Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

TargetLink, a new method for identifying the endogenous target set of a specific microRNA in intact living cells

Yan Xu [ID][a], Yan Chen [ID][a], Daliang Li[a], Qing Liu[a], Zhenyu Xuan[b], and Wen-Hong Li[a]

[a]Department of Cell Biology and of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX, USA; [b]Department of Biological Sciences, Center for Systems Biology, The University of Texas at Dallas, Richardson, TX, USA

ABSTRACT

MicroRNAs are small non-coding RNAs acting as posttranscriptional repressors of gene expression. Identifying mRNA targets of a given miRNA remains an outstanding challenge in the field. We have developed a new experimental approach, TargetLink, that applied locked nucleic acid (LNA) as the affinity probe to enrich target genes of a specific microRNA in intact cells. TargetLink also consists a rigorous and systematic data analysis pipeline to identify target genes by comparing LNA-enriched sequences between experimental and control samples. Using miR-21 as a test microRNA, we identified 12 target genes of miR-21 in a human colorectal cancer cell by this approach. The majority of the identified targets interacted with miR-21 via imperfect seed pairing. Target validation confirmed that miR-21 repressed the expression of the identified targets. The cellular abundance of the identified miR-21 target transcripts varied over a wide range, with some targets expressed at a rather low level, confirming that both abundant and rare transcripts are susceptible to regulation by microRNAs, and that TargetLink is an efficient approach for identifying the target set of a specific microRNA in intact cells. C20orf111, one of the novel targets identified by TargetLink, was found to reside in the nuclear speckle and to be reliably repressed by miR-21 through the interaction at its coding sequence.

Abbreviations: LNA, locked nucleic acid; miRNA, microRNA; Ago, Argonaute protein; SEC-HPLC, size exclusion chromatography – high pressure liquid chromatography; GT buffer, guanidinium thiocyanate buffer; KO, knockout; WT, wild type; UV, ultraviolet; WTX, wild type cells UV-crosslinked; WTnX, wild type cells not UV-crosslinked; KOX, miR-21 knockout cells UV crosslinked; 3′-UTR, 3′-untranslated region

Introduction

microRNAs are small noncoding RNAs that play important roles in regulating gene expression. A specific microRNA can target multiple genes, and can perform diverse roles in development, physiology and/or disease of different cell types.[1] Because of their complex involvement in numerous pathways and in different cell types, it is important to identify targets of a given microRNA in specific cell types and under specific conditions. Despite extensive efforts to develop both computational and experimental approaches to predict or identify miRNA targets, target identification of a given miRNA remains an outstanding challenge in the field.[2] A number of bioinformatics programs predict miRNA targets on a genomic scale exploiting the rule of "seed matching" – conserved Watson-Crick pairing to the 5′ region of miRNA centered on nucleotides 2–7, which is called the miRNA "seed".[3,4] The various algorithms developed thus far, however, often yield different sets of predictions that do not overlap.[3,5] (Fig. S1). Further, computational methods rarely take into the consideration of how the variation in biologic contexts may affect RNA conformation or RNA-protein interaction, both of which can influence the accessibility of RNAs and hence their engagement with the micro-RNA/Argonaute silencing complex[6-11] Established genome-wide experimental methods for identifying targets of endogenous microRNAs include high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (CLIP).[12-14] In this method, the isolated Argonaute protein (Ago) binding sites are analyzed computationally to infer the possible microRNA(s) that could have been responsible for tethering Ago to those binding sites. However, at the miRNome level, it is not always straightforward to pair the isolated Ago-binding sequences with a specific microRNA based on seed matching, especially when considering that non-canonical microRNA::target interactions represent a rather frequent feature of microRNA targeting.[2,15-22] More recently, a crosslinking and ligation approach termed CLASH was reported in which crosslinked miRNA-target sequence are ligated to generate chimeras for sequencing.[22] This method should overcome the ambiguity in assigning the Ago-bound sequence to the microRNA responsible for recruiting the target gene to the RISC complex.[21] A drawback of the method, however, is its relatively low efficiency of chimera generation and

capture.[23,24] In addition to the aforementioned techniques for identifying the native targets of endogenous microRNAs, target discovery based on affinity pulldown after cellular delivery of exogenous microRNA mimetics has also been reported.[25-28] A potential caveat of these methods is that exogenously delivered miRNAs do not necessarily engage with the same set of targets as the endogenous miRNAs.

Despite these important advances in target identification on a miRNome level, there are still pressing needs to develop new experimental methods to improve the efficiency and accuracy of target identification. Moreover, since the majority of studies on microRNA focus on functional analysis of a single microRNA of interest, methods designed to recover the sequences directly bound by a specific endogenous microRNA in intact living cells should greatly facilitate such investigations.

We have developed a new approach, TargetLink, for identifying target genes directly bound by a specific microRNA. TargetLink involves crosslinking microRNA and its target genes in intact live cells by UV illumination. Cell lysates containing the crosslinked microRNA-Argonaute-mRNA ternary complexes were prepared and subjected to affinity purification using locked nucleic acid (LNA) as the capture probe to pull down complementary microRNA and its crosslinked target genes. The affinity-purified RNAs were then deproteinized, converted to cDNAs, and sequenced. Using miR-21 as a test microRNA, we have also developed a data analysis pipeline by comparing the sequenced reads of the crosslinked sample with the control sample without UV illumination, as well as with the crosslinked sample from an isogenic cell line lacking miR-21. Our analysis identified 12 target genes of miR-21 in a human colorectal cancer cell. Remarkably, the expression level of these 12 target genes varied over a wide range, confirming that both abundant and rare transcripts are susceptible to regulation by microRNAs. Among several novel miR-21 targets identified by TargetLink, C20orf111 was reliably repressed by miR-21 through the interaction at its coding region.

Results

TargetLink exploits LNA affinity purification to enrich microRNA and its associated target RNAs

To develop TargetLink, we chose miR-21 as a testing microRNA. miR-21 has emerged as an important microRNA involved in tumorigenesis. A large-scale survey to determine the miRNA signature of 540 tumor samples—including lung, breast, stomach, prostate, colon, and pancreatic tumors and their respective normal adjacent tissues — revealed miR-21 was the only miRNA upregulated in all these tumors.[29] miR-21 dysregulation is now recognized as a rather general feature in a variety of tumors. In a mouse model, overexpression of miR-21 led to a pre-B-cell lymphoma whose maintenance was also dependent on miR-21, demonstrating that mir-21 was a genuine oncogene.[30]

We chose RKO cells, a human colorectal cancer cell line expressing miR-21,[31] as a model system for developing TargetLink. Another reason for choosing this cell was that a miR-21 knockout cell line, RKO[miR-21(−/−)], has been generated by the targeted deletion of mature miR-21 sequence in the primary miR-21 locus[31] (Fig. S2A). The RKO[miR-21(−/−)] cell line hence served as a good control for evaluating the quality and specificity of affinity purification against miR-21.

We devised TargetLink based on the principles of RNA photo crosslinking and antisense purification of crosslinked RNA complexes. Since microRNAs are only ∼22 nt long, we applied locked nucleic acid (LNA) as the affinity probe to pull down microRNA and associated protein complexes. LNA oligonucleotides bind to complementary RNA with very high affinity and specificity to yield remarkably stable LNA/RNA duplexes,[32] and antisense LNA oligonucleotides as short as 8-mer have been successfully used to knock down miRNA.[33] The high stability of LNA::miR-21 duplex would allow us to perform the hybridization and LNA bead washes at high stringency to remove non-specifically bound RNAs. The workflow of TargetLink consists of 5 major steps (Fig. 1A). In the first step, cells were irradiated at 254 nm to induce RNA-protein crosslink to form covalently linked ternary complexes containing microRNA/Argonaute/target[12]. After digesting the cell lysate with DNase and limited RNase, we homogenized the resulting lysate in a potent chaotropic buffer (4 M guanidinium thiocyanate containing 10 mM DTT, or GT buffer) to denature proteins and to disrupt non-covalent protein-RNA complexes. Since UV illumination generated protein-RNA crosslinks with only modest efficiency, we expected that a sizable portion of microRNAs in the cell lysate remained to be uncrosslinked. Anticipating that these free microRNAs would likely lower the efficiency of the subsequent LNA affinity purification of crosslinked miR-21/Ago/target complexes, we devised a 2-step procedure to remove free microRNAs including miR-21: first, the homogenized sample was filtered by membrane diafiltration (molecular weight cutoff of 50 KDa) to enrich materials with molecular weight higher than 50 KDa; second, we applied size exclusion HPLC (SEC-HPLC) to deplete free microRNAs and to enrich high molecular weight proteins and protein-RNA complexes. Samples of each SEC-HPLC fraction were deproteinated using proteinase K and analyzed by Northern blot analysis. miR-21 crosslinked to larger molecules was eluted off from the column very early (8–10 minutes), whereas free miR-21 eluted much later (13–15 minutes)(Fig. 1B, C; and Fig. S2B). This separation efficiently enriched the crosslinked miR-21 complexes by removing a sizable amount of small RNA fragments (Fig. S3).

