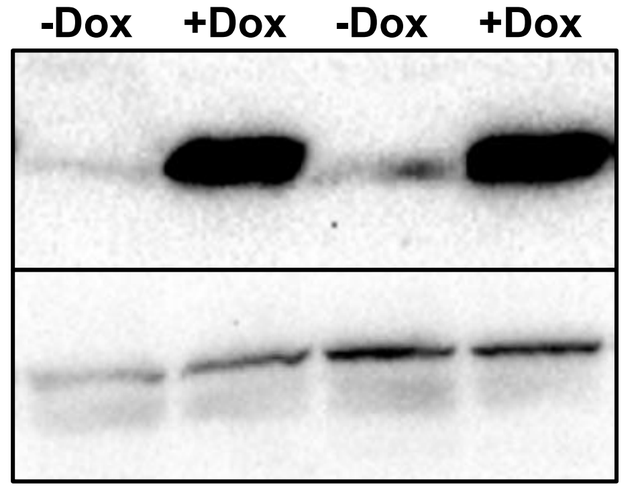
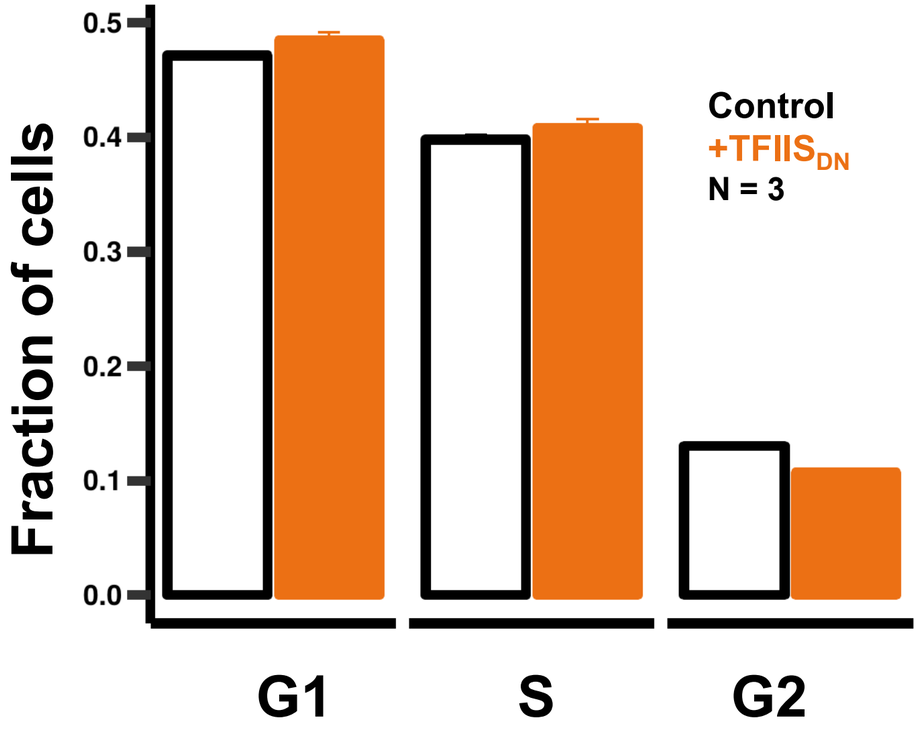
A

```
S. cerevisiae   1    MDsKEVLVHVKNLEKN--KSNDAAVLEILHVLdKEFVPTEKLLRETKVGVEVNKFKK-STNVEISKLVKKMISSWKDAIN  77
M. musculus     1    ME-DEVVRIAKKMDKMVQKKNAAGALDLLKEL-KNIPMTLELLQSTRIGMSVNALRKQSTDEEVTSLAKSLIKSWKKLLD  78
H. sapiens      1    ME-DEVVRFAKKMDKMVQKKNAAGALDLLKEL-KNIPMTLELLQSTRIGMSVNAIRKQSTDEEVTSLAKSLIKSWKKLLD  78

S. cerevisiae   78   KNKRSRQAQQHHQDHAPgNAEDKTTVGESVNGVQQPASSQSDAMKQDKYVSTKPRNSK[8]YHHKLRDQVLKALYDVLAK  162
M. musculus     79   GPSTDKDPEEKKKEPAI-SSQNSPEARESSSSSNVSSRKDETNARDTYVSSFPRAPS   TSDSVRLKCREMLAAALRT  154
H. sapiens      79   GPSTEKDLDEKKKEPAI-TSQNSPEARESTSSGNVSNRKDETNARDTYVSSFPRAPS   TSDSVRLKCREMLAAALRT  154

S. cerevisiae   163  ESEHPPQ--SILHTAKAIESEMNKvnNCDTNEAAYKARYRIIYSNVISKNNPDLKHKIANGDITPEFLATCDAKDLAPAP  240
M. musculus     155  GDDYVAIGADEEELGSQIEEAIYQ--EIRNTDMKYKNRVRSRISNLKDAKNPNLRKNVLCGNIPPDLFARMTAEEMASDE  232
H. sapiens      155  GDDYIAIGADEEELGSQIEEAIYQ--EIRNTDMKYKNRVRSRISNLKDAKNPNLRKNVLCGNIPPDLFARMTAEEMASDE  232

S. cerevisiae   241  LKQKIEEIAKQNLYNAQGATIERSVTDRFTCGKCKEKKVSYYQIQTRSADEPLTTFCTCEACGNRWKFS  309
M. musculus     233  LKEMRKNLTKEAIREHQMAKTGGTQTDLFTCGKCKKNCTYTQVQTRSADEPMTTFVVCNECGNRWKFC  301
H. sapiens      233  LKEMRKNLTKEAIREHQMAKTGGTQTDLFTCGKCKKNCTYTQVQTRSADEPMTTFVVCNECGNRWKFC  301
                                                                  __
```

B



C



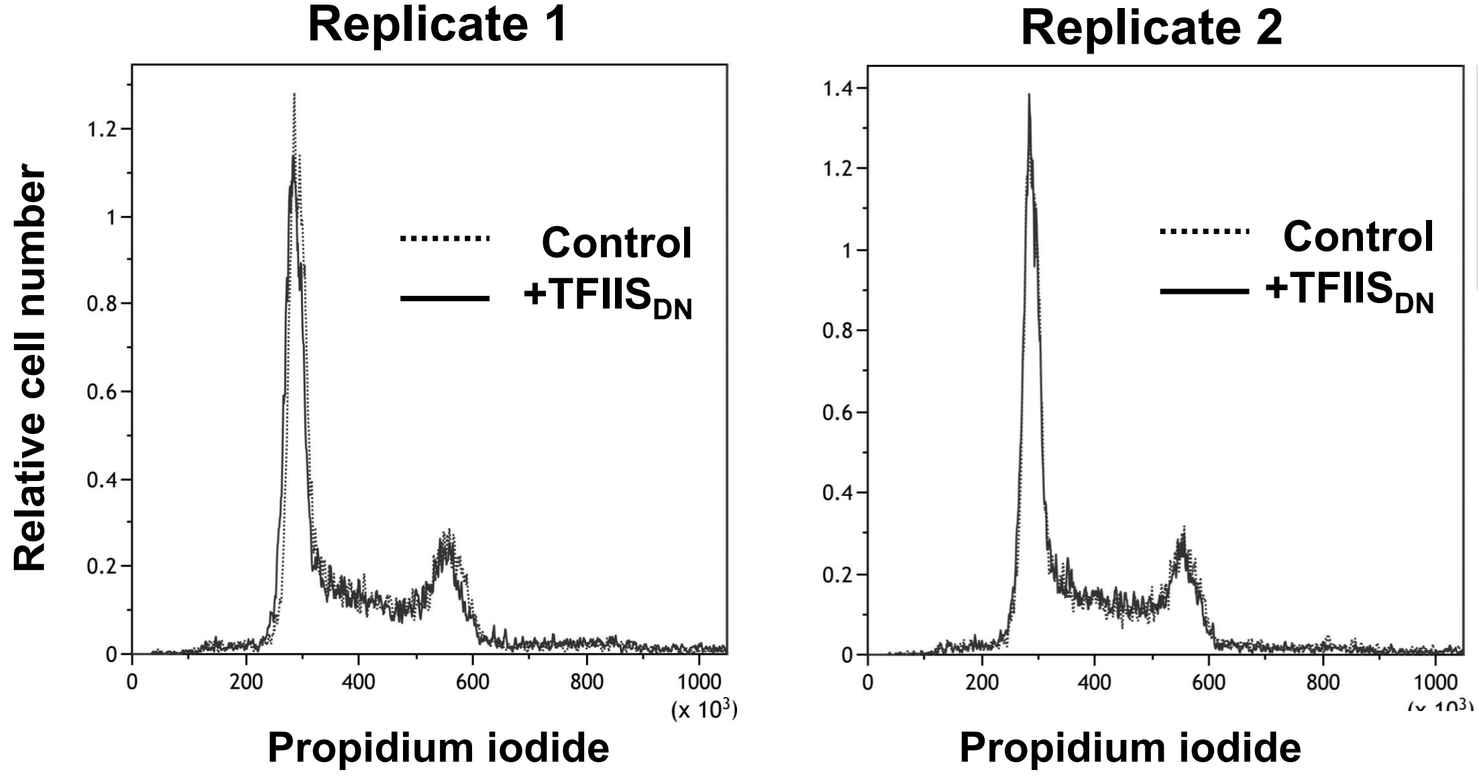


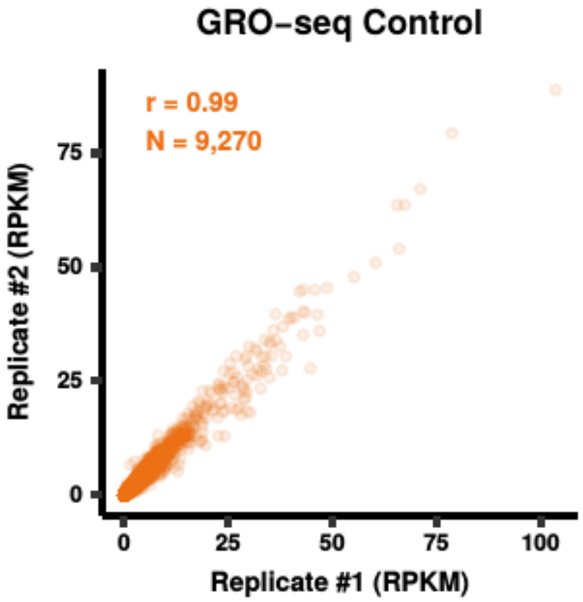
Fig. S1

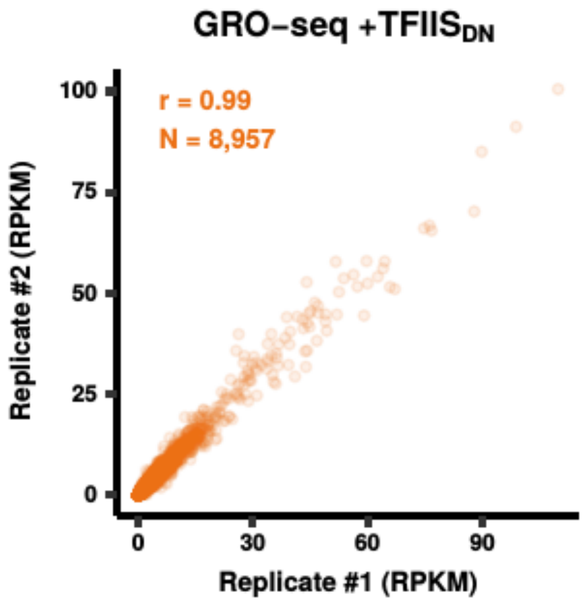
Figure S1, related to Figure 1.

