ISSN 0006-2979, Biochemistry (Moscow), 2007, Vol. 72, No. 11, pp. 1161-1178. © Pleiades Publishing, Ltd., 2007.
Original Russian Text © J. A. Makarova, D. A. Kramerov, 2007, published in Biokhimiya, 2007, Vol. 72, No. 11, pp. 1427-1448.

REVIEW

Noncoding RNAs

J. A. Makarova and D. A. Kramerov*

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, 119991 Moscow, Russia; fax: (495) 135-1405; E-mail: makarova@eimb.ru; kramerov@eimb.ru

Received June 5, 2007

Abstract—There has been substantial progress in studies on non-protein-coding RNAs (ncRNAs) during the last decade. Due to the development of new experimental approaches, a wide variety of such molecules have been found. This review considers the major groups of eukaryotic ncRNAs successfully studied in recent years. In particular, snoRNAs and scaRNAs involved in RNA modification and also miRNAs, siRNAs, tasiRNAs, and piRNAs responsible for RNA silencing are discussed. The transcripts of SINEs and SINE-like genes are considered separately.

DOI: 10.1134/S0006297907110016

Key words: ncRNA, SRA RNA, 7SK RNA, NRSE RNA, snoRNA, scaRNA, miRNA, tasiRNA, piRNA, BC1 RNA, BC200 RNA, 4.5S RNA

The study of non-protein-coding RNAs (ncRNAs) is one of the most actively developing and fascinating trends in molecular biology and biochemistry. The first ncRNAs were found half a century ago. These were ribosomal and transfer RNAs involved in translation. Then small nuclear RNAs (U1, U2, U4, U5, U6, U11, and U12) involved in splicing were revealed. Some more ncRNAs responsible for various functions were discovered later. Thus, 7SL RNA serves a carcass of a signal-recognizing particle and participates in protein transport in endoplasmic reticulum. RNA, which is a component of telomerase, acts as a template for synthesis of telomere repeats. RNAs are also found within RNase P and mitochondrial RNA processing RNase (RNase MRP). RNase P removes the 5′-terminal leader sequence in pre-tRNA. RNase MRP is involved in the processing of pre-rRNA and introduces a break into the first internal transcribed spacer. It also destroys the link between the primer RNA and the newly synthesized DNA in mitochondria [1]. All these RNAs are rather well studied and described in textbooks. Therefore, they will not be considered in the present review.

During recent years, strong evidence has been obtained that the number of ncRNA genes in eukaryotes is higher than the number of known and predicted protein-coding genes. Thus, analysis of the genome databases and cloning of short RNAs (of 20-200 nt in length) from the total cellular RNA or protein-associated ones revealed thousands of new functional RNAs. This very fruitful approach was termed RNomics [2]. Analysis of the genome transcriptional activity suggests that a considerably larger part of the genome than it has been supposed previously is transcribed, and the majority of new loci give rise to long polyadenylated and nonpolyadenylated ncRNAs.

The idea about the extent of genome transcription was mainly a result of development of a new experimental approach termed genomic tiling arrays (GTA). On conducting GTA, synthetic oligonucleotides or PCR products uniformly covering a chosen DNA sequence, e.g. the whole chromosome, are placed onto a chip. They can be arranged at a distance from one to another or overlap as tiling: the beginning of the next element corresponds to the end of the preceding one (repeated sequences are excluded). Most frequently, chips are hybridized with cDNA prepared from polyadenylated RNA. This approach is highly productive, does not depend on current annotations of the genome, and is very sensitive, which allows detection of rare transcripts [3].

Abbreviations: bp) base pair; dsRNA) double-stranded RNA; GTA) genomic tiling arrays; miRNAs) microRNAs; ncRNAs) noncoding RNAs; NRSE) neuron-restrictive silencer element; nt) nucleotides; UTR) untranslated region; piRNAs) Piwi-interacting RNAs; rasiRNAs) repeat-associated siRNAs; RNP) ribonucleoproteins; scaRNAs) small Cajal body-specific RNAs; SINE) short interspersed elements; siRNAs) small interfering RNAs; snRNAs) small nuclear RNAs; snoRNAs) small nucleolar RNAs; SRA) steroid receptor activator; tasiRNAs) transacting siRNAs; TGS) transcriptional gene silencing.

* To whom correspondence should be addressed.

The study by GTA of ten human chromosomes has shown that 10% of the genome sequences initiate cytoplasmic polyadenylated transcripts [4]. Note that the sequences located within mRNA are no more than 2% of the genome [5]. Similar results were obtained by the same approach for chromosomes 20-22 [6-8], the complete human genome ($1.5 \cdot 10^9$ nt, repeated sequences being excluded) [9], and also the genomes of Drosophila [10, 11] and Arabidopsis [12]. Some of the detected transcribed sequences correspond to previously unknown exons or whole protein-coding genes. However, 30-50% of such sequences constitute ncRNAs.

Similar data concerning the extent of genome transcription and the number of ncRNAs have been obtained by other methods [13-18]. In particular, according to data of the FANTOM3 project targeted for creating a complete catalog of mouse transcripts ncRNAs are encoded by ≥22,000 loci [19, 20].

Thus, a great many transcribed sequences remained unnoticed for some years. By analogy with physics, these transcripts were termed "dark matter" [3]. Its constituent ncRNAs may be subdivided into long (ranging from ~400 to tens of thousands of nucleotides) and short (20-400 nt) ones.

LONG ncRNAs

Tens of thousands of long polyadenylated and nonpolyadenylated ncRNAs found in the genomes of human and model organisms remain underexplored. As their transcription level is usually low, many of these RNAs are supposed to perform regulatory functions. There is also an alternative and widely discussed hypothesis [21] that ncRNAs emerge as a result of a nonspecific transcription and perform no function. The low conservativeness of the majority of ncRNAs favors the existence of such "transcriptional noise" [19]. The functional importance of these transcripts is supported by some data. Thus, promoters of ncRNAs are even more conservative than promoters of protein-coding genes [19]. Moreover, the absence of conservativeness does not always suggest that a sequence has no function [22, 23]. Many ncRNAs with known functions, such as Xist and Air, have low conservativeness.

The expression of many ncRNAs is tissue-specific, and they are transcribed during definite stages of development [4, 8, 12, 24]. Expression of many ncRNAs changes upon treatment with retinoic acid [25] and lipopolysaccharides [24]. Some oncological and hereditary diseases are accompanied by disorders in expression of ncRNAs, but it is often unclear whether these disorders are the cause or consequence of the disease [26].

Some examples of run-through transcription of full genomic domains comprising a number of protein-encoding genes have been described [26]. Thus, the domain including three genes of chicken α-globin is transcribed with production of 33,000-nt RNA [27]. This giant RNA is degraded within the nucleus. The run-through transcription of the genomic domain seems to be important for maintaining intense transcription of mRNAs of the α-globin genes themselves.

Many ncRNAs are encoded by the same loci as proteins but are located on the other DNA chain. Moreover, the same locus of both chains can encode different ncRNAs. Both DNA chains are transcribed in such loci and the resulting RNAs are able to form duplexes. In the mouse genome >70% of transcription units are overlapped with the opposite chain transcripts, and in half of the cases ncRNA is at least one member of the pair [20]. Although the size of genomes is sufficient for arrangement of all genes in separate loci, such a pattern of organization is widely distributed and seems to be biologically reasonable [28]. There is an attractive idea that an antisense transcript, e.g. ncRNA, regulates expression of its partner. Because interaction of these transcripts results in dsRNA (double-stranded RNA), this regulation could be realized by RNA interference mechanisms and RNA editing by adenosine deaminases [29]. The transcription of a pair member is also likely to prevent the transcription of its partner. But there are very few experimental data in favor of the regulatory role of antisense transcripts, and it is often difficult to interpret them unambiguously [28]. Thus, inhibition of expression of a pair member by siRNA (small interfering RNA) could be associated with an increase, decrease, or retention of the expression level of the other member [20]. Data confirming the functional significance of noncoding antisense RNAs and their putative regulatory role have been obtained [25]. Using GTA and chromatin immunoprecipitation, the authors mapped binding sites of the transcriptional factors p53, Sp1, and cMyc on human chromosomes 21 and 22; they also revealed 36% of the sites within the 3′-regions of the protein-coding genes, and the location of these sites correlated with the ncRNA distribution. Treatment with retinoic acid caused a coregulation of many antisense RNA/protein gene pairs. Because promoter regions of genes either coding or not coding proteins include the binding sites of the same transcription factors and external influences cause coordinated changes in expression of many of them, antisense RNAs are suggested to be not products of a nonspecific transcription, but to perform biological functions.

The spectrum of ncRNA functions is extremely wide. Thus, the ncRNA Xist and its antisense partner Tsix are involved in dosage compensation processes in mammals providing for the termination of expression of one of X chromosomes in females [30, 31]. In Drosophila ncRNAs roX1 and roX2 are involved in dosage compensation and contribute to increase in expression of the single X chromosome in males [32]. The ncRNA TUG1 is required for formation of retinal photoreceptors [33]. Some ncRNAs are necessary for expression of snoRNAs

(small nucleolar RNAs) encoded in their introns [34-38] but possibly they also have other functions [39]. Some ncRNAs are implicated in the regulation of expression of imprinted genes. Thus, RNA Air is expressed from the paternal allele and is required for repression of transcription of the locus Igf2r [40]. Many long ncRNAs are precursors of short ncRNAs [41, 42]; thus, some miRNAs are formed from exons of genes encoding ncRNAs. At present, the functional importance of several hundreds of long ncRNAs is established [26]. The information concerning them is given in the RNAdb database [43].

Authors of the work [44] were the first who have attempted to systematically study functions of long ncRNAs. They chose 512 conservative ncRNAs and inhibited their expression using siRNAs. This allows them to detect six ncRNAs necessary for cell survival. Another ncRNA termed NRON acted as a repressor of the transcription factor NFAT.

The appearance of systematic approaches for investigating functions of ncRNA seems promising for elucidation whether the majority of ncRNAs is a product of "transcriptional noise" or performs some functions. Consider more in detail SRA (steroid receptor activator) RNA, which is one of few well-characterized long (~1000 nt) ncRNAs with known function. SRA RNA discovered in 1999 is a coactivator of steroid hormone receptors [45]. Steroid hormone receptors belong to the nuclear receptor family, which includes about 50 members. All these receptors are transcription factors, which usually have two activator domains: activator domain 1 (AF-1) on the N-end and activator domain 2 (AF-2) on the C-end. AF-1 can be activated by phosphorylation with MAP kinase, whereas AF-2 is activated by ligands. Maximum transcription activation requires the activation of both domains. The activated receptor acquires the ability of interacting with the principal transcription factors and also with other proteins, the so-called transcription coactivators. Coactivators can possess histone acetylase or chromatin remodeling activities, and recruit RNA polymerase II [46, 47]. Almost all known coactivators are proteins. SRA RNA is a rare exception. This RNA is expressed in the majority of tissues. It is well represented in adrenals, hypophysis, and liver, and scarcely represented in mammary gland, uterus, and ovaries [48]. It is interesting that its expression is increased in malignant tumors of mammary gland, uterus, and ovaries, i.e. in the tumors that often depend on steroid hormones [48, 49]. SRA RNA is polyadenylated and exists in different forms (products of alternative splicing) with the common central part and 5′- and 3′-ends of varied length, and they all can intensify transcription [45]. The size range of this RNA varies from 900 to 2000 nt. SRA RNA has a sophisticated secondary structure, and its integrity is necessary for its functioning [50]. SRA RNA is found *in vivo* in 600-700-kD ribonucleoprotein complexes, which also contain the coactivator of steroid receptors NCoA-1/SRC-1 [45].

SRA RNA selectively strengthens the transcription of steroid hormone target genes [45]. It can also interact with receptors of thyroid hormones, and this interaction increases expression of reporter genes [51].

The mechanism of transcription activation by SRA RNA is not sufficiently clear. SRA RNA is shown to directly bind with the RNA-binding domain of thyroid hormone receptors, which is located near the ligand-binding domain containing the AF-2 motif [51]. As to steroid hormone receptors, SRA RNA can activate the transcription only in the presence of AF-1 domain but does not directly interact with the receptor [45]. If the receptor domain AF-1 is activated by MAP kinase, SRA RNA strengthens the transcription in the absence of the hormone [52]. If the hormone is bound, the receptor undergoes conformational changes, which lead to recruitment by AF-2 of TFIIH and the associated cyclin-dependent kinase CDK7. CDK7 promotes phosphorylation of AF-1 resulting in activation of the latter [53]. Coactivators interacting with one another also bind with each activator domain. This results in formation of a complex and maximal strengthening of the transcription. SRA RNA is involved in the assembly of this complex directly interacting with some coactivators and apparently performing a structural function. Moreover, SRA RNA is required for activity of the complex constituent coactivator p72/p68 [54]. The essential role of SRA RNA is also confirmed by the finding that pseudouridine synthases mPus1p and mPus3p act as coactivators of steroid hormone and retinoic acid receptors [55, 56]. They pseudouridylate SRA RNA at a number of positions, and this seems to cause conformational changes in RNA promoting its interaction with other coactivators [55]. Because many coactivators of transcription contain RNA-binding motifs, discovery of other SRA-like RNA-coactivators seems probable [57].