In the second step of TargetLink, we added a biotinylated LNA probe to the SEC-HPLC purified sample to perform affinity purification. This step is the key to TargetLink and distinguishes it from previous CLIP approaches: by using an LNA probe that is complementary to a specific miRNA, we specifically recover target RNAs of that particular miRNA. Since the LNA::miR-21 duplex is expected to be very stable, we performed the hybridization at high stringency (4M GT buffer, 95°C to 43°C in 1.5 h) to reduce non-specifically bound RNAs. Next, we washed the captured miR-21 complexes thoroughly (4M GT buffer, 43°C for 1 min, 5 repetitions) before eluting it off the bead. The unbound ("wash") and bound ("eluent") fractions were deproteinated by proteinase K, and analyzed by denaturing gel electrophoresis and Northern blot. The results (Fig. 1D) showed that the TargetLink affinity step efficiently enriched for miR-21 (crosslinked) compared with a scramble control LNA oligonucleotide.
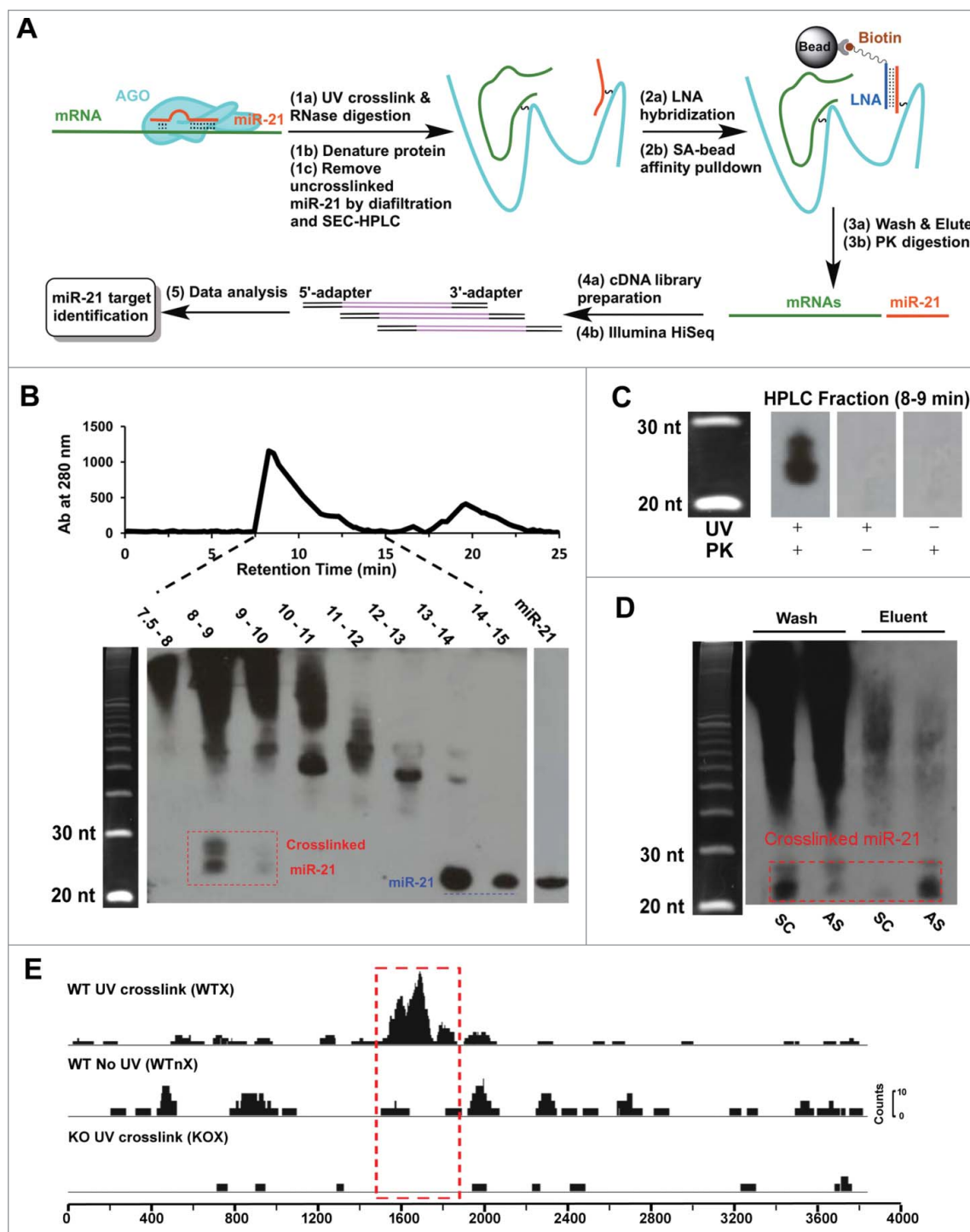
Figure 1. Development of TargetLink for identifying the target genes of a specific microRNA. (A) Workflow of TargetLink illustrated with miR-21 as an example. Key steps include removal of free (uncrosslinked) small RNAs by SEC-HPLC and affinity purification of crosslinked ternary complexes with a biotinylated LNA probe antisense to miR-21. See text for details. (B) Removal of free miR-21 by SEC-HPLC. HPLC eluents (top, monitored at UV absorbance at 280 nm) were fractionated, de-proteinated and analyzed by Northern blot (bottom) using a biotinylated probe against miR-21. Cross-linked products were eluted between 8 min and 10 min; whereas free miR-21 was eluted out after 13 min. (C) Northern blot of the HPLC fraction (8 – 9 min, B) detected the crosslinked miR-21 bands only when using the UV-illuminated and deproteinated sample (proteinase K, or PK). (D) Purification of cross-linked miR-21 (and its associated Ago and mRNA targets) via LNA affinity purification using either a miR-21 antisense LNA (AS) or a scramble control LNA (SC). The majority of crosslinked miR-21 was detected in the bound-and-eluted fraction (lane 4) when using AS LNA but not SC LNA. (E) An example of RNA-Seq reads mapped to a miR-21 target gene (MKNK2) using LNA (miR-21 antisense) affinity purified samples from wild type RKO cell (WT) or miR-21 knockout RKOcell (KO). The highlighted region contained RNA-Seq reads that were selectively enriched in crosslinked WT cells.

Finally, we performed deep sequencing using the amplified cDNA library of the affinity purified RNAs. Sequenced reads with exactly the same sequence were collapsed to a single read and mapped to the NCBI Refseq human transcriptome using

NovoAlign. We then removed the potential PCR duplicates by condensing reads that were aligned to the same position on the RefSeq transcriptome. The resulting file containing the unique mapped reads reflected the distribution of affinity purified RNA fragments

across the transcriptome, and it was used for detecting the enrichment of sequences resulting from LNA affinity purification. Fig. 1E shows an example of the reads mapped to a candidate miR-21 target gene MKNK2 (vide infra). Compared to the control samples such as uncrosslinked wild type cells (WTnX) or crosslinked miR-21 knockout cells (KOX), crosslinked wild type RKO cells (WTX) contained RNA-Seq reads that were selectively enriched in a segment of MKNK2 transcript over the 2 controls.

Data analysis to identify miR-21 targets from the enriched sequences

To identify miR-21 target genes using the deep sequencing data from the affinity purified WTX sample, we devised a data analysis pipeline containing the following steps: (1) First, we divided each of the annotated human RefSeq transcripts into 200-nt bins and summed the number of mapped reads falling within each bins. Bins containing few mapped read were filtered and only bins containing 2 or more unique mapped reads in at least 2 thirds of the biologic replicates were kept (Filter 1, Fig. S4A). (2) Second, to detect the enrichment of the remaining bins resulting from the affinity purification, we compared the number of mapped reads within each bin between the LNA-purified WTX sample and the input sample (HPLC fractions eluted between 8 min and 10 min and contained the crosslinked miR-21, Fig. 1B), or between the LNA-purified WTX sample and the LNA-purified WTnX control. The detail statistics model can be found in the Methods. In brief, for each of the 6 biologic replicates of affinity purified WTX samples, we calculated the P-value of reads enrichment in LNA-purified WTX sample by using the Poisson distribution. The λ values used in Poisson were derived either from the input sample (for calculating the P-value of enrichment against the input, $P^{\text{WTX vs. Input}}$) or from the WTnX control by taking the maximal value out of all replicates (for calculating the P-value of enrichment against the affinity purified uncrosslinked control, $P^{\text{WTX vs. WTnX}}$). Fisher combination method was used to combine P-values of 6 biologic replicates for input or WTnX control separately to obtain the combined P-values of enrichment for each bin (Fig. S4B). (3) Finally, to filter sequences that were enriched through the direct binding to the miR-21 antisense LNA, we similarly calculated the P-values of read enrichment in all the bins from the affinity purified WTnX control against the corresponding input sample ($P^{\text{WTnX vs. Input}}$). Bins with $P^{\text{WTnX vs. Input}} \leq 0.05$ were considered to be enriched and were filtered (Filter 2, Fig. S4A), as such bins from the uncrosslinked samples were likely enriched through the direct binding between mRNAs and the LNA antisense probe.

From the 6 biologic replicates of the affinity purified WTX sample, we obtained a total of 3,497 bins after the above 2 filtering steps. Their P-values of enrichment against the input ($P^{\text{WTX vs. Input}}$) or against the WTnX control ($P^{\text{WTX vs. WTnX}}$) were graphed as a quantile-quantile plot (Fig. 2A, left plot). We set a threshold of $P < 1.23 \times 10^{-7}$ for the significant enrichment on the transcriptome scale (based on a conservative family wise error rate (FWER) of 0.05 for all bins in the RefSeq transcripts) and detected 131 bins containing uniquely mapped reads that were significantly enriched when compared with both the input and with the WTnX control (bins in the top right quadrant, left plot, Fig. 2A). When we processed the RNA-Seq data from the WTnX control samples following the

same procedure, we obtained a total of 2,915 bins after the 2 filtering steps. However, in striking contrast with the WTX sample, only 3 of the filtered bins were found to be significantly enriched when compared with the input and with the WTX sample using the same criteria (left plot, Fig. 2B). This confirmed that the LNA affinity probe selectively pulled down more transcripts in UV crosslinked samples, presumably through the crosslinked ternary complexes containing miR-21/Ago/targets (cf. Fig. 1A).

To examine the effect of UV crosslinking per se on the outcome of LNA affinity pulldown, we similarly purified cell lysates from UV crosslinked RKO$^{\text{miR-21}(-/-)}$ cells (miR-21 knockout crosslink, or KOX) and sequenced the purified RNAs. We then analyzed the RNA-Seq data following the same procedure as before and obtained 4,250 filtered bins. When the non-crosslinked samples (WTnX) were used as the control, 45 bins were found to be significantly enriched, as judged by both $P^{\text{KOX vs. Input}}$ and $P^{\text{KOX vs. WTnX}}$ to be less than 1.23×10^{-7} (left plot, Fig. 2C). Since RKO$^{\text{miR-21}(-/-)}$ cells contained no mature miR-21 or pre-miR-21 (Fig. S2A), this result suggested that UV crosslinking could generate RNA-protein complexes that were more amenable to LNA affinity pulldown than the corresponding uncrosslinked RNA sequences, and hence led to the recovery of more enriched transcripts.

To eliminate such false positives resulting from the non-specific interaction between the LNA probe and the crosslinked RNA-protein complex, we included the crosslinked samples of miR-21 knockout cells as another control (in addition to WTnX control) in our data analysis. Thus, when analyzing the WTX sample, we added KOX as a control: the initially identified 3,497 bins (left plot, Fig. 2A) were further filtered to remove those bins showing enrichment over the input in the KOX sample, i.e., bins with $P^{\text{KOX vs. Input}} \leq 0.05$ were filtered (Filter 3, Fig. S4A). This left us with 2,692 bins (right plot, Fig. 2A). We then recalculated the enrichment P value of each of these 2,692 bins against their corresponding uniquely mapped read counts from both WTnX and KOX samples using the Poisson distribution. The λ used in Poisson was estimated by taking the maximal value from all control libraries, including those using KOX and those using WTnX. Expression levels of each transcript measured by RNA-seq in KOX and WTX samples were used to adjust the λ estimation. We found a total of 13 bins that are significantly enriched over the input and the 2 control samples (bins in the top right quadrant, right plot, Fig. 2A). We identified their corresponding genes as the candidate miR-21 targets.

