

Passenger strand loading in overexpression experiments using microRNA mimics

Rolf Søkilde¹, Inga Newie^{1,2}, Helena Persson¹, Åke Borg^{1,2,3}, and Carlos Rovira^{1,2,3,*}

¹Department of Clinical Sciences, Lund; Division of Oncology and Pathology; Lund University Cancer Center at Medicon Village; Lund, Sweden; ²BioCARE; Strategic Cancer Research Program; Lund University Cancer Center at Medicon Village; Lund, Sweden; ³CREATE Health; Strategic Center for Translational Cancer Research; Lund University Cancer Center at Medicon Village; Lund, Sweden

Keywords: Argonaute protein, microRNA, mimic, siRNA, off-target effect, passenger strand, RNA interference

Abbreviations: AGO, Argonaute; GEO, Gene Expression Omnibus; HITS-CLIP, high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation; miRNA, microRNA; mRNA, mRNA; nt, nucleotide; OAS, oligoadenylate synthetase; PKR, protein kinase R; RISC, RNA-induced silencing complex; TLR3, Toll-like receptor 3

MicroRNAs (miRNAs) are important regulators of gene function and manipulation of miRNAs is a central component of basic research. Modulation of gene expression by miRNA gain-of-function can be based on different approaches including transfection with miRNA mimics; artificial, chemically modified miRNA-like small RNAs. These molecules are intended to mimic the function of a miRNA guide strand while bypassing the maturation steps of endogenous miRNAs. Due to easy accessibility through commercial providers this approach has gained popularity, and accuracy is often assumed without prior independent testing. Our *in silico* analysis of over-represented sequence motifs in microarray expression data and sequencing of AGO-associated small RNAs indicate, however, that miRNA mimics may be associated with considerable side-effects due to the unwanted activity of the miRNA mimic complementary strand.

Introduction

MicroRNAs (miRNAs) are expressed from the genome as primary transcripts (pri-miRNAs) with a hairpin structure that is cleaved by Drosha to produce precursor miRNAs (pre-miRNAs). These are exported to the cytoplasm and further processed by Dicer into 22 nucleotide (nt) double-stranded small RNAs with one strand derived from the 5' (5p) and the other from the 3' (3p) arm of the hairpin stem. Mature miRNAs are loaded into the RNA-induced silencing complex (RISC) that contains the Argonaute (AGO) proteins to repress gene expression by base-pairing to partially complementary target sites in mRNAs (mRNAs). During RISC loading one strand of the 22 nt miRNA duplex is discarded (the “passenger”) while the other is retained (the “guide”) to direct recruitment of the RISC to target mRNAs. For many miRNAs only one of the strands is preferentially selected as guide (either 5p or 3p), while for others both strands are used equally.¹

Various strategies have been adopted to manipulate miRNA levels for functional studies. Overexpression in cell culture can be achieved with vectors expressing the natural miRNA gene or a short hairpin, but the combined effects of 5p and 3p miRNAs may complicate interpretation of the results when both are functional. Artificial miRNAs, so-called miRNA mimics, are aimed at circumventing this problem by transfection of small RNAs chemically similar to processed miRNAs. Although the use of

single-stranded mimics has been suggested,² the most widely used design is a small double-stranded RNA with a guide strand that is identical in nucleotide sequence to the miRNA, while the passenger strand is perfectly complementary to the guide (unlike endogenous animal miRNAs). Unmodified mimics display substantial artifactual incorporation of the passenger strand.³ Different approaches to avoid this include combinations of specific parameters affecting duplex thermodynamics, the use of a passenger strand segmented into 2 shorter pieces and/or chemical modification of the 5' end of the passenger strand to impair interaction with the MID domain of AGO proteins such as C6 amine linkers, NH₂, 5'-O-methyl, alkylamine groups, NHCOCH₃, acetyl groups, fluorescein, biotin, acridine, spacer 18 (PEG), and amide (DMT-Hexa(ethylene glycol)).

