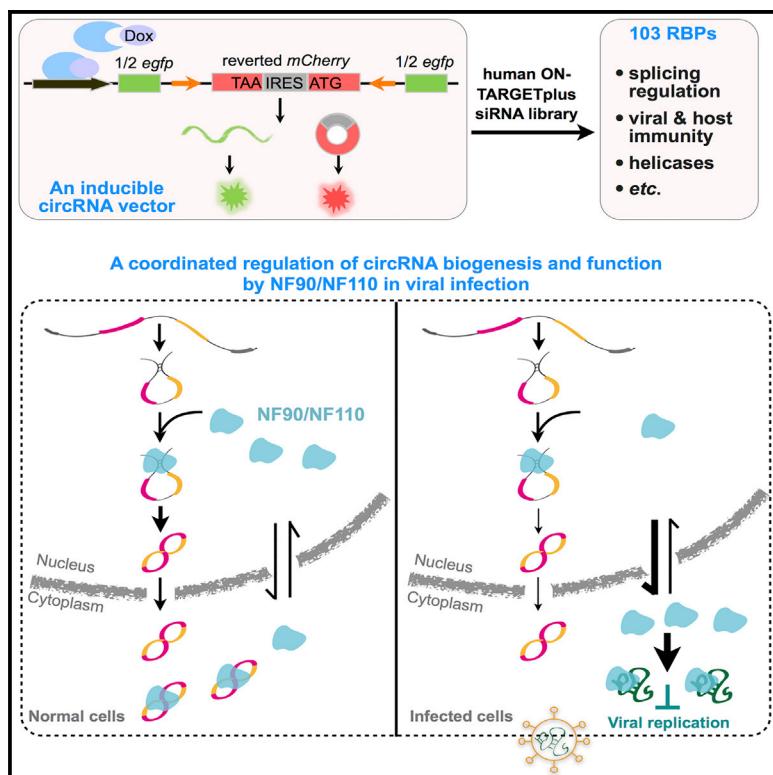


# Molecular Cell

## Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection

### Graphical Abstract



### Authors

Xiang Li, Chu-Xiao Liu, Wei Xue, ..., Run-Wen Yao, Li Yang, Ling-Ling Chen

### Correspondence

liyang@picb.ac.cn (L.Y.), linglingchen@sibcb.ac.cn (L.-L.C.)

### In Brief

Li et al. use genome-wide siRNA screening and an efficient circRNA expression reporter to identify factors involved in circRNA formation. They find that, upon viral infection, circRNA biogenesis and function are regulated by double-stranded RNA-binding domain containing immune factors NF90/NF110.

### Highlights

- Genome-wide siRNA screening identifies immune factors involved in circRNA formation
- NF90/NF110 promote back-splicing in the nucleus and also are components of circRNPs
- Nuclear export of NF90/NF110 upon viral infection leads to reduced circRNA levels
- NF90/NF110 released from circRNPs bind to viral mRNAs for antiviral immune response

# Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection

Xiang Li,<sup>1,4</sup> Chu-Xiao Liu,<sup>1,4</sup> Wei Xue,<sup>2,4</sup> Yang Zhang,<sup>1</sup> Shan Jiang,<sup>1</sup> Qing-Fei Yin,<sup>1</sup> Jia Wei,<sup>2</sup> Run-Wen Yao,<sup>1</sup> Li Yang,<sup>2,3,\*</sup> and Ling-Ling Chen<sup>1,3,5,\*</sup>

<sup>1</sup>State Key Laboratory of Molecular Biology and Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology

<sup>2</sup>Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences University of Chinese Academy of Sciences, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China

<sup>3</sup>School of Life Science and Technology, ShanghaiTech University, 100 Haik Road, Shanghai 201210, China

<sup>4</sup>These authors contributed equally

<sup>5</sup>Lead Contact

\*Correspondence: liyang@picb.ac.cn (L.Y.), linglingchen@sibcb.ac.cn (L.-L.C.)

<http://dx.doi.org/10.1016/j.molcel.2017.05.023>

## SUMMARY

Circular RNAs (circRNAs) generated via back-splicing are enhanced by flanking complementary sequences. Expression levels of circRNAs vary under different conditions, suggesting participation of protein factors in their biogenesis. Using genome-wide siRNA screening that targets all human unique genes and an efficient circRNA expression reporter, we identify double-stranded RNA-binding domain containing immune factors NF90/NF110 as key regulators in circRNA biogenesis. NF90/NF110 promote circRNA production in the nucleus by associating with intronic RNA pairs juxtaposing the circRNA-forming exon(s); they also interact with mature circRNAs in the cytoplasm. Upon viral infection, circRNA expression is decreased, in part owing to the nuclear export of NF90/NF110 to the cytoplasm. Meanwhile, NF90/NF110 released from circRNP complexes bind to viral mRNAs as part of their functions in antiviral immune response. Our results therefore implicate a coordinated regulation of circRNA biogenesis and function by NF90/NF110 in viral infection.

## INTRODUCTION

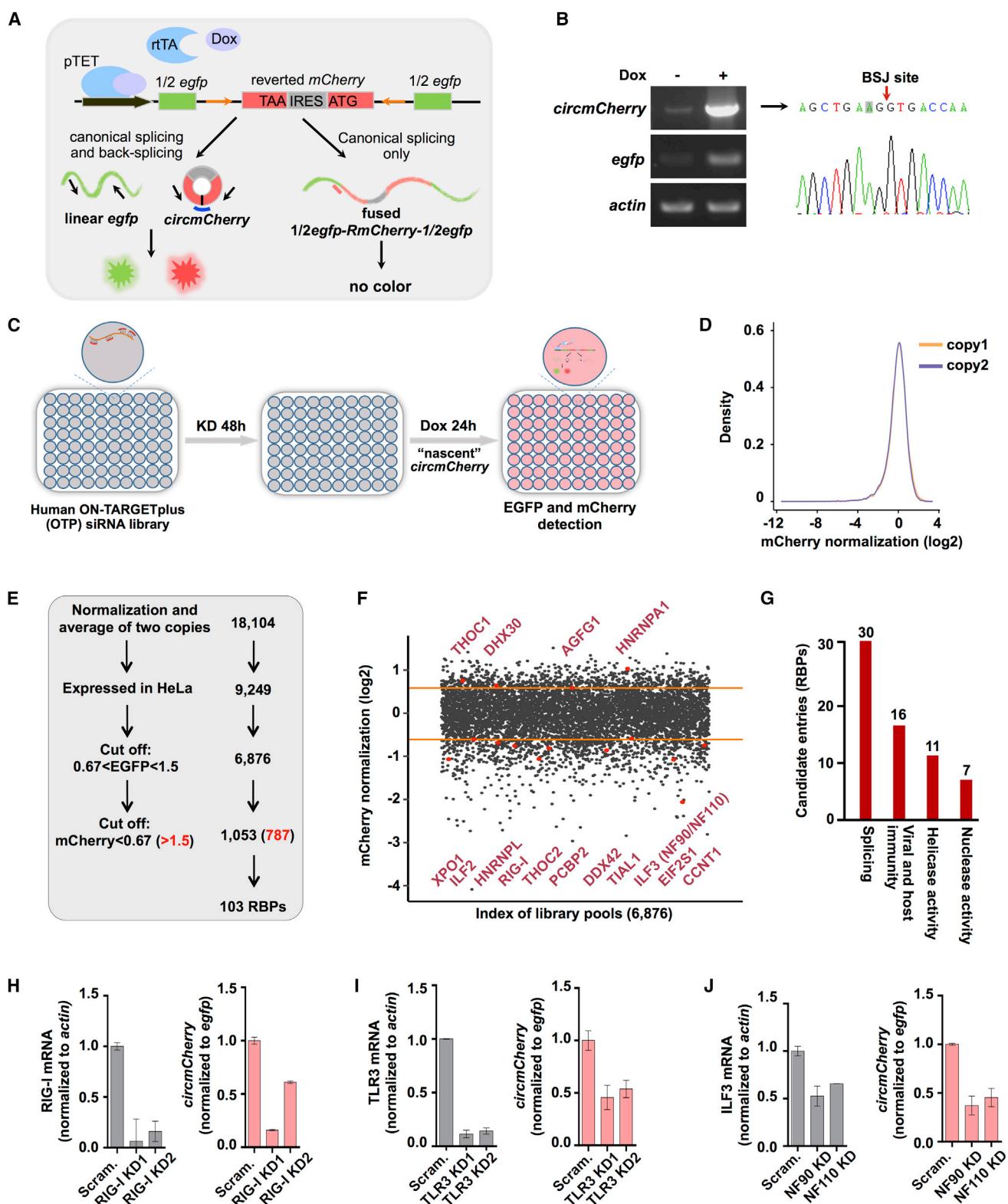
Over ten thousand different circular RNAs (circRNAs) from back-spliced exons have been recently archived across various cell lines, tissues, and species (Dong et al., 2016; Rybak-Wolf et al., 2015; Zhang et al., 2016a). Although generally expressed at low levels, the expression of some circRNAs has been reported to be as high as or even higher than that of their linear counterparts from the same gene loci (Rybak-Wolf et al., 2015; Salzman et al., 2013; You et al., 2015; Zhang et al., 2016b).

Back-splicing is inefficiently catalyzed by the spliceosome (Starke et al., 2015; Zhang et al., 2016b) and can be facilitated by paired intronic sequences that juxtapose circularized exon(s)

(Chen, 2016; Ivanov et al., 2015; Liang and Wilusz, 2014; Yang, 2015; Zhang et al., 2014). Among these paired sequences, orientation-opposite repetitive elements play important roles in promoting back-splicing in general and contribute to alternative back-splicing selection (Dong et al., 2016; Zhang et al., 2016a, 2014). The majority of human circRNAs are back-spliced from exons with inverted repeated *Alus* (IRAlus) across their flanking introns (Dong et al., 2016; Zhang et al., 2016a, 2014). Intriguingly, although having the same *cis*-elements, expression levels of circRNAs from the same loci are diverse in different cell lines and tissues (Dong et al., 2016; Rybak-Wolf et al., 2015; Salzman et al., 2013; Zhang et al., 2016a), suggesting the involvement of other layers of regulation in circRNA expression. Recent studies have revealed that fast-transcribed genes tend to produce more circRNAs (Zhang et al., 2016b) and that multiple RNA-binding proteins (RBPs) can regulate back-splicing and circRNA biogenesis (Ashwal-Fluss et al., 2014; Conn et al., 2015; Ivanov et al., 2015; Kramer et al., 2015; Rybak-Wolf et al., 2015).

It is well established that RBPs are involved in mRNA life cycle essentially in every step, from the initial stage of transcriptional production and processing in the nucleus, to the functional engagement and later on degradation in the cytoplasm (Moore, 2005). It should be the case for circRNAs as well. Recent studies have revealed several hundred known and previously unknown RBPs in human genomes (Castello et al., 2012; He et al., 2016). However, RBPs associated with circRNA processing and function remain to be explored.

Here, we apply a genome-wide siRNA screening that targets all human unique genes with an efficient circRNA expression reporter for profiling proteins that are involved in circRNA processing and function. We identified that over 100 RBPs have the potential to interfere with circRNA expression with this screening, including multiple factors related with host immune responses, such as the human interleukin enhancer binding factor 3 (ILF3). The ILF3 gene produces several isoforms of nuclear factor 90 (NF90) due to alternative splicing and alternative polyadenylation (Patiño et al., 2015). In this study, we uncovered that NF90 and its 110 isoform NF110 (NF110) could promote circRNA production by stabilizing intronic RNA pairs in the nucleus and that they were also components of circRNPs in the cytoplasm. Upon viral infection,



**Figure 1. Identification of Factors Involved in circRNA Biogenesis**

(A) A cell-based circRNA expression reporter system. An inducible circRNA expression reporter cassette, containing a pair of complementary sequences (orange arrows) that are able to form RNA pairing structure across introns (Zhang et al., 2016b) flanking the *mCherry* exon, was stably integrated into the genome of HeLa cells. See main text for details.

(legend continued on next page)

NF90/NF110 were rapidly exported to the cytoplasm, and correspondingly nascent circRNA expression was decreased. Meanwhile, the de-association of NF90/NF110 from circRNPs in the cytoplasm allowed their binding to viral mRNAs to inhibit viral replication. These findings suggest immune response factors in circRNA biogenesis and the involvement of circRNAs in viral infection.

## RESULTS

### A Dox-Inducible *circmCherry*-Expression System for Genome-wide siRNA Screening that Targets All Human Unique Genes

To identify proteins that could potentially affect circRNA biogenesis, we used a Dox-inducible *circmCherry* expression vector that was stably expressed in HeLa cells for genome-wide siRNA screening that targets all human unique genes. This Dox-*circmCherry* expression vector efficiently produces a circRNA from the reverted *mCherry* sequence (*circmCherry*) by back-splicing within a split EGFP gene (Zhang et al., 2016b). An IRES sequence and an ATG start codon were inserted into the middle of the reverted *mCherry* mRNA sequence, allowing *mCherry* protein to be translated only from the back-spliced *circmCherry* (Figure 1A). The *egfp* mRNA could be also produced by joining the split EGFP sequences together by canonical splicing to express EGFP protein. Thus, the canonical splicing and back-splicing work together to allow the expression of EGFP and *mCherry* in response to Dox addition (Figures 1A and 1B).

The great majority of cells have both EGFP produced from canonical splicing and *mCherry* produced from back-splicing (Figures S1A–S1C). siRNAs targeting the back-spliced junction site of *circmCherry* (blue arc line in Figure 1A) could efficiently reduce *mCherry* expression but not the linear *egfp* mRNA splicing and expression (Figures S1A, S1D, and S1E). This observation indicated that the fluorescent densities between *mCherry* and EGFP did not interfere with each other. A fused *egfp* and *mCherry* linear RNA could be alternatively produced (Figure 1A, right bottom). This fused *egfp* and *mCherry* linear RNA did not contribute to *mCherry* or EGFP expression, confirmed by no fluorescent reduction by the treatment of siRNAs (pink thick line in Figure 1A) that target the junction site between the first half *egfp* and the reverted *mCherry* (Figures S1F and S1G). Together, this Dox-inducible circRNA expression system would allow us to

perform the genome-wide siRNA screening (Figure 1C) for identification of potential candidates that regulate circRNA production without affecting host linear mRNAs.

The human ON-TARGETplus (OTP) siRNA libraries were reverse transfected into HeLa cells that stably express Dox-inducible *circmCherry* reporter in multiple 384-well plates. A smart pool combined four different OTP siRNAs that target different regions of the same mRNA to ensure effective silencing of the intended target. Forty-eight hours (hr) after the siRNA transfection, Dox was added to induce the expression of EGFP and *mCherry*. Fluorescence signals of EGFP and *mCherry* were monitored 24 hr post the Dox treatment to ensure that the fluorescence change was due to the Dox-inducible “nascent” *mCherry* and EGFP produced within these 24 hr. Of note, such a time window is important because circRNA is highly stable (Zhang et al., 2016b), and a prolonged Dox treatment may lead to the reduced likelihood to uncover RBPs that control nascent circRNA generation. The whole-genome screening was carried out independently in duplicates. The distribution of fluorescence densities of all screened genes in duplicates was highly correlated (Figure 1D; Table S1).

### Identification of Splicing Regulators and Antiviral Immune Factors in circRNA Biogenesis

We then set up an analysis pipeline (Figure 1E) to identify RBPs that may interfere with circRNA biogenesis. We excluded genes that were not expressed in HeLa cells and genes that also affected EGFP expression in the screening and only selected candidates that influenced the *mCherry* fluorescence with at least a 1.5-fold change (Figure 1E). We applied this relatively low threshold partially because of the fact that the efficiency of back-splicing and translation for a circRNA was expected to be lower than that for a linear RNA (Wang and Wang, 2015; Zhang et al., 2016b). In total, knockdown (KD) of 1,053 and 787 genes by human OTP siRNA libraries was found to reduce or enhance *mCherry* expression with the current cutoff, respectively (Figure 1E; Table S1). Although this screening has the potential to identify factors that affect the life cycle of *circmCherry* RNA and *mCherry* protein, we aimed to focus on potential RBP candidates in the current study, and 103 such proteins were uncovered by gene ontology (GO) with the database for annotation, visualization, and integrated discovery (DAVID) (Huang da et al., 2009).

(B) Verification of the circRNA expression reporter system. Left, expression of *circmCherry* and linear *egfp* RNAs in HeLa cells following 24 hr Dox treatment, revealed by RT-PCR. The primers for RT-qPCR used to detect RNAs are labeled in (A). Right, Sanger sequencing of RT-PCR products of *circmCherry* showing the correct back-splicing junction (BSJ) sequence of *circmCherry*.

(C) A schematic illustration of genome-wide siRNA screening associated with circRNA biogenesis. See main text for details.

(D) Two independent screenings were highly correlated. x axis, log<sub>2</sub> expression of *mCherry* over control siRNA-transfected cells; y axis, the density of OTP siRNA library targets.

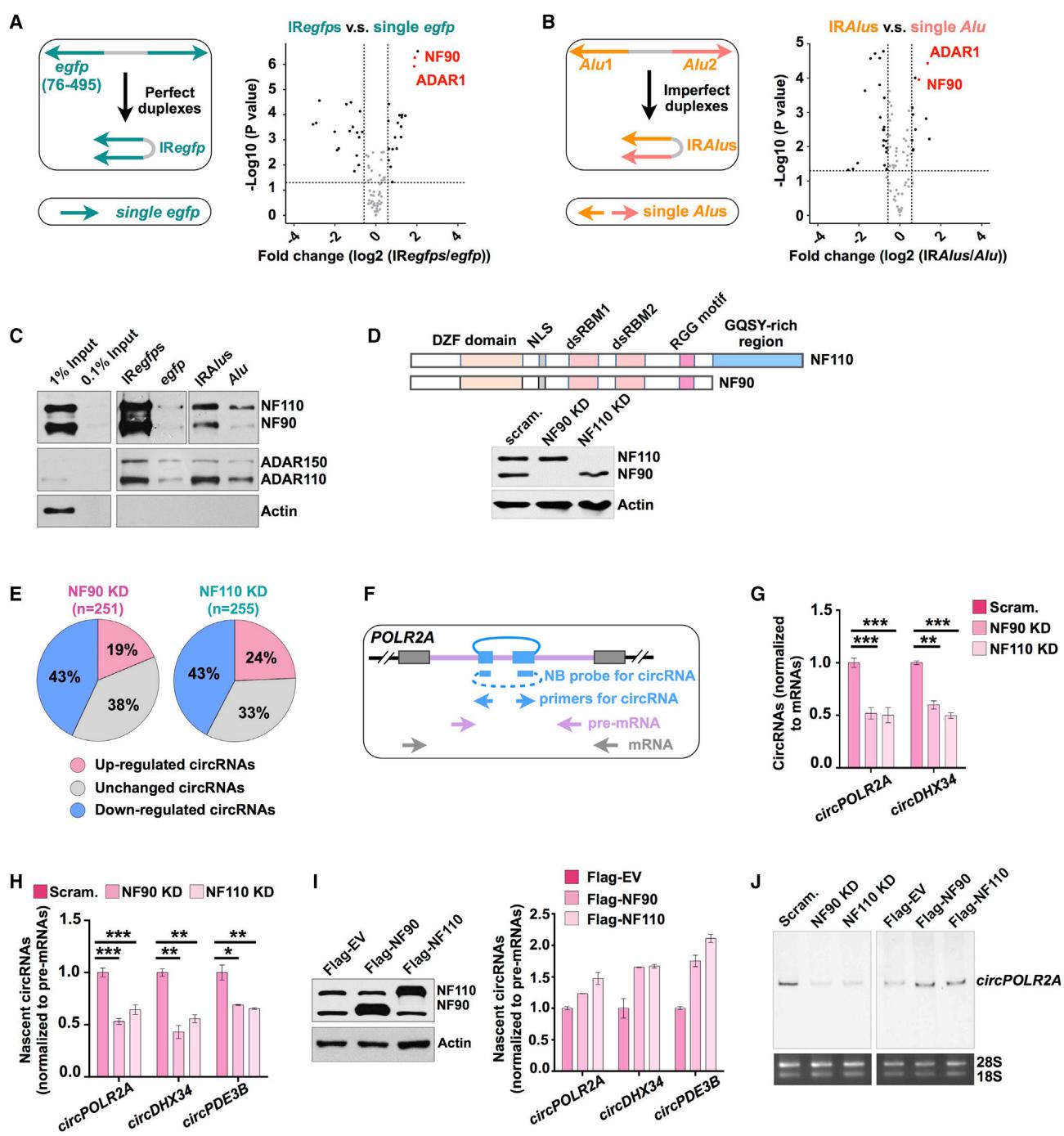
(E) A schematic drawing of the analysis flow. 103 RBPs were identified using DAVID to be potential candidates associated with circRNA biogenesis.

