c-MYC empowers transcription and productive splicing of the oncogenic splicing factor Sam68 in cancer

Cinzia Caggiano[1,5], Marco Pieraccioli[1,5], Valentina Panzeri[1,2], Claudio Sette[1,3] and Pamela Bielli[1,4]

[1]*Laboratory of Neuroembryology, Fondazione Santa Lucia IRCCS, Rome, Italy;*

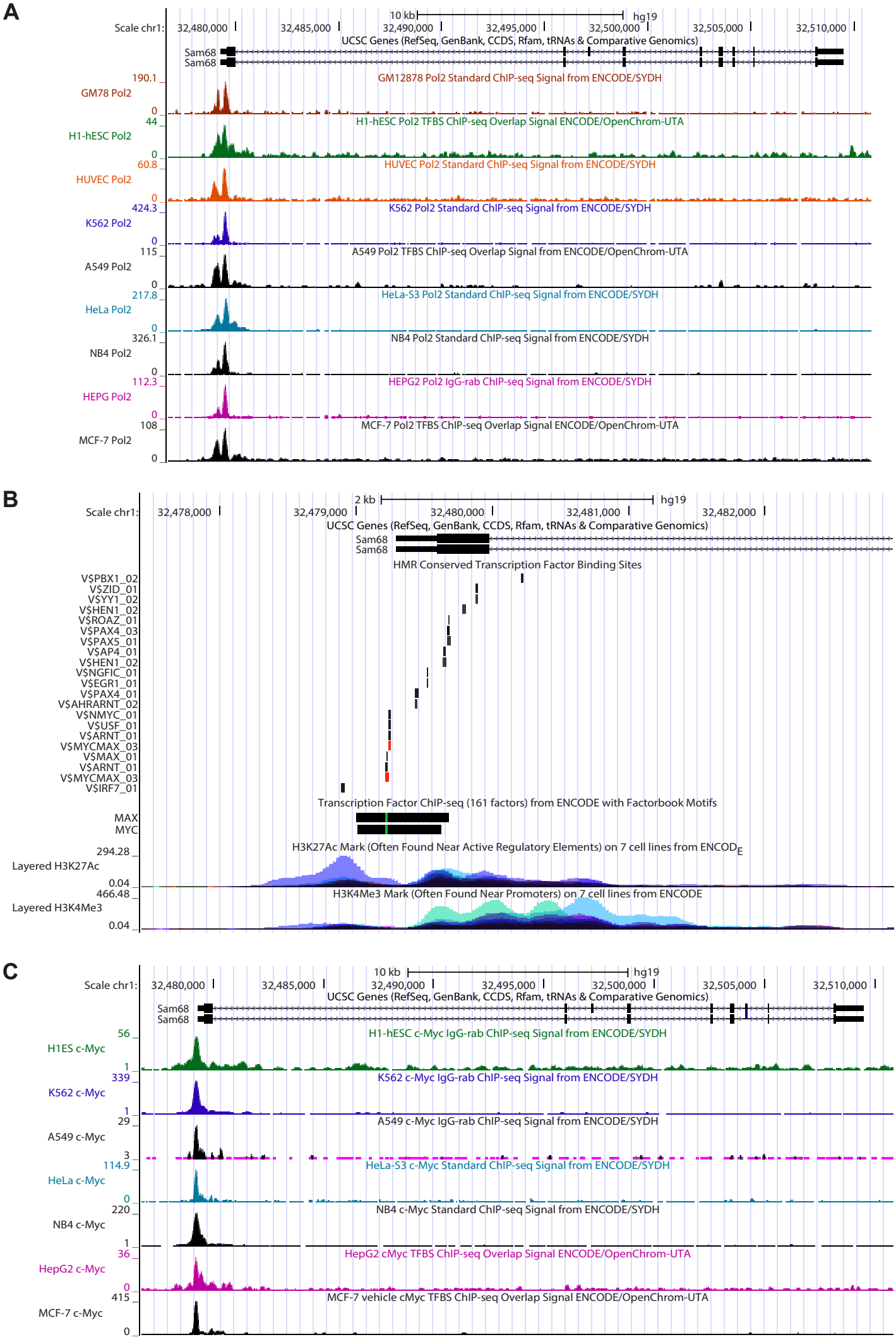
[2]*Department of science medical/chirurgic and translational medicine, University of Rome Sapienza, Rome, Italy;*

[3]*Institute of Human Anatomy and Cell Biology, Catholic University of the Sacred Hearth, Rome, Italy;*

