Briefings in Bioinformatics, 18(5), 2017, 780–788

doi: 10.1093/bib/bbw053
Advance Access Publication Date: 2 July 2016
Paper

Screening and validation of lncRNAs and circRNAs as miRNA sponges

Giuseppe Militello*, Tyler Weirick*, David John, Claudia Döring, Stefanie Dimmeler and Shizuka Uchida

Corresponding author: Shizuka Uchida, Institute of Cardiovascular Regeneration, Centre for Molecular Medicine, Goethe University, 60590 Frankfurt, Germany. Tel.: +49 69 6301 87949; E-mail: heart.lncrna@gmail.com
*These authors contributed equally to this work.

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022

Abstract

Intensive research in past two decades has uncovered the presence and importance of noncoding RNAs (ncRNAs), which includes microRNAs (miRs) and long ncRNAs (lncRNAs). These two classes of ncRNAs interact to a certain extent, as some lncRNAs bind to miRs to sequester them. Such lncRNAs are collectively called 'competing endogenous RNAs' or 'miRNA sponges'. In this study, we screened for lncRNAs that may act as miRNA sponges using the publicly available data sets and databases. To uncover the roles of miRNA sponges, loss-of-function experiments were conducted, which revealed the biological roles as miRNA sponges. *LINC00324* is important for the cell survival by binding to miR-615-5p leading to the derepression of its target *BTG2*. *LOC400043* controls several biological functions via sequestering miR-28-3p and miR-96-5p, thereby changing the expressions of transcriptional regulators. Finally, we also screened for circular RNAs (circRNAs) that may function as miRNA sponges. The results were negative at least for the selected circRNAs in this study. In conclusion, miRNA sponges can be identified by applying a series of bioinformatics techniques and validated with biological experiments.

Key words: circular RNA; lncRNA; microRNA; transcriptome

Introduction

Increasing evidence suggests that protein-coding genes comprise a minority in the human genome, but most of the genome is transcribed as RNAs that are not translated [1]. These RNAs are collectively called 'noncoding RNAs' (ncRNAs). Besides the well-known ribosomal RNAs, transfer RNAs and microRNAs (miRs), there is a new emerging of class of ncRNAs called 'long ncRNAs (lncRNAs)'. LncRNAs are defined broadly, which include ncRNAs that are longer than 200 nt. Although the exact number of lncRNAs in the human genome is unknown, it is estimated to be more than that of protein-coding genes [2]. Given the diverse functions of protein-coding genes and their protein products, it is speculated that lncRNAs also possess various functions [3, 4]. Of these functions, lncRNAs that bind to other macromolecules (i.e. nucleic acids, proteins) are of great interest, as they could be used as a molecular switch to activate or to inhibit biological processes. Given the great therapeutic potential of miRs as diagnostic biomarkers and as a tool to inhibit the translation of proteins [5–7], lncRNAs that bind

Giuseppe Militello is a PhD student at the Institute of Cardiovascular Regeneration (Uchida Lab), who is working with long noncoding RNAs in the skeletal muscle.
Tyler Weirick is a PhD student at the Institute of Cardiovascular Regeneration (Uchida Lab), focused on elucidating the evolutional conservation of long noncoding RNAs.
David John is a PhD student at the Institute of Cardiovascular Regeneration (Uchida Lab), who is developing computational algorithms and pipelines to identify RNA modification events.
Claudia Döring is a bioinformatics scientist and laboratory manager of the RNA lab at the Dr. Senckenberg Institute of Pathology, who is focused on gene expression and next-generation sequencing analysis especially in lymphoma diseases.
Stefanie Dimmeler is the director of the Institutes of Cardiovascular Regeneration.
Shizuka Uchida is an independent junior group leader at the Institute of Cardiovascular Regeneration. His lab ('Cardiovascular Bioinformatics') is interested in elucidating the functions of long noncoding RNAs using dry and wet lab techniques.
Submitted: 27 March 2016; **Received (in revised form):** 12 May 2016

© The Author 2016. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

to miRs are of great interest, as they could be used to control miR functions and therefore may be used for therapeutic purpose. Such lncRNAs are collectively called 'competing endogenous RNAs' or 'miRNA sponges' to highlight their functionality. In this study, the term 'miRNA sponge' will be used. Up until now, only a handful of lncRNAs were found to be bound to miRs, which were experimentally validated. Examples include *H19*, which binds let-7 family members [8] and miR-106a [9]. Moreover, the binding of the following lncRNAs to the respective miRs was shown: *linc-MD1* (to miR-133 and miR-135 [10]), *Linc-RoR* (to miR-145 [11]), *lncRNA-BGL3* (to miR-17, miR-93, miR-20a, miR-20b, miR-106a and miR-106b [12]), *Malat-1* (to miR-133 [13] and miR-9 [14]), *MIAT* (to miR-150-5p [15]), *MRAK088388* (to miR-29b-3p [16]) and *MRAK081523* (to let-7i-5p [16]). Even a database called 'lnCeDB' [17] is set up that allows screening for lncRNAs that might bind to miRs based on predictions. However, given the low expression of most lncRNAs and recent studies showing that over-expression of competing endogenous RNAs has little effect on miR bio-availability and function [18], more vigorous testing and direct evidence that lncRNAs functioning as miRNA sponges is urgently needed before researchers confirm whether such a mechanism of action for lncRNAs is of their common features or not.

Cross-linking immunoprecipitation (CLIP) is a method to analyze the interaction of the target protein to RNAs by using ultraviolet cross-linking followed by immunoprecipitation [19]. The pulled down RNAs can be analyzed at once using RNA sequencing via next-generation sequencing. When such read-out method is used, it is called 'CLIP-seq' or 'HITS-CLIP' (HITS = high-throughput sequencing). Given that the argonaute (AGO) protein family constitutes RNA-induced silencing complex (RISC), which is responsible for RNA interference by loading miRs into its complex [20], researchers have used CLIP-seq to identify the potential targets of RISC by using anti-AGO antibodies. If indeed, the primary function of RISC is to inhibit the translation of mRNAs by binding to their 3′-untranslated regions (UTRs), then such data should not contain any lncRNAs, as their primary definition is that these RNAs are not translated. In reality, the analyzed results of AGO CLIP-seq data contain more regions than simply 3′-UTRs of protein-coding genes. Although some of these regions could be experimental noise (e.g. resulting from the preparation of library for sequencing), it could also be that these regions correspond to those of lncRNAs that function as miRNA sponges. Based on this idea, there are several databases set up, which one could screen for such lncRNAs from CLIP-seq data [21–24]. Although these databases are of use for *in silico* screening of potential miRNA sponges, defined biological validation experiments are missing for these databases.

To provide more direct biological evidence of lncRNAs functioning as miRNA sponges, we screened the published data of CLIP-seq data and validated their presences by biological experiments to prove the usage of the above databases of CLIP-seq data. By using loss-of-function experiments by siRNAs and readout by microarrays, we provided the direct evidence of functional lncRNAs with binding to AGO. In line with our previous publication of circular RNAs (circRNAs) [25], we provide the evidence of circRNAs not functioning as miRNA sponges.

