UDC 577.212.3 + 577.214.5

NON-CODING RNAS AND EPIGENOME: *DE NOVO* DNA METHYLATION, ALLELIC EXCLUSION AND X-INACTIVATION

V. A. HALYTSKIY, S. V. KOMISARENKO

Palladin Institute of Biochemistry, National Academy of Sciences of Ukraine, Kyiv;
e-mail: volha@biochem.kiev.ua

Non-coding RNAs are widespread class of cell RNAs. They participate in many important processes in cells – signaling, posttranscriptional silencing, protein biosynthesis, splicing, maintenance of genome stability, telomere lengthening, X-inactivation. Nevertheless, activity of these RNAs is not restricted to posttranscriptional sphere, but cover also processes that change or maintain the epigenetic information.

Non-coding RNAs can directly bind to the DNA targets and cause their repression through recruitment of DNA methyltransferases as well as chromatin modifying enzymes. Such events constitute molecular mechanism of the RNA-dependent DNA methylation. It is possible, that the RNA-DNA interaction is universal mechanism triggering DNA methylation de novo.

Allelic exclusion can be also based on described mechanism. This phenomenon takes place, when non-coding RNA, which precursor is transcribed from one allele, triggers DNA methylation in all other alleles present in the cell. Note, that miRNA-mediated transcriptional silencing resembles allelic exclusion, because both miRNA gene and genes, which can be targeted by this miRNA, contain elements with the same sequences. It can be assumed that RNA-dependent DNA methylation and allelic exclusion originated with the purpose of counteracting the activity of mobile genetic elements.

Probably, thinning and deregulation of the cellular non-coding RNA pattern allows reactivation of silent mobile genetic elements resulting in genome instability that leads to ageing and carcinogenesis.

In the course of X-inactivation, DNA methylation and subsequent heterochromatinization of X chromosome can be triggered by direct hybridization of 5'-end of large non-coding RNA Xist with DNA targets in remote regions of the X chromosome.

Key words: epigenome, non-coding RNA, miRNA, de novo DNA methylation, allelic exclusion, transposon, silencing, chromatin remodeling, X-inactivation.

Despite the central dogma of molecular biology, not all cell RNAs are translated into protein sequences. Some of non-coding RNAs – transfer RNAs, ribosomal RNAs, small nuclear RNAs, ribonuclease P – have been studied long ago. However, their involvement in protein biosynthesis machinery did not allow to understand that there is a whole world of RNAs. Recently it was determined that at least 90% of the genome undergoes active transcription, whereas only 2% encodes proteins [1].

Small non-coding RNAs (usually, less than 200 nt in length) belong to several types and perform, besides translation and splicing, many other important processes in cells:

- microRNAs (miRNAs) – silencing of cell genes involved in proliferation, differentiation and apoptosis;
- short interfering RNAs (siRNAs) – silencing of alien genetic information;
- Piwi-interacting RNAs (piRNAs) and endogenous short interfering RNAs (endo-siRNAs) – maintenance of genome stability via silencing of mobile genetic elements in germline and early embryogenesis;
- X-inactivation linked RNAs (xiRNAs) – X-inactivation [2];
- promoter-associated RNAs (paRNAs), also known as transcription start site-associated RNAs (TSSaRNAs) and as transcription initiation RNAs (tiRNAs) – transcriptional regulation of downstream genes [3], maintaining accessible chromatin architecture [4];
- enhancer RNAs (eRNAs) – probably, transcription activation [3];
- circular RNAs (circRNAs) – probably, miRNA sequestration [5].

Also, a group of small non-coding RNAs includes centromere repeat associated small interacting RNAs (crasiRNAs), telomere-specific small RNAs (tel-sRNAs) and pyknons; their functions remain unclear [1].

A group of long non-coding RNAs (lncRNAs) consists of long intergenic non-coding RNAs

(lincRNAs), long intronic non-coding RNAs, telomere-associated non-coding RNAs (TERRAs), long non-coding RNAs with dual functions, pseudogene RNAs and transcribed-ultraconserved regions (T-UCRs) [1]. lncRNAs participate in the regulation of embryo development, imprinting, allelic exclusion, recruitment of transcription factors and modulation of their activity. Some of lncRNAs – Xist RNA and Tsix RNA – are involved in X-inactivation [2]. Another important lncRNA – telomerase RNA component (TERC) – serves as template for reverse transcription of telomeric repeats. Sometimes, lncRNAs can be processed into small non-coding RNAs involving in usual RNA interference.

Non-coding RNAs are often associated with diseases [6]. So, tumor growth is tightly intertwined with regular shifts in miRNA expression pattern. More than 50% of miRNA genes are located in fragile chromosomal regions that undergo the amplification, translocation and other damages during the carcinogenesis. Shifts in miRNA pattern facilitate proliferation and surviving [7], impair cytoskeleton and adhesion as well as increase motility of the cancer cells [8]. Also, lncRNA deregulation and abnormal expression is described in some tumors [1].

Because listed non-coding RNAs are involved in the regulation of genetic information expression, the aim of present research is to identify in what ways non-coding RNAs can perform such functions, in particular, how these RNAs affect the cellular epigenetic information (epigenome). As some non-coding RNAs are able to cause gene repression at transcription level, there should be a common pathway for the initiation of epigenetic changes by non-coding RNA.

Interfering RNAs and RNA-dependent DNA methylation

siRNAs were discovered as molecules mediating RNA interference triggered by double-stranded RNA (dsRNA). dsRNA is recognized and cleaved by endoribonuclease III Dicer into fragments of about 22 nucleotides in length [9]. These fragments enter the RISC [10, 11], where dsRNA is untwisted with the aid of Argonaute protein, and one strand of dsRNA remains bound to this protein. This monochain oligonucleotide (aptamer) is the siRNA. After binding to complementary sites on RNA, siRNAs initiate their degradation; as a result, the gene expression is interrupted at the posttranscriptional stage.

Biogenesis of miRNAs, which can also trigger gene silencing, is like. Most miRNAs are transcribed from their own genes that contain gene promoter and regulatory units. However, as much as 40% miRNAs are originated from introns of protein coding genes, whereas approximately 10% – from lncRNA introns [12]. miRNA sequences are synthesized as over 1 kb long primary transcripts (pri-miRNAs). pri-miRNAs form imperfect hairpins which are processed by the nuclear ribonuclease III Drosha into double-stranded pre-miRNAs, about 60 b.p. in length [13]. Pre-miRNAs enter the cytoplasm due to active transport with participation of exportin 5 [14] and subsequently are processed, like siRNA, by ribonuclease Dicer [15]. Then mature miRNAs can silence gene expression at the posttranscriptional level – expression of 60% genes in the human genome can be regulated by miRNAs [16]. The total number of miRNA genes might be closer to a thousand and thus potentially constitutes about 3% of the genome [17].

Nevertheless, siRNAs can initiate gene silencing not only at posttranscriptional but also at transcriptional level. It has been shown that artificial siRNA against CG-containing promoters causes in human cells methylation of DNA and K9 residue of histone H3 *de novo* [18]. This process is dependent on the simultaneous presence of DNA methyltransferases DNMT1 and DNMT3b, but not DNMT2. Assuming this background, we had attempted to define the link between interfering RNAs and epigenome.

Concentration of C(N)G sites in siRNAs and miRNAs. As is known, in the course of evolution the majority of cytidylate residues in CG dinucleotides are lost due to the deamination of 5-methylcytosine to thymine. At present, for instance, frequency of CG dinucleotide in human or mouse genome amounts to 1%, while in rat genome – 1.2% [19, 20]. These levels are more than 4 times lower than expected CG frequency in the random DNA sequence (e.g., since the human genome contains 42% G-C pairs, frequency of CG dinucleotide should be equal to $0.21 \cdot 0.21 = 4.41\%$).

Nevertheless, we found *in silico* that the concentration of CG dinucleotide amounts to 2.85% in siRNA sequences and to $2.58 \pm 0.17\%$ in mature miRNA sequences [21–23]. It was quite intriguing that the concentration of CG sites in miRNAs and siRNAs significantly exceeds the general genome level.

The frequency of CNG trinucleotides (here N designates any nucleotide) amounts to 6.29% in siRNA sequences and to $5.77 \pm 0.25\%$ in mature miRNA sequences [21–23], although the average frequency of CNG trinucleotide in human genome should be equal to $0.21 \cdot (0.21 + 0.21 + 0.29 + 0.29) \cdot 0.21 = 4.41\%$, where 0.21 – level of cytosine as well as guanine, 0.29 – level of adenine

as well as thymine. Thus, CNG trinucleotide was also found significantly more often in siRNA and miRNA sequences than in the random genome sequence.