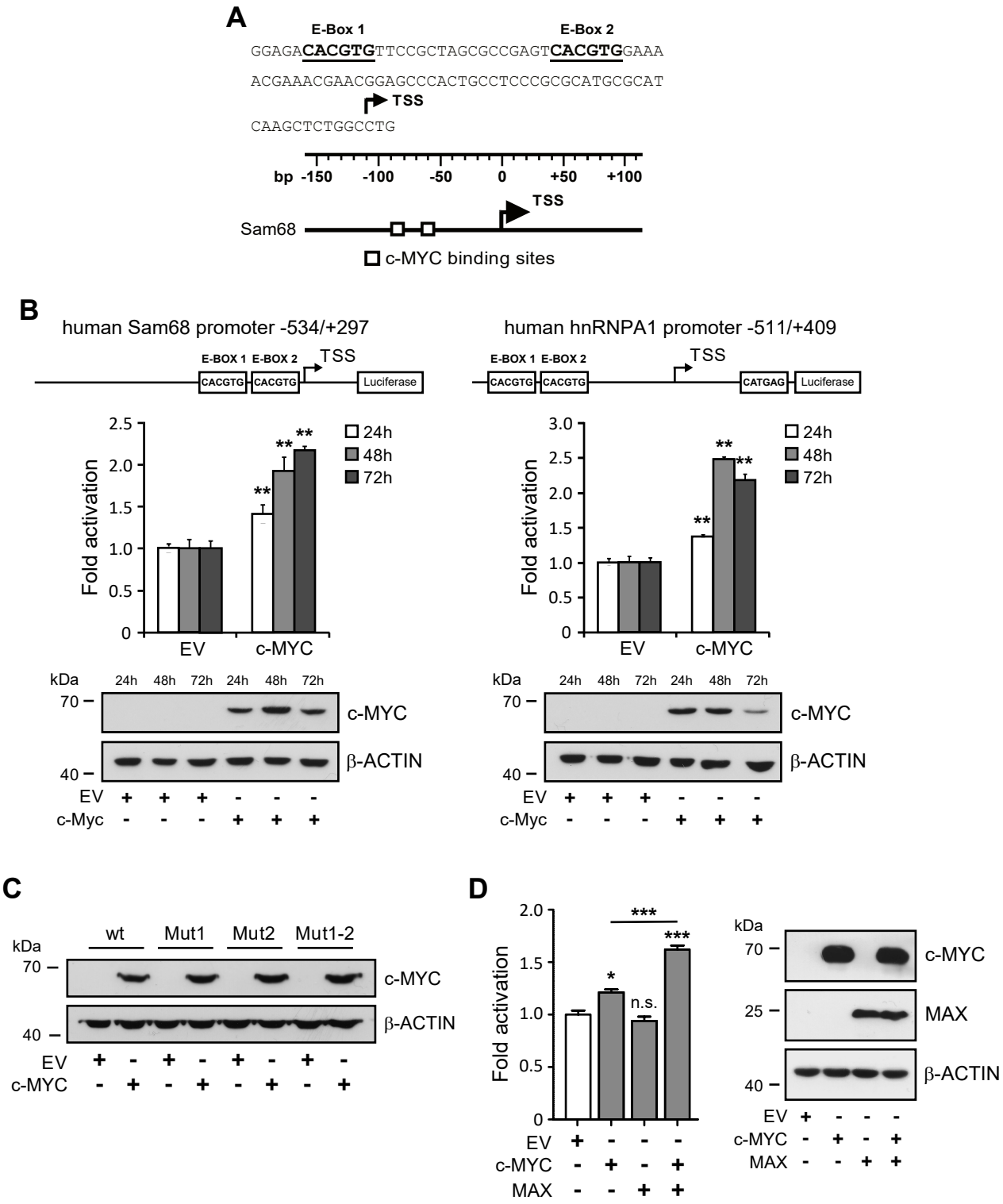
[4]*Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy.*

[5]These authors equally contributed to the work.