Methods

Screening of potential miRNA sponges

The analyzed results (in the BED file format; hg19) of AGO-CLIP-seq data were downloaded from starBase (http://starbase.sysu.edu.cn) [21]. To annotate the CLIP-seq data, annotatePeaks.pl program of HOMER (http://homer.salk.edu/homer/) [26] was used. Then, custom PERL scripts were written to purse the results and to search for official gene symbols and their corresponding annotation from the annotation information provided by Entrez Gene (http://www.ncbi.nlm.nih.gov/gene) and ENSEMBL (http://www.ensembl.org/index.html) databases. For the prediction of miR binding sites, miRanda (http://www.microrna.org/microrna/getDownloads.do) [27] was used.

Culturing of cells, qRT-PCR and siRNAs

Human umbilical vein endothelial cells (HUVEC) were culture as previously [28]. RNAs were isolated with QIAzol Lysis Reagent (Qiagen) followed by the purification with miRNeasy Mini Kit (Qiagen).

HEK-293 and Hs68 cells were cultured in the growth medium containing 10% FBS (Life Technologies) in DMEM with low glucose, pyruvate (Life Technologies) supplemented with antibiotics (penicillin and streptomycin, Sigma-Aldrich). All cells were cultured at 37 °C in a humidified atmosphere containing 5% CO_2. RNA was isolated with TRIzol reagent (Life Technologies) and purified according to the protocol provided by the manufacturer.

After the purification and treatment of RNA with TURBO DNase (Life Technologies), 1 μg of RNA was reverse transcribed with SuperScript VILO Master Mix (Life Technologies). The first-strand cDNA was diluted to the concentration of 5 ng/μl. For quantitative reverse transcription polymerase chain reaction (qRT-PCR), 1 μl (5 ng) of the cDNA template was used with Fast SYBR Green Master Mix (Life Technologies) via StepOne Plus Real-Time PCR System (Applied Biosystem) with the following thermal cycling condition: 95 °C for 20 s followed by 40 cycles of 95 °C for 3 s and 60 °C for 30 s. Relative fold expression was calculated by $2^{-\Delta\Delta Ct}$ using *GAPDH* as an internal control. The primer pairs were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) [29] and *in silico* validated with UCSC In-Silico PCR (https://genome.ucsc.edu/cgi-bin/hgPcr) before extensive testing by experiments for the existence of a single band of the expect size for each primer pair. The list of primer pairs used in this study can be found in Supplementary Table S1.

To isolate nuclear and cytoplasmic RNAs, HEK-293 cells were washed once with 1× PBS and detached from the dish. After two washes with ice-cold PBS, pellets were resuspended in 200 μl of Buffer A (10 nM Tris pH = 8; 140 mM NaCl; 1.5 mM $MgCl_2$; and 0.5% Nonidet P-40) and incubated on ice for 5 min with gently flicking the tube every 90 s. Following the incubation, the suspension was centrifuged at 1000 × g at 4 °C for 3 min. The supernatant (containing the cytoplasmic fraction) was collected and mixed with 700 μl of TRIzol reagent. The cell pellet was washed twice with Buffer A and resuspended in Buffer B (Buffer A + 1% Tween-40; 0.5% Deoxycholate). After centrifugation at 1000 × g at 4 °C for 3 min, the supernatant was discarded. The pellet (containing nuclei) was resuspended in 700 μl of TRIzol reagent.

The following three siRNA duplexes (MISSION, Sigma-Aldrich) were used to silence *LINC00324*: (1) siLINC00324.1 sense CGUUUAUCAGUGGUUGGAAdTdT/antisense UUCCAACCACUGAUAAACGdTdT; and (2) siLINC00324.2 sense GUGUCAAGAUCCCAGGUUAdTdT/antisense UAACCUGGGAUCUUGACACdTdT. For silencing *LOC400043*: (1) siLOC400043.1 sense CAAAGCUGCUGACAGUCAUdTdT/antisense AUGACUGUCAGCAGCUUUGdTdT; and (2) siLOC400043.2 sense GAAGGAACCUGACUUGCUUdTdT/antisense AAGCAAGUCAGGUUCCUUCdTdT. Control siRNA used

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022

was Mission Negative control SIC-002, confidential sequence (Sigma-Aldrich). Transient siRNA transfection (10 nM final concentration) was carried out using RNAiMax (Life Technologies) according to the manufacturer's protocol. Forty-eight hours after the transfection of siRNAs, cells were exposed to TRIzol to extract RNAs.

Detection of circRNAs

All the primer pairs for the detection of circRNAs were designed against the circRNA-specific back-splice sites and checked with the UCSC In-Silico PCR against 'genome assembly' and 'Gencode Genes' for the human genomic and mRNA sequences, respectively. If indeed, the primer pairs were designed correctly, there should not be any matches to both genomic and mRNA sequences via the UCSC In-Silico PCR, which is the case for all of the primer pairs used in this study to detect circRNAs.

For RNase R digestion, 1.5 μg of RNA was incubated with 5 U RNaseR (Epicentre) for 30 min at 37 °C in a 10 μl scale. Following the incubation, 1 μl of TURBO DNase (Ambion) and 1.2 μl of TURBO DNase 10× buffer (Ambion) were added and incubated for additional 30 min at 37 °C to digest the genomic DNA. After the incubation, RNA was purified and resuspended in water. For the synthesis of cDNA, 100 ng of purified RNA was used with SuperScript VILO Master Mix (Life Technologies). The generated cDNA was diluted with water to a final concentration of 1 ng/μl.

Microarray experiments and data analysis

GeneChip®Human Gene 1.0 ST Arrays (Affymetrix) were used according to the manufacturer's protocol and scanned. The CEL files were analyzed through the updated version of noncoder web interface (http://noncoder.mpi-bn.mpg.de) [30] using the pipeline set up for Gene Array Analyzer web interface (http://gaa.mpi-bn.mpg.de) [31]. After the normalization by Robust Multi-array Average (RMA) [32] and the application of moderate t-statistics via the limma package [33], Transcript Cluster IDs that do not match to a gene or to multiple genes were discarded. Then, a standard deviation is calculated across samples. For a gene that matches to multiple Transcript Cluster IDs, the Transcript Cluster ID with the highest standard deviation was kept for further analysis. The results of microarray analysis can be accessed through our noncoder web interface using the user name and password as 'miR_Sponges'.

To detect miRs, GeneChip® miRNA 2.0 Arrays (Affymetrix) were used according to the manufacturer's protocol and scanned. The CEL files were normalized by RMA, and moderate t-statistics was applied via R using affy [34], limma [33] and vsn [35] packages. Because the probe information have not been updated by Affymetrix since 22 December 2010, PERL scripts were written to extract the updated information about mature miRs from miRBase (http://www.mirbase.org) [36] by using the probe sequences given by Affymetrix to pattern match them against 'mature.fa' provided by miRBase (downloaded on 29 April 2015).

All the microarray data in this study were deposited in the Gene Expression Omnibus (GSE73691).

