nature structural & molecular biology

Exon-intron circular RNAs regulate transcription in the nucleus

Zhaoyong Li[1,2,7], Chuan Huang[1,7], Chun Bao[1,3,4,7], Liang Chen[1], Mei Lin[1], Xiaolin Wang[1], Guolin Zhong[1], Bin Yu[1], Wanchen Hu[1], Limin Dai[1], Pengfei Zhu[1], Zhaoxia Chang[1], Qingfa Wu[1], Yi Zhao[5], Ya Jia[3,4], Ping Xu[6], Huijie Liu[1] & Ge Shan[1,2]

Noncoding RNAs (ncRNAs) have numerous roles in development and disease, and one of the prominent roles is to regulate gene expression. A vast number of circular RNAs (circRNAs) have been identified, and some have been shown to function as microRNA sponges in animal cells. Here, we report a class of circRNAs associated with RNA polymerase II in human cells. In these circRNAs, exons are circularized with introns 'retained' between exons; we term them exon-intron circRNAs or EIciRNAs. EIciRNAs predominantly localize in the nucleus, interact with U1 snRNP and promote transcription of their parental genes. Our findings reveal a new role for circRNAs in regulating gene expression in the nucleus, in which EIciRNAs enhance the expression of their parental genes in *cis*, and highlight a regulatory strategy for transcriptional control via specific RNA-RNA interaction between U1 snRNA and EIciRNAs.

One of the central roles of ncRNAs is to regulate gene expression[1–7]. Multiple long ncRNAs such as XIST and HOTAIR control gene expression through the epigenetic modification of chromatin status[8,9], and ncRNAs such as U1 RNA and 7SK RNA modulate transcription via their association with the RNA polymerase II (Pol II) transcription complex[10–13]. Multiple ncRNAs such as roX1, roX2 and HOTAIR regulate gene transcription in *trans*, and each affects the expression of genes located away from the locus where it is transcribed[14]. Other ncRNAs such as linc-HOXA1 and Air function in *cis* and may directly affect the expression of only limited neighboring genes[15,16]. There are also ncRNAs, such as ncRNA-a, that execute their gene-regulatory functions with both *cis*- and *trans*-mediated mechanisms[17].

The existence of circRNAs with covalent linkages in mammalian cells was indicated by EM more than three decades ago[18], and since then circRNAs generated from exons of specific coding genes have been reported sporadically[19–21]. Through high-throughput RNA sequencing and bioinformatic analysis, circRNAs have now been recognized as a large species of RNAs with thousands of members in animal cells[22–28]. circRNAs often show developmental stage–specific and tissue-specific expression, thus suggesting potential regulatory roles[26,27].

Several lines of evidence have indicated that circRNAs are most probably noncoding[21,27,29]. Recent reports have demonstrated that at least two circRNAs are able to function as microRNA sponges, with CDR1as as a sponge for miR-7 and circRNA generated from the *Sry* gene (circSry) as a sponge for miR-138 (refs. 27,28). Consistently with their roles in regulating microRNA functions, CDR1as and circSry are predominantly localized in the cytoplasm[27,28]. Bioinformatic analysis has shown that the majority of circRNAs do not possess multiple binding sites for microRNAs and thus may not function as microRNA sponges[29].

We set out to identify Pol II–associated ncRNAs, speculating that some of these RNAs might regulate gene transcription. We identified a subclass of circRNAs that, to our knowledge, was previously undescribed, and we found that these circRNAs localize in the nucleus. With further biochemistry and molecular characterizations, we provide lines of evidence that some of these circRNAs regulate the Pol II transcription of their parental genes in *cis* via specific RNA-RNA interaction.

RESULTS

A special subclass of circRNAs associated with Pol II

We hypothesized that some ncRNAs are involved in transcriptional regulation and sought to identify Pol II–associated ncRNAs via cross-linking and immunoprecipitation (CLIP) with an antibody to Pol II (**Fig. 1a** and **Supplementary Fig. 1a**). RNA sequencing of Pol II CLIP samples and subsequent bioinformatic analysis revealed that some circRNAs were associated with Pol II in HeLa cells (**Supplementary Tables 1–3**). We identified a total of 111 circRNAs by Pol II CLIP. We noted that, similarly to data described in a previous report[27], the read number was low for most circRNAs identified. We believe that the number of circRNAs deduced from the deep-sequencing data might be underestimated for two reasons: first, the deep-sequencing protocol was not optimized for detecting circRNAs; second, bioinformatic

[1]School of Life Sciences, University of Science and Technology of China, Hefei, China. [2]Chinese Academy of Sciences Key Laboratory of Brain Function and Disease, University of Science and Technology of China, Hefei, China. [3]Department of Physics, Central China Normal University, Wuhan, China. [4]Institute of Biophysics, Central China Normal University, Wuhan, China. [5]Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. [6]National Center for Protein Sciences, Beijing, China. [7]These authors contributed equally to this work. Correspondence should be addressed to G.S. (shange@ustc.edu.cn).

Received 10 November 2014; accepted 19 December 2014; published online 9 February 2015; corrected after print 19 August 2016; doi:10.1038/nsmb.2959

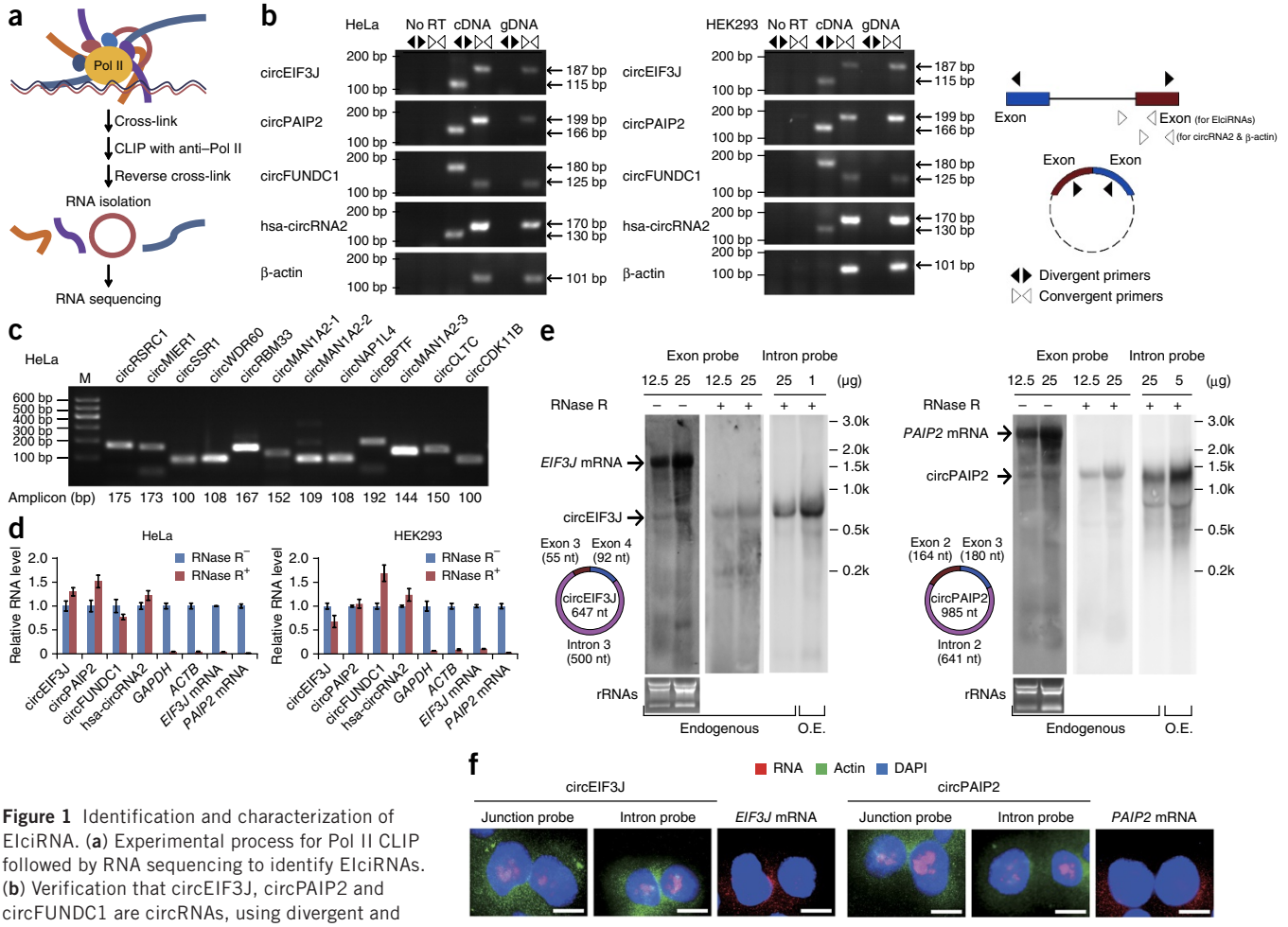

Figure 1 Identification and characterization of EIciRNA. (**a**) Experimental process for Pol II CLIP followed by RNA sequencing to identify EIciRNAs. (**b**) Verification that circEIF3J, circPAIP2 and circFUNDC1 are circRNAs, using divergent and convergent primers. hsa-circRNA2, known circRNA positive control; β-actin mRNA, negative control; gDNA, genomic DNA; RT, reverse transcription. (**c**) RT-PCR products with divergent primers showing circularization of the other 12 circRNAs in HeLa cells. M, size marker. (**d**) Real-time PCR showing resistance of circEIF3J, circPAIP2 and circFUNDC1 to RNase R digestion. hsa-circRNA2 and *GAPDH*, *ACTB*, *EIF3J* and *PAIP2* mRNAs are positive or negative controls. Error bars, s.e.m. from triplicate experiments. (**e**) Northern blots of *EIF3J* and *PAIP2* transcripts showing that both circEIF3J and circPAIP2 are EIciRNAs, on the basis of hybridization with exon and intron probes. O.E., RNA from cells transfected with the circRNA-overexpression plasmids. K denotes 1,000 nt. Uncropped images of northern blots are shown in **Supplementary Data Set 1**. (**f**) Representative FISH images ($n = 150$ cells for each circRNA) of circEIF3J and circPAIP2, showing localization of both circRNAs to the nuclei of HEK293 cells. Immunostaining for β-actin and FISH for *EIF3J* and *PAIP2* mRNA is shown for comparison. Scale bars, 10 μm. Split-channel images are shown in **Supplementary Figure 2c**. For **e** and **f**, exon probe is complementary to the exonic sequences in the EIciRNA (the 5′ exon), junction probe is complementary to the junction sequences, and intron probe is complementary to the intron sequences in the EIciRNA.

