RNA stability and translation depends on specific *cis*-acting sequences (often within 3'-UTR).

Motif finder

Letter

Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes

Edward Yang,[1,6] Erik van Nimwegen,[4,6] Mihaela Zavolan,[2] Nikolaus Rajewsky,[5] Mark Schroeder,[2] Marcelo Magnasco,[3] and James E. Darnell Jr.[1,7]

[1]*Laboratory of Molecular Cell Biology,* [2]*Laboratory of Computational Genomics,* [3]*Laboratory of Mathematical Physics, and* [4]*Center for the Study of Physics and Biology, The Rockefeller University, New York, New York 10021-6399, USA;* [5]*Department of Biology and Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, USA*

Although mRNA decay rates are a key determinant of the steady-state concentration for any given mRNA species, relatively little is known, on a population level, about what factors influence turnover rates and how these rates are integrated into cellular decisions. We decided to measure mRNA decay rates in two human cell lines with high-density oligonucleotide arrays that enable the measurement of decay rates simultaneously for thousands of mRNA species. Using existing annotation and the Gene Ontology hierarchy of biological processes, we assign mRNAs to functional classes at various levels of resolution and compare the decay rate statistics between these classes. The results show statistically significant organizational principles in the variation of decay rates among functional classes. In particular, transcription factor mRNAs have increased average decay rates compared with other transcripts and are enriched in "fast-decaying" mRNAs with half-lives <2 h. In contrast, we find that mRNAs for biosynthetic proteins have decreased average decay rates and are deficient in fast-decaying mRNAs. Our analysis of data from a previously published study of *Saccharomyces cerevisiae* mRNA decay shows the same functional organization of decay rates, implying that it is a general organizational scheme for eukaryotes. Additionally, we investigated the dependence of decay rates on sequence composition, that is, the presence or absence of short mRNA motifs in various regions of the mRNA transcript. Our analysis recovers the positive correlation of mRNA decay with known AU-rich mRNA motifs, but we also uncover further short mRNA motifs that show statistically significant correlation with decay. However, we also note that none of these motifs are strong predictors of mRNA decay rate, indicating that the regulation of mRNA decay is more complex and may involve the cooperative binding of several RNA-binding proteins at different sites.

[Supplemental material is available online at www.genome.org, and also at http://genomes.rockefeller.edu/~yange.]

articolo (in parte discusso nella lezione 10 (4-dic)

from Yang et al., 2003.

HepG2 cells

Actinomycin 2-3 hours

RNA extraction, labelling and → Affymetrix microarrays

Repeated on primary fibroblasts

Decay rates estimated for 5,245 genes.

Group of genes either short lived or long lived explored for ontology

...
To study the rates of mRNA degradation ("decay") in human cells, we measured changes in mRNA levels following application of the RNA polymerase inhibitor Actinomycin D with Affymetrix U95Av2 high-density oligonucleotide arrays.

In this way, we obtained decay rate estimates for 5245 accessions

Combining the decay rate for all probe sets present in the initial and final conditions, we find that the median half-life in both cell types was ~10 h

Comparing this median half-life with the median half-lives of transcripts in yeast and bacteria, it appears that the half-life of the mRNA pool of a cell scales roughly in proportion to the length of the cell cycle: cell cycle lengths of 20, 90, and 3000 min correspond to median half-lives of 5, 21, and 600 min, respectively, for *E. coli, S. cerevisiae, a*nd human HepG2/Bud8 cells

from Yang et al., 2003.

As indicated in the cumulative distribution plot, a small percentage (~5%) of expressed transcripts have “fast” decay rates (which we define as *r > 0.5 h1* or a half-life < 2 h). A similar percentage of rapidly decaying genes was observed when we re-ran an HepG2 experiment with U95B arrays, which are predominantly expressed sequence tags.

....

Although total length of cDNA did not correlate with decay rate, we did find evidence that mRNAs with 3-UTR sequence >1 kb decayed at a significantly faster rate than shorter 3-UTRs

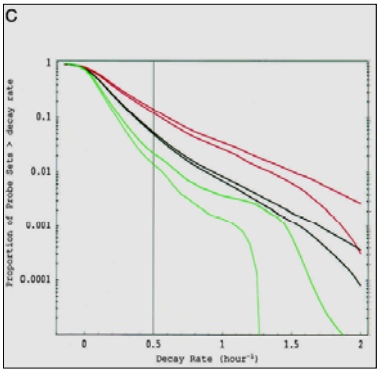


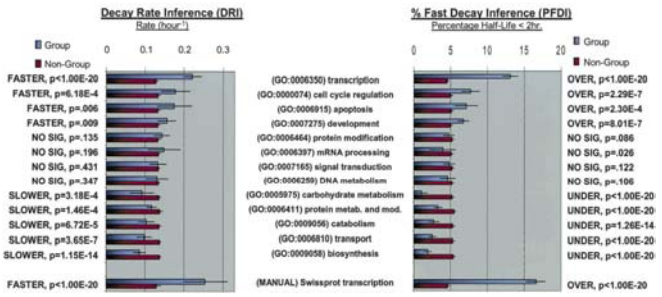
Figure 1 Functional analysis of decaying transcripts in human cells.

(C) Reverse cumulative distribution of decay rates for probe sets in different functional classes (HepG2 experiments). Decay rate r is shown horizontally, while vertically the fraction of probe sets with decay rates higher than r is plotted on a logarithmicsc ale. The pairs of lines show the 98% posterior probability intervals for the fraction at each value of r. (Red) GO process transcription; (black) all probe sets; (green) biosynthesis. The gray line indicates the decay rate r = 0.5 h□1, which is our cutoff for fast decay in PFDI.

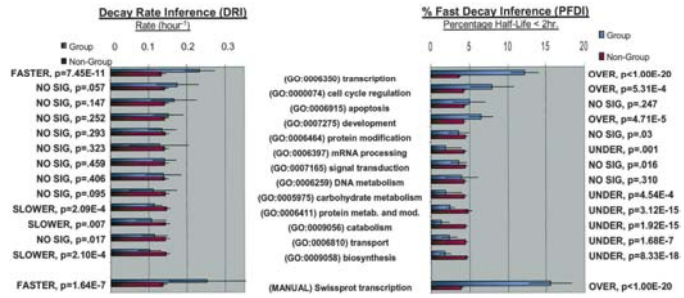
Decay rates were studied in functional groups, i.e. in G.O. categories.

from Yang et al., 2003.

A. HepG2



B. Bud8



from Yang et al., 2003.