SRA RNAs interacts with at least two transcription corepressors, SHARP and SLIRP. The binding of SHARP prevents SRA RNA entering into the receptor-associated complex of coactivators [58], whereas SLIRP competes with coactivator NCoA-1 for binding with SRA RNA [59].

Note that starts of alternative transcription result in SRA RNA isoforms capable of encoding a protein. This protein was detected in human mammary gland tumors [60] and termed SRAP. The increased expression of this protein in the human mammary gland tumor cells MCF-7 decreases the transcriptional activity of estrogen receptors [61]; however, functions of SRAP are still unclear. Thus, the gene of SRA RNA is a rare example of a gene simultaneously encoding protein and RNA products.

SHORT ncRNAs

7SK RNA. This RNA was described as early as in the 1970s [62], but its function has been established only

recently. Human 7SK RNA is 331 nt in length, located in the nucleus, transcribed by RNA polymerase III, and conservative in vertebrates.

7SK RNA contributes to inhibition of a transcription elongation factor P-TEFb. P-TEFb consists of cyclin-dependent kinase 9 (CDK9) and cyclin T1 (CycT1) and is necessary for expression of genes of the majority of proteins: in its absence RNA polymerase II synthesizes only short 5′-terminal sequences of pre-mRNA.

The transcriptional and kinase activities of P-TEFb are inhibited by a complex consisting of 7SK RNA and the protein HEXIM1 linked with its 5′-terminal hairpin. The 3′-terminal hairpin of 7SK RNA is linked with CycT1, whereas another component of P-TEFb, CDK9, interacts with HEXIM1, which leads to inhibition of the P-TEFb activity. Thus, 7SK RNA is responsible for association of P-TEFb and HEXIM1. Note that only the active form of P-TEFb (with CDK9 phosphorylated at T186) can interact with the 7SK RNA–HEXIM1 complex. Therefore, this complex removes from turnover an active P-TEFb capable of providing for elongation of transcription. As a result, transcription intensity can be regulated rapidly [63].

NRSE RNA. This double-stranded 20-nt RNA is located in the nucleus [64]. Its nucleotide sequence is identical to the sequence disposed in promoter regions of many neuron-specific genes and is called RE1/NRSE (repressor element 1/neuron-restrictive silencer element) [65, 66]. The RE1 (its length is 19-24 nt) binds with the transcription factor REST/NRSF (RE1 silencing transcription factor/neuron restrictive silencer factor) [67, 68]. This factor recruits corepressors, in particular, histone deacetylase, which prevents transcription of genes with RE1-containing promoters [69]. In mature neurons, REST is not expressed, and in all other cells it inhibits expression of neuron-specific genes [65, 70-72]. However, later REST expression was found in some types of mature neurons [72-74].

Mechanisms of RE1-containing gene expression in the presence of REST are still unclear. According to some data, RE1-containing promoters are not linked with REST in mature neurons [71], but other authors believe that such a linkage exists [64, 72]. This difference can be caused by different affinities of REST for diverse RE1-containing promoters [66].

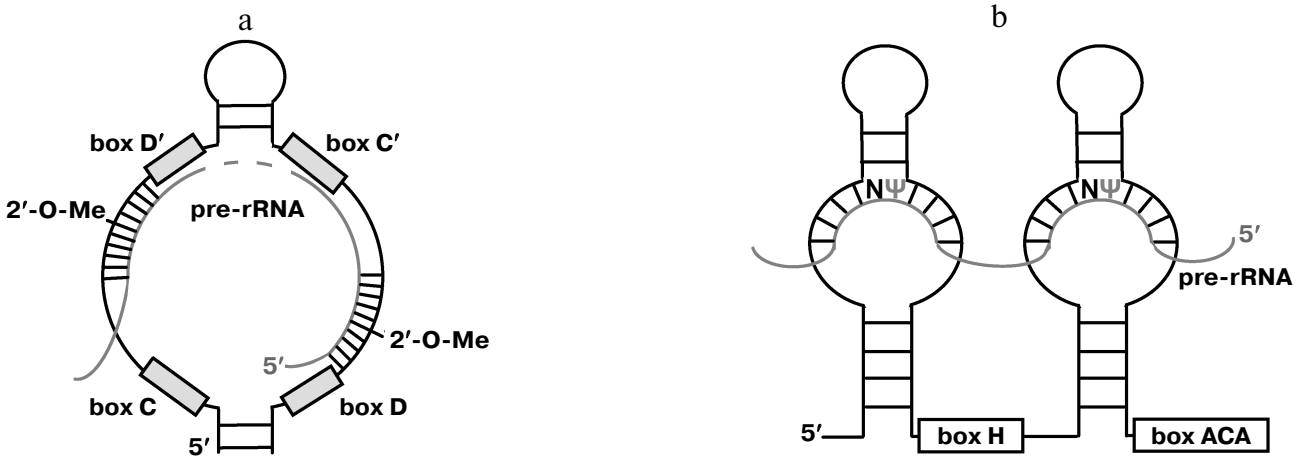
The authors detecting the retained binding between REST and promoters of target genes have found that these genes are de-repressed and even activated by NRSE RNA [64]. This RNA is not found in neurons, but it can be detected in cultured hippocampus stem cells differentiating by the neuronal pathway and also in the part of the hippocampus which shelters neurogenesis.

Expression of NRSE RNA in neuronal stem cells provokes their differentiation by the neuronal pathway and is accompanied by activation of RE1-containing gene transcription. Thus, NRSE RNA is a coactivator of transcription [75]. Although the mechanism of action of NRSE RNA is unknown, it has been shown to associate with REST, which is bound with the RE1 sequence of the promoter. It seems that REST can bind NRSE RNA because its sequence is identical to that of RE1. It is supposed that the association with NRSE RNA prevents the binding with REST of transcription corepressors and promotes transcription activation by REST possibly as a result of NRSE RNA-induced conformational changes even in the absence of NRSE RNA [64]. Further studies are likely to define more exactly the role of NRSE RNA and the transcription factor REST in regulation of the RE1-containing gene transcription.

SnoRNAs. Small nucleolar RNAs (snoRNAs) comprise a large group of ncRNAs: mammals have >200 snoRNAs. The functions of snoRNAs were determined in the 1990s, although they were discovered about 40 years ago. SnoRNAs are involved in cutting of pre-rRNAs and stipulate the two most frequent modifications of RNA: 2′-O-methylation and pseudouridylation. In vertebrates ribosomal RNA contains ~100 modifications of each kind [76]. Similar RNAs are found in archaebacteria and termed sRNA (sno-like) [77].

Based on conservative elements of the nucleotide sequence, two families of snoRNA are distinguished: C/D and H/ACA (box C/D and box H/ACA snoRNAs). RNA from RNase MRP also belongs to snoRNAs. This RNase introduces a break in the site A3 of the first internal transcribed spacer of pre-rRNA that causes release of 5.8S rRNA [1].

The snoRNAs of the C/D family determine the RNA nucleotide that will be 2′-O-methylated. The snoRNAs are ~70 nt in length and contain the so-called boxes C (UGAUGA), D (CUGA), and, usually also their copies C′ and D′ which can be degenerate [78]. Conservative elements are arranged in the order of 5′-C-D′-C′-D-3′, and boxes C and D are brought together due to complementary interactions of the snoRNA terminal nucleotides (Scheme 1, panel (a)). The boxes C and D and terminal hairpin form a structure termed C/D-motif [79]. This motif functions as a binding site of four C/D-RNP (ribonucleoprotein) core proteins: NOP56, NOP58, 15.5-kD protein, and fibrillarin [80, 81]. Association with the core proteins is necessary for the nucleolar location of snoRNAs and protects their ends against degradation [82]. The so-called antisense elements (10-21-nt sequences) complementary to a fragment of one of the cellular RNAs and capable of interacting with it are located towards the 5′-end from box D and/or D′. As a result of this interaction, the RNA nucleotide located in the produced double helix and separated by four nucleotides from the sequence D and/or D′ is 2′-O-methylated [83]. The methylation is realized by a C/D-snoRNP core protein—fibrillarin. The majority of C/D-snoRNAs contains one antisense element and determines 2′-O-methylation of one nucleotide, but



Nucleolar RNAs of the C/D (a) and H/ACA (b) families. 2′-O-Me) 2′-O-methylated nucleotide; ψ) pseudouridylated nucleotide

Scheme 1

some of them contain two elements and can modify two nucleotides.

Almost all snoRNAs of the H/ACA family can determine a nucleotide of RNA destined for pseudouridylation. These 100-150-nt RNAs are present in the cell within RNP particles containing four core proteins: GAR1, dyskerin, NHP2, and NOP10 [84, 85]. H/ACA RNAs form a conservative secondary structure, which includes two hairpins with the so-called boxes H (ANANNA) and ACA (Scheme 1, panel (b)). One of the hairpins (or less often both) contains an antisense element consisting of two 3-10-nt stretches complementary to a fragment of one of the cellular RNAs [86]. Due to interaction between the antisense element and RNA target, the nucleotide predetermined for modification is exposed in the produced single-stranded "window", where it is recognized by pseudouridine synthase dyskerin realizing the modification (Scheme 1, panel (b)). The majority of H/ACA-snoRNAs contains one antisense element and promotes modification of one nucleotide, but some of them contain two such elements and are involved in modification of two nucleotides.

Some RNAs of the families C/D (U3, U8, U14, U22) and H/ACA (U17/E1/snR30, E2, E3, snR10) are involved in the cutting of pre-rRNA apparently acting as RNA chaperons [87-89]. Almost all other known snoRNAs control 2′-O-methylation and pseudouridylation of rRNAs and small nuclear RNAs [90] that stabilize their secondary structures [91]. Targets of some snoRNAs are still unknown [92-94]. Probably, mRNAs will be such targets. Thus, the snoRNA MBII-52 has been recently found to regulate the alternative splicing of the 5-HT2C serotonin receptor gene mRNA [95]. The deletion of this snoRNA gene contributes to pathogenesis of the severe hereditary disease Prader–Willi syndrome. The structure and functions of snoRNAs are considered in detail in [96].

Note that the snoRNA genes can be differently arranged in the genome. In vertebrates, nearly all snoRNA genes are located within introns of other genes, which are called host genes, and one intron contains only one snoRNA gene [97]. SnoRNAs are processed during splicing of host gene pre-mRNAs [98] or (less often) are cut from the introns by endonucleases [99]. In this case, the production of mRNA and nucleolar RNA is alternative. Nearly all host genes encode proteins, but some of them produce noncoding RNAs [34, 39].

The majority of snoRNA genes of *Drosophila melanogaster* [100] and monocotyledonous plants [101] are organized as clusters located within introns of host genes. In dicotyledonous plants, clusters of snoRNA genes are transcribed from their own promoters, and the produced polycistronic RNA is processed by endonucleases [101]. Finally, all taxons of eukaryotes contain some snoRNA genes which are independent transcription units. The organization of snoRNA genes is considered in more detail in [102].

scaRNAs. Some RNAs of the C/D and H/ACA families are disposed not in the nucleus but inside Cajal bodies and are termed small Cajal body-specific RNAs (scaRNAs) [103]. Cajal bodies (CBs), or coiled bodies, are oval or round bodies 0.3-0.5 μm in diameter located in the nucleus [104]. In particular, CBs are suggested to take part in the maturation and assembly of snRNPs and snoRNPs.

Small nuclear RNAs (snRNAs) involved in the splicing are targets for scaRNAs. In addition to C/D and H/ACA scaRNAs, CBs also contain 250-300-nt scaRNAs comprising conservative sequences of both C/D and H/ACA RNAs. Such chimeric scaRNAs are associated with core proteins of C/D and H/ACA RNAs and are involved in both 2′-O-methylation and pseudouridylation [103, 105].

No elements have been found in the C/D family scaRNAs that would be capable of promoting their presence in CBs, and it is still unclear how they are kept there. Each hairpin forming the secondary structure of H/ACA and chimeric scaRNAs contains the sequence UGAG termed the CAB-box [106]. The CAB-box is responsible for location of these RNAs in Cajal bodies.