analysis could count only junction sequences as circRNA reads, thus missing all other sequences enclosed within circRNAs. We confirmed the enrichment of the 15 most abundant circRNAs by Pol II RNA immunoprecipitation (RNA IP) and subsequent real-time quantitative PCR (**Table 1** and **Supplementary Fig. 1b**–**e**). Some of the circRNAs, such as circCLTC, demonstrated cell-specific expression in HeLa and HEK293 cells (**Supplementary Fig. 1f**). For the 15 RNAs, we confirmed that they were circular by RNase R digestion and divergent reverse-transcription PCR (RT-PCR; **Fig. 1b**–**d**), according to previously described methodology[24]. We further analyzed two of these circRNAs, circEIF3J and circPAIP2, by northern blotting and RT-PCR (**Fig. 1b**,**e**). The data demonstrated that introns between circularized exons were retained in both cases (**Fig. 1e** and **Supplementary Fig. 2a**). In actuality, all 15 circRNAs analyzed appeared to contain intronic sequences when examined by RT-PCR with divergent primers corresponding to the 5′ exon and upstream intron closest to the 3′ exon in the circRNA (**Supplementary Fig. 2b**). The intron-retention property of these circRNAs is distinct from microRNA-sponge circRNAs and the other circRNAs characterized experimentally, such as hsa-circRNA6, hsa-circRNA2 and hsa-circRNA9 (refs. 27,28); these circRNAs are also formed from exon back-splicing, but they consist exclusively of exonic sequences[27]. We have therefore termed our newly identified intron-containing circRNAs EIciRNAs. We estimated the approximate circEIF3J and circPAIP2 copy number per cell to be ~31 and ~22, respectively, in HeLa cells (copy-number analysis in Online Methods). From northern blots, we also calculated that the ratio of EIciRNA to parental mRNA is ~8.9% (circEIF3J/*EIF3J* mRNA) and ~9.4% (circPAIP2/*PAIP2* mRNA). Fluorescence *in situ* hybridization (FISH) revealed that circEIF3J and circPAIP2 are localized exclusively in the nucleus (**Fig. 1f** and **Supplementary Fig. 2c**–**g**). The nuclear localization of these EIciRNAs and their association with Pol II suggest that they might be involved in transcriptional regulation.

Table 1 The top 15 circRNAs enriched in the RNA-sequencing pool from Pol II CLIP

Name	Parental gene	Junction	Linear distance of head-to-tail junction (nt)
circEIF3J	*EIF3J*	Exon 3-4	647
circPAIP2	*PAIP2*	Exon 2-3	985
circRSRC1	*RSRC1*	Exon 2-3	1,889
circFUNDC1	*FUNDC1*	Exon 4-5(P)	3,364
circMIER1	*MIER1*	Exon 6-9	5,102
circSSR1	*SSR1*	Exon 2-3	6,480
circWDR60	*WDR60*	Exon 2-4	6,837
circRBM33	*RBM33*	Exon 3-5	8,042
circMAN1A2-1	*MAN1A2*	Exon 2-4	12,646
circMAN1A2-2	*MAN1A2*	Exon 2-5	18,464
circNAP1L4	*NAP1L4*	Exon 2-14	27,979
circBPTF	*BPTF*	Exon 23-29	30,550
circMAN1A2-3	*MAN1A2*	Exon 2-6	40,140
circCLTC	*CLTC*	Exon 2-30	41,533
circCDK11B	*CDK11B*	UPR-exon2	64,072

Parental gene sequences involved in head-to-tail splicing to form junctions and the linear distances of the head-to-tail junctions are shown. Exon 4-5(P) indicates that part of the known exon 5 sequence was involved in the circularization. UPR, the upstream region of the *CDK11B* gene.

EIciRNAs can be overexpressed with their flanking sequences

Previous bioinformatic analysis showed that sequences flanking exons forming circRNAs were more likely to contain complementary Alu elements, and there was also indication that flanking inverted repeats might mediate the circularization of circSry[30,31]. The genomic regions of 4 of the 15 circRNAs contained flanking Alu complementary pairs, and four others contained flanking complementary sequence pairs other than Alu (**Fig. 2a**). We constructed plasmids with DNA sequences corresponding to circEIF3J and circPAIP2 without flanking sequences (retaining the 5′ splicing site with the conserved AG and 3′ splicing site with the conserved GT); with their endogenous flanking sequences (which include the complementary Alu pairs); or with a 1-kb complementary repeat (detailed plasmid information in Online Methods). The endogenous flanking sequences resulted in the production of circRNAs at relative amounts of ~548-fold for circEIF3J and ~33-fold for circPAIP2 in HeLa cells; 1-kb complementary repeats produced a fold overexpression of ~2,310 for circEIF3J and ~123 for circPAIP2 (**Fig. 2b**). DNA sequences corresponding to circEIF3J and circPAIP2 without flanking sequence but with just the 5′ and 3′ splicing sites also resulted in the production of circRNAs, although with the lower relative amounts of ~35-fold for circEIF3J and ~3-fold for circPAIP2 (**Fig. 2b**). These constructs also had similar circRNA-overexpression effects in HEK293 cells (**Fig. 2c**). The overexpression plasmids also resulted in higher amounts of the corresponding circRNA in the nucleus, and the overexpressed circRNAs were also intron retaining (**Fig. 2d** and **Supplementary Fig. 3a**). From the data from these artificial plasmids, it appeared that flanking sequences could facilitate circularization of the RNA, and the circRNA sequences could possess internal circularization characteristics. During the process of publishing this work, two groups reported that flanking repeat sequences could facilitate circularization of the RNA[32,33] and that the circRNA sequences could possess internal

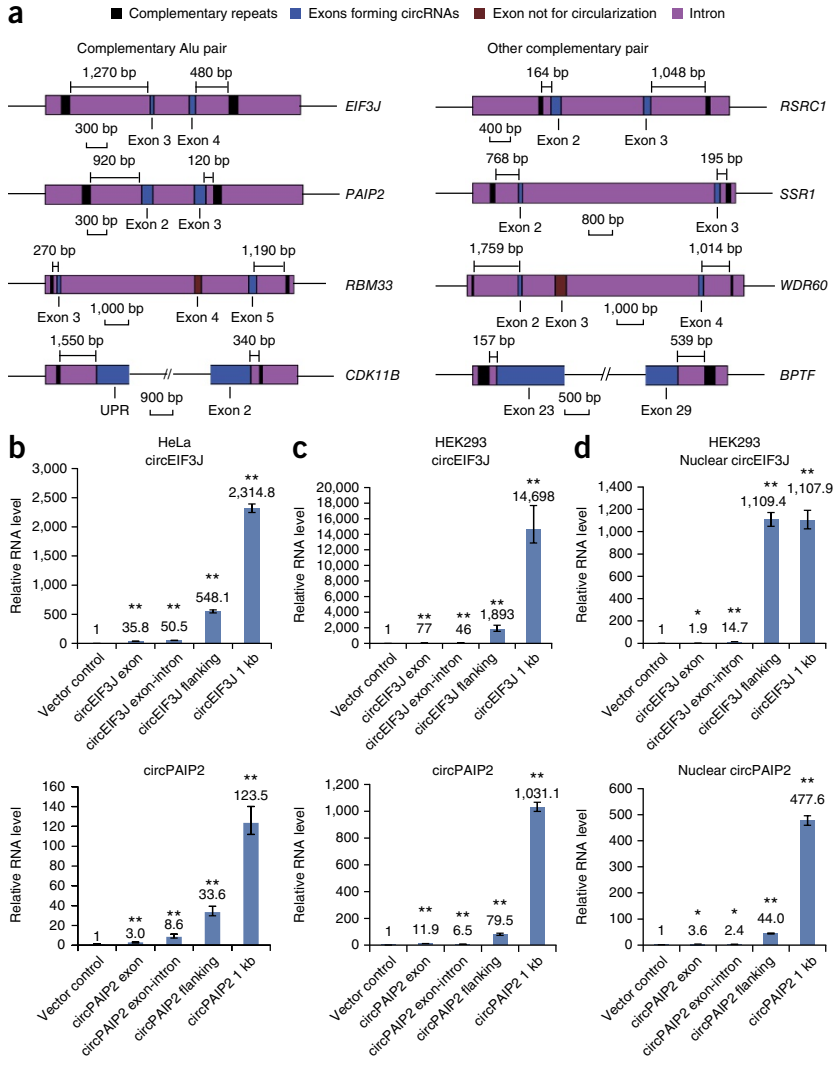

Figure 2 Sequences related to circularization. (**a**) Schematics showing that genomic regions of 4 out of 15 circRNAs contain flanking Alu complementary pairs, and another four contain flanking complementary sequences other than an Alu element. The linear distances from the repeat sequences to the exons forming circRNAs are labeled. Only those complementary repeats that had at least one part of the two complementary elements within 1.5 kb of the flanking region were counted. (**b**) Overexpression of circEIF3J and circPAIP2 with various constructs in HeLa cells. Details about the overexpression plasmids are in Online Methods. (**c**) As in **b** for HEK293 cells. (**d**) As in **b**, measured in HEK293 nuclei. In **b**–**d**, error bars, s.e.m. from triplicate transfections. $^{*}P < 0.05$; $^{**}P < 0.01$ by two-tailed Student's *t* test.

