides that affect cancer, it is important to dwell on the issue of drug resistance. Indeed, the de-regulation of TEs is the most important mechanism for the development of drug resistance, since TEs are able to rebuild global regulatory gene networks [96]. This is due to the influence of TEs on epigenetic regulation [97], due to which cell clones become resistant to chemotherapy used to treat cancer. In particular, it was revealed that TEs alter gene expression and may impact response to cisplatin therapy in ovarian cancer [98]. Of the lncRNAs listed in Table **3**, LINC-PINT is involved in drug resistance, which suppresses cisplatin resistance in gastric cancer by inhibiting autophagy activation *via* epigenetic silencing of ATG5 by EZH2 [99]. Also, LINC-PINT attenuates paclitaxel resistance in triple-negative breast cancer cells *via* targeting the RNA-binding protein NONO [100] and regulates laryngeal carcinoma cell stemness and chemoresistance through miR-425-5p/PTCH1/SHH axis [101]. In addition, Anisomycin inhibits the activity of human ovarian cancer stem cells *via* regulating antisense RNA NCBP2-AS2/MEK/ERK/STAT3 signaling. Therefore, mutations or changes in the expression of this lncRNA may be a

factor in drug resistance to anisomycin [102]. LINC00467 mediates the 5-fluorouracil resistance in breast cancer cells [103], and knockdown of this lncRNA contributed to Axitinib sensitivity in hepatocellular carcinoma through miR-509-3p/PDGFRA axis [104]. LINC00675 activates the androgen receptor axis signaling pathway to promote castration-resistant prostate cancer progression [105].

Transcriptional suppression of Dicer by the HOXB-AS3/EZH2 complex dictates sorafenib resistance and cancer stemness [106]. MAGI2-AS3 suppresses castration-resistant prostate cancer proliferation and migration *via* the miR-106a-5p/RAB31 axis [107]. Long non-coding RNA LINC00665 promotes gemcitabine resistance of Cholangiocarcinoma cells *via* regulating EMT and stemness properties through miR-424-5p/BCL9L axis [108] and induces acquired resistance to gefitinib through recruiting EZH2 and activation PI3K/AKT pathway in non-small-cell lung cancer [109]. CRNDE induces cisplatin resistance through the SRSF1/TIA1 signaling pathway in ovarian cancer [110] and is involved in radiation resistance in hepatocellular carcinoma *via* the SP1/PDK1 axis [111]. MALAT1 mediates cisplatin resistance *via* miR-101-3p/VEGF-C pathway in bladder cancer [112], modulates oxaliplatin resistance of gastric cancer *via* sponging miR-22-3p [113], mediates doxorubicin resistance of hepatocellular carcinoma by regulating miR-3129-5p/Nova1 axis [114], regulates the resistance of breast cancer cells to paclitaxel *via* the miR-497-5p/SHOC2 axis [115], enhances gemcitabine resistance in non-small cell lung cancer cells by directly affecting miR-27a-5p/PBOV1 axis [116].

Recent studies suggest that lncRNAs may influence drug resistance in cancer by targeting specific proteins. For example, exosomal lncRNA LOC85009 inhibits docetaxel resistance in lung adenocarcinoma by regulating ATG5-induced autophagy *via* the USP5/USF1 axis, suggesting that LOC85009 might be a potential target to reverse docetaxel resistance in the treatment of lung adenocarcinoma [117]. There were identified the role of different isoforms (by alternative polyadenylation) of XBP1 in cisplatin resistance of lung adenocarcinoma cells. Due to decreased usage of the distal polyadenylation site (PAS) in 3'UTR, the XBP1-UL level was lowered in cancer cells and decreased further in cisplatin-resistant cells. XBP1 transcripts with shorter 3'UTR (XBP1-US) were more stable and presented stronger potentiation on DDP resistance. DDP-induced nuclear lncRNA LINC00221 facilitated CPSF6-induced proximal PAS choice in the pre-XBP1 3'end, influencing progression and cisplatin resistance in lung adenocarcinoma [118].

Some lncRNAs are also precursors for microRNAs, so their DNA coding regions are both lncRNA genes and microRNA genes. Moreover, the transcription products of such genes also have the potential to be translated into functional peptides. Examples are lncRNA MIR497HG [32], lncRNA MIR22HG [33], lncRNA MIR155HG [31]. Pri-miRNA molecules can also be translated into functional peptides called miPEPs [29, 51, 52] due to the presence of short ORFs that allow them to interact with ribosomes [51]. Peptides and proteins encoded by pri-microRNAs are expressed in plant and animal cells, regulating the growth and development of normal and cancer cells [30]. MiPEP (miRNA encoded peptide) peptides, which are formed during the translation of pri-microRNAs, are involved in the regulation of both PCGs [52] and their own microRNA genes. The functional activity of miPEP has been proven. They are used in agriculture to improve plant productivity. For example, miPEP172c enhances the expression of miR172c, which promotes the formation of soybean root nodules and stimulates symbiosis with rhizobia [51].

In *Arabidopsis*, a functional miPEP165a was identified and it was proven that more than each of the 50 pri-microRNAs of this plant contains a short ORF. Pri-miR171b was detected in *Medicago truncatula* [29]. A functional peptide miPEP31 was identified, formed during translation of pri-miRNA-31, which suppresses the expression of miR-31, and enhances the induction of regulatory T lymphocytes, acting as a transcriptional repressor [119]. According to a 2023 systematic review of the scientific literature, miPEPs are currently known from various plants and animals, including miPEP-156a in *Brassica rapa,* miPEP397a in *Brassica oleacera,* miPEP-164b, miPEP165a, miPEP-397a, miPEP858a in *Arabidopsis thaliana,* miPEP164b in *Barbarea vulgaris,* miPEP-164c, miPEP-171d1, miPEP-172d, miPEP3635b in *Vitis vinifera,* miPEP171a/b/c/d/e/f in *Medicago truncatula,* miPEP171b in *Lotus japonicas,* miPEP171i y *Oryza sativa,* miPEP171e in *Solanum lycopersicum,* miPEP172c in *Glycine max,* miPEP-31 in mouse, miPEP8 in *Drosophila* [120]. MiPEP408, which is involved in the response to arsenic stress and sulfur assimilation in *Arabidopsis*, was also recently described [121].

In addition to the involvement of human miPEPs in the antiviral response [122, 123], microRNAs are also encoded in viral genomes. Moreover, virus-derived miRNAs (v-miR) function by targeting host and virus-encoded transcripts and are critical in shaping host-pathogen interaction. V-miRs can function in an autocrine and paracrine fashion to counter antiviral pathways in infected cells and likely modulate immune responses in bystander cells [124]. Since in evolution viruses and TEs were characterized by mutual transformations [77], it can be assumed that viral microRNAs can also be translated, which would become the basis for a new direction in antiviral therapy since miPEPs are characterized by inhibition of the expression of their own microRNAs [30]. Indeed, viruses have been found to regulate their proteins by microRNAs formed from their genomes. For example, when Epstein Barr Virus multiplies, miR-BART-1-5p, miR-BART3, miR-BART5, miR-BART16, miR-BART17-5p are formed, the targets of which are Latent membrane protein 1 (LMP1, a membrane protein expressed by EBV). LMP1 acts as a functional homolog of CD40, a constitutively active TNF receptor. miR-BART16 suppresses the type I interferon signaling by directly targeting and downregulating the CREB binding protein (CREBBP or CBP), which is a well-characterized transcriptional co-activator of interferon signaling [124]. Expression products of TEs are also interferon stimulators [125], which indicates the evolutionary relation-

ship of exogenous viruses and TEs, as well as the systems of their interaction with the human body at various levels. The study of microRNAs of viruses and their interactions with their own proteins and peptides, as well as with host molecules, is promising for the development of new methods of antiviral and antitumor therapy due to the possibility of using convenient tools for these - specific peptides.

The ability to translate into functional peptides is characterized not only by lncRNAs and pri-microRNAs, but also by other known RNA molecules, such as snoRNAs [69]. It demonstrates the enormous number of different peptide molecules that have not yet been discovered, functioning both specifically for individual tissues or organs and for pathological processes such as cancer. That is, the regulatory systems of eukaryotic genomes are much more complex than the classical concepts prevailing in genetics. TEs could be the oldest universal system providing evolutionary mutual regulatory tools for all known organisms, starting from the origin of life. This assumption is supported by the above data on the nonrandom processing of mature tRNAs, rRNAs, snRNAs, and snoRNAs characteristic of TEs. The ncRNAs formed from them are used to interfere with TEs [126-131]. Table **4** presents data on the formation of both mature microRNAs with their specific functions and peptides from pri-microRNA molecules (which can serve as long ncRNAs).

Since the peptides encoded by pri-miRNAs influence the expression of self-miRNAs involved in carcinogenesis, as well as signaling pathways similar to self-miRNAs, it is promising to study the role of such peptides in cancer drug resistance. For this purpose, the influence of miRNAs on cancer drug resistance can be initially investigated. LncRNAs can also be used as tools to modulate microRNA expression. For example, it has been shown that HDAC1/4-mediated silencing of microRNA-200b (miR-200b) enhances docetaxel-resistance of human lung adenocarcinoma cells. There were confirmed that MARCKSL1-2 alleviated docetaxel resistance in lung adenocarcinoma cells by abolishing the inhibitory effect of HDAC1 on miR-200b *via* the recruitment of SUZ12 [135]. Tumor-suppressor miR-22 (which pri-miR-22 is translated into antiviral miPEP22 [122]) was found silenced in drug-resistant colorectal cancers [136]. MiR-155 (which pri-miR-155 is translated into miPEP-155 [31]) takes part in a sequence of bioprocesses that contribute to the development of breast cancer drug resistance, including repression of FOXO3a, enhancement of epithelial-to-mesenchymal transition (EMT) and mitogen-activated protein kinase (MAPK) signaling, reduction of RhoA, and affecting the length of telomeres [137]. Therefore, it can be assumed that miPEP-155 can be used to treat chemoresistant breast tumors by inhibiting miR-155. MiR-31 (which pri-miR-31 is translated into miPEP-31 [119]) confers Oxaliplatin resistance by targeting LATS2 in colorectal cancer [138]. It can be assumed that miPEP-31 can be used for the treatment of chemoresistant colorectal cancer by inhibiting miR-31. MiR-497 (which pri-miR-497 is translated into miPEP-497 [32]) mediates Tamoxifen resistance *via* PI3K/AKT signaling in breast cancer [139]. Therefore, miPEP-497 could also be used to overcome Tamoxifen resistance in breast cancer. It can be assumed that peptides will be found for other microRNAs involved in carcinogenesis that can be used as tools in the complex treatment of chemoresistant cancer.

6. FUNCTIONALITY OF MRNA OF PROTEIN-CODING GENES AS A PROPERTY OF TRANSPOSABLE ELEMENTS

Introns of PCGs are also processed to form functional ncRNAs [140]. This confirms their evolutionary relationship with TEs, which are key sources of small ncRNAs that mature through transcript processing [33]. Like lncRNAs, pre-mRNA molecules of protein-coding genes can form specific secondary and tertiary structures that affect the processing and stability of these molecules [141]. This suggests an ancient origin of the mRNA sequences translated into peptides [142-145] from TEs. The abundance of lncRNA domains with transposon sequences is due to the fact that TEs form spatial structures involved in biological processes [13]. Algorithms using minimum free energy and maximum expected accuracy are used to predict RNA secondary structures with comparative evolutionary methods. Using these methods, it was shown that the secondary structure of pre-mRNA can enhance or suppress its own splicing depending on the sequence features in the introns. Even after splicing, the secondary structure of RNAs can affect their regulation by RNAi. The tertiary structure also influences pre-mRNA splicing. For example, a G-quadruplex can enhance or suppress splicing by creating or shading RNA-protein binding sites [141]. In addition, REs are able to control the recognition of splice sites by forming specific spatial hairpin structures [146].