Since different designs are openly discussed in the academic scientific literature,⁴ custom-made miRNA mimics can be readily synthesized, but these require laborious testing and validation. In response to this, many companies offer convenient off-the-shelf solutions. These are designed according to proprietary formulations and the provided information usually only relates to the activation of cellular sensors for foreign RNA such as PKR, TLR3 and OAS (interferon induction) but not to side-effects caused by unintended sequence-specific targeting of transcripts. Although evidence for the accuracy of these reagents is limited and their chemical nature is normally not disclosed to users, commercial miRNA mimics have been applied in a large number of

*Correspondence to: Carlos Rovira; Email: carlos.rovira@med.lu.se

Submitted: 10/10/2014; Revised: 02/07/2015; Accepted: 02/13/2015

http://dx.doi.org/10.1080/15476286.2015.1020270

studies due to their easy accessibility and the possibility of selecting the desired guide strand. Here we show that effects of some commercial miRNA mimics may vary greatly depending on the synthesis batch, leading to undesired side-effects due to passenger strand loading. Thus caution is advisable when using commercial miRNA mimics, stressing the need for independent testing following the standard procedures of scientific research.

Results

During functional studies of the miRNA hsa-miR-4728-3p⁵ we used several batches of the same commercially available miRNA mimic in overexpression studies in cell culture. To determine the target genes for this miRNA, we also performed microarray expression analysis and searched for differential expression between cells transfected with an hsa-miR-4728-3p mimic or a non-targeting control mimic. The ranked gene list of differentially expressed genes was further analyzed with SylArray⁶ and MixMir,⁷ algorithms used to identify miRNA targeting effects from gene expression data based on identification of enriched sequence motifs among differentially expressed genes.

For the first batch, the most enriched motifs among down-regulated genes corresponded to targets of hsa-miR-4728-3p, as expected (compare Newie et al.⁵ and Fig. 1A). In contrast, the results from overexpression using a second batch of the same mimic displayed a clear enrichment for target genes of hsa-miR-486-3p, suggesting a co-regulatory loop between the 2 miRNAs (Fig. 1B). To verify this change, we analyzed small RNA sequencing data for the tested cell line. Expression of hsa-miR-486-3p was very low and constant in both control and treated cells (on average 1 read per library), and therefore unlikely to account for down-regulation of genes containing target sites for this miRNA. However, a sequence alignment between hsa-miR-486-3p and a reverse complement of the hsa-miR-4728-3p mimic revealed a clear sequence similarity (Fig. 1C)

corresponding to an offset 6-mer site⁸ with 6 contiguous matches to nucleotides 3–8. This suggests a sequence-specific artifact derived from unintentional loading of the mimic passenger strand. Upon inquiry, the manufacturer revealed only that their mimics are double-stranded RNA.

To clarify this, we performed Argonaute high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation⁹ (AGO HITS-CLIP), determining the entire complement of miRNAs loaded into RISC after transfection with the second, problematic mimic batch. Since the microarray data shown in Figure 1 suggested large batch-dependent variation for this mimic, we compared the second batch with a third and fourth production batch. The hsa-miR-4728-3p guide strand was clearly associated with AGO2 in all tested batches of the mimic. However, the passenger strand was also detected in the AGO HITS-CLIP experiments, but with varying ratios of 250, 15, and 10%, respectively, compared to the guide strand (Fig. 2). Coincidentally, the largest ratio of passenger to guide strand incorporation was found for the problematic mimic batch where the SylArray analysis identified enrichment of targets for hsa-miR-486-3p (Fig. 1B). These results confirm the suspicion that accidental passenger strand loading can be substantial and explain the enrichment profiles observed in Figure 1B.

We next searched for similar side-effects in independently produced data, to determine whether this is a general problem with miRNA mimics or particular to our experiments. To estimate the extent of miRNA mimic usage, we searched PubMed for: (miRNA OR microRNA) AND (mimic or mimics). At the time of writing this manuscript the search produced 1475 hits, 79 of them marked as reviews, although this search will miss authentic matches as well as include non-relevant publications. Furthermore, the database miRTarBase¹⁰ (release 4.4) for validated miRNA target genes is based on 2636 papers, giving another indication of the extent of functional studies that have been done for miRNAs. Most of these will be perturbation experiments with a sizable fraction of them using miRNA

