iScience, Volume 23

Supplemental Information

Biosynthesis of Circular RNA

ciRS-7/CDR1as Is Mediated

by Mammalian-wide Interspersed Repeats

Rei Yoshimoto, Karim Rahimi, Thomas B. Hansen, Jørgen Kjems, and Akila Mayeda

Figure S1. Evidence for 'Back-splicing' rather than 'Intra-lariat splicing' to produce circular ciRS-7 (Related to Figure 1)

(A) Two proposed pathways for circRNA biosynthesis. Left panel: the generation by an intra-lariat splicing of lariat precursor (L1) that is excised by the preceding exon(s)-skipped splicing of pre-mRNA (P). Right panel: the circRNA (C) generation by a direct back-splicing event that is promoted by the looped pre-mRNA (P′) formed by base-paring of the flanking intronic complementary sequences. Blue and red bars indicate the antisense oligoribonucleotides (ASOs), numbered 1–6 in panels C and D, annealed to prevent the following pathways (Exonskipping, Intra-lariat splicing, Back-splicing, or Second-splicing).

(B) Detection of intermediates (P', L1) and products (C, M) from the ciRS-7 precursor (P). Total RNA from human cerebral cortex was treated with (+) or without (–) RNase R to distinguish opened linear RNA from closed RNA. The purified RNA was analyzed by RT–PCR with indicated primers (arrowheads).

(C) Schematic representation of the ciRS-7 precursor and the alternatively spliced isoforms (see panel B and Figure S2). Bars indicate ASOs (1–6) to block specific alternative splicing event.

(D) The quantified data of splicing prevention using ASOs (1–6; see panel C). SH-SY5Y cells were electroporated with each ASO. After 24 h incubation, RT–PCR analysis was performed with ciRS-7 primers and the control GAPDH primers. The detected ciRS-7 expression levels were normalized to GAPDH (ciRS-7/GAPDH) and plotted as ratios to the value of control cells treated with GFP-ASO. Means ± standard deviation (SD) are given for three independent experiments (***P < 0.001).

Figure S2. Identified Alternative Splice Sites Might Cause Skipping of the ciRS-7 Exon

(Related to Figures 1, S1)

Modified screeenshot of human brain RNA-Seq data (GSE59612) mapped on the the human hg19 genome sequence is indicating the assembled ciRS-7 precursor transcript. The junction containing reads revealed the active alternative 5′ and 3′ splice sites within the exons 5 and 6.

Figure S3. CRISPR/Cas9-Mediated Deletion of the Flanking Alternative Splice Sites in ciRS-7 Gene Supports 'Back-Splicing Pathway' in Biosynthesis of ciRS-7

(Related to Figures 1, S1)

(A) Schematic structures of the ciRS-7 precursor and its alternatively spliced isoforms (orange box indicates ciRS-7 exon). The positions of the guide RNAs (gRNA1–gRNA5) targetting the alternative splice sites used in 'Intra-lariat splicing pathway' (Figure S1A), PCR primers for detecting deleted sites (filled triangles), and those for detecting undeleted sites (open triangeles) are also indicated.

(B) The targeted ciRS-7 genomic deletions in HEK-293 cells were verified by genomic PCR. The indicated three pairs of gRNAs were used to delete the alternative splice sites. PCR primers indicated in panel A were used for detecting deleted sites (open triangles) and non-deleted sites (filled triangles). PCR primers for the *NEAT1* gene were used as control.

(C) The deletion of either the alternative 5′ splice sites (between gRNA2 and gRNA3) or the 3′ splice sites (between gRNA4 and gRNA5) outside of ciRS-7 exon barely prevented the generation of ciRS-7. The ciRS-7 production was analyzed by quantitative RT–PCR with ciRS-7 primers and the control GAPDH primers. The ciRS-7 expression levels were normalized to the control expression level of GAPDH (ciRS-7/GAPDH). The results were plotted as ratios to the value of control wild-type cells. Means ± standard deviation (SD) are given for three independent experiments (**P < 0.01).

