Genomics & Informatics Vol. 5(3) 107-112, September 2007

Statistical Analysis of Gene Expression in Innate Immune Responses: Dynamic Interactions between MicroRNA and Signaling Molecules

Vincent Piras[1#], Kumar Selvarajoo[1#], Naoki Fujikawa[1], Sangdun Choi[2], Masaru Tomita[1], Alessandro Giuliani[3] and Masa Tsuchiya[1*]

[1]Institute for Advanced Biosciences, Keio University, 14-1 Baba-cho, Tsuruoka, Yamagata 997-0035, Japan, [2]Department of Molecular Science and Technology, Ajou University, Suwon 443-749, Korea, [3]Istituto Superiore di Sanita', Environment and Health Department, Viale Regina Elena 299, 00161, Rome, Italy

Abstract

MicroRNAs (miRNAs) are known to negatively control protein-coding genes by binding to messenger RNA (mRNA) in the cytoplasm. In innate immunity, the role of miRNA gene silencing is largely unknown. In this study, we performed microarray-based experiments using lipopolysaccharide (LPS)-stimulated macrophages derived from wild-type, MyD88 knockout (KO), TRIF KO, and MyD88/TRIF double KO mice. We employed a statistical approach to determine the importance of the commonality and specificity of miRNA binding sites among groups of temporally co-regulated genes. We demonstrate that both commonality and specificity are irrelevant to define *a priori* groups of co-down-regulated genes. In addition, analyzing the various experimental conditions, we suggest that miRNA regulation may not only be a late-phase process (after transcription) but can also occur even early (1h) after stimulation in knockout conditions. This further indicates the existence of dynamic interactions between miRNA and signaling molecules/transcription factor regulation; this is another proof for the need of shifting from a 'hard-wired' paradigm of gene regulation to a dynamical one in which the gene co-regulation is established on a case-by-case basis.

Keywords: TLR4 innate immunity, co-regulated genes, miroRNA, dynamic regulation

*Corresponding author: E-mail tsuchiya@ttck.keio.ac.jp
Tel +81-235-29-0829, Fax +81-235-29-0829
Accepted 6 August 2007
#These authors contributed equally to this work.

Introduction

Recognition of lipopolysaccharide (LPS) by Toll-like receptor (TLR) 4 of macrophages triggers the innate immune response by activating transcription factors such as NF-kB and AP-I, resulting in the induction of large numbers of genes, predominantly those that encode proinflammatory cytokines (Hirotani *et al.*, 2005). Although transcriptional activation and resultant gene expression studies have been performed in the innate immune system, the discovery of microRNAs (miRNA) adds a layer of complexity for the understanding of overall biological regulation.

miRNA are approximately 20-25-nucleotide, small RNAs that bind to the 3'-untranslated region (3'-UTR) of messenger RNA (mRNA) in the cytoplasm, and either negatively regulate the expression of the mRNA or repress translation into the corresponding protein (Bartel, 2004). Extensive studies on miRNAs during the last few years (Lai, 2002; Bagga *et al.*,2005; Lim, 2005; Baskerville *et al.*, 2005; Taganov *et al.*, 2006; Sood *et al.*, 2006) have made possible the discovery of hundreds of miRNAs. This has resulted in the development of databases of miRNA target predictions in 3'-UTR regions of mRNAs (Lewis *et al.*, 2003; Lai, 2004; Krek *et al.*, 2005; Bentwich *et al.*, 2005; Grün *et al.*, 2005; Rajewsky, 2006), facilitating high-throughput analysis, such as the design of gene regulatory networks in terms of shared miRNAs (Shalgi *et al.*, 2007; Tsang *et al.*, 2007).