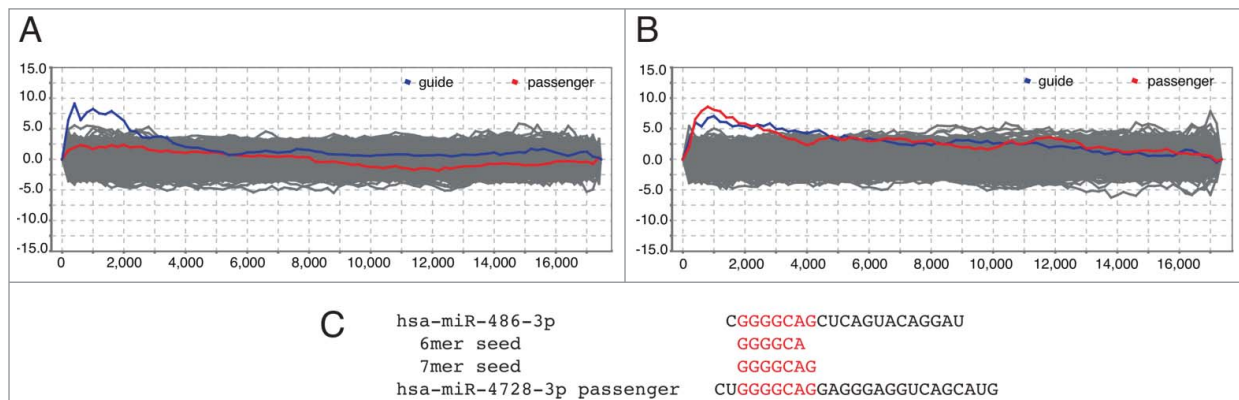


Figure 1. SylArray enrichment analysis for sequence motifs in mRNA expression profiles after transient transfection of MCF10A cells with a hsa-miR-4728-3p mimic compared to a non-targeting control shows markedly different profiles for different batches of miRNA mimic. In (A) only sequence motifs complementary to the guide strand (blue) are enriched. In (B) the array data shows enrichment of a motif complementary to the seed sequence of hsa-miR-486-3p (red). This motif is also present in the putative passenger strand of the hsa-miR-4728-3p mimic. The x-axis is sorted by t-statistic for the pairwise comparison of differential expression with the most negative value (i.e. the most significantly downregulated transcript) on the far left. The y-axis shows $-\log_{10}$ of the P-value for enrichment and \log_{10} for depletion. An alignment of hsa-miR-4728-3p and miR-486-3p is shown in (C).

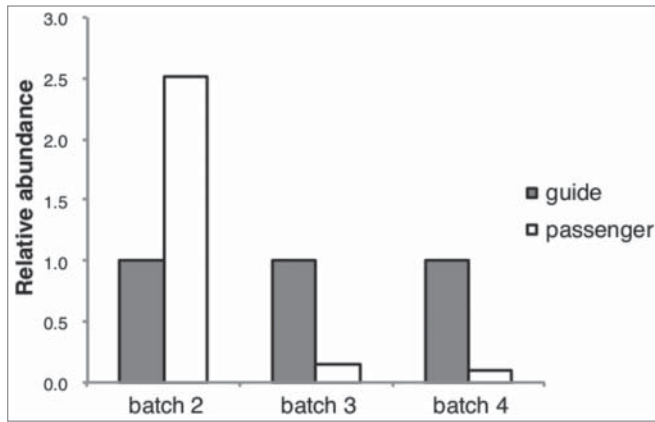


Figure 2. The ratio of passenger to guide strand incorporated into the RNA-induced silencing complex varies considerably between different batches of hsa-miR-4728-3p mimic, as demonstrated by Argonaute high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (AGO HITS-CLIP). Batch 2 was also used for the motif enrichment analysis shown in **Figure 1B**, batches 3 and 4 are additional batches from the same manufacturer.

mimics, but unfortunately we could not find detailed experimental information within the database.

The Gene Expression Omnibus (GEO) contains a total of 31 datasets that involve single miRNA mimics (rather than e.g. siRNA or pre-miRNA chemistry or a complex mixture of mimics) and provide expression data from microarray studies (Table S1). We analyzed these data sets for enrichment of 6mer motifs with MixMir,⁷ another motif discovery algorithm based on linear models, and found significant enrichment (Bonferroni-corrected $P < 0.05$) of the passenger strand for 5 out of 38 mimics (13%). For 11 mimics (29%) there was no significant enrichment of either guide or passenger strand motifs, but this could of course depend on a range of factors in addition to mimic design. We noted that the miRNA mimics used in these datasets were biased to a few manufacturers, which limits the extent of the comparison. An example of a study where off-target effects were detected is Mackiewicz et al.,¹¹ where mimics for hsa-miR-34a were used (Fig. 3). In this particular case, the associated GEO data (GSE21832) compared commercial mimics for the same miRNA from 2 different vendors. MixMir did not identify

any enriched motifs for the passenger strand in the data set where the Dharmacon mimic was used, while the first and second most enriched motifs for the Qiagen mimic matched the guide and passenger strands, respectively (Table 1). This confirms our results and indicates that off-targeting by the passenger strand may be a general hazard of commercial miRNA mimics.