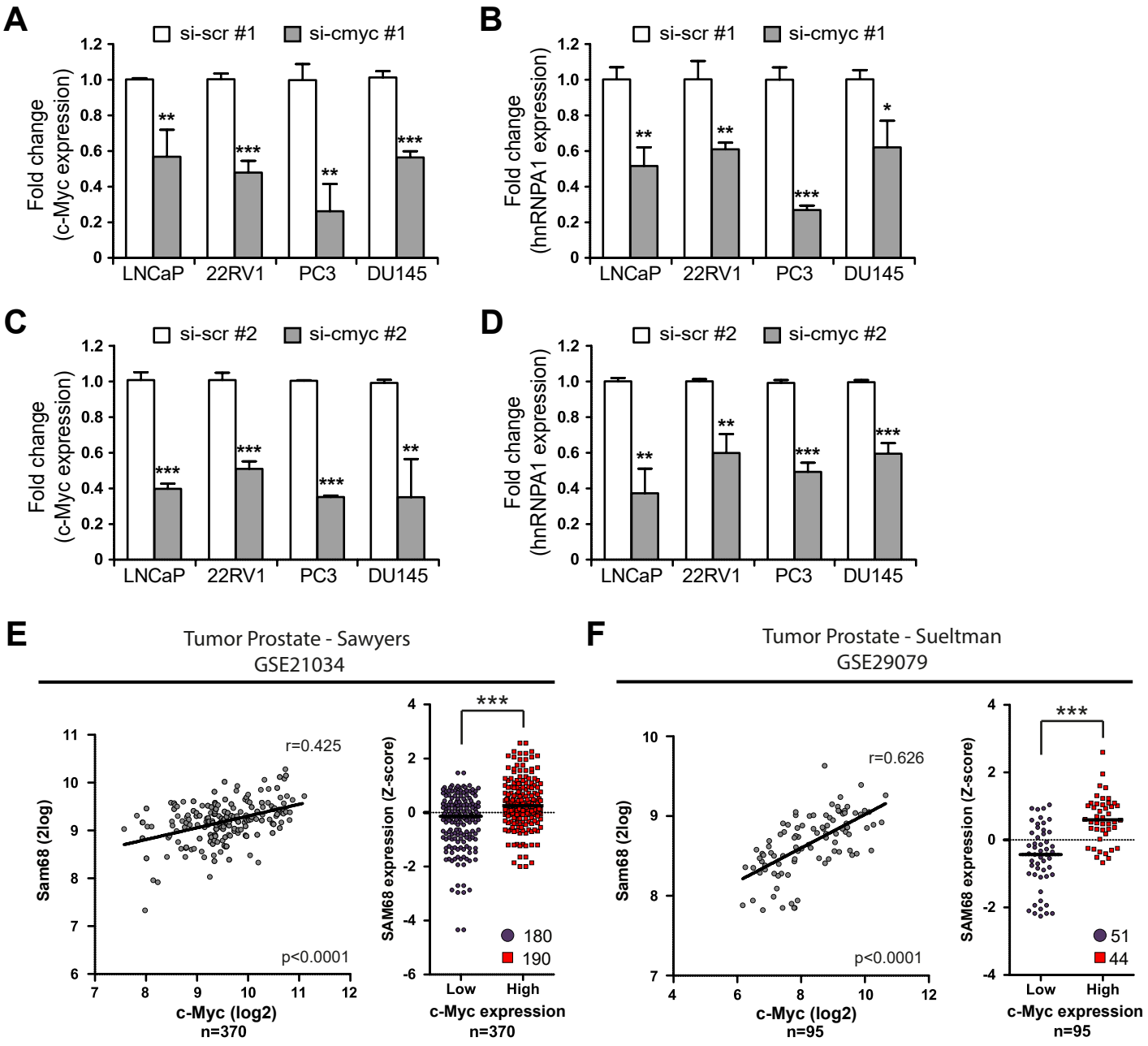
SUPPLEMENTAL INFORMATION


A
Scale chr1:
10 kb
hg19
32,480,000
32,485,000
32,490,000
32,495,000
32,500,000
32,505,000
32,510,000
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
Sam68
Sam68
GM12878 Pol2 Standard ChIP-seq Signal from ENCODE/SYDH
GM78 Pol2
H1-hESC Pol2 TFBS ChIP-seq Overlap Signal ENCODE/OpenChrom-UTA
H1-hESC Pol2
HUVEC Pol2 Standard ChIP-seq Signal from ENCODE/SYDH
HUVEC Pol2
K562 Pol2 Standard ChIP-seq Signal from ENCODE/SYDH
K562 Pol2
A549 Pol2 TFBS ChIP-seq Overlap Signal from ENCODE/OpenChrom-UTA
A549 Pol2
HeLa-S3 Pol2 Standard ChIP-seq Signal from ENCODE/SYDH
HeLa Pol2
NB4 Pol2 Standard ChIP-seq Signal from ENCODE/SYDH
NB4 Pol2
HEPG2 Pol2 IgG-rab ChIP-seq Signal from ENCODE/SYDH
HEPG Pol2
MCF-7 Pol2 TFBS ChIP-seq Overlap Signal ENCODE/OpenChrom-UTA
MCF-7 Pol2
B
Scale chr1:
2 kb
hg19
32,478,000
32,479,000
32,480,000
32,481,000
32,482,000
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
Sam68
Sam68
HMR Conserved Transcription Factor Binding Sites
V$PBX1_02
V$ZID_01
V$YY1_02
V$HEN1_02
V$ROAZ_01
V$PAX4_03
V$PAX5_01
V$AP4_01
V$HEN1_02
V$NGFIC_01
V$EGR1_01
V$PAX4_01
V$AHRARNT_02
V$NMYC_01
V$USF_01
V$ARNT_01
V$MYCMAX_03
V$MAX_01
V$ARNT_01
V$MYCMAX_03
V$IRF7_01
Transcription Factor ChIP-seq (161 factors) from ENCODE with Factorbook Motifs
MAX
MYC
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCOD$_E$
Layered H3K27Ac
H3K4Me3 Mark (Often Found Near Promoters) on 7 cell lines from ENCODE
Layered H3K4Me3
C
Scale chr1:
10 kb
hg19
32,480,000
32,485,000
32,490,000
32,495,000
32,500,000
32,505,000
32,510,000
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
Sam68
Sam68
H1-hESC c-Myc IgG-rab ChIP-seq Signal from ENCODE/SYDH
H1ES c-Myc
K562 c-Myc IgG-rab ChIP-seq Signal from ENCODE/SYDH
K562 c-Myc
A549 c-Myc IgG-rab ChIP-seq Signal from ENCODE/SYDH
A549 c-Myc
HeLa-S3 c-Myc Standard ChIP-seq Signal from ENCODE/SYDH
HeLa c-Myc
NB4 c-Myc Standard ChIP-seq Signal from ENCODE/SYDH
NB4 c-Myc
HepG2 cMyc TFBS ChIP-seq Overlap Signal ENCODE/OpenChrom-UTA
HepG2 c-Myc
MCF-7 vehicle cMyc TFBS ChIP-seq Overlap Signal ENCODE/OpenChrom-UTA
MCF-7 c-Myc

