SUPPLEMENTAL ITEMS

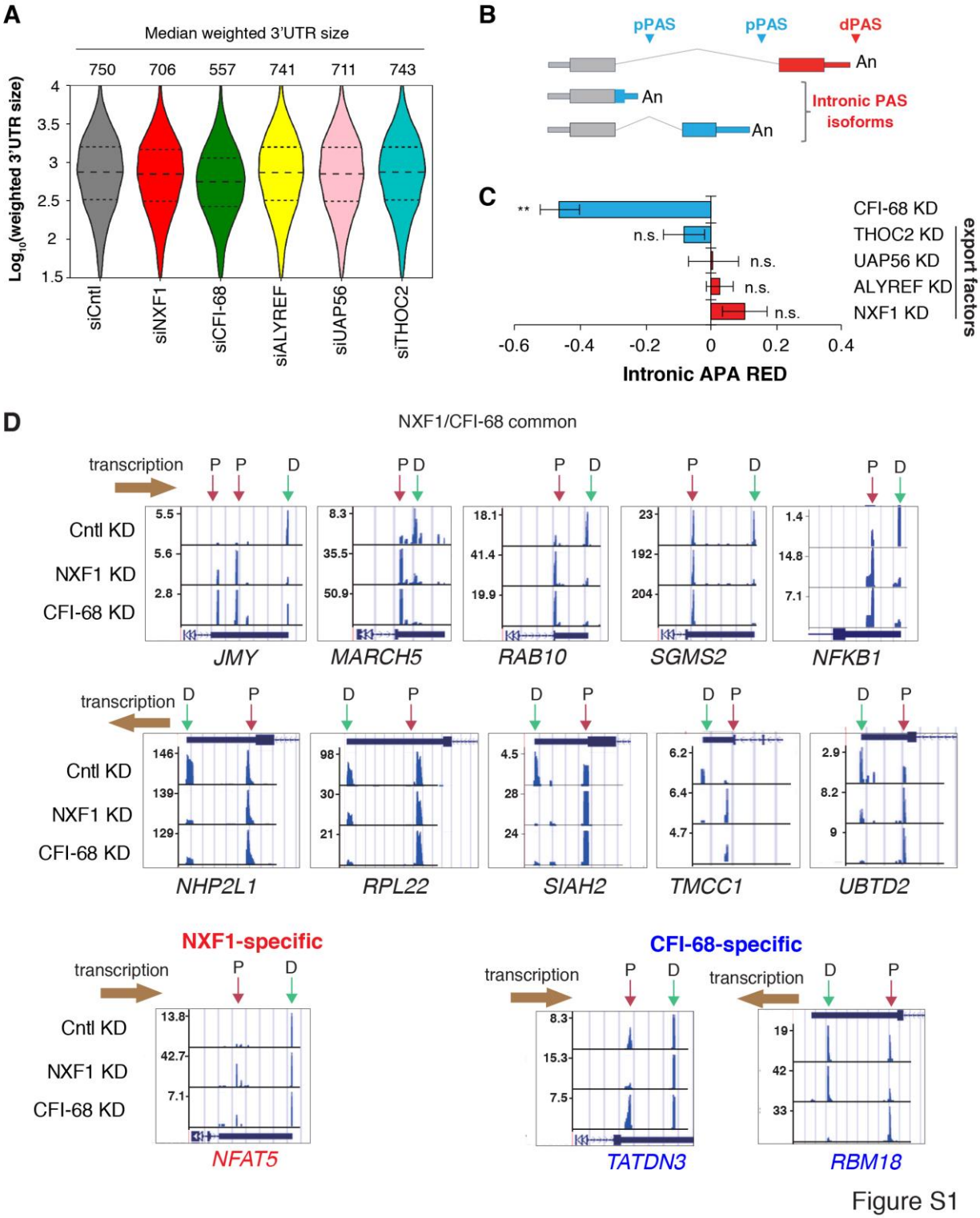

Figure S1

Figure S1. APA regulation by export factors and CFI-68, Related to Figures 1 and 2.

(A) The average 3'UTR length of a gene weighted over all isoforms.

(B) Schematic of intronic APA. All intronic APA isoforms were grouped as the pPAS group, and all last exon APA isoforms were grouped as the dPAS group.

(C) Intronic APA RED in different KD samples. RED, relative expression difference between dPAS and pPAS groups. Error bars are standard deviation based on random sampling of data for 20 times. Statistically significant level was indicated based on FDR cutoffs. (** FDR < 5%, * FDR < 10%, n.s., not significant).

(D) Genes with 3'UTR shortening in both NXF1 KD and CFI-68 KD cells, in CFI-68 KD cells only, and in NXF1 KD cells only.