Appropriate chemical modification of small RNA duplexes can be used to reduce passenger strand loading compared to unmodified RNA. Chen et al.¹² performed extensive characterization of an siRNA, comparing 4 different designs; 1) unmodified guide and passenger strands, 2) 5'-O-methylation of the passenger strand, 3) 5'-O-methylation of the guide strand, and 4) 5'-O-methylation of both strands. This modification is expected to prevent intracellular addition of the 5' phosphate group that is recognized by AGO during RISC loading. Through FLAG-AGO2 pull-down, *in vitro* RISC cleavage and reporter gene assays the authors showed that 5'-O-methylation inhibited both RISC loading and function of the modified strand. We analyzed this data for enrichment of 6mer motifs using MixMir. The unmodified duplex displayed significant enrichment of seed site motifs from both guide and passenger strands (Bonferroni-corrected $P < < 0.01$). As expected, 5'-O-methylation of either guide or passenger strand resulted in significant enrichment of motifs exclusively corresponding to the other, unmodified strand (Bonferroni-corrected $P < < 0.01$ in both experiments).

Discussion

Double-stranded small RNA mimics are widely used for transient overexpression of miRNAs. This straightforward technique bypasses the natural mechanism of miRNA biogenesis and is even introduced as a new therapeutic option for modulating miRNA function.¹³ Problems such as RNA degradation, elicitation of an innate immune response, and low permeability of cell membranes can be addressed by chemical modifications.¹⁴ Additions of 2'-O-methyl or 2'-O-fluorine to the 2'-OH of the sugar ring for instance increase RNA stability.

Another potential complication is unintentional loading of the mimic passenger strand into the RISC. This strand is fully or partially complementary to the guide strand and most often not

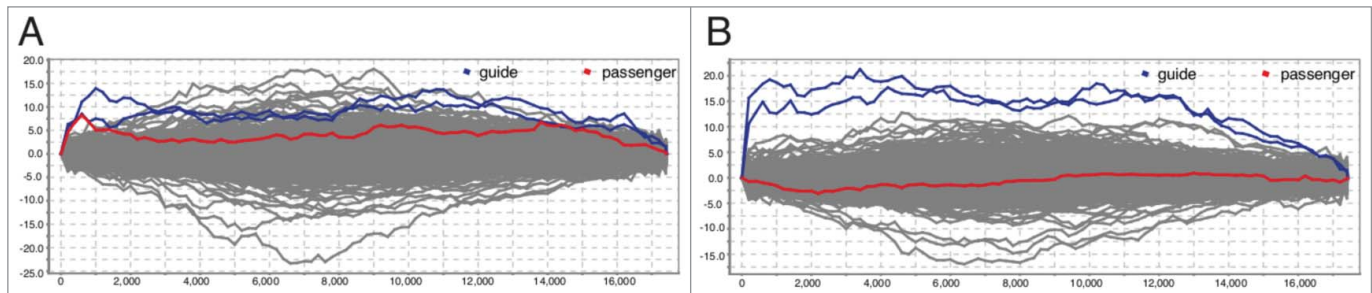


Figure 3. SylArray enrichment analysis for sequence motifs in mRNA expression profiles after transient transfection of MDA-MB-231 cells with an hsa-miR-34a-5p mimic compared to a non-targeting control (GSE21832) shows markedly different profiles for miRNA mimics from (A) Qiagen and (B) Dharmacon.

Table 1. MixMir enrichment analysis for 6mer motifs in mRNA expression profiles after transient transfection of MDA-MB-231 cells with an hsa-miR-34a-5p mimic compared to a non-targeting control (GSE21832) shows significant enrichment of motifs matching both guide (rank 1) and passenger (rank 2) strands of the Qiagen mimic

Rank	Motif	Adj. P-value	Coefficient	N 3' UTRs
1	GGCAGT	0.0000	0.272	5571
2	ACCAGC	0.0016	0.275	5557
3	TGGCAG	0.0027	0.369	7496
4	TAAAAA	0.0088	0.413	11790
5	AAAAAT	0.0127	0.409	11866
6	CAGATT	0.0192	0.299	6002
7	TTCAAG	0.0278	0.359	7277
8	AAAAAG	0.0590	0.467	10749
9	CATTGG	0.0788	0.185	3876
10	TTAAAA	0.0843	0.423	11630

identical to the “star” strand of the endogenous miRNA gene. Experimental evidence shows that *ad hoc* chemical formulations allow specific selection of only one of the 2 strands of the mimic, minimizing off-target effects from the passenger strand. Henry et al.¹⁵ for instance have described a mimic design optimized to increase stability and decrease passenger strand loading where they use a 5' phosphorylated guide strand with a 2 nt 3' overhang together with a blunt-ended passenger strand with a 3' terminal deoxynucleotide. The 5' end of the passenger strand also had an SpC3 modification, which might further interfere with RISC loading. 2'-O-methylation at selected positions in both guide and passenger strands was used to increase mimic stability.

