Systematic analysis of posttranscriptional gene expression

Adam R. Morris, Neelanjan Mukherjee and Jack D. Keene

Recent systems studies of gene expression have begun to dissect the layers of regulation that underlie the eukaryotic transcriptome, the combined consequence of transcriptional and posttranscriptional events. Among the regulatory layers of the transcriptome are those of the ribonome, a highly dynamic environment of ribonucleoproteins in which RNA-binding proteins (RBPs), noncoding regulatory RNAs (ncRNAs) and messenger RNAs (mRNAs) interact. While multiple mRNAs are coordinated together in groups within the ribonome of a eukaryotic cell, each individual type of mRNA consists of multiple copies, each of which has an opportunity to be a member of more than one modular group termed a posttranscriptional RNA operon or regulon (PTRO). The mRNAs associated with each PTRO encode functionally related proteins and are coordinated at the levels of RNA stability and translation by the actions of the specific RBPs and noncoding regulatory RNAs. This article examines the methods that led to the elucidation of PTROs and the coordinating mechanisms that appear to regulate the RNA components of PTROs. Moreover, the article considers the characteristics of the dynamic systems that drive PTROs and how mRNA components are bound collectively in physical ‘states’ to respond to cellular perturbations and diseases. In conclusion, these studies have challenged the extent to which cellular mRNA abundance can inform investigators of the functional status of a biological system. We argue that understanding the ribonome has greater potential for illuminating the underlying coordination principles of growth, differentiation, and disease.

Eukaryotic gene expression is a complex process that integrates myriad signals to coordinate production of thousands of gene products in effectively precise temporal and spatial patterns. Its inherent complexity has been addressed in recent years using a systems approach, thus much of systems biology is devoted to the study of gene expression. It is generally believed that improper regulation of gene expression can lead to many human defects and disorders. Thus, understanding the underlying mechanisms of gene expression has become an active subfield of genetic diseases and medicine. For example, the advent of microarray technology was crucial for the study of global gene expression and has led to advances in understanding and categorizing human disease. As discussed below, while traditional microarray approaches have limitations for understanding some aspects of gene expression such as transcription rates, novel applications have emerged, which provide insight into underlying mechanisms of gene coordination at the posttranscriptional level.

Many studies of global gene expression focus solely on the transcriptome, and the only factor assessed is mRNA abundance, which is but a single aspect of gene expression. Several studies have suggested that measuring global mRNA levels to assess the transcriptome using microarrays can be misleading, as gene expression has multiple layers that manifest themselves in both the nucleus and the cytoplasm after transcription has ended. A prime example of the importance of the posttranscriptional
environment was recently demonstrated in a study from the laboratory of David Baltimore, which showed that mRNA stability strongly influences gene expression induction kinetics during the inflammatory response, in some cases overriding the effects of transcription. Regulation of transcription is also a relatively slow process for responding to cellular perturbations; in the sea urchin embryo, the average gene is translated seven times more rapidly than it is transcribed (56 min for transcription versus 8 min for translation). In addition, some human genes can take as long as 16 h to be transcribed. In order to rapidly produce a protein, it is advantageous to increase the rate of translation immediately, without waiting for transcription and subsequent mRNA export and cytoplasmic regulation.

Even without waiting for transcription and subsequent mRNA export and cytoplasmic regulation. The underlying posttranscriptional environment was recently demonstrated in a study from the laboratory of David Baltimore, which showed that mRNA stability strongly influences gene expression induction kinetics during the inflammatory response, in some cases overriding the effects of transcription. Regulation of transcription is also a relatively slow process for responding to cellular perturbations; in the sea urchin embryo, the average gene is translated seven times more rapidly than it is transcribed (56 min for transcription versus 8 min for translation). In addition, some human genes can take as long as 16 h to be transcribed. In order to rapidly produce a protein, it is advantageous to increase the rate of translation immediately, without waiting for transcription and subsequent mRNA export and cytoplasmic regulation.

Moreover, repression of transcription can be a relatively slow means to cease protein production if not coupled to rapid mRNA decay and/or a decrease in translation. In addition, many studies have shown that transcription may be more stochastic than was previously believed, and therefore the newly synthesized mRNA populations created by transcription may be altered correspondingly in the posttranscriptional environment. The underlying mechanisms of posttranscriptional regulation (PTR) are determined by many factors that may bind to and regulate an mRNA after transcription and up to and during translation, including RNA-binding proteins (RBPs) and small noncoding RNAs, such as microRNAs. Recent studies have shown that understanding PTR on a global level provides insights into the coordination of gene expression and its implications for disease.

The life of every copy of an mRNA involves multiple points for regulation (Figure 1), including splicing, polyadenylation, transport from the nucleus, localization within the cytoplasm, translation, and decay (for a review of the life of an mRNA we recommend). In most of the cases where PTR has been shown to be important, the regulation involved one or multiple RNA-binding proteins. Thus, RNA-binding proteins have key roles in posttranscriptional gene expression, coordinating many aspects of the life of an mRNA.

Regulation of gene expression at the posttranscriptional level involves both control and coordination. While control describes an individual interaction that results in a specific outcome, coordination describes a process of integrating multiple control functions to achieve a higher level of harmonized outcome. Historically, the study of PTR has focused on the one-on-one small-scale control functions, which may result in profound outcomes, but does not address overall RNA coordination. For example, an RBP or microRNA may affect an mRNA sequence element within a Luciferase reporter, demonstrating control of the expression of that RNA. In addition, those same trans-acting factors may also be found to alter the phenotype of a cell or organism, and could be presumed to do so by affecting the same single mRNA target from which the sequence present in the reporter system was derived in vivo. However, the phenotypic change is just as likely to result from the combined effects of that trans-acting RBP or microRNA on coordinating multiple mRNA targets. In addressing this issue, advances in molecular biological techniques and detection methods have allowed study of control on a wider basis, often global, thus leading to a greater understanding of RNA coordination. Gao et al. (1994) found that the ELAV/HuB RBP can target multiple mRNAs in vitro using total brain mRNA and suggested that this could represent a coordinating function for posttranscriptional gene regulation. Subsequently, Tenenbaum et al. (2000) demonstrated that the HuR and HuB RBPs target multiple mRNAs in vivo during neuronal differentiation in mouse embryonic carcinoma P19 cells. Similar multiple targeting interactions by microRNAs were predicted using computational algorithms, and it is generally assumed today that microRNAs, like RBPs, can target and affect multiple mRNAs in living cells. While RBPs have been shown in many studies to target functionally related mRNAs, such conclusions have not emerged for microRNAs. This is consistent with the very broad target predictions of microRNAs; however, microRNAs have been reported to have profound effects on phenotypes. Interestingly, the actions of specific microRNAs have been shown to fine-tune the production of multiple proteins, possibly acting as a multi-targeted mRNA rheostat, but functional coherence has not been demonstrated among the affected proteins. Indeed, coordination of PTR has best been demonstrated by identifying the genome-wide mRNAs associated with particular RBPs in ribonucleoprotein (RNP) particles, and therefore the role of RBPs in coordinating gene expression is the focus of this review.