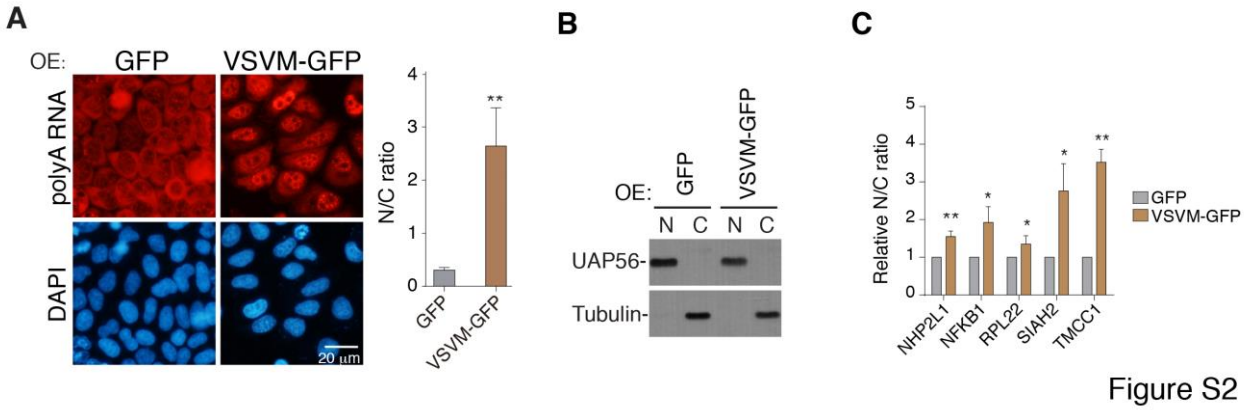

Figure S2. Overexpression of VSV-M blocked mRNA export, Related to Figure1.

(A) FISH analysis examining the nucleocytoplasmic distribution of polyA RNAs in control and VSV-M overexpression cells. Data are presented as mean ± s.d..

(B) Western blot analysis examining the purity of nuclear and cytoplasmic fractions. UAP56 and Tubulin served as nuclear and cytoplasmic markers, respectively.

(C) RT-qPCR analysis examining nucleocytoplasmic distribution changes of indicated mRNAs upon VSV-M overexpression. Data are presented as mean ± s.d..

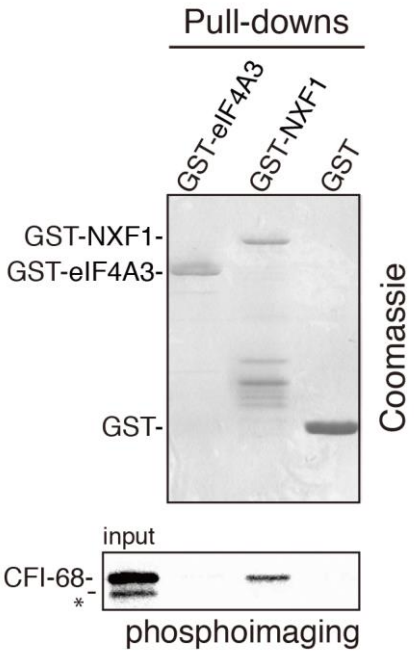

Figure S3

Figure S3. NXF1 directly interacts with CFI-68, Related to Figure 2.

Upper panel, Coomassie staining to detect GST-fused proteins. Lower panel, *in vitro*-translated ^{35}S-labeled CFI-68 pulled down by GST-eIF4A3, GST-NXF1 or GST as visualized by autoradiography. * indicates a nonspecific band.

