Models of microRNA-Target Coordination

Neil R. Smalheiser and Vetle I. Torvik

Psychiatric Institute

University of Illinois at Chicago,

Chicago, IL 60612

Corresponding Author:
Neil R. Smalheiser , M.D., Ph.D.
Assistant Professor in Psychiatry
UIC Psychiatric Institute M/C 912
1601 W. Taylor Street, room 285
Chicago, IL 60612
Phone: (312) 413-4581
E-mail: Nsmalheiser@psych.uic.edu

1. Introduction

The number of microRNAs appears to be ever-growing, as more intensive sequencing of small RNAs reveals a large population of sequences that are expressed at low abundance, or in a tissue- or stage-specific manner. Computational studies have also predicted the existence of thousands of candidate microRNA precursor hairpin structures throughout mammalian genomes (reviewed in Bentwich, 2005). The number of predicted potential targets per microRNA is also steadily increasing, with the recognition that binding of a 7-mer seed at the 5'-end of a microRNA may be sufficient to regulate a target mRNA functionally (Doench and Sharp, 2004; Farh et al., 2005; Stark et al., 2005; Lim et al., 2005; Sood et al., 2006). But how are microRNAs and their targets coordinated – if at all?

2. A Random Model

One recent paper proposes that microRNAs arise whenever a RNA hairpin structure happens to be transcribed, that happens to be competent for processing by Drosha and Dicer (Svoboda and Cara, 2006). Most of these microRNAs will have no function at all, at least not initially: They will bind to a relatively large number of putative target regions at random (a 7-mer sequence will bind randomly every 4^7 = 16, 384 bases on average), and those target regions that happen to be associated with a useful phenotypic response will tend

to be retained over evolutionary time whereas those mRNAs that show deleterious responses will become relatively depleted in target sequences (Svoboda and Cara, 2006). Such a random model predicts that no coordination exists initially between microRNAs and their targets, but that coordinated sets of microRNAs and targets emerge over time via natural selection.

3. Cases in which microRNAs and their Targets have an Intrinsic Relationship

On the other hand, at least 4 situations have been described in which microRNA precursors do have a close intrinsic relationship with their target sequences: **a)** In plants, several microRNA precursors arise from inverted duplicated repeat sequences, so that the microRNA naturally targets the same mRNA from whence it was derived (Allen et al., 2004). **b)** Several mammalian microRNAs are encoded on the opposite strand from their target mRNA, so that they comprise a perfect antisense match to the target (Smalheiser, 2003; Royo et al., 2006). **c)** Several human microRNAs have been proposed to arise from within processed pseudogenes, one of which arises from the antisense strand of a beta-5 tubulin gene and hence would be expected to target the normal cognate beta-5 tubulin mRNA (Devor, 2006). **d)** A set of 4 human microRNA precursors were described that derived entirely from MIR/LINE2 genomic repeats (Smalheiser and Torvik, 2005). In these cases, two different segments of MIR repeats had become apposed in opposite orientations along a chromosome, forming a hairpin structure that was transcribed and processed as microRNAs (fig. 1). These

microRNAs would be expected to target MIR/LINE2 repeats inserted into the 3'-UTR regions of mRNAs in the proper (opposite) orientation.

It is worth emphasizing that genomic repeats are well-represented among the sequences that are predicted to be good microRNA precursor hairpin candidates (Sewer et al., 2005). Indeed, repeat sequences are routinely observed during small RNA cloning projects, but most, if not all, have been interpreted as rasiRNAs arising from dsRNA (Aravin et al., 2003; Chen et al., 2005). Any microRNAs derived from genomic repeats that form precursor hairpins would have been excluded from miRBase or other databases because they do not map unambiguously to one or a few distinct sites within the genome. Thus, it is possible that the contribution of genomic repeats to microRNA gene evolution has been underestimated. It is also conceivable that some "conventional" microRNA precursors have evolved originally from this pool of potential microRNA precursor hairpins formed by genomic repeats, but that progressive changes occurring over evolution have obscured their origins. As well, transposable elements can contribute promoter sequences that can affect transcription of nearby genes (e.g., Nigumann et al., 2002; Jordan et al., 2003), possibly including microRNA precursors (Smalheiser, 2003). Thus, genomic repeats may influence the probability of a microRNA precursor hairpin being transcribed, even if the repeat sequences are not themselves incorporated within the microRNA precursor.

