

RNomics: Identification and Function of Small Non-Protein-coding RNAs in Model Organisms

A. HÜTTENHOFER

Innsbruck Biocenter, Division of Genomics and RNomics, Innsbruck Medical University, 6020 Innsbruck, Austria

In the recent past, our knowledge on small non-protein-coding RNAs (ncRNAs) has exponentially grown. Different approaches to identify novel ncRNAs that include computational and experimental RNomics have led to a plethora of novel ncRNAs. A picture emerges, in which ncRNAs have a variety of roles during regulation of gene expression. Thereby, many of these ncRNAs appear to function in guiding specific protein complexes to target nucleic acids. The concept of RNA guiding seems to be a widespread and very effective regulatory mechanism. In addition to guide RNAs, numerous RNAs were identified by RNomics screens, lacking known sequence and structure motifs; hence no function could be assigned to them as yet. Future challenges in the field of RNomics will include elucidation of their biological roles in the cell.

Cells from all organisms known to date contain two different kinds of RNAs: mRNAs, which are translated into proteins, and ncRNAs, which function on the level of the RNA and are not translated into proteins (Eddy 2001; Mattick 2001; Hüttenhofer et al. 2002). Sizes of ncRNAs range from very large, for example, about 17 kb as Xist RNA, to extremely small (21–23 nucleotides) as microRNAs (miRNAs). In general, the sizes of the many functional ncRNAs, known up to now, vary from about 20 to 500 nucleotides, well below the size of the majority of mRNAs.

The early view of ncRNAs was that they were relics of a primordial “RNA world” in which RNA served both as the carrier of genetic information and as the catalytic agent. The current view of the RNA world is far more complex. True catalytic RNAs (so-called “ribozymes”) are in fact quite rare. Instead, most ncRNAs perform their cellular duties by a range of mechanisms that are not directly catalytic. For example, a few ncRNAs, such as SRP-RNA (Halic et al. 2004; Halic and Beckmann 2005), appear to function as obligate cofactors of catalytic protein complexes. Some ncRNAs, such as 7SK RNA (Yang et al. 2001), 6S RNA (Wassarman and Storz 2000), CsrB and CsrC RNAs (Dubey et al. 2005), and perhaps *Air* RNA (Pauler et al. 2005), act as genetic regulators by means of antagonistic competition for protein-binding sites. Others serve a structural role or act as scaffolds onto which catalytic proteins can assemble. Hence, the numerous ncRNAs might be classified according to their functions: catalytic RNAs, guide RNAs, catalytic cofactor RNAs, antisense RNAs, protein-binding site–antagonists/agonist RNAs, or templating RNAs, for example.

It has been postulated that up until now, many ncRNAs in genomes of model organisms have escaped detection and that in fact in higher eukaryotes ncRNAs outnumber protein-coding mRNAs (Mattick 2004, 2005; Mattick and Makunin 2005). This paper focuses on methods for the identification of novel ncRNA species in various model organisms in the recent past. We summarize these