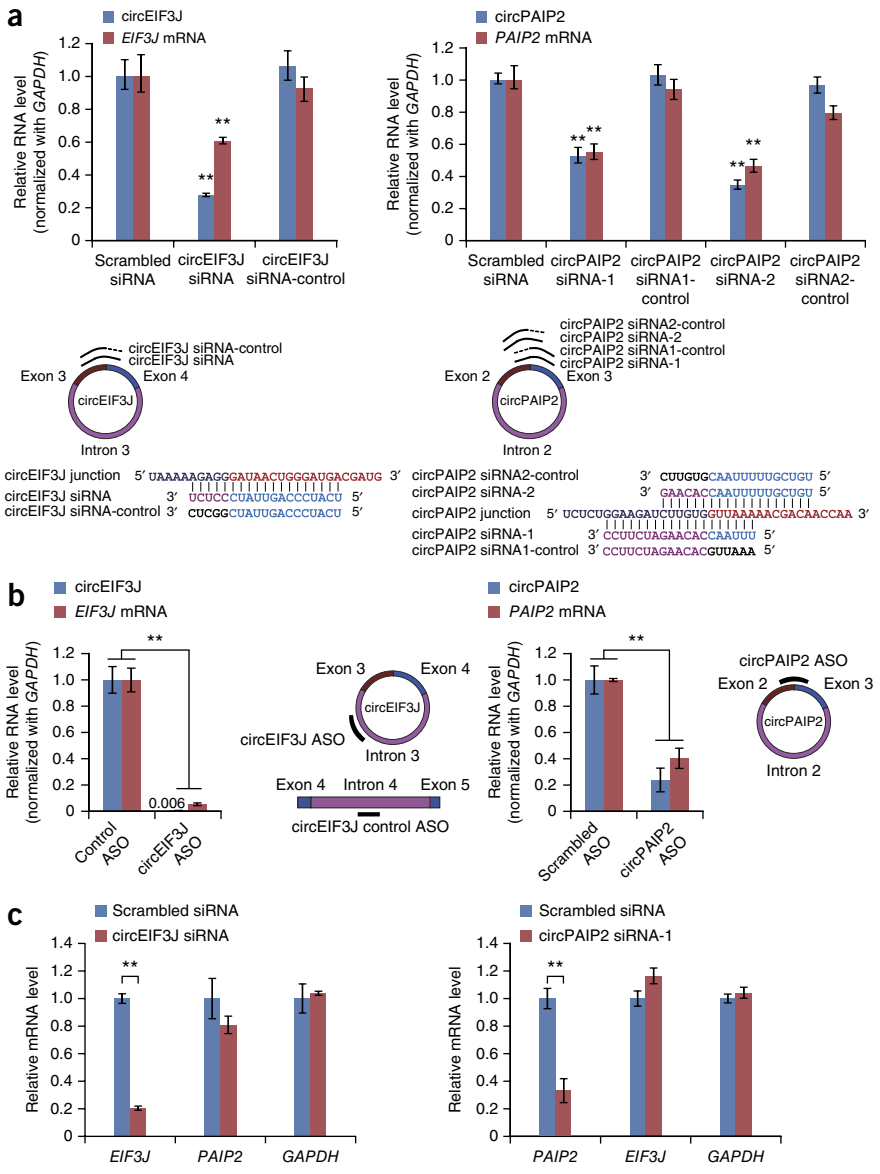

Figure 3 EIciRNAs have *cis* regulatory effects. (**a**) Decrease in mRNA levels of the parental genes after knockdown of circEIF3J or circPAIP2 with siRNA. The siRNAs targeted circEIF3J and circPAIP2 at junction sequences. The blue and purple sequences below the histograms show the corresponding 5′ and 3′ exon sequences forming the junction; mismatched sequences in the corresponding control siRNA are shown in black. (**b**) Decrease in mRNA levels of the parental genes after knockdown of circEIF3J or circPAIP2 with ASOs. For circEIF3J, the ASO targets intron sequences in the EIciRNA, and the control ASO targets a downstream intron. For circPAIP2, the ASO targets the EIciRNA junction (sequences specified in **Supplementary Fig. 3c**). (**c**) Nuclear run-on experiments showing a specific decrease in transcription of the corresponding parental gene after circRNA knockdown. The two genes were cross-examined, and *GAPDH* is a negative control. Throughout figure, error bars, s.e.m. from triplicate transfections. $^{**}P < 0.01$ by two-tailed Student's *t* test.

circularization characteristics[33]. We later used these overexpression plasmids to investigate whether EIciRNAs could regulate the expression of their parental genes in *trans*.

EIciRNAs regulate expression in *cis*

We started to evaluate roles of EIciRNAs by knocking down their expression. Knockdown of circEIF3J and circPAIP2 with short interfering RNAs (siRNAs) or RNase H–based antisense oligonucleotides (ASOs)[34] targeting EIciRNA resulted in a decrease in the mRNA levels of the parental genes in HeLa and HEK293 cells, respectively (**Fig. 3a**,**b** and **Supplementary Fig. 3b**,**c**). We also examined the 5′ neighboring genes of *EIF3J* and *PAIP2* and found that knockdown of circEIF3J and circPAIP2 had no effect on the neighboring genes (**Supplementary Fig. 3d**). To further examine whether the decrease in mRNA levels resulted from a decrease in the transcription of the parental genes, we performed nuclear run-on experiments. We knocked down circEIF3J and circPAIP2 with their corresponding siRNAs and then extracted cell nuclei for run-on experiments. We found that knockdown of circEIF3J and circPAIP2 resulted in lower *EIF3J* and *PAIP2* transcription levels, respectively (**Fig. 3c** and **Supplementary Fig. 3e**). In addition, in nuclear run-on experiments, knockdown of *EIF3J* and *PAIP2* mRNA with siRNA had no effect on the transcription of *EIF3J* and *PAIP2*, respectively (**Supplementary Fig. 3f**). Additionally, knockdown of *EIF3J* or *PAIP2* mRNA with short hairpin RNA (shRNA) or siRNA had no effect on the levels of circEIF3J or circPAIP2 (**Supplementary Fig. 4a**–**c**). Overexpression of circEIF3J and circPAIP2 from plasmids also had no substantial effect on the levels of *EIF3J* and *PAIP2* mRNA (**Supplementary Fig. 4d**,**e**). RNA-DNA double FISH revealed that circEIF3J and circPAIP2 each colocalized with the genomic loci of their corresponding parental genes in more than half of the cells (**Fig. 4a**,**b**). This finding is in contrast with genes that flank *EIF3J* or *PAIP2* at the 5′ and 3′ ends and the genomic locus for *GAPDH*, which demonstrated no colocalization with the two EIciRNAs (**Fig. 4c**). Collectively, these data suggest that circEIF3J and circPAIP2 may regulate the expression of their parental genes in *cis*. However, the localization of these two EIciRNAs in the nucleus was not confined to their parental gene loci, thus indicating potential *trans* effects of these EIciRNAs on loci other than their parental genes. In this study, we chose to focus on the *cis* effects of EIciRNAs on their parental genes.

EIciRNAs interact with Pol II, U1 snRNP and gene promoters

Using pulldown assays with biotin-labeled oligonucleotides complementary to specific EIciRNA sequences, we analyzed the proteins and RNAs coprecipitating with the two EIciRNAs. Furthermore, to investigate potential interactions between EIciRNAs and chromatin, we also examined genomic DNA coprecipitated (in chromatin isolation by RNA purification (ChIRP) experiments) with specific EIciRNAs (**Fig. 5a** and **Supplementary Fig. 5a**; efficiency and specificity of these pulldowns in **Supplementary Fig. 5a**). Pulldown with either circEIF3J or circPAIP2 yielded not only Pol II but also the U1A and U1C proteins and U1 small nuclear RNA (snRNA; **Fig. 5b**,**c** and **Supplementary Fig. 5b**–**f**). U2 snRNA could also be pulled down with the circRNAs, but it appears that the association between

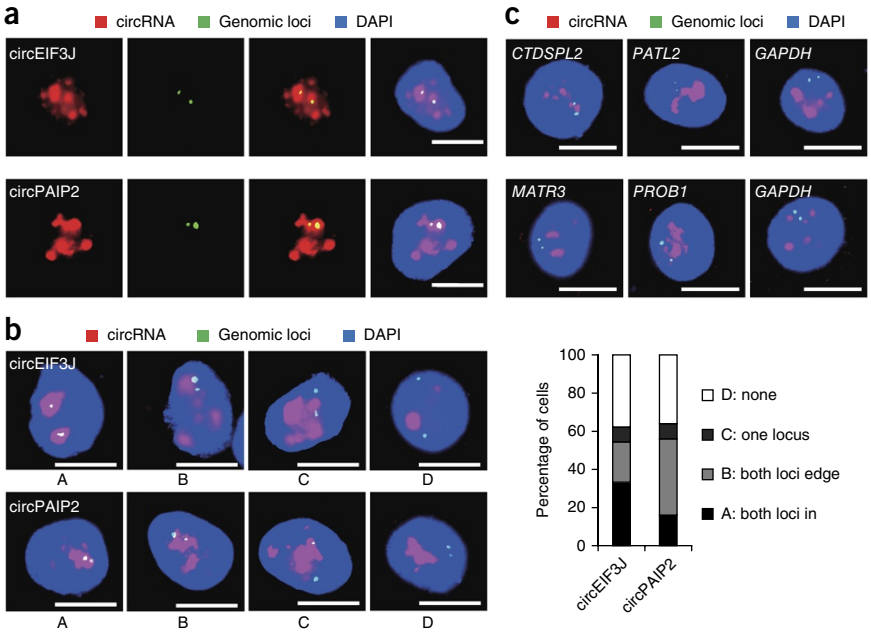

Figure 4 Double FISH for circRNA and its parental or neighboring gene loci. (**a**) Representative FISH images ($n = 30$ for circEIF3J and $n = 16$ for circPAIP2) showing colocalization of circRNAs (detected with a junction probe) and the corresponding parental gene loci. DAPI, 4′,6-diamidino-2-phenylindole stain. (**b**) FISH images demonstrating four categories in the colocalization of circRNA and parental gene loci: A, both gene loci are localized inside the circRNA signal; B, both gene loci are localized at the edge of, and colocalized with, circRNA; C, one gene locus is, and the other one is not, colocalized with circRNA; D, neither locus is colocalized with circRNA. Images shown in **a** belong to the category A. Statistical analysis for the four categories is shown in bar graph. Both parental gene loci are colocalized (categories A + B) with the circRNA in more than half of the cells, and the percentages are ~54.4% for circEIF3J and ~56% for circPAIP2 ($n = 90$ cells for circEIF3J and 100 cells for circPAIP2). (**c**) Representative double FISH images ($n = 40$ cells for each image panel) showing no circRNA colocalization with the genomic loci of *GAPDH* and corresponding 5′ or 3′ neighboring genes. For *EIF3J*, 5′ gene *CTDSPL2* and 3′ gene *PATL2*; for PAIP2, 5′ gene *MATR3* and 3′ gene *PROB1*. The *CTDSPL2* and *PROB1* genomic loci were occasionally just at the edge of the corresponding circRNA without colocalization. The circRNA FISH signal was detected with a junction probe in HEK293. Scale bars, 10 μm.