The functionality of PCG mRNA molecules is evidenced by the formation of specific circular RNAs (circRNAs) from gene introns, which have multiple regulatory abilities and can also be translated to form functional peptides and proteins. CircRNAs are widely expressed in the genomes of various eukaryotes because they are formed by “forward back splicing,” which results in the formation of covalent, closed-loop RNA structures without 3′-5′ polarity. They do not contain PolyA tails and are expressed at relatively low levels. CicrRNAs, like lncRNAs, can function as microRNA sponges and are involved in the regulation of PCGs and neuronal development, and the mechanisms of carcinogenesis [143]. In addition, circRNAs bind to target gene promoters, recruit transcription factors and modulate the expression of PCGs [142]. The peptides encoded by circRNAs are of greatest interest in humans due to the discovery of their involvement in the development of malignant neoplasms, since they can be used as targets for antitumor therapy (Table **5**) [142-145]. However, one of the first human circRNAs capable of translation was circ-ZNF609, which is involved in the regulation of skeletal muscle differentiation. Unlike the protein product of the core PCGs (transcription factor, zinc finger protein) from which this circRNA is generated, the translation product of this ncRNA does not contain zinc finger domains [147]. Since TEs in evolution were the sources of the

Table 4. The role of pri-microRNAs as sources of functional microRNAs and peptides.

Precursor Name	miRNA Name/ Function [Reference]	Peptide Name (Length)/ Function [Reference]
pri-miR-147b	miRNA-147b targets the NDUFA4 mRNA with immune-dampening effects and simultaneously enhances RIG-I/MDA-5-mediated viral immunity [123]	MOCCI (83) regulation of antiviral immune responses, paralog of the NDUFA4 subunit of cytochrome C oxidase (Complex IV), replaces NDUFA4 in Complex IV during inflammation to lower mitochondrial membrane potential, leading to cyto-protection and dampened immune response [123]
pri-miR-22 (lnc MIR22HG)	miRNA-22/ inhibits osteosarcoma cell proliferation, migration and invasion through the regulation of Twist1 expression [132]	miPEP22 (57) antiviral reactions [122]
pri-miR-7-3 (lnc MIR7-3HG)	miRNA-7-3/ suppresses autophagy and apoptosis in Hela and lung carcinoma cell lines through the downregulation of autophagy and Beclin 1 regulator 1 [133]	miPEP7-3 (125) protects pancreatic beta cells from DEX-induced dysfunction by activating the PI3K/AKT pathway [133]
pri-miR-155 (lnc MIR155HG)	miRNA-155/ antitumor activity of lymphocytes and myeloid cells [134]	miPEP155 (17) interacts with the adenosine 5'-triphosphate binding domain of heat shock-related protein 70 (HSC70), a chaperone required for antigen transport and presentation in dendritic cells [31]
pri-miR-200a	miRNA-200a/ tumor supressor [30]	miPEP-200a (187) regulates the expression of vimentin, and inhibits the epithelial-mesenchymal transition of prostate cancer cells, suppressing their migration [30]
pri-miR-200b	miRNA-200b/ tumor supressor [30]	miPEP-200b (54) regulates the expression of vimentin, and inhibits the epithelial-mesenchymal transition of prostate cancer cells, suppressing their migration [30]
pri-miR-31	miRNA-31/ negative regulator of regulatory T lymphocytes [119]	miPEP-31 (48) promotes the induction of regulatory T lymphocytes, inhibiting the expression of miRNA-31 [119]
pri-miR-34a	miRNA-34a/ tumor suppressor inhibits the expression of 700 target genes [7]	miPEP-133 (133) in mitochondria interacts with mitochondrial heat shock protein, prevents the interaction of HSPA9 with proteins, enhances the transcriptional activity of p53, and stimulates miR-34a expression [7]
pri-miR-497 (lnc MIR497HG)	miRNA-497/ led to downregulation of genes encoding cell cycle activators, as expected, such as CDC25A, CDK6, and Cyclin E [32]	miPEP497 (21) promotes the proliferation of tumor cells, and inhibits cervical cancer [32]

Table 5. Circular RNAs and peptides formed during their translation.

circRNA Name	Peptide Expressed Tissue/ Reference	Peptide Name/Length	Peptide Function
circ-AKT3	Glioblastoma/ [153]	AKT-174aa/174	Decreased the cell proliferation and radiation resistance
circCAPG	Triple-negative breast cancer/ [154]	CAPG-171aa/171	Activates MEKK2-MEK1/2-ERK1/2 pathway
circ-SHPRH	Glioblastoma/ [155]	SHPRH-146aa/146	Protects full-length circ-SHPRH from degradation by the ubiquitin proteasome, ubiquitinates the nuclear antigen of proliferating cells PCNA, inhibits cancer
circ-FBXW7	Glioblastoma/ [143]	FBXW7-185aa/185	Inhibits proliferation and acceleration of the cell cycle, suppresses cancer development
circ-FNDC3B	Colon cancer cells/ [145]	circ-FNDC3B-218aa/218	Inhibits Snail expression, promotes the tumor suppressor effect of FBP1, suppresses proliferation, invasion and migration of cancer cells
circ-Gprc5a	Bladder cancer/ [142]	Circ-Gprc5a/11	In combination with the membrane protein, Gprc5a activates GPCR signaling, promotes the development of cancer
circ-PPP1R12A	Colon cancer/ [144]	circ-PPP1R12A -73aa/73	Stimulates proliferation, migration and invasion of cancer cells by activating Hippo-YAP signaling pathways
circ-β-catenin	Liver cancer/ [156]	β-catenin-370aa/370	Activates the Wnt pathway by stabilizing full-length β-catenin, stimulates cancer development
circZKSCAN1	Hepatocellular carcinoma/ [157]	circZKSaa	Promoted the interaction of FBXW7 with the mammalian target of rapamycin (mTOR) to promote the ubiquitination of mTOR, thereby inhibiting the PI3K/AKT/mTOR pathway, cancer suppressor

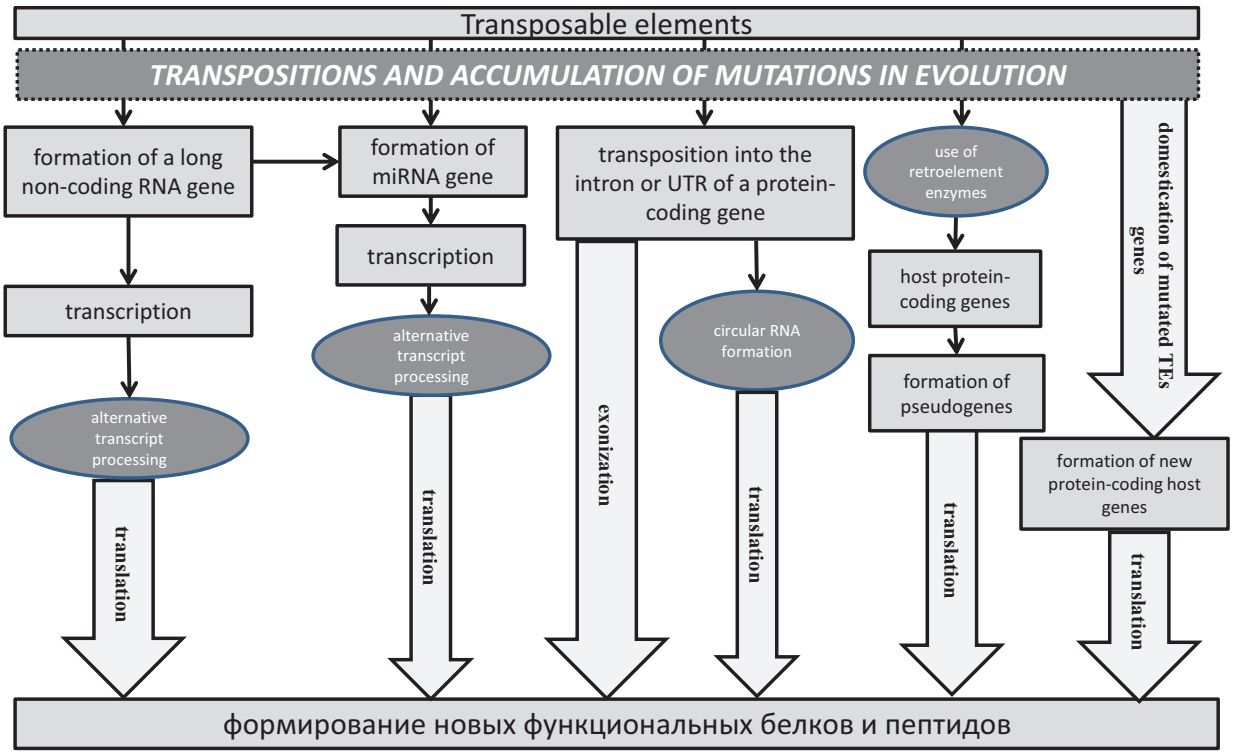

Fig. (6). Scheme of various ways transposable elements participate in the origin of proteins and peptides in eukaryotic evolution (transpositions and accumulation of mutations in evolution are highlighted with a dotted line to indicate the duration of the process in evolution). (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

emergence of spliceosomal introns, as well as their variability [35], it is logical to assume that the ability of circRNAs to be translated into functional peptides also reflects the universal properties of TEs and is due to the emergence of circRNAs from TEs. Indeed, circRNAs, like lncRNAs, are characterized by tissue-specific expression. Among all complementary sequences, human SINEs (especially Alu) have been shown to contribute the most to the formation of circRNAs and their diversity. Accordingly, different distributions of REs between species lead to an increase in the complexity of circRNA expression during the evolution of species [148]. This is due to the role of Alu as a sources of splicing acceptors and inhibitors of mRNA translation. Alu are the primary targets of the RNA editing enzyme ADAR, and the formation of Alu exons is suppressed by the ribonucleoprotein HNRNPC. The nuclear RNA helicase DHX9 was found to specifically bind to Alu inverted repeats within PCGs mRNA. An increase in the number of Alu-containing genes that form circRNAs occurs with the loss of DHX9, which acts as a nuclear RNA resolvase that neutralizes the threat of transcriptional and translational disruptions caused by Alu insertions [149]. It has been proposed that circRNAs of 100–1000 nucleotides in length, produced by small self-cleaving RNAs or called hammer-type ribozymes, are encoded by non-autonomous RE retrozymes [150]. This new family of REs appears to be widespread in eukaryotic genomes. The resulting expression of these REs containing hammer motifs can be easily detected by RNA blotting [151]. It is possible that the wide distribution of Alu in the genomes of eukaryotes (occupying 11% of the human genome) is due to the ability to form circRNAs involved in the regulation of gene expression during the adaptive evolution of organisms. Indeed, the formation of circRNAs requires the pairing of RNA nucleotides between sequences flanking back splicing sites due to the presence of Alus in gene introns [152].

