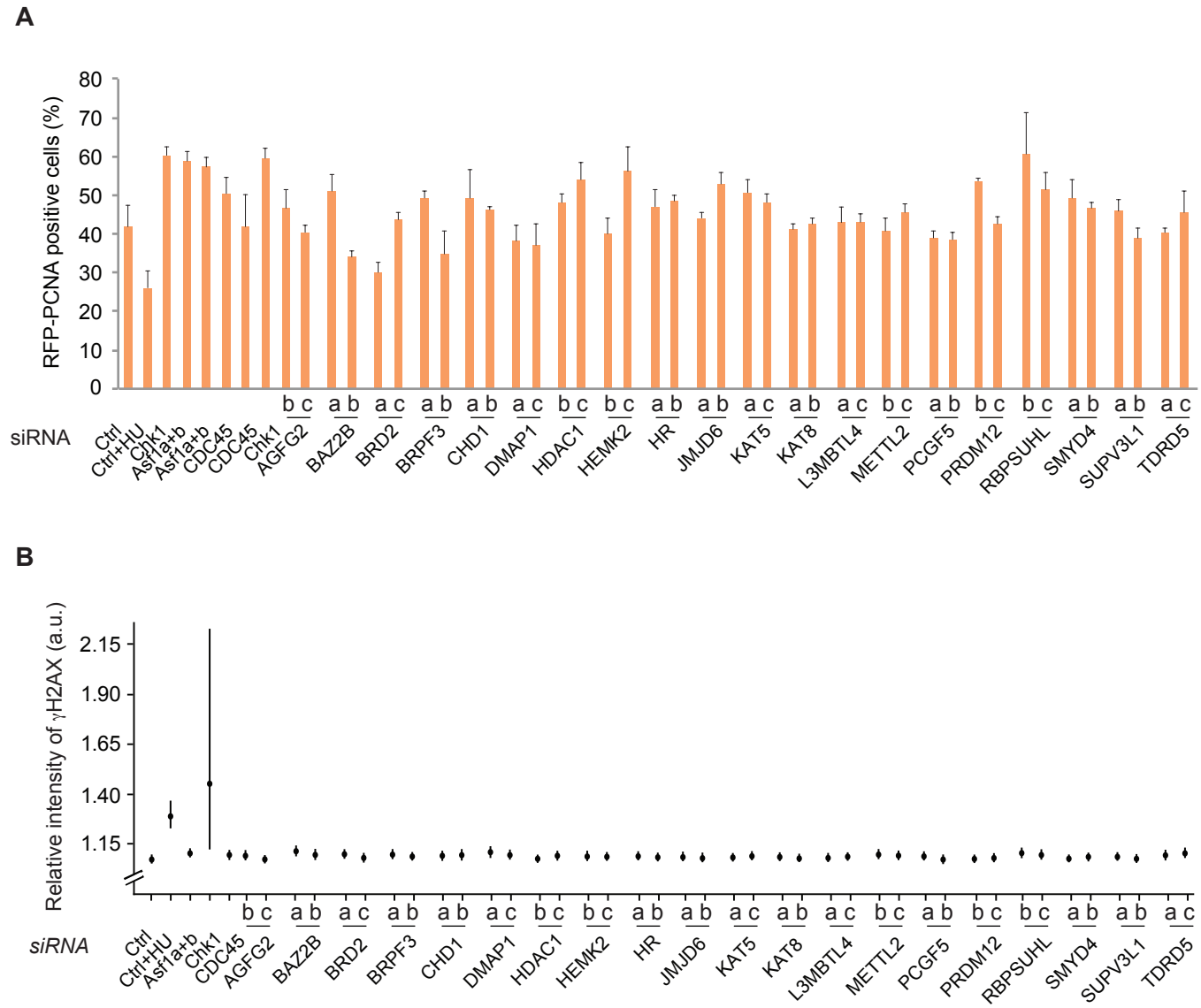
BRPF3-HBO1 regulates replication origin activation and histone H3K14 acetylation

Yunpeng Feng [1,2,10], Arsenios Vlassis [1,10], Céline Roques [3,10], Marie-Eve Lalonde [3], Cristina González-Aguilera [1], Jean-Philippe Lambert [4], Sung-Bau Lee [1,5], Xiaobei Zhao [6,7], Constance Alabert [1], Jens V. Johansen [1], Eric Paquet [3], Xiang-Jiao Yang [8], Anne-Claude Gingras [4,9], Jacques Côté [3*] and Anja Groth [1*]

