Supplemental Method 1. The detailed methods of circular RNA sequencing.

NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, USA) were used according to manufacturer' constructions. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer(5X). First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150~200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then adaptor-ligated cDNA at 37°C for 15 min followed by 5 min at 95 °C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumia) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina Hiseq platform and 125 bp/150 bp paired-end reads were generated. Clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and low-quality reads from raw data. At the same time, Q20, Q30 and GC content the clean data were calculated. All the downstream analyses were based on the clean data with high quality. Data analysis contained quality control, mapping to reference genome, and circRNA identification.

Supplemental Method 2. Data analysis of circular RNA sequencing.

Raw data (raw reads) of FASTQ format were firstly processed through in-house perl

scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and low quality reads from raw data. At the same time, Q20, Q30 and GC content the clean data were calculated. All the downstream analyses were based on the clean data with high quality. Reference genome and gene model annotation files were downloaded from genome website directly. Index of the reference genome was built using Bowtie v2.2.3 and paired-end clean reads were aligned to the reference genome using Hisat2 v2.0.5. We selected Hisat2 as the mapping tool for that Hisat2 can generate a database of splice sets based on the gene model annotation file and thus a better mapping result than other non-splice mapping tools. HTSeq was used to count the reads numbers mapped to each gene. And then FPKM of each gene was calculated based on the length of the gene and reads count mapped to this gene. FPKM, expected number of Fragments Per Kilobase of transcript sequence per Millions base pairs sequenced, considers the effect of sequencing depth and gene length for the reads count at the same time, and is currently the most common used method for estimating gene expression levels (Trapnell, Cole, et al., 2010). Differential expression analysis was performed using the edgeR package. The edgeR is one of the most popular Bioconductor packages for assessing differential expression in RNA-seq data. It is based on the negative binomial (NB) distribution and it models the variation between biological replicates through the NB dispersion parameter. This method is immediately able to handle complex experimental designs. Genes with an adjusted P-value <0.05 found by DESeq were assigned as differentially expressed (Robinson MD, et al., 2010). (For DESeq2 with biological replicates) Differential expression analysis of two conditions/groups (two biological replicates per condition) was performed using the DESeq2 R package. DESeq2 provide statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting P-values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted P-value <0.05 found by DESeq2 were assigned as differentially expressed. (For DEGSeq without biological replicates) Differential expression analysis of two

conditions was performed using the DEGSeq R package. The P values were adjusted using the Benjamini & Hochberg method. Corrected P-value of 0.005 and log2(Fold change) of 1 were set as the threshold for significantly differential expression. Function enrichment analysis of differentially expressed genes was implemented by the clusterProfiler R package, including Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome pathway analysis. The GO project provides the most comprehensive resource currently available for computable knowledge regarding the functions of genes and gene products, mainly covering three aspects of biology: cell components, molecular functions and biological processes. GO terms with corrected P value less than 0.05 were considered significantly enriched by differential expressed genes. KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-through put experimental technologies (http://www.genome.jp/kegg/). Reactome is a collection database of articles on human reactions and biological pathways by experts and peer reviewed.