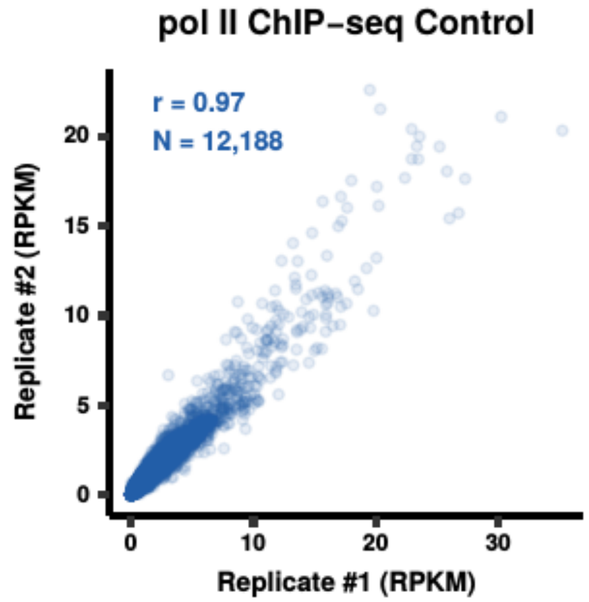
(A) Sequence alignment of yeast, mouse, and human TFIIS. The motif that interacts with the pol II trigger loop and bridge helix is highlighted in yellow, with the residues mutated in $TFIIS_{DN}$ underlined.

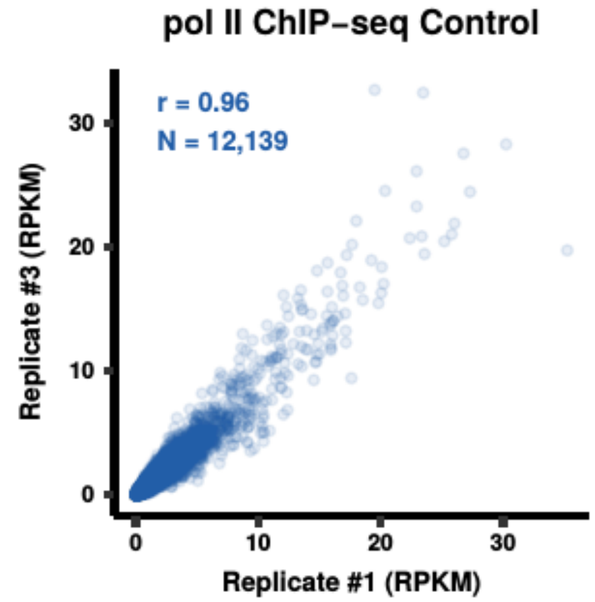
(B) Anti-TFIIS (TCEA1) Western blot of HEK293 extracts before and shown after induction of $TFIIS_{DN}$ with 2 μg/ml doxycycline (+Dox) for 24 hrs. Two biological replicates are shown. CstF77 is a loading control. Note strong over-expression of $TFIIS_{DN}$ relative to endogenous TFIIS.

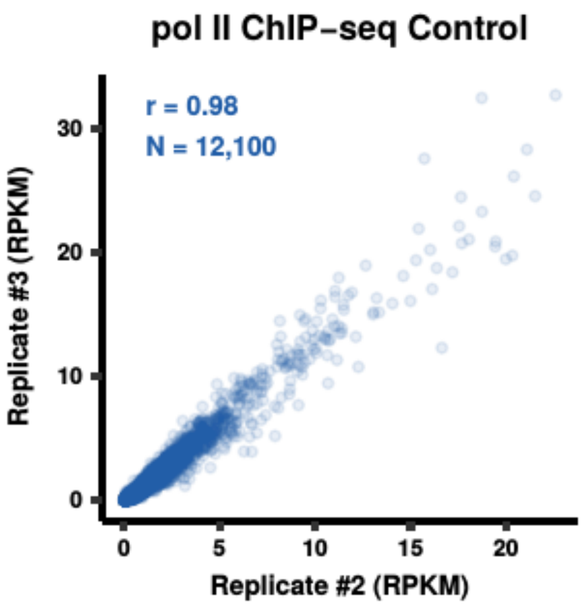
(C) Expression of $TFIIS_{DN}$ does not affect the cell cycle distribution as measured by propidium iodide flow cytometry. The fraction of cells in G1, S, and G2 phases of the cell cycle is shown before and after expression of $TFIIS_{DN}$ for 24 hours in HEK293 cells. The mean is shown for three biological replicates with the error bars representing the SEM (left). The cell cycle distribution is shown before and after expression of $TFIIS_{DN}$ for two biological replicates. Cells were incubated in Krishan stain for 72 hours before performing flow cytometry.

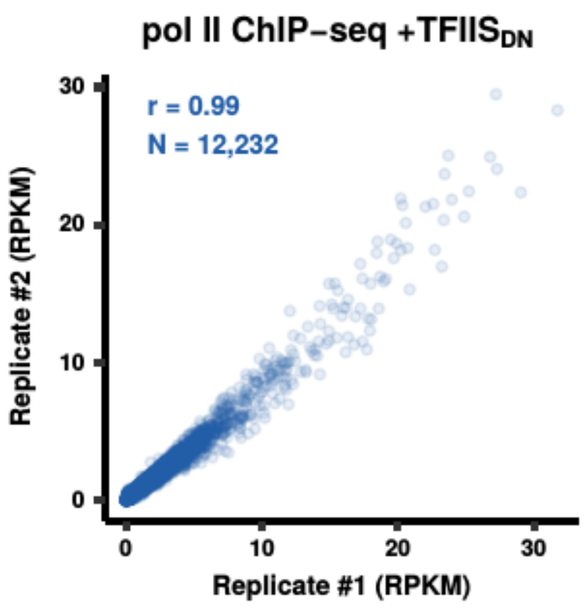

GRO-seq Control
r = 0.99
N = 9,270
Replicate #2 (RPKM)
Replicate #1 (RPKM)
0
25
50
75
100

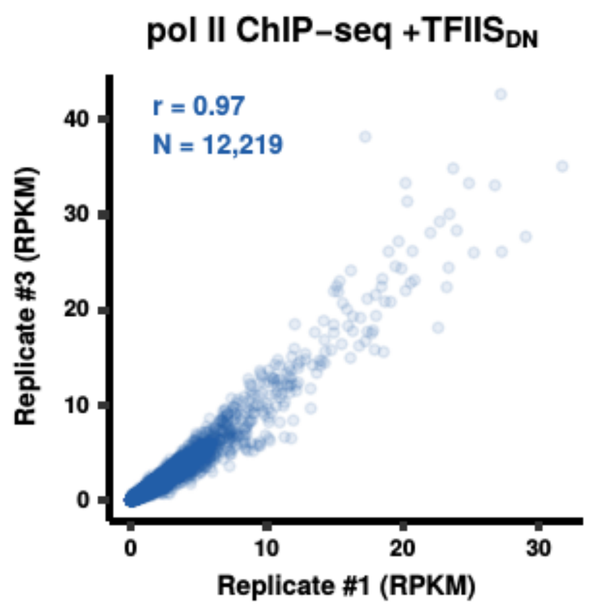

GRO-seq +TFIIS$_{DN}$
r = 0.99
N = 8,957
Replicate #2 (RPKM)
Replicate #1 (RPKM)
0
30
60
90

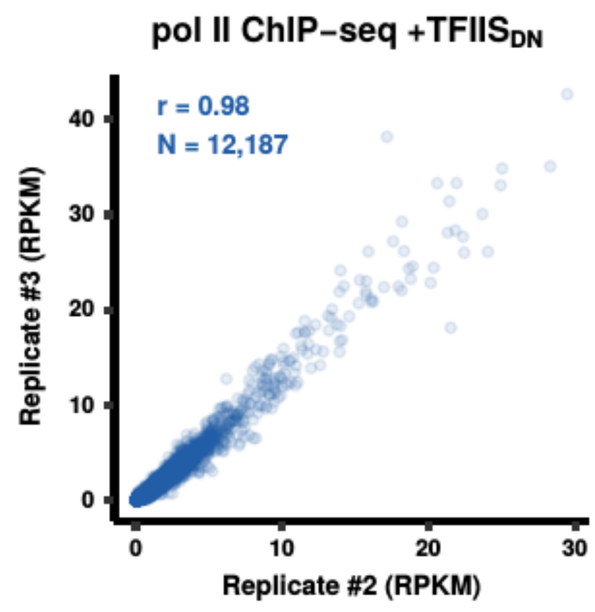

pol II ChIP-seq Control
r = 0.97
N = 12,188
Replicate #2 (RPKM)
Replicate #1 (RPKM)
0
5
10
15
20

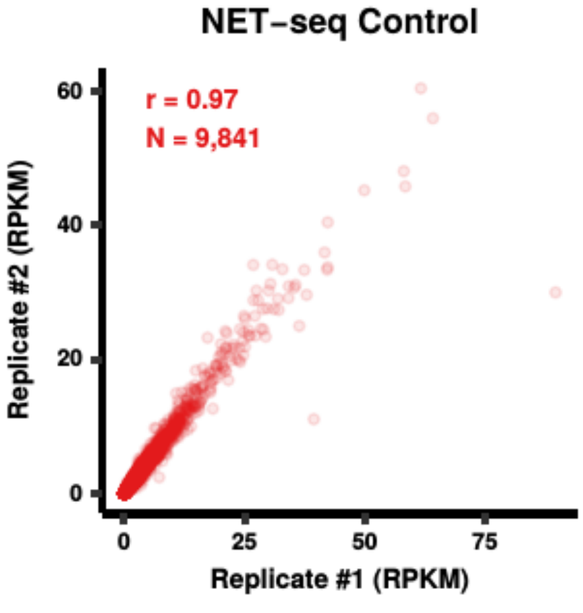

pol II ChIP-seq Control
r = 0.96
N = 12,139
Replicate #3 (RPKM)
Replicate #1 (RPKM)
0
10
20
30


pol II ChIP-seq Control
r = 0.98
N = 12,100
Replicate #3 (RPKM)
Replicate #2 (RPKM)
0
5
10
15
20


pol II ChIP-seq +TFIIS$_{DN}$
r = 0.99
N = 12,232
Replicate #2 (RPKM)
Replicate #1 (RPKM)
0
10
20
30


pol II ChIP-seq +TFIIS$_{DN}$
r = 0.97
N = 12,219
Replicate #3 (RPKM)
Replicate #1 (RPKM)
0
10
20
30
40


pol II ChIP-seq +TFIIS$_{DN}$
r = 0.98
N = 12,187
Replicate #3 (RPKM)
Replicate #2 (RPKM)
0
10
20
30
40


NET-seq Control
r = 0.97
N = 9,841
Replicate #2 (RPKM)
Replicate #1 (RPKM)
0
20
40
60
0
25
50
75