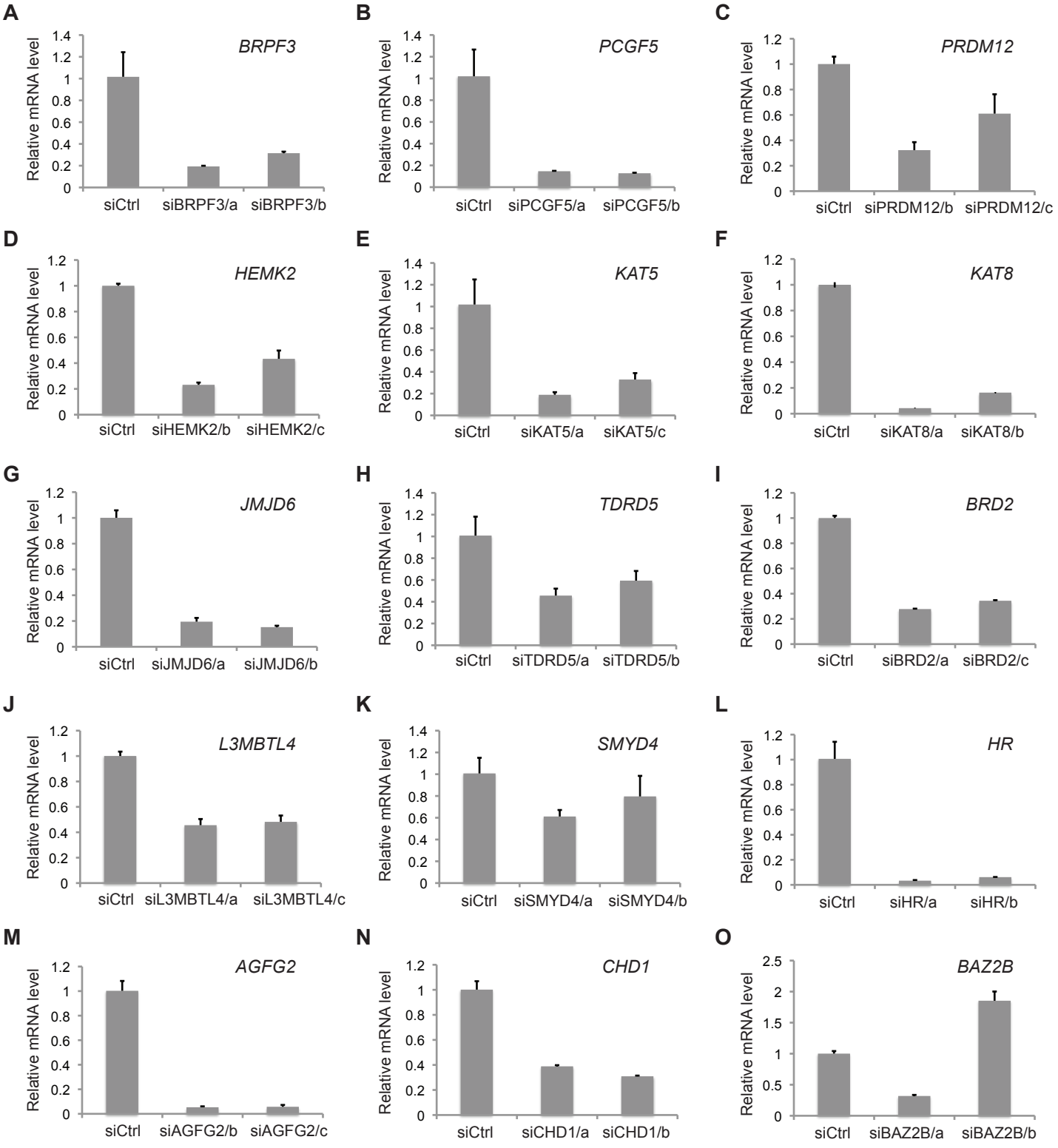
Appendix

Appendix Figure S1



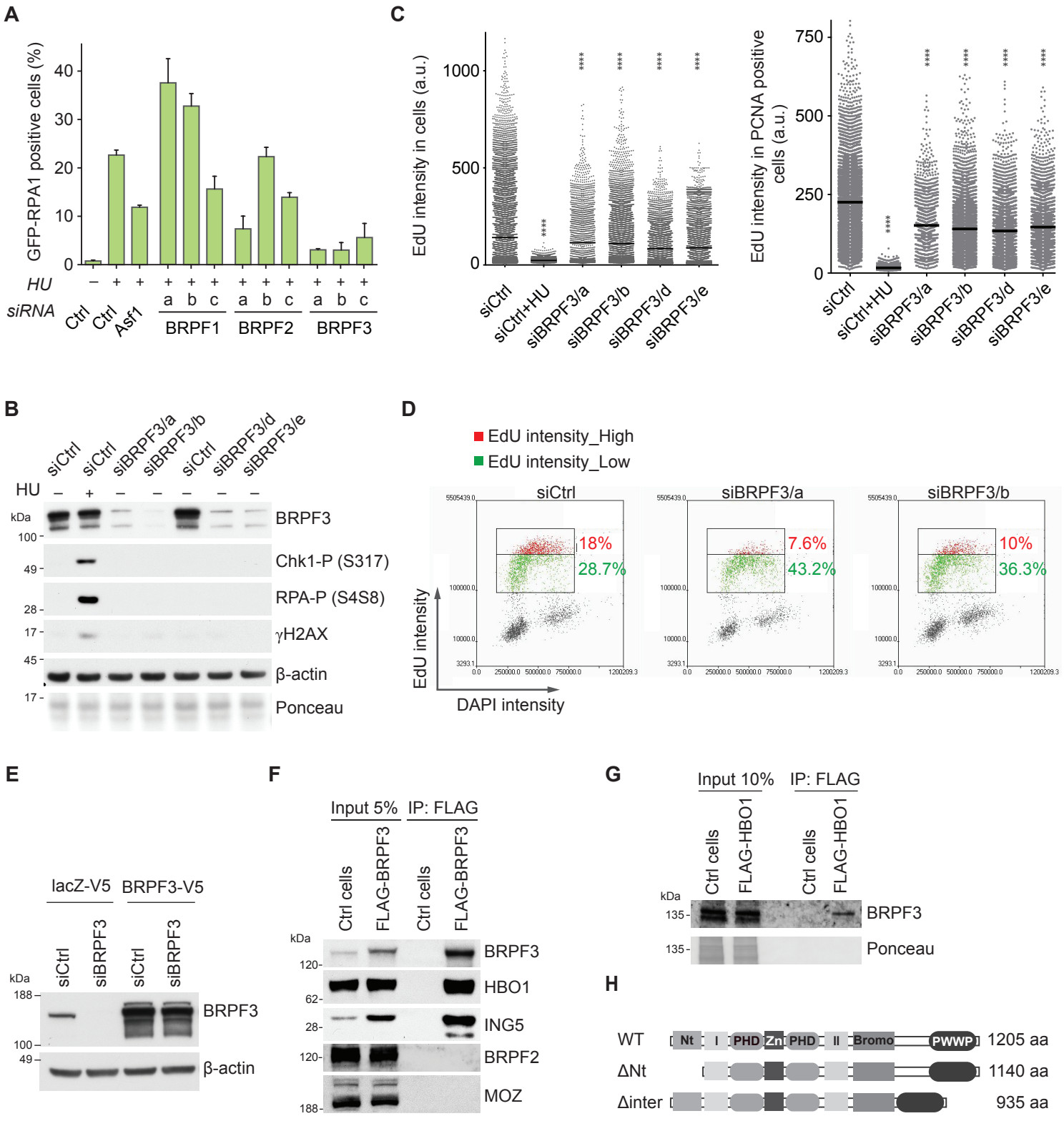
Appendix Figure S1. **Functional analyses of top genes identified in siRNA screen.** A, High-content imaging analysis of chromatin-bound PCNA. Reporter RFP-PCNA U-2-OS cells were treated with siRNAs targeting 20 candidate genes, pre-extracted and imaged. Error bars, SD; n = 6 technical replicas. B, High-content imaging analysis of γH2AX. Reporter RFP-PCNA U-2-OS cells were treated with siRNAs targeting 20 candidate genes, stained for γH2AX and imaged. S-phase cells positive for RFP-RCNA were included in the analysis. Median with interquartile range of relative intensity per cell is shown, n > 4000. One representative experiment out of two biological replicas is shown.

