directed towards improvement of agricultural qualities, perhaps these goals can be combined to increase tolerance to temperature extremes, salinity, flooding, or insect pests in plants capable of pollutant detoxification or, more importantly for value-enhancement – transfer of phyto-remediative traits to elite plant cultivars having the highest biomass or agricultural productivity. Obviously, concerns about contaminant uptake and accumulation will limit the use of phyto-crops for food or human contact products, so every effort must be made to identify parent compound fate and toxicity for these applications. However, as observed with the development of chemo-preventative enriched, Se-hyperaccumulating plants, opportunities exist to combine pollutant decontamination capabilities with beneficial human and ecological health qualities in engineered plants.

References

1 Bizily, S.P. *et al*. (1999) Phytoremediation of methylmercury pollution: merB expression in *Arabidopsis thaliana* confers resistance to organomercurials. *Proc. Natl. Acad. Sci. U. S. A.* 96, 6808–6813
2 Doty, S.L. *et al*. (2000) Enhanced metabolism of halogenated hydrocarbons in transgenic plants containing mammalian cytochrome P450 2E1. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6287–6291
3 French, C.E. *et al*. (1999) Biodegradation of explosives by transgenic plants expressing pentaerythritol tetranitrate reductase. *Nat. Biotechnol.* 17, 491–494
4 Ellis, D.R. *et al*. (2004) Production of Se-methylselenocysteine in transgenic plants expressing selenocysteine methyltransferase. *BMC Plant Biol.* 4, 1
5 Baker, A.J.M. and Brooks, R.R. (1989) Terrestrial higher plants which hyper-accumulate metallic elements: a review of their distribution, ecology and phytochemistry. *Biorecovery* 1, 81–126
6 Baker, A.J.M. *et al*. (1994) The possibility of *in situ* heavy metal decontamination of polluted soils using crops of metal-accumulating plants. *Res. Conserv. Recyc.* 11, 41–49
7 Li, Y.M. *et al*. (2003) Development of the technology for commercial phytoextraction of nickel: economic and technical considerations. *Plant Soil* 249, 107–115
8 LaCoste, C. *et al*. (2001) Uptake of thallium by vegetables: Its significance for human health, phytoremediation, and phytomining. *J. Plant Nutr.* 24, 1205–1215
9 Guerinot, M.L. and Salt, D.E. (2001) Fortified foods and phytoremediation. Two sides of the same coin. *Plant Physiol.* 125, 164–167
10 Lahner, B. *et al*. (2003) Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. *Nat. Biotechnol.* 21, 1215–1221
11 Maser, P. *et al*. (2001) Phylogenetic relationships within cation transporter families of *Arabidopsis*. *Plant Physiol.* 126, 1646–1667
12 de Souza, M.P. *et al*. (1998) Rate-limiting steps in selenium assimilation and volatilization by Indian mustard. *Plant Physiol.* 117, 1487–1494
13 Singh, M. and Singh, N. (1977) Effect of sulfur and selenium on sulfur-containing amino-acids and quality of oil in Raya (*Brassica juncea* Coss) in normal and sodic soil. *Ind. J. Plant Physiol.* 20, 56–62
14 Garifullina, G.F. *et al*. (2003) Expression of a mouse selenocysteine lyase in *Brassica juncea* chloroplasts affects selenium tolerance and accumulation. *Physiol. Plant.* 118, 538–544
15 Berken, A. *et al*. (2002) Genetic engineering of plants to enhance selenium phytoremediation. *Crit. Rev. Plant Sci.* 21, 567–582
16 Banuelos, G.S. *et al*. (2002) Biotransfer possibilities of selenium form plants used in phytoremediation. *Int. J. Phytoremediation* 4, 315–329
17 Brown, T.A. and Shrift, A. (1981) Exclusion of selenium from proteins in selenium-tolerant *Astragalus* species. *Plant Physiol.* 67, 1951–1953
18 Neuhierl, B. and Bock, A. (1996) On the mechanism of selenium tolerance in selenium-accumulating plants: purification and characterization of a specific selenocysteine methyltransferase from cultured cells of *Astragalus bisulcatus*. *Eur. J. Biochem.* 239, 235–238
19 Combs, G.F. and Gray, W.P. (1998) Chemopreventative agents. *Selenium. Pharmacol. Ther.* 79, 179–192
20 Medina, D. *et al*. (2001) Se-methylselenocysteine: a new compound for chemoprevention of breast cancer. *Nutr. Cancer* 40, 12–17
21 Orser, C.S. *et al*. (1999) Brassica plants to provide enhanced human mineral nutrition: selenium phytoenrichment and metabolic transformation. *J. Med. Food* 1, 253–261
22 Pickering, I.J. *et al*. (2003) Chemical form and distribution of selenium and sulfur in the selenium hyperaccumulator *Astragalus bisulcatus*. *Plant Physiol.* 131, 1–8
23 Pulford, I.D. *et al*. (2002) Heavy metal uptake by willow clones from sewage sludge-treated soil: the potential for phytoremediation. *Int. J. Phytoremediation* 4, 59–72
24 Bungay, H. (2004) Confessions of a bioenergy advocate. *Trends Biotechnol.* 22, 67–71
25 Pilon-Smits, E. and Pilon, M. (2002) Phytoremediation of metals using transgenic plants. *Crit. Rev. Plant Sci.* 21, 439–456
26 Rugh, C.L. (2001) Mercury detoxification with transgenic plants and other biotechnological br eakthroughs for phytoremediation. *In Vitro Cell. Dev. Biol. Plant* 37, 321–325