It is also very significant that only 15.69% siRNA sequences and 20.01% mature miRNA sequences contain none of CG or CNG sites [23, 24].

Localization of C(N)G sites in siRNAs and miRNAs. Also, siRNA and miRNA sequences have been examined in order to discovery of localization pattern of CG dinucleotides and CNG trinucleotides in miRNA sequences. As a result, certain regularities have been found in this pattern.

Although the frequency of CG sites exceeds the average genomic level along whole siRNA sequence, peaks of their concentration are found close to the nucleotides 4, 11 and 16 from 5′-end. CNG sites are congregated around the nucleotides 4 and 13 [25].

In miRNAs, one of the greatest peaks of CG site concentration is located close to the nucleotide 10 from 5′-end [23]. The level of CG sites exceeds in this peak more than four the average genomic frequency of this dinucleotide. Other evolutionary conservative peaks are located around the nucleotides 16 and 21 in miRNA sequence. Peaks of CNG site concentration correspond with nucleotides 4 and 8 [23].

Approximately half of miRNA sequences contain two or more C(N)G sites. The distance between these sites amounts mostly to 1 or 3-9 nucleotides [23].

So, the peaks of C(N)G sites concentration are located not only within seed region, but also along the full length of miRNA sequence. Nevertheless, these sites are disposed in the same miRNA molecule mainly within one turn of double helix.

Interfering RNAs as pointers of RNA-dependent DNA methylation. Above described investigations indicate that CG dinucleotide in siRNAs and miRNAs is present more often than it should be found in random sequence. CG sites are found in about half of the siRNAs and miRNAs molecules. Overwhelming majority of siRNAs and miRNAs contain also CNG trinucleotide that is also present more often than in random sequence. At the same time, only one fifth part of miRNA sequences and less than one sixth part of siRNA sequences contains none of CG and CNG sites. This indicates that CG dinucleotides and CNG trinucleotides in interfering RNAs should have essential biological role.

On this basis, we had proposed that siRNAs and miRNAs should directly scan the DNA strands unwinding during transcription (this explains dependence of RNA-mediated transcriptional silencing on functioning of RNA polymerase) [21, 22, 24]. Most likely, interfering RNAs interact with the nontemplate (coding) DNA strand. Therefore, siRNAs and miRNAs are also complementary to sequence of RNA transcript and can hybridize with it in the course of usual post-transcriptional silencing.

On encounter of complementary DNA site, siRNA or miRNA should bind to it directly (Fig. 1). This should involve the cellular DNA methyltransferases in cytosine methylation *de novo* in CG dinucleotides and CNG trinucleotides, which are bound to complementary sites in the interfering RNA (Fig. 1) [22]. The maintenance methyltransferase DNMT1 can methylate the complementary DNA strand (template) already after the removal of interfering RNA from RNA-DNA duplex and renewal of the double-stranded DNA structure. It is also possible that these events involve histone deacetylases and histone methylases that remove active chromatin marks and, respectively, apply marks of silenced chromatin; however, more likely, the components of chromatin remodeling complexes are recruited indirectly because of recognition of the methylated DNA (Fig. 1) [24]. C(N)G site localization pattern (see above) can be evidence of participation of whole miRNA or siRNA sequence in proposed mechanism of RNA-dependent DNA methylation [23].

At first sight, suggested mechanism can lead only to methylation of sites which DNA is unwinding by RNA polymerase, i.e., within the transcribed DNA. If so, how are methylated the regulatory elements in non-transcribed DNA regions? This collision can be solved by assumption that the epigenetic marks are able to be spread throughout the entire DNA domain – in particular, because processes of histone deacetylation and methylation can target not only the nucleosomes connected directly with the methylated DNA sites, but also the adjacent ones (Fig. 1). Then the DNA methyltransferases recognize methylated histones and thereafter methylate DNA associated with them; this leads again to deacetylation and methylation of histones in adjacent nucleosomes and so on [24].

As a result of these events, the gene recognized by siRNA or miRNA is redirected to heterochromatin and becomes steadily inactivated. In addition, if the recognized motif is present in several genes, all of them can be repressed simultaneously [24].

Recent research confirms that non-coding RNAs can directly hybridize with DNA, causing involvement of the *de novo* DNA methyltransferase DNMT3B [26]. Small non-coding RNAs can bind

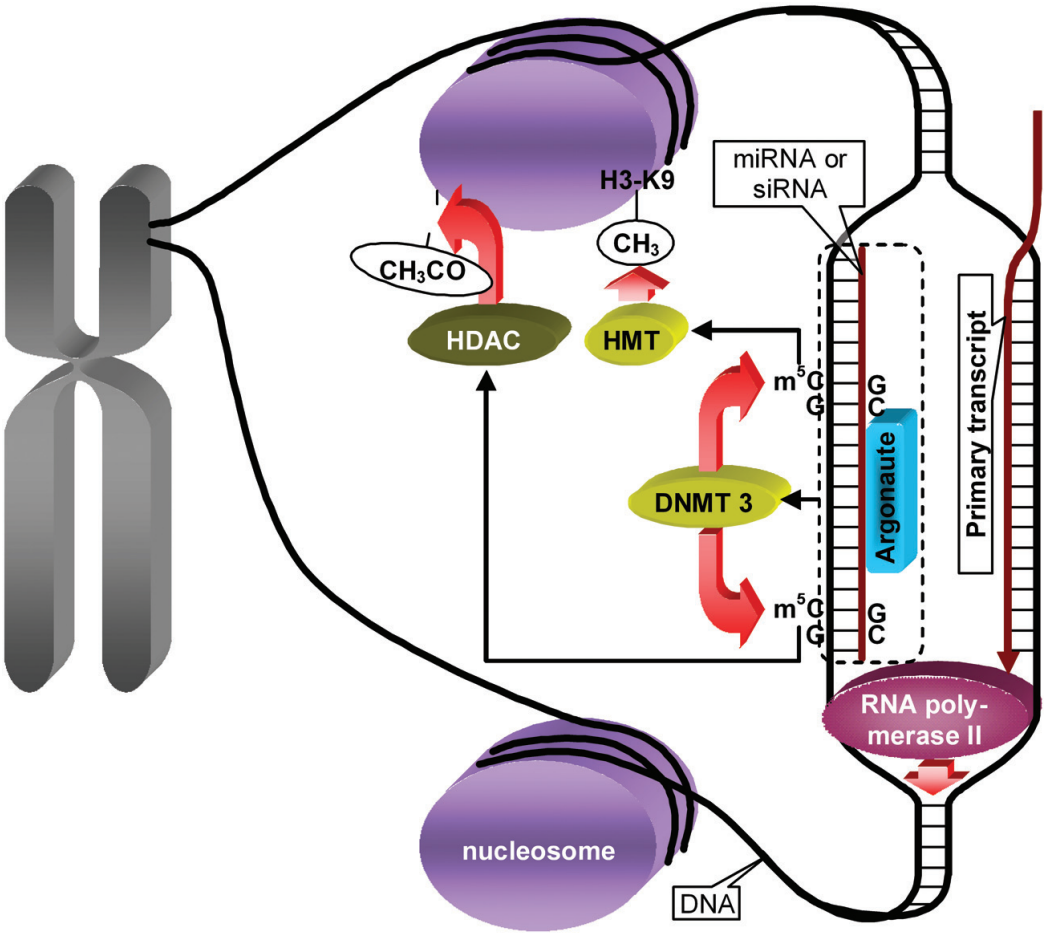

Fig. 1. Initiation of DNA methylation de novo by interfering RNAs

to target DNA sequences, forming of RNA:DNA heteroduplexes or RNA:DNA:DNA triplexes; in addition, lncRNAs can recognize specific features of the chromatin surface. These events should cause involvement of the chromatin modifying complexes [27].

We have also suggested that the RNA-induced transcriptional silencing complex (RITS), which is discovered in some organisms [28, 29], is assembled already after the events described above. RITS function is to initiate cleavage of RNA transcript of the gene targeted by siRNA molecule. Probably, the first element, which should bind with the siRNA-Argonaute complex (already after hybridization siRNA with complementary DNA sequence and methylation of DNA and histone H3), is the chromodomain-containing protein Chp1 (Fig. 2) [24]. Protein Tas3 joins the complex after (since Tas3 is able to bind RNA [30], it is likely, that the key role of this protein is the search of the RNA transcript which synthesis entails the recognition of the gene sequence by siRNA). Here the RITS assembling is completed, after which it binds RNA-directed RNA polymerase (RdRP) and aims it to dsRNA synthesis on the transcript that still is bound to RNA polymerase. Thus, initiation of degradation of the primary transcript (and the rest of RNA molecules which had been transcribed earlier from this gene) requires, first, recognition of the complementary DNA motif by siRNA and, second, interaction of the Chp1 protein chromodomain with methyl mark on the residue K9 of histone H3. This should prevent unauthorized degradation of the transcript [24].