**RIBONOMICS-GLOBAL ANALYSIS OF mRNPS**

The global analysis of mRNA and protein components of RNPs has been termed ribonomics because
it explores the ribonome, the total RNP content of a cell including proteins, mRNAs, and noncoding regulatory RNAs. The majority of ribonomic experiments have employed ribonucleoprotein-immunoprecipitation-microarray (RIP-Chip) and more recently RIP-Seq, when deep sequencing procedures are employed in place of microarrays. RIP-Chip involves immunoprecipitation of RNP complexes, typically through use of an antibody to one of the RBP components, extraction of the associated RNAs, and identification of this RNA population on a microarray (Figure 2). Many variations of this method have been employed, such as using a recombinant protein with a physical (e.g. epitope) tag as well as an antibody against the endogenous protein for immunoprecipitation. RIP-Chip has proven successful in various tissues and species, from yeast to mammals, as demonstrated in Table 1. One of the main benefits of RIP-Chip is that it
is designed to recover RNP complexes, allowing identification of components other than mRNAs, such as other regulatory or RNA-processing proteins, as well as small noncoding RNAs. A recent study by Tuschl and colleagues encompassed an in-depth characterization of protein and mRNA components of Ago RNPs, demonstrating the usefulness of RIP-Chip type experiments for identifying both the protein and RNA components of RNPs. Other recent advances made by our lab in RIP-Chip analysis, as discussed below, also demonstrate that RIP-Chip is an ideal method for studying remodeling of RNPs during the dynamic processes of posttranscriptional gene expression.

A less commonly used adaptation of RIP-Chip based on a step of ultraviolet (UV) cross-linking prior to immunoprecipitation, termed cross-linking and immunoprecipitation (CLIP) or high-throughput sequencing (HITS)-CLIP, can also be used to determine global mRNAs associated with RNA-binding proteins. The various modifications of CLIP are designed to identify the specific RNA sequences bound to an RBP of interest, through either cloning or more recently through deep sequencing. UV irradiation forms covalent bonds (albeit inefficiently) between proteins and RNAs that are in direct contact, and thus the RNA that was not in direct contact can be digested away after cross-linking, allowing identification of the exact interacting site. The use of cross-linking also allows some protein–RNA interactions to be retained during the extensive purification steps of CLIP. After UV treatment and immunoprecipitation, RNA–protein complexes are separated by SDS–polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose. Then, the protein component is digested away with proteinase K and the remaining RNA is cloned through use of linker ligation and reverse transcriptase polymerase chain reaction or HITS.

CLIP can be useful for determining the exact RNA sequences with which some RBPs interact; however, treatment with UV radiation has been shown to physically and chemically alter the RNP infrastructure, thus results obtained via CLIP may not reflect the true RNP state existing prior to UV treatment. In addition, CLIP is ill-suited for recovering entire RNP complexes because of the extensive purification protocol and is significantly more labor intensive than standard RIP-Chip, although new iterations of CLIP utilizing deep sequencing may avoid some of this additional effort. UV cross-linking also has sequence bias at the level of both RNA and protein, with the greatest photoreactivity occurring between pyrimidines and certain amino acids, such as Cys, Lys, Phe, Trp, and Tyr. The use of cross-linking has obvious advantages for studying RBP-binding sites; however, because of the caveats involved with UV cross-linking, a similar method that increases the cross-linking efficiency or avoids the drawbacks of UV irradiation would prove to be a valuable tool.

Ribonomics analysis has led to many insights into both control and coordination of PTR. For the purpose of this review, the remainder of this section will focus mainly on those insights that pertain to the field of systems biology.

1. One key aspect in understanding how RBPs can coordinate PTR is the insight that RBPs typically associate with multiple mRNAs. The range of mRNAs that a typical RBP associates with is very wide, from a large majority of mRNA species to only a few; however, it appears that most RBPs are associated with multiple species of mRNA. This concept was demonstrated by Gao et al. and Tenenbaum et al., and while confirmed in dozens of published studies, was recently supported most convincingly by Patrick Brown’s laboratory in a ribonomic study of 46 RBPs in yeast, which showed that almost all of the RBPs studied were associated with multiple species of mRNA.