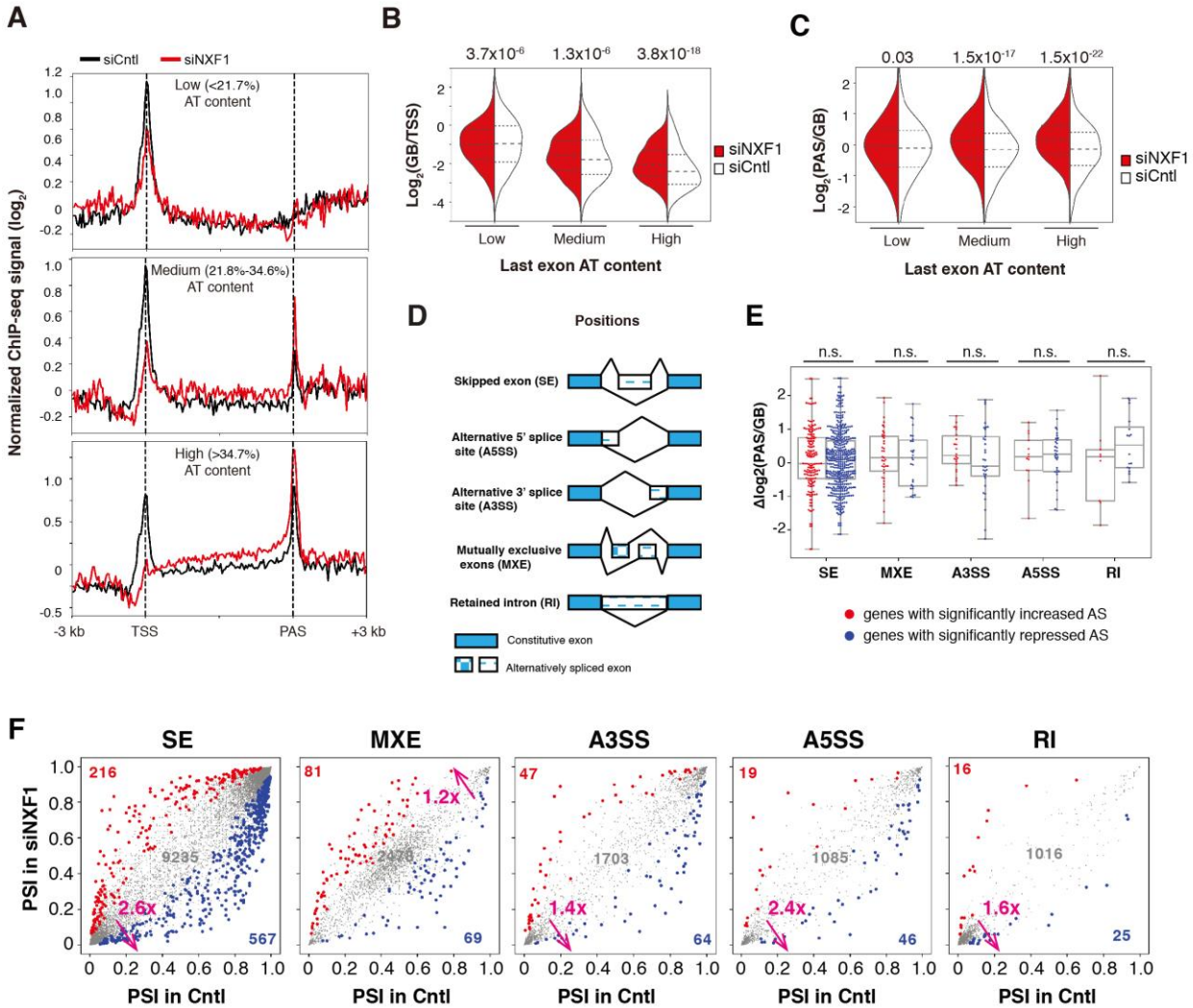
A
siCntl
siNXF1
Low (<21.7%) AT content
Medium (21.8%-34.6%) AT content
High (>34.7%) AT content
Normalized ChIP-seq signal (log2)
-3 kb
TSS
PAS
+3 kb
B
3.7x10-6
1.3x10-6
3.8x10-18
Log2(GB/TSS)
Low
Medium
High
Last exon AT content
siNXF1
siCntl
C
0.03
1.5x10-17
1.5x10-22
Log2(PAS/GB)
Low
Medium
High
Last exon AT content
siNXF1
siCntl
D
Positions
Skipped exon (SE)
Alternative 5' splice site (A5SS)
Alternative 3' splice site (A3SS)
Mutually exclusive exons (MXE)
Retained intron (RI)
Constitutive exon
Alternatively spliced exon
E
n.s.
Δlog2(PAS/GB)
SE
MXE
A3SS
A5SS
RI
genes with significantly increased AS
genes with significantly repressed AS
F
SE
MXE
A3SS
A5SS
RI
PSI in siNXF1
PSI in Cntl
216
9235
567
2.6x
81
2476
69
1.2x
47
1703
64
1.4x
19
1085
46
2.4x
16
1016
25
1.6x
G
stable ex:
Flag-Cntl
Flag-NXF1
Flag-NXF1-
NXF1-
Flag-NXF1-
Flag-Cntl-
GAPDH-

Figure S4

Figure S4. NXF1 impacts RNP II distribution in gene sets, Related to Figures 4 and 5.

(A) Metagene analysis of normalized RNAPII signals (log2(IP/input) along genes with low (top), medium (middle), and high (bottom) AT contents. The contribution of each gene is normalized (ChIP-seq signal of each position of a gene is divided by the signal sum of the gene).

(B) RNAP II signals in gene body (GB) vs. transcription start site (TSS) in different AT content groups.

(C) RNAP II signal in PAS region (PAS) vs. gene body (GB) in different AT content groups.

(D) Schematic of different types of alternative splicing (AS) analyzed here.

(E) Difference of RNAPII signals in PAS region vs. gene body, $\Delta\log_2$(PAS/GB), in genes with significantly increased (red) and decreased (blue) indicated AS events in NXF1 KD cells.

(F) AS changes in siNXF1-treated cells. Events with significant increase (red), decrease (blue), and no change (gray) are indicated. The numbers of events and ratios between events are shown.

(G) Western analysis of the level of Flag-NXF1 in stable expression cells.