Vendors of commercially available miRNA mimics typically do not release details regarding the chemical formulation, but ensure that off-targeting by the passenger strand will be insignificant, if at all present. However, using AGO HITS-CLIP for transiently transfected cells, we have shown that the passenger strand of a commercial miRNA mimic can indeed be loaded into the RISC. By microarray expression analysis we have also demonstrated that it can proceed to silence target genes, occasionally even to a larger extent than the desired guide strand. Due to the secrecy regarding the design and synthesis of these mimics, we can only speculate about the functional reasons for these off-target effects.

Our report does not aim for a comparative analysis between mimic designs and providers. We have focused on the use of miRNA mimics only in combination with global expression profiling and the datasets available in GEO are biased toward very few mimic providers. Still, we found independent evidence for passenger strand targeting for 5 out of 38 mimics from the 31 GEO data sets analyzed (13%). However, it appears that this problem does not affect all commercial miRNA mimics equally. Interestingly, Thomson et al. found that their commercial modified miRNA mimics were not significantly associated with AGO complexes.³

Here we observed that the specificity of guide *versus* passenger strand selection varied substantially for different production batches, suggesting that the problem with these mimics is linked to oligonucleotide synthesis. Such intolerable batch-to-batch

variation most likely results from insufficient quality control, but details about formulation and preparation of commercial reagents are confidential. Based on our results, we would like to raise awareness of the potential dangers of using commercially available miRNA mimics without further testing. Even if the same sequence from the same manufacturer has been tested before, proper controls should always be performed alongside current experiments to ensure scientifically sound and reproducible results. Furthermore, as has been suggested before,¹⁶ commercial providers ought to disclose complete information regarding their reagents in accordance with current standards for scientific publication.

Materials and Methods

Cell culture and miRNA mimics

MCF10A were purchased from ATCC and cultured as reported previously.¹⁷ Transfections were performed with Lipofectamine 2000 (Life Technologies) following the manufacturer's instructions. All miRNA mimics were purchased from Qiagen and were transfected at 25 nM.

Microarray expression analysis

RNA was extracted with TRIzol (Life Technologies) according to the manufacturer's instructions. RNA quantity and quality were assessed with NanoDrop ND 1000 spectrophotometer (NanoDrop Tech) and LabChip GX (Perkin Elmer) before loading the samples on Human HT-12 v3.0 or v4.0 Expression Bead-Chips (Illumina) in 4–6 biological replicates. Raw data for the batch without passenger off-targeting is available at the GEO repository under accession number GSE55822. All data were imported and normalized using the Base server (<http://base.thep.lu.se>). Array identifiers sorted by t-value were analyzed using SylArray.⁶ SylArray was performed using default parameters with purging of low complexity sequences and including all available words, not just miRNA seeds. MixMir⁷ was run using default settings for log₂ fold changes after averaging probes for the same RefSeq gene identifier. The 3' UTR sequences were obtained from Ensembl BioMart and filtered to retain only the longest sequence for each RefSeq. The MixMir output was analyzed for mimic motif enrichment using custom Perl scripts.

AGO HITS-CLIP and next-generation sequencing

AGO HITS-CLIP was performed as described.⁹ Cells were UV irradiated in a Stratalinker (Stratagene) twice with 400 mJ/cm². Antibodies were monoclonal AGO2 4G8 (Wako Chemicals). Sequencing libraries from purified RNA were prepared with NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs) according to the manufacturer's instructions and sequenced on an Illumina HiSeq sequencer in paired-end mode with 2 × 101 cycles. The numbers of sequences matching the mature miRNA (guide strand) and its reverse complement (putative passenger strand) were counted allowing 2 mismatches at the 5' and 3' ends.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

governmentally supported strategic cancer research program shared between the universities of Lund and Gothenburg.

Funding

This research was supported by the Swedish Research Council, the Swedish Cancer Society and the Fru Berta Kamprad Cancer Foundation. RS and IN were also supported by BioCARE, a

Supplemental Material

Supplemental data for this article can be accessed on the publisher's website.

