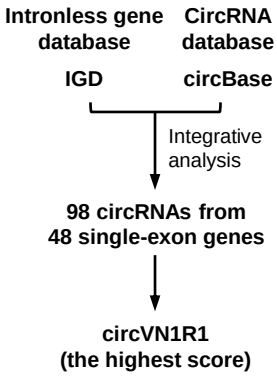
Molecular Cell, Volume *84*

Supplemental information

ZC3H14 facilitates backsplicing by binding to exon-intron boundary and 3′ UTR

Qiqi Li, Gang Yang, Bingbing Ren, Xu Liu, Li-Qin Tang, Qinghua Shi, Ge Shan, and Xiaolin Wang

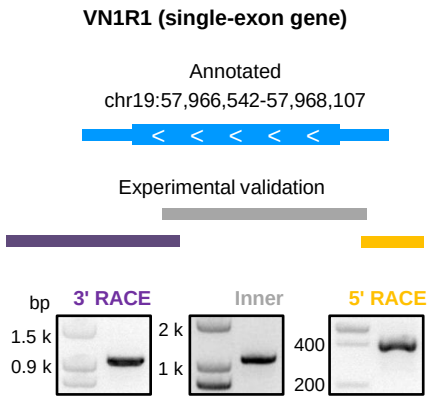
A



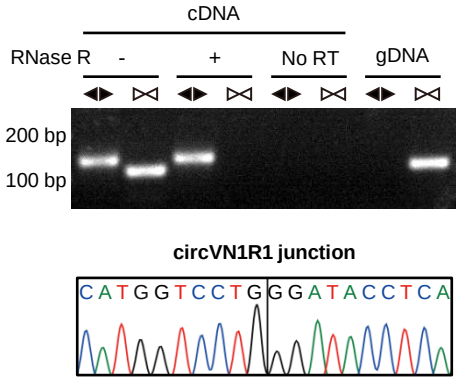
B

CircRNA ID	SE-gene	Position (hg19)	Strand	Length	Score
hsa_circ_0000963	VN1R1	chr19:57967020-57967550	-	530	239
hsa_circ_0001946	CDR1	chrX:139865339-139866824	+	1485	161
hsa_circ_0097431	TAS2R14	chr12:11273608-11276922	-	3314	49
hsa_circ_0001587	HIST1H2BE	chr6:26158779-26184405	-	25626	14
hsa_circ_0001583	HIST1H1C	chr6:26056122-26056299	-	177	13
hsa_circ_0097336	TAS2R14	chr12:11199618-11276922	-	77304	12
hsa_circ_0001593	HIST1H2AH	chr6:27114917-27115102	-	185	10
hsa_circ_0122427	RAP2B	chr3:152886168-152886425	+	257	9
hsa_circ_0001589	HIST1H1D	chr6:26234499-26234709	-	210	9
hsa_circ_0097432	TAS2R14	chr12:11273608-11281962	-	8354	9

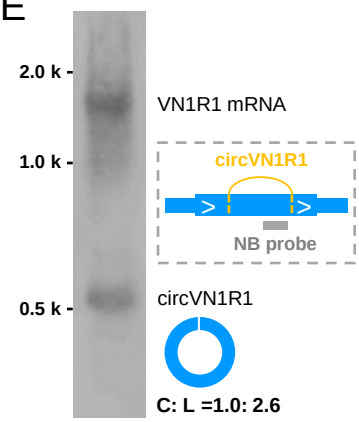
C



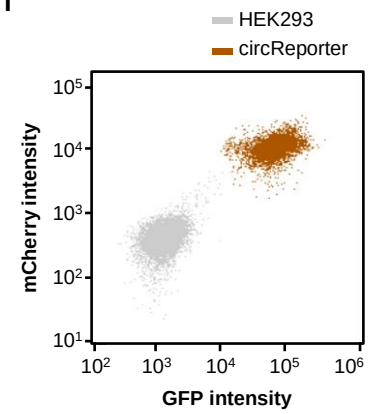
D



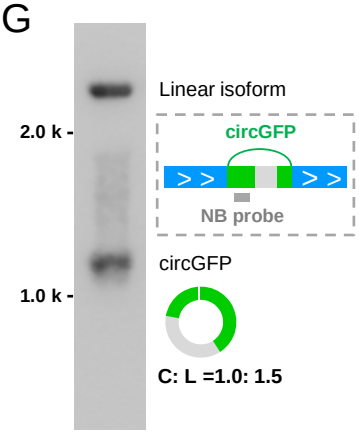
E



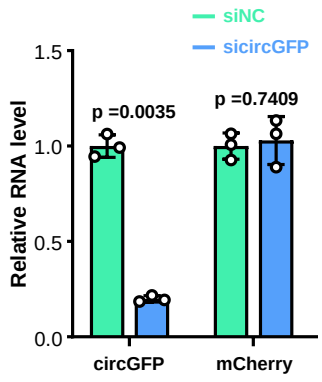
F



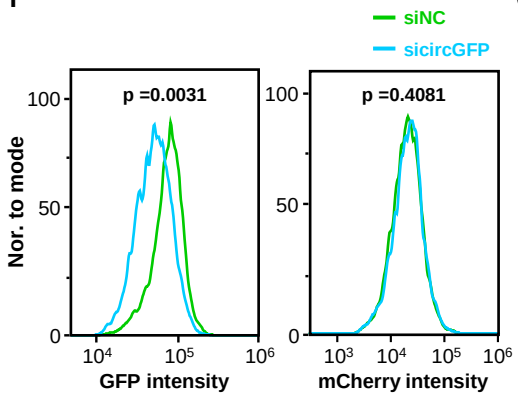
G



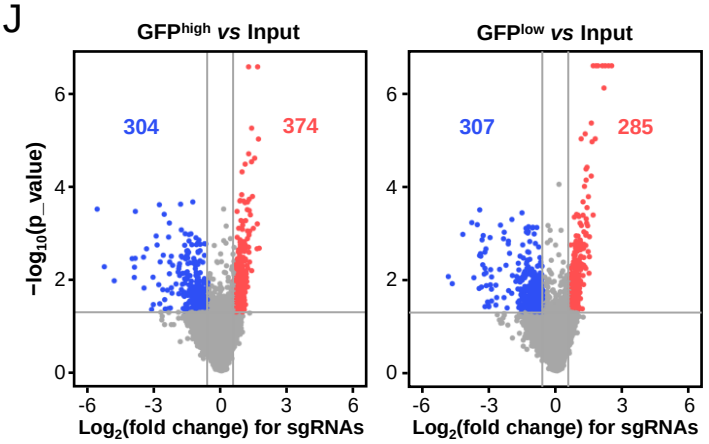
H



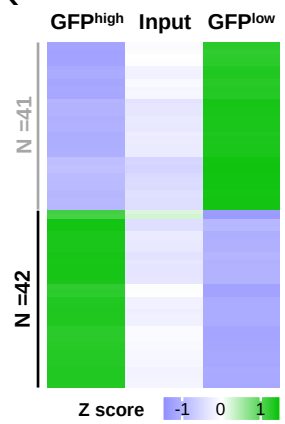
I



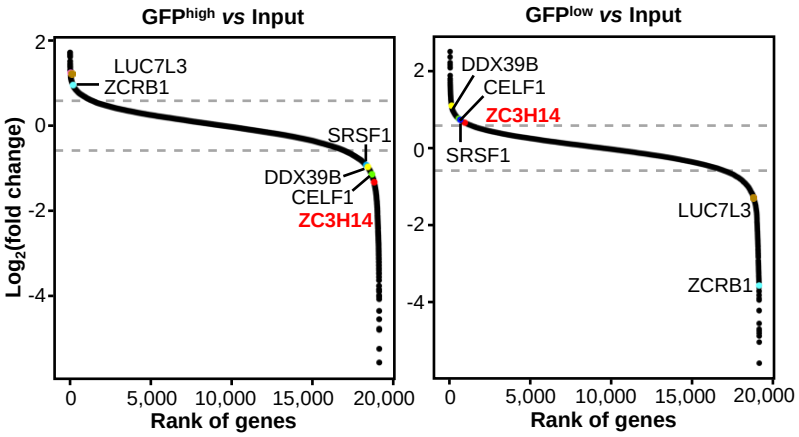
J



K



L



M

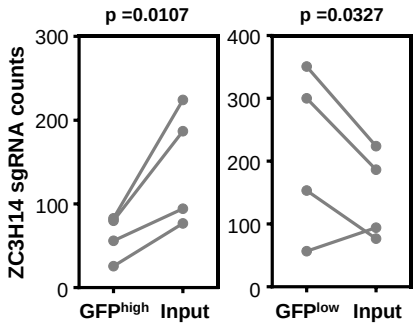


Figure S1. Construction of the circReporter and genome-wide CRISPR screen for regulators of circRNA biogenesis, related to Figure 1

(A) The workflow to identify circRNAs derived from single-exon genes. The human IGD contains 687 identified SE genes. For an individual circRNA, the "score" was defined as the average backsplicing junction (BSJ) reads from all RNA sequencing (RNA-seq) datasets collected in circBase.