(F) Average distribution of fluorescence densities of *mCherry* from duplicates. Black and red dots represent 6,876 screened genes shown in (E). Cutoff values shown as orange lines indicate 1.5- or 0.67-fold change of fluorescence densities of *mCherry*. Viral and host immunity-related genes were labeled as red dots. y axis, log<sub>2</sub> expression of *mCherry* relative to the OTP non-targeting control (negative control) (GE Dharmacon); x axis, index of library pools.

(G) Analysis of RBPs identified in (E).

(H–J) Validation of RBP candidates by shRNAs. Knockdown (KD) of RIG-I (H), TLR3 (I), and ILF3 (J) in HeLa cells, followed by the examination of *circmCherry* abundance by RT-qPCRs. Left, the expression of candidate mRNAs was detected after KD. Right, the expression of *circmCherry* RNAs (normalized to linear *egfp* mRNAs) was reduced after KD following by 24 hr Dox treatment.

See also Figure S1; Table S1.



**Figure 2.** NF90 and NF110 Are Associated with circRNA Biogenesis

(A) Identification of RBPs interacting with perfect dsRNA duplexes *in vitro*. Left, a schematic drawing of the biotin-labeled inverted repeated *egfp* fragments (IRegfps, perfect duplexes) and single *egfp* fragment made from *in vitro* transcription (IVT). Right, identification of IRegfps-binding proteins *in vitro*. Nuclear extracts of PA1 cells were incubated with individual biotin-IRegfps or biotin-*egfp*, followed by anti-biotin pull-down. Proteins enriched in the biotin-IRegfps, but not biotin-*egfp* controls, revealed by mass spectrometry (MS) were further analyzed. NF90 and ADAR1 (red dots) were top two proteins identified to interact with IRegfps.

(B) Identification of RBPs interacting with imperfect dsRNA duplexes *in vitro*. Left, a schematic drawing of the biotin-labeled inverted repeated *Alu* elements (IRAlus, imperfect duplexes) or single *Alu* made from IVT. Right, identification of IRAlus-binding proteins *in vitro*. See (A) for details. NF90 and ADAR1 (red dots) were top two proteins identified to interact with IRAlus.

(C) Interaction of NF90/NF110 with RNAs. The biotinylated RNA pull-down assays in (A) and (B) followed by western blotting (WB) with antibodies for anti-NF90/NF110 and anti-ADAR1 (both its p110 and p150 isoforms) serves as positive controls for dsRNA-pull-down assays.

(legend continued on next page)

Further analyses revealed that 30 from 103 RBPs have reported roles in splicing regulation, 16 are involved in immune responses, 11 have helicase activities, and 7 exhibit nuclease activities (Figures 1F, 1G and S1H). These factors might influence the back-splicing efficiency, circRNA stability, export, and circRNA translation of this reporter, and then subsequently affected mCherry expression (Figure 1A). Random selection of several target RBPs that have reported roles in splicing regulation (Figure S1I) and innate immune responses (Figures 1H–1J and S1J) by shRNA treatment further verified outcomes from genome-wide screening. In all validations, we found reduced *circmCherry* expression, after normalized to linear *egfp* mRNAs (Figures 1H–1J, S1I, and S1J), implying that these factors are more likely to affect circRNA processing.

### NF90/NF110 Affect circRNA Formation Genome-wide

Indeed, additional *trans*-factors must participate in circRNA biogenesis because the expression of circRNA in the same locus varies among different cell lines and tissues (Dong et al., 2016; Salzman et al., 2013; Zhang et al., 2016a). Splicing regulators were anticipated to be one group of factors involved in back-splicing (Figures 1G and S1H), as suggested by recent studies (Ashwal-Fluss et al., 2014; Conn et al., 2015). Surprisingly, a group of immune response factors (Figures 1F and 1G) also appeared to affect circRNA biogenesis (Figures 1H–1J and S1J). Among them, ILF3 was particularly interesting to us. NF90/NF110 produced from the *ILF3* gene (Patiño et al., 2015) were reported to be required for efficient host immune defense (Hara-shima et al., 2010; Isken et al., 2003; Pfeifer et al., 2008). In our studies, NF90 was also identified from an independent set of experiments that aimed to seek double-stranded (ds) RNA-associating proteins that may directly participate in circRNA formation (Figures 2A and 2B, Table S2).

The biotin-labeled RNA pull-down assays revealed that NF90 and human ADAR1 could be repeatedly precipitated by dsRNA duplexes (dsRNAs) formed from either inverted repeated *egfp* fragments (IRegfps) (Figure 2A) or IRAIus (Figure 2B), but not single-stranded RNAs (ssRNAs), as revealed by mass spectrometry (MS). The dsRNA-binding protein ADAR1 (Chen et al., 2015) was reported to suppress circRNA production (Ivanov et al., 2015; Rybak-Wolf et al., 2015). NF90/NF110 were known to have

dsRNA-binding activity as well (Urcuqui-Inchima et al., 2006). Western blotting (WB) further confirmed MS results. NF90, NF110, and ADAR1 all strongly bound to ds IRegfps and IRAIus (Figure 2C). Interestingly, compared to NF90, NF110 also appeared to bind single *Alu* sequence (Figure 2C). These results indicated that NF90/110 could directly regulate back-splicing by associating with transient dsRNAs formed across circle-forming exons.

We then examined the effect of NF90 and NF110 on circRNA formation in cells. The polyadenylated (poly(A)<sup>+</sup>) RNAs collected from scramble, NF90, or NF110 shRNA-treated stable KD HeLa cells (Figure 2D) followed by RNA-seq revealed that the expression of the great majority of genes remained unchanged upon NF90 or NF110 depletion (Figures S2A and S2B; Table S3). The non-polyadenylated and ribosomal-RNA-depleted RNAs (poly(A)-/Ribo-) were also collected from the same cells and were subjected to deep sequencing (Figure S2C). After it was normalized to its linear isoform, we found that about 250 high-confidently expressed circRNAs (with reads spanning back-splicing junction [BSJ]  $\geq 5$ ) in NF90 or NF110 stable KD cells (Figure S2D). The expression of 43% of high-confidence circRNAs was reduced, of about 30% was not changed, and of about 20% was increased, upon NF90 or NF110 depletion (Figure 2E; Table S4). Importantly, about 85% (~213/251 or 255) of these high-confidence circRNAs could be detected in both NF90 or NF110 stable KD cells (Figure S2E), and 50% of these overlapped circRNAs (213) showed reduced expression in each KD (Figure S2F). Furthermore, among these downregulated circRNAs, 70% were detected in both samples (Figure S2G), suggesting that NF90 and NF110 can target the same circRNA-producing genes. Random selection of circRNAs by RT-qPCR further confirmed this genome-wide analysis (Figures 2F, 2G, and S3A).

It is known that the final mRNA levels are balanced between their production and degradation (Rabani et al., 2011). However, for many circRNAs, the detection of their expression at the steady-state level is largely due to accumulation (Zhang et al., 2016b). Thus, the nascent circRNA level is more accurate than the steady-state level to access the effect of *trans*-factors on circRNA formation. We then asked whether NF90/NF110 could affect nascent circRNA production. Indeed, it was the case. Stable KD of NF90 or NF110 reduced the level of newly back-spliced

(D) Generation of NF90 KD or NF110 KD stable cell lines. Top, a schematic illustration of NF90 and NF110. Bottom, WB confirms the stable KD of NF90 or NF110 by shRNAs targeting the unique 3' UTRs of NF90 and NF110 in HeLa cells, but not in scramble shRNA-treated HeLa cells.

(E) Genome-wide expression of circRNAs in NF90 KD or NF110 KD HeLa cells. A total of 251 or 255 high-confidence circRNAs (BSJ, reads  $\geq 5$ ) were identified by the poly(A)-/Ribo- RNA-seq in NF90 KD or NF110 KD cells, respectively. The expression of each circRNA was normalized to its linear RNA isoform under each condition.

(F) A schematic illustration of primer sets and the northern blotting (NB) probe to detect a circRNA, its linear isoform, or pre-mRNA in (G)–(J), exemplified by the circRNA-producing gene *POLR2A*.

(G) NF90 KD or NF110 KD reduced circRNA expression. The relative abundance of each circRNA at the steady state was normalized to its spliced mRNA isoform, as revealed by RT-qPCR. Primer sets were shown in (F).

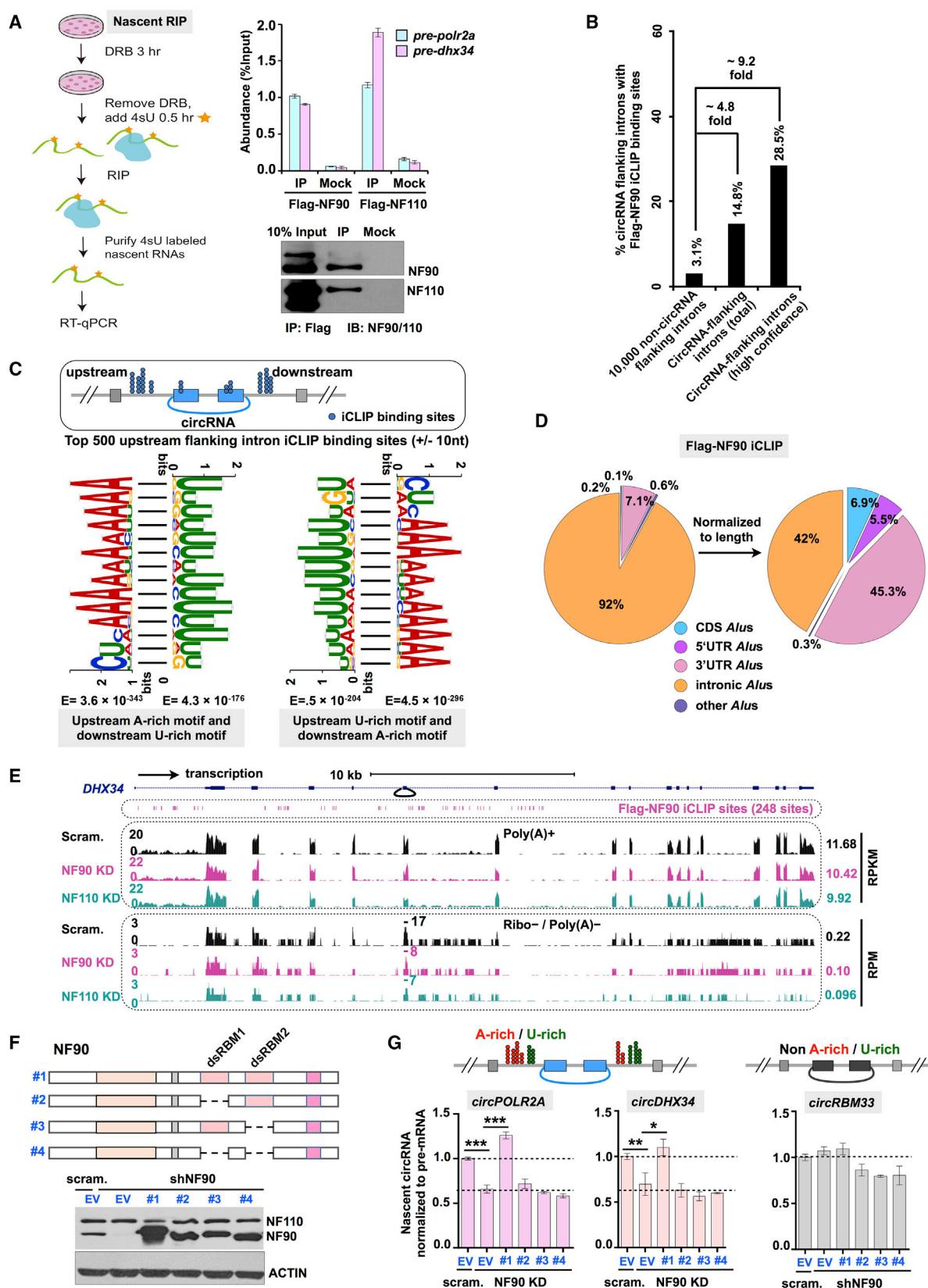
(H) NF90 KD or NF110 KD reduced nascent circRNA production. Nascent RNAs labeled with 4sU were purified from HeLa cells (Zhang et al., 2016b) and were subjected to RT-qPCR with primer sets that recognize the circRNA or pre-mRNA at each *POLR2A*, *DHX34*, and *PDE3B* gene locus shown in (F). The relative abundance of each circRNA was normalized to its pre-mRNA.

(I) Overexpression of NF90 or NF110 enhanced nascent circRNA formation (right). Left, WB of HeLa cells that stably expressed the empty vector (Flag-EV), Flag-NF90, or Flag-NF110.

(J) NB analyses confirmed that both NF90 and NF110 could promote endogenous circPOLR2A biogenesis.

In (G)–(H), error bars represent SD in triplicate experiments. \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ .

See also Figures S2 and S3; Tables S2, S3, and S4.



(legend on next page)

circRNAs (Figures 2H and S3B). Overexpression of Flag-tagged proteins (Flag-NF90 or Flag-NF110) also enhanced the production of nascent circRNAs, but not that of pre-mRNAs (Figures 2I and S3C). Such a positive regulation of NF90/NF110 on circRNA back-splicing was also confirmed by northern blotting (NB) with one abundant circRNA, *circPOLR2A* (Figure 2J).

Since both NF90 and NF110 likely promoted the same circRNA generation (Figures 2F–2J, S2G, and S3A–S3C), we then asked whether a double depletion of NF90/NF110 could further enhance this effect. Interestingly, we found that the stable double KD of NF90/NF110 cell lines could not be developed because such KD led to severe cell death (data not shown). Transient KD of both NF90/NF110 (Figure S3D) followed by the examination of steady-state (Figure S3E) and nascent (Figure S3F) levels of circRNA expression revealed a similar level of reduction when compared to those in NF90 or NF110 single KD cells (Figures 2G and 2H). Together, these results strongly suggested that NF90/NF110 could act as *trans*-factors to regulate nascent circRNA processing and that they exhibited a similar role in facilitating back-splicing.

### NF90/NF110 Promote circRNA Processing by Stabilizing Flanking Intronic RNA Pairs

It is well accepted that the orientation-opposite intronic complementary sequences flanking circularized exons promote circRNA biogenesis by forming RNA duplexes that juxtapose the splice sites (Chen, 2016). *In vitro* RNA precipitation assays have suggested that NF90/NF110 could strongly bind to dsRNA duplexes (Figures 2A and 2B). Within cells, nascent RNA immunoprecipitation assays (RIP) allowed to capture nascent pre-mRNAs associated with NF90/NF110 further revealed the interaction of NF90/NF110 with pre-mRNAs of the circRNA-forming

*POLR2A* and *DHX34* genes (Figure 3A). Of note, the formation of *circPOLR2A* was known to be dependent on flanking intronic sequences (Zhang et al., 2014), and *circDUX34* also contained such *cis*-elements (see below, Figure 3E). We then speculated that NF90/NF110 might directly bind to circRNA-flanking *cis*-complementary sequences and stabilize RNA pairs to stimulate exon circularization.

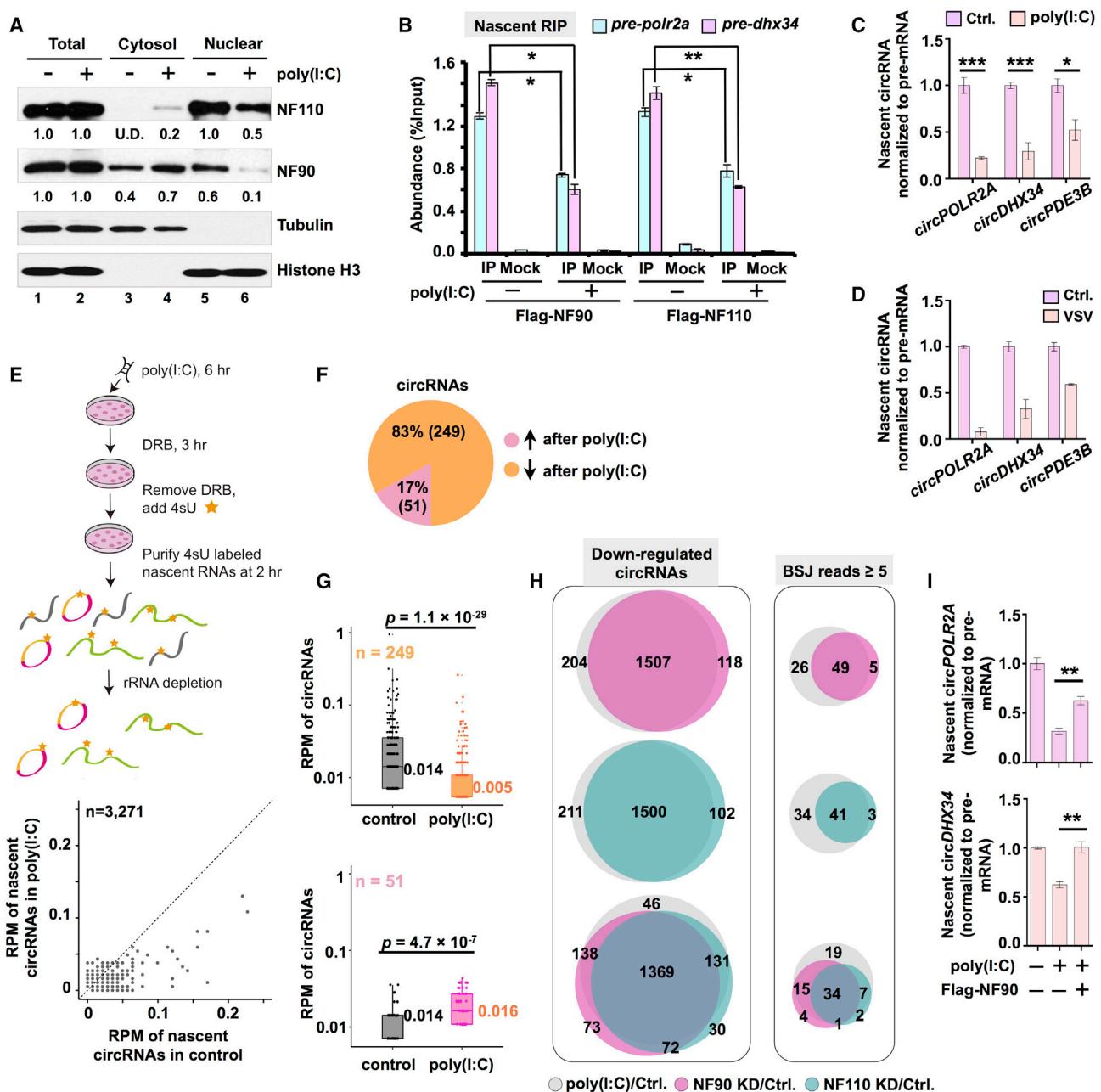
To test this idea, we characterized the *in vivo* binding pattern of NF90 at the single nucleotide resolution by performing the individual-nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP). The iCLIP was carried out with anti-Flag antibodies in Flag-NF90 stable expressing HeLa cells in duplicates (Figures S4A–S4C). It has been shown that NF90 is an AU-rich element (ARE)-binding protein that can bind to the ARE-containing 3' UTRs of mRNAs (Kuwano et al., 2010; Pei et al., 2008). The iCLIP confirmed the binding of NF90 to 3' UTRs of mRNAs with a preference to binding AREs (Figures S4D and S5A). Further analysis revealed that 6.6% of Flag-NF90 iCLIP reads were located in 3' UTRs and that 91.7% were in introns (Figure S5B). Although normalization to the total length of introns and 3' UTRs showed a higher density of Flag-NF90 binding to 3' UTRs (Figure S5B), the striking overall distribution of NF90-binding sites in introns suggests a previously underestimated role of NF90 in pre-mRNA processing.

