Global and Local Competition between ExogenouslyIntroduced microRNAs and Endogenously Expressed microRNAs

Doyeon Kim¹,², Jongkyu Kim¹,², and Daehyun Baek¹,²,³,*

It has been reported that exogenously introduced microRNA (exo-miRNA) competes with endogenously expressed miRNAs (endo-miRNAs) in human cells, resulting in a detectable upregulation of mRNAs with endo-miRNA target sites (TSs). However, the detailed mechanisms of the competition between exo- and endo-miRNAs remain uninvestigated. In this study, using 74 microarrays that monitored the whole-transcriptome response after introducing miRNAs or siRNAs into HeLa cells, we systematically examined the derepression of mRNAs with exo-and/or endo-miRNA TSs.

We quantitatively assessed the effect of the number of endo-miRNA TSs on the degree of mRNA derepression. As a result, we observed that the number of endo-miRNA TSs was significantly associated with the degree of derepression, supporting that the derepression resulted from the competition between exo- and endo-miRNAs. However, when we examined whether the site proficiency of exo-miRNA TSs could also influence mRNA derepression, to our surprise, we discovered a strong positive correlation. Our analysis indicates that site efficiencies of both exo- and endo-miRNA TSs are important determinants for the degree of mRNA derepression, implying that the derepression of mRNAs in response to exo-miRNA is more complex than that currently perceived.

Our observations may lead to a more complete understanding of the detailed mechanisms of the competition between exo- and endo-miRNAs and to a more accurate prediction of miRNA targets. Our analysis also suggests an interesting hypothesis that long 3′-UTRs may function as molecular buffer against gene expression regulation by individual miRNAs.

INTRODUCTION

MicroRNA (miRNA) is an approximately 22-nucleotide-long single-stranded RNA that regulates its target gene expression via the RNA-induced silencing complex (RISC). RISC consists of miRNA-loaded Argonaute and other proteins, which recognize specific targets by forming complementary base-pairing with mRNA targets and mediate mRNA destabilization and/or translational repression (Baek et al., 2008; Huttner and Si-mard, 2008; Pratt and MacRae, 2009). Although miRNA mainly downregulates its target gene expression, recent studies have reported that many genes are also upregulated in response to exogenously introduced miRNAs (exo-miRNA) and siRNAs (Castanotto et al., 2007; Khan et al., 2009). These upregulated mRNAs include target sites (TSs) for endogenously expressed miRNAs (endo-miRNAs), and the competition between exo-and endo-miRNAs was proposed as a model that may explain the observed pattern of mRNA derepression (Castanotto et al., 2007; Khan et al., 2009).

However, the previous model did not consider the potential interplay between exo- and endo-miRNAs. Here, we hypothesized that exo-miRNA TSs as well as endo-miRNA TSs may play a role in determining the degree of derepression, since both of their regulatory activities are mediated by the same RISC machinery. The potential coupling between exo- and endo-miRNAs has not been carefully studied. In this study, we used 74 publicly available microarrays that measured the whole-transcriptome response after introducing miRNAs or siRNAs into HeLa cells to investigate the competition between exo- and endo-miRNAs.

Based on the observations that we have made using the large dataset of microarrays, we propose a new model that may better explain the complicated aspects of the coupled competition. Our new model for the competition between exo- and endo-miRNAs may be useful for the development of a miRNA target prediction tool with higher accuracy. We also report an interesting hypothesis that long 3′-UTRs might serve as effective molecular buffer that resists gene expression regulation by individual miRNAs.

MATERIALS AND METHODS

Reference mRNAs
To obtain a set of unique cDNAs, a database of human mRNA
Global and Local Competition between Endo- and Exo-miRNAs
Doyeon Kim et al.

sequences was constructed using the RefSeq database (Pruitt et al., 2000) and the hg19 human genome (Lander et al., 2001) from the UCSC Genome Browser (http://genome.ucsc.edu). The processing and filtering methods from a previous study by Baek et al. were applied in order to obtain the representative isoform of each miRNA (Baek et al., 2008). Briefly, these filtering methods include discarding any miRNAs that contains the wrong open reading frames (ORFs) or that are candidates for nonsense-mediated decay. These non-redundant cDNAs were used as the reference cDNAs in the analysis.