(B) The list of 10 SE-gene-derived circRNAs with the top scores. Score for a circRNA was defined as the average BSJ from all RNA-seq datasets collected in circBase. SE, single-exon; BSJ, backsplicing junction.

(C) Rapid amplification of cDNA ends (RACE) of the VN1R1 mRNA in HEK293 cells. The 3' RACE, 5' RACE, and inner amplifications of VN1R1 were indicated in gel images.

(D) Amplification of *circVN1R1* with divergent primers followed by Sanger sequencing in HEK293 cells. cDNA, complementary DNA; gDNA, genomic DNA; RT, reverse transcription.

(E) Northern blotting for *circVN1R1* and *VN1R1* mRNA in HEK293 cells. The probe and the ratio of *circVN1R1* to linear *VN1R1* (C: L) were shown.

(F) GFP and mCherry intensities of circReporter cells examined by fluorescence-activated cell sorting (FACS).

(G) Northern blotting of *circGFP* and linear GFP in circReporter cells. The probe and the ratio of *circGFP* to linear *GFP* (C: L) were shown.

(H) RT-qPCR analysis demonstrating circGFP and mCherry RNA levels in circReporter cells transfected with siRNA against BSJ of circGFP (sicircGFP). Relative levels were normalized to 18S rRNA. Data are shown as mean ± SD from three independent experiments. siNC, negative control siRNA with scrambled sequences.

(I) GFP and mCherry intensities examined by flow cytometry in circReporter cells upon circGFP knockdown.

(J) Volcano plots showing regulators with differentially enriched/depleted sgRNAs (Effective sgRNAs ≥3, fold change >1.5 or <0.67, p-value <0.05) in GFP^{high} and GFP^{low} cells, compared to unsorted cells (Input). Blue and red dots represent regulators with significantly depleted and enriched sgRNAs, respectively.

(K) Heatmap demonstrating the distribution of sgRNAs for the 83 overlapped regulators in GFP^{high}, GFP^{low}, and Input cells.

(L) Ranks of 6 RBPs (ZC3H14, ZCRB1, SRSF1, LUC7L3, DDX39B, and CELF1) analyzed by MAGeCK in GFP^{high} and GFP^{low} cells.

(M) The ZC3H14 sgRNA counts in GFP^{high}, GFP^{low}, and Input cells. The p-values were generated by the MAGECK algorithm.

For E and G, the gray scale statistics of Northern blotting were performed by ImageJ. For H and I, p-values were calculated by two-tailed Student's *t*-test.

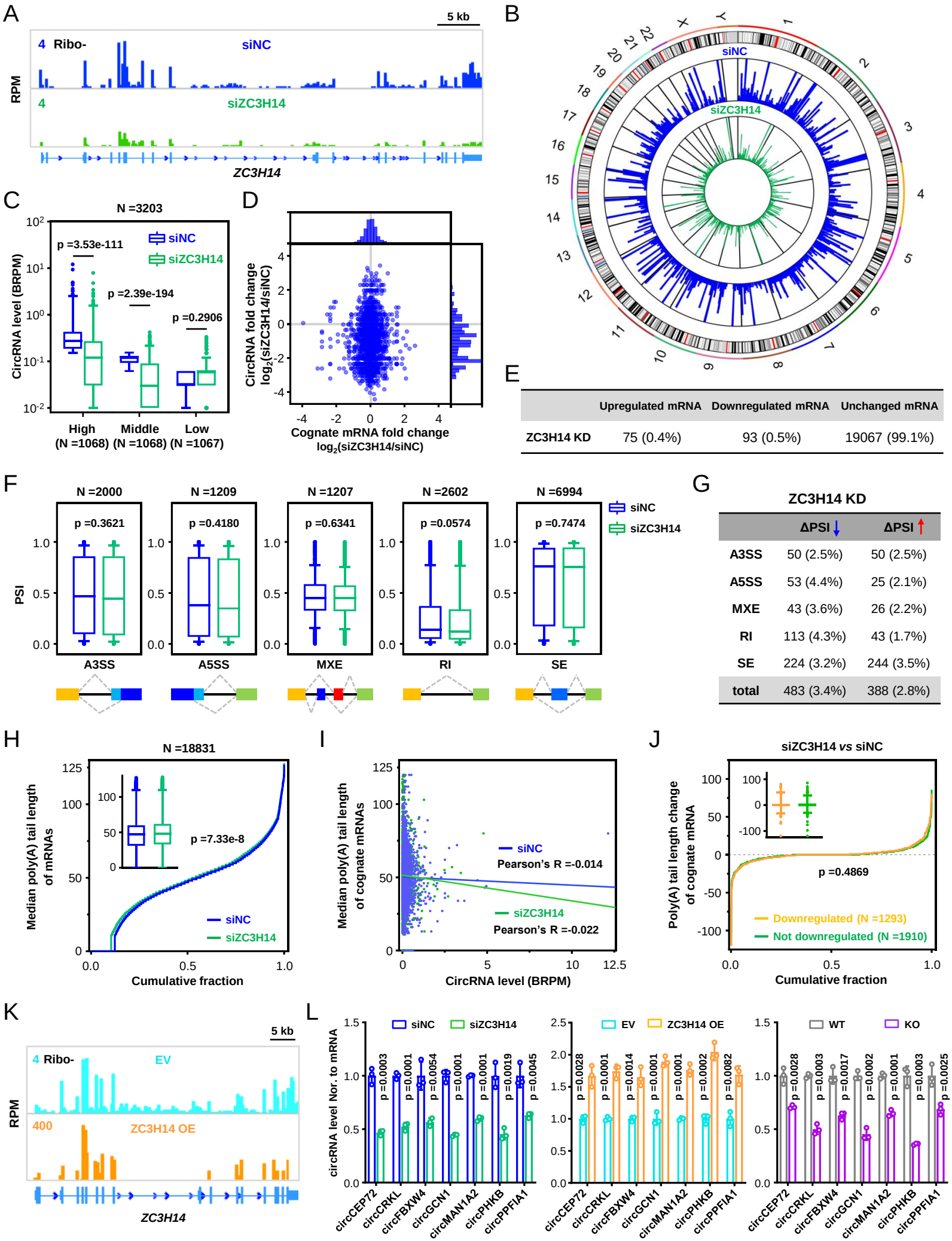

A
5 kb
4 Ribo-
siNC
RPM
4
siZC3H14
ZC3H14
B
siNC
siZC3H14
C
N =3203
p =3.53e-111
p =2.39e-194
p =0.2906
siNC
siZC3H14
CircRNA level (BRPM)
High (N =1068)
Middle (N =1068)
Low (N =1067)
D
CircRNA fold change log2(siZC3H14/siNC)
Cognate mRNA fold change log2(siZC3H14/siNC)
E
Upregulated mRNA
Downregulated mRNA
Unchanged mRNA
ZC3H14 KD
75 (0.4%)
93 (0.5%)
19067 (99.1%)
F
N =2000
N =1209
N =1207
N =2602
N =6994
p =0.3621
p =0.4180
p =0.6341
p =0.0574
p =0.7474
siNC
siZC3H14
PSI
A3SS
A5SS
MXE
RI
SE
G
ZC3H14 KD
ΔPSI ↓
ΔPSI ↑
A3SS 50 (2.5%) 50 (2.5%)
A5SS 53 (4.4%) 25 (2.1%)
MXE 43 (3.6%) 26 (2.2%)
RI 113 (4.3%) 43 (1.7%)
SE 224 (3.2%) 244 (3.5%)
total 483 (3.4%) 388 (2.8%)
H
N =18831
p =7.33e-8
siNC
siZC3H14
Median poly(A) tail length of mRNAs
Cumulative fraction
I
siNC
Pearson's R =-0.014
siZC3H14
Pearson's R =-0.022
Median poly(A) tail length of cognate mRNAs
CircRNA level (BRPM)
J
siZC3H14 vs siNC
p =0.4869
Downregulated (N =1293)
Not downregulated (N =1910)
Poly(A) tail length change of cognate mRNA
Cumulative fraction
K
5 kb
4 Ribo-
EV
RPM
400
ZC3H14 OE
ZC3H14
L
siNC
siZC3H14
EV
ZC3H14 OE
WT
KO
circRNA level Nor. to mRNA
circCEP72
circCRKL
circFBXW4
circGCN1
circMAN1A2
circPHKB
circPPFIA1

Figure S2. The effects of ZC3H14 on circRNA, mRNA, alternative splicing and poly(A)-tail length, related to Figure 2

(A) The wiggle-tracks of Ribo-minus RNA-seq revealing the ZC3H14 mRNA expression in HEK293 cells upon ZC3H14 knockdown.

(B) Circos plots showing the genomic distribution of circRNAs identified from Ribo-minus RNA-seq of HEK293 cells treated with siNC or siZC3H14. Human reference genome hg19 was used, and circRNAs (total BSJ ≥2) were chosen for analysis. The outer tracks represent the cytoband ideogram of the chromosome.