Remarkably, NF90 selectively bound to flanking introns of circularized exons (Figure 3B). About 15% of the circRNA-forming exons were flanked by NF90 iCLIP-binding sites, and the number was up to 29% among highly expressed circRNAs. In contrast, only 3% of introns flanking randomly selected non-circRNA-forming exons contained NF90 iCLIP-binding sites (Figure 3B). Motif analyses revealed that Flag-NF90 preferred to bind clusters of A-rich or U-rich sequences (Figure 3C), most of which

### Figure 3. NF90 Promotes circRNA Formation via Its dsRNA-Binding Activity

- (A) Association of circRNA-producing pre-mRNAs with NF90 and NF110. Left, a schematic illustration of the nascent RNA immunoprecipitation assay (nascent RIP). Top right, nascent RIP was performed in HeLa cells stably expressed Flag-NF90 or Flag-NF110 using anti-Flag or anti-IgG antibodies followed by RT-qPCRs with primers recognizing pre-mRNAs. The percentage of RIP-enriched circRNAs relative to input was calculated under different conditions. Bottom right, WB showed the precipitation with each indicated antibody in either Flag-NF90- or Flag-NF110-expressing HeLa cells.
- (B) Flag-NF90 was enriched in flanking introns of circularized exons than control introns. Overlapped NF90-binding sites in two independent iCLIP-seq replicates were analyzed. The percentage of all circRNA-flanking introns with NF90 iCLIP-binding sites was 4.8-fold higher than that of control introns (compared 14.8% with 3.1%). The percentage of the high-confidence circRNA-flanking introns with NF90 iCLIP-binding sites was 9.2-fold higher than that of control introns (compared 28.5% with 3.1%).
- (C) Motif analysis of Flag-NF90 binding sites in circRNA-flanking introns. Motif (by MEME Suite v4.11.2) probability was computed from enriched 14-mers in circRNA flanking introns with Flag-NF90 iCLIP-binding sites. The context sequences (+/- 10 nt) from top 500 iCLIP-binding sites were used for the motif alignment. Two predominant motifs show A-rich (and U-rich) located in upstream flanking intron with a corresponding U-rich (and A-rich) motif within paired downstream flanking intron.
- (D) The genomic distribution of Flag-NF90 iCLIP-binding sites in *Alus*. The Flag-NF90 associated RNAs crosslink clusters was associated with intronic *Alus* (left); the genomic distribution of Flag-NF90 iCLIP-binding sites in *Alus* after normalization to length (right). Overlapped NF90-binding sites in two independent iCLIP-seq replicates were analyzed.
- (E) An example of circRNA produced from *DHX34* in HeLa cells. Top, Flag-NF90 iCLIP-binding sites in flanking introns of the single circularized exon (*circDUX34*) were shown as thin bars in magenta. Middle, wiggle-tracks of poly(A)+ RNA-seq showed that the expression of the linear *dux34* mRNA (each RPKM was labeled) remained largely unchanged upon NF90 KD or NF110 KD. Bottom, wiggle-tracks of poly(A)-/Ribo- RNA-seq showed that the expression of *circDUX34* reduced upon NF90 KD or NF110 KD. BSJ reads for *circDUX34* were marked underneath each circularized exon, and the RPM values were shown in different cell lines.
- (F) A schematic drawing of NF90 truncations used in (G).
- (G) The two dsRNA RBMs of NF90 are required for circRNA production. Top, a schematic illustration of a circRNA with (left, such as *circPOLR2A* or *circDUX34*) or without (right, such as *circRBM33*) A/U-rich flanking complementary sequences. Bottom, the decreased production of *circPOLR2A* and *circDUX34* in NF90 KD cells could be rescued by re-introduction of NF90 WT (#1) but not truncations (#2–#4), but *circRBM33* production was not affected by NF90 WT or mutants. Nascent RNAs were collected to evaluate the effect of NF90 and its truncations on newly produced circRNAs by back-splicing. Error bars represent SD in triplicate experiments. \*p ≤ 0.05, \*\*p ≤ 0.01, \*\*\*p ≤ 0.001.

See also Figures S4 and S5.



**Figure 4. Cytoplasmic Export of NF90/NF110 upon Viral Infection Correlates with Decreased circRNA Expression**

(A) NF90 and NF110 were translocated to the cytoplasm upon poly(I:C) treatment. The subcellular localizations of NF90 and NF110 were assessed by WB. The relative abundance of tubulin and histone H3 in cytoplasmic or nuclear fractions indicated a successful subcellular fractionation.

(B) NF90- or NF110-associated nascent pre-mRNAs were significantly decreased in cells upon poly(I:C) stimulation. HeLa cells stably expressing Flag-NF90 or Flag-NF110 were treated with poly(I:C) for 6 hr, followed by nascent RIP (Figure 3A). Bar plots represented the normalized nascent pre-mRNAs immunoprecipitated by anti-Flag over anti-IgG antibody, respectively, under each examined condition.

(C and D) CircRNA formation was rapidly reduced upon poly(I:C) stimulation (C) or VSV infection (D) in HeLa cells. Nascent RNAs were collected 6 hr of poly(I:C) transfection or 24 hr of VSV infection and quantitated with RT-qPCR by normalizing to each corresponding pre-mRNA. Primer sets were illustrated in Figure 2F.

(E) A genome-wide alteration of nascent circRNA production upon a rapid poly(I:C) treatment. Top, a schematic drawing to illustrate the collection of nascent RNAs upon poly(I:C) treatment. Bottom, nascent RNA-seq followed by the identification of nascent circRNAs (Zhang et al., 2016b) revealed a general reduction of nascent circRNAs upon poly(I:C) treatment.

(F) Reduction of nascent circRNA production upon poly(I:C) treatment. The production of over 83% high-confidence circRNAs was reduced upon poly(I:C) treatment.

(legend continued on next page)

were located to *Alus* in introns (Figure S5C) and could potentially form RNA pairs surrounding circRNA-forming exons. Consistent with this observation, counting the number of NF90-binding sites to 3' UTR *Alus* or intronic *Alus* revealed a dramatic enrichment of its distribution to intronic *Alus* (92% intronic *Alus* versus 7.1% 3' UTR *Alus*) (Figure 3D, left). Normalization to total lengths of NF90-binding *Alus* in different genomic regions revealed a comparably high NF90-binding intensity in introns and 3' UTRs (Figure 3D, right). One example of *circDHX34* was shown to illustrate our analysis (Figure 3E). Of note, while the expression of the DHX34 gene remained largely unchanged, *circDHX34* was reduced ~2-fold upon the depletion of NF90 or NF110.

Together, these results clearly indicated that the binding of NF90/NF110 to dsRNAs formed during nascent pre-mRNA processing could stabilize such transient RNA duplexes and promote back-splicing of a subset of circRNAs. In support of this view, we expressed the wild-type (WT) or a number of NF90 truncations lacking either one dsRNA-binding motif (dsRBM) or both dsRBMs individually in scramble or NF90 stable KD cell lines, followed by the examination of nascent circRNA formation. We found that re-introduction of only WT Flag-NF90, but not these NF90 truncations, into NF90 KD cells could rescue the nascent *circPOLR2A* and *circDHX34* expression (Figures 3F and 3G). In contrast, a circRNA (*circRBM33*), which does not contain A- or U-rich complementary motifs in its flanking introns, did not respond to NF90 KD or re-introduction of WT or NF90 truncations (Figure 3G, right panel).

Finally, in agreement with the view that NF90/NF110 have a similar role in regulating back-splicing for some circRNAs (Figures 2 and S3), re-introduction of WT Flag-NF110, but not its truncations, into NF90 KD cells rescued the expression of *circPOLR2A* and *circDHX34*, but not *circRBM33* (Figures S5D and S5E).

### CircRNA Expression Was Decreased upon Poly(I:C) Treatment, Corresponding to the Export of NF90/NF110 to the Cytoplasm

As a shuttling protein between the nucleus and the cytoplasm, NF90/NF110 are exported to the cytoplasm upon infection and become associated with viral transcripts to suppress viral replication (Harashima et al., 2010; Pfeifer et al., 2008). We observed that NF90/NF110 were rapidly exported to the cytoplasm upon 6 hr of poly(I:C) stimulation to mimic viral infection (Figures 4A and S6A). Nascent RNA RIP with anti-Flag antibodies performed in Flag-NF90 or Flag-NF110 stably expressing HeLa cells revealed a reduced association between the circRNA-forming

pre-mRNAs and Flag-NF90/NF110 after the poly(I:C) treatment (Figure 4B). Correspondingly, the reduced production of nascent circRNAs was observed upon 6 hr of poly(I:C) treatment (Figure 4C). Vesicular stomatitis virus (VSV) infection of cells also led to a similar decrease of nascent circRNA expression (Figure 4D), while their cognate mRNAs were largely unaffected (data not shown).

Genome-wide analyses of nascent circRNA expression (Figure S6B) showed that circRNA expression was decreased upon 6 hr of poly(I:C) treatment (Figure 4E). For instance, among the top 300 high-confidence nascent circRNAs identified in both samples before and after poly(I:C) stimulation, the production of over 80% was reduced (Figures 4F and 4G; Table S5), while only about 20% circRNAs slightly increased in their expression (Figures 4F and 4G; Table S5). The reduced expression of randomly selected nascent circRNAs could be verified by RT-qPCR (Figure S6C).

Remarkably, the downregulated circRNAs upon poly(I:C) treatment largely overlapped with the decreased circRNAs in response to NF90 KD or NF110 KD at the nascent level (Figure 4H; Table S6). For instance, the production of 1,625 nascent circRNAs was reduced in NF90 KD cells (the pie in magenta in Figure 4H; Table S6), and the expression of 1,711 nascent circRNAs was downregulated in cells treated with poly(I:C) (the pie in gray in Figure 4H; Table S5). Approximately 90% of these downregulated circRNAs at the nascent level were overlapped in these conditions (Figure 4H, upper panel). Together, these results indicate that the export of NF90/NF110 to the cytoplasm upon viral infection might reduce their ability to bind pre-mRNA in the nucleus and subsequently suppress the generation of a subset of circRNAs.

To confirm that this observed circRNA reduction upon poly(I:C) treatment is dependent on the nucleocytoplastic redistribution of NF90/NF110 (Figures 4A and S6A), we overexpressed NF90 in cells followed by poly(I:C) stimulation. As expected, the poly(I:C)-stimulation-caused circRNA reduction could be noticeably rescued by overexpression of NF90 for examined circRNAs (Figure 4I). Furthermore, it has been shown that the NF90/NF110 export depends on PKR activation and phosphorylation upon infection (Harashima et al., 2010; Shiina and Nakayama, 2014). Consistent with these studies, we found that NF90/NF110 were no longer exported to the cytoplasm upon poly(I:C) treatment (Figure S6A, middle and right panels) in PKR-depleted cells (Figure S6D). Although the PKR depletion led to circRNA downregulation (Figure S6E), importantly, upon poly(I:C) treatment, the level of the reduced circRNA expression

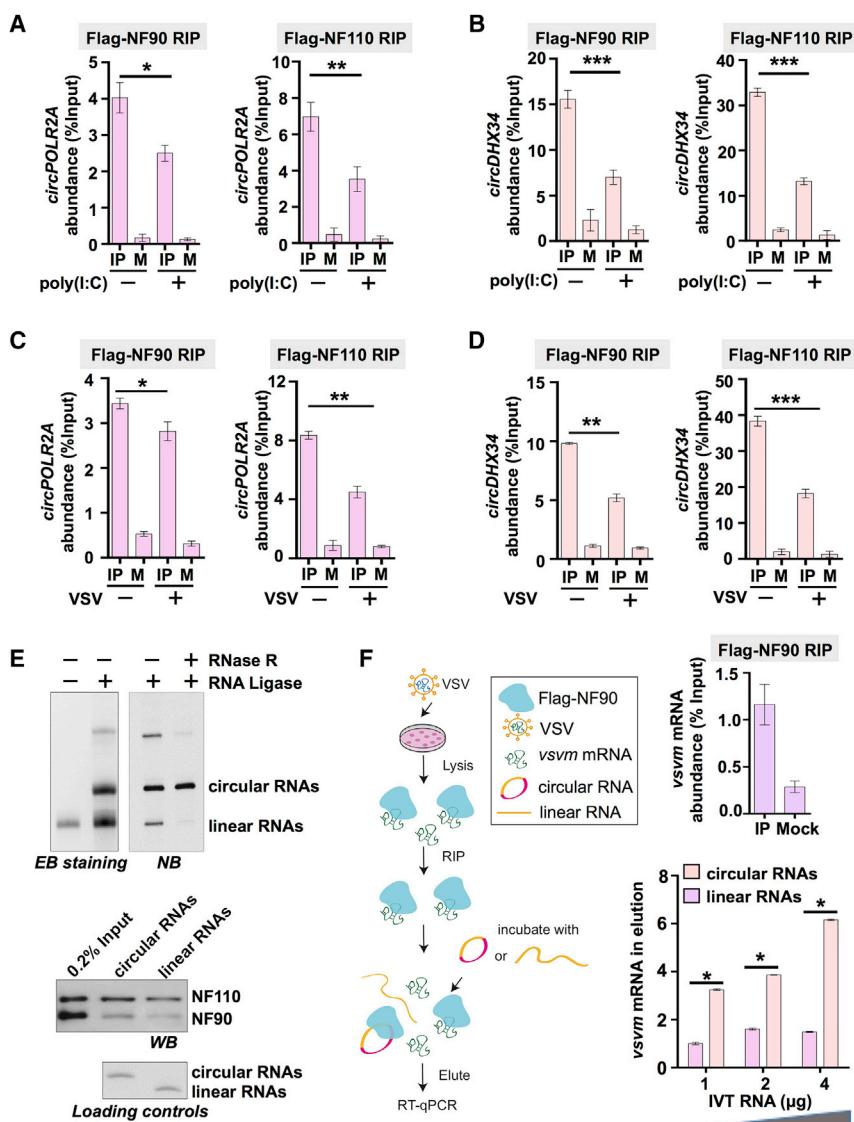
(G) The average expression levels of two groups of high-confidence circRNAs in control and poly(I:C)-stimulated cells shown in (F). p value, Wilcoxon rank-sum test.

(H) Genome-wide analyses revealed a strong correlation between the reduced circRNA formation upon poly(I:C) treatment and the decreased circRNA production in NF90- or NF110-depleted cells at the nascent level. The downregulated nascent circRNAs upon poly(I:C) treatment or NF90 or NF110 depletion were compared with one another. The high-confidence circRNAs were selected by BSJ reads  $\geq 5$  in at least one sample. The expression of each circRNA was normalized to its linear RNA isoform.

(I) Overexpression of NF90 could largely rescue the circRNA reduction in response to poly(I:C) stimulation. Nascent RNAs labeled with 4sU (Zhang et al., 2016b) were collected in control or NF90 overexpressed HeLa cells 6h post poly(I:C) transfection. Each nascent circRNA abundance was normalized to each corresponding pre-mRNA by RT-qPCR.

In (B), (C), and (I), error bars represent SD in triplicate experiments. \*p  $\leq 0.05$ , \*\*p  $\leq 0.01$ , \*\*\*p  $\leq 0.001$ .

See also Figure S6; Tables S5 and S6.



**Figure 5. Released NF90/NF110 from circRNP Complexes Bind to Viral mRNAs**

(A and B) NF90- and NF110-bound *circPOLR2A* (A) or *circDHX34* (B) was reduced in poly(I:C)-treated cells. HeLa cells stably expressing Flag-NF90 were treated with or without poly(I:C) followed by RIP using anti-Flag (IP) or anti-IgG (Mock, M) antibodies. The percentage of RIP-enriched circRNAs relative to input was calculated under different conditions.

(C and D) NF90- and NF110-bound *circPOLR2A* (C) or *circDHX34* (D) was reduced in VSV infected cells. See (A) for details.

(E) NF90 and NF110 preferred to bind to circRNAs but not linear RNAs with the same sequences in vitro. Top, an in-vitro-transcribed linear RNA (204 nt, from *circXPO1*) was circularized by RNA ligase. Linear and circularized RNA products were resolved on urea denaturing PAGE gel and visualized by ethidium bromide staining (top left). The in-vitro-ligated RNA circles were further enriched by RNase R digestion (as revealed by NB) following by purification (top right). Bottom, an equal amount of biotin-labeled linear or circRNAs was purified and subjected to RNA pull-down assays by incubation with HeLa cell extracts, followed by WB with antibodies to NF90 and NF110. Note that NF90 exhibited stronger preference than NF110 to interact with circular but not linear RNAs.

(F) Circular but not linear RNAs competed with vsvm mRNAs binding to NF90. Left, a schematic illustration of the circular or linear RNAs binding competition assay with vsvm mRNAs to NF90. HeLa cells with Flag-NF90 stable expression were infected by VSV; NF90 bound with vsvm mRNAs were purified by RIP with anti-Flag antibodies (top right), followed by incubation with the in-vitro-synthesized linear or circRNAs used in (E); and vsvm mRNAs dis-associated from Flag-NF90 were quantified by RT-qPCRs. Bottom right, increased amount of circRNAs, but not linear RNAs with the same sequences, competed with vsvm mRNAs for binding Flag-NF90.

In (A), (B), (C), (D), and (F), error bars represent SD in triplicate experiments. \*p ≤ 0.05, \*\*p ≤ 0.01, \*\*\*p ≤ 0.001.

See also Figure S7.

detected in PKR KD cells (which have the restricted NF90/NF110 nuclear export, Figure S6A) could be noticeably recovered (Figure S6F). Together, these studies suggest that NF90/NF110 play an important role in back-splicing regulation and that their nucleocytoplasmic re-distribution represents one mechanism for the reduced circRNA formation upon viral infection.

### NF90/NF110 Were Released from circRNP Complexes upon Viral Infection, Allowing Their Binding to Viral mRNAs

It has been shown that dsRNA-binding proteins, such as NF90/NF110 (Urcuqui-Inchima et al., 2006), could bind to structured RNAs in a sequence-independent manner (Larcher et al., 2004; Patiño et al., 2015; Ryter and Schultz, 1998). As circRNAs were hypothesized to be highly structured (Lasda and Parker, 2014), we speculated that in addition to facilitating circRNA formation in the nucleus (Figures 2, 3, and S2–S5), NF90/NF110 may also

interact with mature circRNAs. RIP assays revealed that both NF90 and NF110 were associated with circRNAs, such as *circPOLR2A* and *circDHX34* (Figures 5A–5D) in untreated cells, indicating that NF90/NF110 were components of some circRNP complexes. Intriguingly, the NF90/NF110-interacting circRNAs were noticeably reduced upon poly(I:C) treatment (Figures 5A and 5B) or VSV infection (Figures 5C and 5D). In the control, we found that the association of NF90/NF110 with linear mRNA isoforms was not altered before and after poly(I:C) treatment or VSV infection (Figures S7A and S7B).