One of the chief goals in analyzing innate immune signaling is to discover co-regulated genes using mRNA expression data. It has been mostly assumed that similar patterns in temporal mRNA expression profiles usually suggest possessing common transcription factors (Yeo *et al.*, 2007). We performed a similar analysis on the basis of post-transcriptional miRNA regulation of co-regulated genes. If miRNA-based regulation has a major role in the establishment of gene regulatory circuits, we expect the temporally co-regulated genes to share a higher proportion of common miRNA binding sites than we would expect by chance. To test this hypothesis, we focused on the temporal mRNA expression profiles of wild-type macrophages during LPS stimulation, paying particular attention to the profiles of downregulated genes, since miRNAs negatively regulate

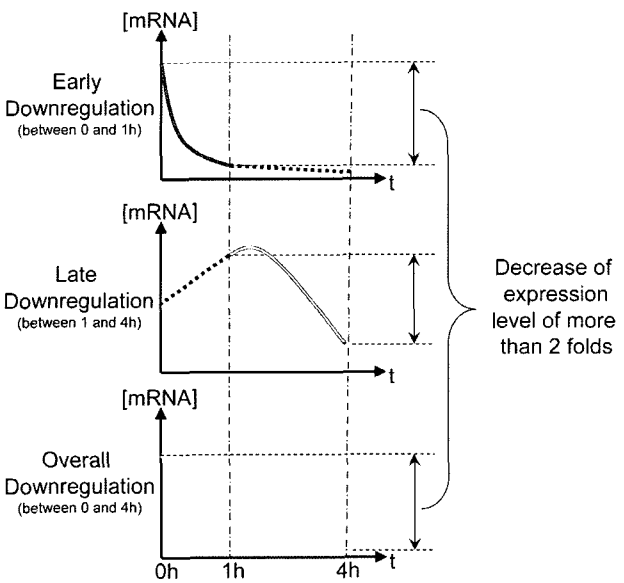

Fig. 1. The three profiles used to define downregulated genes. This classification allows us to distinguish various dynamic downregulation profiles by miRNAs. Classification is done according to time after stimulation; 1) 0-1 h is considered early downregulation profile (EDR), 2) 1-4 h, late downregulation profile (LDR) and 3) 0-4 h, overall downregulation profile (ODR). The downregulated genes show a 2-fold or more decrease in expression level for all profiles.

mRNA expression (Fig. 1) (Lim *et al.*, 2005).

We developed two criteria to check the validity of this hypothesis: 1) the number of common miRNA binding sites shared by a group of temporally co-regulated genes; ie, the commonality of miRNA binding sites, and 2) the number of genes that are targeted by a specific miRNA in a group of temporally co-regulated genes; ie, the specificity of miRNA binding sites. We performed statistical analyses by comparing the commonality and specificity scores in groups of co-downregulated genes with the same measures relative to groups of randomly-chosen genes. We showed that both commonality and specificity are low and are not especially higher for groups of co-downregulated genes when compared with groups of random genes. This result indicates that both commonality and specificity are not significant features to define groups of *a priori* co-downregulated genes, and that miRNA regulation is not a hard-wired counterpart of effective functional regulation circuits in cells and can adapt dynamically to external stimuli. Further analysis of single knockout conditions revealed that miRNA regulation may be involved not only at late phases (after transcription) but also at early times. Our result suggests that 1) miRNA regulation in wild-type conditions is repressed at early times in order not to interfere with the response to stimuli, and 2) the interplay between miRNA and signaling molecules/transcription factor regulation exists from early signaling processes.

Materials and Methods

Experiments and Microarray dataset

We utilized microarray data obtained from time-series experiments (0, 1, and 4 h) performed on peritoneal macrophages from wild-type, MyD88$^{-/-}$, TRIF$^{-/-}$, and MyD88$^{-/-}$TRIF$^{-/-}$ mice treated with 100 ng/ml LPS (*Salmonella minnesota* Re595, Sigma). Affymetrix mouse expression array A430 microarray chips were used for gene expression detection. The microarray dataset obtained from these experiments contains expression levels for 22690 Affymetrix probe set IDs, for a total of 12 expression level measurements. The experimental details can be found in Hirotani *et al.*, 2005.

Databases

We used the Pictar database (http://pictar.bio.nyu.edu) for target predictions of mouse miRNAs based on conservation in mammals (human, chimpanzee, mouse, rat, and dog) (Baskerville *et al.*, 2005). The Pictar database uses algorithms in sequences and evolutionarily-conserved data among vertebrate miRNA target predictions to determine to their reliability. The initial dataset of miRNAs contains 264 different miRNAs.