0167-7799/$ - see front matter
doi:10.1016/j.tibtech.2004.08.003

Exploring the post-transcriptional RNA world with DNA microarrays

Vishwanath R. Iyer

Center for Systems and Synthetic Biology, and Institute for Cellular and Molecular Biology, University of Texas at Austin, 1 University Station A4800, Austin, TX 78712-0159, USA

Genomic approaches are valuable for understanding the complex layer of gene regulation that involves the control of RNA processing, localization and stability. Recent work provides a prime example of the power of unbiased microarray-based methods to discover unexpected functions for proteins in the RNA world. The challenges ahead relate to extending such approaches to larger genomes and to integrating this type of information with that generated by standard expression profiling.

Corresponding author: Vishwanath R. Iyer (vishy@mail.utexas.edu).

Although gene expression is often regulated by transcription factors at the level of transcription initiation, the subsequent steps of RNA processing, turnover, subcellular localization and entry into the translation machinery strongly influence the extent of protein translation and the function of encoded proteins. Such post-transcriptional steps therefore have marked effects on the expression and function of genes in processes as diverse as cytokinesis, early embryonic development and neuronal function [1].

When trying to infer the global phenotypes of cells from large-scale mRNA expression profiling data, it is important to be aware of this intervening layer of gene regulation. Most post-transcriptional events are mediated by the association of RNAs with specific proteins or macromolecular protein complexes. Comprehensive determination of the RNA targets of RNA-binding proteins is therefore likely to be important in deciphering the complex events at this level of gene regulation.

The La protein is a conserved eukaryotic protein that is thought to be important in the realm of post-transcriptional regulation and, as we discuss here, a recent study by Inada and Guthrie [2] provides a prime example of the use of a genomic approach to elucidate the targets and potential function of such an RNA-binding protein.

Ribonomics with cDNA microarrays

cDNA microarrays have been heavily used for quantitative mRNA profiling, but there are increasing examples of the varied use of cDNA microarrays to follow the fates of mRNAs in the cell after they are made, rather than to measure only their steady-state levels. One objective is to determine the binding targets of proteins that interact with RNAs at any point during the lifetime of the RNA. Protein–RNA interactions represent one of the most abundant categories of molecular interactions in cells, and the total number of RNA-interacting proteins rivals that of other categories such as transcription factors and signaling molecules, even if one excludes the hundreds of proteins that are integral components of the spliceosome and ribosome [3,4].

Proteins can interact with RNA from the time that they are transcribed, and they affect transcriptional efficiency, capping, 3′-end processing, splicing, nuclear export, subcellular localization, translation and turnover of RNA [5]. The sheer diversity, cell- and tissue-specificity, and conservation of RNA-binding proteins has led to the notion that primary transcripts, rather than advancing smoothly through each of the subsequent RNA processing steps, participate in a complex network of regulatory processes at the post-transcriptional level [6]. Clearly, identifying the RNA targets of specific RNA-binding proteins is likely to be at least as informative and important with regard to understanding global gene regulation as is measuring changes in steady-state levels of RNAs in response to cellular signals.

