National Science Review
1: 184–189, 2014

BIOLOGY & BIOCHEMISTRY

Special Topic: Frontiers in RNA Research

The globalization of messenger RNA regulation

Jack D. Keene

Transcriptional regulation has been a dominant area of research in gene expression for more than 50 years. However, the past 20 years of research in biology brought new focus on post-transcriptional events, namely RNA regulation. This emphasis on understanding the regulatory functions of RNAs emerged with the identification of *cis*- and *trans*-regulatory regions in pre-mRNAs and mature mRNAs, non-coding regulatory RNAs, and ribozymes; and identification of the proteins that bind them. Moreover, the cDNA cloning of regulatory RNA-binding proteins (RBPs) and their interactions with a myriad of non-coding RNAs (ncRNAs) as well as coding mRNAs began to accelerate in the second half of the 1980s. For example, the previously discovered transfer RNAs, uridylate-rich small nuclear RNAs, and snoRNAs, as well as viral ncRNAs, were examined at all levels of RNA processing. The standard mode of investigation at that time focused on specific molecular interactions between two entities, in this case, a single type of RBP and a specific RNA. These discoveries provided an abundant amount of information, both descriptive and mechanistic, to the rapidly expanding field of RNA biology. Indeed, these studies laid the groundwork for the invention of new methodologies that allowed investigations into global mechanisms that underlie protein–RNA and RNA–RNA interactions that continue today.

GLOBAL mRNA TARGETING *IN VITRO*

The global expansion of cellular messenger RNA targeting in the 1990s began with the demonstration that RBPs could bind to multiple mRNAs that have related RBP binding sites and, possibly, similar functional relationships to one another. As little genomic sequence information was available at that time, it was difficult to identify sequences in mRNAs that were related to one another, even using computational searches. One exception was the AU-rich elements (AREs) found in the 3′ untranslated regions (UTRs) of mRNAs encoding lymphokines and proto-oncogenes by Shaw and Kamen [1] to function as mRNA destabilizers. Subsequently, Levine, King, and co-workers [2] cDNA-cloned a neuronal RBP, Hel-N1 (HuB) that bound to ARE mRNAs, and showed it to target multiple mRNAs by using *in vitro* selection procedures based on both randomized [2] and naturally occurring [3] RNA libraries prepared from human brain 3′ UTRs, in order to identify over 100 neuronal HuB mRNA targets. Gao *et al.* [3] suggested that HuB RBPs may coordinately regulate these multiply targeted mRNAs to affect neuronal growth and differentiation, and indicated the importance of identifying multiple RNA targets *in vivo*. It has since been established that all four ELAV/Hu family members similarly target multiple ARE-containing mRNAs encoding growth regulatory proteins, including cytokines and proto-oncogenes. It should be noted that many other RBPs including AUF1 and tristetraprolin have since been found to interact with ARE sequences, and these were later examined for global multi-targeting functions as well.

A similar study by Butt *et al.* [4] used *in vitro* SELEX (Systematic Evolution of Ligands by Exponential Enrichment) of randomized sequences to examine RNAs binding to the iron response element RBPs, IRP1 (aconitase), and IRP2 that were shown previously to bind and regulate specific stem-loop binding elements in mRNAs encoding enzymes involved in heme biosynthesis, including ferritin, the transferrin receptor, and the erythroid 5-aminolevulinate synthetase [5]. Indeed, the iron response system is a classical mechanistic example of RNA regulation. With respect to global mRNA targeting and regulation, both the Gao [3] and the Butt [4] studies used *in vitro* selection methods, although different experimental designs, to suggest that specific RBPs could broadly expand post-transcriptional regulation across mRNA targets, and potentially allow different mRNA subsets to respond dynamically to differing conditions of growth or differentiation. These were early brush strokes on the canvas leading to eventual formalization of the post-transcriptional RNA operon/regulon model of global RNA coordination [6].

GLOBAL mRNA TARGETING *IN VIVO*

These early attempts to survey multiple targeting by RBPs occurred during some of early phases of the human genome project and the advent of microarray technologies [7]. These advances created an opportunity to take RNA–protein interactions to a global level, and the RNA-immunoprecipitation-sequencing (RIP-chip/seq) procedure was invented to

identify RNA targets of RBPs globally *in vivo* [8,9]. Data from this first RIP-chip study by Tenenbaum *et al.* [8] used the neuronal HuB and ubiquitous HuR RBPs in retinoic acid induced P19 cell extracts, and in both cases, the identified RNA targets were consistent with mRNA sequences identified by Gao *et al.* [3] using *in vitro* selection of brain mRNAs. Subsequent work in many labs using DNA arrays and high-throughput sequencing to study global post-transcriptional events such as RNA splicing, stability, translation, and localization of multiple RNAs have used RIP procedures and crosslinking procedures that approach single nucleotide resolution (reviewed in [10]).