SMALL RNAs INVOLVED IN GENE SILENCING

In 1998, double-stranded RNAs (dsRNAs) were found to repress gene expression. Later, short single-stranded RNAs were shown to be leaders during this process. The mechanism of gene expression inhibition by these RNAs was called RNA interference and also RNA silencing. This mechanism is found in representatives of all large taxons of eukaryotes: vertebrates and invertebrates, fungi, plants, and protozoa. Expression can be repressed at the transcription level and after it. In all cases a similar set of proteins and short RNAs of 21-32 nt in length are required. Such RNAs include microRNAs (miRNAs), small interfering RNAs (siRNAs), *trans*-acting siRNAs (tasiRNAs), and recently detected RNAs associated with the Piwi family proteins (piRNAs).

miRNAs. The miRNAs of 22 nt in length are found in plants, animals, and even some viruses.

miRNA genes. Genes of some miRNAs are independent transcription units, whereas genes of many other miRNAs are clustered. In both cases miRNAs are transcribed as components of longer capped and polyadenylated precursors. About half of the known mammalian miRNAs are encoded by protein gene introns and exons of ncRNA genes. Some genes of miRNAs are located within 3′-untranslated regions (UTR) of protein-encoding genes. In this case, the production of mRNA and miRNA are alternative processes [107]. All the transcripts containing miRNA sequences are assigned to primary miRNAs (pri-miRNAs) [42].

Some copies of pseudogenes [108] and mobile elements of mammals [109] have been shown to encode miRNAs. Copies of mobile elements are components of 3′-UTRs of many genes, so miRNAs originating from the repeats (more often LINE2) can regulate the cellular level of many unrelated mRNAs. This pathway of origination can produce taxon-specific miRNAs. Thus, a miRNA miR333 is found only in rodents as expected because it is encoded by the copy of a rodent-specific short retroposon B2 [109]. This is an interesting example of the arising of new genes and illustrates the adapting by the cell of "egoistic" DNA for performing useful functions.

Viral genomes also encode miRNAs. After the first detection of such miRNA in 2004, its number is continuously increasing. Viral miRNAs can contribute to repression of both viral and possibly host cell gene expression [110-112].

Maturation of miRNAs. Within pri-miRNA there is a ~70-nt region which includes an miRNA sequence and can form a hairpin. The hairpin is cut off by nuclear endonuclease Drosha complexed with an RNA-binding protein Pasha/DGCR8 (names of the Drosophila and mammalian orthologs, respectively). Under the influence of the protein Exportin5 the hairpin, now called pre-miRNA, is transferred into the cytoplasm, where it is treated by endonuclease Dicer associated with the RNA-binding protein Loqs/TRBP (names of the Drosophila and human orthologs, respectively). Both Drosha and Dicer belong to the RNase III family, the members of which have endonuclease activity toward dsRNA [113]. The actions of these two RNases result in dsRNA of ~22 nt in length that contains two unpaired nucleotides on the 3′-ends. This dsRNA is a component of an RNA-induced silencing complex (RISC), which includes some proteins with an unidentified functions and a protein of the Ago family, which is the major component of RISC [114]. The Ago family proteins are members of the large protein group Argonaute involved in RNA silencing. These proteins are found in all eukaryotes studied (except the yeast *Saccharomyces cerevisiae*) and even in archae, where their function remains a mystery [115]. A protein of the Ago family associates with dsRNA and destroys the complex with retention of miRNA within RISC and degradation of the complementary chain. The miRNA processing is considered more comprehensively in other reviews [42, 114, 116-118].

Mechanisms of regulation of mRNA expression by miRNA. Within RISC, miRNA can complementarily interact with mRNA. This interaction results either in degradation of the target mRNA or in repression of its translation.

mRNA is degraded when miRNA is completely complementary to the target RNA or contains two-three unpaired nucleotides. RISC also has to contain an Ago protein possessing nuclease activity [119]. Such activity is inherent in many proteins of this family, but only Ago2 of four human Ago proteins displays it [114]. Within RISC, possessing the endonuclease activity Ago cuts the bond between mRNA nucleotides located opposite the 10th and 11th nucleotides of miRNA. The resulting products are destroyed by systems of mRNA degradation [114, 119].

A number of animal and viral miRNAs and also all known miRNAs of plants are fully or almost fully complementary to target mRNAs and therefore induce their degradation [42, 116]. MiRNA binding sites are usually located in 3′-UTRs in animals and in the coding region in plants [120], which seems to allow the most efficient termination of protein synthesis [42].

The majority of animal miRNAs are not fully complementary to target mRNAs and therefore cause no degradation, but they repress translation. Usually only the 5′-region of miRNA (nucleotides 2-7/8, the so-called

seed region) remains complementary to the target mRNA, whereas the other part can contain unpaired nucleotides [42, 114]. Due to the small size of the seed region, one miRNA can have many target mRNAs. Just this may be an explanation of high conservativeness of the majority of known miRNAs: the necessity to interact with many targets strongly limits the variability of miRNAs [26]. However, the current algorithms of miRNA detection in the genome databases are usually based on searches for conservative sequences, so non-conservative miRNAs, if they exist, cannot be revealed by the bioinformatics approaches. In fact, upon removal of limitations on conservativeness, a considerable number of primate-specific miRNAs has been detected [23].

Repression mechanisms of target mRNAs are still unclear. Based on the available data, repression is supposed to occur during different stages of translation, both during the initiation and later [119, 121]. In some cases the translation can be recommenced, but the repression is often accompanied by a pronounced degradation of target mRNAs. The degradation is not mediated by RNA interference because miRNA and mRNA are not completely complementary. Because the stability of mRNA can be associated with its translation status, the degradation of target mRNAs seems to be due to repression of their translation. mRNAs with translation terminated due to interaction with miRNAs are accumulated in the so-called processing bodies (P-bodies) [122, 123]. P-bodies are places of accumulation and degradation of mRNAs whose translation has been repressed. P-bodies contain proteins involved in these processes, in particular, those of the Ago family, and also miRNAs. In some cases, e.g. under stress conditions, individual mRNAs can leave P-bodies and recommence their involvement in translation.

It seems that miRNAs direct deadenylation of mRNAs. In Drosophila, the miRNA-associated protein dAgo1 of the Ago family interacts with a P-body component—RNA-binding protein GW182 [122], which in turn recruits deadenylase. Then the cap seems to be removed by a decapping enzyme consisting of two subunits: Dcp1 and Dcp2. Then mRNA is 5′-3′-degraded by exonuclease Xrn1 [121]. But it is still unclear why some targets of miRNAs are degraded while others undergo only repression of the translation.

Functions of miRNAs. More than 300 miRNAs have been experimentally detected in humans [124], and together with those predicted by bioinformatics approaches their number is 800-1000 [23, 125]. From 10 [126] to 30% [127] of all protein-coding genes are believed to be regulated by miRNAs. The known miRNAs usually regulate gene expression at the post-transcriptional level, but can also repress transcription of target genes [128].

Gene expression is differently regulated by miRNAs in plants and animals: in animals miRNAs seem to subtly regulate activities of hundreds of genes and strongly decrease expression only of some genes [129]. In these relatively rare genes, 3′-UTRs contain a number of binding sites of one or several miRNAs, which enhances repression of translation. 3′-UTRs of genes subjected to "fine tuning" usually contain one miRNA binding site [130]. Note that housekeeping genes have short 3′-UTRs with a low density of miRNA binding sites. Possibly just this allows their transcripts to essentially escape the miRNA-mediated repression of translation [110, 130]. In plants target mRNAs usually contain one miRNA binding site [120], and interaction with miRNAs results in degradation of the target mRNAs.

In contrast to targets of miRNAs in animals, those in plants may be rather reliably predicted because the miRNAs are usually fully complementary to target mRNAs. About half of all predicted targets of miRNAs in Arabidopsis are transcription factors, although they are only 6% of protein-coding genes [42]. As a rule, one miRNA regulates expression of several transcription factors of the same family [120, 131]. miRNAs and miRNA-controlled transcription factors regulate the main stages of plant development, such as beginning of flowering and formation of vegetative and generative organs [116]. Note that mRNAs of DCL1 (Dicer homolog in plants) and an AGO1 protein of the Ago family are targets of miRNAs. This suggests a feedback mechanism that allows miRNAs to determine the level of their own production [42].

Animal miRNAs have many more different target RNAs and are necessary for normal embryogenesis and tissue differentiation. They are also involved in programmed cell death. Expression of the majority of miRNAs is tissue-specific, and many of them are expressed only in the nervous system and are required for its normal functioning [42, 110].

Point mutations of miRNAs and their targets are supposed to weaken or strengthen expression of target mRNAs and also induce new targets. These effects can vary in manifestations from appearance of different phenotypes to development of severe diseases. The first data supporting these hypotheses are presented in a review [110]. Because miRNAs are involved in numerous vitally important processes, disorders in their expression can lead to tumorigenesis. Expression of many miRNAs is usually decreased in malignant tumors of different origin [132]. This is in agreement with the current concept that the expression of miRNA is connected with differentiation: expression of miRNAs is decreased in tumor cells with a lost profile of tissue-specific transcription. MiRNAs can act as both oncosuppressors and oncogenes. In the first case, they lower expression of oncogenes and in the other case that of oncosuppressors. The association of disorders in miRNA expression with malignancies is discussed in more detail in reviews [133, 134].

The miRNA transcription level is promising for diagnosis of malignancies, and the accuracy of such diagnosis is markedly higher than on using the current mRNA

markers [132]. Moreover, in some cases the profile of miRNA expression can be used to predict disease outcome [133]. The involvement of miRNAs in tumorigenesis can be used for development of new approaches for treatment. Some research groups have created oligonucleotides complementary to miRNAs and capable of *in vivo* lowering of their activities [135-138]. This is associated with an increase in expression of target mRNAs. These oligonucleotides seem to be candidates for a new generation of drugs.

The siRNAs are of 21-25 nt in length, and they are produced from longer dsRNAs. These RNAs can have an origin from viral infections, genetic constructions introduced into the genome, long hairpins within transcripts, and bidirectional transcription of mobile elements [114].

DsRNAs are cut by RNase Dicer into 21-25-nt fragments with protruding 3′-ends of two nucleotides, and then one of the chains becomes a component of RISC and directs the cutting of homologous RNA molecules. RISC comprises siRNAs corresponding to both plus- and minus-chains of dsRNAs [42, 139]. The origin of siRNA is different from that of miRNAs. MiRNAs are encoded by their own genes and are cut from a hairpin formed by the precursor. SiRNAs do not have their own genes and are fragments of longer RNAs. SiRNAs direct the cutting of a target RNA because they are fully complementary. In plants, fungi, and the nematode *Caenorhabditis elegans* the gene expression is repressed by siRNAs with involvement of RNA-dependent RNA polymerases [140]. In *C. elegans*, siRNAs not only direct the cutting of homologous RNA but also serve as primers for RNA-dependent RNA polymerase, which completes the second chain using the target RNA as a template [141]. The resulting dsRNA is cut by Dicer with production of new, so-called secondary siRNAs. This is the manner of signal amplification.

The siRNAs can regulate gene expression in two ways. First, as mentioned above, they direct the cutting of target RNAs. This phenomenon is termed quelling in fungi, post-transcriptional gene silencing (PTGS) in plants, and RNA interference in animals [114]. These processes occur with involvement of 21-23-nt siRNAs. Second, siRNAs can repress transcription of genes containing sequences homologous to siRNAs [142]; this phenomenon termed transcriptional gene silencing (TGS) is found in the yeast *Schizosaccharomyces pombe* [143], animals [144, 145], and plants [146]. In plants TGS is realized through ~24-nt siRNAs [146]. The siRNAs direct histone modifications and DNA methylation, which results in production of heterochromatin and repression of transcription [114]. TGS is best studied in *S. pombe* [143]. In *S. pombe*, siRNAs integrate into a RISC-like protein complex termed RITS (RNA-induced initiation of transcriptional gene silencing). As in the case of RISC, within this complex siRNA interacts with a protein of the Ago family. The siRNAs are suggested to be capable of directing this complex towards the gene containing a siRNA-homologous fragment. Then RITS proteins recruit methyltransferases that results in formation of heterochromatin within the locus encoding the target gene of siRNA, and this terminates active expression of the gene. Mechanisms of TGS are considered in review [147], and formation pathways, functions, and history of discovery of siRNAs are described in reviews [117, 118, 148-150].

The siRNAs contribute to cell defense against viruses, repression of transgenes, regulation of some genes, and formation of the centromeric chromatin. The siRNAs also repress expression of mobile elements at both transcriptional and post-transcriptional levels.

The siRNAs with sequences corresponding to fragments of different genomic repeats including mobile elements are found in various organisms, such as *Trypanosoma brucei* [151], *C. elegans* [152], *D. melanogaster* [153], *Arabidopsis thaliana* [146, 154, 155], *Danio rerio* [156], and *Mus musculus* [157]. These siRNAs are often called repeat-associated small interfering RNAs (rasiRNAs). In particular, they are produced upon the cutting by RNase Dicer of dsRNAs synthesized on a bidirectional transcription of mobile elements (the antisense chain is produced when a mobile element incorporates next to the cell promoter). The length of such rasiRNAs is 21-23 nt [139, 157].

