A systems view of mRNP biology

Haley Hieronymus and Pamela A. Silver[1]

Department of Systems Biology, Harvard Medical School and the Dana-Farber Cancer Institute,
Boston, Massachusetts 02115, USA

Gene expression occurs through a complex mRNA–protein (mRNP) system that stretches from transcription to translation. Gene expression processes are increasingly studied from global perspectives in order to understand their pathways, properties, and behaviors as a system. Here we review these beginnings of mRNP systems biology, as they have emerged from recent large-scale investigation of mRNP components, interactions, and dynamics. Such work has begun to lay the foundation for a broader, integrated view of mRNP organization in gene expression.

The central processes of gene expression, from mRNA biogenesis to translation and degradation, revolve around messenger ribonucleoprotein (mRNP) complexes. These mRNP complexes consist of individual transcripts bound by a changing repertoire of proteins that mediate and couple the post-transcriptional events of gene expression. The mRNP lifecycle begins with transcription, when a host of proteins cotranscriptionally bind the nascent mRNA to form a growing mRNP (Fig. 1). Transcription itself involves mRNA-binding proteins, such as those that prevent cotranscriptional formation of RNA:DNA hybrids and promote transcriptional elongation (Huertas and Aguilera 2003). Other mRNA-binding proteins (RBPs) mediate cotranscriptional 5′ end capping, splicing, and editing; 3′ cleavage and polyadenylation; and quality control of mRNAs within nascent mRNP complexes (Gott and Emeson 2000; Neugebauer 2002; Reed 2003). Many of these processes can take place post-transcriptionally as well (Gott and Emeson 2000). A range of mRNA export factors are also cotranscriptionally recruited to transcripts and subsequently mediate mRNP translocation through the nuclear pores (Vinciguerra and Stutz 2004). Once in the cytoplasm, some mRNPs are localized to specific cellular regions by the signal recognition particle or by various motor proteins that bind mRNAs via mRNP adaptor proteins (Singer 2003; Van de Bor and Davis 2004). Transcripts are often translationally repressed during local-

ization (Huang and Richter 2004), and many mRNA-binding proteins involved in translation and translational regulation also affect mRNA transport and stability (Kuersten and Goodwin 2003). Ultimately, mRNAs undergo exonuclease-mediated degradation by normal, nonsense-mediated (NMD), and nonstop decay (NSD) pathways (Parker and Song 2004). This straightforward mRNP lifecycle gains complexity through the multiplicity and regulation of different mRNP pathways in gene expression processes (Vinciguerra and Stutz 2004). In this sense, mRNPs make up a complex, dynamic gene expression system rather than a single linear pathway. The study of this system constitutes the emerging field of mRNP systems biology.

Systems biology, at its most fundamental level, is the study of how biological systems function and give rise to emergent biological properties (Levesque and Benfey 2004). Such systems are typically defined as sets of elements that specifically interact to produce coherent emergent behaviors. Systems biology therefore has several defining steps (Fig. 2; Kitano 2002; Aitchison and Galitski 2003). The first involves discovery and integration of information about the system's component elements, processes, and dynamics under normal and perturbed states. This is often achieved using high-throughput techniques that yield large-scale or comprehensive information about the system features (Ge et al. 2003). The system can then be modeled to provide a framework for hypothesis testing. Finally, hypotheses from this model are iteratively generated and tested. This process allows testing and refinement of the model's underlying assumptions and subsequent biological discovery. Through this cycle, systems biology aims to determine how biological processes act as whole systems, rather than collections of a few model components.

Systems-level investigation of mRNP biology has thus far focused primarily on the first and, to a lesser extent, second aims of systems biology: system component characterization and modeling of mRNP processes. First, the physical and functional interactions among mRNP proteins, transcripts, and their parent genes are increasingly being examined on a system-wide scale at most points in gene expression. Progress has also been made in measuring the kinetics of mRNP processes such as transport and mRNA degradation. Additionally, rudimentary systems-level modeling of transcriptional regulation and mRNA degradation has been undertaken. Here we will
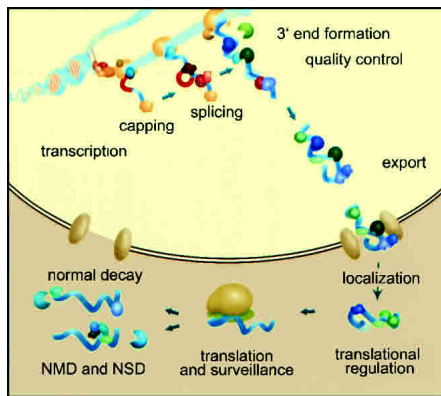
Figure 1. The mRNP lifecycle. mRNAs traverse the gene expression process from synthesis to translation and degradation in the form of dynamic protein-bound messenger ribonucleoprotein complexes (mRNPs). Within an mRNP complex, a transcript is bound by a changing set of proteins that mediate the cotranscriptional and post-transcriptional events that make up a transcript's lifecycle. Transcripts first undergo 5′ end capping, splicing in many cases, 3′ cleavage and polyadenylation, mRNA quality control by the nuclear exosome, and export factor recruitment. They are then exported to the cytoplasm, where some undergo specific subcellular localization. Transcripts are eventually translated, often in a regulated manner, and degraded. Degradation may occur through normal decay or through nonsense-mediated and nonstop decay pathways upon detection of these errors.

review these beginnings of mRNP system biology and the insights they provide into the mRNP lifecycle.

Systems approaches in mRNP biology

Systems-level mRNP biology has begun to emerge from quantitative genomic, proteomic, and microscopy-based investigation. These approaches have yielded information about mRNA and protein state, protein–protein and protein–nucleic acid interactions, protein localization, and dynamics of mRNP systems and processes. Since large-scale proteomic studies have generally focused on whole proteome interactions rather than those of mRNP systems in particular, this review primarily focuses on genomic and kinetic investigation of mRNPs in gene expression.

Global mRNA state has been extensively quantified at the level of the transcriptome, transcription, translation, and mRNA decay by high-throughput approaches. Transcriptomes, or whole-genome steady-state mRNA populations, are now routinely characterized by microarray and other high-throughput techniques (e.g., Velculescu et al. 1997; Holstege et al. 1998; Shoemaker et al. 2001; Kapranov et al. 2002; Kampa et al. 2004). Less commonly, large-scale transcriptional state has been established directly by microarray analysis of nuclear run-on products and indirectly from microarray-determined mRNA levels and degradation rates (Holstege et al. 1998; Fan et al. 2002). Genome-wide translational state has increasingly been characterized by microarray analysis

of transcripts associated with actively transcribing polyribosomes, as compared with total or nonpolyribosomal transcripts (e.g., Johannes et al. 1999; Kuhn et al. 2001; Kash et al. 2002; Arava et al. 2003; Preiss et al. 2003; Rajasekhar et al. 2003). Global mRNA degradation has also been profiled by monitoring changes in transcript abundance after transcriptional inhibition (e.g., Holstege et al. 1998; Bernstein et al. 2002; Fan et al. 2002; Wang et al. 2002; Grigull et al. 2004).

The RNA–protein interactions of mRNPs are commonly identified by high-throughput analysis of mRNAs that coimmunoprecipitate with particular mRNA-binding proteins. In this approach, an mRNA-binding protein of interest is immunoprecipitated, and the associated mRNA is isolated (Penalva et al. 2004). RNA–protein cross-linking is not routinely used with genome-wide RNA coimmunoprecipitation because of concerns about the effects of cross-linking reversal conditions on RNA quality; however, ultraviolet and chemical cross-linking approaches have been recently used for stringent identification of some mRNP target transcripts (Ule et al. 2003; Gilbert et al. 2004). Following coimmunoprecipitation, interacting transcripts are quantified by microarray or sequencing. For comparative microarray analysis, coimmunoprecipitated mRNAs and total mRNAs are often competitively hybridized to the same microarray.
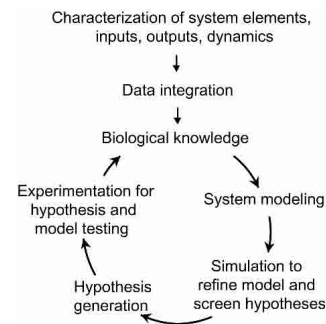

Figure 2. A systems biology approach. Systems biology begins with a foundational understanding of a system's structural and dynamic elements. It often uses utilizes high-throughput or massively parallel approaches to identify and characterize these elements, the inputs that affect them, and the effects to which they give rise. Basic genomic and proteomic approaches have begun to quantify the mRNA and protein components of cells and subcellular mRNP systems. Large-scale identification of mRNA or DNA that associate with RNA-binding proteins, as well as proteomic approaches, have begun to define the interactions within and among mRNPs in almost all events in gene expression. Integration of data from different approaches provides a more robust understanding of the system. This understanding of system components and dynamics then allows modeling of the system, and the model can in turn be tested computationally against existing data for refinement. For example, computational modeling of nonsense-mediated and general mRNA degradation pathways has been carried out through this iterative process; it has subsequently been used to test hypotheses about recognition of nonsense-containing transcripts in nonsense-mediated decay. Thus, the model can then be used to generate and formulate hypotheses that, when tested experimentally, give rise to further understanding of the system.

The resulting ratio of immunoprecipitated to total mRNA indicates the relative levels of protein binding across the transcriptome. These ratios are generally further normalized to the mean signal level, to the levels of internal standards (e.g., housekeeping genes), or to the levels of external standards added to the samples, yielding binding levels relative to these standards. Another complementary approach is to hybridize the immunoprecipitated sample to a mock or negative control immunoprecipitation. This method does not yield the protein's relative binding levels across the transcriptome, but instead establishes the set of transcripts that are bound by the protein above background binding levels. One caveat to these approaches is that they may not comprehensively identify all in vivo mRNA interactions of mRNP proteins, since RNA coimmunoprecipitation without cross-linking may not capture the full or specific repertoire of mRNAs bound throughout the cell and the mRNP lifecycle.

Since many nuclear mRNA-binding proteins are recruited to transcripts during transcription and can be cross-linked to DNA at the corresponding transcriptional loci, the protein–DNA interactions of mRNPs are increasingly being identified by coimmunoprecipitation of chromatin with specific mRNP proteins (ChIP) and subsequent microarray analysis. This approach is analogous to the RNA immunoprecipitation approach used to identify large-scale mRNA–protein interactions, except that the DNA and proteins are chemically cross-linked prior to immunoprecipitation. Such cross-linking captures both direct interactions and indirect mRNA-mediated interactions (Abruzzi et al. 2004). This method therefore provides a snapshot of a protein's chromatin interactions, or "genomic location", in living cells (e.g., Lieb et al. 2001; Yu et al. 2004). A disadvantage of this approach is that the chromatin targets of an mRNP protein may not precisely map to its mRNA targets due to post-transcriptional events.