Selection of initial list of genes

Refseq RNA IDs are used as genes identifiers. As one Refseq RNA ID can correspond to one or more Affymetrix IDs, the expression level of one gene is the highest Affymetrix expression row of Affymetrix expressions in microarray dataset corresponding to the same Refseq RNA IDs. Only genes having at least one miRNA binding site were considered. Unlike previous studies (Hirotani *et al.*, 2005), in which only genes having a signaling intensity of more than 100 points in expression level were considered, in our study we are also considering expression changes for genes showing relatively low expression levels, as the previous threshold may hide many important processes. As a result, we obtained a list of 2969 genes for the analysis.

Co-downregulated genes

As miRNAs are known to be negative regulators of mRNA (Lim *et al.*, 2005), for our study, we focused on genes showing a decrease in temporal expression levels characterized by one of the following three expression profiles: 1) early downregulation profile (EDR) defined by a 2-fold decrease in expression level between 0 and 1 h, 2) late downregulation profile (LDR) between 1 and 4 h,

and 3) overall downregulation profile (ODR) between 0 and 4 h (Fig. 1) for each experimental condition (wild-type, MyD88 KO, TRIF KO, or double KO). For each of these profiles (EDR, LDR, and ODR), we selected gene pairs exhaustively and calculated their temporal expression Pearson correlation coefficient, taking the *n* genes (*n* = 5, 10, 15,...40) that have the highest correlations. Among the identified group of genes, we eliminated repeating groups, that is, groups that contain more than 80% the same genes.

To check whether co-regulated genes share a higher proportion of common miRNA binding sites, we defined commonality and specificity *scores:*

Commonality score

For each group of *n* genes, we can extract *p* pairs of genes:

$$p = \frac{n(n-1)}{2} \quad (1)$$

For each pair of genes, we computed the commonality score by taking the ratio of the number of miRNA binding sites shared by both genes to the total number of miRNA binding sites in both genes:

$$Commonality_{pair} = \frac{\text{No. of miRNA binding sites shared by both genes}}{\text{Total no. of miRNA binding sites in both genes}} \quad (2)$$

By averaging the ratios of all the pairs, we defined the average miRNA binding sites sharing ratio for the group:

$$Commonality_{group} = \frac{1}{p}\Sigma_{i=1}{}^{p} \quad (3)$$

Specificity score

It is possible that commonality *scores* can be low for genes that possess a large number of miRNA binding sites. To avoid this case, we defined a measure for the sharing of specific miRNA binding sites in co-regulated gene pairs, called the *specificity score.*

For each group of *n* genes, we calculated the specificity *score* for each individual miRNA binding site:

$$Specificity_{bindingsite} = \frac{\text{No. of genes in the grouphaving specific miRNA binding site}}{n} \quad (4)$$

We next defined the Maximum miRNA Binding Site Specificity Score (MMBSS) to represent group score, which is the specificity score of the most highly-represented miRNA binding sites among a group of genes:

$$Specificity_{group} = \max(specificity_{binding\ site}) \quad (5)$$

Statistical Significance

To check for eventual statistically significant differences between randomly chosen gene groups (control) made by all the couples from *n* genes (*n* = 5 to 40) and test groups (eg, EDR) with the same numerosity, we carried out commonality and specificity computations and checked for 95% confidence interval. To determine significant differences between the two populations, we further performed *t*-test on the populations made by control and test groups.

Results and Discussion

We evaluated commonality and specificity scores for the three profiles of early downregulated genes (EDR), late downregulated genes (LDR), and overall downregulated genes (ODR), and found surprisingly both commonality and specificity scores were very low. The commonality score was approximately 2% and the specificity score was on average about 25% (since we are using only one miRNA with the highest MMBSS, 25% can be considered very low). These scores are similar to those relative to randomly selected genes (Fig. 2A and Fig. 3). This result indicates that the sharing of miRNA binding sites (static view) is not a suitable criterion to identify dynamically co-regulated genes. We also observe, from Fig. 2A, that the commonality score of EDR, regardless of the number of genes in each group, is always lower than randomly-selected genes, and scores of LDR and ODR are always higher. In randomly-selected genes, we expect the group to contain a mixture of both miRNA-regulated and -unregulated genes. Hence, focusing on the lower commonality scores of EDR compared with randomly-selected genes, this could be due to two possible scenarios: 1) miRNA regulation is not active at early time points (0-1 h), or 2) miRNAs in the EDR group of genes are suppressed by signaling molecules such as transcription factors (Fig. 4).