To estimate the false discovery rate of this data analysis pipeline, we applied the same procedure to look for bins that were significantly enriched in either WTnX or KOX sample, using the other 2 groups of samples as the control. For the WTnX sample, we compared uniquely mapped reads within each bin to the corresponding mapped reads found in WTX and KOX samples. This yielded 2,593 filtered bins by removing bins with $P^{\text{WTnX vs. Input}} \leq 0.05$ or $P^{\text{KOX vs. Input}} \leq 0.05$. Among them, 2 bins were found to be significantly enriched over the other 2 samples (right plot, Fig. 2B). For the KOX control sample, we found only one bin to be significantly enriched when compared against the WTnX and WTX samples (right plot, Fig. 2C). Judging from the number of significantly enriched bins found in the experimental group (WTX) and in the other
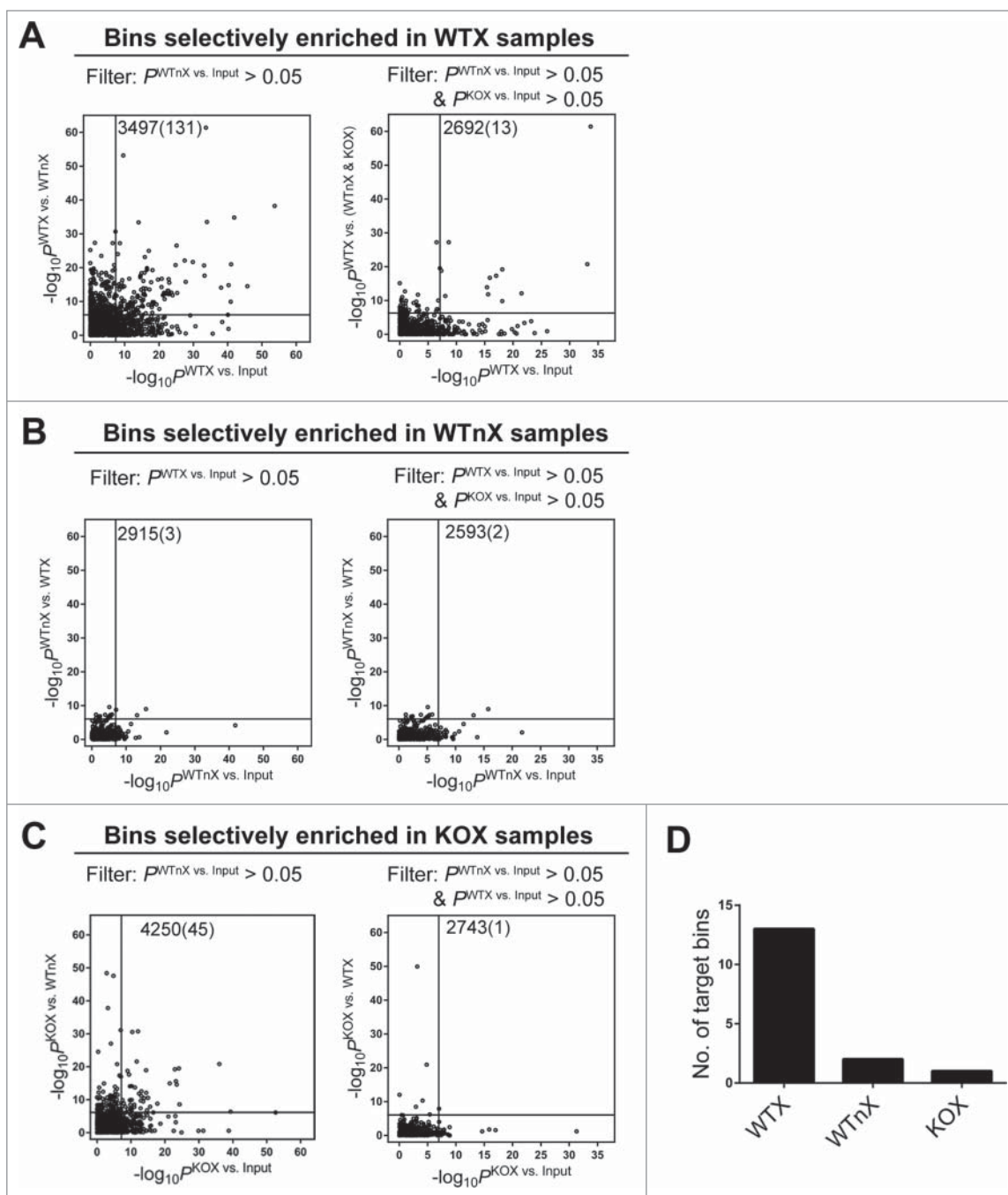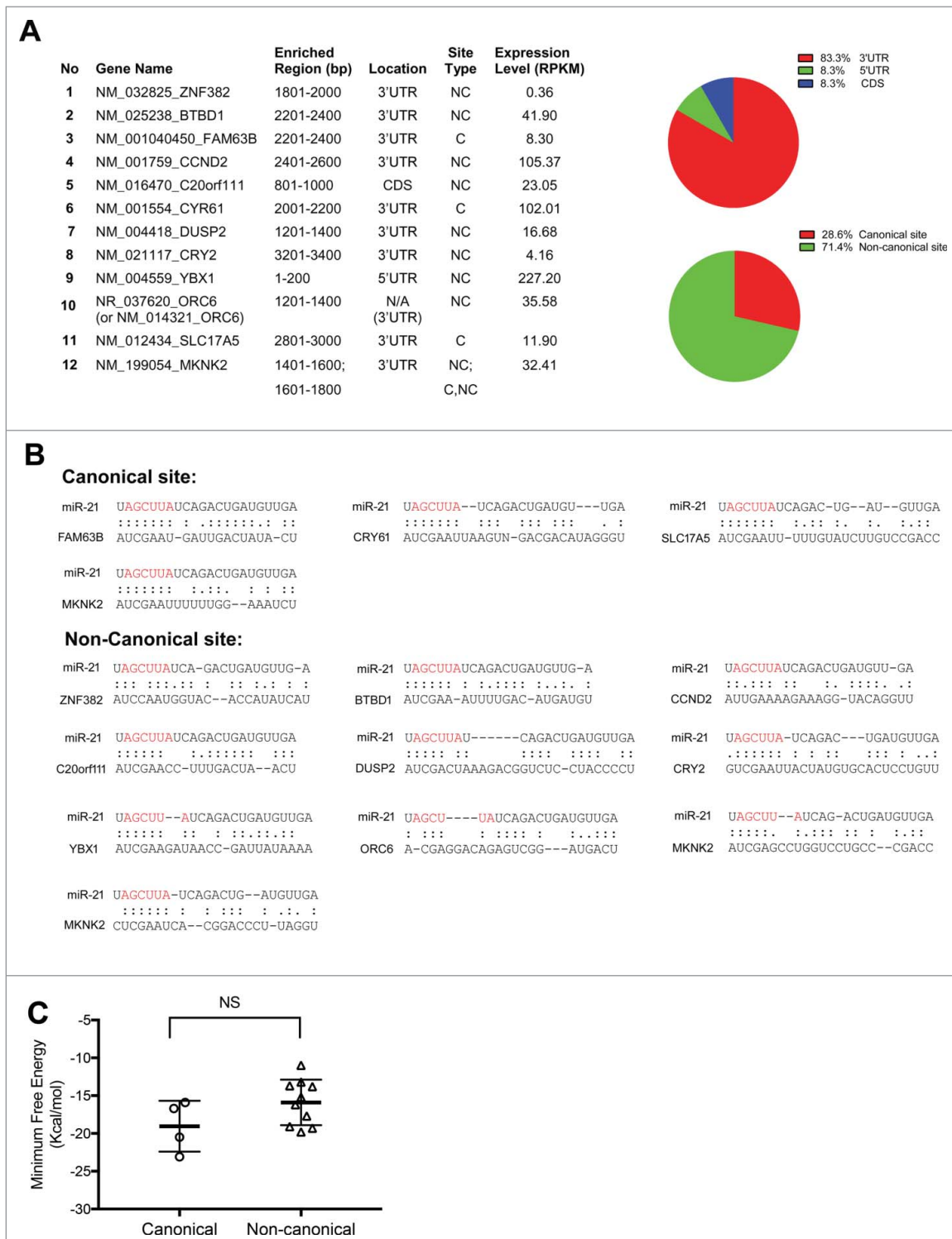
Figure 2. Identification of miR-21 target bins through differential enrichment analyses. (A– C) The -log₁₀(P values of enrichment) of filtered bins were calculated against the corresponding inputs (x-axis), or against control samples (y-axis). Bins selectively enriched in WTX samples (A), WTnX samples (B), and KOX samples (C) were plotted. The vertical and horizontal lines marked the threshold for the significant enrichment ($P = 1.23 \times 10^{-7}$. See text). (D) Number of identified candidate target bins from 3 samples (WTX, WTnX, and KOX) following the same data analysis pipeline.

2 controls (WTnX and KOX, Fig. 2D), we estimated that the upper bound of the false discovery rate to be 16%.

Validation of identified miR-21 targets

The miR-21 target sequences identified from WTX sample consisted of 13 bins from 12 genes (Fig. 3A). One gene, MKNK2, was found to contain 3 miR-21 target sites in its 3′ UTR over 2 adjacent bins (a stretch of 400 nucleotides). The rest of the genes appeared to contain a single target site in their respective enriched bins. Among all the identified target sites, only one was mapped to the coding sequence (CDS) and another to 5′ UTR, with the rest of the target sites residing at 3′ UTR (Fig. 3A). Sequence alignment of the identified target sites with miR-21 revealed that less than 30% of the target sites interacted with miR-21 via the canonical seed pairing. The majority of the target sites, however, interacted with miR-21 non-canonically (Fig. 3B). To address whether the observed prevalence of non-canonical pairing was due to difference in the thermodynamic stabilities between canonical and non-canonical interactions, we computed the minimal free energy (MFE) of each miR-21:: target site duplex using the program RNA-Hybrid.[34] The result

A

No	Gene Name	Enriched Region (bp)	Location	Site Type	Expression Level (RPKM)
1	NM_032825_ZNF382	1801-2000	3'UTR	NC	0.36
2	NM_025238_BTBD1	2201-2400	3'UTR	NC	41.90
3	NM_001040450_FAM63B	2201-2400	3'UTR	C	8.30
4	NM_001759_CCND2	2401-2600	3'UTR	NC	105.37
5	NM_016470_C20orf111	801-1000	CDS	NC	23.05
6	NM_001554_CYR61	2001-2200	3'UTR	C	102.01
7	NM_004418_DUSP2	1201-1400	3'UTR	NC	16.68
8	NM_021117_CRY2	3201-3400	3'UTR	NC	4.16
9	NM_004559_YBX1	1-200	5'UTR	NC	227.20
10	NR_037620_ORC6 (or NM_014321_ORC6)	1201-1400	N/A (3'UTR)	NC	35.58
11	NM_012434_SLC17A5	2801-3000	3'UTR	C	11.90
12	NM_199054_MKNK2	1401-1600;	3'UTR	NC;	32.41
		1601-1800		C,NC	

- 83.3% 3'UTR
- 8.3% 5'UTR
- 8.3% CDS

- 28.6% Canonical site
- 71.4% Non-canonical site

B

Canonical site:

```
miR-21    UAGCUUAUCAGACUGAUGUUGA
          ::::::: : .:::::::.. ::
FAM63B    AUCGAAU-GAUUGACUAUA-CU
```

```
miR-21    UAGCUUA--UCAGACUGAUGU---UGA
          ::::::: ::: ::: :::   . :
CRY61     AUCGAAUUAAGUN-GACGACAUAGGGU
```

```
miR-21    UAGCUUAUCAGAC-UG--AU--GUUGA
          ::::::: :.:: :. :. :.::
SLC17A5   AUCGAAUU-UUUGUAUCUUGUCCGACC
```

```
miR-21    UAGCUUAUCGACUGAUGUUGA
          ::::::: :.::. : : ::
MKNK2     AUCGAAUUUUUUGG--AAAUCU
```

Non-Canonical site:

```
miR-21    UAGCUUUCA-GACUGAUGUUG-A
          ::: :::::.:  .:.. :.. :
ZNF382    AUCCAAUGGUAC--ACCAUAUCAU
```

```
miR-21    UAGCUUAUCAGACUGAUGUUG-A
          ::.::: ::  . :::: :..:. :
BTBD1     AUCGAA-AUUUUGAC-AUGAUGU
```

```
miR-21    UAGCUUAUCAGACUGAUGUU-GA
          ::.::: :: .. :::::. .:
CCND2     AUUGAAAAGAAAGG-UACAGGUU
```

```
miR-21    UAGCUUAUCGACUGAUGUUGA
          ::::::  :.::::: :::
C20orf111 AUCGAACC-UUUGACUA--ACU
```

```
miR-21    UAGCUUAU------CAGACUGAUGUUGA
          ::::: :: :::: :::: ::
DUSP2     AUCGACUAAAGACGGUCUC-CUACCCCU
```

```
miR-21    UAGCUUA-UCAGAC---UGAUGUUGA
          .::::: : : :: ::: : ..:
CRY2      GUCGAAUUACUAUGUGCACUCCUGUU
```

```
miR-21    UAGCUU--AUCAGACUGAUGUUGA
          ::::::. :: ::.::.::
YBX1      AUCGAAGAUAACC-GAUUAUAAAA
```

```
miR-21    UAGCU----UAUCAGACUGAUGUUGA
          : :::   : :::: :  ...:::
ORC6      A-CGAGGACAGAGUCGG---AUGACU
```

```
miR-21    UAGCUU--AUCAG-ACUGAUGUUGA
          :::::. :.::: :: : :.::
MKNK2     AUCGAGCCUGGUCCUGCC--CGACC
```

```
miR-21    UAGCUUA-UCAGACUG--AUGUUGA
          ::::::: : : ::: : .:. :
MKNK2     CUCGAAUCA--CGGACCCU-UAGGU
```

