Biochemistry (Moscow), Vol. 70, No. 4, 2005, pp. 406-423. Translated from Biokhimiya, Vol. 70, No. 4, 2005, pp. 493-513.
Original Russian Text Copyright © 2005 by Tchurikov.

REVIEW

Molecular Mechanisms of Epigenetics

N. A. Tchurikov

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, 119991 Moscow, Russia; fax: (7-095) 135-1405; E-mail: tchurikov@eimb.ru

Received October 26, 2004
Revision received November 17, 2004

Abstract—The main epigenetic mechanisms in regulation of gene expression are discussed. The definition of epigenetics and its specific mechanisms including DNA methylation and gene imprinting, modifications of nucleosomal histones associated with silencing or activation of gene transcription, RNA interference, chromosomal silencing, and the role of mobile elements are discussed.

Key words: epigenetics, imprinting, histone modifications, chromosomal silencing, heterochromatin, PcG proteins, forum domains, RNA interference, siRNA, miRNA, RISC, RITS, mobile elements

Epigenetics is a relatively new field of molecular biology dealing with regulatory mechanisms of gene activity and inheritance that are independent of changes in the nucleotide sequence of DNA. Studies of recent years have shown that gene expression changes during differentiation and these changes are transmitted to daughter cells through mitosis. Similarly, changes in gene expression arising during ontogeny can be inherited through generations, i.e., transmitted through meiosis. In both cases these changes are not related with changes in DNA sequence. Thus, epigenetics includes two interrelated fields. One investigates hereditable mechanisms in regulation of gene expression during development, and the other elucidates mechanisms of hereditary transmission of this regulation for individual genes by germ line cells. In both cases the hereditable mechanisms that determine gene expression are not related with changes in the DNA text.

At first glance, it would seem that no genetic changes can occur without corresponding changes in the DNA sequence. In fact, from the time of the "central dogma" and discovery of the genetic code hereditable changes are known to be inevitably related with changes in the DNA text. However, recent studies forced us to admit the obvious fact that not only the DNA sequence is inherited, i.e., transmitted in the chain of cell divisions, but also the structures involved with it. The mode of gene expression can be encoded not only by the nucleotide sequence of DNA but also by chromosomal structures resulting on interaction of DNA duplexes with one another and also with proteins, RNA molecules, and, possibly, other macromolecules. Some of these structures are produced or changed as a result of DNA methylation or protein modifications. Consequently, the genetic code is the first fundamental level in the transmission of hereditary information. On the second level informational abilities of the hereditable structures are used, which include DNA, RNA, proteins, as well as chemical groups modifying nucleotides in DNA or amino acids in proteins. Therefore, this code was named as epigenetic [1]. During mitosis, cells transmit elements of the structural organization of gene expression and thereby the state of differentiation.

The progeny obtains from the parents not pure DNA but DNA as a part of a structured genome, and moreover, as a constituent of a fertilized cell that exists under strictly controlled conditions. Some programs responsible for reading the information encoded in DNA are in epigenomic structures, or the epigenome (an organism as a whole is an extreme example of an epigenome). These structures are retained during mitosis, play an important role in development, and are inherited in the set of cell generations. During differentiation, new structures appear and are retained, and this is how tissues and organs are developed from stem cells.

During meiosis epigenomic structures are largely erased, and this is an essential programmed change. However, some of these structures are conserved for still unclear reasons, and a new program of expression of some genes is inherited in the set of generations. This phenomenon is also named epigenetic inheritance or epigenetic memory. Thus, epigenetic inheritance is assured by mechanisms maintaining programs of development. And not only the genome but also the specific program of gene expression and cell development are inherited.

0006-2979/05/7004-0406 ©2005 Pleiades Publishing, Inc.

Mechanisms of epigenetic inheritance and epigenetic markers of silencing or activation of gene expression

No.	Epigenetic mechanism	Effect	References
1	DNA methylation	silencing	[1]
2	Lysine methylation in histones	silencing (H1; H3-K9, H3-K27, H4-K20)/transcriptional activation (H3-K4, H3-K79)	[31, 32, 34, 94]
3	Histone acetylation	transcriptional activation	[36]
4	Histone phosphorylation	transcriptional activation	[95]
5	RNA interference: RISC; siRNA, microRNA	silencing	[44]
6	Chromosomal silencing or gene activation via binding protein complexes (PcG, trx) on boundaries of forum domains	silencing/transcriptional activation	[83, 86, 98]
7	Chromosomal silencing with involvement of small RNAs: RITS	silencing	[103]
8	Transposition of mobile elements: insulators, promoters, enhancers	silencing/transcriptional activation	[109]

This review is not designed to exhaustively describe all data published on the epigenetic machineries. It is more attractive to set out critically the results illustrating major trends in epigenetics, with special consideration in more detail the mechanisms of RNA interference and chromosomal silencing, including some data of the author.

Epigenetics investigates mechanisms of inheritance which are not associated with changes in the gene sequence, but epigenetic mechanisms themselves are controlled by genes whose products encode the enzymes needed for DNA methylation, histone modifications, RNA regulatory apparatus, and also for numerous regulatory proteins which produce dynamical chromosomal structures, both local and lengthy. So far only a small fraction of genes which control mechanisms of epigenetic regulation have been identified. Therefore, review of the literature at this stage of our understanding of epigenetic inheritance can concern only a small part of the real epigenetic mechanisms. But it is still more interesting to propose and experimentally test hypotheses in this field of biology. And this is the reason for presenting in this review some unpublished data and hypotheses of the author.

The major known epigenetic mechanisms are listed in the table. Most of them are interrelated: thus, small RNAs can induce DNA methylation and histone modifications and, thus, change the local structure of chromosomes affecting the binding of structural and regulatory proteins to DNA.

DNA METHYLATION

Epigenetic inheritance, i.e., transmission of the gene expression state from parental into daughter cells often occurs through DNA methylation. This is a very economical regulatory machinery of gene expression. It involves introduction of epigenetic marks by methylation of specific cytosine positions in DNA (addition of methyl group into the carbon-5 position in the cytosine ring, 5mC). Usually cytosines of CpG dinucleotides are methylated. Some organisms, e.g., *C. elegans*, have virtually no methylated cytosines. Other organisms, such as *Drosophila*, have very little 5mC, and more often in CpT dinucleotides than in CpG [1]. Methylated areas occupy in DNA large domains separated by large domains of unmethylated DNA. Differences in the DNA methylation rate in different genomes seem to be associated with formation of epigenetic memory not only by methylation but also by other mechanisms [2].

The level of DNA methylation sharply decreases during early embryogenesis of mammals but later recovers due to so-called *de novo* methylation [3]. One of the most interesting features of the DNA methylation pattern is the existence of so-called CpG islands, or GC-enriched regions, which contain many unmethylated CpG dinucleotides in 5′-regions of various genes [1]. Some of them are methylated during the development and this leads to the stable silencing of these genes. Such methylation is a part of the developmental program and

also provides for the inactivation of one of the X chromosomes of females.

During DNA replication, a new chain (5′ CpG 3′) is synthesized with unmethylated dinucleotides located opposite methylated positions of the old DNA chain (5′ 5mCpG 3′). The DNA-methylating enzyme, DNMT1, methylates the new DNA chain only in those positions that were methylated earlier. Thus the methylation pattern is reproduced during the replication. However, experimental data have shown that the reproduction lacks strict precision, which is 95% per CpG site per cell division [4]. Nevertheless, even if methylation of a particular dinucleotide fails to be reproduced by this mechanism, the status of the CpG island methylation/demethylation will be reliably retained. This is also important for maintenance of allele-specific DNA methylation for imprinting [5].

And what is the origin of the methylation pattern, i.e., what determines the fate of a CpG-islet under consideration during *de novo* methylation? From one viewpoint, it depends on accessibility of DNA regions. Indeed, not all regions of the genome are equally accessible for DNA methyltransferases. Thus, tightly packed heterochromatin regions or inactivated X chromosome are not easily accessible. Transcriptional activators, which change local structures of chromatin, such as SNF2, the chromatin-remodeling factor, etc., affect the pattern of DNA methylation. From another viewpoint, the DNA sequence itself defines the choice of a target to be methylated. Methylation is extended around this center and is stopped by some barriers that delimit the methylated and unmethylated CpG islands [6]. Such centers are represented by DNA repeats, but it is still unclear whether they are a direct cause which determines the methylation or only a link in the chain of events. Moreover, unusual DNA structures, RNA interference, and cognate mechanisms that involve the recognition of RNA or DNA sequences by RNA molecules are also considered as putative triggers of methylation mechanisms [7].

Methylation of DNA is closely associated with gene silencing. However, it is not its cause but only its consequence. Thus, silencing of retroviral genes is observed two days after the infection, whereas regions of their promoters are methylated significantly later, after 15 days [8]. Similarly, silent genes in the inactivated X chromosome are methylated after the gene silencing and not before it. The methylation mechanism is suggested to recognize silent genes, and methylation of their promoter regions seems to be required for their irreversible inactivation in somatic cells, beginning from the development stage under consideration [1]. Protein complexes containing DNA methyltransferases can recognize modified nucleosomal histones H3-K9. The methylation prevents binding of transcriptional factors in the promoter regions and, by contrast, gives a signal for binding proteins that provide for more reliable gene repression.

In contrast, regions of unmethylated DNA are associated with actively expressed genes. Promoters of these genes do not undergo the *de novo* methylation because it is not associated with active chromatin structures that are involved in acetylation of nucleosomal histones or methylation of H3-K4.

The epigenetic memory ensured by DNA methylation is believed to be an evolutionary analog of another mechanism of inheritable cell memory, gene repression or activation with involvement of PcG–trx protein complexes [1]. These mechanisms will be discussed is a later section.

The inability to precisely reproduce methylated or unmethylated state of specific CpG islands changes gene expression and phenotype in plants and animals. But mechanisms responsible for maintaining or changing methylation in the locus of interest are still unclear in detail.

Sites of methylation are investigated by various techniques using bisulfite PCR. This elegant approach is based on earlier biochemical data on nucleotide modifications in tRNA [9]. This approach includes modifications of small amounts of DNA by bisulfite, which converts unmethylated cytosines to uracils. Primers chosen for PCR should have between them a CpG island, i.e., a GC-enriched area (GC > 50%) of 200-300 bp in length containing a sufficient number of CpG dinucleotides (the ratio between the number observed at random distribution and the expected number should be >0.6). Moreover, to be sufficiently different from primers of the initial DNA, texts of the modified DNA primers should have a sufficient number of cytosines. Depending on the approach used to analyze the PCR products, such as sequencing, restriction, or an MSP (Methylation-Specific PCR), the primers are chosen following additional rules [10, 11]. There are different techniques using sequencing of the product or analysis of symmetrical methylation positions in different DNA chains [12]. The latter technique includes ligation on the 3′-terminus of a restricted fragment of the genomic DNA of a synthetic single-stranded linker (Fig. 1), which links the complementary DNA chains of the CpG island under study. Upon denaturation of such a product, the island chains occur on different sides of the linker. The so-called pyrosequencing initially proposed for analysis of short DNAs or RNAs [13] is also applicable to analyze methylation of specific CpG positions. The approach also uses bisulfite PCR [14].