Appendix Figure S2



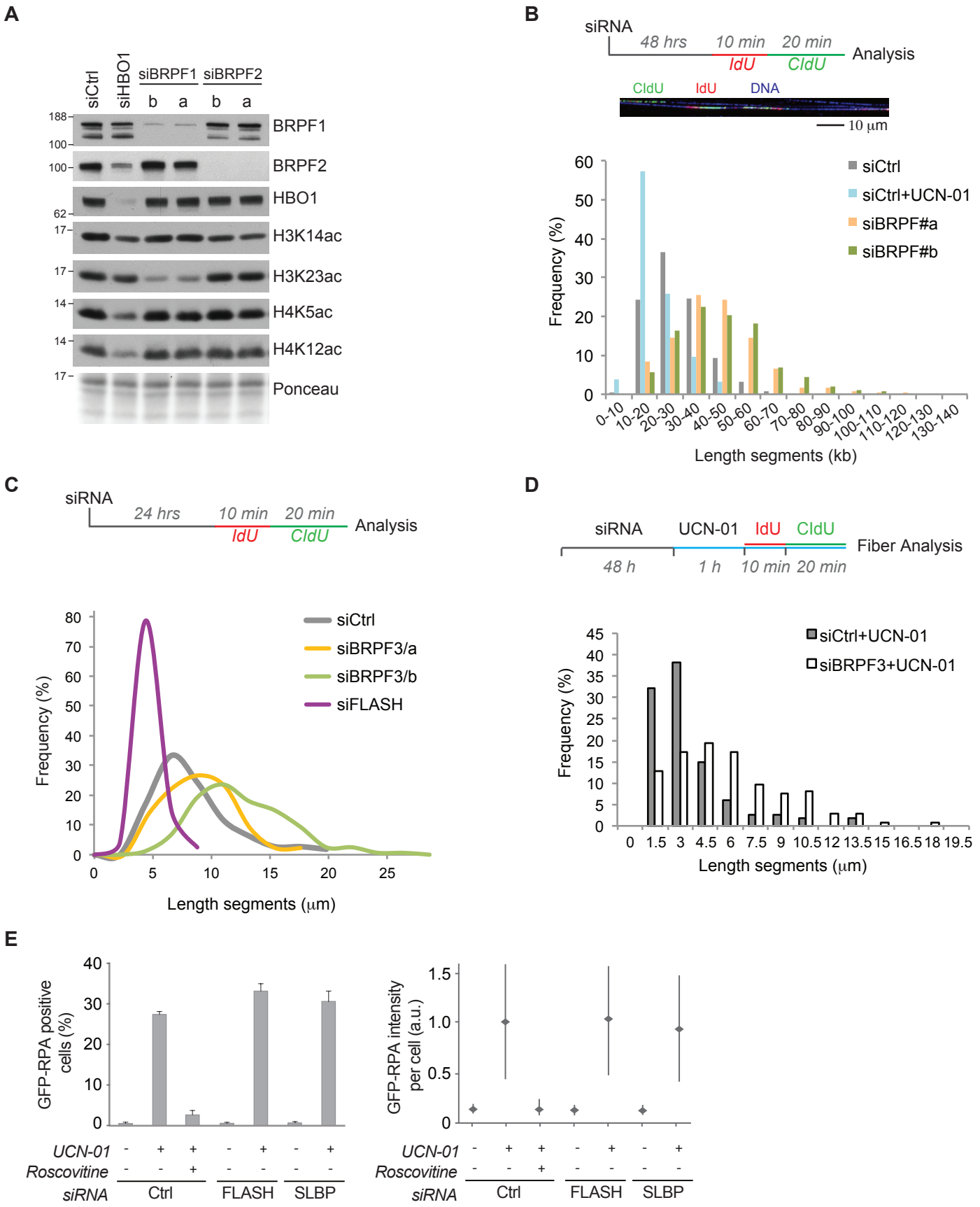
Appendix Figure S2. **Validation of siRNA targeting efficiency for selected hits from the siRNA screen.** A-O, Quantification of mRNA levels by RT-qPCR 48 hrs after siRNA transfection. mRNA levels were normalized to GAPDH. Error bars, SD; n = 3 technical replicas. One representative experiment out of two biological replicas is shown.