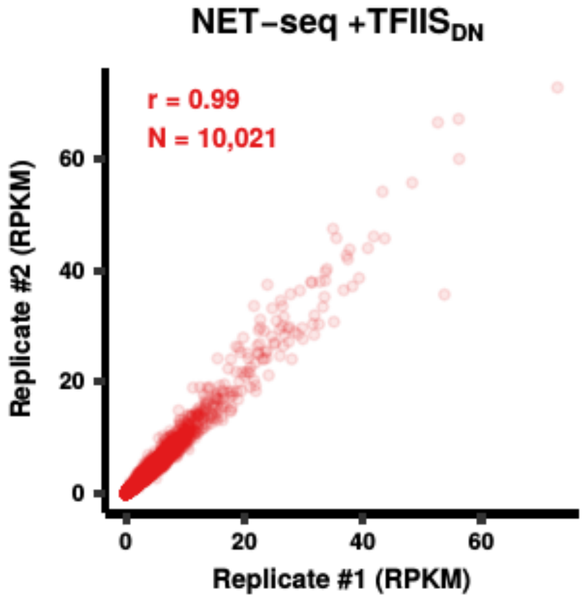
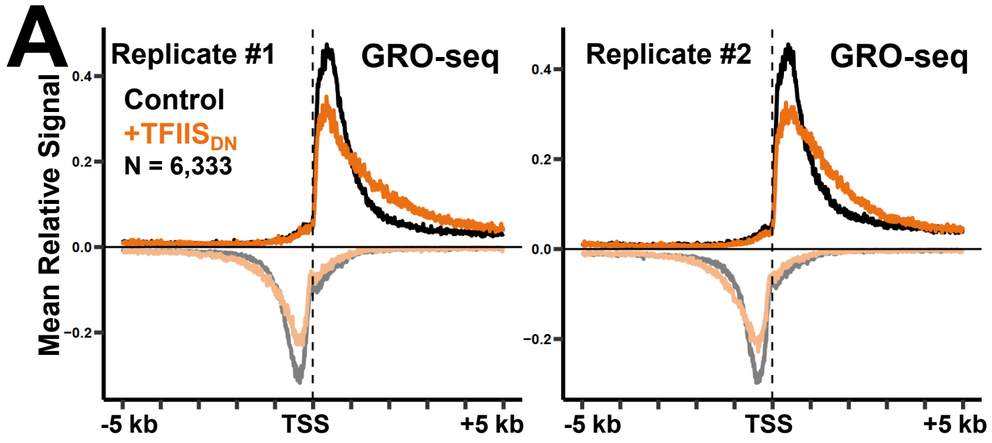
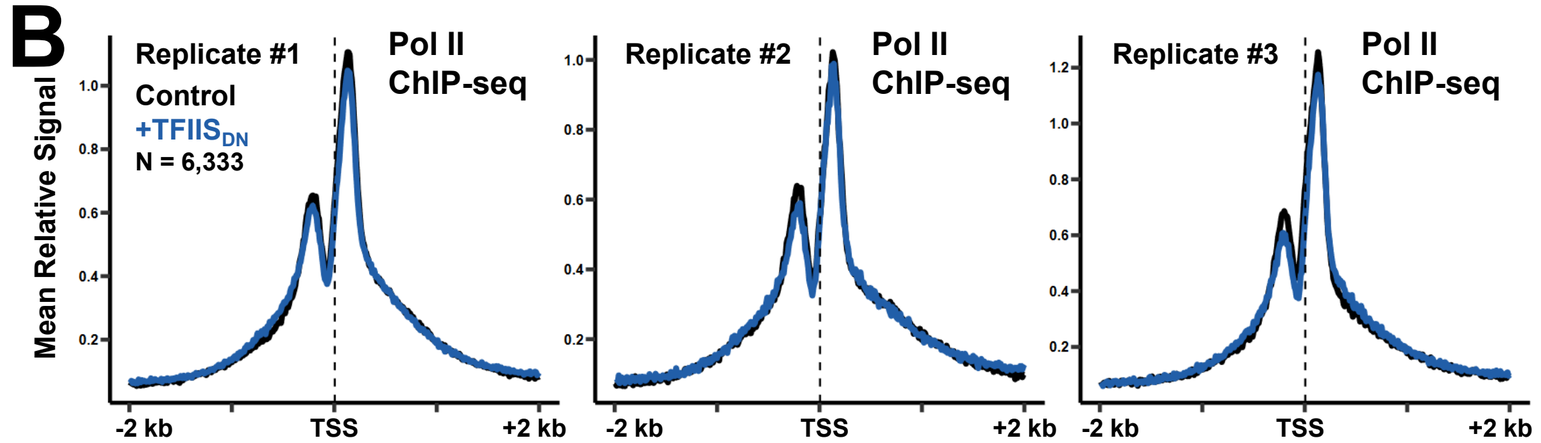

NET-seq +TFIIS$_{DN}$
r = 0.99
N = 10,021
Replicate #2 (RPKM)
Replicate #1 (RPKM)
0
20
40
60

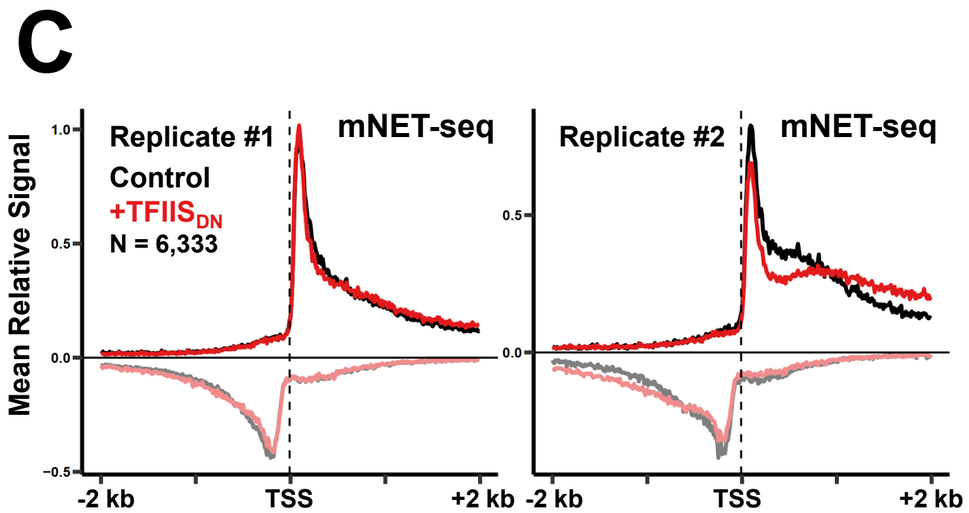
Fig. S2

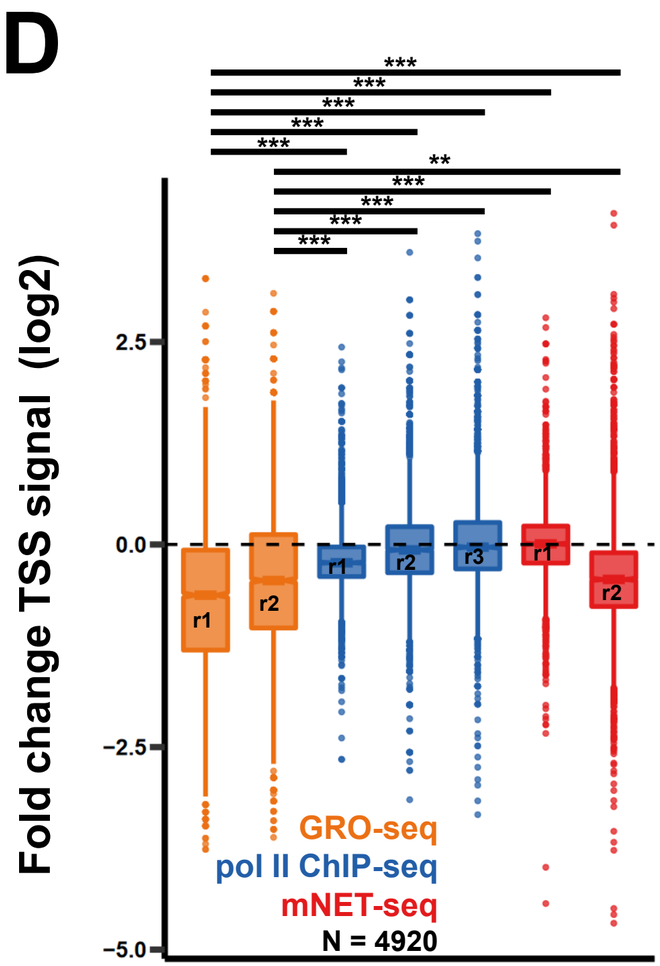
Figure S2, related to figures 1 and 2.

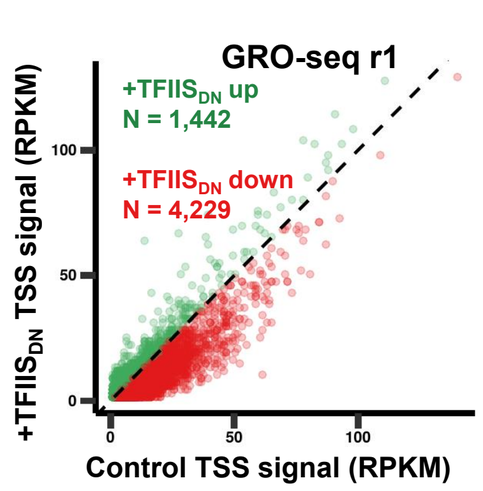
Scatter plots comparing the signal within each gene for each biological replicate used in Fig 1. Genes are included that are >1 kb long, separated by >2 kb, have at least one mapped read and are present in both datasets being compared. The Pearson correlation coefficient is shown.

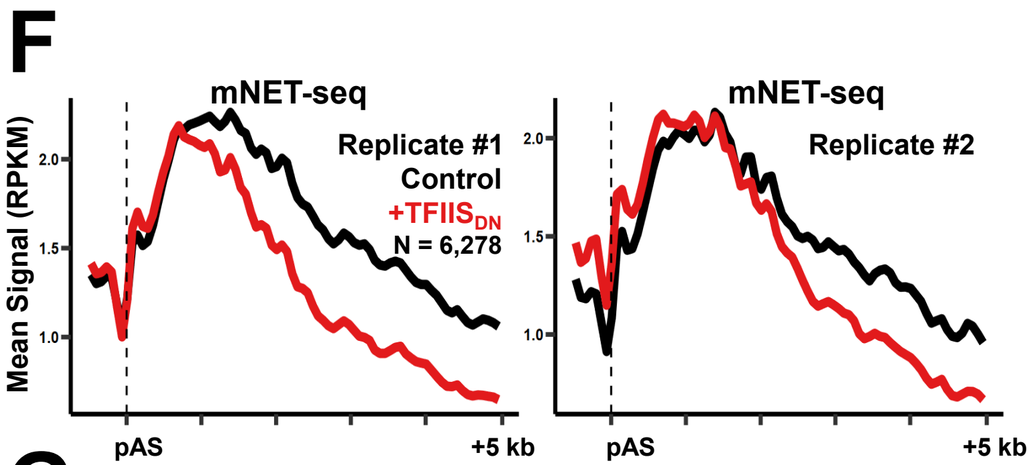

A
Replicate #1
GRO-seq
Control
+TFIIS$_{DN}$
N = 6,333
Replicate #2
GRO-seq
Mean Relative Signal
-5 kb
TSS
+5 kb

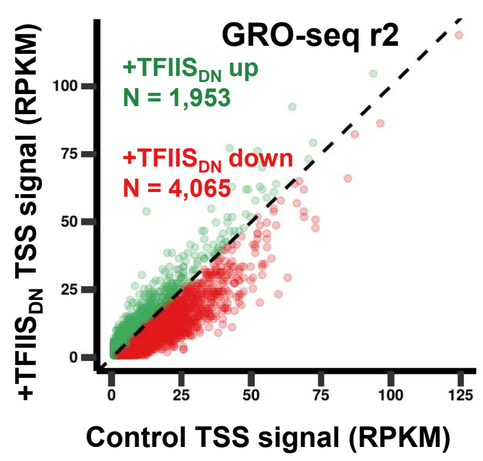

B
Replicate #1
Pol II
ChIP-seq
Control
+TFIIS$_{DN}$
N = 6,333
Replicate #2
Replicate #3
Mean Relative Signal
-2 kb
TSS
+2 kb


C
Replicate #1
mNET-seq
Control
+TFIIS$_{DN}$
N = 6,333
Replicate #2
Mean Relative Signal
-2 kb
TSS
+2 kb


D
Fold change TSS signal (log2)
GRO-seq
pol II ChIP-seq
mNET-seq
N = 4920


E
GRO-seq r1
+TFIIS$_{DN}$ up
N = 1,442
+TFIIS$_{DN}$ down
N = 4,229
+TFIIS$_{DN}$ TSS signal (RPKM)
Control TSS signal (RPKM)


F
mNET-seq
Replicate #1
Control
+TFIIS$_{DN}$
N = 6,278
Replicate #2
Mean Signal (RPKM)
pAS
+5 kb


GRO-seq r2
+TFIIS$_{DN}$ up
N = 1,953
+TFIIS$_{DN}$ down
N = 4,065
+TFIIS$_{DN}$ TSS signal (RPKM)
Control TSS signal (RPKM)