Figure S4. The ciRS-7 Exon Is Not Flanked by Inverted *Alu* **Elements**

(Related to Figure 1)

The locations of the human genomic *Alu* elements are shown together with the assembled ciRS-7 precursor transcript (see Figure S2 above). The *Alu* elements (bars) were extracted from UCSC Repeat Masker database (+ and – are the same and opposite strand of the ciRS-7 exon, respectively). Analysis of 3′ rapid amplification of cDNA ends (RACE) detected the 3' terminal of ciRS-7 precursor with canonical poly (A) tail (from downstream of chrX: 139873517), indicating that the transcript is terminated around the end of exon 6.

Figure S5. Deletion of Flanking MIRs Causes Accumulation of Precursor

(Related to Figure 2)

(A) Total RNA (used in Figure 2B) was quantified by RT–qPCR. The ciRS-7 precursor expression levels were normalized to GAPDH (Precursor/GAPDH) and compared to wild-type (WT) plasmid-expressing HEK293 cells. Means ± SD are given for three independent experiments. Means ± SD are given for three independent experiments $(***P < 0.001, **P < 0.01, *P < 0.05).$

(B) Total RNA (used in Figure 2B), with (+) or without (–) RNase R treatment, was analyzed by Northern blotting.

A MIR-dependent circRNAs

B MIR-independent circRNAs

Figure S6. Additional Human circRNAs Were Identifed with Flanking Inverted MIR Elements (Related to Figure 3)

The loci of two MIR-dependent circRNAs **(A)** and three MIR-independent circRNAs **(B)** are shown. The positions of the host genes (with red numbered exons), circRNA exons (orange boxes), and the identified inverted MIR elements (purple bars) are indicated. Red numbers indicate the Smith-Waterman (SW) alignment scores of these MIR elements. The SW alignment scores of MIRs in MIR-dependent human ciRS-7 (see Figure 1A) are 945 (upstream) and 1077 (downstream).

Figure S7. CRISPR/Cas9-Mediated Deletion of Either Upstream or Downstream MIRs Aborts the Production of circCDK8

(Related to Figure 3)

(A) Schematic genomic structure of circCDK8 locus with the flanking inverted MIR sequences. The positions of guide RNAs (gRNA1–gRNA4) to delete each MIR element are indicated with red vertical arrowheads. PCR primers for detecting deleted sites are indicated with filled triangles.

(B) The genomic deletions of the flanking MIR elements in HEK293 cells were verified by genomic PCR. The indicated two pairs of gRNAs were used to delete the 5′ and 3′ MIR elements. PCR primers indicated in panel A were used for detecting the MIR-deleted sites (open triangles in KO#1 and KO#2) and non-deleted sites (filled triangles in Control).

(C) The effects of the MIR deletions (5′MIR KO and 3′MIR KO) on the circCDK8 production was quantified by qRT– PCR. The ciRS-7 expression levels were normalized to the control expression level of GAPDH (circCDK8/GAPDH). Values are relative to the value of undeleted parental cells (WT). Means ± standard deviation (SD) are given for three independent experiments (*P < 0.05).

Figure S8. Inverted Sequences Are Less Complementary in MIRs than in *Alu***s**

(Related to Figures 3–6)

RNA secondary structures were predicted by RNAcofold of the Vienna RNA package. The cofold structure of MIRdependent circRNAs **(A)**, MIR-independent circRNAs (Cf. mouse ciRS-7, see Figure 1A) **(B)**, and *Alu*-dependent circRNAs **(C)** are shown. The corrected thermal stability (–kcal/mol) of each MIR was calculated by the RNAup program.

Table S1 List of all the synthetic oligonucleotides used in the experiments (Related to Figures 1–3, 5, 6, S1, S3, S5, S7)

TRANSPARENT METHODS

Analysis of RNA-Seq Datasets

We used RNA-Seq datasets from non-neoplastic brain tissue (GSE59612; Gill et al., 2014). Obtained data were aligned to reference human genome hg19 using HISAT2 (Kim et al., 2015) and the aligned sequence reads were assembled by StringTie (Pertea et al., 2015). Repeated sequences were downloaded from the RepeatMasker track in the UCSC table browser (https://genome.ucsc.edu/cgi-bin/hgTables) and analyzed with BEDtools (http://bedtools.readthedocs.io).

