Supplemental Experimental Procedures

RBL1 promoter construct stable transfection

Pools of T98G cells containing an integrated wildtype or E2F binding site mutant RBL1 promoter construct were created using the site-specific recombinase phiC31 integrase and a modified luciferase reporter plasmid essentially as previously described (Hillman et al., 2012). To generate the luciferase reporter plasmid, the minimum 35bp attB sequence (GGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCG) was cloned into pGL4.21 (Promega) downstream of the puromycin selection cassette between restriction sites BstBI and SalI, creating pGL4.21-attB. Next, an RBL1 promoter fragment, corresponding to nucleotides -287 to +42 relative to the annotated transcription start site (NCBI RefSeq mRNA NM_002895.3), was PCR-amplified from human genomic DNA and cloned into pGL4.21-attB between KpnI and EcoRV restriction enzyme sites. The -287/+42 RBL1 promoter containing the tandem E2F binding site mutations described by Burkhart et al. (2010) was prepared as synthetic dsDNA (IDT) and cloned into pGL4.21-attB as described above. T98G cells in 35mm dishes were co-transfected with 1μg of a codon-optimized phiC31 expression vector (Raymond et al., 2007) (Addgene plasmid #13795) and 100ng of either pGL4.21-attB RBL1-WT or pGL4.21-attB RBL1-E2F mutant plasmids using FugeneHD (Promega) transfection reagent. Approximately 24 hours post-transfection cells were split into 10cm dishes and selected for stable integration with 1μg/ml puromycin. After 10 days of selection the several hundred surviving colonies were trypsinized and RBL1-WT or RBL1-E2F mutant cells were maintained as separate pools under continuous puromycin selection.

Retroviral shRNA transduction and siRNA transfection

Non-silencing control, THAP11, and HCF-1 pSuper.Retro.Puro retroviral shRNA expression vectors have been described previously (Parker et al., 2012). ZNF143, E2F1, E2F4, and DP1 shRNAs were cloned into BglII/HindII sites of pSuper.Retro.Puro as described (Parker et al., 2012) using oligonucleotides containing the following shRNA target sequences: ZNF143, GGACATGCTACAAGAGTAA; E2F1, GGGAGAAGTCACGCTATGA; E2F4, GGATTTACGACATTACCAA; DP1, GAGGAGACTTGAAAGAATA. Two shRNAs targeting HPV18 E7 (E7-A, GGAGTTAATCATCAACATT; E7-B, GGAAGAAAACGATGAAATA) were cloned into pSuper.Retro.Puro as described above. All retroviral shRNA constructs are available at the plasmid repository Addgene (http://www.addgene.org /Debu_Chakravarti/). VSV-G pseudotyped retrovirus was produced and used to spin-infect cells as previously described (Parker et al., 2012). Two days post-transduction, cells were split into media containing 2μg/ml puromycin and selected for at least two days or as indicated.

For knockdown using siRNA, HeLa cells were transfected with 30nM non-silencing control siRNA (Thermo Scientific) or 15nM each of THAP11 and ZNF143 siRNA using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions. Media was changed 24 hours post transfection, and transfected HeLa cells were harvested for RNA and protein isolation after 72 hour incubation. THAP11 and ZNF143 siRNA oligonucleotides corresponding to the shRNA targeting sequences were synthesized by Thermo Scientific and are as follows: ZNF143 sense, GGACAUGCUACAAGAGUAAUU; ZNF143 antisense,

UUACUCUUGUAGCAUGUCCUU; THAP11 sense, UGAUGGAAGUGAAGAUGAAUU; THAP11 antisense, UUCAUCUUCACUUCCAUCAUU.

Immunoblots and Quantitative RT-PCR

Whole cell extracts were prepared from cells expressing shRNA four days post-transduction using modified RIPA buffer (20mM Tris-HCl pH 7.6, 150mM NaCl, 1mM EDTA, 1mM EGTA, 1% IGEPAL CA-630, 1% sodium deoxycholate, 0.25% SDS). Extracts were clarified by centrifugation at 20,000 x *g* for 15 minutes at 4°C and protein concentrations determined by BCA assay (Pierce). Thirty micrograms of whole cell extract was separated by SDS-PAGE, transferred to nitrocellulose, and immunoblotted as described (Parker et al., 2012). Antibodies and working dilutions used for immunoblot are as follows: THAP11 (R&D Systems #MAB5727, 1:1000), HCF-1 (Bethyl Laboratories #A301-399A, 1:5000), ZNF143 (Novus Biologicals #H00007702-M01, 1:2000), E2F1 (Millipore #05-379, 1:2000), DP1 (Santa Cruz Biotechnology #sc-610x, 1:5000), GAPDH (Sigma #G-9545, 1:10,000), HPV18 E7 (Santa Cruz Biotechnology #sc-1590, 1:500). Immunoblots were developed by chemiluminescence using ECL Prime Western Blotting Detection Reagent (GE Healthcare).