Known and novel motifs were examined: for each of these, DRI and PDFI is calculated and deviation from global estimated.

HepG2 motif summary

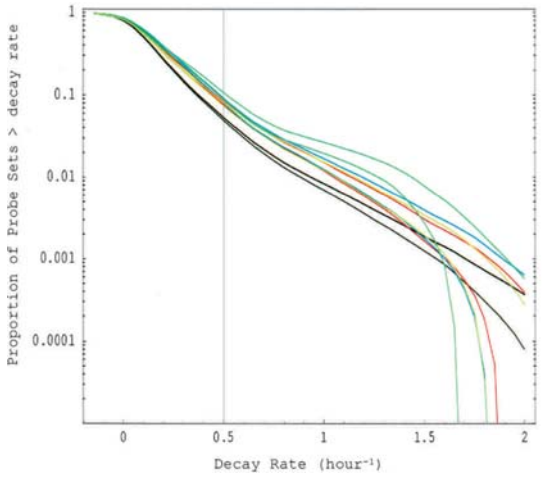
		DRI				PFDI			
		p-value for In-Group rate change				p-value for In-Group Representation change			
Motif	**Sequence**	5'UTR	ORF	3'UTR	WHOLESEQ	5'UTR	ORF	3'UTR	WHOLESEQ
Described Motifs									
1	[AT]ATTTA[AT]		2.02E-12	0	0	2.95E-06	0	0	0
2A	ATTTATTTATTTATTTATTTA		na				na		
2B	ATTTATTTATTTATTTA		na				na		
2C	[AT]ATTTATTTATTTA[AT]		na				na		
2D	[AT]{2}ATTTATTTA[AT]{2}		na				na	0.009	
2E	[AT]{4}ATTTA[AT]{4}		3.50E-04	1.13E-13	1.38E-14		1.70E-14	0	0
MEG	TTATTTATT			2.57E-07	1.00E-07			7.50E-20	1.00E-20
MEGSHORT	TATTTAT		5.03E-04	0	0	8.75E-04	0.001	0	0
Undescribed Motifs									
H1	TTTTTTT			0	3.50E-20	2.33E-04	0.001	0	0
H2	TTTTTAAA	3.83E-04	4.84E-04	0	0	1.93E-10	2.03E-10	0	0
H3	TTGTAAATA			4.77E-11	6.96E-10		4.43E-05	0	0
B1	TTTTAAAT		1.48E-06	1.52E-11	1.57E-13		0	0	0
B2	TTTTAATTT				0.005	*0.007*	0.008	7.24E-04	1.94E-04
B3	AAATATTTT		0.004	6.05E-09	5.11E-10		6.33E-05	0	0
B4	AATATTTTT			3.43E-09	6.77E-10		7.94E-05	2.00E-20	0
H-1	CCGCCTC		0.005					*2.86E-05*	
H-2	CCAGCCTC		1.93E-08		4.18E-04	*2.89E-04*	0	*8.56E-11*	9.39E-07
B-1	GGGCCTGG					*0.004*	*0.007*		*5.65E-04*
B-2	CCCAGCCCC		7.72E-04				2.96E-06	*1.78E-07*	

from Yang et al., 2003.

Bud8 motif summary

		DRI				PFDI			
		p-value for In-Group rate change				p-value for In-Group Representation change			
Motif	**Sequence**	5'UTR	ORF	3'UTR	WHOLESEQ	5'UTR	ORF	3'UTR	WHOLESEQ
Described Motifs									
1	[AT]ATTTA[AT]		0.010	1.92E-13	6.14E-12		0.002	0	0
2A	ATTTATTTATTTATTTATTTA		na	na			na	na	
2B	ATTTATTTATTTATTTA		na	na			na	na	
2C	[AT]ATTTATTTATTTA[AT]		na				na	*0.005*	*9.96E-04*
2D	[AT]{2}ATTTATTTA[AT]{2}		na				na		
2E	[AT]{4}ATTTA[AT]{4}			5.50E-06	2.18E-05			4.22E-16	1.44E-13
MEG	TTATTTATT			0.001	0.004		*3.38E-06*	4.44E-11	1.17E-08
MEGSHORT	TATTTAT			8.05E-10	1.40E-07			0	2.22E-15
Undescribed Motifs									
H1	TTTTTTT			3.32E-07	1.22E-06		*0.001*	0	0
H2	TTTTTAAA			1.39E-07	9.70E-09	0.009		0	0
H3	TTGTAAATA			3.45E-06	7.95E-06		6.20E-05	3.44E-16	2.70E-15
B1	TTTTAAAT		0.004	0.001	1.50E-04		3.78E-08	7.61E-09	2.83E-10
B2	TTTTAATTT			0.005	0.003		0.009	1.06E-05	1.37E-06
B3	AAATATTTT							7.31E-06	2.78E-06
B4	AATATTTTT			0.003	0.007			7.70E-07	3.96E-06
H-1	CCGCCTC						0.009		0.008
H-2	CCAGCCTC						0.001	*0.001*	
B-1	GGGCCTGG					*1.84E-04*	0.007		
B-2	CCCAGCCCC			*0.003*				*8.51E-05*	

from Yang et al., 2003.



from Yang et al., 2003.

(*C) Reverse cumulative distribution of decay rates for* probe sets from genes that contain particular sequence motifs in their 3-UTR (HepG2 experiment). Decay rate *r is shown horizontally, while vertically the fraction of probe sets with decay rates higher than r is* plotted on a logarithmicsc ale. The pairs of lines show the 98% posterior probability intervals for the fraction at each value of *r. (Red) Motif 1; (blue) motif MEGSHORT; (green) Motif 2E; (light green) Motif H1;* (black) all probe sets. “Described” AU-rich decay motifs (1–2E, MEG, MEGSHORT) and “undescribed” motifs were derived from the sources mentioned in Methods.

....
Together, these observations show that the examined RNA motifs correlate with shifts in the distribution of decay rates, but that they do not reliably predict turnover behavior. It thus seems that the regulation of mRNA decay is more complicated and might involve combinatorial interactions, that is, cooperative binding between different RNAbinding proteins that bind at different sites in the mRNA. This might also explain why the effect on decay rate is context dependent for certain motifs (such as H-2).

from Yang et al., 2003.

Motifs imparting shorter half-life may represent protein-binding elements or miRNA targets (other ?).

microRNA biochemistry was initially attributed to translational effects however microarray studies have shown that a substantial part of 3’-UTR target-containing mRNAs are indeed downregulated “also” at the RNA level.

