Figure S1. Kabir et al.

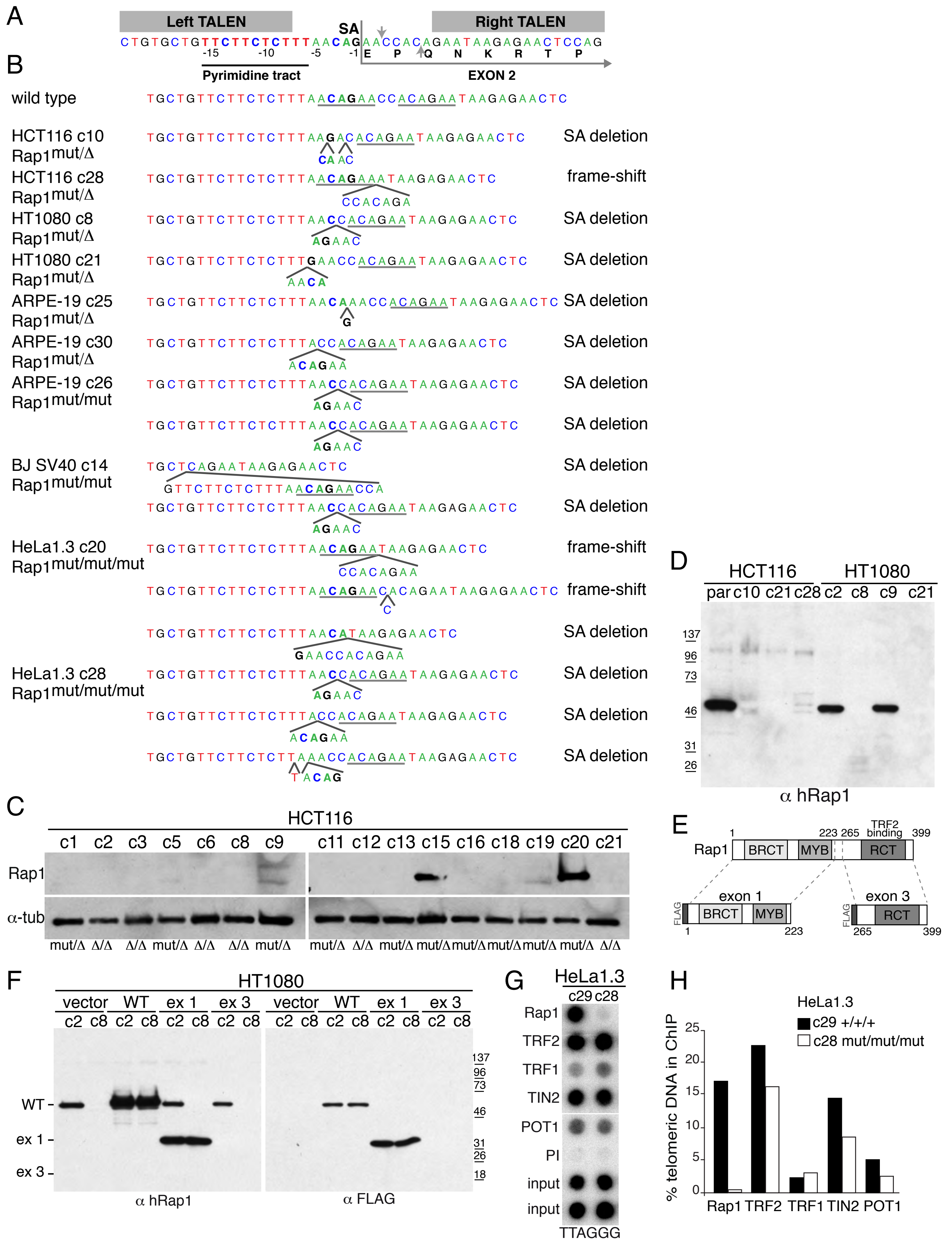


Figure S2. Kabir et al.