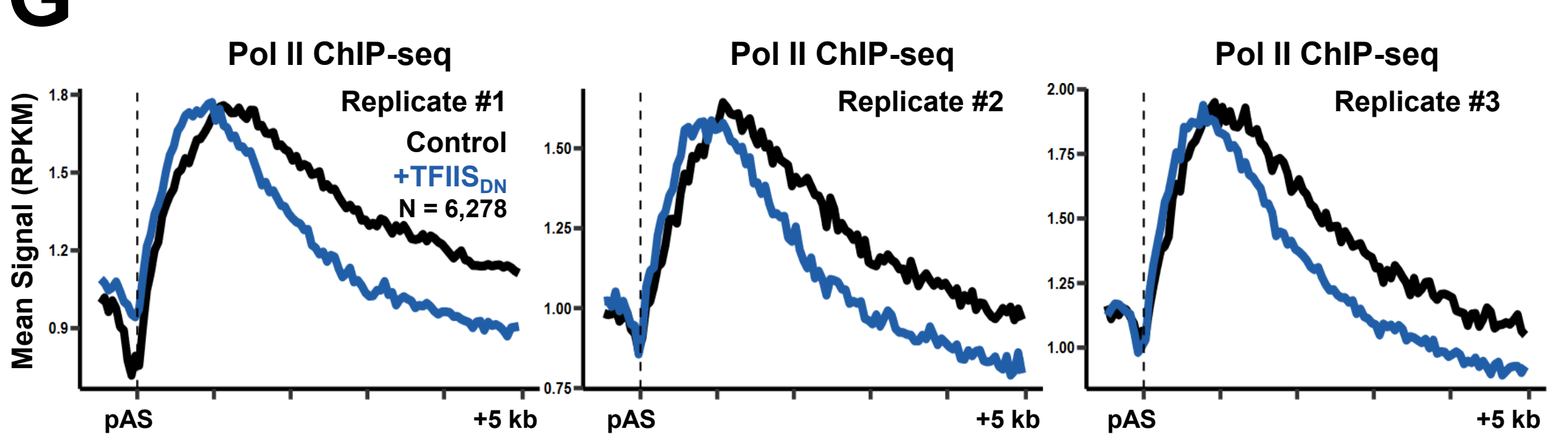

G
Pol II ChIP-seq
Replicate #1
Control
+TFIIS$_{DN}$
N = 6,278
Replicate #2
Replicate #3
Mean Signal (RPKM)
pAS
+5 kb

Fig. S3

Figure S3, related to figures 1 and 2.

(A) Metaplots of relative GRO-seq signals (10 bp bins) for each biological replicate used in Fig. 1A for genes >5 kb long and >2kb separated. The relative signal for each gene was calculated by dividing the signal in each bin by the total signal within the region plotted. The relative signal was then averaged across the plotted genes. Negative values correspond to anti-sense signal.

(B) Metaplots of relative pol II ChIP-seq signals as in B for each biological replicate used in Fig. 1B.

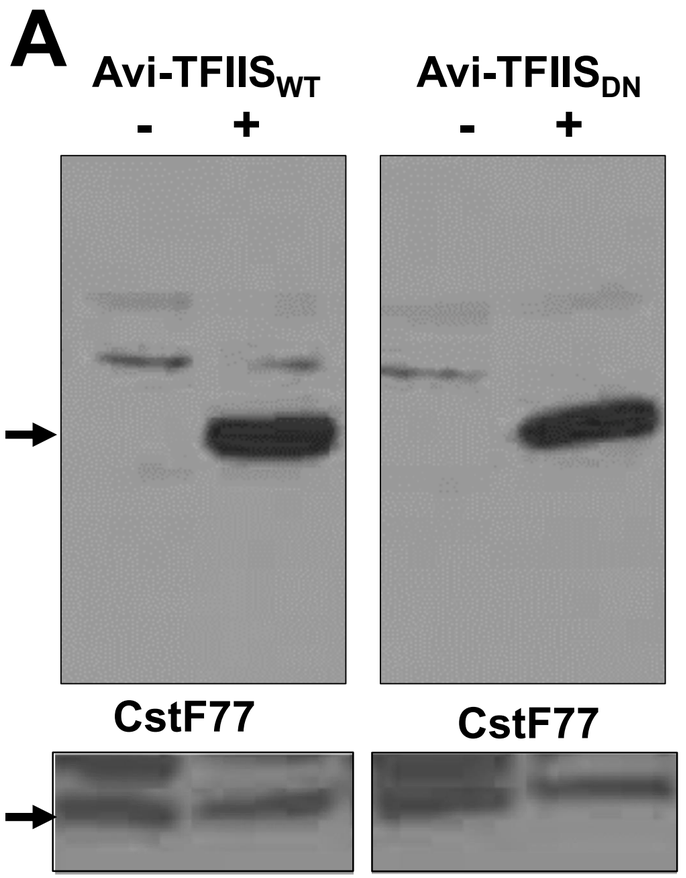
(C) Metaplots of relative pol II mNET-seq signals as in B for each biological replicate used in Fig. 1C.

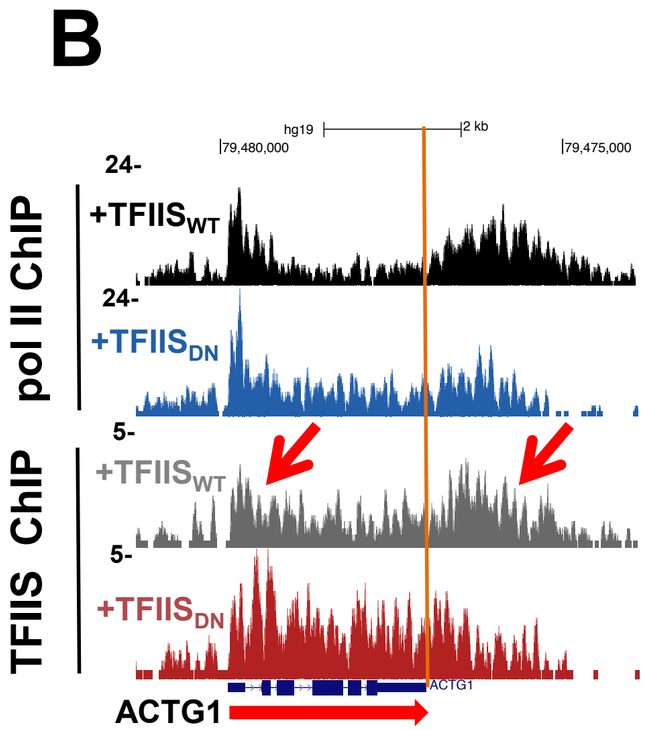
(D) Comparison of the fold change in GRO-seq, pol II ChIP-seq, and mNET-seq signal after expression of $TFIIS_{DN}$ for the region from the TSS to +300 bp for each biological replicate used in Fig. 1D for genes >1 kb long and separated by >2 kb. ** $p < 0.01$, *** $p < 10^{-3}$ Welch two sample t-test, corrected for multiple testing using the Bonferroni-Holm method.

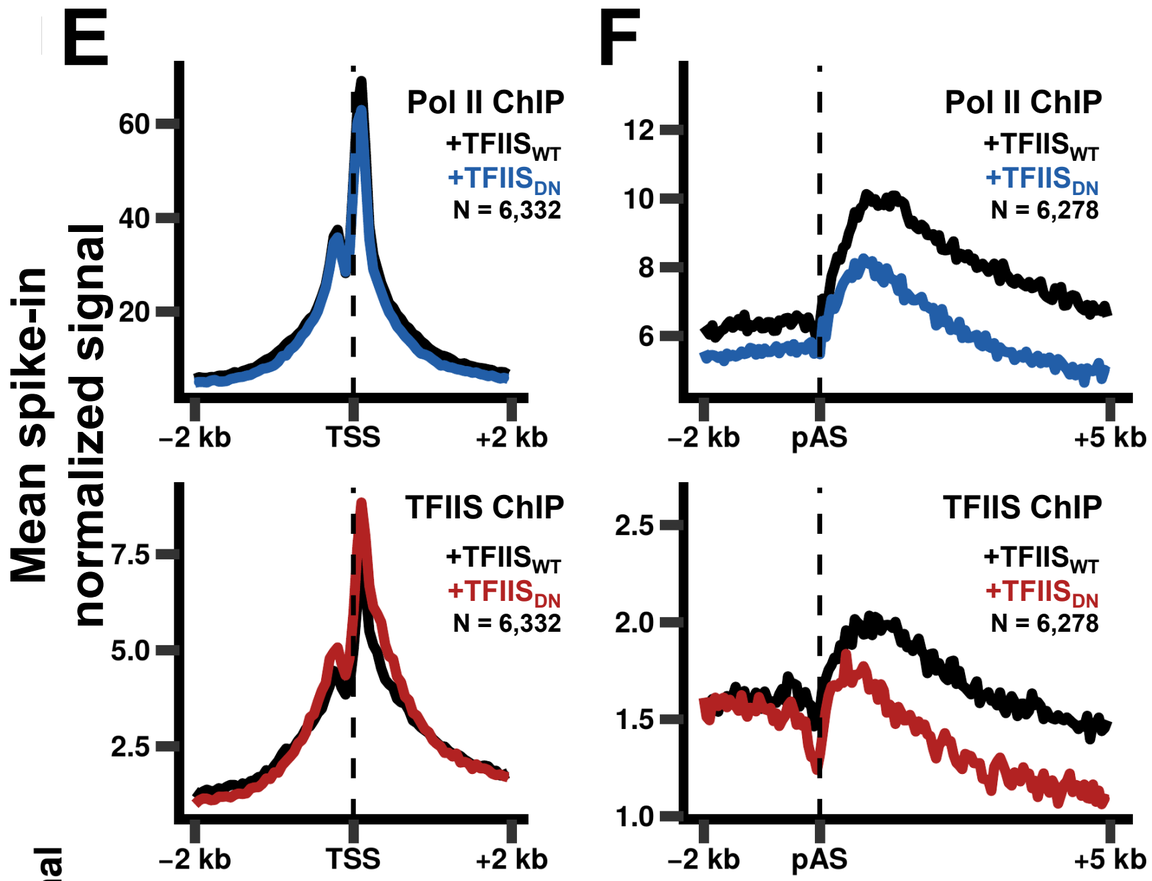
(E) Most genes show a reduction in GRO-seq signal in the region downstream of the TSS. Scatter plots comparing GRO-seq signal in the region from the TSS to +300 bp downstream for uninduced cells and cells expressing $TFIIS_{DN}$. Genes are shown that are >1 kb long, >2 kb separated, contain at least one mapped read in the region, and are present in both datasets.

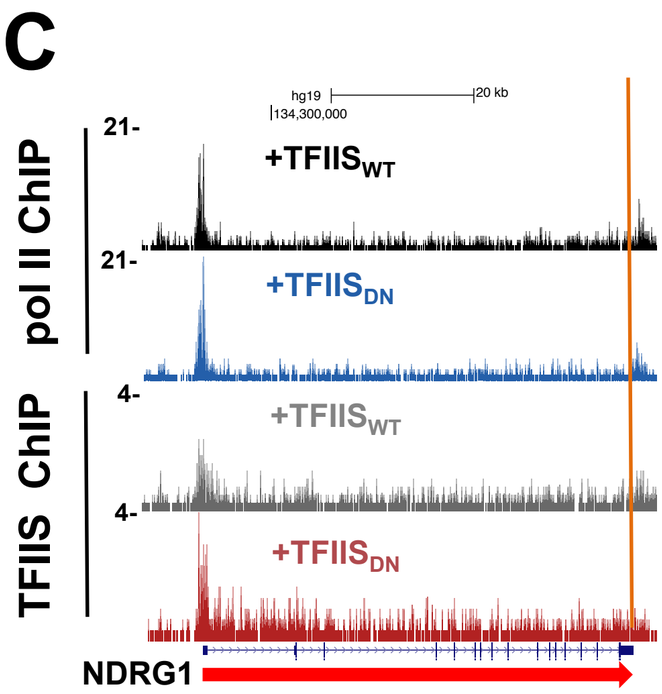
(F) Metaplots of mNET-seq signals (50 bp bins) for each biological replicate used in Fig. 2D. Data were plotted for genes >1 kb long and separated by >5 kb.

(G) Metaplots of pol II ChIP-seq signals (50 bp bins) for each biological replicate used in Fig. 2B for genes shown in F.