Imprinting. Imprinting is an introduction of epigenetic marks into the gene that results in its monoallelic expression, i.e., expression of the gene of one of the parents. Such a parent-dependent expression is a significant exception to classical genetics. This genomic imprinting is also often called gametic or parental imprinting. Its mechanism is unclear, but the parent-specific DNA methylation during the development of germ line cells is associat-

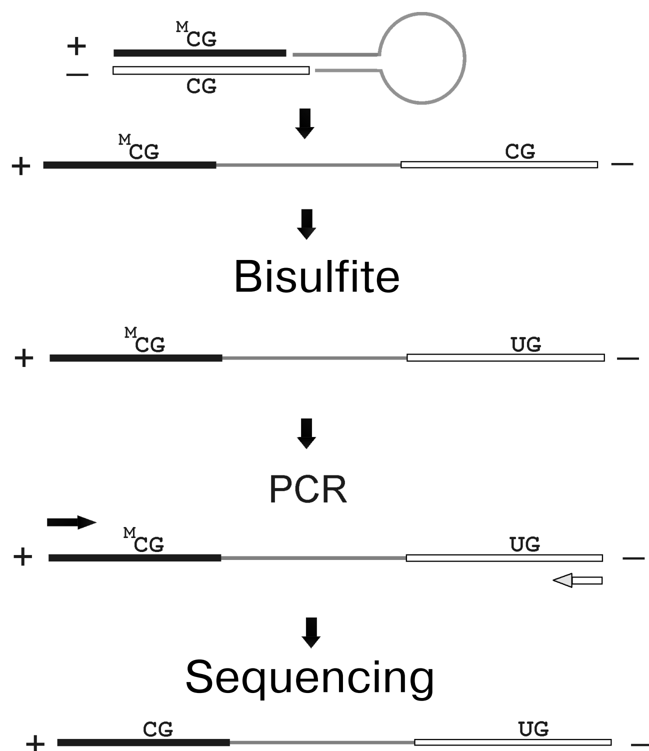

Fig. 1. Scheme illustrating analysis of symmetric methylated positions in complementary DNA by hairpin-bisulfite PCR [12]. In the initial DNA, a sticky end is produced on the 3′-terminus by digestion with a restriction enzyme. The hairpin-producing oligonucleotide has a similar sticky end. The sticky ends of the same type are annealed and ligated, then DNA is denatured and unmethylated cytosines are modified into uracils. Considering the modification, a pair of primers is chosen to amplify DNA. Cytosines in the PCR product correspond to the methylated positions.

ed with the parent-specific silencing of appropriate genes. During the organism's development, the imprinting (i.e., the acquired epigenetic marks) is retained in somatic cells and, conversely, erased in germ line cells [15]. In the course of gametogenesis, the imprinted genes retain the marks only of one parent, either of paternal or maternal allele. Imprinting prevents parthenogenesis in mammals. Disorders in imprinting can result in congenital defects of development. Some families of *de novo* methyltransferases are involved only in maternal imprinting, e.g., transferases of the Dnmt3 family [16].

Imprinted genes play an important role in development. They are usually located in large domains. The allelic repression of such a gene cluster is regulated by the ICR (Imprinting Control Region), which is a direct target for methylation. The ICR is methylated in one of the parental alleles. The ICR-surrounding DNA sequences and non-histone proteins are involved in this differential methylation during gametogenesis. The state of imprinting is maintained in somatic cells by mechanisms of DNA methylation and histone modification and by the *Polycomb* group proteins. Particular mechanisms of imprinting regulation by ICR vary in different domains. In some domains unmethylated ICRs act as insulators preventing interaction between promoter and enhancer. In other domains, small RNAs interact with the ICR region and trigger the binding of chromatin-modifying complexes [17]. The latter mechanism has similarities with the X chromosome inactivation.

An ICR can be located in gene introns. Thus, the locus Kcnq1 of mouse contains an ICR in the tenth intron. This locus and six adjacent genes are regulated by this ICR, which is methylated in the maternal allele and unmethylated in the paternal allele, thus manifesting the parent-dependent expression. The unmethylated ICR is related with the two-directional silencing and triggering of expression of the anti-sense RNA of the locus. And just this RNA was found to cause the silencing. However, in this case the silencing is not associated with RNA interference and has another character, because dsRNA corresponding to these genes fails to affect the activity of this ICR [18].

Specific mechanisms of gene silencing through DNA methylation are very different. Expression of exogenous transgenes in plants can result in epigenetic silencing of the homologous endogenes on the transcriptional level through DNA hypermethylation in the promoter region. Mechanisms initiating and maintaining DNA hypermethylation are studied better than those preventing methylation and permitting the retention of a high transcriptional level. In this regard, the ROS1 locus of *Arabidopsis*, which inhibits silencing, is of interest. This locus encodes DNA glycosylase/lyase, which provides for the reversibility of the transgene-induced silencing. The methylated DNA is glycosylated and cleaved in a sequence-specific way by the locus product, then the methylated base is excised by repair mechanisms, and the initial DNA text is recovered without methylated C [19]. Thus, the initial gene transcription is restored.

Epigenetic regulation is responsible for temporal, tissue-specific, and parent-dependent gene expression. The gene *eed*, the mouse *Polycomb* group gene provides for repression of the imprinted X chromosome. Damage to this gene reactivates the repressed X chromosome [1]. This finding suggests a relation between different mechanisms of epigenetic memory. Studies on the role of the *eed* gene in the regulation of autosomal imprinted loci have revealed that embryos with the $eed^{-/-}$ genotype (i.e., with the damaged genes in both chromosomes) contain transcripts of the repressed paternal alleles, notwithstanding the retained parent-dependent methylation in these embryos. Methylation was changed in specific islands, the so-called differentially methylated regions (DMR) of silent loci. Thus, the *eed* gene represents a new class of trans-factors, which regulate the parent-dependent expression of imprinted loci [20].

Hairpin-bisulfite PCR (see above) provides information on the extent of methylation in symmetric positions of complementary DNA chains. Obviously, the 5′-3′ CpG island in the complementary chain also has the corresponding sequence 5′-3′ CpG. Using this approach, the maintenance of the methylated state was studied in a CpG island of the *FMR1* gene in human lymphocytes [21]. In the hypermethylated CpG island of the gene in the inactivated X chromosome 96% methylation per site is retained per cell division. The efficiency of *de novo* methylation was about 17%. Note, that the same island in the active X chromosome is hypomethylated and retains only 1% of *de novo* methylation. Thus, the methylation state is maintained in the islands during many cell divisions.

Not only the mechanisms of reliable reproduction of epigenetic marks are important, but also the mechanisms of their recognition by proteins. An interesting example of such recognition is the chromatin protein CTCF, containing 11 DNA-recognizing zinc fingers. This protein binds to a particular ICR that possesses features of distant insulators. The binding of this protein depends on the state of DNA methylation in the ICR. Point nucleotide mutations in one of the ICR alleles disturb its methylation and the CTCF binding with DNA. But this occurs only if the ICR has undergone maternal meiosis [22]. Consequently, there are important factors responsible for both the recognition and reproduction of gametic epigenetic marks. The importance of recognition of epigenetic marks is also supported by data obtained on human and mouse homologous genes. The imprinted *Grb10* gene of mouse displays a maternal expression. A similar human gene, *GRB10*, has a biallelic expression. This difference in expression is caused not by varied methylation of CpG islands in these genes but by different recognition of similar epigenetic marks [23].

HISTONE MODIFICATIONS

Specific histone modifications are important epigenetic marks involved in inheritable changes of gene expression not affecting the DNA sequence. Dimethylation of lysine (K9) in the histone H3 is recognized by a chromodomain of HP1 protein. The binding of HP1, which is a product of the *SU(VAR)2-5* gene, recruits other proteins which change local chromosomal structures and gene expression. Amino acid residues of the HP1 chromodomain have been found which are involved in recognition of the methylated K9 in H3 [24].

The histone code hypothesis. The histone code hypothesis suggests that gene expression is regulated through modifications of histone tails. These modifications give rise to a number of specific binding sites for proteins that regulate gene expression. Thus, acetylation of histones results in binding of bromodomain-containing transcriptional activators, including the SWI/SWF chromatin-remodeling complex and the transcriptional factor TFIID. According to this hypothesis, the recruitment of each protein is caused by a specific histone modification. It was suggested that histone acetylation should result in generation of charged patches recognizable by the bromodomain rather than generating binding sites for particular proteins. A thorough analysis of histone acetylation patterns during activation of the human *IFN-β* gene has shown that specific sites of histone acetylation do recruit transcriptional activators in a certain order [25]. These data support the hypothesis.

Methylation of nucleosomal histones and gene silencing. Switching from acetylation to methylation of K9 in the H3 histone (H3-K9) is significant for chromosomal silencing because it is a link in the shutdown mechanism of gene expression. A microarray of human GpC islands can be used to detect targets of gene silencing. To do this, DNA probes prepared by chromatin immunoprecipitation (ChIP) are used. Such samples are prepared after cross-linking with formaldehyde of chromatin proteins in their *in vivo* locations. Then the DNA is subjected to fragmentation and precipitated with antibodies specific to some covalently bound proteins. The procedure results in a heterogeneous population of DNA fragments which consists of sequences interacting with the chosen chromatin protein. Such DNA samples can be amplified by PCR. Thus, the genes regulated by acetylation/methylation of H3-K9 were revealed with antibodies to methylated or acetylated H3-K9. Combined microarray, ChIP, and PCR approaches allowed isolation of such genes in the human genome [26].

Histone modifications are important for epigenetic silencing of ribosomal RNA (rDNA) genes. In cells with active metabolism about half of rDNA genes are epigenetically inactivated. A nucleolar remodeling complex (NoRC) containing a SWI/SNF-like regulator has been shown to provide for inheritance of an active or repressed state of these genes [27]. This complex is associated with the nuclear matrix and nuclear actin. The binding of the complex recruits methyltransferases and deacetylases. Histones are methylated and deacetylated in the region of rDNA gene promoters. This results in establishing silent structural state characteristic for heterochromatin. On binding in the region of rDNA gene promoters, NoRC interacts with the transcription-terminating factor (TTF-1). After the binding, NoRC displaces the nucleosome, which packs the promoter region to repress the activity of RNA polymerase I. The nucleosome is displaced due to its interaction with the N-terminus of the H4 histone [28]. Histone deacetylase (HDA6) is required to enhance the methylation induced by double-stranded RNA (dsRNA), i.e., gene expression is regulated by interrelated RNA signals and mechanisms of histone modification [29].

