Supplemental Materials

The Supplemental Materials section includes the following items.

1. Supplemental Materials and Methods
2. Supplemental Figures and Figure Legends
3. Supplemental Tables
Supplemental Materials and Methods

Cell lines and cell culture

BJ human foreskin fibroblasts were purchased from ATCC and maintained in minimum essential medium (MEM, Corning, NY) supplemented with 10% FBS, 1 mM nonessential amino acids, and antibiotics [47,71]. LinX-A retroviral packaging cells [32] were grown in Dulbecco’s modified Eagle medium (Corning, NY) supplemented with 10% FBS and antibiotics. Primary human small airway epithelial cells (HSAEC, Life Cell Technology, MD) were grown in Epithelial Airway Medium Complete Kit (Lifeline, MD). Primary human pancreatic epithelial cells (Cell Biologics, IL) were grown in epithelial cell medium kit (Cell Biologics, IL). HEK-293T cells were grown in GlutaMAX™ Dulbecco’s modified Eagle medium (Gibco, MD) supplemented with 10% FBS and antibiotics. Cell lines are authenticated twice a year by expression profiling and by mycoplasma tests using Mycoplasma Detection Kit (R&D System).

Plasmids.

BabePuro-Ha-rasV12, -MEK1Q56P, -MKK3E, -MKK6E, -MKK4E, -MKK7D, -Raf-1 CAAX, -MEK1E(+), -MEK1-AA, -Myristoylated p110 and -AKT expression vectors were described previously [47]. cDNAs encoding the long and the short isoform of MZF1 were purchased from Origene (Rockville, MD) and Genescript (Piscataway, NJ), respectively, and subcloned into the lentiviral pLV vector (Biosettia, CA) according to manufacturer’s protocol.

Oligonucleotides for small hairpin RNAs (shRNAs) targeting both isoforms of MZF1 (sh-T1, 5’ GAAAGAGGAGTCAGAGGTTAC-3’ and sh-T2, 5’- GCAGGTGAAAGAGGAGTCAGA -3’), the long isoform of MZF1 (sh-L, 5’-CTCCGCAGGTTCCAGGTTGT-3’), and the short isoform of MZF1 (sh-S, 5’-ATCTTCTCCAGTTGAGGCAGGGA-3’) were designed, synthesized and cloned
into the pLV-H1-EF1α-puro vector according to the manufacture’s protocol (Biosettia, CA). Vectors encoding shRNAs for CHD7, c-Jun and Ets1, and the pBabePuro expression vector for TAM67, a dominant negative c-Jun mutant lacking the transactivation domain, have been reported before [32,47]. pBabe-Neo-hTERT and the retroviral luciferase reporter vector pBabeBlast-Luc were described previously [32].

**Retrovirus- and lentivirus-mediated gene transduction.**

Recombinant retroviruses and lentiviruses were packaged and transduced into cells as previously described [32]. Transduced cells were purified with 80 (BJ) or 50 (HSAECs and primary human pancreatic epithelial cells) µg/mL of hygromycin B, 1.5 (BJ) or 2 (HSAECs and primary human pancreatic epithelial cells) µg/mL of puromycin or 5 µg/mL of blasticidin (BJ).

**Analysis of senescence.**

HSAEC and primary human pancreatic epithelial cells were immortalized by transduction with pBabe-Neo-hTERT and selected in 400 µg/mL of G418 for 8 days. Growing BJ cells (at population doublings between 25 and 33), and hTERT-immortalized HSAEC (at population doublings 10-20) and primary human pancreatic epithelial cells (at population doublings 7-20) were used for experiments. Senescence was induced by transducing these cells with pBabe-WZLHygro- or pBabe-puro-Ha-rasV12 via retroviral infection as described previously [32,47]. Infected cells were cultured in medium containing 80 (BJ) or 50 (HSAECs and primary human pancreatic epithelial cells) µg/ml of hygromycin, or 1.5 (BJ) or 2 (HSAECs and primary human pancreatic epithelial cells) µg/mL of puromycin for 5-7 days to obtain OIS cells.
Senescence cells were analyzed by the rate of growth and the expression of the senescence-associated β galactosidase (SA-β-Gal) senescence marker as described previously [32,47]. Cell growth rate was calculated with the following formula, population doublings (PD)=\log(N2/N1)/\log2, where N1 is the number of cells seeded and N2 is the number of cells recovered. For SA-β-Gal staining, cells were washed twice with PBS (pH7.4) containing 2 mM MgCl₂ and fixed at room temperature for 15 min in PBS containing 0.5% glutaraldehyde. After washing twice with PBS (pH7.4) containing 2 mM MgCl₂, cells were incubated with staining solution [PBS pH6.0, 2 mM MgCl₂, 5 mM K₃Fe(CN)₆, 5 mM K₄Fe(CN)₆, 1 mg/ml X-gal] at 37 C for 4-6h. SA-β-Gal positive cells were counted in random field in each well. At least 200 cells were counted. Each experiment was performed in triplicate.