It should be noted that viruses also encode functional circRNAs, which indicates the role of TEs in the emergence and evolution of circRNAs (due to the evolutionary relationship of viruses and transposable elements [77]). For example, EBV-encoded circBART2.2 was found to be highly expressed in nasopharyngeal carcinoma where it upregulated PD-L1 expression and inhibited T-cell function by binding the helicase domain of RIG-I and activating transcription factors IRF3 and NF-κB, resulting in tumor immune escape. Therefore, circBART2.2 may become a promising target for the treatment of nasopharyngeal carcinoma [158]. The study of circRNAs of other oncogenic viruses may become the basis for the development of new approaches to the treatment of malignant tumors. Such data are especially relevant for the development of drug-resistant cancer treatment methods. Multiple circRNAs encoded by the KSHV genome are expressed in KSHV-infected endothelial cells and primary effusion lymphoma cells. The KSHV circRNAs are located within ORFs of viral lytic genes, are up-regulated upon the induction of the lytic cycle, and alter cell growth. There were revealed new host-virus interactions of circRNAs: human antiviral circRNAs are activated in response to KSHV infection, and viral circRNA expression is induced in the lyt-

ic phase of infection [159]. Oncogenic human papillomaviruses generate circRNAs, some of which encompass the E7 oncogene (circE7) that is translated to produce E7 oncoprotein. Specific disruption of circE7 in CaSki cervical carcinoma cells reduces E7 protein levels and inhibits cancer cell growth both *in vitro* and in tumor xenografts [160]. Analysis carried out in 2021 identified 28,754 circRNAs encoded MERS-CoV, 720 circRNAs encoded SARS-CoV1 and 3437 circRNAs encoded SARS-CoV2. Viral circRNAs exhibit low expression levels, but they up-regulated genes related to mRNA splicing and processing in the early stage of viral infection, and regulated genes involved in diverse functions including cancer, metabolism, autophagy, viral infection in the late stage of viral infection [161]. SARS-CoV-2 infection induces epigenetic changes in the LTR69 subfamily of endogenous retroviruses [162]. Since TEs and exogenous viruses are evolutionarily interrelated [77], it can be assumed that the effect of such current viral infections as COVID-19 on the human body may be due to the effect of virally encoded ncRNAs on TEs. Identifying such relationships would make it possible to find new ways to treat viral infections using non-coding RNAs of viral origin as targets.

CONCLUSION

This article presents the role of TEs as sources of lncRNA, microRNA, circRNA and protein-coding genes in human genome evolution (Fig. **6**). Since TEs have been universal opposing units of life since the emergence of life on Earth, the multifunctionality of their transcripts became the basis for the emergence of complex systems for the functioning of the human genome. TEs inserted into genomes pose a potential danger to the body; therefore, in the course of evolution, protective systems have been developed in cells aimed at TE silencing at the transcriptional and post-transcriptional levels. As a result, epigenetic systems of genome regulation arose, which, in parallel with the fight against TEs, made it possible to complicate the functioning and significantly diversify the life forms of eukaryotes. Since many ncRNA genes and introns of protein-coding genes are derived from TEs, their transcripts, in addition to splicing, are processed to form various RNA molecules capable of translation to form peptides and proteins. This review presents such peptides and proteins that may be new therapeutic targets for the treatment of cancer and viral infections since they are involved in carcinogenesis and antiviral response. The targets of such peptides may also be TEs, which must be taken into account when developing treatment methods to avoid possible side effects associated with exposure to TEs, which also play a role in carcinogenesis and antiviral response. This is especially relevant for the development of therapies for drug-resistant cancer. An analysis of the scientific literature showed that among the ncRNAs translated into peptides, MALAT1, CRNDE, LINC00665, MAGI2-AS3, HOXB-AS3, LINC00675, LINC00467, LINC-PINT have an effect on drug resistance. It can be assumed that peptides can be used as tools to influence such lncRNAs in overcoming drug resistance. The most promising targets in such methods are multifunctional lncRNAs, such as MALAT1, which are involved in drug resistance of various cancers through interaction with various microRNAs.

AUTHOR'S CONTRIBUTIONS

The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

LIST OF ABBREVIATIONS

cDNA	=	complementary DNA
circRNAs	=	circular RNAs
EMT	=	Epithelial-to-mesenchymal Transition
lncRNAs	=	Long Non-coding RNAs
MALAT1	=	Metastasis-associated Lung Adenocarcinoma Transcript 1
MAPK	=	Mitogen-activated Protein Kinase
ncRNA	=	non-coding RNA
ORF	=	Open Reading Frame
PCGs	=	Protein-coding Genes
RBPs	=	RNA-binding Proteins
RNPs	=	Ribonucleoproteins
TEs	=	Transposable Elements
UTR	=	Untranslated Region

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Nurk, S.; Koren, S.; Rhie, A.; Rautiainen, M.; Bzikadze, A.V.; Mikheenko, A.; Vollger, M.R.; Altemose, N.; Uralsky, L.; Gershman, A.; Aganezov, S.; Hoyt, S.J.; Diekhans, M.; Logsdon, G.A.; Alonge, M.; Antonarakis, S.E.; Borchers, M.; Bouffard, G.G.; Brooks, S.Y.; Caldas, G.V.; Chen, N.C.; Cheng, H.; Chin, C.S.; Chow, W.; de Lima, L.G.; Dishuck, P.C.; Durbin, R.; Dvorkina, T.; Fiddes, I.T.; Formenti, G.; Fulton, R.S.; Fungtammasan, A.; Garrison, E.; Grady, P.G.S.; Graves-Lindsay, T.A.; Hall, I.M.; Hansen, N.F.; Hartley, G.A.; Haukness, M.; Howe, K.; Hunkapiller, M.W.; Jain, C.; Jain, M.; Jarvis, E.D.; Kerpedjiev, P.; Kirsche, M.; Kolmogorov, M.; Korlach, J.; Kremitzki, M.; Li, H.; Maduro, V.V.; Marschall, T.; McCartney, A.M.; McDaniel, J.; Miller, D.E.; Mullikin, J.C.; Myers, E.W.; Olson, N.D.; Paten, B.; Peluso, P.; Pevzner, P.A.; Porubsky, D.; Potapova, T.; Rogaev, E.I.; Rosenfeld, J.A.; Salzberg, S.L.; Schneider, V.A.; Sedlazeck,

F.J.; Shafin, K.; Shew, C.J.; Shumate, A.; Sims, Y.; Smit, A.F.A.; Soto, D.C.; Sović, I.; Storer, J.M.; Streets, A.; Sullivan, B.A.; Thibaud-Nissen, F.; Torrance, J.; Wagner, J.; Walenz, B.P.; Wenger, A.; Wood, J.M.D.; Xiao, C.; Yan, S.M.; Young, A.C.; Zarate, S.; Surti, U.; McCoy, R.C.; Dennis, M.Y.; Alexandrov, I.A.; Gerton, J.L.; O'Neill, R.J.; Timp, W.; Zook, J.M.; Schatz, M.C.; Eichler, E.E.; Miga, K.H.; Phillippy, A.M. The complete sequence of a human genome. *Science,* **2022**, *376*(6588), 44-53. http://dx.doi.org/10.1126/science.abj6987 PMID: 35357919

[2] Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G.; Lagarde, J.; Veeravalli, L.; Ruan, X.; Ruan, Y.; Lassmann, T.; Carninci, P.; Brown, J.B.; Lipovich, L.; Gonzalez, J.M.; Thomas, M.; Davis, C.A.; Shiekhattar, R.; Gingeras, T.R.; Hubbard, T.J.; Notredame, C.; Harrow, J.; Guigó, R. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.,* **2012**, *22*(9), 1775-1789. http://dx.doi.org/10.1101/gr.132159.111 PMID: 22955988

[3] Makałowski, W.; Gotea, V.; Pande, A.; Makałowska, I. Transposable elements: Classification, identification, and their use as a tool for comparative genomics. *Methods Mol. Biol.,* **2019**, *1910*, 177-207. http://dx.doi.org/10.1007/978-1-4939-9074-0_6 PMID: 31278665

[4] de Koning, A.P.J.; Gu, W.; Castoe, T.A.; Batzer, M.A.; Pollock, D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.,* **2011**, *7*(12), e1002384. http://dx.doi.org/10.1371/journal.pgen.1002384 PMID: 22144907

[5] Mustafin, RN Functional dualism of transcripts of transposons in the evolution of eukaryotic genomes. *Russ. J. Dev. Biol.,* **2018**, *49*(6), 339-355. http://dx.doi.org/10.1134/S1062360418070019

[6] Mustafin, R.N.; Khusnutdinova, E.K. The role of reverse transcriptase in the origin of life. *Biochemistry (Mosc.),* **2019**, *84*(8), 870-883. http://dx.doi.org/10.1134/S0006297919080030 PMID: 31522669

[7] Kang, M.; Tang, B.; Li, J.; Zhou, Z.; Liu, K.; Wang, R.; Jiang, Z.; Bi, F.; Patrick, D.; Kim, D.; Mitra, A.K.; Yang-Hartwich, Y. Identification of miPEP133 as a novel tumor-suppressor microprotein encoded by miR-34a pri-miRNA. *Mol. Cancer,* **2020**, *19*(1), 143. http://dx.doi.org/10.1186/s12943-020-01248-9 PMID: 32928232

[8] Fitzgerald, K.A.; Caffrey, D.R. Long noncoding RNAs in innate and adaptive immunity. *Curr. Opin. Immunol.,* **2014**, *26*, 140-146. http://dx.doi.org/10.1016/j.coi.2013.12.001 PMID: 24556411

[9] Fico, A.; Fiorenzano, A.; Pascale, E.; Patriarca, E.J.; Minchiotti, G. Long non-coding RNA in stem cell pluripotency and lineage commitment: Functions and evolutionary conservation. *Cell. Mol. Life Sci.,* **2019**, *76*(8), 1459-1471. http://dx.doi.org/10.1007/s00018-018-3000-z PMID: 30607432

[10] Long, Y.; Wang, X.; Youmans, D.T.; Cech, T.R. How do lncRNAs regulate transcription? *Sci. Adv.,* **2017**, *3*(9), eaao2110. http://dx.doi.org/10.1126/sciadv.aao2110 PMID: 28959731

[11] Mustafin, R.N. The Relationship between transposons and transcription factors in the evolution of eukaryotes. *J. Evol. Biochem. Physiol.,* **2019**, *55*(1), 14-23. http://dx.doi.org/10.1134/S0022093019010022

[12] Kapusta, A.; Kronenberg, Z.; Lynch, V.J.; Zhuo, X.; Ramsay, L.; Bourque, G.; Yandell, M.; Feschotte, C. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet.,* **2013**, *9*(4), e1003470. http://dx.doi.org/10.1371/journal.pgen.1003470 PMID: 23637635

[13] Johnson, R.; Guigó, R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA,* **2014**, *20*(7), 959-976. http://dx.doi.org/10.1261/rna.044560.114 PMID: 24850885

[14] Hadjiargyrou, M.; Delihas, N. The intertwining of transposable elements and non-coding RNAs. *Int. J. Mol. Sci.,* **2013**, *14*(7), 13307-13328. http://dx.doi.org/10.3390/ijms140713307 PMID: 23803660

[15] Gerdes, P.; Richardson, S.R.; Mager, D.L.; Faulkner, G.J. Transposable elements in the mammalian embryo: Pioneers surviving through stealth and service. *Genome Biol.,* **2016**, *17*(1), 100-116. http://dx.doi.org/10.1186/s13059-016-0965-5 PMID: 27161170

[16] Lu, X.; Sachs, F.; Ramsay, L.; Jacques, P.É.; Göke, J.; Bourque, G.; Ng, H.H. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nat. Struct. Mol. Biol.,* **2014**, *21*(4), 423-425. http://dx.doi.org/10.1038/nsmb.2799 PMID: 24681886

[17] Honson, D.D.; Macfarlan, T.S. A lncRNA-like Role for LINE1s in Development. *Dev. Cell,* **2018**, *46*(2), 132-134. http://dx.doi.org/10.1016/j.devcel.2018.06.022 PMID: 30016617

[18] Ramsay, L.; Marchetto, M.C.; Caron, M.; Chen, S.H.; Busche, S.; Kwan, T.; Pastinen, T.; Gage, F.H.; Bourque, G. Conserved expression of transposon-derived non-coding transcripts in primate stem cells. *BMC Genomics,* **2017**, *18*(1), 214-226. http://dx.doi.org/10.1186/s12864-017-3568-y PMID: 28245871