On the N-tail of the H3 histone, lysines 9 and 27 (K9 and K27) are methylation sites associated with the epige-

netic repression. These amino acid residues are located in the context with a protein motif ARKS. The proteins PcG and HP1 containing the so-called chromodomain bind with high selectivity to these motifs both *in vivo* and *in vitro* if their lysines are methylated. Nevertheless, these proteins are different in the pattern of their location on polytene chromosomes of *Drosophila*. The recombinant HP1 protein with the chromodomain of the PcG protein artificially installed instead of its own chromodomain can fully reproduce binding sites of the PcG protein. The opposite exchange of domains results in a similar effect, i.e., the chromodomains determine the location of these proteins in chromosomes. The structure of the PcG chromodomain bound with the H3 peptide containing trimethylated lysine 27 was compared by crystallography with the structure of chromodomain HP1 complex with the peptide containing the trimethylated lysine 9 [30]. In the first complex the methylated target of the H3 histone was recognized in the region of the ARKS motif and in the preceding region containing five amino acid residues. Along with the above-described methylation sites of nucleosomal histones, trimethylated H3-K9 and H4-K20 are silencing marks of human pericentromeric chromatin [31].

The nature of vernalization. Data that raise in memory some episodes of the recent history of Russian biology suggest that vernalization is associated with epigenetic silencing. It is known that for reliable blossoming various plants need a sufficiently long-term cooling. Such an exposure (in nature it is winter) decreases the content of the FLC protein, which prevents an out-of-time florescence of plants under conditions of middle and high latitudes. The FLC expression is decreased because of activation of genes *VRN1* and *VRN2* responsible for vernalization. The first of these genes encodes a DNA-binding protein, and the other is homologous to PcG proteins involved in gene silencing in animals. The vernalization, i.e., the long-term cold exposure, results in dimethylation of H3 histone (K9 and K27) in the FLC locus [32]. In *vrn2* mutants, K27 is not dimethylated, whereas in K9 dimethylation is lost only in double mutants, *vrn1* and *vrn2*. Thus, epigenetic memory of winter cold is realized through the histone code. This ancient epigenetic code is responsible for silencing in both plants and animals. Proteins PcG are also involved in cold imprinting of plants. One of the genes responsible for maintaining the cold memory for several weeks after short-term cold is a member of the PcG family [33].

Methylation of H3-K4 and H3-K79. Most data on histone methylation describe mono-, di-, and trimethylation of the histone H3 lysines in positions 9 and 27. These modifications are strictly related with uncovering of gene silencing. However, methylation in the same N-terminal tail of the H3 histone but at the lysine 4 is associated with transcriptional activation [34]. Covalent modifications of histones indicating sites of active transcription of euchromatic genes include hyperacetylation of the H3 and H4 histones, as well as hypermethylation of H3-K4 and H3-K79. Conversely, the corresponding positions in these histones are hypoacetylated and hypomethylated in inactive genes [35].

Histone acetylation. Histone acetyltransferases (HAT) are constituents of protein complexes specific to definite genes; they vary in the architecture of promoters and are located in different chromosomal contexts. These complexes contain subunits interacting with a TATA-binding protein and specific transcriptional factors in the promoter region [36]. The acetylation level of the H3 and H4 histones in yeast chromosomes in the promoter region was higher than in the encoding areas [37]. Damage to the *Eaf3* gene encoding a subunit in protein complexes of histone acetylase NuA4 and histone deacetylase Rpd3 dramatically changes the genomic profile of acetylation. Especially sharp changes occur in the acetylation pattern of histone H4.

HAT is involved in activation of gene expression with involvement of the so-called chromatin-remodeling complex SWI/SNF. This complex binds to gene promoter areas via sequence-specific transcriptional activators and uses the energy of ATP to induce local decondensation of chromosomal structures. Only after this the basal transcription apparatus can be bound. In the absence of activator, the complex can remain in the promoter region for a long time, but the presence of HAT is required. This is due to direct binding of the SWI/SNF complex to acetylated histones, recognizing them by the bromodomain, which is a conserved sequence specific for transcriptional activators. Thus, acetylation of nucleosomal histones in the promoter region is an epigenetic mark of transcriptional activation [38]. Mutation in the globular part of the H3 histone results in defects in the transcription. A careful analysis has shown this part of the histone is important for the SWI/SNF association with chromatin [39].

Ubiquitination of histones and other variants. Covalent histone modifications in the N-tails are important for gene expression in eukaryotes. Histone acetylation needed to activate expression of some genes occurs with involvement of acetyltransferases in strictly definite chromosomal regions. In yeast, some genes are also activated through modifications in the histone C-tail. Proteins that are to be degraded are usually marked with a short protein ubiquitin [40]. This process is named ubiquitination. By this mechanism the H2B histone is modified in the promoter regions of some inducible genes, and this activates gene expression [41]. But in this case the modification is not a “kiss of death”. The process is reversible, and the induction is followed by deubiquitination of this histone. Some details of this mechanism are also known. The reverse process occurs with involvement of the SAGA complex, which contains ubiquitin protease. A protein Rad6 catalyzes ubiquitination. The Rad6 asso-

ciation with promoters is provided for by E3 ligase and Gal4 activator.

Ubiquitination of the H2B histone, along with the earlier detected methylation of the H3 histone in the K4 and K79 positions, indicates regions of active chromatin. Thus, some sites of methylation are associated with silencing and others with activation (see table). This and other examples demonstrate that gene expression is regulated by various enzymes, which are responsible for a large spectrum of reactions with involvement of chemical groups, DNA, RNA, and protein molecules.

The state of chromatin can be marked not only by varied covalent histone modifications, but also by histone variants. Chromatin in the active gene area can be marked not only by methylation of the H3 histone in the K4 and/or K79 positions, but also by the presence of the H3.3 histone. The latter is associated with a replication-independent nucleosomal assembly. It is suggested that for nucleosomal assembly H3 and its variant H3.3 are used in replicative forks at actively transcribed genes [42].

RNA REGULATION

RNA interference (RNAi) is a powerful machinery of inheritable gene silencing unassociated with changes in the DNA text and triggered by formation of RNA–RNA duplexes (dsRNA) in the cell. In the first stage, the duplex is cut by an enzyme endoribonuclease, Dicer, which is a member of the RNase III family. The cleavage of the duplex by Dicer is accompanied by uptake of energy from ATP and results in production of dsRNA fragments 21-23 bp in length with 5′-phosphorylated and 3′-OH dinucleotide overhangs (Fig. 2). Such a fragment is termed small interfering RNA (siRNA) duplex. In the second stage siRNA fragments are transferred into protein complex RISC (RNA interfering silencing complex). A protein of this complex, the so-called Argonaute-2 (Ago-2), untwists the siRNA duplex using the energy of ATP. Ago-2 of *Drosophila* is homologous to proteins RDE-1 of *C. elegans* and Qde-2 of *Neurospora* from the so-called PPD family of proteins, which contain PAZ and Piwi domains. Upon separation of the siRNA chains, only one of them remains in the complex held by Argonaute. This single-stranded siRNA functions as a sequence-specific guide [43-47]. It is annealed with complementary sequences in other RNAs. Another endoribonuclease is involved in RNA interference. It is a constituent of RISC and degrades the mRNA target chosen by siRNA in the boundaries of complementary region forming 3′ dinucleotide overhangs. Therefore, products of RISC degradation are also of 21-23 nucleotides in length. The degradation of mRNA targeted by siRNA leads to silencing, i.e., to switching off of the gene expression on the post-transcriptional stage. Thus, if any mRNA is involved in production of dsRNA, the cell triggers a mechanism of its specific degradation of the given mRNA.

The ability for complementary recognition is the main feature of nucleic acid molecules, and this provides for the specificity of RNA interference. At the dawn of molecular biology the discovery of complementarity of nucleic acids easily solved the mysterious mechanism of copying genetic information. The same property of nucleic acids is used by the cell for a strictly selective mRNA "weeding" regulating gene activities.

In 1998 experiments on *C. elegans* resulted in the discovery of RNA interference, and it was suggested that siRNA should be amplified inside the cell [43, 47]. Studies on plants have shown that siRNA can be transferred from one part of the plant to another. In fact, even trace amounts of siRNA can be involved in degradation of excess mRNA targets far from the location of siRNA injection. Later, studies on plants and fungi confirmed that the RNA-dependent RNA polymerase (RdRP) can use siRNA as a primer for synthesis of long dsRNA on the mRNA target. This dsRNA can be degraded by Dicer and provide for complexes with new siRNAs able to degrade the given mRNA. But experiments on human and *Drosophila* failed to confirm these expectations.

Some sensors are suggested to exist that are capable of recognizing aberrant transcripts, which are to be degraded earlier, even before generation of dsRNA. Thus, aberrant single-stranded RNAs can be recognized by such a sensor, and this triggers initially the conversion of such RNAs to double-stranded RNAs due to synthesis on them of complementary RNAs and, later, the RNA interference itself. Uncompleted transcripts, products of splicing errors, molecules deprived of the poly(A) tail or incapable of translation, can be recognized as aberrant. RdRP seems to be a candidate for being such a sensor.

Thus, RNA interference, which is a powerful and ancient mechanism, has two important functions. First, it is responsible for physiological regulation of gene activities and, second, for defense [46]. In fact, viral infection and expansion of mobile elements are accompanied by appearance of foreign RNAs parasitizing the transcriptional and translational apparatuses of the cell. That is why RNA interference is also believed to be the most ancient immunity, immunity on the level of RNA molecules. This immunity can recognize foreign molecules and eliminate them.

RNA interference with siRNA as the recognizing molecule occurs not only post-transcriptionally. Thus, generation of dsRNA in plants can be also accompanied by transcriptional silencing, which uses the above-considered DNA methylation mechanisms in the promoter region indicating the sites for transcriptional repression. Moreover, the mechanism of transcriptional (i.e., promoter) silencing was also found to convert dsRNA into siRNA. There are also genetic findings that suggest still more extensive use of recognitions with involvement of