Total RNA was prepared from shRNA expressing cells four days post-transduction or from siRNA transfected cells 3 days post-transfection using Qiagen RNeasy Mini Kit. Total RNA (350ng) was reverse transcribed using qScript cDNA Synthesis Kit (Quanta Biosciences) according to manufacturer's instructions. Quantitative PCR was performed on diluted cDNA using an Applied Biosystems ABI PRISM 7900HT 384-well real time PCR machine in a final volume of 20μl using SYBR

green PCR master mix (Applied Biosystems). HPV18 E6/E7 transcripts were detected using the primer sequences described by Magaldi et al.

Immunoprecipitation

SW620 cells in 15cm tissue culture dishes were rinsed three times with ice-cold PBS, scraped into PBS and collected by centrifugation at 500 x g for five minutes at 4°C. Cells were washed once with Buffer A (10mM HEPES-KOH pH 7.6, 10mM KCl, 1.5mM MgCl2) and then resuspended in 2.5 pellet cell volumes (PCV) of Buffer A plus 0.34M sucrose and Complete protease inhibitors (Roche). Cytoplasmic membranes were lysed by addition of 2.5 PCV of Buffer A, 0.34M sucrose, 0.2% Triton X-100 while gently mixing the cells by vortexing at half-maximum setting. Cells were incubated on ice for 10 minutes and nuclei were isolated by centrifugation at 2000 x g for 5 minutes at 4°C. Nuclei were washed once with Buffer A without sucrose, and resuspended in 1 PCV of Buffer C (20mM HEPES-KOH pH 7.6, 420mM NaCl, 1.5mM MgCl2, 0.2mM EDTA, 25% glycerol) supplemented with Complete protease inhibitors (Roche). Nuclei were extracted for one hour at 4°C with gentle inversion. Nuclear extracts were clarified by centrifugation (20,000 x g for 15 minutes at 4°C), diluted with one volume of Buffer C without glycerol or NaCl, and adjusted to 0.2% Triton X-100. Nuclear extracts were re-clarified by centrifugation to remove precipitates formed by dilution and immunoprecipitations were performed using the following antibodies: THAP11 (R&D Systems #MAB5727, 2μg), HCF 1 (Bethyl Laboratories #A301-399A, 1μg). Immunoprecipitations were performed for two hours at 4°C with inversion. Protein G Dynabeads (20μl) were added and immunoprecipitation continued for an additional two hours. Beads were then washed four times with binding buffer (20mM HEPES-KOH pH

7.6, 150mM NaCl, 1.5mM MgCl2, 0.2mM EDTA, 0.2% Triton X-100) and bound proteins eluted by boiling in 2x Laemmli buffer. Immunoprecipitated proteins were resolved by SDS-PAGE and immunoblotted with HCF-1 and THAP11 antibodies as described. Blots were developed by enhanced chemiluminescence using an appropriate anti- light chain specific HRP-conjugated secondary antibody (Jackson Immunoresearch).

Flow Cytometry

Cell cycle profile was determined by propidium iodide (PI) staining flow cytometry. Transduced cells (~ 1×10^6) were harvested by trypsinization and collected by centrifugation at 500 × *g* for 5 min at 4°C. Cell pellets were washed twice with ice-cold PBS, and resuspended in cell fixation buffer (200µL ice-cold PBS + 800µL ice-cold ethanol). After overnight fixation at -20°C, cells were recovered by centrifugation (500 × *g* for 5 min at 4°C), washed twice using ice-cold PBS and incubated with 1mL PI staining buffer (50µg/mL propidium iodide, 0.1% Triton-X, 2mg RNAse A) at 37°C for 30 minutes. Propidium iodide stained cells were analyzed by flow cytometry using a BD LSRFortessa Analyzer (BD Biosciences).