Review allegata a Lezione 6

Small silencing RNAs: an expanding universe

Megha Ghildiyal and Phillip D. Zamore

Abstract | Since the discovery in 1993 of the first small silencing RNA, a dizzying number of small RNA classes have been identified, including microRNAs (miRNAs), small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs). These classes differ in their biogenesis, their modes of target regulation and in the biological pathways they regulate. There is a growing realization that, despite their differences, these distinct small RNA pathways are interconnected, and that small RNA pathways compete and collaborate as they regulate genes and protect the genome from external and internal threats.

Table 1 | **Types of small silencing RNAs**

Name	Organism	Length (nt)	Proteins	Source of trigger	Function	Refs
miRNA	Plants, algae, animals, viruses, protists	20–25	Drosha (animals only) and Dicer	Pol II transcription (pri-miRNAs)	Regulation of mRNA stability, translation	93–95, 200–202,226
casiRNA	Plants	24	DCL3	Transposons, repeats	Chromatin modification	38,44,51, 52,61–63
tasiRNA	Plants	21	DCL4	miRNA-cleaved RNAs from the *TAS* loci	Post-transcriptional regulation	64–68
natsiRNA	Plants	22	DCL1	Bidirectional transcripts induced by stress	Regulation of stress-response genes	71,72
		24	DCL2			
		21	DCL1 and DCL2			
Exo-siRNA	Animals, fungi, protists	~21	Dicer	Transgenic, viral or other exogenous dsRNA	Post-transcriptional regulation, antiviral defense	4,5,8,227
	Plants	21 and 24				
Endo-siRNA	Plants, algae, animals, fungi, protists	~21	Dicer (except secondary siRNAs in *C. elegans*, which are products of RdRP transcription, and are therefore not technically siRNAs)	Structured loci, convergent and bidirectional transcription, mRNAs paired to antisense pseudogene transcripts	Post-transcriptional regulation of transcripts and transposons; transcriptional gene silencing	75–79,82, 83,86,87, 200,201, 228
piRNA	Metazoans excluding *Trichoplax adhaerens*	24–30	Dicer-independent	Long, primary transcripts?	Transposon regulation, unknown functions	157, 163–169, 177,202
piRNA-like (soma)	*Drosophila melanogaster*	24–30	Dicer-independent	In *ago2* mutants in *Drosophila*	Unknown	76
21U-RNA piRNAs	*Caenorhabditis elegans*	21	Dicer-independent	Individual transcription of each piRNA?	Transposon regulation, unknown functions	114, 173–175
26G RNA	*Caenorhabditis elegans*	26	RdRP?	Enriched in sperm	Unknown	114

ago2, Argonaute2; casiRNA, *cis*-acting siRNA; DCL, Dicer-like; endo-siRNA, endogenous small interfering RNA; exo-siRNA, exogenous small interfering RNA; miRNA, microRNA; natsiRNA, natural antisense transcript-derived siRNA; piRNA, Piwi-interacting RNA; Pol II, RNA polymerase II; pri-miRNA, primary microRNA; RdRP, RNA-dependant RNA polymerase; tasiRNA, *trans*-acting siRNA.

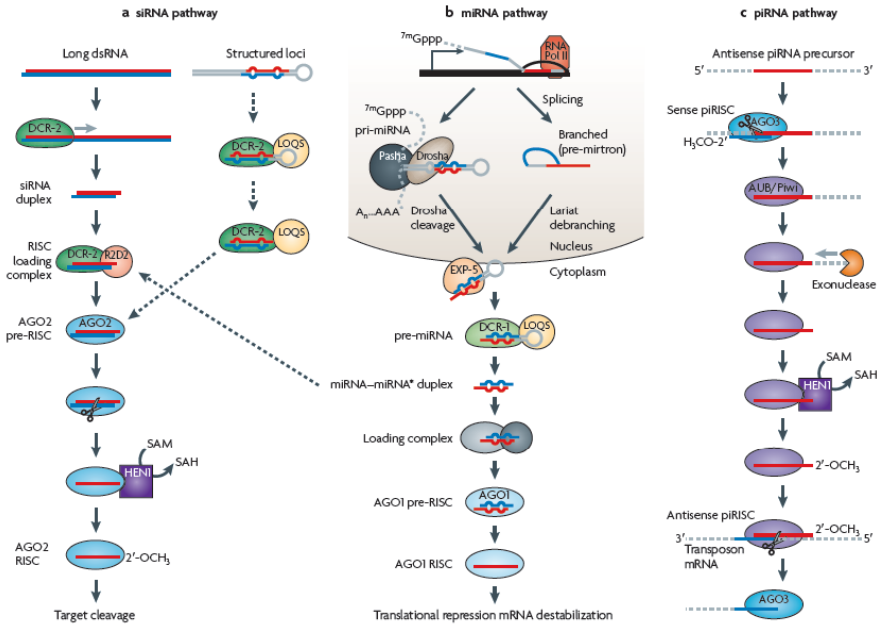
Both exo-siRNA and endo-siRNA

Variety and sources vary among different organisms

siRNA pathway: Ago2 is the main actor

Diversity among organisms: C. elegans has 27 different Argonaute proteins, D. melanogaster has 5, A. thaliana has 10.
Numbers of Dicer also vary, mammals have a single Dicer

Choice between guide and passenger strains of siRNA are selected based on thermodynamic stability of 5' end (the higher → guide)
The cut is between nucleotides paired to positions 10 and 11 of the guide (AGO2).



Small RNA silencing pathways in Drosophila.

Plants exhibit a surprising diversity of small RNA types and the proteins that generate them.

In plants, inverted-repeat transgenes or coexpressed sense and antisense transcripts produce two sizes of siRNAs: 21 and 24 nucleotides. The DCl4-produced 21-mers typically associate with AGo1 and guide mRNA cleavage. The 24-mers associate with AGo4 (in the major pathway) and AGo6 (in the surrogate pathway), and promote the formation of repressive chromatin.

In plants, single-stranded sense transcripts from tandemly repeated or highly expressed single-copy transgenes are converted to dsRNA by RDR6, a member of the RNA-dependent RNA polymerase (RdRP) family that transcribe ssRNAs from an RNA template. RDR6 and RDR1 also convert viral ssRNA into dsRNA, initiating an antiviral RNAi response.