U2 snRNA and circRNA may be indirect (**Supplementary Fig. 5d**). We repeated these experiments with another set of biotin-labeled oligonucleotides complementary to exonic sequences in EIciRNA and obtained similar results (**Supplementary Fig. 5e,f**). Sites within the promoter and first-exon regions of the parental genes also coprecipitated (**Fig. 5d** and **Supplementary Fig. 6a,b**). Conversely,

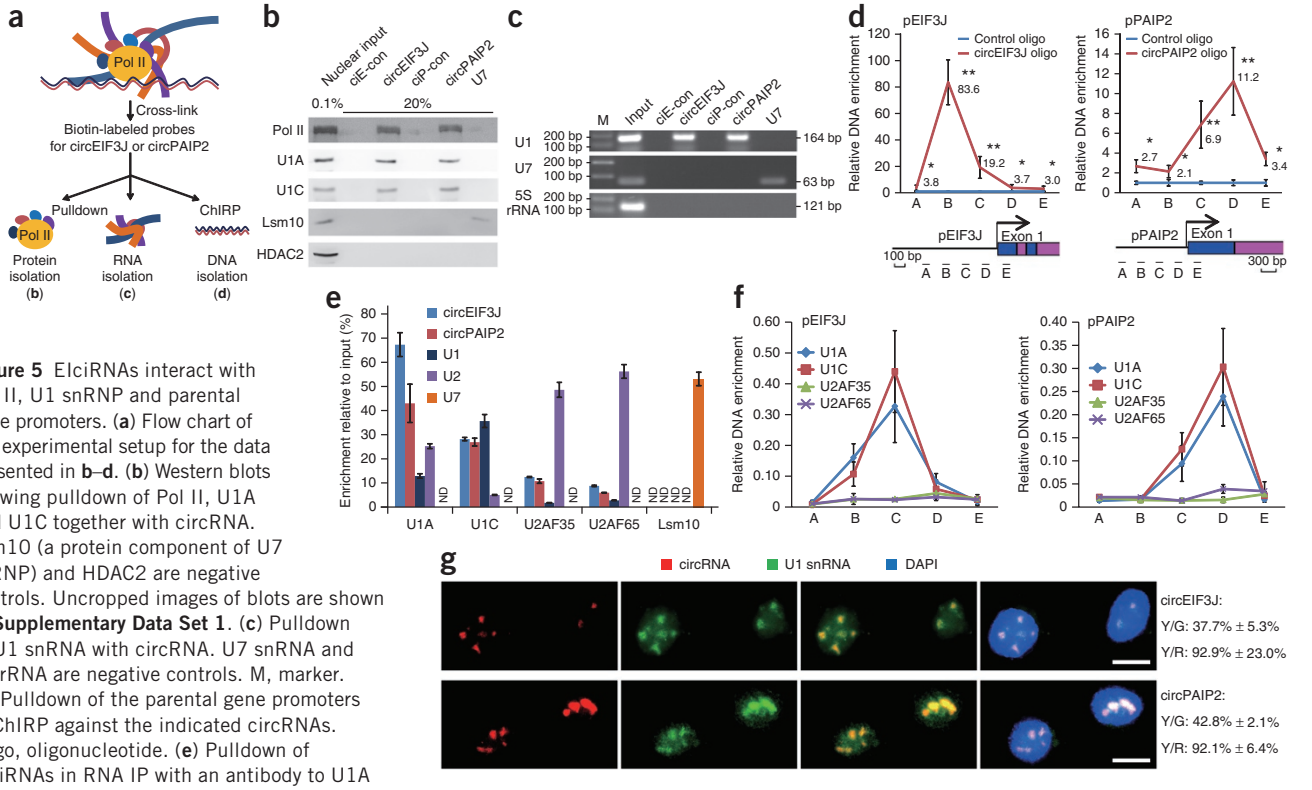

Figure 5 EIciRNAs interact with Pol II, U1 snRNP and parental gene promoters. (**a**) Flow chart of the experimental setup for the data presented in **b**–**d**. (**b**) Western blots showing pulldown of Pol II, U1A and U1C together with circRNA. Lsm10 (a protein component of U7 snRNP) and HDAC2 are negative controls. Uncropped images of blots are shown in **Supplementary Data Set 1**. (**c**) Pulldown of U1 snRNA with circRNA. U7 snRNA and 5S rRNA are negative controls. M, marker. (**d**) Pulldown of the parental gene promoters in ChIRP against the indicated circRNAs. Oligo, oligonucleotide. (**e**) Pulldown of EIciRNAs in RNA IP with an antibody to U1A or U1C. Results for U2AF35, U2AF65 and Lsm10 are shown for comparison. ND, not detected (as described in Online Methods). (**f**) Pulldown of the parental gene promoters by ChIP assay with antibodies to U1A or U1C. Results for U2AF35 and U2AF65 are shown for comparison. For **d**–**f**, error bars, s.e.m. from triplicate experiments. $^*P < 0.05$, $^{**}P < 0.01$ by two-tailed Student's t test. (**g**) Representative FISH images ($n = 80$ cells for each circRNA) showing colocalization of the majority of the circRNAs (detected with a junction probe) with U1 snRNA (Y/R ratio) in HEK293 cells. Y/G, ratio of yellow to green; Y/R, ratio of yellow to red. Scale bars, 10 μm. More information about EIciRNA and U1 snRNA colocalization is provided in **Supplementary Figure 7**.

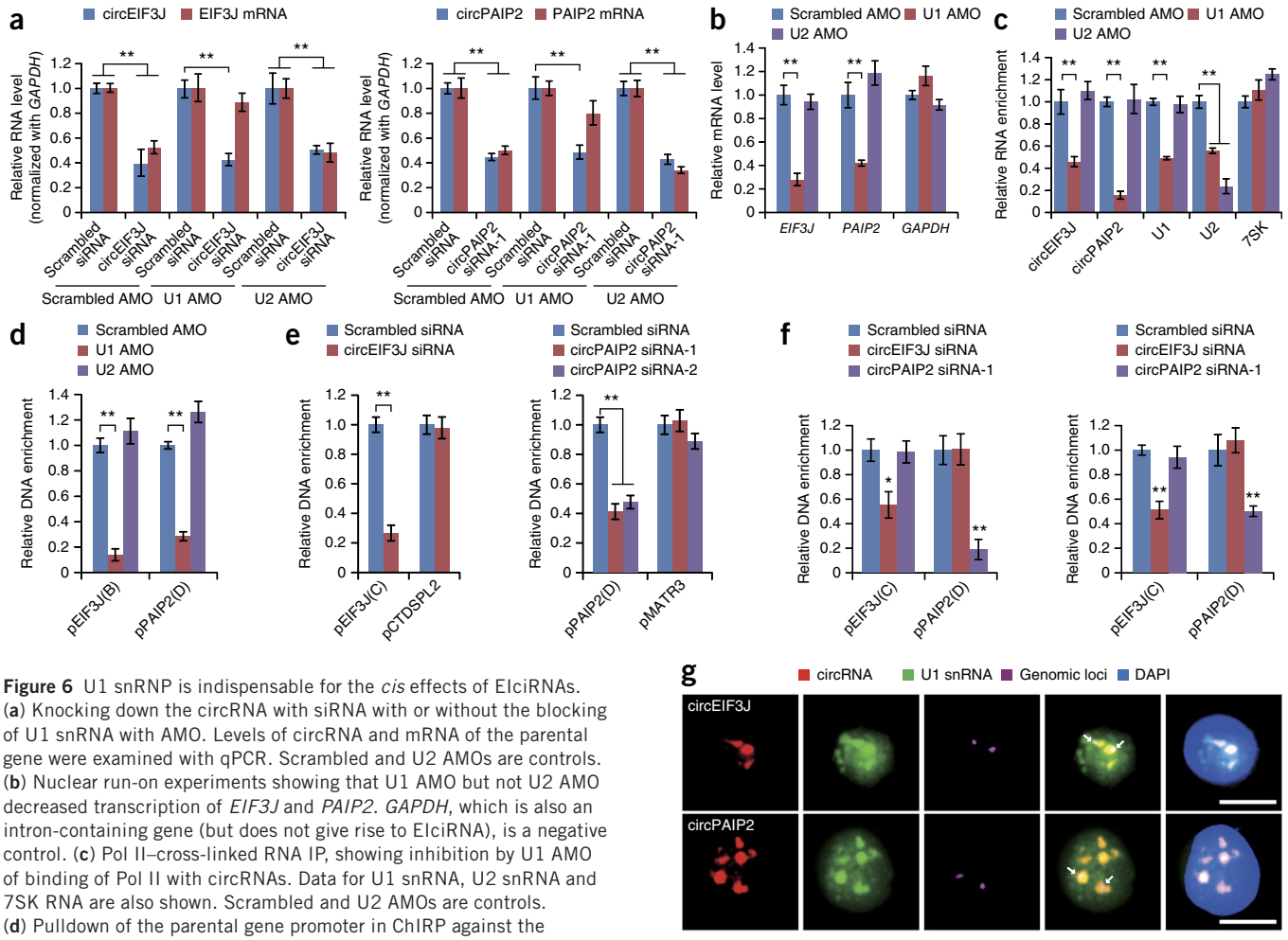

Figure 6 U1 snRNP is indispensable for the *cis* effects of EIciRNAs. (**a**) Knocking down the circRNA with siRNA with or without the blocking of U1 snRNA with AMO. Levels of circRNA and mRNA of the parental gene were examined with qPCR. Scrambled and U2 AMOs are controls. (**b**) Nuclear run-on experiments showing that U1 AMO but not U2 AMO decreased transcription of *EIF3J* and *PAIP2*. *GAPDH*, which is also an intron-containing gene (but does not give rise to EIciRNA), is a negative control. (**c**) Pol II–cross-linked RNA IP, showing inhibition by U1 AMO of binding of Pol II with circRNAs. Data for U1 snRNA, U2 snRNA and 7SK RNA are also shown. Scrambled and U2 AMOs are controls. (**d**) Pulldown of the parental gene promoter in ChIRP against the corresponding circRNA with or without the blocking of U1 snRNA with AMO. pEIF3J(B), the B site shown in **Figure 3d**; pPAIP2(D), the D site shown in **Figure 3d**. Scrambled and U2 AMOs are controls. (**e**) Pol II ChIP assay showing reduced Pol II binding to the parental gene promoters upon circRNA knockdown via siRNA. Data are also shown for promoters of corresponding neighboring genes (*CTDSPL2* for *EIF3J*; *MATR3* for *PAIP2*). pEIF3J(C), the C site shown in **Figure 3d**. (**f**) U1A and U1C ChIP assay showing reduced binding of U1A and U1C to the promoters of corresponding parental genes upon circRNA knockdown via siRNA. For **e** and **f**, more details are shown in **Supplementary Figure 8a**,**b**. (**g**) FISH images showing localization of the two parental gene loci (white arrow) to regions enriched for EIciRNA (detected with junction probe) and U1 snRNA in ~65.4% of cells for *EIF3J* and ~51.3% cells for *PAIP2* ($n = 52$ cells for EIF3J, 39 cells for PAIP2). Additional information about circRNA, U1 snRNA and parental-gene-loci FISH images is shown in **Supplementary Figure 7b**. $^{*}P < 0.05$; $^{**}P < 0.01$ by two-tailed Student's *t* test. Error bars, s.e.m. from triplicate experiments.