**Western blot analysis.**

Western blot was performed as described [32]. Up to 40 µg total protein was loaded. Primary antibodies were from Abcam (MZF1, CHD7, c-Fos, phospho-c-Fos-S374, ETS1, and phospho-ETS1-T38), Cell Signaling (phospho-p53-Ser15), Santa Cruz (Ras C-20, p16<sup>INK4A</sup>, p21<sup>WAF1</sup> C-19, p53 FL-393, phospho-c-Jun-S73, c-Jun, and β-actin). Secondary antibodies were from Cell Signaling. Signals were detected using enhanced chemiluminescence and captured by the FluorChem HD2 imaging system (Proteinsimple).

**RNA isolation and quantitative Real-time PCR (qRT-PCR).**

RNA was extracted from BJ cells using TRizol (ThermoFisher Scientific) and subjected to reverse transcription using iScript™ Reverse Transcription Supermix (Bio-Rad). qRT-PCR was performed in triplicates with gene-specific primers (Table S3) and SsoAdvanced™ SYBR Green
Supermix (Bio-Rad) in a Bio-Rad CFX96 REAL TIME SYSTEM following manufacturer’s protocols. β-actin was used as internal control.

**Luciferase reporter assay.**

The retroviral luciferase reporter for the p16\(^\text{INK4A}\) promoter has been reported previously [32]. The luciferase reporter for the p16\(^\text{INK4A}\) promoter with point mutations in the MZF1 binding sites (sites 1, 2 and 3) was generated by 2-step PCR [72] using the wild type p16\(^\text{INK4A}\) promoter as template and primers F1-Mlu1 and mut MZF1-R1 (used in the first step PCR), and mut MZF1-F2 and mut MZF1-R2-Bgl2 (used in the second step PCR) (Table S3). The luciferase reporter for the p16\(^\text{INK4A}\) promoter with deletions in the MZF1 binding sites (sites 1, 2 and 3) was also generated by 2-step PCR using the wild type p16\(^\text{INK4A}\) promoter as template and primers F1-Mlu1 and del MZF1-R1 (used in the first step PCR), and del MZF1-F2 and del MZF1-R2-Bgl2 (used in the second step PCR) (Table S3). The sequences of the point mutations in the p16\(^\text{INK4A}\) promoter are shown in Table S1. The entire consensus binding sequence shown in Table S1 was deleted in each of the deletion mutants. These reporters were cloned into pGL2 (Promega) and a self-inactivating retroviral luciferase reporter vector pBabe-PGK-Blast-SIN-Luc [32] for transient transfection in 293T cells and for creating BJ cells containing stably integrated luciferase reporters, respectively.

To construct the 2.5 kb and 1.5 kb luciferase reporters for the MZF1 promoter, a 2.5 kb MZF1 promoter fragment (nucleotide -2512 to +131 relative to the transcriptional start site, or TSS) and a 1.5 kb MZF1 promoter fragment (nucleotide -1500 to +130 relative to the TSS), respectively, were amplified by PCR using genomic DNA from 293T cells and primers listed in Table S3, and cloned into pGL2 (Promega) or pBabe-PGK-Blast-SIN-Luc [32]. The luciferase reporters containing the 1.5 kb MZF1 promoter with point mutations in the c-Jun or Ets1 binding sites
(MUT-c-Jun-A, MUT- c-Jun-B, MUT- c-Jun-A-B and MUT ETS1-A) were generated by site-directed mutagenesis using primers c-Jun-A-mutant, c-Jun-B-mutant, and ETS1-A mutatant (Table S3) and the QuikChange Site-Directed Mutagenesis Kit (Agilent), following manufacture’s protocols. The sequences of the point mutations in the MZF1 promoter are shown in Table S2. For transient transfection, 5×10⁵/well of 293T cells were seeded into 12-well plates and transfected the next day with 0.15 µg of a pGL2-p16INK4A promoter reporter and 1.5 µg of pcDNA3 or pcDNA3-flag-CHD7, or pLv, pLv-MZF1-long isoform or pLv-MZF1-short isoform using 4.0 µl of Lipofectamine-2000 (ThermoFisher) according to manufacturer’s instruction. Alternatively, 0.15 µg of wild type 1.5kb or 2.5kb MZF1 promoter reporter and 1.5 µg of pcDNA3 or pcDNA3-Ras were cotransfected into 293T cells.