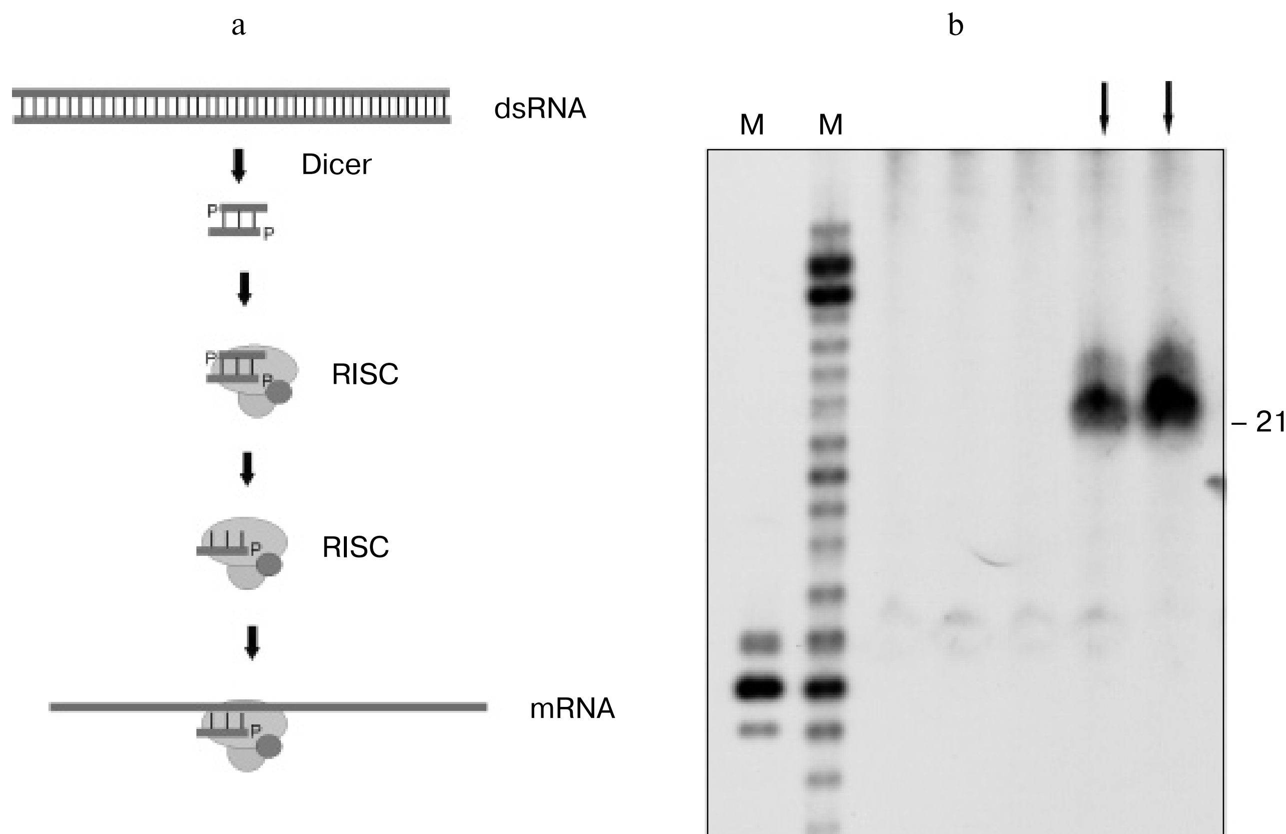

Fig. 2. Stages of RNA interference and electrophoretic siRNA pattern. a) In the first stage of RNA interference, dsRNA is recognized and cut by Dicer onto fragments of 21-23 bp in length. The resulting "dice" are recognized by the protein complex RISC. The protein Ago2 of the complex untwists RNA chains and holds one of them. Then the complex guided by siRNA recognizes the complementary sequence in the corresponding mRNA and cuts it, excising the fragment complementary to siRNA. b) The pattern of siRNA fragments in denaturing polyacrylamide gel. Arrows indicate the lanes containing fractionated products of RNA interference. M, marker (RNA). The position which corresponds to 21 nucleotides is shown.

RNA molecules. In experiments on *Neurospora* a mechanism was detected which supervised not only RNA molecules but also unpaired DNA during meiosis. Meiotic silencing by unpaired DNA (MSUD) inhibited expression of genes in the regions containing in the diploid phase unpaired sister chromosomes. The genome-controlling system interpreted such unpaired chromosomes as foreigners capable of damaging the cell. It was suggested that a mechanism should exist providing for synthesis of dsRNA on such genes and triggering the silencing with involvement of siRNA molecules. It has also been shown that proteins involved in RNA interference are associated with chromatin. Possibly, siRNA not only finds the mRNAs to be degraded but can also detect suitable sequences in chromosomes, and all this is due to the ability of nucleic acid molecules for complementary recognition.

siRNAs are suggested to recruit a typical component of heterochromatin, the HP1 protein, into both heterochromatin regions and puff areas, which manifest active transcription [48, 49]. The recruitment is indirect and includes targeting of histone-modifying enzymes at specific complementary siRNA sequences. As a result, the chromodomain of HP1 specifically binds to modified nucleosomal histones. Mutations in the chromodomain disturb the recruitment of HP1. But as HP1 is recruited into the areas of active transcription, it can be considered as a multifunctional component of chromosomes.

In plants and yeasts, RNA interference is involved in production of the silent chromatin structures, i.e., in the chromosomal silencing which results in transcriptional repression [50]. The RNA interference occurs with involvement of sequences homologous to the corresponding chromosomal region. This type of silencing is also called transcriptional. Mechanisms of RNA interference are suggested to be also used in production of centromeric heterochromatin.

Small RNAs is a universal recognition tool applied in various cell machineries. Recently a mechanism similar to RNA interference has been found to be used in elimi-

nation of DNA during formation of the *Tetrahymena* somatic macronucleus, which is an important stage in the development of this organism. Sequences of the parental macronucleus epigenetically control rearrangements during formation of the new macronucleus. This is the first indication that RNA regulation can directly change the organization of DNA sequences [51]. Elimination of DNA in *Tetrahymena* is accompanied by histone modification and accumulation of chromodomain-containing proteins, which mark regions of heterochromatin for its subsequent elimination.

Over some decades geneticists have been going from character to gene, but the so-called reverse genetics have just the opposite problem: to elucidate functions of numerous genes in the sequenced genomes. siRNAs is a powerful tool for artificial gene silencing used as a new approach in reverse genetics. Switching off genes by RNA interference is now widely used to elucidate the role of various encoding sequences. Moreover, RNA interference is promising in gene therapy for switching off crucial genes involved in the development of a disease. Unlike labor-consuming gene knockout, this approach does not damage the gene itself but only changes its regulation. To solve such problems of RNA technology, it is necessary to reveal the possible role of the nucleotide text promoting the most effective degradation of the mRNA target. Careful analysis of various siRNAs has resulted in some rules for choosing the sequence to attack mRNA. According to these rules, the middle of the encoding region of mRNA or its 5′-noncoding areas are not good targets, whereas the most effective is degradation of mRNA fragments which contain G or C bases in position 11 and T base in position 19. It is also important to concurrently use 4-5 siRNAs [52]. Moreover, siRNAs should have a low G/C composition and lack inverted repeats [53].

MicroRNA. The gene *let-7* was the first one found of a large family of genes which encode RNAs providing not for protein synthesis, but for special regulatory mechanisms responsible for expression of many genes in eukaryotes. RNAs of such genes produce an imperfect hairpin, which is recognized and cut by Dicer onto short fragments of 21-25 bp in length, or microRNAs (miRNA(s)). In *C. elegans*, miRNA of the *let-7* gene can bind to six sites of the 3′ non-coding sequence (3′ UTS) of the *lin-41* gene mRNA area and induce its degradation in sites of partially complementary RNA–RNA duplex. Disorders in the function of 3′ UTS prevent expression of the *lin-41* gene on the translational level [54]. Careful studies on RNA–RNA duplexes of the given mRNA and *let-7* miRNA revealed two conservative regions in 3′ UTS. They were termed *let-7* complementary sites (LCS). Note that the sequence located between two LCS of 27 bp in length is also necessary for gene silencing. Mutations in LCS affecting the complementation with *let-7* miRNA can partially recover the function of 3′ UTS.

MicroRNAs are signal molecules that play a crucial role in development. A number of homeodomain-containing genes, which are transcriptional factors, are regulated by miRNAs [55], and the regulatory mechanism is determined by the extent of the miRNA–mRNA complementation. In animals, miRNAs are cut by enzymes Drosha and Dicer onto incomplete miRNA–miRNA duplexes [56]. Different RISCs vary in functions and contain different variants of Argonaute, which directly bind to miRNA. It is reasonably to think that mechanisms of RNA silencing are essential for both functioning of the eukaryotic cell and the development of the organism [57, 58].

Regulation of RNA interference. It should be noted that RNA interference itself is regulated during development. This is shown by studies on RNA interference involving mobile elements of *Drosophila*. The non-LTR F-element has homologous a short retro-element, suffix, which has both sense and antisense transcripts at all stages of *Drosophila* development. But suffix-specific siRNAs were detected among products of the expected RNA interference only at a single stage. This suggests that some factors inhibit RNA interference at other stages of the development [59]. Recently it was shown that the pattern of siRNA expression also changes during both development and physiological processes [56]. Numerous tissue-specific miRNAs were found using a microarray containing 245 human and mouse miRNAs [60].

Arguments in favor of regulation of RNA interference were also provided by data on RNA interference inhibition in plants. Viruses can virtually always are able to escape defense mechanisms of the host. A similar new example is presented by suppression of RNA interference in plants. Viruses have been found which encode the proteins capable of binding to RNA and inhibiting RNA interference [61].

Role of parallel RNA–RNA duplexes. Nucleic acid molecules can produce parallel duplexes. A family of symmetric RNA sequences capable of producing antiparallel and parallel duplexes is presented in Fig. 3. Expression of a small RNA complementary in parallel orientation to the *lon* gene mRNA in *E. coli* resulted in a powerful gene-specific *in vivo* silencing [62] that seemed to be associated with translational inhibition because it was not accompanied by degradation of the *lon* gene mRNA. These investigations also showed that the gene-specific silencing caused either by expression of the antisense RNA or by parallel complementary RNA gradually decreases by the 50th generation. And this is not associated with mutations in the constructs expressing small RNAs, because DNA of the constructs isolated from the cells from 50-60 generations could re-induce powerful silencing in new cells. This indicates the existence of plastic cellular regulatory mechanisms of silencing, and this has to be taken into account in bioengineering.

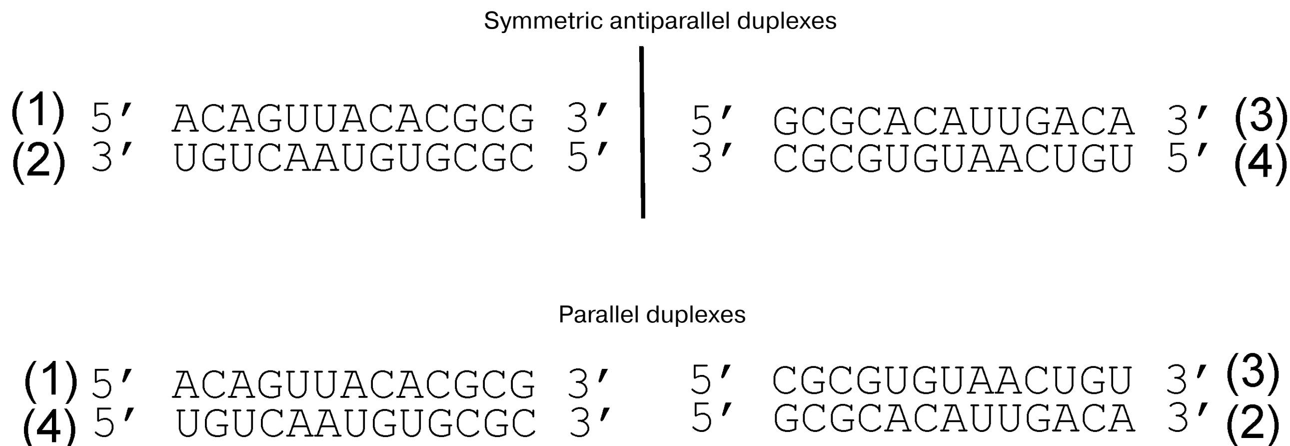

Fig. 3. Family of symmetric RNA sequences capable of producing antiparallel or parallel duplexes [62]. Four different sequences in the 5′-3′ direction are shown.