The functional effects of mRNP proteins on their mRNA targets are typically identified by microarray-based profiling of cells lacking the wild-type proteins. While total mRNA levels are most commonly examined, global translational state in translation factor mutants has also been examined by polysomal transcript analysis (e.g., Johannes et al. 1999; Baron-Benhamou et al. 2003). The transcripts exported by specific nuclear export factors have also been identified by genome-wide microarray comparison of cytoplasmic to total transcript levels in knockdown cells (e.g., Herold et al. 2003; Rehwinkel et al. 2004); the transcripts with significantly decreased cytoplasmic levels require the particular export factor for normal export. A significant drawback of these approaches is that functional targets of mRNPs may be direct or indirect. Most notably, it is difficult to rule out secondary effects of removing or mutating proteins central to gene expression on mRNA populations.

One fundamental variable in large-scale mRNP analyses is the method used to define the nucleic acid population that is bound or affected by an RNA-binding protein from profiling data. This issue poses particular difficulties in analysis of nucleic acid–protein binding interactions. Since comparisons of bound chromatin to total chromatin are generally normalized to the mean binding level, only the binding level values relative to this mean binding level are determined. These relative data must be used to somehow define bound populations. One approach is to determine whether the binding levels show a bimodal distribution, a distribution with two modes that represent a subpopulation of enriched targets with one binding level mode and a second unbound population with another mode (e.g., Lieb et al. 2001). Cutoffs can then be used to isolate the enriched subpopulation to various degrees of confidence. However, many binding profiles do not have a bimodal distribution (e.g., Lee et al. 2002; Yu et al. 2004). Therefore, another common approach is to define the bound genes as those that show enrichment of the immunoprecipitated protein relative to the mean binding level at various confidence levels. High confidence levels are often used to avoid false positives at the risk of increasing false negatives. The bound population therefore depends on the method used to define it and the confidence levels chosen. To begin to address this issue, some analyses have used expression profiling data to help identify bound genes (Lee et al. 2002; Bar-Joseph et al. 2003). One recent study has used expression profiling data to define correlated expression patterns found in genes bound at high confidence and then used those patterns to identify genes with the same patterns in the binding genes of one or more factors at lower confidence levels (Bar-Joseph et al. 2003). This method substantially decreased the false negative rate without increasing the false positive rate. Nonetheless, false positives and negatives remain a thorny problem for systems-level investigations in mRNP biology.

Transcription

The first step in the mRNP lifecycle is transcription and concurrent cotranscriptional mRNA processing and packaging. Systems-level investigation of transcription has taken several main avenues. First, transcriptomes and their properties have been established by mRNA expression profiling of many systems under hundreds of conditions (e.g., DeRisi et al. 1997; Velculescu et al. 1997; Wodicka et al. 1997; Holstege et al. 1998; Shoemaker et al. 2001; Kapranov et al. 2002; Kampa et al. 2004). Genome-wide transcriptome characterization not only supplies fundamental data about biological state but also provides information that is useful to understanding the mechanisms of mRNP biology. For example, genome-wide mRNA transcription and degradation rates (Holstege et al. 1998; Wang et al. 2002; Grigull et al. 2004) have been widely used to analyze the properties of genes and transcripts bound by specific mRNP proteins in yeast (e.g., Hieronymus and Silver 2003; Casolari et al. 2004; Yu et al. 2004). Second, the transcriptional roles and interactions of many transcription factors have been determined by chIP profiling and mutant expression profiling (e.g., Holstege et al. 1998; Lee et al.

2000; Ren et al. 2000, 2002; Iyer et al. 2001; Shen et al. 2003; Zeitlinger et al. 2003; Cawley et al. 2004). Last, networks of transcriptional regulatory and binding interactions have been modeled to reveal the connectivity of the transcriptional machinery (Thieffry et al. 1998; Guelzim et al. 2002; Lee et al. 2002; Milo et al. 2002; Shen-Orr et al. 2002; Bar-Joseph et al. 2003; Covert et al. 2004; Harbison et al. 2004). Yeast transcriptional regulatory networks contain patterns, or motifs, shared by neuronal networks and electronic circuits, suggesting that these motifs may reflect a common information processing function (Milo et al. 2002). In all, these types of investigation into genome-wide transcriptional interactions and networks provide a framework for future study of post-transcriptional processes. While extensive review of such work is beyond the scope of this article, large-scale transcriptional studies are nonetheless central to the current analyses and future aims of mRNP systems biology.

Splicing

Splicing is a RNP-mediated processing step central to the expression of many genes. mRNA sequence information is transformed through this process to generate functional and regulatory diversity not directly encoded by the genome. Our understanding of the spliced transcriptome is far from complete, since global mRNA analysis approaches generally do not provide a comprehensive picture of splicing or mRNA splice variants. However, recent high-throughput studies have begun to expand our understanding of splicing and the spliced transcriptome. These large-scale splicing studies have three broad aims: (1) identifying the splicing patterns and spliced isoforms of all mRNAs under various conditions, including different tissue and cell types, developmental stages, and disease states; (2) understanding regulated alterations in the spliced transcriptome upon these changing conditions; and (3) understanding and characterizing the roles of different splicing factors in generating the spliced transcriptome. A long-term goal is to model splicing networks and their regulation.

The genome-wide splicing patterns of human transcripts have recently been established using several new types of microarrays (Shoemaker et al. 2001; Johnson et al. 2003). Splicing arrays that probe every consecutive exon junction have been used to predict the splicing patterns of nearly a third of the human transcriptome in 50 cell types (Johnson et al. 2003). Exon arrays that probe all predicted human exons (Shoemaker et al. 2001) and tiling arrays that probe regular, closely spaced intervals across chromosomes 21 and 22 (Shoemaker et al. 2001; Kapranov et al. 2002) have also been used to predict human alternative splicing patterns by various analytical approaches (e.g., Wang et al. 2003). These approaches predict that at least three-fourths of multiexon genes are alternatively spliced in humans. Revealingly, similar tissues have been found to have similar large-scale splicing patterns, while splicing in cell lines clusters separately (Johnson et al. 2003). The transcripts with the highest level of alternative splicing were found in cell lines and were enriched for transcripts involved in cell communication and enzyme regulation. These results indicate that alternative splicing patterns show specificity in different cell types and across different functional categories of proteins for the first time. However, exon junction, exon, and tiling arrays provide only composite pictures of each gene's alternative splicing because they cannot distinguish whether the various splicing events they predict occur in the same or different transcripts. Several other promising approaches have recently been developed to profile such alternative splicing, including splicing patterns of single mRNA molecules, but they have not yet been scaled to examine large numbers of genes (Yeakley et al. 2002; Zhu et al. 2003).

Recent work has begun to establish the effects and interactions of the conserved splicing machinery on a genome-wide scale. Splicing arrays representing each yeast intron, exon junction, and 3′ exon were used to examine the splicing effects of 18 nonessential splicing factors in *Saccharomyces cerevisiae* (Clark et al. 2002). Because both intron and exon sequences were probed, different types of splicing phenotypes could be observed in *S. cerevisiae* mutants of these splicing factors. For example, yeast Dbr1 is required for lariat debranching but not splicing itself, and loss of Dbr1 resulted in genome-wide increases in intron levels but had no effect on exon-junction levels (Clark et al. 2002). Interestingly, no fixed relationship was found between splicing factors' effects on unspliced and spliced mRNA levels (Clark et al. 2002). Clustering of these effects uncovered both similar and differential specificities of the profiled splicing factors. The splicing patterns of different transcript groups showed different dependencies on the yeast mRNA splicing factors studied, suggesting some degree of splicing specificity by these splicing factors. The relative similarities between these splicing factors' effects were also revealed by clustering (Clark et al. 2002), though few of these similarities are recapitulated by other hierarchical clustering approaches (Barrass and Beggs 2003). This work has nonetheless uncovered a previously unknown diversity of splicing factor effects in yeast.

Splicing factor targets have also been comprehensively identified by chromatin and mRNA coimmunoprecipitation, followed by high-throughput analysis (Kotovic et al. 2003; Ule et al. 2003; Blanchette et al. 2004). The *Drosophila* mRNA splicing and export factor dU2AF[50] appears to have both intron-containing and intron-less targets, as determined by microarray analysis of mRNAs affected by dU2AF[50] knockdown and of coimmunoprecipitating mRNA (Blanchette et al. 2004). However, the most highly bound 3% of mRNAs were greatly enriched for those of intron-containing genes. Similarly, the genomic targets of the *S. cerevisiae* U1 snRNP protein Prp42 were overwhelmingly intron-containing genes (Kotovic et al. 2003). Last, the mRNA targets and mRNA-binding sites of the neuronal splicing regulator Nova-1 were identified by a new sequencing-based UV-cross-linking and immunoprecipitation (CLIP) approach (Ule et al. 2003). The Nova-bound mRNA included novel

and known target transcripts (Buckanovich and Darnell 1997) that were functionally enriched for inhibitory synapse components, generating new hypotheses about the role of Nova proteins as autoantigens in neuronal ataxia (Ule et al. 2003). Since up to 15% of disease-causing mutations are thought to affect splicing (Krawczak et al. 1992), these large-scale approaches have the potential to not only identify the targets of the splicing machinery, but also to help uncover mechanisms of disease.

mRNP transport and localization

Nuclear mRNP export networks

After transcription and processing, mRNAs are transported from the nucleus to the cytoplasm by proteins that bind and export mRNAs as mRNP complexes (Stutz and Izaurralde 2003). Until recently, little was known about the mRNA composition of these export-competent mRNPs or the specificity of their bound mRNA export factors. To begin to address these issues, the genes cotranscriptionally bound and mRNAs bound and exported by a range of mRNA export factors, including Hpr1, Tho2, Sub2/UAP56, Yra1, Nab2, Npl3, Mex67/Nxf1, Nxf2, Nxf3, and p15, have been globally identified (Herold et al. 2003; Hieronymus and Silver 2003; Blanchette et al. 2004; Rehwinkel et al. 2004; Rodriguez-Navarro et al. 2004; Yu et al. 2004). These approaches have begun to illuminate the specificity of mRNA export factor interactions with DNA and RNA and the specificity of the mRNP export process in general.

mRNA export factors at all stages of export, from transcriptional coupling to translocation, show specificity for transcript and gene subpopulations (Herold et al. 2003; Hieronymus and Silver 2003; Rehwinkel et al. 2004; Yu et al. 2004). This export factor specificity often manifests as preferential association of an export factor to some transcripts or genes relative to others. *Drosophila* THO complex members Hpr1 and Tho2 are essential for nuclear export of a fifth and a seventh of all transcripts respectively, as shown by a drop in the transcripts' cytoplasmic levels upon Tho2 and Hpr1 knockdown (Rehwinkel et al. 2004). Consistent with this export specificity, yeast Tho2 is preferentially recruited to approximately a sixth of the genome (Yu et al. 2004). Yra1, an mRNA export factor that binds Tho2, preferentially associates with a sixth to an eighth of the genome (Yu et al. 2004) and an eighth of the transcriptome in yeast (Hieronymus and Silver 2003). Mex67/NXF1, an export factor thought to facilitate mRNP translocation through the nuclear pore, binds almost a fifth of transcripts in yeast. *Drosophila* NXF1 exports a fifth to a half of the transcriptome, as determined by cytoplasmic mRNA analysis after knockdown for 2 d, and a third to three-fourths of the transcriptome by analysis after knockdown for 4 d (Herold et al. 2003). Nearly identical populations of transcripts show mRNA export dependence on NXF1, the NXF1 coreceptor p15, and the Sub2 homolog UAP56 (Herold et al. 2003). Interestingly, NXF2, NXF3, and the transport factor Crm1 do not ap-

pear to be essential for mRNA export in *Drosophila* cells. Together these studies provide growing evidence for specificity in mRNA export and suggest that bulk mRNA export mutant phenotypes do not provide information about the number of mRNAs exported by the mutated protein (Rehwinkel et al. 2004).