This sequence of RITS complex assembly and function predicts that siRNAs and miRNAs, which contain none of CG dinucleotides or CNG trinucleotides, should be involved only in the usual posttranscriptional silencing.

Molecular basis of allelic exclusion. Explanation of molecular mechanism of allelic exclusion as well as other phenomena with the gene dosage effect represents a very complicated problem. Not all genes undergo the allelic exclusion, but even if

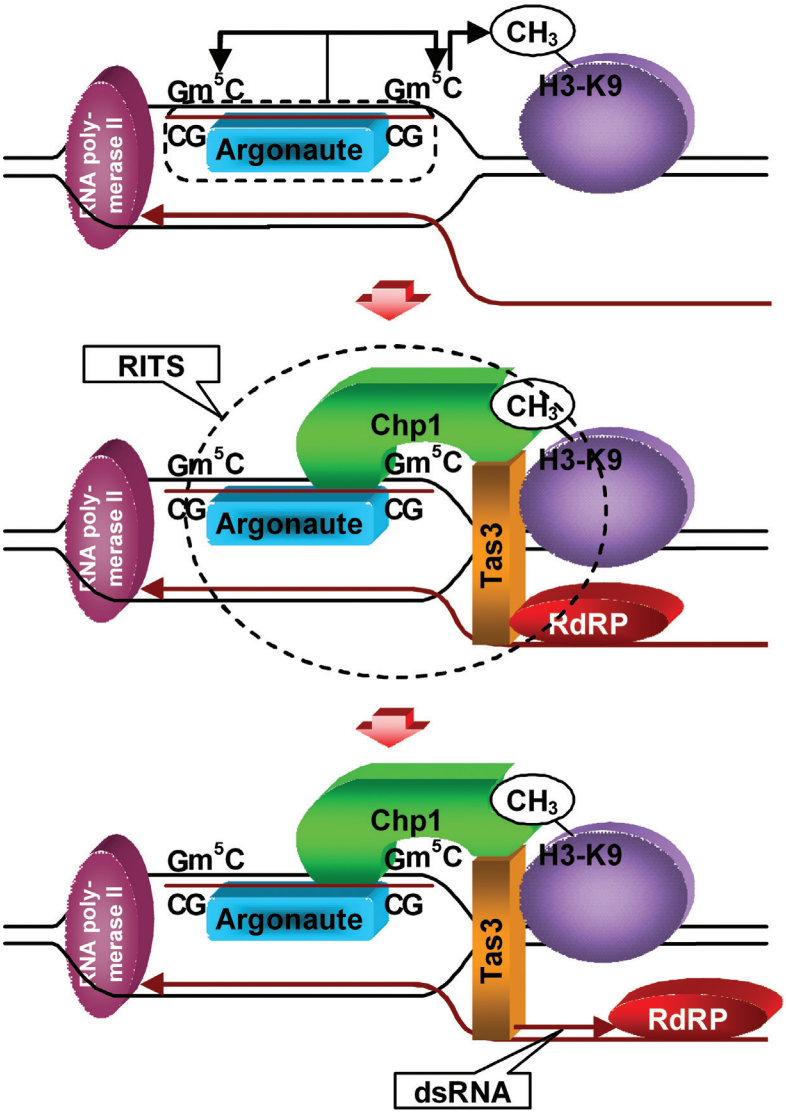

Fig. 2. Assembling of the RITS complex

it is characteristic of certain gene, allelic exclusion will be constantly maintained in spite of largest fluctuations of gene expression level and regulatory influences. Regardless of the number of alleles, no more and no less than one of them will remain active in the cell as a result of the allelic exclusion.

Nevertheless, the molecular mechanism of allelic exclusion should be surprisingly simple when supposing that the non-template DNA strand can be also transcribed and its transcript encodes precursor of miRNA, which mature form can subsequently target this DNA strand [24, 31]. Bidirectional transcription is found, at least sometimes, in genes undergoing allelic exclusion.

Probably, the concentration of miRNA transcribed from only one allele is in all cases insufficient to overcome the threshold of initiation of the above-proposed process of RNA-dependent DNA methylation leading to the repression of this allele *de novo*. The presence of such threshold is expected, because RNA polymerase movement along transcribing DNA strand should release it from the bound miRNAs (possibly, such removal or cleavage of miRNA should occur with certain delay, as the miRNA-DNA hybrid is somewhat more stable than double-stranded DNA) [24]. Thereat, miRNAs have no time to initiate the epigenetic modifications.

If more than one allele are active in cell for some reason, miRNA concentration overcomes the above mentioned threshold sooner or later. In this case, RNA polymerase has no time to remove miRNAs before they trigger the DNA methylation and further remodeling of chromatin within

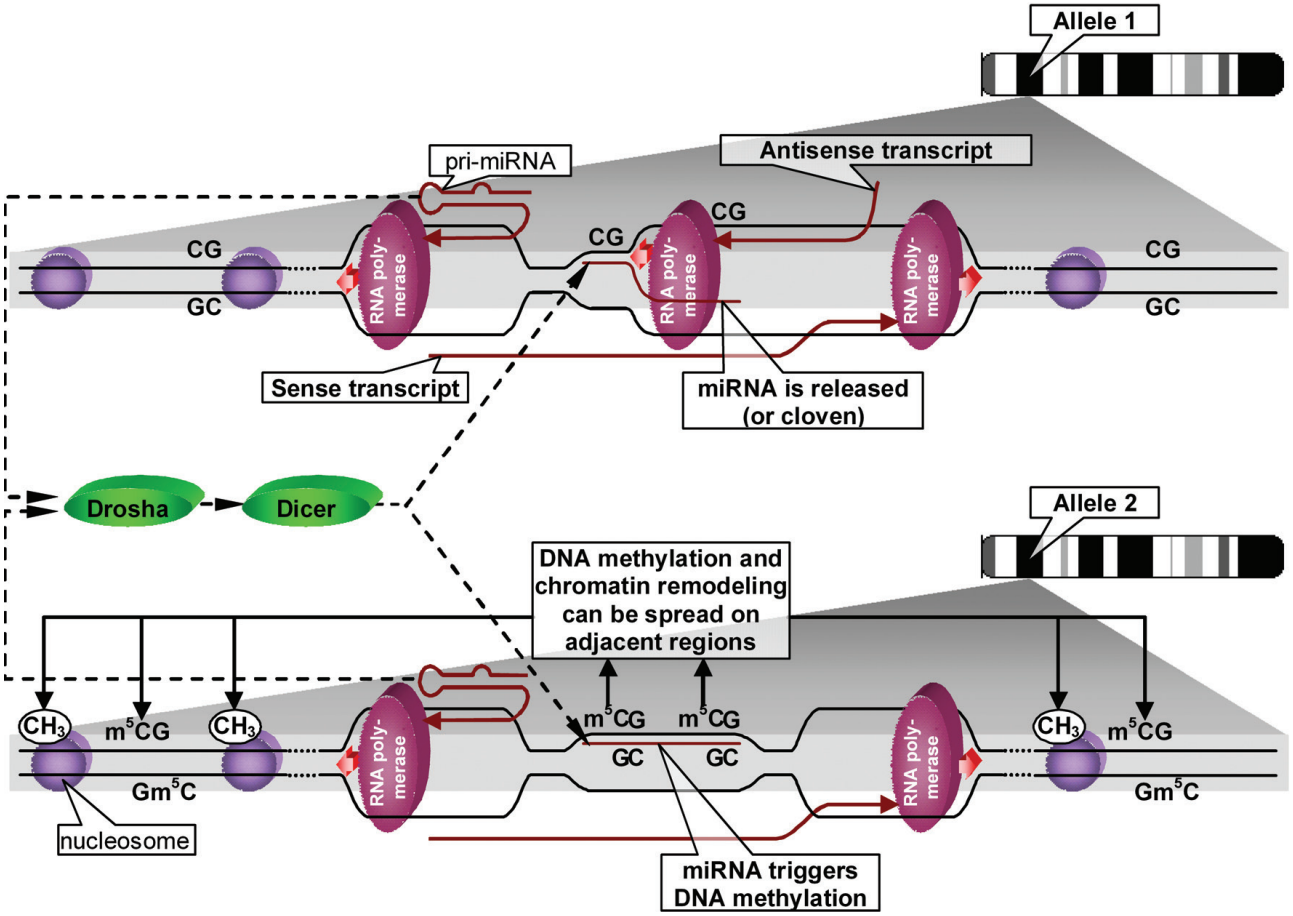

Fig. 3. Mechanism of allelic exclusion: allele 2 undergo the repression, allele 1 remains active

bound targets (Fig. 3). As a result, both strands of DNA in the marked allele undergo the repression, whereas the other allele remains active. These events constitute the molecular mechanism of allelic exclusion [24].

