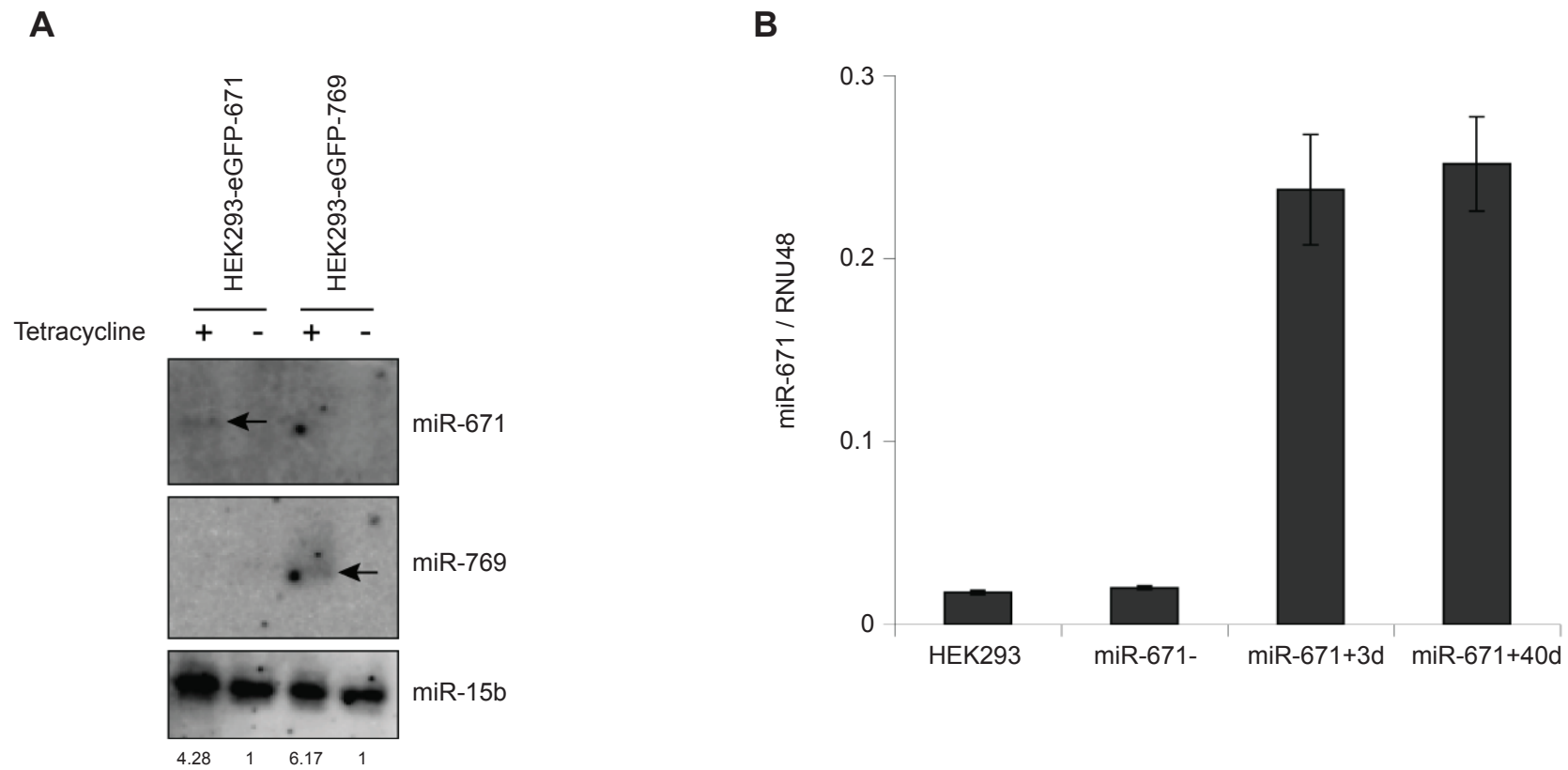
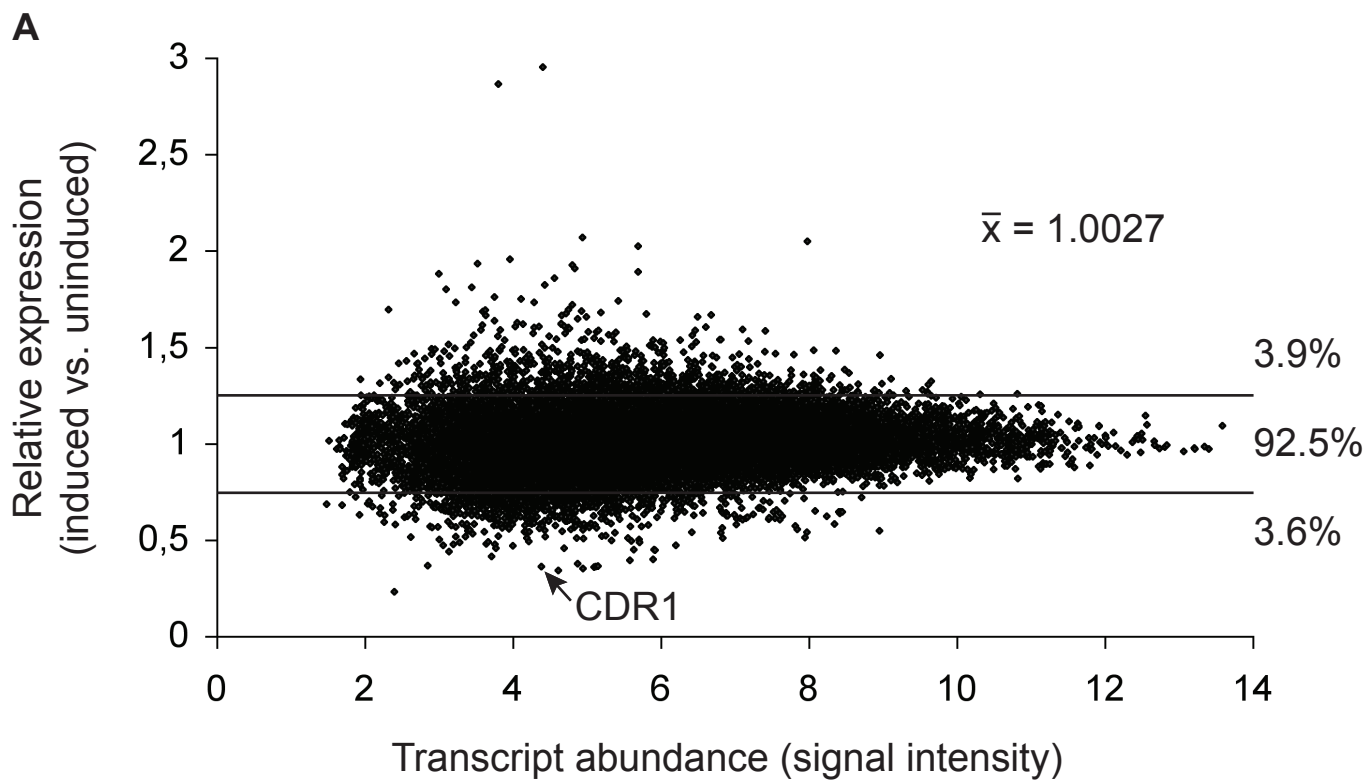