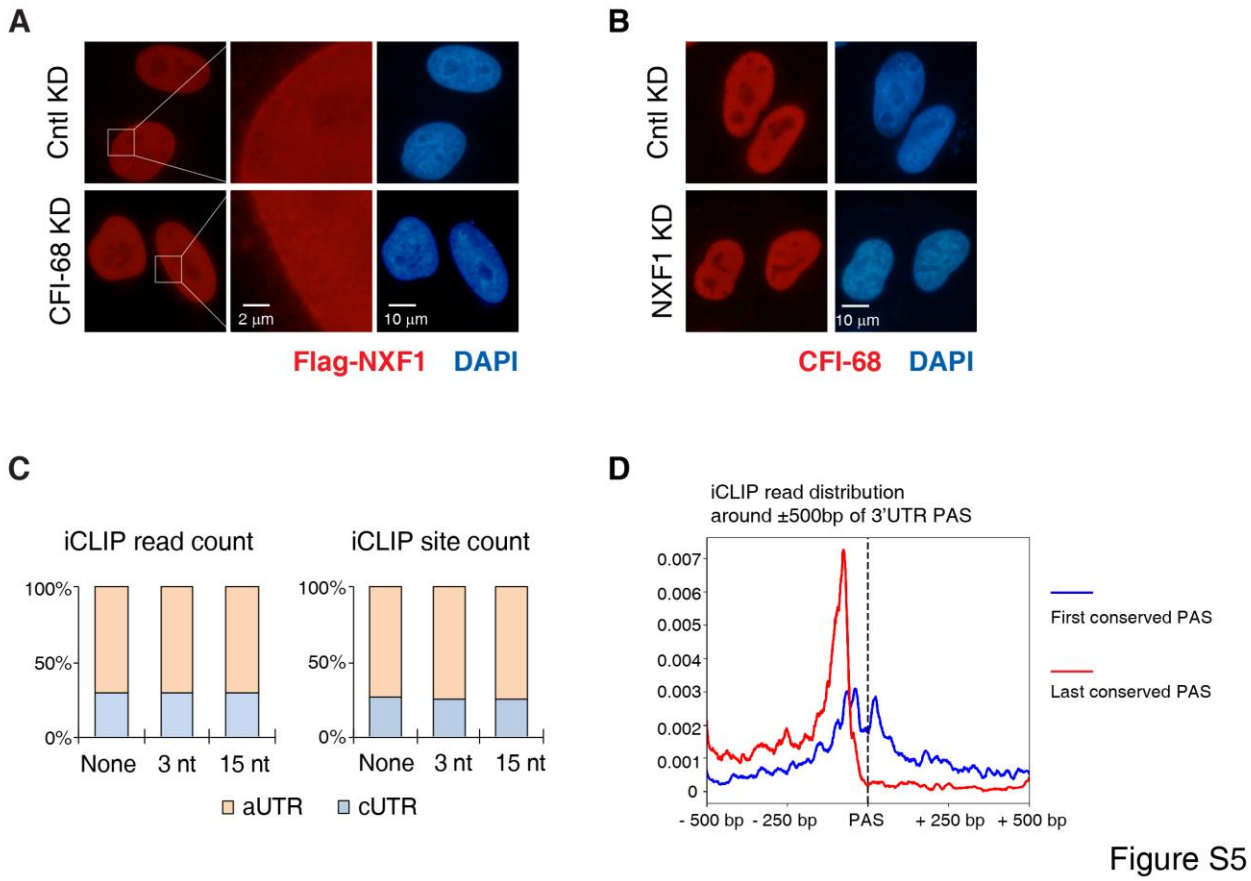

Figure S5. NXF1 promotes nuclear export of long 3'UTR isoforms, Related to Figure 6.

(A) Examination of the localization of Flag-NXF1 in control and CFI-68 KD cells. The higher magnification shows its localization at nuclear pores.

(B) Examination of CFI-68 localization in control and NXF1 KD cells.

(C) Percentages of NXF1 iCLIP sites (left panel) and number of iCLIP reads (right panel) in cUTRs (blue) and aUTRs (pink) using different window sizes (none, ±3 bp and ±15 bp) for clustering iCLIP reads.

(D) Normalized iCLIP reads mapped to ±500 bp region of the first conserved (blue) and last conserved (red) 3'UTR PASs.

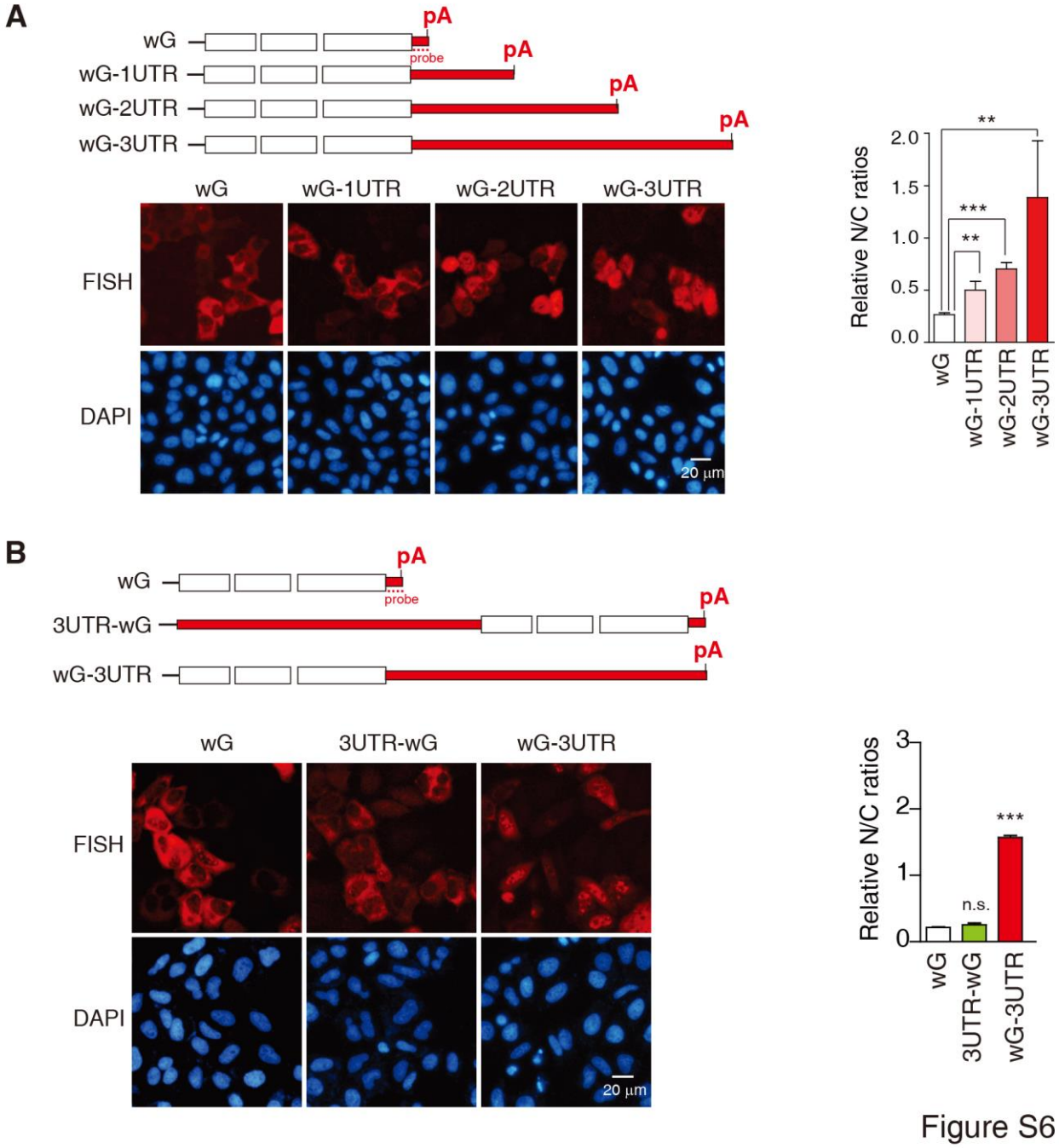
A
wG
wG-1UTR
wG-2UTR
wG-3UTR
pA
probe
wG
wG-1UTR
wG-2UTR
wG-3UTR
FISH
DAPI
20 μm
Relative N/C ratios
2.0
1.5
1.0
0.5
0.0
**
***
**
wG
wG-1UTR
wG-2UTR
wG-3UTR
B
wG
3UTR-wG
wG-3UTR
pA
probe
wG
3UTR-wG
wG-3UTR
FISH
DAPI
20 μm
Relative N/C ratios
3
2
1
0
n.s.
***
wG
3UTR-wG
wG-3UTR