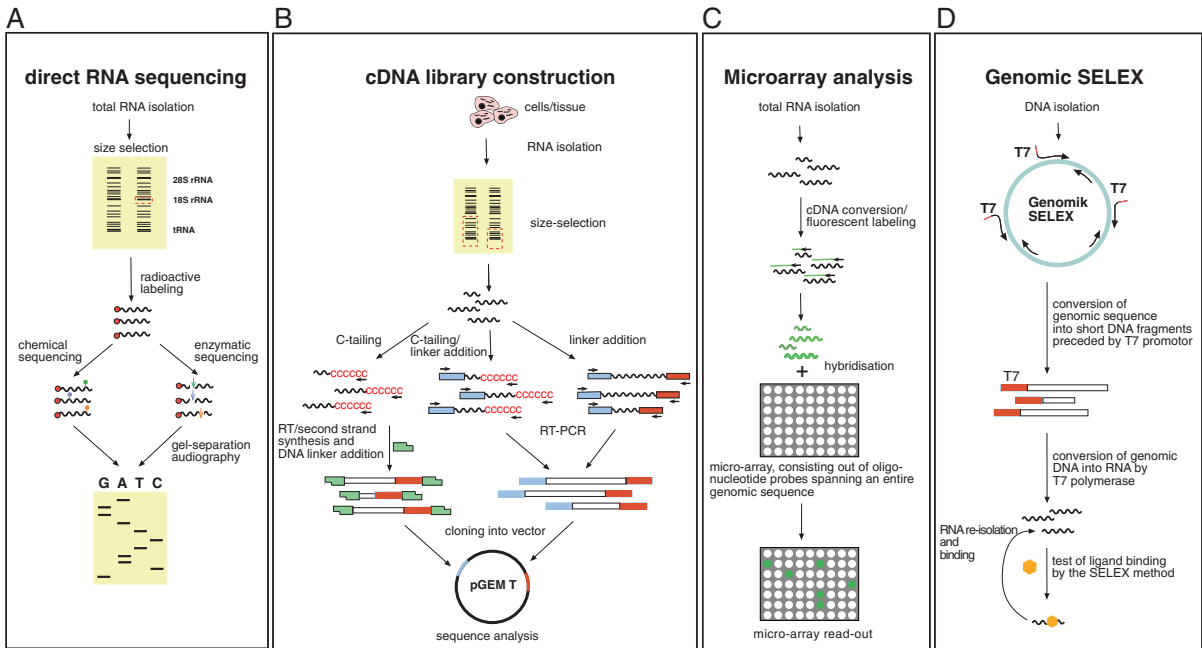
methods designated as “RNomics” (Filipowicz 2000; Hüttenhofer et al. 2002). Subsequently, we also try to elucidate why in the period following the “early RNA world” many of these ncRNAs still have essential roles in the current “protein world.”

METHODS FOR IDENTIFICATION OF ncRNAs IN MODEL ORGANISMS

The term “experimental RNomics” has been coined for identification of novel ncRNAs (Filipowicz 2000; Hüttenhofer et al. 2001). Four different methods currently exist (Fig. 1): (1) RNA sequencing (enzymatically or chemically) as the most traditional method to reveal novel ncRNA species; (2) the parallel cloning of many ncRNA by generating specialized cDNA libraries; (3) the use of microarrays to predict ncRNAs that are expressed under a given experimental condition; and (4) “genomic SELEX” and its potential application to select ncRNA candidates from the sequence space represented by the genome of an organism of interest.

Identification of ncRNAs by Chemical or Enzymatic Sequencing

In the very early days of ncRNA research, that is, some 35–40 years ago, single ncRNA species (e.g., ribosomal RNAs, tRNAs, or viral RNAs) were selected by size separation of total RNA on denaturing gels, followed by visualization and excision of specific bands, ideally representing single ncRNA species (Fig. 1A). Thus, for its identification, the ncRNA of interest must be present in high amounts, that is, visible as a distinct band in an ethidium-bromide-stained polyacrylamide gel, exposed to UV light. Subsequently, RNAs were radiolabeled and sequenced by chemical or enzymatic sequencing methods (Sanger et al. 1965; Brownlee et al. 1972; Donis-Keller et al. 1977; Ehresmann et al. 1977; Peattie 1979).