(C) The circRNA levels of distinct groups in HEK293 cells treated with siNC or siZC3H14. Three groups of circRNAs (High, Middle and Low) were evenly divided based on their ranked expression levels (BRPM) in siNC-treated cells. The BRPM ranges for High, Middle and Low groups are 0.15-12.13, 0.06-0.15 and 0.00-0.06, respectively.

(D) Scatterplots displaying fold changes in expressions of each circRNA (y-axis) and its cognate mRNA (x-axis) in HEK293 cells upon ZC3H14 knockdown. The individual distribution profiles of circRNA and mRNA are shown above and right of the scatterplots, respectively.

(E) Differentially expressed mRNAs in HEK293 cells upon ZC3H14 knockdown.

(F) The PSI of five alternative splicing events analyzed by rMATS upon ZC3H14 knockdown. N represents the number of each kind of alternative splicing we detected from poly(A)+ RNA-seq. A3SS, alternative 3' splice sites; A5SS, alternative 5' splice sites; MXE, mutually exclusive exons; RI, retained introns; SE, skipped exons; PSI, percent spliced in.

(G) Differentially expressed alternative splicing events in HEK293 cells upon ZC3H14 knockdown. The blue and red arrows represent the downregulated and upregulated alternative splicing events, respectively.

(H) Cumulative fraction curves and boxplots showing median poly(A)-tail length of mRNAs in HEK293 cells treated with siNC or siZC3H14. N represents the number of analyzed mRNAs.

(I) The correlation analysis of circRNA levels and median poly(A)-tail lengths of cognate mRNAs in siNC- and siZC3H14-treated HEK293 cells.

(J) Poly(A) length changes of cognate mRNAs for downregulated or not-downregulated circRNAs upon ZC3H14 knockdown. Downregulated circRNAs indicate circRNAs with significantly decreased levels upon ZC3H14 knockdown. Not-downregulated circRNAs indicate circRNAs with increased and unaltered levels upon ZC3H14 knockdown.

(K) The wiggle-tracks of Ribo-minus RNA-seq showing the ZC3H14 mRNA expression in HEK293 cells upon ZC3H14 overexpression (OE).

(L) RT-qPCR showing expression levels of 7 randomly selected high-reliable circRNAs in HEK293 cells upon ZC3H14 knockdown, OE, or KO. The circRNA levels were normalized to the corresponding cognate mRNAs. Data are shown as mean ± SD from three independent experiments. P-values were calculated by two-tailed Student's *t*-test.

For C, F, H and J, p-values were calculated by the Mann-Whitney U test.

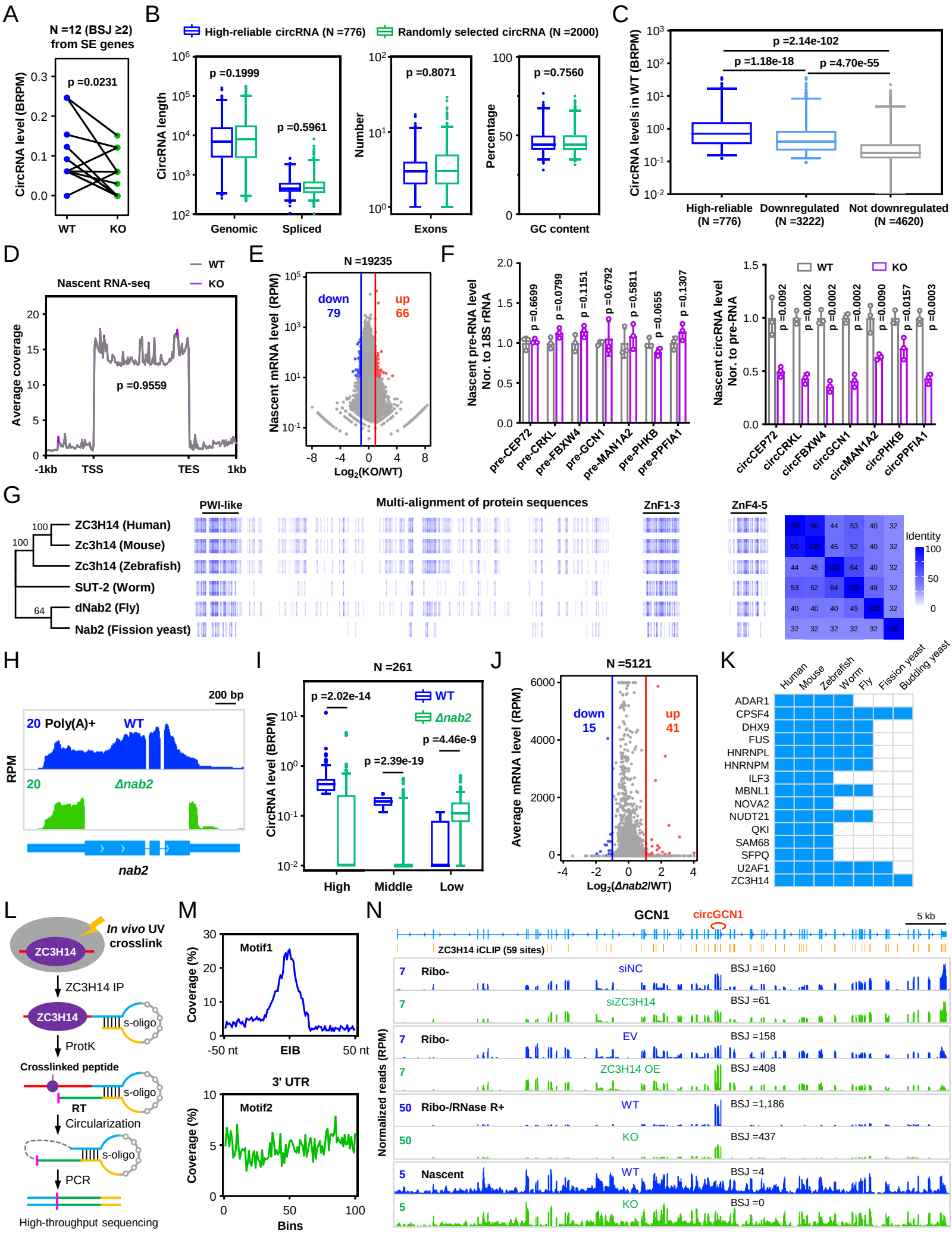

A
N =12 (BSJ ≥2)
from SE genes
CircRNA level (BRPM)
p =0.0231
WT
KO
B
High-reliable circRNA (N =776)
Randomly selected circRNA (N =2000)
CircRNA length
p =0.1999
p =0.5961
Genomic
Spliced
Number
p =0.8071
Exons
Percentage
p =0.7560
GC content
C
CircRNA levels in WT (BRPM)
p =2.14e-102
p =1.18e-18
p =4.70e-55
High-reliable (N =776)
Downregulated (N =3222)
Not downregulated (N =4620)
D
Nascent RNA-seq
WT
KO
Average coverage
p =0.9559
-1kb
TSS
TES
1kb
E
N =19235
down 79
up 66
Nascent mRNA level (RPM)
Log2(KO/WT)
F
Nascent pre-RNA level Nor. to 18S rRNA
p =0.6699
p =0.0799
p =0.1151
p =0.6792
p =0.5811
p =0.0655
p =0.1307
pre-CEP72
pre-CRKL
pre-FBXW4
pre-GCN1
pre-MAN1A2
pre-PHKB
pre-PPFIA1
Nascent circRNA level Nor. to pre-RNA
WT
KO
p =0.0092
p =0.0002
p =0.0002
p =0.0002
p =0.0090
p =0.0157
p =0.0003
circCEP72
circCRKL
circFBXW4
circGCN1
circMAN1A2
circPHKB
circPPFIA1
G
PWI-like
Multi-alignment of protein sequences
ZnF1-3
ZnF4-5
ZC3H14 (Human)
Zc3h14 (Mouse)
Zc3h14 (Zebrafish)
SUT-2 (Worm)
dNab2 (Fly)
Nab2 (Fission yeast)
Identity
H
200 bp
20 Poly(A)+
WT
20
Δnab2
RPM
nab2
I
N =261
p =2.02e-14
p =2.39e-19
p =4.46e-9
WT
Δnab2
CircRNA level (BRPM)
High
Middle
Low
J
N =5121
down 15
up 41
Average mRNA level (RPM)
Log2(Δnab2/WT)
K
Human
Mouse
Zebrafish
Worm
Fly
Fission yeast
Budding yeast
ADAR1
CPSF4
DHX9
FUS
HNRNPL
HNRNPM
ILF3
MBNL1
NOVA2
NUDT21
QKI
SAM68
SFPQ
U2AF1
ZC3H14
L
In vivo UV crosslink
ZC3H14
ZC3H14 IP
s-oligo
ProtK
Crosslinked peptide
RT
Circularization
PCR
High-throughput sequencing
M
Motif1
Coverage (%)
-50 nt
EIB
50 nt
3' UTR
Motif2
Bins
N
GCN1
circGCN1
5 kb
ZC3H14 iCLIP (59 sites)
Ribo-
siNC
BSJ =160
siZC3H14
BSJ =61
EV
BSJ =158
ZC3H14 OE
BSJ =408
Ribo-/RNase R+
WT
BSJ =1,186
KO
BSJ =437
Nascent
BSJ =4
BSJ =0
Normalized reads (RPM)

Figure S3. ZC3H14 regulates circRNA biogenesis in mammalian cells and yeast, related to Figures 2 and 3

(A) Expression levels of SE-gene-derived circRNAs in Ribo-minus /RNase R+ RNA-seq of ZC3H14 WT and KO HEK293 cells.