Figure S6

Figure S6. Inefficient nuclear export of reporter mRNAs with long 3'UTRs is not due to nuclear retention signals, Related to Figure 7.

(A) Nuclear export efficiency of β-globin reporter mRNAs decreases as 3'UTR length increases. Top, schematic of β-globin constructs with different 3'UTR lengths. Bottom, FISH to detect the nucleocytoplasmic distribution of β-globin reporter mRNAs at 12 hr time point after transfection. DAPI staining serves as nuclear marker. The graph on the right shows nucleus/cytoplasm (N/C) ratios of different reporter mRNAs. Data are presented as mean ± s.d..

(B) Insertion of the same sequence in 3'UTR, but not in 5'UTR, decreased mRNA nuclear export efficiency. Top, schematic of β-globin constructs. Bottom, FISH to detect the nucleocytoplasmic distribution of β-globin reporter mRNAs at 8 hr time point after transfection. DAPI staining serves as nuclear marker. The graph on the right shows N/C ratios of different reporter mRNAs. Data are presented as mean ± s.d..

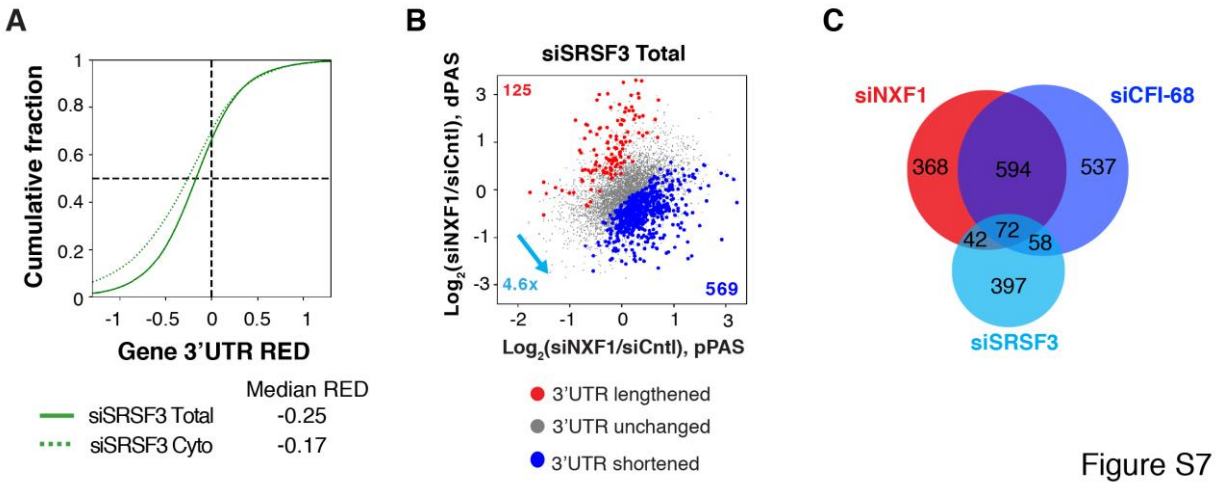

Figure S7. The APA events regulated by NXF1 and CFI-68 is not highly overlapped with those regulated by SRSF3, Related to Figure 1.

(A) Cumulative distribution of 3'UTR RED values based on cytoplasmic RNA (dotted line) or total RNA (solid line) data from siSRSF3-treated cells.

(B) 3'UTR APA changes in siSRSF3 cells using total RNA. Genes with significant 3'UTR lengthening (red), 3'UTR shortening (blue) or no change (gray) are indicated. The numbers of genes in each group and ratio of number of lengthening genes to that of shortening genes are shown.