Identification of ncRNA function

Figure 1. Experimental approaches to identify ncRNAs. Four experimental approaches (A–D) to identify candidates for ncRNAs in model organisms are shown. (A) Identification of ncRNAs by chemical or enzymatic sequencing of extracted abundant RNAs. (B) Identification of ncRNAs by cDNA cloning and sequencing; three different methods are indicated to reverse-transcribe ncRNAs, usually lacking poly(A) tails, into cDNAs (e.g., by C-tailing, C-tailing and linker addition, or linker addition only, followed by RT-PCR). (C) Identification of ncRNAs by microarray analysis. DNA oligonucleotides covering the sequence space of an entire genome are spotted onto glass slides, to which fluorescently labeled samples derived from cellular RNA are hybridized. (D) Identification of ncRNAs by genomic SELEX. By random priming, the sequence of a genome is converted into short PCR fragments containing a T7 promoter at their 5' ends. Subsequently, in vitro transcription by means of T7 RNA polymerase converts this genomic sequence of an organism into RNA fragments, which can then be assayed for function, such as binding to a specific protein or small chemical ligand, by SELEX.

Identification of ncRNAs by Specialized cDNA Libraries

The second method for identification of novel ncRNA species involves the generation of cDNA libraries (Fig. 1B), in analogy with expressed sequence tag libraries (EST libraries) for identification of mRNAs (Gerhold and Caskey 1996; Ohlrogge and Benning 2000). The original mRNA cloning method is based on reverse transcription of mRNAs from an organism by an oligo (dT) primer and second-strand synthesis, resulting in a cDNA library that ideally represents all protein-coding transcripts of a genome. Compared to these conventional EST libraries, the main difference for ncRNA library approaches is the source and treatment of the cloned RNA.

Since most mRNAs are more than 500 nucleotides in length, but many ncRNAs are considerably smaller, RNAs in the size range of about 20–500 nucleotides are first isolated. This fraction is usually depleted in EST libraries as it will not be present in poly(A)⁺ mRNA. The isolation of small-sized RNAs is achieved by size separation of total RNA (either from the entire organism at different developmental stages or from an individual organ) by denaturing polyacrylamide gel electrophoresis.

In many cases, these size- or antibody-selected RNAs lack polyadenylated tails. In general, there are three different methods to reverse-transcribe ncRNAs into cDNA as a prerequisite for cloning and sequencing (Fig. 1B): Prior to reverse transcription, RNAs are ligated to short oligonucleotide linkers (RNA or DNA) at their 5' and 3' ends (Hüttenhofer et al. 2004). Alternatively, at their 3' ends, RNAs can be tailed by poly(A) polymerase, employing either ATP or CTP (Martin and Keller 1998). Reverse transcription of linker-ligated and/or tailed RNA species is followed by polymerase chain reaction (PCR), employing primers complementary to linker sequences. Subsequently, cDNA fragments are cloned into standard vector systems and sequenced.

Microarray Analysis

Microarrays have become the preferred method to monitor the levels of many transcripts in parallel and often at the whole-genome level (Fig. 1C). Microarrays, also known as DNA chips or expression arrays, are glass (or silicon) slides onto whose surface DNA probes have been printed in a grid-like arrangement. To date, single-

stranded DNA oligonucleotides of 25–70 in length are the predominant type of DNA probe on commercial microarrays, although double-stranded PCR products may also serve as probes.

To analyze the entire level of cellular transcripts, samples are prepared from total RNA of an organism. The samples used for microarray hybridization can be the extracted RNA, the converted cDNA, or the cRNA; in any case, these probes will generally be labeled with fluorescent dyes such as Cy3 or Cy5. For more details on the various labeling protocols that are currently being used, see references in Stoughton (2005).

Genomic SELEX

Many ncRNAs form ribonucleoprotein particles (RNPs) at various time points in their life cycle. Such RNA-binding proteins may help an ncRNA fold into its active conformation, shield it from nucleases prior to exerting its function, or promote its annealing with target RNAs up to guiding a protein to its proper target. Other ncRNAs interact with proteins to directly regulate their activity.

The techniques discussed so far allow us to identify ncRNAs from the pool of expressed cellular RNAs after copurification with proteins; that is, by cloning, direct sequencing, or microarray analysis. Given that many such proteins bind their RNA ligands in a nanomolar range, it should also be possible to select RNA ligands from the pool of ncRNAs that an organism can possibly express even without isolating their *in vivo* transcripts.