RNA immunoprecipitation

Magna RIP Kit (Millipore) was used according to the manufacturer's protocol. The sub-confluent HEK-293 cells were fixed with 1% formaldehyde in PBS at room temperature for 10 min. Cross-linking reaction was stopped by adding 590 μl of 2.5 M Glycine. Fixed cells were subsequently harvested and resuspended in RNA immunoprecipitation (RIP) lysis buffer supplemented with protease/RNase inhibitors. Lysate was obtained using a dounce homogenizer on ice (dounce 10 times for releasing nuclei) followed by incubation on ice for 15 min. The equal volume of RIP lysis buffer was added to the cellular pellet. From the solution, 10 μl (10%) of the lysate was removed and stored as an 'input'. For each RIP reaction, 100 μl of the lysate were mixed with 5 μg of rabbit anti-IgG (negative control provided with the kit) or anti-AGO2 antibody (Millipore, #03-110) previously conjugated with Protein A/G magnetic beads (provided with the kit). After overnight incubation at +4 °C, RNA–protein immune-complexes were extensively washed with RIP Wash Buffer (provided with the kit). The cross-linking was reversed by incubation with proteinase K. The immune-precipitated RNA was purified through Phenol:Chloroform:Isoamyl Alcohol (125:24:25) isolation. The purified immune-precipitated RNA was treated with DNase I and reverse transcribed using SuperScript VILO Master Mix.

Statistics

Data are presented as mean ± S.E.M. unless otherwise indicated. Two-sample, two-tail, heteroscedastic Student's *t*-test was performed to calculate a *P*-value via Microsoft Excel.

Results

Screening of potential miRNA sponges

To identify lncRNAs acting as miRNA sponges, we screened CLIP-seq data of HEK-293 (Human Embryonic Kidney 293) cells, as this cell line is one of the most widely used *in vitro* model [37]. From the analyzed results of starBase [21], AGO-bound peaks were annotated via HOMER [26] (Figure 1A; Supplementary Table S2). As expected, the most abundant annotation class of bound regions is 3′-UTR (37.18% of the total number of annotated peaks). Of note, the percent distribution of 'noncoding' is only 4.72%. This is owing to the fact that known ncRNAs are also included in other annotation classes according to their exons, introns, etc. Next, the identified AGO-bound peaks were separate for each gene by using the annotation that corresponds to their exons and 3′/5′-UTRs (Supplementary Table S3). The gene with most AGO-bound peaks is *KLHL15* with 623 bound peaks. When the distribution of biotypes was calculated, the most abundant ones are protein-coding genes; a finding that is consistent with the current understanding that miRs bind primarily to the 3′-UTR of protein-coding genes to inhibit their translation (10 518 protein-coding genes; 234 pseudogenes; 316 ncRNAs; 1 small nucleolar RNAs). In the biotype provided by Entrez Gene, ncRNAs include miRs and their precursors. Based on these results, we conclude that CLIP-seq data contain a limited number of AGO-bound peaks for lncRNAs, which suggests that lncRNAs degradation is not generally controlled by AGO/miR binding.

Biological validation of lncRNAs functioning as miRNA sponges

From the above results of AGO-bound regions and their corresponding biotypes, lncRNAs were further filtered based on the annotation provided by the ENSEMBLE database (GRCh38, version 79). Of note, miRs and other small RNAs (e.g. SCARNAs; <200 nt) were removed from the further analysis. As a result, 261 lncRNAs were selected (Supplementary Table S4). In general, the longer the transcript itself, more AGO-bound regions could be found. To address this point, first, in the list of

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022

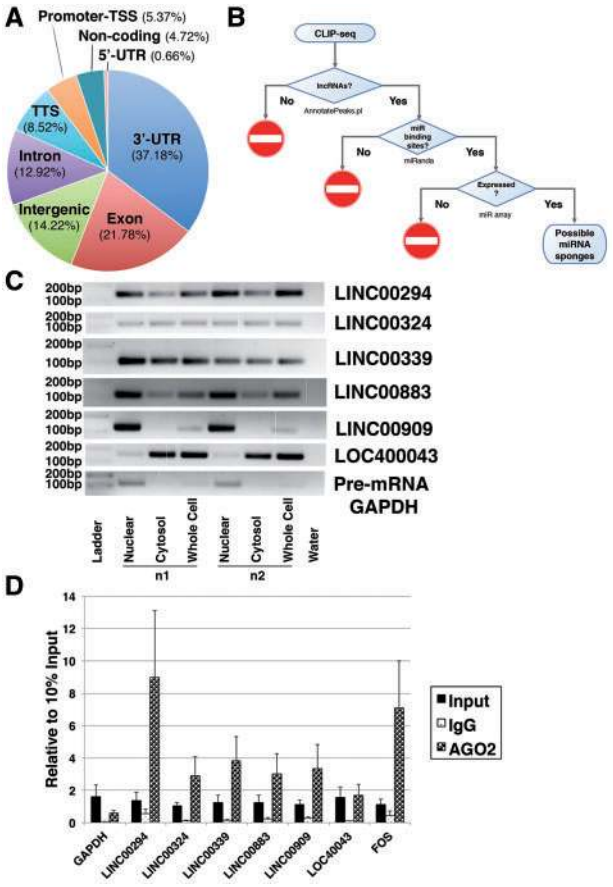

Figure 1. Screening of miRNA sponges. (**A**) Distribution of annotation classes of AGO-bound regions. (**B**) Flow chart of screening of potential miRNA sponges. (**C**) Cellular localization of potential miRNA sponges. cDNAs were prepared from whole cell, nuclear and cytoplasm fractions of HEK-293 cells after the DNase I treatment. Pre-mRNA of *GAPDH* (OTTHUMT00000268059; targeting the intron between its exons 2 and 3) was used as a control. $n = 2$. (**D**) AGO RIP-PCR of potential miRNA sponges. The results of quantitative RT-PCR (qRT-PCR) were normalized to the Ct value of 10% input for each condition. $n = 3$. Two types of negative controls were used: (1) anti-IgG antibody samples as a negative control for RIP assay; and (2) a primer pair against *GAPDH* (designed against exons 1∼3 of OTTHUMT00000268059) as a negative control for AGO binding. *FOS* is used as a positive control, whose primer pair was provided with the RIP kit. A colour version of this figure is available at BIB online: https://academic.oup.com/bib.

potential miRNA sponges, the number of AGO-bound regions was normalized to the length of the transcript that is the longest of the corresponding gene (Supplementary Table S4). Next, using the bound regions, potential miR binding sites were predicted via miRanda [27] (Supplementary Table S5). For a miR to bind, it must be expressed in the target cell. To this end, we performed miR array experiment using HEK-293 cells. To obtain a threshold value to determine the expressed miRs, we further performed miR array experiment using human foreskin fibroblast cell line Hs68 and HUVEC. Using the average RMA-normalized values of miRs among three cell lines (= 2.33), 237 miRs were selected (Supplementary Table S6). Through the results of miR arrays, the predicted miR binding sites in the lncRNAs were further filtered (Figure 1B). This list is considered as a final list for the experimental validations below, which contains 87 lncRNAs (Supplementary Table S7). Of note, this list contains the already known miRNA sponge *MALAT1* [13, 14].