2. When the multiple mRNAs associated with an RBP are analyzed, it is typically found that the encoded proteins share a functional relationship. This is another property of the posttranscriptional environment that has been revealed through ribonomic analysis; the mRNAs associated with an RBP are functionally related, for example, in encoding a macromolecular complex, signaling cascade, or developmental process. This property, arising from the discovery that RBPs are multi-targeted, contributed to the proposal of the RNA operon/regulon model. This model, demonstrated in Figure 1, describes how RBPs coordinate PTR by associating with multiple, functionally related mRNAs and coordinating their protein production or RNA stability, analogous to the way DNA operons allow coordinated protein production in bacteria. The posttranscriptional RNA operon or regulon (PTRO) model proffered the coordinated spatial and temporal production of functionally related proteins, even when the genes encoding these proteins are found dispersed throughout the genome. RNA operons also accommodate the multi-functionality of eukaryotic proteins,
<table>
<thead>
<tr>
<th>Cell type/line</th>
<th>Additional biological conditions</th>
<th>Name</th>
<th>RNA target(s) summary</th>
<th>Motif</th>
<th>Notes</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEK 293</td>
<td>—</td>
<td>Ago1</td>
<td>~600 mRNAs</td>
<td>—</td>
<td>E, S</td>
<td>(37)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>—</td>
<td>Ago1</td>
<td>Mostly miRNAs, but also snorNA, mRNAs, noncoding RNAs, transposons</td>
<td>—</td>
<td>E, H</td>
<td>(38)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>—</td>
<td>Ago1</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>—</td>
<td>Ago2</td>
<td>~600 mRNAs</td>
<td>—</td>
<td>E, S</td>
<td>(37)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>—</td>
<td>Ago2</td>
<td>Mostly miRNAs, but also snorNA, mRNAs, noncoding RNAs, transposons</td>
<td>—</td>
<td>E, H</td>
<td>(38)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>miR122 overexpression</td>
<td>Ago2</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>miR-124 Overexpression</td>
<td>Ago2</td>
<td>—</td>
<td>miR-124 seed</td>
<td>T, M</td>
<td>(39)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>Overexpression miR-1 and miR-124</td>
<td>Ago2</td>
<td>miRNAs, 56 miR-1 targets, 288 miR-124 targets</td>
<td>miR-1 and miR-124 seeds</td>
<td>T, M</td>
<td>(40)</td>
</tr>
<tr>
<td>Jurkat</td>
<td>—</td>
<td>Ago2</td>
<td>~75% miRNA, 0.7% mRNA</td>
<td>—</td>
<td>E, H</td>
<td>(41)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>—</td>
<td>Ago3</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>Jurkat</td>
<td>—</td>
<td>Ago3</td>
<td>~88% miRNA, 2.3% tRNA, 1.5% mRNA</td>
<td>—</td>
<td>E, H</td>
<td>(41)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>—</td>
<td>Ago4</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>K562</td>
<td>—</td>
<td>alphaCP2</td>
<td>Cytoskeleton, transcription factors, proto-oncogenes, and signaling</td>
<td>CU-rich</td>
<td>E, M</td>
<td>(42)</td>
</tr>
<tr>
<td>HeLa</td>
<td>—</td>
<td>AUF1</td>
<td>58% of targets were AUF1 targets</td>
<td>—</td>
<td>E, M</td>
<td>(43)</td>
</tr>
<tr>
<td>HeLa</td>
<td>—</td>
<td>AUF1</td>
<td>Pre-mRNA and mRNAs</td>
<td>Long AU-rich stem with four internal bulges</td>
<td>E, M</td>
<td>(44)</td>
</tr>
<tr>
<td>MCF10A</td>
<td>MCT-1 overexpression</td>
<td>AUF1</td>
<td>Cell cycle and MAPK signaling in nontransformed</td>
<td>—</td>
<td>E, M</td>
<td>(45)</td>
</tr>
<tr>
<td>HEK 293</td>
<td>—</td>
<td>Dicer</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>Dopaminergic neuroblastoma cells</td>
<td>—</td>
<td>DJ1</td>
<td>Mitochondrial, oxidative damage, and PI3K/AKT pathway</td>
<td>GG/CC rich</td>
<td>E, M</td>
<td>(46)</td>
</tr>
<tr>
<td>HESC</td>
<td>—</td>
<td>Fox2</td>
<td>~7% of genes in hESCs</td>
<td>UGCAUG</td>
<td>E, H, X</td>
<td>(47)</td>
</tr>
<tr>
<td>HeLa</td>
<td>—</td>
<td>GW182 (TNRC6A)</td>
<td>18 mRNAs</td>
<td>—</td>
<td>E, M</td>
<td>(48)</td>
</tr>
<tr>
<td>HeLa</td>
<td>—</td>
<td>HuR</td>
<td>57% of targets were AUF1 targets</td>
<td>—</td>
<td>E, M</td>
<td>(43)</td>
</tr>
<tr>
<td>MCF10A</td>
<td>MCT-1 overexpression</td>
<td>HuR</td>
<td>Cell cycle and MAPK signaling upon transformation</td>
<td>—</td>
<td>E, M</td>
<td>(45)</td>
</tr>
<tr>
<td>MCF7</td>
<td>MCT-1 overexpression</td>
<td>HuR</td>
<td>Ras, PI3K, angiogenesis, and hypoxia pathways</td>
<td>—</td>
<td>E, M</td>
<td>(49)</td>
</tr>
<tr>
<td>Cell type/line</td>
<td>Additional biological conditions</td>
<td>Name</td>
<td>RNA target(s) summary</td>
<td>Motif</td>
<td>Notes</td>
<td>Refs</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------</td>
<td>------</td>
<td>-----------------------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>RKO</td>
<td>ActinomycinD</td>
<td>HuR</td>
<td>—</td>
<td>U-Rich 12- to 20-nt long stem loop</td>
<td>E, M</td>
<td>(50)</td>
</tr>
<tr>
<td>HEK 293</td>
<td></td>
<td>IGF2BP1</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>alphaT3-1</td>
<td></td>
<td>KSRP</td>
<td>12 mRNAs (many RBPs)</td>
<td>—</td>
<td>T, M</td>
<td>(51)</td>
</tr>
<tr>
<td>HeLa</td>
<td>IL-1 stimulated</td>
<td>KSRP</td>
<td>~11% genes expressed</td>
<td>—</td>
<td>T, M</td>
<td>(52)</td>
</tr>
<tr>
<td>T98G</td>
<td>Co-cultivated with PY4.1</td>
<td>PABP</td>
<td>Human T98G-specific translational profile</td>
<td>—</td>
<td>T, M</td>
<td>(53)</td>
</tr>
<tr>
<td>HEK 293</td>
<td></td>
<td>PABPC4</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>HeLa</td>
<td></td>
<td>PTB</td>
<td>Intracellular transport, vesicle trafficking and apoptosis</td>
<td>YYYYUCUUYYYY</td>
<td>E, M</td>
<td>(54)</td>
</tr>
<tr>
<td>HeLa</td>
<td></td>
<td>Pum1</td>
<td>Cell cycle and transcription factors</td>
<td>UGUHAUA</td>
<td>E, M</td>
<td>(55)</td>
</tr>
<tr>
<td>HeLa</td>
<td></td>
<td>Pum1</td>
<td>Nucleic acid binding and p53 pathway</td>
<td>UGUHAUA</td>
<td>E, M</td>
<td>(55)</td>
</tr>
<tr>
<td>HeLa</td>
<td></td>
<td>Pum1</td>
<td>Cell cycle and transcription factors</td>
<td>UGUHAUA</td>
<td>E, M</td>
<td>(56)</td>
</tr>
<tr>
<td>HEK 293</td>
<td></td>
<td>QK1</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>3T3</td>
<td></td>
<td>SAM68</td>
<td>Cell migration, cell cycle, and transcriptional regulation</td>
<td>UAAA, UUUA, OR Poly(U)</td>
<td>E, S, X</td>
<td>(57)</td>
</tr>
<tr>
<td>HEK 293</td>
<td></td>
<td>SFRS1</td>
<td>RNA-processing factors</td>
<td>Purine-rich octamer with GAAGAA core</td>
<td>E, H, X</td>
<td>(58)</td>
</tr>
<tr>
<td>HeLa</td>
<td></td>
<td>SLBP</td>
<td>Replication-dependant histones</td>
<td>Stem loop</td>
<td>T, M</td>
<td>(60)</td>
</tr>
<tr>
<td>RKO</td>
<td></td>
<td>TIA1</td>
<td>—</td>
<td>U-rich 30–37-nt bipartite element</td>
<td>E, M</td>
<td>(61)</td>
</tr>
<tr>
<td>RKO</td>
<td></td>
<td>TIAR</td>
<td>Many targets dissociate upon UVC</td>
<td>C-rich with stem and loop with additional loop</td>
<td>E, M</td>
<td>(62)</td>
</tr>
<tr>
<td>HEK 293</td>
<td></td>
<td>TNRC6A</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>HEK 293</td>
<td></td>
<td>TNRC6B</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>HEK 293</td>
<td></td>
<td>TNRC6C</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(35)</td>
</tr>
<tr>
<td>Human dendritic cells</td>
<td></td>
<td>TTP</td>
<td>Protein synthesis</td>
<td>AAUAAA</td>
<td>E, M</td>
<td>(63)</td>
</tr>
<tr>
<td>HeLa</td>
<td></td>
<td>U2AF-65kD</td>
<td>Spliced cell cycle and transcription factors</td>
<td>UUUNYYYY</td>
<td>E, M</td>
<td>(54)</td>
</tr>
<tr>
<td>P19</td>
<td></td>
<td>EIF4E</td>
<td>A subset of expressed mRNAs</td>
<td>—</td>
<td>E, M</td>
<td>(7)</td>
</tr>
<tr>
<td>ES cells</td>
<td></td>
<td>Esg1</td>
<td>Cell cycle, chromatin remodeling, and transcription factors</td>
<td>—</td>
<td>E, M</td>
<td>(64)</td>
</tr>
<tr>
<td>Brain</td>
<td></td>
<td>FMRP</td>
<td>—</td>
<td>G-Quartet</td>
<td>E, M</td>
<td>(65)</td>
</tr>
<tr>
<td>Cell type/line</td>
<td>Additional biological conditions</td>
<td>Name</td>
<td>RNA target(s) summary</td>
<td>Motif</td>
<td>Notes</td>
<td>Refs</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
<td>------</td>
<td>-----------------------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>P19</td>
<td>Retinoic Acid (RA) treated</td>
<td>HuB</td>
<td>Early response genes</td>
<td>UUUAUUU</td>
<td>E, M</td>
<td>(7)</td>
</tr>
<tr>
<td>P19</td>
<td>RA treated</td>
<td>HuR</td>
<td>—</td>
<td>—</td>
<td>E, M</td>
<td>(7)</td>
</tr>
<tr>
<td>Brain</td>
<td>—</td>
<td>JERKY</td>
<td>Ribosomal proteins and cellular stress</td>
<td>—</td>
<td>T, M</td>
<td>(66)</td>
</tr>
<tr>
<td>Brain</td>
<td>—</td>
<td>Nova</td>
<td>Also bound 3’UTRs to regulate polyadenylation</td>
<td>YCAY rich regions</td>
<td>E, H, X</td>
<td>(67)</td>
</tr>
<tr>
<td>Brain</td>
<td>—</td>
<td>Nova</td>
<td>Neuronal synapse, neuronal inhibition</td>
<td>YCAY rich regions</td>
<td>E, S, X</td>
<td>(68)</td>
</tr>
<tr>
<td>P19</td>
<td>—</td>
<td>PABP</td>
<td>A subset of expressed mRNAs</td>
<td>—</td>
<td>E, M</td>
<td>(7)</td>
</tr>
<tr>
<td>PY4.1</td>
<td>Co-cultivated with T98G and 4TI</td>
<td>PABP</td>
<td>Mouse endothelial PY4.1-specific translational profiles</td>
<td>—</td>
<td>T, M</td>
<td>(53)</td>
</tr>
<tr>
<td>24 Central Nervous System (CNS) cell populations</td>
<td>—</td>
<td>Rpl10a</td>
<td>Translational profiles for CNS cell populations</td>
<td>—</td>
<td>T, M</td>
<td>(69)</td>
</tr>
<tr>
<td>Striatonigral and striatopallidal medium spiny neurons</td>
<td>Cocaine stimulated</td>
<td>Rpl10a</td>
<td>GABA-A receptors upon cocaine stimulation</td>
<td>—</td>
<td>T, M</td>
<td>(70)</td>
</tr>
<tr>
<td>RAW 264.7</td>
<td>LPS activation</td>
<td>TTP</td>
<td>Development</td>
<td>UUUAUUUAUU, AUUUA</td>
<td>E, M</td>
<td>(71)</td>
</tr>
<tr>
<td>Testis</td>
<td>—</td>
<td>Dazl</td>
<td>—</td>
<td>UUCUUCUGUU CUU</td>
<td>E, M</td>
<td>(72)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Aco1</td>
<td>Saccharomyces Cerevisiae</td>
<td>Mitochondrial-encoded RNAs</td>
<td>—</td>
<td>T, M</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Atg8</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Bfr1</td>
<td>Proteins localized to nucleolus and nucleus</td>
<td>UGAAAAWUUUY</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Bud27</td>
<td>Component of RNA polymerase I complex</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Cbc2</td>
<td>Intron-containing transcripts and potential CUTs</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Cbf5</td>
<td>H/ACA snoRNAs and potentially some mRNAs</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Dhh1</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Gbp2</td>
<td>Nucleolus and chromosome, ribosome biogenesis, and chromatin modification</td>
<td>UGAAAAWUUUUYA</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Hrb1</td>
<td>Intron-containing transcripts</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Idh1</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Khd1</td>
<td>Cell wall and Bud tip</td>
<td>(6) CNN repeats</td>
<td>T, M</td>
<td>(74)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Khd1</td>
<td>Cell wall, plasma membrane, Bud, ER proteins</td>
<td>HWNCAUUWY</td>
<td>T, M</td>
<td>(73)</td>
</tr>
</tbody>
</table>