Injections into *Drosophila* embryos of excess RNAs complementary in the parallel direction to mRNA of the gene *Kruppel*, which controls formation of anterior segments in larvae, resulted in formation of only a part of these segments, i.e., to development of Kr phenocopies [63]. The silencing mechanism in this case is suggested to be associated with generation of parallel RNA–RNA duplexes between mRNA of the gene and the injected RNA. The parallel duplexes seem to be recognized by an extremely sensitive system responsible for monitoring RNA molecules of the cell. The cellular sensor triggers the specific degradation of mRNA involved in the generation of such duplexes. It is unclear whether “parallel silencing” can occur during *in vivo* expression of such small RNAs or *Drosophila* and in human cells.

CHROMOSOMAL SILENCING

Chromosomal compartments. There are two kinds of structures in eukaryotic chromosomes: large regions of transcriptionally active decondensed chromatin (euchromatin) and large regions of transcriptionally inactive condensed chromatin (heterochromatin (HCh)). These structural–functional chromosomal compartments were discovered in cytogenetic studies [64]. Later it was found that euchromatic genes transferred by inversions into HC regions became epigenetically inactivated (position effect variegation). Thus HC was defined genetically [65]. Recently molecular definitions for HC were suggested. It was shown that this type of chromosomal structure is determined by both DNA sequence and the attached proteins and their modifications, or by small RNA molecules in protein complexes. Heterochromatin regions are specified by methylation of the H3-K9 and H3-K27. As it has been said, these H3 modifications are recognized by chromodomain of the HP1 protein, and later other proteins are bound that leads to a condensed state of the corresponding chromosomal region.

Generation of HC and inactivation of euchromatic regions could be determined by common mechanisms. Such is inactivation of the X chromosome in mammals. It seems that the paternal or maternal chromosome is chosen at random, but then this choice is maintained epigenetically [66]. The above-mentioned position effect is also exemplified by heterochromatization of genes located in euchromatin.

Regions of HC are poor in genes and contain numerous repeats, such as satellites, mobile elements, etc. [65]. Heterochromatin is specified by the late replication in the S-phase and a decreased incidence of meiotic recombinations. It is still unclear what DNA sequences and proteins are involved in the initial stage of heterochromatization in the germ line and somatic cells. It is also unclear which DNA sequences and proteins restrict the expansion of heterochromatin structures at their boundaries within euchromatin regions. However, formation of heterochromatin seems to be associated not with some specific DNA sequences, but rather with different repeats. Aberrant transcripts of these repeats are likely to trigger RNA interference, and small RNAs can target chromatin modifiers (e.g., methyltransferases) at given chromosomal regions [29]. There are only preliminary data on features of the barrier between HC and euchromatin. Thus, methylation of H3-K4 and inverted repeats are found in yeast on the borderline of the HC region [50]. Deletions of the inverted repeats promote the expansion of HC, which is accompanied by the methylation of H3-K9, binding of Swi6 (an analog of *Drosophila*’s HP1), and demethylation of H3-K4. The locus of chicken β-globin is also surrounded by a barrier containing methylated H3-K4 [67].

In addition to large regions, HC is also represented by its small islands distributed in chromosomes. In areas of this so-called intercalary HC [68], mobile elements are found, which are also typical for large regions of HC [69]. The role of HC islands in the epigenetic regulation of gene expression will be considered below.

Expression of the hormone-induced reporter gene located in HC of *Drosophila* began to be repressed at the gastrula stage, approximately 1 h after appearance of the cytologically detectable HC.

A strong repression was observed until the third stage of larval development, and it was weakened during further development. These data show that HC activity is regulated during development [70]. Heterochromatin has several functions: centromeric, gene silencing, and nucleus organization. The condensed state of pericentromeric HC is associated with a regular nucleosomal packaging, which partially depends on satellite repeats. Methylation of the H3 histone N-tail is an epigenetic mark, which, on one hand, upsets acetylation and phosphorylation of histone tails but, on the other hand, acts as a signal for HP1 recognition. The choice of either the activated state or silencing of heterochromatic genes seems to depend on the balance between factors of high order chromosomal structures and transcriptional factors that bind to regulatory sequences and activate gene expression [71].

The HP1 of *Drosophila* is a key component of HC needed for the classic position effect exemplified by stable genetic silencing. Less is known about the family of human HP1 proteins, which can also be located in euchromatin. In this case, they are recruited in the expressing loci after association with them of special repressors and are held by methylated H3-K9. The transgene model suggested that HP1 binding leads to epigenetically inheritable silencing similar to that observed in the *Drosophila* position effect [72].

Role of DNA sequences in formation of heterochromatin. For a long time telomeric repeats were believed to be necessary and sufficient *cis*-elements to provide for functioning of telomeres. Thus, ring chromosomes of yeast lacking telomeric repeats retained structure and meiotic function. The Tazl proteins, which usually bind to telomeric repeats, and heterochromatic Swi6 (an analog of HP1 of *Drosophila*) in the absence of repeats, bind to subtelomeric regions [73]. During the pre-meiotic stage, these chromosomes interact with the spindle pole body (SPB), which is an equivalent of a centrosome in yeast. The Tazl binding by subtelomeric regions and interaction of these chromosomes with SPB are inheritable, but the character of this inheritance is achieved via different strategies. Taz1 is recruited by specific sequence in the subtelomeric regions, i.e., by a genetic mechanism, whereas interaction with the centrosome is not related to DNA sequences, i.e., is epigenetic. Functional chromosomal domains are huge highly organized molecular complexes that include many proteins and a definite DNA sequence. These complexes, in particular, have a great store of reliability for retaining structure and function. Nevertheless, when such a complex is destroyed because of absence of a crucial constituent, it cannot be recovered by subsequent addition of this constituent. In fact, deletion of the *taz* gene in the cells by ring chromosomes disturbs their interaction with SPB, and the subsequent addition of the operating gene does not recover this interaction. Thus, the reliability store of telomeric complexes is provided for by both genetic and epigenetic mechanisms. Earlier, epigenetic mechanisms were shown to be important for maintaining the centromere function in *Drosophila* [74].

Role of HP1 in formation of heterochromatin. HP1 is an indispensable structural component of silent chromatin in telomeres and centromeres. The position effect is associated with extension of heterochromatization of these regions onto euchromatic genes. HP1 is bound by the sites of modifications of nucleosomal histones H3-K9. This methylation is realized by the methyltransferase SU(VAR)3-9. During the position effect, HP1 recruits this enzyme, and this enlarges the region of histone methylation onto adjacent nucleosomes and, consequently, HP1 is bound along the chromatin thread. Recently, the HP1 binding was shown to induce the extension of gene silencing in both directions. HP1 was fused with the DNA-binding domain of the *E. coli* lacI repressor. The construct was expressed in a *Drosophila* strain containing two reporter genes controlled by the heat shock promoter and located 1.9 and 3.7 kb downstream from repeats of the lac-operator. The initial recruitment of HP1 in this experiment was determined not by methylation of the target nucleosomes but by binding of lacI–HP1 to the lac-operator repeats. This binding at the distance of 1.9 and 3.7 kb upstream of the reporter genes caused their silencing. The binding of nucleosomes in this region suggested the extension of the HP1 binding in both directions from the binding sites of the chimeric protein lacI–HP1. In the mutant Su(var)3-9, the first reporter gene remained silent, whereas silencing of the other gene located 3.7 kb downstream was terminated. Thus, the HP1 silencing not always depends on the methyltransferase SU(VAR)3-9 [75].

Proteins PcG and trx. A hereditable repressed state of chromosomal regions arises in the course of development, in particular, when it is necessary to switch off expression of homeotic genes during the later stages of development [76]. This type of silencing needs *Polycomb* group (PcG) proteins [77], which produce at least two types of protein complexes. One of the complexes has histone deacetylase activity and inhibits ATP-dependent remodeling of local chromosomal structures [78-80]. Recently a relation was found between the silencing caused by binding of PcG complexes and the H3-K27 methylation [81, 82].

In contrast, proteins of the trithorax group (trx) can activate transcription [83]. In addition to regulation of

homeotic genes, protein complexes PcG and trx also regulate activities of other genes essential for development [84, 85].

These complexes, both activating and repressive, bind to specific DNA sequences, the so-called *Polycomb* response element (PRE) or trithorax response element (TRE) sequences. These sequences are needed for silencing or activation of adjacent genes [86, 87]. Among PcG proteins, only the pleiohomeotic protein (PHO) is known to directly bind to DNA in the PRE region. Two trx proteins, GAGA factor and Zeste, are known to directly bind to TRE [83, 88]. The proteins of the complexes were named based on genetic studies on *Drosophila*. Later analogs of these proteins have been found in all eukaryotes studied. The complexes of PcG repressors and trx activators are suggested to be formed directly on PRE or TRE sequences. Upon the binding of PHO, Zeste, or GAGA proteins, other components of the complex are recruited [89]. Investigation of this process in detail has shown that the complex containing the GAGA factor initially displaces nucleosomes with involvement of two other protein complexes that promote the passage of RNA polymerase II across nucleosomes. This leads to accessibility of the DNA-recognizing site. Only afterwards the GAGA-containing complex interacts with transcriptional complexes or recruits other proteins needed for transcriptional activation or silencing.

Different PcG and trx complexes bind *in vitro* to GAGA motifs in some PREs. The GAGA factor is present in both repressive and activating complexes. Binding to PRE, trx complexes stimulate transcription and prevent recovery to the repressed state. A decrease in the concentration of trx complexes in some regions of *Drosophila* imaginal disks is accompanied by recovery of silencing. Thus, the binding of trx complexes does not remove epigenetic marks of silencing. In some cases, trx complexes coactivate the expression of operating genes but cannot initiate their transcription. Association of histone acetylase with PRE during embryogenesis fails to induce silencing but results in the loss of epigenetic marks in the imaginal disks, i.e., deacetylation of histones is necessary for appearance of epigenetic markers of the repressed state [90].

The genes encoding the *Polycomb* group proteins repress a number of key genes in development. In this connection the human gene *CBX7* is of interest. Primary human cultured cells, unlike human embryonic stem or tumor cells, have a restricted lifespan. After a definite number of divisions, they enter an irreversible growth-arresting state termed replicative senescence. Detection of genes responsible for immortalization is very important for oncology. Analysis of cDNAs specific for the cells in the replicative senescence phase resulted in isolation of the gene *CBX7* encoding a PcG protein. This gene prolongs the lifespan of cells repressing the Ink4a/Arf locus expression. Functioning of the *CBX7* gene is repressed in replicative senescence. Switching off the gene by artificial RNA interference through expression of the specific siRNA activates the Ink4a/Arf locus, and this inhibits growth of normal cells [91]. Thus, epigenetic mechanisms can control the lifespan of cells.

In *Drosophila*, PcG proteins produce two complexes, the *Polycomb* repressive complex 1 (PRC1) and ESC–E(Z). The biochemical function of the majority of PcG proteins is unclear. Seven genes of this group positively or negatively regulate the transcription of one another, and this makes it difficult to analyze them genetically [92]. The transcriptional repressor SCM of *Drosophila* is a PcG protein and a constituent of protein complexes. Similarly to another protein of the same group, polyhomeotic (PH), it has the SPM domain. The proteins PH and SCM bind to each other through these domains. PH is a core component of the repressive complex PRC1, which also includes SCM [93]. In human cells, the PcG 2 and 3 complexes contain histone lysine methyltransferase (HLMT), which is active toward histones H1 and H3 [94].