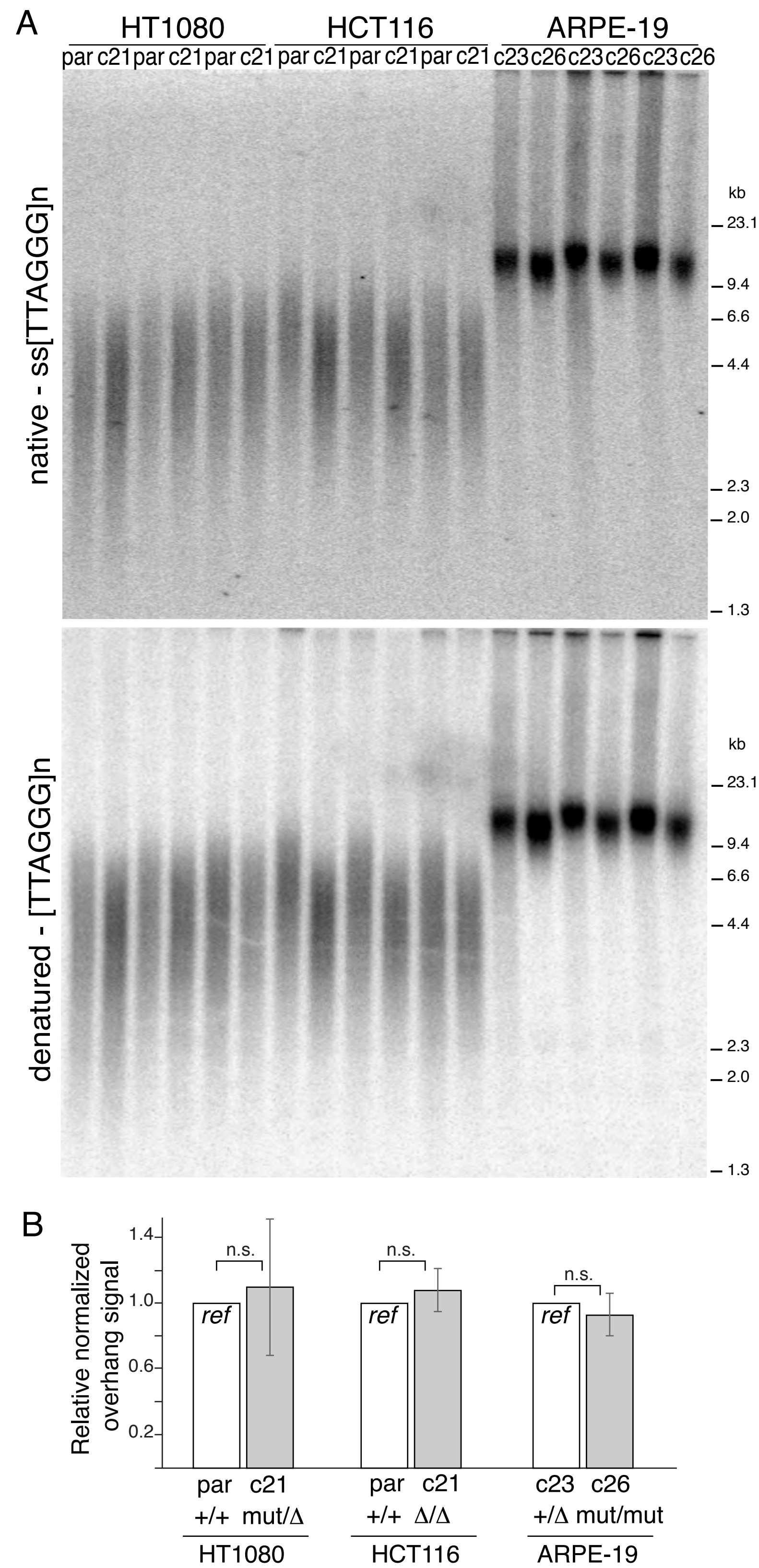
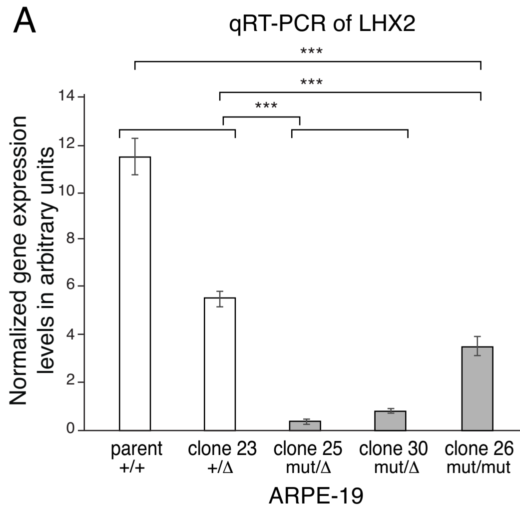
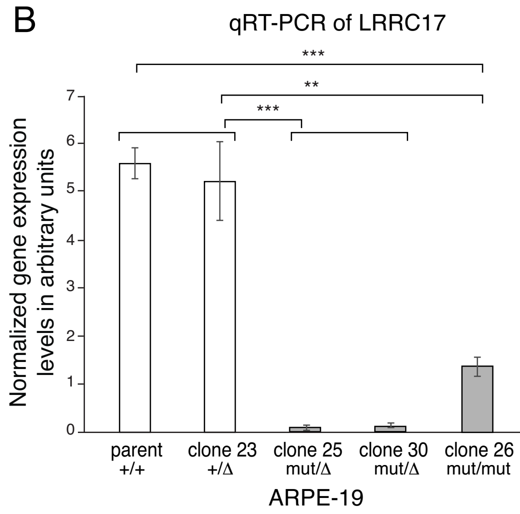
