

Computational prediction of RNA structural motifs involved in posttranscriptional regulatory processes

Michal Rabani*, Michael Kertesz*, and Eran Segal*[†]

*Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel

Edited by Sean R. Eddy, Howard Hughes Medical Institute, Janelia Farm, Ashburn, VA, and accepted by the Editorial Board August 6, 2008 (received for review April 12, 2008)

Messenger RNA molecules are tightly regulated, mostly through interactions with proteins and other RNAs, but the mechanisms that confer the specificity of such interactions are poorly understood. It is clear, however, that this specificity is determined by both the nucleotide sequence and secondary structure of the mRNA. Here, we develop RNApromo, an efficient computational tool for identifying structural elements within mRNAs that are involved in specifying posttranscriptional regulations. By analyzing experimental data on mRNA decay rates, we identify common structural elements in fast-decaying and slow-decaying mRNAs and link them with binding preferences of several RNA binding proteins. We also predict structural elements in sets of mRNAs with common subcellular localization in mouse neurons and fly embryos. Finally, by analyzing pre-microRNA stem-loops, we identify structural differences between pre-microRNAs of animals and plants, which provide insights into the mechanism of microRNA biogenesis. Together, our results reveal unexplored layers of posttranscriptional regulations in groups of RNAs and are therefore an important step toward a better understanding of the regulatory information conveyed within RNA molecules. Our new RNA motif discovery tool is available online.

bioinformatics | motif prediction | posttranscriptional regulation | RNA secondary structure | SCFGs

RNA molecules undergo diverse posttranscriptional regulation of gene expression, including regulation of RNA transport and localization, mRNA translation, and RNA decay (1–3). In many cases, such posttranscriptional regulation occurs through elements on the mRNA molecule that interact with the hundreds of RNA binding proteins (RBPs) that exist in the cell (4). A well-known example is the iron-responsive element (IRE), a secondary structure RNA motif located on UTRs of members of the iron metabolism and transport pathway (5). The binding of the RBPs Irp1 and Irp2 to IRE elements affects the translation rate of the mRNA, and by that coordinates the response to changing levels of iron in the environment. Other examples include a 118-nucleotide stem-loop structure through which mRNAs are transported to the yeast bud tip by the RBP She2 (6) and the RBP Sbp2 that is involved in mediating UGA redefinition from a stop codon to selenocysteine by binding specific stem-loop structures, termed selenocystein insertion site (SECIS) elements, in the 3' UTR of selenoproteins (7). In other cases, elements on the mRNA molecule interact with other RNAs that direct the regulatory effect. For example, the recognition and binding affinity of a microRNA to its mRNA target is determined by both the sequence and structure of the target mRNA (8–11).

The examples above suggest that the posttranscriptional regulation of mRNAs is determined not only by its linear nucleotide sequence but also by its secondary structure. Thus, a key goal is to understand the involvement of mRNA secondary structures in such regulation. One approach is to identify recurring patterns, termed *motifs*. However, linear sequence motifs, which are commonly found in DNA sequences, are not suitable in this case. Instead, we wish to identify motifs that combine primary and secondary structural elements and are therefore better suited to describe functional elements in RNA molecules.

Here we develop RNApromo (RNA prediction of motifs), a new computational method to identify short structural RNA motifs in sets of long unaligned RNAs. Using our method, we identify putative motifs in sets of mRNAs with substantial experimental evidence for a common posttranscriptional regulation, and we support these findings with cross-validation analysis. In some cases, sequence conservation of the putative motifs provides strong independent support for our findings. The identified motifs include motifs for mRNAs with similar decay rates, mRNAs that are bound by the same RBP, and mRNAs with a common cellular localization. Additionally, analysis of pre-microRNA structures identified differences between animals and plants in the sizes of stem and loop structures of pre-microRNAs, suggesting that animals and plants use different mechanisms for microRNA biogenesis.

Results and Discussion

A New Computational Scheme for RNA Motif Discovery. Several tools exist for finding local structural motifs in a set of long input RNAs. Many existing works use free energy minimization considerations for predicting local motifs, by applying one of three major schemes: use of a sequence-based local alignment to build a motif consensus structure (12, 13); identification of common structures in the RNAs' predicted minimal free energy folds (14–16); or simultaneous alignment of the RNAs and prediction of their secondary structure (17–20). Recently, some graph theoretical techniques were also proposed for this task (21, 22). Yet some of the most successful approaches for motif discovery are based on advanced probabilistic models, which are highly suitable to capture the observed variation in the input set.