It is also to be noted that the miRNA gene transcript during its own synthesis could not induce the DNA methylation itself, because the pri-miRNA aspires to fold in hairpin configuration leading to fast break of its connection with the transcribed DNA chain (Fig. 3).

Obviously, the repression process targets the allele that, all other things being equal, is less actively released from the bound miRNAs, i.e., is weaker transcribed from the DNA strand which transcript is the miRNA precursor (for instance, due to some peculiarities of its chromatin or adjacent chromosome regions). It can also be suggested that the repression-initiated threshold of miRNA concentration is not standard value for each of alleles, but is flexibly tuned. The more active is transcription and, consequently, separation of the miRNAs from DNA targets, the higher is this threshold (Fig. 3). An important consequence of this is that, if the second allele begins its expression in the cell, the repression-determining threshold of miRNA concentration for this allele can become significantly lower than for the already completely active allele. In the extreme case, concentration of the miRNA transcribed from the active allele, being subthreshold one for repression of this allele itself, at the same time turns out to be already superthreshold for the alleles beginning the expression, which leads to their silencing. On the other hand, owing to the tunable threshold, the allelic exclusion can be established at the very different levels of expression of the involved genes [24].

It can be expected that the non-random pattern of exclusion of some paternal and maternal alleles in interspecies hybrids is certainly associated with the initially determined inequality of their expression levels, for instance due to differences in nucleotide sequences of the allele regulatory elements of genes belonging to different species.

In the other extreme case, if all alleles are approximately equally active and accordingly can

begin simultaneously the labeling with the repressive chromatin marks, the functioning of alleles will be affected by counteraction of stimuli activating gene expression (these stimuli are able to recruit the chromatin remodeling complexes that remove markers of the silent chromatin) and of RNA-dependent DNA methylation triggered due to accumulation of miRNAs. Under these conditions the miRNA concentration evidently should aspire to reach the value near the threshold for the present transcription intensity; but since this state is unsteady or unachievable, the random fluctuations at application and removal of epigenetic marks can determine in the end, which of the alleles will undergo the repression, and which will remain active [24].

The palindromic structure of miRNA precursors reveal that the genes originate from mobile genetic elements [32]. It is likely that the mechanisms that mediate the allelic exclusion and other gene dosage effect appeared in the evolution exactly with the purpose of transcriptional repression of transposons. Probably, the cells had allowed expression of only one copy of transposable element, transcript of which was used for recognition of all other copies of this element and for involving of the DNA methyltransferases and chromatin remodeling complexes in repression of these copies [24].

Recent studies confirm the involvement of RNA interference in transposon repression – transcripts of *LINE1* retrotransposons (see below), being double-stranded, can be processed into siRNAs by endonuclease III DICER; moreover, the presence of these siRNAs suppresses retrotransposons activity as well as causes degradation of *LINE1* transcripts, i.e. posttranscriptional silencing [33]. It is also shown that retrotransposons are reactivated in the lack of DNA methyltransferases, particularly, Dnmt3L that is involved in the *de novo* DNA methylation [34] or Dnmt1 that is responsible for maintenance of DNA methylation pattern [35].

As a result, the cells with the aid of interfering RNAs suppress functioning of transposons, maintain the cell genome stability and counteract the horizontal gene transfer. Also, this mechanism protects the cells from consequences of the excessive copying of some genes, caused by mobile genetic elements. The similar mechanisms using siRNAs prevent also the multiplication of the alien genetic information invading the cell. siRNAs silence this information or, if it is represented only as RNA, trigger its total cleavage. Sometimes, these mechanisms target the genes transferred purposefully into cells with the aid of biotechnological manipulations.

miRNA-directed allelic exclusion in immunoglobulin gene loci. It has been suggested that miRNAs can cause allelic exclusion in the immunoglobulin gene loci in the same way [36]. *In silico* search in miRNA database miRBase shows that genes of four human miRNAs – miR-4539, miR-4507, miR-4538, miR-4537 – are clustered within immunoglobulin heavy chain locus (IGH@) of chromosome 14, position 106325653-106323741 [37]. This cluster is located in a spacer between sequences encoding last J-fragment and first exon of Cμ-fragment on the same strand (negative). Also, miRNA candidate sequence AB019440.1 is transcribed from positive strand in V-fragments region of IGH@.

The gene of another human miRNA, miR-650, is found within immunoglobulin lambda chain locus (IGL@) on positive strand of chromosome 22, position 23165270-23165365. This gene overlaps exon 1 of IGLV2-8-001 gene and intron 1 of the gene encoding lincRNA D87024.2. Six miRNA candidate sequences – snoU13.380, D86994.2, D86994.1, D87015.1, D86998.1 and D87024.1 – are also transcribed from sequences located in IGL@ [37].

Therefore, localization of miRNAs genes in immunoglobulin gene loci corresponds to the predictions from hypothesis of miRNA-directed allelic exclusion. Furthermore, each mature sequence of miR-650, miR-4539, miR-4507, miR-4538 and miR-4537 contains 2-4 C(N)G sites. Also, exons of listed miRNA candidate sequences contain several C(N)G sites [37]. These facts confirm that miRNAs can cause allelic exclusion through the methylation of complementary DNA sequences.

Loss of interfering RNAs allows transposon reactivation, ageing and carcinogenesis. Significant part of eukaryotic DNA consists of transposons – the nucleotide sequences that can change their localization within the genome. Part of transposons in the human genome is 45% [38], which corresponds to 3-4 million copies. There are two basic classes of transposons: DNA-transposons, which accounts for 3% of nuclear DNA, and retroelements. Retroelements, in their turn, are subdivided into 3 sub-classes: long interspersed nuclear elements (*LINEs*), short interspersed nuclear elements (*SINEs*) and retrovirus-like elements with long terminal repeats (*LTRs*). The latter also include endogenous retroviruses [39]. *LINEs* occupy 21% of nuclear DNA, and *SINEs* and *LTRs* occupy 13% and 8%, respectively. The most common among transposons of *LINE* class are *LINE1* retrotransposons, and among *SINEs* – *Alu* sequences (1 million copies per genome). Most of transposons in genome were damaged in the course of evolution,

nevertheless, some defective transposons can transpose at present, using proteins encoded by remaining viable transposons [40].

Transposons play an important role in the evolution, since they can move DNA fragments that contain coding and regulatory sequences within the genome and even between genomes of different species, which results in assembling of new genes and in the rearrangement of sites that regulate their expression. Probably, transposon activity has led to the formation of linear chromosomes, and telomeres are reduced transposon sequences. This assumption is supported by observation that, at least sometimes, retrotransposons assume telomere function [41]. Addition of new telomeric repeats with the use of transposed RNA-template as well as telomerase that is a reverse transcriptase still resembles some stages of retrovirus replication; considerable similarity was found between sequences of retrotransposon reverse transcriptase and telomerase reverse transcriptase – TERT-subunit of telomerase [42].

However, at the level of individual cells and organisms, transposons are primarily the genomic parasites that pose a serious threat as they can cause both the genome destabilization through mass lesions [34] and single mutations, including mutations that provoke tumor growth and other diseases [43–46]. In particular, in chronic myeloid leukaemia *Alu* elements cause the reciprocal translocation of cell proto-oncogene *c-abl* under control of B-cell receptor gene promoter, outwardly manifested in Philadelphia chromosome appearance [47]. It is also evidenced that some endogenous retroviruses are still active [48, 49].

Twenty years ago, Murray had proposed that with time the number of transposons increases exponentially and they can kill a cell line or organism because of inactivation of essential cell gene through insertion of transposon sequences into this gene [50]. Although there is no such gene, we think that other aspects of this hypothesis mirror correctly the molecular mechanism of ageing.

Expression of transposons can be suppressed through methylation of their DNA [51] as well as through influence of interfering RNAs, in particular miRNAs [52]. Above it was shown how the interfering RNAs can induce methylation of transposon DNA in proposed process of RNA-dependent DNA methylation.