To search for potential MIR-dependent circRNAs, we used 'SupTable1- LOW Backsplice Set.xls' file (Supplemental Table S1 in Jeck et al., 2013) that contains the back-splice sites' information along with the chromosome location, the starting/ending positions in the chromosome, the corresponding strand, and the spliced reads per billion mapping (SRPBM) of identified 7771 circRNAs. The list of mouse circRNAs were obtained from circBase (Glažar et al., 2014) and the UCSC-LiftOver tool (https://genome.ucsc.edu/cgibin/hgLiftOver) was used to find mouse orthologous circRNAs.

The box plots were constructed using the R/Bioconductor package (http://www.bioconductor.org). For statistical comparisons of groups, Wilcoxon rank-sum tests were used to calculate P-values.

Possible RNA pairings between inverted MIRs were predicted using the RNAcofold program from the Vienna RNA package (Lorenz et al., 2011). Thermodynamic stability of the basepaired MIRs was calculated with the RNAup program in the Vienna RNA package using '–b' option to include the probability of unpaired regions (Lorenz et al., 2011). Exceptionally weak MIR-MIR base-pairings that the RNAup program could not calculate their thermodynamic stabilities were set to zero values ('All circRNAs' in Fig. 4D).

RT–PCR Assays

RT–PCR analysis was performed essentially as previously described (Yoshimoto et al., 2017). Human cerebral cortex total RNA was purchased from Clontech (CLN 636561). From culture cells, total RNA was isolated with a Nucleospin RNA kit (Machery Nagel) according to the manufacturer's instructions. 'On-column' DNase I digestion was performed to remove contaminated DNA. Purified RNA was reverse-transcribed using PrimeScript II (Takara Bio) with oligo-dT and random primers, and cDNA was amplified by PCR (20–30 cycles) with Ex-Taq (Takara Bio) and specific primers (Table S1). PCR-amplified products were analyzed by 8% polyacrylamide gel electrophoresis.

For nested PCR, the first PCR reaction mixture (30 cycles) was purified with a PCR cleanup column (Takara Bio) and the eluate was used for the second PCR reaction (30 cycles). The purified RNase R (1 µg) was added to digest 1 µg of total RNA to remove linear RNA in a 20 µL reaction mixture at 30ºC for 30 min as described previously (Suzuki et al., 2006).

To perform quantitative PCR (qPCR), a real-time PCR instrument (Eco Real-Time PCR System, Illumina) was used with the same primers (Table S1) that were used in the regular PCR. For the qPCR reactions, KAPA Taq PCR kit was used according to the manufacturer's protocol (Kapa Biosystems).

For the RT-qPCR analysis in Figure 2E, DNase-treated RNA (1 µg) was reverse-transcribed using a Superscript VILO cDNA Synthesis Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Obtained cDNA was mixed with SYBR Green I Master (Roche Molecular Systems) and analyzed on a real-time PCR instrument (LightCycler 480, Roche Molecular Systems) according to the manufacturer's protocol.

Northern Blot Analyses

Total RNAs (5 µg each) were separated by electrophoresis on 1% agarose containing 3% formaldehyde, rinsed twice in distilled water for 10 min, denatured in 7.5 mM NaOH for 20 min, neutralized in 20 \times SSC containing 3 M NaCl and 0.3 M sodium citrate (pH 7.0) for 20 min, and blotted overnight on a nylon membrane (RPN82B, GE Healthcare Life Sciences) followed by UV irradiation at 254 nm with 120 mJ/cm² (CL-1000, Funakoshi). DIG-labeled probes were hybridized in PerfectHyb solution (HYB-101, Toyobo) overnight at 68ºC, and hybridized probes were detected by alkaline phosphatase-conjugated anti-DIG antibodies (#11093274910, Roche) and CDP-star (#CDP-RO, Roche). The chemiluminescence signals were observed by imaging analyzer (ImageQuant LAS 500, GE Healthcare Life Sciences).