Stochastic context-free grammars (SCFGs) are a class of probabilistic models proposed for modeling common sequence and structure in a set of input RNAs (23, 24), which replaced the thermodynamic considerations in several existing RNA motif discovery tools (25, 26). However, currently available SCFG applications optimize the model's parameters by essentially considering all possible secondary structures of the input sequences. This approach is successful mainly when it is possible to exploit covariation between organisms to infer the secondary structure at the motif position. However, when one wishes to identify *short* motifs, using data from a single organism, covariation data cannot be used. Moreover, enumerating all possible secondary structures results in a rather high time complexity of the algorithm, making it unfeasible to scan large RNA sets or long RNA sequences.

Here, we devised a new SCFG-based method for finding local motifs in a set of unaligned RNAs, which restricts the search space to a predefined and limited number of structures for each input RNA. Yet which structures should be considered by the algorithm?

Author contributions: M.R., M.K., and E.S. designed research, performed research, contributed new analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. S.R.E. is a guest editor invited by the Editorial Board.

[†]To whom correspondence should be addressed. E-mail: eran@weizmann.ac.il.

This article contains supporting information online at www.pnas.org/cgi/content/full/0803169105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA

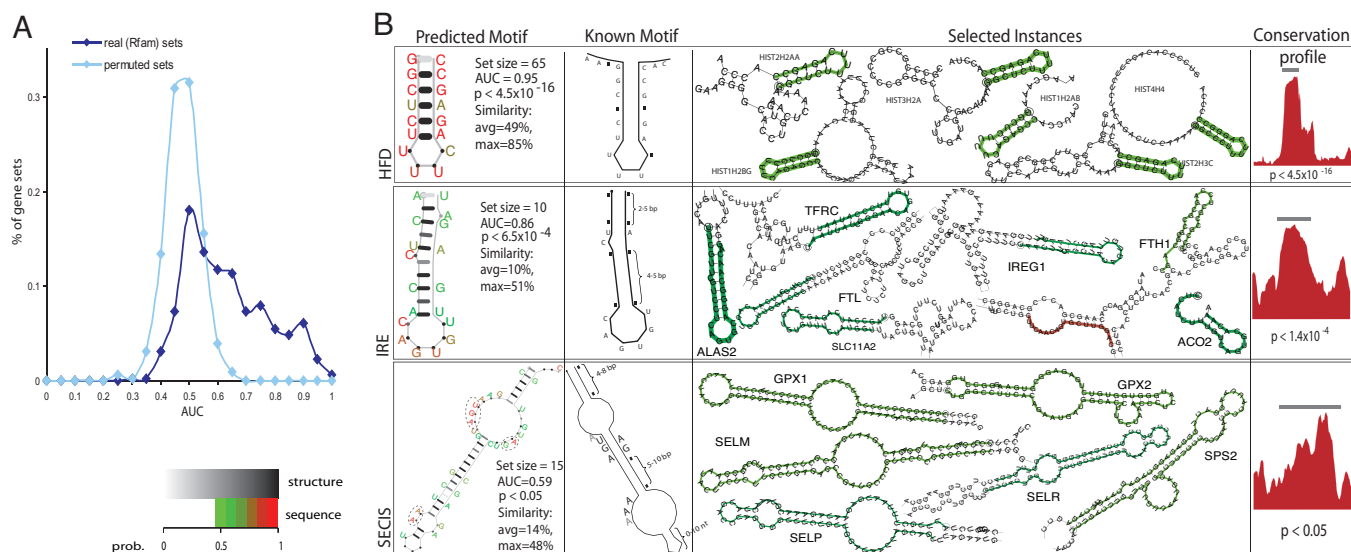


Fig. 1. Validation of motif discovery scheme. (A) Distribution of AUC scores for motifs in the true (dark blue) vs. permuted (light blue) sets (using ViennaRNA predicted folds). AUC scores of permuted sets are distributed around 0.5, whereas AUC for motif-containing sets are usually higher. (B) Prediction of three known human motifs: HFD (Top), IRE (Middle), and SECIS (Bottom). Motifs are represented using a structural logo of the motif's most probable structure. Specific positions are color coded according to their probability (green-to-red scale for sequence, and gray scale for structure). Shown is the predicted motif logo and the known consensus structure (Left), with several examples of correctly classified motifs (Right). The motif position is annotated in green, and the 5' end of the motif is circled. The FTH1 UTR (in the IRE part), which was misfolded at the IRE loop, is shown in red. Sequence conservation profile (average across all of the motif instances) is also shown (Rightmost Column).