© 2009 John Wiley & Sons, Inc.
<table>
<thead>
<tr>
<th>Cell type/line</th>
<th>Additional biological conditions</th>
<th>Name</th>
<th>RNA target(s) summary</th>
<th>Motif</th>
<th>Notes</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>—</td>
<td>—</td>
<td>Lhp1</td>
<td>RNA Pol II transcripts and ribosomal proteins</td>
<td>—</td>
<td>T, M</td>
<td>(75)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Mex67</td>
<td>Integral membrane proteins, cell wall, RNA-binding protein, ribosomal proteins</td>
<td>—</td>
<td>T, M</td>
<td>(76)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Mex67</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Msl5</td>
<td>Intron-containing transcripts</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Myo4</td>
<td>Bud localized, stress sensing/response</td>
<td>Four loops separated by short stem</td>
<td>T, M</td>
<td>(77)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Myo4</td>
<td>Ash1, IST2 and other Bud localized mRNAs</td>
<td>—</td>
<td>T, M</td>
<td>(78)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nab2</td>
<td>Nucleus, nucleic acid metabolism</td>
<td>DRARAMGMD</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nab2</td>
<td>Transcription</td>
<td>AAAAAG</td>
<td>T, M</td>
<td>(79)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nab3</td>
<td>snoRNAs, potential cryptic unstable transcripts (CUTs)</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nab4</td>
<td>Amino acid metabolism</td>
<td>UAUUAUA</td>
<td>T, M</td>
<td>(79)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nab6</td>
<td>Cell wall</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nce102</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nop13</td>
<td>5.8S and 35S ribosomal RNAs</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nop56</td>
<td>C/D box snoRNAs and potentially many mRNAs</td>
<td>UUUUCUUU</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Npl3</td>
<td>Cytosolic and mitochondrial ribosome, oxidative phosphorylation, intron-containing transcripts</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Npl3</td>
<td>Ribosomal proteins</td>
<td>—</td>
<td>T, M</td>
<td>(79)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nrd1</td>
<td>snoRNAs, potential CUTs</td>
<td>UUCUGUW</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nrp1</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Nsr1</td>
<td>Ribosomal proteins, ribosomal internal transcribed regions</td>
<td>GGGWAACGGW</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Pab1</td>
<td>Ribosomal proteins and intron-containing transcripts</td>
<td>WUAUAUA</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Pbp2</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Pin4</td>
<td>Novel sequence element in 3′-UTRs</td>
<td>UUUAUGA</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Pub1</td>
<td>Ribosomal proteins</td>
<td>U-rich degenerate stabilize elements</td>
<td>T, M</td>
<td>(80)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Pub1</td>
<td>Localized ER, vacuole and plasma membranes, cell wall, intron-containing transcripts</td>
<td>HUUUUU</td>
<td>T, M</td>
<td>(73)</td>
</tr>
</tbody>
</table>
# TABLE 1