In order to further evaluate these two hypotheses, we investigated the commonality score for other conditions (TRIF KO, MyD88 KO, MyD88/TRIF DKO conditions, Fig. 2). If miRNA does not possess early regulation (0-1 h), the commonality score of EDR is not expected to change. However, looking firstly at the TRIF KO (Fig. 2B), we notice that the commonality score of genes showing an EDR profile was increased. In addition, the commonality score for LDR was reduced, showing the reverse trend from wild-type conditions. This could imply that miRNA regulation has increased at early times and decreased at later times, suggesting that under wild-type conditions, TRIF-dependent

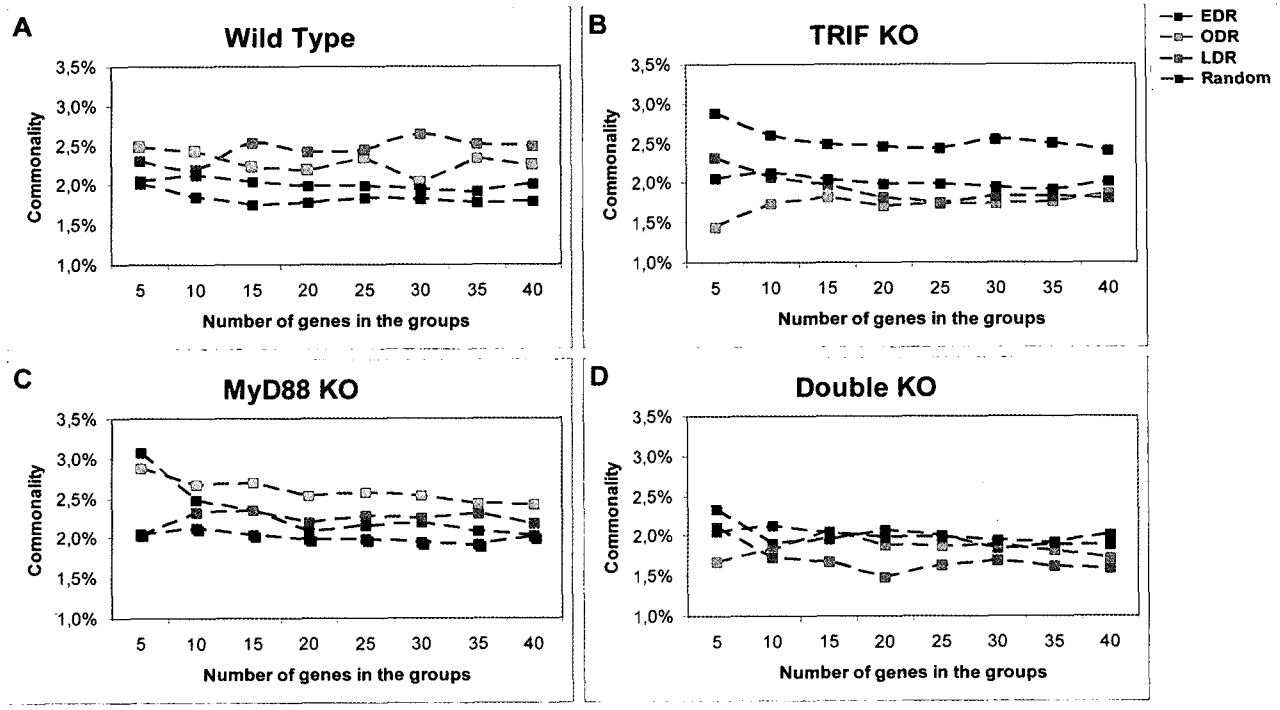

Fig. 2. Commonality scores for all profiles, in wild-type and KO conditions, for various group sizes of co-regulated genes. Each curve represents the average commonality score of groups of *n* co-downregulated genes and groups of *n* randomly-chosen genes (*n* varying from 5 to 40) for wild-type, MyD88 KO, TRIF KO, and MyD88/TRIF DKO conditions. Downregulated genes show at least a 2 fold decreasein expression level. Random population is constituted of 100 groups of randomly-selected genes. Groups of genes share less than 80% common genes.