To generate stable p16INK4A or MZF1 promoter reporter cell lines, BJ cells were transduced with pBabe-PGK-Blast-SIN-Luc-p16INK4A or pBabe-PGK-Blast-SIN-Luc-MZF1, respectively. If needed, these reporter cell lines were transduced with shRNAs for MZF1, ETS1 or c-Jun, TAM67 or appropriate vector controls, and with Ha-rasV12 or vector control. Cell lysates were collected using Passive Lysis Buffer (E1941, Promega) 48h after transient transfection or on day-9 or -10 after transduction of Ha-rasV12 or vector control. Luciferase activity was detected using Luciferase Assay System (E1500, Promega) on GloMax®-Multi+ Microplate Multimode Reader (Promega), and normalized to protein concentrations as determined using a BCA kit (ThermoFisher Scientific). Each experiments were performed in triplicates.

ChIP PCR

BJ cells were crosslinked with 1.5 mM Ethylene glycol bis (EGS) for 30 min at room temperature, followed by 1% formaldehyde for 10 min. The crosslinking was stopped by incubation with 0.125
M glycine for 5 min at room temperature. Cells were washed by phosphate-buffered saline (PBS), dislodged from plates in 3 ml hypotonic buffer (25 mM HEPES [pH 7.8], 1.5 mM MgCl2, 10 mM KCl, and 1 mM DTT), passed through a 25.5-gauge needle for 5 times, and spun down at 4,000 rpm for 5 min. Cells were re-suspend in 0.3 ml of sonication buffer (50mM Tris-HCl, 10mM EDTA, 1% SDS, and Roche Complete protease inhibitor cocktail) and sonicated 15 times, each for 20 second at 70% duty to get 200-500 bp DNA fragment.

10% of each sample was saved as total input. A 1 mg quantity of chromatin was incubated with 5 μg of an anti-MZF1 (Abcam), anti-CHD7 (Abcam), anti-c-Jun (Santa Cruz) or anti-Ets1 (Abcam) antibody or normal rabbit or mouse IgG (Santa Cruz) at 4°C overnight, followed by incubation with 60 μl (bead volume) of agarose beads (Santa Cruz) at 4°C for 4 h. The beads were washed sequentially once with each of the following buffers, buffer 1 (50 mM HEPES [pH 7.9], 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% NaDOC, and 0.1% SDS), buffer 2 (50 mM HEPES [pH 7.9], 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% NaDOC, and 0.1% SDS), buffer 3 (20 mM Tris-HCl [pH 8], 250 mM LiCl, 1 mM EDTA, 1% NP-40, and 1% NaDOC), and Tris-EDTA (TE; pH 8). After the final wash, protein-DNA-bead complexes were suspended in 300 μl TE buffer (10mM Tris-HCl, 1mM EDTA) and treated with 10 μg/ml of RNaseA at 37°C for 15 min. Crosslink was reversed by addition of NaCl to 300 mM and incubation at 65°C for overnight. After addition of 2.5 μl of 20% SDS and 36 μg of proteinase K, the reaction mixtures were incubated at 37°C for an additional 2 to 4 h. After pelleting of the beads, DNA in the supernatant was extracted with phenol-chloroform, precipitated with ethanol, and dissolved in double-distilled water. DNA precipitated by chromatin immunoprecipitation (ChIP) was quantified by real-time PCR using primers amplifying the targeted region. Primes for the amplification of the -110 to +50 region of p16INK4A promoter, and the -252 ~ -379 region and the -740 ~ -830 region of the MZF1
promoter were listed in Table S3. All ChIP-qPCR data were normalized to the input-qPCR data following Percent Input Method [32, 36].

**TCGA data acquisition and analysis**

Gene expression data and clinical information of each cancer type were downloaded from TCGA data portal (https://portal.gdc.cancer.gov/). Only cancers with more than 5 matched normal samples were retained in the downstream analysis, including BLCA, BRCA, COAD, ESCA, HNSC, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, PRAD, READ, STAD, THCA and UCEC (Fig. S6). A total of 20531 genes (protein coding and noncoding) were included in the TCGA IlluminaHiSeq RNASeqV2 data. We used level 3 gene expression data, which were derived from the reads per kilobase of transcript per million reads mapped (RPKM). The gene expression values were logarithmically transformed (base 2) prior to further analysis. MZF1 gene expression of tumor and matched normal samples were compared by paired t-tests with p-values provided. CDKN2A (encoding p16\(^{INK4A}\)) expression in PAAD were compared between tumor and unpaired normal by Students’ t-test. Correlations between MZF1 and CDKN2A/p16\(^{INK4A}\) gene expression in LUAD, LUSC and PAAD were determined by Pearson correlation analysis, with correlation coefficient (R) and significance level (p) provided.