4. Cases in which microRNAs hit genomic repeats or regulatory elements within 3'-UTRs

Another way that microRNAs may interact with multiple mRNA targets in a coordinated fashion is by interacting with specific sequences within 3'-UTRs such as those that regulate mRNA transport, stability and polyadenylation. **a)** For example, a recent report proposes that miR-15 binds to AU-rich (ARE) elements and thereby mediates functional inhibition of translation (Jing et al., 2005). Messenger RNAs bearing ARE elements comprise one of several distinct functional classes that are regulated by sets of RNA-binding proteins (Barreau et al., 2006), hence they are functionally as well as structurally related to each other. Recently, AU-rich elements within 3'-UTRs have been proposed to be originally derived from insertions of Alu elements (An et al., 2004). **b)** As well, we have recently predicted that Alu elements in 3'-UTRs are likely microRNA targets for a set of 28 human microRNAs that share the 5'-seed sequence AAGUGC (Smalheiser and Torvik, manuscript submitted for publication). There is experimental evidence that mRNAs expressing Alu elements in their 3'-UTRs tend to be involved in cellular growth and differentiation, and are regulated as a class via translational control (Krichevsky et al., 1999; Vidal et al., 1993; Stuart et al., 2000; Vila et al., 2003; Spence et al., 2006); thus, microRNA targeting could potentially provide a mechanism for the translational control. **c)** Our previous analysis suggesting that certain human microRNAs may target MIR/LINE2 repeats within 3'-UTRs (Smalheiser and Torvik, 2005) constitutes another

example in which human microRNAs appear to target genomic repeats. From the standpoint of evolution, microRNAs that happen to hit upon genomic repeats or widely scattered elements should be particularly prominent (strongly favored or quickly lost) because they will hit an entire group of mRNAs that may already share certain functions or are already regulated in a coordinated manner.

5. Discussion

The ultimate origin of microRNA genes remains obscure in most cases. One study has shown that microRNA genes are not located randomly through the genome, but that about half are located at fragile chromosomal sites (Calin et al., 2004). Another study indicates that microRNA genes and (some of) their mRNA gene targets in *C. elegans* tend to be encoded near each other on the same chromosome (Inaoka et al., 2006). However, the meaning of these clues, in terms of microRNA gene biogenesis, is difficult to decipher at this point. Our studies have provided evidence that transposable elements have contributed both to microRNA precursor gene sequences and to their target sequences within 3'-UTRs. It is too early to assess whether this represents a minor phenomenon, or is indicative of a major role for these elements in shaping microRNA evolution and in coordinating microRNAs with their targets. We hope that this brief chapter will stimulate further research on this issue.

REFERENCES:

Allen, E., Xie, Z., Gustafson, A.M., Sung, G.H., Spatafora, J.W. and Carrington, J.C. (2004). Evolution of microRNA genes by inverted duplication of target gene sequences in Arabidopsis thaliana. *Nature Genetics,* **36**, 1282-1290.

An, H.J., Lee, D., Lee, K.H. and Bhak, J. (2004). The association of Alu repeats with the generation of potential AU-rich elements (ARE) at 3' untranslated regions. *BMC Genomics*, **5**, 97.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J. and Tuschl, T. (2003). The small RNA profile during *Drosophila melanogaster* development. *Developmental Cell*, **5**, 337-350.

Barreau, C., Paillard, L. and Osborn,e H.B. (2006). AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Research,* **33**, 7138-7150.

Bentwich, I. (2005). Prediction and validation of microRNAs and their targets. *Federation of the European Biochemical Society Letters,* **579**, 5904-5910.

Calin, G.A., Sevignani, C., Dumitru, C.D, Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M. and Croce, C.M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences USA,* **101**, 2999-3004.

Chen, P.Y., Manninga, H., Slanchev, K., Chien, M., Russo, J.J., Ju, J., Sheridan, R., John, B., Marks, D.S., Gaidatzis, D., Sander, C., Zavolan, M. and Tuschl,

T. (2005). The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes & Development,* **19**, 1288-1293.

Devor, E.J. (2006). Primate microRNAs miR-220 and miR-492 lie within processed pseudogenes. *J Hered*., **97**, 186-190.

Doench, J.G. and Sharp, P.A. (2004). Specificity of microRNA target selection in translational repression. *Genes & Development,* **18**, 504-511.