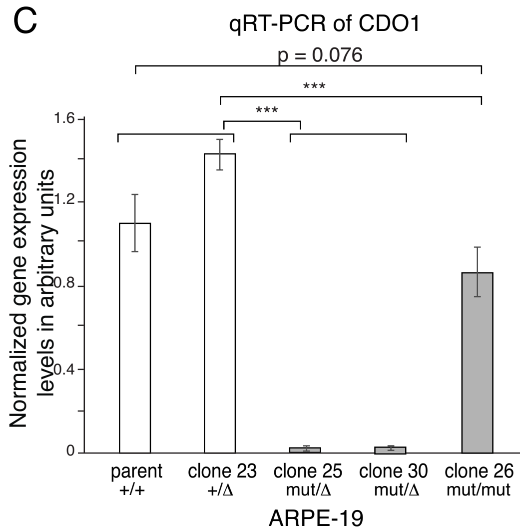
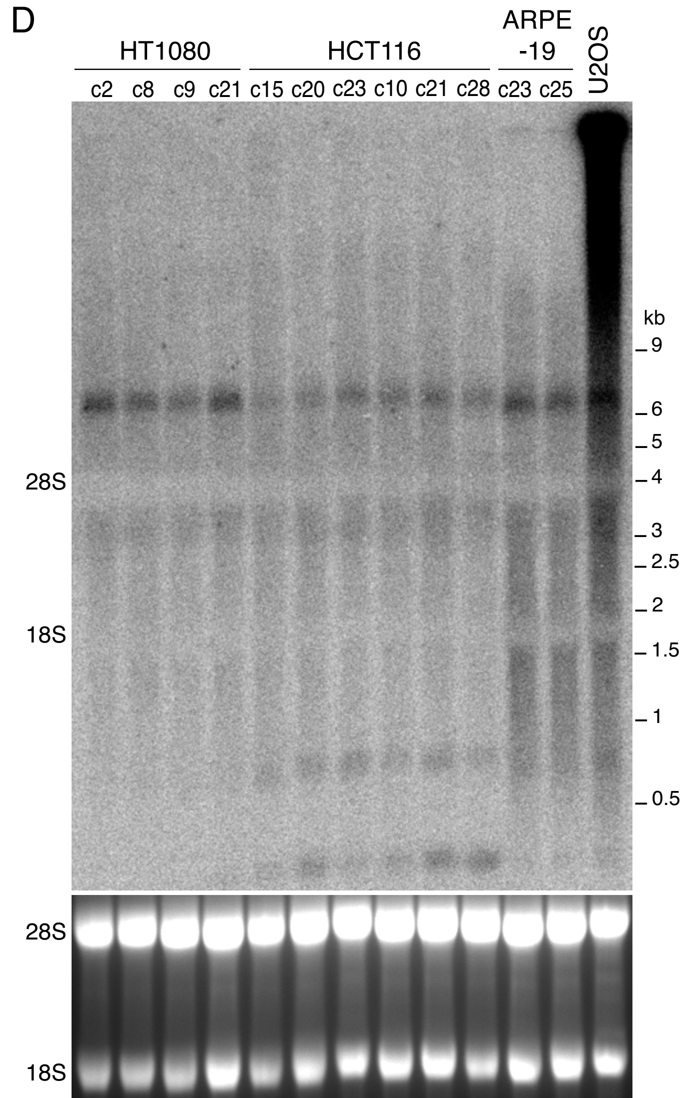