For clinical information, the patient survival data, quantified as ‘days to death’ (overall survival time), ‘days to recurrence’ (disease-free survival time) or ‘days to last follow up’ (censored) were extracted and converted to numeric values for subsequent survival analysis. Samples without survival information were excluded. To study the clinical relevance of a particular gene, tumor samples were first ranked according to the expression of this gene. Then, the survival time of the top and bottom half of samples was extracted and compared by log-rank test. All the preprocessing,
statistical analysis and visualization procedures were implemented using R version 3.6.3. All plots were created with the ggplot2 package unless otherwise stated.

**GEO data analysis**

Gene expression data in cell lines were downloaded from Gene Expression Omnibus (GEO), including 8 RNA-seq data (GSE63577, GSE55949, GSE42509, GSE118494, GSE108278, GSE86546, GSE101750 and GSE74324) and 1 array data (GSE54402), which compare senescent cells and proliferating cells. GSE63577 and GSE42509 data were generated with human fibroblast BJ cells, while others were obtained with human fibroblast IMR-90 cells. Expression levels of CDKN2A/p16\(^{INK4A}\) and the transcription factors predicted to bind to the CHD7 binding site on the p16\(^{INK4A}\) promoter was compared between growing and senescence cells by one-tailed t-test, with p-value < 0.05 deemed significant.

Three microarray gene expression datasets in GEO, GSE101448, GSE62452 and GSE71729, were used to compare expression levels of MZF1 between tumor and adjacent normal tissues from PAAD patients. Logarithmically transformed (base 2) MZF1 expression levels in tumor and adjacent normal was compared by one-tailed t-test, with p-value < 0.05 deemed significant.

In situ Hi-C data of GSE118494, stored as normalized matrices of Hi-C intensity in both growing (n=2 replicates) and senescence (n=3) statuses, were also downloaded and analyzed. In the processed Hi-C data, each chromosome has been segmented into bins of 40-kb, and the Hi-C matrix for each chromosome represents the bin-wise interaction strength. We determined the chromosome where MZF1 or CDKN2A resides and picked up the corresponding Hi-C matrix, located the bins harboring them, focused on their flanking 30 bins and eventually extracted a 60 x 60 submatrix for subsequent analysis. Since MZF1 is located on the end of chromosome 19, only
31 bins were selected. We calculated the Hi-C difference between growing (G) and senescence (OIS) conditions by subtracting the average of 2 growing replicates from the average of 3 OIS replicates, yielding a difference heat map of bin-wise interactions. DNA interactions between the MZF1 or CDKN2A gene with its neighboring regions was extracted from the heat map and illustrated with bar plot for visualization. ChIP-seq peak profiles of CAP-H2, CAP-H2-FLAG, and SMC1 in regions surrounding the MZF1 and CDKN2A genes in both G and OIS conditions were also extracted and compared.

Correlations between MZF1 expression levels and overall survival of NSCLC patients were analyzed with GEO data using Kaplan-Meier Plotter, an online tool that incorporates GEO, EGA and TCGA database for assessing effect of genes on cancer survival ([https://kmplot.com/analysis](https://kmplot.com/analysis)) [48, 49].

**Analysis of single cell RNA-seq data from PAAD**

Gene-by-cell expression matrix was downloaded from published depository [42]. The original matrix includes 24005 genes and 57530 single cells collected from 24 primary PAAD tumors and 11 normal control pancreases, annotates 10 different cell types. Since ductal cells are the primary origin of PAAD tumor cells, we only focused on the two types of ductal cells, which were annotated as type 1 and type 2 ductal cells, representing normal and malignant ductal epithelial cells, respectively. To avoid confounding factors, 10 of the 24 PAAD patients with diabetes were excluded from the downstream analysis. Low quality cells (< 200 genes/cell, < 100 cells/gene and > 10% mitochondrial genes) were also excluded. After these preprocessing steps, we obtained a new gene-by-cell matrix with 18527 genes and 16084 cells, covering 14 PAAD tumors and 11 healthy controls for downstream analysis.
The R package Seurat [43] was applied to implement the downstream analysis and visualization. We performed nonlinear dimensional reduction analysis with the t-distributed stochastic neighbor embedding (t-SNE) technique. Cells were mapped to 2D t-SNE space and colored in different ways, including cell type (type 1 and type 2 ductal cells), sample source (PAAD vs. control), and gene expression levels of MZF1, CDKN2A, and RB1. Differences in expression of MZF1, CDKN2A, and RB1 between type 1 and type 2 ductal cells were determined by Student’s t-test. Type 2 ductal cells were divided into the RB1-normal group in which RB1 expression was higher than or equal to the median value of type 1 ductal cells and the RB1-low group in which RB1 expression was lower than the median of type 1 ductal cells. To test whether the cells with higher MZF1 expression were more likely to express higher level of CKDN2A in RB1-normal type 2 ductal cells, we designed a Fisher’s exact test as follows: first, we divided cells into two equally-sized parts by the MZF1 expression level with the median expression of these cells as threshold; then we counted the number of cells by CKDN2A expression level with the median expression as threshold. A similar approach was used to test the correlation between RB1 and CKDN2A expression levels in type 2 cells. We divided cells into two equally-sized parts by the CKDN2A expression level with the median expression of these cells as threshold; then we counted the number of cells by RB1 expression level with the median expression as threshold. With this procedure we obtained a 2-by-2 contingency table for the RB1-normal type 2 ductal cells and total type 2 cells, respectively, and a Fisher’s exact test was performed on each table separately.
**Statistical Analyses**