The specificity of mRNA export factors often reflects functional specificity for transcripts that encode proteins with related functions. The export factors Tho2, Yra1, Npl3, and Mex67 all preferentially associate with mRNA and/or DNA targets enriched for targets of specific functional classes. For example, Npl3-bound genes represent many functions, including ribosomal biogenesis, reproduction and development, and stimulus perception (Yu et al. 2004). Yra1's DNA and RNA targets are enriched for targets encoding cell wall proteins, carbohydrate metabolism genes, and translation factors (Hieronymus and Silver 2003; Yu et al. 2004). Genes upregulated upon *Drosophila* Tho2 knockdown are significantly enriched for DNA repair proteins and mRNA export factors, likely reflecting compensation for the hyper-recombination and mRNA export phenotypes of Tho2 depletion (Rehwinkel et al. 2004). System-wide investigation into mRNA export factor targets has therefore revealed functional organization in mRNA export.

While the basis of mRNA export factor specificity is not known, this specificity is linked to transcription in several ways. mRNA binding of the cotranscriptionally recruited export factor Yra1 correlates with transcriptional frequency on a whole-genome scale, while binding of Mex67 does not (Hieronymus and Silver 2003). Similarly, the DNA association of export factors Yra1, Tho2, Npl3, and Nab2 shows varied levels of correlation with transcriptional frequency (Yu et al. 2004). Systematic analysis of mRNA export factor binding therefore suggests that transcriptional recruitment may play a role in establishing mRNP specificity for whole transcript populations.

RNA–protein interactions in mRNA localization

After nuclear export, many mRNAs are localized to subcellular regions by various mechanisms as an efficient method of localizing protein expression (Tekotte and Davis 2002). Recent studies have begun to characterize the extent and mechanisms of mRNA localization using genomic approaches (Diehn et al. 2000; Takizawa et al. 2000; Tenenbaum et al. 2000; Lerner et al. 2003; Shepard et al. 2003; Lopez de Silanes et al. 2004). A starting point for many of these studies has been the systematic identification of the mRNAs bound and localized by specific localization factors, such as the yeast She proteins that asymmetrically localize transcripts to the bud-tip of daughter cells (Takizawa et al. 2000; Shepard et al. 2003). More than 20 transcripts were found to be bud-tip localized by She1, She2, and She3 using coimmunoprecipitation and microarray analysis (Takizawa et al. 2000; Shepard et al. 2003). Subsequent characterization of one such transcript, IST2, found that its She-dependent localization is maintained by a septin membrane diffusion

barrier at the bud neck; this work thereby demonstrated multiple mechanisms of localization establishment and retention (Takizawa et al. 2000). Proteins involved in translational regulation and stability as well as mRNA localization, such as the Hu/ELAV family, have also been studied by these and other genomic approaches as discussed below (Diehn et al. 2000; Tenenbaum et al. 2000; Lerner et al. 2003; Lopez de Silanes et al. 2004; see Translation and Its Regulation). Thus, genome-wide analysis of mRNA localization factors' targets and interactions has begun to expand our understanding of mRNA localization mechanisms.

Quantitative kinetics and dynamics of mRNP transport

In addition to mapping mRNP component interactions, systems-level understanding of mRNA export and localization requires quantitative knowledge of transport kinetics. The kinetics of mRNP movement has been studied at several levels: intranuclear movement, nuclear-cytoplasmic export, and cytoplasmic movement and localization. A number of approaches have been used to observe and quantify these forms of mRNP movement in living cells (Singer 2003). To observe endogenous mRNPs in live cells, fluorescent proteins such as GFP have been used to tag mRNA transport proteins directly and mRNAs indirectly through insertion of RNA hairpin sequences that bind proteins such as GFP-U1A or GFP-MS2 (e.g., Bertrand et al. 1998; Brodsky and Silver 2000; Rook et al. 2000; Calapez et al. 2002; Carrero et al. 2003; Forrest and Gavis 2003; Fusco et al. 2003).

Intranuclear mRNA kinetics have recently been quantified in real time from the point of transcription onward (Janicki et al. 2004; Shav-Tal et al. 2004). These kinetics were established in a mammalian cell system that enables in vivo visualization of transcription and transcript movement; the transcription locus was visualized via a YFP-lac repressor bound to hundreds of upstream *lac* operator sequences, and the transcribed mRNA was tagged with MS2 hairpins and visualized by bound YFP-MS2 (Janicki et al. 2004). Using this system, mRNA synthesis is detected immediately upon transcriptional induction, though the gene locus remains condensed for several minutes post-induction. mRNA levels show an initial linear increase that slows over time, possibly due to inhibition of further transcription by accumulated transcripts. By 150 min, the mRNA localizes to the transcriptional locus as well as to a particulate pattern throughout the nucleoplasm (Janicki et al. 2004). The mRNA particles travel through the nucleoplasm at 0.3–0.8 µm/sec in a pattern that fits a simple diffusion model (Shav-Tal et al. 2004). Their diffusion coefficient decreases linearly with temperature (Shav-Tal et al. 2004), consistent with simple diffusion and known nuclear kinetics of mRNP proteins (Calapez et al. 2002). Interestingly, this temperature relationship depends on the mRNA binding ability of the tested mRNP protein PABP2 (Calapez et al. 2002). Two types of nuclear mRNA diffusion were observed: simple diffusion where

mRNP movement does not deviate from the diffusion model, and partially corralled diffusion where the mRNP movement appears directionally hindered (Shav-Tal et al. 2004). Since these studies are carried out with a single artificial transcript, it remains to be seen whether other mRNAs exhibit different or diverse nuclear kinetics.

In contrast to work on intranuclear transport, elucidation of nuclear export kinetics has focused on protein movement rather than mRNA movement (Ribbeck and Gorlich 2001; Siebrasse and Peters 2002; Gorlich et al. 2003). Investigation of protein export has led to fruitful computational modeling of nuclear protein transport (Smith et al. 2002; Gorlich et al. 2003; Kustanovich and Rabin 2004) and generated several theoretical models for protein translocation mechanisms (Rout et al. 2000; Ribbeck and Gorlich 2001; Kustanovich and Rabin 2004). However, little is known about the kinetics of endogenous mRNA translocation. mRNA export is assumed to not be rate limiting in the overall mRNA localization process, since mRNA does not accumulate at the nuclear periphery under normal conditions. Nuclear export of two transcripts, LT-α and β-globin, has also been shown to occur within 2–5 min with first-order reaction kinetics (Audibert et al. 2002). In the future, it will be interesting to see whether mRNP translocation fits current models of receptor-bound protein translocation, since mRNA translocation has been proposed to occur through general mRNA export receptors that associate with mRNA-binding adaptors (Stutz and Izaurralde 2003).

The kinetics of cytoplasmic mRNA localization have been established for specific transcripts in yeast, *Drosophila*, and mammalian cells by in vivo visualization (Bertrand et al. 1998; Rook et al. 2000; Fusco et al. 2003). Early work developed systems for visualizing cytoplasmic movements of MS2 stem loop-containing mRNA constructs using GFP-MS2 (e.g., Beach et al. 1999; Takizawa and Vale 2000). A yeast mRNA construct containing the ASH1 3'-UTR and MS2 stem loops was subsequently found to move to the bud-tip at speeds consistent with known myosin V motor speeds (Bertrand et al. 1998). In mammalian hippocampal neurons, a dendritically localized mRNA construct with the calcium/calmodulin-dependent kinase IIα 3'-UTR was found to move at roughly one-tenth this speed (Rook et al. 2000). While the bud-tip-localized mRNA's movement was largely directional (Bertrand et al. 1998), the calcium/calmodulin-dependent kinase IIα construct had three distinct classes of movement: oscillatory movement and directional movement toward and away from the cell body (Rook et al. 2000). The mRNA speed did not change upon neuronal depolarization, but movement down the dendrites doubled and oscillatory motion decreased. Kinetic studies have therefore provided a window into dynamic mRNA movements and their biological regulation in vivo.

In addition to the kinetics of localized mRNAs, the movements of single nonlocalized mRNA molecules have also been recently characterized (Fusco et al. 2003). Nonlocalized mRNAs have been hypothesized to diffuse through the cytoplasm; however, a lacZ-MS2 mRNA

construct containing the SV40 or human growth hormone reporter 3′-UTR sequences exhibited static behavior, diffusion, and directed movement at speeds an order of magnitude slower than seen in neurons. Interestingly, microtubules appeared to be required for the "corralling" of diffusive movement. These studies are the first to establish the kinetics of mRNP movement and provide data for future modeling of mRNPs in the nucleus, cytoplasm, and between.

Translation and its regulation

Global translational state and regulation

Translational state and its regulation have recently begun to be defined on a genomic scale (Diehn et al. 2000; Kuhn et al. 2001; Kash et al. 2002; Arava et al. 2003; Lerner et al. 2003; Toyoda et al. 2003; Qin and Sarnow 2004). Global translational state has increasingly been assayed through microarray analysis of polyribosome-associated mRNAs relative to total or nonribosome-associated mRNA levels. This approach has been used to determine the genome-wide translation levels of yeast mRNAs (Arava et al. 2003). For almost all translated genes, the majority of mRNA molecules were actively translated. Transcripts were bound by ribosomes at one-fifth the maximal ribosome density on average, consistent with translational initiation being rate limiting.