To avoid ambiguity in the annotation itself, long intergenic ncRNAs (lincRNAs), which are far away from protein-coding genes (in comparison with others, such as sense overlapping lncRNAs), were chosen for the validation experiments. In the final list of 87 lncRNAs with the predicted miR binding sites (Supplementary Table S7), there are 25 lincRNAs. From these lincRNAs, the following six lincRNAs were chosen based on their high expression levels in human kidney using our previously released C-It-Loci knowledge database [38]: *LINC00294*, *LINC00324*, *LINC00339*, *LINC00883*, *LINC00909* and *LOC400043*. First, the subcellular localization of potential miRNA sponges was evaluated. To rule out the possible contamination by genomic DNA after the DNase I treatment, PCR experiment was performed using RT-samples and primer pair targeting pre-mRNA of *GAPDH* (Supplementary Figure S1). Although *LINC00909* is expressed exclusively in the nucleus of HEK-293 cells, all the other lncRNAs are found in the nucleus and cytoplasm (Figure 1C). Next, the binding between the selected lincRNAs and AGO was confirmed by RIP followed by quantitative RT-PCR experiment (RIP-PCR) (Figure 1D).

If indeed, the selected lincRNAs function as miRNA sponges, they should have biological functions. To elucidate this point, two of the above selected lincRNAs were selected for further analysis and silenced using siRNAs followed by profiling for gene expression changes using microarrays. In the case of *LINC00324*, which is expressed both in the nucleus and cytoplasm of HEK-293 cells (Figure 1C), an efficient silencing was achieved by two siRNAs [89.97 ± 3.55% down with $P = 0.0009$ and 67.44 ± 2.77% down with $P = 0.0060$ for siLINC00324.1 and siLINC00324.2, respectively, compared with scramble control (siScr)] (Figure 2A). When molecular profiles were examined by microarrays (Supplementary Table S8), 68 up- and 94 down-regulated genes were selected at the threshold ratio of above 1.4-fold and $P < 0.05$ (Figure 2B). To these differentially expressed genes, Gene Ontology (GO) analysis was performed, which indicated the enrichment of GO terms related to metabolism (Figure 2C) and to cell death (Figure 2D) for up- and down-regulated genes, respectively.

In the similar manner, *LOC400043*, which is highly expressed in the cytoplasm of HEK-293 cells (Figure 1C), was efficiently silenced by two siRNAs (90.08 ± 0.48% down with $P = 0.0029$ and 93.23 ± 0.72% down with $P = 0.0028$ for siLOC400043.1 and siLOC400043.2, respectively, compared with siScr) (Figure 3A). Compared with the microarray result of the silencing of *LINC00324*, more differentially expressed genes were selected when the same threshold values were applied (above 1.4-fold and $P < 0.05$): 178 up- and 416 down-regulated genes (Supplementary Table S9). When GO analysis was performed, GO terms related to protein localization and transport are enriched for up-regulated genes (Figure 3C), while terms related to metabolism are enriched for down-regulated genes (Figure 3D), which suggests different biological effects after silencing of *LOC400043* compared with that of *LINC00324*.

The expected primary function of miRNA sponges is to bind to miRs by sequestering the otherwise available miRs for the translational inhibition of target protein-coding genes. Thus, we quantified for changes in miRs using miR microarrays on silencing of *LINC00324* or *LOC400043* using the same RNAs used for gene microarrays above (Figures 2 and 3). Among 1130 mature miRs on the miR array, 85 miRs were regulated by *LINC00324* (11 up- and 74 down-regulated miRs) (Figure 4A; Supplementary Table S10) and 84 miRs after *LOC400043* silencing (28 up- and 56 down-regulated miRs) (Figure 4B; Supplementary Table S11). Based on the previous screening (Figure 1B; Supplementary Table S7) combined with other miR prediction tools, there are 13 miRs with predicted binding sites in *LINC00324* and 5 miRs