*In vitro* circular or linear RNA pull-down assays revealed that NF90 preferred to bind circularized RNAs, compared to linear RNAs that have the same sequences (Figure 5E). NF110 also exhibited similar strength as NF90 for circRNA binding; however, it was also associated with linear RNA isoforms (Figure 5E). This finding was consistent with the earlier observation that both proteins strongly interacted with dsRNAs but only NF110 could bind

to ssRNAs *in vitro* (Figure 2C). These observations thus indicated that NF90/NF110 both interacted with circular and structured RNAs, but NF110 appeared to bind RNAs more promiscuously than NF90, probably due to the GQSY motif in the C-terminal region of NF110.

To further examine whether the strong binding capability of circRNAs to NF90 could compete with its interaction with viral mRNAs, we performed a competition assay to examine the binding capability between the *in-vitro*-transcribed (IVT) linear mRNAs or circRNAs to the partially purified Flag-NF90 from cells infected with VSV (Figure 5F, left panel). In these cells, Flag-NF90 could efficiently precipitate *vsvm* mRNAs (Figure 5F, top right panel). Incubation of the *in-vitro*-circularized RNAs strongly competed with the binding of *vsvm* mRNAs with Flag-NF90, as revealed by the fact that more soluble *vsvm* mRNAs could be readily detected with increased amount of circRNAs as competitors (Figure 5F, bottom right panel). In the control, we could not detect the competitive effect of the IVT linear RNAs on the binding with Flag-NF90 (Figures S7A and S7B). In addition, we performed a reciprocal competition assay by incubating *in-vitro*-transcribed *vsvm* or *egfp* mRNAs with Flag-NF90-bound endogenous circRNAs (Figure S7C, left panel). Consistently, we found that *vsvm* mRNAs, but not *egfp* mRNAs, competed with Flag-NF90-bound *circDHX34*, as revealed by the fact that more soluble *circDHX34* could be readily detected with increased amount of *vsvm* mRNAs as competitors (Figure S7C, top right). Also, we could not detect the competitive effect of *vsvm* mRNAs on the binding of linear *dhx34* mRNAs with Flag-NF90 (Figure S7C, bottom right).

### Overexpression of circRNA Competes for NF90/NF110 Binding with Viral mRNAs

Results shown in Figure 5 suggest that competition of circRNAs with viral mRNAs to NF90 binding was sequence independent. We reasoned that overexpression of any circRNAs would be able to interact with NF90/NF110, sequestering them from binding to viral mRNAs. To test this possibility, we transfected individual vectors into HeLa cells for either non-*circPOLR2A* or *circPOLR2A* overexpression (Figure 6A, #1 or #2, respectively) (Zhang et al., 2014). Of note, both vectors could produce high levels of *egfp* mRNAs (Figures 6A and 6B). Meanwhile, a reduced binding of NF90/NF110 to *vsvm* mRNA was observed only in *circPOLR2A*, but not non-*circPOLR2A*, overexpressed cells (Figure 6B, right panel).

Similar observations were found by taking advantage of the Dox-inducible *circmCherry* stable expression HeLa reporter cells (Figure 1A). Such cells were first treated with Dox to induce *circmCherry* expression followed by VSV infection (Figure 6C). Compared to cells that did not express *circmCherry*, Dox-treated cells exhibited high levels of *circmCherry* expression (Figure 6C). Meanwhile, *vsvm* mRNAs associated with NF90/NF110 were reduced (Figure 6C). Correspondingly, overexpression of different circRNAs could facilitate viral replication (Figure 6D). These results together support a model that the down-regulation of circRNAs upon viral infection and the reduced association of NF90/NF110 with circRNP complexes subsequently allow these proteins' binding to viral mRNAs for efficient host defense (Figure 6E).

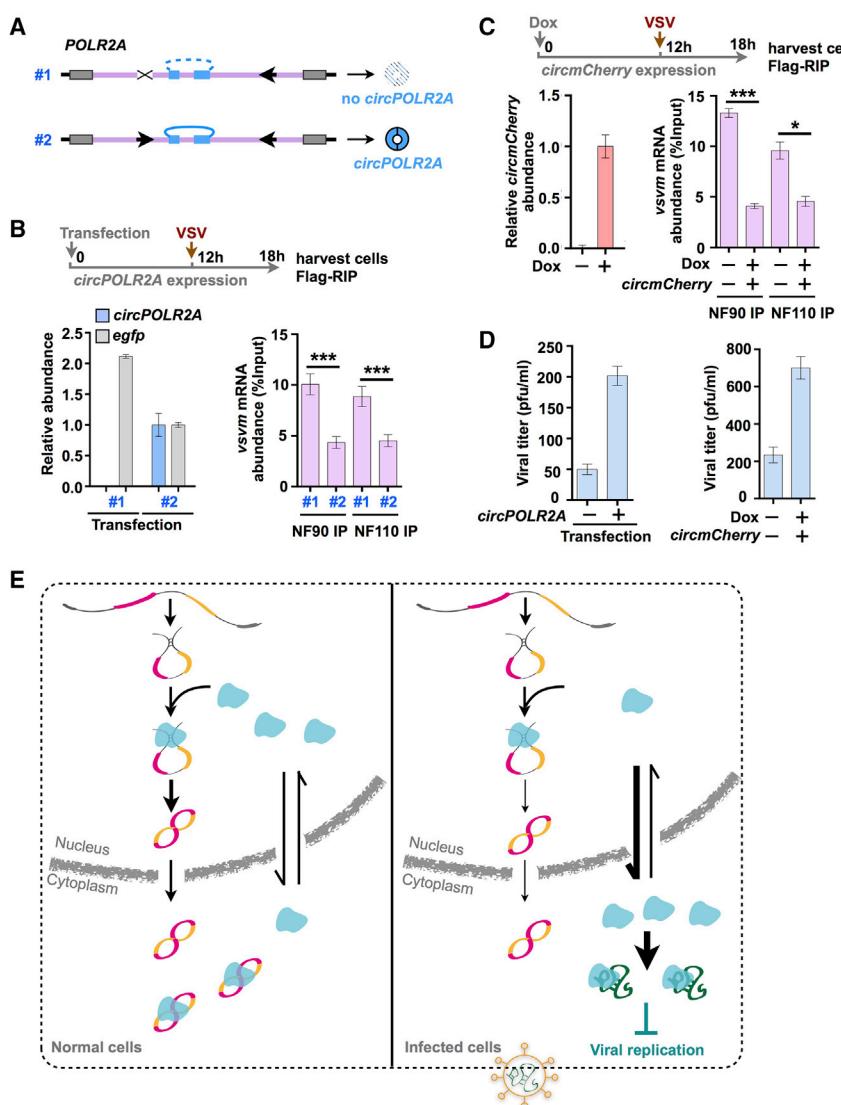
## DISCUSSION

Back-splicing is facilitated by the orientation-opposite complementary sequences residing in the flanking introns of the circularized exon(s) (Chen, 2016). Although having the same *cis*-elements, circRNA expression is in a cell-type-specific or tissue-specific manner (Dong et al., 2016; Salzman et al., 2013; Zhang et al., 2016a). In endogenous conditions, the number of potential RNA pairs formed across or within flanking introns, the distance between each potential pair of complementary sequences, and additional *trans*-factors associated with these *cis*-elements that stabilize or melt RNA pairs can affect circRNA generation.

Here, we applied a genome-wide siRNA library that targets all human genes to screen the potential *trans*-factors that affect back-splicing with an efficient circRNA expression reporter (Figure 1). As expected, we found that many splicing factors were involved in back-splicing regulation (Figures 1E and S1H). Conspicuously, factors related with host immunity were identified to affect circRNA production (Figures 1E–K and S1J). These factors include RIG-I (Figure 1I) that detects dsRNAs to trigger anti-viral cytokine production (Yoneyama and Fujita, 2008), TLR3 (Figure 1J) that recognizes negatively charged dsRNAs to activate immune responses during infection (Choe et al., 2005), multi-functional proteins ILF2 (Figure S1J) and ILF3 (Figure 1K) that are known to play important roles in antiviral immune responses (Harashima et al., 2010; Patiño et al., 2015; Pfeifer et al., 2008; Shi et al., 2007; Wen et al., 2014). Although how exactly these identified proteins could affect circRNA biogenesis remains largely unknown at the moment, they might directly or indirectly affect circRNA processing and even function. In the current study, we uncovered that NF90/NF110 produced from the ILF3 gene directly modulated back-splicing (Figures 2, 3, and S2–S5) and also coordinated with circRNA production in response to viral infection (Figures 4, 5, 6, and S6). Our data suggest that dsRNA-binding proteins NF90/NF110 (Figures 2A and 2B) strongly bind to and stabilize transient RNA pairs (Figure 3) formed between intronic complementary sequences, leading to the enhanced circRNA biogenesis in the nucleus (Figures 2F–2J and S3–S5). Such a stabilization of intronic RNA pairs by NF90/NF110 could be post-transcriptional, as most circRNA processing occurred post-transcriptionally (Zhang et al., 2016b).

Over ten thousand different circRNAs were reported in various organisms. Except for a few circRNAs, by far the function of the majority of circRNAs still remained to be explored (Chen, 2016). Here, we found that generation of many nascent circRNAs was rapidly downregulated upon poly(I:C) treatment or VSV infection (Figures 4C–4H). Such a reduction of newly synthesized circRNAs was correlated with the disrupted formation of nascent circRNAs in response to NF90/NF110 depletion (Figures 4H and 4I), suggesting that the nuclear export of NF90/NF110 upon infection (Figures 4A and S6A) (Larcher et al., 2004) contributed in part to the decreased circRNA production. Of note, the subcellular redistribution upon poly(I:C) treatment appeared to be different between NF90 and NF110 (Figure 4A). This raises a possibility that their different export may have variable effects on circRNA biogenesis during infection, which requires further investigation.

In addition to the involvement of NF90/NF110 in circRNA biogenesis, some circRNAs were found to be associated with



**Figure 6. Overexpression of circRNAs Competes for NF90/NF110 Binding with Viral mRNAs**

(A) A schematic drawing of vectors used in (B) for transfection to express *circPOLR2A*. Of note, both vectors produce linear spliced *egfp* mRNAs, but only #2 produces *circPOLR2A* via back-splicing enhanced by complementary sequences (black arrow) flanking *circPOLR2A* exons (Zhang et al., 2014).

(B) Overexpression of *circPOLR2A* impaired NF90/NF110 binding to viral mRNAs. Top, a schematic drawing to show the experimental procedure. Bottom left, RT-qPCR showed the expression of *circPOLR2A* and *egfp* mRNAs in transfected cells. Bottom right, NF90- or NF110-bound VSV mRNAs were reduced in *circPOLR2A* overexpressing cells. The percentage of Flag-NF90 or Flag-NF110 RIP-enriched VSV mRNAs relative to input was calculated.

(C) Dox-induced overexpression of *circmCherry* impaired NF90/NF110 binding to viral mRNAs. Top, a schematic drawing to show the experimental procedure. Bottom left, RT-qPCR showed the expression of *circmCherry* upon Dox induction. Bottom right, NF90- or NF110-bound VSV mRNAs were reduced in cells upon Dox induction for *circmCherry* overexpression. The percentage of Flag-NF90 or Flag-NF110 RIP-enriched VSV mRNAs relative to input was calculated.

(D) Overexpression of circRNAs facilitated viral replication. Viral replication was 3- to 4-fold higher in HeLa cells with overexpressed *circPOLR2A* (left) or induced *circmCherry* (right).

(E) A proposed model for the coordinated regulation of circRNA biogenesis and function by NF90/NF110 in viral infection. See text for details.

In (B) and (C), error bars represent SD in triplicate experiments. \*p ≤ 0.05, \*\*\*p ≤ 0.001.

NF90/NF110 (Figure 5). This observation indicates that the formation of NF90/NF110-circRNP accumulations in the cytoplasm may influence the host immune response. We found that circRNAs, but not their linear RNA isoforms, competed with viral mRNAs for binding to NF90/NF110 and that binding circRNAs to NF90/NF110 appeared to be in a sequence-independent manner (Figures 5 and 6). This finding is in agreement with the fact that many dsRNA-binding proteins, including NF90/NF110, bind to short or long dsRNA species independent of primary sequences (Chen et al., 2015; Larcher et al., 2004; Patiño et al., 2015; Ryter and Schultz, 1998). Hence, it is possible that circRNAs may also associate with other such proteins for functions. The topological structure of circRNAs is a mystery, and future studies are warranted to address how the unusually structured RNA circles may confer their strong binding to NF90/NF110 and other dsRNA-binding proteins.

Finally, as it is well known that circRNA expression is generally low (Chen, 2016), we speculate that many circRNAs, but not one specific circRNA, may act as a group in response to immune

response. In this scenario, circRNAs together can be used as a molecular reservoir of NF90/NF110 for prompt immune response upon viral infection; in the non-infected condition, the association of NF90/NF110 with circRNAs may keep them from non-specific immune responses in cells. Future work that aims to address additional immune-response-related factors (Figures 1 and S1) in circRNA biogenesis and ultimately function has a potential to uncover new roles of circRNA in immunity.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human Cell Lines

## ● METHOD DETAILS

- Cell culture and cell transfection
- Plasmid constructions and generation of stable cell lines
- Lentivirus production and cell infection
- RNA isolation, RT-PCR and RT-qPCR
- Genome-wide siRNA screening of human genes associated with circRNA biogenesis
- Functional annotation of RBPs
- Polyadenylated/Non-polyadenylated RNA separation, Ribosomal RNA depletion and RNA-seq
- Northern blotting
- Subcellular fractionation
- Immunofluorescence (IF)
- Metabolic labeling of nascent RNA with 4sU and nascent RNA purification
- Native and nascent RNA Immunoprecipitation (RIP)
- Individual-nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP)-seq
- In vitro RNA transcription, circularization and purification
- Cell lysate preparation and ds- or ss-RNA pull-down assay
- Cell lysate preparation and circular/linear RNA pull-down assay
- Circular/linear RNAs binding competition assay
- vsvm mRNA and circRNA binding competition assay
- Virus infection and viral titer determination

## ● QUANTIFICATION AND STATISTICAL ANALYSIS

- RNA-seq analyses
- iCLIP-seq analyses
- Statistical analyses

## ● DATA AND SOFTWARE AVAILABILITY

### SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2017.05.023>.

### AUTHOR CONTRIBUTIONS

Conceptualization, L.Y. and L.-L.C.; Methodology, X.L., C.-X.L., W.X., and L.-L.C.; Software, X.L., C.-X.L., W.X., and L.Y.; Validation, X.L. and C.-X.L.; Investigation, X.L., C.-X.L., Y.Z., S.J., Q.-F.Y., J.W., and R.-W.Y.; Data Curation, X.L., C.-X.L., W.X., L.Y., and L.-L.C.; Writing – Original Draft, X.L., C.-X.L., W.X., and L.-L.C.; Writing – Review & Editing, L.Y. and L.-L.C.; Visualization, X.L., C.-X.L., and L.-L.C.; Supervision, L.Y. and L.-L.C.; Funding Acquisition, L.Y. and L.-L.C.

### ACKNOWLEDGMENTS

We are grateful to B. Tian for reading the manuscript, F. Hou for preparing VSV viral particles, Z. Bai for initiating RNA circularization assay, and all lab members for discussion. This work was supported by CAS (XDB19020104), MOST (2016YFA0100701 and 2014CB964802), and NSFC (91440202 and 91540115).

Received: December 13, 2016

Revised: March 31, 2017

Accepted: May 19, 2017

Published: June 15, 2017

## REFERENCES

- Ashwal-Fluss, R., Meyer, M., Pamudurti, N.R., Ivanov, A., Bartok, O., Hanan, M., Evental, N., Memczak, S., Rajewsky, N., and Kadener, S. (2014). circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* 56, 55–66.
- Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36.
- Bolte, S., and Cordelières, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* 224, 213–232.
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steimetz, L.M., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406.
- Chen, L.L. (2016). The biogenesis and emerging roles of circular RNAs. *Nat. Rev. Mol. Cell Biol.* 17, 205–211.
- Chen, T., Xiang, J.F., Zhu, S., Chen, S., Yin, Q.F., Zhang, X.O., Zhang, J., Feng, H., Dong, R., Li, X.J., et al. (2015). ADAR1 is required for differentiation and neural induction by regulating microRNA processing in a catalytically independent manner. *Cell Res.* 25, 459–476.
- Choe, J., Kelker, M.S., and Wilson, I.A. (2005). Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science* 309, 581–585.
- Conn, S.J., Pillman, K.A., Touibia, J., Conn, V.M., Salmanidis, M., Phillips, C.A., Roslan, S., Schreiber, A.W., Gregory, P.A., and Goodall, G.J. (2015). The RNA binding protein quaking regulates formation of circRNAs. *Cell* 160, 1125–1134.
- Dong, R., Ma, X.K., Chen, L.L., and Yang, L. (2016). Increased complexity of circRNA expression during species evolution. *RNA Biol.* 1–11.
- Harashima, A., Guettouche, T., and Barber, G.N. (2010). Phosphorylation of the NFAR proteins by the dsRNA-dependent protein kinase PKR constitutes a novel mechanism of translational regulation and cellular defense. *Genes Dev.* 24, 2640–2653.
- He, C., Sidoli, S., Warneford-Thomson, R., Tatomer, D.C., Wilusz, J.E., Garcia, B.A., and Bonasio, R. (2016). High-resolution mapping of RNA-binding regions in the nuclear proteome of embryonic stem cells. *Mol. Cell* 64, 416–430.
- Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Isken, O., Grassmann, C.W., Sarisky, R.T., Kann, M., Zhang, S., Grosse, F., Kao, P.N., and Behrens, S.E. (2003). Members of the NF90/NFAR protein group are involved in the life cycle of a positive-strand RNA virus. *EMBO J.* 22, 5655–5665.
- Ivanov, A., Memczak, S., Wyler, E., Torti, F., Porath, H.T., Orejuela, M.R., Piechotta, M., Levanon, E.Y., Landthaler, M., Dieterich, C., and Rajewsky, N. (2015). Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* 10, 170–177.
- Kim, D., and Salzberg, S.L. (2011). TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome Biol.* 12, R72.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36.
- König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* 17, 909–915.
- Kramer, M.C., Liang, D., Tatomer, D.C., Gold, B., March, Z.M., Cherry, S., and Wilusz, J.E. (2015). Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev.* 29, 2168–2182.
- Kuwano, Y., Pullmann, R., Jr., Marasa, B.S., Abdelmohsen, K., Lee, E.K., Yang, X., Martindale, J.L., Zhan, M., and Gorospe, M. (2010). NF90 selectively represses the translation of target mRNAs bearing an AU-rich signature motif. *Nucleic Acids Res.* 38, 225–238.