C

Figure 3. miR-21 target genes and target sites identified by TargetLink in RKO cells. (A) List of target genes and the location of the enriched bins for each gene. ORC6 is annotated to have a non-coding and a coding transcript isoforms in RefSeq. NC: non-canonical; C: canonical. Pie charts showed the distribution of enriched bins (top) and the canonical/noncanonical sites (bottom). (B) Sequence complementarity of the identified target sites with miR-21. The seed sequence of miR-21 is highlighted in red. Vertical dots indicate Watson-Crick pairing or G-U wobble pairing. (C) Comparison of minimal free energies of miR-21::target site hybrids shown in (B).

showed that there was not a significant difference in the mean MFE between the canonical and non-canonical pairings identified by TargetLink (Fig. 3C). Moreover, MFE values of all 12 duplexes except one (miR-21::ZNF382, Table S1) were below −12 Kcal/mol, a cutoff value of MFE previously used for the computational identification of microRNA targets.[35] Of the 4 canonical sites that we identified, they all contain adenosine at the 3′-terminus to complement with miR-21 at position 1 (A-U pairing, Fig. 3B). For the non-canonical sites, all except 2 also contained adenosine at the 3′-termini, reflecting a strong bias of having adenosine at this position.[36] This adenosine preference was explained by the recent structural and biochemical

study showing that Ago2 binds adenosine at this position with 3-fold higher affinity than other nucleotides.[37]

To validate the identified targets, we examined their regulation by miR-21 using qRT-PCR to quantify their changes in expression upon manipulating miR-21 level. Inhibiting miR-21 in RKO cells (WT) caused significant increases ($P < 0.05$) in the expression of 8 out of 12 targets (Fig. 4A). The remaining 4 targets were found to be expressed at about the same level between cells treated with a miR-21 inhibitor or with a control oligonucleotide. Conversely, restoring miR-21 level in RKO$^{miR-21(-/-)}$ cells by using a miR-21 mimic led to significant downregulation ($P < 0.05$) of nearly 70% of the identified targets (Fig. 4B). The remaining targets, with the exception of ZNF382, were found to

be refractory to the exogenous miR-21 mimetic. This inverse correlation of miR-21 and its bound transcripts confirmed that a large portion of the identified interactions from UV crosslinking were indeed functional. The remaining 3 or 4 of the identified targets showing marginal or no response under the conditions tested probably reflected the posited role of microRNA in fine tuning gene expression.[38] Further, in addition to promoting transcript degradation, microRNA also regulates gene expression through translational repression. To evaluate the effect of miR-21 on these targets more thoroughly, we used the luciferase reporter assay after fusing the miR-21 target sequences to 3′ UTR of firefly luciferase. Among the 12 identified targets, we focused on the protein coding transcripts with miR-21 target
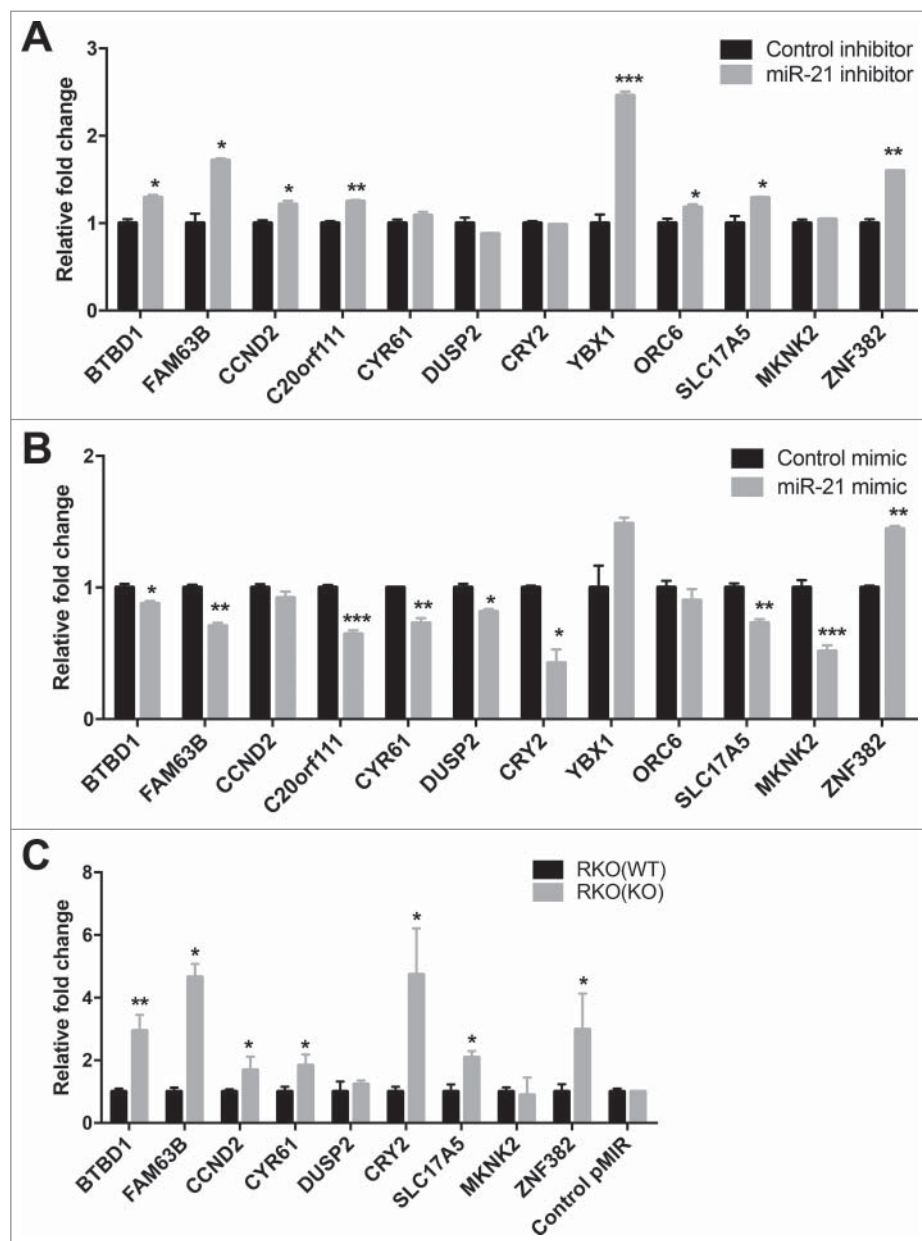
Figure 4. Validation of candidate miR-21 targets. (A) Comparison of the target expression (normalized to GAPDH) by qRT-PCR in RKO cells (WT) treated with a miR-21 inhibitor or a control oligonucleotide. (B) Comparison of the normalized target expression in RKO$^{miR-21(-/-)}$ cells (KO) either treated with a miR-21 mimetic or a control. The expression level of each gene in the control sample was arbitrarily set to 1. (C) Luciferase reporter assay of the identified miR-21 target sequences. Reporter vectors were constructed by inserting miR-21 target sequences to 3′ UTR of Firefly luciferase, whose expression was normalized against the co-transfected Renilla luciferase. The expression ratios of the reporters between KO and WT cells were further normalized against a control reporter (pMIR) lacking miR-21 target site at 3′ UTR. Mean ± SEM is plotted (N = 3). *$P < 0.05$; ** $p < 0.005$; ***$p < 0.0005$.

sites localized at their 3′ UTRs (Fig. 3A). This led to 9 luciferase reporters corresponding to this sub-set of targets (Fig. S5). Seven of these reporters showed a significant (P < 0.05) de-repression of luciferase expression in the absence of miR-21, and the remaining 2 reporters were found to be expressed at about the same level between WT and KO cells (Fig. 4C). The results confirmed that the majority of the identified target sites conferred potent inhibitory effect on the expression of their host genes through miR-21, providing further evidence validating the identified target sites.

Among the 12 identified targets, C20orf111 (also named as OSER1) encodes a protein of unknown function and was the only gene whose miR-21 target site was mapped to the coding region (Fig. 3A). To further validate that the identified interaction at the coding region was functional, we compared the expression of C20orf111 between wild type RKO cells and the isogenic miR-21 knockout RKO cells. At both the transcript and protein level, the expression of C20orf111 was repressed by miR-21 (Fig. 5), confirming that the target site of miR-21 located in the coding region of C20orf111 was indeed functional. Interestingly, staining of OSER1 by immunofluorescence revealed that the protein was localized exclusively in the cell nucleus, and further analysis by dual-color immunofluorescence confirmed that OSER1 was colocalized with SC35, a nuclear speckle marker protein.[63] Nuclear speckles are sub-nuclear structures important for RNA splicing and RNA processing. The localization of C20orf111 to this organelle suggests its potential role in the formation/maintenance of nuclear speckles and/or in RNA processing. Since our result provided unequivocal evidence showing that miR-21 directly repressed C20orf111 at both transcript and protein levels, it raised the interesting hypothesis for future investigation concerning the possible involvement of miR-21 in RNA processing through C20orf111.

Discussion

We have developed a new technique, TargetLink, for identifying the direct target sequences of a specific microRNA (miR-21) in fully intact living cells. Distinguishing features of TargetLink include: (1) Employing LNA oligonucleotides as the affinity probe to pull down crosslinked protein-RNA complexes. The high binding affinity and stability of LNA::RNA duplex
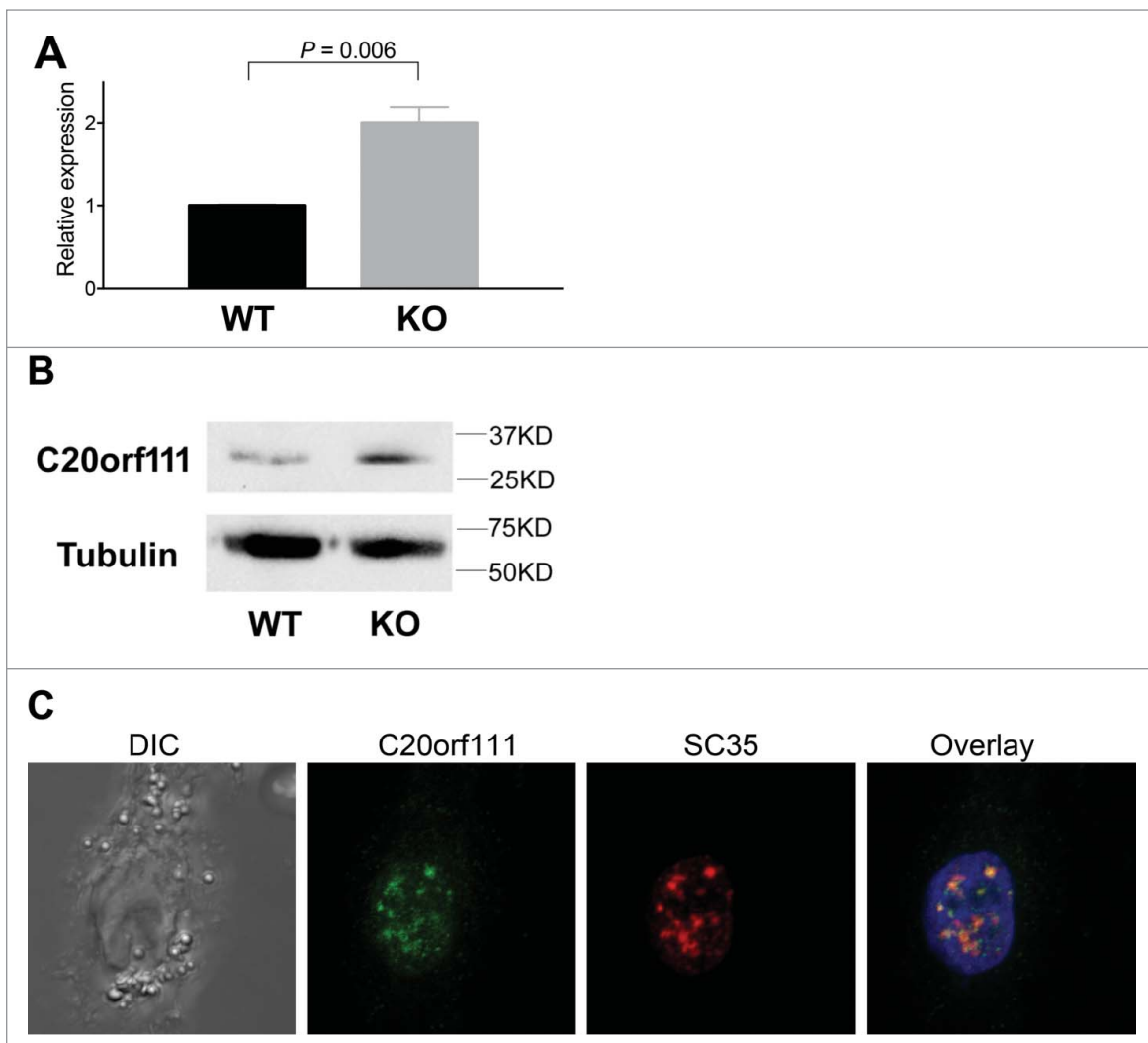
Figure 5. Repression of C20orf111 by miR-21. (A) Quantification of C20orf111 expression by qPCR in WT or miR-21 knockout RKO cells (N = 3, normalized against GAPDH). Expression in wild type cells was arbitrarily set to 1. (B) Quantification of protein abundance by Western blot. (C) C20orf111 is colocalized with SC35 in the nuclear speckle. The cell nucleus is marked with DAPI in overlay.