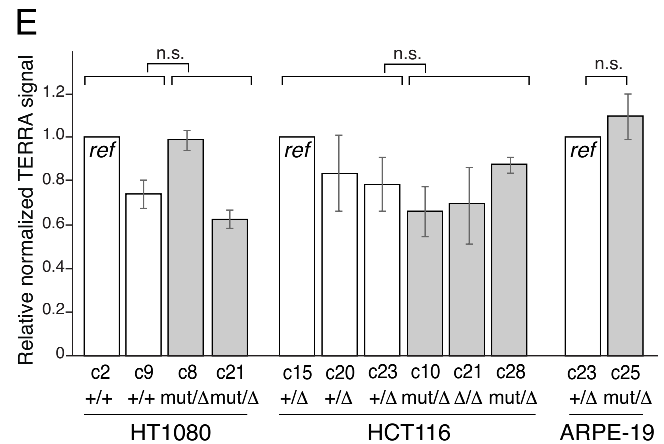
Figure S3. Kabir et al.











Supplemental Information

Supplemental Figures and Figure Legends

Figure S1, Related to Figure 1.

TALEN-induced mutations in the *TERF2IP* locus.

(A) Schematic illustrating TALEN binding sites, their predicted cutting site (grey arrows), the pyrimidine tract, splice acceptor (SA), and the location of exon 2 in the *TERF2IP* locus. A short repeat (underlined in grey) is frequently mutated in the mutant alleles.

(B) Relevant sequences of the indicated clones. Predicted consequences of the mutations are indicated on the right.

(C) Immunoblotting of protein from targeted homozygous and heterozygous HCT116 clones showing absence of Rap1 in a large proportion of heterozygous clones. Genotypes indicated below.

(D) Immunoblots of HCT116 and HT1080 KO clones probed with anti-hRap1 showing the absence of detectable truncated Rap1 proteins.

(E) Schematic of retroviral constructs expressing FLAG-tagged exon 1 (ex 1) and FLAG-tagged exon 3 (ex 3).

(F) Immunoblotting for the expression of constructs in (C) in HT1080 KO and WT clones. The protein fragment encoded by exon 1 is expressed and detected by the Rap1 antibody. The protein fragment encoded by exon 3 contains epitopes that the Rap1 antibody can recognize, but lack of detectable signal indicates that it is not expressed.

(G) Telomeric ChIP of HeLa 1.3 WT and KO cells.

(H) Quantification of the percent of telomeric DNA recovered in each ChIP. Methods as in Figure 1D and E.

Figure S2, Related to Figure 3.

No change in the telomeric overhangs upon Rap1 loss.

(A) Telomeric DNA analysis of WT and KO clones in three different cell lines. Top, in-gel detection of native telomeric restriction fragments with a C-strand telomeric probe revealing the G-strand overhang signals; bottom, same gel rehybridized after in situ denaturation of the DNA, revealing the total telomeric DNA in each lane. Three biological replicates of each cell line were run adjacent to each other on the same gel.

(B) Overhang signals were normalized to total telomeric signals in each lane and plotted as a ratio compared to the WT overhang signal of the corresponding cell line. Error bars represent SDs. Lack of statistical significance derived from two-tailed paired t-tests using the three independent experiments conducted with each cell line.

Figure S3, Related to Figure 4.

Validation of Rap1 as a transcriptional regulator, but not of TERRA.

(A-C) Quantitative RT-PCR illustrates differential expression of three genes ((A) LHX2, (B) LRRC17, (C) CDO1) in Rap1 WT and KO ARPE-19 cells. Gene expression was normalized to GAPDH and mean expression level for each gene as determined by the ΔCt method from 3 independent replicates is graphed in arbitrary units. Error bars represent SDs. Significance was calculated by two-tailed unpaired T-Tests.

(D) Northern blot hybridized with a telomeric probe (Sty11) showing TERRA levels of Rap1 WT and KO clones in the HT1080, HCT116 and ARPE-19 cell lines. U2OS serves as a positive control for TERRA expression. Ethidium bromide staining of ribosomal RNA serves as a loading control.

(E) TERRA signals were normalized using the 18S ribosomal RNA and plotted as a ratio compared to the WT TERRA signal of the corresponding cell line. Error bars represent SDs. Two-tailed unpaired t-tests of 3 independent experiments illustrates lack of statistical significance between Rap1 WT and KO TERRA expression levels.

Table S1. Microarray expression analysis for ARPE-19, related to Figure 4.

Cell Line	Gene (chromosome) [LogFC][a]		
KO v WT	LHX2 (chr9)	LRRC17 (chr7)	CDO1 (chr5)
c25 v c23	-3.3	-5.1, -3.4	-3.5
c30 v c23	-2.7	-5.1, -3.4	-3.8
c25 v par	-4.2	-4.6, -3.1	-3.3
c30 v par	-3.7	-4.5, -3.0	-3.5

[a],Multiple LogFC values reflect data from multiple probes for the corresponding gene.

Table S2. Microarray expression analysis for HT1080, related to Figure 4.