Analysis of microarrays
To observe the change in mRNA levels on a large scale, the published data of a previous study with 74 microarrays was used in our analysis (Anderson et al., 2008; Birmingham et al., 2006; Garcia et al., 2011; Grimson et al., 2007; Jackson et al., 2006a; 2006b; Lim et al., 2005; Schwarz et al., 2006). These array data were generated to measure mRNA level changes after exogenous miRNA or siRNA was transfected (Lander et al., 2001) into human HeLa cells. The changes in mRNA expression level of our reference cDNAs were assigned according to the measured values depicted on the log2 scale from the previous study (Garcia et al., 2011). Then, the association between mRNA expression level changes and the characteristics in the 3'-UTR such as the miRNA TSs or its length were analyzed.

Statistical analysis and identification of target sites
In order to measure the number of endo-miRNA TSs from each mRNA, ten highly expressed miRNA in the human HeLa cell line were selected from the published endo-miRNA expression profile data (Landgraf et al., 2007). The number of endo-miRNA TSs was obtained by counting the number of 8mer, 7mer-m8, 7mer-A1, and 6mer sites, respectively on each 3'-UTR. mRNAs without 6-8mer sites were regarded as "no site" miRNAs. For regression analyses, the simple linear regression was performed using the 'lm()' command in R package (http://www.r-project.org). To compare the slopes of trend lines from different mRNA subsets, ANCOVA (analysis of covariance) was performed using the 'anova()' command in R package.

Target proficiency analysis
To investigate the effect of exo-miRNA TSs on the mRNA derepression, miRNAs with ≥1 exo-miRNA 8mer, 7mer-m8, or 7mer-A1 sites in the 3'-UTR were chosen and analyzed. After measuring the proficiency of exo-miRNA 7-8mer sites with 'context+ score' (Garcia et al., 2011), miRNAs were sorted based on their 3'-UTR length and compared with the predicted proficiency of two miRNAs, whose 3'-UTR lengths were close (length difference of < 50 nucleotides). Then, these were separated into two groups depending on their relative proficiency of exo-miRNA 7-8mer sites. In order to ensure that the two groups were not influenced by other factors, other than exo-miRNA TS proficiency, the P values for the 3'-UTR length, number of endo-miRNA TSs, and site proficiency of endo-miRNA 7-8mer sites were calculated with Wilcoxon's rank-sum test using 'scipy' package in Python (http://www.scipy.org). The data were fitted with a simple linear regression and the slopes of trend lines were compared with ANCOVA.

Analysis of derepression in mRNAs without exo- or endo-miRNA TSs
To investigate whether the degree of derepression changes based on the different criteria for the selection of endo-miRNAs, we analyzed miRNAs without exo-miRNA 6-8mer sites in the 3'-UTR. From the published endo-miRNA expression profile data in HeLa cells (Landgraf et al., 2007), the 10 most highly expressed endo-miRNAs were selected and a subset of miRNAs was made by discarding any miRNAs that contained 6-8mer sites for the 10 endo-miRNAs in the 3'-UTR. The analysis was expanded to include the top 20 and the top 30 highly endo-miRNAs as well in order to create a more robust set of miRNAs without exo- or endo-miRNA TSs. The correlation between the fold-change and the 3'-UTR length on a log scale was observed and they were graphically depicted with the mean fold-change and the log scaled 3'-UTR length of 15,000 miRNAs at each point. Then the data were fitted with the simple linear regression.

RESULTS
The target site proficiency of endogenously expressed miRNAs influences mRNA derepression
Using a large dataset of 74 microarrays that monitored the whole-transcriptome response after introducing miRNAs or siRNAs into the HeLa cell line, we systematically examined the derepression of miRNAs that have 3'-UTR TSs for the endo-miRNAs. According to previous studies (Castanotto et al., 2007; Khan et al., 2009), exo-miRNA or siRNA competes with endo-miRNAs and consequently miRNAs that have endo-miRNA TSs are detectably upregulated. Consistent with the previous reports, we observed that miRNAs with 3'-UTR 7-8mer sites for endo-miRNAs exhibited a significant upregulation (Fig. 1A).