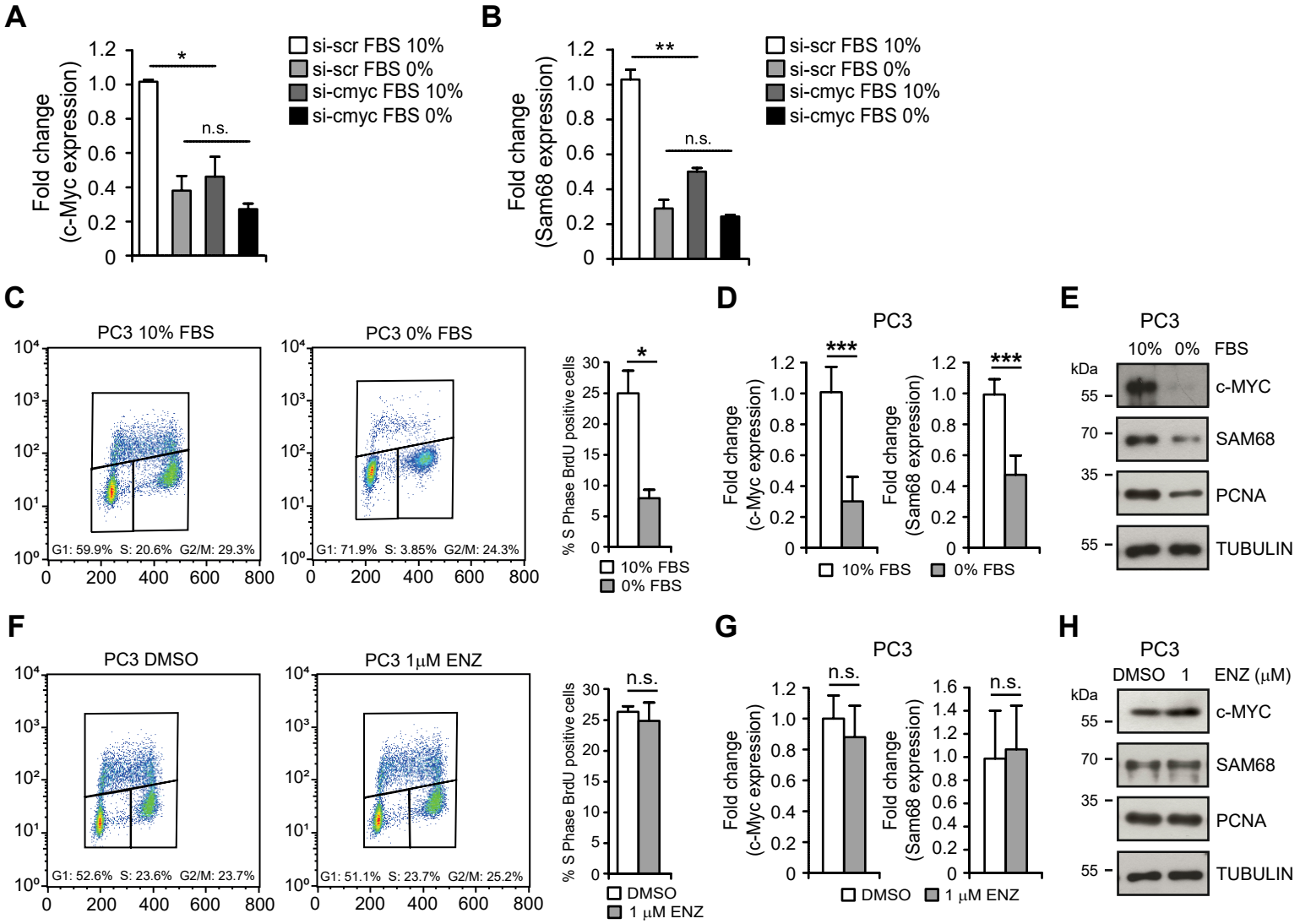
Supp Fig. 1


A
E-Box 1
E-Box 2
GGAGACACGTGTTCCGCTAGCGCCGAGTCACGTGGAAA
ACGAAACGAACGGAGCCCACTGCCTCCCGCGCATGCGCAT
TSS
CAAGCTCTGGCCTG
bp -150 -100 -50 0 +50 +100
TSS
Sam68
c-MYC binding sites
B
human Sam68 promoter -534/+297
human hnRNPA1 promoter -511/+409
E-BOX 1 E-BOX 2
TSS
CACGTG
CACGTG
CATGAG
Luciferase
Fold activation
24h
48h
72h
EV
c-MYC
kDa
c-MYC
β-ACTIN
C
wt
Mut1
Mut2
Mut1-2
D
n.s.
MAX