Recent studies have shown that transposons are suppressed in germline and some somatic cells by two other classes of interfering RNAs: piRNAs [53] and endo-siRNAs that are generated in the course of processing of double-stranded transcripts in early embryogenesis [54, 55]. It seems plausible that both piRNAs and endo-siRNAs come from the same RNA sequences that form double-stranded regions which are processed by different enzyme systems. Although it is not yet entirely clear why some cells produce piRNAs, and other ones produce endo-siRNAs, the global strategy of cells is understandable. Since transposons usually contain the bidirectional transcribed genes, and recombination of some classes of transposons requires the inverted repeats, transposon transcripts should form double-stranded regions. So, it is quite necessary for the cell to suppress as far as possible the expression of all DNA sequences producing transcripts that contain such regions.

It is believed that cells can involve all interfering RNAs, including piRNAs and endo-siRNAs, in the above proposed mechanism of RNA-dependent DNA methylation.

At the next stages of embryo development, piRNAs and endo-siRNAs disappear because they should initiate repression of the sequences from which precursors of these piRNAs and endo-siRNAs were transcribed. This allows the expression of the silenced genes, if they are activated legitimately in response to specific signals. At the same time, in principle, silencing of transposons must be kept over cell generations both by maintaining DNA methylation and by reciprocal renewal of repressive marks on DNA and histones.

However, a gradual decrease in the DNA methylation level is observed with age [56], and this means that over time the cells slowly lose repressive chromatin marks. Probably, this occurs because the accuracy of maintenance of DNA methylation is not absolute although it is no less than 99% [57]. Assuming that the accuracy of this process is 99%, it can be estimated that only $0.99^{50} \approx 0.605 \approx 61\%$ of the initial amount of methyl marks will remain in DNA after 50 cell generations.

We supposed that miRNAs expressed in the cells can subject the sequences, initially repressed by piRNA and endo-siRNA, to transcriptional and posttranscriptional silencing again [58]. However, it is still difficult to judge how well cell miRNAs can override the spectrum of diversity of these sequences. It is only known that miRNAs are able to repress on posttranscriptional level at least 60% of cell genes [16].

miRNAs often silence the stage-specific genes responsible for normal course of the cell differentiation [59]. Therefore, cell differentiation, starting with the early stages, requires repression of some miRNA genes, otherwise these miRNAs would prevent expression of stage-specific genes. On this basis, it has been suggested that, at some instant, the maturing cells must lose the ability to

renew repressive marks on the sequences of some transposons, if these marks are completely erased for any reason [58]. Moreover, it is also possible that this takes place not only in maturing but in all other body cells, including stem cells, because the production of piRNAs and endo-siRNAs is terminated earlier than the cells begin the differentiation that requires repression of miRNA genes.

Anyway, DNA demethylation, which can occur either spontaneously or in the course of repair of methylated DNA, or due to direct activation of gene expression caused by binding of stage-specific transcriptional factors, makes possible reactivation of some transposons in the cells which have lost interfering RNA that recognize sequences of these transposons. This should lead to dramatic increase of DNA damage level. Activated in response, the DNA repair systems, in particular, BRCA1-dependent mechanisms based on homologous recombination should cause not only the DNA recovery, but also illegitimate recombination in telomere caps, which leads to the conversion of T-loops into T-circles (telomeric caps represent ready pre-recombination Holliday junctions) – see Fig. 4 [58]. It predetermines the accelerated erosion of telomeres that are most accessible to the recombinases (this explains the unequal erosion of different telomeres in the cell) and, finally, apoptosis. In the case of severe DNA damage the cells may undergo apoptosis before the critical telomere

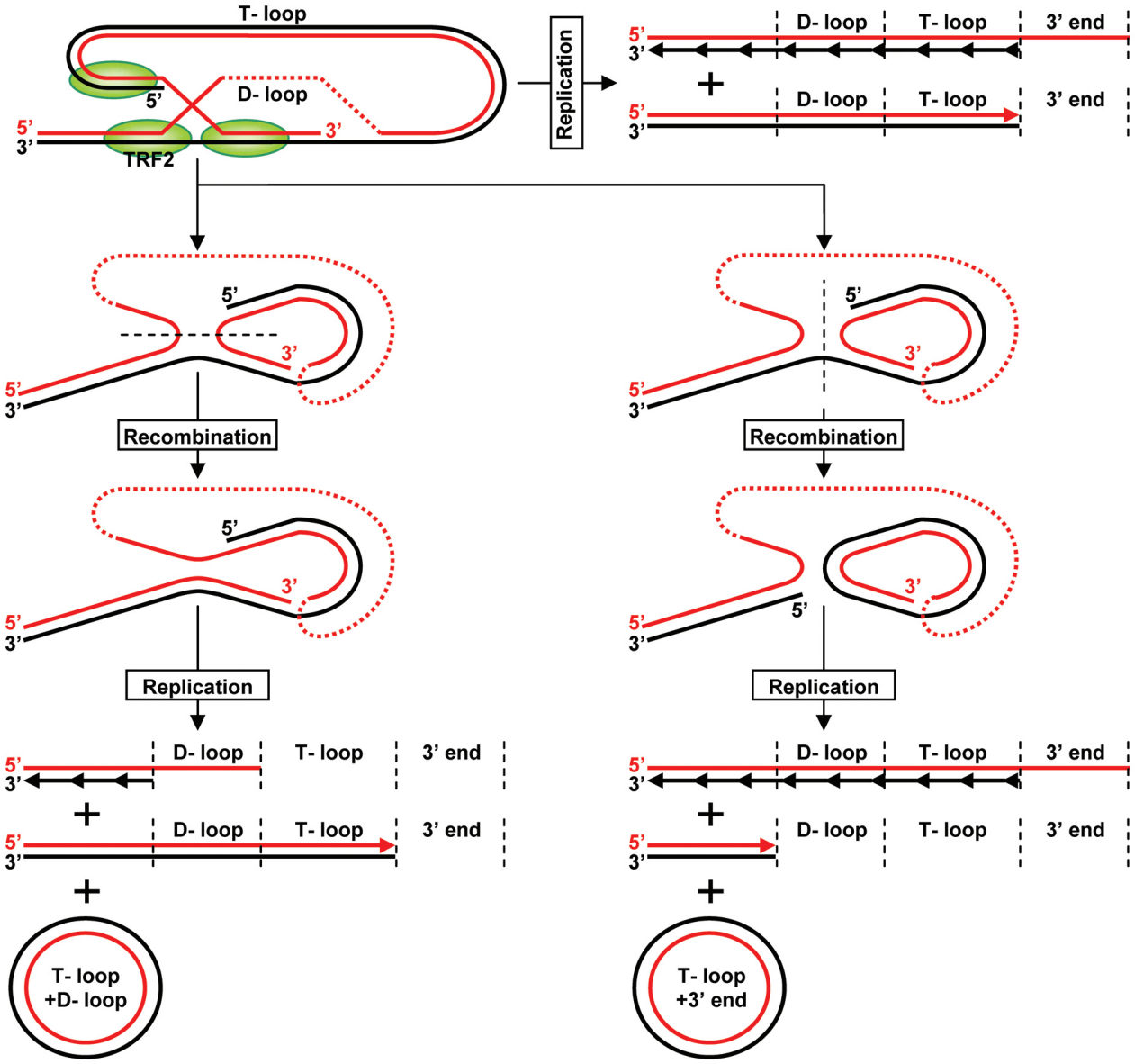

Fig. 4. Telomere and recombinational loss of telomeric DNA

erosion occurs. With age the increasing number of cells must reach a threshold of reactivation of dormant transposons; subsequent apoptosis of the majority of these cells leads to the ageing of the organism [58]. In addition, genome destabilization can cause not only apoptosis, but also tumor transformation of cells; thereby, transposon derepression determines the correlation between ageing and cancer appearance [60].

Species with higher genome stability possess greater life span and lower cancer risk because their cells reach the threshold of dormant transposon reactivation later than cells of species with low genome stability. This mirrors in amount of transposon-mediated mutations – in particular, the estimated frequency of retrotranspositions in mice is more than 60-fold greater than in humans [61]. Nevertheless, species with higher genome stability require steady ecological niche because their variability is low.

miRNA-mediated indirect impact on DNA methylation and chromatin remodeling. Recently, we have found *in silico* that miRNA can also indirectly influence epigenome through posttranscriptional silencing of genes encoding the *de novo* DNA methyltransferases, histone modifying enzymes as well as components of chromatin remodeling complexes. In particular, miRNAs, hyperexpression of which is essential for abnormal proliferation and surviving of cancer cells (miR-18a/b, miR-19, miR-21, miR-29a, miR-155, miR-181, miR-206, miR-210 and miR-221/222), can silence genes encoding histone deacetylases HDAC1/2/4/6/7/8/9 and SIRT1/3/5/7. Also, these miRNAs can silence *ASH1L*, *DOT1L*, *EHMT1/2*, *EZH1*, *MLL*, *MLL2/3/5*, *NSD1*, *PRDM2*, *SET*, *SETBP1*, *SETD1A/B*, *SETD2/3/5/6/7/8*, *SETDB1/2*, *SMYD1/2/4/5*, *SUV39H1/2* and *SUV420H1/2* genes encoding histone methyltransferases. Other targets of some of these miRNAs are genes encoding *de novo* DNA methyltransferases DNMT3A, DNMT3B and DNMT3L.