We hypothesized that if the observed upregulation or derepression was due to the competition between exo- and endo-miRNAs, the number of TSs should be associated with the degree of derepression. Thus, to quantitatively assess the effect of the number of 3'-UTR TSs for endo-miRNAs on the degree of derepression, the correlation between them was measured. Since many species of endo-miRNAs are expressed in HeLa cells, it is difficult to identify miRNAs without endo-miRNA TSs. While a reasonable way to tackle this problem would be to take highly expressed endo-miRNAs into account, it is uncertain to determine how many endo-miRNAs should be considered. Regardless of the number of endo-miRNAs that we chose, we noticed a strong positive correlation between the number of endo-miRNA TSs and their 3'-UTR lengths (Fig. 1B), implying that the 3'-UTR length can serve as a good predictor for the number of endo-miRNA TSs. As a result, we found that the degree of derepression was significantly associated with both the number of endo-miRNA TSs and the 3'-UTR length (Figs. 1C and 1D).

To further test whether the observed derepression was influenced by the presence or absence of exo-miRNA 3'-UTR TSs, we collected a subset of miRNAs without exo-miRNA 3'-UTR 6-8mer sites and measured the correlation between the degree of derepression and the number of endo-miRNA 6-8mer sites. These miRNAs exhibited a clear upregulation in comparison to the entire miRNAs probably due to the absence of exo-miRNA TSs (compare Figs. 1C vs. 1E and 1D vs. 1F). Furthermore, a significant positive correlation was observed for this subset of miRNAs, indicating that endo-miRNA TSs may induce the mRNA derepression independently of the presence or absence of exo-miRNA TSs on the 3'-UTR (Figs. 1E and 1F).

Then we observed the correlation between mean fold-change and 3'-UTR length on a log scale among miRNAs without exo- or endo-miRNA 6-8mer sites to verify that miRNAs without...
**Fig. 1.** The effect of the global competition between exo- and endo-miRNAs on derepression of endo-miRNA targets. (A) Among mRNAs without exo-miRNA 6-8mer sites, mRNAs with endo-miRNA 7-8mer sites (red) are significantly upregulated than those without endo-miRNA 6-8mer sites (blue). To count the number of endo-miRNA TSSs, top 10 highly expressed endo-miRNAs in HeLa cell were considered (Landgraf et al., 2007). (B) Correlation between 3′-UTR length and the number of endo-miRNA 7-8mer sites. Top 10 (red), 20 (green), and 30 (magenta) highly expressed endo-miRNAs were considered to measure the number of endo-miRNA 7-8mer sites in each case. (C) Using all the mRNAs included in our 74 microarray data, fold-change values of mRNAs in response to exo-miRNAs or siRNAs were correlated with the number of endo-miRNA 7-8mer sites. (D) Fold-change values were correlated with 3′-UTR length. Otherwise as in (C). (E, F) For a subset of mRNAs without exo-miRNA 6-8mer sites, the same correlation analyses as in (C, D) were repeated. (G) For mRNAs without exo-miRNA 6-8mer sites (blue), the correlation between the mRNA fold-change and the 3′-UTR length on a log scale was analyzed (thick lines) and the simple linear regression lines are depicted (thin lines). After discarding mRNAs with 6-8mer sites in the 3′-UTR for the top 10 (red), 20 (green), and 30 (magenta) highly expressed endo-miRNAs, the same correlation analysis was conducted. Each point represents mean fold-change and the log scaled 3′-UTR length of 15,000 mRNAs.
A pair of mRNAs with the smallest difference in their 3'-UTR length was classified as "weaker exo-miRNA 7-8mer sites" group, while the other was classified as the "stronger exo-miRNA 7-8mer sites" group, based on the sum of context+ scores of exo-miRNA 7-8mer sites was nearly identical 3'-UTR length. Otherwise these TSs are not influenced by miRNA or siRNA transfection. As we took a larger number of endo-miRNAs into consideration, the slopes of the linear regression lines were closer to zero, indicating that mRNAs without exo- or endo-miRNA TSs are not affected by miRNA or siRNA transfection (Fig. 1G). These mRNAs were slightly downregulated probably because we only considered TSs in 3'-UTR while some ORF TSs of the exo-miRNA can also be functional (Baek et al., 2008; Forman et al., 2008).

Taken together, these results support the previous observation that the competition between exo- and endo-miRNAs may cause the derepression of endo-miRNA target mRNAs and the range of the derepression is broad enough to influence exo-miRNA targets as well as non-targets.