Apoptosis was determined by Annexin V/PI staining flow cytometry. Adherent cells were harvested by trypsinization and collected by centrifugation (500 × *g* for 5 min at 4°C), while floating cells were directly collected from medium by centrifugation (500 × *g* for 5 min at 4°C). Adherent and floating cells were combined, washed using ice-cold PBS, and resuspended in Annexin binding buffer (10mM HEPES-KOH pH 7.6, 140mM NaCl, 5mM $CaCl_2$). Final concentration was adjusted to 1×10^6 cells/mL. Next, 100µL of cell suspension was incubated with 5µL AlexaFluor488 conjugated Annexin V (Invitrogen) and 0.5µL propidium iodide (1mg/mL) for 15 min at room temperature. After

the incubation period, each sample was gently diluted with 400µL of Annexin binding buffer on ice and analyzed by flow cytometry using a Beckman Coulter Epics XL Analyzer (BD Biosciences).

In Silico Analyses

To obtain the Thap11 DNA binding matrix, we performed *de novo* motif discovery analysis on previously-published mouse Thap11 ChIP-Seq data (Dejosez et al., 2010) using MEME software v4.8.1 (Bailey et al., 2009; Bailey and Elkan, 1994). The derived DNA binding matrix was used as input to MAST software v4.8.1 (Bailey et al., 2009; Bailey and Gribskov, 1998) to search for potential THAP11 binding sites within 400bp of ENCODE HA-E2F1 binding sites in HeLa-S3 cells (Consortium, 2011). The potential THAP11 binding sites were ranked by their similarity to the THAP11 DNA binding matrix and mapped to the nearest gene using R ChIPPeakAnno package (Zhu et al., 2010).

ChIP-seq data was retrieved from the NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/sra): HCF-1 N [SRX092569], ZNF143 [SRX159127], THAP11 [SRX092574], E2F1 [SRX150563], ZNF143 (ENCODE) [SRX150515]. SRX092569, SRX159127, and SRX092574 data were previously published by Michaud et al. (2013); SRX150563 and SRX150515 were deposited to SRA by the ENCODE consortium. FastQ data was extracted using `fastq-dump` (v2.3.4, SRA Toolkit, http://eutils.ncbi.nih.gov/Traces/sra/?view=software) and concatenated into a single file where multiple sequencing files were provided per sample. Sequencing reads were mapped to the human hg19 genome using bowtie (v1.0.0, http://bowtie-bio.sourceforge.net) with `-v 0 -m 1 -p 4 -t -S hg19` parameters followed by HOMER's (v4.3, http://homer.salk.edu) `makeTagDirectory` with `-tbp 1` flag. Peaks

were called using HOMER's `findPeaks.pl` (`-style factor`, FDR of 0.001) and using respective ChIP-seq inputs for background filtering. Co-localized peaks within 500bp of each other (from peak centers) were found using `intersectBed` (bedtools, `-wa` flag). Distances between peaks were calculated from peak centers for co-localized peak pairs. Chromatin occupancy correlation analysis was performed only on co-localized peak pairs. Given such peak pairs, HCF-1 and corresponding transcription factor ChIP-seq tags in the 1000bp region surrounding the transcription factor were counted using HOMER's `annotatePeaks.pl` with `–size 1000 –log` parameters. Pearson product-moment correlation coefficient was calculated on inverse $\log_2$-transformed values using GraphPad. E2F1 peaks lacking THAP11 and ZNF143 peaks within 500bp were determined using `intersectBed` with `–v` flag. These peaks were then intersected with HCF-1 peaks using `intersectBed` with `–f 0.5 –wa` flags to obtain E2F1/HCF-1 overlapping peaks devoid of nearby THAP11/ZNF143 binding.

Supplemental References

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol *2*, 28-36.

Magaldi, T.G., Almstead, L.L., Bellone, S., Prevatt, E.G., Santin, A.D., and DiMaio, D. (2012). Primary human cervical carcinoma cells require human papillomavirus E6 and E7 expression for ongoing proliferation. Virology *422*, 114-124.

Zhu, L.J., Gazin, C., Lawson, N.D., Pages, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics *11*, 237.