(B) The comparison of genomic length, spliced length, exon number, and GC content for high-reliable and randomly selected circRNAs.

(C) The abundance comparison of three circRNA groups in WT HEK293 cells. Downregulated indicates significantly decreased circRNAs upon ZC3H14 KO. Not downregulated indicates unaltered and upregulated circRNAs upon ZC3H14 KO. P-values were calculated by one-way analysis of variance (ANOVA).

(D) The nascent RNA-seq coverage showing transcriptional curve in WT and ZC3H14 KO HEK293 cells. TSS, transcription start sites; TES, transcription end sites.

(E) Differentially expressed nascent mRNAs in HEK293 cells upon ZC3H14 KO. Blue and Red dots represent significantly downregulated and upregulated nascent mRNAs, respectively.

(F) RT-qPCR showing nascent levels of 7 randomly selected high-reliable circRNAs and their cognate pre-RNAs in HEK293 cells upon ZC3H14 KO. Data are shown as mean ± SD from three independent experiments.

(G) The conservation analysis for ZC3H14 across species. Phylogenetic diagram of ZC3H14 protein in multiple species (left). The amino acid sequences of ZC3H14 protein from distinct species were used to generate a phylogenetic tree with MEGA 7.0 using the Neighbor-Joining (NJ) method. Numbers on branches indicate the NJ bootstrap support values. A multi-alignment of ZC3H14 protein sequences analyzed by ClustalW (middle). The conserved domains of ZC3H14 were shown. The identity of ZC3H14 protein sequences across species was shown to the right.

(H) The wiggle-tracks of poly(A)+ RNA-seq revealing the *nab2* mRNA expression in *nab2* KO *S. pombe* strain.

(I) The circRNA levels of distinct groups in WT and Nab2 KO *S. pombe* strains. Three groups (High, Middle and Low) were evenly divided based on circRNA levels in WT *S. pombe* strain. The BRPM ranges for High, Middle and Low groups are 0.28-11.82, 0.12-0.28 and 0.00-0.12, respectively.

(J) Differentially expressed mRNAs in Nab2 KO *S. pombe* strain. Blue and Red dots represent downregulated and upregulated mRNAs, respectively.

(K) Conservation analysis of the 14 known regulators and ZC3H14 involved in circRNA biogenesis across species. Blue box indicates the presence of homologue, and empty box presents the absence of homologue.

(L) Illustration of the experimental procedure of ZC3H14 iCLIP-seq.

(M) The coverage profile of ZC3H14 binding Motif1 at up/down-stream 50 nt of EIBs (top). The coverage profile of ZC3H14 binding Motif2 in 3' UTR (bottom).

(N) An example of circRNA (*circGCN1*, with the highest BSJ in WT HEK293 cells) generated from *GCN1*. Genomic structure of the *GCN1* gene and ZC3H14 iCLIP binding sites across the *GCN1* gene were shown.

For A and F, p-values were calculated by two-tailed Student's *t*-test. For B, D, and I, p-values were calculated by the Mann-Whitney U test.

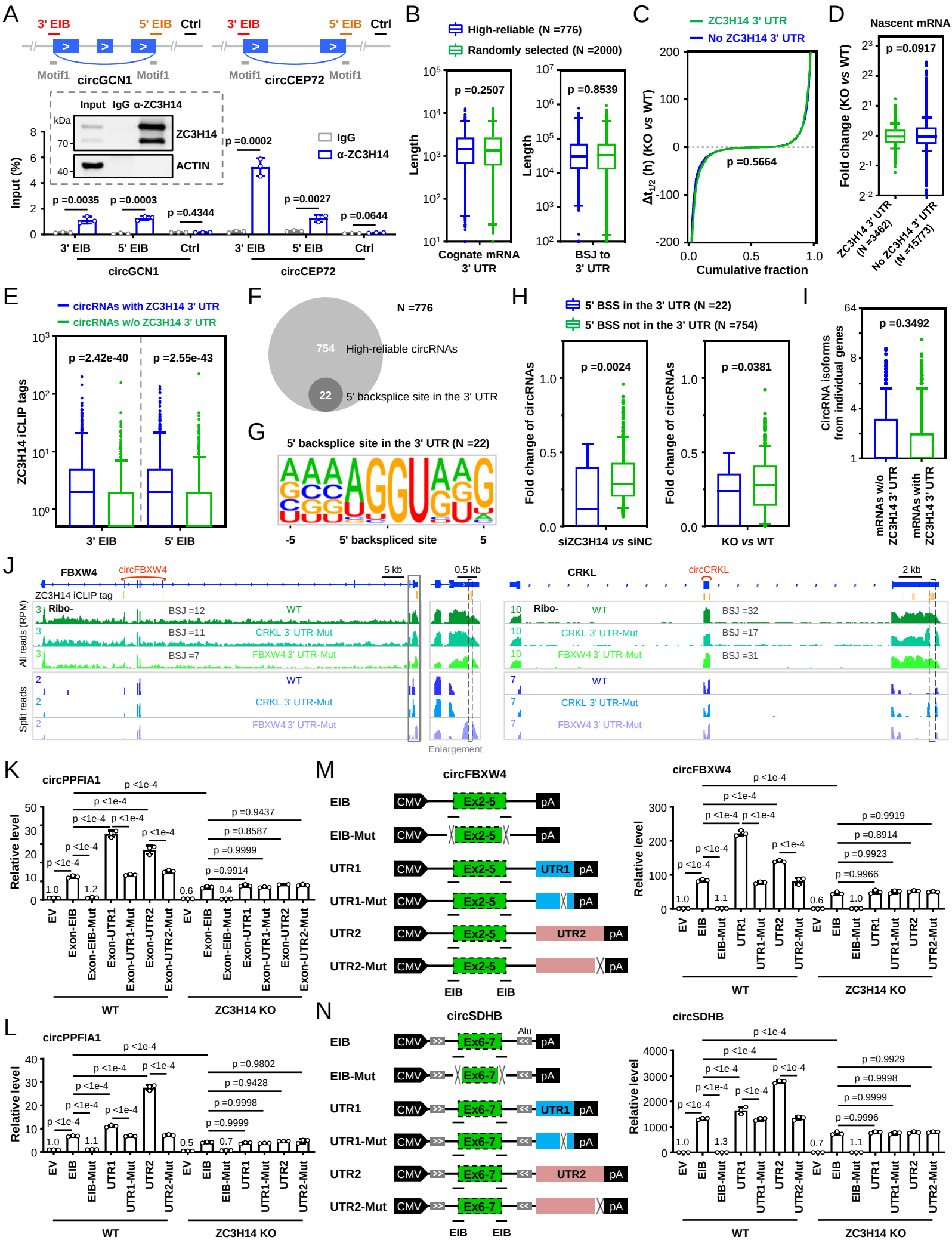

A
3' EIB
5' EIB
Ctrl
Motif1
circGCN1
circCEP72
Input IgG α-ZC3H14
kDa
ZC3H14
ACTIN
Input (%)
p =0.0035
p =0.0003
p =0.4344
p =0.0002
p =0.0027
p =0.0644
IgG
α-ZC3H14
B
High-reliable (N =776)
Randomly selected (N =2000)
p =0.2507
p =0.8539
Length
Cognate mRNA 3' UTR
BSJ to 3' UTR
C
ZC3H14 3' UTR
No ZC3H14 3' UTR
p =0.5664
Δt1/2 (h) (KO vs WT)
Cumulative fraction
D
Nascent mRNA
p =0.0917
Fold change (KO vs WT)
ZC3H14 3' UTR (N =3462)
No ZC3H14 3' UTR (N =15773)
E
circRNAs with ZC3H14 3' UTR
circRNAs w/o ZC3H14 3' UTR
p =2.42e-40
p =2.55e-43
ZC3H14 iCLIP tags
F
N =776
754
High-reliable circRNAs
22
5' backsplice site in the 3' UTR
G
5' backsplice site in the 3' UTR (N =22)
-5
5' backspliced site
5
H
5' BSS in the 3' UTR (N =22)
5' BSS not in the 3' UTR (N =754)
p =0.0024
p =0.0381
Fold change of circRNAs
siZC3H14 vs siNC
KO vs WT
I
p =0.3492
CircRNA isoforms from individual genes
mRNAs w/o ZC3H14 3' UTR
mRNAs with ZC3H14 3' UTR
J
FBXW4
circFBXW4
5 kb
0.5 kb
CRKL
circCRKL
2 kb
ZC3H14 iCLIP tag
All reads (RPM)
Split reads
Ribo-
WT
BSJ =12
CRKL 3' UTR-Mut
BSJ =11
FBXW4 3' UTR-Mut
BSJ =7
BSJ =32
BSJ =17
BSJ =31
Enlargement
K
circPPFIA1
Relative level
EV
Exon-EIB
Exon-EIB-Mut
Exon-UTR1
Exon-UTR1-Mut
Exon-UTR2
Exon-UTR2-Mut
WT
ZC3H14 KO
M
circFBXW4
EIB
EIB-Mut
UTR1
UTR1-Mut
UTR2
UTR2-Mut
CMV
Ex2-5
pA
L
circPPFIA1
N
circSDHB
Ex6-7
Alu