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022

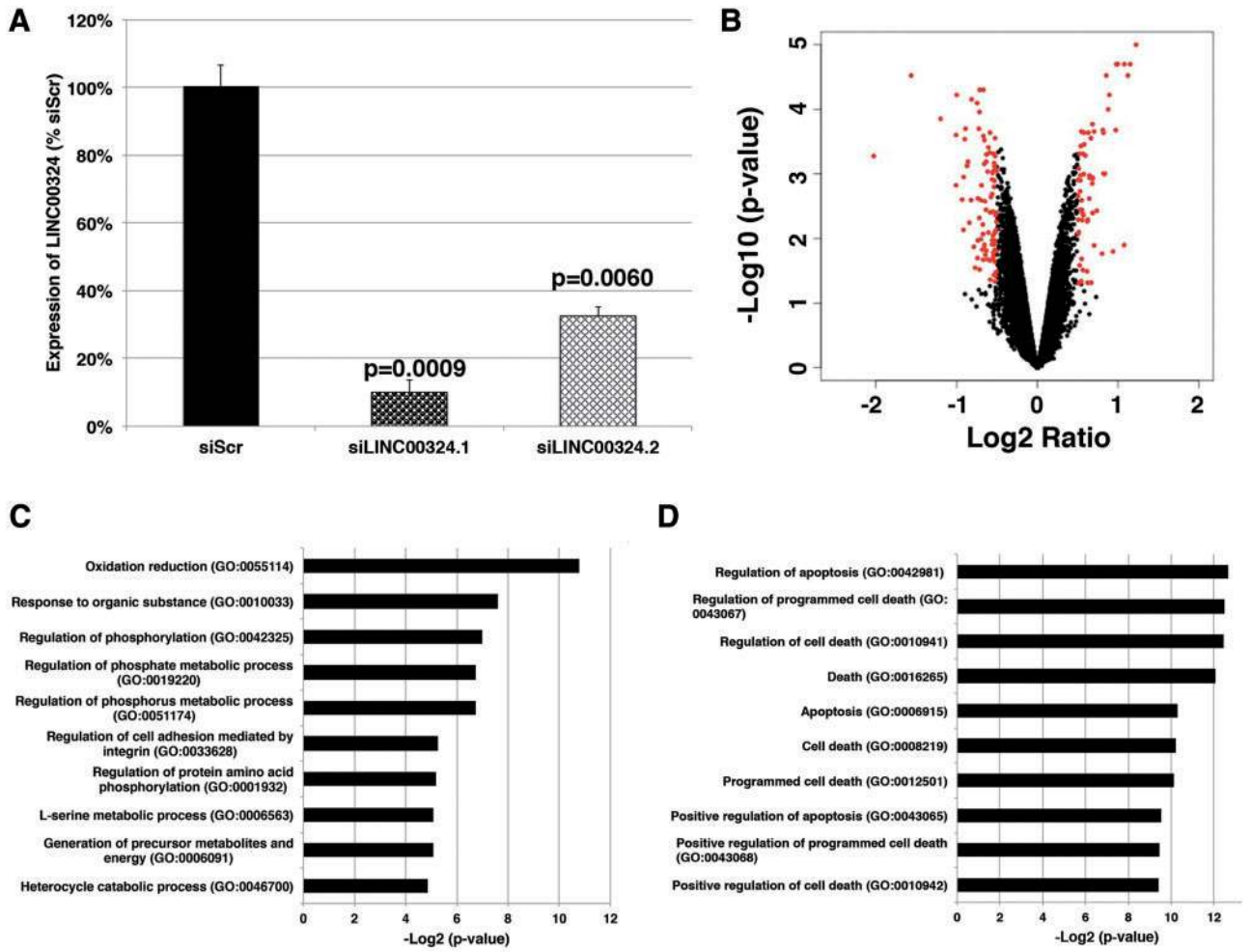

Figure 2. Silencing of *LINC00324*. (**A**) Efficiency of silencing quantified by qRT-PCR. $n = 3$. (**B**) Volcano plot of gene array results ($n = 2$). Genes whose expressions selected at 1.4-fold threshold with $P < 0.05$ are indicated in lighter color. (**C** and **D**) Top 10 GO terms in (**C**) up- and (**D**) down-regulated genes. All GO terms are categorized under biological processes. A colour version of this figure is available at BIB online: https://academic.oup.com/bib.

for *LOC400043* (Figure 4C). Among miRs that might be bound to *LINC00324*, hsa-miR-3656, hsa-miR-4448, hsa-miR-4449, hsa-miR-5572, hsa-miR-6751-5p and hsa-miR-8078 are not on the miR array. In the case of *LOC400043*, all predicted miRs are on the array. When their expression levels were examined on silencing of the corresponding miRNA sponges, some of them are down-regulated (Figure 4C and D). If indeed, a miR sponge function to capture miRs, then up-regulation of miRs is expected, which is not the case here. The down-regulated miRs include miR-615-5p (31.60% down compared with siScr with $P = 0.0113$), miR-675-5p (30.55% down; $P = 0.0477$) and miR-4281 (41.64% down; $P = 0.0002$) on silencing of *LINC00324*; whereas miR-28-3p (61.84% down; $P = 0.0281$) and miR-96-5p (41.60% down; $P = 0.0003$) are selected for *LOC400043* at the threshold ratio of above 1.4-fold and $P < 0.05$.

Surprised by the above findings, we screened miRNA:gene interactions from DIANA-TarBase v7.0 (http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index) [39] and miRTarBase 2016 (http://miRTarBase.mbc.nctu.edu.tw/) [40], which resulted in the identifications of 239 target protein-coding genes for miR-615-5p; 142 for miR-675-5p; 30 for miR-4281; 41 for miR-28-3p; and 991 for miR-96-5p. Next, the expression changes of the target protein-coding genes were screened from the previously performed gene array results. For this purpose, only those differentially expressed genes were chosen (Supplementary Tables S8 and S9). In the case of *LINC00324*, although no differentially expressed genes being identified as targets of miR-675-5p and miR-4281, two following differentially expressed genes were found for miR-615-5p: '*ADCYAP1R1*' (1.58-fold down-regulated with $P = 0.0143$) and '*BTG2*' (1.87-fold down-regulated with $P = 0.0003$). Because *BTG2* is known for its pro-apoptotic function [41], its down-regulation in siLINC000324 compared with siScr controls could be explained in relation to miR-615-5p and *BTG2* as the enrichment of GO terms related to cell death was observed (Figure 2D). In the case of *LOC400043*, more differentially expressed target genes are identified: two genes for miR-28-3p and 43 genes for miR-96-5p (Supplementary Table S12), where there are 23 up- and 21 down-regulated genes. When these differentially expressed genes are examined carefully, the list includes one chromatin remodeling factor (*JMJD1C*), three transcription factors (*CREB3L2*, *RELA* and *THAP2*) and two transcription co-factors (*CARM1* and *GREB1*) based on the information provided by the AnimalTFDB (http://www.bioguo.org/AnimalTFDB/index.php) [42] (Figure 4E). This fits nicely to the results of gene arrays, which showed the enrichment of GO terms related to various biological processes, including metabolisms for down-regulated genes (Figure 3C and D). To clarify the above predictions, further functional studies are necessary.