Appendix Figure S3



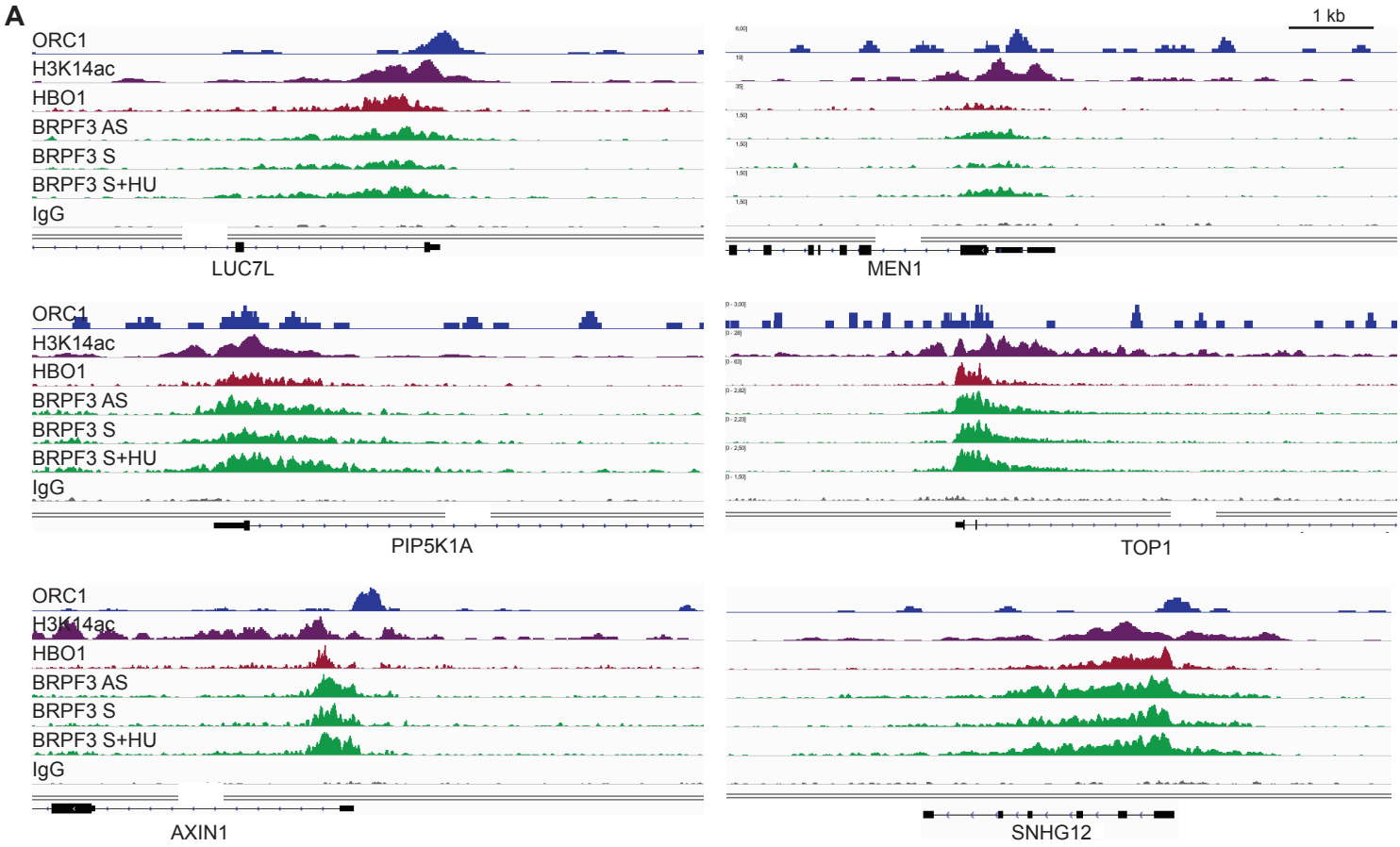
Appendix Figure S3. **BRPF3 is required for DNA replication and interacts with HBO1.** A, Analysis of chromatin-bound GFP-RPA in cells depleted for BRPF1, 2 or 3. Data are from the siRNA screen as described in Fig. 1A, each of genes was targeted by three independent siRNAs (a, b and c). siRNA against ASF1 was used as positive control. The averages and standard error of GFP-RPA positive cells from two biological replica-screens are shown. B, Western blot analysis of total extracts from siRNA treated U-2-OS cells. Cells treated 2 h with HU were included as a positive control for checkpoint signaling. C-D, DNA synthesis rates measured by high content image analysis of EdU incorporation. (C) Dot plots illustrating the distribution of EdU intensities in total cells (*left*) or PCNA-positive cells (*right*). siRNA treated U-2-OS cells pulsed 15 min with EdU. n > 1000. Mann-Whitney: **** $P < 10^{-4}$. (D) Dot plot of EdU and DAPI intensities. Cells were treated as in C. n > 1000. Replicating cells were gated according to their EdU intensities and shown as percentage of the total population. Red (High EdU intensity); Green (Low EdU intensity). E, BRPF3 levels in stable cell lines used for the complementary analysis described in Fig. 2B. F, Immunoprecipatation of FLAG- BRPF3 from stable U-2-OS cell lines expressing lacZ-V5 (ctrl) and FLAG-BRPF3. One representative experiment out of two is shown. G, Immunoprecipatation of FLAG-HBO1 from transiently transfected U-2-OS cells. H, Illustration of BRPF3 protein domains and deletion constructs used in Fig. 2D.