Farh, K.K., Grimson, A., Jan, C., Lewis, B.P., Johnston, W.K., Lim, L.P., Burge, C.B. and Bartel, D.P. (2005). The widespread impact of mammalian microRNAs on mRNA repression and evolution. *Science,* **310**, 1817-1821.

Inaoka, H., Fukuoka, Y. and Kohane, I.S. (2006). Lower expression of genes near microRNA in C. elegans germline. *BMC Bioinformatics*, **7**, 112.

Jing, Q., Huang, S., Guth, S., Zarubin, T., Motoyama, A., Chen, J., Di Padova, F., Lin, S.C., Gram, H.. and Han, J. (2005). Involvement of microRNA in AU-rich element-mediated mRNA instability. *Cell*, **120**, 623-634.

Jordan, I.K, Rogozin, I.B., Glazko, G.V. and Koonin, E.V. (2003). Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends in Genetics,* **19**, 68-72.

Krichevsky, A. M., Metzer, E. and Rosen, H. (1999). Translational control of specific genes during differentiation of HL-60 cells*. Journal of Biological Chemistry,* **274**, 14295-14305.

Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelte, J.M., Castle, J., Bartel, D.P., Linsley, P.S. and Johnson, J.M. (2005). Microarray analysis

shows that some microRNAs downregulate large numbers of target mRNAs. *Nature,* **433**, 769-773.

Nigumann, P., Redik, K., Matlik, K. and Speek, M. (2002). Many human genes are transcribed from the antisense promoter of L1 retrotransposon. *Genomics*, **79**, 628-634.

Royo, H., Bortolin, M.L., Seitz, H. and Cavaille, J. (2006). Small non-coding RNAs and genomic imprinting. Cytogenet Genome Research, **113**, 99-108.

Sewer, A., Paul, N., Landgraf, P., Aravin, A., Pfeffer, S., Brownstein, M.J., Tuschl, T., van Nimwegen, E. and Zavolan, M. (2005). Identification of clustered microRNAs using an ab initio prediction method. *BMC Bioinformatics*, **6**, 267.

Smalheiser, N. R. (2003). EST analyses predict the existence of a population of chimeric microRNA precursor-mRNA transcripts expressed in normal human and mouse tissues. *Genome Biology*, **4**, 403.

Smalheiser, N.R. and Torvik, V.I. (2005). Mammalian microRNAs derived from genomic repeats. *Trends in Genetics,* **21**, 322-326.

Sood, P., Krek, A., Zavolan, M., Macino, G. and Rajewsky, N. (2006). Cell-type-specific signatures of microRNAs on target mRNA expression. *Proceedings of the National Academy of Sciences USA,* **103**, 2746-2751.

Spence, J., Duggan, B.M., Eckhardt, C., McClelland, M. and Mercola, D. (2006). Messenger RNAs under differential translational control in Ki-ras-transformed cells. *Molecular Cancer Research,* **4**, 47-60.

Stark, A., Brennecke, J., Bushati, N., Russell, R.B. and Cohen, S.M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell,* **123**, 1133-1146.

Stuart, J.J., Egry, L.A., Wong, G.H. and Kaspar, R.L. (2000). The 3' UTR of human MnSOD mRNA hybridizes to a small cytoplasmic RNA and inhibits gene expression. *Biochemical Biophysical Research Communications,* **274**, 641-648.

Svoboda, P. and Cara, A.D. (2006). Hairpin RNA: a secondary structure of primary importance. *Cell Molecular Life Sciences,* Mar 29; [Epub ahead of print]

Vidal, F., Mougneau, E., Glaichenhaus, N., Vaigot, P., Darmon, M. and Cuzin, F. (1993). Coordinated posttranscriptional control of gene expression by modular elements including Alu-like repetitive sequences. *Proceedings of the National Academy of Sciences USA,* **90**, 208-212.

Vila, M.R., Gelpi, C., Nicolas, A., Morote, J., Schwartz, S. Jr., Schwartz, S. and Meseguer, A. (2003). Higher processing rates of Alu-containing sequences in kidney tumors and cell lines with overexpressed Alu-mRNAs. *Oncology Reports,* **10**, 1903-1909.

Figure 1. Genomic structure of human LINE-2 derived microRNA precursors.

Information was downloaded and edited from the UCSC Genome Browser. Each of the precursors resides within a mRNA intron, and each flanks the junction of two L2 repeats in opposite orientation (darker shading indicates less divergence from the L2 consensus).

(Reprinted from Smalheiser and Torvik, 2005 with permission.)