Ideally, experimental information about RNA structure can be used, and only the few structures that are consistent with such structural data are considered. Even though such experimental structural information currently exists for only a small number of cases, such information may soon be available. Alternatively, using existing thermodynamic-based secondary structure prediction programs (14, 27) (with ~50%–90% prediction accuracy [28]) for predicting a small set of thermodynamically stable folds and restricting the algorithm to those, allows both reduction of the search space and integration of thermodynamic considerations into our model, which are not fully embedded into standard covariance model applications. Certainly, other structure prediction tools, such as those based on probabilistic models (29), can also be used to derive a set of highly probable folds as an input to our method.

Our algorithm, called RNAPromo, takes as input a set of RNA sequences assumed to share a motif, and their suggested secondary structures. The algorithm first identifies specific and relatively short candidate structures that appear in as many inputs as possible. These candidates are then used as seeds for a probabilistic inference algorithm that refines the predicted motif using statistical estimation (see supporting information (SI) for a full description).

To examine the performance of RNAPromo, we first tested its ability to identify known RNA structural motifs in a large collection of validation sets from the Rfam database (30). We used a fivefold cross-validation scheme, in which we partition the input set into five parts, learn a model from each of the possible combinations of four sets, and use this model to assign likelihood scores to the RNAs that were held out while learning it. We use the standard receiver operating characteristic (ROC) curve and its associated area under the curve (AUC) measure to evaluate the significance of the input RNAs' likelihood scores compared with shuffled sequences. We filtered each set to include only sequences with <90% sequence similarity, because high sequence similarity may produce high AUC scores even when a functional motif is not present. The maximal sequence similarity in our validation sets ranges from 30% to 90% (see SI for details). Although high AUC scores indicate that the input RNAs share a biological signal, we still have to validate that this signal results from a common motif. To test that, we applied the same motif discovery scheme to a collection of random sets that

include randomly selected sequences from all of the validation sets. Because we do not expect to identify motifs in these random sets, their AUC scores should be close to 0.5, indicating that in the absence of a common motif in the input set, no signal is detected.

The results (Fig. 1A) show that whereas the AUC scores of the random sets are indeed distributed around 0.5, with only 5% of the scores above 0.6, more than 60% of the AUC scores of the true sets are above 0.6. This indicates that in most cases, RNAPromo indeed detects the biological signal when it is present, yet it rarely detects such a signal if there is no common motif in the input RNAs. Similar results were obtained using a different method to predict the structure of the input RNAs (29) (see SI).

Repeating the same analysis with other available tools (26) shows that RNAPromo is compatible with these tools in terms of results but requires shorter running times (see SI). This result is notable, given that the Rfam database includes motif instances from several organisms, and in such a setting, tools that exploit covariation information have an advantage over RNAPromo.

Next, we tested our method's ability to identify and correctly describe known biological motifs when covariation data is minimal. We analyzed three well-known motifs from human: histone-fold domain (HFD), iron-responsive element (IRE), and seleno-cystein insertion site (SECIS), each sharing a different level of sequence similarity. Although three examples cannot provide global statistics, they can still allow a better evaluation of our method's abilities in such a setting.