This approach, termed genomic SELEX (Singer et al. 1997), is based on the *in vitro* generation of RNA species that are derived from a library of an organism's entire genomic DNA (Fig. 1D). The generated RNA pool will undergo successive rounds of association with a given RNA-binding protein, partitioning, and reamplification. As a result, RNA sequences that are stringently bound by the protein partner will be enriched. Once the sequence of the bound RNAs is determined, this information can be used to search for matches in the genome, and so predicted genomic regions can then be tested for the expression of unknown ncRNAs. Genomic SELEX has been successfully applied to select mRNA-binding partners of proteins (Shtatland et al. 2000; Kim et al. 2003), but studies that focused on ncRNAs have not yet been published for any organism.

Alternative Methods

Alternatively to biochemical methods, genetic and bioinformatic tools are also employed to identify ncRNAs in model organisms. Some of the first chromosomally encoded regulatory ncRNAs, for example, MicF, DsrA, and RprA of *Escherichia coli*, were discovered in the course of a genetic screen (Mizuno et al. 1983; Sledjeski and Gottesman 1995; Majdalani et al. 2001). Similarly, genetics also discovered the founding member, *lin-4* RNA, of the ever-growing class of eukaryotic miRNAs (Lee et al. 1993). For a more detailed review of genetic and biocomputational routes

to ncRNA discovery, see Eddy (2002), Vogel and Sharma (2005), and Washietl et al. (2005).

FUNCTIONS OF IDENTIFIED ncRNAs

The above experimental approaches, as well as computational RNomics approaches, have identified thousands of novel ncRNA species in various model organisms from *E. coli* to *Homo sapiens* in the recent past. It thus appears that the number of ncRNAs in the present protein world is very significant and may in fact outnumber protein-coding mRNAs, especially in higher eukaryotes (Mattick 2005). So why is it that such large numbers of ncRNAs still exist—in a post-RNA-world age—and what is the function of all of these RNA species? Why have these functions not been exerted and overtaken by proteins instead?

Although for the majority of newly discovered ncRNA candidates their function is currently unknown, it appears that many of the ncRNAs with an already known function belong to just a few classes, each composed of hundreds to thousands of members, such as miRNAs or siRNAs (short interfering RNAs). These ncRNAs all seem to exhibit a common function, namely, the guiding of protein–enzyme complexes to nucleic acids; hence, they are designated as guide RNAs. By transcriptional or post-transcriptional mechanisms, these guide RNAs are able to regulate and fine-tune gene expression by interacting with either DNA or RNA.

Chimeric RNA–Protein Enzymes

Guide RNAs function as part of a catalytic RNP complex in which the RNA performs the task of substrate recognition and a protein component performs the catalysis. Because the two essential components belong to different classes of macromolecules, I refer to these RNA-guided proteins as “chimeric RNP enzymes.” In general, chimeric RNP enzymes contain an unvarying protein-based enzyme portion (consisting of one or more proteins) that associates with different small guide RNAs which target the complex to its substrate by antisense complementarity (Fig. 2A).

The range of catalytic “payloads” that are guided by these RNAs is strikingly wide and includes endonucleases, polymerases, DNA-, RNA-, and histone methyltransferases, and many more. Guide RNAs belong to a few large families (see above), three of which, siRNA/miRNAs, small nucleolar RNAs (snoRNAs), and gRNAs, contain hundreds of representatives.

