Meeting report

Regulatory RNAs and the demise of 'junk' DNA

Frank J Slack

Address: Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA.
Email: frank.slack@yale.edu

Published: 28 September 2006

Genome ***Biology*** 2006, **7:**328 (doi:10.1186/gb-2006-7-9-328)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2006/7/9/328

A report of the meeting 'Regulatory RNAs', the 71st Cold Spring Harbor Symposium on Quantitative Biology, Cold Spring Harbor, USA, 31 May-5 July 2006.

A growing body of work suggests that genes for noncoding RNAs make up a substantial class of genes in all organisms, with increasing organismal complexity correlated with an increasing complexity of noncoding RNAs. Many of these noncoding RNAs appear to have regulatory functions and these were the subject of this year's annual Cold Spring Harbor Symposium. Among the most exciting themes of the meeting were the evidence for significant amounts of hitherto undiscovered transcription in genomes and the discovery of novel classes of noncoding RNAs with thousands of members. In this report I review a few of these highlights.

The tenets of the 'central dogma' have required revision over the past few decades as biologists have begun to appreciate that RNA performs many functions once thought solely to be the domain of proteins. Apart from its well established roles as messenger, ribosomal component, and transfer RNA, it is now clear that RNA can have a key role in regulating gene expression. Noncoding regulatory RNAs - RNAs that are not translated into protein - include the small nuclear RNAs (snRNAs), the small nucleolar RNAs (snoRNAs), the *XIST* RNA that mediates mammalian X-chromosome silencing, microRNAs, riboswitches, and the RNA component of the enzyme telomerase. These RNAs direct such diverse processes as gene silencing, transcriptional and translational control, imprinting, and dosage compensation. These discoveries have electrified the biological community as we try to understand the extent of the 'RNA world' and how regulatory RNAs work in controlling gene expression. We are fast learning that large portions of the genome that do not code for proteins are in fact transcribed, and that these regions, previously thought to be 'junk', may be useful after all (Figure 1).

Transcription, transcription everywhere

Whole-genome tiling microarrays offer a relatively unbiased and sensitive approach to detecting rare transcripts and, along with the sequencing of expressed sequence tags (ESTs), are providing ample evidence for an abundance of unsuspected transcription from mammalian genomes that involves both protein-coding and noncoding sequences. Using tiling arrays, both Michael Snyder (Yale University, New Haven, USA) and Thomas Gingeras (Affymetrix, Santa Clara, USA) showed startling evidence for significantly more transcription from the human genome than was previously appreciated, much of it regulated. Snyder reported that there are at least twice as many transcription units as previously thought, and that about one third of these are conserved in mammals, indicating that they are biologically relevant. Gingeras reported that about half of all characterized protein-coding genes in the regions in the human cells he studies use at least one alternative transcription start site that is an average of 100 kb from the previously annotated gene. He also revealed the existence of 450,000 new short transcripts in the human genome, or one approximately every 3,000 nucleotides. John Mattick (University of Queensland, Brisbane, Australia), an early advocate of the complexity of noncoding RNA, also reported the detection of thousands of new noncoding transcripts from human and mouse cells and showed that the expression of many of these changes during development, which suggests that they might have some function.

These observations of low-level, regulated transcription across much of the mammalian genome demonstrate that much more RNA is being generated in our cells than we previously thought. With the mounting evidence of large amounts of transcription, the question becomes: are all these transcripts functional? No one yet has the answer, but it seems hard to fathom why evolution would have selected for such extensive transcription if it were useless or wasteful.