References

1. Kuchenbauer F, Mah SM, Heuser M, McPherson A, Ruschmann J, Rouhi A, Berg T, Bullinger L, Argiropoulos B, Morin RD, et al. Comprehensive analysis of mammalian miRNA* species and their role in myeloid cells. *Blood* 2011; 118:3350-8; PMID:21628414; <http://dx.doi.org/10.1182/blood-2010-10-312454>
2. Chorn G, Klein-McDowell M, Zhao L, Saunders MA, Flanagan WM, Willingham AT, Lim LP. Single-stranded microRNA mimics. *RNA* 2012; 18:1796-804; PMID:22912485; <http://dx.doi.org/10.1261/rna.031278.111>
3. Thomson DW, Bracken CP, Szubert JM, Goodall GJ. On measuring miRNAs after transient transfection of mimics or antisense inhibitors. *PloS One* 2013; 8: e55214; PMID:23358900; <http://dx.doi.org/10.1371/journal.pone.0055214>
4. Deleavey GF, Damha MJ. Designing chemically modified oligonucleotides for targeted gene silencing. *Chem Biol* 2012; 19:937-54; PMID:22921062; <http://dx.doi.org/10.1016/j.chembiol.2012.07.011>
5. Newie I, Sokilde R, Persson H, Grabau D, Rego N, Kvist A, von Stedingk K, Axelson H, Borg A, Vallon-Christersson J, et al. The HER2-encoded miR-4728-3p regulates ESR1 through a non-canonical internal seed interaction. *PloS One* 2014; 9:e97200; PMID:24828673; <http://dx.doi.org/10.1371/journal.pone.0097200>
6. van Dongen S, Abreu-Goodger C, Enright AJ. Detecting microRNA binding and siRNA off-target effects from expression data. *Nat Methods* 2008; 5:1023-5; PMID:18978784; <http://dx.doi.org/10.1038/nmeth.1267>
7. Diao L, Marçais A, Norton S, Chen KC. MixMir: microRNA motif discovery from gene expression data using mixed linear models. *Nucleic Acids Res* 2014; 42(17):e135; PMID:25081207
8. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009; 19:92-105; PMID:18955434; <http://dx.doi.org/10.1101/gr.082701.108>
9. Ule J, Jensen K, Mele A, Darnell RB. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* 2005; 37:376-86; PMID:16314267; <http://dx.doi.org/10.1016/j.ymeth.2005.07.018>
10. Hsu SD, Tseng YT, Shrestha S, Lin YL, Khaleel A, Chou CH, Chu CF, Huang HY, Lin CM, Ho SY, et al. miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res* 2014; 42:D78-85; PMID:24304892; <http://dx.doi.org/10.1093/nar/gkt1266>
11. Mackiewicz M, Huppi K, Pitt JJ, Dorsey TH, Ambs S, Caplen NJ. Identification of the receptor tyrosine kinase AXL in breast cancer as a target for the human miR-34a microRNA. *Breast Cancer Res Treat* 2011; 130:663-79; PMID:21814748; <http://dx.doi.org/10.1007/s10549-011-1690-0>
12. Chen PY, Weinmann L, Gaidatzis D, Pei Y, Zavolan M, Tuschl T, Meister G. Strand-specific 5'-O-methylation of siRNA duplexes controls guide strand selection and targeting specificity. *RNA* 2008; 14:263-74; PMID:18094121; <http://dx.doi.org/10.1261/rna.789808>
13. Bouchie A. First microRNA mimic enters clinic. *Nat Biotechnol* 2013; 31:577; PMID:23839128; <http://dx.doi.org/10.1038/nbt0713-577>
14. Behlke MA. Chemical modification of siRNAs for in vivo use. *Oligonucleotides* 2008; 18:305-19; PMID:19025401; <http://dx.doi.org/10.1089/oli.2008.0164>
15. Henry JC, Azevedo-Pouly AC, Schmittgen TD. MicroRNA replacement therapy for cancer. *Pharm Res* 2011; 28:3030-42; PMID:21879389; <http://dx.doi.org/10.1007/s11095-011-0548-9>
16. Git A. Research tools: a recipe for disaster. *Nature* 2012; 484:439-40; PMID:22538584; <http://dx.doi.org/10.1038/484439a>
17. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 2006; 10:515-27; PMID:17157791; <http://dx.doi.org/10.1016/j.ccr.2006.10.008>