pulldown with the U1A or U1C proteins coprecipitated a substantial amount of EIciRNAs; pulldown with Lsm10 (a protein component of U7 small nuclear ribonucleoprotein (snRNP) or auxiliary factors (U2AF65 and U2AF35) for U2 recognition of the 3′ splice sites did not result in substantial enrichment in EIciRNAs (**Fig. 5e**). U1A and U1C but not U2AF65 or U2AF35 interacted with the promoter regions of corresponding parental genes, as demonstrated by chromatin immunoprecipitation (ChIP; **Fig. 5f**). ChIP results with antibodies to U1A, U1C, U2AF35 or U2AF65 for the promoter of the neighboring gene (*CTDSPL2* for EIF3J and *MATR3* for PAIP2) showed no enrichment (**Supplementary Fig. 6c**), thus indicating that the U1 snRNP binds to the promoters of some genes, such as *EIF3J* and *PAIP2*, but not to the promoters of genes such as *CTDSPL2* and *MATR3*. Our results (**Fig. 5d**,**f**) indicate that EIciRNA and the U1 snRNP (U1A and U1C) occupy a region ~300 bp upstream of the transcriptional start site of the parental genes, suggesting that EIciRNAs, U1 snRNP and Pol II might interact with each other at promoter regions of the parental genes. Dual RNA FISH for EIciRNA and U1 snRNA revealed that the majority of circEIF3J or circPAIP2 colocalized with U1 snRNA in the nucleus, further indicating potential interaction between them (**Fig. 5g** and **Supplementary Fig. 7**).

U1 snRNP is indispensable for the *cis* effects of EIciRNAs

It may be possible that U1 snRNP mediates the function of EIciRNAs. Indeed, we found that blocking U1 snRNA with a U1 antisense morpholino (AMO) abolished the effects of EIciRNA knockdown on the mRNA levels of the parental genes, whereas a U2 AMO had no such effect (**Fig. 6a**). Furthermore, U1 AMO but not U2 AMO decreased transcription of *EIF3J* and *PAIP2* in nuclear run-on experiments (**Fig. 6b**). The U1 AMO also substantially decreased the association between Pol II and the individual EIciRNAs examined (**Fig. 6c**) but not the interaction between Pol II and 7SK RNA, which was a negative control[11,12]. The interaction between EIciRNAs and the promoters of the parental genes was also reduced upon U1 AMO treatment (**Fig. 6d**). Conversely, knockdown of EIciRNA with siRNA decreased the binding of Pol II and specific U1 snRNP proteins (U1A and U1C) to their parental

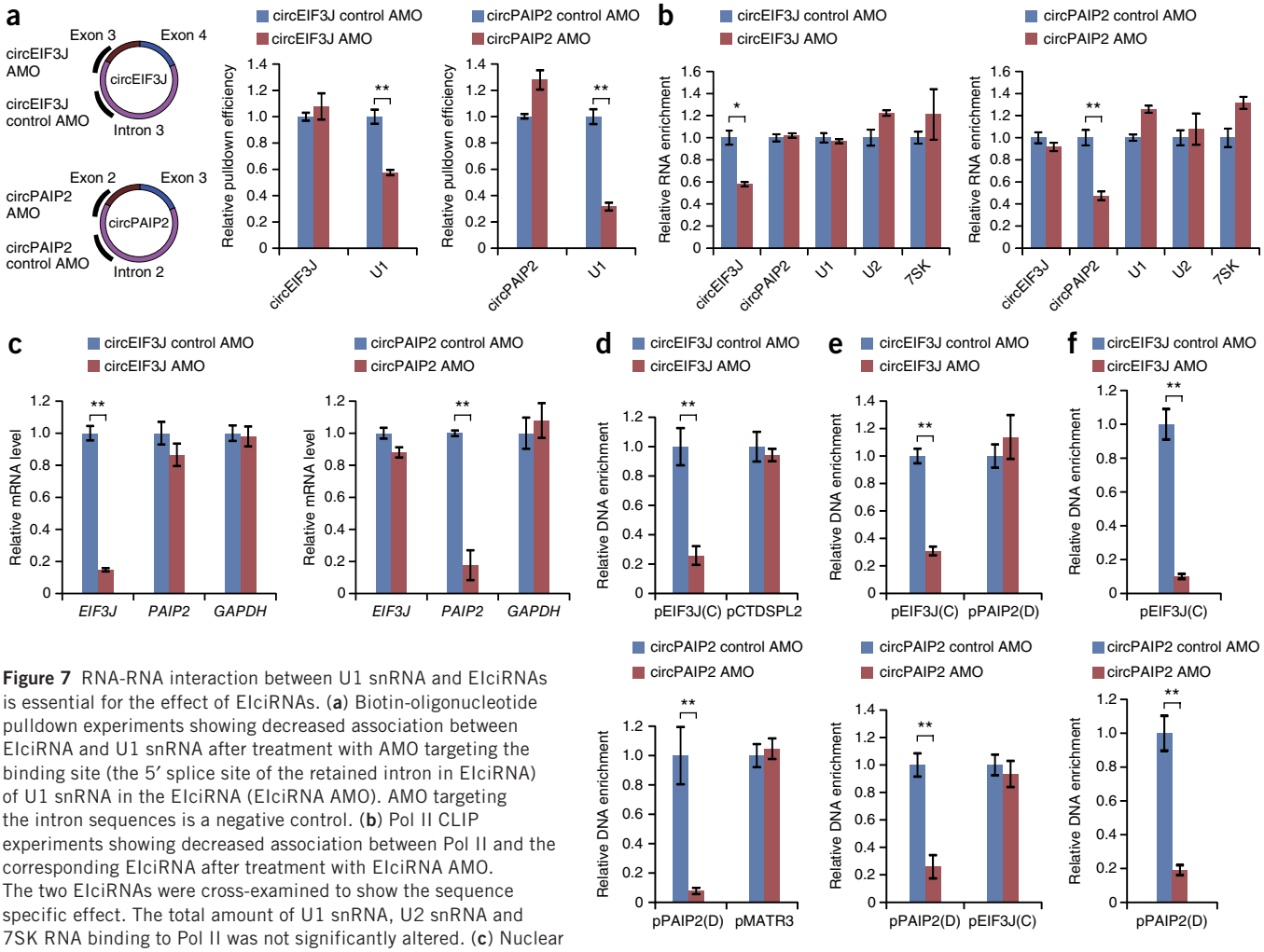

Figure 7 RNA-RNA interaction between U1 snRNA and EIciRNAs is essential for the effect of EIciRNAs. (**a**) Biotin-oligonucleotide pulldown experiments showing decreased association between EIciRNA and U1 snRNA after treatment with AMO targeting the binding site (the 5′ splice site of the retained intron in EIciRNA) of U1 snRNA in the EIciRNA (EIciRNA AMO). AMO targeting the intron sequences is a negative control. (**b**) Pol II CLIP experiments showing decreased association between Pol II and the corresponding EIciRNA after treatment with EIciRNA AMO. The two EIciRNAs were cross-examined to show the sequence specific effect. The total amount of U1 snRNA, U2 snRNA and 7SK RNA binding to Pol II was not significantly altered. (**c**) Nuclear run-on experiments showing decreased transcription of the corresponding parental gene after treatment with EIciRNA AMO. *GAPDH* is a negative control, and the two genes were cross-examined. (**d**) Pol II CHIP showing impeded binding of Pol II to the promoter of the corresponding parental gene after treatment with EIciRNA AMO. (**e**) U1C CHIP showing decreased binding of U1 snRNP to the promoter of the corresponding parental gene after treatment with EIciRNA AMO. The two genes were cross-examined. (**f**) EIciRNA ChIRP showing attenuation of the association between the EIciRNA and the parental gene promoter after treatment with the corresponding EIciRNA AMO. Throughout figure, error bars, s.e.m. from triplicate experiments. $^{*}P < 0.05$; $^{**}P < 0.01$ by two-tailed Student's *t* test.

gene promoters (**Fig. 6e**,**f** and **Supplementary Fig. 8a**,**b**). Pol II binding to the gene bodies of the parental genes was also decreased upon the knockdown of the corresponding circRNA (**Supplementary Fig. 8c**,**d**). In contrast, mRNA knockdown with siRNA had no effect on Pol II binding at the gene promoter (**Supplementary Fig. 8e**). FISH images showed that EIciRNA and U1 snRNA are concentrated around parental gene loci in more than half of the cells (**Fig. 6g** and **Supplementary Fig. 7b**). Together, these data indicate that U1 snRNP mediates the effects of EIciRNA on the expression of parental genes.