Figure S4. ZC3H14 binds with exon-intron boundaries and 3' UTRs, related to Figures 3 and 4

(A) ZC3H14 RIP followed by RT-qPCR analysis revealing the association of 3'&5' EIBs for *circGCN1* and *circCEP72* with ZC3H14 in HEK293 cells. Ctrl, intronic region without ZC3H14 iCLIP signals. P-values were calculated by two-tailed Student's *t*-test.

(B) Boxplots showing 3' UTR length of cognate mRNAs (left) and distance of BSJ to 3' UTR (right) between high-reliable and randomly selected circRNAs.

(C) Comparison of half-life changes of mRNAs with or without ZC3H14 3' UTR upon ZC3H14 KO.

(D) The comparison of changes in nascent mRNA levels between those with ZC3H14 3' UTR and without ZC3H14 3' UTR upon ZC3H14 KO.

(E) The comparison of ZC3H14 iCLIP tags in 3' and 5' EIBs of circRNAs between cognate mRNAs with ZC3H14 3' UTR and those without ZC3H14 3' UTR.

(F) Venn diagrams showing the high-reliable circRNAs and circRNAs with 5' backsplice site in 3' UTR.

(G) Motif analysis of 5 nt up/down-stream of the 5' BSS in cognate mRNA 3' UTRs of the 22 circRNAs.

(H) Comparison of expression changes between circRNAs with 5' backsplice site (BSS) in 3' UTR and the other high-reliable circRNAs upon ZC3H14 knockdown or KO.

(I) Comparison of circRNA isoforms from individual genes between cognate mRNAs with ZC3H14 binding 3' UTR and those without ZC3H14 binding 3' UTR.

(J) Schematic illustration of *FBXW4* (left) and *CRKL* (right) genes. Genomic structure of *FBXW4*/*CRKL* gene and ZC3H14 iCLIP tags were shown (top). The wiggle-tracks of Ribo-minus RNA-seq revealing the expression of *FBXW4*/*CRKL* mRNA and *circFBXW4*/*circCRKL* in WT cells, CRKL 3' UTR-Mut cells, and FBXW4 3' UTR-Mut cells. The isoforms generated from *FBXW4 and CRKL* genes were shown using the split reads (bottom). The dotted lines indicate the deleted regions in the corresponding 3' UTRs.

(K) Overexpression of *circPPFIA1* with six constructs (corresponding to Figure 4B) in WT and ZC3H14 KO cells. *CircPPFIA1* levels were normalized to mRNA of Neomycin gene (the resistance gene of plasmids).

(L) Overexpression of *circPPF1A1* normalized to Neomycin mRNA with six constructs (corresponding to Figure 4C) in WT and ZC3H14 KO cells examined by RT-qPCR.

(M) Overexpression of *circFBXW4* normalized to Neomycin mRNA with six constructs in WT and ZC3H14 KO cells. Intronic flanking sequences (~300 bp) were included in the constructs. Solid lines above the label of EIB represent the ZC3H14 binding EIBs. The EIB mutation and depletion in 3' UTR were also marked.

(N) Overexpression of *circSDHB* normalized to Neomycin mRNA with six constructs in WT and ZC3H14 KO cells. Part of endogenous flanking sequences including the complementary *Alu* element pairs were included. Solid lines above the label of EIB represent the ZC3H14 binding EIBs.

For A and K-N, data are shown as mean ± SD from three independent experiments. For B, C, D, E, H, and I, p-values were calculated by the Mann-Whitney U test. For K-N, p-values were calculated by one-way analysis of variance (ANOVA).

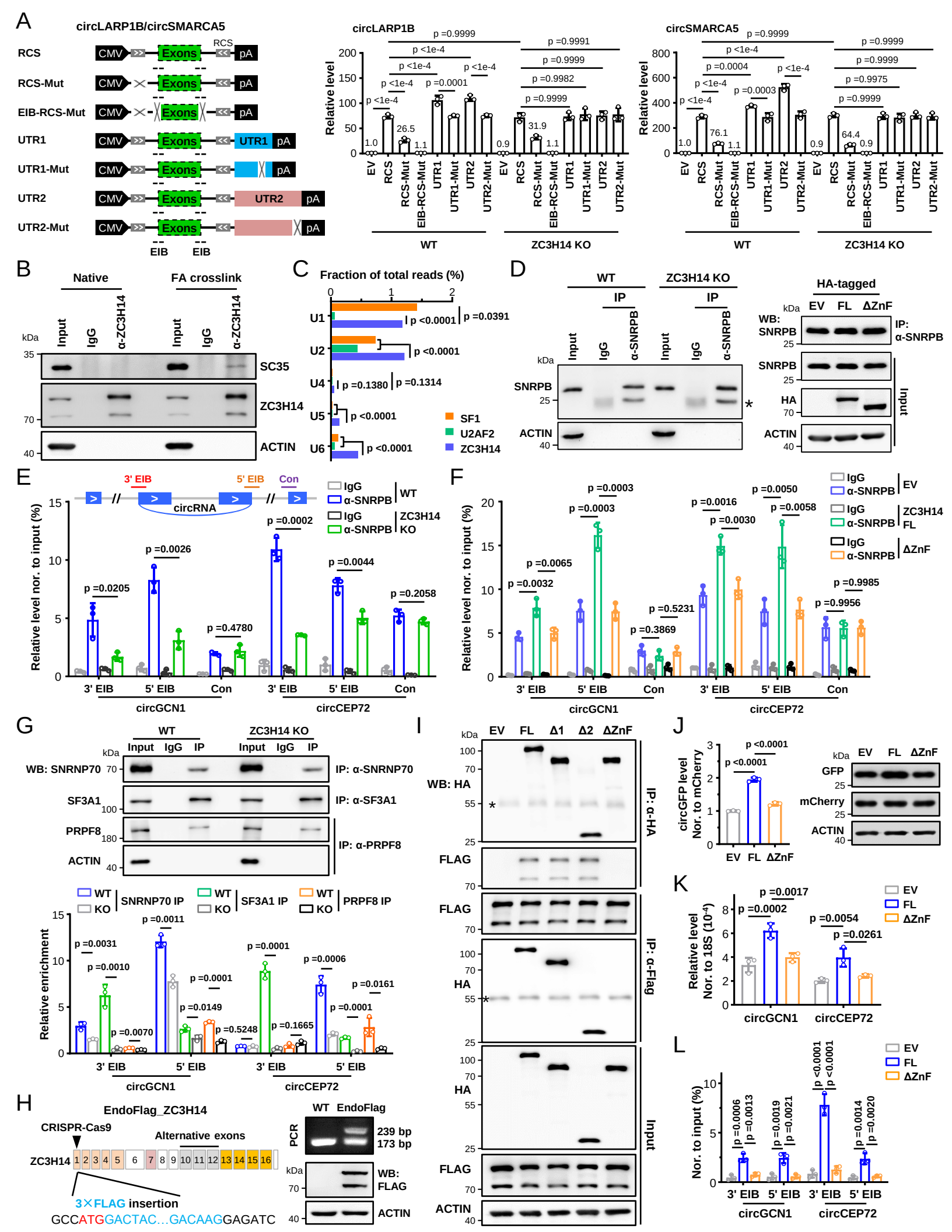

A
circLARP1B/circSMARCA5
RCS
RCS-Mut
EIB-RCS-Mut
UTR1
UTR1-Mut
UTR2
UTR2-Mut
CMV
Exons
pA
EIB
circLARP1B
circSMARCA5
Relative level
WT
ZC3H14 KO
B
Native
FA crosslink
Input
IgG
α-ZC3H14
SC35
ZC3H14
ACTIN
C
Fraction of total reads (%)
U1
U2
U4
U5
U6
SF1
U2AF2
ZC3H14
D
IP
α-SNRPB
SNRPB
HA-tagged
EV
FL
ΔZnF
WB: SNRPB
IP: α-SNRPB
HA
ACTIN
E
3' EIB
5' EIB
Con
circRNA
IgG
α-SNRPB
WT
ZC3H14 KO
Relative level nor. to input (%)
circGCN1
circCEP72
F
ZC3H14 FL
G
WT
ZC3H14 KO
Input
IgG
IP
WB: SNRNP70
SF3A1
PRPF8
IP: α-SNRNP70
IP: α-SF3A1
IP: α-PRPF8
SNRNP70 IP
SF3A1 IP
PRPF8 IP
Relative enrichment
H
EndoFlag_ZC3H14
CRISPR-Cas9
Alternative exons
ZC3H14
3×FLAG insertion
GCCATGGACTAC…GACAAGGAGATC
WT EndoFlag
PCR
239 bp
173 bp
WB: FLAG
ACTIN
I
EV
FL
Δ1
Δ2
ΔZnF
WB: HA
FLAG
IP: α-HA
IP: α-Flag
Input
J
circGFP level Nor. to mCherry
GFP
mCherry
ACTIN
K
Relative level Nor. to 18S (10⁻⁴)
L
Nor. to input (%)