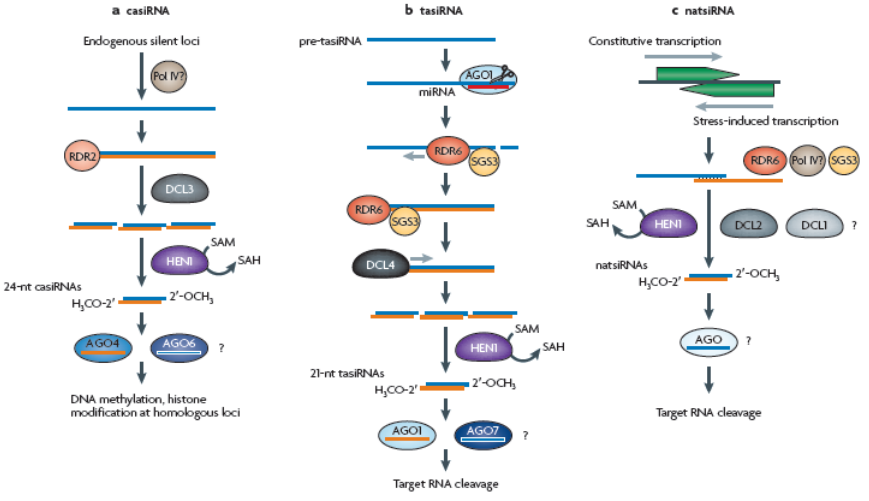


Figure 2 | **Plant endogenous small interfering RNA (endo-siRNA) biogenesis.** *Cis*-acting siRNAs (casiRNAs), *trans*-acting siRNAs (tasiRNAs) and natural antisense transcript-derived siRNAs (natsiRNAs) are derived from distinct loci. Several of the proteins involved in their biogenesis are genetically redundant, whereas others have specialized roles.

endo-siRNA

The first endo-siRNAs were detected in plants and *C. elegans, and the recent discovery of endo-siRNAs* in flies and mammals suggests that endo-siRNAs are ubiquitous among higher eukaryotes.

In many cases plant and worm endo-siRNA dependt upon RDRP activity.

The genomes of flies and mammals do not seem to encode such RdRP proteins, so the recent discovery of endo-siRNAs in flies and mice was unexpected.

The first mammalian endo-siRNAs to be reported corresponded to the long interspersed nuclear element (l1) retrotransposon and were detected in cultured human cells (2006).

More recently, endo-siRNAs have been detected in somatic and germ cells of *Drosophila species and in* mouse oocytes.
(most done by AGO2 immunoprecipitation followed by RNA-Seq)

Fly endo-siRNAs derive from transposons, heterochromatic sequences, intergenic regions, long RNA transcripts with extensive structure and, most interestingly, from mRNAs.

A subset of fly endo-siRNAs derives from ‘structured loci’, RNA transcripts of which can fold into long intramolecularly paired hairpins (intramolecular information) others from pseudogenes (trans).

endo-siRNAs have also been identified in mouse oocytes (Tam et al., Nature 453: 534-538, 2008; Watanabe et al., Nature 453: 539-543, 2008).

As in flies, mouse endo-siRNAs are 21 nucleotides, Dicer-dependent and derived from a variety of genomic sources

A subset of mouse oocyte endo-siRNAs maps to regions of protein-coding genes that are capable of pairing to their cognate pseudogenes, and to regions of pseudogenes that are capable of forming inverted repeat structures. (Interestingly, pseudogenes can no longer encode proteins, but they drift from their ancestral sequence more slowly than would be expected if they were simply junk DNA).

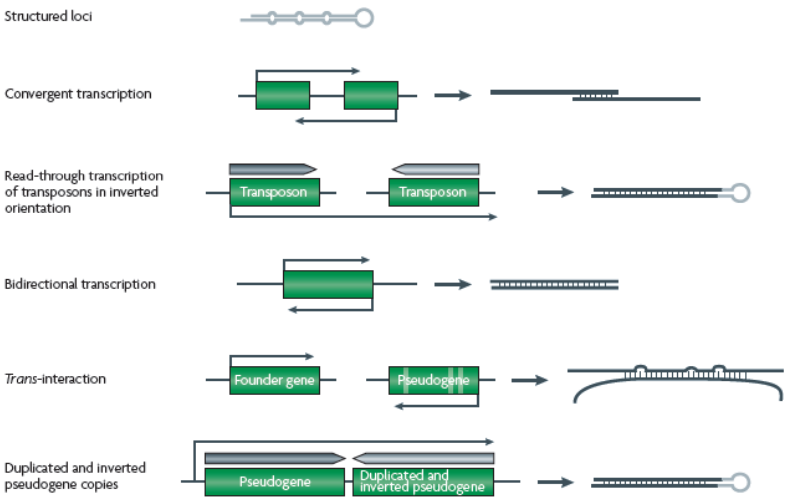


Figure 3 | **Genomic sources of dsRNA triggers for endogenous small interfering RNAs (endo-siRNAs) in flies and mammals**

piRNA are small RNAs associated to the Piwi-subfamily of Argonaute proteins.

piRNAs were first proposed to ensure germline stability by repressing transposons when Aravin and colleagues discovered in flies a class of longer small RNAs (~25–30 nucleotides) associated with silencing of repetitive elements

Mammalian piRNAs can be divided into pre-pachytene and pachytene piRNAs, according to the stage of meiosis at which they are expressed in developing spermatocytes. like piRNAs in flies, pre-pachytene piRNAs predominantly correspond to repetitive sequences and are implicated in silencing transposons, such as l1 and intracisternal A-particle

miRNA as mRNA regulators

- > 3000 miRNAs have been identified in vertebrates, flies, worms, plants and viruses
- 1000 or more miRNAs are predicted to function in humans, regulating 30% of human genes
  - target mRNAs and biological function have been assigned to many miRNAs, including cell proliferation control (miR-181 and let-7); hematopoietic B-cell lineage fate (miR-181); B-cell survival (miR15a and miR16-1); brain patterning (miR430); pancreatic cell insulin secretion (miR375); adipocyte development (miR-143), and many other
  - expression of many miRNAs is specific to particular tissues or developmental stages
  - deregulation of miRNA important in pathology (cancer, metabolic diseases)

In Cancer:

- altered patterns of miRNA expression may affect cell cycle and survival programs
- germ-line and somatic mutations in miRNAs and polymorphisms in the target mRNAs contribute to cancer predisposition and progression