All guide RNA families enable an efficient form of modularity, in which multiple substrates can be processed by a single protein complex. Another type of modularity has recently been observed within the miRNA/siRNA family. Not only can multiple RNAs target a single protein catalytic complex to multiple substrates, but different catalytic enzymes can be transported to different substrates by means of similar RNP complexes (Fig. 2B). This is observed, for example, in plants, where an miRNA/siRNA-Dicer-Argonaute complex can guide either a DNA methyltransferase (Kawasaki and Taira

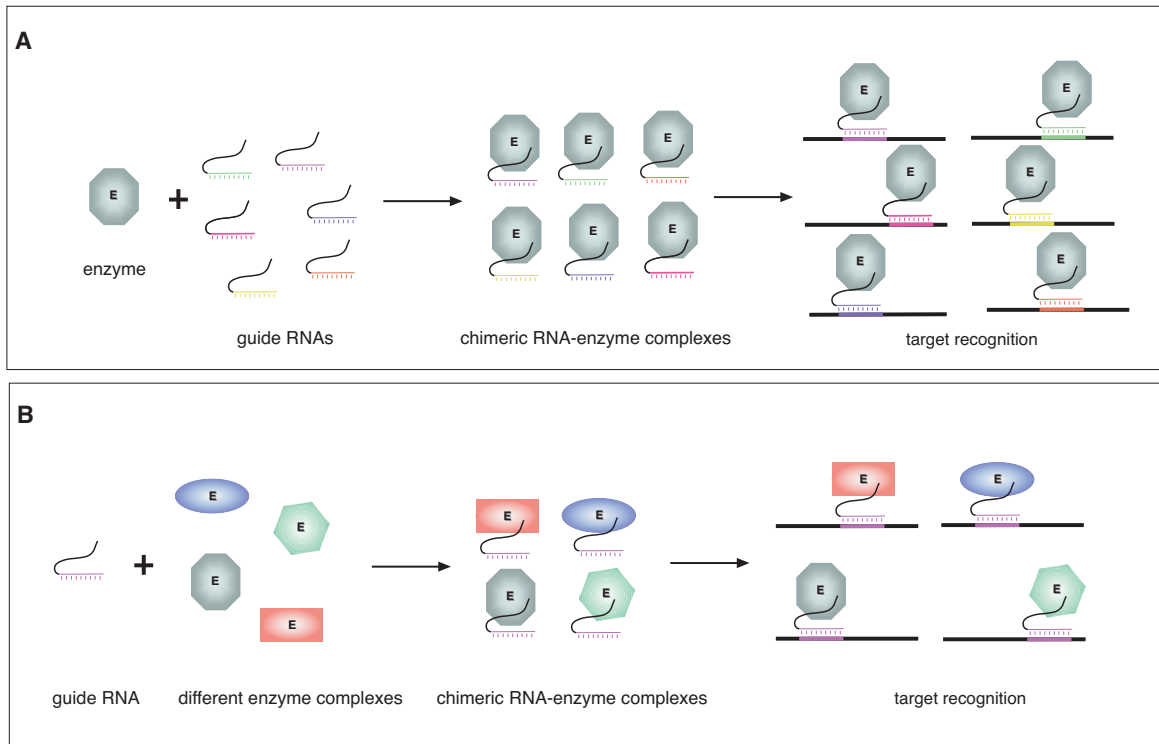


Figure 2. The concept of RNA guiding. (A) One non-sequence-specific enzyme complex (E), consisting of one to several proteins, binds to many different small guide RNAs, resulting in chimeric RNA-enzyme complexes that recognize their targets by Watson-Crick base-pairing and thus guide the enzymatic complex to different substrates (target recognition). (B) One guide RNA can recognize different enzyme complexes (E): One class of guide RNAs (such as miRNAs) might bind to different enzyme complexes (E) and thus is able to guide different enzymatic reactions.

2004; Matzke and Birchler 2005) or a histone methyltransferase (Volpe et al. 2002), or an RNA endonuclease (Meister and Tuschl 2004; Tomari and Zamore 2005). What determines the choice of protein partner remains unknown. A very similar siRNA/miRNA RNP complex guides mRNA translation inhibition in animals (Liu et al. 2004) and DNA endonucleases in *Tetrahymena* (Mochizuki and Gorovsky 2004).