Figure S5. ZC3H14 interacts with spliceosome and dimerizes, related to Figures 4 and 5

(A) Overexpression of *circLARP1B* and *circSMARCA5* with seven constructs in WT and ZC3H14 KO cells. Part of endogenous flanking sequences including RCSs were included. For *circLARP1B*, RCS is annotated *Alu* element. Dotted lines above the label of EIB represent the EIBs not bound by ZC3H14. CircRNA levels were normalized to Neomycin mRNA. RCS, reverse complementary sequences.

(B) The interaction between ZC3H14 and SC35 under native and formaldehyde (FA) crosslinking conditions.

(C) Fraction of total ZC3H14 iCLIP reads mapping to snRNAs. U2AF2 and SF1 (iCLIP data: GSE220186) are positive controls that bind to snRNAs. P values were calculated by Chi-square test.

(D) SNRPB IP in WT and ZC3H14 KO cells. The bands indicated by the asteroid were the light chain of the antibody (left). SNRPB IP in HEK293 cells transfected with HA-tagged full-length (FL) or ZnF-truncated (ΔZnF) ZC3H14 (right).

(E) SNRPB RIP followed by RT-qPCR analysis revealing the bindings of spliceosome to 3' and 5' EIBs of *circGCN1* and *circCEP72* upon ZC3H14 KO in HEK293 cells. The anti-SNRPB antibody is well-accepted to pull down spliceosome. Primers used for RT-qPCR were indicated for circularized exons (top). Con (control) represents EIB around non-circularized exon.

(F) SNRPB RIP followed by RT-qPCR analysis in HEK293 cells overexpressed with full-length or truncated (ΔZnF) ZC3H14 proteins demonstrating the bindings of spliceosome to 3' and 5' EIBs of *circGCN1* and *circCEP72*.

(G) SNRNP70, SF3A1, and PRPF8 RIP followed by RT-qPCR analysis revealing their bindings to 3' and 5' EIBs of *circGCN1* and *circCEP72* upon ZC3H14 KO in HEK293 cells. The IP efficiency was examined by western blotting.

(H) Construction of EndoFlag_ZC3H14 HEK293 cell line mediated by CRISPR-Cas9. The 3×FLAG sequence (blue) was inserted behind the start codon (ATG) of ZC3H14 (left). Validation of the heterozygous EndoFlag_ZC3H14 HEK293 cells (right).

(I) Co-IP assays of EndoFlag_ZC3H14 HEK293 cells transfected with HA-tagged full-length or truncated ZC3H14 plasmids. Full-length and truncated forms of ZC3H14 were determined by western blots with anti-HA or anti-FLAG antibodies. ACTIN was a loading control for western blot. Δ1, Δ2 and Δ3 represent ZC3H14 proteins with the deletions of PWI-like domain, middle region and zinc finger domain, respectively. The bands indicated by the asteroid were the heavy chain of the antibody.

(J) The *circGFP* RNA (normalized to *mCherry* RNA) and GFP protein levels in circReporter cells transfected with indicated constructions. ACTIN was a loading control for western blotting. EV, empty vector; FL, full-length ZC3H14; ΔZnF, ZC3H14 with zinc finger domain deleted.

(K) The expression levels of *circGCN1* and *circCEP72* (normalized to 18S rRNA) in HEK293 cells transfected with indicated constructions. EV, empty vector; FL, full-length ZC3H14; ΔZnF, ZC3H14 with zinc finger domain deleted.

(L) RNA immunoprecipitation (RIP) of full-length (FL) or ΔZnF ZC3H14 followed by RT-qPCR analysis revealing the association of 3' & 5' EIBs for *circGCN1* and *circCEP72* with FL or ΔZnF ZC3H14 in HEK293 cells.

For A, E, F, G and J-L, data are shown as mean ± SD from three independent experiments. For E and G, p-values were calculated by two-tailed Student's *t*-test. For A, F and J-L, p-values were calculated by one-way analysis of variance (ANOVA).