Following the original discovery of microRNAs by Ambros and Ruvkun in 1993, Tuschl, Bartel, and others rapidly expanded the field by discovering hundreds of microRNAs in various species (reviewed in [11]). Subsequently in 2003, Lewis *et al.* [12] predicted that these small RNAs of 21–22 nucleotides can also target hundreds of mRNAs via their complementary seed sequences of 6–8 nucleotides. Thus, there were two aspects of microRNA globalization: the discovery of more microRNAs and the discovery of their myriad mRNA targets. The fact that both microRNAs and RBPs bind to similar and overlapping mRNA subsets enriches the combinatorial possibilities of globally coordinated RNA regulation. Since that time, many other ncRNAs have been discovered that are also candidates for global combinatorial RNA regulation. Moreover, in most cases, the *trans*-acting microRNAs and the *trans*-acting RBPs can dynamically cooperate or compete in coordinating the fates of multiple mRNAs, a concept that is now well substantiated within the post-transcriptional RNA operon/regulon concept [6,9,13].

FUNCTIONAL COHERENCE OF DYNAMIC GLOBAL RNA REGULATION

Among the biologically significant findings derived from global analysis of RNA regulation is that the mRNA targets of many RBPs not only have related sequence elements in common, but they also coordinately regulate mRNA targets at all levels of processing, export, and localization [8,9]. For example, multiple mRNAs can be coordinately spliced, stabilized or destabilized, exported in groups, or translated into proteins that have similar functions [6,9,13]. These findings were the basis for the proposal that messenger ribonucleoproteins may represent posttranscriptional RNA operons and regulons, a concept that has been widely substantiated in many laboratories [6,13–16]. The chief tenet of the concept is that certain RBPs, in cooperation with ncRNAs, can dynamically coordinate global regulation of distinct classes of mRNAs encoding proteins that are in the same pathway or macromolecular complex, and in turn, coordinately regulate their splicing, stability, export, and localization. This idea of global dynamic coordination of post-transcriptional outcomes applies to essentially all mRNAs across large numbers of overlapping transcripts. And given that each mRNA type has multiple copies in the cell, each mRNA has the potential to join with different groups of functionally related transcripts in time and space. This mode of global regulation logically provides wide flexibility for post-transcriptional events to evolve emergent properties via dynamic responses of mRNA subsets to cellular signals in an efficient and economical manner. Indeed, the RNA regulon concept is about the combinatorial use of RNAs.

Among the recent advances in this field are the dozens of 'developmental' RNA regulons discovered in trypanosomes that dynamically change during different periods of the life cycle, for example, differentiation in the blood steam by coordinating the expression and stability of RNAs encoding metabolic as well as structural proteins (reviewed in [15]). In addition, the early discoveries of the PUF (PUm-Fbf) family of RNA regulons in Pat Brown's lab at Stanford have continued to advance up the phylogenetic scale to include mammalian Pum1 and Pum2 RBPs (reviewed in [13]), and developmental RNA regulons among the STAR-GLD-1 and GLD2 RBP family in *C. elegans* (discussed in [16]). Thus, the field is advancing toward better understanding the biological consequences of dynamic changes in RNA subpopulations in more complex developing biological systems.

FUTURE GOALS AND CHALLENGES

While many examples of RNA regulons have been demonstrated in yeasts, worms, flies, trypanosomes, mammals, and many other species [6,9,14–16], several challenging aspects remain to be understood. Primary among them is the 'combinatorial' assembly of RBPs and ncRNAs that govern the networks of RNAs that are dynamically coordinated as RNA regulons. Likewise, we need to better understand the signals that activate or repress, or segregate specific subsets of RNAs to serve changing stages of development. In the broader view, understanding how global RNA networks feed forward and feed backward between transcription and translation (and RNA stability) will enrich our knowledge of cellular balance and adaptability. Clearly, every transcription factor must be translated and every posttranscriptional factor must be transcribed, and a homeostatic system of this type requires dynamic responses when the cell is perturbed or a developmental switch is activated [6,8,9,14–16].

Moreover, better methods are needed to discriminate transient from more stable interactions between RBPs and ncRNAs and their global mRNA targets, and to determine how the order of binding governs the sequential assembly of functional ribonucleoproteins. The RIP-chip/seq procedures originally used to reveal RNA regulons have been used successfully for probabilistic quantification of dynamically changing cellular perturbation, but do not necessarily identify precise binding sites to the nucleotide. Crosslinking procedures, in contrast, can in some cases locate binding sites, but due to low efficiencies, crosslinking-induced sequence biases, and lack of saturation in deep sequencing, data are not easily amenable to quantifying dynamic changes in global mRNA targets. Indeed, comparative quantitation will be especially important

when addressing dynamic RNA–protein interactions in developing tissues and organs. These limitations of both RIP-chip/seq and UV crosslinking have been addressed in cultured cells by integrating RIP-chip/seq data with PAR-CLIP (Photoactivatable Ribonucleoside Crosslinking and Immunoprecipitation) UV crosslinking data obtained under the same conditions in the same cell system. Ideally, however, entirely new methods will be developed that avoid the limitations of the classical RIP and CLIP approaches, while taking advantage of their respective advantages.