The genomic strategy for determining the RNA partners of RNA-binding proteins involves immunoprecipitation of the protein of interest along with its associated RNA, fluorescent labeling of the enriched RNA (as cDNA), and finally microarray hybridization in conjunction with an appropriate reference probe (Figure 1). This approach was first used independently in the laboratories of Ron Vale [7] and Jack Keene [8] and was termed 'ribonomics' by the latter. Variations of this method have been subsequently used to identify the targets of more than a dozen RNA-binding proteins (see Gerber *et al.* [9] and references therein).

The function of La in the cell

A prime example of the power of ribonomics has been provided recently by Maki Inada and Christine Guthrie [2] in their analysis of the function of the La protein in yeast. La is a ubiquitous, nuclear RNA-binding protein that is conserved among eukaryotes. It is known to associate with the 3′-UUU-OH containing ends of the primary transcripts of RNA polymerase III, including all tRNAs and other small RNAs, and is thought to be involved in the processing and assembly of many of its target RNAs into ribonucleoprotein complexes. The interaction of La with RNAs is mediated by its La motif and its 'RNA recognition' motif (RRM) [10]. La also associates with the 5′ ends of some mRNAs that contain a 5′-terminal oligopyrimidine end or internal ribosome entry sites (IRES), and can potentially influence their translation [11,12]. This diverse set of properties indicates that La has a complex role, and its conservation suggests that it is important in eukaryotic cells; nevertheless, the yeast La gene (*LHP1*) is not essential for viability under normal growth conditions.

In Inada and Guthrie's [2] ribonomic approach to identify the direct binding targets of yeast La protein, a Myc-tagged Lhp1 protein was immunoprecipitated with its associated

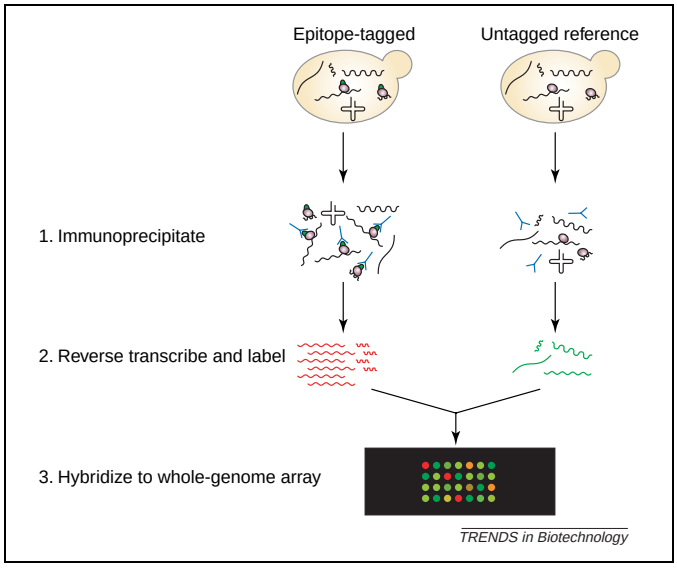

Figure 1. Strategy for identifying the targets of RNA-binding proteins. An approach used in yeast cells is shown here as an example. Proteins of interest, along with their associated RNAs, are immunoprecipitated from cellular extracts with a specific antibody (left). The reference or control sample can consist of either a parallel sample immunoprecipitated from an untagged strain (right) or, alternatively, total RNA isolated from the extract but not subjected to immunoprecipitation (not shown). A variation of this method involves crosslinking the RNA-binding protein to its target RNAs before the immunoprecipitation step [13]. The immunoprecipitated and reference samples are reverse transcribed to cDNA, differentially labeled with fluorescent dyes, and hybridized to a whole-genome microarray representing all potential targets. The fluorescent ratio of each spot on the microarray is a measure of the relative association of the transcript represented by that spot with the protein of interest.

RNAs and an untagged strain was used as the hybridization reference sample. What sets this work apart is the use of a truly whole-genome microarray containing all yeast open reading frames, annotated small RNAs and all intergenic regions to analyze the RNA precipitated with Lhp1. Within the technical limits of the immunoprecipitation procedure, such an approach should thus identify every possible target of Lhp1 in yeast under the conditions of the experiment and provide important clues about its function.

A major class of La targets identified in their work comprised several RNA polymerase III targets, including several tRNAs, U6 small nuclear RNA and *SCR1*, among others. A second class of targets was the small nucleolar RNAs (snoRNAs). Inada and Guthrie identified about 20 annotated snoRNAs among the most highly enriched targets, thereby considerably expanding the repertoire of known La targets. The use of a whole-genome array containing all intergenic regions enabled them to identify at least three novel snoRNA targets before they were annotated as such, raising the possibility that some of the enrichment signals that they observed from other intergenic regions also represented novel unannotated transcripts. Obviously, these targets would have remained undetected had they used a microarray consisting of only previously annotated genes.