Cell Line	Gene (chromosome) [LogFC][a]							
KO v WT	ATP9A (chr20)	CDCP1[b] (chr3)	CYP2J2 (chr1)	FAIM3 (chr1)	MGC 39900	NELL2 (chr12)	PTGR1 (chr9)	TERF2IP (chr16)
c8 v c2	3.0	3.4, 1.8, 2.5, -	2.0	-3.4	-2.1	3.0	2.0, 1.7	-2.9
c21 v c2	3.1	3.7, 1.9, 2.4, 3.0	2.2	-3.6	-2.3	2.2	2.1, 1.9	-2.3
c8 v c9	3.8	3.6, 1.8, 2.3, -	1.6	-2.0	-2.7	2.9	2.0, 1.7	-2.9
c21 v c9	3.9	3.9, 1.8, 2.3, 2.8	3.5	-2.2	-2.9	2.2	2.1, 1.9	-2.3

[a],Multiple LogFC values reflect data from multiple probes for the corresponding gene. [b] CDCP1 was identified as a Rap1 associated locus by ChIP-seq in a subclone of HT1080 (Yang et al., 2011).

Table S3. Microarray expression analysis for HCT116, related to Figure 4.

Cell Line	Gene (chromosome) [LogFC] [a]		
KO v WT	BMP4 (chr14)	SLC2A3 (chr12)	SUSD2 (chr22)
c10 v c15	-2.0	-3.4	-2.6
c21 v c15	-2.1	-2.7	-2.2
c28 v c15	-1.9	-2.3	-2.1
c10 v c20	-2.0	-2.7	-2.0
c21 v c20	-2.0, -2.5	-2.0	-1.6
c28 v c20	-1.8	-1.5	-1.5
c10 v c23	-1.8	-2.9	-2.8
c21 v c23	-1.9, -2.4	-2.1	-2.5
c28 v c23	-1.7	-1.7	-2.3

[a],Multiple LogFC values reflect data from multiple probes for the corresponding gene.

Supplemental Experimental Procedures

TALENs and *TERF2IP* Targeting Construct

The heterodimeric TALEN pair for *TERF2IP* targeting was constructed using the following RVD sequences. LEFT2: 5'-HD-NG-NN-NG-NN-HD-NG-NN-NG-NG-HD-NG-NG-HD-NG-HD-NG, RIGHT1: 5'-HD-NG-NN-NN-NI-NN-NG-NG-HD-NG-HD-NG-NG-NI-NG-NG-3'. The PGK Neomycin cassette from the PL451 vector (NCI) was liberated using restriction enzymes *Nhe*I (5') and *Bst*BI (3') and ligated into *Nhe*I- and *Bst*BI- digested pSL301 (cloning vector from Invitrogen). The pEF Blasticidin cassette from plasmid pEF/Bsd (Life Technologies) was released using *Nhe*I (5') and *Eco*RI (3') and ligated into *Nhe*I- and *Eco*RI- digested pSL301. The 5' and 3' homology arms were PCR-amplified with restriction site overhangs from genomic SV40LT BJ fibroblast DNA. Primers for PCR of the 5' arm were as follows: 5'-ATGCGGTACCTTGCCCAAACTCCTGTCTTCTTAGGGC-3' and 5'-GCATGCTAGCAGAGAAGAACAGCACAGATTAGCAATAGCC-3'. Primers for PCR of the 3' arm were 5'-ATGCTTCGAACTAGATTTACTCATTATTTTTTCCCTACC-3' and 5'-GCATTTCGAACCTGTAATCCCAGCACTTTGGGAG-3'. The resulting 600 bp 5' homology arm ends 7 bp from the intron 1/exon 2 junction and has *Kpn*I and *Nhe*I sites on the 5' and 3' ends, respectively. The resulting 578 bp 3' homology arm starts 32 bp from the exon 2/intron 2 junction and has *Bst*BI restriction sites at both ends. The homology arms were cloned into the relevant restriction sites in pSL301 containing either the PGK Neomycin or pEF Blasticidin. The 3' homology arm insertion was screened for orientation and the donor constructs were sequenced using the following primers: T7, T3, 5'-GCTCGCGTCGTGCAGGACGT-3' (PGK internal primer), and 5'- GCTGTGCTCGACGTTGTCAC-3' (Neomycin internal primer).

Cell Culture

HCT116, HT080, ARPE-19, and HeLa1.3 cells were grown in Dulbecco modified Eagle medium (DMEM) supplemented with L-glutamine, penicillin-streptomycin, nonessential amino acids and 15% bovine calf serum (BCS) (HyClone). SV40-large T transformed BJ fibroblasts (neo resistant) were cultured in complete DMEM containing 199 medium (4:1) and 10% BCS.