The target site proficiency of exogenously introduced miRNAs also influences mRNA derepression
To examine whether the site type of an exo-miRNA TS can influence mRNA derepression, we investigated mRNAs that included a single 3'-UTR 8mer site of exo-miRNA but no other exo-miRNA TSs. As shown in Fig. 2A, a strong association between the mean mRNA fold-change and 3'-UTR length was observed. mRNAs with longer 3'-UTRs tended to exhibit a weaker response to the exo-miRNA, probably due to a stronger derepression by a larger number of endo-miRNA TSs. For instance, the mean fold-change of mRNAs with short 3'-UTRs (< 500 nucleotides (nt)) was -0.40 while that of mRNAs with long 3'-UTRs (≥ 4,000nt) was only -0.06, demonstrating a strikingly weaker response of the latter compared to the former (fold-change difference = -0.40–(-0.06) = -0.34).

However, when looking at mRNAs that included a single 3'-UTR 7mer-m8 site for exo-miRNA but no other exo-miRNA TSs, the mean fold-change between the short and long 3'-UTR groups (< 500 nt vs. ≥ 4,000 nt) was relatively close (Fig. 2A, fold-change difference = -0.21–(-0.03) = -0.18). Similarly, the mean fold-change differences became smaller as we looked at weaker site types (Fig. 2A, fold-change differences = -0.13, -0.02, and -0.06 for 7mer-A1, 6mer, and no site, respectively).

A linear regression between the mean mRNA fold-change and the 3'-UTR length on a log scale was performed to test whether the slopes were significantly different depending on the site type of exo-miRNA TSs (Fig. 2B). We found that all four site types presented significantly different regression slopes in comparison to the no site group (P = 1.0 × 10^-16, < 1.0 × 10^-16, < 1.0 × 10^-16, and 2.0 × 10^-12 for 8mer, 7mer-m8, 7mer-A1, and 6mer, respectively, F-test).

To further control for potential confounding factors such as the difference in the site proficiency for endo-miRNA TSs, we chose two groups of mRNAs that have nearly identical 3'-UTR lengths, but one of the group of mRNAs had a higher site proficiency of exo-miRNA 7-8mer sites than the other. Thus, we aimed to control for the overall site proficiency of exo-miRNA TSs and to assess the effect of the site proficiency of exo-miRNA TSs on the degree of mRNA derepression. Indeed, the median length, the number of endo-miRNA 7-8mer sites, and the sum of context+ scores of endo-miRNA 7-8mer sites between the two groups were not significantly different (P = 0.99, 0.77, and 0.85, respectively, Wilcoxon's rank-sum test). On the other hand, the group with stronger exo-miRNA TSs showed a significantly stronger correlation between mean mRNA fold-change and 3'-UTR length in comparison to the other group with weaker exo-miRNA TSs (Fig. 2C, P = 1.4 × 10^-11, F-test). Our analysis indicates the site proficiency of not only endo-miRNA TSs but also exo-miRNA TSs are important determinants for the degree of derepression of endo-miRNA target, implying that the mRNA derepression in response to exo-miRNAs

**Fig. 2.** The effect of the local competition between exo- and endo-miRNAs on derepression of endo-miRNA targets. (A) For mRNAs with a single exo-miRNA 8mer site, the correlation between mRNA fold-change and 3'-UTR length was analyzed (red). The same correlation analysis was repeated for 7mer-m8 (orange), 7mer-A1 (green), 6mer (blue), and no site (magenta). (B) mRNA fold-change values were correlated with log values of 3'-UTR length (thick lines) and their simple linear regression lines (thin lines) are shown together. Otherwise as in (A). (C) For each pair of mRNAs that have nearly identical 3'-UTR, one with a stronger proficiency (measured by the sum of context+ scores) of exo-miRNA 7-8mer sites was classified as the “stronger exo-miRNA 7-8mer sites” group, while the other was classified as “weaker exo-miRNA 7-8mer sites” group. A pair of mRNAs with the smallest difference in their 3'-UTR length was iteratively chosen until the 3'-UTR length difference became longer than 50 nucleotides. For each of the “stronger exo-miRNA TSs” and “weaker exo-miRNA TSs” group, mRNA fold-change values were correlated with log values of 3'-UTR length. Otherwise as in (B).
Fig. 3. A proposed model to explain the effect of global and local competition between exo- and endo-miRNAs on the derepression of endo-miRNA targets. (A) mRNAs without exo- or endo-miRNA TSs on their 3'-UTRs are not affected by miRNA or siRNA transfection. (B) mRNAs with endo-miRNA TSs but no exo-miRNA TS show a moderate degree of derepression caused by the global competition between exo- and endo-miRNAs. (C) mRNAs with both exo- and endo-miRNA TSs show a strong derepression, resulting from both the global and local competition between exo- and endo-miRNAs.
regulation by individual miRNAs.