U1 snRNA-EIciRNA interaction is essential for the effect

We next examined where the U1 snRNA binding site is located in the EIciRNAs. On the basis of sequence complementarity, there might be only one U1 snRNA–binding site in each of the two EIciRNAs, potentially at the 5′ splice site of the retained intron. Sterically blocking this site with AMO decreased interaction between U1 snRNA and EIciRNA, as examined by EIciRNA pulldown with antisense biotin-labeled oligonucleotides (**Fig. 7a**). Interactions between Pol II and EIciRNA examined with Pol II CLIP also decreased upon AMO blocking of the U1-binding site (**Fig. 7b**). Furthermore, AMO blocking of the binding site of U1 snRNA in the corresponding EIciRNA decreased the transcription of the parental gene in nuclear run-on experiments (**Fig. 7c**). Consistently with the decreased transcription, the interactions between the parental gene promoter and Pol II were also decreased (**Fig. 7d**). Binding of U1 snRNP to the parental gene promoter, examined with ChIP, and interaction of EIciRNA with the parental gene promoter, examined with ChIRP, were all substantially attenuated by these AMO blockages (**Fig. 7e**,**f**). These data demonstrate that the binding of U1 snRNA to EIciRNA is critical for every aspect as well as the whole process of transcriptional enhancement by EIciRNAs. Thus the specific RNA-RNA interaction between U1 snRNA and EIciRNA via the U1-binding site in EIciRNA is essential for the transcription-enhancing effect of EIciRNAs.

DISCUSSION

Here we identified a special class of circRNAs as EIciRNAs (for example, circEIF3J and circPAIP2). EIciRNAs might hold factors such as U1 snRNP through RNA-RNA interaction between U1 snRNA and EIciRNA, and then the EIciRNA–U1 snRNP complexes might further interact with the Pol II transcription complex at the promoters of

parental genes to enhance gene expression (model in **Fig. 8**). We speculate that once the transcription of a gene is turned on, the generation of EIciRNA from the gene would further promote the gene's transcription, hence generating positive feedback. Regulation at the transcriptional level by ncRNAs is a fundamental aspect of gene expression[35,36]. The conventional function of U1 snRNA is splicing[37], but additional lines of evidence have shown that U1 snRNP has other roles, such as stimulating early transcriptional events, preventing premature polyadenylation and determining promoter directionality[10,38–42]. Specific RNA-RNA interaction between ncRNAs such as that between U1 snRNA and EIciRNA might be one of the central themes underlying the functional mechanism of ncRNAs[43]. However, our data also do not exclude other potential regulatory roles and mechanisms of EIciRNAs.

Biogenesis of circRNAs has been a recent hotspot of research. Our work (**Fig. 2**) and several publications have shown a link between flanking repeat sequences and biogenesis of circRNAs[30,32,33]. The internal sequences also contribute to circRNA biogenesis, as noticed by us (**Fig. 2b–d**) and also by others[33]. Whether circRNAs are generated cotranscriptionally or post-transcriptionally is still a topic under debate[33,44]. In our model, whether EIciRNAs are generated cotranscriptionally or post-transcriptionally is not an issue for their *cis* function, as long as they are generated at the spot of transcription. To our knowledge, researchers have not previously described any circRNA tissue or cell specificity similar to our observations for circCLTC (**Supplementary Fig. 1f**). The mRNA of the CLTC-encoding gene is expressed in both HeLa and HEK293 cells, although only HeLa cells express circCLTC. Beyond the nucleotide sequences, biogenesis of circRNAs seems to be regulated by some mechanisms with cell and/or gene specificity. It is possible that there is a relationship of competition between the linear mRNA and the circRNA regarding which one is generated from the single linear RNA precursor during splicing[44]. However, once generated, EIciRNAs may modulate the expression of the parental genes transcriptionally to increase levels of both circRNA and mRNA. The two EIciRNAs that we characterized in detail also showed some degree of heterogeneity in their amounts and distributions among individual cells (**Figs. 4**, **5g** and **6g**, and **Supplementary Fig. 7**), and this may indicate certain dynamic features of gene expression and gene regulation.

Recently, an argument has been raised that only very limited circRNAs (maybe just several of them) could act as microRNA sponges[29]. It has also been noted that, according to bioinformatic analysis, ~20% of the thousands of circRNAs in mammalian cells might have retained introns[29]. The majority of circRNAs have low abundance, and doubts have been raised regarding a biological function for most circRNAs[29]. For the *cis* effect, such as we present here for EIciRNAs, the abundance of individual circRNAs does not necessarily need to be high.

By showing that EIciRNAs promote the transcription of parental genes via interaction with U1 snRNP, we provide new insight into a gene-expression fine-tuning mechanism that works via RNA-RNA interaction. The identification of EIciRNAs in this study, together with circRNAs formed exclusively with either exonic or intronic sequences[27–29,45], suggests that there are at least three distinct circRNA populations in animal cells. Also, certain exonic sequences, which have been classically viewed as 'protein-coding' sequences, contribute to the formation of at least two types of 'noncoding' circular transcripts of exonic circRNAs and EIciRNAs. It is also fascinating that exon-only circRNAs may be involved in regulatory functions in the cytoplasm[27,28], whereas the EIciRNAs identified in this study appear to be efficiently retained for transcriptional regulation in the nucleus. Furthermore, we speculate that the functions and related mechanisms of circRNAs may be rather diverse, and therefore further studies will need be performed to rigorously explore the physiological roles of circRNAs.

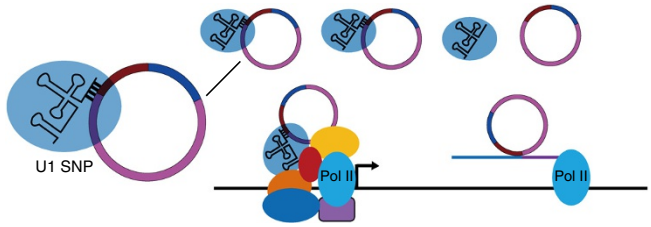

Figure 8 A working model for the *cis* effects of EIciRNAs on the expression of parental genes. EIciRNAs might hold factors such as U1 snRNP through specific RNA-RNA interaction between U1 snRNA and EIciRNA, and then the EIciRNA–U1 snRNP complexes might further interact with the Pol II transcription complex at the promoters of parental genes to enhance gene expression.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. RNA-sequencing data have been deposited in the Gene Expression Omnibus database under accession code GSE64443.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors thank S. Altman, Y. Shi, R. Chen, E. Wang, W.W. Walthall, P. Jin, S. Guang, X. Song, Q. Liu, Y. Mei and M. Wu for discussions and members of the Shan laboratory for discussions and technical support. This work was supported by grants to G.S. (the National Basic Research Program of China, 2011CBA01103 and 2015CB943000; the National Natural Science Foundation of China, 81171074, 91232702 and 31471225; the Chinese Academy of Sciences, 1731112304041 and KJZD-EW-L01-2; and the Fundamental Research Funds for the Central Universities of China, WK2070000034), Z.L. (the National Natural Science Foundation of China, 81372215 and 31301069) and Y.J. (the National Natural Science Foundation of China, 11175068).

AUTHOR CONTRIBUTIONS

G.S. conceived this project, designed experiments and supervised their execution. G.S. wrote the manuscript with the assistance of Z.L. and C.H. Z.L., C.H. and C.B. performed most of the experiments and analyzed most of the data. L.C., M.L., X.W., G.Z., B.Y., W.H., L.D., Y.J., P.X. and H.L. performed some of the experiments and (or) data analysis. X.W., P.Z., Z.C., Q.W. and Y.Z. performed bioinformatic analysis. All authors discussed the results and made comments on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Lee, J.T. Epigenetic regulation by long noncoding RNAs. *Science* **338**, 1435–1439 (2012).
2. Castel, S.E. & Martienssen, R.A. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat. Rev. Genet.* **14**, 100–112 (2013).
3. Ulitsky, I. & Bartel, D.P. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**, 26–46 (2013).
4. Hu, S., Wu, J., Chen, L. & Shan, G. Signals from noncoding RNAs: unconventional roles for conventional pol III transcripts. *Int. J. Biochem. Cell Biol.* **44**, 1847–1851 (2012).
5. Liu, H. *et al. Escherichia coli* noncoding RNAs can affect gene expression and physiology of *Caenorhabditis elegans*. *Nat. Commun.* **3**, 1073 (2012).