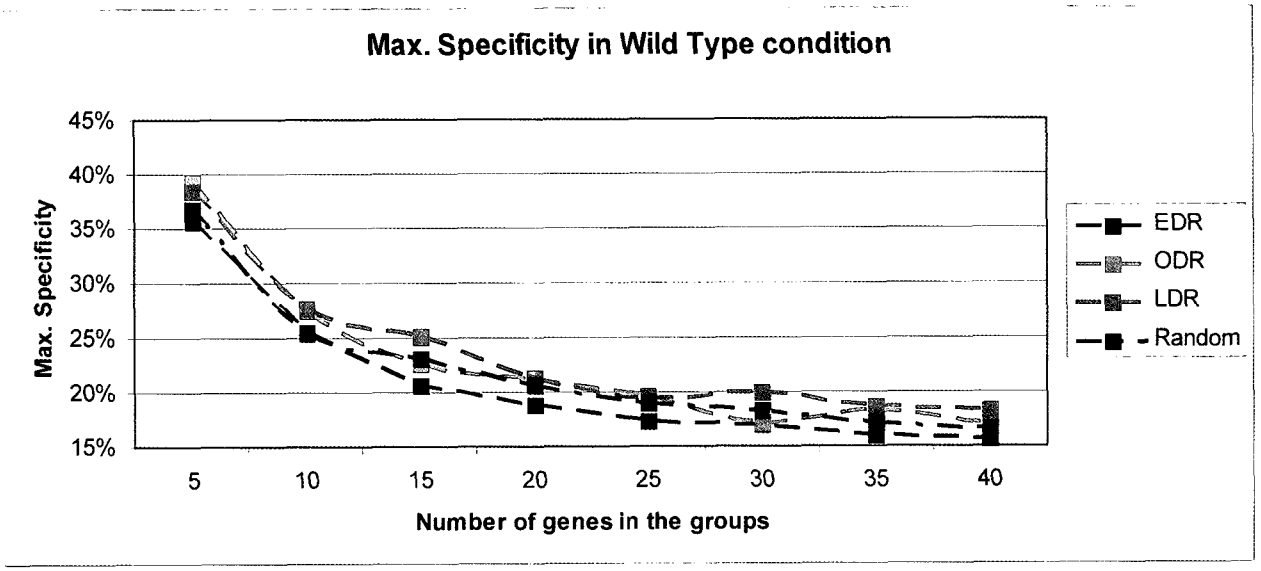

Fig. 3. Maximum specificity scores for all profiles in wild-type conditions for various group sizes of co-regulated genes. Each curve represents the average maximum specificity score of groups of *n* co-downregulated genes and groups of n randomly-chosen genes, for *n* varying from 5 to 40, for wild-type conditions. The same settings (80% overlap threshold, 2-fold decrease for co-downregulated genes, population of 100 groups for random genes) are used here, as in Fig. 2.

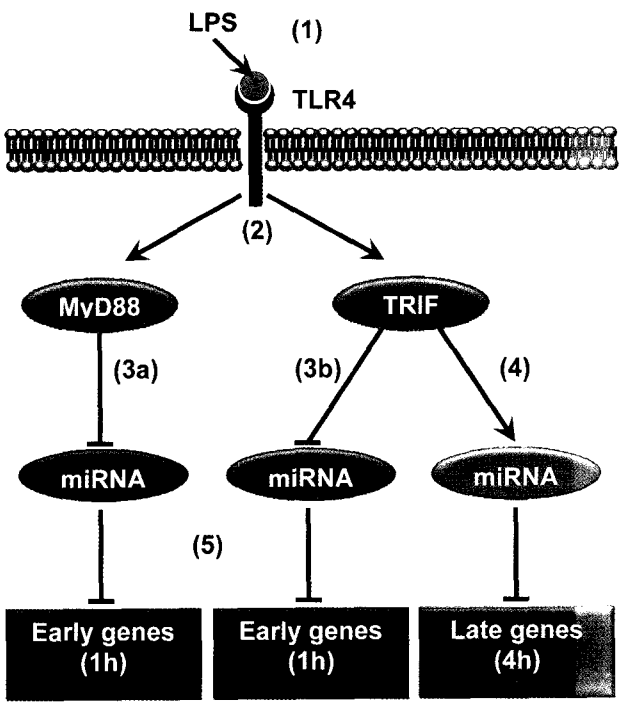

Fig. 4. Schematic representation of the possible interactions between MyD88 /TRIF-dependent pathways and miRNAs. In innate immune responses, Toll-like Receptor 4 (TLR4) is activated by LPS (1), and triggers MyD88- and TRIF-dependent signaling pathways (2), which in turn activate transcription factors that control gene expression. These transcription factors may also repress miRNA genes at early time points (3a, 3b), and activate miRNA genes later time points (4), to dynamically control miRNA-mediated gene silencing (5).

pathways are 1) suppressing early miRNA regulation (0-1 h), and 2) activating late miRNA regulation (Fig. 4).