Supplemental Data

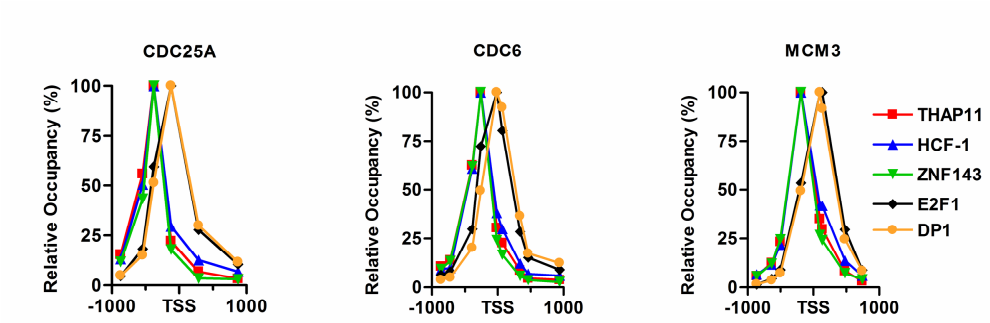

Figure S1 Related to Figure 1: HCF-1 occupancy at E2F target genes correlates with THAP11 and ZNF143 binding. Distribution of THAP11, HCF-1, ZNF143, E2F1, and DP1 on the promoter-proximal regions of E2F1 target genes in T98G cells as determined by ChIP-scanning assays.