(C) Venn diagram showing overlap of genes with shortened 3'UTRs in the NXF1, CF1-68 and SRSF3 KD cells.

Table S1. Statistics of PAS and APA in KD and control samples, Related to Figures 1 and 2.

Sample	No. of PAS reads	No. of genes	No. of genic PASs	No. of PASs per gene	No. of APA genes
siALYREF	2,294,055	11,101	25,349	2.28	6,721
siCFI-68	3,198,896	10,990	28,316	2.58	7,265
siUAP56	3,230,793	10,927	27,976	2.56	7,109
Cntl	4,027,034	11,431	29,524	2.58	7,508
siNXF1	3,828,508	11,415	32,166	2.82	7,770
siTHOC2	3,209,968	11,100	27,125	2.44	7,103

Table S2. Nuclear pore factors present in the NXF1 or Cntl (WDR33) immunoprecipitate, Related to Figures 4 and 5.

Gene Name	MW(kD)	Cntl (WDR33)			NXF1		
		Peptide	Unique	Coverage	Peptide	Unique	Coverage
WDR33	145.18	55	55	45.94	0	0	0
NXF1	70.13	0	0	0	30	30	58.64
Nup210	204.98	0	0	0	35	35	24.01
Nup214	213.49	0	0	0	16	16	9.62
Nup205	227.78	0	0	0	14	14	8.8
Nup93	93.43	0	0	0	7	7	10.18
Nup98	197.46	0	0	0	7	7	4.4
Nup155	155.1	0	0	0	5	5	4.53
Nup88	83.49	0	0	0	3	3	7.15
Nup133	128.9	0	0	0	3	3	4.15
Nup160	162.02	0	0	0	3	3	2.3
Nup85	74.97	0	0	0	2	2	4.88
Nup37	36.68	0	0	0	1	1	4.6
Nup62	53.22	0	0	0	1	1	2.68
Nup43	42.12	0	0	0	1	1	2.63
Nup107	106.31	0	0	0	1	1	1.3
Nup188	195.92	0	0	0	1	1	0.74
Gle1	79.79	0	0	0	2	2	5.43
Rae1	40.94	0	0	0	1	1	7.83

Table S3. siRNAs and Primers used in this study, Related to STAR Methods.

siRNAs used in this study		
Target gene		**Targeting sequence**
Luciferase		AACAGUCGCGUUUGCUACUUU
CFI-68		GACCGAGAUUACAUGGAUA
NXF1		GCGCCAUUCGCGAACGAUUUU
UAP56		AAGGGCUUGGCUAUCACAUUU
URH49		AAAGGCCUAGCCAUCACUUUU
THOC2		GGUUAUGCCAAGCUGAUUG
ALYREF		TGGGAAACTGCTGGTGTCCAA
CFI-25		CCUCUUACCAAUUAUACUU
Cloning primers used in this study		
wS-1UTR	Forward	CCGGATATCGTGCACCTGACTCCTG
wS-1UTR	Reverse	CCGGATATCGTGATACTTGTGGGCCAG
wS-2UTR	Forward	CCGGATATCGTGCACCTGACTCCTGAGGAGAA
wS-2UTR	Reverse	AAGGAAAAAAGCGGCCGCGTGATACTTGTGG
wS-3UTR	Forward	CTAGTCTAGAGTGCACCTGACTCCTGAGGAGAAGTCT
wS-3UTR	Reverse	AGTCGGGCCCGTGATACTTGTGGG
wS-3UTR UGUA	Forward	CCGTGCCTTCCTTGACCC TGTA TGTA TGTA TGTA TGTA TGGAAGGTGCCACTCCCA
wS-3UTR UGUA	Reverse	TGGGAGTGGCACCTTCCATACATACATACATA CATACAGGGTCAAGGAAGGCACGG
wG-1UTR	Forward	CCGGATATCGTGCACCTGACTCCTG
wG-1UTR	Reverse	CCGGATATCGTGATACTTGTGGGCCAG
wG-2UTR	Forward	CCGGATATCGTGCACCTGACTCCTGAGGAGAA
wG-2UTR	Reverse	AAGGAAAAAAGCGGCCGCGTGATACTTGTGG
wG-3UTR	Forward	CTAGTCTAGAGTGCACCTGACTCCTGAGGAGAAGTCT
wG-3UTR	Reverse	AGTCGGGCCCGTGATACTTGTGGG
1UTR-wG	Forward	CCCAAGCTTGTGCACCTGACTCCTGAGGAGAAGT
1UTR-wG	Reverse	CGGGGTACCGTGATACTTGTGGGCCAGGG
2UTR-wG	Forward	CGGGGTACCGTGCACCTGACTCCTGAGGAGAAGTCTG
2UTR-wG	Reverse	CGGGGTACCGTGATACTTGTGGGCCAGGGCA
3UTR-wG	Forward	CCCAAGCTTGTGCACCTGACTCCTGAGG
3UTR-wG	Reverse	CCCAAGCTTGTGATACTTGTGGGCCAGGG
Flag-CFI-68	Forward	CCGGAATTCAATGGCGGACGGCGTGGAC
Flag-CFI-68	Reverse	TCTGATATCCTAACGATGACGATATTC
Flag-NXF1	Forward	CCGGAATTCAATGGCGGACGAGGGGAAG
Flag-NXF1	Reverse	CGGGATCCTCACTTCATGAATGCCAC
HAGE-Flag-NXF1	Forward	CTAGTCTAGAATGGCGGACGAGGGGAAGTCG
HAGE-Flag-NXF1	Reverse	CGCGGATCCTCACTTCATGAATGCCACTT
HAGE-Flag-eIF4A3	Forward	CGACGCGTATGGCGACCACGGCCACGAT
HAGE-Flag-eIF4A3	Reverse	CGGGATCCTCAGATAAGATCAGCAACGT
GST-NXF1	Forward	CGCGGATCCATGGCGGACGAGGGGAAGTCG
GST-NXF1	Reverse	ATAAGAATGCGGCCGCTCACTTCATGAATGCCACTT
RT-qPCR primers used in this study		
Gene Name	Targeting Region	Direction	Sequence