Each experiment was performed in triplicates and repeated at least 3 times. The exact sample size, the statistical methods and statistical significance resulted from the statistical analysis for each experiment was indicated in figure legends.
Supplemental Figures and Figure Legends

Figure S1

Fold Change (SEN/Grow)
**Fig. S1.** Expression of MZF1 and p16\(^{INK4A}\) are increased in senescent cells in GEO datasets.

Fold changes in expression levels of the genes encoding p16\(^{INK4A}\) and 12 transcription factors predicted to bind to the CHD7 binding site on the p16\(^{INK4A}\) promoter (Table S2) were determined in senescence cells versus growing cells in 9 GEO datasets (1 representing replicative senescence and the other 8 representing oncogene-induced senescence). The cell line analyzed in each dataset and number of biological repeats (senescent/growing cells) are indicated. Values are fold changes in senescent over growing cells. * p < 0.05; ** p < 0.01; *** p < 0.001 by Student’s t test.
Figure S2

(A) Population doubling vs. time (days) for different conditions.
(B) % SA-b-gal positive cells for different conditions.
(C) Population doubling vs. time (days) for different conditions.
(D) % SA-b-gal positive cells for different conditions.

**Note:** The text and figures are not translated or described in the image.
Fig. S2. p16$^{\text{INK4A}}$ and p53 are essential for oncogenic ras-induced senescence.

(A) Growth curves of BJ cells transduced with shRNA for p16$^{\text{INK4A}}$ (shp16-423) or a shRNA control (SC) and HaRasV12 (Ras) or vector (WH).

(B) Percentage of SA-β-Gal positive cells in populations in (A) on day-12 of the growth curves.

(C) Growth curves of BJ cells transduced with shRNA for p53 (shp53-854) or a shRNA control (SC) and HaRasV12 (Ras) or vector (WH).

(D) Percentage of SA-β-Gal positive cells in populations in (C) on day 12 of the growth curves.

(A-D) Values are means±SDs for triplicates. ** p < 0.01 vs SC-Ras by Student’s t test.
Figure S3

A. Western blot images showing the expression levels of p16, p-p53(S15), p53, MZF1-L, MZF1-S, and β-actin in different cell lines.

B. Bar graph showing the relative p16 expression levels in pLv, MZF1-L, and MZF1-S cell lines. The expression levels are normalized to β-actin.

C. Line graph showing the population doublings over time in pLv, MZF1-L, and MZF1-S cell lines. The graph includes data points for each cell line at different time points.

D. Bar graph showing the percentage of SA-β-gal positive cells in pLv, MZF1-L, and MZF1-S cell lines. The graph includes data points for each cell line.

E. Western blot images showing the expression levels of p16, MZF1-L, MZF1-S, Ras, p-p53(S15), p53, and β-actin in WH and Ras cell lines.

F. Bar graph showing the relative p16 expression levels in SC, sh-T1, sh-L, sh-S cell lines for WH and Ras conditions. The expression levels are normalized to β-actin.

G. Line graph showing the population doublings over time in SC, sh-T1, sh-L, sh-S cell lines for WH and Ras conditions. The graph includes data points for each condition.

H. Bar graph showing the percentage of SA-β-gal positive cells in SC, sh-T1, sh-L, sh-S cell lines for WH and Ras conditions. The graph includes data points for each condition.
**Fig. S3.** The long isoform of MZF1 mediates oncogenic ras-induced p16INK4A expression and senescence in primary human small airway epithelial cells (HSAECs).

(A) Western blot analysis of HSAEC cells transduced with the long (MZF1-L) or short (MZF1-S) isoform of MZF1 or vector (pLv).

(B) HSAEC cells transduced with the long (MZF1-L) or short (MZF1-S) isoform of MZF1 or vector (pLv) were analyzed for p16INK4A expression by qRT-PCR.

(C) Growth curves of HSAEC cells transduced with the long (MZF1-L) or short (MZF1-S) isoform of MZF1 or vector (pLv).