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022

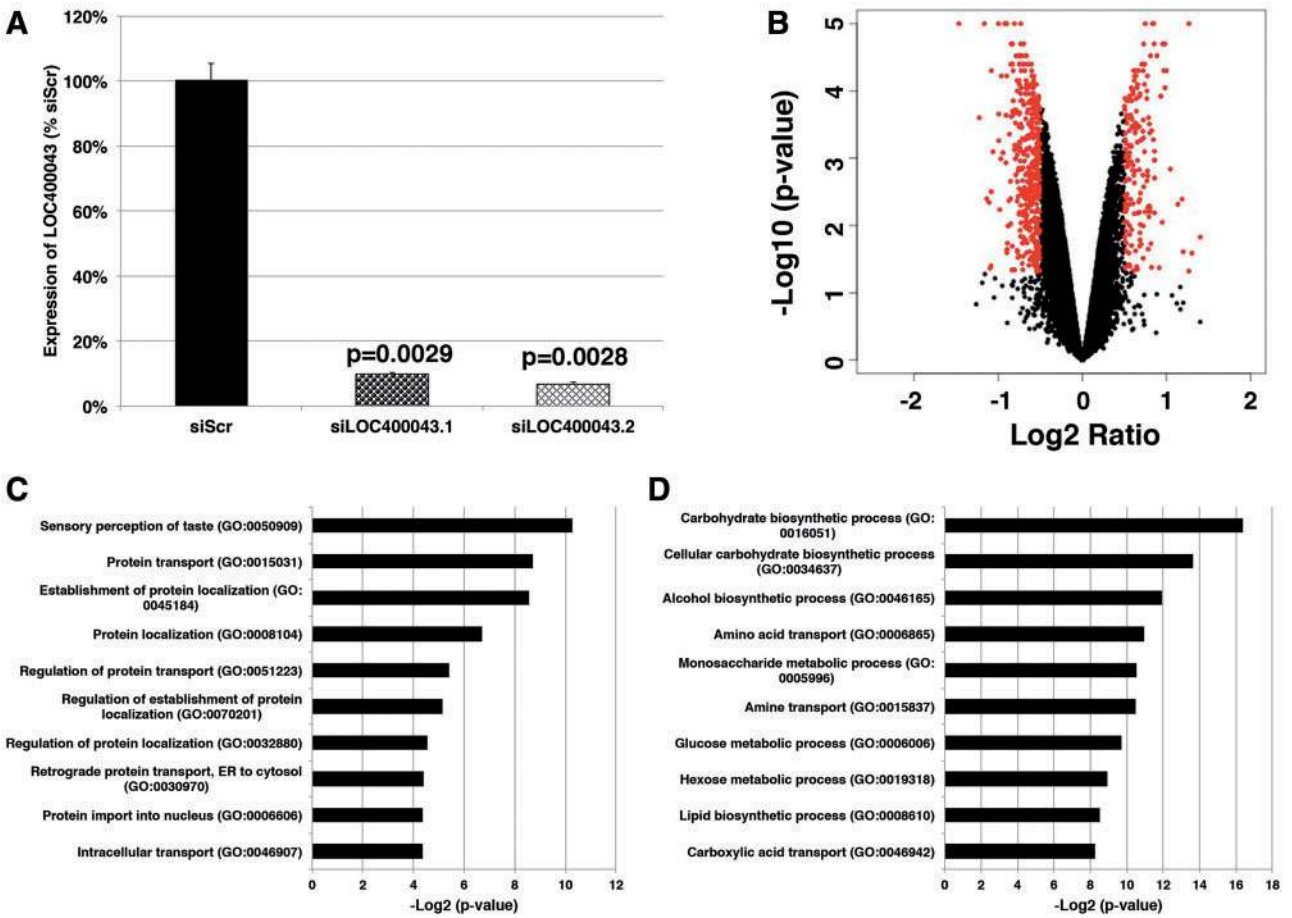

Figure 3. Silencing of *LOC400043*. (**A**) Efficiency of silencing quantified by qRT-PCR. $n = 3$. (**B**) Volcano plot of gene array results ($n = 2$). Genes whose expressions selected at 1.4-fold threshold with $P < 0.05$ are indicated in lighter color. (**C** and **D**) Top 10 GO terms in (**C**) up- and (**D**) down-regulated genes. All GO terms are categorized under biological processes. A colour version of this figure is available at BIB online: https://academic.oup.com/bib.

Screening of miRNA sponges in circRNAs

Recently, we reported the existence of circRNAs in the cardiovascular system [25]. It was reported that some circRNAs could function as miRNA sponges [43–45]. To test this point, the list of circRNAs was downloaded from the circBase database (http://www.circbase.org) [46] and screened for AGO-bound regions. Of 92 375 circRNAs in the circBase database, 58 063 circRNAs are found to own AGO-bound regions (Supplementary Table S13). This high amount of matches is owing to the fact that circRNAs arise during the splicing of protein-coding genes, which might have AGO-binding sites in their sequences. To validate whether the above-identified circRNAs function as miRNA sponges, first, RT-PCR experiment was performed to observe the cellular localization. For this purpose, the following six candidates were chosen based on their high numbers of AGO-bound regions and their presence in the circBase database, which we use the circBase IDs hereafter: *hsa_circ_0000284*, *hsa_circ_0001417*, *hsa_circ_0005939*, *hsa_circ_0007292*, *hsa_circ_0008887* and *hsa_circ_0013647* (Figure 5A). All the primer pairs for the detection of circRNAs were designed against the circRNA-specific back-splice sites. Furthermore, to confirm the circularization of mRNAs, the RNase R digestion was performed to digest linear RNAs and to keep only circRNAs. Among six circRNAs, four are clearly circular in the cytosol while *hsa_circ_0013647* and *hsa_circ_0005939* did not withstand the treatment with RNase R. Next, to prove that four circRNAs function as miRNA sponges, RIP-PCR was performed (Figure 5B). Compared with the linear miRNA sponges (Figure 1D), although the enrichment of AGO binding compared with IgG control was observed for all samples, such enrichment was significantly less than the negative control using the primer pair against exons 1∼3 of *GAPDH*, suggesting that these circRNAs might not be *bona fide* miRNA sponges.

Discussion

In this study, we screened for lincRNAs and circRNAs that function as miRNA sponges using the published data and databases. To validate our bioinformatics screening, we conducted experiments to detect lincRNAs and circRNAs as well as their bindings to RISC. Furthermore, loss-of-function experiments were performed to elucidate the roles of miRNA sponges. In the case of circRNAs, we extended our previous study in endothelial cells [25] to understand a general trend of circRNAs that function as miRNA sponges.

The recent computational study reported that of 7112 human circRNAs identified in the ENCODE data sets, only two circRNAs own more predicted miR binding sites than expected by chance [47]; thus, arguing against the generalization of circRNAs as miRNA sponges. This argument was confirmed at least for our selected circRNAs (Figure 5). However, the situation is different for linear lincRNAs with AGO binding sites (Figure 1).