## Supplementary Information

# Supplementary Figure 1

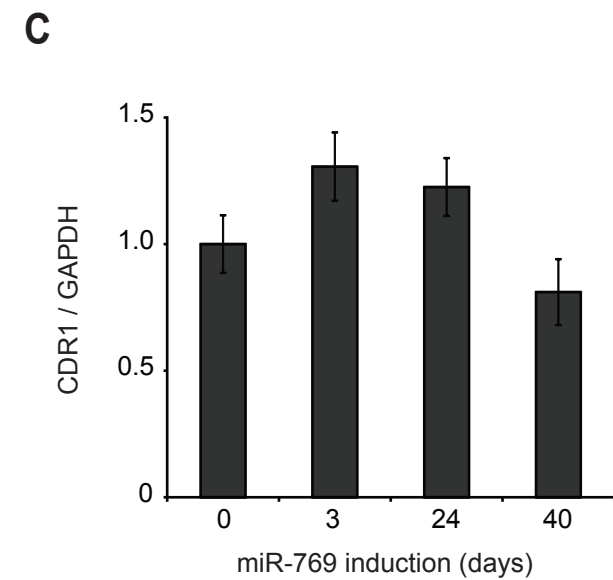


## Supplementary Figure 2



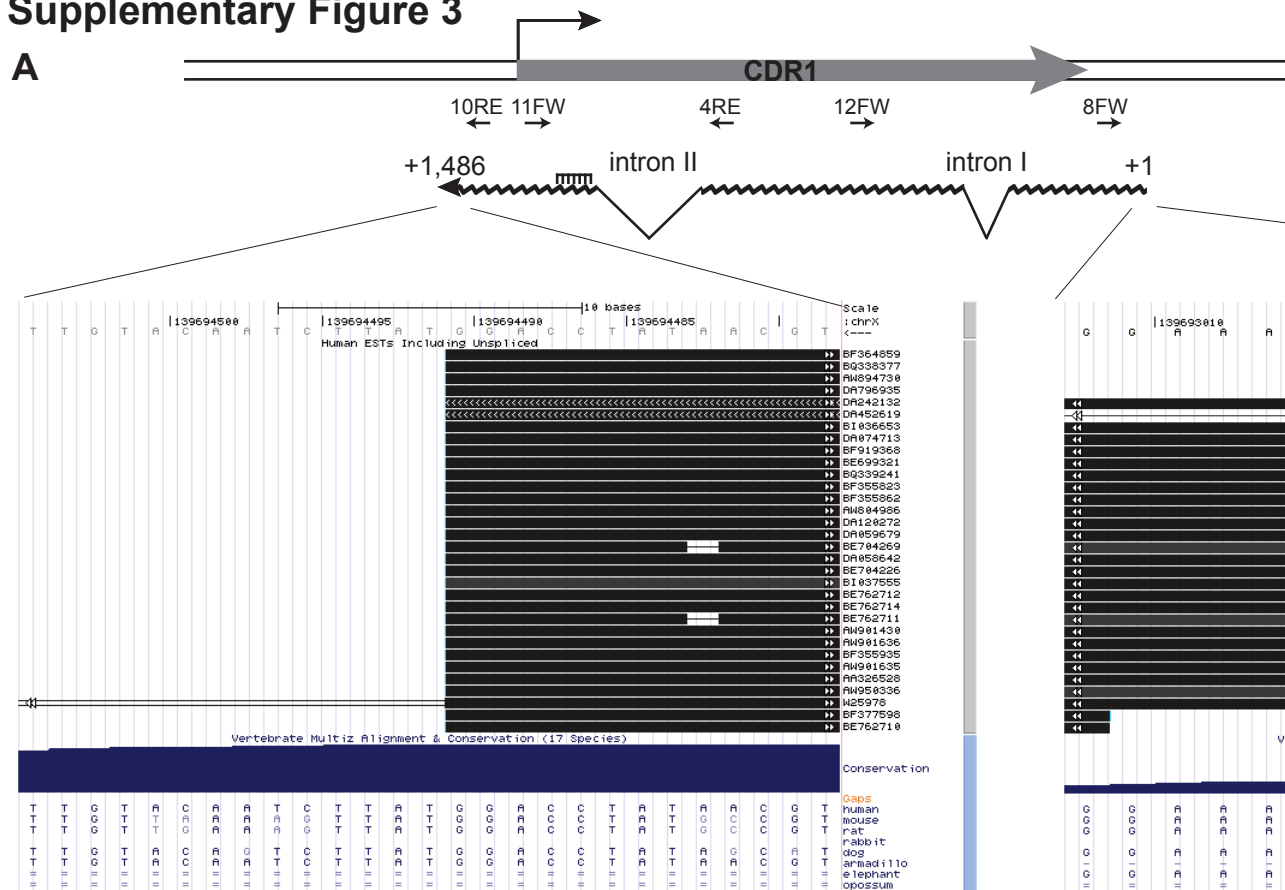
**B** **Relative CDR1 expression**

miR-671		miR-769	
<i>3d</i>	<i>40d</i>	<i>3d</i>	<i>40d</i>
0,47	0,36	1,06	0,83