However, it recently became clear that rasiRNAs are not representatives of a homogenous class: in addition to the above-described short rasiRNAs, longer rasiRNAs (of 24-32 nt in length) are found. Unlike the short rasiRNAs, they are associated with other proteins of the Argonaute group and seem to be produced not upon the cutting of dsRNAs but as a result of processing of a long one-stranded precursor. These longer rasiRNAs belong to the family of recently discovered piRNAs [158-161].

TasiRNAs. In flowering plants and mosses [162] siRNAs are found encoded by their own genes and directing the cutting of mRNAs of other non-homologous genes. Therefore, they are termed transacting siRNAs (tasiRNAs). They are generated by pathways different from those of miRNA and siRNA generation.

There are five genes whose transcription results in polyadenylated and putatively capped precursors of tasiRNAs: TAS1a, TAS1b, TAS1c, TAS2, and TAS3 [163-166]. TAS1a, TAS1b, and TAS1c RNAs are encoded by paralogical loci and are slightly like TAS2 RNA, so all the four loci seem to be paralogs [165]. All five transcripts have no elongated open reading frame and can encode peptides with no more than 50 amino acid residues. Therefore, the authors describing these genes think that TAS RNAs do not encode proteins. However, it should be noted that for plants transcripts are known capable of encoding short functional peptides. Thus, peptides encoded by the ENOD40 gene of Leguminosae contain 12 and 24 amino acid residues [167, 168].

All TAS RNAs serve as targets for miRNAs, which direct their cutting: TAS1a, TAS1b, TAS1c, and TAS2 RNAs interact with miRNA miR173 [165, 166], and TAS3 RNA interacts with miR390 [166]. As a result, TAS RNAs are cut into two parts. The 5′-terminal part of TAS3 RNA and the 3′-terminal parts of other TAS RNAs are transformed into double-stranded forms by RNA-dependent RNA polymerase RDR6 [165, 166, 169]. The resulting dsRNAs are cut into 21-nt fragments with protruding 3′-ends by RNase III Dicer-like 4 (DCL4), which is homologous to animal Dicer [165, 170] (Scheme 2). This results in tasiRNAs. Within the primary transcript, almost all detected tasiRNAs are disposed either tightly or at a distance divisible by 21 nt. Thus, the place of TAS RNA cutting prescribed by miRNA determines the phase of the resulting tasiRNAs.

All TAS RNA genes have two exons. In the case of TAS1a, TAS1b, TAS1c, and TAS2 genes, the site of miRNA cutting and all the detected tasiRNAs (~10 for each TAS RNA) reside within an intron. Thus, unspliced TAS RNAs undergo processing with involvement of RDR6. TAS3 RNA gives rise to two virtually identical tasiRNAs located side by side within the exon. Nearly all found tasiRNAs correspond to fragments of the TAS RNA sequence. However, for each transcript at least one tasiRNA corresponding to the minus-chain is found. Such RNAs can regulate the cell levels of their own precursors [165, 166].

The tasiRNAs seem to be constituents of RISC and, similarly to miRNAs and siRNAs, direct cutting of target genes [171] (Scheme 2). TAS3 RNA products use as targets mRNAs of transcription factors ARF2, ARF3, and ARF4 (auxin responsive factors) [166, 169]. Disorders in production of these tasiRNAs result in phenotypic defects [172]. Some genes with unidentified function act as targets for other tasiRNAs [163-166].

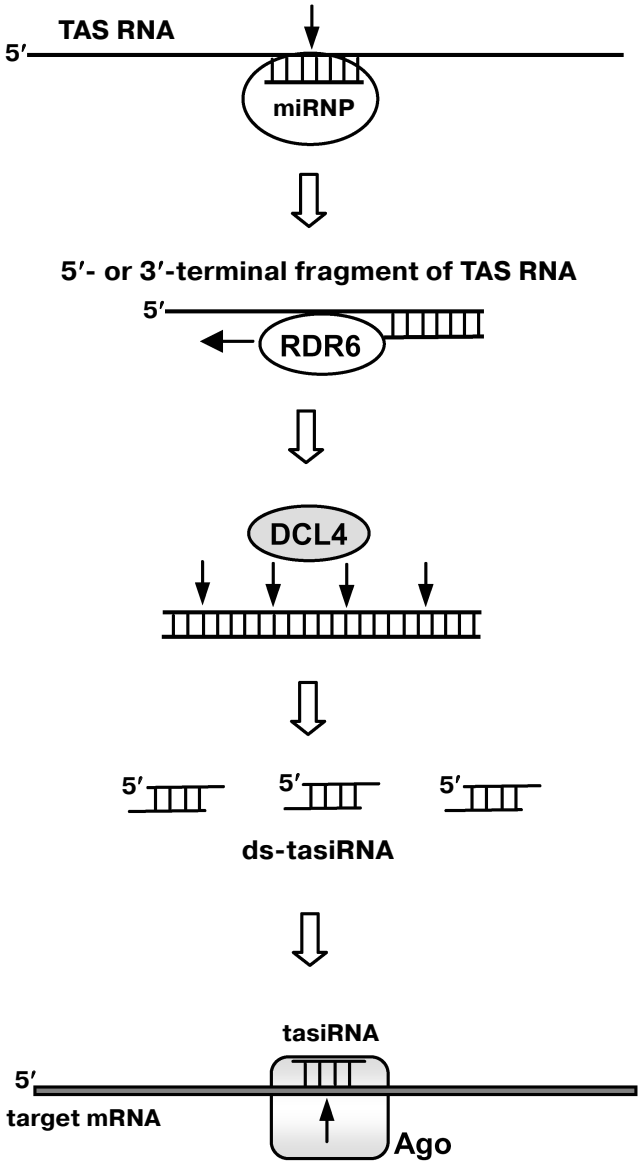
As in the case of miRNAs, TAS RNA and target mRNA retain their similarity only in the limits of a fragment corresponding to the tasiRNA sequence.

TAS1a, TAS1b, TAS1c, and TAS2 RNAs are not conservative (they are found only in *A. thaliana*) [163]. By contrast, TAS3 RNA is found in many mono- and dicotyledonous plants, but the similarity between supposed homologs is observed only in limits of the fragment encoding tasiRNA [166, 169].

In addition to 21-nt RNA, TAS1a and TAS2 RNAs produce several 24-nt siRNAs. Targets of these RNAs are not yet detected, but such small RNAs are known to be involved in transcriptional silencing in plants. It seems that 24-nt tasiRNAs are produced via an alternative pathway, because their processing does not require miRNAs but needs DCL3 [165], which is also a plant homolog of animal Dicer and is necessary for production of longer siRNAs involved in transcriptional silencing [173].

Although the production pathways and functions of tasiRNAs are rather well studied, many questions are still unanswered. Thus, it is unclear why just TAS RNA cutting products act as templates for RDR6 if miRNAs are involved in degradation of many RNAs. In the case of TAS1 and TAS2 RNAs, this can be associated with the miR173 binding site location within an intron [165], and the pathway of tasiRNA production is directed by proteins associated with unspliced RNA. It is also unclear why only one product of the TAS RNA cutting is transformed to the double-strained form and why it is either the 5′-terminal or 3′-terminal region of RNA [171].

The discovery of tasiRNAs has increased the number of known functions of miRNAs: they are shown not only



Mechanism of formation and action of tasiRNAs. Black vertical arrows indicate places of RNA cutting

Scheme 2

to regulate gene expression but also take part in formation of other regulatory RNAs.

PiRNAs. In 2006, five research groups reported the discovery of a new class of small RNAs [157, 160, 161, 174, 175]. These RNAs were found in mice, rats, and humans in association with immunoprecipitated proteins of the Piwi family and were termed piRNAs (Piwi-interacting RNAs). The Piwi family proteins belong to a large group of Argonaute proteins. In animals this group also includes proteins of the Ago family, which realize RNA silencing with involvement of miRNAs and siRNAs [115]. Proteins of the Piwi family are almost exclusively expressed in the germline and are required for maintaining germline stem cells, spermatogenesis, and repression of mobile element activities [159, 176, 177]. Male mice with mutations in each protein of the Piwi family (Mili, Miwi, and Miwi2) have defects in spermatogenesis and are sterile [176, 178, 179].

Similarly to proteins of the Piwi family, in mammals piRNAs are found only in testes. The length of piRNAs is 26-32 nt, i.e. they are longer than miRNAs and siRNAs. Two populations of piRNAs are likely to exist [157]: the first includes 26-28-nt RNAs and is found in association with the Mili protein [160], and the other is associated with the Miwi protein and comprises 28-32-nt RNAs [161, 175]. No piRNAs associated with the third member of the Piwi family (Miwi2) have been detected.

The piRNAs of both populations carry a phosphate group on the 5′-end. In 80-90% of cases U is the first nucleotide on the 5′-end [157, 160, 161, 174, 175]. The last nucleotide on the 3′-end is 2′-O-methylated [180, 181]. The significance of this modification is unknown, but miRNAs and siRNAs in plants are modified likewise, and it seems to increase their stability [116, 182, 183]. This modification in animals may play a similar role, because in mammalian cells 2′-O-methylated synthetic siRNAs are more stable than their unmodified analogs [184]. Thus, synthetic piRNAs 2′-O-methylated at the 3′-end are more stable than the unmodified ones on incubation with extract of mouse testes [180]. Moreover, 2′-O-methylation can promote piRNA binding with the Piwi family proteins because the Ago family proteins of animals have a low affinity for the modified RNAs [185]. It seems that piRNA of Drosophila and *D. rerio* are also 2′-O-methylated [177, 186].

More than 50,000 piRNAs are found in mammals [161], i.e. considerably more than the number of known small RNAs of other classes. Approximately 17% of piRNAs correspond to repeated sequences including mobile elements [160, 161, 174, 175]. Note that the number of piRNAs corresponding to repeats is lower than the proportion of repeated sequences in the genome (17% and ~42%, respectively, in rodents) [174, 175].

Other piRNAs are encoded by unique sequences which are gathered in clusters ranging from 1000 to ~100,000 bp. In 90% of cases the clusters are located within regions not containing annotated genes or repeats, but sometimes they reside within introns and exons, and in the latter case they correspond to the sense-chain of mRNA [161, 174, 175]. Each cluster includes from 10 to ~4500 piRNAs [160, 161, 175]. The piRNAs are encoded only by one DNA chain. Such an organization suggests that piRNA should be produced from long primary transcripts. This hypothesis is supported by existence of testis-specific ESTs and mRNAs corresponding to piRNA loci [157, 160, 174]. Thus processing of piRNAs should be different from processing of miRNAs and siRNAs. The existence of a specific processing manner is confirmed by the absence in the piRNA clusters of complicated secondary structures characteristic for pri-miRNAs [157, 160, 161, 174, 175]. Moreover, only sense sequences of unique piRNAs are found, which suggests the absence of dsRNA precursors [157].

In some cases piRNA clusters are located side by side, but are encoded by different chains (Scheme 3, panel (a)), which suggests bidirectional transcription from a common promoter [160]. The largest clusters of rodent piRNAs have orthologs in humans, although the piRNA sequences in the majority of cases are not alike [160, 161].

The level of piRNA expression changes during maturation of gametes. piRNAs can be detected when spermatocytes (diploid precursors of spermatozoa) enter the first meiosis division prophase, which is called a thick thread stage, or pachytene. Crossing-over occurs during this stage. The piRNA content strongly decreases upon production of spermatids (haploid products of meiosis) and they seem to be absent in the mature sperm [157, 160, 161, 175]. Thus, mammalian piRNAs are found only in maturing male gametes.

Functions of piRNAs and mechanisms of their processing are yet to be discovered. However, some data have been recently obtained as a result of studies on piRNAs of Drosophila. Drosophila has three proteins of the Piwi family: Piwi, Aubergine (Aub), and Ago3. As distinguished from the mammalian proteins, these proteins are found not only in testes but also in ovaries, and their expression is not limited to germline [158, 187-189]. As in mammals, Piwi proteins are required for maintaining germline stem cells [190] and gametogenesis [191]. Moreover, mutations in genes of these proteins result in activation of mobile element movements [192-194].

The piRNAs of Drosophila have been studied in detail [158]. Earlier 24-27-nt rasiRNAs were found in Drosophila testes [153]. Association of rasiRNAs with Piwi, Aub, and Ago3 proteins is shown in [186, 188, 189]. Thus the majority of Drosophila rasiRNAs may be assigned to piRNAs.