allowed us to use a potent denaturing buffer (4 M guanidinium thiocyanate) to wash off non-specific binding events. (2) The technique identifies targets of an endogenous microRNA and does not require infecting cells with tagged proteins or exogenous microRNA mimetics. (3) TargetLink consists of a rigorous and unbiased data processing pipeline to systematically analyze and compare sequencing data between the experimental group and controls. The inclusion of 2 control samples, non-crosslinked wild type cells expressing miR-21 (WTnX) and crosslinked cells devoid of miR-21 (KOX), turned out to be critical for eliminating false positives and for identifying the bona fide target sequences of miR-21 (Fig. 2).

Combining these unique features, we developed TargetLink as a systematic and unbiased approach for identifying the direct targets a specific microRNA in intact cells. Target identification using this approach does not rely on the assumption of seed matching or the location of the target site at 3′-UTR. Using miR-21 as a testing microRNA, we successfully identified 12 target genes (or 13 target sites) of miR-21 in a human colorectal cancer cell (RKO). A number of evidences supported that these sequences were enriched through the specific LNA affinity purification of miR-21/Ago/target ternary complex (compare Fig. 1A): (1) These 14 target sites were only identified in crosslinked cells expressing miR-21, but not in control samples. (2) The expression levels of most identified target genes, measured from RNA-Seq, were fairly close between wild type RKO cells and miR-21 knockout cells (Table S2), arguing that these target sequences are unlikely enriched through the direct interaction with LNA. (3) In fact, by performing local sequence alignment, we found that the sequence similarity of the identified target sites with miR-21 was significantly lower (p < 0.0001) than those highly enriched sequences found in both experimental and control samples (Fig. 6).

A somewhat unexpected finding from this study was that a large portion of the identified miR-21 targets interacted with miR-21 through imperfect seed matching (Fig. 3). We considered a few possible explanations that might account for the outcome. First, in the case of RKO colorectal cancer cells, we cannot formally exclude the possibility that imperfect seed pairing could represent a major type of interaction between miR-21 and its targets. In fact, most of our identified non-canonical sites start with adenosine, a nucleotide known to preferentially engage with the RISC complex during the initial matching process when a microRNA is interrogating its potential targets.[37] Further, a comparison of the thermodynamic stabilities of the canonical and non-canonical duplexes of our target set revealed no significant difference (Fig. 3C). In addition, we also calculated the minimal free energies of miR-21 duplexes of several reported miR-21 targets, most of which were derived from computational prediction and contained canonical sites (Fig. S6). Again, the result showed a rather comparable thermodynamic stability of our target set and the reported miR-21 targets. Together, these results provided no evidence to persuade us to expect that perfect seed pairing ought to be the dominant determinant of target engagement, at least in the case of miR-21. Since the 5′ region of microRNAs is more conserved than the rest of the sequence, seed pairing has become a popular filter for selecting candidate targets in several target prediction programs. However, the presence of a seed site in 3′ UTR does not necessarily ensure engagement of a miRNA,[39] and other factors such as mRNA folding and accessibility of binding sites, sub-cellular localization etc, will determine whether a mRNA is engaged with a microRNA or not. Moreover, since evolution conservation of a candidate miRNA binding site is not expected to be a feature that can be assessed by the RISC complex in situ,[2] potential miR-21 binding sites with either perfect or imperfect seed matching are equally likely to become occupied by the RISC complex as long as the target sites are accessible and can form stable duplexes with microRNAs.

An alternative explanation is that our current version of TargetLink may turn out to be more efficient in recovering non-canonical sequences due to, for example, differential UV crosslinking efficiencies of RNA duplexes containing either perfect or imperfect seed pairing. While we cannot formally exclude this possibility at the moment, we are not aware of a systematic study comparing the efficiency of UV crosslinking between RNA duplexes with different degree of complementary and the protein to which they bind. In fact, a careful examination of miR-21 duplexes with either canonical or non-canonical sites in our target set (Fig. 3), or in the reported miR-21 targets (Fig. S6), revealed no obvious difference in either the extent or the pattern of overall base complementarity or in the thermodynamic stability of the miR-21::target hybrids.

microRNAs may regulate gene expression by inhibiting translation and/or by inducing degradation of their target genes. At the transcript level, the majority of the identified miR-21 targets exhibited an inverse correlation with miR-21 level when the expression of miR-21 was acutely altered (Fig. 4). In most cases, the magnitude of change in miR-21 target expression was modest yet statistically significant, consistent with the notion that microRNAs play a major role in fine tuning gene expression.[38] Compared to the acute inhibition of miR-21, miR-21 deletion in RKO cells resulted in less pronounced effect of de-repressing miR-21 targets, as judged from RNA-Seq data and DeSeq analysis (Table S2). Presumably this reflected the difference in cells' responses to permanent deletion versus acute inhibition of a gene, whereas adaptation and/or compensation in the knockout model may dampen the effect of miR-21 deletion on its target expression. In support of this possibility, we compared the expression levels of several previously reported miR-21 targets in RKO wild type or miR-21 knockout cells. The result was similar to that of our own target set, with few reported targets showing statistically significant ($P_{adj} < 0.05$) difference (Table S3). Compared to qRT-PCR or RNA-Seq data, the luciferase reporter assay using the 3′-UTR harboring miR-21 target sites showed a relatively higher magnitude of de-repression when miR-21 was deleted (Fig. 4C). This robust de-repression liekly resulted from a combined relief from translational inhibition and mRNA degradation.

Among a dozen novel targets identified by TargetLink, C20orf111 harbored a target site at its coding sequence (CDS). The identified target site was functional as C20orf111 was repressed by miR-21 by several assays (Fig. 4A, B and Fig. 5). Despite the fact that 3′-UTR represents a primary location for microRNA targeting, there have been increasing reports of CDS targeting by different miRs.[6,40-44] Similar to our result on C20orf111, these reports also confirmed that microRNA binding sites within CDS were functional in controlling gene expression, and the repression could be mediated by mRNA degradation and/or translational inhibition. Moreover,
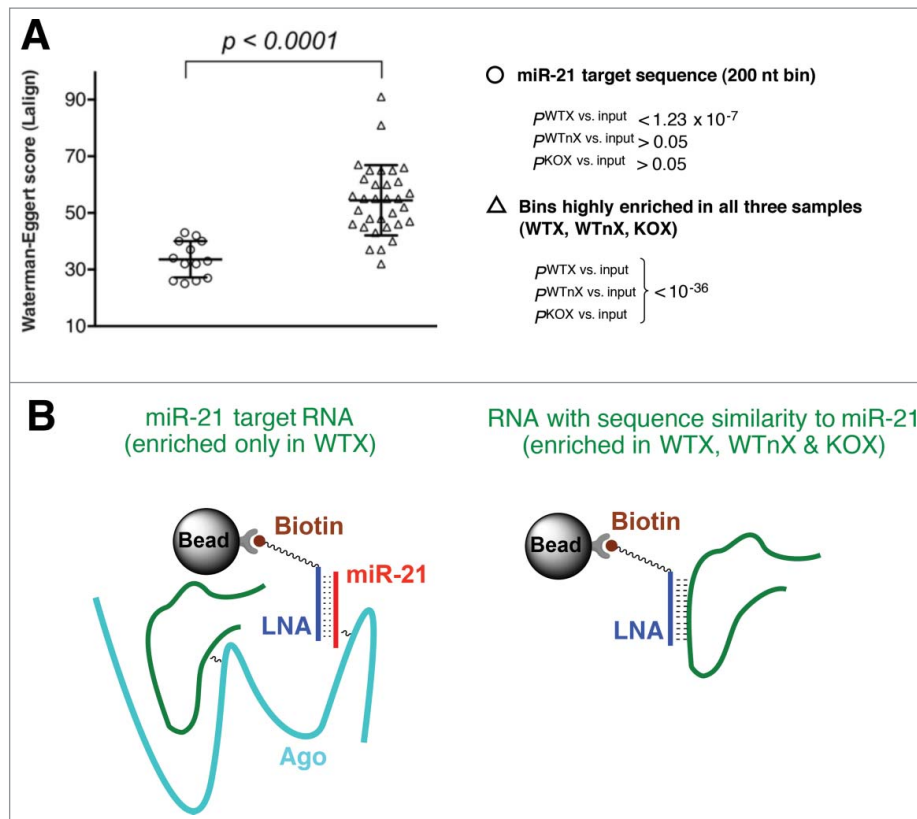
Figure 6. (A) The identified miR-21 target sites showed much lower sequence similarity to miR-21 (measured by the Smith-Waterman algorithm, http://www.ebi.ac.uk/Tools/psa/lalign/nucleotide.html) compared with those sequences highly enriched in both experimental (WTX) and control samples (WTnX and KOX) (p value, unpaired t-test). (B) miR-21 target sequences were enriched through the crosslinked ternary complex consisting of miR-21::Ago::target. In contrast, RNAs with sequence homology to miR-21 were enriched through the direct binding to the antisense LNA.

transcriptome wide CLIP data has revealed that sequences at both CDS and 5′-UTR could be occupied by the Argonaute silencing complex.[12,13] and, at least in Drosophila, preferentially conserved miRNA targeting in the open reading frame appeared to be as widespread as in 3′-UTRs.[45] Together, these observations suggest that functional microRNA targeting in the open reading frame can be widespread and may well represent an underappreciated mode of action of microRNA mediated repression. Future optimization and application of TargetLink should accelerate the pace of identifying such CDS targets in mammalian cells, which in turn will facilitate addressing their functional roles and modes of action in regulating gene expression.