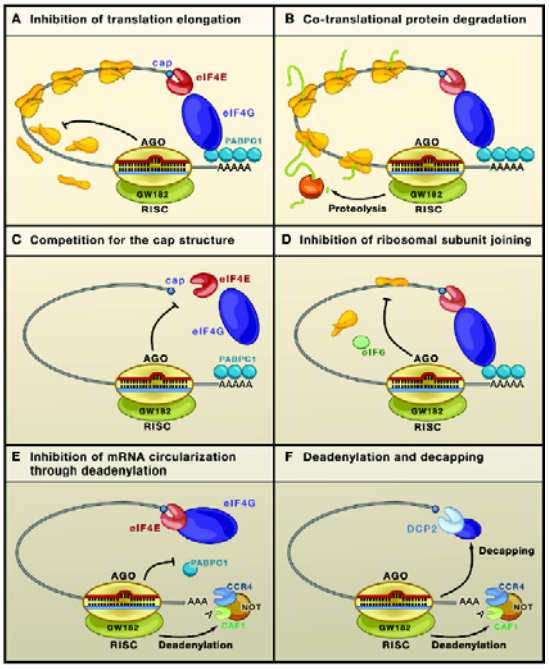


Figure 1. Mechanisms of miRNA-Mediated Gene Silencing

(A) Postinitiation mechanisms. MicroRNAs (miRNAs; red) repress translation of target mRNAs by blocking translation elongation or by promoting premature dissociation of ribosomes (ribosome drop-off).

(B) Cotranslational protein degradation. This model proposes that translation is not inhibited, but rather the nascent polypeptide chain is degraded cotranslationally. The putative protease is unknown.

(C–E) Initiation mechanisms. MicroRNAs interfere with a very early step of translation, prior to elongation. (C) Argonaute proteins compete with eIF4E for binding to the cap structure (cyan dot).

(D) Argonaute proteins recruit eIF6, which prevents the large ribosomal subunit from joining the small subunit.

(E) Argonaute proteins prevent the formation of the closed loop mRNA configuration by an ill-defined mechanism that includes deadenylation.

(F) MicroRNA-mediated mRNA decay. MicroRNAs trigger deadenylation and subsequent decapping of the mRNA target. Proteins required for this process are shown including components of the major deadenylase complex (CAF1, CCR4, and the NOT complex), the decapping enzyme DCP2, and several decapping activators (dark blue circles). (Note that mRNA decay could be an independent mechanism of silencing, or a consequence of translational repression, irrespective of whether repression occurs at the initiation or postinitiation levels of translation.) RISC is shown as a minimal complex including an Argonaute protein (yellow) and GW182 (green). The mRNA is represented in a closed loop configuration achieved through interactions between the cytoplasmic poly(A) binding protein (PABPC1; bound to the 3′ poly(A) tail) and eIF4G (bound to the cytoplasmic cap-binding protein eIF4E).

Since the “seeding” sequence is very short (6-8nt), potential targets for known miRNA can be identified in the 3’-UTR of hundreds of genes each.

However, experiment using either trasfection of miRNA in cultured cells or knock-down of endogenous miRNA function by “antagomir” , followed by gene expression profiling with microarrays, demonstrated that a limited number of targets exist for each miRNA, and that (in fewer cases) new unidentified targets may exist.

Many studies have shown that several context-dependent factors are important :

1) the number of miRNA targets / 3’UTR
2) cooperativity with different miRNA
3) position of the targets
4) Protein-binding sites

MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing

Andrew Grimson,[1,2,4,6] Kyle Kai-How Farh,[1,2,3,4,6] Wendy K. Johnston,[1,2,4] Philip Garrett-Engele,[5] Lee P. Lim,[5,*] and David P. Bartel[1,2,4,*]
[1]Howard Hughes Medical Institute
[2]Department of Biology
[3]Division of Health Sciences and Technology
Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]Whitehead Institute for Biomedical Research. 9 Cambridge Center. Cambridge.
[5]Rosetta Inpharmatics, 401 Terry Avenue N, Seattle, WA 98109, USA
[6]These authors contributed equally to this work.
*Correspondence: lee_lim@merck.com (L.P.L.), dbartel@wi.mit.edu (D.P.B.)
DOI 10.1016/j.molcel.2007.06.017

Articolo

Mammalian microRNAs (miRNAs) pair to 3′UTRs of mRNAs to direct their posttranscriptional repression. Important for target recognition are ~7 nt sites that match the seed region of the miRNA. However, these seed matches are not always sufficient for repression, indicating that other characteristics help specify targeting. By combining computational and experimental approaches, we uncovered five general features of site context that boost site efficacy: AU-rich nucleotide composition near the site, proximity to sites for coexpressed miRNAs (which leads to cooperative action), proximity to residues pairing to miRNA nucleotides 13–16, positioning within the 3′UTR at least 15 nt from the stop codon, and positioning away from the center of long UTRs. A model combining these context determinants quantitatively predicts site performance both for exogenously added miRNAs and for endogenous miRNA-message interactions. Because it predicts site efficacy without recourse to evolutionary conservation, the model also identifies effective nonconserved sites and siRNA off-targets.

five general features of site context that boost site efficacy:

1) AU-rich nucleotide composition near the site,
2) proximity to sites for coexpressed miRNAs (which leads to cooperative action),
3) proximity to residues pairing to miRNA nucleotides 13–16,
4) positioning within the 3'UTR at least 15 nt from the stop codon,
5) positioning away from the center of long UTRs.

A model combining these context determinants quantitatively predicts site performance both for exogenously added miRNAs and for endogenous miRNAmessage interactions.

Because it predicts site efficacy without recourse to evolutionary conservation, the model also identifies effective nonconserved sites and siRNA off-targets.

nature Vol 455|4 September 2008|doi:10.1038/nature07228

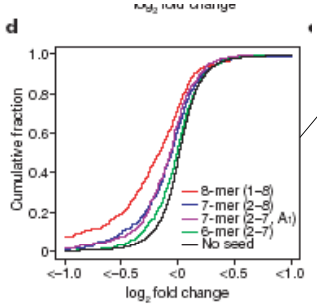
ARTICLES

Widespread changes in protein synthesis induced by microRNAs

Matthias Selbach[1], Björn Schwanhäusser[1]*, Nadine Thierfelder[1]*, Zhuo Fang[1], Raya Khanin[2] & Nikolaus Rajewsky[1]