Down-regulation of other miRNAs (miR-16, miR-122, miR-31, miR-143, miR-145 and miR-320) allows overexpression of genes encoding histone acetyltransferases PCAF, Elp3, ATF2, MOZ, MORF, HBO1 and TIP60. Also, this down-regulation can lead to higher expression of *KDM1A/B*, *KDM2/B*, *KDM3A*, *KDM4A/B/C* and *KDM5A/C/D* genes encoding histone demethylases as well as of *ARID1A/B*, *SMARCC2*, *SMARCD1/2* and *ACTL6A* genes coding components of SWI/SNF chromatin remodeling complex.

This causes an increase of overall level of chromatin acetylation and expression and, therefore, makes possible the reactivation of silent oncogenes and transposons, which can rapidly lead to DNA damage and genome destabilization. Such shifts should underlie the initial stage of carcinogenesis.

Other mechanisms of influence of non-coding RNAs on the epigenome. The great length of lncRNAs makes possible the presence of domains with different functions. lncRNAs can recognize specific DNA motifs and simultaneously recruit proteins – e.g., histone modifying enzymes or transcription factors, aiming them to the DNA targets and causing chromatin remodeling. Also, lncRNAs can serve as scaffolds for assembling of multimolecular complexes binding heterochromatin proteins, e.g., in case of pericentric heterochromatin formation. In addition, lncRNAs can transmit changes of global chromosomal configuration into chromatin modification and vice versa [62, 63].

lncRNAs can sequester miRNAs, weakening the miRNA-mediated posttranscriptional silencing of true targets [62]. In addition, non-coding RNAs can protect some transcripts by masking the miRNA binding sites [63].

X chromosome and *cis*-interacting non-coding RNAs

The number of X chromosomes is different in XX and XY individuals and can increase several times in case of X aneuploidy. Dosage compensation of X-linked genes is achieved due to the X chromosome inactivation, which takes place at early development stages in all cells containing more than one X chromosome. X inactivation leads to repression of most genes in all except one of the X chromosomes in cell. After establishment, inactive state of X chromosome is stably inherited by all next cell generations (excepting germline cells that reactivate own X chromosome).

X inactivation depends on two partially overlapping genes on the X chromosome, *Xist* and *Tsix*, that encode lncRNAs acting in *cis* [2]. *Xist* RNA coats the same X chromosome, from which it is transcribed, and triggers silencing of this chromosome. In contrast, *Tsix* RNA is transcribed from active X chromosome and, being antisense to *Xist* transcript, counteracts its activity, also in *cis*. The key pluripotency transcription factors Nanog, Oct4 and Sox2 can be bound to *Xist* intron 1, resulting in restriction of the *Xist* expression in undifferentiated embryonic stem cells [64].

Recently investigations shown, that *Xist* and *Tsix* transcripts can be processed into xiRNAs – small RNAs that are 24-42 nt in length [2].

Molecular basis of X-inactivation. Our recent studies have shown that large non-coding RNA

Xist can predetermine the X-inactivation through the RNA-dependent DNA methylation [65].

Although the concentration of CG and CNG sites in *Xist* averages 0.8%, that is even less than the general genome level (1%), the localization of these sites reveals extreme irregularity. Of particular interest are: 1) area that includes the first 2000 nucleotides from 5′-end of the *Xist* RNA (within its first exon) and 2) small A region within this area between nucleotides 350 and 770 from 5′-end, where CG concentration exceeds the average level several times, reaching 3% and 5%, respectively. Also, CNG concentration reaches maximum density in this area and in the A region (5.7% and 7.6%, respectively).

In addition, both the X chromosome and other chromosomes contain numerous sites that are homologous to the sequences of the above-mentioned area of *Xist*.

These data support the hypothesis that the 5′-end sequences of the *Xist* RNA directly bind with complementary DNA of other genes, triggering the RNA-dependent DNA methylation [65]. The interaction between *Xist* RNA and target sequences should begin already during synthesis of the *Xist* transcript, before its separation from RNA polymerase, otherwise free *Xist* RNA undergoes folding and rapid cleavage (Fig. 5) [65]. This explains, how *Xist* causes the X-inactivation only in *cis*, as well as why inactive X chromosome is coated with *Xist* transcript. The fact, that not all genes contain sequences complementary to the *Xist* RNA, explains how these genes keep activity in the inactivated X chromosome.

On the active X chromosome, *Tsix* RNA causes *Xist* repression due to involving of the DNA methyltransferases that methylate the *Xist* promoter. It is shown, that this *Tsix*-mediated silencing of *Xist* and other steps in the X inactivation process are independent of the RNA interference mechanism [66]. Therefore, it seems that *Tsix* RNA triggers *Xist* repression again through the *cis*-interaction with DNA target, which initiates the RNA-dependent DNA methylation. This interaction should occur before *Tsix* transcript separation from RNA polymerase.

Because only one X chromosome remains active in cell, X-inactivation requires *Xist* expression from all alleles except one. Note that this situation is the opposite of the allelic exclusion. Therefore,

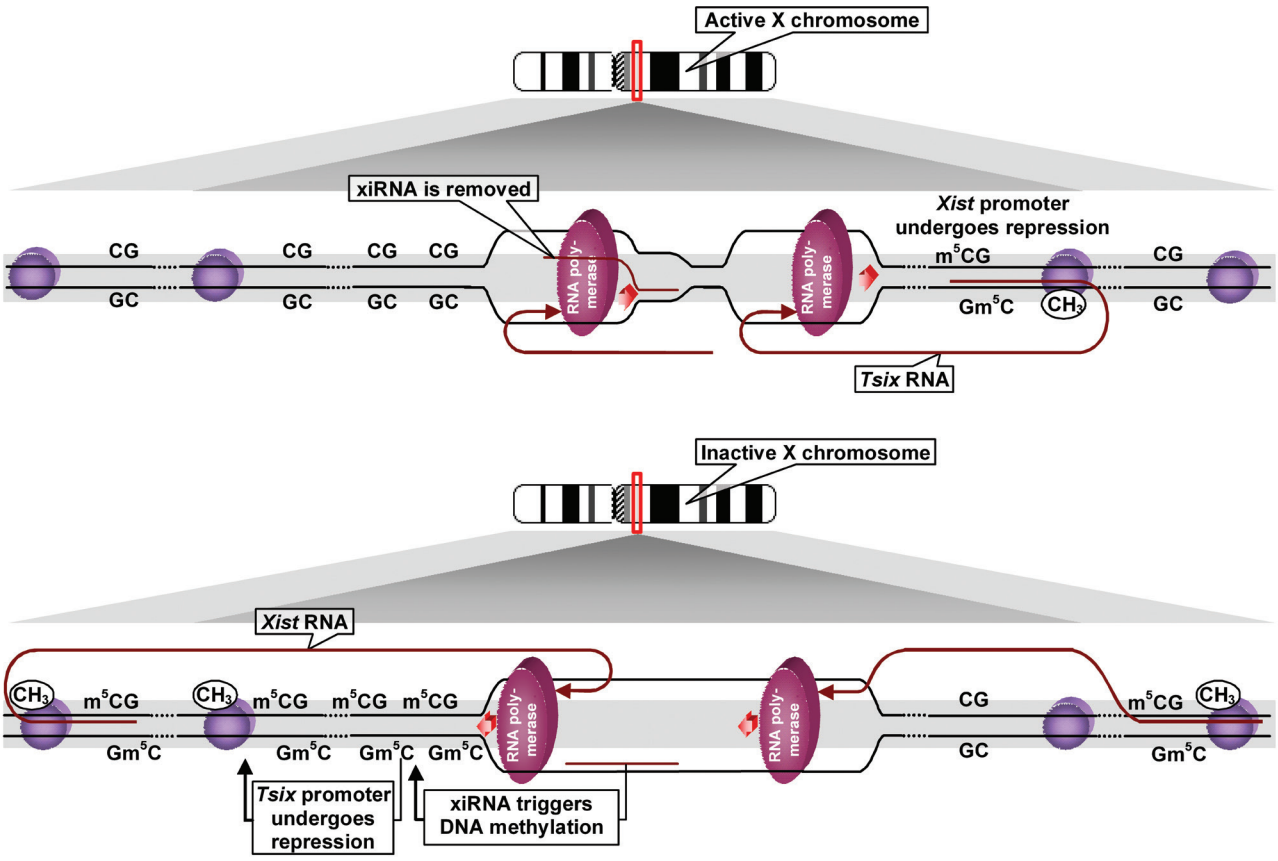

Fig. 5. Mechanism of X-inactivation

allelic exclusion must touch *Xist* repressors (i.e., *Tsix* or positive regulators of this gene) instead of *Xist*. Probably, xiRNAs are involved in this process. They can play the same role that miRNAs play in the allelic exclusion mechanism described above.