The trx protein ASH2 as a constituent of large protein complexes activates development-regulating genes including homeotic genes. In these complexes, ASH2 directly binds to the protein Skittles, which is an enzyme phosphorylating histone H1. These histone modifications can activate transcription [95].

Confocal *in vivo* microscopy has shown that cells containing GFP fused with 21 different nuclear proteins can find their binding sites by scanning three-dimensional chromatin structures, and these features are essential for plasticity of genome expression [96]. The majority of chromatin proteins have a high turnover, with binding times of several seconds. The major part of such a protein is bound, and usually the proteins transiently bind to a given site.

Paramutation. The epigenetic phenomenon of paramutation is figuratively called "chromatin conversation". Combination of two alleles of the same gene results in a hereditary change in the expression of one of these alleles [97]. This so-called trans-sensing occurs between the endogene and transgene. Paramutations are associated with changes in the DNA methylation or with changes in hypersensitivity to nucleases. For the trans-sensing tandem repeats are required. Mutations that disturb paramutations also affect silencing of mobile elements. Some trans-sensing phenomena have been described that involve homologous sequences.

Forum domains. It was mentioned above that epigenetic marks are inheritable changes in gene expression not involving the nucleotide sequence of DNA. However, the marks themselves appear and are erased as a result of activities of genes that conserve unchanged sequences during epigenetic processes. Functions of some genes are associated solely with epigenetic regulatory mechanisms. Consequently, there is no reason for epigenetics to be

contrasted with genetics. For epigenetics DNA is a groundwork, canvas, or a pure paper sheet that can be variously written on. A genome consists of a variety of genes, and some of them are essentially involved in epigenomic mechanisms of inheritance. Moreover, the generation of chromosomes and their functional domains, as well as fundamental principles of their organization, is closely related with epigenetic regulation. A kind of epigenome imprint is also present in the organization of chromosomal DNA.

Domain structure is a basic principle in chromosomal organization that provides for silencing of a gene complex in the domain. Eukaryotic chromosomes consist of consecutive domains arranged along their axis, the so-called forum domains which are units of chromosomal silencing. These domains have the same length of 50-200 kb in chromosomes of various eukaryotes and are easily excised from the chromosomes during their spontaneous degradation (Fig. 4). Binding sites for the proteins involved in chromosomal silencing, such as PcG, trx, etc., are located on the domain boundaries where chromosomes are very brittle and sites of breaks between forum domains are dispersed. Such an organization of chromosomes promotes switching off during the differentiation of large gene cassettes or large gene loci located in the same domain. Boundary regions (BR) between the neighboring forum domains correspond to small islands of intercalary HC [98]. Thus, molecularly the latter could be defined as the genome regions binding with the PcG and trx proteins, which are responsible for the distant regulation of gene activity. It is very likely that heterochromatization of these regions is similar to that in large compartments of HC [99]. This type of regulation needs a relatively homogenous dispersion of BR, i.e., the regions of intercalary HC in chromosomes. The PRE and TRE sequences are likely to be located in BRs. BR include mobile elements. It is yet to be explained why these regions of chromosomes are the most brittle. It is also unclear whether all putative BR are functioning in the cell or are activated during development. Cytogenetic data suggest that the regions of intercalary HC are replicated late and are involved in ectopic pairing of chromosomes. Chromosomal threads connecting these regions in different chromosomes have been described earlier [100].

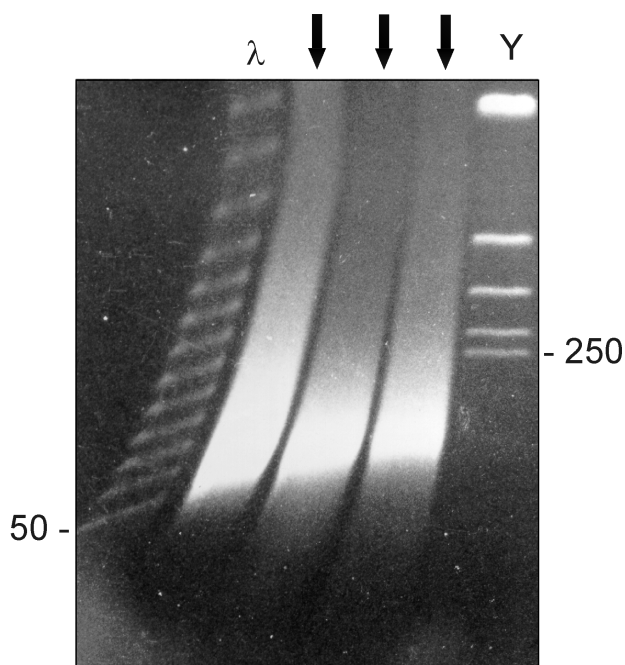

Fig. 4. Fractionation pattern of DNA forum domains after pulse electrophoresis. The arrows indicate the lanes with fractionated DNA of domains. DNA oligomers (λ) and yeast chromosomes (Y) are used as markers. Positions of DNA fragments with length of 50 and 250 kb are shown.

The discovery of forum domains, i.e., functional domains in chromosomes, allowed us to combine earlier cytogenetic data on intercalary HC with new data of molecular biology. Moreover, this discovery combines the previously separate data of molecular biology concerning different types of silencing. Genes of human miRNA are very often located in brittle areas of chromosomes [101]. These areas restrict minimal amplicons, contain important regulatory signals, and are associated with tumor development. Brittle areas of chromosomes are represented by BR, which delimit forum domains. Another example of RNA silencing associated with elimination of *Tetrahymena* chromatin was considered above [51]. The size of DNA domains of the macronucleus of Protozoa coincides with the size of forum domains of other eukaryotes [102]. Clusters of human and mouse imprinted genes are also located in large domains regulated through PcG proteins [17]. Proteins of this group bind to BR, i.e., heterochromatin small islands on the boundaries of domains.

Thus, it is reasonable to believe that basic principles of the genome organization, i.e., the alternation of DNA sequences and structure of functional domains in chromosomes (telomeres, centromeres, large HC regions, euchromatic forum domains delimited by HC islands) are associated with epigenetic regulatory mechanisms of gene activities.

RITS complex. Centromeric silencing in yeast also involves components of RNA interference. However, in this case RNA interference acts on the transcriptional level [103] with involvement of another, although a cognate protein complex RITS (RNA-induced initiation of transcriptional silencing). The existence of RITS complexes suggests a close relation between different epigenetic mechanisms, in this case between those of RNA interference and chromosomal silencing. There are data on association of RNA interference with chromatin modifications [104, 105] and on the role of RNA interference in elimination of DNA [51]. Centromeric silencing in

yeast occurs with proteins that are also involved in RNA interference, such as Dicer, Argonaute, RNA-dependent RNA polymerase, etc. The RITS complex includes a protein Chp1 containing a chromodomain, Ago1 homologous to *Drosophila*'s Argonaute, and also a recently found protein Tas3 [106]. It is suggested that similarly to RISC, RITS with involvement of siRNA recognizes the specific chromatin site and triggers its heterochromatization, i.e., transcriptional silencing. It seems that RITS proteins directly switch on mechanisms of histone modification that lead to silencing. This conclusion is based on disorders in histone modifications caused by mutations in the genes encoding the RITS proteins. It is suggested that hybrids of siRNA–RNA or siRNA–DNA in chromosomes are recognized either with involvement of immature transcripts or specific DNA sequences. Possibly, RITS is also involved in the so-called LTR-silencing (see below). The Ago1 protein of the RITS complex is also required for production of miRNA [107].

MOBILE ELEMENTS AS EPIGENETIC SIGNALS

There are two viewpoints concerning the role of mobile elements in genomes of eukaryotes. According to the first, mobile elements are genomic parasites. A putative manner of the cell defense against the undesirable expression of mobile elements has been considered above on discussing RNA interference. Another viewpoint suggests that mobile elements are "walking" regulatory sequences used by the cell to regulate its genes. In fact, mobile elements contain promoters, insulators, and other regulatory signals that can be found in new areas of the genome and influence expression of the host's adjacent genes.

These opposite viewpoints are incompatible only at first glance. Both viewpoints are supported by facts. In some situations mobile elements can display features which are dangerous for the cell and are neutralized, whereas in other situations they can be used by the cell for regulation of gene functions and serve as a material and factor of evolution. Some data on the positive role of mobile elements, in particular, their significance for epigenetic mechanisms will be considered below.

Mobile elements are centers of initial heterochromatization. The role of DNA sequences in the choice of chromatin structures were considered earlier. Mobile elements exemplify a similar choice of heterochromatic structures. Repeats are associated with generation of silent chromosomal areas. Thus, satellite and telomeric repeats are associated with heterochromatic regions. Mobile elements are shown to be centers of arising and further extension of large zones of silencing.

Large regions of *Drosophila*'s HC (pericentromeric and telomeric HC and the fourth chromosome) are characterized by hypoacetylation of nucleosomal histones, methylation of H3-K9, and binding of HP1. Specific relationships between HP1 and histone methyltransferases suggest a mechanism providing for heterochromatization through methylation of histones. However, it is important to determine how regions for production of large HC areas are chosen initially. Studies on transgenes in the fourth chromosome have revealed DNA sequences with length up to 80 kb, and the loss of these prevents the development of HC structures. Studies on this region of DNA have shown that the mobile element 1360 is a natural target of initial heterochromatization. And mobile elements are suggested to be centers of heterochromatization, which extends in both directions until it reaches euchromatin barriers [108]. Recently the HC production in *Arabidopsis* was found to depend on mobile elements and tandem repeats. This process occurs with involvement of a chromatin-remodeling factor DDM1, which seems to be oriented by siRNA molecules [109].

Mobile elements and regulation of transcription. Mobile elements can act as epigenetic factors significant for regulation of the transcription of the host's genome. Thus, two LTR-containing elements were found at a distance of about 3 kb upstream from *Drosophila white* gene. These elements have different effects on the gene transcription. The ZAM element displays a regulatory effect on gene *white*, strengthening its transcription in the eyes. This effect can be neutralized by insertion of the element Idefix, which acts as an insulator preventing the ZAM effect on gene *white* [110]. Mobile elements epigenetically regulate activities of adjacent genes with involvement of DDM1 and DNA methyltransferases [111]. Different mobile elements are involved in different regulatory mechanisms.

The mobile genes that are transcriptionally active in somatic tissues can be responsible for variability of characters. Silencing of mobile elements occurs during early embryogenesis, but it is not always complete, resulting in mosaic expression of mobile elements and adjacent genes. These activities of mobile elements are corrected in each generation but incompletely, and this leads to inheritable epigenetic changes [112]. RNA interference provides for the control of mobile element expression also in generative cells [113].

Expression of siRNA in plants can induce not only RNA interference but also methylation of DNA of the homologous locus. Methylation, which represses several meiotic genes, is caused by the neighboring LTR sequences of mobile elements in yeast, i.e., LTR silencing is realized [114]. It is suggested that LTR transcripts are recognized by the cell as aberrant, and this triggers local chromosomal silencing.