A
Avi-TFIIS$_{WT}$
- +
Avi-TFIIS$_{DN}$
- +
CstF77
CstF77


B
hg19
2 kb
79,480,000
79,475,000
24-
+TFIIS$_{WT}$
24-
+TFIIS$_{DN}$
pol II ChIP
5-
+TFIIS$_{WT}$
5-
+TFIIS$_{DN}$
TFIIS ChIP
ACTG1
ACTG1


E
F
Mean spike-in normalized signal
Pol II ChIP
+TFIIS$_{WT}$
+TFIIS$_{DN}$
N = 6,332
N = 6,278
TFIIS ChIP
−2 kb
TSS
pAS
+2 kb
+5 kb
Mean TFIIS/pol II signal


C
hg19
20 kb
134,300,000
21-
+TFIIS$_{WT}$
21-
+TFIIS$_{DN}$
pol II ChIP
4-
+TFIIS$_{WT}$
4-
+TFIIS$_{DN}$
TFIIS ChIP
NDRG1

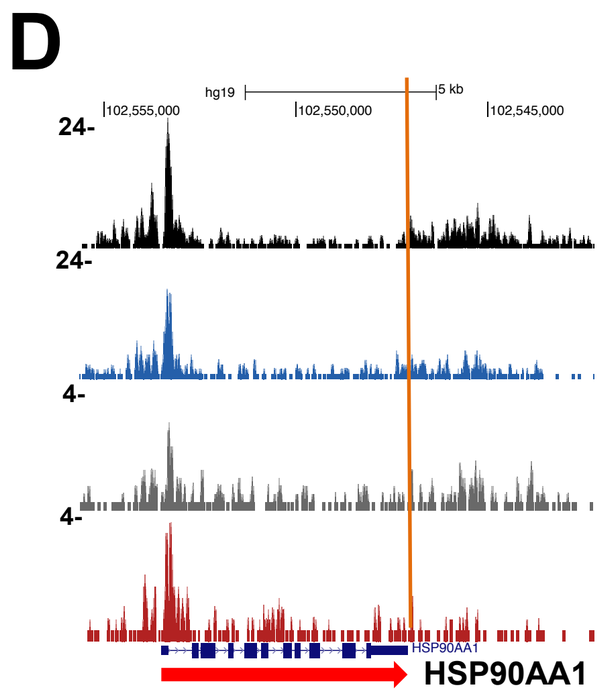

D
hg19
5 kb
102,555,000
102,550,000
102,545,000
24-
24-
4-
4-
HSP90AA1
HSP90AA1

Fig. S4

Figure S4, related to figures 1 and 2.

(A) Western blot of epitope tagged $TFIIS_{WT}$ and $TFIIS_{DN}$ with rabbit anti-Avi-tag antibody before and after induction of HMLE-RAS pCW57bla-$TFIIS_{WT}$ and -$TFIIS_{DN}$ cells with 2 μg /ml doxycycline for 24 hrs.

(B-D) UCSC genome browser screen shots of anti-pol II (pan CTD) and anti TFIIS (Avitag) ChIP-seq signals in HMLE-RAS pCW57bla-$TFIIS_{WT}$ and -$TFIIS_{DN}$ cells induced with 2 μg /ml doxycycline for 24 hrs. Signals were normalized relative to a spike-in of yeast extract expressing Avi-tagged Rpb3 which reacts with both anti-pol II (pan CTD) and anti-Avitag antibodies.

(E) $TFIIS_{DN}$ associates with pol II near the TSS in a similar manner as wild type TFIIS. Metaplots (50 bp bins) showing quantitative ChIP-seq signal for pol II (anti-pan CTD) and TFIIS (anti-Avi-tag) for the region around the TSS after expression of either $TFIIS_{WT}$ or $TFIIS_{DN}$ (upper and middle panels). A spike-in of yeast extract expressing Avi-tagged Rpb3 which reacts with both anti-pol II CTD and anti-Avi-tag antibodies was used to normalize anti-pol II and anti-TFIIS signals for genes shown in Fig. 1F. Average TFIIS/pol II ratios are shown in the lower panel. TFIIS/pol II ratios were calculated for each gene before averaging across all plotted genes.

(F) $TFIIS_{DN}$ associates with pol II near the poly(A) site in a similar manner as wild type TFIIS. Metaplots (50 bp bins) showing quantitative ChIP-seq signal for pol II and TFIIS as in A for the region around the poly(A) site for genes shown in Fig. 2E. Average TFIIS/pol II ratios are shown in the lower panel as in A.

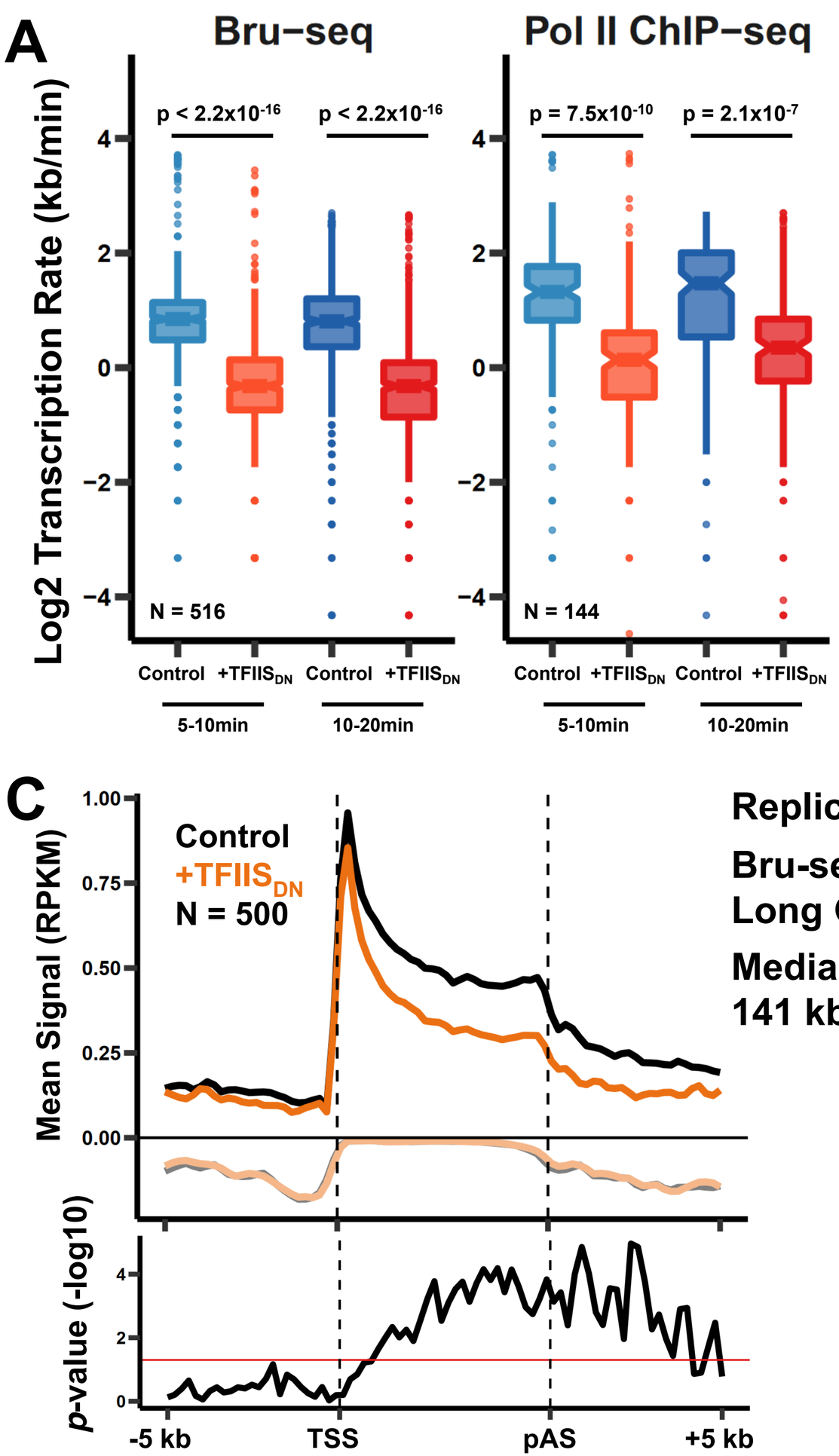

A
Bru−seq
Pol II ChIP−seq
Log2 Transcription Rate (kb/min)
p < 2.2x10-16
p < 2.2x10-16
p = 7.5x10-10
p = 2.1x10-7
N = 516
N = 144
Control
+TFIIS_DN
5-10min
10-20min
B
Control
+TFIIS_DN
N = 500
Replicate #1
Bru-seq
Long Genes
Median:
141 kb
Mean Signal (RPKM)
p-value (-log10)
-5 kb
TSS
pAS
+5 kb
C
Replicate #2
D
Control
+TFIIS_DN
N = 9,912
Bru-seq
All genes
Median:
35 kb

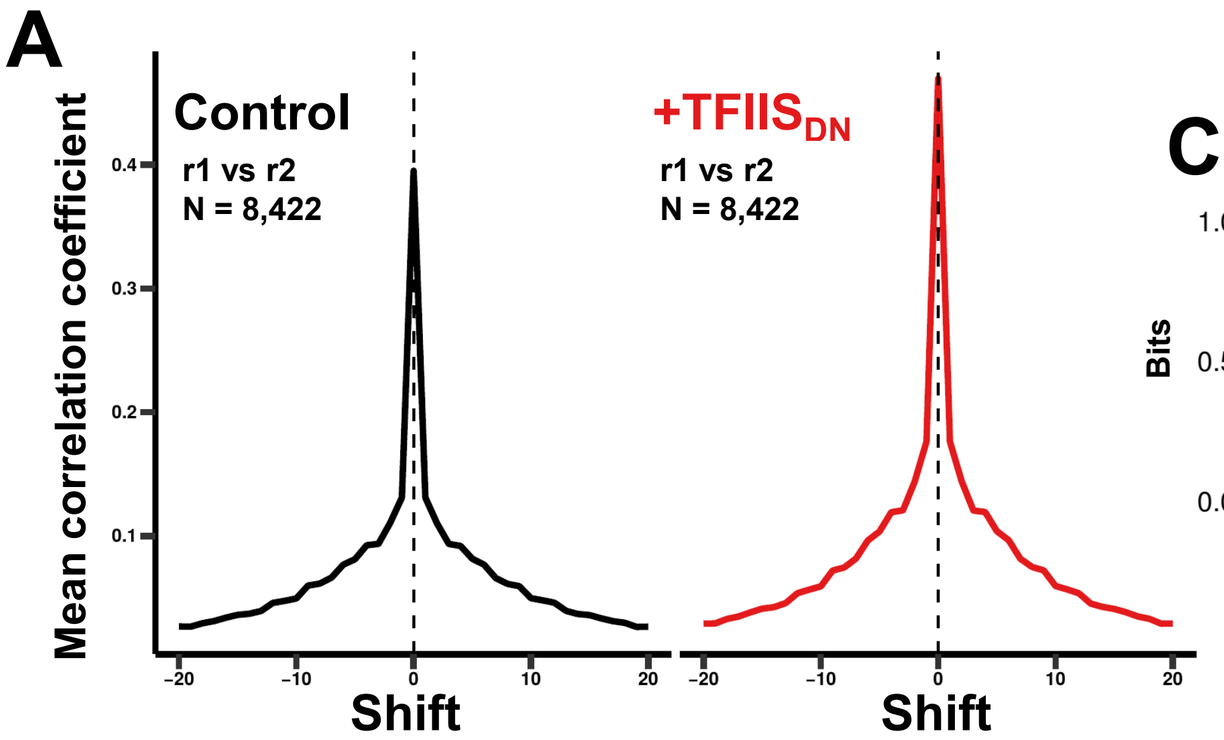
Fig. S5

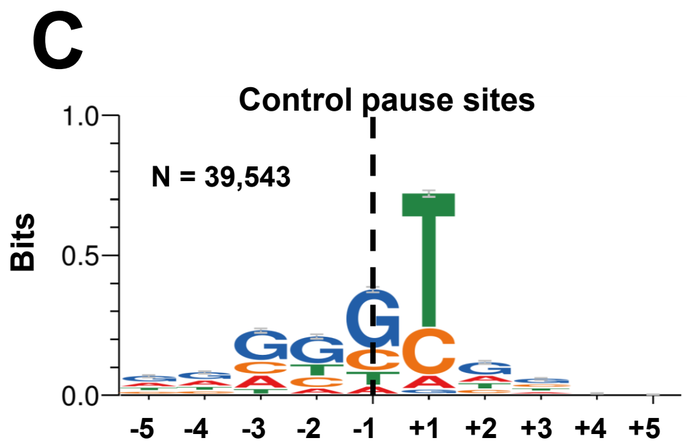
Figure S5, related to figure 3.