Global translational responses to changing metabolic and cellular state have also recently been established for the first time. Rapid transfer of yeast from fermentable glucose to nonfermentable glycerol carbon sources results in a global decrease in translation (Kuhn et al. 2001). However, much of this decrease is accompanied by a decrease in overall mRNA abundance, making it difficult to distinguish between transcriptional, translational, and mRNA stability alterations. Nonetheless, 2% of the transcripts showed a more than twofold decrease in translational level without a corresponding decrease in overall abundance (Kuhn et al. 2001). More than two-thirds of these transcripts encoded ribosomal proteins, and conversely all ribosomal protein transcripts showed translational repression upon carbon shift to some degree. Interestingly, this ribosomal regulation resembles 5′-terminal oligopyrimidine (5′ TOP)-mediated translational repression of mammalian ribosomal protein transcripts upon serum starvation (Keene 2001).

In addition to translational state, fundamental mechanisms of translation have been investigated through genome-wide studies (Johannes et al. 1999; Diehn et al. 2000; Tenenbaum et al. 2000; Grolleau et al. 2002; Baron-Benhamou et al. 2003; Lerner et al. 2003; Preiss et al. 2003; Rajasekhar et al. 2003; Toyoda et al. 2003; Gera et al. 2004; Qin and Sarnow 2004). Translation initiation and localization have been major focuses of such work. Some of the earliest whole-genome translation studies examined the genome-wide requirements for initiation factors eIF4GI and eIF4GII in translation (Johannes et al.

1999; Baron-Benhamou et al. 2003). eIF4G, as part of the eIF4F complex, associates with cap binding eIF4E and increases its affinity for the 5′ m^7GpppN cap (von der Haar et al. 2004). This cap-eIF4E-eIF4G chain of interactions is essential for full cap-dependent translation and consequently has the potential to regulate translation. Transcripts that do not require full eIF4G levels for translation have been identified by microarray profiling of polysome-associated mRNAs in poliovirus-infected cells that have reduced eIF4G concentrations (Johannes et al. 1999). A small percentage of transcripts show heightened polysome association at these lower eIF4G levels, representing transcripts that are potentially translated through cap-independent initiation at internal ribosome entry sites (IRESs). The identified transcripts included mRNAs encoding immediate early factors and mitogen-activated regulators (Johannes et al. 1999), consistent with subsequent work demonstrating preferential polysome association of IRES-containing transcripts during mitosis (Qin and Sarnow 2004).

Besides providing a broad window into translational initiation, genomic approaches have established patterns of translational localization (Diehn et al. 2000; Lerner et al. 2003; Toyoda et al. 2003). Localized translation occurs predominantly at the ER, where proteins are co-translationally threaded into and through the membrane (Nicchitta 2002). Transcripts of such membrane and secretory proteins have been identified by microarray analysis of membrane-associated mRNAs and polyribosomes (Diehn et al. 2000; Lerner et al. 2003; Toyoda et al. 2003). Significantly, transcripts of some cytoplasmic proteins have been found to associate with membrane-bound polysomes, suggesting that membrane-associated translation may not confer a secretory or membrane fate in all cases (Diehn et al. 2000; Lerner et al. 2003).

mRNP interactions in translation and translational regulation

Translational regulation is mediated by a host of mRNA-binding proteins and microRNAs (miRNAs) that make up mRNP and miRNP complexes. These regulatory RNPs often affect transcript stability and localization as well as translation. Until recently, the mRNA targets and composition of such RNPs were identified and studied in small numbers. A more comprehensive understanding of RNP interactions and targets in translation has subsequently emerged from microarray-based analysis of polysomal and coimmunoprecipitating transcripts (e.g., Johannes et al. 1999; Tenenbaum et al. 2000; Brown et al. 2001; Miyashiro et al. 2003; Gerber et al. 2004; Inada and Guthrie 2004).

Translational RNA-binding proteins (RBPs) have widely been found to bind specific transcript subpopulations, shown for the first time through large-scale identification. These RBPs include HuR (Lopez de Silanes et al. 2004), HuB (Tenenbaum et al. 2000), Scp160p (Li et al. 2003), the pumilio-fem-3-binding factor (Puf) family (Gerber et al. 2004), αCP2 (Waggoner and Liebhaber 2003), and Lhp1 (Inada and Guthrie 2004). The mRNA

populations associated with these proteins vary in size from ~1% to 15% of the mRNAs assayed. The bound mRNAs comprise unique but often overlapping sets. Mammalian translational and stability regulator HuB and eIF4E bind a tenth of assayed transcripts relative to a negative control immunoprecipitation in embryonal carcinoma stem cells, whereas the poly(A)-binding protein PABP1 has a more ubiquitous binding pattern consistent with its broad role in mRNA processing and translation (Tenenbaum et al. 2000). Similarly, HuR binds ~15% of transcripts (Lopez de Silanes et al. 2004). The yeast Puf translational repressor proteins also each bind distinct transcript populations with 12% bound by more than one Puf protein (Gerber et al. 2004). The partial redundancy of these translational regulators suggests that they bind transcripts in a combinatorial manner (Tenenbaum et al. 2000; Keene 2001).

Large-scale characterization of translational RBP targets has also identified possible sequence determinants of translational mRNP specificity. The Hu proteins bind individual mRNAs containing 3′ AU-rich elements (AREs), and correspondingly the HuB target transcripts were found to be enriched in U-rich motifs (Tenenbaum et al. 2000). Similar sequence motifs were identified from the HuR-bound transcript population and were then successfully used to predict additional HuR target transcripts (Lopez de Silanes et al. 2004). Puf3-, Puf4-, and Puf5-bound mRNA targets were also shown to contain distinct sequence motifs that selectively bind to the corresponding protein in vitro (Gerber et al. 2004). Nonetheless, these individual sequences are neither universally nor exclusively found in the target mRNA populations. Future work will be needed to help resolve the combinatorial contributions of mRNA sequence, structure, and secondary protein interactions in establishing mRNP specificity.

mRNP binding specificity is hypothesized to generate specificity in translational regulation. This translational specificity might mediate coregulation of mRNAs encoding functional protein classes and thereby functionally organize gene expression (Keene and Tenenbaum 2002). Indeed, translational regulatory proteins specifically bind transcripts encoding functional and colocalized protein classes (Brown et al. 2001; Miyashiro et al. 2003; Waggoner and Liebhaber 2003; Gerber et al. 2004). For example, the Puf proteins each bind transcript groups enriched in mRNAs encoding proteins with common localization and functions: Puf1 and Puf2 targets encode membrane-associated transport and vesicular trafficking proteins, Puf3-bound mRNAs encode mitochondrial translation factors and regulators, and Puf4 and Puf5 targets represent nuclear structure and processing factors (Gerber et al. 2004). In addition to this steady-state functional specificity, dynamic changes in mRNP specificity are induced by cellular changes such as differentiation (Tenenbaum et al. 2000). mRNP specificity and its dynamics in translation therefore act to post-transcriptionally organize gene expression.

The specific mRNA interactions of translational regulators are also central to disease. Fragile X syndrome, the most common form of inherited mental retardation, is caused by lack of the mRNA-binding fragile X mental retardation protein (FMRP). FMRP acts as a neuronal translational repressor that is essential for proper synaptic development (Antar and Bassell 2003). Several hundred mRNA targets bound by rat FMRP have been identified by coimmunoprecipitation from brain extracts and microarray analysis (Brown et al. 2001). Half of these binding targets were found to be translationally regulated by human FMRP, as demonstrated by increases or decreases in their polysomal binding in Fragile X patient cells. Two-thirds of the regulated transcripts contain a FMRP-binding G-quartet structure identified in vitro, and many encode proteins involved in synaptic development and cognition (Darnell et al. 2001). To further understand FMRP binding specificity in vivo, FMRP targets were identified in situ in hippocampal neurons by antibody-positioned RNA amplification (Miyashiro et al. 2003). In this new technique, FMRP-associated transcripts are reverse transcribed in situ using anti-FMRP antibody-targeted probes, and the resulting antibody–cDNA complexes are immunoprecipitated and identified by microarray. The identified FMRP targets require FMRP for normal transcript distribution in the brain and for normal protein expression levels, demonstrating the functional effects of FMRP on its target transcripts. These and other results have led to the model that FMRP suppresses translation of specific mRNAs during and after mRNA localization to the synapse in the absence of proper stimulus. *Drosophila* FMRP has been shown to translationally repress individual homologs of murine FMRP1 target mRNAs. Deletion of one such homolog, *futsch,* suppressed the neuronal defects of *Drosophila* lacking the *FMR1* homolog, while overexpression led to neuromuscular junction phenotypes (Zhang et al. 2001). Together these studies provide an unbiased view of FMRP's molecular function at mRNA and generate further hypotheses as to how FMRP abrogation gives rise to Fragile X phenotypes.

In contrast to the growing body of identified translational mRNP targets, miRNA targets and miRNP composition have proven difficult to identify. While computational and genomic approaches have been applied to miRNA prediction (Bartel 2004) and quantification (Liu et al. 2004; Schmittgen et al. 2004), the identification of miRNP targets has largely remained in the realm of low-throughput investigation. Computational approaches to miRNA target prediction have begun to address this problem and have generated hypotheses about miRNA functions based on the target mRNA functions (Bartel 2004). Extension of RNP analysis approaches, such as large-scale polysomal and coimmunoprecipitating mRNA analysis discussed here, to the problem of miRNP target analysis has great potential to expand global understanding of miRNP biology in the future.

mRNA degradation

mRNA degradation plays an essential role in the mRNP lifecycle as a determinant of steady-state mRNA levels

and a source of mRNA regulation and quality control. General mRNA decay can occur by two pathways: $5' \rightarrow 3'$ exonucleolytic degradation by the exonuclease Xrn1 and $3' \rightarrow 5'$ degradation by the exosome complex (Maquat 2004; Parker and Song 2004). Degradation of aberrant transcripts occurs by nuclear degradation, NMD, and NSD. These pathways and the proteins that mediate them have been extensively delineated through studies of individual transcripts. Global studies have now begun to examine and model these pathways.