A third class of La targets comprised mRNAs, which was intriguing because yeast was not thought to contain mRNAs with the previously known determinants of La binding to mRNAs: namely, 5′-terminal oligopyrimidine (TOP) elements or internal ribosome entry sites. A notable target in this class was the *HAC1* mRNA, which is spliced by a tRNA ligase to generate a functional transcription factor only during induction of the unfolded protein response. Lhp1 did not affect *HAC1* splicing or stability, but it did seem to have an effect on Hac1 protein levels. Furthermore, under conditions in which the unfolded protein response was induced, *LHP1* was indeed found to be essential. This comprehensive ribonomic survey thus suggested a possible biological role for Lhp1 functions under a specific physiological condition.

The other mRNA targets of Lhp1 were enriched for those encoding ribosomal protein genes. The few TOP-containing mRNA targets of La previously defined in human cells also encode ribosomal and other proteins involved in translation. Thus, La might affect the overall process of translation efficiency at many different levels, both directly and indirectly, by targeting ribosomal protein mRNAs, 5S rRNA and tRNAs, as well as snoRNAs that are used to process other RNAs involved in translation. This effect on translation might well be the dominant theme with regard to La function in eukaryotic cells. As Inada and Guthrie point out [2], La itself might be regulated by growth conditions and nutrient signaling in yeast. The full spectrum of its functions in different cells and under different physiological conditions remains, however, an open issue and will need further genome-scale exploration in higher eukaryotic cells.

Future research

The ribonomics approach will undoubtedly play a substantial role in future investigations of La function. It remains to be seen whether La associates with a similarly broad range of RNA targets in higher eukaryotes, and whether the range of its targets changes depending on the physiological state of the cell.

More generally, the important challenge ahead for ribonomics will be to integrate and to reconcile the data on the binding targets of hundreds of RNA-binding proteins with the global views of gene expression derived from only mRNA expression profiling. For example, if all ribosomal protein genes are tightly co-regulated at the level of transcription, but only a subset of them are targeted by a given RNA-binding protein such as La, then what does this imply about the process of ribosome assembly and translation?

Recent work using the ribonomics approach to identify the global targets of the Puf RNA-binding proteins in yeast [9] has revealed a remarkable functional coherence among the distinct sets of RNA targets for each RNA-binding protein; however, these sets of targets do not correspond in any obvious way to clusters of genes co-regulated at the level of transcription. What is the relative importance of effects at the level of transcription versus those at the post-transcriptional steps involving many RNA-binding proteins? The process of addressing such questions will surely yield more accurate views of the genome-wide control of gene expression.

References

1 Keene, J.D. (2003) Posttranscriptional generation of macromolecular complexes. *Mol. Cell* 12, 1347–1349
2 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. U. S. A.* 101, 434–439
3 Hall, K.B. (2002) RNA–protein interactions. *Curr. Opin. Struct. Biol.* 12, 283–288
4 Anantharaman, V. *et al*. (2002) Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res.* 30, 1427–1464
5 Dreyfuss, G. *et al*. (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* 3, 195–205
6 Maniatis, T. and Reed, R. (2002) An extensive network of coupling among gene expression machines. *Nature* 416, 499–506
7 Takizawa, P.A. *et al*. (2000) Plasma membrane compartmentalization in yeast by messenger RNA transport and a septin diffusion barrier. *Science* 290, 341–344
8 Tenenbaum, S.A. *et al*. (2000) Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc. Natl. Acad. Sci. U. S. A.* 97, 14085–14090
9 Gerber, A.P. *et al*. (2004) Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol.* 2, E79
10 Alfano, C. *et al*. (2004) Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nat. Struct. Mol. Biol.* 11, 323–329
11 Wolin, S.L. and Cedervall, T. (2002) The La protein. *Annu. Rev. Biochem.* 71, 375–403
12 Intine, R.V. *et al*. (2003) Differential phosphorylation and subcellular localization of La RNPs associated with precursor tRNAs and translation-related mRNAs. *Mol. Cell* 12, 1301–1307
13 Ule, J. *et al*. (2003) CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302, 1212–1215

0167-7799/$ - see front matter
doi:10.1016/j.tibtech.2004.08.006