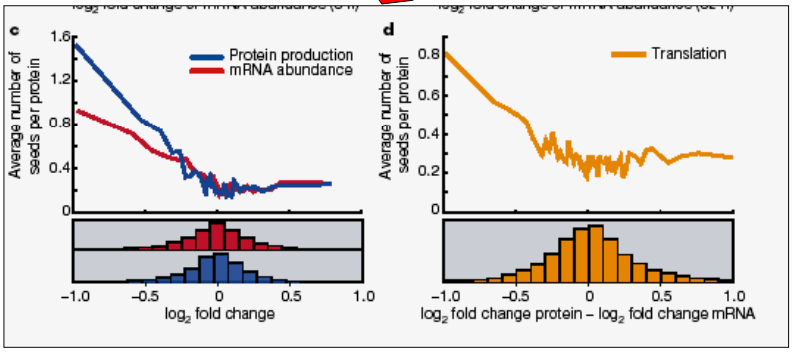
Animal microRNAs (miRNAs) regulate gene expression by inhibiting translation and/or by inducing degradation of target messenger RNAs. It is unknown how much translational control is exerted by miRNAs on a genome-wide scale. We used a new proteomic approach to measure changes in synthesis of several thousand proteins in response to miRNA transfection or endogenous miRNA knockdown. In parallel, we quantified mRNA levels using microarrays. Here we show that a single miRNA can repress the production of hundreds of proteins, but that this repression is typically relatively mild. A number of known features of the miRNA-binding site such as the seed sequence also govern repression of human protein synthesis, and we report additional target sequence characteristics. We demonstrate that, in addition to downregulating mRNA levels, miRNAs also directly repress translation of hundreds of genes. Finally, our data suggest that a miRNA can, by direct or indirect effects, tune protein synthesis from thousands of genes.

Articolo discusso nella lezione 10 (4-dic)



Goodness of seed

After miRNA, global levels of mRA measured on microarray. By comparing protein changes with mRNA level changes



miRNA target sequences alone, as discussed before for “decay” sequences, is not sufficient to “predict” the fate of a specific mRNA.

A more complex network, involving miRNA, RNA binding proteins (and perhaps endosiRNA), together with more classical proteins, is more likely in place to control the life of mRNA.

In fact, some proteins are being discovered that bind RNA close to miRNA sites and regulate either positively or negatively the function of miRNA.

One example is given by the Dnd1 protein (next slide)

Second example comes from the PUM1 PUM2 proteins (homologues of D. melanogaster Pumilio, a translational regulator of *nanos* mRNA in early development) (Galgano et al., 2008).

RNA-Binding Protein Dnd1 Inhibits MicroRNA Access to Target mRNA

Martijn Kedde,[1] Markus J. Strasser,[2] Bijan Boldajipour,[2] Joachim A.F. Oude Vrielink,[1] Krasimir Slanchev,[2,5] Carlos le Sage,[1] Remco Nagel,[1] P. Mathijs Voorhoeve,[1] Josyanne van Duijse,[1] Ulf Andersson Ørom,[3] Anders H. Lund,[3] Anastassis Perrakis,[4] Erez Raz,[2,*] and Reuven Agami[1,*]

[1]The Netherlands Cancer Institute, Division of Tumor Biology, Plesmanlaan 121, 1066CX, Amsterdam, The Netherlands
[2]Max-Planck-Institute for Biophysical Chemistry, Germ Cell Development, Am Fassberg 11, 37070 Goettingen, and Institute for Cell Biology, ZMBE, Center for Molecular Biology of Inflammation, University of Münster, Münster 48149, Germany
[3]Biotech Research and Innovation Centre, University of Copenhagen, Ole Maaloes Vej 5, 2200N, Copenhagen, Denmark
[4]The Netherlands Cancer Institute, Division of Molecular Carcinogenesis, Plesmanlaan 121, 1066CX, Amsterdam, The Netherlands
[5]Present address: Max-Planck-Institute for Immunology, Stuebeweg 51, 79108 Freiburg, Germany.

Cell *131*, 1273–1286, December 28, 2007 ©2007 Elsevier Inc.

SUMMARY

MicroRNAs (miRNAs) are inhibitors of gene expression capable of controlling processes in normal development and cancer. In mammals, miRNAs use a seed sequence of 6–8 nucleotides (nt) to associate with 30 untranslated regions (30UTRs) of mRNAs and inhibit their expression. Intriguingly, occasionally not only the miRNA-targeting site but also sequences in its vicinity are highly conserved throughout evolution. We therefore hypothesized that conserved regions in mRNAs may serve as docking platforms for modulators of miRNA activity. Here we demonstrate that the expression of dead end 1 (Dnd1), an evolutionary conserved RNA binding protein (RBP), counteracts the function of several miRNAs in human cells and in primordial germ cells of zebrafish by binding mRNAs and prohibiting miRNAs from associating with their target sites. These effects of Dnd1 are mediated through uridine-rich regions present in the miRNA-targeted mRNAs. Thus, our data unravel a novel role of Dnd1 in protecting certain mRNAs from miRNA-mediated repression.

OPEN ACCESS Freely available online PLoS one

Comparative Analysis of mRNA Targets for Human PUF-Family Proteins Suggests Extensive Interaction with the miRNA Regulatory System

Alessia Galgano[1], Michael Forrer[1], Lukasz Jaskiewicz[2], Alexander Kanitz[1], Mihaela Zavolan[2], André P. Gerber[1]*

1 Institute of Pharmaceutical Sciences, ETH Zurich, Zurich, Switzerland, 2 Biozentrum, University of Basel, Basel, Switzerland

Abstract

Genome-wide identification of mRNAs regulated by RNA-binding proteins is crucial to uncover post-transcriptional gene regulatory systems. The conserved PUF family RNA-binding proteins repress gene expression post-transcriptionally by binding to sequence elements in 3′-UTRs of mRNAs. Despite their well-studied implications for development and neurogenesis in metazoa, the mammalian PUF family members are only poorly characterized and mRNA targets are largely unknown. We have systematically identified the mRNAs associated with the two human PUF proteins, PUM1 and PUM2, by the recovery of endogenously formed ribonucleoprotein complexes and the analysis of associated RNAs with DNA microarrays. A largely overlapping set comprised of hundreds of mRNAs were reproducibly associated with the paralogous PUM proteins, many of them encoding functionally related proteins. A characteristic PUF-binding motif was highly enriched among PUM bound messages and validated with RNA pull-down experiments. Moreover, PUF motifs as well as surrounding sequences exhibit higher conservation in PUM bound messages as opposed to transcripts that were not found to be associated, suggesting that PUM function may be modulated by other factors that bind conserved elements. Strikingly, we found that PUF motifs are enriched around predicted miRNA binding sites and that high-confidence miRNA binding sites are significantly enriched in the 3′-UTRs of experimentally determined PUM1 and PUM2 targets, strongly suggesting an interaction of human PUM proteins with the miRNA regulatory system. Our work suggests extensive connections between the RBP and miRNA post-transcriptional regulatory systems and provides a framework for deciphering the molecular mechanism by which PUF proteins regulate their target mRNAs.