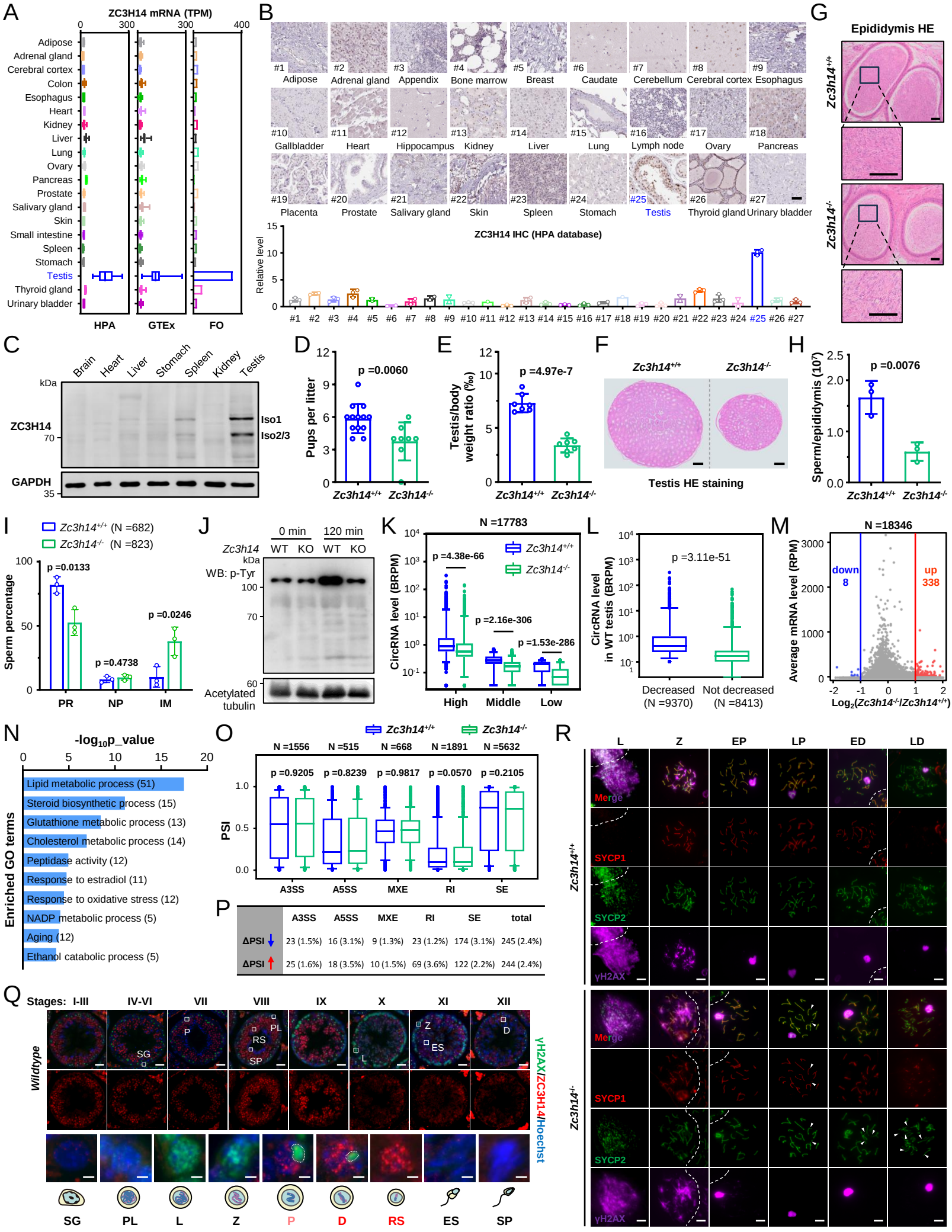

A
ZC3H14 mRNA (TPM)
0
300
300
400
Adipose
Adrenal gland
Cerebral cortex
Colon
Esophagus
Heart
Kidney
Liver
Lung
Ovary
Pancreas
Prostate
Salivary gland
Skin
Small intestine
Spleen
Stomach
Testis
Thyroid gland
Urinary bladder
HPA
GTEx
FO
B
#1 Adipose
#2 Adrenal gland
#3 Appendix
#4 Bone marrow
#5 Breast
#6 Caudate
#7 Cerebellum
#8 Cerebral cortex
#9 Esophagus
#10 Gallbladder
#11 Heart
#12 Hippocampus
#13 Kidney
#14 Liver
#15 Lung
#16 Lymph node
#17 Ovary
#18 Pancreas
#19 Placenta
#20 Prostate
#21 Salivary gland
#22 Skin
#23 Spleen
#24 Stomach
#25 Testis
#26 Thyroid gland
#27 Urinary bladder
ZC3H14 IHC (HPA database)
Relative level
15
10
5
0
#1 #2 #3 #4 #5 #6 #7 #8 #9 #10 #11 #12 #13 #14 #15 #16 #17 #18 #19 #20 #21 #22 #23 #24 #25 #26 #27
G
Epididymis HE
Zc3h14+/+
Zc3h14-/-
C
Brain
Heart
Liver
Stomach
Spleen
Kidney
Testis
kDa
ZC3H14
70
Iso1
Iso2/3
GAPDH
35
D
p =0.0060
Pups per litter
12
9
6
3
0
Zc3h14+/+
Zc3h14-/-
E
p =4.97e-7
Testis/body weight ratio (‰)
12
10
5
0
Zc3h14+/+
Zc3h14-/-
F
Zc3h14+/+
Zc3h14-/-
Testis HE staining
H
p =0.0076
Sperm/epididymis (10^7)
2.5
2.0
1.0
0.0
Zc3h14+/+
Zc3h14-/-
I
Zc3h14+/+ (N =682)
Zc3h14-/- (N =823)
p =0.0133
p =0.4738
p =0.0246
Sperm percentage
100
50
0
PR
NP
IM
J
0 min
120 min
Zc3h14
WT
KO
WT
KO
kDa
WB: p-Tyr
100
70
60
Acetylated tubulin
K
N =17783
Zc3h14+/+
Zc3h14-/-
p =4.38e-66
p =2.16e-306
p =1.53e-286
CircRNA level (BRPM)
High
Middle
Low
L
p =3.11e-51
CircRNA level in WT testis (BRPM)
Decreased (N =9370)
Not decreased (N =8413)
M
N =18346
down 8
up 338
Average mRNA level (RPM)
3000
2000
1000
0
-2
-1
0
1
2
Log2(Zc3h14-/-/Zc3h14+/+)
N
-log10p_value
0
5
10
15
20
Lipid metabolic process (51)
Steroid biosynthetic process (15)
Glutathione metabolic process (13)
Cholesterol metabolic process (14)
Peptidase activity (12)
Response to estradiol (11)
Response to oxidative stress (12)
NADP metabolic process (5)
Aging (12)
Ethanol catabolic process (5)
Enriched GO terms
O
Zc3h14+/+
Zc3h14-/-
N =1556
N =515
N =668
N =1891
N =5632
p =0.9205
p =0.8239
p =0.9817
p =0.0570
p =0.2105
PSI
1.0
0.5
0.0
A3SS
A5SS
MXE
RI
SE
P
A3SS
A5SS
MXE
RI
SE
total
ΔPSI ↓
23 (1.5%)
16 (3.1%)
9 (1.3%)
23 (1.2%)
174 (3.1%)
245 (2.4%)
ΔPSI ↑
25 (1.6%)
18 (3.5%)
10 (1.5%)
69 (3.6%)
122 (2.2%)
244 (2.4%)
R
L
Z
EP
LP
ED
LD
Zc3h14+/+
Zc3h14-/-
Merge
SYCP1
SYCP2
γH2AX
Q
Stages: I-III
IV-VI
VII
VIII
IX
X
XI
XII
Wildtype
SG
P
PL
RS
SP
L
Z
ES
D
γH2AX/ZC3H14/Hoechst
SG
PL
L
Z
P
D
RS
ES
SP

Figure S6. ZC3H14 is highly expressed in human testes and *Zc3h14*$^{-/-}$ mice exhibit abnormal testes and epididymides, related to Figure 6

(A) The ZC3H14 mRNA expression pattern across various human tissues from three databases. TPM, transcript per million.

(B) The tissue expression pattern of ZC3H14 protein. ZC3H14 immunohistochemistry (IHC) images are obtained from the HPA database, and the ZC3H14 IHC signal normalized to the cerebellum is quantified as percentage of positive area, calculated by ImageJ. Scale bar, 50 μm.

(C) ZC3H14 protein levels in various mouse tissues. GAPDH was used as a loading control.

(D) The average number of pups per litter generated from *Zc3h14*$^{+/+}$ or *Zc3h14*$^{-/-}$ male mice (N =3 per group).

(E) Ratios of testis to body weight of adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice (N =7 per group).

(F) Hematoxylin and eosin (H&E) staining of the panoramic testis sections from adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. Scale bar, 500 μm.

(G) H&E staining of the epididymide sections from adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. Scale bar, 50 μm.

(H) The counts of sperm released from the epididymides of adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice (N =3 per group).

(I) The percentage of PR, NP, and IM sperm in adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. PR, progressive motility; NP, nonprogressive motility; IM, immotility.

(J) Western blotting revealing the phosphorylated tyrosine (p-Tyr) levels of *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ sperm in a capacitation medium for 0 min or 120 min. Acetylated tubulin was a loading control.

(K) The circRNA levels of distinct groups in testes of *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. Three groups (High, Middle and Low) were evenly divided based on the corresponding circRNA levels in *Zc3h14*$^{+/+}$ testes.

(L) Boxplots showing the corresponding circRNA levels in *wildtype* testes between those decreased circRNAs and not decreased circRNAs upon ZC3H14 KO.

(M) Differentially expressed mRNAs in mice testes upon ZC3H14 KO. Blue and Red dots represent significantly downregulated and upregulated mRNAs, respectively.

(N) Gene ontology (GO) analysis of differentially expressed mRNAs identified from the testes of *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. The gene number of the corresponding GO term is included in the brackets.

(O) The PSI of five alternative splicing events analyzed by rMATS in *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ testes. A3SS, alternative 3' splice sites; A5SS, alternative 5' splice sites; MXE, mutually exclusive exons; RI, retained introns; SE, skipped exons; PSI, percent spliced in.

(P) Differential alternative splicing events in *Zc3h14*$^{-/-}$ testes. The blue and red arrows represent decreased and increased alternative splicing events, respectively.

(Q) ZC3H14 and γH2AX double immunofluorescence (IF) staining of testicular sections from adult wildtype mice and dynamic ZC3H14 expression pattern enlarged from testicular tubules during spermatogenesis. Nuclei were stained with Hoechst. SG, spermatogonium; PL, preleptotene; L, leptotene; Z, zygotene; P, pachytene; D, diplotene; RS, round spermatid; ES, elongating spermatid; SP, sperm. The white dotted line encircles the sex body. Scale bar, 50 and 5 μm (enlarged areas).

(R) IF staining of SYCP2, SYCP1 and γH2AX in spermatocyte spreads from adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. The white arrowheads mark the abnormal SYCP2or SYCP1 signals in *Zc3h14*$^{-/-}$ mice pachytene or diplotene spermatocytes. The white dotted lines represent the cell boundaries. Scale bars, 10 μm.

For D, E, H and I, p-values were calculated by two-tailed Student's *t*-test. For K, L and O, p-values were calculated by the Mann-Whitney U test.