(D) Percentage of SA-β-Gal positive cells in populations in (C) on day-12 of the growth curves.

(E) Western blot analysis of HSAEC cells transduced with a shRNA control (SC) or shRNAs targeting both isoforms (sh-T1), the long isoform (sh-L) or the short (sh-S) of MZF1 and HaRasV12 (Ras) or vector (WH).

(F) HSAEC cells transduced with a shRNA control (SC) or shRNAs targeting both isoforms (sh-T1), the long isoform (sh-L) or the short (sh-S) of MZF1 and HaRasV12 (Ras) or vector (WH) were analyzed for p16INK4A expression by qRT-PCR.

(G) Growth curves of HSAEC cells transduced with a shRNA control (SC) or shRNAs targeting both isoforms (sh-T1), the long isoform (sh-L) or the short (sh-S) of MZF1 and HaRasV12 (Ras) or vector (WH) on day-12 of the growth curves.

(H) Percentage of SA-β-Gal positive cells in populations in (G) on day-12 of the growth curves.

(B-D, F-H) Values are means±SDs for triplicates. ns, not significant, p > 0.05; * p < 0.05; ** p < 0.01 vs pLv-con (B-D) or SC-Ras (F-H) by Student’s t test.
Figure S4

(A) Western Blot for p16, p-p53(S15), p53, MZF1-L, and MZF1-S. 

(B) Relative p16 expression. 

(C) Population doublings. 

(D) % SA-β-gal positive cells. 

(E) Western Blot for p16, MZF1-L, MZF1-S, p-p53(S15), p53, Ras, WH, and β-actin. 

(F) Relative p16 expression. 

(G) Population doublings. 

(H) % SA-β-gal positive cells.
Fig. S4. The long isoform of MZF1 mediates oncogenic *ras*-induced *p16*<sup>INK4A</sup> expression and senescence in primary human pancreatic epithelial cells.

(A) Western blot analysis of primary human pancreatic epithelial cells transduced with the long (MZF1-L) or short (MZF1-S) isoform of MZF1 or vector (pLv).

(B) Primary human pancreatic epithelial cells transduced with the long (MZF1-L) or short (MZF1-S) isoform of MZF1 or vector (pLv) were analyzed for *p16*<sup>INK4A</sup> expression by qRT-PCR.

(C) Growth curves of primary human pancreatic epithelial cells transduced with the long (MZF1-L) or short (MZF1-S) isoform of MZF1 or vector (pLv).

(D) Percentage of SA-β-Gal positive cells in populations in (C) on day-12 of the growth curves.

(E) Western blot analysis of primary human pancreatic epithelial cells transduced with a shRNA control (SC) or shRNAs targeting both isoforms (sh-T1), the long isoform (sh-L) or the short (sh-S) of MZF1 and HaRasV12 (Ras) or vector (WH).

(F) Primary human pancreatic epithelial cells transduced with a shRNA control (SC) or shRNAs targeting both isoforms (sh-T1), the long isoform (sh-L) or the short (sh-S) of MZF1 and HaRasV12 (Ras) or vector (WH) were analyzed for *p16*<sup>INK4A</sup> expression by qRT-PCR.

(G) Growth curves of primary human pancreatic epithelial cells transduced with a shRNA control (SC) or shRNAs targeting both isoforms (sh-T1), the long isoform (sh-L) or the short (sh-S) of MZF1 and HaRasV12 (Ras) or vector (WH).

(H) Percentage of SA-β-Gal positive cells in populations in (G) on day-12 of the growth curves.

(B-D, F-H) Values are means±SDs for triplicates. ns, not significant, p > 0.05; * p < 0.05; ** p < 0.01, vs pLv-con (B-D) or SC-Ras (F-H) by Student’s t test.
Figure S5

A. Genome bins for CDKN2A

B. Bar chart showing log2FC for genomic bins

C. Genome bins for MZF1

D. Bar chart showing log2FC for genomic bins

E. Genomic bins for COHESIN II

F. Genomic bins for COHESIN I

Log2FC values for various genomic bins are shown, with differences in expression levels between conditions.
**Fig. S5.** Reduction in contacts between the MZF1 and CDKN2A/p16INK4A genes and their neighboring chromosomal regions and enrichment of condensing II and cohesin (SMC1) in the vicinity of the MZF1 and CDKN2A/p16INK4A genes in cells undergoing oncogene-induced senescence.

(A, C) Heat maps showing differential DNA contact strength surrounding the CDKN2A (A) and MZF1 (C) genes between oncogene-induced senescence (OIS) and growing (G) states. Contact map was derived from Hi-C data measured in OIS (n=3) and growing (n=2) states in the GEO dataset GSE118494. Chromosome 9 (A) and chromosome 19 (C), where the CDKN2A and MZF1 genes are located, respectively, were divided into 40-kb bins. The bin-wise interaction strength are determined and compared between OIS and G cells by Student’s t-tests. The mean difference was calculated to generate the heat map.