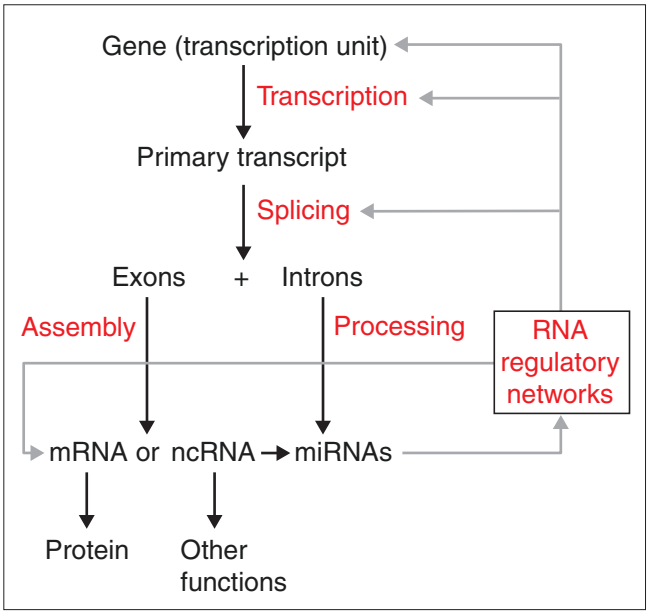

Figure 1
The potential role of RNAs in the regulation of gene expression. For the purposes of this article a gene is loosely defined as a transcription unit that may produce either protein-coding sequence (mRNA) or noncoding RNA (ncRNA), or in some cases both. The flow of genetic information from gene to a functional protein or RNA is indicated by the black arrows. Noncoding RNAs can act to regulate the expression of genes at multiple points (gray arrows). Black text refers to physical entities such as DNA and RNA, while red text refers to activities. Figure courtesy of and adapted from J. Mattick.

Opening out the RNA world

The ramifications of the behavior of some noncoding RNAs are indeed gradually being unraveled. Michel Georges (University of Liège, Belgium) described an example in which a mutation leading to the creation of an miRNA-binding site is responsible for one distinctive characteristic of a breed of sheep. Georges and his group have mapped and cloned a mutation responsible for the large muscle mass of Texel sheep. This mutation maps to the 3′ UTR of the myostatin gene (a repressor of muscle development) and creates a novel predicted binding site for *mir-1* (an miRNA known to be expressed in muscle), which would enable repression of the myostatin gene by this miRNA, allowing increased muscle development. They have found that single-nucleotide polymorphisms (SNPs) in 3′ UTRs that create or destroy miRNA-binding sites are common, and many may alter miRNA-regulated gene expression (see the Patrocles website [http://www.patrocles.org]). Their analysis strongly suggests that miRNA control of gene expression will turn out to be an important driving force in evolution.

Several groups reported the existence of novel classes of noncoding RNAs, and there was a growing sense that like miRNAs and other well established classes of noncoding RNAs, these new RNAs will also prove to be functional. Gregory Hannon (Cold Spring Harbor Laboratory, New York, USA), Thomas Tuschl (Rockefeller University, New York, USA), David Bartel (Whitehead Institute, Boston, USA) and Robert Kingston (Harvard Medical School, Boston, USA) reported the discovery of a new class of thousands of testis-specific, 30-mer RNAs, named piRNAs, that associate with proteins of the Piwi subfamily of Argonaute proteins. The Argonaute proteins are known to associate with miRNAs and function as components of the RNA-induced silencing complex (RISC). Since Piwi proteins are closely related to Argonaute, they were also thought to be involved in gene silencing, but little was known about their function. Now Piwi proteins have themselves been linked to small RNAs, but little is known of the function of these piRNAs. They are perhaps the first of many new classes of RNA awaiting discovery in specific cell types. It will be interesting to determine whether these 30-mers correspond in any way to a collection of 30-mer sequences, named pyknons, that are found repeated throughout the human genome as reported by Isidore Rigoutsos (IBM, Yorktown Heights, USA).

Bartel also discussed two new classes (based on their sequence and size) of small RNAs that have been found in nematodes, and which led him to propose that *C. elegans* may have another few thousand genes. An exciting report was that of David Baulcombe (John Innes Centre, Norwich, UK), who described new noncoding RNAs from the single-celled alga *Chlamydomonas*. Some of these resemble miRNAs, and since miRNAs have previously only been found in multicellular organisms, these findings suggest that miRNAs are not a specialty of multicellularity. The discovery of thousands of additional novel noncoding RNAs, big and small, was reported for many different groups of organisms, including mammals (mouse), insects (*Drosophila*), cellular slime molds (*Dictyostelium*), and flatworms (planarians). The function of most of these RNAs remains unknown and awaits further experimentation, but the impact of all these studies showed us how much there is left to discover in this fast-moving area of biology.

The meeting treated participants to a feast of regulatory mechanisms and new genomic discoveries. With the evidence for massive amounts of transcription in the genome and for new functional classes of noncoding RNAs, I was left with the impression that our understanding of the regulatory RNA world is still in its infancy. The situation was nicely summed up by Mattick in his comment "...and this is probably just the tip of the iceberg". The next few years are certain to bring new discoveries that will provide a greater appreciation and understanding of the role of RNA in regulating our genomes. Perhaps it is time to bid farewell to the term 'junk' DNA - we knew not your true nature.