Genome-wide mRNA degradation patterns have been established in a number of systems, including *Escherichia coli* (Bernstein et al. 2002; Selinger et al. 2003), *S. cerevisiae* (Holstege et al. 1998; Wang et al. 2002; Grigull et al. 2004), *Arabidopsis* (Gutierrez et al. 2002), and human cells (Fan et al. 2002; Raghavan et al. 2002; Yang et al. 2003). These studies most commonly define the mRNA decay rates of thousands of genes by profiling changes in mRNA levels upon transcriptional inhibition (e.g., Grigull et al. 2004). To avoid potential effects of transcriptional inhibition on mRNA stability, microarray profiling of nuclear run-on products has also recently been used (Fan et al. 2002). mRNA stability has not been found to correlate with ORF or operon length, mRNA abundance, or ribosome density (Bernstein et al. 2002; Wang et al. 2002). Intriguingly, mRNA stability is instead often correlated among transcripts encoding constituents of a common macromolecular complex or functional class. Decay rates are highly similar among mRNAs encoding components within stoichiometric complexes such as the nucleosome, proteasome, and ribosome (Wang et al. 2002). mRNAs encoding central metabolism proteins have lower degradation rates than average (Bernstein et al. 2002; Wang et al. 2002; Yang et al. 2003; McCarroll et al. 2004). Transcripts encoding transcriptional machinery and factors (Wang et al. 2002; Yang et al. 2003; McCarroll et al. 2004) as well as translation and ribosome biogenesis machinery (Wang et al. 2002; Grigull et al. 2004) are rapidly turned over. The high turnover of transcriptional and translational machinery suggests that mRNA degradation may facilitate rapid regulation of these central gene expression processes.

Large-scale studies of mRNA decay have lent support to hypotheses that differential mRNA degradation plays a general role in regulation of gene expression. Individual gene expression has been shown to be affected by mRNA degradation induced by a variety of cellular signals and conditions (Thomson et al. 1999). On a large scale, cellular signaling and stresses such as heat shock and UV irradiation have been shown to induce alterations in genome-wide mRNA degradation patterns (Fan et al. 2002; Raghavan et al. 2002). Approximately one-tenth of transcripts are found to be destabilized and 3% are stabilized under stress conditions, for example (Fan et al. 2002). Thus, while transcription is widely accepted as a major global regulator of gene expression, mRNA degradation also seems to have a significant regulatory role.

The molecular foundation of global mRNA degradation specificity is not well understood. While human mRNA decay rates correlate with the presence of AREs (Raghavan et al. 2002; Yang et al. 2003), AREs are not found comprehensively nor exclusively in rapidly decaying mRNAs (Bernstein et al. 2002; Yang et al. 2003). New sequence motifs have been found in rapidly decaying transcript populations but cannot account for the stability differences of all transcripts (Yang et al. 2003). It is currently unknown how sequence elements establish the global mRNA stability patterns or how the interplay between mRNA stability, translation, and localization affects these patterns. In the future, systems-level investigation of mRNA degradation pathways may identify target populations that define sequence and other determinants of mRNA decay.

Mechanistic insights into mRNA degradation are beginning to emerge from global expression and stability analysis of degradation machinery mutants (Lelivelt and Culbertson 1999; He et al. 2003; Grigull et al. 2004). Specific stability factors, such as the yeast Hu homolog Pub1, deadenylation factors Ccr4 and Pan2, and translation/stability factor Puf4, regulate the stability of transcripts enriched in translational machinery mRNAs (Grigull et al. 2004). The general yeast decapping enzyme Dcp1 and the $5' \rightarrow 3'$ exonuclease Xrn1 surprisingly affect the levels of only 15% and 20% of transcripts upon deletion, respectively (Lelivelt and Culbertson 1999; He et al. 2003). This relatively small target transcript population suggests that $3' \rightarrow 5'$ degradation, rather than $5' \rightarrow 3'$ Xrn1-mediated degradation, may be the central mRNA decay pathway in yeast. Alternatively, expression profiling of knockout mutants might not reveal the full set of $5' \rightarrow 3'$ degradation targets if transcription or other degradation pathways compensate for loss of $5' \rightarrow 3'$ degradation in the mutants.

The global effects of NMD on the yeast transcriptome have also been established using mutant profiling. Deletion of NMD factors Upf1, Nmd2, and Upf3 each alter the levels of a tenth of the yeast transcriptome, with almost complete overlap between these factors (Lelivelt and Culbertson 1999; He et al. 2003). Their targets are enriched for transcripts of cellular defense proteins and underenriched in transcripts encoding protein synthesis machinery (Lelivelt and Culbertson 1999; He et al. 2003), in contrast to Ccr4 and Pan2 deadenylase targets (Grigull et al. 2004). A third of the yeast Upf targets represent structural classes of transcripts, such as transposable element, pseudogene, and bicistronic transcripts. These Upf target classes include intron-containing genes (He et al. 2003), despite the observation that the yeast Upf3 deletion mutant shows little accumulation of unspliced transcripts by splicing array (Clark et al. 2002). Since codeletion of Xrn1 with any of the Upf proteins has a similar effect on mRNA levels as deletion of Xrn1 alone, the general Xrn1 pathway may mediate $5' \rightarrow 3'$ degradation of the majority of NMD targets (He et al. 2003). In all, steady-state global studies have begun to define the scope and mechanisms of NMD. Similar studies will undoubtedly establish the contributions of different mRNA degradation pathways to total cellular mRNA decay in the future.

An ultimate goal of investigations into global mRNA degradation is the reconstruction of decay regulatory networks. In the past several years, both general and nonsense-mediated degradation have been modeled computationally (Cao and Parker 2001, 2003). General mRNA degradation has been modeled to include transcription, deadenylation, deadenylation-independent and -dependent decapping, and $3' \to 5'$ and $5' \to 3'$ cytoplasmic degradation (Cao and Parker 2001). This model assumes zero-order transcription, first-order decay steps, and lack of coupling between degradation steps. Nonetheless, this framework fairly accurately models the degradation of PGK1 and MFA2 transcripts using a combination of estimated and experimentally determined rate constants. The computational model was subsequently extended to NMD by introducing degradation initiated by aberrant mRNA recognition and followed by rapid decapping, accelerated deadenylation, and degradation upon varying levels of deadenylation (Cao and Parker 2003). This NMD model recapitulates degradation of nonsense-containing PGK1 constructs, if nonsense-containing transcript recognition is assumed to occur with near perfect efficiency. This suggests that the "leaky-surveillance" NMD model, in which nonsense-containing transcripts are recognized with limited efficiencies, does not accurately reflect experimental data for PGK1. Since these computational models were constructed using a theoretical framework generated from studies of transcripts such as PGK1, it will be interesting to see whether they consistently model other transcripts. A major hurdle to testing these computational models more extensively is their requirement for transcript-specific rate constants for each degradation step. High-throughput approaches may someday provide such degradation kinetics for large-scale modeling of mRNA decay.

An emerging mRNP systems biology

mRNP systems biology has recently begun to emerge from studies of mRNP components and dynamics, particularly on a large scale. This approach aims to provide an understanding of post-transcriptional gene expression and its pathways as a whole system. Many emergent properties of post-transcriptional mRNP processes may only be apparent from such a systems-level perspective.

One such theme that has dominated recent large-scale mRNP characterization is that mRNP proteins and processes have specificity for transcript and gene populations. Indeed, large-scale profiling approaches tend to reveal any differences or specificity in the nucleic acid populations by their very nature. mRNP specificity suggests that mRNA-binding machinery organizes the flow of transcripts through central gene expression processes (Fig. 3). Such a model is a fundamental shift from the canonical view that almost all transcripts are routed through these processes in a generic or unregulated manner. The significance of this specificity is highlighted by its functional nature: Many of the genes and transcript targets of mRNP proteins or processes are enriched for
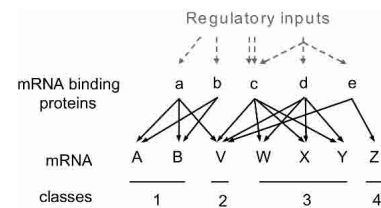

Figure 3. Post-transcriptional organization of gene expression by mRNPs. mRNA-binding proteins that mediate gene expression processes from mRNA synthesis to degradation (lowercase) bind mRNAs (uppercase) in a combinatorial manner. Some proteins bind or regulate their own mRNA, as shown here for proteins a and b. The specific combination of interactions may dictate the path a transcript takes through post-transcriptional gene expression. The transcript populations bound by particular mRNA-binding protein(s) are often enriched for transcripts encoding proteins of specific functional and other characteristic classes. In this way, protein–mRNA interactions structurally and functionally organize gene expression. It has further been suggested that regulatory inputs (gray) regulate binding of RNA-binding proteins (RBPs) to their target transcripts and thereby regulate RBP-mediated organization of gene expression.

transcripts encoding functional or interacting groups of proteins. mRNP proteins also show specificity for their own genes and transcripts in processes such as nuclear export and translation, suggesting widespread autoregulation by mRNA-binding proteins (e.g., Hieronymus and Silver 2003; Waggoner and Liebhaber 2003).

mRNP-binding proteins further show combinatorial interactions with genes and transcripts across the genome. Combinations of RBPs may act to define mRNA subpopulations that encode more functionally coherent groups of transcripts, as is seen in the specific combinatorial binding of both Puf4 and Puf5 to the histone core transcripts (Gerber et al. 2004). mRNA-binding proteins therefore form intricate post-transcriptional RNA–protein networks, similar to the transcriptional DNA–protein networks generated by DNA-binding proteins. This combinatorial mRNP specificity supports the hypothesis that mRNA-binding proteins post-transcriptionally coordinate transcript populations that encode functionally coherent protein sets, much as DNA-binding proteins regulate such functional gene classes (Keene 2001; Keene and Tenenbaum 2002).

It has been hypothesized that mRNPs not only organize but also regulate functional sets of transcripts and thereby mediate functionally specific post-transcriptional regulation (Keene and Tenenbaum 2002). Small-scale studies have demonstrated that handfuls of functionally related mRNAs are regulated by specific RBPs during $3'$ end formation, mRNA export, and translation (Saavedra et al. 1996; Dominski and Marzluff 1999; Mendez and Richter 2001; Lei and Silver 2002; Eisenstein and Ross 2003). On a larger scale, widespread post-transcriptional regulation of larger functionally enriched transcript populations by RBPs has yet to be definitively established. Cellular changes and signaling regulate post-transcriptional processing of large and functionally enriched transcript groups, as evidenced by translational

effects of Ras, Akt, and mTOR signaling (Grolleau et al. 2002; Preiss et al. 2003; Rajasekhar et al. 2003; Gera et al. 2004) and mRNA stability changes upon various cellular stresses (Fan et al. 2002). Cellular state changes are also known to alter the mRNA-binding profiles of specific mRNA-binding proteins such as HuB (Tenenbaum et al. 2000). However, regulated changes in mRNA-binding protein specificity have not yet been shown to cause subsequent alterations in post-transcriptional processing of the target mRNA populations. Additional work is therefore needed to demonstrate that cellular changes regulate functional classes of transcripts through RBP specificity and to investigate the broader cellular functions of this specificity. Protein modification of RBPs may play a significant role in such regulation (Yu et al. 2004). Systems-level investigation of mRNP biology has nonetheless revealed emergent properties of mRNP systems in the organization and potential regulation of gene expression.