<table>
<thead>
<tr>
<th>Cell type/line</th>
<th>Additional biological conditions</th>
<th>Name</th>
<th>RNA target(s) summary</th>
<th>Motif</th>
<th>Notes</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>—</td>
<td>—</td>
<td>Puf1</td>
<td>Membrane-associated proteins</td>
<td>UAAUAUUAAAU (73)</td>
<td>T, M</td>
<td>(81)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Puf2</td>
<td>Membrane-associated proteins</td>
<td>UAAUAUUW (73)</td>
<td>T, M</td>
<td>(81)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Puf3</td>
<td>Mitochondrial proteins</td>
<td>CHUGUAWAUA</td>
<td>T, M</td>
<td>(81)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Puf4</td>
<td>Nucleolar ribosomal RNA-processing factors</td>
<td>WHUGUAHAWUA</td>
<td>T, M</td>
<td>(81)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Puf5</td>
<td>Chromatin modifiers and spindle pole body</td>
<td>UGUAAYAWUA</td>
<td>T, M</td>
<td>(81)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Rib2</td>
<td>Several tRNAs</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Rna15</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Scp160</td>
<td>Cell wall, plasma membrane, nuclear membrane, ER proteins, and ribosome biogenesis and transport</td>
<td>UGAAAAAUUU</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Sgn1</td>
<td>RPS20</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>She2</td>
<td>Bud localized, stress sensing/response</td>
<td>Two loops separated by short stem</td>
<td>T, M</td>
<td>(77)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>She2</td>
<td>Ash1, IST2 and other Bud localized mRNAs</td>
<td>—</td>
<td>T, M</td>
<td>(78)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>She3</td>
<td>Bud localized, stress sensing/response</td>
<td>Three loops separated by short stem</td>
<td>T, M</td>
<td>(77)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>She3</td>
<td>Ash1, IST2 and other Bud localized mRNAs</td>
<td>—</td>
<td>T, M</td>
<td>(78)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Ski2</td>
<td>Ski3 and Ski7 of the exosome</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Sof1</td>
<td>C/D box snoRNAs including U3 snoRNA</td>
<td>AKUCAUUCCUU</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Ssd1</td>
<td>Proteins localized to the cell wall and Bud</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Tdh3</td>
<td>Mitochondrial-encoded mRNAs, many nuclear-encoded mRNAs</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Upf1</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Vts1</td>
<td>Carbohydrate transport</td>
<td>UKWCGRGNN</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Yll032c</td>
<td>Transport, development, cell signaling</td>
<td>AAUACCY</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Ypl184c</td>
<td>Proteins localized to the cell wall, ER, plasma membrane; mRNAs containing uORFs and IRES</td>
<td>YUUCWBUYH</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Yra1</td>
<td>Cell wall, RNA-binding proteins, ribosomal proteins, carbohydrate metabolism</td>
<td>—</td>
<td>T, M</td>
<td>(76)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>Yra2</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(73)</td>
</tr>
<tr>
<td>Cell type/line</td>
<td>Additional biological conditions</td>
<td>Name</td>
<td>RNA target(s) summary</td>
<td>Motif</td>
<td>Notes</td>
<td>Refs</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------</td>
<td>------</td>
<td>-----------------------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>Leaf</td>
<td></td>
<td>RPL18</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(83)</td>
</tr>
<tr>
<td>Leaf</td>
<td></td>
<td>RPL23a</td>
<td>—</td>
<td>—</td>
<td>T, M</td>
<td>(83)</td>
</tr>
<tr>
<td>Stroma</td>
<td></td>
<td>CRP1</td>
<td>Photosystem I core complex</td>
<td>Identical 7-mer and 11-mer separated by 51 nts</td>
<td>T, M</td>
<td>(84)</td>
</tr>
<tr>
<td>Leaf</td>
<td></td>
<td>PPR4</td>
<td>rps12 Pre-mRNA</td>
<td>—</td>
<td>T, M</td>
<td>(84)</td>
</tr>
<tr>
<td>Leaf</td>
<td></td>
<td>PPR5</td>
<td>Intron of unspliced plastid tRNA precursor</td>
<td>pre-trnG-UCC</td>
<td>—</td>
<td>E, M</td>
</tr>
<tr>
<td>Intestine</td>
<td>Intestinal expression of Pab1</td>
<td>Pab1</td>
<td>Intestinal translational profiles, chromosomal clustering</td>
<td>—</td>
<td>T, M</td>
<td>(86)</td>
</tr>
<tr>
<td>Muscle</td>
<td>Muscle expression of Pab1</td>
<td>Pab1</td>
<td>Muscle translational profiles, chromosomal clustering</td>
<td>—</td>
<td>T, M</td>
<td>(87)</td>
</tr>
<tr>
<td>S2</td>
<td>miR1-KO</td>
<td>Ago1</td>
<td>89 and 108 mRNAs respectively</td>
<td>~30% contained miR-1 seed</td>
<td>T, M</td>
<td>(88)</td>
</tr>
<tr>
<td>Photoreceptors</td>
<td></td>
<td>PABP</td>
<td>Different classes of photoreceptor genes</td>
<td>—</td>
<td>T, M</td>
<td>(89)</td>
</tr>
<tr>
<td>Adult Ovary</td>
<td></td>
<td>Pum</td>
<td>Transcriptional regulation, V-ATPase, cyclins</td>
<td>UGUHAUAUA</td>
<td>T, M</td>
<td>(90)</td>
</tr>
<tr>
<td>Embryo</td>
<td></td>
<td>Pum</td>
<td>Germ cell development and anterior-posterior axis patterning</td>
<td>UGUHAUAUA</td>
<td>T, M</td>
<td>(90)</td>
</tr>
<tr>
<td>Embryo</td>
<td></td>
<td>U2AF</td>
<td>Nuclear export of intron-less mRNAs</td>
<td>—</td>
<td>E, M</td>
<td>(91)</td>
</tr>
<tr>
<td>S2</td>
<td></td>
<td>Hrp36</td>
<td>Combinatorial regulation by Hrp38 and Hrp40</td>
<td>—</td>
<td>E, M</td>
<td>(92)</td>
</tr>
<tr>
<td>S2</td>
<td></td>
<td>Hrp38</td>
<td>Combinatorial regulation by Hrp36 and Hrp40</td>
<td>—</td>
<td>E, M</td>
<td>(92)</td>
</tr>
<tr>
<td>S2</td>
<td></td>
<td>Hrp40</td>
<td>Combinatorial regulation by Hrp36 and Hrp38</td>
<td>—</td>
<td>E, M</td>
<td>(92)</td>
</tr>
<tr>
<td>S2</td>
<td></td>
<td>Hrp48</td>
<td>Distinct subset from Hrp36, Hrp38 and Hrp40</td>
<td>—</td>
<td>E, M</td>
<td>(92)</td>
</tr>
<tr>
<td>Trypanosome</td>
<td></td>
<td>UBP1</td>
<td>Metabolism and glycoproteins</td>
<td>25–30 bp stem loop</td>
<td>E, M</td>
<td>(93)</td>
</tr>
<tr>
<td>Trypanosome</td>
<td></td>
<td>RBP3</td>
<td>Ribosomal proteins</td>
<td>25–30 bp stem loop</td>
<td>E, M</td>
<td>(93)</td>
</tr>
<tr>
<td>Trypanosome</td>
<td></td>
<td>DRBD3</td>
<td>Membrane proteins</td>
<td>U-rich</td>
<td>E, M</td>
<td>(94)</td>
</tr>
</tbody>
</table>