6. Meng, L.F., Chen, L., Li, Z.Y., Wu, Z.X. & Shan, G. Environmental RNA interference in animals. *Chin. Sci. Bull.* **58**, 4418–4425 (2013).
7. Fu, X. Non-coding RNA: a new frontier in regulatory biology. *Natl. Sci. Rev.* **1**, 190–204 (2014).
8. Wutz, A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat. Rev. Genet.* **12**, 542–553 (2011).
9. Rinn, J.L. *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311–1323 (2007).
10. Kwek, K.Y. *et al.* U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nat. Struct. Biol.* **9**, 800–805 (2002).
11. Yang, Z., Zhu, Q., Luo, K. & Zhou, Q. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**, 317–322 (2001).
12. Nguyen, V.T., Kiss, T., Michels, A.A. & Bensaude, O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414**, 322–325 (2001).
13. Ji, X. *et al.* SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* **153**, 855–868 (2013).
14. Názer, E. & Lei, E.P. Modulation of chromatin modifying complexes by noncoding RNAs *in trans*. *Curr. Opin. Genet. Dev.* **25**, 68–73 (2014).
15. Maamar, H., Cabili, M.N., Rinn, J. & Raj, A. linc-HOXA1 is a noncoding RNA that represses Hoxa1 transcription in *cis*. *Genes Dev.* **27**, 1260–1271 (2013).
16. Nagano, T. *et al.* The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**, 1717–1720 (2008).
17. Lai, F. *et al.* Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **494**, 497–501 (2013).
18. Hsu, M.T. & Coca-Prados, M. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* **280**, 339–340 (1979).
19. Nigro, J.M. *et al.* Scrambled exons. *Cell* **64**, 607–613 (1991).
20. Cocquerelle, C., Daubersies, P., Majerus, M.A., Kerckaert, J.P. & Bailleul, B. Splicing with inverted order of exons occurs proximal to large introns. *EMBO J.* **11**, 1095–1098 (1992).
21. Capel, B. *et al.* Circular transcripts of the testis: determining gene Sry in adult mouse testis. *Cell* **73**, 1019–1030 (1993).
22. Dixon, R.J., Eperon, I.C., Hall, L. & Samani, N.J. A genome-wide survey demonstrates widespread non-linear mRNA in expressed sequences from multiple species. *Nucleic Acids Res.* **33**, 5904–5913 (2005).
23. Shao, X., Shepelev, V. & Fedorov, A. Bioinformatic analysis of exon repetition, exon scrambling and trans-splicing in humans. *Bioinformatics* **22**, 692–698 (2006).
24. Suzuki, H. *et al.* Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res.* **34**, e63 (2006).
25. Al-Balool, H.H. *et al.* Post-transcriptional exon shuffling events in humans can be evolutionarily conserved and abundant. *Genome Res.* **21**, 1788–1799 (2011).
26. Salzman, J., Chen, R.E., Olsen, M.N., Wang, P.L. & Brown, P.O. Cell-type specific features of circular RNA expression. *PLoS Genet.* **9**, e1003777 (2013).
27. Memczak, S. *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**, 333–338 (2013).
28. Hansen, T.B. *et al.* Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384–388 (2013).
29. Guo, J.U., Agarwal, V., Guo, H. & Bartel, D.P. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.* **15**, 409 (2014).
30. Jeck, W.R. *et al.* Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19**, 141–157 (2013).
31. Dubin, R.A., Kazmi, M.A. & Ostrer, H. Inverted repeats are necessary for circularization of the mouse testis Sry transcript. *Gene* **167**, 245–248 (1995).
32. Zhang, X.O. *et al.* Complementary sequence-mediated exon circularization. *Cell* **159**, 134–147 (2014).
33. Liang, D. & Wilusz, J.E. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev.* **28**, 2233–2247 (2014).
34. Wheeler, T.M. *et al.* Targeting nuclear RNA for *in vivo* correction of myotonic dystrophy. *Nature* **488**, 111–115 (2012).
35. Goodrich, J.A. & Kugel, J.F. Non-coding-RNA regulators of RNA polymerase II transcription. *Nat. Rev. Mol. Cell Biol.* **7**, 612–616 (2006).
36. Geisler, S. & Coller, J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat. Rev. Mol. Cell Biol.* **14**, 699–712 (2013).
37. Moore, M.J. & Proudfoot, N.J. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136**, 688–700 (2009).
38. Damgaard, C.K. *et al.* A 5′ splice site enhances the recruitment of basal transcription initiation factors *in vivo*. *Mol. Cell* **29**, 271–278 (2008).
39. Furger, A., O'Sullivan, J.M., Binnie, A., Lee, B.A. & Proudfoot, N.J. Promoter proximal splice sites enhance transcription. *Genes Dev.* **16**, 2792–2799 (2002).
40. Almada, A.E., Wu, X., Kriz, A.J., Burge, C.B. & Sharp, P.A. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363 (2013).
41. Berg, M.G. *et al.* U1 snRNP determines mRNA length and regulates isoform expression. *Cell* **150**, 53–64 (2012).
42. Kaida, D. *et al.* U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–668 (2010).
43. Cech, T.R. & Steitz, J.A. The noncoding RNA revolution: trashing old rules to forge new ones. *Cell* **157**, 77–94 (2014).
44. Ashwal-Fluss, R. *et al.* circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* **56**, 55–66 (2014).
45. Zhang, Y. *et al.* Circular intronic long noncoding RNAs. *Mol. Cell* **51**, 792–806 (2013).

ONLINE METHODS

Cross-linking immunoprecipitation (CLIP) assay. CLIP was carried out as previously described with some modifications[46]. Briefly, the cultured cells were irradiated in a UV cross-linker (254 nm, 400 mJ/cm^2, 1 min) and then harvested in ice-cold lysis buffer (10 mM HEPES, pH 7.4, 200 mM NaCl, 30 mM EDTA and 0.5% Triton-X 100), 100 units/ml RNasin Plus RNase Inhibitor (Promega), 1.5 mM DTT, and 1× protease-inhibitor cocktail (Sangon). Cells were sonicated for 15 min with a Bioruptor (Diagenode), the cell suspension was centrifuged at 12,000*g* for 5 min at 4 °C, and the supernatant was collected and precleared with protein G–Sepharose 4 Fast Flow suspension (GE Amersham) for 1 h at 4 °C (Input). Anti–Pol II antibody (Abcam, cat. no. ab5095) or IgG (as control) was added to couple antigen for 1 h at 4 °C, and then Protein G–Sepharose 4 Fast Flow suspension was added to perform CLIP for at least 3 h at 4 °C. The antibody–Protein G bead complexes were washed five times with lysis buffer, and one-fifth volume of solution after the last wash was saved for western blots. Then the antibody–Protein G bead complexes were resuspended in 50 µl elution buffer (100 mM Tris, pH 7.8, 10 mM EDTA, and 1% SDS) and were reverse cross-linked with 30 µg of proteinase K at 65 °C for 1 h; this was followed with phenol/chloroform (pH 4.2) extraction to obtain RNA. To prepare CLIP RNA for deep sequencing, HeLa cells were pretreated with 20 µg/ml α-amanitin (Sigma) for 6 h before harvesting, and anti-RNA polymerase II CTD phospho-S2 antibody, ChIP grade (Abcam, cat. no. ab5095) was used for CLIP. α-amanitin was used to inhibit transcription, and it presumably would reduce the association of newly synthesized RNAs with Pol II. For later verification of enrichment of circRNAs, CLIP assays were performed without α-amanitin. To analyze whether the presence of genomic DNA was necessary for the association between Pol II and circRNAs, we added DNase I (Promega) to the cell suspension after sonication and then followed this with Pol II immunoprecipitation. For AGO2 CLIP, a monoclonal antibody to AGO2 from Sigma (cat. no. SAB4200085) was used. For these and all the other antibodies used, validation is provided on the manufacturers' websites.

RNA sequencing and bioinformatic analysis. For high-throughput sequencing, CLIP RNA samples were prepared according to the manufacturer's instructions and were applied to an Illumina GAIIx system for 80-nt single-end sequencing. In total, we obtained 8,975,835 and 5,088,901 reads from the control (IgG-CLIP) library and the sample (Pol II-CLIP) library, respectively. Reference genome hg19 was downloaded from the UCSC genome browser (http://genome.ucsc.edu/)[47]. Reads shorter than 50 nt were filtered out first, and the remaining reads were used to predict circRNAs according to the approach established by Rajewsky and colleagues[27]. In brief, we aligned reads to the reference genome and discarded reads that aligned contiguously and full length to the genome; from the remaining reads, we extracted 20-mers from both ends and aligned them independently to find unique anchor positions, and then we extended the anchor alignments to detect the breakpoints flanked by GU/AG splice sites.

Northern blot. Digoxin-labeled RNA probes were prepared with DIG Northern Starter Kit (Roche) with the corresponding PCR products as templates for T7 transcription according to the manual. Total RNA with or without RNase R digestion and RiboRuler High Range RNA Ladder (Thermo Scientific) were loaded on a 2% agarose gel containing 1% formaldehyde and were run for 1.5 h in MOPS buffer. In the O.E. lanes, total RNA was from overexpression with the circEIF3J_exon/intron or circPAIP2_exon/intron plasmids. RNA was transferred onto Hybond-N$^+$ membranes (GE Healthcare) by capillary transfer. Hybridization was performed at 62 °C overnight. Membranes were stringently washed twice in 0.1× SSC and 0.1% SDS at 62 °C for 30 min, and detection was performed according to the manual (Roche, DIG Northern Starter Kit). Images were taken with an ImageQuant LAS4000 Biomolecular Imager (GE Healthcare). Original images of northern blots can be found in **Supplementary Data Set 1**.

Plasmids and plasmid construction. All plasmids were constructed with restriction-enzyme digestion and ligation or with recombinant methods. Oligos for primers for all plasmid construction, probe preparation, siRNA, ASO, AMO, and biotin-oligo are listed in **Supplementary Table 4**. All plasmids were sequenced for confirmation. The shRNA plasmids for knockdown of *EIF3J* mRNA (shEIF3J-1, TRCN0000062016; shEIF3J-2, TRCN0000062017) and *PAIP2* mRNA (shPAIP2-1, TRCN0000219946; shPAIP2-2, TRCN0000153174) with negative shRNA control (SHC002) were obtained from the MISSION shRNA Library (Sigma). For the overexpression of circRNAs (**Fig. 2**), circEIF3J_exon and circPAPI2_exon are plasmids with inserts corresponding to the two exons forming the circRNA (with 5′-AG and 3′-GT included in the insertion). circEIF3J_exon/intron and circPAPI2_exon/intron are plasmids with insertions corresponding to the exon/intron sequences forming the circRNA (with 5′ AG and 3′ GT included in the insertion). circEIF3J_flanking and circPAPI2_flanking are plasmids with insertions corresponding to sequences forming the circRNA plus the 5′ and 3′ flanking genomic sequences. circEIF3J_1 kb and circPAPI2_1 kb are plasmids with insertions corresponding to sequences including the 5′ flanking genomic sequences (~1 kb), the sequences forming the circRNA, and reverse-complementary sequences of the 1-kb 5′ flanking sequence. The vector was pEGFP-C1 for these overexpression constructs. Further information about these plasmids can be obtained upon request.

Quantification of circRNA copy number per cell. DNA fragments corresponding to circEIFP3J or circPAIP2 were amplified with cDNA, and then purified dsDNA was used to plot a standard curve by real-time PCR. circRNA copy number in HeLa cells was calculated as follows. Total RNA was extracted from 1.0×10^6 cells, and cDNA was then synthesized. The copy number per cell of each circRNA was calculated on the basis of cell numbers and the Ct value from real-time PCR using the synthesized cDNA.

Fluorescence *in situ* hybridization (FISH). RNA probes were transcribed by the TranscriptAid T7 High Yield Transcription Kit (Thermo Scientific), with the corresponding insertion in the T vector as a template for transcription, and were labeled with Alexa Fluor546, Fluor488, or Fluor647 with the ULYSIS Nucleic Acid Labeling Kit (Invitrogen), which added a Fluor on every G of the probe to amplify the fluorescence intensity. Cells and RNA probes were denatured at 80 °C for 10 min and then incubated at 42 °C for 15–17 h with human Cot-1 DNA (Life Technologies, final concentration 30 ng/µl). Slides were washed with 2× SSC at 45 °C for 10 min. For DNA-RNA double FISH, DNA probes were prepared with genomic PCR and then labeled with a ULYSIS Nucleic Acid Labeling Kit; then hybridization was performed with the same conditions as for RNA FISH. For FISH with RNase R digestion, fixed cells were washed in PBS, treated with RNase R at 37 °C for 30 min and then fixed again before being subjected to the regular FISH protocol. FISH signal for circRNA was detected with junction probe if not specified.