[19] Arendt, T.; Ueberham, U.; Janitz, M. Non-coding transcriptome in brain aging. *Aging (Albany NY),* **2017**, *9*(9), 1943-1944. http://dx.doi.org/10.18632/aging.101290 PMID: 28898200

[20] Lapp, H.E.; Hunter, R.G. The dynamic genome: Transposons and environmental adaptation in the nervous system. *Epigenomics,* **2016**, *8*(2), 237-249. http://dx.doi.org/10.2217/epi.15.107 PMID: 26791965

[21] Anderson, D.M.; Anderson, K.M.; Chang, C.L.; Makarewich, C.A.; Nelson, B.R.; McAnally, J.R.; Kasaragod, P.; Shelton, J.M.; Liou, J.; Bassel-Duby, R.; Olson, E.N. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell,* **2015**, *160*(4), 595-606. http://dx.doi.org/10.1016/j.cell.2015.01.009 PMID: 25640239

[22] Nelson, B.R.; Makarewich, C.A.; Anderson, D.M.; Winders, B.R.; Troupes, C.D.; Wu, F.; Reese, A.L.; McAnally, J.R.; Chen, X.; Kavalali, E.T.; Cannon, S.C.; Houser, S.R.; Bassel-Duby, R.; Olson, E.N. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science,* **2016**, *351*(6270), 271-275. http://dx.doi.org/10.1126/science.aad4076 PMID: 26816378

[23] Zhang, J.; Mujahid, H.; Hou, Y.; Nallamilli, BR.; Peng, Z. Plant Long ncRNAs: A new frontier for gene regulatory control. *Am. J. Plant Sci.,* **2013**, *4*(5), 1038-1045. http://dx.doi.org/10.4236/ajps.2013.45128

[24] Levine, M.T.; Jones, C.D.; Kern, A.D.; Lindfors, H.A.; Begun, D.J. Novel genes derived from noncoding DNA in *Drosophila melanogaster* are frequently X-linked and exhibit testis-biased expression. *Proc. Natl. Acad. Sci. USA,* **2006**, *103*(26), 9935-9939. http://dx.doi.org/10.1073/pnas.0509809103 PMID: 16777968

[25] Cai, J.; Zhao, R.; Jiang, H.; Wang, W. De novo origination of a new protein-coding gene in *Saccharomyces cerevisiae* . *Genetics,* **2008**, *179*(1), 487-496. http://dx.doi.org/10.1534/genetics.107.084491 PMID: 18493065

[26] Xie, C.; Zhang, Y.E.; Chen, J.Y.; Liu, C.J.; Zhou, W.Z.; Li, Y.; Zhang, M.; Zhang, R.; Wei, L.; Li, C.Y. Hominoid-specific de novo protein-coding genes originating from long non-coding RNAs. *PLoS Genet.,* **2012**, *8*(9), e1002942. http://dx.doi.org/10.1371/journal.pgen.1002942 PMID: 23028352

[27] Ruiz-Orera, J.; Messeguer, X.; Subirana, J.A.; Alba, M.M. Long non-coding RNAs as a source of new peptides. *eLife,* **2014**, *3*, e03523. http://dx.doi.org/10.7554/eLife.03523 PMID: 25233276

[28] Guo, L.; Zhao, Y.; Yang, S.; Zhang, H.; Wu, Q.; Chen, F. An integrated evolutionary analysis of miRNA–lncRNA in mammals. *Mol. Biol. Rep.,* **2014**, *41*(1), 201-207. http://dx.doi.org/10.1007/s11033-013-2852-4 PMID: 24186852

[29] Lauressergues, D.; Couzigou, J.M.; Clemente, H.S.; Martinez, Y.; Dunand, C.; Bécard, G.; Combier, J.P. Primary transcripts of microRNAs encode regulatory peptides. *Nature,* **2015**, *520*(7545), 90-93. http://dx.doi.org/10.1038/nature14346 PMID: 25807486

[30] Fang, J.; Morsalin, S.; Rao, V.; Reddy, E.S. Decoding of non-coding DNA and non-coding RNA: Pri-micro RNA-encoded novel peptides regulate migration of cancer cells. *J. Pharm. Sci. Pharmacol.,* **2017**, *3*(1), 23-27. http://dx.doi.org/10.1166/jpsp.2017.1070

[31] Niu, L.; Lou, F.; Sun, Y.; Sun, L.; Cai, X.; Liu, Z.; Zhou, H.; Wang, H.; Wang, Z.; Bai, J.; Yin, Q.; Zhang, J.; Chen, L.; Peng, D.; Xu, Z.; Gao, Y.; Tang, S.; Fan, L.; Wang, H. A micropeptide encoded by lncRNA MIR155HG suppresses autoimmune inflam-

mation *via* modulating antigen presentation. *Sci. Adv.,* **2020**, *6*(21), eaaz2059.
http://dx.doi.org/10.1126/sciadv.aaz2059 PMID: 32671205

[32] Prel, A.; Dozier, C.; Combier, J.P.; Plaza, S.; Besson, A. Evidence that regulation of Pri-miRNA/miRNA expression is not a general rule of miPEPs function in humans. *Int. J. Mol. Sci.,* **2021**, *22*(7), 3432.
http://dx.doi.org/10.3390/ijms22073432 PMID: 33810468

[33] Wei, G.; Qin, S.; Li, W.; Chen, L.; Ma, F. MDTE DB: A database for microRNAs derived from transposable element. *IEEE/ACM Trans Comput Biol Bioinform.,* **2016**, *13*(6), 1155-1160.
http://dx.doi.org/10.1109/TCBB.2015.2511767

[34] Schrader, L.; Schmitz, J. The impact of transposable elements in adaptive evolution. *Mol. Ecol.,* **2019**, *28*(6), 1537-1549.
http://dx.doi.org/10.1111/mec.14794 PMID: 30003608

[35] Kubiak, M.R.; Makałowska, I. Protein-coding genes' retrocopies and their functions. *Viruses,* **2017**, *9*(4), 80.
http://dx.doi.org/10.3390/v9040080 PMID: 28406439

[36] Abascal, F.; Tress, M.L.; Valencia, A. Alternative splicing and co-option of transposable elements: The case of TMPO/LAP2α and ZNF451 in mammals. *Bioinformatics,* **2015**, *31*(14), 2257-2261.
http://dx.doi.org/10.1093/bioinformatics/btv132 PMID: 25735770

[37] He, Z.; Chen, O.; Phillips, N.; Pasquesi, G.I.M.; Sabunciyan, S.; Florea, L. Predicting Alu exonization in the human genome with a deep learning model. *bioRxiv,* **2024**, 2024.01.03.574099.
http://dx.doi.org/10.1101/2024.01.03.574099.

[38] Tan, S.; Cardoso-Moreira, M.; Shi, W.; Zhang, D.; Huang, J.; Mao, Y.; Jia, H.; Zhang, Y.; Chen, C.; Shao, Y.; Leng, L.; Liu, Z.; Huang, X.; Long, M.; Zhang, Y.E. LTR-mediated retroposition as a mechanism of RNA-based duplication in metazoans. *Genome Res.,* **2016**, *26*(12), 1663-1675.
http://dx.doi.org/10.1101/gr.204925.116 PMID: 27934698

[39] Zhu, Z.; Tan, S.; Zhang, Y.; Zhang, Y.E. LINE-1-like retrotransposons contribute to RNA-based gene duplication in dicots. *Sci. Rep.,* **2016**, *6*(1), 24755.
http://dx.doi.org/10.1038/srep24755 PMID: 27098918

[40] Sakai, H.; Mizuno, H.; Kawahara, Y.; Wakimoto, H.; Ikawa, H.; Kawahigashi, H.; Kanamori, H.; Matsumoto, T.; Itoh, T.; Gaut, B.S. Retrogenes in rice (*Oryza sativa* L. ssp. *japonica*) exhibit correlated expression with their source genes. *Genome Biol. Evol.,* **2011**, *3*, 1357-1368.
http://dx.doi.org/10.1093/gbe/evr111 PMID: 22042334

[41] Grandi, F.C.; Rosser, J.M.; Newkirk, S.J.; Yin, J.; Jiang, X.; Xing, Z.; Whitmore, L.; Bashir, S.; Ivics, Z.; Izsvák, Z.; Ye, P.; Yu, Y.E.; An, W. Retrotransposition creates sloping shores: A graded influence of hypomethylated CpG islands on flanking CpG sites. *Genome Res.,* **2015**, *25*(8), 1135-1146.
http://dx.doi.org/10.1101/gr.185132.114 PMID: 25995269

[42] Du, Z.Q.; Yang, C.X.; Rothschild, M.F.; Ross, J.W. Novel microRNA families expanded in the human genome. *BMC Genomics,* **2013**, *14*(1), 98-105.
http://dx.doi.org/10.1186/1471-2164-14-98 PMID: 23402294

[43] Hoen, D.R.; Bureau, T.E. Discovery of novel genes derived from transposable elements using integrative genomic analysis. *Mol. Biol. Evol.,* **2015**, *32*(6), 1487-1506.
http://dx.doi.org/10.1093/molbev/msv042 PMID: 25713212

[44] Alzohairy, A.M.; Gyulai, G.; Jansen, R.K.; Bahieldin, A. Transposable elements domesticated and neofunctionalized by eukaryotic genomes. *Plasmid,* **2013**, *69*(1), 1-15.
http://dx.doi.org/10.1016/j.plasmid.2012.08.001 PMID: 22960324

[45] Kopera, H.C.; Moldovan, J.B.; Morrish, T.A.; Garcia-Perez, J.L.; Moran, J.V. Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase. *Proc. Natl. Acad. Sci. USA,* **2011**, *108*(51), 20345-20350.
http://dx.doi.org/10.1073/pnas.1100275108 PMID: 21940498

[46] Casacuberta, E. Drosophila: Retrotransposons making up telomeres. *Viruses,* **2017**, *9*(7), 192.
http://dx.doi.org/10.3390/v9070192 PMID: 28753967

[47] Kipling, D.; Warburton, P.E. Centromeres, CENP-B and tigger too. *Trends Genet.,* **1997**, *13*(4), 141-145.
http://dx.doi.org/10.1016/S0168-9525(97)01098-6 PMID: 9097724

[48] Zdobnov, E.M.; Campillos, M.; Harrington, E.D.; Torrents, D.; Bork, P. Protein coding potential of retroviruses and other transposable elements in vertebrate genomes. *Nucleic Acids Res.,* **2005**, *33*(3), 946-954.
http://dx.doi.org/10.1093/nar/gki236 PMID: 15716312

[49] Feschotte, C. Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.,* **2008**, *9*(5), 397-405.
http://dx.doi.org/10.1038/nrg2337 PMID: 18368054

[50] de Souza, F.S.J.; Franchini, L.F.; Rubinstein, M. Exaptation of transposable elements into novel cis-regulatory elements: Is the evidence always strong? *Mol. Biol. Evol.,* **2013**, *30*(6), 1239-1251.
http://dx.doi.org/10.1093/molbev/mst045 PMID: 23486611

[51] Couzigou, J.M.; André, O.; Guillotin, B.; Alexandre, M.; Combier, J.P. Use of micro RNA -encoded peptide mi PEP 172c to stimulate nodulation in soybean. *New Phytol.,* **2016**, *211*(2), 379-381.
http://dx.doi.org/10.1111/nph.13991 PMID: 27105382

[52] Lv, S.; Pan, L.; Wang, G. Commentary: Primary transcripts of micrornas encode regulatory peptides. *Front. Plant Sci.,* **2016**, *7*, 1436.
http://dx.doi.org/10.3389/fpls.2016.01436 PMID: 27713758

[53] Volff, J.N. Turning junk into gold: Domestication of transposable elements and the creation of new genes in eukaryotes. *BioEssays,* **2006**, *28*(9), 913-922.
http://dx.doi.org/10.1002/bies.20452 PMID: 16937363