Advantages of RNA Guiding

Given that RNA or DNA target recognition can also be accomplished by proteins alone, why is the RNA-guided enzyme mechanism so widely used? A possible explanation is based on the observation that an RNA-guided enzyme system requires only one (non-sequence-specific) protein for its enzymatic activity. Sequence specificity, and thereby target recognition, is accomplished by the small ncRNA component of the RNP complex. This strategy both limits the amount of the genome that must be allocated to encode the required genes and facilitates the evolution of novel targets for the complex.

Moreover, evolutionary mechanisms that generate novel targets for protein-only enzymes are necessarily more complex. This is because new genes rarely arise de novo, but rather by gene duplication, followed by mutation of the duplicated copy (Long et al. 2003). To accomplish recognition of new target sites, a sophisticated mutation mechanism would be required. This

mechanism would most likely require multiple point mutations, changing several amino acids, in order to modify the RNA-binding domain to target the new site. Since many protein mutations within RNA-binding domains would be expected to often result in loss of function, this would be a highly inefficient means of generating novel target diversity.

In contrast, RNA-guided systems avoid both the problem of requiring multiple protein enzymes to catalyze reactions that involve multiple substrates and the difficulties of evolving additional enzymes for new target sites. A single RNA-guided protein catalytic complex can perform numerous modifications or cleavages simply by associating with the appropriate guide RNA. Since guide RNA genes are generally much shorter than protein-coding genes, significant gains in genomic coding efficiency are possible. In addition, the energetic cost of synthesizing a protein molecule is much higher than that for an RNA molecule.

Furthermore, the RNA-guided enzyme system has the potential to expand its repertoire of target sites simply by duplicating the gene for the RNA guide and incorporating single nucleotide mutations within its antisense sequence. Such single-base antisense mutations will more often generate a new set of target sites and will rarely lead to loss of functionality of the RNP complex, as compared to mutations in protein genes. To achieve this flexibility of target selection with protein-only enzymes would likely require a much

larger number of nucleotide substitutions, with an increased likelihood that at least one of these mutations caused a loss of function.

The General Concept of RNA Guiding Is Widespread in Biology

The first RNAs to be called guide RNAs were those found in kinetoplast mitochondria of trypanosomes, which guide the insertion or deletion of U residues into mitochondrial pre-mRNAs (Stuart et al. 2005). I suggest that the concept of guide RNAs is far more widespread than initially anticipated and can be extended to snoRNAs, si/miRNAs, and even small nuclear RNAs (snRNAs). Indeed, two additional ncRNA families have recently been identified in *Caenorhabditis elegans* (Deng et al. 2006). Whether these ncRNAs represent new guide RNA families is unknown, although it seems likely that our list of guide RNA families is still incomplete. The large families of guide RNAs thereby outnumber the relatively few representatives of catalytic ncRNAs, to which most attention has been drawn in recent years. In evolutionary terms, the concept of RNA guiding has proved to be a very powerful way of generating genetic diversity because new target sites can be generated by gene duplication of guide RNAs genes and mutation of their antisense elements.

CONCLUSIONS: CHALLENGES FOR THE FUTURE

After establishing the biological roles for some novel identified ncRNA species, such as guide RNAs (e.g., snoRNAs, miRNAs, and gRNAs), the future challenge will reside in the analysis of the function of the many other ncRNAs especially in higher eukaryotes, for which the function has not been determined, up until now. Although the number of ncRNAs in eukaryotes has been proposed to be larger than that for protein-coding genes, it still has to be demonstrated, however, whether all of these predicted ncRNA are really biologically functional. Thus, in the future, high-throughput techniques will be required to study the function of thousands of proposed ncRNA candidates in different model organisms.