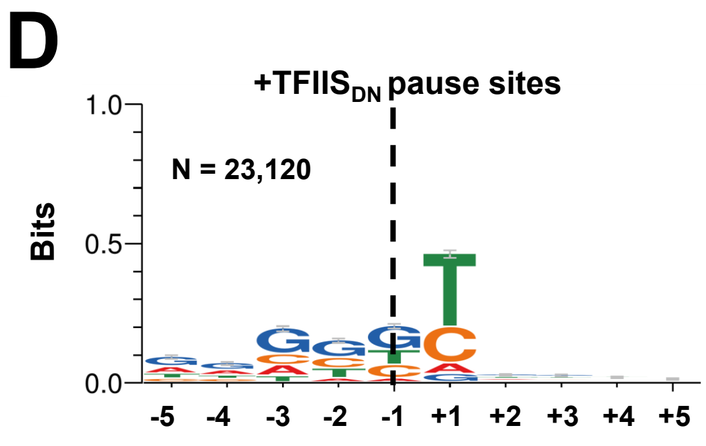
(A) Transcription rates were calculated using Bru-seq data shown in Fig. 3A (left panel) or pol II ChIP-seq data shown in Fig. 3B (right panel) for uninduced and $TFIIS_{DN}$-expressing cells by comparing the 5 min and 10 min or 10 min and 20 min timepoints after DRB removal.

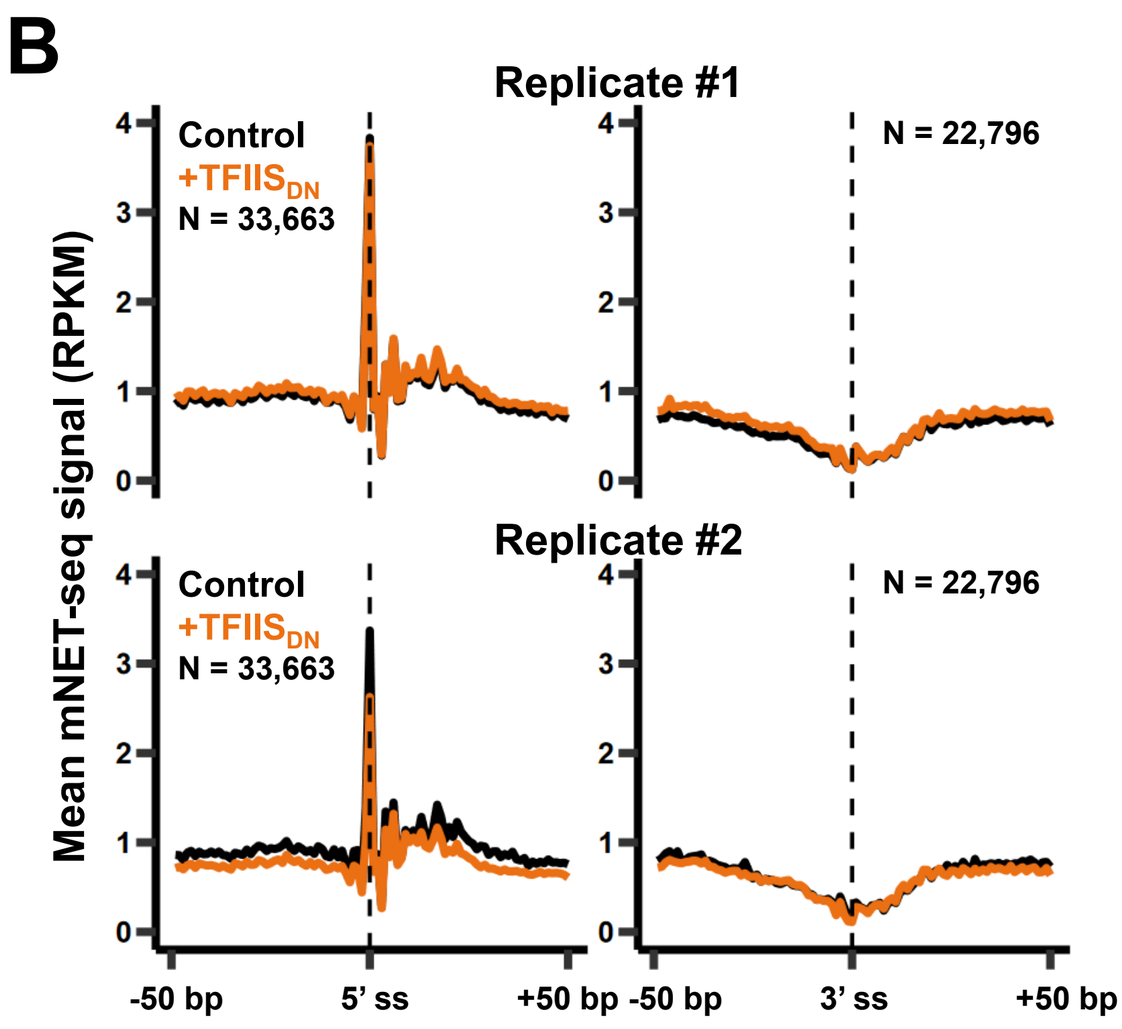
(B, C) Metaplots of Bru-seq signals for uninduced and $TFIIS_{DN}$-expressing cells for each biological replicate used in Fig. 3E. Data were plotted for 200bp bins for the regions from -5 kb to the TSS and from the poly(A) site to +5 kb downstream, and 30 bins of variable length for the region from the TSS to the poly(A) site. Negative values correspond to anti-sense signal. Genes from the 3^{rd} and 4^{th} quartiles from Fig. 3D are plotted. p-values were calculated by comparing the signals for uninduced and $TFIIS_{DN}$-expressing cells at each bin using the Welch two sample t-test (bottom panel).

(D) Metaplots of Bru-seq signals for genes >1 kb long and >2 kb separated. The mean signal was calculated for each bin for each individual replicate. The mean signal was then averaged for the replicates and plotted, with the shaded region representing the SEM for two biological replicates. Negative values correspond to anti-sense signal.


A
Control
r1 vs r2
N = 8,422
+TFIIS$_{DN}$
r1 vs r2
N = 8,422
Mean correlation coefficient
0.4
0.3
0.2
0.1
−20
−10
0
10
20
Shift
Shift


C
Control pause sites
N = 39,543
Bits
1.0
0.5
0.0
-5 -4 -3 -2 -1 +1 +2 +3 +4 +5


D
+TFIIS$_{DN}$ pause sites
N = 23,120
Bits
1.0
0.5
0.0
-5 -4 -3 -2 -1 +1 +2 +3 +4 +5


B
Replicate #1
Control
+TFIIS$_{DN}$
N = 33,663
N = 22,796
Replicate #2
Control
+TFIIS$_{DN}$
N = 33,663
N = 22,796
Mean mNET-seq signal (RPKM)
4
3
2
1
0
-50 bp
5' ss
+50 bp
-50 bp
3' ss
+50 bp

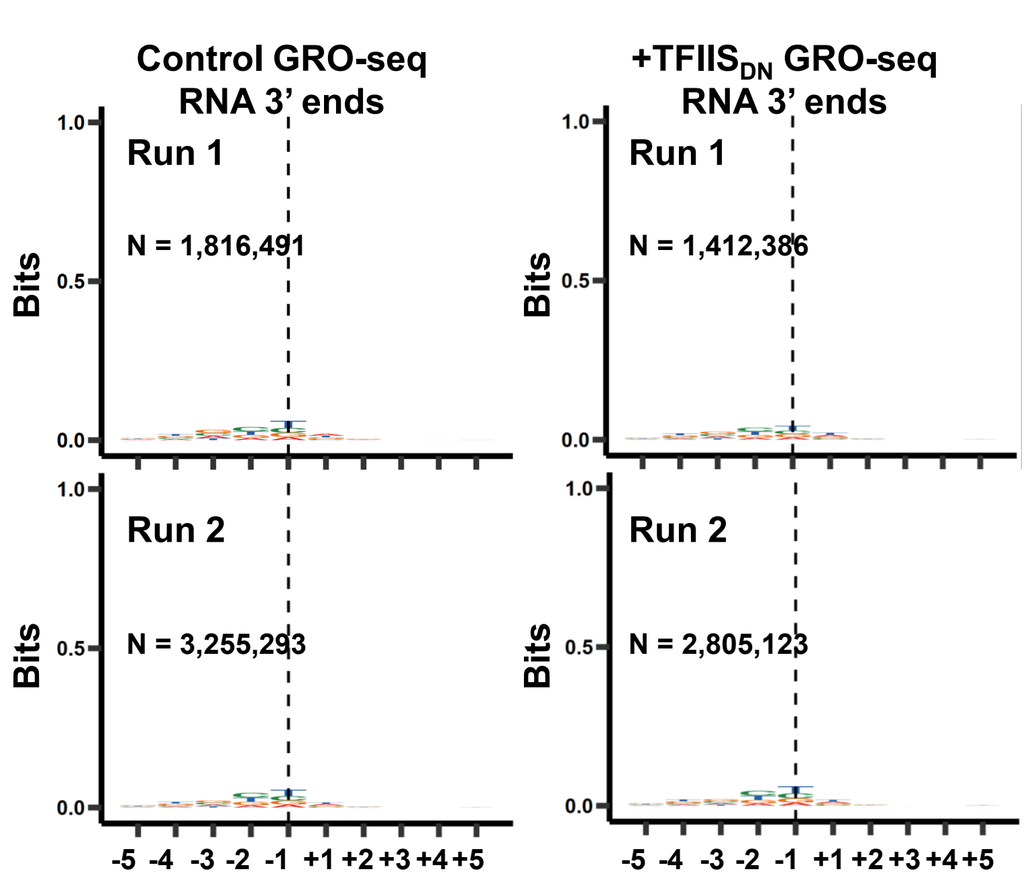

E
Control GRO-seq
RNA 3' ends
+TFIIS$_{DN}$ GRO-seq
RNA 3' ends
Run 1
N = 1,816,491
N = 1,412,386
Run 2
N = 3,255,293
N = 2,805,123
Bits
1.0
0.5
0.0
-5 -4 -3 -2 -1 +1 +2 +3 +4 +5

Fig. S6

Figure S6, related to figure 4.