TMCC1	Common	Forward	GTGAAGGTGTGGTGGATA
TMCC1	Common	Reverse	CGAATGAGTGAGGCAATC
TMCC1	aUTR	Forward	GTCGCATAACCAGTTCCA
TMCC1	aUTR	Reverse	CACAGAATGAGGCTAAGAGG
SIAH2	Common	Forward	AATACCGTCCCTACTCCT
SIAH2	Common	Reverse	GAAGGGTGGTAATGCTCT
SIAH2	aUTR	Forward	GACTGTCAGCAGATTCCT
SIAH2	aUTR	Reverse	CTATTAGCCAGCCATCCA
MARCH5	Common	Reverse	GAGCCGACCATTATTCCT
MARCH5	Common	Forward	AACAGTACAGCCAGAGTG
MARCH5	aUTR	Forward	CAGAATCAGCAACTCAAGG
MARCH5	aUTR	Reverse	GGACATCACATGGGTTAAAG
NFKB1	Common	Forward	GAAGATGTGGTGGAGGAT
NFKB1	Common	Reverse	GGGTGGTCAAGAAGTAGT
NFKB1	aUTR	Forward	GCATTCCTTCTGACCACA
NFKB1	aUTR	Reverse	GGCACATCAAGTGACTCT
JMY	Common	Forward	GTGGGAGAAGGAAGAGTC
JMY	Common	Reverse	GAGGCAGTTCAGAGGTAA
JMY	aUTR	Forward	AAGTCTCTCAACCCATCC
JMY	aUTR	Reverse	CTGCCAACAATACTCTGC
UBTD2	Common	Forward	GCACAGACACAGTATTCC
UBTD2	Common	Reverse	CACAGGTTGGCTCACTAT
UBTD2	aUTR	Forward	CACCCTTCTGGAACACTA
UBTD2	aUTR	Reverse	GACCTTCTAAGCCTGGAT
SGMS2	Common	Forward	ATAGTGGGACGCAGATTC
SGMS2	Common	Reverse	CAGGCACAGGTAGAGTAG
SGMS2	aUTR	Forward	GGAAGTGTGAGAAGGAGT
SGMS2	aUTR	Reverse	GCAACTGAGTCAAAGGAG
RPL22	Common	Forward	TGTTAGCAACTACGCGCAAC
RPL22	Common	Reverse	GACCATCGAAAGGAGCAAGA
RPL22	aUTR	Forward	TCAATGCTCAGTCCTACC
RPL22	aUTR	Reverse	TGACGGCAGAGAAGAAAT
NHP2L1	Common	Forward	CTGAGGCTGATGTGAATCC
NHP2L1	Common	Reverse	ATGACTGCTGAACGAGGT
NHP2L1	aUTR	Forward	AACAGAGATGGTGGAGTC
NHP2L1	aUTR	Reverse	GGATGCTCACAAGAATGG
RAB10	Common	Forward	CACCATCACAACCTCCTA
RAB10	Common	Reverse	GTCGTCCATATCACACTTG
RAB10	aUTR	Forward	GAGGCTGAGTTTAGGACA
RAB10	aUTR	Reverse	TATGTGAGCCAAGAGGTC
NFAT5	Common	Forward	CTGTTGTTGCTGCTGGAT
NFAT5	Common	Reverse	GGCTGGAAGAGGTGGTAA
NFAT5	aUTR	Forward	TGACCTGTTCATCTGTGG
NFAT5	aUTR	Reverse	GAGCAAAGACTGAATGGC
TATDN3	Common	Forward	TCCACCAGAAGACCAAAG
TATDN3	Common	Reverse	TCTAGTCCAACCTCTCCA
TATDN3	aUTR	Forward	CCGAGTTCTGGCAGGATA
TATDN3	aUTR	Reverse	TCAGTTCAGGTGGTGGTT