# Supplementary Figure 3

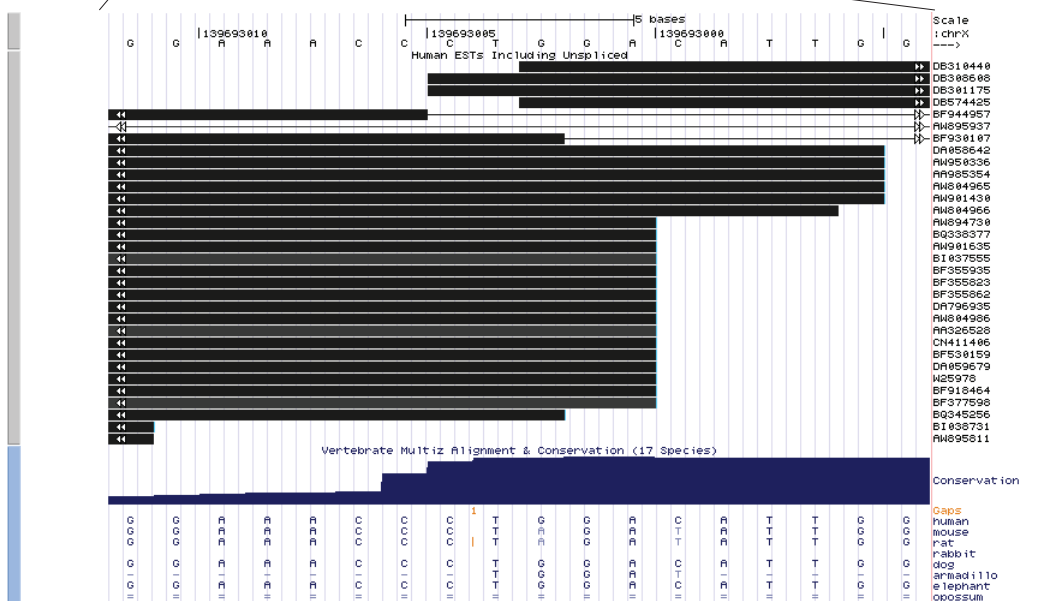
**A**



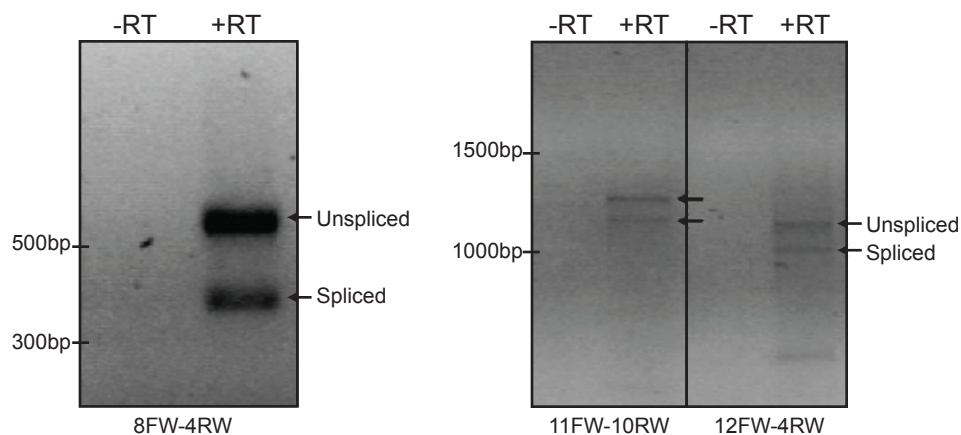
**B**

NAS aligning ESTs

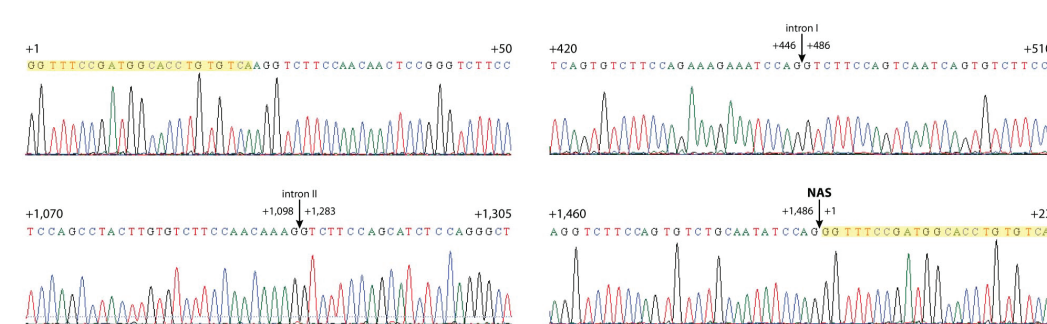
AA326528	BE704226	DA058642
AW804986	BF355823	DA059679
AW894730	BF355862	DA074713
AW901430	BF355935	DA120272
AW901635	BF364859	DA242132