[54] Kapitonov, V.V.; Jurka, J. Harbinger transposons and an ancient HARBI1 gene derived from a transposase. *DNA Cell Biol.,* **2004**, *23*(5), 311-324.
http://dx.doi.org/10.1089/104454904323090949 PMID: 15169610

[55] Abrusán, G.; Zhang, Y.; Szilágyi, A. Structure prediction and analysis of DNA transposon and LINE retrotransposon proteins. *J. Biol. Chem.,* **2013**, *288*(22), 16127-16138.
http://dx.doi.org/10.1074/jbc.M113.451500 PMID: 23530042

[56] Duan, C.G.; Wang, X.; Xie, S.; Pan, L.; Miki, D.; Tang, K.; Hsu, C.C.; Lei, M.; Zhong, Y.; Hou, Y.J.; Wang, Z.; Zhang, Z.; Mangrauthia, S.K.; Xu, H.; Zhang, H.; Dilkes, B.; Tao, W.A.; Zhu, J.K. A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation. *Cell Res.,* **2017**, *27*(2), 226-240.
http://dx.doi.org/10.1038/cr.2016.147 PMID: 27934869

[57] Sinzelle, L.; Izsvák, Z.; Ivics, Z. Molecular domestication of transposable elements: From detrimental parasites to useful host genes. *Cell. Mol. Life Sci.,* **2009**, *66*(6), 1073-1093.
http://dx.doi.org/10.1007/s00018-009-8376-3 PMID: 19132291

[58] Wang, J.; Vicente-García, C.; Seruggia, D.; Moltó, E.; Fernandez-Miñán, A.; Neto, A.; Lee, E.; Gómez-Skarmeta, J.L.; Montoliu, L.; Lunyak, V.V.; Jordan, I.K. MIR retrotransposon sequences provide insulators to the human genome. *Proc. Natl. Acad. Sci. USA,* **2015**, *112*(32), 4428-4437.
http://dx.doi.org/10.1073/pnas.1507253112

[59] Malfavon-Borja, R.; Feschotte, C. Fighting fire with fire: Endogenous retrovirus envelopes as restriction factors. *J. Virol.,* **2015**, *89*(8), 4047-4050.
http://dx.doi.org/10.1128/JVI.03653-14 PMID: 25653437

[60] Malik, H.S.; Henikoff, S.; Eickbush, T.H. Poised for contagion: Evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res.,* **2000**, *10*(9), 1307-1318.
http://dx.doi.org/10.1101/gr.145000 PMID: 10984449

[61] Mallet, F.; Bouton, O.; Prudhomme, S.; Cheynet, V.; Oriol, G.; Bonnaud, B.; Lucotte, G.; Duret, L.; Mandrand, B. The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology. *Proc. Natl. Acad. Sci. USA,* **2004**, *101*(6), 1731-1736.
http://dx.doi.org/10.1073/pnas.0305763101 PMID: 14757826

[62] Dupressoir, A.; Marceau, G.; Vernochet, C.; Bénit, L.; Kanellopoulos, C.; Sapin, V.; Heidmann, T. Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc. Natl. Acad. Sci. USA,* **2005**, *102*(3), 725-730.
http://dx.doi.org/10.1073/pnas.0406509102 PMID: 15644441

[63] Cornelis, G.; Vernochet, C.; Malicorne, S.; Souquere, S.; Tzika, A.C.; Goodman, S.M.; Catzeflis, F.; Robinson, T.J.; Milinkovitch, M.C.; Pierron, G.; Heidmann, O.; Dupressoir, A.; Heidmann, T. Retroviral envelope *syncytin* capture in an ancestrally diverged mammalian clade for placentation in the primitive Afrotherian ten-

recs. *Proc. Natl. Acad. Sci. USA,* **2014**, *111*(41), E4332-E4341. http://dx.doi.org/10.1073/pnas.1412268111 PMID: 25267646

[64] Heidmann, O.; Vernochet, C.; Dupressoir, A.; Heidmann, T. Identification of an endogenous retroviral envelope gene with fusogenic activity and placenta-specific expression in the rabbit: A new "syncytin" in a third order of mammals. *Retrovirology,* **2009**, *6*(1), 107. http://dx.doi.org/10.1186/1742-4690-6-107 PMID: 19943933

[65] Joly-Lopez, Z.; Bureau, T.E. Exaptation of transposable element coding sequences. *Curr. Opin. Genet. Dev.,* **2018**, *49*, 34-42. http://dx.doi.org/10.1016/j.gde.2018.02.011 PMID: 29525543

[66] Bernard, D.; Méhul, B.; Thomas-Collignon, A.; Delattre, C.; Donovan, M.; Schmidt, R. Identification and characterization of a novel retroviral-like aspartic protease specifically expressed in human epidermis. *J. Invest. Dermatol.,* **2005**, *125*(2), 278-287. http://dx.doi.org/10.1111/j.0022-202X.2005.23816.x PMID: 16098038

[67] Toll-Riera, M.; Bosch, N.; Bellora, N.; Castelo, R.; Armengol, L.; Estivill, X.; Mar Alba, M. Origin of primate orphan genes: A comparative genomics approach. *Mol. Biol. Evol.,* **2008**, *26*(3), 603-612. http://dx.doi.org/10.1093/molbev/msn281 PMID: 19064677

[68] Lu, S.; Zhang, J.; Lian, X.; Sun, L.; Meng, K.; Chen, Y.; Sun, Z.; Yin, X.; Li, Y.; Zhao, J.; Wang, T.; Zhang, G.; He, Q.Y. A hidden human proteome encoded by 'non-coding' genes. *Nucleic Acids Res.,* **2019**, *47*(15), 8111-8125. http://dx.doi.org/10.1093/nar/gkz646 PMID: 31340039

[69] van Heesch, S.; Witte, F.; Schneider-Lunitz, V.; Schulz, J.F.; Adami, E.; Faber, A.B.; Kirchner, M.; Maatz, H.; Blachut, S.; Sandmann, C.L.; Kanda, M.; Worth, C.L.; Schafer, S.; Calviello, L.; Merriott, R.; Patone, G.; Hummel, O.; Wyler, E.; Obermayer, B.; Mücke, M.B.; Lindberg, E.L.; Trnka, F.; Memczak, S.; Schilling, M.; Felkin, L.E.; Barton, P.J.R.; Quaife, N.M.; Vanezis, K.; Diecke, S.; Mukai, M.; Mah, N.; Oh, S.J.; Kurtz, A.; Schramm, C.; Schwinge, D.; Sebode, M.; Harakalova, M.; Asselbergs, F.W.; Vink, A.; de Weger, R.A.; Viswanathan, S.; Widjaja, A.A.; Gärtner-Rommel, A.; Milting, H.; dos Remedios, C.; Knosalla, C.; Mertins, P.; Landthaler, M.; Vingron, M.; Linke, W.A.; Seidman, J.G.; Seidman, C.E.; Rajewsky, N.; Ohler, U.; Cook, S.A.; Hubner, N. The translational landscape of the human heart. *Cell,* **2019**, *178*(1), 242-260.e29. http://dx.doi.org/10.1016/j.cell.2019.05.010 PMID: 31155234

[70] Li, X.L.; Pongor, L.; Tang, W.; Das, S.; Muys, B.R.; Jones, M.F.; Lazar, S.B.; Dangelmaier, E.A.; Hartford, C.C.R.; Grammatikakis, I.; Hao, Q.; Sun, Q.; Schetter, A.; Martindale, J.L.; Tang, B.; Jenkins, L.M.; Robles, A.I.; Walker, R.L.; Ambs, S.; Chari, R.; Shabalina, S.A.; Gorospe, M.; Hussain, S.P.; Harris, C.C.; Meltzer, P.S.; Prasanth, K.V.; Aladjem, M.I.; Andresson, T.; Lal, A. A small protein encoded by a putative lncRNA regulates apoptosis and tumorigenicity in human colorectal cancer cells. *eLife,* **2020**, *9*, e53734. http://dx.doi.org/10.7554/eLife.53734 PMID: 33112233

[71] Ge, Q.; Jia, D.; Cen, D.; Qi, Y.; Shi, C.; Li, J.; Sang, L.; Yang, L.; He, J.; Lin, A.; Chen, S.; Wang, L. Micropeptide ASAP encoded by LINC00467 promotes colorectal cancer progression by directly modulating ATP synthase activity. *J. Clin. Invest.,* **2021**, *131*(22), e152911. http://dx.doi.org/10.1172/JCI152911 PMID: 34591791

[72] Wu, S.; Zhang, L.; Deng, J.; Guo, B.; Li, F.; Wang, Y.; Wu, R.; Zhang, S.; Lu, J.; Zhou, Y. A novel micropeptide encoded by Y-Linked LINC00278 links cigarette smoking and AR signaling in male esophageal squamous cell carcinoma. *Cancer Res.,* **2020**, *80*(13), 2790-2803. http://dx.doi.org/10.1158/0008-5472.CAN-19-3440 PMID: 32169859

[73] Barczak, W.; Carr, S.M.; Liu, G.; Munro, S.; Nicastri, A.; Lee, L.N.; Hutchings, C.; Ternette, N.; Klenerman, P.; Kanapin, A.; Samsonova, A.; La Thangue, N.B. Long non-coding RNA-derived peptides are immunogenic and drive a potent anti-tumour response. *Nat. Commun.,* **2023**, *14*(1), 1078. http://dx.doi.org/10.1038/s41467-023-36826-0 PMID: 36841868

[74] Xiao, W.; Halabi, R.; Lin, C.H.; Nazim, M.; Yeom, K.H.; Black, D.L. The lncRNA *Malat1* is trafficked to the cytoplasm as a localized mRNA encoding a small peptide in neurons. *Genes Dev.,* **2024**, *38*(7-8), 294-307. http://dx.doi.org/10.1101/gad.351557.124 PMID: 38688681

[75] Zapata, J.C.; Campilongo, F.; Barclay, R.A.; DeMarino, C.; Iglesias-Ussel, M.D.; Kashanchi, F.; Romerio, F. The human immunodeficiency virus 1 ASP RNA promotes viral latency by recruiting the Polycomb Repressor Complex 2 and promoting nucleosome assembly. *Virology,* **2017**, *506*, 34-44. http://dx.doi.org/10.1016/j.virol.2017.03.002 PMID: 28340355

[76] Bidet, K.; Dadlani, D.; Garcia-Blanco, M.A. G3BP1, G3BP2 and CAPRIN1 are required for translation of interferon stimulated mRNAs and are targeted by a dengue virus non-coding RNA. *PLoS Pathog.,* **2014**, *10*(7), e1004242. http://dx.doi.org/10.1371/journal.ppat.1004242 PMID: 24992036

[77] Mustafin, R.N. The hypothesis of the origin of viruses from transposons. *Mol. Gen. Microbiol. Virol.,* **2018**, *33*(4), 223-232. http://dx.doi.org/10.3103/S0891416818040067

[78] Vachon, V.K.; Conn, G.L. Adenovirus VA RNA: An essential pro-viral non-coding RNA. *Virus Res.,* **2016**, *212*, 39-52. http://dx.doi.org/10.1016/j.virusres.2015.06.018 PMID: 26116898

[79] Cao, S.; Moss, W.; O'Grady, T.; Concha, M.; Strong, M.J.; Wang, X.; Yu, Y.; Baddoo, M.; Zhang, K.; Fewell, C.; Lin, Z.; Dong, Y.; Flemington, E.K. New noncoding lytic transcripts derived from the Epstein-Barr virus latency origin of replication, *oriP*, are hyperedited, bind the paraspeckle protein, NONO/p54nrb, and Support viral lytic transcription. *J. Virol.,* **2015**, *89*(14), 7120-7132. http://dx.doi.org/10.1128/JVI.00608-15 PMID: 25926645

[80] Szafron, L.M.; Balcerak, A.; Grzybowska, E.A.; Pienkowska--Grela, B.; Felisiak-Golabek, A.; Podgorska, A. The novel gene journal pre-proof 25 CRNDE encodes a nuclear peptide (CRNDE-P) which is overexpressed in highly proliferating tissues. *PLoS One,* **2015**, *10*(5), e0127475. http://dx.doi.org/10.1371/journal.pone.0127475 PMID: 25978564

[81] D'Lima, N.G.; Ma, J.; Winkler, L.; Chu, Q.; Loh, K.H.; Corpuz, E.O.; Budnik, B.A.; Lykke-Andersen, J.; Saghatelian, A.; Slavoff, S.A. A human microprotein that interacts with the mRNA decapping complex. *Nat. Chem. Biol.,* **2017**, *13*(2), 174-180. http://dx.doi.org/10.1038/nchembio.2249 PMID: 27918561

[82] Guo, B.; Wu, S.; Zhu, X.; Zhang, L.; Deng, J.; Li, F.; Wang, Y.; Zhang, S.; Wu, R.; Lu, J.; Zhou, Y. Micropeptide CIP 2A- BPencoded by LINC 00665 inhibits triple-negative breast cancer progression. *EMBO J.,* **2020**, *39*(1), e102190. http://dx.doi.org/10.15252/embj.2019102190 PMID: 31755573

[83] Wang, Y.; Wu, S.; Zhu, X.; Zhang, L.; Deng, J.; Li, F. LncRNA-encoded polypeptide ASRPS inhibits triple-negative breast cancer angiogenesis. *J Exp Med.,* **2020**, *217*(3), jem.20190950. http://dx.doi.org/10.1084/jem.20190950.