Concerted regulation of genes. RNA interference involving a short retro element of *Drosophila* suffix (SINE) is concurrent with a specific degradation of the F element (LINE) transcripts, which contain the sequence of the suffix on the 3′-terminus [59]. Suffix was found in

3′-noncoding areas of some mRNAs of *Drosophila*. Therefore, other genes can also be targets for RNA interference. Thus, expression of some genes located in different chromosomal regions can be disturbed. Sequences of *Alu*, which is the most widespread SINE in the human genome, are frequently present in mRNA. Sense and antisense transcripts of this element can trigger RNA interference and, consequently, degradation of the cognate sequences in different mRNAs. Such a distribution of

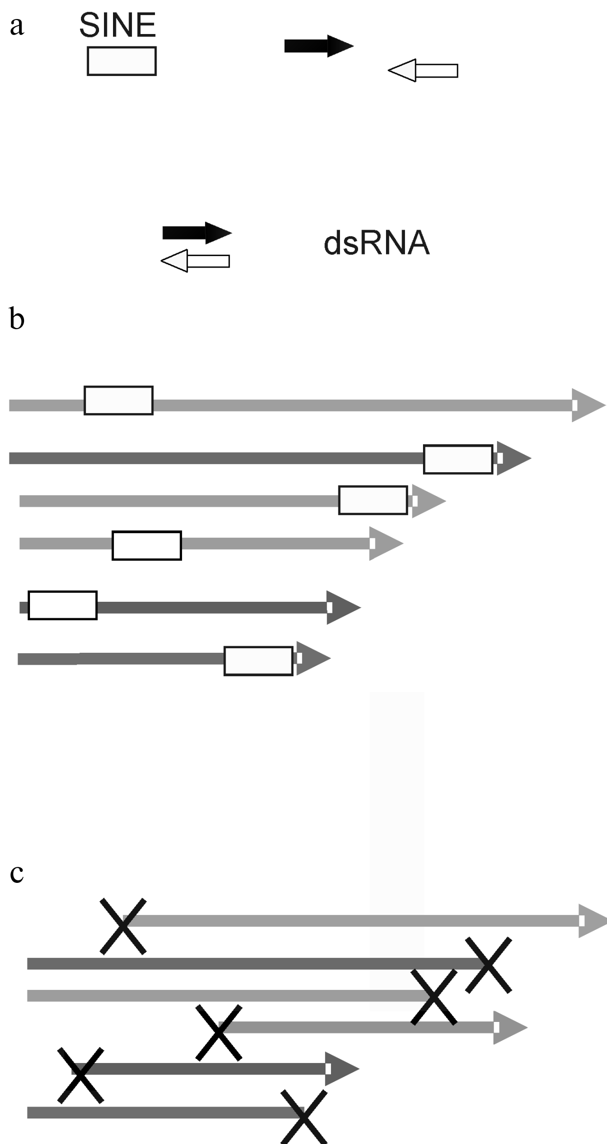

Fig. 5. Scheme illustrating the concerted regulation of expression of genes that contain in the 5′- or 3′-noncoding areas the sequence of the same SINE. a) Transcripts of both SINE chains can produce dsRNA, which triggers RNA interference; b) SINE sequences in different mRNAs are targets of RNA interference; c) as a result of RNA interference, mRNA sequences lose functionally important regions and the ability to be translated that leads to synchronous silencing of all genes marked by a particular SINE.

SINE sequences in the genome can provide for a synchronous silencing of expression of all genes that contain sequences of this retro-element in the 5′- or 3′-noncoding sequences (Fig. 5). The concerted silencing of the genes, similarly to the synchronous switching off of the genes in a forum domain with involvement of PcG proteins, can also play a role in differentiation. In forum domains expression of the neighboring genes are switched off, whereas the concerted silencing mechanism can switch off genes independently of their location in chromosomes.

Our understanding of epigenetic machineries is still far from complete. Some concepts are schematic. Thus, the functional division of proteins onto activators and repressors seems to be oversimplified, and the following examples confirm this conclusion. Swi/Snf is a chromatin-remodeling transcription-activating complex, but it also can repress transcription of some genes. In particular, it can repress the transcription of ribosomal RNA genes by RNA polymerase II [115]. The recognition of sites of silencing by the HP1 domain is another example. It has been discussed above that the HP1 silencing does not always depend on methyltransferase SU(VAR)3-9 [75], although the methylated H3 is known to be specifically recognized by a domain of this protein.

The transcriptional activator, often co-localized with RNA polymerase II, can also be present in areas specific for repressive chromatin and containing HP1 protein (Georgieva et al., unpublished data). Similarly, protein SuUR, which has several domains specific for transcriptional activators, occurs only in the regions of intercalary heterochromatin—in BR, on boundaries of forum domains [98].

Mechanisms of erasing epigenetic marks during meiosis are virtually uninvestigated. These marks determine which epigenetically encoded characters can be inherited in the next generation. There are only sparse data on erasing the imprinted marks of DNA methylation in germ line cells. Regulation of the mobile element transposition in these cells are to be studied. It is unclear what occurs with the chromosomal structures during meiosis, what is retained of numerous complexes RISC, RITS, PcG, trx, various siRNAs, miRNAs, and many other components of epigenetic mechanisms. In other words, the molecular organization of chromosomes, nucleus, and also of cytoplasmic elements significant for determination of gene expression, i.e., the epigenome, undergo significant changes programmed by meiosis. Studies on these events will allow us to characterize the reliability of mechanisms responsible for erasing the marks and evaluate the possibility of transmission of some epigenetic characters via meiosis and the role of this transmission in evolution. Studies of “nulling” and uncovering of epigenetic programs of development are of importance for various lines of studies, both basic and applied.

It is also difficult to answer how epigenetic regulatory mechanisms are controlled. Studies on homeotic genes

during the last 20 years have failed to result in clear comprehension of the circular chain of events in their regulation. It seems that the search for the primary source needed for triggering epigenetic mechanisms and generation of some or other chromosomal structures will take a long time.

The author is grateful to A. Olovnikov for his suggestion to write this review and to S. Georgieva for presenting unpublished data. Some experimental data obtained by the author were supported by the Program on Molecular and Cell Biology, Russian Academy of Sciences.

REFERENCES

1. Bird, A. (2002) *Genes Dev.*, **16**, 6-21.
2. Colot, V., and Rossignol, J. L. (1999) *Bioessays*, **21**, 402-411.
3. Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H., and Razin, A. (1992) *Genes Dev.*, **6**, 705-714.
4. Riggs, A. D., Xiong, Z., Wang, L., and LeBon, J. M. (1998) *Novartis Found Symp.*, **214**, 214-225.
5. Jaenisch, R. (1997) *Trends Genet.*, **13**, 323-329.
6. Selker, E. U. (1999) *Cell*, **97**, 157-160.
7. Bender, J. (2001) *Cell*, **106**, 129-132.
8. Niwa, O., Yokota, Y., Ishida, H., and Sugahara, T. (1983) *Cell*, **32**, 1105-1113.
9. Hayatsu, H., Wataya, Y., Kai, K., and Iida, S. (1970) *Biochemistry*, **9**, 2858-2865.
10. Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L., and Paul, C. L. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1827-1831.
11. Li, L. C., and Dahiya, R. (2002) *Bioinformatics*, **18**, 1427-1431.
12. Riggs, A. D., and Xiong, Z. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 4-5.
13. Ronaghi, M. **(**2001**)** *Genome Res.*, **11**, 3-11.
14. Fazzari, M. J., and Greally, J. M. (2004) *Nat. Rev. Genet.*, **5**, 446-455.
15. Falls, J. G., Pulford, D. J., Wylie, A. A., and Jirtle, R. L. (1999) *Am. J. Pathol.*, **154**, 635-647.
16. Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004) *Nature*, **429**, 900-903.
17. Delaval, K., and Feil, R. (2004) *Curr. Opin. Genet. Dev.*, **14**, 188-195.
18. Thakur, N., Tiwari, V. K., Thomassin, H., Pandey, R. R., Kanduri, M., Gondor, A., Grange, T., Ohlsson, R., and Kanduri, C. (2004) *Mol. Cell. Biol.*, **24**, 7855-7862.
19. Gong, Z., Morales-Ruiz, T., Ariza, R. R., Roldan-Arjona, T., David, L., and Zhu, J. K. (2002) *Cell*, **111**, 803-814.
20. Mager, J., Montgomery, N. D., de Villena, F. P., and Magnuson, T. (2003) *Nat. Genet.*, **33**, 502-527.
21. Laird, C. D., Pleasant, N. D., Clark, A. D., Sneeden, J. L., Hassan, K. M., Manley, N. C., Vary, J. C., Jr., Morgan, T., Hansen, R. S., and Stoger, R. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 204-209.
22. Pant, V., Mariano, P., Kanduri, C., Mattsson, A., Lobanenkov, V., Heuchel, R., and Ohlsson, R. (2003) *Genes Dev.*, **17**, 586-590.
23. Arnaud, P., Monk, D., Hitchins, M., Gordon, E., Dean, W., Beechey, C. V., Peters J., Craigen, W., Preece, M., Stanier, P., Moore, G. E., and Kelsey, G. (2003) *Hum. Mol. Genet.*, **12**, 1005-1019.
24. Nielsen, P. R., Nietlispach, D., Mott, H. R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A. G., Murzina, N. V., and Laue, E. D. (2002) *Nature*, **416**, 103-107.
25. Agalioti, T., Chen, G., and Thanos, D. (2002) *Cell*, **111**, 381-392.
26. Gilbert, C., Kristjuhan, A., Winkler, G. S., and Svejstrup, J. Q. (2004) *Mol. Cell*, **14**, 457-464.
27. Santoro, R., Li, J., and Grummt, I. (2002) *Nat. Genet.*, **32**, 393-396.
28. Strohner, R., Nemeth, A., Nightingale, K. P., Grummt, I., Becker, P. B., and Langst, G. (2004) *Mol. Cell. Biol.*, **24**, 1791-1798.
29. Aufsatz, W., Mette, M. F., Winden, J., Matzke, M., and Matzke, A. J. M. (2002) *EMBO J.*, **21**, 6832-6841.
30. Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003) *Genes Dev.*, **17**, 1870-1881.
31. Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004) *Genes Dev*., **18**, 1251-1262.
32. Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., and Dean, C. (2004) *Nature*, **8**, 164-167.
33. Kohler, C., and Grossniklaus, U. (2002) *Curr. Biol.*, **12**, R129-131.
34. Schneider, R., Bannister, A. J., Myers, F. A., Thorne, A. W., Crane-Robinson, C., and Kouzarides, T. (2004) *Nat. Cell. Biol.*, **6**, 73-77.
35. Schubeler, D., MacAlpine, D. M., Scalzo, D., Wirbelauer, C., Kooperberg, C., Leeuwen, F., Gottschling, D. E., O'Neill, L. P., Turner, B. M., Delrow, J., Bell, S. P., and Groudine, M. (2004) *Genes Dev.*, **18**, 1263-1271.
36. Jacobson, S., and Pillus, L. (2004) *Mol. Cell. Biol.*, **24**, 6029-6039.
37. Reid, J. L., Moqtaderi, Z., and Struhl, K. (2004) *Mol. Cell. Biol.*, **24**, 757-764.
38. Hassan, A. H., Prochasson, P., Neely, K. E., Galasinski, S. C., Chandy, M., Carrozza, M. J., and Workman, J. L. (2002) *Cell*, **111**, 369-379.
39. Duina, A. A., and Winston, F. (2004) *Mol. Cell. Biol.*, **24**, 561-572.
40. Ciechanover, A., Finley, D., and Varshavsky, A. (1984) *J. Cell. Biochem.*, **24**, 27-53.
41. Cheng-Fu, K., Hillyer, C., Tsukuda, T., Henry, K., Berger, S., and Osley, M. A. (2004) *Genes Dev*., **18**, 184-195.
42. McKittrick, E., Gafken, P. R., Ahmad, K., and Henikoff, S. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 1525-1530.
43. Sharp, P. A. (1999) *Genes Dev*., **13**, 139-141.
44. Zamore, P. D. (2002) *Science*, **296**, 1265-1269.
45. Tuschl, T. (2002) *Nat. Biotechnol.*, **20**, 446-448.
46. Plasterk, R. H. A. (2002) *Science*, **296**, 1263-1265.
47. Fire, A. (1998) *Nature*, **391**, 806.
48. Piacentini, L., Fanti, L., Berloco, M., Perrini, B., and Pimpinelli, S. (2003) *J. Cell. Biol.*, **161**, 707-714.
49. Kellum, R. (2003) *J. Cell. Biol.*, **161**, 671-672.
50. Schramke, V., and Allshire, R. (2004) *Curr. Opin. Genet. Dev.*, **14**, 174-180.