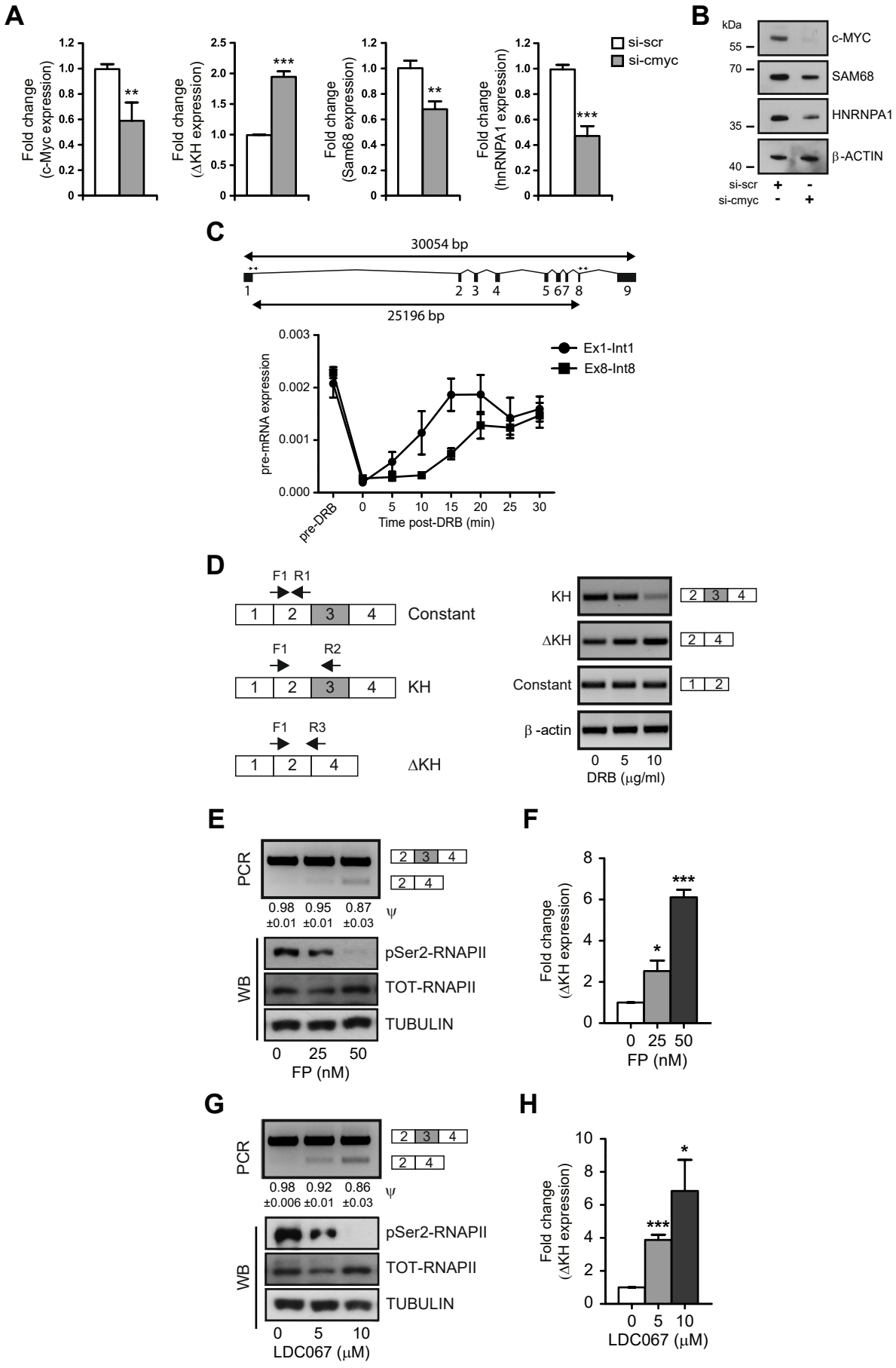
Supp Fig. 2


A
si-scr #1
si-cmyc #1
Fold change (c-Myc expression)
LNCaP
22RV1
PC3
DU145
B
Fold change (hnRNPA1 expression)
C
si-scr #2
si-cmyc #2
D
E
Tumor Prostate - Sawyers
GSE21034
r=0.425
p<0.0001
Sam68 (2log)
c-Myc (log2)
n=370
SAM68 expression (Z-score)
Low
High
c-Myc expression
180
190
F
Tumor Prostate - Sueltman
GSE29079
r=0.626
n=95
51
44

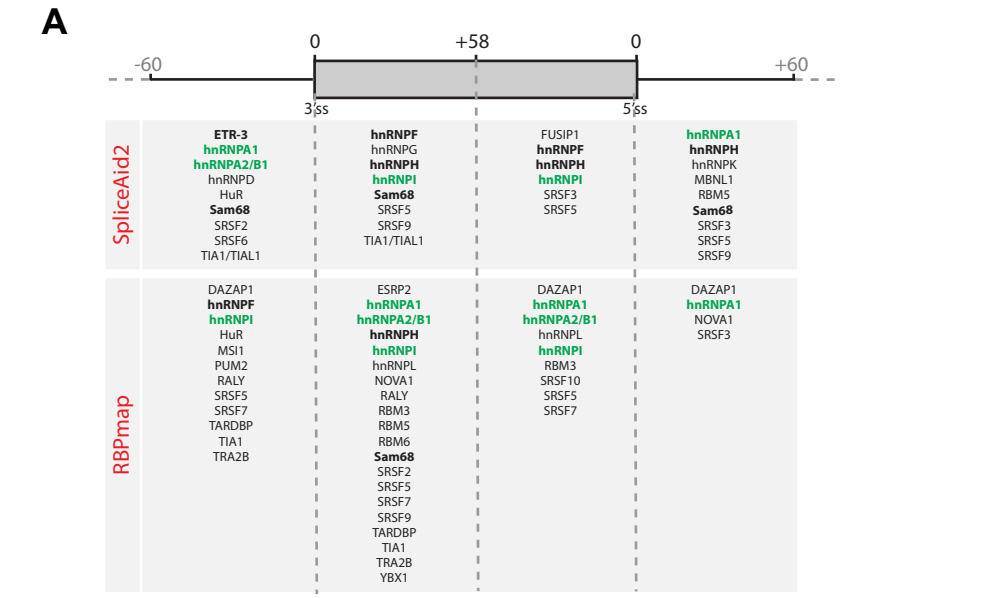
Supp Fig. 3

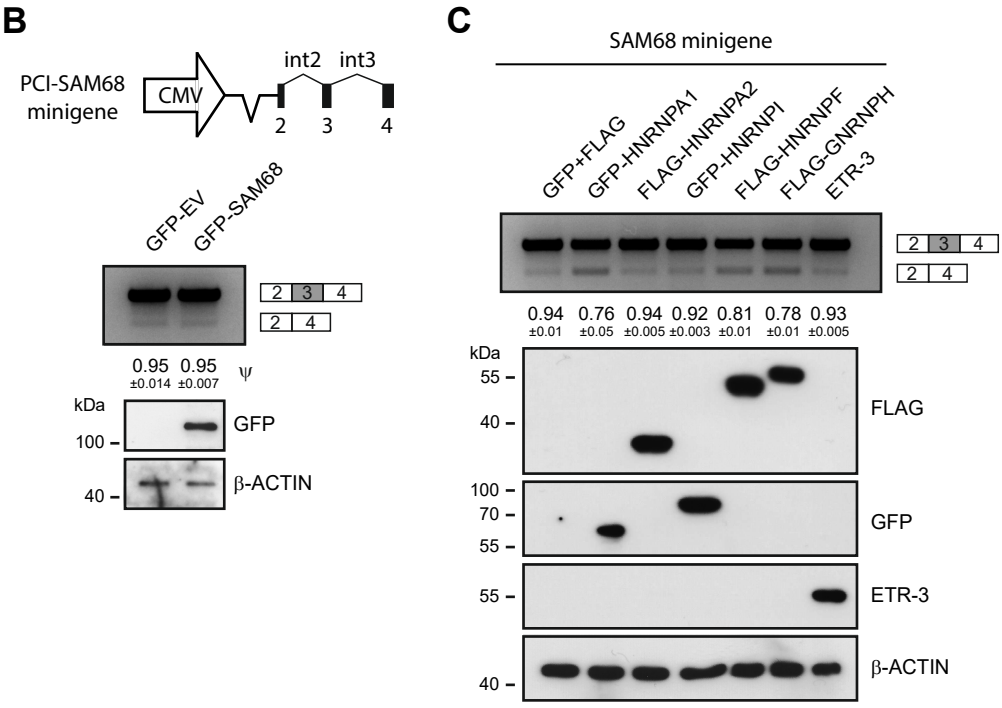

A
Fold change (c-Myc expression)
si-scr FBS 10%
si-scr FBS 0%
si-cmyc FBS 10%
si-cmyc FBS 0%
*
n.s.
B
Fold change (Sam68 expression)
**
n.s.
C
PC3 10% FBS
PC3 0% FBS
G1: 59.9% S: 20.6% G2/M: 29.3%
G1: 71.9% S: 3.85% G2/M: 24.3%
% S Phase BrdU positive cells
10% FBS
0% FBS
D
PC3
***
E
PC3
10% 0% FBS
kDa
c-MYC
SAM68
PCNA
TUBULIN
F
PC3 DMSO
PC3 1μM ENZ
G1: 52.6% S: 23.6% G2/M: 23.7%
G1: 51.1% S: 23.7% G2/M: 25.2%
DMSO
1 μM ENZ
G
H
DMSO 1 ENZ (μM)