[84] Zhang, Z.; Yi, Y.; Wang, Z.; Zhang, H.; Zhao, Y.; He, R.; Luo, Y.; Cui, Z. LncRNA MAGI2-AS3-encoded polypeptide restrains the proliferation and migration of breast cancer cells. *Mol. Biotechnol.,* **2024**, *66*(6), 1409-1423. http://dx.doi.org/10.1007/s12033-023-00801-3 PMID: 37358745

[85] Polycarpou-Schwarz, M.; Groß, M.; Mestdagh, P.; Schott, J.; Grund, S.E.; Hildenbrand, C.; Rom, J.; Aulmann, S.; Sinn, H.P.; Vandesompele, J.; Diederichs, S. The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation. *Oncogene,* **2018**, *37*(34), 4750-4768. http://dx.doi.org/10.1038/s41388-018-0281-5 PMID: 29765154

[86] Huang, J.Z.; Chen, M.; Chen, D.; Gao, X.C.; Zhu, S.; Huang, H.; Hu, M.; Zhu, H.; Yan, G.R. A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. *Mol. Cell,* **2017**, *68*(1), 171-184.e6. http://dx.doi.org/10.1016/j.molcel.2017.09.015 PMID: 28985503

[87] Meng, N.; Chen, M.; Chen, D.; Chen, X.H.; Wang, J.Z.; Zhu, S.; He, Y.T.; Zhang, X.L.; Lu, R.X.; Yan, G.R. Small protein hidden in lncRNA LOC90024 promotes "cancerous" RNA splicing and tumorigenesis. *Adv. Sci. (Weinh.),* **2020**, *7*(10), 1903233. http://dx.doi.org/10.1002/advs.201903233 PMID: 32440474

[88] Zhu, S.; Wang, J.Z.; Chen, D.; He, Y.T.; Meng, N.; Chen, M.; Lu, R.X.; Chen, X.H.; Zhang, X.L.; Yan, G.R. An oncopeptide regulates m^6A recognition by the m^6A reader IGF2BP1 and tumorigenesis. *Nat. Commun.,* **2020**, *11*(1), 1685.

http://dx.doi.org/10.1038/s41467-020-15403-9 PMID: 32245947

[89] Zheng, W.; Guo, Y.; Zhang, G.; Bai, J.; Song, Y.; Song, X.; Zhu, Q.; Bao, X.; Wu, G.; Zhang, C. Peptide encoded by lncRNA BVES-AS1 promotes cell viability, migration, and invasion in colorectal cancer cells via the SRC/mTOR signaling pathway. *PLoS One,* **2023**, *18*(6), e0287133.
http://dx.doi.org/10.1371/journal.pone.0287133 PMID: 37347740

[90] Pang, Y.; Liu, Z.; Han, H.; Wang, B.; Li, W.; Mao, C.; Liu, S. Peptide SMIM30 promotes HCC development by inducing SRC/YES1 membrane anchoring and MAPK pathway activation. *J. Hepatol.,* **2020**, *73*(5), 1155-1169.
http://dx.doi.org/10.1016/j.jhep.2020.05.028 PMID: 32461121

[91] Xu, W.; Deng, B.; Lin, P.; Liu, C.; Li, B.; Huang, Q.; Zhou, H.; Yang, J.; Qu, L. Ribosome profiling analysis identified a KRAS-interacting microprotein that represses oncogenic signaling in hepatocellular carcinoma cells. *Sci. China Life Sci.,* **2020**, *63*(4), 529-542.
http://dx.doi.org/10.1007/s11427-019-9580-5 PMID: 31240521

[92] Zhang, M.; Zhao, K.; Xu, X.; Yang, Y.; Yan, S.; Wei, P.; Liu, H.; Xu, J.; Xiao, F.; Zhou, H.; Yang, X.; Huang, N.; Liu, J.; He, K.; Xie, K.; Zhang, G.; Huang, S.; Zhang, N. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat. Commun.,* **2018**, *9*(1), 4475.
http://dx.doi.org/10.1038/s41467-018-06862-2 PMID: 30367041

[93] Du, B.; Zhang, Z.; Jia, L.; Zhang, H.; Zhang, S.; Wang, H.; Cheng, Z. Micropeptide AF127577.4-ORF hidden in a lncRNA diminishes glioblastoma cell proliferation *via* the modulation of ERK2/METTL3 interaction. *Sci. Rep.,* **2024**, *14*(1), 12090.
http://dx.doi.org/10.1038/s41598-024-62710-y PMID: 38802444

[94] Song, H.; Wang, J.; Wang, X.; Yuan, B.; Li, D.; Hu, A.; Guo, Y.; Cai, S.; Jin, S.; Zhou, Y.; Li, Q.; Chen, G.; Gao, H.; Zheng, L.; Tong, Q. HNF4A-AS1-encoded small peptide promotes self-renewal and aggressiveness of neuroblastoma stem cells *via* eEF1A1-repressed SMAD4 transactivation. *Oncogene,* **2022**, *41*(17), 2505-2519.
http://dx.doi.org/10.1038/s41388-022-02271-4 PMID: 35318442

[95] Sun, L.; Wang, W.; Han, C.; Huang, W.; Sun, Y.; Fang, K.; Zeng, Z.; Yang, Q.; Pan, Q.; Chen, T.; Luo, X.; Chen, Y. The oncomicropeptide APPLE promotes hematopoietic malignancy by enhancing translation initiation. *Mol. Cell,* **2021**, *81*(21), 4493-4508.e9.
http://dx.doi.org/10.1016/j.molcel.2021.08.033 PMID: 34555354

[96] Kermi, C.; Lau, L.; Asadi Shahmirzadi, A.; Classon, M. Disrupting mechanisms that regulate genomic repeat elements to combat cancer and drug resistance. *Front. Cell Dev. Biol.,* **2022**, *10*, 826461.
http://dx.doi.org/10.3389/fcell.2022.826461 PMID: 35602594

[97] Mustafin, R.N.; Khusnutdinova, E.K. Non-coding parts of genomes as the basis of epigenetic heredity. *Vavilov Journal of Genetics and Breeding,* **2017**, *21*(6), 742-749.
http://dx.doi.org/10.18699/10.18699/VJ17.30-o

[98] Mombach, D.M.; Mercuri, R.L.V.; da Fontoura Gomes, T.M.F.; Galante, P.A.F.; Loreto, E.L.S. Transposable elements alter gene expression and may impact response to cisplatin therapy in ovarian cancer. *Carcinogenesis,* **2024**, bgae029.
http://dx.doi.org/10.1093/carcin/bgae029 PMID: 38722203

[99] Zhang, C.; Kang, T.; Wang, X.; Wang, J.; Liu, L.; Zhang, J.; Liu, X.; Li, R.; Wang, J.; Zhang, J. LINC-PINT suppresses cisplatin resistance in gastric cancer by inhibiting autophagy activation *via* epigenetic silencing of ATG5 by EZH2. *Front. Pharmacol.,* **2022**, *13*, 968223.
http://dx.doi.org/10.3389/fphar.2022.968223 PMID: 36091809

[100] Chen, J.; Zhu, M.; Zou, L.; Xia, J.; Huang, J.; Deng, Q.; Xu, R. Long non-coding RNA LINC-PINT attenuates paclitaxel resistance in triple-negative breast cancer cells *via* targeting the RNA-binding protein NONO. *Acta Biochim. Biophys. Sin. (Shanghai),* **2020**, *52*(8), 801-809.
http://dx.doi.org/10.1093/abbs/gmaa072 PMID: 32632453

[101] Yuan, Z.; Xiu, C.; Liu, D.; Zhou, G.; Yang, H.; Pei, R.; Ding, C.; Cui, X.; Sun, J.; Song, K. Long noncoding RNA LINC-PINT regulates laryngeal carcinoma cell stemness and chemoresistance through miR-425-5p/PTCH1/SHH axis. *J. Cell. Physiol.,* **2019**, *234*(12), 23111-23122.
http://dx.doi.org/10.1002/jcp.28874 PMID: 31131448

[102] Ling, L.; Wen, Y.; Xiong, Y.; Liu, X.; Chen, J.; Liu, T.; Zhang, B. Anisomycin inhibits the activity of human ovarian cancer stem cells *via* regulating antisense RNA NCBP2-AS2/MEK/ERK/STAT3 signaling. *J. Gene Med.,* **2024**, *26*(1), e3571.
http://dx.doi.org/10.1002/jgm.3571 PMID: 37483091

[103] Li, L.; Zhang, Y.; Zhan, Y.; Zhong, Y.; Li, X. LINC00467 mediates the 5-fluorouracil resistance in breast cancer cells. *In Vitro Cell. Dev. Biol. Anim.,* **2023**, *60*(1), 80-88.
http://dx.doi.org/10.1007/s11626-023-00832-9 PMID: 38127229

[104] Li, W.; He, Y.; Chen, W.; Man, W.; Fu, Q.; Tan, H.; Guo, H.; Zhou, J.; Yang, P. Knockdown of LINC00467 contributed to Axitinib sensitivity in hepatocellular carcinoma through miR-509-3p/PDGFRA axis. *Gene Ther.,* **2021**, *28*(10-11), 634-645.
http://dx.doi.org/10.1038/s41434-020-0137-9 PMID: 32221502

[105] Yao, M.; Shi, X.; Li, Y.; Xiao, Y.; Butler, W.; Huang, Y.; Du, L.; Wu, T.; Bian, X.; Shi, G.; Ye, D.; Fu, G.; Wang, J.; Ren, S. LINC00675 activates androgen receptor axis signaling pathway to promote castration-resistant prostate cancer progression. *Cell Death Dis.,* **2020**, *11*(8), 638.
http://dx.doi.org/10.1038/s41419-020-02856-5 PMID: 32801300

[106] Tseng, C.F.; Chen, L.T.; Wang, H.D.; Liu, Y.H.; Shiah, S.G. Transcriptional suppression of Dicer by HOXB-AS3/EZH2 complex dictates sorafenib resistance and cancer stemness. *Cancer Sci.,* **2022**, *113*(5), 1601-1612.
http://dx.doi.org/10.1111/cas.15319 PMID: 35253323