Using RNAPromo, we detected a motif in all three sets (Fig. 1B): HFD (AUC = 0.95, $P < 5 \times 10^{-16}$), IRE (AUC = 0.86, $P < 7 \times 10^{-4}$), and SECIS (AUC = 0.59, $P < 0.05$). The predicted consensus structures are highly similar to those described in the literature and usually match the known motif positions (94% of the HFD instances, 60% of the IRE instances, and 47% of the SECIS instances). In the misclassified cases, the motif position was not folded into the correct structure by the prediction algorithm. Note that the SECIS consensus structure includes two noncanonical G-A base pairs, which standard folding algorithms cannot predict. Nonetheless, RNAPromo identifies a motif in this set and predicts its structure quite accurately. The predicted SECIS motif obviously

structural context within which these sequence elements are embedded determines their activity: a small loop will induce destabilization, whereas a long U-rich loop will stabilize the mRNA.

Many studies have shown that much of the posttranscriptional regulation of mRNAs in the cell occurs through their interactions with the hundreds of different cellular RBPs. It is therefore possible that the motifs we identified in the fast- and slow-decaying mRNAs are bound by specific proteins that modulate mRNA stability. In an attempt to identify such proteins, we selected several available sets of genome-wide measurements of mRNAs bound by the same RBP and tried to identify common motifs in them.

Proteins from the Puf family of RBPs have been reported to bind UGUR motifs located in the 3' UTR of their targets and thereby repress gene expression by affecting mRNA stability and translation rate. A recent study in yeast (32) measured the set of RNAs bound by five RBPs from the Puf family, resulting in groups of 40–220 bound mRNAs per protein. By applying a DNA motif discovery tool to the 3' UTR primary sequence of the mRNAs bound by each Puf protein, the investigators were able to identify distinct 10-nucleotide RNA sequence motifs containing the UGUR element in mRNAs interacting with Puf3, Puf4, and Puf5. Applying RNApromo to the targets of each of the five Puf proteins, we identify a significant motif in the 3' UTRs of Puf3 ($AUC = 0.57$, $P < 2 \times 10^{-3}$), Puf4 ($AUC = 0.58$, $P < 8 \times 10^{-5}$), and Puf5 ($AUC = 0.59$, $P < 4 \times 10^{-5}$) targets. The three motifs are somewhat similar (Fig. 2C) and include an AU-rich sequence that is folded into a stem-loop structure with a relatively short loop. For the Puf5 motif, we find a significant level of sequence conservation ($P < 10^{-4}$), as an independent support for the predicted motif. Moreover, the Puf5 consensus structure includes the UGU element, which is part of the previously proposed Puf5 sequence motif (in 63% of the predicted Puf5 target sites).

The yeast polyU binding protein, Pub1, an embryonic lethal abnormal visual (ELAV)-like RBP with mammalian homologues, is known to play key roles in cellular mRNA decay. Pub1 was shown to bind with high affinity different target mRNAs with AU-rich sequences in their 3' UTR and stabilize them (33). In a recent study, a genome-wide measurement of Pub1 targets identified 368 target transcripts. Applying RNApromo to the 3' UTRs of Pub1 targets, we predict a significant motif ($AUC = 0.6$, $P < 2 \times 10^{-5}$), which is also highly conserved ($P < 5 \times 10^{-10}$). The predicted motif is AU rich and folds into a stem-loop structure that includes a long and highly U-rich loop (Fig. 2C).

Looking at the motifs we predict for these RBPs, we notice their similarity to the mRNA decay motifs: in targets of Pub1, a protein that is known to stabilize its targets, we identify a very similar motif to the slow-decay motif, whereas in targets of the Puf proteins, which were suggested to increase mRNA degradation rates, we identify a motif with a similar structure to the fast-decay motif. Indeed, fast-decaying mRNAs (half-life <6 min) are enriched with Puf protein targets ($P < 8 \times 10^{-16}$), whereas slow-decaying mRNAs (half-life >40 min) are enriched for Pub1 targets ($P < 2 \times 10^{-3}$). Moreover, looking at the decay rates of the top 20% of targets of each protein (scored by the predicted motifs), we see (Fig. 2C) that the half-lives of the Puf targets are lower (18 ± 13 min on average) than the Pub1 targets (28 ± 32 min on average).

Finally, Sam68 is a human RBP involved in cell growth regulation, which was shown to bind to AU-rich motifs within a stem-loop context (34). Analyzing its eight known targets, we predict a stem-loop motif ($AUC = 0.73$, $P < 2 \times 10^{-3}$). On the basis of the previous observations and the structure of this motif, we can hypothesize that Sam68 promotes the degradation of its mRNA targets. If true, this would indicate that this destabilization motif is conserved between human and yeast.