In Drosophila, as distinguished from mammals, ~80% of piRNAs correspond to repeated sequences of the genome: different mobile elements and DNA repeats of heterochromatin [158]. As in mammals, piRNA genes are

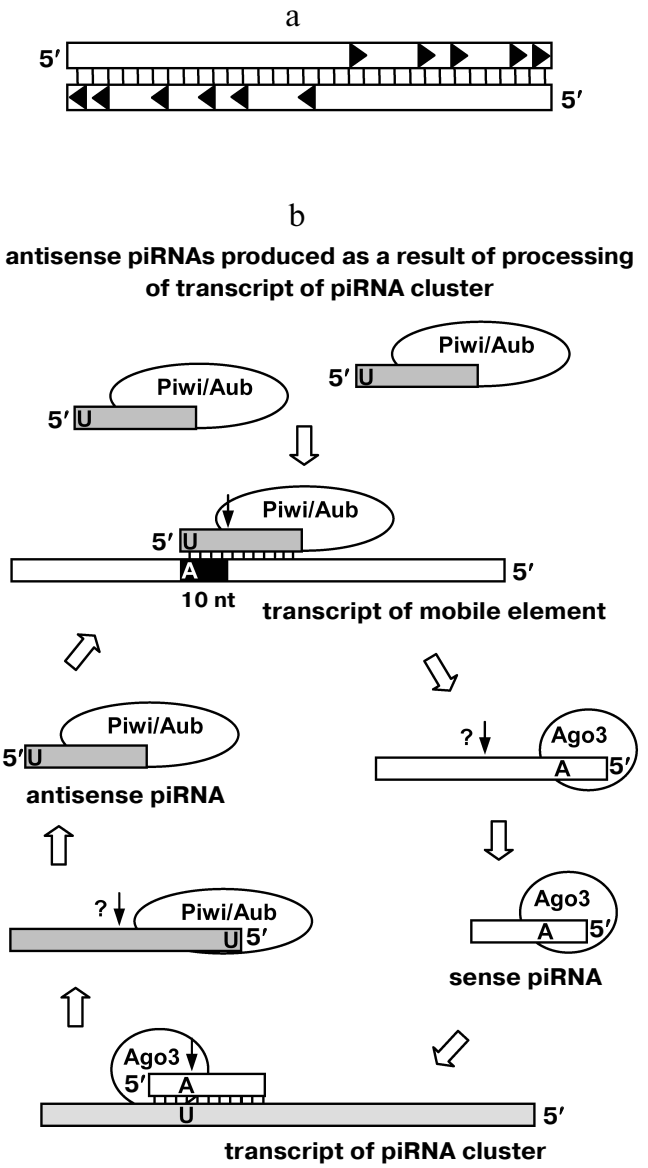
clustered, and a significant part of piRNAs corresponding to the repeats is also transcribed from a small number of loci.

The piRNAs of Drosophila are involved in repression of mobile element activities in the germline [186, 188]. The locus *flamenco* determining the repression of some mobile elements and containing their multiple copies has been shown to encode a large cluster of piRNAs [158]. Mutations in this locus are manifested by a decrease in expression of the encoded piRNAs and increase in the number of transcripts of a *flamenco*-controlled retro-element gypsy.

A model of repression of mobile element activities under the influence of piRNAs has been proposed [158, 189]. It was shown that in about 80% of cases Piwi and Aub proteins bind piRNAs complementary to mobile element sequences, i.e. antisense piRNAs, whereas Ago3 binds the sequences complementary to sense piRNAs [158]. Similar results are reported other works [186, 188, 189].

Molecules with complementary 5′-terminal regions are detected in populations of antisense- and sense piRNAs, and within a supposed duplex their 5′-ends are separated by ten nucleotides. In 83 and 72% of RNA molecules bound with Piwi and Aub, respectively, U is the first nucleotide on the 5′-end, and only 37% of Ago3-bound RNA molecules have U on the 5′-end.

To explain these findings, the following model has been proposed [158, 189]. Antisense piRNAs are produced from a transcript of a piRNA cluster, associate with Piwi or Aub, and direct the cutting of mobile element transcripts. Proteins of the Argonaute class, which include Piwi, Aub, and Ago3, cut the phosphodiester bond in a target RNA located opposite to the 10th and 11th nucleotides of the guide RNA. As a result, the 5′-end of one of the two produced fragments is separated from the 5′-end of antisense piRNA by just ten nucleotides (Scheme 3, panel (b)). This fragment seems to associate with Ago3, and a new sense piRNA is generated. The mechanism of its 3′-end processing is still unknown. By analogy with generation of siRNA, this sense piRNA can be called secondary. The Ago3 complex with the secondary piRNA cuts a target, which is more likely represented by a primary transcript of piRNA cluster that contains antisense sequences of mobile elements. In fact, transcripts of piRNA clusters have been found [158]. The processing mechanism of 3′-ends of new piRNAs is also unknown. The resulting antisense piRNAs can both perform silencing of mobile elements and direct generation of new sense piRNAs. This amplification allows the silencing to be strengthened in response to an increase in expression of mobile elements. This model takes into account the known properties of Piwi, Aub, and Ago3 proteins, in particular, their endonuclease activity [188, 189]. This model also explains why 5′-ends of Ago3-bound piRNAs are not enriched with U. Moreover, it predicts that the secondary piRNAs will be enriched with adenyl residues in position 10, because this position is complementary to the 5′-terminal U of antisense piRNA. In fact, in 73% of piRNAs bound with Ago3 the 10th nucleotide is represented by A.



PiRNA clusters within the genome (a) and the supposed mechanism of formation of secondary piRNAs (b). Black triangles indicate piRNA genes. Vertical arrows show places of RNA cutting. The sign "?" indicates action sites of nucleases that have not been established. A 10-nt region of complementary interaction between antisense piRNA and the resulting sense piRNA is shown in black within the mobile element transcript

Scheme 3

However, many questions are still without answers. Thus, it is unclear what determines the initial asymmetry in the piRNA binding by Piwi proteins, because the majority of loci encoding piRNAs contain transposon copies in both orientations.

Evidence in favor of involvement of mammalian piRNAs in repression of mobile element activities has been obtained [159]. Another population of piRNAs associated with Mili protein was detected by the same authors in mouse testes. As distinguished from the earlier described piRNAs, these RNAs were found in spermatogonia and later were present simultaneously with piRNAs that became detectable during the meiosis I pachytene. Prepachytenic piRNAs are also encoded by clusters, but the fraction of these molecules corresponding to DNA repeats (mainly mobile elements) is higher than the fraction of pachytenic piRNAs and reaches 35%.

Mutations in the Mili-encoding gene cause a five-to-ten-fold increase in expression of retro-elements L1 and IAP. Note that piRNAs corresponding to the L1 and IAP elements are enriched with complementary molecules, and in the supposed duplex their 5′-ends are separated by exactly 10 nucleotides. Moreover, sense piRNAs corresponding to the L1 and IAP sequences are enriched with A in position 10. Therefore, it is supposed that in mammals Piwi proteins and the piRNAs associated with them also contribute to the conservative mechanism of mobile element repression, and efficiency of this mechanism is additionally increased due to amplification of the initial signal.

A similar mechanism of repression of mobile element activities is also found in fishes (*Danio rerio*). But specific features are also revealed. Thus, piRNAs are found in both testes and ovaries of *D. rerio*, and their processing seems to differ from that in mammals.

A considerable part of piRNAs, especially in mammals, is not associated with mobile elements; thus, these RNAs are supposed to perform other functions that are still unclear [195, 196].

SMALL RNAs ASSOCIATED WITH PROTEIN-CODING GENES

The detection of several new groups of small RNAs was published in June 2007 [41]. In the course of studies on the human genome by GTA (repeated sequences excluded) the authors revealed many undescribed short RNAs (of 20-200 nt in length), and ~70% of them can be detected by Northern-blot hybridization and RT-PCR. Many of these RNAs are conservative, which suggests their biological significance. Sequences corresponding to ~40% of short RNAs are also found inside longer (>200 nt) polyadenylated RNAs that are encoded by the same loci as short RNAs. These longer RNAs are likely to be precursors of the short RNAs.

Sequences of many short RNAs were shown to correspond to promoter regions of protein-coding genes; therefore, they were termed promoter-associated small RNAs (PASRs). Short RNAs with sequences corresponding to 3′-terminal regions of the genes are also found. These RNAs are termed termini-associated small RNAs (TASRs). Both groups of RNAs are associated with ~50% of protein-coding genes, and the level of PASR expression is proportional to the expression level of the associated protein-coding gene. Therefore, the loci encoding the protein genes are suggested to also encode sets of ncRNAs. Similar data were obtained in another work [197]: ncRNAs were found which corresponded to the 5′-terminal region of the protein-coding genes. Thus, protein-encoding genes can be encircled by numerous ncRNAs that seem to perform regulatory functions [41].

SMALL RNAs ARE TRANSCRIPTS OF SINEs AND RELATED GENES

Genomes of mammals and other eukaryotes contain a huge number (10^5-10^6) of copies of DNA repeated sequences, which are called short retroposons or SINEs [198]. Their length is no more than 500 bp. SINEs multiply due to retroposition, which includes transcription of these elements by RNA polymerase III and reverse transcription coupled with integration of the synthesized DNA into a new region of the genome. During this process the role of reverse transcriptase is played by a polypeptide encoded by another type of mobile genetic elements, long interspersed elements (LINEs). Owing to accumulated mutations, SINE copies are different in 5-30% of the bases. The totality of such copies forms a family of SINEs. A genome usually includes no more than four SINE families, and each of them is specific for only one or several families (or orders) of organisms.

Two SINE families that were described first (B1 of rodents [199] and Alu of primates [200]) originated by evolution from 7SL RNA, a small cytoplasmic RNA involved in synthesis of secreted proteins [201]. The loss of a 182-nt stretch from the central region of the 7SL RNA nucleotide sequence was a crucial event in the origin of these SINEs. Alu consists of two sequences homologous to 7SL RNA and oriented similarly. Thus, Alu is a dimeric ~300-bp SINE with similar but not identical left and right monomers. B1 is a monomeric SINE, and its nucleotide sequence is more alike the left monomer of Alu than the right one [202].

Almost all other SINE families descend from tRNAs [198]. On the 5′-end of such SINEs there is a stretch slightly like the tRNA sequence. Nearby a stretch of unknown origin is usually located, which is specific for every SINE family. On the 3′-end of SINEs there is an A-rich tail (in mammals and plants) or a stretch consisting of some very short direct repeats (in reptiles and fishes). Some properties of SINEs and other mobile genetic elements suggest their selfish or parasitic character. Thus, they are characterized by a narrow taxonomic distribution, the ability to rapidly multiply within a genome and to destroy the gene functioning upon integration into them [198]. However, data have accumulated (although sometimes contradictory) indicating that small RNAs produced

as a result of SINE transcription by RNA polymerase III can perform definite functions in cells. Consider in more detail the putative functions of these small RNAs.

Various stresses (heat shock, viral infection, etc.) are associated with a multiple increase in the level of small RNAs transcribed from SINEs [203, 204]. These small RNAs were supposed to be important for cell survival under stress conditions. To test this hypothesis, human cells were cotransfected by two different plasmids, one of which bears an Alu-element and the other a reporter gene encoding luciferase [205]. Introduction into the cell of exogenous Alu-element (human SINE) not only increased the level of RNA transcribed from Alu by polymerase III (in the initial cells this level is close to zero), but also considerably strengthened synthesis of luciferase, although the content of its mRNA did not increase [205]. It was concluded that the translation had been stimulated by an increased content of Alu RNA. By analogy with small viral VAI RNA, the effect of Alu RNA on the translation was supposed to be mediated through double-stranded RNA-activated protein kinase R (PKR). Both Alu RNA and VAI RNA were shown to bind to PKR, which decreased the PKR activity [205]. This binding resulted in lowering of the PKR-catalyzed phosphorylation of the translation initiation factor eIF-2a, which in turn activated the translation. For stimulation of the translation, a stretch in the right monomer of Alu is essential, which has a nucleotide sequence specific for this monomer and lacking in the left monomer [206]. A high level of Alu RNA in the cells subjected to stress was supposed to be of help for survival of these cells due to activation of translation by the above-described pathway including inhibition of PKR and activation of eIF-2a [205].

However, this scheme is far from being confirmed by the further works of the same authors. First, the study on individual copies of Alu have shown that the transcription of many of them under heat shock conditions is not increased, but decreased [207], i.e. many Alu seem to be uninvolved in the cell response to stress. Second, the cotransfection of cells with the reporter gene and Alu resulted in the Alu-activated translation of not all mRNAs but of only mRNA transcribed from the reporter gene, and this effect disappears 22 h after the cotransfection [206]. This as yet mysterious phenomenon may be somehow associated with the ability of initiation factor eIF4G to discriminate new and old mRNAs [208]. Third, experiments with cells isolated from mice with knocked-out *PKR* gene have shown that the effect of Alu RNA (similarly to that of VAI RNA) on the translation does not depend on the PKR activity [206]. Thus, further studies are needed to elucidate the mechanism and confirm the biological significance of the translation stimulation under the influence of SINE transcripts.

A small RNA generated by the transcription of B2 copies (a tRNA-related SINE from genomes of mice and rodents of four related families [209]) can *in vitro* bind to RNA polymerase II and effectively inhibit it [210, 211]. This pathway is supposed to switch off transcription of the great bulk of genes, which is observed under conditions of heat shock. But it remains unclear how the genes encoding heat shock proteins can escape repression, considering that their transcription is pronouncedly activated at elevated temperature.