- Larcher, J.C., Gasmi, L., Viranačken, W., Eddé, B., Bernard, R., Ginzburg, I., and Denoulet, P. (2004). If3 and NF90 associate with the axonal targeting element of Tau mRNA. *FASEB J.* **18**, 1761–1763.
- Lasda, E., and Parker, R. (2014). Circular RNAs: diversity of form and function. *RNA* **20**, 1829–1842.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079.
- Liang, D., and Wilusz, J.E. (2014). Short intronic repeat sequences facilitate circular RNA production. *Genes Dev.* **28**, 2233–2247.
- Moore, M.J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. *Science* **309**, 1514–1518.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods* **5**, 621–628.
- Patiño, C., Haenni, A.L., and Urcuqui-Inchima, S. (2015). NF90 isoforms, a new family of cellular proteins involved in viral replication? *Biochimie* **108**, 20–24.
- Pei, Y., Zhu, P., Dang, Y., Wu, J., Yang, X., Wan, B., Liu, J.O., Yi, Q., and Yu, L. (2008). Nuclear export of NF90 to stabilize IL-2 mRNA is mediated by AKT-dependent phosphorylation at Ser647 in response to CD28 costimulation. *J. Immunol.* **180**, 222–229.
- Pfeifer, I., Elsby, R., Fernandez, M., Faria, P.A., Nussenzveig, D.R., Lossos, I.S., Fontoura, B.M., Martin, W.D., and Barber, G.N. (2008). NFAR-1 and -2 modulate translation and are required for efficient host defense. *Proc. Natl. Acad. Sci. USA* **105**, 4173–4178.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.
- Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., Gnirke, A., Nusbaum, C., Hacohen, N., Friedman, N., et al. (2011). Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat. Biotechnol.* **29**, 436–442.
- Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R., et al. (2015). Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* **58**, 870–885.
- Ryter, J.M., and Schultz, S.C. (1998). Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.* **17**, 7505–7513.
- Salzman, J., Chen, R.E., Olsen, M.N., Wang, P.L., and Brown, P.O. (2013). Cell-type specific features of circular RNA expression. *PLoS Genet.* **9**, e1003777.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682.
- Shi, L., Qiu, D., Zhao, G., Corthesy, B., Lees-Miller, S., Reeves, W.H., and Kao, P.N. (2007). Dynamic binding of Ku80, Ku70 and NF90 to the IL-2 promoter in vivo in activated T-cells. *Nucleic Acids Res.* **35**, 2302–2310.
- Shi, Y., Yuan, B., Qi, N., Zhu, W., Su, J., Li, X., Qi, P., Zhang, D., and Hou, F. (2015). An autoinhibitory mechanism modulates MAVS activity in antiviral innate immune response. *Nat. Commun.* **6**, 7811.
- Shiina, N., and Nakayama, K. (2014). RNA granule assembly and disassembly modulated by nuclear factor associated with double-stranded RNA 2 and nuclear factor 45. *J. Biol. Chem.* **289**, 21163–21180.
- Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L.H., and Bindereif, A. (2015). Exon circularization requires canonical splice signals. *Cell Rep.* **10**, 103–111.
- (UniProt) The UniProt Consortium (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169.
- Urcuqui-Inchima, S., Castaño, M.E., Hernandez-Verdun, D., St-Laurent, G., 3rd, and Kumar, A. (2006). Nuclear Factor 90, a cellular dsRNA binding protein inhibits the HIV Rev-export function. *Retrovirology* **3**, 83.
- Wang, Y., and Wang, Z. (2015). Efficient backsplicing produces translatable circular mRNAs. *RNA* **21**, 172–179.
- Wen, X., Huang, X., Mok, B.W., Chen, Y., Zheng, M., Lau, S.Y., Wang, P., Song, W., Jin, D.Y., Yuen, K.Y., and Chen, H. (2014). NF90 exerts antiviral activity through regulation of PKR phosphorylation and stress granules in infected cells. *J. Immunol.* **192**, 3753–3764.
- Wu, H., Yin, Q.F., Luo, Z., Yao, R.W., Zheng, C.C., Zhang, J., Xiang, J.F., Yang, L., and Chen, L.L. (2016). Unusual processing generates SPA LncRNAs that sequester multiple RNA binding proteins. *Mol. Cell* **64**, 534–548.
- Yang, L. (2015). Splicing noncoding RNAs from the inside out. *Wiley Interdiscip. Rev. RNA* **6**, 651–660.
- Yang, L., Duff, M.O., Graveley, B.R., Carmichael, G.G., and Chen, L.L. (2011). Genomewide characterization of non-polyadenylated RNAs. *Genome Biol.* **12**, R16.
- Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.D., and Gage, F.H. (2009). An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat. Struct. Mol. Biol.* **16**, 130–137.
- Yin, Q.F., Yang, L., Zhang, Y., Xiang, J.F., Wu, Y.W., Carmichael, G.G., and Chen, L.L. (2012). Long noncoding RNAs with snoRNA ends. *Mol. Cell* **48**, 219–230.
- Yin, Q.F., Chen, L.L., and Yang, L. (2015). Fractionation of non-polyadenylated and ribosomal-free RNAs from mammalian cells. *Methods Mol. Biol.* **1206**, 69–80.
- Yoneyama, M., and Fujita, T. (2008). Structural mechanism of RNA recognition by the RIG-I-like receptors. *Immunity* **29**, 178–181.
- You, X., Vlatkovic, I., Babic, A., Will, T., Epstein, I., Tushev, G., Akbalik, G., Wang, M., Glock, C., Quedenau, C., et al. (2015). Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat. Neurosci.* **18**, 603–610.
- Zhang, X.O., Wang, H.B., Zhang, Y., Lu, X., Chen, L.L., and Yang, L. (2014). Complementary sequence-mediated exon circularization. *Cell* **159**, 134–147.
- Zhang, X.O., Dong, R., Zhang, Y., Zhang, J.L., Luo, Z., Zhang, J., Chen, L.L., and Yang, L. (2016a). Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* **26**, 1277–1287.
- Zhang, Y., Xue, W., Li, X., Zhang, J., Chen, S., Zhang, J.L., Yang, L., and Chen, L.L. (2016b). The biogenesis of nascent circular RNAs. *Cell Rep.* **15**, 611–624.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE   | SOURCE             | IDENTIFIER                        |
|---|--------------------|-----------------------------------|
| Antibodies  |                    |                                   |
| Rabbit monoclonal anti-ILF3   | Abcam              | Cat#Ab92355; RRID: AB_2049804     |
| β-Actin   | Sigma              | Cat#A3854; RRID: AB_262011        |
| FLAG  | Sigma              | Cat#F1804; RRID: AB_262044        |
| FLAG  | Shanghai Genomics  | Cat#SG4110-16                     |
| Cy3-AffiniPure Goat anti-Rabbit IgG(H+L)                            | Jackson            | Cat#111-165-045; RRID: AB_2338003 |
| Anti-ADAR1  | Santa Cruz         | Cat#sc-19077; RRID: AB_2257912    |
| Anti-Histone H3   | Abcam              | Cat#Ab1791; RRID: AB_302613       |
| Anti-Tubulin  | Santa Cruz         | Cat#sc-23948; RRID: AB_628410     |
| Bacterial and Virus Strains   |                    |                                   |
| VSV-GFP   | (Shi et al., 2015) | N/A                               |
| Chemicals, Peptides, and Recombinant Proteins                       |                    |                                   |
| Poly(I:C)   | Sigma              | Cat#P9582                         |
| 4-thiouridine (4sU)   | Sigma              | Cat# T4509                        |
| DRB   | Sigma              | Cat# D1916                        |
| doxycycline (Dox)   | Clontech           | Cat#631311                        |
| T4 Polynucleotide Kinase  | NEB                | Cat#M0201                         |
| CircLigase II   | Epicenter          | Cat#CL9021K                       |
| RNase R   | Epicenter          | Cat# MRNA092                      |
| T4 RNA Ligase I   | NEB                | Cat#M0204L                        |
| EZ-link biotin-HPDP   | Pierce             | Cat#21341                         |
| cComplete ULTRA Tablets, Mini, EASYPack Protease Inhibitor Cocktail | Roche              | Cat#000000005892970001            |
| Ribonucleoside Vanadyl Complex                                      | NEB                | Cat#S1402S                        |
| Dynabeads Protein G   | Invitrogen         | Cat#1003D                         |
| Dynabeads MyOne Streptavidin C1                                     | Invitrogen         | Cat#65001                         |
| X-tremeGENE 9   | Roche              | Cat#6366236001                    |
| Lipofectamine 2000 Reagent  | Thermo             | Cat#11668019                      |
| Hoechst 33342   | Sigma              | Cat#B2261-25MG                    |
| Glutathione Sepharose 4B  | GE healthcare      | Cat#17-0756-01                    |
| VECTASHIELD Antifade Mounting Medium                                | Vector Lab         | Cat#H-1000                        |
| ProLong Gold Antifade Mountant                                      | ThermoFisher       | Cat#P36930                        |
| MNase   | NEB                | Cat#M0247S                        |
| Glycerol  | ABCO               | Cat#G46055                        |
| HEPES   | ABCO               | Cat#H33755                        |
| TWEEN 20  | ABCO               | Cat#P87875                        |
| Triton X-100  | ABCO               | Cat#X10010                        |
| Agarose   | ABCO               | Cat#A47902                        |
| Bovine Serum Albumin  | ABCO               | Cat#A23088                        |
| Critical Commercial Assays  |                    |                                   |
| DNA-free kit  | Ambion             | Cat#AM1907                        |
| DIG Northern Starter Kit  | Roche              | Cat#12039672910                   |
| Mut Express MultiS Fast Mutagenesis Kit                             | Vazyme             | Cat#C213-01                       |
| Hieff Clone One Step Cloning Kit                                    | Yeasen             | Cat#10905ES25                     |
| 2 × T5 Super PCR Mix  | TSINGKE            | Cat#TSE005                        |

(Continued on next page)

**Continued**

| REAGENT or RESOURCE  | SOURCE                   | IDENTIFIER  |
|--|--------------------------|---|
| SuperScript III Reverse Transcriptase                                  | Invitrogen               | Cat#18080044  |
| RiboMAX Large Scale RNA Production System                              | Promega                  | Cat# P1300  |
| Dual-Glo Luciferase Assay System                                       | Promega                  | Cat#E2980   |
| Pierce Silver Stain for Mass Spectrometry kit                          | Thermo                   | Cat#24600   |
| NuPAGE Novex 4-12% Bis-Tris Gel  | Invitrogen               | Cat#NP0321BOX   |
| Human ON-TARGETplus (OTP) siRNA libraries                              | Dharmacon                | Cat#G-105005-05   |
| StarPrep Gel Extraction Kit StarPrep                                   | GenStar Co.LTD           | Cat#D205-04   |
| Deposited Data   |                          |   |
| Raw and analyzed data  | This paper               | GEO:GSE90874  |
| Mendeley data  | This paper               | <a href="http://dx.doi.org/10.17632/gkn6f863t9.1">http://dx.doi.org/10.17632/gkn6f863t9.1</a>                             |
| Experimental Models: Cell Lines  |                          |   |
| PA1  | ATCC                     | Cat#CRL-1572  |
| HeLa   | ATCC                     | Cat#CCL-2   |
| HEK293   | ATCC                     | Cat#CRL-1573  |
| Oligonucleotides   |                          |   |
| siRNA target sequences: circmCherry<br>siRNA1: GCTGAAGGTGACCAAGGGT     | This paper               | N/A   |
| siRNA target sequences: circmCherry<br>siRNA2: GCCAAGCTGAAGGTGACCA     | This paper               | N/A   |
| siRNA target sequences: circmCherry<br>siRNA3: CCAAGCTGAAGGTGACCAA     | This paper               | N/A   |
| siRNA target sequences: egfp-circmCherry<br>siRNA: GGCTACGTCCAGGTGACCA | This paper               | N/A   |
| Other primers and shRNA target sequences, see <a href="#">Table S7</a> | This paper               | N/A   |
| Recombinant DNA  |                          |   |
| P23-tight TRE promoter-1/2egfp-circmCherry-1/2 egfp                    | This paper               | N/A   |
| P23-Flag-NF90  | This paper               | N/A   |
| P23-Flag-NF110   | This paper               | N/A   |
| P23-Flag-NF90-dsRBM1-mut   | This paper               | N/A   |
| P23-Flag-NF90-dsRBM2-mut   | This paper               | N/A   |
| P23-Flag-NF90-dsRBM1+2-mut   | This paper               | N/A   |
| P23-Flag-NF110-dsRBM1-mut  | This paper               | N/A   |
| P23-Flag-NF110-dsRBM2-mut  | This paper               | N/A   |
| P23-Flag-NF110-dsRBM1+2-mut  | This paper               | N/A   |
| pZW1-circPOLR2A  | (Zhang et al., 2014)     | Addgene Plasmid # 73449   |
| pZW1-circPOLR2A-mut  | (Zhang et al., 2014)     | N/A   |
| Software and Algorithms  |                          |   |
| GraphPad Prism   | GraphPad Software        | <a href="http://www.graphpad.com/scientificsoftware/prism/">http://www.graphpad.com/scientificsoftware/prism/</a>         |
| TopHat v2.0.9  | (Kim et al., 2013)       | <a href="http://ccb.jhu.edu/software/tophat/index.shtml">http://ccb.jhu.edu/software/tophat/index.shtml</a>               |
| TopHat-Fusion v2.0.9   | (Kim and Salzberg, 2011) | <a href="http://ccb.jhu.edu/software/tophat/fusion_index.shtml">http://ccb.jhu.edu/software/tophat/fusion_index.shtml</a> |
| CIRCExplorer 1.1.10  | (Zhang et al., 2014)     | <a href="https://github.com/YangLab/CIRCExplorer">https://github.com/YangLab/CIRCExplorer</a>                             |
| DAVID Bioinformatics Resources (v6.7)                                  | (Huang da et al., 2009)  | <a href="https://david-d.ncifcrf.gov/">https://david-d.ncifcrf.gov/</a>   |
| UniProt  | (UniProt, 2017)          | <a href="http://www.uniprot.org/">http://www.uniprot.org/</a>   |

(Continued on next page)

**Continued**

| REAGENT or RESOURCE      | SOURCE   | IDENTIFIER  |
|--------------------------|--|---|
| MEME Suite v4.11.2       | (Bailey and Elkan, 1994)   | <a href="http://meme-suite.org/tools/meme">http://meme-suite.org/tools/meme</a>                   |
| R v.3.2.2                | <a href="https://www.r-project.org">https://www.r-project.org</a>          | <a href="https://www.r-project.org">https://www.r-project.org</a>                                 |
| bedtools v2.19.0         | (Quinlan and Hall, 2010)   | <a href="http://bedtools.readthedocs.io/en/latest/">http://bedtools.readthedocs.io/en/latest/</a> |
| Samtools Version: 0.1.18 | (Li et al., 2009)  | <a href="http://samtools.sourceforge.net/">http://samtools.sourceforge.net/</a>                   |
| RPKM                     | (Mortazavi et al., 2008)   | N/A   |
| RPM                      | (Zhang et al., 2014)   | N/A   |
| iCLIP analyses           | (Chen et al., 2015; König et al., 2010; Wu et al., 2016; Yeo et al., 2009) | N/A   |

**CONTACT FOR REAGENT AND RESOURCE SHARING**

Further information and requests for reagents may be directed to, and will be fulfilled by, the corresponding author, Ling-Ling Chen ([linglingchen@sibcb.ac.cn](mailto:linglingchen@sibcb.ac.cn)).

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

**Human Cell Lines**

Human cell lines including HeLa, HEK293 and PA1 cells were purchased from the American Type Culture Collection (ATCC; <https://www.atcc.org/>).

**METHOD DETAILS**

**Cell culture and cell transfection**

HEK293FT, HeLa and PA1 cells were cultured using standard protocols from ATCC. Plasmid or poly(I:C) transfection was carried out using X-treme GENE 9 (Roche) or Lipofectamine 2000 Reagent (Thermo) for HeLa and HEK293FT cells according to the manufacturer's protocols, with 70% ~80% transfection efficiency in general.

**Plasmid constructions and generation of stable cell lines**

A tet-on circRNA vector that expresses the back-spliced *circmCherry* RNA (Zhang et al., 2016b) was modified by placing a stop codon in front of IRES sequence, followed by a start codon (see Figure 1 for details). Thus, the mCherry protein could be only transcribed from the back-spliced *circmCherry*. The construct was then stably transfected into HeLa rtTA cells.

To knock down NF90 or NF110, target sequences for NF90, NF110 and a scramble sequence were individually cloned into pLKO.1-TRC vector between the Age I and EcoR I sites. HeLa cells were infected by lentiviral shRNAs to generate stable cell lines with NF90 or NF110 knockdown, respectively. The specificity of NF90 or NF110 knockdown was obtained by the specific shRNAs that target the unique 3'-UTR region of NF90 or NF110. NF90 and NF110 transient double knockdown was obtained by the shRNA that target the common CDS region of NF90 or NF110.

To overexpress NF90 or NF110, the N-terminal FLAG -fused NF90 or NF110 sequences were amplified with Phanta Super-Fidelity DNA Polymerase (Vazyme Biotech) from HeLa cells and cloned into the p23-phage vector between the NheI and Xhol sites. The plasmids were stably transfected into HeLa cells, respectively. HeLa cells stably transfected with the empty FLAG vector were used as controls. Primers for plasmid constructions were listed in Table S7. All constructs were confirmed by Sanger sequencing.

**Lentivirus production and cell infection**

To produce lentiviral particles,  $1.25 \times 10^7$  HEK293FT cells in a 15-cm dish were co-transfected with 20  $\mu$ g pLKO.1 shRNA construct, 15  $\mu$ g of psPAX2 and 10  $\mu$ g pMD2.G. The supernatant containing viral particles was harvested twice at 48 and 72 hr after transfection, and filtered through Millex-GP Filter Unit (0.22  $\mu$ m pore size, Millipore). Viral particles were then concentrated about 100-fold by sucrose gradient ultracentrifugation, resuspended in PBS containing 0.1% BSA, and stored at  $-80^{\circ}\text{C}$  until use. To infect HeLa cells with lentivirus, cells were incubated with culture medium containing 10  $\mu\text{L}$  concentrated lentivirus and 5  $\mu\text{g}/\text{ml}$  polybrene (Sigma) at  $37^{\circ}\text{C}$  for 1 hr. To increase the knockdown efficiency, infected cells were under several days of puromycin selection. Knockdown efficiency of proteins was evaluated by western blotting (WB).

**RNA isolation, RT-PCR and RT-qPCR**

Total RNAs from cultured cells were extracted with Trizol Reagent (Life technologies) according to the manufacturer's protocol. RNAs were further treated with DNase I (Ambion, DNA-free kit) and cDNAs were reverse transcribed with SuperScript III (Invitrogen) and

applied for PCR and/or qPCR analysis. *Actin* mRNA was examined as an internal control for normalization. The relative expression of each examined gene was determined with triplicate experiments. Primers for PCRs and qPCRs were listed in [Table S7](#).

### Genome-wide siRNA screening of human genes associated with circRNA biogenesis

Human ON-TARGETplus (OPT) siRNA library (GE Dharmacon) targeting 18,104 genes was used to screen potential genes associated with circRNA expression according to the manufacturer's protocol. All steps involved in multiple dispersing of samples were performed by Multidrop Combi Reagent Dispenser (Thermo scientific). Briefly, prior to cell transfection, RNAiMAX transfection reagent (0.1  $\mu$ L IMAX in each well) was added to the 384-well siRNA screening plates in which each well contains four siRNAs targeting one single gene. After 20 min, HeLa reporter cells were seeded at a density of 500 cells/well. Cells were cultured for 48 hr to knock down gene expression and then treated with 1  $\mu$ g/ml Dox for 24 hr to induce expression of mCherry and EGFP. After washed with PBS using LS405 microplate (BioTek), the fluorescence of mCherry and EGFP were measured, recorded and analyzed using Operetta (PerkinElmer). For accurate and unbiased screening, two independent experiments were performed and assessed. Analyses of two screening are listed in [Table S1](#).

### Functional annotation of RBPs

Potential RNA binding proteins (RBPs) from high throughput screening of HeLa reporter cells were identified by criteria listed in [Figure 1E](#) and DAVID Bioinformatics Resources (v6.7) ([Huang da et al., 2009](#)). Further classification of 103 RBPs was manually clustered with UniProt Knowledgebase ([UniProt, 2017](#)) (<http://www.uniprot.org/>).