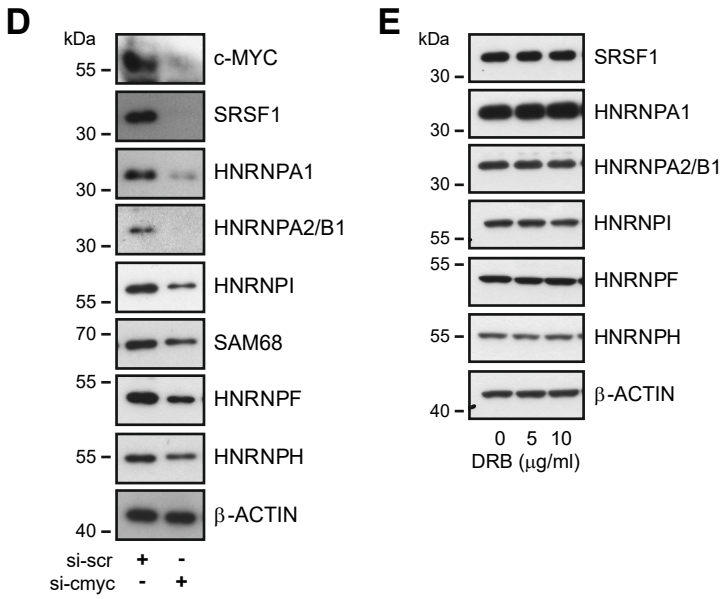
Supp Fig. 4


A
Fold change (c-Myc expression)
Fold change (ΔKH expression)
Fold change (Sam68 expression)
Fold change (hnRNPA1 expression)
si-scr
si-cmyc
B
kDa
55
70
35
40
c-MYC
SAM68
HNRNPA1
β-ACTIN
si-scr + -
si-cmyc - +
C
30054 bp
25196 bp
1 2 3 4 5 67 8 9
pre-mRNA expression
Ex1-Int1
Ex8-Int8
pre-DRB
Time post-DRB (min)
D
F1 R1
Constant
F1 R2
KH
F1 R3
ΔKH
KH
ΔKH
Constant
β -actin
DRB (μg/ml)
E
PCR
0.98 ±0.01
0.95 ±0.01
0.87 ±0.03
ψ
WB
pSer2-RNAPII
TOT-RNAPII
TUBULIN
FP (nM)
F
Fold change (ΔKH expression)
FP (nM)
G
PCR
0.98 ±0.006
0.92 ±0.01
0.86 ±0.03
ψ
WB
pSer2-RNAPII
TOT-RNAPII
TUBULIN
LDC067 (μM)
H
Fold change (ΔKH expression)
LDC067 (μM)

Supp Fig. 5


A
0
+58
0
-60
+60
3'ss
5'ss
SpliceAid2
ETR-3
hnRNPA1
hnRNPA2/B1
hnRNPD
HuR
Sam68
SRSF2
SRSF6
TIA1/TIAL1
hnRNPF
hnRNPG
hnRNPH
hnRNPI
Sam68
SRSF5
SRSF9
TIA1/TIAL1
FUSIP1
hnRNPF
hnRNPH
hnRNPI
SRSF3
SRSF5
hnRNPA1
hnRNPH
hnRNPK
MBNL1
RBM5
Sam68
SRSF3
SRSF5
SRSF9
RBPmap
DAZAP1
hnRNPF
hnRNPI
HuR
MSI1
PUM2
RALY
SRSF5
SRSF7
TARDBP
TIA1
TRA2B
ESRP2
hnRNPA1
hnRNPA2/B1
hnRNPH
hnRNPI
hnRNPL
NOVA1
RALY
RBM3
RBM5
RBM6
Sam68
SRSF2
SRSF5
SRSF7
SRSF9
TARDBP
TIA1
TRA2B
YBX1
DAZAP1
hnRNPA1
hnRNPA2/B1
hnRNPL
hnRNPI
RBM3
SRSF10
SRSF5
SRSF7
DAZAP1
hnRNPA1
NOVA1
SRSF3


B
PCI-SAM68 minigene
CMV
int2
int3
2
3
4
GFP-EV
GFP-SAM68
2 3 4
2 4
0.95 ±0.014
0.95 ±0.007
ψ
kDa
100
GFP
40
β-ACTIN
C
SAM68 minigene
GFP+FLAG
GFP-HNRNPA1
FLAG-HNRNPA2
GFP-HNRNPI
FLAG-HNRNPF
FLAG-GNRNPH
ETR-3
2 3 4
2 4
0.94 ±0.01
0.76 ±0.05
0.94 ±0.005
0.92 ±0.003
0.81 ±0.01
0.78 ±0.01
0.93 ±0.005
kDa
55
40
FLAG
100
70
55
GFP
55
ETR-3
40
β-ACTIN


D
kDa
55
c-MYC
30
SRSF1
30
HNRNPA1
30
HNRNPA2/B1
55
HNRNPI
70
SAM68
55
HNRNPF
55
HNRNPH
40
β-ACTIN
si-scr + -
si-cmyc - +
E
kDa
30
SRSF1
30
HNRNPA1
30
HNRNPA2/B1
55
HNRNPI
55
HNRNPF
55
HNRNPH
40
β-ACTIN
0 5 10
DRB (μg/ml)

Supp Fig. 6

Supplemental figure legends

Figure S1. c-Myc binds Sam68 promoter region. Related to Figure 1. A) UCSC Genome Browser snapshot of publicly accessible RNAPII ChIP-seq profiles of Sam68 gene locus including an intergenic region ~20 Kbp upstream of 5'end from indicated human cell lines. B) UCSC Genome Browser snapshot of publicly accessible putative binding sites for transcription factors on Sam68 promoter. Putative c-MYC and MAX binding sites are indicated in red. ChIP-seq profiles for MAX, c-MYC, H3K27Ac and H3K4Me3 on region between -2000/+2500bp from Sam68 TSS are also shown. C) UCSC Genome Browser snapshot of publicly accessible ChIP-seq profiles for c-MYC carried out on Sam68 locus in indicated cell lines. ChIP-seq control is omitted for visual clarity.