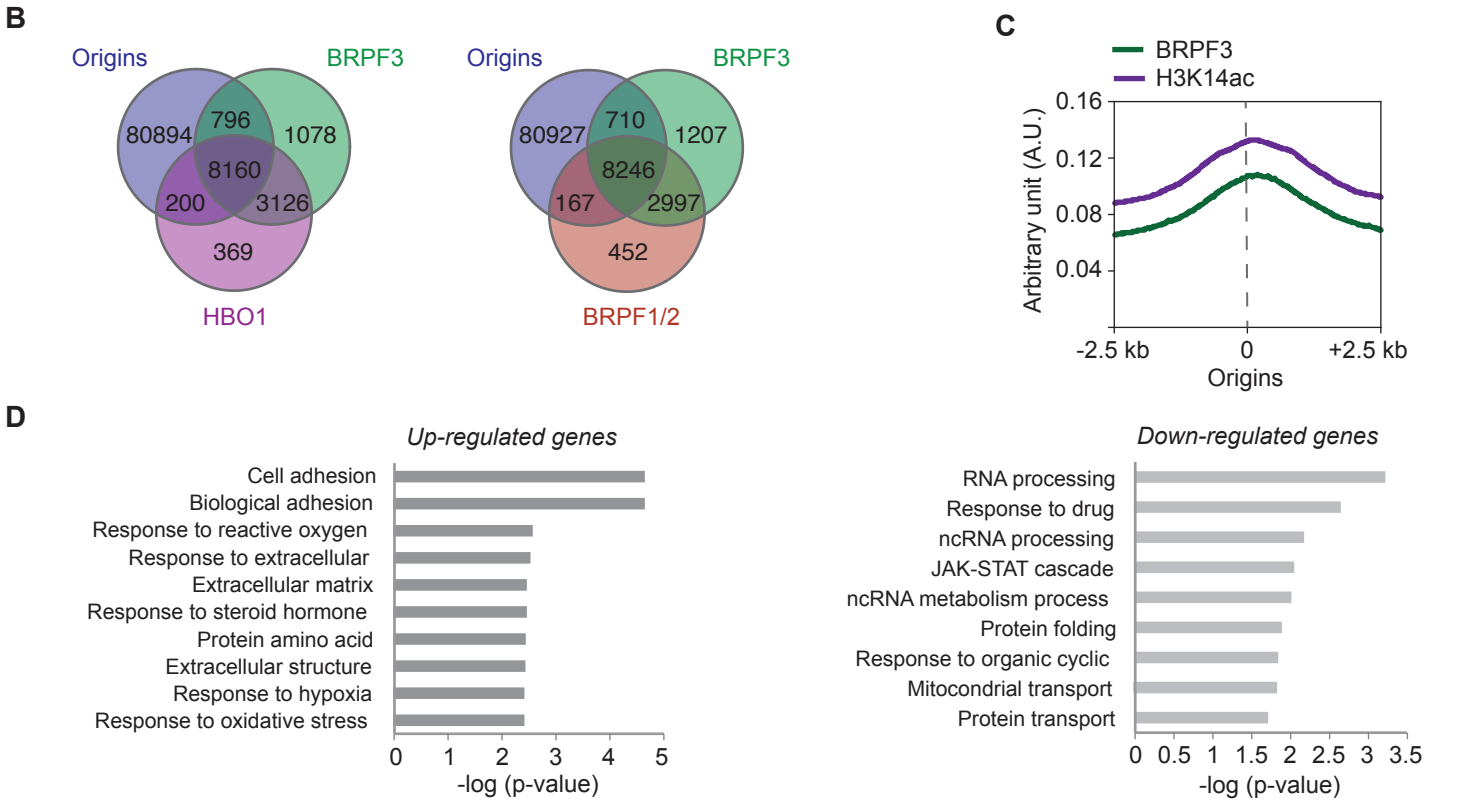
Appendix Figure S4



Appendix Figure S4. **BRPF3 regulates replication fork progression and origin activation.** A, Western analysis of total extracts from siRNA treated U-2-OS cells. One representative experiment out of two biological replicas is shown. B, Molecular DNA combing. (*top*) Experimental design and representative images of combed DNA. IdU, green; CIdU, red; DNA, blue. (*bottom*) Size distribution of CIdU track length. n > 300. One representative experiment out of two biological replicas is shown. C-D, Fiber assay analysis of DNA replication in BRPF3 depleted cells. Fork elongation rate and density were analysed in the absence or presence of UCN-01. (C) (*top*) Experimental setup. (*bottom*) Frequency distribution of CIdU track length. n > 110. siRNA against FLASH shown to decrease fork speed (Mejlvang et al., 2014) was included as a control. (D) (*top*) Experimental setup. (*bottom*) Frequency distribution of inter-CIdU-track distances. n > 110. E, Analysis of chromatin-bound GFP-RPA. siRNA treated cells were incubated 2 h with UCN-01 (300 nM) as indicated, pre-extracted and imaged. Roscovitine (25 nM) was added immediately before UCN-01 as a positive control. siRNA against FLASH and SLBP shown to decrease fork speed without reducing origin density (Mejlvang et al., 2014) was included as a negative control. (*left*) Percentage of GFP-RPA positive cells scored after pre-extraction. Error bars, SD, n = 12. (*right*) GFP-RPA intensity per cell. Median with interquartile range is shown, n > 6000. a.u., arbitrary unit.