For each study, authors’ results are reported in appropriate columns. Motifs represent the single most significant motif reported. For the column ‘Notes’: E, endogenous; T, tagged; S, sequencing; H, high-throughput sequencing; M, microarray; and X, cross-linking.
allowing a single gene to participate in multiple operons.

A compelling example of a mammalian PTRO with a likely role in balancing inflammation was discovered by Paul Fox and Barsanjit Mazumder. An RNP known as the IFN–gamma-activated inhibitor of translation (GAIT) complex that contains a phosphorylated form of ribosomal protein L13a binds to and reduces translation of the mRNAs encoding the inflammation protein ceruloplasmin (Cp), vascular endothelial growth factor (VEGF), zipper interacting protein (ZIP) kinase, and death associated protein (DAP) kinase. It was hypothesized that translation of other mRNAs encoding pro-inflammatory proteins was also silenced by the GAIT complex, a putative posttranscriptional operon. To test this hypothesis, the authors used genome-wide translational profiling to discover an array of numerous chemokine mRNAs encoding proteins involved in inflammation that are silenced in a similar manner to the Cp mRNA. As predicted, many of these mRNAs involved in the GAIT PTRO contain a cis-hairpin element that was shown to mediate silencing of translation. These mRNAs were also responsive to knockdown of the ribosomal protein L13a protein, a key component of the GAIT complex whose phosphorylation by the ZIP kinase is the switch to silence translation of these functionally related mRNAs. Interestingly, the ZIP kinase itself is activated by phosphorylation by the other GAIT-targeted protein kinase, DAP. The authors postulate that the γIFN response is transcriptionally activated but then subsequently subdued by feedback from the GAIT PTRO, thus optimizing a chronic inflammatory response and preventing an over-reaction of the immune system.