Of the 12 miR-21 targets identified in this study, the expression level of these genes varied over a wide range, from less than 1 RPKM (ZNF382) to over 100 RPKM (CCND2, CYR61 and YBX1, Fig. 3A and Table S2). The fact that LNA affinity purification used here was able to pulldown miR-21 targets of relatively low abundance lends support to the efficiency of target enrichment using this technique. However, it was also thought that a specific microRNA could potentially be involved in interacting with more than tens if not hundreds of mRNAs in situ,[38] so it is quite likely that our current procedure may have missed some of the bona fide miR-21 targets. We envision optimizing several experimental variables for the future improvement on the sensitivity of target discovery using this approach. First, there should be room for dialing down the

stringency of affinity purification to increase the yield of RNA recovery, which should help recovering more target sequences. Of course, lowering the stringency of affinity purification is expected to pull down more non-specific RNAs. However, we do not expect to see a rise in false positives as our data analysis pipeline detects targets through a series of cross comparisons between experimental group (WTX) and controls (WTnX and KOX). Non-specifically recovered RNAs or RNAs that bind to LNA probe directly (instead of through crosslinked miR-21/Ago complex)) will be filtered off because they would be enriched in all the samples after affinity purification (Fig. 6). Second, we plan to test and adopt some recent advancements in cDNA library preparation tailored for UV-crosslinked RNAs.[46,47] and/or for low input RNA samples [48] So far we have been using the commercial library preparation kit intended for the standard RNA-Seq, which is likely to be sub-optimal for sequencing crosslinked and affinity purified RNAs. By applying these very recent procedures tailored for processing UV-crosslinked RNAs, we anticipate major improvements in the efficiency of library preparation. Last but not least, to increase the yield of RNA-protein crosslinking, we may attempt other means such as chemical crosslinking [49] or using photoactive nucleoside analogs.[13] Our current version of TargetLink has already yielded a dozen targets of miR-21, and hence represents a major advancement in the efficiency of target discovery compared with numerous previous studies in which only one or a few miR-21 targets were reported in each case (Fig. S6).[31,50-53]

Moreover, the 12 target genes identified here should provide a valuable reference for benchmarking the efficiency and specificity of future optimization and improvement.

As it currently stands, the workflow of TargetLink includes isogenic RKO cells devoid of miR-21 as an important control during data analysis to effectively filter off false positives (Fig. 2). Comparing the enriched reads between the un-crosslinked sample

The miR-21 antisense LNA probe used for the affinity purification contained (from 3′- to 5′-end) the biotin affinity label, 2 repeats of hydrophilic hexaethyloxy glycol linker, and an 18mer-oligonucleotide complementary to the last 18 nucleotides of miR-21 (UUAUCAGACUGAUGUUGA). The LNA had the following composition (LNAs are represented by +A, +G, +C, and +T; and DNAs are represented by A, G, C, T):

LNA Probe	Composition
miR-21 antisense (Used for affinity purification)	5′ +TC+A+AC+A+TC+AG+TC+TG+A+T+A+A-(EG$_6$)$_2$-biotin 3′
Scramble LNA (Used as a control for affinity purification)	5′ +TC+T+A+T+T+TCG+A+T+T+AGG+AG+A-(EG$_6$)$_2$-biotin 3′
miR-21 antisense (Used for Northern blot)	5′ +TC+A+AC+A+TC+AG+TC+TG+A+T+A+AGC+T+A-EG$_6$-biotin 3′

(WTnX) and the crosslinked sample from miR-21 knockout cells (KOX) revealed that UV-crosslinking by itself was able to generate RNA::protein complexes that were amenable to LNA affinity pulldown, and such an enrichment occurred independent of miR-21 (Fig. 2B, C, left). When analyzing bins selectively enriched in KOX samples, nearly all the false positives (Fig. 2C, left) were eliminated by using WTX as the control (Fig. 2C, right). Likewise, when analyzing bins selectively enriched in WTX samples, adding KOX to the control group reduced the number of selectively enriched bins by roughly 90% (Fig. 2A). With the recent rapid advancements in genetic engineering technologies including TALEN and CRISPR-Cas9, it is becoming increasingly straightforward to delete small non-coding RNAs in animal cells.[54-56] Moreover, as transgenic animals devoid of specific microRNAs are becoming increasingly accessible from different species including mouse,[57] Drosophila [58] and nematode,[59] we anticipate that TargetLink will gain broad applications for defining the target set of a specific microRNA in different biologic systems including cultured cells and freshly dissected primary tissues. Finally, as discussed earlier, as we continue to optimize the experimental procedures of TargetLink to increase RNA recovery from affinity purification, to adjust processing crosslinked RNA for library preparation and to enhance crosslinking efficiency, we speculate that we ought to be able to further broaden the application of TargetLink by omitting knockout cells as a future enhancement of this method.

Material and methods

A. Preparation of locked nucleic acid (LNA) probes for affinity purification

Monomers of LNA phosphoramidites were purchased from Exiqon. DNA phosphoramidites and 3′-protected Biotin Serinol CPG were purchased from Glen Research. DMT-hexaethyloxy-glycol phosphoramidite was purchased from ChemGenes Corporation. Anhydrous solvents and common reagents were purchased from Aldrich or VWR. LNA probes were synthesized in-house by the standard solid phase phosphoramidite chemistry from 3′- to 5′-end on an ASM 800 DNA Synthesizer (Biosset).[60] The synthesized LNA probes were purified by PAGE as described in the Supplement.

TargetLink protocol

Materials and reagents

UV cross-linking.
(a) Cold PBS, Stratalinker (Stratagene, Model 1800), 15 cm tissue culture dishes.

RNA digestion (partial) and DNA digestion.
(a) Cell lysis buffer: 50 mM Tris-HCl, pH 7.4; 100 mM NaCl; 1 mM MgCl$_2$; 0.1 mM CaCl$_2$; 1% NP-40; 0.5% sodium deoxycholate; 0.1% SDS; protease inhibitor and ANTI-RNase (both added fresh).
(b) ANTI-RNase (Ambion, AM2692), RNase I (Ambion, AM2295), Turbo DNase (Ambion, AM2239), proteinase inhibitor cocktail (Calbiochem, 535140), 1M DTT, Thermomixer R (Eppendorf).

Diafiltration and HPLC-SEC.
(a) Ultracel-10K (Millipore, UFC901024) and 50K centrifugal filters (Millipore, UFC905024).
(b) Regenerated cellulose membrane 0.22 μm (Corning, Cat. # 431222).
(c) Agilent Technologies 1200 series HPLC system; BioSep-SEC-s2000 column (Phenomenex).
(d) HPLC running buffer: 4 M guanidinium thiocyanate (GT) and 25 mM Nacitrate, pH 7.0.

LNA affinity purification.
(a) 1X GT buffer: 4 M guanidinium thiocyanate; 25 mM Nacitrate, pH 7.0; 0.5%(w/v) sodium N-lauroyl sarcosinate.
(b) LNA antisense probe against miR-21.
(c) NanoLink Streptavidin Magnetic beads (Solulink, M10178), Magnetic stand (Invitrogen), Heating block (BenchMark).
(d) TE buffer, yeast tRNA (Roche, Cat # 14042122).
(e) Elution buffer: 0.1 M NaOH.
(f) Neutralizing buffer: 1 M glacial acetic acid.

Protein digestion and RNA isolation.
(a) Proteinase K (Roche, 14724600).
(b) 1X PK buffer a: 100 mM Tris-HCl, pH 7.4; 50 mM NaCl; 10 mM EDTA.

(c) 1X PK buffer b: PK buffer 1 + 7 M urea.
(d) RNA extraction: Saturated phenol pH 6.6 (Ambion, AM9712), chloroform, glycoblue (Ambion, AM9515), linear acrylamide (Ambion, AM9520), 3M sodium acetate pH 5.5 (Ambion, AM9740), isopropanol, 75% ethanol, 100% ethanol.

Library preparation.
(a) Superscript® III Reverse Transcriptase (Invitrogen, 18080).
(b) KAPA Library Amplification Kit (Kapa Biosystems, KK2611)
(c) Agencourt AMPure XP(Beckman, A63881)
(d) TruSeq stranded mRNA Library Prep Kit (Illumina, RS-122–2101).

RKO Cell. Wild type human colorectal RKO cells (WT, catalog # HD PAR-059) and isogenic miR-21 knockout RKO cells (RKO$^{miR-21(-/-)}$ or KO, catalog # HD R05–013) were purchased from Horizon Discovery Ltd (www.horizondiscovery.com). Cells were cultured in RPMI 1640 including 2mM L-glutamine, 10% FBS, penicillin (100 U/mL) and streptomycin (100 μg/mL). Cells were cultured as a monolayer at 37°C in a humidified atmosphere with 5% CO_2, and passaged every 3–5 d when reaching ∼80% confluence.

Experimental procedure
The whole procedure takes about 5 d to complete and consists of 3 major steps:

Cell lysate preparation. Cells are irradiated with UV to crosslink proteins and nucleic acids. Cells are then lysed and the lysate is digested with RNase and DNase. The digested cell lysate is subjected to diafiltration and SEC-HPLC purification to remove uncrosslinked small RNAs.

LNA affinity purification and RNA extraction. The above cell lysate is then affinity purified by LNA. The eluent containing crosslinked ternary complexes (miR-21/protein/mRNA) are deproteinated and extracted with phenol/chloroform to yield RNAs.

Library preparation for deep sequencing. We prepared cDNA library of the purified RNAs using the commercial library preparation kit with modifications.
A step-by-step description of the experimental procedure is included in the Supplement.

TargetLink analyses

Mapping of sequenced reads

Filtering and pre-processing of raw reads. The sequenced reads from each library were first processed by CIMS package (http://zhanglab.c2b2.columbia.edu/index.php/CIMS_Documentation) to filter the low quality reads and to collapse identical copies.[61] We then used cutadapt (version 1.3) to trim residual (if any) Illumina sequencing adapters. Only reads at least 20 nt long are kept for further analysis.

Mapping. The pre-processed reads were mapped to NCBI human Refseq transcriptome using novoalign with the setting of –t 85 –l 25 -s. The longest isoform of each transcript was used for base numeration. To further remove PCR duplicates, we applied SAMtools to collapse multiple reads that have been aligned to start at the same position on the RefSeq to a single read. The final unique mapped reads were used to calculate the read counts within 200-nt bins of individual Refseq transcripts. For reads spanning the junction of adjacent bins, we only count the number of reads in a bin if those reads have their 5′-end mapped inside the bin. So when a read overlaps 2 adjacent bins, only one bin will be counted.

Detecting significantly enriched bins and miR-21 target sites
We split each transcript into consecutive 200-nt bins, and counted the number of unique mapped reads falling into each bin. For the 6 biologic replicates using crosslinked wild-type RKO cells (WTX), only bins containing 2 or more unique mapped reads in at least 2 thirds of the biologic replicates were kept. we then calculated the p-value of seeing the observed count in each remaining bin by using Poisson distribution. The λ used in Poisson was estimated from all control libraries, including those using crosslinked miR-21 knockout RKO cells (KOX) and those using wild-type RKO cells without crosslinking (WTnX) by using the following formula:

$$\lambda_{iw} = N_i * \frac{\mathbf{max}}{j = 1, 2, \ldots, J} \left(\frac{c_{jw}}{n_j} \right),$$

Whereas N_i is the total number of uniquely mapped reads in the ith WTX experiment, n_j is the total number of uniquely mapped reads in the jth control experiment (total 6 control replicates, 3 from KOX, and 3 from WTnX), and c_{jw} is the number of reads in wth bin from the jth control experiment. Since we took the maximal value from all the control samples, λ_{iw} is a very conservative estimation of the background count for the wth bin. After obtaining the p-values from all WTX replicates, we combined them by using Fisher combination method to calculate the statistic x_w which follows chi-square distribution with degree of freedom 12 (for 6 replicates of WTX samples):

$$x_w = (-2) * \sum_i \ln(p_{iw})$$

The Bonferroni correction was applied to chi-square p-values to achieve the list of candidate target bins (Fig. 3A).
To evaluate the false discovery rate of applying this data analysis pipeline, we applied the same procedure to examine the number of bins selectively enriched in the control samples (Fig. 3B, 3C). When processing WTnX replicates, only bins containing 2 or more unique mapped reads in at least 2 thirds of the WTnX replicates were kept. We then compared the remaining bins in each of WTnX replicates with all the replicates from WTX and KOX (Fig. 3B, WTX and KOX serving as controls for

WTnX). The same procedure was applied to KOX replicates (Fig. 3C, WTX and WTnX serving as controls for KOX).