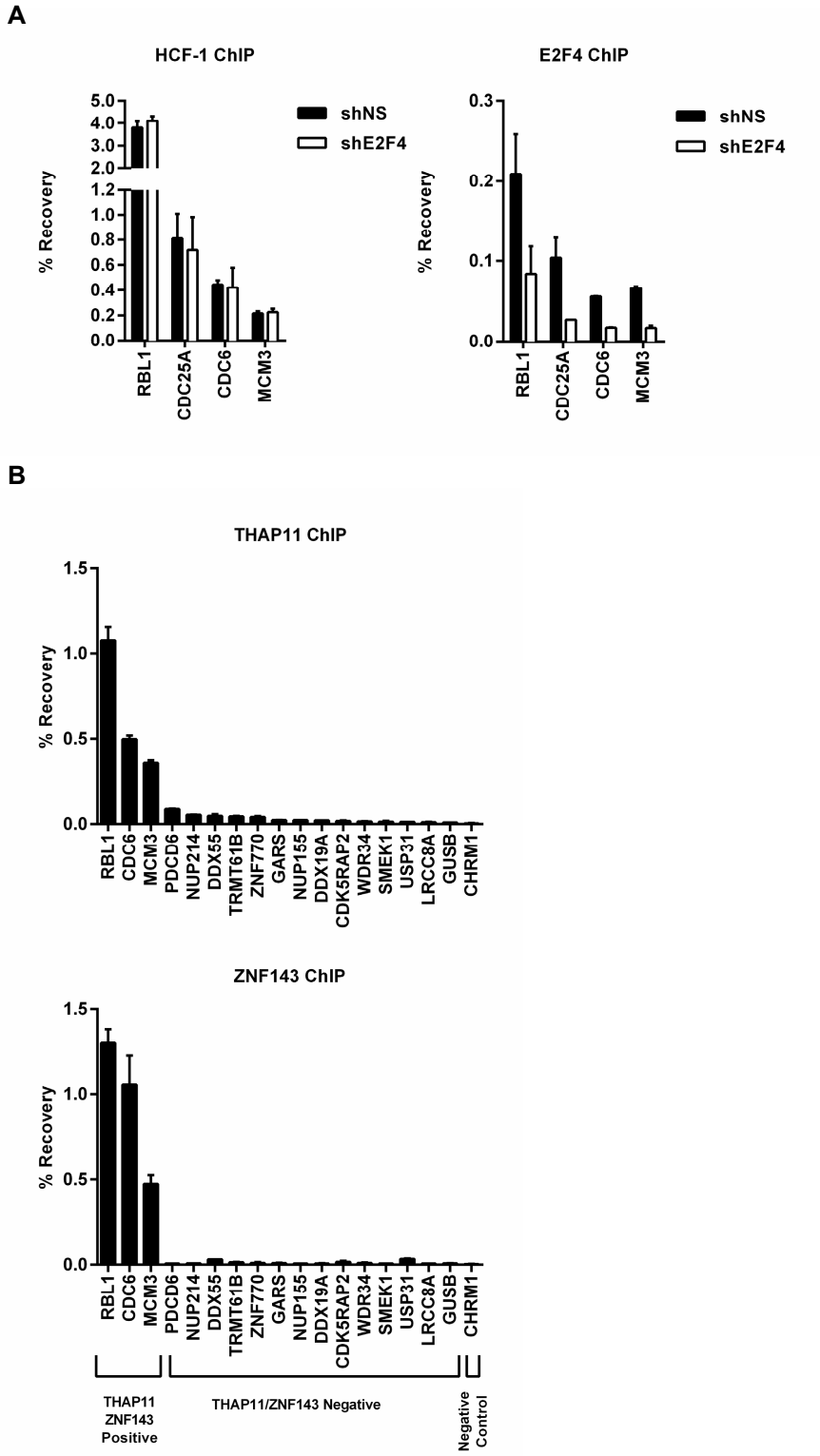

Figure S2 Related to Figure 4: E2Fs are dispensable for HCF-1 promoter occupancy. (A) HCF-1 and E2F4 occupancy at the indicated promoters was determined by ChIP assays in HeLa cells expressing either control (shNS) or E2F4 shRNA. (B) Validation of HCF-1/E2F1-bound promoters as THAP11- and ZNF143-unbound was performed by ChIP. In panels A and B, values represent mean ± standard deviation of duplicate PCR reactions from a single experiment performed at least three times with similar results.

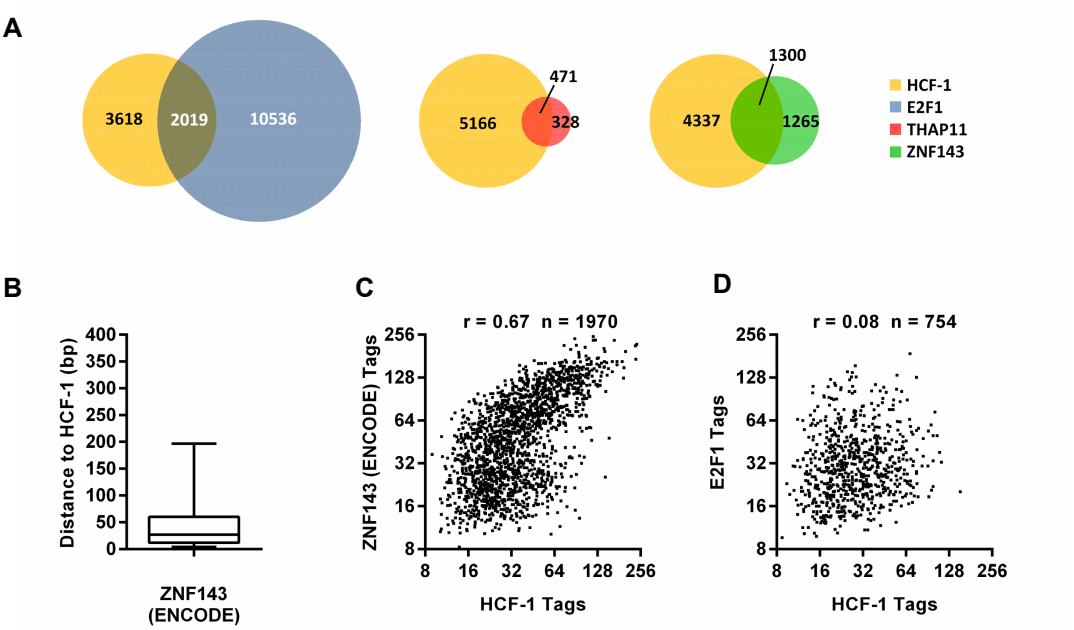

Figure S3 Related to Figure 5: THAP11 and ZNF143, but not E2F1, genome-wide chromatin occupancy correlates HCF-1 binding. (A) Co-localized ChIP-seq peaks were determined as peaks with centers located no more than 500 bases apart. (B) Absolute distances between HCF-1 and ZNF143 (ENCODE) peak centers are shown as box-and-whisker plots. Whiskers indicate 10% and 90% of the population. (C) Correlation of ZNF143 (ENCODE) tag counts with HCF-1 tag counts in overlapping regions. (D) Correlation of E2F1 tag counts with HCF-1 tag counts in overlapping regions devoid of ZNF143 or THAP11 peaks. (C and D) n = total number of peak pairs. r = Pearson product-moment correlation coefficient.