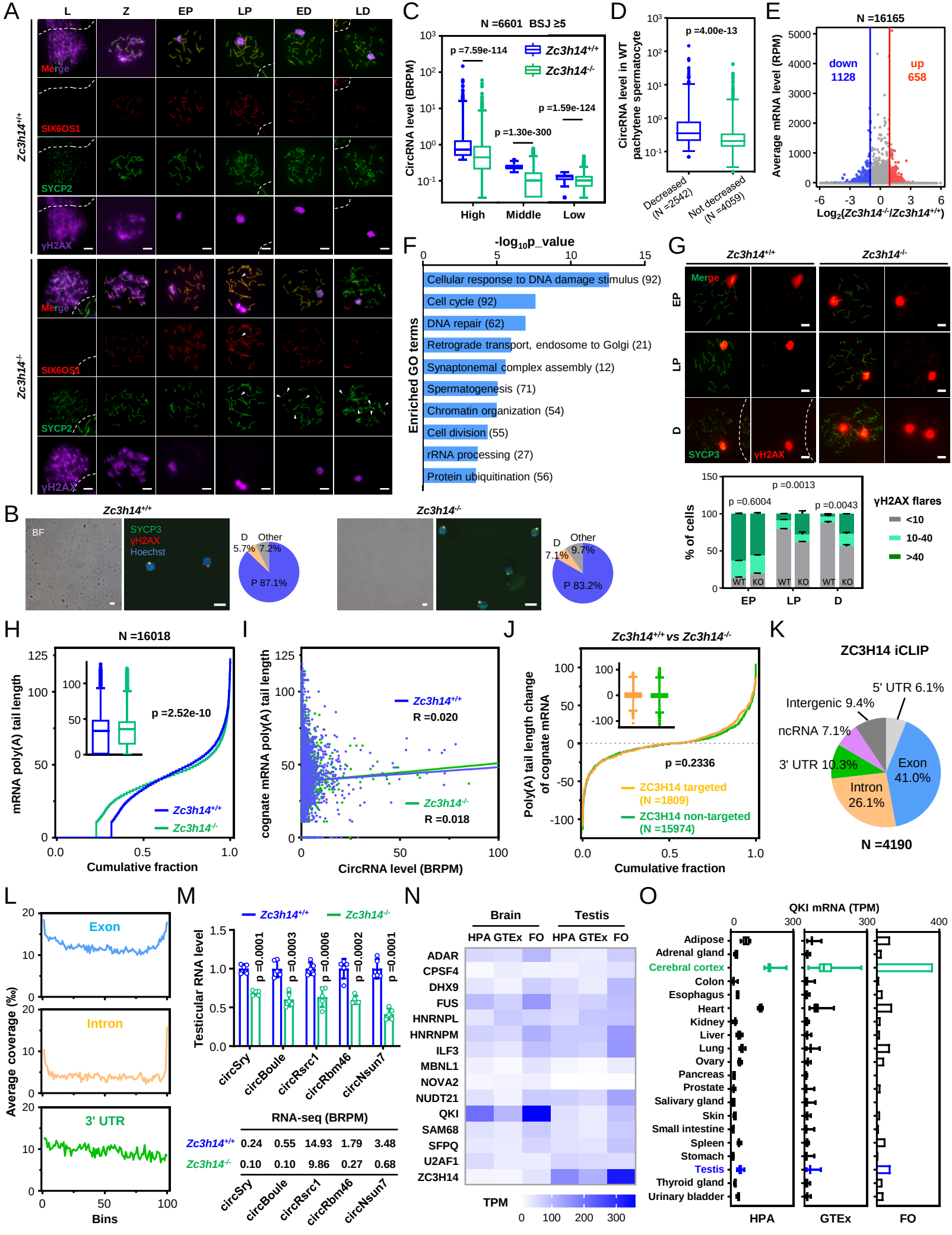

A
L
Z
EP
LP
ED
LD
Zc3h14+/+
Zc3h14-/-
Merge
SIX6OS1
SYCP2
γH2AX
B
Zc3h14+/+
Zc3h14-/-
BF
SYCP3
γH2AX
Hoechst
D 5.7%
Other 7.2%
P 87.1%
D 7.1%
Other 9.7%
P 83.2%
C
N =6601 BSJ ≥5
CircRNA level (BRPM)
p =7.59e-114
p =1.30e-300
p =1.59e-124
Zc3h14+/+
Zc3h14-/-
High
Middle
Low
D
CircRNA level in WT pachytene spermatocyte
p =4.00e-13
Decreased (N =2542)
Not decreased (N =4059)
E
N =16165
Average mRNA level (RPM)
down 1128
up 658
Log2(Zc3h14-/-/Zc3h14+/+)
F
-log10p_value
Enriched GO terms
Cellular response to DNA damage stimulus (92)
Cell cycle (92)
DNA repair (62)
Retrograde transport, endosome to Golgi (21)
Synaptonemal complex assembly (12)
Spermatogenesis (71)
Chromatin organization (54)
Cell division (55)
rRNA processing (27)
Protein ubiquitination (56)
G
Zc3h14+/+
Zc3h14-/-
EP
LP
D
Merge
SYCP3
γH2AX
p =0.0013
p =0.6004
p =0.0043
% of cells
WT
KO
γH2AX flares
<10
10-40
>40
H
N =16018
mRNA poly(A) tail length
p =2.52e-10
Zc3h14+/+
Zc3h14-/-
Cumulative fraction
I
cognate mRNA poly(A) tail length
Zc3h14+/+
R =0.020
Zc3h14-/-
R =0.018
CircRNA level (BRPM)
J
Zc3h14+/+ vs Zc3h14-/-
Poly(A) tail length change of cognate mRNA
p =0.2336
ZC3H14 targeted (N =1809)
ZC3H14 non-targeted (N =15974)
Cumulative fraction
K
ZC3H14 iCLIP
5' UTR 6.1%
Intergenic 9.4%
ncRNA 7.1%
3' UTR 10.3%
Exon 41.0%
Intron 26.1%
N =4190
L
Exon
Intron
3' UTR
Average coverage (‰)
Bins
M
Zc3h14+/+
Zc3h14-/-
Testicular RNA level
p =0.0001
p =0.0003
p =0.0006
p =0.0002
p =0.0001
circSry
circBoule
circRsrc1
circRbm46
circNsun7
RNA-seq (BRPM)
Zc3h14+/+ 0.24 0.55 14.93 1.79 3.48
Zc3h14-/- 0.10 0.10 9.86 0.27 0.68
N
Brain
Testis
HPA GTEx FO HPA GTEx FO
ADAR
CPSF4
DHX9
FUS
HNRNPL
HNRNPM
ILF3
MBNL1
NOVA2
NUDT21
QKI
SAM68
SFPQ
U2AF1
ZC3H14
TPM
O
QKI mRNA (TPM)
Adipose
Adrenal gland
Cerebral cortex
Colon
Esophagus
Heart
Kidney
Liver
Lung
Ovary
Pancreas
Prostate
Salivary gland
Skin
Small intestine
Spleen
Stomach
Testis
Thyroid gland
Urinary bladder
HPA
GTEx
FO

Figure S7. The effects of ZC3H14 in pachytene spermatocytes and its molecular mechanism in mouse germ cells, related to Figures 6 and 7

(A) IF staining of SYCP2, SIX6OS1 and γH2AX in spermatocyte spreads from adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. The white arrowheads mark the abnormal SYCP2 or SIX6OS1 signals in *Zc3h14*$^{-/-}$ mice pachytene or diplotene spermatocytes. The white dotted lines represent the cell boundaries. Scale bars, 10 μm.

(B) Verification of sorted pachytene spermatocytes from adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice testes. The percentage of pachytene spermatocytes in all sorted cells was also indicated. Scale bars, 20 μm.

(C) The circRNA levels of distinct groups in pachytene spermatocytes from adult *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice testes. Three groups (High, Middle and Low) were evenly divided based on circRNA levels in pachytene spermatocytes from adult *Zc3h14*$^{+/+}$ mice testes.

(D) Boxplots showing the circRNA levels in wildtype pachytene spermatocytes between the decreased circRNAs and not decreased circRNAs upon ZC3H14 KO.

(E) Differentially expressed mRNAs in pachytene spermatocytes of *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. Blue and Red dots represent significantly downregulated and upregulated mRNAs, respectively.

(F) Gene ontology (GO) analysis of differentially expressed mRNAs identified from the pachytene spermatocytes of *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. The gene number of the corresponding GO term is included in the brackets.

(G) IF staining of SYCP3 and γH2AX with long exposure in spermatocyte spreads from *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice. The white dotted lines represent the cell boundaries. The histogram showing the percentages of three categories of γH2AX flares on autosome at the EP, LP, and D stages in *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mice (right). EP, early pachytene; LP, late pachytene; D, diplotene. Scale bars, 10 μm.

(H) Cumulative fraction curves and boxplots showing median poly(A)-tail length of mRNAs in *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ mouse male germ cells. N represents the number of analyzed mRNAs.

(I) The correlation analysis of circRNA levels and median poly(A)-tail lengths of cognate mRNAs in *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ testes. R represents Spearman's correlation coefficient.

(J) Poly(A)-tail length changes of cognate mRNAs for ZC3H14 targeted and ZC3H14 non-targeted circRNAs in mouse germ cells upon ZC3H14 KO. ZC3H14 targeted circRNAs are circRNAs with decreased levels in *Zc3h14*$^{-/-}$ testes and pachytene spermatocytes. ZC3H14 non-targeted circRNAs are the remaining circRNAs in testes.

(K) The genomic distribution of ZC3H14 binding sites in mouse germ cells. N represents the total ZC3H14 binding sites.

(L) Coverage profile of ZC3H14 binding sites for exon, intron, and 3' UTR in mouse germ cells. Each subject was divided into 100 equal bins and coverage levels were summarized.

(M) RT-qPCR analysis showing levels of 5 circRNAs (normalized to 18S rRNA) in *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ testes (N =5 per group). The BRPM of 5 circRNAs from RNA-seq data of *Zc3h14*$^{+/+}$ and *Zc3h14*$^{-/-}$ testes were shown (bottom).

(N) The mRNA levels of 14 reported circRNA regulators and ZC3H14 in human brain and testis from three databases. TPM, transcript per million.

(O) QKI mRNA expression pattern in 20 human tissues from three databases.

For C, D, H and J, p-values were calculated by the Mann-Whitney U test. For G and M, p-values were calculated by two-tailed Student's t-test.