Understanding the complicated nature of the competition between exo- and endo-miRNAs may help us draw a more complete picture of miRNA targeting rules. For instance, one would be able to computationally model the competition between exo- and endo-miRNAs as well as the coupling between 3'UTR length and the efficacy of miRNA targeting. These future efforts may lead to improve the accuracy of miRNA target prediction and reduce off-targeting by siRNA and shRNA.

ACKNOWLEDGMENTS

We thank J. Park for his critical comments on our manuscript and other Baek laboratory members for their helpful suggestions. This work was supported by the Institute for Basic Science (EM1402), by the Basic Science Research Program (2011-0014523), and the Bio & Medical Technology Development Program (2012M3A9D1054622) of the National Research Foundation funded by the Ministry of Science, ICT & Future Planning, and by the TJ Park Science Fellowship of POSCO TJ Park Foundation.

REFERENCES

Anderson, E.M., Birmingham, A., Baskerville, S., Reynolds, A., Maksimova, E., Leake, D., Fedorov, Y., Karpilow, J., and Khvorova, A. (2008). Experimental validation of the importance of seed complement frequency to siRNA specificity. RNA 14, 853-861.

Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. Nature 455, 64-71.

Birmingham, A., Anderson, E.M., Reynolds, A., Ilseley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J., et al. (2006). 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets (vol 3, pg 199, 2006). Nat. Methods 3, 487-487.

Castanotto, D., Sakurai, K., Lingeman, R., Li, H., Shively, L., Aggaard, L., Soffer, H., Gatignol, A., Riggs, A., and Rossi, J.J. (2007). Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. Nucleic Acids Res. 35, 5154-5164.

Forman, J.J., Legesse-Miller, A., and Coller, H.A. (2008). A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. Proc. Natl. Acad. Sci. USA 105, 14879-14884.

Garcia, D.M., Baek, D., Shin, C., Bell, G.W., Grimson, A., and Bartel, D.P. (2011). Weak seed-pairing stability and high target-site abundance decrease the proficiency of let-6 and other microRNAs. Nat. Struct. Mol. Biol. 18, 1139-1146.

Grimson, A., Fath, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol. Cell 27, 91-105.

Hultvagner, G., and Simard, M.J. (2008). Argonaute proteins: key players in RNA silencing. Nat. Rev. Mol. Cell Biol. 9, 22-32.

Jackson, A.L., Burchard, J., Leake, D., Reynolds, A., Schelter, J., Guo, J., Johnson, J.M., Lim, L., Karpilow, J., Nichols, K., et al. (2006a). Position-specific chemical modification of siRNAs reduces “off-target” transcript silencing. RNA 12, 1197-1205.

Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L., and Linsley, P.S. (2006b). Widespread siRNA “off-target” transcript silencing mediated by seed region sequence complementarity. RNA 12, 1179-1187.

Khan, A.A., Betel, D., Miller, M.L., Sander, C., Leslie, C.S., and Marks, D.S. (2006). Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. Nat. Biotechnol. 27, 549-555.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature 409, 860-921.

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., et al. (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 129, 1401-1414.

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.

Pratt, A.J., and MacRae, I.J. (2009). The RNA-induced silencing complex: a versatile gene-silencing machine. J. Biol. Chem. 284, 17897-17901.

Prutt, K.D., Katz, K.S., Sicotte, H., and Maglott, D.R. (2000). Introducing RefSeq and LocusLink: curated human genome resources at the NCBI. Trends Genet. 16, 44-47.

Schwarz, D.S., Ding, H.L., Kennington, L., Moore, J.T., Schelter, J., Burchard, J., Linsley, P.S., Aronin, N., Xu, Z.S., and Zamore, P.D. (2006). Designing siRNA that distinguish between genes that differ by a single nucleotide. PLoS Genet. 2, 1307-1318.