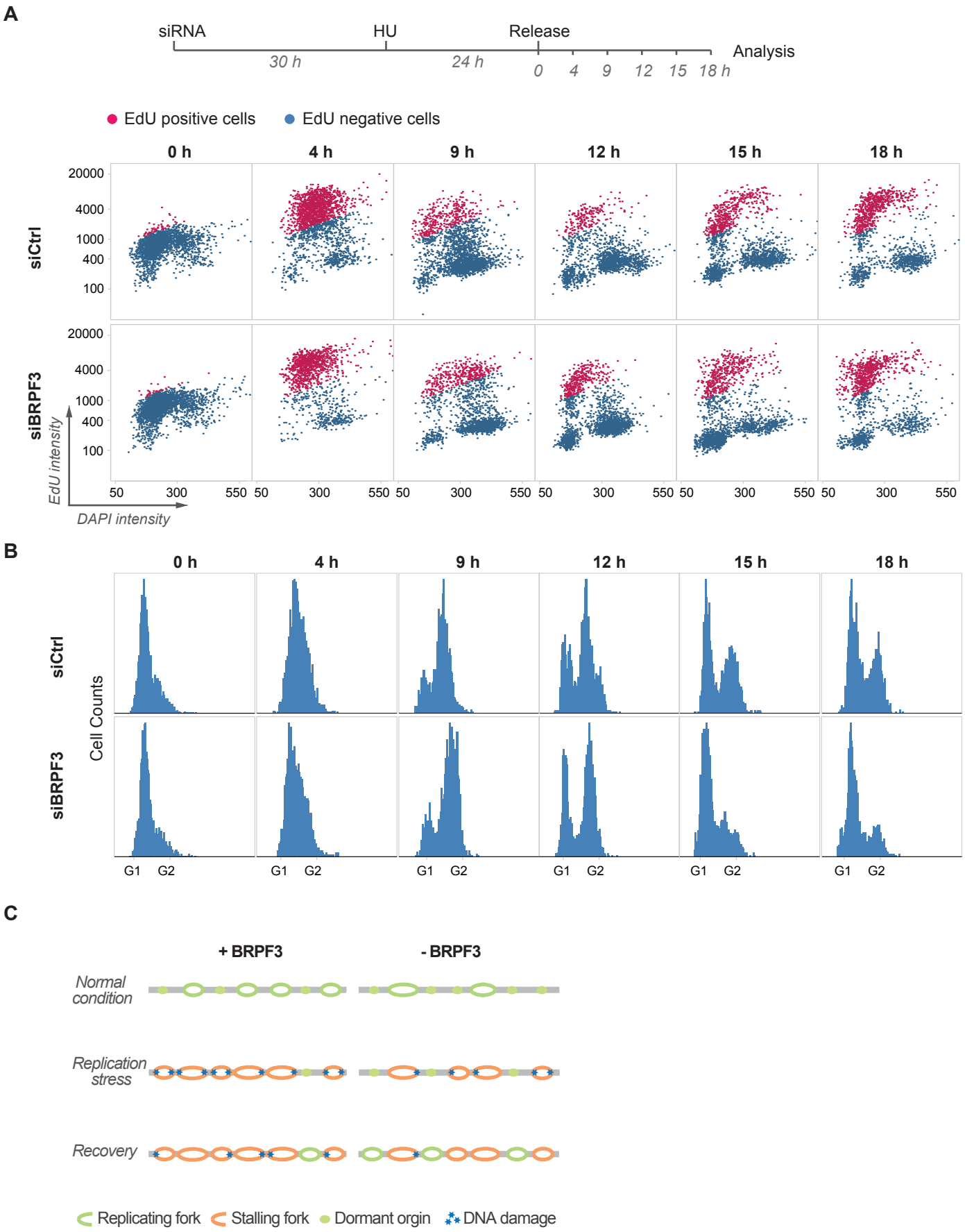
Appendix Figure S5





Appendix Figure S5. **Enrichment of BRPF3, HBO1 and H3K14ac around replication origins.** BRPF3 enrichment signal obtained by ChIP-seq with asynchronous (AS), synchronized (S) and released into HU (S+HU) RKO cells, on PIP5K1A, MEN1, TOP1, AXIN1, LUC7L and SNHG12 replication origins. Signals were compared to the published enrichment signals for ORC1 obtained in HeLa cells (Dellino et al., 2013, GSE37583), HBO1 signal in RKO cells (Avvakumov et al., 2012, GSE33221), and to the H3K14ac enrichment signal obtained in IMR90 (GSM521881). The negative control IgG signal shown is from the S+HU chromatin. B, Venn diagrams illustrating the overlap of genome binding sites between (*left*) BRPF3, HBO1 and replication origins, and (*right*) BRPF3, BRPF1/2 and replication origins in human cells. Data from replication origins defined previously by short nascent strand (SNS) technique were obtained from http://pbil.univ-lyon1.fr/members/fpicard/research.html (Hela table) (Picard et al., 2014). C, Average profiles of BRPF3 and H3K14ac ChIP-seq signal on ± 2.5 kb around the center of replication origins defined in B. D, Gene Ontology (GO) analyses of up and down-regulated genes with altered expression levels after 48 h of BRPF3 siRNA treatment. GO analyses has been done using DAVID (Huang et al., 2009). GO terms with p-value < 0.05 and gene count ≥ 5 were considered significant. 10 first significant categories sorted by p-value are shown.

Appendix Figure S6



Appendix Figure S6. **Recovery of BRPF3 depleted cells from replication stress.** A-B, Time course analysis of EdU incorporation and cell cycle progression of BRPF3-depleted cells released from HU. Following siRNA transfection, U-2-OS cells were treated with HU (1 mM) for 24 h and then released into S phase. Cells were pulsed with EdU for 15 min prior to sample collection and analyzed by high-content imaging for EdU and DAPI. (A) Experimental design (*top*). Dot plots of EdU and DAPI intensities (*bottom*). Cells recovering DNA synthesis (increased EdU incorporation compared to HU treated cells) are gated and shown in purple. (B) Cell cycle profiles based on DAPI staining of DNA. C, Model depicting the role of BRPF3 in DNA replication. In the absence of BRPF3 origin firing is impaired and, to compensate, forks move faster. In response to replication stress, dormant origins fail to fire in the absence of BRPF3 and cells experience less DNA damage. Upon recovery, cells lacking BRPF3 are now more apt to complete S phase because they suffered a reduced load of DNA damage.

Appendix Table S1

Appendix Table S1. **Sequences of primers**

Sequences of primers used in RT-qPCR

Genes	Forward primer	Reverse primer
BRPF3	5' GGCCTCCTGCACAATGGCGT	5' TGCGGATGCTGGTCTTGCGG
HEMK2	5' ACTGGCAGGAGTGGAAATATGCCT	5' GGTACAAGCTGCTGCCTCAGGGT
PRDM12	5' AAGCATGAGGACTTCCACCCGGC	5' GACTGGCTGAAGCGGCGGTT
SMYD4	5' TGTGCTGCCTTAGGAGACTGGCAA	5' TGCTCAGGGCTTCGGGTACTGC
JMJD6	5' GTACCAGGAGGCTGGTGGCATGTTG	5' CAACCGAGTCTGCGAGGACTGC
KAT5	5' TTCAGCGTCATTTGACCAAGTGTGA	5' GCCAAAAGACACAGGTTCTGGGAAT
PCGF5	5' TGGGCAATCAGGGGACAATGTAGT	5' ACCATTGCACAGCACATCCAACTCA
TDRD5	5' TCCTCCCTGAGGTTCCTCAAGTGC	5' CCCCACTAATGGCTTCAAACCGCA
AGFG2	5' CCCTTTCACAGCTCCCGCCG	5' GGAAGCCAGGCCCAGCACTG
L3MBTL4	5' GACGGACGCTTGGAGCAAGCTG	5' CAGCTCAACAGGTGCTGCGACA
BAZ2B	5' TCCCAATGGTCAGCCACCCAGT	5' CGCCACCATTCTGCACCTGGAA
CHD1	5' GCAAAACAGTTGCAGACCCGTGC	5' TCCTCTTTGAACTTCCCGCACCA
HR	5' CGGGGAGGGGCTCTGGTCTC	5' CACCATCTGGAGAAAGCGGCG
KAT8	5' GCCGGCGACCGGATAGCAC	5' AGCCGCCGGTTAAAGCCCAC