Even with these insights, significant work is still needed to build a systems-level understanding of mRNP interactions and biology. Definition of mRNP system elements and dynamics, a key aim in systems biology, is far from realized. Large-scale analysis of mRNP systems is lacking in several technical areas, such as in vivo nucleic acid–protein interactions, protein–protein interactions, and the dynamics of these mRNP interactions and processes. Though information about in vivo RNA–protein interactions has been elusive heretofore, newer techniques including protein–RNA cross-linking and antibody-positioned amplification may help address this deficit (Ule et al. 2003; Gilbert et al. 2004). Tiling arrays additionally hold the promise of fine-scale, subgenic mapping of nucleic acid–protein interactions and global mechanistic insights into RBP recruitment. Similarly, new and existing large-scale proteomic data have the potential to elucidate mRNP protein–protein interaction networks (e.g., Bader et al. 2004). For example, proteomic characterization has shown that *S. cerevisiae* RNA transport complexes appear more dispersed and connected to complexes of other functions than those in mRNA metabolism and translation (Gavin et al. 2002). To date, many proteomic investigations have revealed how the proteins involved in mRNA synthesis, splicing, processing, transport, and degradation interact (e.g., Uetz et al. 2000; von Mering et al. 2002) and form complexes that act on mRNA (e.g., Zhou et al. 2002; Aloy et al. 2004). The challenge now is to sift large-scale proteomic data for a better understanding of the protein interactions in mRNP systems and to integrate these data with data from other approaches. mRNP biology also needs to move beyond static analysis of component interactions and look at the dynamics of post-transcriptional processes and their interactions on a large scale. A full understanding of mRNP biology will require knowledge of how different inputs affect and regulate these systems and their component elements.

A second barrier to understanding mRNP systems lies in the lack of characterization of most mRNA-binding proteins (Issel-Tarver et al. 2002; Gerber et al. 2004). The

interactions and functions of only a few mRNP proteins have been systematically examined in most gene expression processes; in other processes such as 3′ end processing, they have gone almost entirely unexplored on a system-wide basis (Hajarnavis et al. 2004). Nonetheless, 10 of the >500 predicted yeast mRNA-binding proteins (She2, Khd1, Scp160p, Yra1, Mex67, and Puf1 through Puf5) bind almost 40% of the transcriptome (Gerber et al. 2004), and a tenth of known mRNA export factors localize to 41% of the genome (Yu et al. 2004). As the repertoire of known nucleic acid–protein interactions expands, the coverage and distribution of mRNA-binding proteins over the genome, transcriptome, and functional groups can begin to be more fully evaluated. This and other work may also reveal the extent of and differences in functional specificity among mRNP proteins.

Last, mRNP systems have gone largely unexplored on many biological scales. mRNP biology is generally examined in cell populations and therefore reflects the workings of the average cell (Levsky and Singer 2003). Extension of single cell approaches may someday define mRNP interactions and compositions within individual cells and even within individual mRNP complexes. Recent and emerging large-scale approaches for subcellular protein localization (Kumar et al. 2002; Huh et al. 2003; Mootha et al. 2003) may also provide new information about localized protein function in mRNP biology. In addition, little is known about genome-wide mRNP biology at tissue or organismal levels. However, work is underway to examine the expression patterns of almost all RNA-binding proteins in murine brain (E. Minet, A. McKee, and P. Silver, unpubl.). Such studies are needed to understand the properties of mRNP systems in individual cells, tissues, and multicellular organisms.

Integration of systems-level data, another key element in systems biology approaches, have also been untapped in mRNP biology (Levesque and Benfey 2004). Integration of data on mRNP protein–protein and protein–nucleic interactions, protein and transcript levels, protein localization, and mRNP dynamics is vital to gaining a robust understanding of mRNP system biology (Ge et al. 2003; Hegde et al. 2003). At a basic level, attempts at data integration suffer from a lack of convenient and consistent post-publication data sharing and annotation conventions. Meaningful data integration would also benefit from increased standardization of experimental and basic analytical methods. New analytical methods are also needed for combining diverse types of systems-level data. Ultimately, a major challenge of mRNP systems biology is to integrate information provided by many experimental approaches.

Beyond these basic foundations of mRNP systems biology, modeling and model testing has been minimal in mRNP systems biology. Only mRNA synthesis and degradation, processes at opposite ends of the mRNP lifecycle, have been modeled (Cao and Parker 2001, 2003; Wei et al. 2004). Even then, comprehensive transcriptional networks are rarely modeled kinetically, and kinetic mRNA turnover models have only been tested for a few transcripts rather than the entire system. More

extensive modeling of these and other mRNP system processes will require a better understanding of the kinetics, inputs, and outputs of these processes.

In sum, recent investigation into the dynamics, interactions, and global effects of mRNPs has begun to lay the foundation for mRNP systems biology. This work has revealed unexpected emergent properties of mRNP systems, including post-transcriptional organization of gene expression and function. In the future, mRNP systems biology needs to delve into the iterative cycle of system modeling, model testing, hypothesis generation, and experimentation that will expand and refine our understanding of these systems. mRNP systems biology, though still in its infancy, holds the potential to revolutionize our understanding of mRNP systems and their roles throughout biology.

Acknowledgments

References

Abruzzi, K.C., Lacadie, S., and Rosbash, M. 2004. Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. *EMBO J.* **23:** 2620–2631.

Aitchison, J.D. and Galitski, T. 2003. Inventories to insights. *J. Cell Biol.* **161:** 465–469.

Aloy, P., Bottcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., Gavin, A.C., Bork, P., Superti-Furga, G., Serrano, L., et al. 2004. Structure-based assembly of protein complexes in yeast. *Science* **303:** 2026–2029.

Antar, L.N. and Bassell, G.J. 2003. Sunrise at the synapse: The FMRP mRNP shaping the synaptic interface. *Neuron* **37:** 555–558.

Arava, Y., Wang, Y., Storey, J.D., Liu, C.L., Brown, P.O., and Herschlag, D. 2003. Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **100:** 3889–3894.

Audibert, A., Weil, D., and Dautry, F. 2002. In vivo kinetics of mRNA splicing and transport in mammalian cells. *Mol. Cell. Biol.* **22:** 6706–6718.

Bader, J.S., Chaudhuri, A., Rothberg, J.M., and Chant, J. 2004. Gaining confidence in high-throughput protein interaction networks. *Nat. Biotechnol.* **22:** 78–85.

Bar-Joseph, Z., Gerber, G.K., Lee, T.I., Rinaldi, N.J., Yoo, J.Y., Robert, F., Gordon, D.B., Fraenkel, E., Jaakkola, T.S., Young, R.A., et al. 2003. Computational discovery of gene modules and regulatory networks. *Nat. Biotechnol.* **21:** 1337–1342.

Baron-Benhamou, J., Fortes, P., Inada, T., Preiss, T., and Hentze, M.W. 2003. The interaction of the cap-binding complex (CBC) with eIF4G is dispensable for translation in yeast. *RNA* **9:** 654–662.

Barrass, J.D. and Beggs, J.D. 2003. Splicing goes global. *Trends Genet.* **19:** 295–298.

Bartel, D.P. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116:** 281–297.

Beach, D.L., Salmon, E.D., and Bloom, K. 1999. Localization and anchoring of mRNA in budding yeast. *Curr. Biol.* **9:** 569–578.

Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S., and Cohen, S.N. 2002. Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci.* **99:** 9697–9702.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. 1998. Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2:** 437–445.

Blanchette, M., Labourier, E., Green, R.E., Brenner, S.E., and Rio, D.C. 2004. Genome-wide analysis reveals an unexpected function for the *Drosophila* splicing factor U2AF(50) in the nuclear export of intronless mRNAs. *Mol. Cell* **14:** 775–786.

Brodsky, A.S. and Silver, P.A. 2000. Pre-mRNA processing factors are required for nuclear export. *RNA* **6:** 1737–1749.

Brown, V., Jin, P., Ceman, S., Darnell, J.C., O'Donnell, W.T., Tenenbaum, S.A., Jin, X., Feng, Y., Wilkinson, K.D., Keene, J.D., et al. 2001. Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* **107:** 477–487.

Buckanovich, R.J. and Darnell, R.B. 1997. The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol. Cell. Biol.* **17:** 3194–3201.

Calapez, A., Pereira, H.M., Calado, A., Braga, J., Rino, J., Carvalho, C., Tavanez, J.P., Wahle, E., Rosa, A.C., and Carmo-Fonseca, M. 2002. The intranuclear mobility of messenger RNA binding proteins is ATP dependent and temperature sensitive. *J. Cell Biol.* **159:** 795–805.

Cao, D. and Parker, R. 2001. Computational modeling of eukaryotic mRNA turnover. *RNA* **7:** 1192–1212.

———. 2003. Computational modeling and experimental analysis of nonsense-mediated decay in yeast. *Cell* **113:** 533–545.

Carrero, G., McDonald, D., Crawford, E., de Vries, G., and Hendzel, M.J. 2003. Using FRAP and mathematical modeling to determine the in vivo kinetics of nuclear proteins. *Methods* **29:** 14–28.

Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., and Silver, P.A. 2004. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* **117:** 427–439.

Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A.J., et al. 2004. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116:** 499–509.

Clark, T.A., Sugnet, C.W., and Ares Jr., M. 2002. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* **296:** 907–910.

Covert, M.W., Knight, E.M., Reed, J.L., Herrgard, M.J., and Palsson, B.O. 2004. Integrating high-throughput and computational data elucidates bacterial networks. *Nature* **429:** 92–96.

Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warren, S.T., and Darnell, R.B. 2001. Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* **107:** 489–499.

DeRisi, J.L., Iyer, V.R., and Brown, P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278:** 680–686.

Diehn, M., Eisen, M.B., Botstein, D., and Brown, P.O. 2000. Large-scale identification of secreted and membrane-associated gene products using DNA microarrays. *Nat. Genet.* **25:** 58–62.

Dominski, Z. and Marzluff, W.F. 1999. Formation of the 3′ end of histone mRNA. *Gene* **239:** 1–14.

Eisenstein, R.S. and Ross, K.L. 2003. Novel roles for iron regulatory proteins in the adaptive response to iron deficiency. *J. Nutr.* (Suppl. 1) **133:** 1510S–1516S.

Fan, J., Yang, X., Wang, W., Wood III, W.H., Becker, K.G., and Gorospe, M. 2002. Global analysis of stress-regulated mRNA turnover by using cDNA arrays. *Proc. Natl. Acad. Sci.* **99:** 10611–10616.

Forrest, K.M. and Gavis, E.R. 2003. Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in *Drosophila*. *Curr. Biol.* **13:** 1159–1168.