AW901636	BF919368	DA452619
AW950336	BQ339241	DA796935
BE699321	CN411406	



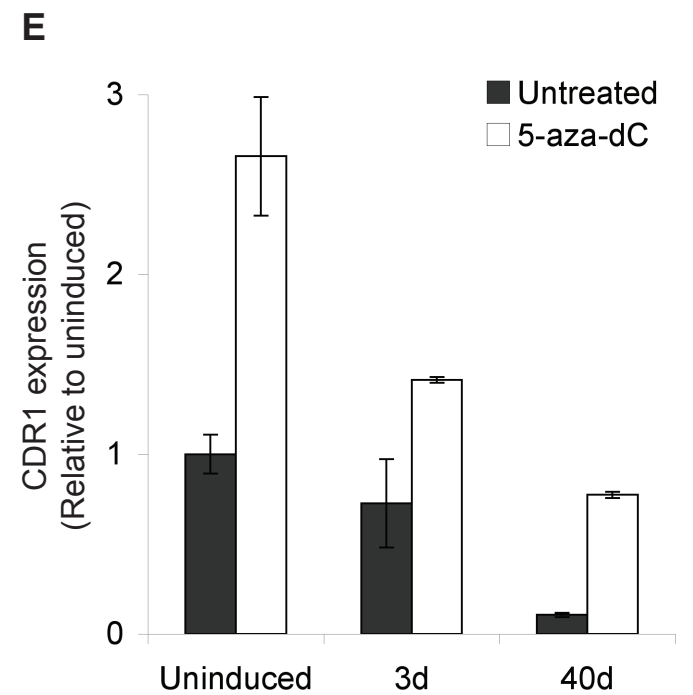
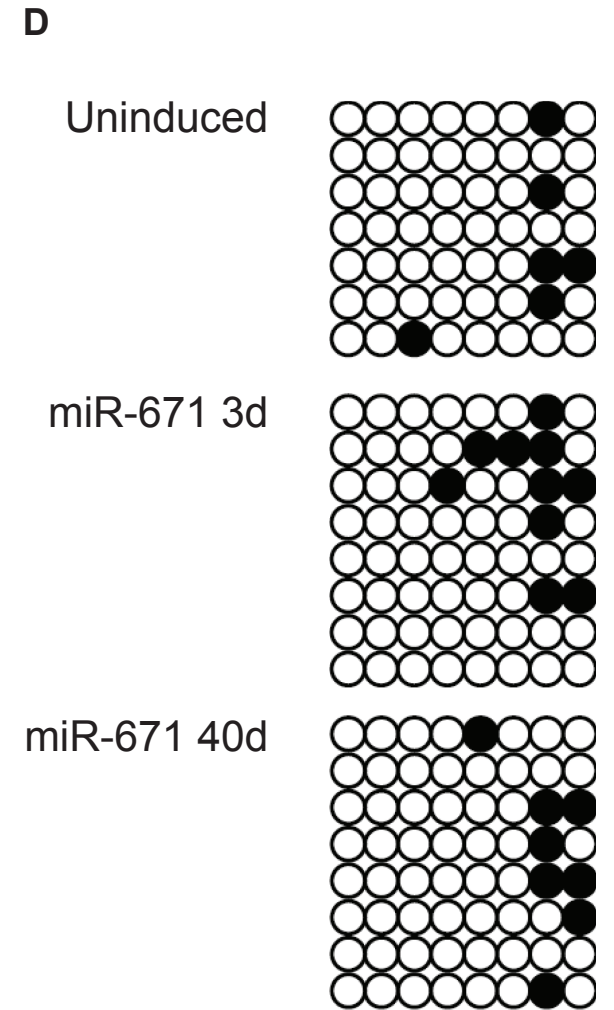
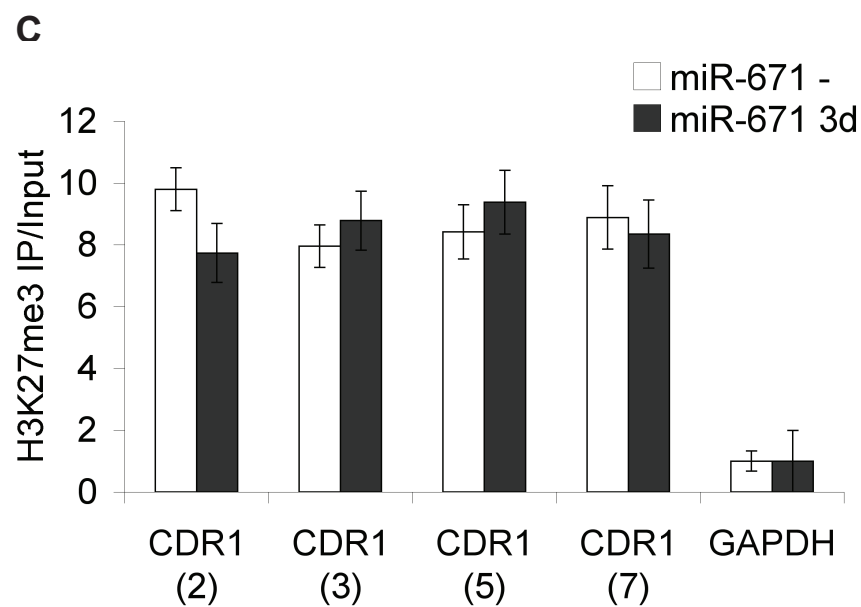
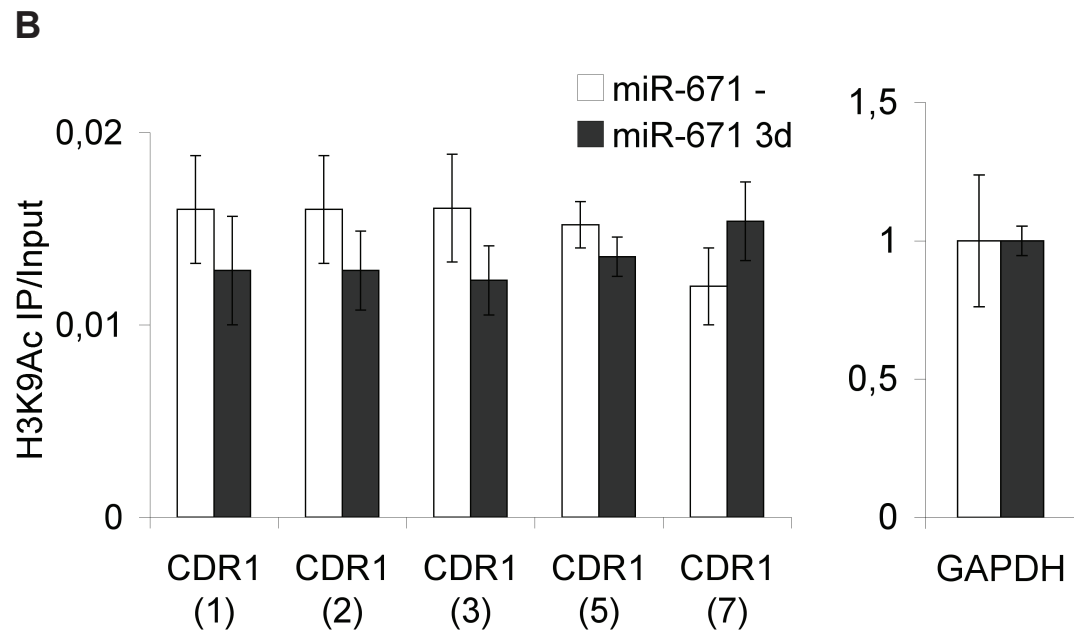
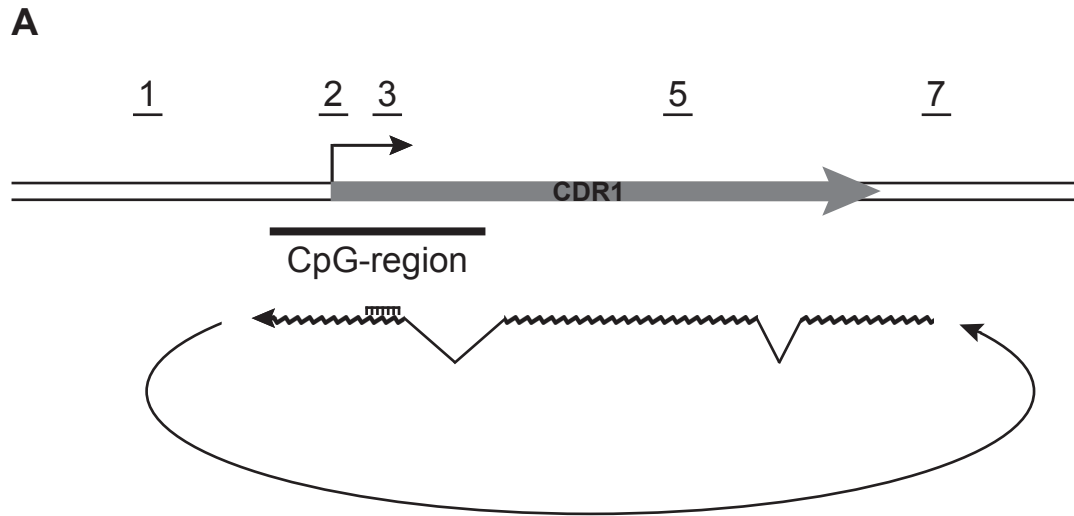
**C**