RBM18	Common	Forward	TACCGAATACCACCTCCT
RBM18	Common	Reverse	CTCTGCTTCCTGCTTAGT
RBM18	aUTR	Forward	GTAGCGGTCTAAGTGGAA
RBM18	aUTR	Reverse	CATTACAGTCTTGGTGCC
3’ end RT-qPCR primers			
3’ END RT PCR Primer	Reverse	CAAGCAGAAGACGGCATACGAGA TTTTTTTTTTTTTTTTTTTTTTTTVN	
P7 primer	Reverse	CAAGCAGAAGACGGCATACGAGA	
NFKB1 short	Forward	TGTCCCTCTGCTACGTTC	
NFKB1 long	Forward	AATGGTATTTTCCCCCTT	
SIAH2 short	Forward	TTTCGACACAGCCATAGCA	
SIAH2 long	Forward	AGTAGCTGGTGTGAAAGAC	
ChIP qPCR primers			
RAB10 1	Forward	AGGTAAGACCTGTGGGAGGA	
RAB10 1	Reverse	CGGGAAGCAAGATGTCTGGA	
RAB10 2	Forward	TAAGGAAGCTGGGCATGGTG	
RAB10 2	Reverse	AGGCTGGAATGCAATGGCTA	
RAB10 3	Forward	GCAGCAGTATGTCCCAGAGG	
RAB10 3	Reverse	GGCCAAACTGTACCAATGGC	
RAB10 4	Forward	ACGCCCGTAATCCTGACATC	
RAB10 4	Reverse	CACCCTGCCCAGCTAATTCT	
RAB10 5	Forward	GGTTGGTGTGGAGCCTTCTT	
RAB10 5	Reverse	TGGCTGATTGGTTCCAGGAC	
RAB10 6	Forward	ATGCCAGGAAATGTCTACA	
RAB10 6	Reverse	ACAACTACAATCACAGGAGA	
TMCC1 1	Forward	GCTAACTTCTGGGACGCCAT	
TMCC1 1	Reverse	GGCCCGAGGGACTGTTTAG	
TMCC1 2	Forward	GCCCAGCCATGTCCAAGTAA	
TMCC1 2	Reverse	AGGGACACCTGTGGTAGAGT	
TMCC1 3	Forward	GTTGATGGGCTCACCACTGA	
TMCC1 3	Reverse	CTTGTGCATGCTCTCTTGCC	
TMCC1 4	Forward	CCTCAACCCTGACCTCACAG	
TMCC1 4	Reverse	CAGATTTGGCAGGGAGGTGT	
TMCC1 5	Forward	GGATCAAACGAGTGCTGTGC	
TMCC1 5	Reverse	AGACCAGAGTCTCCAGGCAT	
TMCC1 6	Forward	GTGCTCTTCCCTCAGACACC	
TMCC1 6	Reverse	CCTCGCTGCTAAGGTGACAT	
TMCC1 7	Forward	TCCTGGGCTAAGTGCATGTG	
TMCC1 7	Reverse	GCCCAAGACCATCTGCCTAA	
TMCC1 8	Forward	TGGATCTCTCTGGCTTTC	
TMCC1 8	Reverse	ACTGTAGTAGGACTGTGC	
In	Forward	AATCCAGGTGAGCCTGCTTC	
In	Reverse	AACCAAGTGGTGAGGTTGGG	

Table S4. Statistics of 3'READS+ data, Related to STAR Methods.

Sample	Type	Raw reads	Uniquely mapped	PAS reads
siALYREF, replicate 1	Total RNA	57,435,571	13,010,769	2,294,055
siCFI-68, replicate 2	Total RNA	40,261,637	9,518,029	3,198,896
siUAP56, replicate 1	Total RNA	35,440,912	8,376,341	3,230,793
siCntl, replicate 1	Total RNA	60,610,165	15,131,137	4,027,034
siNXF1, replicate 1	Total RNA	40,360,020	9,955,632	3,828,508
siTHOC2, replicate 1	Total RNA	45,878,421	10,945,403	3,209,968
siCFI-68, Cyto, replicate 1	Cytoplasmic RNA	21,152,231	14,408,786	6,693,981
siCFI-68, Nucl, replicate 1	Nuclear RNA	31,482,862	23,230,027	11,179,827
siCntl_Cyto, replicate 1	Cytoplasmic RNA	16,677,137	13,148,681	7,777,427
siCntl, Nucl, replicate 1	Nuclear RNA	21,250,652	15,294,373	7,060,556
siCntl, Cyto, replicate 2	Cytoplasmic RNA	44,042,288	29,539,173	14,419,979
siCntl, Nucl, replicate 2	Nuclear RNA	29,716,185	20,024,422	9,538,146
siNXF1, Cyto, replicate 1	Cytoplasmic RNA	25,206,878	18,596,275	9,921,956
siNXF1, Nucl, replicate 1	Nuclear RNA	24,514,576	17,891,038	9,093,806
siCFI-68, replicate 2	Total RNA	13,367,780	10,824,155	4,936,959
siCntl, replicate 2	Total RNA	12,001,518	10,089,105	4,672,731
siNXF1, replicate 2	Total RNA	12,774,551	10,615,251	5,188,485

Table S5. Statistics of ChIP-seq data, Related to STAR Methods.

Sample	Raw reads	Uniquely mapped reads
siCntl_input	57,598,942	51,347,503
siCntl_RNAPII IP	35,997,408	26,629,189
siNXF1_input	53,163,456	47,202,800
siNXF1_RNAPII_IP	51,981,394	15,667,833

Table S6. Statistics of nascent RNA-seq data, Related to STAR Methods.

Sample	Raw Reads	Uniquely mapped reads	Unique 3' ends
siCntl	100,354,084	47,701,646	17,669,456
siNXF1	128,379,794	75,903,155	32,446,571

Unique 3' ends are unique genomic positions corresponding to the 3' ends of reads.

Table S7. Statistics of total RNA RNA-seq data, Related to STAR Methods.

Sample	Raw reads	Uniquely mapped reads
siCntl	20,176,630	17,447,827
siNXF1	24,891,840	21,766,043