Frequency of CG and CNG sites in *Tsix* is even less than in *Xist* and amounts to 0.6%. The greatest peak of the concentration of these sites (2.4% and 8.2%, respectively) is localized near the region, overlapping *Xist* sequence, although this region contains other peaks. Perhaps because of this feature, the repression wave, triggered as a result of xiRNA binding in the bidirectionally transcribed region, should be directed toward the *Tsix* promoter. As a result, *Xist* remains active, whereas *Tsix* undergoes the allelic exclusion.

Interfering RNA can cause gene silencing due to direct hybridization with the target DNA sequences. This involves DNA methyltransferases in methylation *de novo* of cytosine within DNA targets. DNA methylation recruits histone deacetylases as well as histone methyltransferases, which cause chromatin remodeling and gene repression. Such events constitute molecular mechanism of the RNA-dependent DNA methylation. It is possible, that not only miRNAs and siRNAs, but also endo-siRNAs and piRNAs can cause gene silencing in that way.

Allelic exclusion is the ancient mechanism counteracting the activity of mobile genetic elements. The molecular basis of the allelic exclusion is that the miRNA (or another non-coding RNA), which precursor is transcribed from one allele, triggers DNA methylation *de novo* in all other alleles present in the cell. Immunoglobulin gene loci, which, as is known, undergo the allelic exclusion, contain genes of non-coding RNAs that are miRNAs or miRNA candidate.

The lack of interfering RNAs possibly predetermine the progressive DNA demethylation that allows reactivation of silent transposons and subsequent genome instability. Activated in response, the DNA repair systems should cause also illegitimate recombination in telomere caps, resulting in quick exhaustion of telomeres due to conversion of T-loops into T-circles. Therefore, the telomere length is integral indicator of genomic stability in normal cells. Consequently, ageing is a result of activity of mechanisms that causes apoptosis of cells that have a higher risk of transformation.

X-inactivation can be also mediated through the RNA-dependent DNA methylation. In this case, large non-coding RNA *Xist* should hybridize during own synthesis with DNA targets in remote regions of X chromosome. Obviously, the RNA-DNA interaction is universal mechanism triggering the DNA methylation *de novo* and the chromatin repression.

НЕКОДУЮЧІ РНК І ЕПІГЕНОМ: МЕТИЛУВАННЯ ДНК *DE NOVO*, АЛЕЛЬНЕ ВИКЛЮЧЕННЯ ТА Х-ІНАКТИВАЦІЯ

В. А. Галицький, С. В.Комісаренко

Інститут біохімії ім. О. В. Палладіна
НАН України, Київ;
e-mail: volha@biochem.kiev.ua

Некодуючі РНК є надзвичайно поширеним класом клітинних РНК. Вони беруть участь у багатьох важливих процесах у клітині – сигналінгу, посттранскрипційному сайленсингу, біосинтезі протеїну, сплайсингу, забезпеченні стабільності геному, подовженні теломер, Х-інактивації. Проте сфера активності цих РНК не обмежується посттранскрипційним рівнем, а охоплює також процеси, що змінюють епігенетичну інформацію чи відповідають за її збереження.

Некодуючі РНК можуть безпосередньо зв'язуватися з мішенями у ДНК, спричинюючи їх репресію через залучення ДНК-метилтрансфераз та хроматинмодифікуючих ензимів. Ці події є молекулярним механізмом РНК-залежного метилування ДНК. Імовірно, РНК–ДНК-взаємодія є універсальним механізмом ініціації метилування ДНК *de novo*.

Алельне виключення також може здійснюватись за допомогою зазначеного механізму. Це явище буде мати місце у тому разі, коли некодуюча РНК, попередник якої транскрибується з одного алеля, спричинює метилування ДНК усіх інших алелів, наявних у клітині. Слід зазначити, що мікроРНК-опосередкований транскрипційний сайленсинг теж нагадує алельне виключення, оскільки і гени мікроРНК, і гени, репресія яких може бути зумовлена цими мікроРНК, містять ділянки з одними і тими самими послідовностями. Можна припустити, що РНК-залежне метилування ДНК і алельне виключення виникли в процесі еволюції з метою протидії активності мобільних генетичних елементів.

Імовірно, перешкоди на шляху реактивації дрімаючих мобільних генетичних елементів знімаються внаслідок скорочення і дерегуляції набору клітинних некодуючих РНК, і це веде до нестабільності геному, котра зумовлює старіння та канцерогенез.

У процесі Х-інактивації, метилування ДНК та подальша гетерохроматинізація Х-хромосоми можуть бути спричинені прямою гібридизацією 5′-кінця довгої некодуючої РНК Xist із ДНК-мішенями у віддалених регіонах Х-хромосоми.

Ключові слова: епігеном, некодуюча РНК, мікроРНК, метилування ДНК *de novo*, алельне виключення, транспозон, сайленсинг, ремоделювання хроматину, Х-інактивація.

НЕКОДИРУЮЩИЕ РНК И ЭПИГЕНОМ: МЕТИЛИРОВАНИЕ ДНК *DE NOVO*, АЛЛЕЛЬНОЕ ИСКЛЮЧЕНИЕ И Х-ИНАКТИВАЦИЯ

В. А. Галицкий, С. В. Комисаренко

Институт биохимии им. А. В. Палладина
НАН Украины, Киев;
e-mail: volha@biochem.kiev.ua

Некодирующие РНК являются чрезвычайно распространенным классом клеточных РНК. Они участвуют во многих важных процессах в клетке – сигналинге, посттранскрипционном сайленсинге, биосинтезе протеина, сплайсинге, обеспечении стабильности генома, удлинении теломер, Х-инактивации. Однако сфера активности этих РНК не ограничивается посттранскрипционным уровнем, а охватывает также процессы, изменяющие эпигенетическую информацию или отвечающие за ее сохранение.

Некодирующие РНК могут непосредственно связываться с мишенями в ДНК, вызывая их репрессию путем привлечения ДНК-метилтрансфераз и хроматинмодифицирующих энзимов. Эти события являются молекулярным механизмом РНК-зависимого метилирования ДНК. Вероятно, РНК–ДНК-взаимодействие является универсальным механизмом инициации метилирования ДНК *de novo*.

Аллельное исключение также может осуществляться при помощи указанного механизма. Это явление будет иметь место в случае, когда некодирующая РНК, предшественник которой транскрибируется с одного аллеля, вызывает метилирование ДНК всех аллелей, имеющихся в клетке. Следует отметить, что микроРНК-опосредованный транскрипционный сайленсинг также напоминает аллельное исключение, поскольку и гены микроРНК, и гены, репрессия которых может быть вызвана этими микроРНК, содержат участки с одними и теми же последовательностями. Можно предположить, что РНК-зависимое метилирование ДНК и аллельное исключение возникли в ходе эволюции с целью противодействия активности мобильных генетических элементов.

Вероятно, препятствия на пути реактивации дремлющих мобильных генетических элементов снимаются из-за сокращения и дерегуляции набора клеточных некодирующих РНК, и это ведет к нестабильности генома, которая предопределяет старение и канцерогенез.

В процессе Х-инактивации, метилирование ДНК и дальнейшая гетерохроматинизация Х-хромосомы могут быть вызваны прямой гибридизацией 5′-конца длинной некодирующей РНК Xist с ДНК-мишенями в отдаленных регионах Х-хромосомы.

Ключевые слова: эпигеном, некодирующая РНК, микроРНК, метилирование ДНК *de novo*, аллельное исключение, транспозон, сайленсинг, ремоделирование хроматина, Х-инактивация.