ID is a family of SINEs originated from alanine $tRNA^{(CGC)}$ and specific for the order of rodents. ID-element consists only of a tRNA-related sequence and an A-enriched tail. With ID used as a hybridization probe, a small cytoplasmic 160-nt RNA was found in the rodents' brain, which was called BC1 [212]. It occurred to be a transcript of an ID copy, which was termed BC1 RNA gene [213]. This RNA is very conservative in rodents. Many ID copies, in turn, are produced as a result of BC1 RNA retroposition due to its expression not only in the nervous tissue but also in the germline cells [213]. Knockout of the BC1 gene does not destroy the normal development of mice but weakens their explorative behavior and increases anxiety [214].

Small RNAs specific for the nervous tissue and testes are found in primates: these are BC200 in human [215] and anthropoids and G22 in lemurs of the Lorisidae family [216]. BC200 contains a monomeric Alu-element, whereas G22 RNA is formed by a sequence of a common dimeric Alu. Both RNAs have elongated A-rich stretches on their 3′-ends. The *BC200* and *G22* genes are supposed to have arisen independently as a result of integration of different Alu elements into the same region of the genomes of anthropoid and Lorisidae ancestors, respectively. The DNA sequence located upstream of the *BC200* and *G22* genes can determine the tissue specificity of their expression [217].

It seems that the three RNAs (BC1, BC200, and G22) perform similar functions. They are shown to be transferred into dendrites but not into neuronal axons [217, 218]. For transporting these RNAs, the protein PURα is important, which complexes with them. These complexes are produced under the influence of G-enriched stretches located in the 5′-terminal regions of these RNAs [219]. BC1 and BC200 RNAs are located in the regions of dendrites where mRNA translation occurs and seem to be involved in its regulation. BC1 can bind through its A-enriched 3′-terminal part to poly(A)-binding protein and the translation initiation factor eIF4A. Just this seems to explain the ability of BC1 to inhibit the translation [220]. There are also data [221], although slightly contradictory [220], indicating that translation of specific mRNAs is regulated under the influence of a BC1 RNA complex with a fragile X mental retardation protein (FMRP). This complex produced as a result of interaction of the *N*-end of FMRP with the 5′-terminal part of BC1 is supposed to bind to some RNAs due to complementary coupling between BC1 RNA and mRNA and induce inhibition of translation of this mRNA [221].

In all tissues of mice and rats, two small RNAs ($4.5S_H$ and $4.5S_I$) genetically related with SINEs and synthesized by RNA polymerase III are additionally found. The nucleotide sequence of $4.5S_H$, although shorter, is very much like that of the B1 element. The genes of $4.5S_H$ RNA are constituents of numerous long tandem repeats and, undoubtedly, are originated from B1 [222, 223]. The major part of $4.5S_H$ is located within the nucleus and is short-lived. Some molecules of $4.5S_H$ are associated with mRNA via complementary interactions [222]. The nucleotide sequence of $4.5S_I$ is slightly similar to those of the B2 and ID elements. Therefore, the genes of $4.5S_I$ seem to originate from one of these SINEs [209]. $4.5S_I$ RNA located within the nucleus is long-lived. $4.5S_H$ is found not only in rodents of the Muridae family but also in representatives of five other related families [223]. $4.5S_I$ is detected in rodents of only four families [224]. Thus, these RNAs are characterized by a narrow taxonomic distribution and recent appearance in the evolution of rodents. The nucleotide sequences of both RNAs are very conservative, which suggests some functions which are still unknown.

It is known that miRNAs are produced from long RNAs synthesized by RNA polymerase II. However, four human miRNAs were recently shown to be cut from small RNAs synthesized by RNA polymerase III as a result of transcription of Alu-element copies and the sequences flanking them from the 3′-end [225]. The miRNAs are produced from stretches located inside these 3′-terminal sequences of the transcripts. It is supposed that the human genome comprises dozens of such miRNA genes transcribed in Alu due to the presence of the promoter for RNA polymerase III.

All the above-presented examples show that at least individual copies of SINEs can become functional and act as genes of small noncoding RNAs. As a rule, these RNAs have a narrow taxonomic distribution; therefore, they are termed stenoRNAs [224]. These RNAs are supposed to be essential for arising of taxon-specific morphological, physiological, and behavioral features.

The detection of tens of thousands of new ncRNAs during the last decade has changed the current concepts about the role of RNA in the cell. RNAs are shown to perform surprisingly many functions and use earlier unknown mechanisms. The ncRNAs are involved in gene transcription regulation, splicing, and regulation of mRNA degradation. They are also involved in translation and its regulation, in processing and modification of ribosomal RNAs, in defense against viral infections and mutagenic activity of mobile genetic elements, etc. RNAs have obviously pressed proteins on their pedestal of the principal molecules responsible for vital activity of cells. Most likely, the world of ncRNAs only starts its opening for researchers, and the significance of these RNAs will be assessed in full measure only in the future.

We are grateful to N. S. Vassetzsky for his help.

This work was supported by the Russian Foundation for Basic Research (project No. 05-04-49553).

REFERENCES

1. Van Eenennaam, H., Jarrous, N., van Venrooij, W. J., and Pruijn, G. J. (2000) *IUBMB Life*, **49**, 265-272.
2. Huttenhofer, A., Kiefmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., Bachellerie, J. P., and Brosius, J. (2001) *EMBO J.*, **20**, 2943-2953.
3. Johnson, J. M., Edwards, S., Shoemaker, D., and Schadt, E. E. (2005) *Trends Genet.*, **21**, 93-102.
4. Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G., Sementchenko, V., Piccolboni, A., Bekiranov, S., Bailey, D., Ganesh, M., Ghosh, S., Bell, I., Gerhard, D. S., and Gingeras, T. (2005) *Science*, **308**, 1149-1154.
5. International Human Genome Sequencing Consortium (2004) *Nature*, **431**, 931-945.
6. Schadt, E. E., Edwards, S. W., GuhaThakurta, D., Holder, D., Ying, L., Svetnik, V., Leonardson, A., Hart, K. W., Russell, A., Li, G., Cavet, G., Castle, J., McDonagh, P., Kan, Z., Chen, R., Kasarskis, A., Margarint, M., Caceres, R. M., Johnson, J. M., Armour, C. D., Garrett-Engele, P. W., Tsinoremas, N. F., and Shoemaker, D. D. (2004) *Genome Biol.*, **5**, R73.
7. Rinn, J. L., Euskirchen, G., Bertone, P., Martone, R., Luscombe, N. M., Hartman, S., Harrison, P. M., Nelson, F. K., Miller, P., Gerstein, M., Weissman, S., and Snyder, M. (2003) *Genes Dev.*, **17**, 529-540.
8. Kampa, D., Cheng, J., Kapranov, P., Yamanaka, M., Brubaker, S., Cawley, S., Drenkow, J., Piccolboni, A., Bekiranov, S., Helt, G., Tammana, H., and Gingeras, T. R. (2004) *Genome Res.*, **14**, 331-342.
9. Bertone, P., Stolc, V., Royce, T. E., Rozowsky, J. S., Urban, A. E., Zhu, X., Rinn, J. L., Tongprasit, W., Samanta, M., Weissman, S., Gerstein, M., and Snyder, M. (2004) *Science*, **306**, 2242-2246.
10. Stolc, V., Gauhar, Z., Mason, C., Halasz, G., van Batenburg, M. F., Rifkin, S. A., Hua, S., Herreman, T., Tongprasit, W., Barbano, P. E., Bussemaker, H. J., and White, K. P. (2004) *Science*, **306**, 655-660.
11. Manak, J. R., Dike, S., Sementchenko, V., Kapranov, P., Biemar, F., Long, J., Cheng, J., Bell, I., Ghosh, S., Piccolboni, A., and Gingeras, T. R. (2006) *Nat. Genet.*, **38**, 1151-1158.
12. Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J., Southwick, A. M., Wu, H. C., Kim, C., Nguyen, M., Pham, P., Cheuk, R., Karlin-Newmann, G., Liu, S. X., Lam, B., Sakano, H., et al. (2003) *Science*, **302**, 842-846.
13. Chen, J., Sun, M., Lee, S., Zhou, G., Rowley, J. D., and Wang, S. M. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 12257-12262.
14. Saha, S., Sparks, A. B., Rago, C., Akmaev, V., Wang, C. J., Vogelstein, B., Kinzler, K. W., and Velculescu, V. E. (2002) *Nat. Biotechnol.*, **20**, 508-512.
15. Jongeneel, C. V., Delorenzi, M., Iseli, C., Zhou, D., Haudenschild, C. D., Khrebtukova, I., Kuznetsov, D., Stevenson, B. J., Strausberg, R. L., Simpson, A. J., and Vasicek, T. J. (2005) *Genome Res.*, **15**, 1007-1014.

16. Semon, M., and Duret, L. (2004) *Trends Genet.*, **20**, 229-232.
17. Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., et al. (2004) *Nat. Genet.*, **36**, 40-45.
18. Washietl, S., Hofacker, I. L., Lukasser, M., Huttenhofer, A., and Stadler, P. F. (2005) *Nat. Biotechnol.*, **23**, 1383-1390.
19. Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V. B., Brenner, S. E., Batalov, S., Forrest, A. R., Zavolan, M., Davis, M. J., Wilming, L. G., Aidinis, V., Allen, J. E., et al. (2005) *Science*, **309**, 1559-1563.
20. Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B., and Wahlestedt, C. (2005) *Science*, **309**, 1564-1566.
21. Dennis, C. (2002) *Nature*, **418**, 122-124.
22. Pang, K. C., Frith, M. C., and Mattick, J. S. (2006) *Trends Genet.*, **22**, 1-5.
23. Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E., Sharon, E., Spector, Y., and Bentwich, Z. (2005) *Nat. Genet.*, **37**, 766-770.
24. Ravasi, T., Suzuki, H., Pang, K. C., Katayama, S., Furuno, M., Okunishi, R., Fukuda, S., Ru, K., Frith, M. C., Gongora, M. M., Grimmond, S. M., Hume, D. A., Hayashizaki, Y., and Mattick, J. S. (2006) *Genome Res.*, **16**, 11-19.
25. Cawley, S., Bekiranov, S., Ng, H. H., Kapranov, P., Sekinger, E. A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A. J., Wheeler, R., Wong, B., Drenkow, J., Yamanaka, M., Patel, S., Brubaker, S., Tammana, H., Helt, G., Struhl, K., and Gingeras, T. R. (2004) *Cell*, **116**, 499-509.
26. Mattick, J. S., and Makunin, I. V. (2006) *Hum. Mol. Genet.*, R17-R29.
27. Razin, S. V., Rynditch, A., Borunova, V., Ioudinkova, E., Smalko, V., and Scherrer, K. (2004) *J. Cell. Biochem.*, **92**, 445-457.
28. Werner, A., and Berdal, A. (2005) *Physiol. Genom.*, **23**, 125-131.
29. Kumar, M., and Carmichael, G. G. (1998) *Microbiol. Mol. Biol. Rev.*, **62**, 1415-1434.
30. Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R., and Willard, H. F. (1991) *Nature*, **349**, 38-44.
31. Lee, J. T., Davidow, L. S., and Warshawsky, D. (1999) *Nat. Genet.*, **21**, 400-404.
32. Franke, A., and Baker, B. S. (1999) *Mol. Cell*, **4**, 117-122.
33. Young, T. L., Matsuda, T., and Cepko, C. L. (2005) *Curr. Biol.*, **15**, 501-512.
34. Tycowski, K. T., Shu, M. D., and Steitz, J. A. (1996) *Nature*, **379**, 464-466.
35. Smith, C. M., and Steitz, J. A. (1998) *Mol. Cell. Biol.*, **18**, 6897-6909.
36. Pelczar, P., and Filipowicz, W. (1998) *Mol. Cell. Biol.*, **18**, 4509-4518.
37. Tanaka, R., Satoh, H., Moriyama, M., Satoh, K., Morishita, Y., Yoshida, S., Watanabe, T., Nakamura, Y., and Mori, S. (2000) *Genes Cells*, **5**, 277-287.
38. Bortolin, M. L., and Kiss, T. (1998) *RNA*, **4**, 445-454.
39. Makarova, J. A., and Kramerov, D. A. (2005) *Gene*, **363**, 51-60.
40. Rougeulle, C., and Heard, E. (2002) *Trends Genet.*, **18**, 434-437.
41. Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., Stadler, P. F., Hertel, J., Hackermueller, J., Hofacker, I. L., Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccolboni, A., Sementchenko, V., Tammana, H., and Gingeras, T. R. (2007) *Science*, **316**, 1484-1488.
42. Du, T., and Zamore, P. D. (2005) *Development*, **132**, 4645-4652.
43. Pang, K. C., Stephen, S., Engstrom, P. G., Tajul-Arifin, K., Chen, W., Wahlestedt, C., Lenhard, B., Hayashizaki, Y., and Mattick, J. S. (2005) *Nucleic Acids Res.*, **33**, D125-D130.
44. Willingham, A. T., Orth, A. P., Batalov, S., Peters, E. C., Wen, B. G., Aza-Blanc, P., Hogenesch, J. B., and Schultz, P. G. (2005) *Science*, **309**, 1570-1573.
45. Lanz, R. B., McKenna, N. J., Onate, S. A., Albrecht, U., Wong, J., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (1999) *Cell*, **97**, 17-27.
46. Aranda, A., and Pascual, A. (2001) *Physiol. Rev.*, **81**, 1269-1304.
47. Glass, C. K., and Rosenfeld, M. G. (2000) *Genes Dev.*, **14**, 121-141.
48. Lanz, R. B., Chua, S. S., Barron, N., Soder, B. M., DeMayo, F., and O'Malley, B. W. (2003) *Mol. Cell. Biol.*, **23**, 7163-7176.
49. Murphy, L. C., Simon, S. L., Parkes, A., Leygue, E., Dotzlaw, H., Snell, L., Troup, S., Adeyinka, A., and Watson, P. H. (2000) *Cancer Res.*, **60**, 6266-6271.
50. Lanz, R. B., Razani, B., Goldberg, A. D., and O'Malley, B. W. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 16081-16086.
51. Xu, B., and Koenig, R. J. (2004) *J. Biol. Chem.*, **279**, 33051-33056.
52. Deblois, G., and Giguere, V. (2003) *J. Steroid. Biochem. Mol. Biol.*, **85**, 123-131.
53. Chen, D., Riedl, T., Washbrook, E., Pace, P. E., Coombes, R. C., Egly, J. M., and Ali, S. (2000) *Mol. Cell*, **6**, 127-137.
54. Watanabe, M., Yanagisawa, J., Kitagawa, H., Takeyama, K., Ogawa, S., Arao, Y., Suzawa, M., Kobayashi, Y., Yano, T., Yoshikawa, H., Masuhiro, Y., and Kato, S. (2001) *EMBO J.*, **20**, 1341-1352.
55. Zhao, X., Patton, J. R., Davis, S. L., Florence, B., Ames, S. J., and Spanjaard, R. A. (2004) *Mol. Cell*, **15**, 549-558.
56. Zhao, X., Patton, J. R., Ghosh, S. K., Fischel-Ghodsian, N., Shen, L., and Spanjaard, R. A. (2007) *Mol. Endocrinol.*, **21**, 686-699.
57. Lonard, D. M., and O'Malley, B. W. (2006) *Cell*, **125**, 411-414.
58. Shi, Y., Downes, M., Xie, W., Kao, H. Y., Ordentlich, P., Tsai, C. C., Hon, M., and Evans, R. M. (2001) *Genes Dev.*, **15**, 1140-1151.
59. Hatchell, E. C., Colley, S. M., Beveridge, D. J., Epis, M. R., Stuart, L. M., Giles, K. M., Redfern, A. D., Miles, L.