***TERF2IP* Targeting and Cell cloning**

All cell lines were transfected in 10 cm dishes at a density of $3x10^6$ cells using Lipofectamine 2000 (Life Technologies) with 4 μg of each TALEN construct and 20 μg of the donor construct. 48 hours after transfection, all cell lines, except for HeLa1.3, were plated in selection medium in 10 cm plates at varying densities ranging from 3,900 to 500,000 cells (using two-fold dilutions). HeLa1.3 cells were plated in 10 cm dishes using two-fold dilutions starting from 8000 cells down to 75 cells. G418 was used at 1 mg/ml to select neomycin-resistant HCT116 cells, at 900 μg/ml for HT1080 cells, and at 800 μg/ml for ARPE-19 cells. Blasticidin was used at 5 μg/ml for HeLa1.3 selection and at 2.5 μg/ml for SV40LT BJ selection. Clones emerged at a frequency of approximately 1 clone per 500 plated HCT116 cells, 1 clone per 2,600 plated HT1080 cells, 1 clone per 7,800 plated ARPE-19 cells, 1 clone per 125 plated HeLa1.3 cells, and 1 clone per 62,500 plated SV40LT BJ cells. The media was not changed after initial plating. Clones were picked 12 days later for all cell lines except the SV40LT BJ clones, which were picked 3 weeks after plating. Approximately 60-70 clones were picked for each cell line using cloning cylinders from plates that contained well-spaced clones and the cells were transferred into 24 well plates. After reaching confluence, half of the cells in each well were harvested to extract genomic DNA, while the remaining cells were expanded into 6 well plates. Approximately 30-40 clones were screened per cell line.

Genotyping and Sequencing

Genotyping PCR used the following primers: F1: 5'- GTGGATTGTGGTACGT GGCCCAGATCTGCC-3'; R1: 5'-TAACATACCACAACCTCCTCAAACTCCCGG-3'; R2: 5'-CATCTGCACGAGACTAGTGAGACGTGCTAC-3'. PCR was performed in 25 μl containing 50 ng of DNA, 0.2 μM of each primer, 0.1 μM dNTPs, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl pH 8.0, and 0.5 U of TaKaRa Taq polymerase. Conditions were as follows: 95°C for 4 min, 25 cycles of 95°C for 30 sec, 62°C for 30 sec, 72°C for 45 sec and final extension at 72°C for 5 min. Clones identified as heterozygous by PCR genotyping and southern blotting but lacking Rap1 protein on western blots were sequenced by PCR amplifying the remaining 'WT' allele with primers F1 and R1 using the PCR conditions described above. The PCR products were eluted from agarose gels and sequenced with primers F1 and R1. The relevant *TERF2IP* PCR product from clones that lacked detectable Rap1 protein but appeared to be untargeted (homozygous 'WT') according to PCR genotyping and southern blotting were PCR amplified using primers F1 and R1, gel-eluted and cloned by TA cloning (Life Technologies). A minimum of 8 resulting TA clones per cell line were sequenced to identify mutations in both alleles. The latter procedure was applied to clones from the ARPE-19, SV40LT BJ, and HeLa1.3 cell lines.

Immunoblotting

Immunoblotting was performed as previously described (Takai et al., 2010). Briefly, cells were harvested by trypsinization, resuspended in Laemmli buffer (100mM Tris-HCl pH 6.8, 200 mM DTT, 3% SDS, 20% glycerol, 0.05% bromophenol blue) at 10,000 cells per μl, denatured for 5 min at 95°C, sheared with an insulin needle, and resolved on SDS/PAGE gels using 100,000 cells per lane. Blots were blocked in 5% non-fat dry milk/PBS+0.1% Tween20. Antibodies used were as follows: Rap1 (765, rabbit polyclonal); alpha-tubulin (Sigma T9026). The hRap1 Ab was generated using full-length hRap1 recombinant protein as an antigen and reacts with multiple Rap1 domains.

Telomeric ChIP

Telomeric ChIP was conducted as previously described (Loayza and de Lange, 2003). The following antibodies or crude sera were used: Rap1 (765, rabbit polyclonal); TRF2 (647, rabbit polyclonal); TRF1 (371, rabbit polyclonal); TIN2 (864, rabbit polyclonal); POT1 (Abcam ab124784); TPP1 (1151; rabbit polyclonal); H3K9me1 (Abcam ab9045); H3K9me2 (Abcam ab1220); H3K9me3 (Abcam ab8898); HP1alpha (Abcam ab77256); HP1beta (Abcam ab10478); HP1gamma (Abcam ab10480); Acetyl Histone H4 (Millipore 06-598).