(B, D) Bar plot highlighting the difference in contacts between the CDKN2A (B) and MZF1 (D) genes and their neighboring regions between the OIS and growing states.

(E-F) Distributions of endogenous (CAP-H2) and ectopically expressed (CAP-H2-FLAG) condensin II subunit CAP-H2 and SMC1 (a subunit of cohesion) binding in regions surrounding the CDKN2A (E) and MZF1 (F) genes in OIS (S) and growing (G) cells. G represents values from growing IMR90 cells; S1 and S2 represent values from two replicates of oncogene induced senescent cells for endogenous and FLAG-tagged CAP-H2; S represents values from oncogene induced senescent cells for SMC1. Values are normalized ChIP-seq signal intensity.
**Figure S6**

(A) Heatmap showing Log2(MZF1) expression across various cancer types.

(B) Scatter plots showing the correlation between Log2(CDKN2A) and Log2(MZF1) for LUAD, LUSC, and PAAD.

(C) Box plots comparing Log2(CDKN2A) expression in Normal and Tumor samples for LUAD and LUSC.

(D) Box plots comparing Log2(CDKN2A) expression in Normal and Tumor samples for PAAD.

*P<0.05, **P<0.01, ***P<0.001
Fig. S6. MZF1 expression is reduced in multiple cancer types in the TCGA database.

(A) Comparison of MZF1 expression levels between tumor tissues and paired normal tissues in 16 cancer types based on RNA sequencing data in the TCGA database. The cancer types with significant downregulation of MZF1 expression are highlighted in red.

(B) Pearson correlation analysis of MZF1 and p16INK4A expression levels in NSCLC (LUAD and LUSC) and PAAD based on the data from the TCGA database. Correlation coefficient (R) and p value from Pearson correlation analysis are indicated.

(C-D) p16INK4A expression is not downregulated in tumor tissues as compared to normal tissues in NSCLC (LUAD and LUSC) (C) and PAAD (D) based on analysis of the data from the TCGA database.

(A, C-D) Box plots represent gene expression levels with relative intensity (log2) of RNA sequencing data. Horizontal lines represent median values. The bottom and top of the box are the first and third quartiles, respectively, with whiskers extending to 1.5 Interquartile range of the lower and upper quartile respectively. The blue hollow circles represent outliers. ns, not significant, p > 0.05; * p < 0.05; ** p < 0.01; *** p < 0.001 vs normal tissues by Student’s t test.
Supplemental Tables

**Table S1.** Locations of consensus binding sites for the indicated transcription factors at the CHD7 binding site (-112→+20) of the p16\(^{INK4A}\) promoter as predicted by the JASPAR database. Mutant sequences used in the study are shown in red.

| Matrix   | Name   | Start | End   | Predicted sequence | Mutant sequence |
|----------|--------|-------|-------|--------------------|-----------------|
| MA0056.1 | MZF1   | -43   | -38   | TGGGGA             | GTTTTC          |
| MA0056.1 | MZF1   | +28   | +33   | GGGGGA             | TTTTTT          |
| MA0056.1 | MZF1   | -94   | -89   | GGGGGA             | TTTTTT          |
| MA0161.7 | NFIC   | -103  | -98   | CTGGCA             |                 |
| MA0499.1 | Myod1  | -52   | -40   | CCCACCTGCCCCCC     |                 |
| MA0597.1 | THAP1  | -105  | -97   | CTGCCAGCA          |                 |
| MA0522.2 | TCF3   | -49   | -40   | CCCACCTGCC         |                 |
| MA0830.1 | TCF4   | -49   | -40   | CCCACCTGCC         |                 |
| MA0103.2 | ZEB1   | -47   | -39   | CCCACCTG           |                 |
| MA0599.1 | KLF5   | -42   | -33   | GCTCCTCCCC         |                 |
| MA0670.1 | NFIA   | -19   | -10   | CTTGCCAACG         |                 |
| MA0670.1 | NFIX   | -19   | -11   | CTTGCCAAC          |                 |
| MA0161.1 | NFIC   | -17   | -12   | TTGGCA             |                 |
| MA0076.2 | ELK4   | +7    | +17   | CTGCTTCCG          |                 |
| MA0765.1 | ETV5   | +8    | +17   | GCGGAAGCA          |                 |
Table S2. Locations of the c-Jun and ETS1 binding sites in the -1500 ~ +130 region of the MZF1 promoter as predicted by the JASPAR database. Mutant sequences used in the study are shown in red.