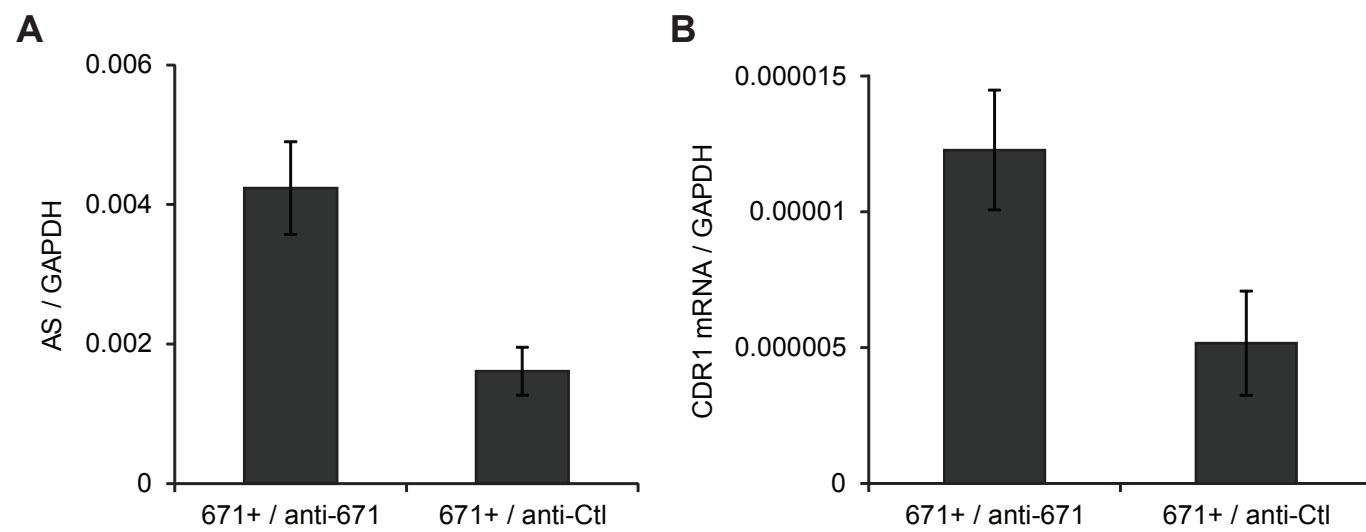
**D**



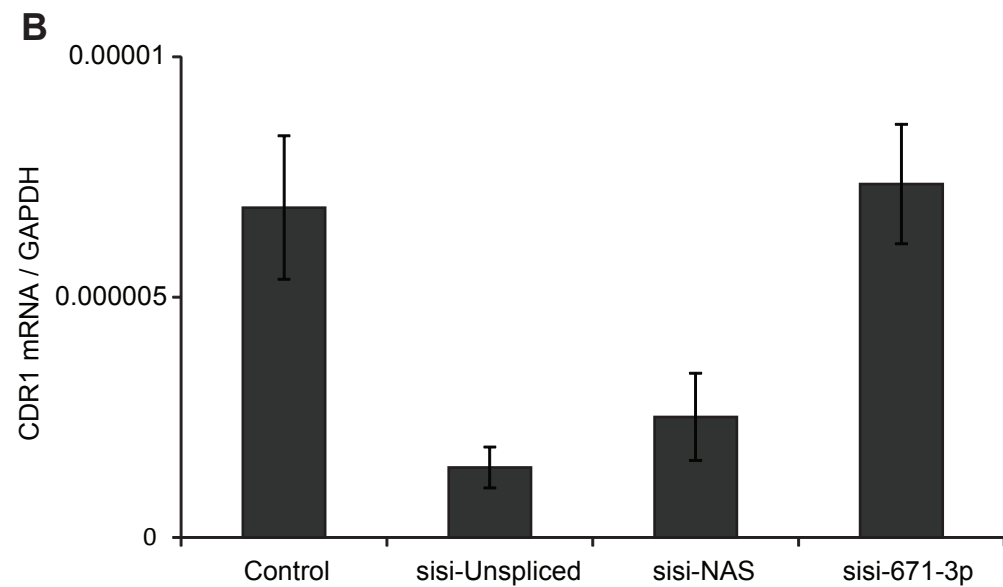
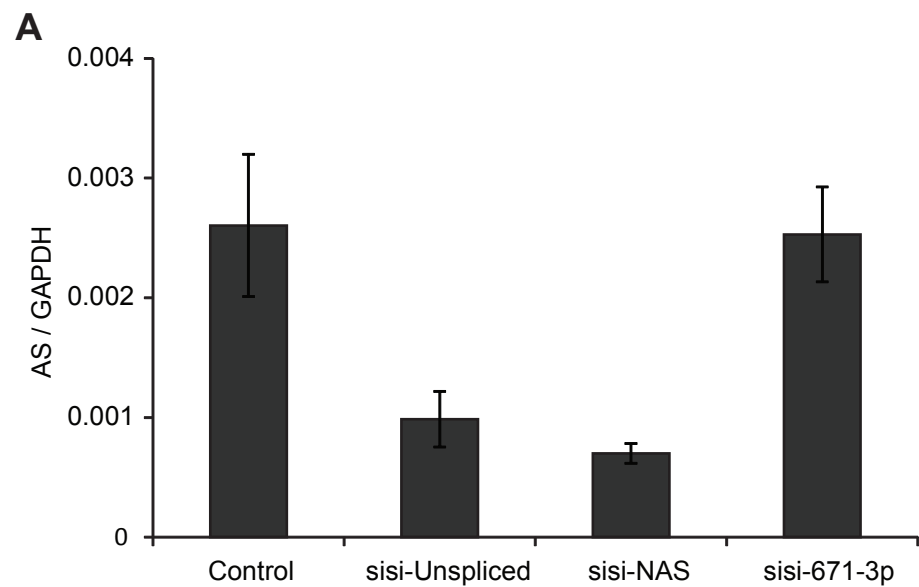
# Supplementary Figure 4



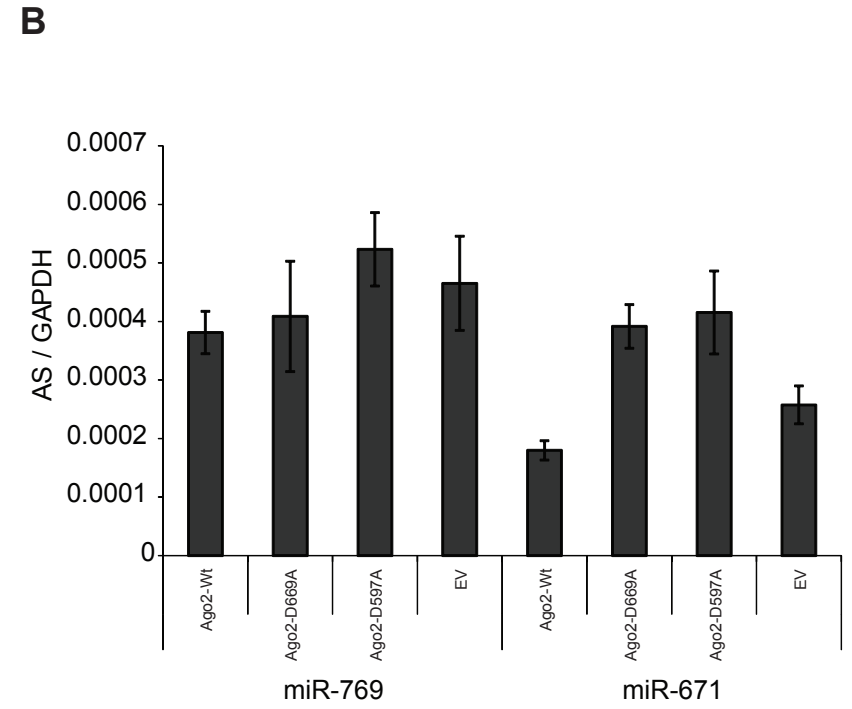
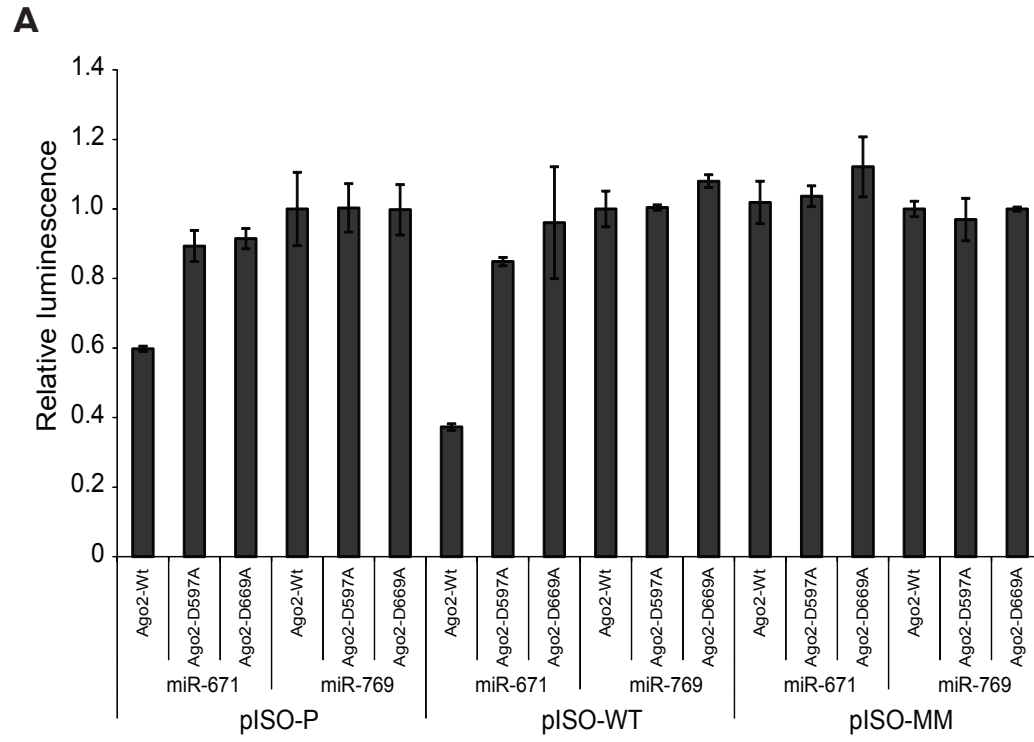
## Supplementary Figure 5