E., Barker, A., MacDonald, L. M., Arthur, P. G., Lui, J. C., Golding, J. L., McCulloch, R. K., Metcalf, C. B., Wilce, J. A., Wilce, M. C., Lanz, R. B., O'Malley, B. W., and Leedman, P. J. (2006) *Mol. Cell*, **22**, 657-668.
60. Emberley, E., Huang, G. J., Hamedani, M. K., Czosnek, A., Ali, D., Grolla, A., Lu, B., Watson, P. H., Murphy, L. C., and Leygue, E. (2003) *Biochem. Biophys. Res. Commun.*, **301**, 509-515.
61. Chooniedass-Kothari, S., Hamedani, M. K., Troup, S., Hube, F., and Leygue, E. (2006) *Int. J. Cancer*, **118**, 1054-1059.
62. Zieve, G., and Penman, S. (1976) *Cell*, **8**, 19-31.
63. Zhou, Q., and Yik, J. H. (2006) *Microbiol. Mol. Biol. Rev.*, **70**, 646-659.
64. Kuwabara, T., Hsieh, J., Nakashima, K., Taira, K., and Gage, F. H. (2004) *Cell*, **116**, 779-793.
65. Schoenherr, C. J., and Anderson, D. J. (1995) *Science*, **267**, 1360-1363.
66. Bruce, A. W., Donaldson, I. J., Wood, I. C., Yerbury, S. A., Sadowski, M. I., Chapman, M., Gottgens, B., and Buckley, N. J. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 10458-10463.
67. Kraner, S. D., Chong, J. A., Tsay, H. J., and Mandel, G. (1992) *Neuron*, **9**, 37-44.
68. Mori, N., Schoenherr, C., Vandenbergh, D. J., and Anderson, D. J. (1992) *Neuron*, **9**, 45-54.
69. Huang, Y., Myers, S. J., and Dingledine, R. (1999) *Nat. Neurosci.*, **2**, 867-872.
70. Chong, J. A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J. J., Zheng, Y., Boutros, M. C., Altshuller, Y. M., Frohman, M. A., Kraner, S. D., and Mandel, G. (1995) *Cell*, **80**, 949-957.
71. Ballas, N., Grunseich, C., Lu, D. D., Speh, J. C., and Mandel, G. (2005) *Cell*, **121**, 645-657.
72. Sun, Y. M., Greenway, D. J., Johnson, R., Street, M., Belyaev, N. D., Deuchars, J., Bee, T., Wilde, S., and Buckley, N. J. (2005) *Mol. Biol. Cell*, **16**, 5630-5638.
73. Palm, K., Belluardo, N., Metsis, M., and Timmusk, T. (1998) *J. Neurosci.*, **18**, 1280-1296.
74. Calderone, A., Jover, T., Noh, K. M., Tanaka, H., Yokota, H., Lin, Y., Grooms, S. Y., Regis, R., Bennett, M. V., and Zukin, R. S. (2003) *J. Neurosci.*, **23**, 2112-2121.
75. Cao, X., Yeo, G., Muotri, A. R., Kuwabara, T., and Gage, F. H. (2006) *Annu. Rev. Neurosci.*, **29**, 77-103.
76. Weinstein, L. B., and Steitz, J. A. (1999) *Curr. Opin. Cell. Biol.*, **11**, 378-384.
77. Omer, A. D., Lowe, T. M., Russell, A. G., Ebhardt, H., Eddy, S. R., and Dennis, P. P. (2000) *Science*, **288**, 517-522.
78. Kiss-Laszlo, Z., Henry, Y., and Kiss, T. (1998) *EMBO J.*, **17**, 797-807.
79. Samarsky, D. A., Fournier, M. J., Singer, R. H., and Bertrand, E. (1998) *EMBO J.*, **17**, 3747-3757.
80. Gautier, T., Berges, T., Tollervey, D., and Hurt, E. (1997) *Mol. Cell. Biol.*, **17**, 7088-7098.
81. Filipowicz, W., and Pogacic, V. (2002) *Curr. Opin. Cell. Biol.*, **14**, 319-327.
82. Kiss, T. (2004) *J. Cell. Sci.*, **117**, 5949-5951.
83. Tollervey, D. (1996) *Science*, **273**, 1056-1057.
84. Watkins, N. J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M., and Luhrmann, R. (1998) *RNA*, **4**, 1549-1568.
85. Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gelugne, J. P., and Caizergues-Ferrer, M. (1998) *EMBO J.*, **17**, 7078-7090.
86. Ganot, P., Bortolin, M. L., and Kiss, T. (1997) *Cell*, **89**, 799-809.
87. Gerbi, S. A., Borovjagin, A. V., and Lange, T. S. (2003) *Curr. Opin. Cell. Biol.*, **15**, 318-325.
88. Tollervey, D., and Kiss, T. (1997) *Curr. Opin. Cell. Biol.*, **9**, 337-342.
89. Mishra, R. K., and Eliceiri, G. L. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 4972-4977.
90. Tycowski, K. T., You, Z. H., Graham, P. J., and Steitz, J. A. (1998) *Mol. Cell*, **2**, 629-638.
91. Peculis, B. (1997) *Curr. Biol.*, **7**, R480-R482.
92. Runte, M., Huttenhofer, A., Gross, S., Kiefmann, M., Horsthemke, B., and Buiting, K. (2001) *Hum. Mol. Genet.*, **10**, 2687-2700.
93. Vitali, P., Royo, H., Seitz, H., Bachellerie, J. P., Huttenhofer, A., and Cavaille, J. (2003) *Nucleic Acids Res.*, **31**, 6543-6551.
94. Kiss, A. M., Jady, B. E., Bertrand, E., and Kiss, T. (2004) *Mol. Cell. Biol.*, **24**, 5797-5807.
95. Kishore, S., and Stamm, S. (2006) *Science*, **311**, 230-232.
96. Makarova, J. A., and Kramerov, D. A. (2007) *Mol. Biol.*, **41**, 246-259.
97. Maxwell, E. S., and Fournier, M. J. (1995) *Annu. Rev. Biochem.*, **64**, 897-934.
98. Kiss, T., and Filipowicz, W. (1995) *Genes Dev.*, **9**, 1411-1424.
99. Caffarelli, E., Arese, M., Santoro, B., Fragapane, P., and Bozzoni, I. (1994) *Mol. Cell. Biol.*, **14**, 2966-2974.
100. Huang, Z. P., Zhou, H., Liang, D., and Qu, L. H. (2004) *J. Mol. Biol.*, **341**, 669-683.
101. Brown, J. W., Echeverria, M., and Qu, L. H. (2003) *Trends Plant. Sci.*, **8**, 42-49.
102. Makarova, J. A., and Kramerov, D. A. (2007) *Genetika*, **43**, 149-158.
103. Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., and Kiss, T. (2002) *EMBO J.*, **21**, 2746-2756.
104. Gall, J. G. (2000) *Annu. Rev. Cell. Dev. Biol.*, **16**, 273-300.
105. Jady, B. E., and Kiss, T. (2001) *EMBO J.*, **20**, 541-551.
106. Richard, P., Darzacq, X., Bertrand, E., Jady, B. E., Verheggen, C., and Kiss, T. (2003) *EMBO J.*, **22**, 4283-4293.
107. Cai, X., Hagedorn, C. H., and Cullen, B. R. (2004) *RNA*, **10**, 1957-1966.
108. Devor, E. J. (2006) *J. Hered.*, **97**, 186-190.
109. Smalheiser, N. R., and Torvik, V. I. (2005) *Trends Genet.*, **21**, 322-326.
110. Kloosterman, W. P., and Plasterk, R. H. (2006) *Dev. Cell*, **11**, 441-450.
111. Scaria, V., Hariharan, M., Maiti, S., Pillai, B., and Brahmachari, S. K. (2006) *Retrovirology*, **3**, 68.
112. Qi, P., Han, J. X., Lu, Y. Q., Wang, C., and Bu, F. F. (2006) *Cell. Mol. Immunol.*, **3**, 411-419.
113. Carmell, M. A., and Hannon, G. J. (2004) *Nat. Struct. Mol. Biol.*, **11**, 214-218.
114. Tomari, Y., and Zamore, P. D. (2005) *Genes Dev.*, **19**, 517-529.
115. Parker, J. S., and Barford, D. (2006) *Trends Biochem. Sci.*, **31**, 622-630.
116. Vaucheret, H. (2006) *Genes Dev.*, **20**, 759-771.
117. Kotel'nikov, R. N., Shpiz, S. G., Kalmykova, A. I., and Gvozdev, V. A. (2006) *Mol. Biol.*, **40**, 595-608.
118. Vilgelm, A. E., Chumakov, S. P., and Prasolov, V. S. (2006) *Mol. Biol.*, **40**, 387-403.