[107] Yang, G.; Li, T.; Liu, J.; Quan, Z.; Liu, M.; Guo, Y.; Wu, Y.; Ou, L.; Wu, X.; Zheng, Y. lncRNA MAGI2-AS3 suppresses castration-resistant prostate cancer proliferation and migration *via* the miR-106a-5p/RAB31 axis. *Genomics,* **2023**, *115*(2), 110599.
http://dx.doi.org/10.1016/j.ygeno.2023.110599 PMID: 36889366

[108] Lu, M.; Qin, X.; Zhou, Y.; Li, G.; Liu, Z.; Geng, X.; Yue, H. Long non-coding RNA LINC00665 promotes gemcitabine resistance of cholangiocarcinoma cells *via* regulating EMT and stemness properties through miR-424-5p/BCL9L axis. *Cell Death Dis.,* **2021**, *12*(1), 72.
http://dx.doi.org/10.1038/s41419-020-03346-4 PMID: 33436545

[109] Liu, X.; Lu, X.; Zhen, F.; Jin, S.; Yu, T.; Zhu, Q.; Wang, W.; Xu, K.; Yao, J.; Guo, R. LINC00665 induces acquired resistance to gefitinib through recruiting EZH2 and activating PI3K/AKT pathway in NSCLC. *Mol. Ther. Nucleic Acids,* **2019**, *16*, 155-161.
http://dx.doi.org/10.1016/j.omtn.2019.02.010 PMID: 30889481

[110] Wu, J.; Ni, X.; Yu, Z.; Wu, S.; Liu, Z. CRNDE inducing cisplatin resistance through SRSF1/TIA1 signaling pathway in ovarian cancer. *Pathol. Res. Pract.,* **2022**, *235*, 153957.
http://dx.doi.org/10.1016/j.prp.2022.153957 PMID: 35653925

[111] Chen, K.Y.; Zhu, S.G.; He, J.W.; Duan, X.P. LncRNA CRNDE is involved in radiation resistance in hepatocellular carcinoma *via* modulating the SP1/PDK1 axis. *Neoplasma,* **2022**, *69*(4), 918-930.
http://dx.doi.org/10.4149/neo_2022_211230N1853 PMID: 35652619

[112] Liu, P.; Li, X.; Cui, Y.; Chen, J.; Li, C.; Li, Q.; Li, H.; Zhang, X.; Zu, X. LncRNA-MALAT1 mediates cisplatin resistance *via* miR-101-3p/VEGF-C pathway in bladder cancer. *Acta Biochim. Biophys. Sin. (Shanghai),* **2019**, *51*(11), 1148-1157.
http://dx.doi.org/10.1093/abbs/gmz112 PMID: 31650173

[113] Zhang, Z.; Li, M.; Zhang, Z. lncRNA MALAT1 modulates oxaliplatin resistance of gastric cancer *via* sponging miR-22-3p. *OncoTargets Ther.,* **2020**, *13*, 1343-1354.
http://dx.doi.org/10.2147/OTT.S196619 PMID: 32104001

[114] Cao, Y.; Zhang, F.; Wang, H.; Bi, C.; Cui, J.; Liu, F.; Pan, H. LncRNA MALAT1 mediates doxorubicin resistance of hepatocellular carcinoma by regulating miR-3129-5p/Nova1 axis. *Mol. Cell. Biochem.,* **2021**, *476*(1), 279-292.
http://dx.doi.org/10.1007/s11010-020-03904-6 PMID: 32965597

[115] Shi, C.; Ren, S.; Zhao, X.; Li, Q. lncRNA *MALAT1* regulates the resistance of breast cancer cells to paclitaxel *via* the miR-497-5p/*SHOC2* axis. *Pharmacogenomics,* **2022**, *23*(18), 973-985.
http://dx.doi.org/10.2217/pgs-2022-0077 PMID: 36420706

[116] Chen, W.; Tan, X.; Yang, Q.; Fang, Z.; Xu, Y. MALAT1 enhances gemcitabine resistance in non-small cell lung cancer cells

by directly affecting miR-27a-5p/PBOV1 axis. *Cell. Signal.,* **2022**, *94*, 110326.
http://dx.doi.org/10.1016/j.cellsig.2022.110326 PMID: 35367362

[117] Yu, Z.; Tang, H.; Chen, S.; Xie, Y.; Shi, L.; Xia, S.; Jiang, M.; Li, J.; Chen, D. Exosomal LOC85009 inhibits docetaxel resistance in lung adenocarcinoma through regulating ATG5-induced autophagy. *Drug Resist. Updat.,* **2023**, *67*, 100915.
http://dx.doi.org/10.1016/j.drup.2022.100915 PMID: 36641841

[118] Zhu, C.; Xie, Y.; Li, Q.; Zhang, Z.; Chen, J.; Zhang, K.; Xia, X.; Yu, D.; Chen, D.; Yu, Z.; Chen, J. CPSF6-mediated XBP1 3'UTR shortening attenuates cisplatin-induced ER stress and elevates chemo-resistance in lung adenocarcinoma. *Drug Resist. Updat.,* **2023**, *68*, 100933.
http://dx.doi.org/10.1016/j.drup.2023.100933 PMID: 36821972

[119] Zhou, H.; Lou, F.; Bai, J.; Sun, Y.; Cai, W.; Sun, L.; Xu, Z.; Liu, Z.; Zhang, L.; Yin, Q.; Zhang, J.; Gao, Y.; Wang, Z.; Niu, L.; Cai, X.; Deng, S.; Wang, H.; Xia, L.; Ginhoux, F.; Li, Q.; Wang, H. A peptide encoded by pri-miRNA-31 represses autoimmunity by promoting T_{reg} differentiation. *EMBO Rep.,* **2022**, *23*(5), e53475.
http://dx.doi.org/10.15252/embr.202153475 PMID: 35343645

[120] Ormancey, M.; Thuleau, P.; Combier, J.P.; Plaza, S. The essentials on microRNA-encoded peptides from plants to animals. *Biomolecules,* **2023**, *13*(2), 206.
http://dx.doi.org/10.3390/biom13020206 PMID: 36830576

[121] Kumar, R.S.; Sinha, H.; Datta, T.; Asif, M.H.; Trivedi, P.K. microRNA408 and its encoded peptide regulate sulfur assimilation and arsenic stress response in Arabidopsis. *Plant Physiol.,* **2023**, *192*(2), 837-856.
http://dx.doi.org/10.1093/plphys/kiad033 PMID: 36682886

[122] Razooky, B.; Obermayer, B.; O'May, J.; Tarakhovsky, A. Viral infection identifies micropeptides differentially regulated in smORF-Containing lncRNAs. *Genes (Basel),* **2017**, *8*(8), 206.
http://dx.doi.org/10.3390/genes8080206 PMID: 28825667

[123] Lee, C.Q.E.; Kerouanton, B.; Chothani, S.; Zhang, S.; Chen, Y.; Mantri, C.K.; Hock, D.H.; Lim, R.; Nadkarni, R.; Huynh, V.T.; Lim, D.; Chew, W.L.; Zhong, F.L.; Stroud, D.A.; Schafer, S.; Tergaonkar, V.; St John, A.L.; Rackham, O.J.L.; Ho, L. Coding and non-coding roles of MOCCI (C15ORF48) coordinate to regulate host inflammation and immunity. *Nat. Commun.,* **2021**, *12*(1), 2130.
http://dx.doi.org/10.1038/s41467-021-22397-5 PMID: 33837217

[124] Ahmad, I.; Valverde, A.; Siddiqui, H.; Schaller, S.; Naqvi, A.R. Viral microRNAs: Interfering the interferon signaling. *Curr. Pharm. Des.,* **2020**, *26*(4), 446-454.
http://dx.doi.org/10.2174/1381612826666200109181238 PMID: 31924149

[125] De Cecco, M.; Ito, T.; Petrashen, A.P.; Elias, A.E.; Skvir, N.J.; Criscione, S.W.; Caligiana, A.; Brocculi, G.; Adney, E.M.; Boeke, J.D.; Le, O.; Beauséjour, C.; Ambati, J.; Ambati, K.; Simon, M.; Seluanov, A.; Gorbunova, V.; Slagboom, P.E.; Helfand, S.L.; Neretti, N.; Sedivy, J.M. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature,* **2019**, *566*(7742), 73-78.
http://dx.doi.org/10.1038/s41586-018-0784-9 PMID: 30728521

[126] Ender, C.; Krek, A.; Friedländer, M.R.; Beitzinger, M.; Weinmann, L.; Chen, W.; Pfeffer, S.; Rajewsky, N.; Meister, G. A human snoRNA with microRNA-like functions. *Mol. Cell,* **2008**, *32*(4), 519-528.
http://dx.doi.org/10.1016/j.molcel.2008.10.017 PMID: 19026782

[127] Jacob, M.D.; Audas, T.E.; Mullineux, S.T.; Lee, S. Where no RNA polymerase has gone before. *Nucleus,* **2012**, *3*(4), 315-319.
http://dx.doi.org/10.4161/nucl.20585 PMID: 22688644

[128] Li, Z.; Ender, C.; Meister, G.; Moore, P.S.; Chang, Y.; John, B. Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Res.,* **2012**, *40*(14), 6787-6799.
http://dx.doi.org/10.1093/nar/gks307 PMID: 22492706

[129] Kumar, P.; Anaya, J.; Mudunuri, S.B.; Dutta, A. Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol.,* **2014**, *12*(1), 78.
http://dx.doi.org/10.1186/s12915-014-0078-0 PMID: 25270025

[130] Venkatesh, T.; Suresh, P.S.; Tsutsumi, R. tRFs: miRNAs in disguise. *Gene,* **2016**, *579*(2), 133-138.
http://dx.doi.org/10.1016/j.gene.2015.12.058 PMID: 26743126

[131] Martinez, G.; Choudury, S.G.; Slotkin, R.K. tRNA-derived small RNAs target transposable element transcripts. *Nucleic Acids Res.,* **2017**, *45*(9), 5142-5152.
http://dx.doi.org/10.1093/nar/gkx103 PMID: 28335016

[132] Ruan, Q.; Wang, C.; Wu, Y.; Zhu, Q. Exosome microRNA-22 inhibiting proliferation, migration and invasion through regulating Twist1/CADM1 axis in osteosarcoma. *Sci. Rep.,* **2024**, *14*(1), 761.
http://dx.doi.org/10.1038/s41598-023-50612-4 PMID: 38191892

[133] Mao, X.; Zhou, J.; Kong, L.; Zhu, L.; Yang, D.; Zhang, Z. A peptide encoded by lncRNA MIR7-3 host gene (MIR7-3HG) alleviates dexamethasone-induced dysfunction in pancreatic β-cells through the PI3K/AKT signaling pathway. *Biochem. Biophys. Res. Commun.,* **2023**, *647*, 62-71.
http://dx.doi.org/10.1016/j.bbrc.2023.01.004 PMID: 36731335

[134] Michaille, J.J.; Awad, H.; Fortman, E.C.; Efanov, A.A.; Tili, E. *miR-155* expression in antitumor immunity: The higher the better? *Genes Chromosomes Cancer,* **2019**, *58*(4), 208-218.
http://dx.doi.org/10.1002/gcc.22698 PMID: 30382602

[135] Jiang, M.; Qi, F.; Zhang, K.; Zhang, X.; Ma, J.; Xia, S.; Chen, L.; Yu, Z.; Chen, J.; Chen, D. MARCKSL1–2 reverses docetaxel-resistance of lung adenocarcinoma cells by recruiting SUZ12 to suppress HDAC1 and elevate miR-200b. *Mol. Cancer,* **2022**, *21*(1), 150.
http://dx.doi.org/10.1186/s12943-022-01605-w PMID: 35864549