Aligning target sites with miRNA

To align the identified target bins with miR-21 (Fig. 4B), we developed a dynamic programming method by modifying Smith-Waterman algorithm (Lalign) with no penalty for gaps outside of miRNA alignment, and score 5 for match, −4 for mismatch, −8 for gap-opening and −2 for gap-extension in the alignment.

Northern blot of HPLC fractions to detect crosslinked miR-21

After running SEC-HPLC to separate the crosslinked miR-21 from free miR-21, we concentrated the collected fractions about 10-fold by diafiltration (Ultracel-10K centrifugal filter, 4°C, 4000 g). Take 25 μL of concentrate and dilute it to 200 μL with 1X PK buffer a containing 2 mg/mL proteinase K. Shake the mixture at 1,100 rpm, 37°C for 3 h. Add 400 μL of 1X PK buffer b and shake the mixture at 37°C, 1100 rpm for another 20 min. To extract the RNA, add 400 μL phenol and 133 μL of CHCl$_3$ to the mixture and shake it at 37°C, 1100 rpm for 20 min. Spin the mixture at 4°C and 14,800 rpm for 15 min. Transfer the top aqueous phase to a 1.5 mL tube. Add 2 μL of linear acrylamide (5 mg/mL) and 0.25 μL glycoblue (15 μg/μL) and vortex briefly. Add 3-volume of isopropanol to precipitate the RNA by leaving the tube at −20°C overnight. Collect the RNA pellet by centrifugation at 4°C and 14,800 rpm for 15 min. Wash the pellet with 1 mL of 75% ethanol. Spin down the pellet again at 4°C and 14,800 rpm for 5 min. Air-dry the pellet for 5 min and add 10 μL of RNase-free water to dissolve RNA. Add 10 μL of RNA loading buffer (Ambion, Cat. No. AM8547) and boil the mixture at 95°C for 5 min.

The above RNA mixture was fractionated by PAGE using a 15% denaturing polyacrylamide gel containing 8 M urea and 20% formamide. RNA was then transferred to the nitrocellulose membrane and fixed to the membrane by UV crosslinking. Membranes were probed with 2 nM of biotin-labeled LNA-modified oligonucleotides: 5′-+TC+A+AC+A+TC+AG+TC +TG+A+T+A+AGC+T+A-biotin-(PEG$_6$)-3′ (+ denotes LNA). We then detected the biotinylated probe on the membrane using the Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific, Cat. No. 89880).

RNA-Seq

RNA was isolated with Trizol (Invitrogen, 15596–018) by a chloroform extraction, assisted by phase lock tubes. The extracted RNA was precipitated by isopropanol and redissolved in 20 μL of RNase free water. After digesting the sample with TURBO^{TM} DNase (Ambion, AM2239), we extracted RNA (in 100 μL water) with chloroform/phenol and precipitated RNA with 250 μL of 100% ethanol and 10 μL of 3M NaOAc at −80°C overnight.

We used the Ribo-Zero^{TM} rRNA removal kit (Epicenter, MRZH116) to remove the ribosome RNA following the instructions from the manufacturer. The purified RNA was used for the cDNA library preparation using the TruSeq Stranded mRNA Library Prep Kit (Illumina, RS-122–1201). The sample was sequenced at 50 cycles, single read on an Illumina HiSeq 2000 platform.

Sequencing reads from 3 wild type RKO cell populations and 2 RKO$^{mir-21(−/−)}$ cell populations were mapped using TopHat2 (Version 2.0.10) to the human genome version 19 (hg19). The average mappability was 78 $\pm$ 3%. Cufflinks2 (Trapnell 2010, Version 2.2.1) was used to calculate expression level (Read Per Kilo basepairPer Million, RPKM) of each transcript and gene. DESeq [62] was used to test significance of transcripts with differential expression between wild-type and mir-21 knockout cells.

Luciferase reporter constructs and dual-luciferase assay

Firefly luciferase reporters of the miR-21 targets were constructed by inserting the sequences near the identified miR-21 target sites (Fig. S5) into pMIR-REPORT^{TM} miRNA Expression Reporter Vector System (LifeTechnologies). The primers used for the amplifying these sequences were listed in Table S3. The amplified fragments were digested with appropriate restriction enzymes (Table S3) and inserted into the 3′ UTR of pMIR-REPORT.

The firefly luciferase reporter was transfected into RKO (WT) or RKO$^{miR-21(−/−)}$ cell using Lipofectamine 2000 (Invitrogen, 11668019). Another expression vector encoding Renilla luciferase was co-infected for normalization of transfection efficiency. Firefly and Renilla luciferase activities were measured using the Dual-Luciferase® Reporter assay system (Promega, E1960) in a Fluoroskan Ascent^{TM} Microplate Fluorometer (Thermo Fisher Scientific, 5210470). The expression ratios of the reporters between KO and WT cells were normalized against a control reporter (pMIR) lacking miR-21 target site at 3 UTR.

Measuring target transcripts after infecting cells with miR-21 mimic or inhibitor

To acutely manipulate miR-21 level in cells and to measure its effect on the expression level of its target transcripts, we either infected RKO cells with a miR-21 inhibitor (Ambion, 4464084) or transfected RKO$^{miR-21(−/−)}$ cells with a miR-21 mimic (Dharmacon, C-300492–03). Both agents were used at a dose of 40 picomoles per well in a 24−well microtiter plate. The miR-21 inhibitor or mimic was mixed with 120 μL of serum free medium in a well. A total of 79.2 μl of serum free medium containing 0.8 μl of RNAi-MAX (Invitrogen, 1641636) was then added to each well. After 10−min incubation at room temperature, ~40,000 cells in suspension were then added to each well in 400 μL of culture medium (15% FBS) to a final total volume of 600 μL. Cells were then cultured at 37°C/5% CO$_2$ for 72 hours. The total RNA from each well was isolated with Trizol, and cDNA was generated by the Quanti-Tect Reverse Transcription Kit (Qiagen, 205311). qPCR was performed using iTaq^{TM} Universal SYBR Green Supermix (Bio-rad, 172–5121) according to the manufacturer's protocol. Relative expression was evaluated by the comparative CT method and normalized to the expression of GAPDH. The expression levels of the target transcripts after miR-21 inhibition (in RKO cells) or overexpression (in RKO$^{miR-21(−/−)}$ cells)were normalized against values from the corresponding cells infected with control oligos supplied

by the vendors. Primers were listed in Table S3 (**Supplementary Information**).

Western blot analysis of C20orf111 (OSER1) expression

WT or miR-21 knockout RKO cells were washed once in phosphate-buffered saline, and lysed in radioimmuno-precipitation buffer (150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris-HCl, pH 8, 2 mM EDTA). The cell lysate containing 5 μg of protein was loaded to a 12% SDS-PAGE gel, separated by electrophoresis, and transferred to a PVDF membrane. Primary antibody: Rabbit anti- C20orf111 (Novus, NBP1–91725) was diluted 1:200; Mouse anti-β-tubulin (DSHB, E7) was diluted 1:1000. Secondary antibody: HRP Goat Anti-rabbit (Bio-rad, 170–6515) was diluted 1:3000; HPR Goat Anti-mouse (Bio-rad, 170–6516) was diluted 1:3000.

Cellular localization of C20orf111 by immunofluorescence

RKO cells were washed with cold phosphate-buffered saline (PBS), fixed with 4% PFA at room temperature for 15 min, permeabilized with 0.1% Triton X-100 for 10 min, and washed 3 times with PBS. Cells were blocked with 10% donkey serum, incubated sequentially with primary antibody (Ab) and secondary Ab for 1 hour each. Cell nuclei were marked with DAPI (1 μg/mL) for 20 min before imaging on a confocal microscope (Zeiss LSM780). Primary Ab for C20orf11: Rabbit Anti-C20orf111 (1:50, Novus Cat # NBP1–91725); Secondary Ab: Alexa Fluor 488 labeled donkey anti-rabbit (1:100, Jackson Cat # 711–545–152). Primary Ab for SC35: Mouse anti-SC35 (1:500, Abcam Cat # ab11826); Secondary Ab: Cy3-labeled donkey anti-mouse (1:100, Jackson Cat # 715–165–151).

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

ORCID

Yan Xu 🆔 http://orcid.org/0000-0002-3595-1549
Yan Chen 🆔 http://orcid.org/0000-0003-4719-3762

References

1. Ambros V. The evolution of our thinking about microRNAs. Nat Med 2008; 14:1036-40; PMID:18841144; http://dx.doi.org/10.1038/nm1008-1036

2. Cloonan N. Re-thinking miRNA-mRNA interactions: intertwining issues confound target discovery. Bioessays 2015; 37:379-88; PMID:25683051; http://dx.doi.org/10.1002/bies.201400191

3. Ekimler S, Sahin K. Computational Methods for MicroRNA Target Prediction. Genes 2014; 5:671-83; PMID:25153283; http://dx.doi.org/10.3390/genes5030671

4. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009; 136:215-33; PMID:19167326; http://dx.doi.org/10.1016/j.cell.2009.01.002

5. Leoni G, Tramontano A. A structural view of microRNA-target recognition. Nucleic Acids Res 2016; 44:e82; PMID:26825463; http://dx.doi.org/10.1093/nar/gkw043

6. Elcheva I, Goswami S, Noubissi FK, Spiegelman VS. CRD-BP protects the coding region of betaTrCP1 mRNA from miR-183-mediated degradation. Mol Cell 2009; 35:240-6; PMID:19647520; http://dx.doi.org/10.1016/j.molcel.2009.06.007

7. Gan HH, Gunsalus KC. Tertiary structure-based analysis of microRNA-target interactions. RNA 2013; 19:539-51; PMID:23417009; http://dx.doi.org/10.1261/rna.035691.112

8. Zhuang R, Rao JN, Zou T, Liu L, Xiao L, Cao S, Hansraj NZ, Gorospe M, Wang JY. miR-195 competes with HuR to modulate stim1 mRNA stability and regulate cell migration. Nucleic Acids Res 2013; 41:7905-19; PMID:23804758; http://dx.doi.org/10.1093/nar/gkt565

9. Wu X, Chesoni S, Rondeau G, Tempesta C, Patel R, Charles S, Daginawala N, Zucconi BE, Kishor A, Xu G, et al. Combinatorial mRNA binding by AUF1 and Argonaute 2 controls decay of selected target mRNAs. Nucleic Acids Res 2013; 41:2644-58; PMID:23303783; http://dx.doi.org/10.1093/nar/gks1453

10. Liu C, Rennie WA, Carmack CS, Kanoria S, Cheng J, Lu J, Ding Y. Effects of genetic variations on microRNA: target interactions. Nucleic Acids Res 2014; 42:9543-52; PMID:25081214; http://dx.doi.org/10.1093/nar/gku675

11. Gan HH, Gunsalus KC. Assembly and analysis of eukaryotic Argonaute-RNA complexes in microRNA-target recognition. Nucleic Acids Res 2015; 43:9613-25; PMID:26432829; http://dx.doi.org/10.1093/nar/gkv990

12. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 2009; 460:479-86; PMID:19536157; http://dx.doi.org/10.1038/nature08170

13. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 2010; 141:129-41; PMID:20371350; http://dx.doi.org/10.1016/j.cell.2010.03.009

14. Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol 2010; 17:909-15; PMID:20601959; http://dx.doi.org/10.1038/nsmb.1838

15. Chi SW, Hannon GJ, Darnell RB. An alternative mode of microRNA target recognition. Nat Struct Mol Biol 2012; 19:321-7; PMID:22343717; http://dx.doi.org/10.1038/nsmb.2230