1. *Sana J., Faltejskova P., Svoboda M. et al.* // J. Transl. Med. – 2012. – **10**. – P. 103.
2. *Ogawa Y., Sun B. K., Lee J. T.* // Science. – 2008 – **320**, N 5881. – P. 1336–1341.
3. *Kaikkonen M. U., Lam M. T., Glass C. K.* // Cardiovasc. Res. – 2011. – **90**, N 3. – P. 430–440.
4. *Seila A. C., Calabrese J. M., Levine S. S. et al.* // Science. – 2008. – **322**, N 5909. – P. 1849–1851.
5. *Hansen T. B., Jensen T. I., Clausen B. H. et al.* // Nature. – 2013. – **495**, N 7441. – P. 384–388.
6. *Esteller M.* // Nat. Rev. Genet. – 2011. – **12**, N 12. – P. 861–874.
7. *Halytskiy V.* // Eur. J. Cancer. – 2011. – **47**, Suppl., N 1. – P. 111.
8. *Halytskiy V.* // The Breast. – 2013. – **22**, Suppl., N 1. – P. S21.
9. *Bernstein E., Caudy A. A., Hammond S. M. et al.* // Nature. – 2001. – **409**, N 6818. – P. 363–366.
10. *Zamore P. D.* // Science. – 2002. – **296**, N 5571. – P. 1265–1269.
11. *Tijsterman M., Plasterk R. H. A.* // Cell. – 2004. – **117**, N 1. – P. 1–3.
12. *Rodriguez A., Griffiths-Jones S., Ashurst J. L. et al.* // Genome Res. – 2004. – **14**, N 10A. – P. 1902–1910.
13. *Lee Y., Ahn C., Han J. et al.* // Nature. – 2003. – **425**, N 6956. – P. 415–419.
14. *Lund E., Guttinger S., Calado A. et al.* // Science. – 2004. – 303, N 5654. – P. 95–98.
15. *He L., Hannon G. J.* // Nat. Rev. Genet. – 2004. – **5**, N 7. – P. 522–531.

16. *Friedman R. C., Farh K. K., Burge C. B. et al.* // Genome Res. – 2009. – **19**, N 1. – P. 92–105.
17. *Lewis B. P., Burge C. B., Bartel D. P.* // Cell. – 2005. – **120**, N 1. – P. 15–20.
18. *Kawasaki H., Taira K.* // Nature. – 2004. – **431**, N 7005. – P. 211–217.
19. *Jabbari K., Bernardi G.* // Gene. – 2004. – **333**. – P. 143–149.
20. *Scarano E., Iaccarino M., Grippo P. et al.* // Proc. Natl. Acad. Sci. USA. – 1967. – **57**, N 5. – P. 1394–1400.
21. *Halytskiy V. A.* // Acta Biochim. Pol. – 2007. – **54**, Suppl., N 2. – P. 5.3.
22. *Halytskiy V. A.* // Eur. J. Cancer Suppl. – 2007. – **5**, N 4. – P. 75.
23. *Halytskiy V. A., Komisarenko S. V.* // Biopolym. Cell. – 2011. – **27**, N 6. – P. 466–472.
24. *Галицкий В. А.* // Цитология. – 2008. – **50**, № 4. – С. 277–286.
25. *Halytskiy V. A.* // Ukr. Biochem. J. – 2009. – **81**, № 4 (special issue). – P. 162.
26. *Schmitz K. M., Mayer C., Postepska A. et al.* // Genes Dev. – 2010. – **24**, N 20. – P. 2264–2269.
27. *Bonasio R., Tu S., Reinberg D.* // Science. – 2010. – **330**, N 6004. – P. 612–616.
28. *Verdel A., Jia S., Gerber S. et al.* // Science. – 2004. – **303**, N 5658. – P. 672–676.
29. *Verdel A., Moazed D.* // FEBS Lett. – 2005. – **579**, N 26. – P. 5872–5878.
30. *Bühler M., Verdel A., Moazed D.* // Cell. – 2006. – **125**, N 5. – P. 873–886.
31. *Halytskiy V. A.* / XX International Congress of Genetics. Abstract book. – Berlin, Germany, 2008. – P. 119–120.
32. *Piriyapongsa J., Mariño-Ramírez L., Jordan I. K.* // Genetics. – 2007. – **176**, N 2. – P. 1323–1337.
33. *Soifer H. S., Zaragoza A., Peyvan M. et al.* // Nucleic Acids Res. – 2005. – **33**, N 3. – P. 846–856.
34. *Bourc'his D., Bestor T. H.* // Nature. – 2004. – **431**, N 7004. – P. 96–99.
35. *Jackson-Grusby L., Beard C., Possemato R. et al.* // Nat. Genet. – 2001. – **27**, N 1. – P. 31–39.
36. *Halytskiy V. A., Komisarenko S. V.* // Biopolym. Cell. – 2009. – **25**, N 1. – P. 12–26.
37. *Halytskiy V. A.* // Acta Biochim. Pol. – 2011. – **58**, Suppl., N 1. – P. 8.
38. *Lander E. S., Linton L. M., Birren B. et al.* // Nature. – 2001. – **409**, N 6822. – P. 860–921.
39. *Prak E. T., Kazazian H. H., Jr.* // Nat. Rev. Genet. – 2000. – **1**, N 2. – P. 134–144.
40. *Brouha B., Shustak J., Badge R. M. et al.* // Proc. Natl. Acad. Sci. USA. – 2003. – **100**, N 9. – P. 5280–5285.
41. *Pardue M. L., Rashkova S., Casacuberta E. et al.* // Chromosome Res. – 2005. – **13**, N 5. – P. 443–453.
42. *Curcio M. J., Belfort M.* // Proc. Natl. Acad. Sci. USA. – 2007. – **104**, N 22. – P. 9107–9108.
43. *Kazazian H. H., Jr.* // Curr. Opin. Genet. Dev. – 1998. – **8**, N 3. – P. 343–350.
44. *Collier L. S., Largaespada D. A.* // Genome Biol. – 2007. – **8**, Suppl., N 1. – P. S5.
45. *Ostertag E. M., Goodier J. L., Zhang Y. et al.* // Am. J. Hum. Genet. – 2003. – **73**, N 6. – P. 1444–1451.
46. *Iskow R. C., McCabe M. T., Mills R. E. et al.* // Cell. – 2010. – **141**, N 7. – P. 1253–1261.
47. *Jeffs A. R., Benjes S. M., Smith T. L. et al.* // Hum. Mol. Genet. – 1998. – **7**, N 5. – P. 767–776.
48. *Stoye J. P.* // Curr. Biol. – 2001. – **11**, N 22. – P. R914–916.
49. *Rowe H. M., Jakobsson J., Mesnard D. et al.* // Nature. – 2010. – **463**, N 7278. – P. 237–240.
50. *Murray V.* // Mutat. Res. – 1990. – **237**, N 2. – P. 59–63.
51. *Yoder J. A., Walsh C. P., Bestor T. H.* // Trends Genet. – 1997. – **13**, N 8. – P. 335–340.
52. *Tabara H., Sarkissian M., Kelly W. G. et al.* // Cell. – 1999. – **99**, N 2. – P. 123–132.
53. *Aravin A. A., Sachidanandam R., Bourc'his D. et al.* // Mol. Cell. – 2008. – **31**, N 6. – P. 785–799.
54. *Girard A., Hannon G. J.* // Trends Cell Biol. – 2008. – **18**, N 3. – P. 136–148.
55. *Malone C. D., Hannon G. J.* // Cell. – 2009. – **136**, N 4. – P. 656–668.
56. *Vanyushin B. F., Nemirovsky L. E., Klimenko V. V. et al.* // Gerontologia. – 1973. – **19**, N 3. – P. 138–152.
57. *Riggs A. D., Xiong Z.* // Proc. Natl. Acad. Sci. USA. – 2004. – **101**, N 1. – P. 4–5.
58. *Галицкий В. А.* // Цитология. – 2009. – **51**, № 5. – С. 388–397.
59. *Georgantas R. W., Hildreth R., Morisot S. et al.* // Proc. Natl. Acad. Sci. USA. – 2007. – **104**, N 8. – P. 2750–2755.
60. *Halytskiy V. A.* // Eur. J. Cancer Suppl. – 2010. – **8**, N 5. – P. 127.
61. *Kazazian H. H., Jr., Moran J. V.* // Nat. Genet. – 1998. – **19**, N 1. – P. 19–24.
62. *Wang K. C., Chang H. Y.* // Mol. Cell. – 2011. – **43**, N 6. – P. 904–914.

63. *Guil S., Esteller M.* // Nat. Struct. Mol. Biol. – 2012. – **19**, N 11. – P. 1068–1075.
64. *Navarro P., Oldfield A., Legoupi J. et al.* // Nature. – 2010. – **468**, N 7322. – P. 457–460.
65. *Halytskiy V.* / Joint Conference of HGM 2013 and 21st International Congress of Genetics "Genetics & Genomics of Global Health and Sustainability". Abstract Book. – Singapore, 2013. – P. 282–283.
66. *Nesterova T. B., Popova B. C., Cobb B. S. et al.* // Epigenetics Chromatin. – 2008. – **1**, N 1. – P. 2.