## Supplementary Figure 6

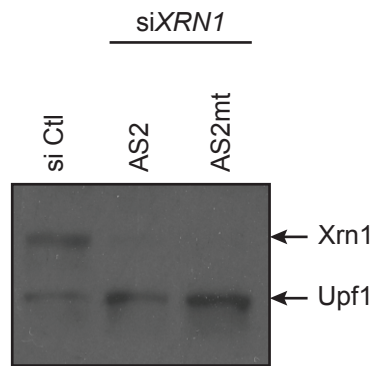


# Supplementary Figure 7

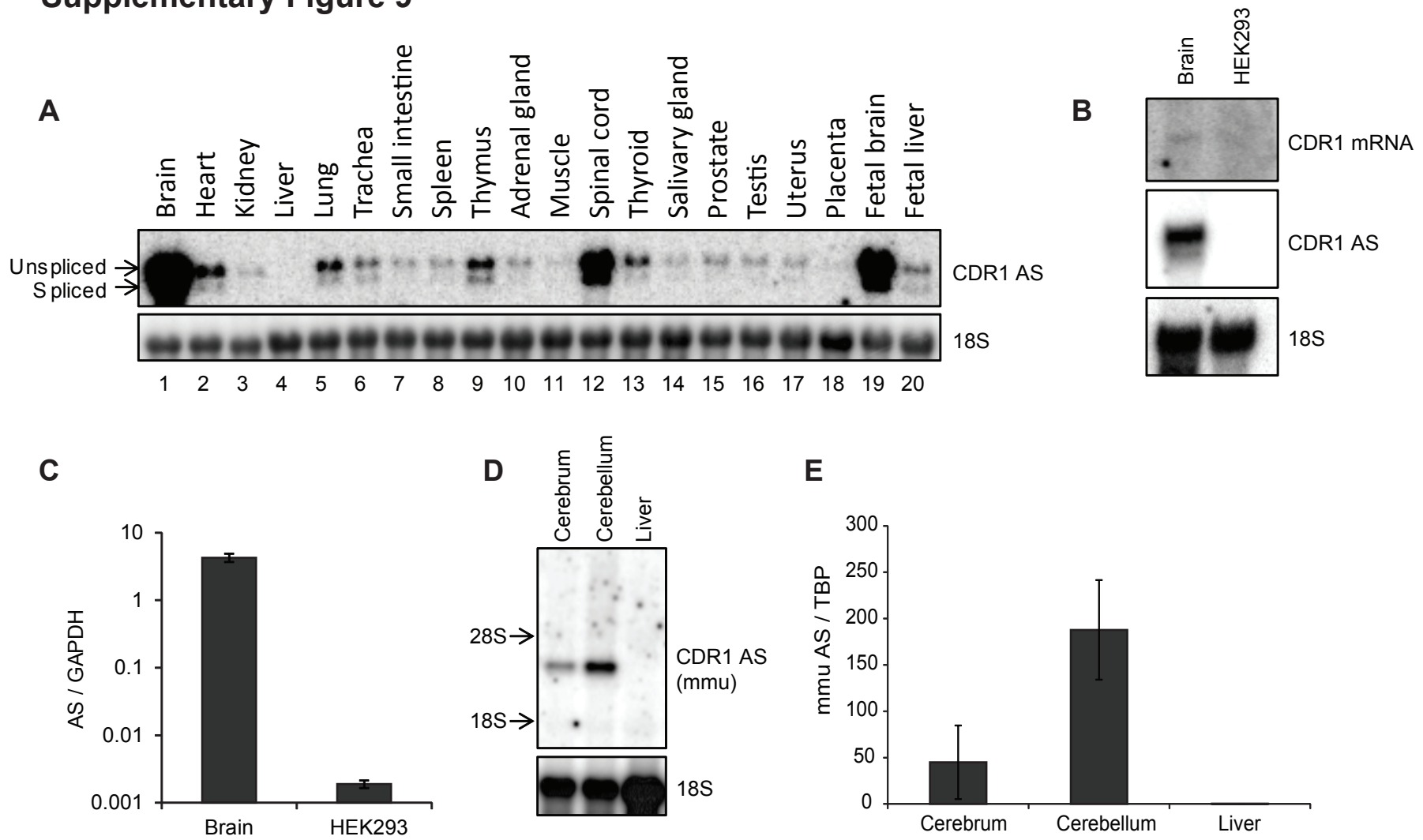




## Supplementary Figure 8



# Supplementary Figure 9



## Supplementary Table 1

### Vector construction

CDR1 AS ts RE GTGCCATCGGAAACCCCTGGATATTGCAGACACTGGA  
 CDR1 AS ts FW TCCAGTGTCTGCAATATCCAGGGTTCCGATGGCAC  
 CDR1 AS FW TTCAGAAGCTTCTATAGCTTCCAAAAATCCG  
 CDR1 AS1 RE TTCAGCTCGAGCTGGATATTGCAGACACTG  
 CDR1 AS2 RE TTCAGCTCGAGAAGACCTGGATTTCTTTCTGG  
 CDR1 AS3 FW TTCAGAAGCTTCAGGGTTCCGATGGCAC  
 CDR1 mRNA FW TTCAGAAGCTTTTGTGGAAGACATGAAGATGCTG  
 CDR1 mRNA RE TTCAGCTCGAGCCAGATAATCCTGAGCTTCCAGAA  
 CDR1 mt FW TGAGACCTCTGGTTCTGAGCTACCTGGAGATATAGGAAGACATGG  
 CDR1 mt RE AGCTCAGAACCAGAGGTCTCAGCATCTGCTCGTCTTCCA

pISO - 671 P target FW ACCTCCAGCCCTCCAGGGCTTCCT  
 pISO - 671 P target RE CTAGAGGAAGCCCTGGAGGGCTGGAGGTAGCT  
 pISO - 671 AS target FW CTTCAGCATCTCCAGGGCTTCCA  
 pISO - 671 AS target RE CTAGTGAAGCCCTGGAGATGCTGGAAGAGCT  
 pISO - 671 MM target FW CTTCAGCATGAGGAGGGCTTCCA  
 pISO - 671 MM target RE ctagTGAAGCCCTCCTCATGCTGGAAGagct  
 miR-671 FW TATGTCGCGGCCCGCCCTGTCACACAGAGATCAC  
 miR-671 RE TATGTCGTCGACAAAGAAAGGTAATGAGGTTAGGGCC  
 miR-769 FW TATGTCGCGGCCCGCAGATTGGGACTCAGGCCTAG  
 miR-769 RE TATGTCGTCGACCTCATCCAGAATTGCCACACTAC  
 RL FW TTCAGCTTAAGAGATCTGCGGCCGGAATTCTCATGGATCCATGACTTCGAAAGTTTATGATCCAG  
 RL RE CTGAACCTCGAGTTATTGTTTCATTTTTGAGAAGCTCGC

### Northern

miR-671 Probe ACCTCCAGCCCTCCAGGGCTTCCT  
 miR-769-5p Probe AGCTCAGAACCAGAGGTCTCA  
 U6 snRNA Probe GAATTTGCGTGTCTCCTTGCAGCGGGCCATGCTAA  
 Lysine tRNA Probe CTGATGCTTACCGACTGAGCTATCCGGGC  
 15b Probe TGTAACCATGATGTGCTGCTA  
 CDR1 AS Probe AGACCCGGAGTTGTTGGAAGACCTTGACACAGGTGCCATCGGAAACCCTGGATATTGCAGACACTGGAAGACCTGAATGT  
 NeoR Probe AGGTAGCCGGATCAAGCGTATGCAGCCCGCCATTGCATCAGCCATGATGGATACCTTTCT  
 eGFP Probe GCTCGATGCGGTTTCCAGGGTGTCCCTCGAACTTCACTCGGCGCGGGTCTTGTAGTTGCCGCTCCTTGAAGAAG  
 18S Probe TTACCCGGCTGTGGCACCAGACTTGCCTCCAATGGATCCTCGTTAAAGGATTTAAAGTGGACTCTTCCAAATTACAG  
 CDR1 mRNA Probe CAGTGTCTTCCAGAAAGAAATCCAGGTCTTCCAGTCAGTCAGTGTCTTCCAGAAAAATCTACGTC  
 mmu CDR1 AS Probe ACTTAGATTTCTTGGAAAGACTGGATTTTTCTGGAAGACATGGATTTCTGGAAGACA