IF-FISH

IF-FISH was conducted as previously described (Dimitrova and de Lange, 2006). Briefly, cells grown on coverslips were fixed for 10 min in 2% paraformaldehyde/3% sucrose at room temperature, followed by three 5 min PBS washes. Coverslips were incubated in blocking solution (1 mg/ml BSA, 3% goat serum, 0.1% TritonX-100, 1 mM EDTA in PBS) for 30 min, followed by incubation with primary antibodies diluted in blocking solution for 1 h at rot. Primary antibodies used were: 53BP1 (Novus 100-304) and Lamin A (Sigma L1293). Cells were washed three times for 5 min with PBS and then incubated with secondary antibodies diluted in blocking solution for 30 min at rt. The secondary antibody used was RRX-anti-rabbit (Jackson 711-295-152). Coverslips were dehydrated with 70%, 95% and 100% ethanol and allowed to dry. Hybridizing solution (70% formamide, 0.5% blocking reagent from Roche, 10 mM Tris-HCl pH 7.2, FITC-OO-(CCCTAA)3 PNA probe from Applied Biosystems) was added to each coverslip and denatured at 80ºC for 5 min, followed by a 2 h incubation at rt. Two 15-min washes in 70% formamide/10 mM Tris-HCl pH 7.2 and three 5-min washes with PBS were performed. DNA was stained with DAPI in the PBS washes and coverslips were mounted using antifade reagent ProLong Gold from Life Technologies. Images were captured using a Zeiss AxioPlan II microscope with a Hamamatsu C4742-95 camera using Volocity software from Perkin Elmer. Distances of telomeres from nuclear membrane were calculated using Image J software.

Telomeric FISH and CO-FISH

Telomeric FISH and CO-FISH were conducted as previously described (van Steensel et al., 1998; Celli et al., 2006). Briefly, colcemid was added to cells 2 hours prior to harvest. Cells were collected by trypsinization, swollen in 0.075 M KCl and fixed overnight at 4ºC in methanol:acetic acid (3:1). Metaphase spreads were dropped on glass slides and aged overnight. Slides were hybridized with FITC-OO-$[CCCTAA]_3$ PNA probe in hybridizing solution, denatured at 80ºC for 5 min and incubated for 2 h at rt. Two 15-min washes in 70% formamide/10 mM Tris-HCl pH 7.2 and three 5 min washes with 0.1 M Tris-HCl pH 7.2/0.15 M NaCl/0.08% Tween20 were performed. DAPI was added to last wash for DNA stain. Slides were dehydrated in 70%, 95%, and 100% ethanol and mounted using ProLong Gold antifade from Life Technologies. Images were captured using a Zeiss AxioPlan II microscope with a Hamamatsu C4742-95 camera using Volocity software from Perkin Elmer. For CO-FISH, BrdU:BrdC (3:1) was added 14 h prior to harvest. Harvesting and metaphase conditions were as described for FISH. Slides were treated with 0.5 mg/ml RNase A diluted in PBS, stained with 0.5 μg/ml Hoechst 33258, exposed to 5400J/m^2 of UV light and subsequently digested with 800 U of Exonuclease III from Promega for 10 min at room temperature. Slides were rinsed with PBS, dehydrated with 70%, 95% and 100% ethanol and sequentially hybridized with TAMRA-OO-$[TTAGGG]_3$ and FITC-OO-$[CCCTAA]_3$ for 2 h each at room temperature, without denaturation. Washing, mounting and capture conditions were as described for FISH.

Genomic Blotting, Telomere Overhang and Telomere Length Analysis

Cells were harvested by trypsinization, washed with PBS and either pelleted and frozen at -80ºC (for telomere length analysis) or processed immediately (for genotyping and telomere overhang analysis) for genomic DNA collection. Genomic DNA was extracted as previously described (de Lange et al., 1990). DNA for genotyping was digested with *EcoR*I, quantitated by fluorometry using Hoechst 33258 and 10 μg was loaded on a 0.7% agarose gel run in 0.5X

TBE. DNA for telomere overhang and length analysis was digested with *Mbo*I and *Alu*I, quantified using Hoechst, and 1 μg was run on 0.7% agarose gels in 0.5XTBE. For genomic blots used for genotyping and telomere length analysis, the gels were depurinated with 0.5N HCl, denatured and neutralized using standard Southern blotting procedures and transferred as previously described (de Lange et al., 1990). Blots were probed with a *TERF2IP* 5' arm probe (indicated in Fig. 1a, Klenow-labeled using random primers and α-^{32}P-dCTP) or a Sty11 probe (de Lange, 1992) for genotyping and telomere length analysis, respectively. For telomere overhang analysis, gels were dried and probed with a [CCCTAA]4 end-labeled with Polynucleotide kinase and γ-^{32}P-ATP as previously described (Karlseder et al., 2002). Gels and membranes were exposed to PhosphorImager screens and quantified with ImageQuant software.