Overall, our results suggest a role for mRNA secondary structure in controlling mRNA stability. The detailed secondary structure of the well known AU-rich elements may play a significant role in determining their function as stabilizing or

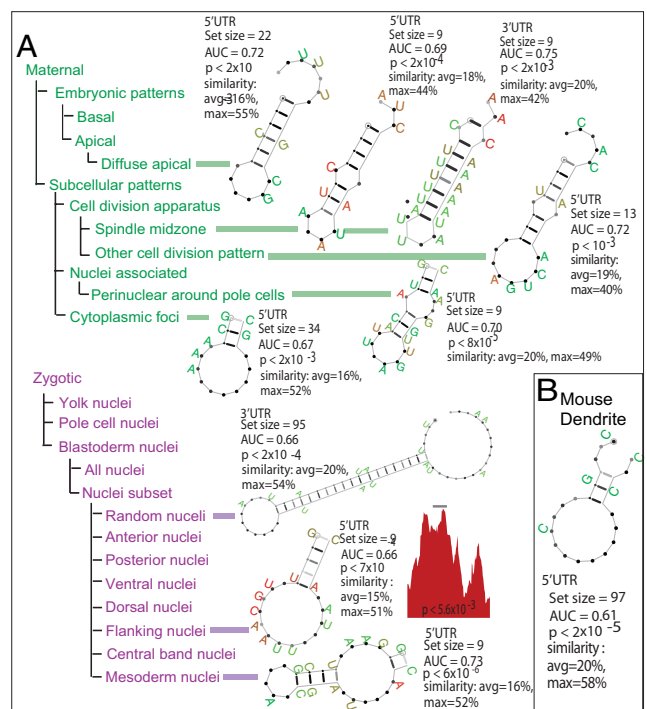


Fig. 3. Localization motifs. (A) Fly embryonic localization patterns. Consensus structure logos of the statistically significant RNA motifs. The location of each set in the localization annotation hierarchy is indicated. (B) RNA motif predicted for mRNAs that are localized to mouse dendrites.

destabilizing elements. By binding different RBPs, AU-rich elements with different structure can change the mRNA stability and therefore affect its translation rates.

Motifs Involved in mRNA Localization. RNA transport and local translation have now been documented in vertebrates, invertebrates, and unicellular organisms, enabling cells to control gene expression at small localized regions. Several studies suggest that cellular localization of mRNAs is specified by RNA motifs on the localized mRNA and bound by RBPs involved in trafficking (5). The RBPs involved in several localization processes were identified experimentally and shown to interact with elements composed of both sequence and structure (6). We thus applied our motif discovery scheme to several sets of mRNAs that share similar localization patterns.

Recently, a large study on mRNA localization during fly embryonic development was performed (35). This study provides a large set of colocalized mRNA transcripts, both to specific parts of the embryo and to distinct cellular compartments. Although some of the mechanisms for mRNA localization during development are known and involve morphogen gradients that induce local transcription of mRNA, these are not active at the subcellular level or for maternal mRNAs. Therefore, it is possible that some of the documented localization events are a result of signals on the mRNA itself. Applying RNApromo to the 94 sets of colocalized mRNAs, we detect significant motifs in nine sets of colocalized maternal and embryonic transcripts (Fig. 3A). Interestingly, whereas mRNA stability motifs are located on 3' UTRs, most predicted localization motifs are located on the 5' UTR of the transcripts.

Of the nine predicted motifs, five (56%) are identified in sets of maternal transcripts with a common cellular localization, although such sets constitute only 7% of the localization patterns ($P < 5 \times 10^{-5}$). One interesting example is the motif predicted in maternal transcripts localized to the spindle midzone, the central area of the spindle where microtubules from opposite poles overlap. This

in other genomic transcripts. To test this hypothesis, we applied RNApromo to stem-loops that appear in arbitrary genomic transcripts. Strikingly, we find that in plants, the length of the stem and loop of arbitrary stem-loop-containing transcripts is very similar to the lengths of the plant pri-microRNAs, whereas in animals, stem-loops of arbitrary transcripts and stem-loops of animal pri-microRNAs have markedly different lengths. Thus, our results suggest that in animals, the nuclear RNase Droscha recognizes and cleaves its pri-microRNA targets, perhaps by measuring the length of their loop, among other features. Longer loops, which are characteristic of pri-microRNAs, enable the enzyme to specifically recognize these transcripts. These results further suggest that in plants, Droscha recognition is not part of the microRNA biogenesis process, and thus a different mechanism may exist.