Antisense Oligoribonucleotide-Mediated Splicing Repression

Antisense 2′-*O*-methyl-modified phosphorothioate oligoribonucleotides (ASOs) were purchased from Integrated DNA Technologies (Table S1). These chemically modified ASOs were electroporated into SH-SY5Y cells (Gene Pulser Xcell, Bio-Rad). Fully confluent SH-SY5Y cells grown on 10 cm plate in D-MEM/Ham's-F12 medium (Fujifilm Wako) were trypsinized and the washed cell pellets were suspended in 1 mL OPTI-MEM medium (Thermo Fisher Scientific). This cell suspension (0.2 mL) plus ASO (2.5–10 µM final concentration) were transferred into 0.4 cm cuvette (BEX) and electroporated at 200 V for 20 ms square-wave pulses. After electroporation, cell suspensions were transferred into 6-well plates supplemented with 1.8 mL D-MEM/Ham's-F12 that was cultured for 24 h before RT–PCR analysis.

Construction and Expression of circRNA-Reporter Plasmids

The expression plasmids for ciRS-7, circCDK8, circTMEM109, circZMYNDR8, and circSRGAP2C were constructed by subcloning the PCR-amplified fragments into FLAGpcDNA3 vector using BamHI and XhoI sites (Yoshimoto et al., 2009). The expression plasmids for circSPNS1 and circHIPK3 were constructed by subcloning the PCR-amplified fragments into FLAG-pcDNA3 plasmid using an In-Fusion HD cloning kit (Takara Bio) according to the manufacturer's protocol. The deletion of MIR element (∆5′MIR, ∆3′MIR, ∆5′3′MIR) of these expression plasmids were made by KOD -Plus- Mutagenesis Kit (TOYOBO) according to the manufacturer's protocol. The expression plasmid for ciRS-7-5′MIR-RC was constructed by subcloning the PCR-amplified 5′MIR fragment into pcDNA3-ciRS-7∆5′MIR vector using BglII and BamHI sites. The expression plasmids for ciRS-7 with flanking repeat elements derived from either circCDK8, circZMYND8, or circHIPK3, were constructed by subcloning the PCRamplified ciRS-7 fragment into pcDNA3-circCDK8, pcDNA3-circZMYNDR8, or pcDNA3 circHIPK3 plasmids, respectively, using an In-Fusion HD cloning kit.

To establish HEK293 cells stably expressing ciRS-7, PCR-amplified a ciRS-7 fragment was subcloned into pcDNA5 FRT-TO vector (Thermo Fisher Scientific) using BamHI and XhoI sites, and transfected into Flp-In T-REx 293 cells (Thermo Fisher Scientific) along with Flp recombinase vector pOG44. The transfected cells were treated with 1 µg/mL DOX for 24 h and total RNA was prepared for the RT–PCR analysis. To inhibit the splicing reaction, 100 ng/mL SSA was added 18 h after DOX treatment.

Mouse N2A cells were transiently transfected with these expression plasmids using Lipofectamine 2000 (Thermo Fischer Scientific) according to the manufacturer's instructions. The transfected cells were incubated for 24 h and total RNA was prepared for RT–PCR analysis.

Targeted Genomic Deletion of the Splice Sites in the ciRS-7 Locus and of MIRs in the circCDK8 Locus

Cas9 RNP was introduced with the Alt-R CRISPR-Cas9 System (Integrated DNA Technologies) according to the manufacturer's instructions. The gRNAs in Figures S3A and S7A and primer sequences are listed in Table S1.

HEK293 cells were seeded into a 96-well plate at a density of 4×10^4 cells/well with DMEM/F-12 medium (Fujifilm Wako Pure Chemical) containing 10% fetal bovine serum (FBS, Sigma-Aldrich) and 0.5% Gibco penicillin-streptomycin mixture (Thermo Fisher Scientific). The cells were transiently transfected with 0.75 pmol of Cas9 RNP using 1.2 µL of Lipofectamine RNAiMAX (Life Technologies) diluted up to 50 µL of OPTI-MEM (Thermo Fisher Scientific). After a 20 min incubation at room temperature, the transfection solution was added dropwise to the cells. At 48 h post-transfection, the cells were trypsinized and filtered through a 40 µm Corning cell strainer (Thermo Fisher Scientific), seeded onto a 10-cm culture plate, and grown for 2 weeks. Each colony (~96 clones per construct) were picked, trypsinized, and seeded onto 96-well plate with two replicates. Genomic DNA was extracted and PCR was performed to verify the deleted region.