Fusco, D., Accornero, N., Lavoie, B., Shenoy, S.M., Blanchard, J.M., Singer, R.H., and Bertrand, E. 2003. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr. Biol.* **13:** 161–167.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415:** 141–147.

Ge, H., Walhout, A.J., and Vidal, M. 2003. Integrating 'omic' information: A bridge between genomics and systems biology. *Trends Genet.* **19:** 551–560.

Gera, J.F., Mellinghoff, I.K., Shi, Y., Rettig, M.B., Tran, C., Hsu, J.H., Sawyers, C.L., and Lichtenstein, A.K. 2004. AKT activity determines sensitivity to mammalian target of rapamycin (mTOR) inhibitors by regulating cyclin D1 and c-myc expression. *J. Biol. Chem.* **279:** 2737–2746.

Gerber, A.P., Herschlag, D., and Brown, P.O. 2004. Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol.* **2:** (3) E79.

Gilbert, C., Kristjuhan, A., Winkler, G.S., and Svejstrup, J.Q. 2004. Elongator interactions with nascent mRNA revealed by RNA immunoprecipitation. *Mol. Cell* **14:** 457–464.

Gorlich, D., Seewald, M.J., and Ribbeck, K. 2003. Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *EMBO J.* **22:** 1088–1100.

Gott, J.M. and Emeson, R.B. 2000. Functions and mechanisms of RNA editing. *Annu. Rev. Genet.* **34:** 499–531.

Grigull, J., Mnaimneh, S., Pootoolal, J., Robinson, M.D., and Hughes, T.R. 2004. Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Mol. Cell. Biol.* **24:** 5534–5547.

Grolleau, A., Bowman, J., Pradet-Balade, B., Puravs, E., Hanash, S., Garcia-Sanz, J.A., and Beretta, L. 2002. Global and specific translational control by rapamycin in T cells uncovered by microarrays and proteomics. *J. Biol. Chem.* **277:** 22175–22184.

Guelzim, N., Bottani, S., Bourgine, P., and Kepes, F. 2002. Topological and causal structure of the yeast transcriptional regulatory network. *Nat. Genet.* **31:** 60–63.

Gutierrez, R.A., Ewing, R.M., Cherry, J.M., and Green, P.J. 2002. Identification of unstable transcripts in *Arabidopsis* by cDNA microarray analysis: Rapid decay is associated with a group of touch- and specific clock-controlled genes. *Proc. Natl. Acad. Sci.* **99:** 11513–11518.

Hajarnavis, A., Korf, I. and Durbin, R. 2004. A probabilistic model of 3′ end formation in *Caenorhabditis elegans*. *Nucleic Acids Res.* **2432:** 3392–3399.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., et al. 2004. Transcriptional regulatory code of a eukaryotic genome. *Nature* **431:** 99–104.

He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., and Jacobson, A. 2003. Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5′ to 3′ mRNA decay pathways in yeast. *Mol. Cell* **12:** 1439–1452.

Hegde, P.S., White, I.R., and Debouck, C. 2003. Interplay of transcriptomics and proteomics. *Curr. Opin. Biotechnol.* **14:** 647–651.

Herold, A., Teixeira, L., and Izaurralde, E. 2003. Genome-wide analysis of nuclear mRNA export pathways in *Drosophila*. *EMBO J.* **22:** 2472–2483.

Hieronymus, H. and Silver, P.A. 2003. Genome-wide analysis of RNA–protein interactions illustrates specificity of the mRNA export machinery. *Nat. Genet.* **33:** 155–161.

Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95:** 717–728.

Huang, Y.S. and Richter, J.D. 2004. Regulation of local mRNA translation. *Curr. Opin. Cell Biol.* **16:** 308–313.

Huertas, P. and Aguilera, A. 2003. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* **12:** 711–721.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. 2003. Global analysis of protein localization in budding yeast. *Nature* **425:** 686–691.

Inada, M. and Guthrie, C. 2004. Identification of Lhp1p-associated RNAs by microarray analysis in *Saccharomyces cerevisiae* reveals association with coding and noncoding RNAs. *Proc. Natl. Acad. Sci.* **101:** 434–439.

Issel-Tarver, L., Christie, K.R., Dolinski, K., Andrada, R., Balakrishnan, R., Ball, C.A., Binkley, G., Dong, S., Dwight, S.S., Fisk, D.G., et al. 2002. *Saccharomyces* genome database. *Methods Enzymol.* **350:** 329–346.

Iyer, V.R., Horak, C.E., Scafe, C.S., Botstein, D., Snyder, M., and Brown, P.O. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409:** 533–538.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., et al. 2004. From silencing to gene expression: Real-time analysis in single cells. *Cell* **116:** 683–698.

Johannes, G., Carter, M.S., Eisen, M.B., Brown, P.O., and Sarnow, P. 1999. Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. *Proc. Natl. Acad. Sci.* **96:** 13118–13123.

Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R., and Shoemaker, D.D. 2003. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* **302:** 2141–2144.

Kampa, D., Cheng, J., Kapranov, P., Yamanaka, M., Brubaker, S., Cawley, S., Drenkow, J., Piccolboni, A., Bekiranov, S., Helt, G., et al. 2004. Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Res.* **14:** 331–342.

Kapranov, P., Cawley, S.E., Drenkow, J., Bekiranov, S., Strausberg, R.L., Fodor, S.P., and Gingeras, T.R. 2002. Large-scale transcriptional activity in chromosomes 21 and 22. *Science* **296:** 916–919.

Kash, J.C., Cunningham, D.M., Smit, M.W., Park, Y., Fritz, D., Wilusz, J., and Katze, M.G. 2002. Selective translation of eukaryotic mRNAs: Functional molecular analysis of GRSF-

1, a positive regulator of influenza virus protein synthesis. *J. Virol.* **76:** 10417–10426.

Keene, J.D. 2001. Ribonucleoprotein infrastructure regulating the flow of genetic information between the genome and the proteome. *Proc. Natl. Acad. Sci.* **98:** 7018–7024.

Keene, J.D. and Tenenbaum, S.A. 2002. Eukaryotic mRNPs may represent posttranscriptional operons. *Mol. Cell* **9:** 1161–1167.

Kitano, H. 2002. Systems biology: A brief overview. *Science* **295:** 1662–1664.

Kotovic, K.M., Lockshon, D., Boric, L., and Neugebauer, K.M. 2003. Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. *Mol. Cell. Biol.* **23:** 5768–5779.

Krawczak, M., Reiss, J., and Cooper, D.N. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: Causes and consequences. *Hum. Genet.* **90:** 41–54.

Kuersten, S. and Goodwin, E.B. 2003. The power of the 3′ UTR: Translational control and development. *Nat. Rev. Genet.* **4:** 626–637.

Kuhn, K.M., DeRisi, J.L., Brown, P.O., and Sarnow, P. 2001. Global and specific translational regulation in the genomic response of *Saccharomyces cerevisiae* to a rapid transfer from a fermentable to a nonfermentable carbon source. *Mol. Cell. Biol.* **21:** 916–927.

Kumar, A., Agarwal, S., Heyman, J.A., Matson, S., Heidtman, M., Piccirillo, S., Umansky, L., Drawid, A., Jansen, R., Liu, Y., et al. 2002. Subcellular localization of the yeast proteome. *Genes & Dev.* **16:** 707–719.

Kustanovich, T. and Rabin, Y. 2004. Metastable network model of protein transport through nuclear pores. *Biophys. J.* **86:** 2008–2016.

Lee, T.I., Causton, H.C., Holstege, F.C., Shen, W.C., Hannett, N., Jennings, E.G., Winston, F., Green, M.R., and Young, R.A. 2000. Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* **405:** 701–704.

Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., et al. 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298:** 799–804.

Lei, E.P. and Silver, P.A. 2002. Protein and RNA export from the nucleus. *Dev. Cell* **2:** 261–272.

Lelivelt, M.J. and Culbertson, M.R. 1999. Yeast Upf proteins required for RNA surveillance affect global expression of the yeast transcriptome. *Mol. Cell. Biol.* **19:** 6710–6719.

Lerner, R.S., Seiser, R.M., Zheng, T., Lager, P.J., Reedy, M.C., Keene, J.D., and Nicchitta, C.V. 2003. Partitioning and translation of mRNAs encoding soluble proteins on membrane-bound ribosomes. *RNA* **9:** 1123–1137.

Levesque, M.P. and Benfey, P.N. 2004. Systems biology. *Curr. Biol.* **14:** R179–R180.

Levsky, J.M. and Singer, R.H. 2003. Gene expression and the myth of the average cell. *Trends Cell. Biol.* **13:** 4–6.

Li, A.M., Watson, A., and Fridovich-Keil, J.L. 2003. Scp160p associates with specific mRNAs in yeast. *Nucleic Acids Res.* **31:** 1830–1837.

Lieb, J.D., Liu, X., Botstein, D., and Brown, P.O. 2001. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein–DNA association. *Nat. Genet.* **28:** 327–334.

Liu, C.G., Calin, G.A., Meloon, B., Gamliel, N., Sevignani, C., Ferracin, M., Dumitru, C.D., Shimizu, M., Zupo, S., Dono, M., et al. 2004. An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc. Natl. Acad. Sci.* **101:** 9740–9744.

Lopez de Silanes, I., Zhan, M., Lal, A., Yang, X., and Gorospe, M. 2004. Identification of a target RNA motif for RNA-binding protein HuR. *Proc. Natl. Acad. Sci.* **101:** 2987–2992.

Maquat, L.E. 2004. Nonsense-mediated mRNA decay: Splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.* **5:** 89–99.

McCarroll, S.A., Murphy, C.T., Zou, S., Pletcher, S.D., Chin, C.S., Jan, Y.N., Kenyon, C., Bargmann, C.I., and Li, H. 2004. Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nat. Genet.* **36:** 197–204.

Mendez, R. and Richter, J.D. 2001. Translational control by CPEB: A means to the end. *Nat. Rev. Mol. Cell Biol.* **2:** 521–529.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. 2002. Network motifs: Simple building blocks of complex networks. *Science* **298:** 824–827.

Miyashiro, K.Y., Beckel-Mitchener, A., Purk, T.P., Becker, K.G., Barret, T., Liu, L., Carbonetto, S., Weiler, I.J., Greenough, W.T., and Eberwine, J. 2003. RNA cargoes associating with FMRP reveal deficits in cellular functioning in Fmr1 null mice. *Neuron* **37:** 417–431.

Mootha, V.K., Bunkenborg, J., Olsen, J.V., Hjerrild, M., Wisniewski, J.R., Stahl, E., Bolouri, M.S., Ray, H.N., Sihag, S., Kamal, M., et al. 2003. Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell* **115:** 629–640.