### Polyadenylated/Non-polyadenylated RNA separation, Ribosomal RNA depletion and RNA-seq

Polyadenylated and non-polyadenylated RNA separation was carried out as described ([Yang et al., 2011; Yin et al., 2015](#)). Nascent RNA purification was carried out as described ([Zhang et al., 2016b](#)). Briefly, prior to construct non-polyadenylated or nascent RNA-seq library, rRNA was depleted as described ([Yang et al., 2011](#)). Ribosomal RNA removal was carried out using RiboMinus Human/Mouse Transcriptome Isolation Kit (Invitrogen) according to the manufacturer's protocol. RNA-seq libraries were prepared using Illumina TruSeq RNA Sample Prep Kit V2 and subjected to deep sequencing with Illumina HiSeq 2500 at CAS-MPG Partner Institute for Computational Biology Omics Core, Shanghai, China.

### Northern blotting

Northern blotting (NB) was performed according to the manufacturer's protocol (DIG Northern Starter Kit, Roche). Digoxigenin (Dig)-labeled antisense and sense riboprobes were made using RiboMAX Large Scale RNA Production Systems (Promega). In brief, 5  $\mu$ g total RNAs or 0.5  $\mu$ g in vitro synthesized linear or circular RNAs were resolved on denaturing urea polyacrylamide gel, transferred to nylon membrane (Roche) and UV-crosslinked using standard manufacturer's protocol. Membrane was then hybridized with specific Dig-labeled riboRNA probes. NB probes are listed in [Table S7](#).

### Subcellular fractionation

Nuclear and cytoplasmic fractionation was carried out as described ([Yin et al., 2012](#)) with slight modifications. Briefly,  $3 \times 10^6$  HeLa cells with or without Poly(I:C) treatment were washed with PBS, scraped into 15 ml-conical tubes and collected by centrifugation at 1,000 rpm for 2 min at room temperature. Cells were resuspended in 300  $\mu$ L lysis buffer (DPBS containing 1% NP-40, 0.5% sodium deoxycholate, 5 mM EDTA, 1 mM DTT, 1 mM PMSF, 2 mM VRC, 15% glycerol, 1X proteinase inhibitor cocktail), incubated on ice for 5 min and then examined under an optical microscope to ensure that membranes of most of cells were disrupted. Cells were centrifuged at 4,000 X g for 1 min at 4°C and the supernatant was saved as cytoplasmic fractionations. The precipitated nuclei were further washed once with lysis buffer, resuspended in 300  $\mu$ L lysis buffer, sonicated and centrifuged at 13,000 rpm for 10 min at 4°C. The supernatant was saved as nuclear extracts.

### Immunofluorescence (IF)

For all experiments, HeLa cells or HEK293 cells seeded on 18  $\times$  18 mm glass coverslips were fixed by 4% paraformaldehyde, permeabilized with 0.5% Triton X-100. For Immunofluorescence, anti-IIF3 (1:200) and anti-Flag (1:200) antibodies were used, and DAPI was used as a nuclear marker. Slides were mounted with VECTASHIELD Antifade Mounting Medium (Vector Lab) and imaged on a DeltaVision Elite imaging system (Applied Precision Imaging/GE Healthcare). Raw images were processed with softWoRx 6.0 and ImageJ/Fiji ([Schindelin et al., 2012](#)). Deconvolution images were generated by softWoRx 6.0 using the enhanced ratio method. Volume and distance measurements were performed using the “3D Object Counter” plugin for ImageJ/Fiji ([Bolte and Cordelières, 2006](#)).

### Metabolic labeling of nascent RNA with 4sU and nascent RNA purification

Metabolic labeling of newly synthesized RNAs was performed as described ([Zhang et al., 2016b](#)) with slight modifications. Cells with different treatments were incubated with 100  $\mu$ M DRB (Sigma, D1916) in the medium for 3 hr to block Pol II transcription. Transcription was reinitiated after removal of DRB and newly transcribed RNA was labeled with 200  $\mu$ M 4sU (Sigma, T4509). TRizol was added to cells to stop transcription and to purify total RNAs. 100–140  $\mu$ g 4sU-labeled RNAs were incubated with 0.2 mg/ml EZ-link biotin-HPDP (Pierce, 21341, dissolved in dimethylformamide (DMF, Sigma, D4551) at a concentration of 1 mg/ml) in biotinylation buffer

(10 mM Tris pH 7.4, 1 mM EDTA) for 1.5 hr at room temperature with rotation. Biotinylated RNAs were extracted twice by chloroform to remove unbound biotin-HPDP and precipitated using equal volume of isopropanol and 1:10 volume of 5 M NaCl. RNAs were precipitated at 13,000 rpm for 15 min at 4°C, washed twice with 75% ethanol and resuspended in 150 µL RNase-free water. RNA quality was evaluated by electrophoretical analysis to exclude RNA degradation. Biotinylated 4sU-labeled RNAs were incubated with 150 µL streptavidin beads (Invitrogen) for 20 min at room temperature. Beads were washed four times with 0.9 mL washing buffer (100 mM Tris pH 7.4, 10 mM EDTA, 1 M NaCl, 0.1% Tween 20, pre-warmed at 65°C), followed by four washes with 0.9 mL washing buffer (room temperature). Nascent RNAs were eluted twice with 100 µL 0.1 M dithiothreitol (DTT) and precipitated in 600 µL ice-cold ethanol.

#### Native and nascent RNA Immunoprecipitation (RIP)

Cells growing in 10 cm dishes were rinsed twice with ice-cold PBS, harvested in 10 mL ice-cold PBS and then centrifuged at 1,000 rpm for 5 min at 4°C. Cells were resuspended in 1 mL RIP buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 0.5% Igepal, 1 mM PMSF, 1X protease inhibitor cocktail (Roche) and 2 mM VRC) and subjected to three rounds of gentle sonication. Cell lysates were centrifuged at 12,000 rpm for 15 min at 4°C and the supernatants were precleared with 15 µL Dynabeads Protein G (Invitrogen) to get rid of non-specific binding. Then, the precleared lysates were used for IP with anti-Flag antibodies (Sigma). IP was carried out for 2 hr at 4°C. The beads were washed three times with high salt buffer and two times with the same RIP buffer, followed by extraction with elution buffer (100 mM Tris, pH 6.8, 4% SDS, and 10 mM EDTA) at room temperature for 10 min. One-third of the eluted sample was used for WB and the remaining was used for RNA extraction. The RNA enrichment was assessed by RT-qPCR. Primers are listed in [Table S7](#).

For Nascent RIP, cells were incubated with 100 µM DRB (Sigma, D1916) for 3 hr to block Pol II transcription. Transcription was reinitiated after removal of DRB and newly transcribed RNAs were labeled with 200 µM 4sU for 30 min (Sigma, T4509), as described above. The native RIP was performed with anti-Flag antibodies (Sigma), followed by the purification of 4sU-labeled RNAs from RIP products and RT-qPCR analyses with primers listed in [Table S7](#).

#### Individual-nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP)-seq

iCLIP-seq and data analyses were performed as described ([Chen et al., 2015; Wu et al., 2016](#)) with slight modifications. HeLa cells expressing Flag-tagged NF90 were UV-crosslinked, resuspended in lysis buffer and sonicated for 20 s x 6 times on 25% duty. Supernatants were incubated with anti-Flag antibodies (Shanghai Genomics) or control IgG-conjugated beads for 2 hr at 4°C. Beads with protein-RNA complexes were treated with 5 Kunz U/ml of micrococcal nuclease in the presence of 5 µg yeast tRNAs for 5 min at 37°C, followed by dephosphorylation of RNA 3' ends with PNK. Then pre-adenylated linker L3 (3' modification with biotin) was linked to 3' end of RNA using T4 RNA ligase 1 at 16°C overnight. 5' end of RNA was labeled with [ $\gamma$ -<sup>32</sup>P] ATP using PNK. Protein-RNA complexes were eluted with 1x NuPAGE loading buffer (Invitrogen) at 70°C for 10 min, resolved on 4%-12% NuPAGE Bis-Tris gel (Invitrogen), transferred to a nitrocellulose membrane and exposed to a Fuji film at -80°C. Bands corresponding to specific protein-RNA complexes were excised and then digested with proteinase K. RNAs were purified and cDNAs were made using Super-script III (Invitrogen) with Rclip primers. cDNAs were separated on 6% denaturing urea polyacrylamide gel, then three bands at 120-200 nt (high), 85-120 nt (medium) and 70-85 nt (low) ([Figure S3](#)) were excised, purified and circularized by Circligase II (Epicenter). Single-stranded cDNAs were hybridized with Cut-Oligo and then digested by BamH I. Linearized cDNAs were PCR amplified and purified. The obtained iCLIP libraries were subjected to high throughput deep sequencing.

#### In vitro RNA transcription, circularization and purification

Biotinylated RNAs were produced using RiboMax large RNA production system (Promega) according to the manufacturer's protocol with slight modifications. Briefly, 1 µg PCR-amplified T7-DNA fragments were incubated with 1 µL T7 RNA polymerase enzyme mix, 0.5 mM biotin-14-CTP, 1 mM GTP and 5 mM each of other NTPs. 5 mM GMP was supplemented in the reaction to produce 5'-monophosphate RNA that is required for subsequent RNA circularization. The reaction was carried out for 2 hr at 37°C, followed by DNase I treatment for 30 min at 37°C to remove DNA templates. Transcribed RNAs were precipitated with ethanol, washed with 70% ethanol and resuspended in RNase-free water. 50 µg linear RNAs were incubated with T4 RNA ligase 1 (NEB) in 500 µL reaction for overnight at 16°C according to the manufacturer's protocol. Circularized RNA products were treated with RNase R as described ([Zhang et al., 2014](#)), concentrated by ethanol precipitation, resolved on denaturing urea polyacrylamide gel and then visualized by Ethidium bromide staining. The corresponding to circular RNA band was excised and circular RNAs were purified.

#### Cell lysate preparation and ds- or ss-RNA pull-down assay

Biotinylated RNA pull-down was performed as described ([Wu et al., 2016](#)) with slight modifications. In brief, 4 µg biotinylated RNAs were heated for 5 min at 65°C in PA buffer (10 mM Tris HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 100 mM NH4Cl) and slowly cooled down to room temperature. To prepare cell lysate for pull-down with single-stranded (ss) RNA (*efgp* or *A/u*) or double-stranded (ds) inverted repeated RNAs (*IRefgfs* or *IRAlus*), 1 × 10<sup>7</sup> PA1 cells were resuspended in 2 mL PBS, 2 mL nuclear isolation buffer (1.28 M sucrose, 40 mM Tris-HCl pH 7.5, 20 mM MgCl<sub>2</sub>, 4% Triton X-100) and 6 mL DEPC-water, followed by incubation on ice for 15 min. Nuclei were pelleted by centrifugation at 1,000 rpm for 5 min at 4°C, and resuspended in 1 mL binding buffer (100 mM HEPES pH 7.0, 50 mM KCl, 10% glycerol, 1 mM EDTA, 1 mM DTT, 0.5% Triton X-100) supplemented with tRNA (0.1 mg/ml), heparin (0.5 mg/ml) and RNasin

(1unit/ $\mu$ L). The nuclei were sonicated and centrifuged at 13,000 rpm for 10 min at 4°C. The supernatant was pre-cleared with Streptavidin Dynabeads (Invitrogen) for 30 min at 4°C, followed by incubation with folded RNAs for 2 hr and with beads for 1 hr at 4°C. Beads were washed five times with the binding buffer, and boiled in 1 x sample buffer for 10 min. The retrieved proteins were subjected to NuPAGE 4%–12% Bis-Tris Gel (Invitrogen). Then the gel was stained using Pierce Silver Stain for Mass Spectrometry kit (Thermo Fisher Scientific) according to the manufacturer's instructions. After silver staining, specific bands were cut and sent to Core Facility of Molecular Biology (Institute of Biochemistry and Cell Biology, Shanghai, CAS) for mass spectrometry (MS) analysis. Identified proteins by MS are listed in [Table S2](#).

#### Cell lysate preparation and circular/linear RNA pull-down assay

Biotinylated RNA pull-down was performed as described ([Wu et al., 2016](#)). Briefly, 1  $\mu$ g biotinylated RNAs (purified by PAGE) were heated for 5 min at 65°C in PA buffer (10 mM Tris HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 100 mM NH4Cl) and slowly cooled down to room temperature. To prepare cell lysate for pull-down with linear or circular RNAs, 5  $\times$  10<sup>6</sup> HeLa cells were resuspended with 1 mL binding buffer (10mM HEPES pH7.0, 50Mm KCl, 10% glycerol, 1Mm EDTA, 1Mm DTT, 0.5% Triton X-100, heparin 0.3 mg/ml), sonicated and centrifuged at 13,000 rpm for 10 min at 4°C. The supernatant was pre-cleared with Streptavidin Dynabeads (Invitrogen) for 30 min at room temperature, followed by incubation with folded RNAs for 30 min and with beads for 15 min at room temperature. Beads were washed five times with the binding buffer, and boiled in 1 x sample buffer for 10 min. The retrieved proteins were subjected to WB.

#### Circular/linear RNAs binding competition assay

HeLa cells with Flag-NF90 stable expression were infected by VSV for 16 hr. NF90 bound VSVM mRNAs were precipitated by native RIP with anti-Flag antibodies (Sigma). The in vitro synthesized linear or circular RNAs were heated for 5 min at 65°C in PA buffer and slowly cooled down to room temperature. Then, different amounts of folded circular or linear RNAs were added to incubate with Flag-NF90-VSVM mRNAs complexes for 2 hr at 4°C in 0.2 mL RIP buffer. The supernatant was collected and used for RNA extraction with Trizol. The relative abundance of VSVM mRNA in the supernatant competed from the incubation with different amounts of circular or linear RNAs was assayed by RT-qPCRs.

#### vsVm mRNA and circRNA binding competition assay

HeLa cells with Flag-NF90 stable expression were used for RIP assay. NF90 bound RNAs (circRNAs or mRNAs) were precipitated by native RIP with anti-Flag antibodies (Sigma). The in vitro synthesized vsVm or egfp were heated for 5 min at 65°C in PA buffer and slowly cooled down to room temperature. Then different amounts of folded vsVm or egfp were incubated with Flag-NF90-RNA complexes for 2 hr at 4°C in 0.2 mL RIP buffer. The supernatant was collected and used for RNA extraction with Trizol. The relative abundance of circRNAs or mRNAs in the supernatant competed from the incubation with vsVm or egfp was assayed by RT-qPCRs.

#### Virus infection and viral titer determination

HeLa cells with overexpressed *circPOLR2A* ([Figure 6B](#)) or induced *circmCherry* ([Figure 6C](#)) were infected with VSV-GFP (m.o.i.: 0.01) for 12 hr. Then supernatant (with released virus) was collected for plaque assay in HEK293FT cells as described ([Pfeifer et al., 2008](#)) with slight modification. Briefly, 293FT cells were cultured in 6-well plates till spreading to nearly 100%, and replaced by FBS-free culture medium with 100  $\mu$ L supernatant (containing virus). After 3 hr incubation, culture medium was replaced with soft-agar mixture (1% agar), which is solidified at room temperature. After solidified, 293FT cells were cultured for 24 hr and colonies with GFP fluorescence were counted with fluorescent microscope. Virus yields were quantitated as PFU of released virus per ml of culture medium.

## QUANTIFICATION AND STATISTICAL ANALYSIS

#### RNA-seq analyses

For prediction of steady-state and nascent circRNAs: each steady-state and nascent RNA-seq dataset was first mapped using TopHat ([Kim et al., 2013](#)) (TopHat v2.0.9, parameters:–microexon-search -g 1 -a 6 -m 2) and aligned to GRCh37/hg19 human reference genome with the UCSC Genes annotation (Human: hg19 knownGene.txt updated at 2013/6/30). The unmapped reads were then realigned to hg19 human genome using TopHat-Fusion ([Kim and Salzberg, 2011](#)) (TopHat v2.0.9, parameters:–fusion-search–keep-fasta-order–bowtie1–no-coverage-search). The steady-state and nascent circRNAs were predicted using CIRCExplorer with existing gene annotations (Human hg19 knownGene.txt updated at 2013/6/30 and refFlat.txt updated at 2013/10/13) as described ([Zhang et al., 2014](#)).

For gene expression analysis: gene expression of mRNAs was determined by RPKM ([Mortazavi et al., 2008](#)) from poly(A)<sup>+</sup> RNA-seq. Refseq genes with RPKM  $\geq$  1 were selected for comparison (Human hg19 refFlat.txt updated at 2013/10/13). Fold change (FC) between NF90 (or NF110) KD and scramble control was used to determine upregulated (FC  $\geq$  2), unchanged (0.5 < FC < 2) and downregulated (FC  $\leq$  0.5) genes. Expression of Refseq genes in HeLa cells is listed in [Table S3](#).

For circRNA expression analysis: the expression of each circRNA was evaluated by RPM (mapped back-splicing Reads Per Million mapped reads) of reads that span circRNA back-splicing junction (BSJ) site ([Zhang et al., 2014](#)). Highly expressed (high-confidence) circRNAs were determined with reads spanning BSJs  $\geq$  5. FC between NF90 KD (or NF110 KD) and scramble was used to determine

upregulated ( $FC \geq 1.5$ ), unchanged ( $0.667 < FC < 1.5$ ) and downregulated ( $FC \leq 0.667$ ) genes at steady-state poly(A)–/Ribo– RNA-seq or nascent Ribo– 4sU-seq, respectively. The distribution of RPMs of circRNAs between NF90 KD (or NF110 KD) and scramble was plotted. Note that RPMs for NF90 KD (or NF110 KD) circRNAs were lower than those of scramble circRNAs. The high-confidence circRNAs at the steady-state and nascent levels in HeLa cells are listed in **Tables S5** and **S6**, respectively.

### iCLIP-seq analyses

We applied iCLIP-seq analysis pipeline according to previous publications with some modification (Chen et al., 2015; König et al., 2010; Wu et al., 2016; Yeo et al., 2009). In brief, iCLIP-seq reads were separated according to 4-nt experimental barcodes, followed by trimming relevant sequencing adapters. The 5-nt random barcodes were registered and removed before mapping to the human reference genome (GRCh37/hg19) by allowing two mismatch with TopHat (TopHat v2.0.9, parameters:—microexon-search -a 6 -m 2). Peak calling and the false discovery rate (FDR) calculating were determined as previously described (König et al., 2010).

To identify iCLIP binding sites, the genomic location of 5' end of each combined read was considered as the cross-link nucleotide, and the binding counts for each cross-link nucleotide were calculated. iCLIP reads associated with protein-coding genes were defined by the region from the annotated gene. Only sense iCLIP reads of associated genes were used for cross-link nucleotide analysis.

To determine the FDR for each cross-link nucleotide, binding counts within  $\pm 5$ nt of each crosslink nucleotide were merged to represent the height of this nucleotide. The background frequency was computed after randomly placing the same number of binding counts within the gene (wgEncodeGencodeCompV19) for 10,000 iterations. Modified FDR (Yeo et al., 2009) for each cross-link nucleotide was computed against the background frequency. For each gene locus, a threshold peak height as the smallest height equivalent to  $FDR < 0.05$  was defined as iCLIP binding sites.

### Statistical analyses

Statistical significance for comparisons of means was assessed by Student's t test for qRT-PCRs. Error bars represent SD in triplicate experiments.  $P$  values below 0.05 were marked by 1 asterisk, while 2 asterisks indicate a  $p$  value  $< 0.01$  and 3 asterisks a  $p$  value  $< 0.001$ . Statistically significant difference was assessed using Wilcoxon rank-sum test with R platform (R v.3.2.2), and statistical significance was set at  $p < 0.05$ . To evaluate the relevant correlations between two group datasets, Pearson correlation coefficient (PCC) was also performed with R platform (R v.3.2.2).