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022

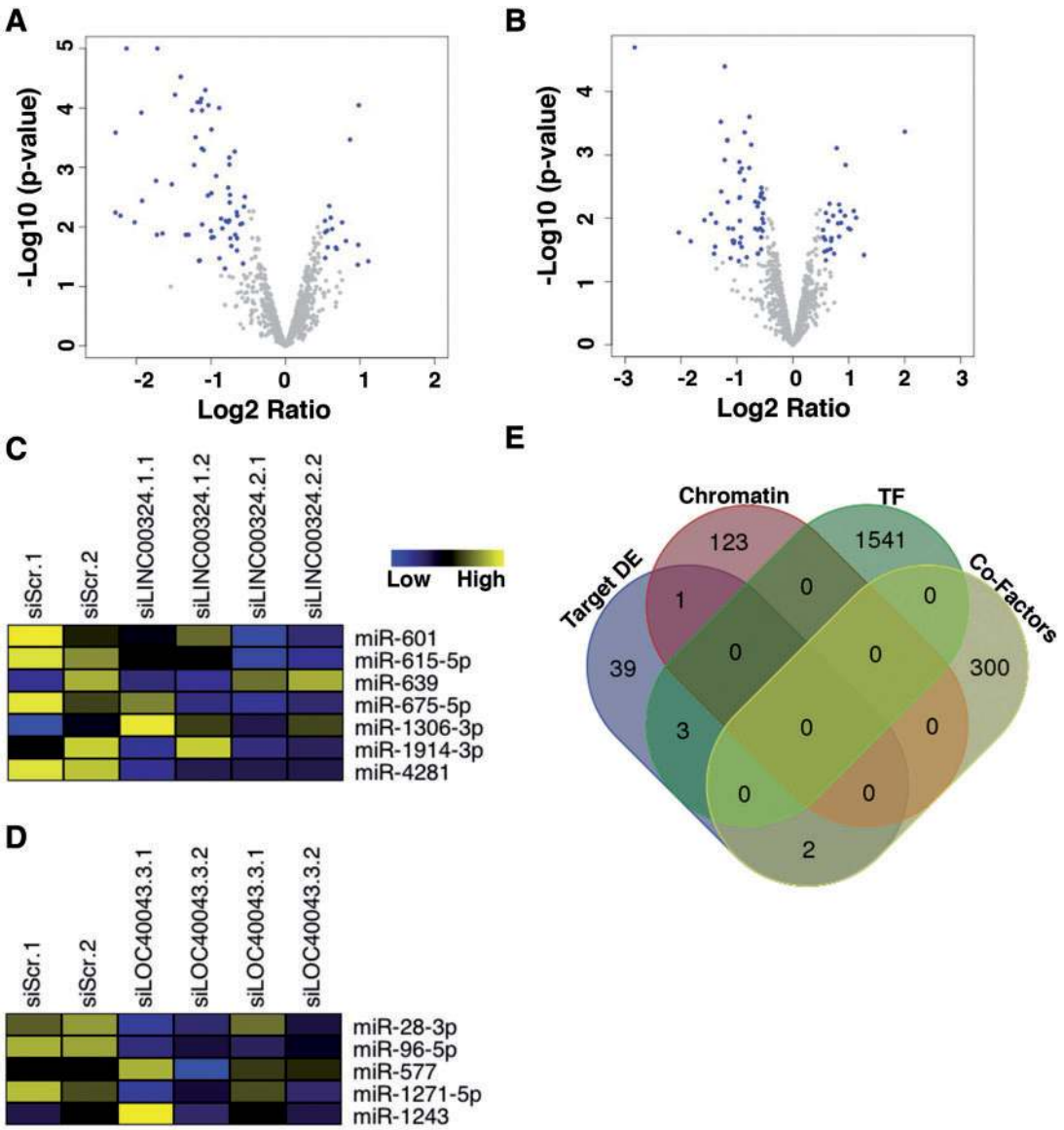

Figure 4. Expressions of miRs and their targets. (**A** and **B**) Volcano plots of miR array results ($n = 2$) on silencing of (**A**) *LINC00324* or (**B**) *LOC400043*. Mature miRs whose expressions selected at 1.4-fold threshold with $P < 0.05$ are indicated in lighter color. (**C** and **D**) Heat maps of miR expressions on silencing of (**C**) *LINC00324* or (**D**) *LOC400043*. (**E**) Venn diagram showing the overlap among target DE genes ('Target DE') and the following information provided by the AnimalTFDB [42]: chromatin remodeling factors ('Chromatin'); transcription factors ('TF'); and co-factors ('Co-Factors').

In contrast to our expectations, silencing of linear lincRNAs, *LINC00324* and *LOC400043*, led to a down-regulation of miRs with predicted binding sites. Although one could not rule out the inaccuracy of prediction of miR binding sites, it is possible that the fundamental concept of direct binding of miRs to miRNA sponges to sequester miRs might not be always correct as demonstrated in this study. This might stem from the current method of the detection of miRs that is not suitable as the binding of miRs to the miR sponge could not be detected via microarrays or any other techniques unless those not-bound miRs must be removed from the sample under investigation, which is hard to achieve with the current technology that is available.

Although miRNA sponges are an attractive way to control the levels of available miRs, a caution must be taken to interpret the binding of RISC to the linear lincRNA, although silencing of potential miRNA sponges in this study showed altered biological processes, which suggest the importance of these lincRNAs. In conclusion, although the canonical mechanism of RISC cannot explain its binding to lncRNAs, more vigorous testing is necessary to understand how such binding could result in the regulation of available miRs through the function of miRNA sponges.

Key Points

- Although several databases for the interaction between miRs and lncRNAs are available, the defined biological proof is not provided for the most of such interactions.
- A series of bioinformatics methods can indicate lncRNAs acting as miRNA sponges.
- Most circRNAs do not function as miRNA sponges.

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022

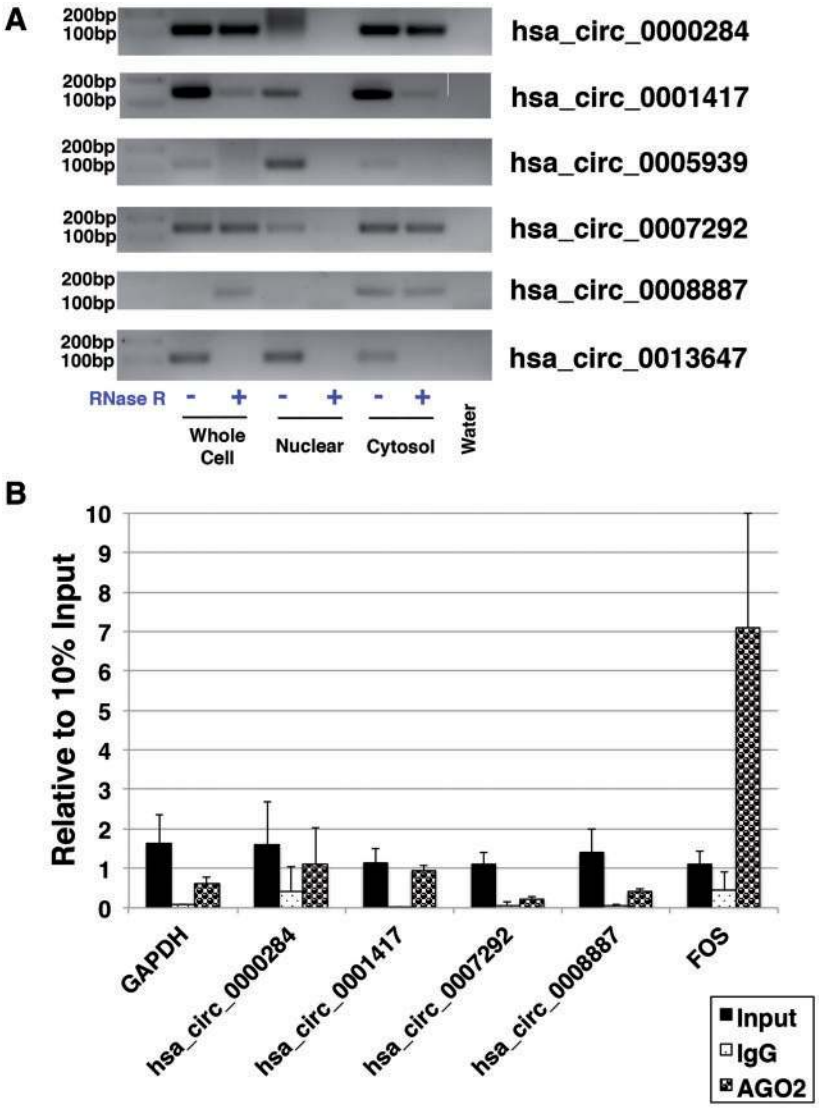

Figure 5. circRNAs and their binding to RISC. (**A**) Circularization of circRNAs. The presence and circularization of circRNAs were tested by subfracturation of HEK-293 cells and treatment of total RNA with RNase R. (**B**) AGO RIP-PCR of potential miRNA sponges. The results of quantitative RT-PCR (qRT-PCR) were normalized to the Ct value of 10% input for each condition. $n = 3$. Two types of negative controls were used: (1) anti-IgG antibody samples as a negative control for RIP assay; and (2) a primer pair against *GAPDH* (designed against exons 1∼3 of OTTHUMT00000268059) as a negative control for AGO binding. *FOS* is used as a positive control, whose primer pair was provided with the RIP kit. A colour version of this figure is available at BIB online: https://academic.oup.com/bib.

Supplementary data

Supplementary data are available online at http://bib.oxfordjournals.org/.

Acknowledgements

We thank Wenjun Jin and Andrea Knau for excellent technical assistances.

Funding

The study was supported by the LOEWE Center for Cell and Gene Therapy (State of Hessen) (to S.U. and S.D.); the German Center for Cardiovascular Research (DZHK) (to S.U. and S.D.); the DFG (SFB834) (to S.U. and S.D.); the Excellence Cluster Cardio-Pulmonary System (ECCPS) (to S.D.); and the ERC Advanced Grant 'Angiolnc' (to S.D.). The authors declare no competing financial interests.