Neugebauer, K.M. 2002. On the importance of being co-transcriptional. *J. Cell. Sci.* **115:** 3865–3871.

Nicchitta, C.V. 2002. A platform for compartmentalized protein synthesis: Protein translation and translocation in the ER. *Curr. Opin. Cell. Biol.* **14:** 412–416.

Parker, R. and Song, H. 2004. The enzymes and control of eukaryotic mRNA turnover. *Nat. Struct. Mol. Biol.* **11:** 121–127.

Penalva, L.O., Tenenbaum, S.A., and Keene, J.D. 2004. Gene expression analysis of messenger RNP complexes. *Methods Mol. Biol.* **257:** 125–134.

Preiss, T., Baron-Benhamou, J., Ansorge, W., and Hentze, M.W. 2003. Homodirectional changes in transcriptome composition and mRNA translation induced by rapamycin and heat shock. *Nat. Struct. Biol.* **10:** 1039–1047.

Qin, X. and Sarnow, P. 2004. Preferential translation of internal ribosome entry site-containing mRNAs during the mitotic cycle in mammalian cells. *J. Biol. Chem.* **279:** 13721–13728.

Raghavan, A., Ogilvie, R.L., Reilly, C., Abelson, M.L., Raghavan, S., Vasdewani, J., Krathwohl, M., and Bohjanen, P.R. 2002. Genome-wide analysis of mRNA decay in resting and activated primary human T lymphocytes. *Nucleic Acids Res.* **30:** 5529–5538.

Rajasekhar, V.K., Viale, A., Socci, N.D., Wiedmann, M., Hu, X., and Holland, E.C. 2003. Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes. *Mol. Cell* **12:** 889–901.

Reed, R. 2003. Coupling transcription, splicing and mRNA export. *Curr. Opin. Cell. Biol.* **15:** 326–331.

Rehwinkel, J., Herold, A., Gari, K., Kocher, T., Rode, M., Ciccarelli, F.L., Wilm, M., and Izaurralde, E. 2004. Genome-wide analysis of mRNAs regulated by the THO complex in *Drosophila melanogaster*. *Nat. Struct. Mol. Biol.* **11:** 558–566.

Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., et al. 2000. Genome-wide location and function of DNA binding proteins. *Science* **290:** 2306–2309.

Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R.A., and Dynlacht, B.D. 2002. E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes & Dev.* **16:** 245–256.

Ribbeck, K. and Gorlich, D. 2001. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* **20:** 1320–1330.

Rodriguez-Navarro, S., Fischer, T., Luo, M.J., Antunez, O., Brettschneider, S., Lechner, J., Perez-Ortin, J.E., Reed, R., and Hurt, E. 2004. Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* **116:** 75–86.

Rook, M.S., Lu, M., and Kosik, K.S. 2000. CaMKIIα 3′ untranslated region-directed mRNA translocation in living neurons: Visualization by GFP linkage. *J. Neurosci.* **20:** 6385–6393.

Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B.T. 2000. The yeast nuclear pore complex: Composition, architecture, and transport mechanism. *J. Cell. Biol.* **148:** 635–651.

Saavedra, C., Tung, K.S., Amberg, D.C., Hopper, A.K., and Cole, C.N. 1996. Regulation of mRNA export in response to stress in *Saccharomyces cerevisiae*. *Genes & Dev.* **10:** 1608–1620.

Schmittgen, T.D., Jiang, J., Liu, Q., and Yang, L. 2004. A high-throughput method to monitor the expression of microRNA precursors. *Nucleic Acids Res.* **32:** (4) e43.

Selinger, D.W., Saxena, R.M., Cheung, K.J., Church, G.M., and Rosenow, C. 2003. Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.* **13:** 216–223.

Shav-Tal, Y., Darzacq, X., Shenoy, S.M., Fusco, D., Janicki, S.M., Spector, D.L., and Singer, R.H. 2004. Dynamics of single mRNPs in nuclei of living cells. *Science* **304:** 1797—1800.

Shen, W.C., Bhaumik, S.R., Causton, H.C., Simon, I., Zhu, X., Jennings, E.G., Wang, T.H., Young, R.A., and Green, M.R. 2003. Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. *EMBO J.* **22:** 3395–3402.

Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U. 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* **31:** 64–68.

Shepard, K.A., Gerber, A.P., Jambhekar, A., Takizawa, P.A., Brown, P.O., Herschlag, D., DeRisi, J.L., and Vale, R.D. 2003. Widespread cytoplasmic mRNA transport in yeast: Identification of 22 bud-localized transcripts using DNA microarray analysis. *Proc. Natl. Acad. Sci.* **100:** 11429–11434.

Shoemaker, D.D., Schadt, E.E., Armour, C.D., He, Y.D., Garrett-Engele, P., McDonagh, P.D., Loerch, P.M., Leonardson, A., Lum, P.Y., Cavet, G., et al. 2001. Experimental annotation of the human genome using microarray technology. *Nature* **409:** 922–927.

Siebrasse, J.P. and Peters, R. 2002. Rapid translocation of NTF2 through the nuclear pore of isolated nuclei and nuclear envelopes. *EMBO Rep.* **3:** 887–892.

Singer, R.H. 2003. RNA localization: Visualization in real-time. *Curr. Biol.* **13:** R673–R675.

Smith, A.E., Slepchenko, B.M., Schaff, J.C., Loew, L.M., and Macara, I.G. 2002. Systems analysis of Ran transport. *Science* **295:** 488–491.

Stutz, F. and Izaurralde, E. 2003. The interplay of nuclear mRNP assembly, mRNA surveillance and export. *Trends Cell Biol.* **13:** 319–327.

Takizawa, P.A. and Vale, R.D. 2000. The myosin motor, Myo4p, binds Ash1 mRNA via the adapter protein, She3p. *Proc. Natl. Acad. Sci.* **97:** 5273–5278.

Takizawa, P.A., DeRisi, J.L., Wilhelm, J.E., and Vale, R.D. 2000.

Plasma membrane compartmentalization in yeast by messenger RNA transport and a septin diffusion barrier. *Science* **290:** 341–344.

Tekotte, H. and Davis, I. 2002. Intracellular mRNA localization: Motors move messages. *Trends Genet.* **18:** 636–642.

Tenenbaum, S.A., Carson, C.C., Lager, P.J., and Keene, J.D. 2000. Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc. Natl. Acad. Sci.* **97:** 14085–14090.

Thieffry, D., Huerta, A.M., Perez-Rueda, E., and Collado-Vides, J. 1998. From specific gene regulation to genomic networks: A global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays* **20:** 433–440.

Thomson, A.M., Rogers, J.T., and Leedman, P.J. 1999. Iron-regulatory proteins, iron-responsive elements and ferritin mRNA translation. *Int. J. Biochem. Cell Biol.* **31:** 1139–1152.

Toyoda, N., Nagai, S., Terashima, Y., Motomura, K., Haino, M., Hashimoto, S., Takizawa, H., and Matsushima, K. 2003. Analysis of mRNA with microsomal fractionation using a SAGE-based DNA microarray system facilitates identification of the genes encoding secretory proteins. *Genome Res.* **13:** 1728–1736.

Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., et al. 2000. A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* **403:** 623–627.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. 2003. CLIP identifies Nova-regulated RNA networks in the brain. *Science* **302:** 1212–1215.

Van de Bor, V. and Davis, I. 2004. mRNA localisation gets more complex. *Curr. Opin. Cell. Biol.* **16:** 300–307.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett Jr., D.E., Hieter, P., Vogelstein, B., and Kinzler, K.W. 1997. Characterization of the yeast transcriptome. *Cell* **88:** 243–251.

Vinciguerra, P. and Stutz, F. 2004. mRNA export: An assembly line from genes to nuclear pores. *Curr. Opin. Cell Biol.* **16:** 285–292.

von der Haar, T., Gross, J.D., Wagner, G., and McCarthy, J.E. 2004. The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat. Struct. Mol. Biol.* **11:** 503–511.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., and Bork, P. 2002. Comparative assessment of large-scale data sets of protein–protein interactions. *Nature* **417:** 399–403.

Waggoner, S.A. and Liebhaber, S.A. 2003. Identification of mRNAs associated with αCP2-containing RNP complexes. *Mol. Cell. Biol.* **23:** 7055–7067.

Wang, Y., Liu, C.L., Storey, J.D., Tibshirani, R.J., Herschlag, D., and Brown, P.O. 2002. Precision and functional specificity in mRNA decay. *Proc. Natl. Acad. Sci.* **99:** 5860–5865.

Wang, H., Hubbell, E., Hu, J.S., Mei, G., Cline, M., Lu, G., Clark, T., Siani-Rose, M.A., Ares, M., Kulp, D.C., et al. 2003. Gene structure-based splice variant deconvolution using a microarray platform. *Bioinformatics* (Suppl. 1) **19:** i315–i322.

Wei, G.H., Liu, D.P., and Liang, C.C. 2004. Charting gene regulatory networks: Strategies, challenges and perspectives. *Biochem. J. (Pt. 1)* **381:** 1–12.

Wodicka, L., Dong, H., Mittmann, M., Ho, M.H., and Lockhart, D.J. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **15:** 1359–1367.

Yang, E., van Nimwegen, E., Zavolan, M., Rajewsky, N., Schroeder, M., Magnasco, M., and Darnell Jr., J.E. 2003. Decay rates of human mRNAs: Correlation with functional

characteristics and sequence attributes. *Genome Res.* **13:** 1863–1872.

Yeakley, J.M., Fan, J.B., Doucet, D., Luo, L., Wickham, E., Ye, Z., Chee, M.S., and Fu, X.D. 2002. Profiling alternative splicing on fiber-optic arrays. *Nat. Biotechnol.* **20:** 353–358.

Yu, M.C., Bachand, F., McBride, A.E., Komili, S., Casolari, J.M., and Silver, P.A. 2004. Arginine methyltransferase affects interactions and recruitment of mRNA processing and export factor. *Genes* & *Dev.* **18:** 2024–2035.

Zeitlinger, J., Simon, I., Harbison, C.T., Hannett, N.M., Volkert, T.L., Fink, G.R., and Young, R.A. 2003. Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling. *Cell* **113:** 395–404.

Zhang, Y.Q., Bailey, A.M., Matthies, H.J., Renden, R.B., Smith, M.A., Speese, S.D., Rubin, G.M., and Broadie, K. 2001. *Drosophila* fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* **107:** 591–603.

Zhou, Z., Licklider, L.J., Gygi, S.P., and Reed, R. 2002. Comprehensive proteomic analysis of the human spliceosome. *Nature* **419:** 182–185.

Zhu, J., Shendure, J., Mitra, R.D., and Church, G.M. 2003. Single molecule profiling of alternative pre-mRNA splicing. *Science* **301:** 836–838.