| Matrix  | Name      | Start | End | Predicted      | Binding   | Mutant sequence          |
|---------|-----------|-------|-----|----------------|-----------|--------------------------|
| MA0099. | FOS::JUN  | -740  | -731| ATGTGTCACT     | c-Jun -A  | GACCCACACTGTTA           |
| MA1141. | FOS::JUN  | -742  | -730| GTATGTGTCAC    |           |                          |
| MA0462. | BATF::JU  | -744  | -734| GGTATGTGTCAG   |           |                          |
| MA0462. | BATF::JU  | -333  | -323| AACATGACTCT    | c-Jun -B  | AGTGCTGAGAG              |
| MA0098. | ETS1      | -886  | -877| GCCGGAAAAT     | ETS1-A    | TTTTTCCCCC               |
| MA0098. | ETS1      | -256  | -247| ACAGGAAAGG     | ETS1-B    |                          |
Table S3. Primer sequences used for qPCR, ChIP-PCR and mutagenesis.

| Genes                  | Primers | Sequence (5’ to 3’)                  |
|------------------------|---------|-------------------------------------|
| **Real time PCR primers** |         |                                     |
| Total MZF1             | Forward | TGAAACTGAGCCTCCAACCTC               |
|                        | Reverse | CTGTAACCTCTGACTCCTCTTTTC            |
| MZF1 long isoform      | Forward | TTTCCGTGCAGCGAGGTG                 |
|                        | Reverse | CTCGACCGAGCCAAAGG                   |
| MZF1 short isoform     | Forward | TTCTCCCCAGGGGCGGAGCGG              |
|                        | Reverse | TCAGCAGGGTTGCTGCGTTG               |
| p16                    | Forward | GCACATTCATGTGGGCAATT               |
|                        | Reverse | GACTCAAGAGAAGCCAGTAACC             |
| p21                    | Forward | CGGAACAAGGAGTCAGACATT              |
|                        | Reverse | AGTGCCAGGGAGAAGACAACACT            |
| ETS1                   | Forward | TCCCCCTATACCTCGGATTAC              |
|                        | Reverse | ATGGGATGGAGCGCTGTGTA               |
| c-Jun                  | Forward | GCTGGAGCGCCTGATAAT                 |
|                        | Reverse | CTGCTCATCTGTCACGTTC                |
| **ChIP-PCR primers**   |         |                                     |
| -119→+50 of p16 promoter | Forward | CGCAGCTCAAACACGCCTT               |
|                        | Reverse | GCACCTAGCGAATGTGGCA                |
| -242→-422 of MZF1      | Forward | GGCCACTATTTTTCTTTTCTCTGT           |
|                        | Reverse | TATTCTTCCTTTCTCCTGTCTCTCT          |
| -729→-900 of MZF1      | Forward | GTCCCGGGTGACTGCGGAAAAATGT          |
|                        | Reverse | AATAGTGACACATACCCAGAAAA            |
| **Mutagenesis primers** |         |                                     |
| F1-p16                 |         | CGACGCGTAAACAGGTATTAGCCTTTAGGATGT |
| mut MZF1-R1            |         | ACGCTTTTGCTGCGAGGCTTTTTCCGCGGGCTGGGAG CAGGGAGCGCGGAGGCCGAGGGGCGTGGGGGCA |
|     | DNA Sequence                                      |
|-----|--------------------------------------------------|
| mut MZF1-F2 | AGGGCGGTGTCGGGCGCGGCGCGCTGGGGGGCAGGGTTTTCGGAGCCCAGT |
|     | CCTCCTTC                                      |
| mut MZF1-R2-Bgl2 | CTTCCGGCTGGTGCCCCCTTTTTCTGACACCAACCTGGG |
|     | GCAGACTTCAGGGGTTGCCACATTCGCTAAGTGCTCGGA       |
|     | GTTAATAGCACCT                                    |
| del MZF1-R1  | ACGCCTTTTGTGGCAGGCCGCGCGGGCTGGGGAGGAGCAGGG  |
|     | AGGCCGGAGGCGGTGTTGGGGGGCA                           |
| del MZF1-F2  | AGGGCGGTGTCGGGGGCGAGGCGAGCAGCCTCCTCTTCTTC        |
| del MZF1-R2-Bgl2 | TGCCTCGGGCGCGCTGGGGAGAGGAGCAGCGG |
| ETS1-mut     | GCGTCCCCGGTACTTTTTTCCCCCTAGTTTTGGCAACG          |
| c-Jun-A-mut  | AAAGTGCTGAGAGCAGTGAGCATG                         |
| c-Jun-B-mut  | CCTTCCCTCCTTCTTTCCCCCCCTTGTCCTATTTT           |