The most informative approach would be the elimination of the RNA itself, or its encoding DNA gene on the genome (e.g., a knockout of the RNA gene), the latter being very time-consuming. As for mRNAs, residing in the cytoplasm of the cell, knockdown by RNA interference techniques has been proven to be a powerful high-throughput tool to study the function of encoded proteins. A similar approach could be applied to ncRNAs localized to the cytoplasm of a cell, however, not to nuclear or nucleolar localized ncRNAs or to ncRNAs located in cell organelles. For these RNA species, other techniques must be established to study their function in a high-throughput manner. Only if the functions of all candidates for ncRNAs have been established, will we know the actual number of regulatory RNA elements present in a cell.

ACKNOWLEDGMENTS

A.H. expresses his gratitude to all past and present lab members from the Division of Genomics and RNomics for their enthusiasm to work on understanding the fascinating world of noncoding RNAs.

REFERENCES

- Brownlee G.G., Cartwright E., McShane T., and Williamson R. 1972. The nucleotide sequence of somatic 5 S RNA from *Xenopus laevis*. *FEBS Lett.* **25**: 8.
- Deng W., Zhu X., Skogerbo G., Zhao Y., Fu Z., Wang Y., He H., Cai L., Sun H., Liu C., et al. 2006. Organization of the *Caenorhabditis elegans* small non-coding transcriptome: Genomic features, biogenesis, and expression. *Genome Res.* **16**: 20.
- Donis-Keller H., Maxam A.M., and Gilbert W. 1977. Mapping adenines, guanines, and pyrimidines in RNA. *Nucleic Acids Res.* **4**: 2527.
- Dubey A.K., Baker C.S., Romeo T., and Babinzke P. 2005. RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA* **11**: 1579.
- Eddy S.R. 2001. Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2**: 919.
- . 2002. Computational genomics of noncoding RNA genes. *Cell* **109**: 137.
- Ehresmann C., Stiegler P., Carbon P., and Ebel J.P. 1977. Recent progress in the determination of the primary sequence of the 16 S RNA of *Escherichia coli*. *FEBS Lett.* **84**: 337.
- Filipowicz W. 2000. Imprinted expression of small nucleolar RNAs in brain: Time for RNomics. *Proc. Natl. Acad. Sci.* **97**: 14035.
- Gerhold D. and Caskey C.T. 1996. It's the genes! EST access to human genome content. *Bioessays* **18**: 973.
- Halic M. and Beckmann R. 2005. The signal recognition particle and its interactions during protein targeting. *Curr. Opin. Struct. Biol.* **15**: 116.
- Halic M., Becker T., Pool M.R., Spahn C.M., Grassucci R.A., Frank J., and Beckmann R. 2004. Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* **427**: 808.
- Hüttenhofer A., Brosius J., and Bachelierie J.P. 2002. RNomics: Identification and function of small, non-messenger RNAs. *Curr. Opin. Chem. Biol.* **6**: 835.
- Hüttenhofer A., Cavaille J., and Bachelierie J.P. 2004. Experimental RNomics: A global approach to identifying small nuclear RNAs and their targets in different model organisms. *Methods Mol. Biol.* **265**: 409.
- Hüttenhofer A., Kiefmann M., Meier-Ewert S., O'Brien J., Lehrach H., Bachelierie J.-P., and Brosius J. 2001. RNomics: An experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.* **20**: 2943.
- Kawasaki H. and Taira K. 2004. Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. *Nature* **431**: 211.
- Kim S., Shi H., Lee D.K., and Lis J.T. 2003. Specific SR protein-dependent splicing substrates identified through genomic SELEX. *Nucleic Acids Res.* **31**: 1955.
- Lee R.C., Feinbaum R.L., and Ambros V. 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**: 843.
- Liu J., Carmell M.A., Rivas F.V., Marsden C.G., Thomson J.M., Song J.J., Hammond S.M., Joshua-Tor L., and Hannon G.J. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**: 1437.
- Long M., Betran E., Thornton K., and Wang W. 2003. The origin of new genes: Glimpses from the young and old. *Nat. Rev. Genet.* **4**: 865.
- Majdalani N., Chen S., Murrow J., St. John K., and Gottesman S. 2001. Regulation of RpoS by a novel small RNA: The characterization of RprA. *Mol. Microbiol.* **39**: 1382.