4. PTR has been shown to occur in all species tested, and although the protein domains and sequence motifs responsible for interactions are well conserved, the identities of the proteins and mRNAs that contain these domains and motifs generally are not. This is another insight gained through ribonomic investigation; PTR evolves in a modular fashion, with the interacting portions of the RBPs and mRNAs being conserved but the identities of mRNAs that are coordinated and the consequences of this regulation differing across species. Thus, evolutionary rewiring has occurred, during which selective pressures led to the conservation of RNA-binding domains and the bound sequence motifs, but allowed these interactions to be used in different functional settings that were best suited for survival. Perhaps the best example of this rewiring can be seen in the case of the PUF (for Pumilio/FBF) proteins, sequence-specific RBPs that exist throughout the eukaryotic lineage. The PUF RNA-binding domain, or PUM-HD, is extremely well conserved between yeast, fly, and human PUF proteins, as is the sequence motif to which this domain binds in global targets. However, ribonomic analysis of PUF proteins from all of these species revealed that there was little conservation of the identities of the associated mRNAs or functions in which these mRNAs were involved, showing how evolutionary rewiring of PTR occurs in a modular fashion.

5. Another aspect of PTR previously revealed through ribonomic analysis was confirmed through studies of PUF proteins; RBPs tend to regulate other regulatory proteins, such as other RBPs and transcription factors. This concept, depicted by the dotted lines in Figure 1, is referred to as the ‘regulators of regulators’ concept and was most elegantly demonstrated in a study by Gorospe and colleagues. This study revealed that a number of RBPs that bind to AU-rich elements in 3′UTRs of mRNAs (ARE-RBPs) can bind to and regulate mRNAs encoding other ARE-RBPs, forming a complex pattern of PTR. In fact, given that so many RBPs target
mRNAs encoding other RBPs, it is likely that
the PTR environment can be extremely robust
by being self-sustaining and causing changes in
gene expression independent of transcriptional
input, at least for limited periods of time.16,22
6. The final insight that will be discussed here
is that the ribonome is highly dynamic; it
responds to environmental signals by altering
RNP contents, both protein and RNA, thus
altering the gene expression program in a cell.
A dramatic visual example of the dynamics of
the posttranscriptional environment can be seen
in the case of stress granules. Stress granules
are large, cytoplasmic aggregates that contain
various RBPs and mRNAs and form upon
response to various types of stress.108 Stress
granules form within minutes of administration
of stress and will also begin to dissociate
within minutes of stress being removed,109
demonstrating the rapid adaptability of PTR.
Unfortunately, it has not been possible to
isolate stress granules in order to analyze their
RNA contents globally. An example of specific
RNA dynamics of PTR was demonstrated by
early RIP-Chip experiments of HuB, mentioned
previously.7 This study showed that a subset
of functionally related mRNA species associated
with HuB changed in a coordinated manner
during neuronal differentiation.

STATES AND POPULATION DYNAMICS

In order to model the dynamics associated with
coordination of PTR, it is useful to think about
mRNAs as existing in physical states in time and
space. For this review, a state represents association
with one or more trans-acting factors, namely RBPs
and small RNAs, thus the states that an mRNA may
exist in are defined by the combinations of trans-acting
factors with which it may associate. These trans-acting
factors will contribute to determining the functional
expression of that mRNA, thus an mRNA existing in
a given state will have a distinct functional outcome
for the expression of that gene. An important aspect of
this model is that changes in mRNA abundance are not
representative of the state of an mRNA (although there
may be a correlation). Thus, measurements of mRNA
abundance, as determined using transcriptomics, do
not reveal PTR events.

It is often assumed that every copy of the same
species of an mRNA undergoes identical PTR, but
this view is not correct.4,7,11,24,98 The multiple states
in which an mRNA may exist at any given time or
location almost ensure that not every copy of a species
of mRNA will have an identical fate, as depicted by
Figure 3. In this example, hypothetical if there are
10 copies of mRNA y in a cell, 5 may be associated
with RNP A (S1) and thus translated for use at the
leading edge of the cell (F1), 3 may be associated
with RNP B (S2) and thus their translated products
will be involved in Wnt signaling (F2), and 2 may
be associated with RNP C (S3) and thus be stored
in an inactive state (F3). When the cell undergoes
a change in gene expression over time, the number
of copies of mRNA y does not change; however,
the state associations are rearranged, thus changing
expression of that gene without changes in mRNA
abundance or transcription rates. When one considers
the many possible steps in the life of an mRNA and
the many potential states that can exist at each step,
it becomes even more apparent that it is unlikely
that every copy of a species of mRNA will have the
same fate.22,98 Allowing different copies of the same
species of mRNA to be regulated differently is also
beneficial in accommodating the multi-functionality
of eukaryotic proteins, allowing one gene to be regulated
in multiple ways in concert as a member of more than
one PTRO.

A drawback of many global studies of PTR is
that results are typically reported as discrete datasets
(X does or does not control expression of Y), losing
much of the information content that could be useful
in inferring more global coordination across multiple
RNA targets of RBPs. In order to address this issue,
we have devised methods to study RBPs globally
and to allow mRNA target data to be analyzed
probabilistically.

An example of this probabilistic modeling
method was demonstrated in our laboratory in a study
of the human RBP Pum1.56 In this study, biological
triplicates of RIP-Chip were performed and t-scores
calculated comparing the Pum1 RIP to the negative
RIP. A histogram of the distribution of these t-scores
revealed that there were two distinct populations of
mRNAs: an enriched, or Pum1-associated, population
and a non-enriched, background population. Utilizing
Gaussian mixture modeling,110 we were able to
objectively define the two populations, and by using
the equations describing these populations, we were
able to assign a log of the odds (LOD) score for each
mRNA. These LOD scores provide a probabilistic
measure of the likelihood of any mRNA being
associated with Pum1 in normally growing HeLa
cells, which was useful in further analysis of Pum1-
associated mRNAs.

While the Pum1 study demonstrates how RIP-
Chip data can be modeled probabilistically, another
FIGURE 3 | mRNA state dynamics. Depicted are the multiple states of a given type of mRNA \( (S_1, S_2, \ldots, S_n) \) defined by association with specific RBPs (different colored circles represent different RBPs). RBP–mRNA association and therefore mRNA states can change over time in response to the environment, resulting in the activation/repression of different functional modules (different colored rectangles—\( f_1, f_2, \ldots, f_n \)), thereby, yielding different phenotypic consequences that vary with cellular conditions (\( C_1, C_2, \ldots, C_n \)). These consequences are due to changes in the proportion of the mRNA (black triangles) existing in a given state, even though the total copy number for that mRNA species (\( m_{y,z,\ldots,n} \)) remains constant across biological conditions.