Targeted Genomic Deletion of MIRs in ciRS-7 Locus

The gRNAs (Figure 2C) are listed in Table S1. The annealed gRNAs were subcloned into the BpiI-cleaved (Thermo Fisher Scientific) site of a modified version of CRSPR vector px458 (#43138, Addgene) without the ITR element that expresses Cas9/EGFP (Højland Knudsen et al., 2018). To increase the efficiency and accuracy of the CRISPR targeting site (MIR elements), single-stranded donor oligodeoxyribonucleotides (ssODN) acting as repair templates were used (Table S1).

SH-SY5Y cells were seeded in six-well plates and grown to 70% confluency (0.8 \times 10⁶) cells/well) in DMEM/F-12 medium containing GlutaMax, 10% FBS, and 0.5% Gibco penicillinstreptomycin mixture (all reagents from Thermo Fisher Scientific). The cells were transiently transfected with 2.5 µg DNA (1 µg each of two gRNA vectors and 0.5 µg ssODN) using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. After 15 min incubation at room temperature, the transfection solution was added dropwise to the cells. The medium was replaced with fresh growth medium at 12 h post-transfection.

At 48 h post-transfection, the cells were prepared for fluorescent-activated cell sorting (FACS) by trypsinization, washing, and re-suspension in PBS with 2% FBS. The cells were FACS sorted in a BD FACSAria III (BD Biosciences) based on their viability and GFP signal quality into 96-well plates (single cell per well) containing 150 µl of conditioned growth medium (30% used and 70% fresh media). Three weeks after the FACS sorting, the grown colonies were passaged into duplicate 96-well plates and one set was used for DNA preparation and genomic PCR with flanking primers to detect the deleted homozygote clones. The verified deleted genomic clones and the appropriate controls were expanded and total RNA was prepared for RT-qPCR analysis (see above).

SUPPLEMENTAL REFERENCES

Gill, B.J., Pisapia, D.J., Malone, H.R., Goldstein, H., Lei, L., Sonabend, A., Yun, J., Samanamud, J., Sims, J.S., Banu, M.*, et al.* (2014). MRI-localized biopsies reveal subtype-specific differences in molecular and cellular composition at the margins of glioblastoma. Proc. Natl. Acad. Sci. USA *111*, 12550–12555.

Glažar, P., Papavasileiou, P., and Rajewsky, N. (2014). circBase: a database for circular RNAs. RNA *20*, 1666– 1670.

Højland Knudsen, C., Ásgrímsdóttir, E.S., Rahimi, K., Gill, K.P., Frandsen, S., Hvolbøl Buchholdt, S., Chen, M., Kjems, J., Febbraro, F., and Denham, M. (2018). A Modified Monomeric Red Fluorescent Protein Reporter for Assessing CRISPR Activity. Front. Cell Dev. Biol. *6*, 54.

Jeck, W.R., Sorrentino, J.A., Wang, K., Slevin, M.K., Burd, C.E., Liu, J., Marzluff, W.F., and Sharpless, N.E. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA *19*, 141–157.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360.

Lorenz, R., Bernhart, S.H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms. Mol. Biol. *6*, 26.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. *33*, 290–295.

Suzuki, H., Zuo, Y., Wang, J., Zhang, M.Q., Malhotra, A., and Mayeda, A. (2006). Characterization of RNase Rdigested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. Nucleic Acids Res. *34*, e63.

Yoshimoto, R., Kaida, D., Furuno, M., Burroughs, A.M., Noma, S., Suzuki, H., Kawamura, Y., Hayashizaki, Y., Mayeda, A., and Yoshida, M. (2017). Global analysis of pre-mRNA subcellular localization following splicing inhibition by spliceostatin A. RNA *23*, 47–57.

Yoshimoto, R., Kataoka, N., Okawa, K., and Ohno, M. (2009). Isolation and characterization of post-splicing lariatintron complexes. Nucleic Acids Res. *37*, 891–902.