Northern Analysis for TERRA

Total RNA was prepared using RNeasy Mini Spin columns (QIAGEN) according to the manufacturer's instructions and northern blot analysis was performed as previously described (Azzalin et al., 2007). Briefly, 20μg of RNA was loaded onto 1.2% formaldehyde agarose gels and separated by gel electrophoresis. RNA was transferred to a Hybond-N+ membrane and crosslinked in a UV Stratalinker. The blot was prehybridized in Church mix at 55°C for 1 hour, followed by overnight hybridization with a Sty11 probe (de Lange, 1992). The blot was washed 3 times for 15 minutes at 55°C with Church wash and then exposed to a Phosphorimager screen for 5 days. Screens were scanned using ImageQuant software and quantified in Image J using the ethidium bromide stained 18S RNA as a loading control for normalization.

Microarray and qRT-PCR Analyses

Total RNA was isolated from cell lines using RNeasy Mini spin columns (QIAGEN) with DNase digestion, according to the manufacturer's instructions. Microarray hybridization and scanning were performed at the Genomics core facility at Rockefeller University, using Whole Human Genome DNA microarrays (Illumina HumanHT-12 v4). The data was analyzed using GeneSpring v12.6. Normalization was performed using quantiles and data was filtered to remove absent genes using flag calls. Experiments for HT1080 and HCT116 cell lines were performed in replicate, using two independent isolations of RNA. Differentially expressed genes were identified after performing moderated T-Tests and applying the Benjamini-Hochberg False Discovery Rate method. A further fold change of 3 or 2.75 was applied to the HT1080 and HCT116 clones respectively to identify genes that were highly transcriptionally deregulated due to the absence of Rap1. Microarrays for ARPE-19 were not performed in replicate and therefore an extremely stringent fold change threshold was applied to remove false negatives and identify differentially expressed genes, which were subsequently validated by qRT-PCR. For qRT-PCR, cDNA was prepared from 1μg of total RNA by using Thermoscript Reverse Transcriptase (Invitrogen). Quantitative PCR reactions were performed using Life Technologies SYBR Green Master Mix on an Applied Biosystems 7900HT Sequence Detection System. Differences between samples calculated using QuantStudio software (Applied Biosystems) using the ΔCT method and were normalized to GAPDH. Two independent isolations of RNA and reverse transcriptase reactions were conducted and the experiment was repeated six times for clones 23, 25, 30 and the parental cell line. The experiment for clone 26 was conducted in triplicate. Data was pooled to derive the mean averages and standard deviations. Significance was calculated using a two-tailed unpaired T-Test.

Supplemental References

Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science *318*, 798-801.

Celli, G. B., Lazzerini Denchi, E., and de Lange, T. (2006). Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination. Nat Cell Biol *8*, 885-890.

de Lange, T. (1992). Human telomeres are attached to the nuclear matrix. Embo J *11*, 717-724.

de Lange, T., Shiue, L., Myers, R. M., Cox, D. R., Naylor, S. L., Killery, A. M., and Varmus, H. E. (1990). Structure and variability of human chromosome ends. Mol Cell Biol *10*, 518-527.

Dimitrova, N., and de Lange, T. (2006). MDC1 accelerates nonhomologous end-joining of dysfunctional telomeres. Genes Dev *20*, 3238-3243.

Karlseder, J., Smogorzewska, A., and de Lange, T. (2002). Senescence induced by altered telomere state, not telomere loss. Science *295*, 2446-2449.

Loayza, D., and de Lange, T. (2003). POT1 as a terminal transducer of TRF1 telomere length control. Nature *424*, 1013-1018.

Takai, K. K., Hooper, S., Blackwood, S., Gandhi, R., and de Lange, T. (2010). In vivo stoichiometry of shelterin components. J Biol Chem *285*, 1457-1467.

van Steensel, B., Smogorzewska, A., and de Lange, T. (1998). TRF2 protects human telomeres from end-to-end fusions. Cell *92*, 401-413.

Yang, D., Xiong, Y., Kim, H., He, Q., Li, Y., Chen, R., and Songyang, Z. (2011). Human telomeric proteins occupy selective interstitial sites. Cell Res *21*, 1013-1027.