These intriguing results are supported by the observation that although in animals cleavage by Droscha is essential for pre-microRNA export to the cytoplasm and further processing by Dicer, most Dicer homologues in plants are already localized to the nucleus (37). Moreover, the existing literature provides no evidence of a plant Droscha homologue and no evidence for accumulation of pre-microRNAs after a Dicer knockout in plants (40). Thus, it is possible that plant pri-microRNAs, in a similar way to siRNAs precursors, are directly recognized and cleaved by Dicer. Indeed, plant pre-microRNAs have long stems that are similar to the long dsRNA precursors that generate siRNAs, and the mature microRNAs, like mature siRNAs, usually induce degradation of their targets rather than a translational arrest.

Conclusions

In summary, we have developed and applied a novel computational tool, RNApromo, that identifies short RNA motifs from the full-length RNA regions in which they are embedded. Unlike

other available tools, RNApromo restricts the motif search to a predefined set of input structures, which allows it to perform well even on sets of mRNAs from a single organism, reduces the time complexity, and allows the use of thermodynamic-based methods for predicting RNA secondary structure.

Using RNApromo, we predict two structurally different and AU-rich motifs in 3' UTRs of fast-decaying and slow-decaying mRNAs and identify proteins that may bind these motifs and modulate the stability of the bound mRNAs. Although the involvement of AU-rich elements in this process is known (31), our analysis reveals the importance of the structural context in which these elements are embedded. Next, we predict motifs in sets of colocalized transcripts in mouse neurons and fly embryos, which can be involved in establishing local cellular concentrations of mRNAs. Finally, we analyze pre-microRNA sequences and reveal that in animals Droscha may recognize its pri-microRNA targets by measuring, among other features, the length of their loop. Moreover, our analysis suggests that plant microRNAs are processed similarly to other siRNAs, in a mechanism that requires only the Dicer enzyme. Overall, the predicted motifs represent novel and experimentally testable findings and demonstrate the potential of RNApromo for uncovering posttranscriptional regulatory mechanisms. Our work thus represents a step toward the long-term goal of making the genome-wide computational prediction of RNA secondary structure elements as routine and robust as is currently practiced for linear DNA elements, thereby revealing some of the mechanisms by which RNA molecules are regulated within the cell.

Methods

See the SI for full details on our methods and on the datasets we used, as well as for additional results. An online version and the implementation of the RNApromo algorithm are available from the authors' website or upon request.