### DATA AND SOFTWARE AVAILABILITY

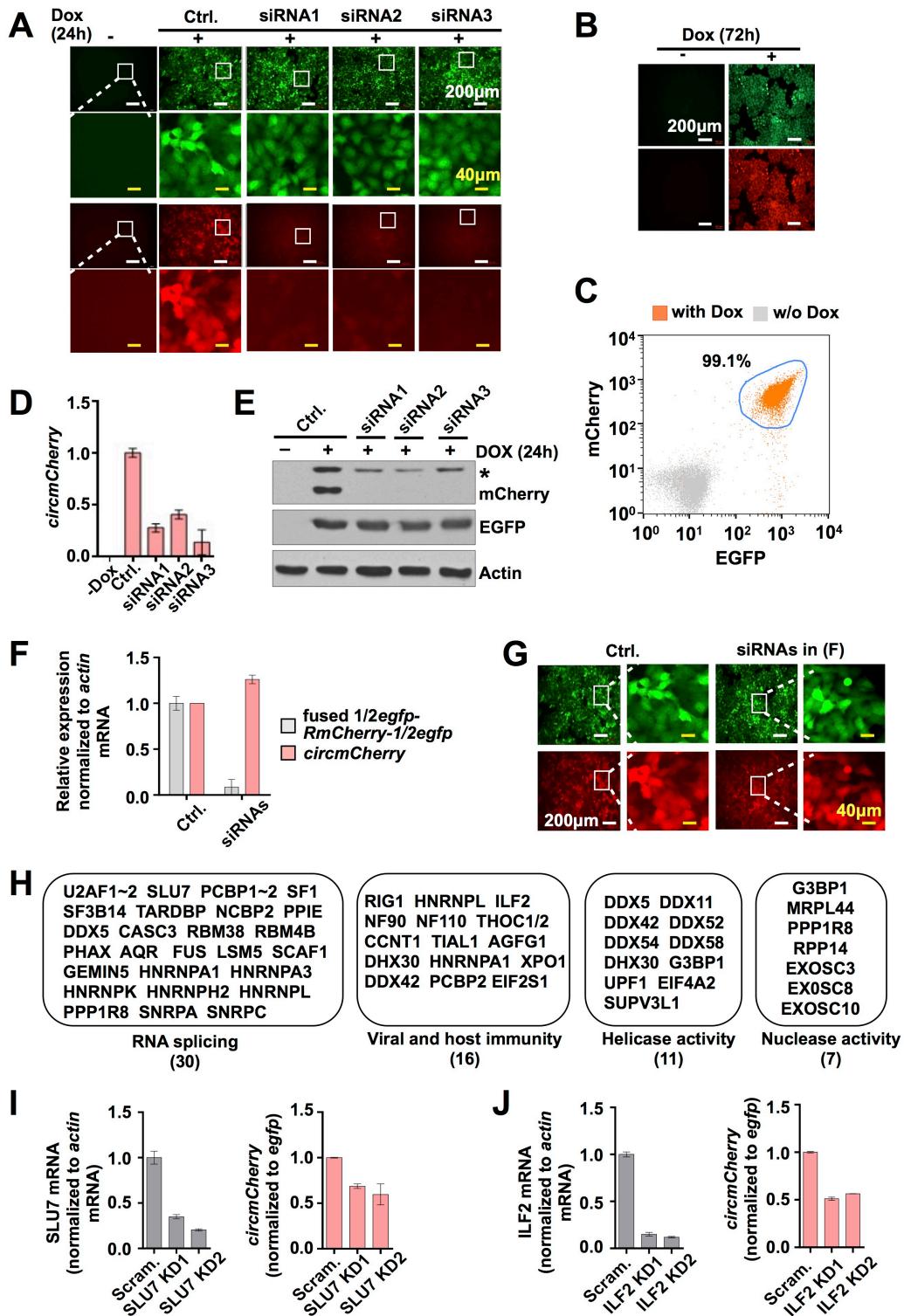
The accession number for the all sequencing data reported in this paper is GEO: GSE90874. Mendeley data have been deposited in the website: <http://dx.doi.org/10.17632/gkn6f863t9.1>.

**Supplemental Information**

**Coordinated circRNA Biogenesis and  
Function with NF90/NF110 in Viral Infection**

**Xiang Li, Chu-Xiao Liu, Wei Xue, Yang Zhang, Shan Jiang, Qing-Fei Yin, Jia Wei, Run-Wen  
Yao, Li Yang, and Ling-Ling Chen**

## Supplemental Figures and Figure Legends



**Figure S1. Verification and characterization of the Dox-inducible circmCherry-expression system for genome-wide siRNA screening (related to Figure 1)**

(A)(D)(E) Verification of the circRNA expression reporter system. After 24 h of Dox treatment, the fluorescence of mCherry or EGFP (A), *circmCherry* RNA (D) and proteins of mCherry and EGFP (E) all increased in HeLa cells stably transfected with the circRNA reporter (shown in Figure 1A). Further transfection of siRNAs targeting the back-spliced junction site of *circmCherry* (blue arc line in Figure 1A) efficiently decreased mCherry fluorescence (A), *circmCherry* RNA (D) and mCherry protein (E),

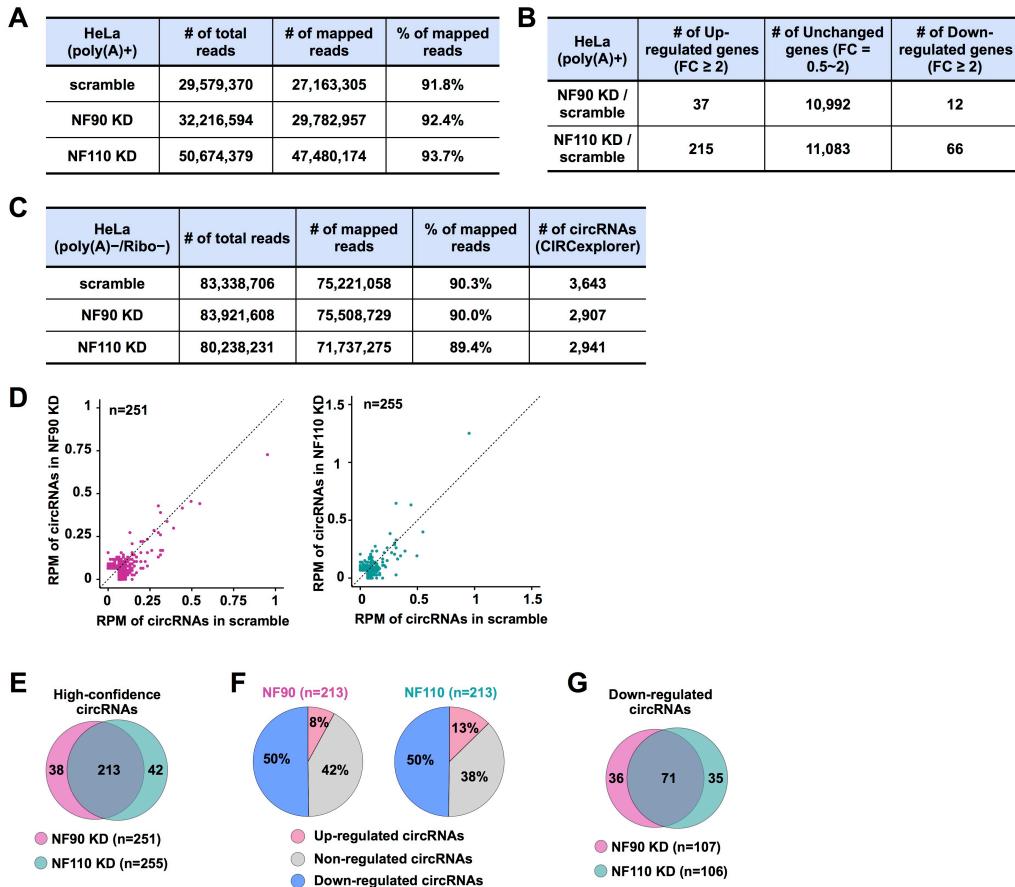
but not *egfp*/EGFP. The asterisk in (E) represents a fused protein produced from fused *1/2egfp-mCherry-1/2egfp* linear RNA (see below, S1F).

(B-C) Both mCherry and EGFP fluorescences are produced from the circRNA expression reporter system (shown in Figure 1A). The expression of both mCherry and EGFP (B) was dramatically increased in transfected HeLa cells after 72 hours of Dox treatment. FACS analysis revealed that the great majority of cells (~ 99%) expressed both mCherry and EGFP at similar fluorescence intensities (C).

(F)(G) The fused *1/2egfp-mCherry-1/2egfp* linear RNA, but not *circmCherry* RNA, was efficiently reduced by siRNAs (pink thick line in Figure 1A) targeting the junction site between the first *1/2 egfp* and the reverted *mCherry*, shown by RT-qPCR (F) and imaging (G).

(H) Lists of RBP candidates involved in RNA splicing, viral and host immunity, helicase activities and nuclease activities (RBP categories shown in Figure 1G).

(I)(J) Validation of RBP candidates by shRNAs. KD of SLU7 (I) and ILF2 (J) in HeLa cells, followed by examination of *circmCherry* abundance by RT-qPCRs. Left panels, expression of targeted genes was detected after KD. Right panels, expression of *circmCherry* RNAs (normalized to linear *egfp* mRNAs) was reduced after KD following by 24 hours of Dox treatment.



**Figure S2. Statistics of RNA-seq samples in NF90 or NF110 KD HeLa cells (related to Figure 2).**

(A) The numbers of total reads, mapped reads and mapped ratio of poly(A)+ RNA-seq in HeLa cells treated with scramble shRNA, or shRNAs targeting NF90 or NF110.

(B) The numbers of up-regulated, unchanged and down-regulated Refseq genes from the steady-state poly(A)+ RNA-seq in NF90 or NF110 stable KD HeLa cells, compared to scramble shRNA treated cells.

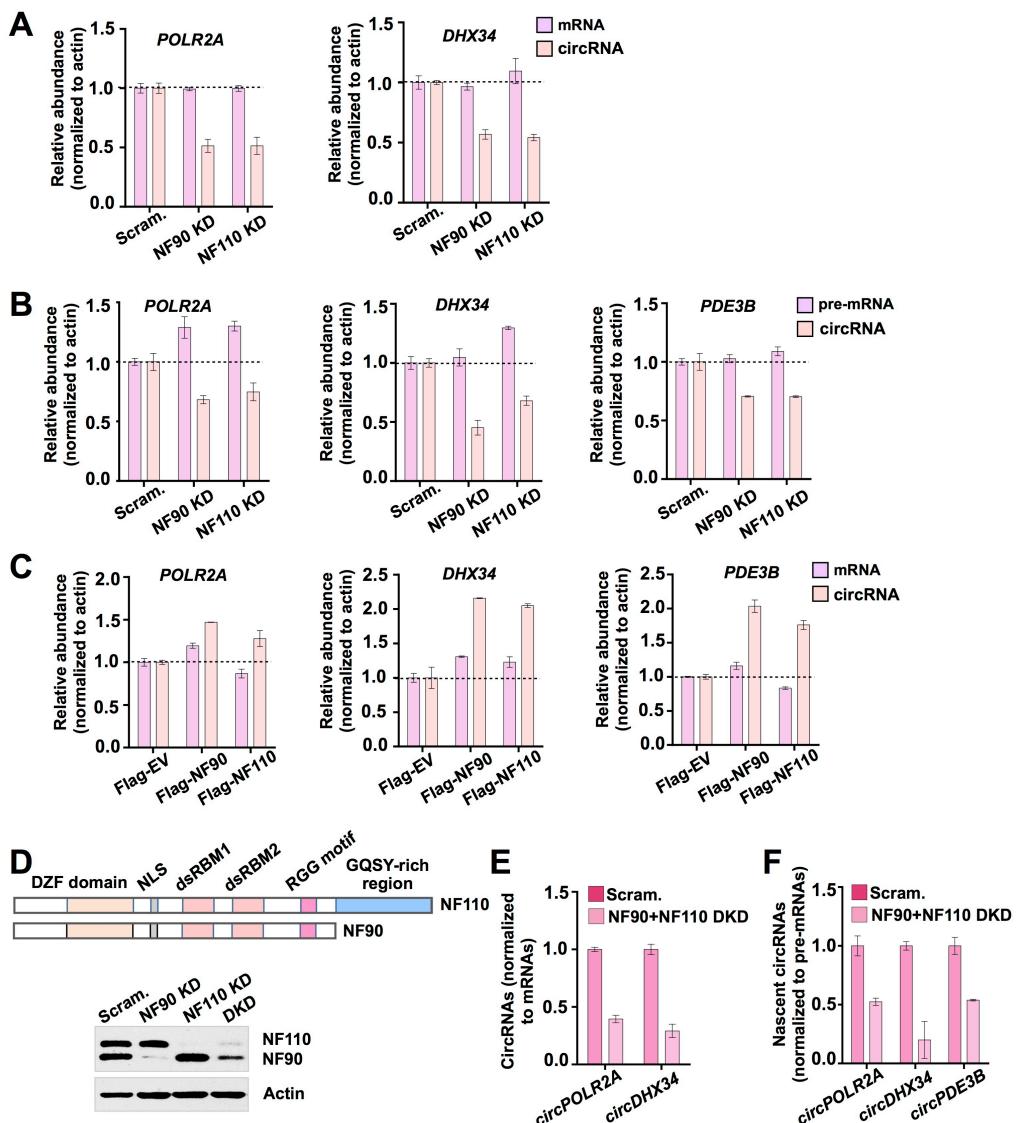
(C) The numbers of total reads, mapped reads, mapped ratio and circRNAs detected by CIRCExplorer (Zhang et al., 2014) are listed in each indicated steady-state RNA-seq dataset, combined from two RNA-seq replicates.

(D) Genome-wide expression of circRNAs in NF90 or NF110 KD HeLa cells. A total of 251 (magenta dots, left panel) or 255 (cyan dots, right panel) high-confidence circRNAs (back splicing junction, BSJ, reads  $\geq 5$ ) were identified by the poly(A)-Ribo-RNA-seq in NF90 or NF110 KD cells, respectively. The expression of each circRNA was normalized to its linear RNA isoform under each condition.

(E) The majority (about 85%, ~213/251 or 255) of high-confidence circRNAs were detected in both NF90 or NF110 KD cells.

(F) 50% of circRNAs detected in S2E were down-regulated in NF90 or NF110 stable KD cell lines.

(G) NF90 and NF110 could target the same circRNA-producing genes. About 70% of the down-regulated circRNAs in S2F were identical circles in NF90 or NF110 KD cells.



**Figure S3. Change of circRNA and corresponding mRNA or pre-mRNA upon NF90 or NF110 depletion or over-expression (related to Figure 2).**

(A) Steady-state level of mRNA and circRNA in NF90 or NF110 stable KD HeLa cells. The relative abundance of each circRNA and corresponding linear mRNA at the steady state was normalized to *actin*. Primer sets were designed to recognize circRNA and spliced mRNA at each *POLR2A* and *DHX34* locus shown in Figure 2F.

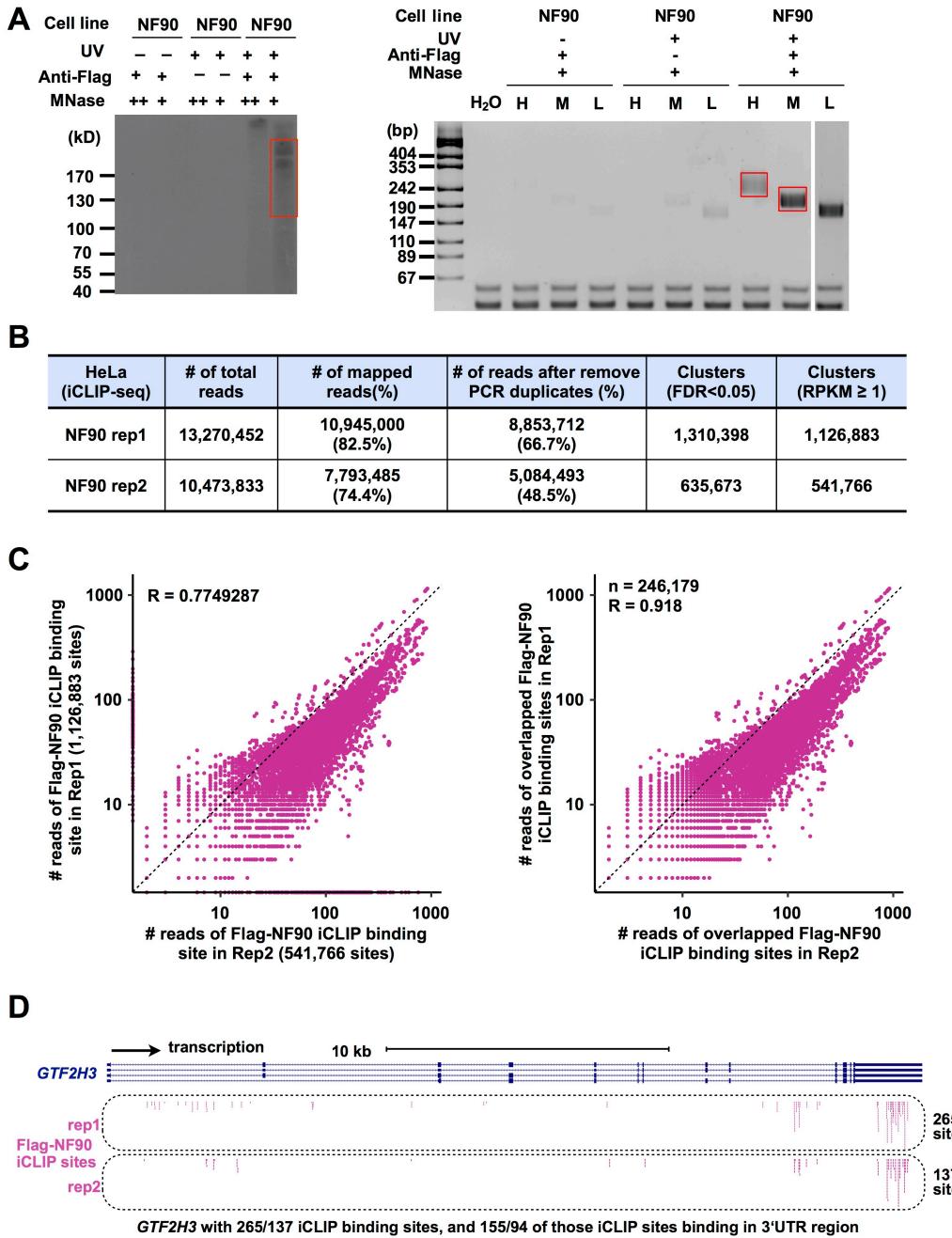
(B) Nascent level of pre-mRNA and circRNA in NF90 or NF110 stable KD HeLa cells. Nascent RNAs labeled with 4sU were purified from HeLa cells (Zhang et al., 2016b) and for RT-qPCR with primer sets that recognize circRNA and pre-mRNA at each *POLR2A*, *DHX34* and *PDE3B* locus shown in Figure 2F. The relative abundance of each circRNA and corresponding pre-mRNA was normalized to pre-RNAs of *actin*.

(C) Steady-state level of mRNA and circRNA in NF90 or NF110 stable over-expression HeLa cells. The relative abundance of each circRNA and corresponding linear mRNA at the steady state was normalized to *actin*.

(D) A schematic drawing of NF90/NF110 and verification of NF90/NF110 transient double KD in HeLa cells by WB. A shRNA targeting the common region of NF90/NF110 was used.

(E) NF90/NF110 double KD reduced circRNA expression. The relative abundance of each circRNA in NF90/NF110 double KD cells was normalized to its spliced mRNA isoform, as revealed by RT-qPCR.