### PCR

CDR1 ssRT ACTTGCTCAGTTCGCTACTAGTCTTCCGGATAATTGG  
 CDR1 ssRT FW AGGCCATTGGAAGATGTG  
 CDR1 ssRT RE ACTTGCTCAGTTCGCTACT  
 GAPDH RT FW GTCAGCCGCATCTTCTTTTG  
 GAPDH RT RE GCGCCAATACGACCAATC  
 7SK RT FW GTCGAAGAGGACGACCAT  
 7SK RT RE GGAGTCTTGAAGCTTGACT  
 BGH RE TAGAAGGCACAGTCGAGG  
 CDR1(1) FW CAACAGTACTCGACTTTGTGTTTTT  
 CDR1(1) RE TTGCAGCAGTCTTATCTGTTT  
 CDR1(2) FW TCTTCCATCAACTGGCTCAA  
 CDR1(2) RE TGCAGACACTGGAAGACCTG  
 CDR1(3) FW CAGTCTTCCATCAACTGGCTCA  
 CDR1(3) RE CTGAATGTCAGAAGATCAGCACACT  
 CDR1(4) FW TGGAAGACGTACCTTTGTTG  
 CDR1(4) RE ATAGCTTCCGAAAAATCCAG  
 CDR1(5) FW CTGGAGGCCATTTGGAAGAT  
 CDR1(5) RE CGTCTTCCCAACAATCCAA  
 CDR1(6) FW GGATTTGGTGAAGACCA  
 CDR1(7) FW CAGGGCCTCCAGAAATATCA  
 CDR1(7) RE AACCTGGACATTGGAACA  
 CDR1(8) FW CTTGACACAGGTGCCATC  
 CDR1(10)RE ACGTCTCCAGTGTGCTGA  
 CDR1(11)FW GGAAGACACTGATTGACTGG  
 CDR1(12)FW AGACCTGGAGGCCATTG  
 mmu NAS FW ACCTTGGTACTGGCACCA  
 mmu NAS RE TGCGCAGTGTCTCCAGT  
 TBP FW AGAGAGCCACGGACAACCTG  
 TBP RE ACTCTAGCATATTTCTTCTGCTGCT

### sisRNA

sisRNA-AS SS 5' GCUUACUUGU  
 sisRNA-AS SS 3' GUUUCCAAUC  
 sisRNA-AS AS UUGGAGACACAAAGUAGGCU  
 sisRNA-S SS 5' GGUUUUGUUG  
 sisRNA-S SS 3' GAAGACGUUUC  
 sisRNA-S AS UACGUUUCCAAACAAUUCU  
 sisi-Unspliced SS 5' UCUUCCAAUC  
 sisi-Unspliced SS 3' AAGCCAUUGU  
 sisi-Unspliced AS ACAUGGUUUUGGUAAGACU  
 sisi-NAS SS 5' CUGGAGAGCC  
 sisi-NAS SS 3' UUUUUUGGAAU  
 sisi-NAS AS UCCAAACAAGGUUUCCAGCA  
 sisi-671-3p AS UCCGGUUUCAGGGCUCCACC  
 sisi-671-3p SS 5' UGGAGCCUUG  
 sisi-671-3p SS 3' AGAACCGGAUU

## Supplementary figure legends

**Supplementary Figure 1: Validation of inducible miR-671 expression.** (A) Northern blot showing tetracycline induced mature miRNA expression in HEK293-eGFP-671 cells and HEK293-eGFP-769 cells. Bands corresponding to mature miRNAs are indicated with arrows. miR-15b is included as a loading control. Calculated intensities of miR-671 and miR-769 relative to miR-15b and normalized to uninduced signal are indicated below. (B) TaqMan® qRT-PCR validation of mature miR-671 expression in wild-type HEK293 FRT/TO cells and uninduced, 3 days and 40 days induced HEK293-eGFP-671 cells (normalized to RNU48).

**Supplementary Figure 2: Microarray expression data.** (A) Affymetrix HU Gene 1.0 ST expression array scatter plot of transcript abundance vs. relative change for uninduced and 3 days induced HEK293-eGFP-671 cells. (B) Relative *CDR1* mRNA expression in 3 days (3d) and 40 days (40d) induced vs. uninduced HEK293-eGFP-671 and HEK293-eGFP-769 cells (Affymetrix gene expression arrays). (C) Quantification of relative *CDR1* expression by qRT-PCR after miR-769 induction as indicated.

**Supplementary Figure 3: NAS specific ESTs and RT-PCR.** (A) UCSC screen dump of the region surrounding the antisense NAS splice sites, with the top and bottom panels reflecting the splice donor (SD) and splice acceptor (SA), respectively. The arrow indicates the splice site position. For SD (top), the additional G-nucleotide in the aligned EST accessions is a misalignment due to the fact that the first exonic nucleotide at the SA is also a G. For SA, the exonic sequence upstream of the SD site (CAATATCCAG, top panel) corresponds to the SA site sequence (CAATGTCCAG), except for a single nucleotide mismatch (underlined). The EST-sequences

covering the mismatched nucleotide all harbor the SD sequence indicative of NAS. **(B)** List of all identified EST accessions covering the NAS splice site. **(C)** NAS specific RT-PCR using *CDRI*(8)FW / *CDRI*(4)RE (left) or *CDRI*(11)FW / *CDRI*(10)RE and *CDRI*(12)FW / *CDRI*(4)RE (right) as primer sets. Primer positions are schematically depicted in (A). Unspliced and spliced annotation refers to the 184 nt optionally spliced traditional intron II proximal to the miR-671 target site. **(D)** Sanger sequencing of RT-PCR products from (C) confirming splicing and NAS in HEK293 cells. Highlighted region indicates exonic sequence immediately downstream of NAS SA.

**Supplementary Figure 4: *CDRI* repression is not associated with epigenetic silencing.** **(A)** Schematic showing the regions in the *CDRI* locus assessed for chromatin changes (numbered 1-5) and DNA methylation (CpG region). **(B, C)** ChIP for active H3K9ac **(B)** and repressive H3K27me3 **(C)** marks in uninduced (white bars) compared 3d induced (grey bars) HEK293-eGFP-671 cells. Enrichment was quantified by qPCR relative to input material for 5 regions in the *CDRI* locus as indicated in **(A)**. **(D)** DNA methylation levels in uninduced, 3d and 40 d induced HEK293-eGFP-671 cells determined by clonal bisulphite sequencing of a region covering 8 CpG dinucleotides spanning the *CDRI* TSS **(A)**. White circles: unmethylated CpG; solid circles: methylated CpG. **(E)** Effect of DNA methylation depletion by 5-aza-dC treatment (48h, 3 $\mu$ M) on *CDRI* expression in uninduced compared 3d and 40d induced HEK293-eGFP-671 cells.

**Supplementary Figure 5. Anti-miR-671-5p treatment of tetracycline induced HEK293-eGFP-671 cells.** **(A-B)** Strand specific qRT-PCR quantifying antisense (A) or mRNA (b) levels from HEK293 cells induced for 3 days and subsequently treated with 100 nM anti-miR or antiControl while kept on tetracycline rich media.

**Supplementary Figure 6: sisiRNA-mediated knockdown.** (A-B) Quantification of AS (A) and *CDRI* mRNA (B) on RNA from HEK293 cells transfected with, sisiRNA targeting within intron II (sisiUnspliced AS), sisiRNA targeting exon-exon junction produced by NAS-splicing (sisiNAS), sisiRNA resembling the mature miR-671-3p sequence (sisi-671-3p) or a control siRNA.