Sequences of primers used in ChIP-qPCR

Genes	Forward primer	Reverse primer
LAMINB2	5' CTCCACCCCCAAGGAAAAAG	5' GGCAGGGTCCCATGCA
SNHG12	5' GCATAGCTGCTGTGGTCAAA	5' CGTTCTGCTGTTCTGTGGAA
LUC7L	5' CAGCGCTTGACAGTCGTTAG	5' ATCTGCACAGCCCTGAGAAT
MCM4	5' TCTGCACTCCGTTCAGCTCCTCTG	5' GAGTGAGGATGCCAGGTCATCTCC

Appendix Table S2

Appendix Table S2. **Plasmid constructs and cloning**

NO.	Construct	Cloning
1	pRev-CMV :3xFlag-BRPF3	Full cDNA of human BRPF3 (NM_015695.2) with N-terminal 3xFlag epitope
2	pRev-CMV: 3xFlag-BRPF3 siBRPF3-a^R	Introduction of silent mutations within the complementary to the siRNAs seed regions in plasmid #1 by site-directed mutagenesis
3	pENTR: CMV: 1xFlag-BRPF3 siBRPF3-a^R	TOPO-reaction between pENTR-D-TOPO vector (Invitrogen) and PCR amplicon Flag-BRPF3 from plasmid #2
4	pENTR: BRPF3 no stop-codon siBRPF3-a^R	TOPO-reaction between pENTR-D-TOPO vector (Invitrogen) and PCR amplicon BRPF3 (no tag) from plasmid #2
5	pLenti6-UbC:1xFlag-BRPF3 (N-terminal Flag tag) siBRPF3-a^R	LR recombination reaction between plasmid #3 and pLenti6/UbC-DEST (Invitrogen)
6	pLenti6-UbC:BRPF3-V5 (C-terminal V5 tag) siBRPF3-a^R	LR recombination reaction between plasmid #4 and pLenti6/UbC-DEST (Invitrogen)
7	pRev-CMV-3xFlag-ΔNt(1-65)BRPF3	
8	V6898:3xFlag-BRPF3Δintern	
9	pAX8	Lentivirus packaging
10	pCMV-VSVG	Lentivirus envelope
11	pLenti6-UbC:*lacZ*-V5	Provided together with the LR recombinase kit (Invitrogen)

Appendix Table S3

Appendix Table S3. **Sequences of siRNAs**

Genes	Sequence (sense strand 5'- 3')	Source and catalog number
siCtrl	GCGCGAUAGCGCGAAUAUA-dTdT	Universal negative control #1 SIC001 (Sigma)
siCtrl	Not available	Ambion Silencer Select negative control #2, 4390846
siBRPF3/a	GCGGGUCUGACUCUGAAUG-dTdT	Sigma (Ambion Silencer sequence)
siBRPF3/b	CCGGCUUGAGAAAGAGUCA-dTdT	Sigma (Ambion Silencer sequence)
siBRPF3/d	CCAACUGCAUGAAGUAUAA-dTdT	Ambion Silencer Select, s25918
siBRPF3/e	GCAUCGUAUCAGCAUCUAU-dTdT	Ambion Silencer Select, s25920
siBRPF2/a	GAAACUAUAGACAAGUUAA-dTdT	Ambion Silencer Select, s24388
siBRPF2/b	GGUUAGAAGCUCAAGGGUA-dTdT	Ambion Silencer Select, s24389
siBRPF1/a	GGAGAUCUUUGAGUACCUA-dTdT	Ambion Silencer Select, s15422
siBRPF1/b	GUUUCUUGGUAUACCGUAA-dTdT	Ambion Silencer Select, s15423
Control pool	Not available	Dharmacon, ON-TARGETplus Non-targeting Pool D-001810-10-20
BRPF3 pool	Not available	Dharmacon, ON-TARGETplus SMARTpool L-025088-01-0005
siHBO1	CAGCUUCGAUAUAAGGAAA-dTdT	Ambion Silencer Select, s255
siING5	GAAAAGAGGAAGAAGAAGU-dTdT	Sigma (seed region from Doyon et al. 2006)
siCDT1	GCGCAAUGUUGGCCAGAUC-UU	Sigma (sequence from {Lovejoy:2006fr}
siCDC45	CGGAUCUCCUUUGAGUAUG-dTdT	Sigma (Ambion Silencer sequence)

Appendix Table S4

Appendix Table S4. **Antibody list**

Antibody target	Company	Product number	Species
BRPF2 (BRD1)	Bethyl	A302-366A	rabbit
BRPF1	Atlas/Sigma	HPA003359	rabbit
BRPF3	Bethyl	A304-082A	rabbit
BrdU	AbD Serotec	OBT0030G	rat
beta-Actin	Sigma	A5316	mouse
CHK1-ph (Ser317)	Cell Signaling	2344s	rabbit
CHK2-ph (Thr68)	Cell Signaling	2661s	rabbit
CDC45	Santa Cruz	sc-55569	mouse
CDC45	Santa Cruz	sc-20685	rabbit
CDT1	Bethyl	A300-786A	rabbit
CIdU	AbCys SA	ABC117-7513	rat
Flag epitope	Sigma	F3165	mouse
HBO1	Bethyl	A302-224A	rabbit
H3	Abcam	ab1791	rabbit
H3	Abcam	ab10799	mouse
H4	Millipore	05-858	rabbit
H3K14ac	Cell Signaling	4318s	rabbit
H4K12ac	Abcam	ab1761	rabbit
H4K5ac	Abcam	ab51997	rabbit
H3K23ac	Active Motif	39131	rabbit
IdU	Becton Dickinson	347580	mouse
ING4	Protein Tech	16188-1-AP	rabbit
ING5	Atlas/Sigma	HPA042685	rabbit
MCM2	Bethyl	A300-122A	mouse
MCM4	BD Biosciences	559544	rabbit
MOF	Bethyl	A300-992A	rabbit
MOZ	Active Motif	39867, 39868	rabbit
ORC1	Bethyl	A301-892A	rabbit
PCNA	Santa Cruz	sc-7907	rabbit
RPA1	Abcam	ab79398	rabbit
RPA2	Thermo Scien.	MS-691	mouse
RPA3	Abcam	ab588	mouse
γH2AX	Millipore	05-636 cl. JBW301	mouse
ssDNA	Chemicon	MAB3868	mouse
V5 epitope	Invitrogen	R960-25	mouse