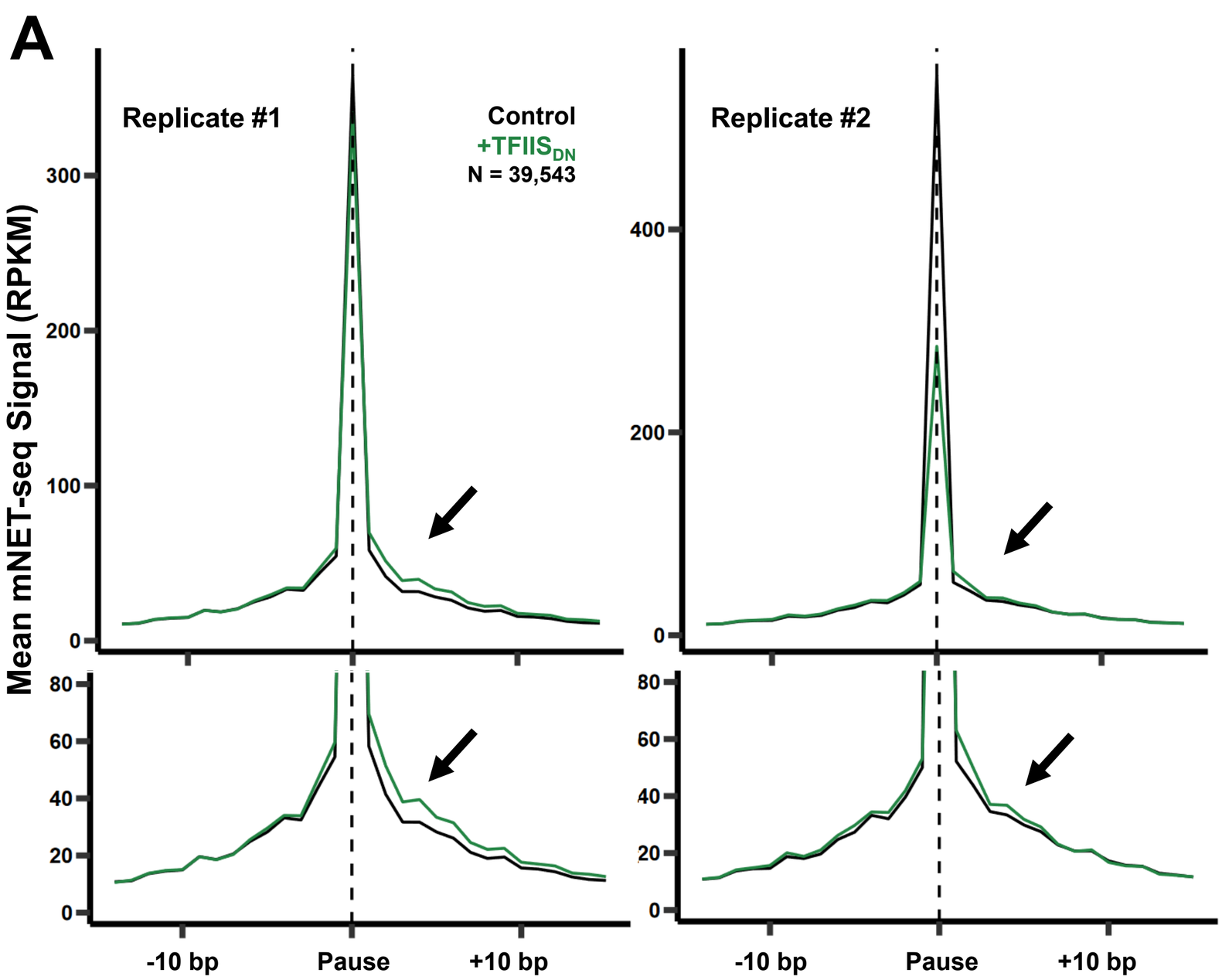
(A) Mean cross-correlation between replicate mNET-seq datasets for genes containing at least one pause site shown in Fig 4E. The analysis was performed by calculating the Pearson's correlation coefficient for each gene after shifting one of the datasets by the amount indicated on the x-axis. The correlation coefficients were then averaged over the entire gene list.

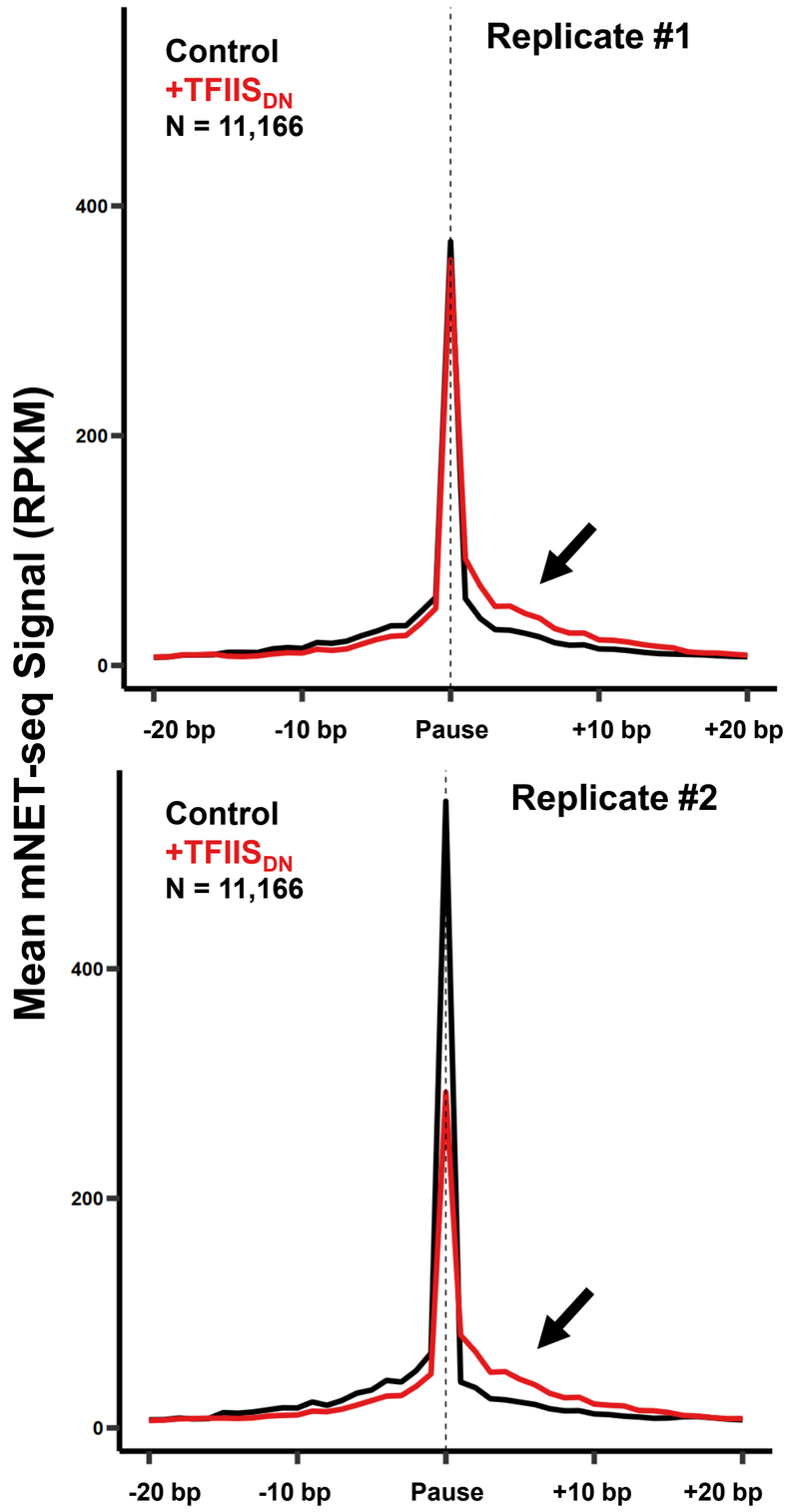
(B) Metaplots of mNET-seq signals around 5' and 3' splice sites (ss, 1 bp bins) for each biological replicate used in Fig. 4C for non-overlapping genes >1 kb long.

(C) Sequence logo generated using WebLogo 3 (http://weblogo.threeplusone.com/create.cgi) to include error bars for pause sites shown in Fig. 4D and E.

(D) Sequence logos for the region surrounding pauses identified in both +TFIIS$_{DN}$ datasets and absent in both uninduced datasets.

(E) Sequence logos for the region surrounding mapped RNA 3' ends from uninduced and +TFIIS$_{DN}$ GRO-seq libraries that were prepared using the same library prep protocol used to generate mNET-seq libraries. Only reads that were long enough to include a 3' adapter and that mapped to positions separated by at least 10 bp were included in the analysis.


A
Replicate #1
Control
+TFIIS$_{DN}$
N = 39,543
Replicate #2
Mean mNET-seq Signal (RPKM)
300
200
100
0
400
200
0
80
60
40
20
0
-10 bp
Pause
+10 bp


B
Control
+TFIIS$_{DN}$
N = 11,166
Replicate #1
Replicate #2
Mean mNET-seq Signal (RPKM)
400
200
0
-20 bp
-10 bp
Pause
+10 bp
+20 bp

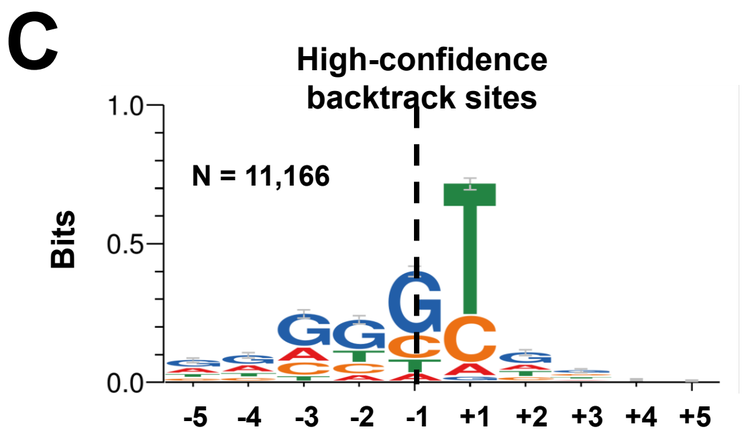

C
High-confidence
backtrack sites
N = 11,166
Bits
1.0
0.5
0.0
-5
-4
-3
-2
-1
+1
+2
+3
+4
+5