References

1. Lander ES, Linton LM, Birren B, *et al*. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921.
2. Mercer TR, Gerhardt DJ, Dinger ME, *et al*. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat Biotechnol* 2012;**30**:99–104.
3. Uchida S, Dimmeler S. Long noncoding RNAs in cardiovascular diseases. *Circ Res* 2015;**116**:737–50.
4. Uchida S, Gellert P, Braun T. Deeply dissecting stemness: making sense to non-coding RNAs in stem cells. *Stem Cell Rev* 2012;**8**:78–86.
5. Beavers KR, Nelson CE, Duvall CL. MiRNA inhibition in tissue engineering and regenerative medicine. *Adv Drug Deliv Rev* 2015;**88**:123–37.

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022

6. Seeger FH, Zeiher AM, Dimmeler S. MicroRNAs in stem cell function and regenerative therapy of the heart. *Arterioscler Thromb Vasc Biol* 2013;**33**:1739–46.
7. Boon RA, Dimmeler S. MicroRNAs in myocardial infarction. *Nat Rev Cardiol* 2015;**12**:135–42.
8. Kallen AN, Zhou XB, Xu J, *et al*. The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol Cell* 2013;**52**:101–12.
9. Imig J, Brunschweiger A, Brummer A, *et al*. miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat Chem Biol* 2015;**11**:107–14.
10. Cesana M, Cacchiarelli D, Legnini I, *et al*. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 2011;**147**:358–69.
11. Zhou X, Gao Q, Wang J, *et al*. Linc-RNA-RoR acts as a "sponge" against mediation of the differentiation of endometrial cancer stem cells by microRNA-145. *Gynecol Oncol* 2014;**133**:333–9.
12. Guo G, Kang Q, Zhu X, *et al*. A long noncoding RNA critically regulates Bcr-Abl-mediated cellular transformation by acting as a competitive endogenous RNA. *Oncogene* 2015;**34**:1768–79.
13. Han X, Yang F, Cao H, *et al*. Malat1 regulates serum response factor through miR-133 as a competing endogenous RNA in myogenesis. *FASEB J* 2015;**29**:3054–64.
14. Leucci E, Patella F, Waage J, *et al*. microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus. *Sci Rep* 2013;**3**:2535.
15. Yan B, Yao J, Liu JY, *et al*. lncRNA-MIAT regulates microvascular dysfunction by functioning as a competing endogenous RNA. *Circ Res* 2015;**116**:1143–56.
16. Song X, Cao G, Jing L, *et al*. Analysing the relationship between lncRNA and protein-coding gene and the role of lncRNA as ceRNA in pulmonary fibrosis. *J Cell Mol Med* 2014;**18**:991–1003.
17. Das S, Ghosal S, Sen R, *et al*. lnCeDB: database of human long noncoding RNA acting as competing endogenous RNA. *PLoS One* 2014;**9**:e98965.
18. Denzler R, Agarwal V, Stefano J, *et al*. Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol Cell* 2014;**54**:766–76.
19. Ule J, Jensen KB, Ruggiu M, *et al*. CLIP identifies nova-regulated RNA networks in the brain. *Science* 2003;**302**:1212–5.
20. Dueck A, Meister G. Assembly and function of small RNA - argonaute protein complexes. *Biol Chem* 2014;**395**:611–29.
21. Li JH, Liu S, Zhou H, *et al*. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res* 2014;**42**:D92–7.
22. Park C, Yu N, Choi I, *et al*. lncRNAtor: a comprehensive resource for functional investigation of long non-coding RNAs. *Bioinformatics* 2014;**30**:2480–5.
23. Yuan J, Wu W, Xie C, *et al*. NPInter v2.0: an updated database of ncRNA interactions. *Nucleic Acids Res* 2014;**42**:D104–8.
24. Yang YC, Di C, Hu B, *et al*. CLIPdb: a CLIP-seq database for protein-RNA interactions. *BMC Genomics* 2015;**16**:51.
25. Boeckel JN, Jae N, Heumuller AW, *et al*. Identification and characterization of hypoxia-regulated endothelial circular RNA. *Circ Res* 2015;**117**:884–90.
26. Heinz S, Benner C, Spann N, *et al*. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010;**38**:576–89.
27. Betel D, Wilson M, Gabow A, *et al*. The microRNA.org resource: targets and expression. *Nucleic Acids Res* 2008;**36**:D149–53.
28. Michalik KM, You X, Manavski Y, *et al*. Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circ Res* 2014;**114**:1389–97.
29. Untergasser A, Cutcutache I, Koressaar T, *et al*. Primer3–new capabilities and interfaces. *Nucleic Acids Res* 2012;**40**:e115.
30. Gellert P, Ponomareva Y, Braun T, *et al*. Noncoder: a web interface for exon array-based detection of long non-coding RNAs. *Nucleic Acids Res* 2013;**41**:e20.
31. Gellert P, Teranishi M, Jenniches K, *et al*. Gene Array Analyzer: alternative usage of gene arrays to study alternative splicing events. *Nucleic Acids Res* 2012;**40**:2414–25.
32. Irizarry RA, Hobbs B, Collin F, *et al*. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;**4**:249–64.
33. Ritchie ME, Phipson B, Wu D, *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;**43**:e47.
34. Gautier L, Cope L, Bolstad BM, *et al*. affy–analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004;**20**:307–15.
35. Huber W, von Heydebreck A, Sultmann H, *et al*. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 2002; **18 (Suppl 1)**:S96–104.
36. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 2014;**42**:D68–73.
37. Stepanenko AA, Dmitrenko VV. HEK293 in cell biology and cancer research: phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. *Gene* 2015;**569**: 182–90.
38. Weirick T, John D, Dimmeler S, *et al*. C-It-Loci: a knowledge database for tissue-enriched loci. *Bioinformatics* 2015;**31**: 3537–43.
39. Vlachos IS, Paraskevopoulou MD, Karagkouni D, *et al*. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Res* 2015;**43**:D153–9.
40. Chou CH, Chang NW, Shrestha S, *et al*. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res* 2016;**44**:D239–47.
41. Mao B, Zhang Z, Wang G. BTG2: a rising star of tumor suppressors (review). *Int J Oncol* 2015;**46**:459–64.
42. Zhang HM, Liu T, Liu CJ, *et al*. AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. *Nucleic Acids Res* 2015;**43**:D76–81.
43. Hansen TB, Wiklund ED, Bramsen JB, *et al*. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J* 2011;**30**:4414–22.
44. Hansen TB, Jensen TI, Clausen BH, *et al*. Natural RNA circles function as efficient microRNA sponges. *Nature* 2013;**495**: 384–8.
45. Memczak S, Jens M, Elefsinioti A, *et al*. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 2013;**495**:333–8.
46. Glazar P, Papavasileiou P, Rajewsky N. circBase: a database for circular RNAs. *RNA* 2014;**20**:1666–70.
47. Guo JU, Agarwal V, Guo H, *et al*. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol* 2014;**15**:409.

Downloaded from https://academic.oup.com/bib/article/18/5/780/2562768 by guest on 20 August 2022