September 2008 | Volume 3 | Issue 9 | e3164

mRNA captured by RIP

Antibodies against PUM1 and PUM2 incubated with mRNA, then immunoprecipitated.

mRNA eluted, RT, labeled and hybridized to Stanford microarrays (cDNA spotted) containing probes for 26,000 transcripts

Bound RNA explored for:
1) function of the encoded protein (Gene Ontology);
2) the presence of a “PUM” recognition nucleotide motif;
3) the presence of other known motifs

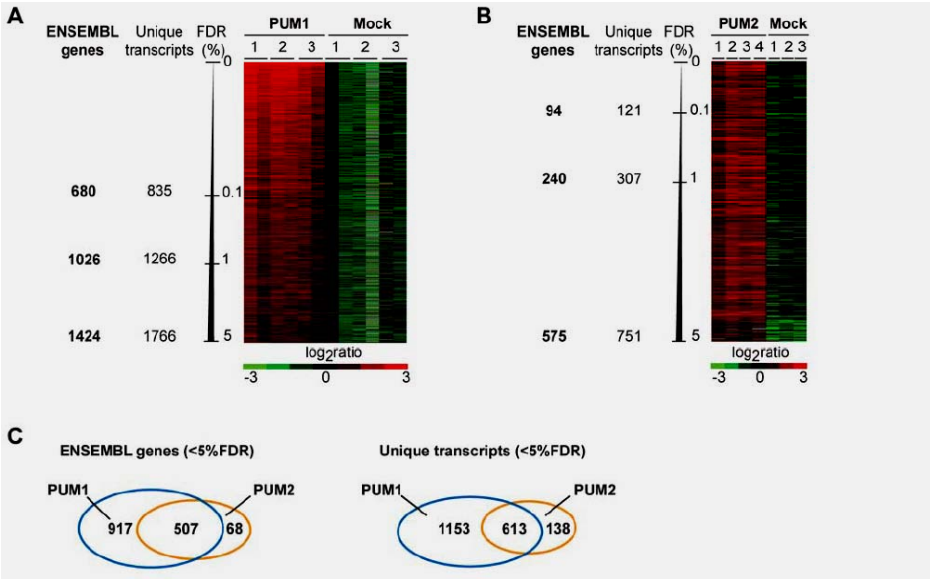


Figure 1. mRNAs specifically associated with human PUM proteins. Rows represent unique transcripts ordered according to increasing FDRs determined by SAM analysis. Columns represent individual experiments. The colour code indicates the degree of enrichment (green-red $\log_2$ ratio scale). (A) mRNAs associated with PUM1. Three experiments with PUM1 protein and three mock experiments both with dye-swap technical replicates are shown. (B) mRNAs associated with PUM2. Four experiments with PUM2 protein and three mock experiments are shown. (C). Venn diagram representing overlap between PUM1 and PUM2 targeted transcripts (right) and the corresponding genes (ENSEMBL, left).

Functional classes significantly associated with mRNAs bound by PUM1 and PUM2, ordere by the significance “p-value”

Table 1. Significantly shared PANTHER and GO annotations among PUM1 and PUM2 mRNA targets.

Category	Term	PUM1 *p*-value	PUM2 *p*-value
Pathway (PANTHER)	Angiogenesis	8×10^{-7}	
	Ras Pathway	1×10^{-6}	
	PDGF signaling pathway	3×10^{-4}	
	T cell activation	5×10^{-4}	
	p53 pathway	1×10^{-3}	9×10^{-3}
	Interleukin signaling pathway	1×10^{-2}	
	EGF receptor signaling pathway	1×10^{-2}	
	B cell activation	1×10^{-2}	
	Parkinson's disease		2×10^{-2}
Biological Process (PANTHER)	Nucleoside, nucleotide and nucleic acid metabolism	1×10^{-19}	1×10^{-6}
	Cell cycle	1×10^{-14}	9×10^{-7}
	mRNA transcription	3×10^{-13}	5×10^{-4}
	Protein phosphorylation	2×10^{-8}	3×10^{-2}
	Intracellular protein traffic	3×10^{-7}	6×10^{-3}
	Intracellular signaling cascade	6×10^{-6}	3×10^{-2}
	Cell proliferation and differentiation	5×10^{-5}	
	Developmental processes	7×10^{-5}	
	Oncogenesis	1×10^{-4}	
	DNA repair	4×10^{-4}	
	MAPKKK cascade	1×10^{-2}	

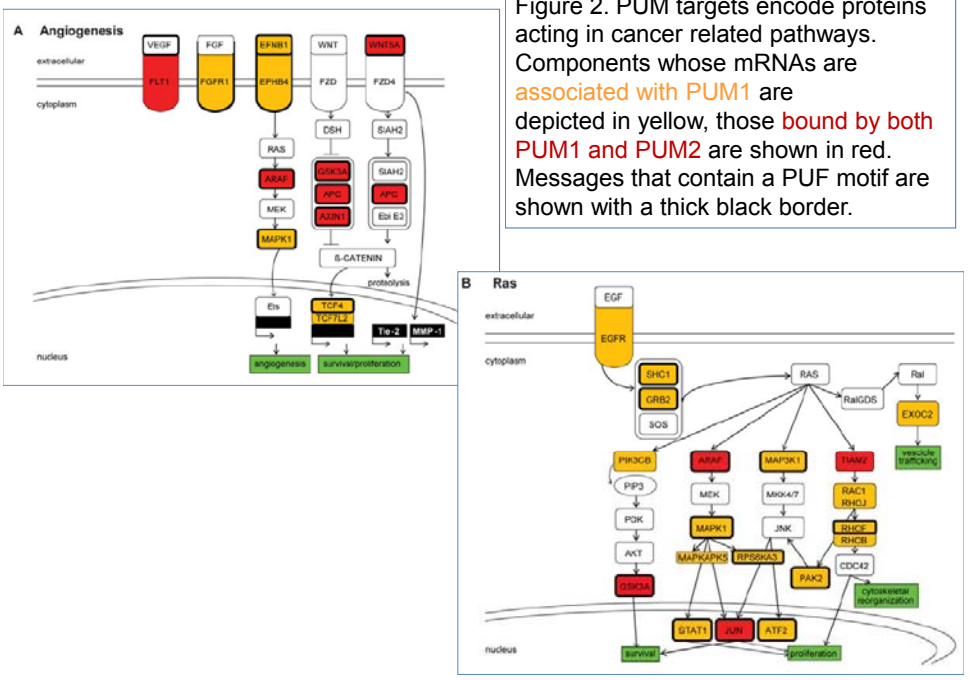


Figure 2. PUM targets encode proteins acting in cancer related pathways. Components whose mRNAs are associated with PUM1 are depicted in yellow, those bound by both PUM1 and PUM2 are shown in red. Messages that contain a PUF motif are shown with a thick black border.