Figure S2. c-MYC induces Sam68 expression. Related to Figure 2. A) Diagram showing c-MYC canonical E-box sequences and their position on Sam68 promoter. TSS: transcription start site (arrow). B) Bar graphs represent luciferase assays performed in HEK293T cells transfected with Sam68 (*left graph*), or hnRNPA1 (*right graph*), promoter reporter vector in combination (c-MYC), or not (EV: empty vector), with c-MYC-pCDNA3 plasmid. Luciferase assays were carried out at the indicated time points. A scheme of Sam68 and hnRNPA1 human promoter (*top schemes*) and representative Western blots showing c-MYC overexpression are also shown (*bottom panels*). C) Representative Western blot analysis showing c-MYC overexpression relative to Figure 2C. D) Bar graph represents luciferase assay performed in HEK293T cells transfected with Sam68 promoter reporter vector in presence of indicated plasmids (EV: empty vector; c-MYC: c-MYC-pCDNA3 vector; MAX: MAX-pCDNA3 vector). Luciferase assay was carried out 24 hours post-transfection. A representative Western blot analysis of c-MYC and MAX overexpression is also shown (*right panel*). B,C,D) β-actin has been used as loading control in western blot analyses. B,D) Bars represent mean ± SD of three biological replicates measured each in triplicate. *P < 0.05, **P < 0.01, ***P < 0.001; n.s., not significant (Student's t-test).

Figure S3. c-MYC regulates Sam68 expression in PCa. Related to Figure 3. A-D) Bar graphs show qPCR analyses of c-Myc (c-Myc; A,C) and hnRNPA1 (hnRNPA1; B,D) expression analysed in PC3 cell line transfected with two different pools of c-Myc (si-cmyc #1 and si-cmyc #2) or control (si-scr #1 and si-scr #2) siRNAs (A-D). Fold change of c-Myc and hnRNPA1 expression relative to Histone 3 expression was calculated by ΔΔCq method. Bars represent mean ± SD of three biological replicates. *P < 0.05, **P < 0.01, ***P< 0.001 (Student's t-test). E,F) Pearson's correlation between Sam68 and MYC expression (E,F *left panels*) and Sam68 expression in c-MYC^{low} (blue circles) and c-MYC^{high} (red squares) patient groups (E,F *right panels*) retrieved from GSE21034-Sawyers (E) and GSE29079-Sueltman (F) datasets. Pearson's correlation coefficient (r) and p-value of the correlation were reported. In right panels statistical significance was calculated by Mann-Whitney test *** p < 0.001.

Figure S4. Down-regulation of c-MYC and Sam68 expression upon cell growth arrest. Related to Figure 4. A,B) Bar graphs show qPCR analyses of c-MYC (A) and Sam68 (B) expression analysed in LNCaP cell line transfected with c-MYC (si-cmyc) or control (si-scr) siRNAs (A,B) and grown

for 3 days in control (FBS 10%) and serum deprivation (FBS 0%) condition. Fold change of c-Myc and Sam68 expression relative to Histone 3 expression was calculated by ΔΔCq method. C-H) Representative dot plot profiles of cytometric analyses showing DNA content versus BrdU incorporation in PC3 cells after 6 days of serum deprivation (C) or 1 μM Enzalutamide (F) conditions. Bar graphs represent the percentage of S-phase BrdU positive cells after 6 days of treatment. qPCR (D,G) and Western blot (E,H) analyses of c-MYC and Sam68 expression are also shown. Fold change of c-Myc and Sam68 expression relative to Histone 3 expression was calculated by ΔΔCq method. A-H) Bars represent mean ± SD of three biological replicates. *P < 0.05, **P < 0.01, ***P < 0.001, n.s., not significant (Student's t-test).

Figure S5. c-MYC regulates alternative splicing of Sam68. Related to Figure 5. A,B) qPCR (A) and representative Western blot (B) analyses showing c-MYC, Sam68-DKH, Sam68 and hnRNPA1 expression related to Figure 5E. In A, fold change of c-MYC, Sam68 and hnRNPA1 expression is relative to Histone 3 expression, while fold change of Sam68-ΔKH expression is relative to Sam68-KH expression. Fold change was calculated by ΔΔCq method. Each value represents the mean ± SD from three independent experiments. P values of Student's t-test are reported, **P < 0.01, ***P<0.001. C) Scheme of Sam68 gene and primers position (*upper scheme*) used for quantitative evaluation (qPCR) of Sam68 exon1-intron1 and exon8-intron8 expression at indicated time point post-DRB release (*lower graph*). D) sqPCR shows *in vivo* splicing assay performed in LNCaP cells treated for 12 hours with suboptimal DRB concentration. A scheme of primer position used to amplify Sam68 isoforms is shown (*right scheme*). E-H) sqPCR (E,G) and qPCR (F,H) analyses showing the Sam68- ΔKH and KH ratio in LNCaP cells treated for 12 hours with suboptimal Flavopiridol (FP) (E,F) and LDC067 (LDC067) (G,H) concentration. Representative western blot (WB) analysis performed to evaluate Serine 2 phosphorylation (pSer2-RNAPII) and total (TOT-RNAPII) RNAPII expression is also shown (E,G). Tubulin has been used as loading control.