(F) NF90/NF110 double KD reduced nascent circRNA production. Nascent RNAs labeled with 4sU were purified from NF90/NF110 transiently double KD HeLa cells (Zhang et al., 2016b) for RT-qPCR with primer sets that recognize the circRNA and pre-mRNA at each indicated gene locus. The relative abundance of each circRNA was normalized to its pre-mRNA.



**Figure S4. Characterization of Flag-NF90 iCLIP (related to Figure 3).**

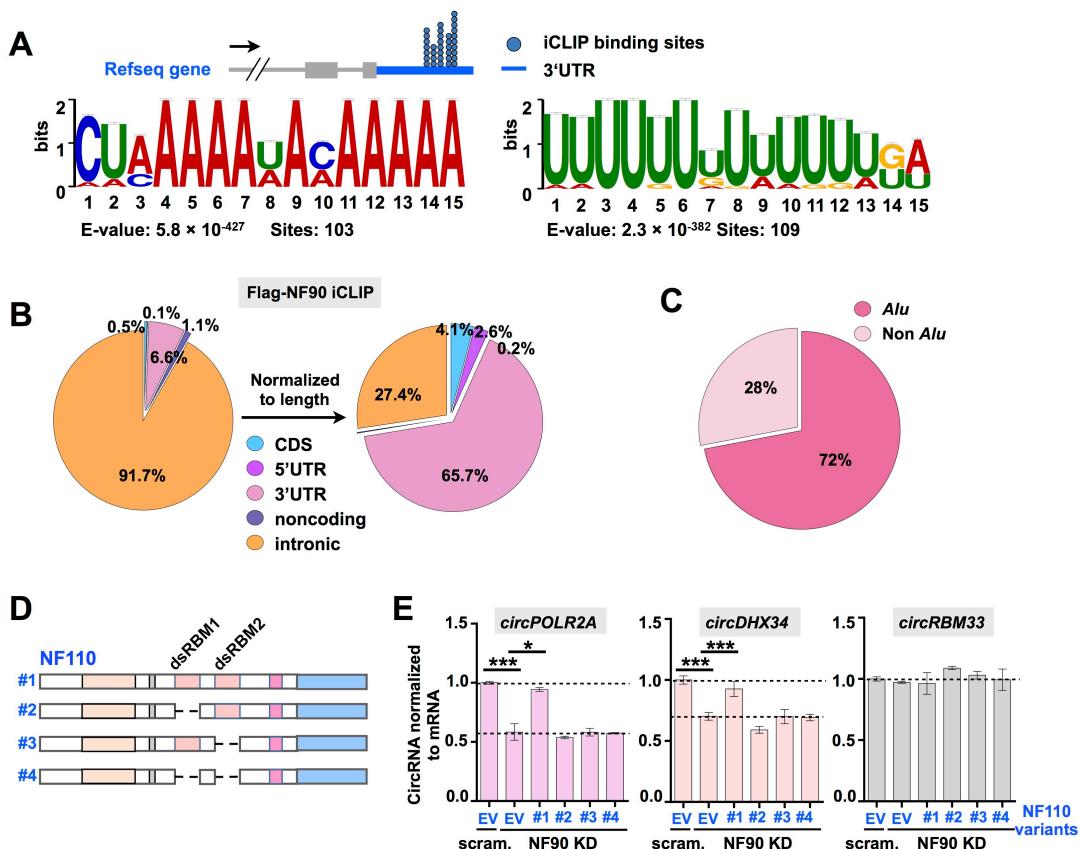
(A) iCLIP-seq library preparation. Left, analysis of cross-linked Flag-NF90-RNA complexes using denaturing gel electrophoresis. Cell lysates were prepared from UVCross-linked HeLa cells stably expressing Flag-NF90, and total RNAs were partially digested using low (+) or high (++) concentration of MNase. Protein-RNA complexes were immuno-purified from cell lysates using an antibody against Flag. The pre-adenylated L3 linker (RNA adapter) was ligated to the 3' ends of RNAs before radioactively labeling the 5' ends. Complexes were size-separated using denaturing gel electrophoresis and transferred to a nitrocellulose membrane. The retardation of protein-RNA complexes was observed compared to the size of the proteins. The radioactive signal disappeared when cells were not cross-linked (-UV). The red boxes marked regions of membranes where were cut out for subsequent purification steps. Right, analysis of PCR-amplified iCLIP cDNA libraries using denaturing gel electrophoresis. RNAs recovered from membrane regions were reverse transcribed and size-purified using denaturing gel electrophoresis (not shown). Three size fractions of cDNAs (High, 120-200 nt, Medium, 85-120 nt, Low, 70-85 nt) were recovered, circularized, linearized and PCR-amplified. PCR

products were separated on denaturing gel and visualized by EB staining. PCR products were absent or weak when cells were not cross-linked (-UV) or not antibody purified. Red boxes indicated the high and medium bands that were combined for high throughput sequencing. One set of data from the duplicated iCLIP-seq library preparation were shown.

(B) The list of total reads, mapped reads, reads after the removal of PCR duplicates and iCLIP enriched clusters of iCLIP-seq for Flag-NF90 in two independent iCLIP experiments.

(C) Correlation of Flag-NF90 iCLIP replicates. Left panel, correlation of all Flag-NF90 iCLIP binding sites between replicate 1 (1,126,883 binding sites) and replicate 2 (541,766 binding sites). Right panel, correlation of overlapped Flag-NF90 iCLIP binding sites in duplicates (246,179 binding sites). Magenta dots present Flag-NF90 iCLIP binding sites identified in replicate 1 and/or replicate 2. Pearson Correlation Coefficient (R value) is  $\sim 0.77$  for all sites (left panel) or  $\sim 0.92$  for overlapped sites (right panel).

(D) Flag-NF90 binding sites were enriched in the 3' UTR of the *GTF2H3* gene. *GTF2H3* has 265/137 iCLIP binding sites, and 155/94 of these iCLIP sites are located in the 3' UTR region in duplicates, respectively.



**Figure S5. Characterization of Flag-NF90 iCLIP (related to Figure 3).**

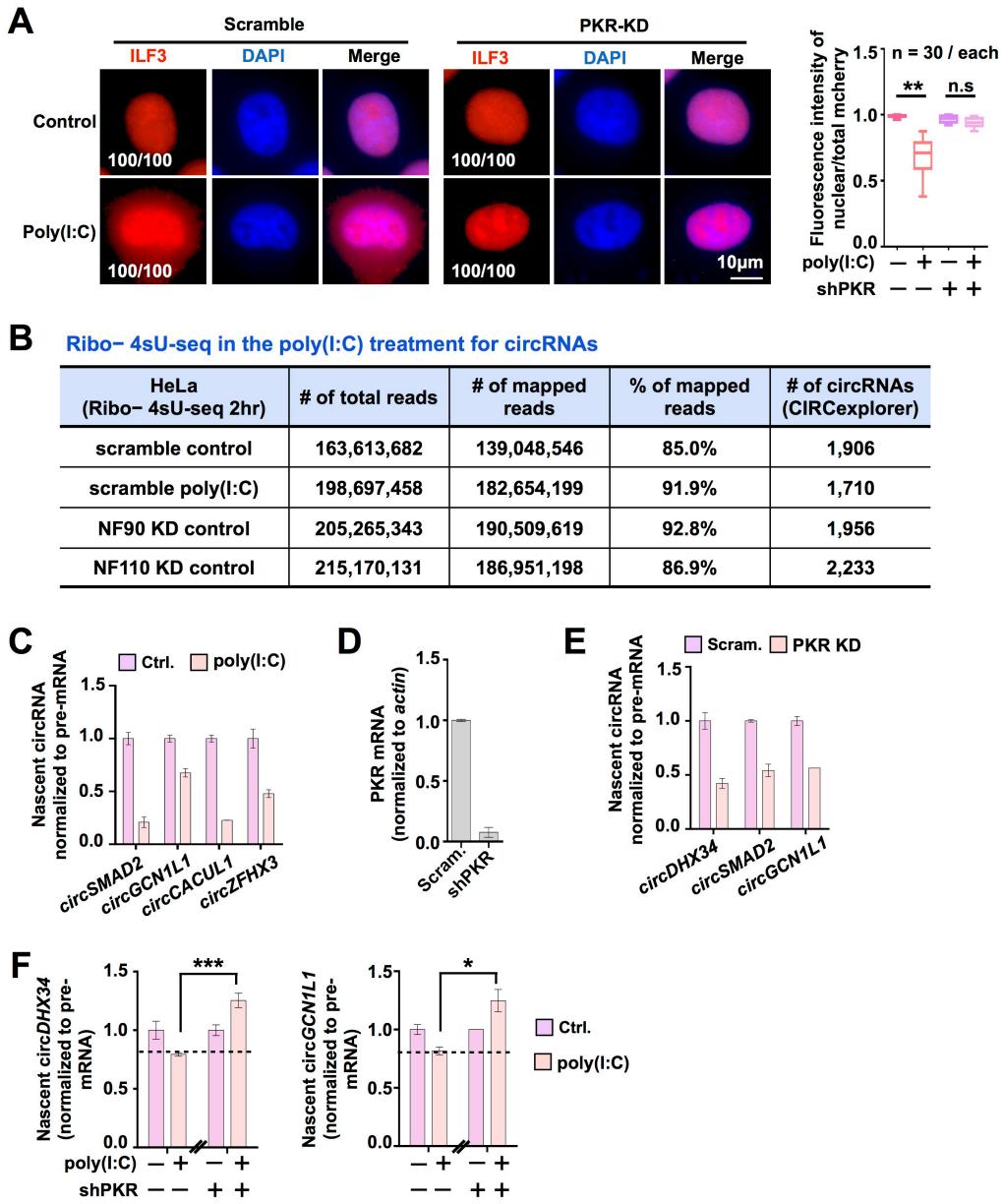
(A) Motif analysis of Flag-NF90 in 3' UTRs. Flag-NF90 binding sites in 3' UTRs were identified by iCLIP. Motif logos (by MEME Suite v4.11.2) probability were computed from aligning enriched 15-mers for Flag-NF90 iCLIP binding circRNA flanking introns. The top 500 Flag-NF90 iCLIP site sequences (+/- 10 nt) of 3' UTR were used to create motifs. Two predominant motifs A-rich and U-rich were shown.

(B) Distribution of the Flag-NF90 binding sites. Left, Flag-NF90 binding sites with iCLIP reads identified from duplicates revealed that NF90 binding sites were enriched in introns. Right, distribution of the Flag-NF90 binding sites after normalization to length.

(C) Distribution of Flag-NF90 binding sites in *Alu* or non-*Alu* sequences.

(D) A schematic drawing of NF110 truncations used in S5E.

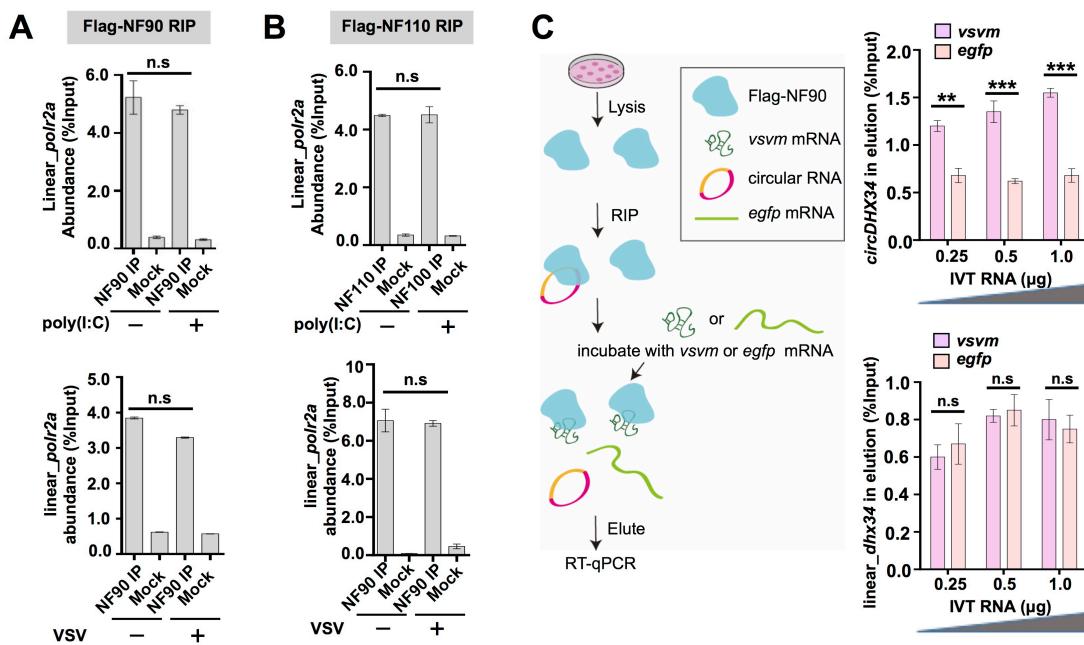
(E) Rescue circRNA reduction in NF90 KD cells by NF110. The decreased expression of *circPOLA2A* and *circDHX34* in NF90 KD cells could be largely rescued by re-introduction of NF110 wt (#1) but not its truncations (#2-#4). Error bars represent SD in triplicate experiments. \* $p \leq 0.05$ , \*\*\* $p \leq 0.001$ .



**Figure S6. Change of circRNA generation upon NF90/NF110 subcellular re-distribution (related to Figure 4).**

(A) PKR was required for NF90 and NF110 translocation to the cytoplasm upon poly(I:C) treatment. Left, re-distribution of NF90/NF110 upon 12 hours of poly(I:C) treatment was detected by immunofluorescence with anti-ILF3 antibodies without (scramble) or with (PKR-KD) PKR depletion. Right, statistical analyses showed about 30% NF90/NF110 in each examined HeLa cell were exported to the cytoplasm upon poly(I:C) treatment. But NF90/NF110 were no longer exported to the cytoplasm upon poly(I:C) treatment in PKR depleted cells. 30 cells were examined in each condition. Raw image 3D stacks were imported into Fiji/ImageJ. The nucleus was masked by Yen method and the whole cell was masked by Triangle method followed by removing the background by NaN background method. Then images were converted from 16 bit to 8 bit. The integrated fluorescence intensity for signal in each nucleus or whole cell was calculated by 3D objects counter.

- (B) The numbers of total reads, mapped reads, mapped ratio and circRNAs detected by CIRCExplorer (Zhang et al., 2014; Zhang et al., 2016) in indicated nascent RNA-seq datasets.
- (C) Validation of sequencing data in Figure 4E. Nascent RNAs labeled with 4sU were purified from HeLa cells upon poly(I:C) treatment for RT-qPCR.
- (D) Validation of PKR knockdown. The relative abundance of PKR mRNA was normalized to *actin*, as revealed by RT-qPCR.
- (E) PKR KD reduced nascent circRNA production. Nascent RNAs labeled with 4sU were purified from HeLa cells and were subjected to RT-qPCR with primer sets that recognize the circRNA and pre-mRNA at each indicated gene locus.
- (F) PKR knockdown impaired the reduction of nascent circRNAs in response to poly(I:C) treatment. Nascent RNAs labeled with 4sU were purified from the scramble or PKR shRNA treated HeLa cells upon poly(I:C) treatment for RT-qPCR.
- In (C)(E)(F), the relative abundance of each circRNA was normalized to its pre-mRNA. Primer sets were illustrated in Figure 2F.
- In (F), Error bars represent SD in triplicate experiments. \* $p \leq 0.05$ , \*\*\* $p \leq 0.001$ .



**Figure S7. NF90/NF110-associated cellular mRNAs remained unchanged upon poly(I:C) treatment or VSV infection (related to Figure 5).**

- (A) NF90-bound linear *POLA2A* mRNAs remained unaltered in poly(I:C) treated or VSV infected cells. HeLa cells stably expressing Flag-NF90 were treated with or without poly(I:C) or VSV followed by RIP using anti-Flag or anti-IgG antibodies. The percentage of RIP-enriched linear isoform of circRNA-producing gene relative to input was calculated under different conditions.
- (B) NF110-bound linear *POLA2A* mRNAs remained unaltered in poly(I:C) treated or VSV infected cells. See (A) for details.
- (C) *vsvm* but not *egfp* mRNA competed with NF90 bound circular RNAs. Left panel, a schematic illustration of *vsvm* or *egfp* mRNA competition assay with NF90 bound RNAs (circular RNAs or linear mRNA isoforms). Right panel, different amounts of *in vitro* synthesized *vsvm* or *egfp* were incubated with Flag-NF90-RNA complexes. The relative abundance of circRNAs (top right) or mRNAs (bottom right) from the supernatant competed from the incubation with *vsvm* or *egfp* was assayed by RT-qPCRs.

In (A-C), Error bars represent SD in triplicate experiments. \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ .

## **Supplemental Tables**

### **Supplemental Table S1. Fluorescence density of reporter genes in the human ON-TARGETplus (OPT) siRNA screening (related to Figure 1).**

Normalized fluorescence densities of mCherry and EGFP from two independent whole genome RNAi screening were individually listed in sheets 1 and 2. Targeted genes that have potential to affect circRNA biogenesis with either positive (mCherry fluorescence < 0.67 and 0.67 < EGFP fluorescence < 1.5) or negative ( fluorescence mCherry > 1.5 and 0.67 < EGFP fluorescence < 1.5) effect were listed in sheets 3 and 4. The 103 known RBPs that have potential to be involved in circRNA expression were listed in sheet 5.

### **Supplemental Table S2. List of protein factors identified by Mass Spectrometry in dsRNA pulldown (related to Figure 2).**

Proteins binding to either perfect (IRegfps v.s single egfp) or imperfect (inverted repeated *Alu* sequences (IRAlus) v.s single *Alu*) dsRNA duplexes were identified by Biotin-labeled dsRNA pulldown and Mass Spectrometry. Gene/protein identities, dsRNA enrichment (fold change) of perfect or imperfect duplexes and relative *p*-value ( $-\log_{10}$ ) were listed.

### **Supplemental Table S3. List of mRNA gene expression in scramble, NF90 KD or NF110 KD cells (related to Figure 2).**

Expression of Refseq genes was individually determined from poly(A)+ RNA-seq in cells with scramble, NF90 KD or NF110 KD by RPKM, listed with gene symbol, genomic location, strand information, gene length and exon number.

### **Supplemental Table S4. List of highly-expressed circRNAs in HeLa cells with NF90 KD or NF110 KD (related to Figure 2).**

Expression of highly-expressed circRNAs (with reads spanning BSJ  $\geq 5$  in at least one sample) was individually determined from poly(A)-Ribo- RNA-seq in cells with scramble, NF90 (sheet 1) KD or NF110 (sheet 2) KD by RPM, listed with circRNA host gene symbol, genomic location, and reads spanning each examined BSJ.

### **Supplemental Table S5. List of nascent circRNAs upon poly(I:C) treatment (related to Figure 4).**

Expression of nascent circRNAs was individually determined from nascent RNA-seq in control or poly(I:C)-treated cells by RPM, listed with circRNA host gene symbol, genomic location, and reads spanning each examined BSJ. All examined nascent circRNAs were listed in sheet 1. Highly-expressed ones were listed in sheet 2. Down-regulated or up-regulated ones with poly(I:C) treatment (reads spanning BSJ  $\geq 1$  in both ctrl and poly(I:C)-treated samples) were listed in sheets 3 and 4, respectively.

### **Supplemental Table S6. List of nascent circRNAs in HeLa cells with NF90 KD or NF110 KD (related to Figure 4).**

Expression of nascent circRNAs was individually determined from nascent RNA-seq in cells with control, NF90 KD or NF110 KD by RPM, listed with circRNA host gene symbol, genomic location, and reads spanning each examined BSJ. Highly-expressed nascent circRNAs were determined with at least five reads spanning the examined BSJs.

### **Supplemental Table S7. List of primer sequences.**