[136] Sukocheva, O.A.; Liu, J.; Neganova, M.E.; Beeraka, N.M.; Aleksandrova, Y.R.; Manogaran, P.; Grigorevskikh, E.M.; Chubarev, V.N.; Fan, R. Perspectives of using microRNA-loaded nanocarriers for epigenetic reprogramming of drug resistant colorectal cancers. *Semin. Cancer Biol.,* **2022**, *86*(Pt 2), 358-375.
http://dx.doi.org/10.1016/j.semcancer.2022.05.012 PMID: 35623562

[137] Yu, D.; Lv, M.; Chen, W.; Zhong, S.; Zhang, X.; Chen, L.; Ma, T.; Tang, J.; Zhao, J. Role of miR-155 in drug resistance of breast cancer. *Tumour Biol.,* **2015**, *36*(3), 1395-1401.
http://dx.doi.org/10.1007/s13277-015-3263-z PMID: 25744731

[138] Hsu, H.H.; Kuo, W.W.; Shih, H.N.; Cheng, S.F.; Yang, C.K.; Chen, M.C.; Tu, C.C.; Viswanadha, V.P.; Liao, P.H.; Huang, C.Y. FOXC1 regulation of miR-31-5p confers oxaliplatin resistance by targeting LATS2 in colorectal cancer. *Cancers (Basel),* **2019**, *11*(10), 1576.
http://dx.doi.org/10.3390/cancers11101576 PMID: 31623173

[139] Tian, Y.; Chen, Z.H.; Wu, P.; Zhang, D.; Ma, Y.; Liu, X.F.; Wang, X.; Ding, D.; Cao, X.C.; Yu, Y. MIR497HG-Derived miR-195 and miR-497 mediate tamoxifen resistance *via* PI3K/AKT signaling in breast cancer. *Adv. Sci. (Weinh.),* **2023**, *10*(12), 2204819.
http://dx.doi.org/10.1002/advs.202204819 PMID: 36815359

[140] Rearick, D.; Prakash, A.; McSweeny, A.; Shepard, S.S.; Fedorova, L.; Fedorov, A. Critical association of ncRNA with introns. *Nucleic Acids Res.,* **2011**, *39*(6), 2357-2366.
http://dx.doi.org/10.1093/nar/gkq1080 PMID: 21071396

[141] Soemedi, R.; Cygan, K.J.; Rhine, C.L.; Glidden, D.T.; Taggart, A.J.; Lin, C.L.; Fredericks, A.M.; Fairbrother, W.G. The effects of structure on pre-mRNA processing and stability. *Methods,* **2017**, *125*, 36-44.
http://dx.doi.org/10.1016/j.ymeth.2017.06.001 PMID: 28595983

[142] Gu, C.; Zhou, N.; Wang, Z.; Li, G.; Kou, Y.; Yu, S.; Feng, Y.; Chen, L.; Yang, J.; Tian, F. circGprc5a promoted bladder oncogenesis and metastasis through Gprc5a-targeting peptide. *Mol. Ther. Nucleic Acids,* **2018**, *13*, 633-641.
http://dx.doi.org/10.1016/j.omtn.2018.10.008 PMID: 30497053

[143] Yang, Y.; Gao, X.; Zhang, M.; Yan, S.; Sun, C.; Xiao, F.; Huang, N.; Yang, X.; Zhao, K.; Zhou, H.; Huang, S.; Xie, B.; Zhang, N. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis. *J. Natl. Cancer Inst.,* **2018**, *110*(3), 304-315.
http://dx.doi.org/10.1093/jnci/djx166 PMID: 28903484

[144] Zheng, X.; Chen, L.; Zhou, Y.; Wang, Q.; Zheng, Z.; Xu, B.; Wu, C.; Zhou, Q.; Hu, W.; Wu, C.; Jiang, J. A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer *via* Hippo-YAP signaling. *Mol. Cancer,* **2019**, *18*(1), 47.

http://dx.doi.org/10.1186/s12943-019-1010-6 PMID: 30925892

[145] Pan, Z.; Cai, J.; Lin, J.; Zhou, H.; Peng, J.; Liang, J.; Xia, L.; Yin, Q.; Zou, B.; Zheng, J.; Qiao, L.; Zhang, L. A novel protein encoded by circFNDC3B inhibits tumor progression and EMT through regulating snail in colon cancer. *Mol. Cancer,* **2020**, *19*(1), 71. http://dx.doi.org/10.1186/s12943-020-01179-5 PMID: 32241279

[146] Kralovicova, J.; Patel, A.; Searle, M.; Vorechovsky, I. The role of short RNA loops in recognition of a single-hairpin exon derived from a mammalian-wide interspersed repeat. *RNA Biol.,* **2015**, *12*(1), 54-69. http://dx.doi.org/10.1080/15476286.2015.1017207 PMID: 25826413

[147] Legnini, I.; Di Timoteo, G.; Rossi, F.; Morlando, M.; Briganti, F.; Sthandier, O.; Fatica, A.; Santini, T.; Andronache, A.; Wade, M.; Laneve, P.; Rajewsky, N.; Bozzoni, I. Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. *Mol. Cell,* **2017**, *66*(1), 22-37.e9. http://dx.doi.org/10.1016/j.molcel.2017.02.017 PMID: 28344082

[148] Dong, R.; Ma, X.K.; Chen, L.L.; Yang, L. Increased complexity of circRNA expression during species evolution. *RNA Biol.,* **2017**, *14*(8), 1064-1074. http://dx.doi.org/10.1080/15476286.2016.1269999 PMID: 27982734

[149] Aktaş, T.; Avşar Ilık, İ.; Maticzka, D.; Bhardwaj, V.; Pessoa Rodrigues, C.; Mittler, G.; Manke, T.; Backofen, R.; Akhtar, A. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. *Nature,* **2017**, *544*(7648), 115-119. http://dx.doi.org/10.1038/nature21715 PMID: 28355180

[150] de la Peña, M. Circular RNAs biogenesis in eukaryotes through self-cleaving hammerhead ribozymes. *Adv. Exp. Med. Biol.,* **2018**, *1087*, 53-63. http://dx.doi.org/10.1007/978-981-13-1426-1_5 PMID: 30259357

[151] Cervera, A.; de la Peña, M. Cloning and detection of genomic retrozymes and their circRNA intermediates. *Methods Mol. Biol.,* **2021**, *2167*, 27-44. http://dx.doi.org/10.1007/978-1-0716-0716-9_3 PMID: 32712913

[152] Welden, J.R.; Stamm, S. Pre-mRNA structures forming circular RNAs. *Biochim. Biophys. Acta. Gene Regul. Mech.,* **2019**, *1862*(11-12), 194410. http://dx.doi.org/10.1016/j.bbagrm.2019.194410 PMID: 31421281

[153] Xia, X.; Li, X.; Li, F.; Wu, X.; Zhang, M.; Zhou, H.; Huang, N.; Yang, X.; Xiao, F.; Liu, D.; Yang, L.; Zhang, N. A novel tumor suppressor protein encoded by circular AKT3 RNA inhibits glioblastoma tumorigenicity by competing with active phosphoinositide-dependent Kinase-1. *Mol. Cancer,* **2019**, *18*(1), 131. http://dx.doi.org/10.1186/s12943-019-1056-5 PMID: 31470874

[154] Song, R.; Guo, P.; Ren, X.; Zhou, L.; Li, P.; Rahman, N.A.; Wołczyński, S.; Li, X.; Zhang, Y.; Liu, M.; Liu, J.; Li, X. A novel polypeptide CAPG-171aa encoded by circCAPG plays a critical role in triple-negative breast cancer. *Mol. Cancer,* **2023**, *22*(1), 104. http://dx.doi.org/10.1186/s12943-023-01806-x PMID: 37408008

[155] Zhang, M.; Huang, N.; Yang, X.; Luo, J.; Yan, S.; Xiao, F.; Chen, W.; Gao, X.; Zhao, K.; Zhou, H.; Li, Z.; Ming, L.; Xie, B.; Zhang, N. A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. *Oncogene,* **2018**, *37*(13), 1805-1814. http://dx.doi.org/10.1038/s41388-017-0019-9 PMID: 29343848

[156] Liang, W.C.; Wong, C.W.; Liang, P.P.; Shi, M.; Cao, Y.; Rao, S.T.; Tsui, S.K.W.; Waye, M.M.Y.; Zhang, Q.; Fu, W.M.; Zhang, J.F. Translation of the circular RNA circβ-catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biol.,* **2019**, *20*(1), 84. http://dx.doi.org/10.1186/s13059-019-1685-4 PMID: 31027518

[157] Song, R.; Ma, S.; Xu, J.; Ren, X.; Guo, P.; Liu, H.; Li, P.; Yin, F.; Liu, M.; Wang, Q.; Yu, L.; Liu, J.; Duan, B.; Rahman, N.A.; Wołczyński, S.; Li, G.; Li, X. A novel polypeptide encoded by the circular RNA ZKSCAN1 suppresses HCC *via* degradation of mTOR. *Mol. Cancer,* **2023**, *22*(1), 16. http://dx.doi.org/10.1186/s12943-023-01719-9 PMID: 36691031

[158] Ge, J.; Wang, J.; Xiong, F.; Jiang, X.; Zhu, K.; Wang, Y.; Mo, Y.; Gong, Z.; Zhang, S.; He, Y.; Li, X.; Shi, L.; Guo, C.; Wang, F.; Zhou, M.; Xiang, B.; Li, Y.; Li, G.; Xiong, W.; Zeng, Z. Epstein–Barr virus–encoded circular RNA CircBART2.2 promotes immune escape of nasopharyngeal carcinoma by regulating PD-L1. *Cancer Res.,* **2021**, *81*(19), 5074-5088. http://dx.doi.org/10.1158/0008-5472.CAN-20-4321 PMID: 34321242

[159] Tagawa, T.; Gao, S.; Koparde, V.N.; Gonzalez, M.; Spouge, J.L.; Serquiña, A.P.; Lurain, K.; Ramaswami, R.; Uldrick, T.S.; Yarchoan, R.; Ziegelbauer, J.M. Discovery of Kaposi's sarcoma herpesvirus-encoded circular RNAs and a human antiviral circular RNA. *Proc. Natl. Acad. Sci. USA,* **2018**, *115*(50), 12805-12810. http://dx.doi.org/10.1073/pnas.1816183115 PMID: 30455306

[160] Zhao, J.; Lee, E.E.; Kim, J.; Yang, R.; Chamseddin, B.; Ni, C.; Gusho, E.; Xie, Y.; Chiang, C.M.; Buszczak, M.; Zhan, X.; Laimins, L.; Wang, R.C. Transforming activity of an oncoprotein-encoding circular RNA from human papillomavirus. *Nat. Commun.,* **2019**, *10*(1), 2300. http://dx.doi.org/10.1038/s41467-019-10246-5 PMID: 31127091

[161] Cai, Z.; Lu, C.; He, J.; Liu, L.; Zou, Y.; Zhang, Z.; Zhu, Z.; Ge, X.; Wu, A.; Jiang, T.; Zheng, H.; Peng, Y. Identification and characterization of circRNAs encoded by MERS-CoV, SARS-CoV-1 and SARS-CoV-2. *Brief. Bioinform.,* **2021**, *22*(2), 1297-1308. http://dx.doi.org/10.1093/bib/bbaa334 PMID: 33757279

[162] Arora, A.; Kolberg, J.E.; Srinivasachar Badarinarayan, S.; Savytska, N.; Munot, D.; Müller, M.; Krchlíková, V.; Sauter, D.; Bansal, V. SARS-CoV-2 infection induces epigenetic changes in the LTR69 subfamily of endogenous retroviruses. *Mob. DNA,* **2023**, *14*(1), 11. http://dx.doi.org/10.1186/s13100-023-00299-1 PMID: 37667401