16. Loeb GB, Khan AA, Canner D, Hiatt JB, Shendure J, Darnell RB, Leslie CS, Rudensky AY. Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. Mol Cell 2012; 48:760-70; PMID:23142080; http://dx.doi.org/10.1016/j.molcel.2012.10.002

17. Betel D, Koppal A, Agius P, Sander C, Leslie C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biol 2010; 11:R90; PMID:20799968; http://dx.doi.org/10.1186/gb-2010-11-8-r90

18. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 2007; 27:91-105; PMID:17612493; http://dx.doi.org/10.1016/j.molcel.2007.06.017

19. Lal A, Navarro F, Maher CA, Maliszewski LE, Yan N, O'Day E, Chowdhury D, Dykxhoorn DM, Tsai P, Hofmann O, et al. miR-24 Inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to "seedless" 3′UTR microRNA recognition elements. Mol Cell 2009; 35:610-25; PMID:19748357; http://dx.doi.org/10.1016/j.molcel.2009.08.020

20. Shin C, Nam JW, Farh KK, Chiang HR, Shkumatava A, Bartel DP. Expanding the microRNA targeting code: functional sites with centered pairing. Mol Cell 2010; 38:789-802; PMID:20620952; http://dx.doi.org/10.1016/j.molcel.2010.06.005

21. Moore MJ, Scheel TK, Luna JM, Park CY, Fak JJ, Nishiuchi E, Rice CM, Darnell RB. miRNA-target chimeras reveal miRNA 3′-end pairing as a major determinant of Argonaute target specificity. Nat Commun 2015; 6:8864; PMID:26602609; http://dx.doi.org/10.1038/ncomms9864

22. Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell 2013; 153:654-65; PMID:23622248; http://dx.doi.org/10.1016/j.cell.2013.03.043

23. Helwak A, Tollervey D. Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). Nat Protoc 2014; 9:711-28; PMID:24577361; http://dx.doi.org/10.1038/nprot.2014.043

24. Hausser J, Zavolan M. Identification and consequences of miRNA-target interactions - beyond repression of gene expression. Nat Rev 2014; 15:599-612; PMID:25022902; http://dx.doi.org/10.1038/nrg3765

25. Orom UA, Lund AH. Isolation of microRNA targets using biotinylated synthetic microRNAs. Methods 2007; 43:162-5; PMID:17889804; http://dx.doi.org/10.1016/j.ymeth.2007.04.007

26. Baigude H, Ahsanullah Li Z, Zhou Y, Rana TM. miR-TRAP: a benchtop chemical biology strategy to identify microRNA targets. Angew Chem Int Ed Engl 2012; PMID:22566243; http://dx.doi.org/10.1002/anie.201201512

27. Tan SM, Lieberman J. Capture and identification of miRNA targets by biotin pulldown and RNA-seq. Methods Mol Biol 2016; 1358:211-28; PMID:26463386; http://dx.doi.org/10.1007/978-1-4939-3067-8_13

28. Imig J, Brunschweiger A, Brummer A, Guennewig B, Mittal N, Kishore S, Tsikrika P, Gerber AP, Zavolan M, Hall J. miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. Nat Chem Biol 2014; 11(2):107-14; PMID:25531890; http://dx.doi.org/10.1038/nchembio.1713

29. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006; 103:2257-61; PMID:16461460; http://dx.doi.org/10.1073/pnas.0510565103

30. Medina PP, Nolde M, Slack FJ. OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. Nature 2010; 467:86-90; PMID:20693987; http://dx.doi.org/10.1038/nature09284

31. Wang P, Zou F, Zhang X, Li H, Dulak A, Tomko RJ, Jr., Lazo JS, Wang Z, Zhang L, Yu J. microRNA-21 negatively regulates Cdc25A and cell cycle progression in colon cancer cells. Cancer Res 2009; 69:8157-65; PMID:19826040; http://dx.doi.org/10.1158/0008-5472.CAN-09-1996

32. Petersen M, Wengel J. LNA: a versatile tool for therapeutics and genomics. Trends Biotechnol 2003; 21:74-81; PMID:12573856; http://dx.doi.org/10.1016/S0167-7799(02)00038-0

33. Obad S, dos Santos CO, Petri A, Heidenblad M, Broom O, Ruse C, Fu C, Lindow M, Stenvang J, Straarup EM, et al. Silencing of microRNA families by seed-targeting tiny LNAs. Nat Genet 2011; 43:371-8; PMID:21423181; http://dx.doi.org/10.1038/ng.786

34. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. Fast and effective prediction of microRNA/target duplexes. RNA 2004; 10:1507-17; PMID:15383676; http://dx.doi.org/10.1261/rna.5248604

35. Hsu JB, Chiu CM, Hsu SD, Huang WY, Chien CH, Lee TY, Huang HD. miRTar: an integrated system for identifying miRNA-target interactions in human. BMC Bioinformatics 2011; 12:300; PMID:21791068; http://dx.doi.org/10.1186/1471-2105-12-300

36. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005; 120:15-20; PMID:15652477; http://dx.doi.org/10.1016/j.cell.2004.12.035

37. Schirle NT, Sheu-Gruttadauria J, MacRae IJ. Gene regulation. Structural basis for microRNA targeting. Science 2014; 346:608-13 PMID:25359968; http://dx.doi.org/10.1126/science.1258040

38. Ebert MS, Sharp PA. Roles for MicroRNAs in Conferring Robustness to Biological Processes. Cell 2012; 149:515-24; PMID:22541426; http://dx.doi.org/10.1016/j.cell.2012.04.005

39. Didiano D, Hobert O. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. Nat Struct Mol Biol 2006; 13:849-51; PMID:16921378; http://dx.doi.org/10.1038/nsmb1138

40. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature 2008; 455:1124-8; PMID:18806776; http://dx.doi.org/10.1038/nature07299

41. Duursma AM, Kedde M, Schrier M, le Sage C, Agami R. miR-148 targets human DNMT3b protein coding region. RNA 2008; 14:872-7; PMID:18367714; http://dx.doi.org/10.1261/rna.972008

42. Reczko M, Maragkakis M, Alexiou P, Grosse I, Hatzigeorgiou AG. Functional microRNA targets in protein coding sequences. Bioinformatics 2012; 28:771-6; PMID:22285563; http://dx.doi.org/10.1093/bioinformatics/bts043

43. Forman JJ, Legesse-Miller A, Coller HA. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. Proc Natl Acad Sci U S A 2008; 105:14879-84; PMID:18812516; http://dx.doi.org/10.1073/pnas.0803230105

44. Takagi S, Nakajima M, Kida K, Yamaura Y, Fukami T, Yokoi T. MicroRNAs regulate human hepatocyte nuclear factor 4alpha, modulating the expression of metabolic enzymes and cell cycle. J Biol Chem 2010; 285:4415-22; PMID:20018894; http://dx.doi.org/10.1074/jbc.M109.085431

45. Schnall-Levin M, Zhao Y, Perrimon N, Berger B. Conserved microRNA targeting in Drosophila is as widespread in coding regions as in 3′UTRs. Proc Natl Acad Sci U S A 2010; 107:15751-6; PMID:20729470; http://dx.doi.org/10.1073/pnas.1006172107

46. Zarnegar BJ, Flynn RA, Shen Y, Do BT, Chang HY, Khavari PA. irCLIP platform for efficient characterization of protein-RNA interactions. Nat Methods 2016; 13:489-92; PMID:27111506; http://dx.doi.org/10.1038/nmeth.3840

47. Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, Blue SM, Nguyen TB, Surka C, Elkins K, et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods 2016; 13:508-14; PMID:27018577; http://dx.doi.org/10.1038/nmeth.3810

48. Sterling CH, Veksler-Lublinsky I, Ambros V. An efficient and sensitive method for preparing cDNA libraries from scarce biological samples. Nucleic Acids Res 2015; 43:e1; PMID:25056322; http://dx.doi.org/10.1093/nar/gku637

49. Toth J, Biggin MD. The specificity of protein-DNA crosslinking by formaldehyde: in vitro and in drosophila embryos. Nucleic Acids Res 2000; 28:e4; PMID:10606672; http://dx.doi.org/10.1093/nar/28.2.e4

50. Meng F, Henson R, Wehbe-Janek H, Ghoshal K, Jacob ST, Patel T. MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. Gastroenterology 2007; 133:647-58; PMID:17681183; http://dx.doi.org/10.1053/j.gastro.2007.05.022

51. Asangani IA, Rasheed SA, Nikolova DA, Leupold JH, Colburn NH, Post S, Allgayer H. MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. Oncogene 2008; 27:2128-36; PMID:17968323; http://dx.doi.org/10.1038/sj.onc.1210856

52. Zhu S, Si ML, Wu H, Mo YY. MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). J Biol Chem 2007; 282:14328-36; PMID:17363372; http://dx.doi.org/10.1074/jbc.M611393200

53. Frankel LB, Christoffersen NR, Jacobsen A, Lindow M, Krogh A, Lund AH. Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. J Biol Chem 2008; 283:1026-33; PMID:17991735; http://dx.doi.org/10.1074/jbc.M707224200

54. Kim YK, Wee G, Park J, Kim J, Baek D, Kim JS, Kim VN. TALEN-based knockout library for human microRNAs. Nat Struct Mol Biol 2013; 20:1458-64; PMID:24213537; http://dx.doi.org/10.1038/nsmb.2701

55. Uhde-Stone C, Sarkar N, Antes T, Otoc N, Kim Y, Jiang YJ, Lu B. A TALEN-based strategy for efficient bi-allelic miRNA ablation in human cells. RNA 2014; 20:948-55; PMID:24717974; http://dx.doi.org/10.1261/rna.042010.113

56. Ho TT, Zhou N, Huang J, Koirala P, Xu M, Fung R, Wu F, Mo YY. Targeting non-coding RNAs with the CRISPR/Cas9 system in human

cell lines. Nucleic Acids Res 2015; 43:e17; PMID:25414344; http://dx.doi.org/10.1093/nar/gku1198

57. Park CY, Jeker LT, Carver-Moore K, Oh A, Liu HJ, Cameron R, Richards H, Li Z, Adler D, Yoshinaga Y, et al. A resource for the conditional ablation of microRNAs in the mouse. Cell Rep 2012; 1:385-91; PMID:22570807; http://dx.doi.org/10.1016/j.celrep.2012.02.008

58. Chen YW, Song S, Weng R, Verma P, Kugler JM, Buescher M, Rouam S, Cohen SM. Systematic study of Drosophila microRNA functions using a collection of targeted knockout mutations. Dev Cell 2014; 31:784-800; PMID:25535920; http://dx.doi.org/10.1016/j.devcel.2014.11.029

59. Miska EA, Alvarez-Saavedra E, Abbott AL, Lau NC, Hellman AB, McGonagle SM, Bartel DP, Ambros VR, Horvitz HR. Most caenorhabditis elegans microRNAs are individually not essential for development or viability. PLoS Genet 2007; 3:2395-403; PMID:18085825; http://dx.doi.org/10.1371/journal.pgen.0030215

60. Zheng G, Cochella L, Liu J, Hobert O, Li WH. Temporal and Spatial Regulation of MicroRNA Activity with Photoactivatable Cantimirs. ACS Chem Biol 2011; 6:1332-8; PMID:21977972; http://dx.doi.org/10.1021/cb200290e

61. Moore MJ, Zhang C, Gantman EC, Mele A, Darnell JC, Darnell RB. Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. Nat Protoc 2014; 9:263-93; PMID:24407355; http://dx.doi.org/10.1038/nprot.2014.012

62. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol 2010; 11:R106; PMID:20979621; http://dx.doi.org/10.1186/gb-2010-11-10-r106

63. Spector DL, Lamond AI. Nuclear speckles. Cold Spring Harb Perspect Biol 2011; 3:a000646; PMID:20926517; http://dx.doi.org/10.1101/cshperspect.a000646