**Supplementary Figure 7: Ago2-mediated cleavage of the *CDRI* antisense.** (A) Firefly luciferase reporters (pISO) with either a perfect miR-671 target (P target), the endogenous *CDRI* antisense target (AS target), or mismatched target (MM target), were co-transfected with a Renilla luciferase expression vector (pcDNA3-RL), miR-671 (pJEBB-671) or miR-769 (pJEBB-769) and Ago2 wild-type or mutant expression vectors as indicated (Ago2-D669A or Ago2-D597A). Data are displayed as relative Firefly/Renilla luminescence. (B) Antisense NAS specific qRT-PCR on RNA from HEK293 cells transiently transfected miR-671 or miR-769 along with empty vector (EV), Ago2 wild-type (Ago2-wt), or Ago2-slicer mutant (Ago2-D669A or Ago2-D597A) expression vectors.

**Supplementary Figure 8: Xrn1 knockdown.** Western blot showing siRNA knockdown efficiency of Xrn1, Upf1 serves as loading control.

**Supplementary Figure 9: Tissue atlas of antisense expression.** (A) Northern blotting using 2 ug total RNA from 20 different human tissues probed with NAS specific probe (upper panel) or 18S load control (lower panel). (B) Northern blotting using 10 ug RNA from human brain and HEK293 probed against *CDRI* mRNA (upper panel), NAS (middle panel) or 18S loading control (lower panel). (C) Quantification of AS expression levels in brain and HEK293 relative to *GAPDH*. (D) Northern blotting using 0.5 ug RNA from mouse cerebrum, cerebellum and liver probed with AS

specific probe (upper panel) or 18S loading control (lower panel). (E) qRT-PCR on mouse RNA from (D) using a NAS-specific primer set relative to *TBP*.

## Supplementary Materials and Methods

### Expression constructs and stable cell lines

HEK293 cells were maintained under standard culture conditions. miR-671 and miR-769 were PCR amplified from human genomic DNA using 671FW/671RE and 769FW/769RE, respectively. Fragments were subcloned into an intron of an eGFP expression vector (pJEBB-671 and pJEBB-769) and subsequently inserted into the pcDNA5 FRT/TO vector (Invitrogen, Carlsbad, CA) (pcDNA5-eGFP-671 and pcDNA5-eGFP-769). Stable cell lines (HEK293-eGFP-671 and HEK293-eGFP-769) were generated by co-transfecting HEK293 Flp-In T-Rex cells (Invitrogen) with pOG44 and pcDNA5-eGFP-671 or pcDNA5-eGFP-769 using calcium phosphate and otherwise according to supplier's protocol. A final concentration of 1 µg/ml tetracycline (Sigma-Aldrich) was added to the media to induce expression.

The AS1, AS3 and CDR1 mRNA expression vectors were generated by PCR on genomic DNA using *CDR1 AS FW/CDR1 AS1 RE*, *CDR1 AS3 FW/CDR1 AS2 RE*, or *CDR1 mRNA FW/CDR1 mRNA1 RE*, respectively, subsequent digest with HindIII/XhoI and ligation into pcDNA3 (Invitrogen). The AS2 antisense expressing vector was generated by overlapping PCR on genomic DNA using *CDR1 AS FW/CDR1 AS ts RE* and *CDR1 AS ts FW/CDR1 AS2 RE*. PCR products

were combined and amplified using *CDR1 AS FW/CDR1 AS2 RE*. The complete product was inserted into pcDNA3 (Invitrogen).

AS1mt and AS2mt were generated from AS1 and AS2, respectively, using overlapping PCR with primersets *CDR1 AS FW/CDR1 mt FW* and *CDR1 AS1 RE/CDR1 mt RE*, or *CDR1 AS FW/CDR1 mt FW* and *CDR1 AS2 RE/CDR1 mt RE*, digested with HindIII and XhoI and inserted into pcDNA3.

Firefly Luciferase reporter vectors were generated by inserting annealed primers (listed in Supplementary Table S1) into XbaI and SacI digested pISO (Lewis et al, 2003). The Renilla expression vector (pcDNA3-RL) was generated by PCR amplifying the Renilla ORF from pRluc-N2 (PerkinElmer, Waltham MA) with RL FW/RL RE and cloning the amplicon into pcDNA3 (Invitrogen).

### **Microarray analyses**

Total RNA was extracted from cells in culture using TRIzol® reagent (Invitrogen) according to standard procedures and stored in aqueous solution at -80°C. RNA quality was assessed on a Bioanalyzer® 2100 microfluidics chip (Agilent, Santa Clara CA), and samples with 28S/18S ratio <1.5 and RIN <9.5 were excluded from microarray analysis. Global gene expression was analysed by Affymetrix (Santa Clara, CA) Human Gene 1.0 ST microarrays performed at the Ramaciotti Centre for Gene Function Analysis (University of New South Wales, Randwick Australia). Raw data was analyzed with the Affymetrix Microarray Suite version 5.1 software.

### **5'RACE**

RACE was performed on total RNA (100 ng) from HEK293 cells subjected to *XRN1* knockdown and miR-671 and AS2 overexpression using the 5' RACE System for Rapid Amplification of



cDNA Ends, Version 2.0 (Invitrogen) adhering to the manufacturer's protocol. BGH RE was used to prime the RT-reaction and CDR1(6)FW and CDR1(8)FW were used in the subsequent nested PCR. The nested PCR product was TOPO-cloned using the TOPO TA Cloning® kit for sequencing (Invitrogen). *XRNI* knockdown was achieved as previously described (Damgaard et al, 2008) using the siRNA duplex: 5'AGAUGAACUUACCGUAGAAAdTdT / 5'UUCUACGGUAAGUUCAUCUdTdT.

### **Luciferase reporter assay**

HEK293 cells were transfected with 0.3 µg pJEBB-671/pJEBB-769, 0.2 µg pISO, and 0.1 µg pcDNA3-RL, and harvested after 48 h using the Dual-Luciferase® Reporter Assay kit (Promega, Madison WI). In case of Ago2 overexpression, 0.4 µg Ago2 expression vector was added to the transfection mix. Luminescence was measured on a BMG FLUOstar luminometer (BMG labtech, Offenburg Germany). Experiments were conducted in biological triplicates.

### **Chromatin IP**

ChIP was performed as described in Damgaard *et al* or using the Magna ChIP™ system (Millipore, Billerica MA) with ChIP-grade antibodies against acetylated H3K9 (07-352, Millipore/Upstate, Billerica MA), and trimethylated H3K27 (07-449, Millipore/Upstate). IP enrichment was quantified by triplicate qPCR using Power SYBR® Green master mix (Applied Biosystems, Foster City CA) on a LightCycler® 480 (Roche, Basel Switzerland).

### **DNA methylation analysis**

Cells grown in monolayer were allowed to settle for 24 hours after seeding, and were then treated with the DNA demethylating agent 5-aza-2'-deoxycytidine (5-Aza-dC) (Sigma-Aldrich) (3.5µM,

48h). The effect on gene expression was subsequently analyzed by qRT-PCR as described above. CpG methylation levels were determined for genomic DNA isolated from cells in culture using the DNeasy® blood & tissue kit (Qiagen, Valencia CA) according to the supplied protocol. Bisulphite conversion of genomic DNA was carried out as previously described and bisulphite specific sequencing primers were designed according to the principles outlined in Clark *et al* (Clark et al, 2006).

### **Statistical analysis**

Statistical significance was determined by the standard two-tailed Student's t test.

### **References**

Clark SJ, Statham A, Stirzaker C, Molloy PL, & Frommer M (2006) DNA methylation: bisulphite modification and analysis. *Nat Protoc* **1**: 2353-2364

Damgaard CK, Kahns S, Lykke-Andersen S, Nielsen AL, Jensen TH, & Kjems J (2008) A 5' splice site enhances the recruitment of basal transcription initiation factors in vivo. *Mol Cell* **29**: 271-278

Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, & Burge CB (2003) Prediction of mammalian microRNA targets. *Cell* **115**: 787-98