119. Valencia-Sanchez, M. A., Liu, J., Hannon, G. J., and Parker, R. (2006) *Genes Dev.*, **20**, 515-524.
120. Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., and Bartel, D. P. (2002) *Cell*, **110**, 513-520.
121. Pillai, R. S., Bhattacharyya, S. N., and Filipowicz, W. (2007) *Trends Cell Biol.*, **17**, 118-126.
122. Anderson, P., and Kedersha, N. (2006) *J. Cell. Biol.*, **172**, 803-808.
123. Sheth, U., and Parker, R. (2003) *Science*, **300**, 805-808.
124. Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., and Enright, A. J. (2006) *Nucleic Acids Res.*, **34**, D140-D144.
125. Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R. H., and Cuppen, E. (2005) *Cell*, **120**, 21-24.
126. John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., and Marks, D. S. (2004) *PLoS Biol.*, **2**, e363.
127. Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005) *Cell*, **120**, 15-20.
128. Bao, N., Lye, K. W., and Barton, M. K. (2004) *Dev. Cell*, **7**, 653-662.
129. Bartel, D. P., and Chen, C. Z. (2004) *Nat. Rev. Genet.*, **5**, 396-400.
130. Stark, A., Brennecke, J., Bushati, N., Russell, R. B., and Cohen, S. M. (2005) *Cell*, **123**, 1133-1146.
131. Kidner, C. A., and Martienssen, R. A. (2005) *Curr. Opin. Plant. Biol.*, **8**, 38-44.
132. Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., and Golub, T. R. (2005) *Nature*, **435**, 834-838.
133. Garzon, R., Fabbri, M., Cimmino, A., Calin, G. A., and Croce, C. M. (2006) *Trends Mol. Med.*, **12**, 580-587.
134. Zhang, B., Pan, X., Cobb, G. P., and Anderson, T. A. (2007) *Dev. Biol.*, **302**, 1-12.
135. Meister, G., Landthaler, M., Dorsett, Y., and Tuschl, T. (2004) *RNA*, **10**, 544-550.
136. Hutvagner, G., Simard, M. J., Mello, C. C., and Zamore, P. D. (2004) *PLoS Biol.*, **2**, E98.
137. Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., and Stoffel, M. (2005) *Nature*, **438**, 685-689.
138. Orom, U. A., Kauppinen, S., and Lund, A. H. (2006) *Gene*, **372**, 137-141.
139. Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001) *Genes Dev.*, **15**, 188-200.
140. Sharp, P. A. (2001) *Genes Dev.*, **15**, 485-490.
141. Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H., and Fire, A. (2001) *Cell*, **107**, 465-476.
142. Bernstein, E., and Allis, C. D. (2005) *Genes Dev.*, **19**, 1635-1655.
143. Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I., and Moazed, D. (2004) *Science*, **303**, 672-676.
144. Kawasaki, H., and Taira, K. (2004) *Nature*, **431**, 211-217.
145. Morris, K. V., Chan, S. W., Jacobsen, S. E., and Looney, D. J. (2004) *Science*, **305**, 1289-1292.
146. Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. (2002) *EMBO J.*, **21**, 4671-4679.
147. Klenov, M. S., and Gvozdev, V. A. (2005) *Biochemistry (Moscow)*, **70**, 1187-1198.
148. Gvozdev, V. A. (2003) *Genetika*, **39**, 151-156.
149. Aravin, A. A., Vagin, V. V., Naumova, N. M., Rozovskii, Ya. M., Klenov, M. S., and Gvozdev, V. A. (2002) *Ontogenez*, **33**, 349-360.
150. Aravin, A. A., Klenov, M. S., Vagin, V. V., Rozovskii, Ya. M., and Gvozdev, V. A. (2002) *Mol. Biol. (Moscow)*, **36**, 240-251.
151. Djikeng, A., Shi, H., Tschudi, C., and Ullu, E. (2001) *RNA*, **7**, 1522-1530.
152. Sijen, T., and Plasterk, R. H. (2003) *Nature*, **426**, 310-314.
153. Aravin, A. A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. (2003) *Dev. Cell*, **5**, 337-350.
154. Mette, M. F., van der Winden, J., Matzke, M., and Matzke, A. J. (2002) *Plant. Physiol.*, **130**, 6-9.
155. Llave, C., Kasschau, K. D., Rector, M. A., and Carrington, J. C. (2002) *Plant. Cell*, **14**, 1605-1619.
156. Chen, P. Y., Manninga, H., Slanchev, K., Chien, M., Russo, J. J., Ju, J., Sheridan, R., John, B., Marks, D. S., Gaidatzis, D., Sander, C., Zavolan, M., and Tuschl, T. (2005) *Genes Dev.*, **19**, 1288-1293.
157. Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006) *Genes Dev.*, **20**, 1732-1743.
158. Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G. J. (2007) *Cell*, **128**, 1089-1103.
159. Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G. J. (2007) *Science*, **316**, 744-747.
160. Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M. J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J. J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., and Tuschl, T. (2006) *Nature*, **442**, 203-207.
161. Girard, A., Sachidanandam, R., Hannon, G. J., and Carmell, M. A. (2006) *Nature*, **442**, 199-202.
162. Talmor-Neiman, M., Stav, R., Klipcan, L., Buxdorf, K., Baulcombe, D. C., and Arazi, T. (2006) *Plant J.*, **48**, 511-521.
163. Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A. C., Hilbert, J. L., Bartel, D. P., and Crete, P. (2004) *Mol. Cell*, **16**, 69-79.
164. Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H. L., and Poethig, R. S. (2004) *Genes Dev.*, **18**, 2368-2379.
165. Yoshikawa, M., Peragine, A., Park, M. Y., and Poethig, R. S. (2005) *Genes Dev.*, **19**, 2164-2175.
166. Allen, E., Xie, Z., Gustafson, A. M., and Carrington, J. C. (2005) *Cell*, **121**, 207-221.
167. Van de Sande, K., Pawlowski, K., Czaja, I., Wieneke, U., Schell, J., Schmidt, J., Walden, R., Matvienko, M., Wellink, J., van Kammen, A., Franssen, H., and Bisseling, T. (1996) *Science*, **273**, 370-373.
168. Rohrig, H., Schmidt, J., Miklashevichs, E., Schell, J., and John, M. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 1915-1920.
169. Williams, L., Carles, C. C., Osmont, K. S., and Fletcher, J. C. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 9703-9708.
170. Xie, Z., Allen, E., Wilken, A., and Carrington, J. C. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 12984-12989.
171. Vaucheret, H. (2005) *Sci. STKE*, **2005**, pe43.
172. Hunter, C., Willmann, M. R., Wu, G., Yoshikawa, M., de la Luz Gutierrez-Nava, M., and Poethig, S. R. (2006) *Development*, **133**, 2973-2981.

173. Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., Jacobsen, S. E., and Carrington, J. C. (2004) *PLoS Biol.*, **2**, E104.
174. Lau, N. C., Seto, A. G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D. P., and Kingston, R. E. (2006) *Science*, **313**, 363-367.
175. Grivna, S. T., Beyret, E., Wang, Z., and Lin, H. (2006) *Genes Dev.*, **20**, 1709-1714.
176. Carmell, M. A., Girard, A., van de Kant, H. J., Bourc'his, D., Bestor, T. H., de Rooij, D. G., and Hannon, G. J. (2007) *Dev. Cell*, **12**, 503-514.
177. Houwing, S., Kamminga, L. M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D. V., Blaser, H., Raz, E., Moens, C. B., Plasterk, R. H., Hannon, G. J., Draper, B. W., and Ketting, R. F. (2007) *Cell*, **129**, 69-82.
178. Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T. W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W., Lin, H., Matsuda, Y., and Nakano, T. (2004) *Development*, **131**, 839-849.
179. Deng, W., and Lin, H. (2002) *Dev. Cell*, **2**, 819-830.
180. Kirino, Y., and Mourelatos, Z. (2007) *Nat. Struct. Mol. Biol.*, **14**, 347-348.
181. Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., and Suzuki, T. (2007) *Nat. Struct. Mol. Biol.*, **14**, 349-350.
182. Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., Steward, R., and Chen, X. (2005) *Science*, **307**, 932-935.
183. Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005) *Curr. Biol.*, **15**, 1501-1507.
184. Czauderna, F., Fechtner, M., Dames, S., Aygun, H., Klippel, A., Pronk, G. J., Giese, K., and Kaufmann, J. (2003) *Nucleic Acids Res.*, **31**, 2705-2716.
185. Ma, J. B., Ye, K., and Patel, D. J. (2004) *Nature*, **429**, 318-322.
186. Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P. D. (2006) *Science*, **313**, 320-324.
187. Cox, D. N., Chao, A., and Lin, H. (2000) *Development*, **127**, 503-514.
188. Saito, K., Nishida, K. M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M. C. (2006) *Genes Dev.*, **20**, 2214-2222.
189. Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M. C. (2007) *Science*, **315**, 1587-1590.
190. Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998) *Genes Dev.*, **12**, 3715-3727.
191. Lin, H., and Spradling, A. C. (1997) *Development*, **124**, 2463-2476.
192. Savitsky, M., Kwon, D., Georgiev, P., Kalmykova, A., and Gvozdev, V. (2006) *Genes Dev.*, **20**, 345-354.
193. Kalmykova, A. I., Klenov, M. S., and Gvozdev, V. A. (2005) *Nucleic Acids Res.*, **33**, 2052-2059.
194. Sarot, E., Payen-Groschene, G., Bucheton, A., and Pelisson, A. (2004) *Genetics*, **166**, 1313-1321.
195. O'Donnell, K. A., and Boeke, J. D. (2007) *Cell*, **129**, 37-44.
196. Lin, H. (2007) *Science*, **316**, 397.
197. Davis, C. A., and Ares, M., Jr. (2006) *Proc. Natl. Acad. Sci. USA*, **103**, 3262-3267.
198. Kramerov, D., and Vassetzky, N. (2005) *Int. Rev. Cytol.*, **247**, 165-221.
199. Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov, A. P., Bayev, A. A., and Georgiev, G. P. (1980) *Nucleic Acids Res.*, **8**, 1201-1215.
200. Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedmann, T., and Schmid, C. W. (1981) *J. Mol. Biol.*, **151**, 17-33.
201. Ullu, E., and Tschudi, C. (1984) *Nature*, **312**, 171-172.
202. Quentin, Y. (1994) *Genetica*, **93**, 203-215.
203. Fornace, A. J., Jr., and Mitchell, J. B. (1986) *Nucleic Acids Res.*, **14**, 5793-5811.
204. Liu, W. M., Chu, W. M., Choudary, P. V., and Schmid, C. W. (1995) *Nucleic Acids Res.*, **23**, 1758-1765.
205. Chu, W. M., Ballard, R., Carpick, B. W., Williams, B. R., and Schmid, C. W. (1998) *Mol. Cell. Biol.*, **18**, 58-68.
206. Rubin, C. M., Kimura, R. H., and Schmid, C. W. (2002) *Nucleic Acids Res.*, **30**, 3253-3261.
207. Li, T. H., and Schmid, C. W. (2001) *Gene*, **276**, 135-141.
208. Novoa, I., and Carrasco, L. (1999) *Mol. Cell. Biol.*, **19**, 2445-2454.
209. Serdobova, I. M., and Kramerov, D. A. (1998) *J. Mol. Evol.*, **46**, 202-214.
210. Allen, T. A., von Kaenel, S., Goodrich, J. A., and Kugel, J. F. (2004) *Nat. Struct. Mol. Biol.*, **11**, 816-821.
211. Espinoza, C. A., Allen, T. A., Hieb, A. R., Kugel, J. F., and Goodrich, J. A. (2004) *Nat. Struct. Mol. Biol.*, **11**, 822-829.
212. Martignetti, J. A., and Brosius, J. (1995) *Mol. Cell. Biol.*, **15**, 1642-1650.
213. Kim, J., Martignetti, J. A., Shen, M. R., Brosius, J., and Deininger, P. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 3607-3611.
214. Lewejohann, L., Skryabin, B. V., Sachser, N., Prehn, C., Heiduschka, P., Thanos, S., Jordan, U., Dell'Omo, G., Vyssotski, A. L., Pleskacheva, M. G., Lipp, H. P., Tiedge, H., Brosius, J., and Prior, H. (2004) *Behav. Brain Res.*, **154**, 273-289.
215. Tiedge, H., Chen, W., and Brosius, J. (1993) *J. Neurosci.*, **13**, 2382-2390.
216. Ludwig, A., Rozhdestvensky, T. S., Kuryshev, V. Y., Schmitz, J., and Brosius, J. (2005) *J. Mol. Biol.*, **350**, 200-214.
217. Khanam, T., Rozhdestvensky, T. S., Bundman, M., Galiveti, C. R., Handel, S., Sukonina, V., Jordan, U., Brosius, J., and Skryabin, B. V. (2007) *Nucleic Acids Res.*, **35**, 529-539.
218. Muslimov, I. A., Santi, E., Homel, P., Perini, S., Higgins, D., and Tiedge, H. (1997) *J. Neurosci.*, **17**, 4722-4733.
219. Johnson, E. M., Kinoshita, Y., Weinreb, D. B., Wortman, M. J., Simon, R., Khalili, K., Winckler, B., and Gordon, J. (2006) *J. Neurosci. Res.*, **83**, 929-943.
220. Wang, H., Iacoangeli, A., Lin, D., Williams, K., Denman, R. B., Hellen, C. U., and Tiedge, H. (2005) *J. Cell. Biol.*, **171**, 811-821.
221. Zalfa, F., Adinolfi, S., Napoli, I., Kuhn-Holsken, E., Urlaub, H., Achsel, T., Pastore, A., and Bagni, C. (2005) *J. Biol. Chem.*, **280**, 33403-33410.
222. Schoeniger, L. O., and Jelinek, W. R. (1986) *Mol. Cell. Biol.*, **6**, 1508-1519.
223. Gogolevskaya, I. K., Koval, A. P., and Kramerov, D. A. (2005) *Mol. Biol. Evol.*, **22**, 1546-1554.
224. Gogolevskaya, I. K., and Kramerov, D. A. (2002) *J. Mol. Evol.*, **54**, 354-364.
225. Borchert, G. M., Lanier, W., and Davidson, B. L. (2006) *Nat. Struct. Mol. Biol.*, **13**, 1097-1101.