- Martin G. and Keller W. 1998. Tailing and 3'-end labeling of RNA with yeast poly(A) polymerase and various nucleotides. *RNA* **4**: 226.
- Mattick J.S. 2001. Non-coding RNAs: The architects of eukaryotic complexity. *EMBO Rep.* **2**: 986.
- . 2004. RNA regulation: A new genetics? *Nat. Rev. Genet.* **5**: 316.
- . 2005. The functional genomics of noncoding RNA. *Science* **309**: 1527.
- Mattick J.S. and Makunin I.V. 2005. Small regulatory RNAs in mammals. *Hum. Mol. Genet.* (spec. no. 1) **14**: R121.
- Matzke M.A. and Birchler J.A. 2005. RNAi-mediated pathways in the nucleus. *Nat. Rev. Genet.* **6**: 24.
- Meister G. and Tuschl T. 2004. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**: 343.
- Mizuno T., Chou M.Y., and Inouye M. 1983. A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (micRNA). *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **59**: 335.
- Mochizuki K. and Gorovsky M.A. 2004. Small RNAs in genome rearrangement in *Tetrahymena*. *Curr. Opin. Genet. Dev.* **14**: 181.
- Ohlrogge J. and Benning C. 2000. Unraveling plant metabolism by EST analysis. *Curr. Opin. Plant Biol.* **3**: 224.
- Pauler F.M., Stricker S.H., Warczuk K.E., and Barlow D.P. 2005. Long-range DNase I hypersensitivity mapping reveals the imprinted Igf2r and Air promoters share cis-regulatory elements. *Genome Res.* **15**: 1379.
- Peattie D.A. 1979. Direct chemical method for sequencing RNA. *Proc. Natl. Acad. Sci.* **76**: 1760.
- Sanger F., Brownlee G.G., and Barrell B.G. 1965. A two-dimensional fractionation procedure for radioactive nucleotides. *J. Mol. Biol.* **13**: 373.
- Shtatland T., Gill S.C., Javornik B.E., Johansson H.E., Singer B.S., Uhlenbeck O.C., Zichi D.A., and Gold L. 2000. Interactions of *Escherichia coli* RNA with bacteriophage MS2 coat protein: Genomic SELEX. *Nucleic Acids Res.* **28**: E93.
- Singer B.S., Shtatland T., Brown D., and Gold L. 1997. Libraries for genomic SELEX. *Nucleic Acids Res.* **25**: 781.
- Sledjeski D. and Gottesman S. 1995. A small RNA acts as an antisilencer of the H-NS-silenced rcsA gene of *Escherichia coli*. *Proc. Natl. Acad. Sci.* **92**: 2003.
- Stoughton R.B. 2005. Applications of DNA microarrays in biology. *Annu. Rev. Biochem.* **74**: 53.
- Stuart K.D., Schnauffer A., Ernst N.L., and Panigrahi A.K. 2005. Complex management: RNA editing in trypanosomes. *Trends Biochem. Sci.* **30**: 97.
- Tomari Y. and Zamore P.D. 2005. Perspective: Machines for RNAi. *Genes Dev.* **19**: 517.
- Vogel J. and Sharma C.S. 2005. How to find small non-coding RNAs in bacteria. *Biol. Chem.* **386**: 1219.
- Volpe T.A., Kidner C., Hall I.M., Teng G., Grewal S.I., and Martienssen R.A. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**: 1833.
- Washielt S., Hofacker I.L., Lukasser M., Hüttenhofer A., and Stadler P.F. 2005. Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. *Nat. Biotechnol.* **23**: 1383.
- Wassarman K.M. and Storz G. 2000. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101**: 613.
- Yang Z., Zhu Q., Luo K., and Zhou Q. 2001. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**: 317.