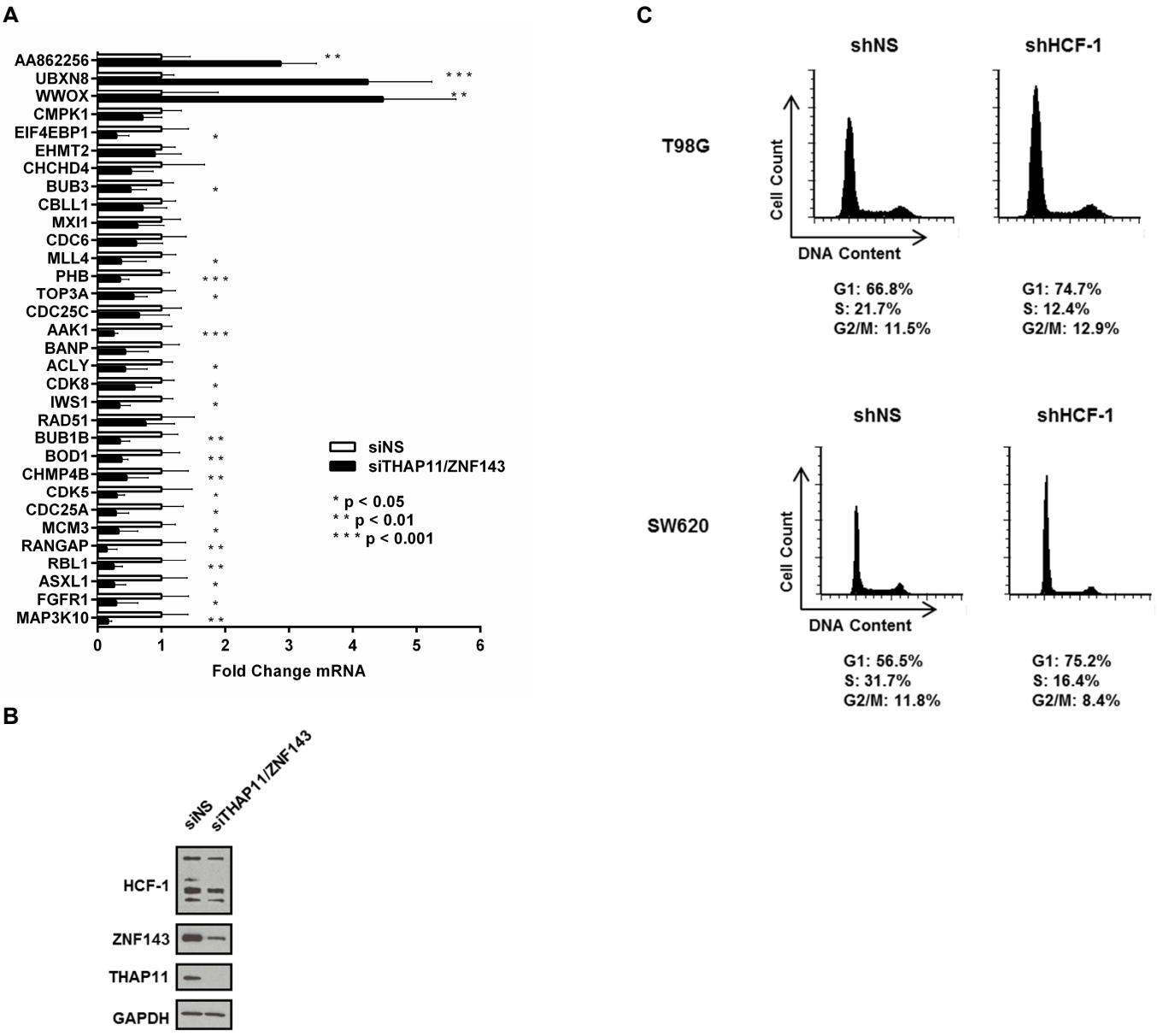

Figure S4 Related to Figure 7: THAP11/ZNF143/HCF-1-dependent gene expression contributes to cell proliferation and cell cycle progression. (A) mRNA expression changes of the same genes in Fig.7A were determined by quantitative RT-PCR in HeLa cells transfected with the control (siNS) or THAP11 and ZNF143 siRNA. Values represent the mean ± standard deviation of four independent experiments. Student's T-test p-values are indicated. (B) Representative immunoblot from cells in panel A. (C) Cell cycle analysis of T98G and SW620 cells collected four days post-transduction with the indicated shRNAs.