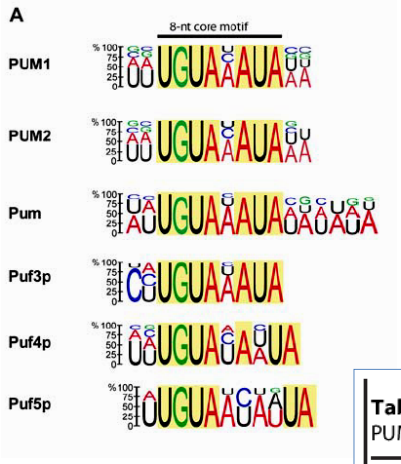


Figure 3. Analysis of an RNA consensus sequence associated with human PUM proteins. (A) PUF consensus motif in 39-UTR sequences associated with PUM1, PUM2, Drosophila Pum and yeast Puf3, Puf4 and Puf5 proteins [35,36]. Height of the letters indicates the probability of appearing at the position in the motif. Nucleotides with less than 10% appearance were omitted.

Table 2. Number of PUF consensus motifs found in human PUMILIO targets (FDR<5%).

	Search option	Sequences	Motifs (%)	*p*-value
PUM1	3′-UTR	1416	977 (69)	$<10^{-100}$
	CDS	1418	187 (13)	$<10^{-100}$
	5′-UTR	1390	24 (2)	5×10^{-2}
PUM2	3′-UTR	571	422 (74)	$<10^{-100}$
	CDS	572	73 (13)	6×10^{-16}
	5′-UTR	558	11 (2)	7×10^{-2}

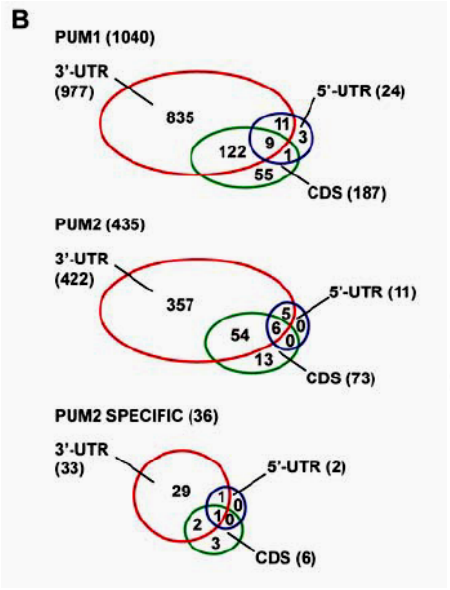
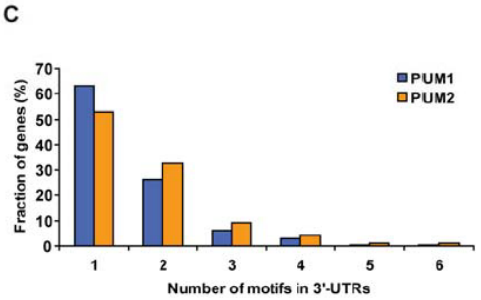


Figure 3. Analysis of an RNA consensus sequence associated with human PUM proteins.

(B) Distribution of PUF consensus motifs.

(C) Number of PUF motifs in the 3'-UTRs of PUM bound messages.



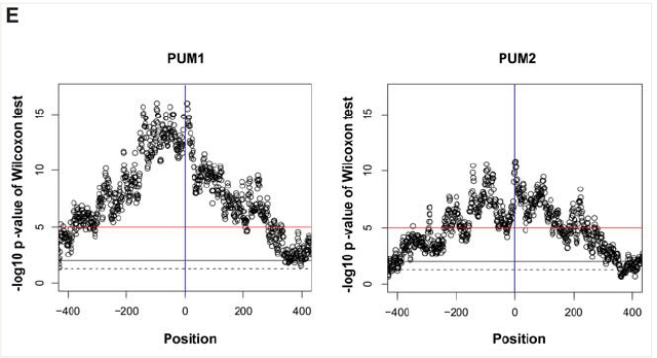


Figure 3. Analysis of an RNA consensus sequence associated with human PUM proteins.
(E) Analysis of PUF motif conservation among PUM1 and PUM2 targets. The x-axis shows the position (relative to the middle of the PUF motif), and the y-axis shows the logarithm (base 10) of the p-value from the Wilcoxon test determining whether conservation scores come from the same distribution for PUM targets and non-targets. The vertical blue line is drawn at position 0 corresponding to the PUF motif. The dashed black line is drawn at a p-value of 0.05, the continuous black line at a p-value of 0.01, and the red line at a p-value of 1025, which is the threshold for significance considering multiple testing.

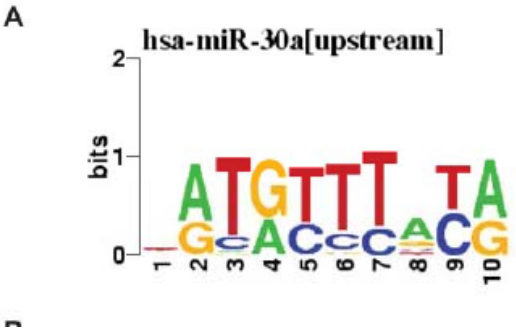


Figure 5. miRNA binding sites are enriched among human PUM targets.
(A) Example of a motif identified using the Phylogibbs motif finding algorithm in the vicinity (400 nucleotides upstream) of high-confidence miR-30a target sites. The x-axis indicates the position of a nucleotide in the inferred motif, and the y-axis gives the information score (bits) at that position. The height of each letter is proportional to the frequency of the respective nucleotide at that particular position in the alignment of inferred sites.

Work from other groups also has shown that binding of PUM1 may be essential to allow an adjacent miRNA site to work properly.

Regulation of mRNA stability derives from a comples compositionality of cis-acting sequences and trans-acting factors, involving both miRNA /Ago1, known mRNA decay regulators (AU-binding proteins), and new proteins as exemplified by Dnd1 and PUM proteins.