- Arava Y, et al. (2003) Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 100:3889–3894.
- Shepard KA, et al. (2003) Widespread cytoplasmic mRNA transport in yeast: Identification of 22 bud-localized transcripts using DNA microarray analysis. *Proc Natl Acad Sci USA* 100:11429–11434.
- Wang Y, et al. (2002) Precision and functional specificity in mRNA decay. *Proc Natl Acad Sci USA* 99:5860–5865.
- Anantharaman V, Koonin EV, Aravind L (2002) Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res* 30:1427–1464.
- Hentze MW, Muckenthaler MU, Andrews NC (2004) Balancing acts: Molecular control of mammalian iron metabolism. *Cell* 117:285–297.
- Olivier C, et al. (2005) Identification of a conserved RNA motif essential for She2p recognition and mRNA localization to the yeast bud. *Mol Cell Biol* 25:4752–4766.
- Krol A (2002) Evolutionarily different RNA motifs and RNA-protein complexes to achieve selenoprotein synthesis. *Biochimie* 84:765–774.
- Kertesz M, et al. (2007) The role of site accessibility in microRNA target recognition. *Nat Genet* 39:1278–1284.
- Robins H, Li Y, Padgett RW (2005) Incorporating structure to predict microRNA targets. *Proc Natl Acad Sci USA* 102:4006–4009.
- Long D, et al. (2007) Potent effect of target structure on microRNA function. *Nat Struct Mol Biol* 14:287–294.
- Zhao Y, Samal E, Srivastava D (2005) Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* 436:214–220.
- Hofacker IL, Fekete M, Stadler PF (2002) Secondary structure prediction for aligned RNA sequences. *J Mol Biol* 319:1059–1066.
- Gautheret D, Lambert A (2001) Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *J Mol Biol* 313:1003–1011.
- Hofacker IL, et al. (1994) Fast folding and comparison of RNA secondary structures. *Monatshfte fur Chemie* 125:167–188.
- Hochsmann M, et al. (2003) Local similarity in RNA secondary structures. *Proc IEEE Comput Soc Bioinform Conf* 2:159–168.
- Pavesi G, et al. (2004) RNAPProfile: An algorithm for finding conserved secondary structure motifs in unaligned RNA sequences. *Nucleic Acids Res* 32:3258–3269.
- Sankoff D (1985) Simultaneous solution of the RNA folding, alignment and protosequence problems. *SIAM J Appl Math* 45:810–825.
- Gorodkin J, Heyer LJ, Stormo GD (1997) Finding the most significant common sequence and structure motifs in a set of RNA sequences. *Nucleic Acids Res* 25:3724–3732.
- Harmanci AO, Sharma G, Mathews DH (2007) Efficient pairwise RNA structure prediction using probabilistic alignment constraints in Dynalign. *BMC Bioinformatics* 8:130.
- Will S, et al. (2007) Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. *PLoS Comput Biol* 3:e65.
- Ji Y, Xu X, Stormo GD (2004) A graph theoretical approach for predicting common RNA secondary structure motifs including pseudoknots in unaligned sequences. *Bioinformatics* 20:1591–1602.
- Hamada M, et al. (2006) Mining frequent stem patterns from unaligned RNA sequences. *Bioinformatics* 22:2480–2487.
- Eddy SR, Durbin R (1994) RNA sequence analysis using covariance models. *Nucleic Acids Res* 22:2079–2088.
- Sakakibara Y, et al. (1994) Stochastic context-free grammars for tRNA modeling. *Nucleic Acids Res* 22:5112–5120.
- Holmes I (2005) Accelerated probabilistic inference of RNA structure evolution. *BMC Bioinformatics* 6:73.
- Yao Z, Weinberg Z, Ruzzo WL (2006) CMfinder—A covariance model based RNA motif finding algorithm. *Bioinformatics* 22:445–452.
- Wuchty S, et al. (1999) Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers* 49:145–165.
- Wiese KC, Hendriks A (2006) Comparison of P-RnaPredict and mfold—Algorithms for RNA secondary structure prediction. *Bioinformatics* 22:934–942.
- Do CB, Woods DA, Batzoglou S (2006) CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics* 22:e90–e98.
- Griffiths-Jones S, et al. (2005) Rfam: Annotating non-coding RNAs in complete genomes. *Nucleic Acids Res* 33:D121–D124.
- Barreau C, Paillard L, Osborne HB (2005) AU-rich elements and associated factors: Are there unifying principles? *Nucleic Acids Res* 33:7138–7150.
- Gerber AP, Herschlag D, Brown PO (2004) Extensive association of functionally and topologically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol* 2:E79.
- Dutta Gupta R, et al. (2005) Global analysis of Pub1p targets reveals a coordinate control of gene expression through modulation of binding and stability. *Mol Cell Biol* 25:4999–5513.
- Itoh M, et al. (2002) Identification of cellular mRNA targets for RNA-binding protein Sam68. *Nucleic Acids Res* 30:5452–5464.
- Lecuyer E, et al. (2007) Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131:174–187.
- Zhong J, Zhang T, Bloch LM (2006) Dendritic mRNAs encode diversified functionalities in hippocampal pyramidal neurons. *BMC Neurosci* 7:17.
- He L, Hannon GJ (2004) MicroRNAs: Small RNAs with a big role in gene regulation. *Nat Rev Genet* 5:522–531.
- Han J, et al. (2006) Molecular basis for the recognition of primary microRNAs by the Droscha-DGCR8 complex. *Cell* 125:887–901.
- Cullen BR (2006) Viruses and microRNAs. *Nat Genet* 38(Suppl):S25–S30.
- Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* 57:19–53.