Probabilistic representation of ribonomic data, for example, in the form of a log likelihood ratio, is advantageous given single-cell stochasticity of gene expression and the resulting heterogeneity for a cell population inherent to most biological experiments.\(^{111,112}\) Furthermore, this representation allows the RBP–mRNA association data to be utilized in either a discrete or continuous manner for predictive network models. In these models, LOD scores could be used as weights for the RBP–mRNA interactions. Integration of ribonomic and other PTR data in a similar fashion will increase our understanding of both cis- and trans-combinatorial PTR.

As mentioned above, a key advantage in modeling data probabilistically is the increased ability to integrate multiple data types. Various methods other than ribonomics exist for studying PTR; for example, global study of alternative splicing has been achieved through splicing microarrays,\(^{113,114}\) and more recently studied along with alternate polyadenylation through the use of deep sequencing technology.\(^{115}\) In addition, analysis of global mRNA decay has been performed by the study from our laboratory.\(^{36}\) reveals one of the true strengths of probabilistic modeling; the ability to observe state changes. In this study, RIP-Chip of another RBP, HuR, was performed at various times after activation of T cells. By creating LOD scores for each timepoint, we were able to define the condition-specific likelihood of an mRNA being associated with HuR, and thus we were able to accurately observe remodeling of the mRNA populations associated with HuR RNPs during the dynamic process of T-cell activation.\(^{36}\) This study demonstrates how creating condition-specific likelihoods of mRNA–RBP association can be used to integrate the same type of data from different cellular conditions, and this type of probabilistic modeling should also allow for more accurate integration of multiple posttranscriptional data types. This study also showed that changes in HuR association were not correlated with changes in total mRNA levels, demonstrating that mRNA abundance is not reflective of mRNA states and that adventitious re-assortment of HuR with different mRNAs after cell lysis does not occur to a significant extent with the RIP-Chip procedure.\(^{7,34}\)
dependent splicing regulation, demonstrating an advantage of integrating multiple data types. These studies revealed potential mRNA decay regulons, sequence motifs shared by mRNAs that decayed with similar rates, and RBPs that may be mediating decay. Global analysis of translational profiles using polysome isolation is another common method to study PTR, often revealing groups of functionally related mRNAs that show changes in translational state during changing cellular conditions. Many of these studies infer which trans-acting factors may be mediating coordination at these various posttranscriptional stages; however, in order to more definitively identify the responsible trans-acting factors, these types of data must be integrated with data from ribonomic studies. This type of integration has been achieved previously, for example, in a study of the RBP FMRP, which is thought to be a repressor of translation. The authors of this study first performed RIP-Chip to identify mRNAs associated with FMRP mRNPs in mouse brains, then performed sucrose gradient fractionation to determine mRNAs that were differentially associated with polysomes in FMRP mutant compared with wild-type cells. This analysis revealed a significant number of the mRNAs identified in the RIP-Chip experiment showed differential association with polysomes in the absence of the FMRP protein. An unexpected result of this analysis was that of the FMRP-associated mRNAs that showed differential polysome association, about half increased and half decreased in association, demonstrating that the effect of an mRNA-binding protein on associated mRNAs in vivo is complex and may not be elucidated simply through ribonomic experiments.

Another demonstration of integrating multiple data types is demonstrated by studies of the RBP NOVA, a splicing factor. In the first study, a CLIP experiment was performed to identify mRNAs that are bound by the NOVA protein in mouse brains. In a follow-up study, splicing microarrays were used to determine alternate patterns of splicing in NOVA knockout mice compared with wild type. Finally, a third study integrated these and other data, resulting in formation of a genome-wide map to predict NOVA-dependent splicing regulation, demonstrating an advantage of integrating multiple data types.

Integration of multiple data types is valuable for determining not only consequences of a trans factor associating with mRNAs, but also how the different steps in the life of an mRNA may be coordinated. This concept is best demonstrated in a study by Hieronymus and Silver, showing that the yeast nuclear export proteins Mex67 and Yra1 show specificity in mRNA association, and genes that are transcriptionally co-regulated are also specifically bound to Yra1. This finding indicates coordination in the different steps of the life of an mRNA, showing that regulation of mRNA export may be linked to regulation of transcription.

Integration of diverse datasets can be difficult, especially when the data come from different studies and different cellular conditions. Many of the methods described above that study global PTR provide data that describe the consequences of mRNAs existing in various states, and thus by integrating these types of data with ribonomic data, it should be possible to model both mRNA states and the consequences of those states. As mentioned previously, one way that this process may be simplified is by probabilistic modeling of data, which allows condition-specific probabilities to be determined for mRNA states, such as RBP association, and definition of the consequences of these states, such as patterns of splicing, translation, or decay. Integrating data from tissue culture and studies using animal models can also be informative, as ribonomic analysis has been performed in various animal models (see Table 1), but many of the techniques to determine functional consequences are most feasibly performed in tissue culture. This integrative approach was demonstrated by the previous example of a study of FMRP which used animal models for ribomics experiments and tissue culture for translational profiling. Conversely, animal models offer advantages over tissue culture in other aspects, such as the ability to identify physiological phenotypes of alteration of PTR and the ability to study PTR in complex behaviors, such as learning and memory.

Integrating ribonomics experiments with other types of experiments can also prove to be a useful approach for identifying targets of miRNAs. In order to determine biochemically associated, physiologically relevant targets of miRNAs, investigators have genetically altered miRNAs (overexpression or knockout), then performed Ago RIPs and compared these with RIPs from control cells, with those mRNAs enriched in the RIPs from the overexpressing cells being ones that are potentially targeted by the miRNA. This method is valuable because it requires no a priori knowledge of ‘rules’ of miRNA targeting or conservation among species to determine targets, instead relying upon a genetic alteration and biochemical enrichment process.
CONCLUSION
Integration of multiple posttranscriptional methods and data types will lead to global models of coordination of PTR that will describe both mRNA states and consequences of these states, as well as the dynamics of the system being studied, in order to accurately model coordination of gene expression at the posttranscriptional level. The insights gained through global study of posttranscriptional control and coordination, as well as advancements in biological and computational methods and technologies, will be crucial in allowing these models to reflect the complexity and organization of the posttranscriptional environment. Modeling will allow better understanding of how posttranscriptional coordination of gene expression is important to human health, and allow exploitation of this knowledge to better treat and prevent human disease.

NOTES
One of the authors (J.D.K.) has a financial relationship with Ribonomics, Inc., and MBL, Inc., which hold licenses to RIP, clip, and other technologies related to topics in this review. The other authors have no known potential conflicts of interest.

REFERENCES