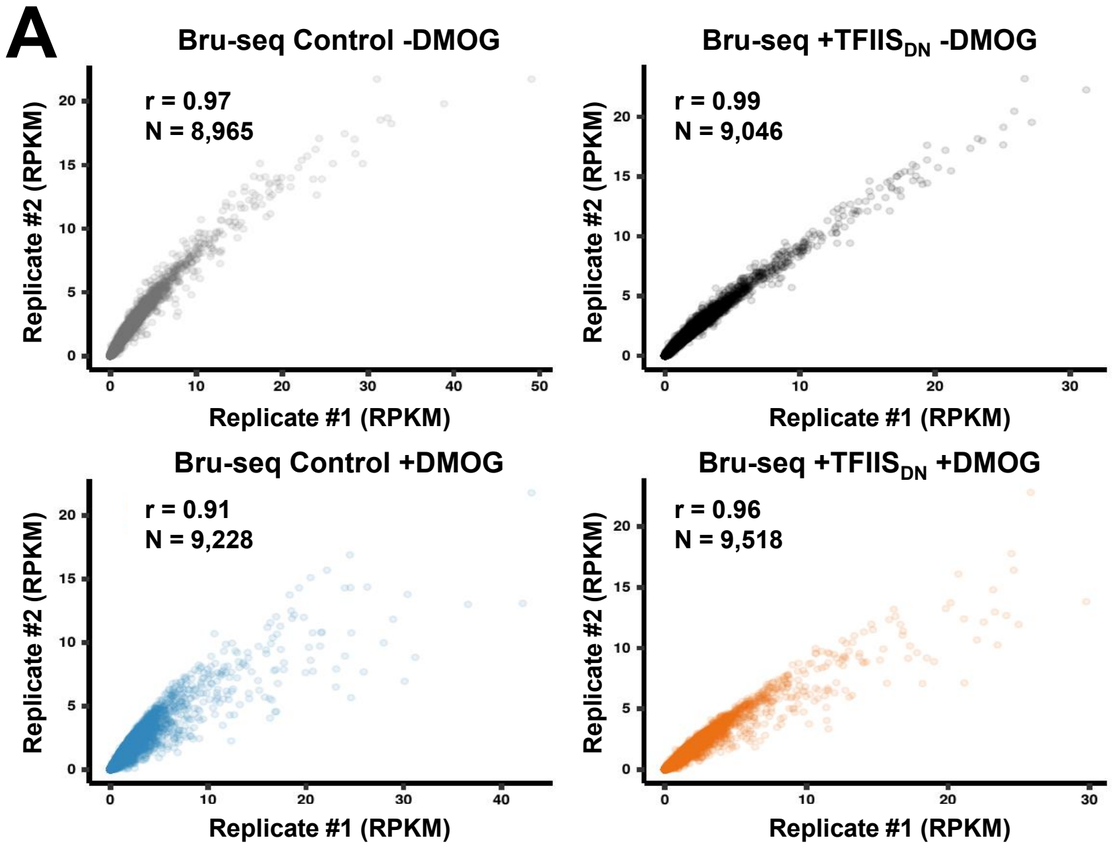
Fig. S7

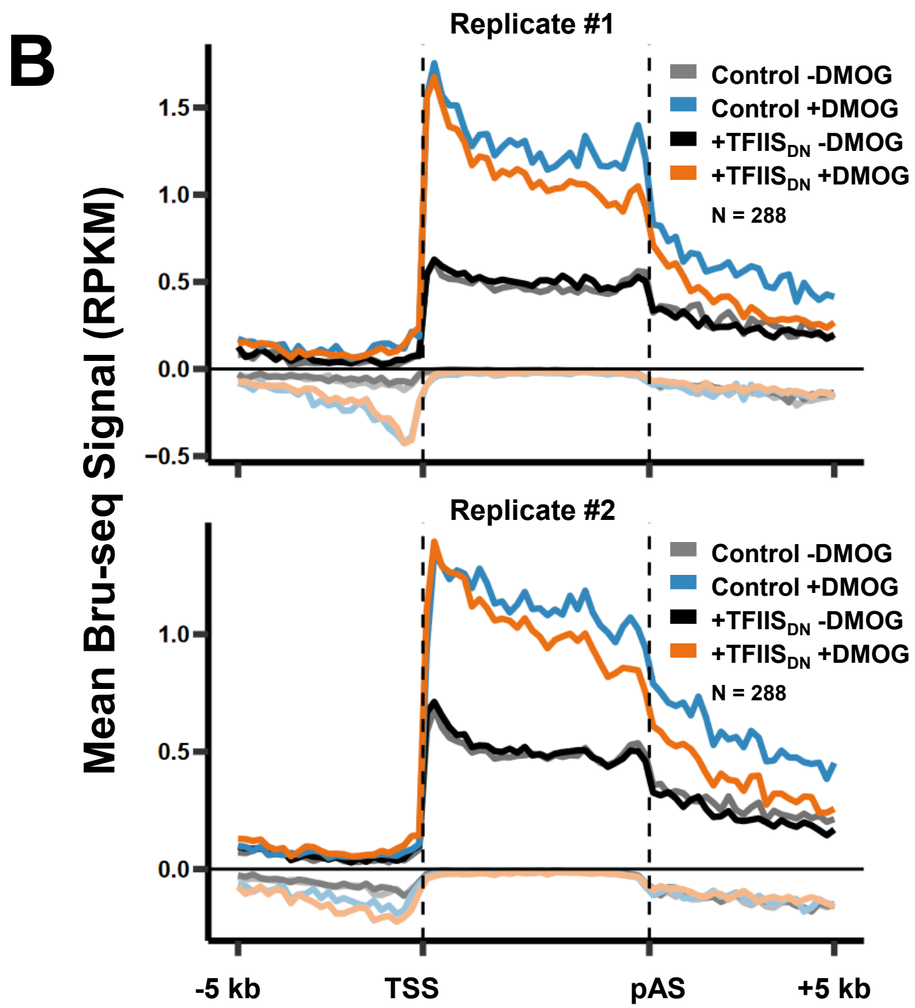
Figure S7, related to figures 4 and 5.

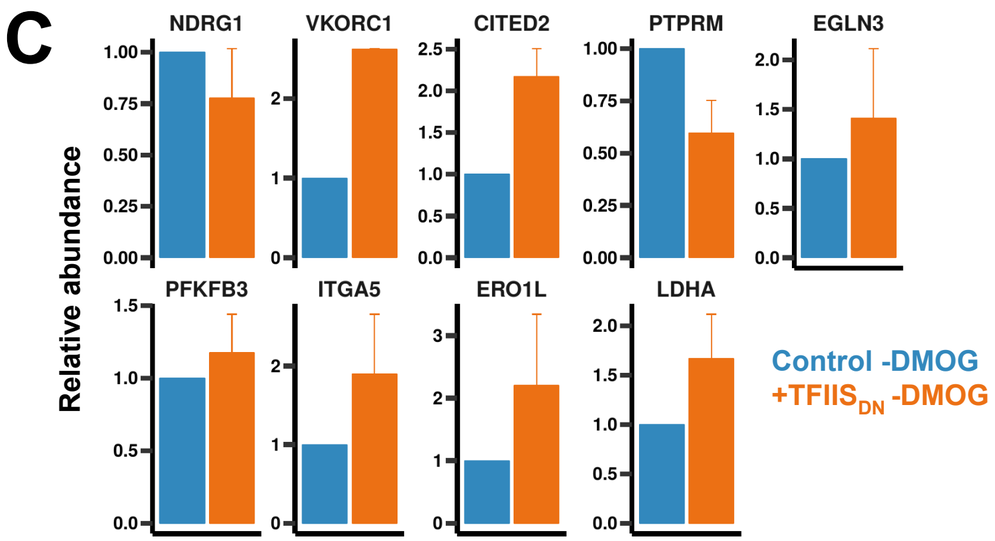
(A) Metaplots of mNET-seq signals (1 bp bins) for uninduced and $TFIIS_{DN}$-expressing cells for each biological replicate used in Fig. 4E. The lower panel shows a zoomed in view of the plots. Data were plotted for pause sites present in both control datasets, separated by at least 30bp and containing signal within 15 bp upstream or downstream of the pause (39,543 pauses).

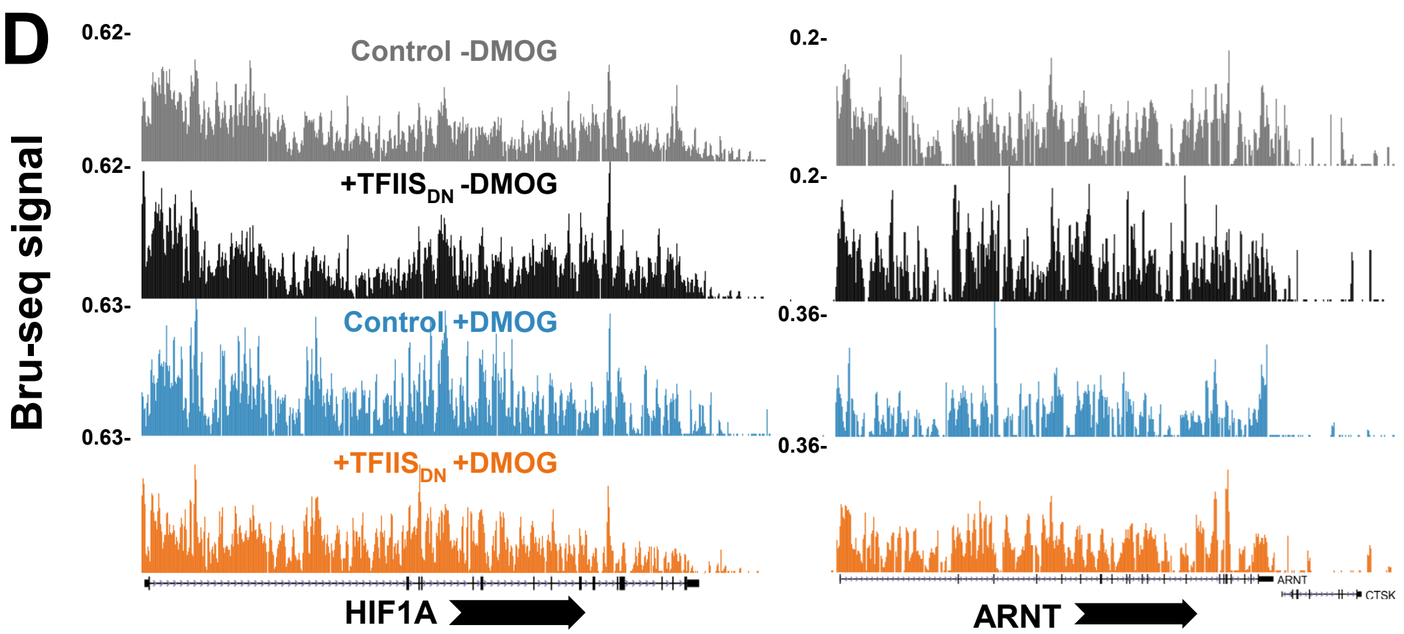
(B) Metaplots of mNET-seq signals (1 bp bins) around pauses identified as high confidence backtracking sites (11,166 sites) for uninduced and $TFIIS_{DN}$-expressing cells for each biological replicate used in Fig. 5B. Data were plotted for high confidence backtrack sites identified in the region from the TSS to +5 kb downstream of the poly(A) site.

(C) Sequence logo for the region surrounding high confidence backtrack sites (11,166 sites).


A
Bru-seq Control -DMOG
r = 0.97
N = 8,965
Bru-seq +TFIIS$_{DN}$ -DMOG
r = 0.99
N = 9,046
Bru-seq Control +DMOG
r = 0.91
N = 9,228
Bru-seq +TFIIS$_{DN}$ +DMOG
r = 0.96
N = 9,518
Replicate #2 (RPKM)
Replicate #1 (RPKM)


B
Replicate #1
Replicate #2
Mean Bru-seq Signal (RPKM)
Control -DMOG
Control +DMOG
+TFIIS$_{DN}$ -DMOG
+TFIIS$_{DN}$ +DMOG
N = 288
-5 kb
TSS
pAS
+5 kb


C
NDRG1
VKORC1
CITED2
PTPRM
EGLN3
PFKFB3
ITGA5
ERO1L
LDHA
Relative abundance
Control -DMOG
+TFIIS$_{DN}$ -DMOG


D
Bru-seq signal
Control -DMOG
+TFIIS$_{DN}$ -DMOG
Control +DMOG
+TFIIS$_{DN}$ +DMOG
HIF1A
ARNT
CTSK

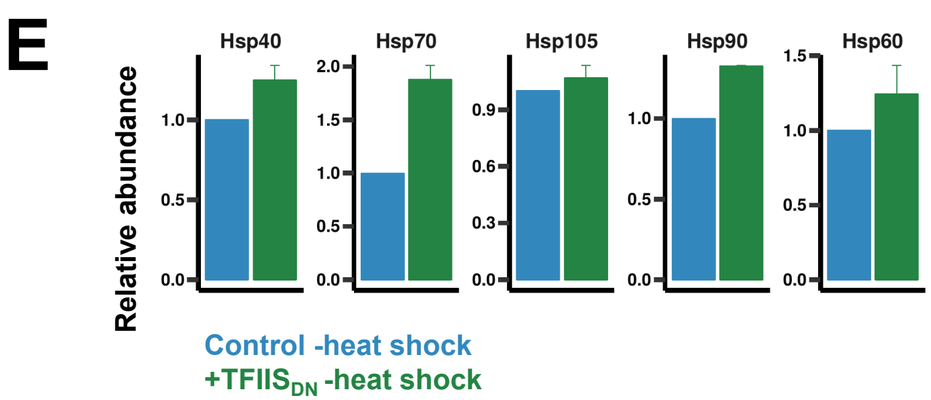

E
Hsp40
Hsp70
Hsp105
Hsp90
Hsp60
Relative abundance
Control -heat shock
+TFIIS$_{DN}$ -heat shock

Fig. S8

Figure S8, related to figure 6.

(A) Scatter plots comparing the Bru-seq signal within each gene for each biological replicate used in Fig 6A. Genes are included that are >1kb long, separated by >2kb, have at least one mapped read and are present in both datasets being compared. The Pearson correlation coefficient is shown.

(B) Metaplots of Bru-seq signals for each biological replicate used in Fig. 6A for cells treated +/- 2 mM DMOG for 16 hours. Data were plotted for 200 bp bins for the regions from -5 kb to the TSS and from the poly(A) site to +5 kb downstream, and 30 bins of variable length for the region from the TSS to the poly(A) site for genes with increased signal +DMOG (p-value < 0.01). Negative values correspond to anti-sense signal.

(C) qRT-PCR showing the relative baseline levels of uninduced hypoxia-responsive transcripts shown in Fig. 6B and C after expression of $TFIIS_{DN}$ for 24 hours. Error bars represent the SEM for at least two biological replicates.

(D) UCSC genome browser screen shots showing Bru-seq signal for HIF1A (left) and the HIF binding partner, ARNT (right) before and after expression of $TFIIS_{DN}$.

(E) qRT-PCR showing the relative baseline levels of uninduced heat shock transcripts shown in Fig. 6E after expression of $TFIIS_{DN}$ for 24 hours. Error bars represent the SEM for two biological replicates.