51. Mochizuki, K., and Gorovsky, M. A. (2004) *Curr. Opin. Genet. Dev.*, **14**, 181-187.
52. Hsieh, A. C., Bo, R., Manola, J., Vazquez, F., Bare, O., Khvorova, A., Scaringe, S., and Sellers, W. R. (2004) *Nucleic Acids Res.*, **32**, 893-901.
53. Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W. S., and Khvorova, A. (2004) *Nat. Biotechnol.*, **22**, 326-330.
54. Vella, M. C., Choi, E. Y., Lin, S. Y., Reinert, K., and Slack, F. J. (2004) *Genes Dev.*, **18**, 132-137.
55. Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A., and Timmermans, M. C. (2004) *Nature*, **428**, 84-88.
56. He, L., and Hannon, G. J. (2004) *Nat. Rev. Genet.*, **5**, 522-531.
57. Hutvagner, G., and Zamore, P. D. (2002) *Curr. Opin. Genet. Dev.*, **12**, 225-232.
58. Schwarz, D. S., and Zamore, P. D. (2002) *Genes Dev.*, **16**, 1025-1031.
59. Tchurikov, N. A., and Kretova, O. V. (2003) *Genetika*, **39**, 300-304.
60. Liu, C. G., Calin, G. A., Meloon, B., Gamliel, N., Sevignani, C., Ferracin, M., Dumitru, C. D., Shimizu, M., Zupo, S., Dono, M., Alder, H., Bullrich, F., Negrini, M., and Croce, C. M. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 9740-9744.
61. Baulcombe, D. C., and Molnar, A. (2004) *Trends Biochem. Sci.*, **29**, 279-281.
62. Tchurikov, N. A., Chistyakova, L. G., Zavilgelsky, G. B., Manukhov, I. V., Chernov, B. K., and Golova, Y. B. (2000) *J. Biol. Chem.*, **275**, 26523-26529.
63. Tchurikov, N. A., Shostak, N. G., Okladnova, O. V., and Chernov, B. K. (2000) *Genetika*, **36**, 149-154.
64. Heits, E. (1928) *Jahrb. Wiss. Botanik*, **69**, 762-818.
65. Grewal, S. I. S., and Elgin, S. C. R. (2002) *Curr. Opin. Genet. Dev.*, **12**, 178-187.
66. Schoenherr, C. J., Levorse, J. M., and Tilghman, S. M. (2003) *Nat. Genet.*, **33**, 66-69.
67. Dhillon, N., and Kamakaka, R. T. (2002) *Curr. Opin. Genet. Dev.*, **12**, 188-192.
68. Kaufmann, B. P. (1939) *Proc. Natl. Acad. Sci. USA*, **25**, 571-577.
69. Georgiev, G. P., Ilyin, Y. V., Ryskov, A. P., Tchurikov, N. A., and Yenikolopov, G. N. (1977) *Science*, **28**, 394-397.
70. Lu, B. Y., Ma, J., and Eissenberg, J. C. (1998) *Development*, **125**, 2223-2234.
71. Dillon, N., and Festenstein, R. (2002) *Trends Genet.*, **18**, 252-258.
72. Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigoe, K., and Rauscher, F. J., III (2003) *Genes Dev.*, **17**, 1855-1869.
73. Sadaie, M., Naito, T., and Ishikawa, F. (2003) *Genes Dev.*, **17**, 2271-2282.
74. Maggert, K. A., and Karpen, G. H. (2001) *Genetics*, **158**, 1615-1628.
75. Danzer, J. R., and Wallrath, L. L. (2004) *Development*, **131**, 3571-3580.
76. Kennison, J. A. (1995) *Annu. Rev. Genet.*, **29**, 289-303.
77. Pirrotta, V. (1998) *Cell*, **93**, 333-336.
78. Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J. R., Wu, C.-T., Bender, W., and Kingston, R. E. (1999) *Cell*, **98**, 37-46.
79. Van der Vlag, J., and Otte, A. P. (1999) *Nat. Genet.*, **23**, 474-478.
80. Lavigne, M., Francis, N. J., King, I. F. G., and Kingston, R. E. (2004) *Mol. Cell*, **13**, 415-425.
81. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002) *Science*, **298**, 1039-1043.
82. Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E., and Simon, J. A. (2002) *Cell*, **111**, 197-208.
83. Simon, J. A., and Tamkun, J. W. (2002) *Curr. Opin. Genet. Dev.*, **12**, 210-218.
84. Gerasimova, T. G., and Corces, V. G. (1998) *Cell*, **92**, 511-521.
85. Vazquez, M., Moore, L., and Kennison, J. A. (1999) *Development*, **126**, 733-742.
86. Orlando, V., Jane, E. P., Chinwalla, V., Harte, P. J., and Paro, R. (1998) *EMBO J.*, **17**, 5141-5150.
87. Horard, B., Tatout, C., Poux, S. P., and Pirrotta, V. (2000) *Mol. Cell. Biol.*, **20**, 3187-3197.
88. Pagans, S., Ortiz-Lombardia, M., Espinas, M. L., Bernues, J., and Azorin, F. (2002) *Nucleic Acids Res.*, **30**, 4406-4413.
89. Lehmann, M. (2004) *Trends Genet.*, **20**, 15-22.
90. Pirrotta, V., Poux, S., Melfi, R., and Pilyugin, M. (2003) *Genetics*, **117**, 191-197.
91. Gil, J., Bernard, D., Martinez, D., and Beach, D. (2004) *Nat. Cell. Biol.*, **l6**, 67-72.
92. Ali, J. Y., and Bender, W. (2004) *Mol. Cell. Biol.*, **24**, 7737-7747.
93. Peterson, A. J., Mallin, D. R., Francis, N. J., Ketel, C. S., Stamm, J., Voeller, R. K., Kingston, R. E., and Simon, J. A. (2004) *Genetics*, **167**, 1225-1239.
94. Kirmizis, A., Bartley, S. M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R., and Farnham, P. J. (2004) *Genes Dev.*, **18**, 1592-1605.
95. Cheng, M. K., and Shearn, A. (2004) *Genetics*, **167**, 1213-1223.
96. Phair, R. D., Scaffidi, P., Elbi, C., Vecerova, J., Dey, A., Ozato, K., Brown, D. T., Hager, G., Bustin, M., and Misteli, T. (2004) *Mol. Cell. Biol.*, **24**, 6393-6402.
97. Chandler, V. L., and Stam, M. (2004) *Nat. Rev. Genet.*, **5**, 532-544.
98. Tchurikov, N. A., Kretova, O. V., Chernov, B. K., Golova, Y. B., Zhimulev, I. F., and Zykov, I. A. (2004) *J. Biol. Chem.*, **279**, 11705-11710.
99. Zhimulev, I. F. (2003) in *General and Molecular Genetics* (Belyaeva, E. S., and Akif'ev, A. P., eds.) [in Russian], Sibirskoe Universitetskoe Izdatel'stvo, Novosibirsk, pp. 344-347.
100. Anan'ev, E. V., and Barskii, V. E. (1985) in *Electron Microscopic Map of Polytene Chromosomes of Drosophila Salivary Glands* (Evgen'ev, M. B., ed.) [in Russian], Nauka, Moscow, pp. 11-14.
101. Calin, G. A., Liu, C. G., Sevignani, C., Ferracin, M., Felli, N., Dumitru, C. D., Shimizu, M., Cimmino, A., Zupo, S., Dono, M., Dell'Aquila, M. L., Alder, H., Rassenti, L., Kipps, T. J., Bullrich, F., Negrini, M., and Croce, C. M. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 11755-11760.
102. Popenko, V. I., Novikova, E. G., Tchurikov, N. A., Viktorova, L. S., and Cherni, N. E. (1998) *Mol. Biol. (Moscow)*, **32**, 323-331.

103. Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., and Martienssen, R. A. (2002) *Science*, **297**, 1833-1837.
104. Pal-Bhadra, M., Bhadra, U., and Birchler, J. A. (2002) *Mol. Cell*, **9**, 315-327.
105. Zilberman, D., Cao, X., and Jacobsen, S. E. (2003) *Science*, **299**, 716-719.
106. Ekwall, K. (2004) *Mol. Cell*, **13**, 304-305.
107. Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M. C. (2004) *Genes Dev.*, **18**, 1655-1666.
108. Sun, F.-L., Haynes, K., Simpson, C. L., Lee, S. D., Collins, L., Wuller, J., Eissenberg, J. C., and Elgin, S. C. R. (2004) *Mol. Cell. Biol.*, **24**, 8210-8220.
109. Lippman, Z., Gendrel, A. V., Black, M., Vaughn, M. W., Dedhia, N., McCombie, W. R., Lavine, K., Mittal, V., May, B., Kasschau, K. D., Carrington, J. C., Doerge, R. W., Colot, V., and Martienssen, R. (2004) *Nature*, **430**, 471-476.
110. Conte, C., Dastugue, B., and Vaury, C. (2002) *Mol. Cell. Biol.*, **22**, 1767-1777.
111. Lippman, Z., and Martienssen, R. (2004) *Nature*, **431**, 364-370.
112. Whitelaw, E., and Martin, D. I. (2001) *Nat. Genet.*, **27**, 361-365.
113. Aravin, A. A., Naumova, N. M., Tulin, A. V., Vagin, V. V., Rozovsky, Y. M., and Gvozdev, V. A. (2001) *Curr. Biol.*, **10**, 1017-1027.
114. Schramke, V., and Allshire, R. (2003) *Science*, **301**, 1069-1074.
115. Dror, V., and Winston, F. (2004) *Mol. Cell. Biol.*, **24**, 8227-8235.