In sum, it is still early in this rapidly expanding field of global protein–RNA interactions, but the implications are exciting given the enormous numbers of small regulatory RNAs and RBPs that are being discovered that affect posttranscriptional regulation and epigenetic changes. It is a bold proposition to suggest, but investigations of global RNA dynamics may provide researchers access to entirely novel solutions for understanding and manipulating human diseases.

ACKNOWLEDGEMENTS

I thank Matt Friedersdorf and Jeff Blackinton for critically reading the manuscript, and I apologize to many colleagues for not being able to cite their valued contributions to this field.

FUNDING

This work was supported by grants from the National Cancer Institute (R01CA157268 to J.D.K.) and the National Science Foundation (0842621 to J.D.K.).

Jack D. Keene
Department of Molecular Genetics & Microbiology, Duke University Medical Center, USA
E-mail: keene001@mc.duke.edu

REFERENCES

1. Shaw, G and Kamen, R. *Cell* 1986; **46**: 659–67.
2. Levine, TD, Gao, F-B, King, PH *et al. Mol Cell Biol* 1993; **13**: 3494–504.
3. Gao, FB, Carson, C, Levine, TD *et al. Proc Natl Acad Sci USA* 1994; **91**: 11207–11.
4. Butt, J, Kim, HY, Basilion, JP *et al. Proc Natl Acad Sci USA* 1996; **93**: 4345–9.
5. Klausner, RD, Rouault, TA and Harford, JB. *Cell* 1993; **72**: 19–28.
6. Keene, JD. *Nat Rev Genet* 2007; **8**: 533–43.
7. Schena, M, Shalon, D, Davis, RW *et al. Science* 1995; **270**: 467–70.
8. Tenenbaum, SA, Carson, CC, Lager, PJ *et al. Proc Natl Acad Sci USA* 2000; **97**: 14085–90.
9. Keene, JD. *Proc Natl Acad Sci USA* 2001; **98**: 7018–24.
10. Milek, M, Wyler, E and Landthaler, M. *Semin Cell Dev Biol* 2012; **23**: 206–12.
11. Bartel, DP. *Cell* 2004; **116**: 281–97.
12. Lewis, BP, Shih, IH, Jones-Rhoades, MW *et al. Cell* 2003; **115**: 787–98.
13. Halbeisen, RE, Galgano, A, Scherrer, T *et al. Cell Mol Life Sci* 2008; **65**: 798–813.
14. Bronicki, LM and Jasmin, BJ. *RNA* 2013; **19**: 1019–37.
15. De Gaudenzi, JG, Carmona, SJ, Aguero, F *et al. Peer J* 2013; **1**: e118.
16. Wright, JE, Gaidatzis, D, Senften, M *et al. EMBO J* 2011; **30**: 533–45.

doi: 10.1093/nsr/nwu004
Advance access publication 14 May 2014

PLANT & ANIMAL SCIENCE

Special Topic: Frontiers in RNA Research

Non-coding RNAs as potent tools for crop improvement

Renyi Liu[1,*] and Jian-Kang Zhu[1,2]

Since the seminal discovery of small interfering RNAs (siRNAs) [1] more than 10 years ago, significant progress has been made in understanding plant non-coding RNAs (ncRNAs). The first discovery was the different types of small ncRNAs, including microRNAs (miRNAs) and various types of siRNAs [2]. These small RNAs differ in their origins and biogenesis, but their final products function in a similar way, i.e. to silence genes of complimentary sequences. The second breakthrough was the identification and functional characterization of major components of the biogenesis pathway of small RNAs. An interesting discovery in this area is the coupling of miRNA biogenesis and mRNA maturation through proteins such as SE, ABH1/CPB80, CPB20, STA1, and SIC [3–5]. Lastly, functional roles of small RNAs in RNA-directed DNA methylation, viral defense, transposon suppression, abiotic and biotic stress responses, DNA double-strand break repair, and plant development have been demonstrated [6]. These advancements in basic research have greatly increased our knowledge of plant ncRNAs and facilitated the effective design of ncRNA-based strategies for crop improvement.

ncRNAs function by repressing the expression of endogenous or exogenous genes at the transcriptional, post-transcriptional, or translational levels in a sequence-specific manner. Therefore, ncRNAs can be used to specifically control the expression of target genes. Before the biogenesis of siRNAs and their mechanisms of expression regulation had been fully understood, RNA silencing was already being exploited in the early 1990s to produce desired crops by introducing antisense or sense transcripts of target genes to transgenic plants. The first commercial crop that was genetically modified based on RNA silencing was the Flavr Savr tomato [7]. The antisense transcript of polygalacturonase (PG) was introduced in tomato to suppress the