For MyD88 KO conditions, we observed, compared with wild-type conditions, that the commonality scores showed that EDR profiles are similar to that of LDR profiles (Fig. 2C). The increase of EDR and ODR compared with random genes at early time points for MyD88 KO conditions suggests that MyD88-dependent pathways are repressing miRNA genes at early times. Analyzing MyD88/TRIF DKO conditions confirms that disabling both pathways removes miRNA regulation, since EDR is similar to random genes, and LDR and ODR profiles are lower compared with random genes (Fig. 2D).

Conclusion

We have shown that miRNA sharing among temporally co-regulated genes is low, indicating that miRNA regulatory processes are not hard-wired. Our result demonstrates that approaches that use static miRNA binding sites to suggest miRNA regulatory processes is clearly not sufficient. Moreover, analyzing various experimental conditions in innate immunity, we suggest that miRNA regulation may not only be involved at late phases (after transcription) but also at early times upon cellular stimulation. This last point may suggest the existence of dynamic interactions between miRNA and signaling molecules/transcription factor regulation and will be investigated next (Piras *et al*, in preparation).

Acknowledgments

We thank Mohamed Helmy for technical support. This work was supported by Japan Science and Technology Agency/Core Research for Evolutional Science and Technology JST CREST and the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), and Research fund by Yamagata Prefecture and Tsuruoka City.

References

Hirotani. T., Yamamoto, M., Kumagai, Y., Uematsu, S., Kawase, I., Takeuchi, O., and Akira, S. (2005). Regulation of lipopolysaccharide-inducible genes by MyD88 and Toll/IL-1 domain containing adaptor inducing IFN-beta. *Biochem Biophys Res Commun.* 328, 383-392.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281-297.

Taganov, K.D., Boldin, M.P., Chang, K.J., and Baltimore, D. (2006). NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc. Natl. Acad. Sci. USA* 103, 12481-12486.

Sood, P., Krek, A., Zavolan, M., Macino, G., and Rajewsky, N. (2006). Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc. Natl. Acad. Sci. USA* 103, 2746-2751.

Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* 122, 553-563.

Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, 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.

Lai, E.C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat. Genet.* 30, 363-364.

Baskerville, S. and Bartel, D.P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with

neighboring miRNAs and host genes. *RNA* 11, 241-247.

Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., and Rajewsky, N. (2005). microRNA target predictions in vertebrates. *Nat Genet.* 37, 495-500.

Lewis, B.P., Shih, I.H, Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. *Cell* 115, 787-798.

Lai, E.C. (2004). Predicting and validating microRNA targets. *Genome Biol.* 5, 115.

Rajewsky, N. (2006). microRNA target predictions in animals. *Nat. Genet.* 38, Suppl:S8-13.

Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E., Sharon, E., Spector, Y., and Bentwich, Z. (2005). Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet.* 37, 766-770.

Grün, D., Wang, Y.L., Langenberger, D., Gunsalus, K.C., and Rajewsky, N. (2005). microRNA target predictions across seven Drosophila species and comparison to mammalian targets. *PLoS Compu.t Biol.* 1, e13.

Shalgi, R., Lieber, D., Oren, M., and Pilpel, Y. (2007). Global and Local Architecture of the Mammalian microRNA-Transcription Factor Regulatory Network. *PLoS Comput Biol.* 3, e131.

Tsang, J., Zhu, J., and van Oudenaarden, A. (2007). MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol Cell.* 26, 753-767.

Yeo, Z.X., Wong, S.T., Arjunan, S.N.V., Piras, V., Tomita, M., Selvarajoo, K., Giuliani, A., and Tsuchiya, M. (2007). Sequential logic model deciphers dynamic transcriptional control of gene expressions. *PLoS One* 2, e776.