Figure S6. hnRNP F modulates alternative splicing of Sam68. Related to Figure 6. A) Schematic representation of *in silico* analyses performed using SpliceAid2 and RBPmap prediction tools to identified putative splicing factor binding sites on Exon 3 of Sam68 RNA. Analysed region includes the last 60bps of intron 2 (black line), sequence of exon 3 (grey box), and the first 60bps of intron 3 (black line). The binding sites position of the most recurrent splicing factors identified and of ETR-3 (black bold), and splicing factors target of c-MYC (green bold) are indicated. B,C) sqPCR analyses of *in vivo* splicing assay of Sam68 minigene performed in HEK293T transfected with GFP-SAM68 (B) or with splicing factors identified in A (C). GFP (GFP) and FLAG (FLAG) empty vectors have been transfected as control plasmids. Representative western blot analysis performed to evaluate the expression of transfected splicing factors is also shown (B,C, *lower panel*). A scheme of Sam68 minigene (B, *upper scheme*) and the Percent of Spliced-In Index (PSI; ψ) are also reported (B,C). D,E) Western blot analyses performed to evaluate the expression of the indicated splicing factors in LNCaP cells depleted (si-cmyc) or not (si-scr) for c-Myc (D) or cultured for 12 hours in presence of suboptimal doses of DRB (E).

Supplementary Table S1.

List of oligonucleotides used in this study.

Oligonucleotides used for sq- and q-PCR analyses of ChIP and CLIP experiments	
Name	Sequence (5' to 3')
human ChIP Sam68 promoter MYCBS F	*GCAGACAAAGACATTCCACAGC*
human ChIP Sam68 promoter MYCBS R	*CCAGAGCTTGATGCGCATG*
human ChIP hnRNPA1 promoter MYCBS F	*GGTTCACTGCCTACTCCTGCC*
human ChIP hnRNPA1 promoter MYCBS R	*AGTCGCTTGGCAAAGCATGA*
human ChIP CyclinB1 promoter MYCBS F	*CTCTCCAGGTGGCCGCTGCAGCTG*
human ChIP CyclinB1 promoter MYCBS R	*CGGCTGCCGGTCGCAGAGAATGC*
human ChIP Nucleolin promoter MYCBS F	*GCGGGAAAGACAGAGTCACTGAG*
human ChIP Nucleolin promoter MYCBS R	*TCCCTCTGGAGATTCCAGGACC*
human ChIP 16q22 F	*CTACTCACTTATCCATCCAGGCTAC*
human ChIP 16q22 R	*ATTTCACACACTCAGACATCACAG*
human CLIP Sam68 ex3 F	*GCAGGAAGAGACTGGTGCAA*
human CLIP Sam68 int3 R	*TTCTGGGCAAGTAGAAGGCA*
Oligonucleotides used for cloning and mutagenesis	
Name	Sequence (5' to 3')
human intergenic F MluI	*TCTAACGCGTTTTCGGCATTTTTCACAGAGATGT*
human intergenic R XhoI	*CAGGCTCGAGCATGCCAGTCTCTTTGTAAGCC*
human Sam68 promoter F MluI	*TCTAACGCGTGTGTTCCCTGATGTCTGGGACC*
human Sam68 promoter R XhoI	*CAGGCTCGAGGGATGTGCGATCCAAGGAGCG*
human hnRNPA1 promoter F MluI	*TCTAACGCGTGCATGCGCAAAGCTAGGACAAAC*
human hnRNPA1 promoter R XhoI	*CAGGCTCGAGACTCGCTCACCAAGGCAAGG*
human Sam68 promoter mut5A F MluI	*TCTAACGCGTCTCAATTTCTCCCTCCAGGTTTCC*
human Sam68 promoter mut5B F MluI	*TCTAACGCGTGACCCTCTCGGCAGTCTTTTACTTC*
human Sam68 promoter mut5C F MluI	*TCTAACGCGTGAGAAACCCAAGGGCGGTGC*
human Sam68 promoter mut3A F XhoI	*CAGGCTCGAGGCACCGCCCTTGGGTTTCTC*
human Sam68 promoter mut3B F XhoI	*CAGG CTCGAG GCCAGAGCTTGATGCGCATG*
human Sam68 promoter mut1 F	*AAGGGGAGGAGAAAAGTATTCCGCTAGCGC*
human Sam68 promoter mut1 R	*GCGCTAGCGGAATACTTTTCTCCTCCCCTT*
human Sam68 promoter mut2 F	*CTAGCGCCGAGTAAAGTTGAAAACGAAACG*
human Sam68 promoter mut2 R	*CGTTTCGTTTTCAACTTTACTCGGCGCTAG*
human MAX F HindIII	*TCTAAAGCTTAGCGATAACGATGACATCGAG*
human MAX R BamHI	*TCTAGGATCCTTAGCTGGCCTCCATCCGG*
human Sam68 minigene ex2 F EcoRI	*CGGAATTCAAATTGAGAAGATTCAGAAAGGAG*
human Sam68 minigene ex4 R SalI	*GCGTCGACCGGTACTAGAAATTTCTTG*
Oligonucleotides used for sqPCR and qPCR	
Name	Sequence (5' to 3')
human H3 F	*GTGTCATCCATGCCAAACGG*
human H3 R	*GTGGCGAGATAGCCCTCCTA*
human Sam68 F1	*TGACGGCAGAAATTGAGAAG*
human Sam68 R1	*GACAGGTATCAGCACTCGCTC*
human Sam68 R2	*CCAAGAATCTTCCCCACAAA*
human Sam68 R3	*GCAGATCCATATTCAAGTGGG*
human Myc F	*AATGAAAAGGCCCCCAAGGTAGTTATCC*
human Myc R	*GTCGTTTCCGCAACAAGTCCTCTTC*
human hnRNPA1 F	*GCCCTGTCAAAGCAAGAGATGGCTAG*
human hnRNPA1 R	*CGACCACTGAAGTTTCCTCCACGACCG*
human Sam68 prox ex1 F	*CAGAGAACAAGTACCTGCCCG*
human Sam68 prox int1 R	*CACGGGAGGCCCCCTTAC*
human Sam68 dist ex8-int8 F	*ATGGACATGGGGAGGTTCAAG*

human Sam68 dist int8 R	*CAGCATTCTCCCAGGCACTT*
pCI F	*GGTGTCCACTCCCAGTTCAA*