Immunofluorescence staining. For β-actin immunostaining, slides after FISH were incubated with antibody against β-actin (Abcam, ab8227, 1:100 dilution) for 4 h; this was followed with incubation with Alexa Fluor555–labeled secondary antibody (Life Technologies, ab150074).

Confocal microscopy. Fluorescence signals were gathered with an Andor iXonEM+ DV897K EM CCD camera mounted on an Andor Revolution XD laser confocal microscope system (Andor Technology), with the Andor IQ 10.1 software.

Image processing and quantification. Cell images were processed with ImageJ image-acquisition software, and color channels were also merged by ImageJ. Quantification of band intensity from northern blotting was also performed with ImageJ. The area (same size as the band below) just above the band of measurement was used as background subtraction.

Cell culture and transfection of plasmids, siRNA, ASO, and AMO. HEK293 and HeLa cells were maintained under standard culture conditions with DMEM plus 10% FBS at 37 °C and 5% CO_2. Plasmid transfection was conducted with Lipofectamine 2000 (Invitrogen) according to the supplier's protocol. To make sure that the transfection efficiency was about the same, real-time PCR for plasmid DNA with *GAPDH* genomic DNA as an endogenous loading control was performed with the transfected cells. Transfection of siRNA was conducted with Lipofectamine 2000 or Oligofectamine (Invitrogen) according to the standard protocol. 2-*O*-methyl RNA/DNA antisense oligonucleotides (ASOs), which were modified by changing the five nucleotides at the 5′ and 3′ ends into 2′-*O*-methyl ribonucleotides, were synthesized (RiboBio). All bases of ASOs were converted into phosphorothioate oligonucleotides[31]. Antisense Morpholino oligonucleotides (AMOs), including U1 AMO, U2 AMO, and scrambled AMO, were synthesized at

Gene Tools. ASO and AMO treatments were performed with electroporation with the Nucleofector system (Lonza) according to the manufacturer's instructions. To minimize side effects of AMO, cells were harvested for analysis or downstream experiments 8 h after AMO transfection. To maximize the knockdown efficiency, ASO transfection was performed twice: 36 h after the first transfection, the second electroporation was performed, and this was followed with cell harvest 12 h later. The final concentrations were: ASO, 5 μM; U1 AMO and U2 AMO, 75 μM; and AMO against EIciRNAs and the corresponding controls, 7.5 μM.

Nuclear run-on assay. For nuclei isolation, cells were rinsed with PBS and harvested in ice-cold hypotonic solution (150 mM KCl, 4 mM MgOAc, and 10 mM Tris-HCl, pH 7.4) and were pelleted by centrifugation. Then pellets were resuspended in lysis buffer (150 mM KCl, 4 mM MgOAc, 10 mM Tris-HCl, pH 7.4, and 0.5% NP-40). The crude nuclei were then prepared by sucrose density gradient centrifugation. The nuclear run-on protocol was modified from Guang *et al.*[48]. The nuclear run-on mixture (10 mM ATP, CTP, GTP, BrUTP, and the crude nuclei) was incubated at 28 °C for 5 min in the presence of RNase inhibitor (Promega). The RNA was isolated by TRIzol reagent (Life Technologies) per the manufacturer's instructions, and DNA was eliminated by DNase I (Takara) treatment. Nascent transcripts were immunoprecipitated by anti-BrdU antibody (Abcam, cat. no. ab1893) and converted to cDNA for real-time PCR assay.

Western blot. For western blots, whole cell lysates and IP mixtures were separated on SDS-PAGE gels and then transferred to PVDF membranes (Millipore). Membranes were processed according to the ECL western blotting protocol (GE Healthcare). The following antibodies were used in western blots: anti-RNA polymerase II CTD phospho-S2 antibody (Abcam, cat. no. ab5095), anti-U1A (Santa Cruz Biotechnology, cat. no. SC-101149), anti-U1C (Santa Cruz Biotechnology, cat. no. SC-101549), anti-U2AF35 (Santa Cruz Biotechnology, cat. no. SC-19961), anti-U2AF65 (Santa Cruz Biotechnology, cat. no. SC-48804), anti-Lsm10 (Abcam, cat. no. ab180128) and anti-HDAC2 (Cell Signaling Technologies, cat. no. 5113). Images were taken with an ImageQuant LAS4000 Biomolecular Imager (GE Healthcare). Original images of western blots can be found in **Supplementary Data Set 1**.

Chromatin immunoprecipitation (ChIP). ChIP was carried out as described previously, with modifications[49]. Cells were fixed on plates with 1% formaldehyde for 10 min at room temperature. Cross-linking was then quenched with 0.125 M glycine for 5 min, and the cells were pelleted at 800*g*. Cell pellets were resuspended in 1 ml of SDS lysis buffer (1% (w/v) SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.1) containing complete protease inhibitor cocktail (Roche) and were incubated for 10 min on ice. Cell extracts were sonicated for 15 min with a Bioruptor (Diagenode) to obtain up to 500-bp DNA fragments. A 100-μl sample of the supernatant was saved as input. The remaining was diluted 1:10 in ChIP dilution buffer (0.01% (w/v) SDS, 1.1% (v/v) Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1, and 167 mM NaCl) containing protease inhibitors. The chromatin solution was precleared, immunoprecipitated with antibody to Pol II (Thermo Scientific, cat. no. 01671887), U1A (Santa Cruz Biotechnology, cat. no. SC-101149), U1C (Santa Cruz Biotechnology, cat. no. SC-101549), U2AF35 (Santa Cruz Biotechnology, cat. no. SC-19961), and U2AF65 (Santa Cruz Biotechnology, cat. no. SC-48804). The immune complexes were eluted in 1% (w/v) SDS and 50 mM $NaHCO_3$, and cross-links were reversed for 6 h at 65 °C. Samples were digested with proteinase K for 1 h at 45 °C, and the DNA was extracted with phenol/chloroform/isoamyl alcohol. Eluted DNA was subjected to quantitative real-time PCR (qPCR) for the detection of enriched genomic DNA regions with the corresponding PCR primer pairs.

Biotin-oligo pulldown of RNA. Biotin-oligo pulldown was carried out as previously described, with some modifications[50]. Briefly, log-phase cells were cross-linked with 1% glutaraldehyde or formaldehyde in PBS for 10 min at room temperature, and cross-linking was then quenched with 0.125 M glycine for 5 min. The cells were pelleted and resuspended in swelling buffer (0.1 M Tris, pH 7.0, 10 mM KOAc, and 15 mM MgOAc, with freshly added 1% NP-40, 1 mM DTT, complete protease inhibitor, and 0.1 U/μl RNase inhibitor) for 10 min on ice. Cell suspensions were then homogenized and pelleted at 2,500*g* for 5 min. Nuclei were further lysed in nuclear lysis buffer (50 mM Tris, pH 7.0, 10 mM EDTA, and 1% SDS; with freshly added 1 mM DTT, complete protease inhibitor, and 0.1 U/μl RNase inhibitor) on ice for 10 min and were sonicated until most chromatin had solubilized and DNA was in the size range of 100–500 bp. Chromatin was diluted in two times volume with hybridization buffer (750 mM NaCl, 1% SDS, 50 mM Tris, pH 7.0, 1 mM EDTA, 15% formamide, 1 mM DTT, protease inhibitor, and 0.1 U/μl RNase inhibitor). Biotin-DNA oligos (100 pmol) were added to 3 ml of diluted chromatin, which was mixed by end-to-end rotation at 37 °C for 4 h. M-280 Streptavidin Dynabeads (Life Technologies) were washed three times in nuclear lysis buffer, which was blocked with 500 ng/μl yeast total RNA and 1 mg/ml BSA for 1 h at room temperature, then washed three times again in nuclear lysis buffer before being resuspended. 100 μl washed/blocked Dynabeads was added per 100 pmol of biotin-DNA oligos, and the whole mix was then rotated for 30 min at 37 °C. Beads were captured by magnets (Life Technologies) and washed five times with 40× the volume of Dynabeads with wash buffer (2× SSC, 0.5% SDS, and 0.1 mM DTT and PMSF (fresh)). Beads were then subjected to RNA elution, DNA elution, or protein elution.

PCR reactions. Total RNA was extracted from cells with TRIzol reagent (Invitrogen) according to the manufacturer's procedures. Complementary DNA for quantitative reverse-transcription PCR (qRT-PCR) was synthesized from total RNA with the GoScript Reverse Transcription System (Promega) according to the supplied protocol, with random hexamer primers or oligo dT. Quantitative PCR (qPCR) was performed with Platinum SYBR Green qPCR Supermix UDG (Invitrogen) on a PikoReal Real-Time PCR System (Thermo Scientific) according to standard procedures. For semiquantitative PCR and semiquantitative RT-PCR, PCR cycle numbers were set between 20 and 25 to avoid saturation of PCR reactions. In real-time PCR data shown in the figures, ND (not detected) refers to a Ct value >35; Ct values for all inputs were 21–23.

Statistical analysis. The values reported in the graphs represent averages of three independent experiments or the actual number of cells stated in the figure legends, with error bars showing s.e.m., if applicable. After analysis of variance by *F* test, the statistical significance and *P* value were evaluated by Student's *t* test.

46. Luo, Y. *et al.* Fragile X mental retardation protein regulates proliferation and differentiation of adult neural stem/progenitor cells. *PLoS Genet.* **6**, e1000898 (2010).
47. Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
48. Guang, S. *et al.* Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. *Nature* **465**, 1097–1101 (2010).
49. Listerman, I., Sapra, A.K. & Neugebauer, K.M. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat. Struct. Mol. Biol.* **13**, 815–822 (2006).
50. Chu, C., Qu, K., Zhong, F.L., Artandi, S.E. & Chang, H.Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678 (2011).

doi:10.1038/nsmb.2959

Corrigendum: Exon-intron circular RNAs regulate transcription in the nucleus

Zhaoyong Li, Chuan Huang, Chun Bao, Liang Chen, Mei Lin, Xiaolin Wang, Guolin Zhong, Bin Yu, Wanchen Hu, Limin Dai, Pengfei Zhu, Zhaoxia Chang, Qingfa Wu, Yi Zhao, Ya Jia, Ping Xu, Huijie Liu & Ge Shan

***Nat. Struct. Mol. Biol.* 22, 256–264 (2015); published online 9 February 2015; corrected after print 19 August 2016**

In the version of this article initially published, the convergent primers depicted in the schematic in Figure 1b were incorrectly placed. The error has been corrected in the HTML and PDF versions of the article.