APPENDIX SUPPLEMENTARY METHODS

ChIP-seq analysis. Library preparation for sequencing were done as previously described (Avvakumov et al, 2012) Samples were sequenced by 50 bp single-reads on either a Genome Analyzer platform (HBO1, BRPF1/2 and input) or a HiSeq 2000 platform (BRPF3, IgG) (Illumina). Raw sequences were mapped using Bowtie version 0.12.8 on build hg18 or hg19 of the human genome and deposited in the GEO database under accession number GSE63722. HBO1, BPRF1/2 and input data were previously deposited under accession numbers GSE47190 and GSE33221 (Avvakumov et al, 2012). Read alignments in sam format were converted to bam files using samtools (Langmead et al, 2009). Samtools was then successively applied to remove duplicated reads, sort, and index bam files. Peak calling was performed using MACS version 1.4.0 by comparing AS-BRPF3, HU-BRPF3 and M-BRPF3 to M-HU IgG and HBO1 and BRPF1/2 to the signal from the input (Zhang et al, 2008). We identified intersection between two sets of significant peaks in hg19 coordinates using bedtools "intersect" function (Quinlan & Hall, 2010). In case where we needed to compare hg18 and hg19 peaks, we used the liftover tool from the UCSC genome browser to convert hg18 into hg19 coordinates (Kent et al, 2002). The ORC1 ChIP-seq data used for comparison was obtained from Dellino et al. (GSE37583) (Dellino et al, 2013) and the SNS origins used were obtained from Picard et al. (http://pbil.univ-lyon1.fr/members/fpicard/research.html) (Picard et al, 2014). We used the Bioconductor package VennDiagram (Chen & Boutros, 2011) to display the overlaps between the different experiments and IGV for ChIP-seq signal visualization in genomic context (Thorvaldsdottir et al, 2013). The ChIP-seq data presented in the form of heatmaps were generated using deepTools (Ramirez et al, 2014).

Mass spectrometry analysis of purified native BRPF1/3 complexes. Purified proteins to be analyzed were resolved on 12% SDS-PAGE gel, digested in-gel with the protease trypsin as previously described (Lambert et al, 2010) and stored at -80°C until analysis. Five microliters of the digested peptides (corresponding to half of the total sample) were used per analysis. Each sample was directly loaded at 400 nL/min onto a 75 μm x 12 cm emitter packed with 3 μm ReproSil-Pur C_{18}-AQ (Dr. Maisch HPLC GmbH, Germany). The peptides were eluted from the column over a 90 min gradient generated by a NanoLC-Ultra 1D plus (Eksigent, Dublin CA) nano-pump

and analyzed on a TripleTOF 5600 instrument (AB SCIEX, Concord, Ontario, Canada). The gradient was delivered at 200 nL/min, starting from 2% acetonitrile with 0.1% formic acid to 35% acetonitrile with 0.1% formic acid over 90 minutes, followed by a 15 min clean-up at 80% acetonitrile with 0.1% formic acid, and a 15 min equilibration period back to 2% acetonitrile with 0.1% formic acid for a total of 120 min. To minimize carryover between each sample, the analytical column was washed for 3 h by running an alternating sawtooth gradient from 35% acetonitrile with 0.1% formic acid to 80% acetonitrile with 0.1% formic acid, holding each gradient concentration for 5 min. Analytical column and instrument performance were verified after each sample by loading 30 fmol BSA tryptic peptide standard (Michrom Bioresources Inc. Fremont, CA) with 60 fmol α-Casein tryptic digest and running a short 30 min gradient. TOF MS calibration was performed on BSA reference ions before running the next sample in order to adjust for mass drift and verify peak intensity. The instrument method was set to a discovery mode which consisted of one 250 ms MS1 TOF survey scan from 400–1300 Da followed by twenty 100 ms MS2 candidate ion scans from 100–2000 Da in high sensitivity mode. Only ions with a charge of 2+ to 4+ which exceeded a threshold of 200 cps were selected for MS2, and former precursors were excluded for 10 seconds after the first occurrence. The mass spectrometry data were stored, searched and analysed using the ProHits laboratory information management system (LIMS) platform (Liu et al, 2010). Within ProHits, the resulting WIFF files were first converted to an MGF format using AB SCIEX MS Data Converter (V1.3 beta) and ProteoWizard (v3.0.4468) and then searched using Mascot (v2.3.02). The spectra were searched with the RefSeq database (version 57, April 18th, 2014) acquired from NCBI against a total of 36113 human and adenovirus sequences supplemented with "common contaminants" from the Max Planck Institute (http://maxquant.org/downloads.htm) and the Global Proteome Machine (GPM; http://www.thegpm.org/crap/index.html). The database parameters were set to search for tryptic cleavages, allowing up to 4 missed cleavage sites per peptide with a mass tolerance of 50 ppm for precursors with charges of 2+ to 4+ and a tolerance of +/- 0.15 amu for fragment ions. Carbamidomethylation of cysteine was selected as a fixed modification while variable modifications searched were acetylation (lysine and protein N-terminus), asparagine deamidation and methionine oxidation. An ions score cut-off of 27 was chosen to produce a false-positive rate of less than 1% in the MS

data. Furthermore, to be considered significant a protein required at least two bold red peptides and a Mascot score greater or equal to 70.

All MS files used in this study and the Mascot search results were deposited at MassIVE (http://massive.ucsd.edu). The MassIVE ID is MSV000078943, the password until publication is FengBRPF3 and the MassIVE link for download is http://massive.ucsd.edu/ProteoSAFe/status.jsp?task=69a9252ac0634e7fa66a3b98b2844948

SUPPLEMENTARY REFERENCES

Chen H, Boutros PC (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12:** 35

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. *Genome Res* **12:** 996-1006

Lambert JP, Fillingham J, Siahbazi M, Greenblatt J, Baetz K, Figeys D (2010) Defining the budding yeast chromatin-associated interactome. *Mol Syst Biol* **6:** 448

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841-842

Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42:** W187-191

Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14:** 178-192

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137