Supplementary Materials and Methods

Histology and Immunohistochemistry

Liver tissue sections were deparaffinized in xylene and rehydrated through an ethanol gradient. For histological analysis, sections were stained with hemotoxylin and eosin (H&E) using standard techniques. For analysis of Ki67 and phospho-γH2AX expression, liver sections were initially boiled in Antigen Retrieval Solution (Dako North America, Inc., Real Carpinteria, CA) using a microwave (5 min at 100% power followed by 20 min at 30% power). Sections were then cooled to room temperature, and endogenous peroxidase activity was quenched by treatment in 0.3% H_2O_2 in methanol (Dako North America, Inc.). Samples were blocked for 30 min in 5% goat serum in PBS. Primary Ki67 (rat monoclonal, Dako North America, Inc.) or phospho-γH2AX (mouse monoclonal, Millipore, Bedford, MA) antibody was diluted 1:250 in blocking buffer and incubated for 1 hour at room temperature. Following washing with PBS, biotinylated secondary antibody and streptavidin-peroxidase conjugate were applied according to Vectastain Elite avidin-biotin complex kit instructions (Vector Laboratories, Burlingame, CA). Staining was developed in 3,3' diaminobenzidine (DAB) solution (Vector Laboratories) for 1-2 min and quenched with H_2O. Slides were counterstained with hematoxylin, dehydrated through graded ethanols and xylene, and mounted using permount (Surgipath Medical Industries, Inc., Richmond, IL). Phospho-histone H3 Serine10 (rabbit polyclonal, Upstate Biotechnology, Billerica, MA) was detected as described above, except using rhodamine-conjugated donkey anti-rabbit secondary antibody (Invitrogen Corporation, Carlsbad, CA) and counterstained with 4,6-diamidino-2-phenylindole (DAPI) (Sigma-Aldrich). For analysis of BrdU incorporation, a Zymed BrdU Staining kit (Invitrogen) was utilized according to manufacturer's instructions.

Sample Numbers for Mouse Liver Tissue Microarray Analyses

	Rb+p53+ Untreated	Rb-p53- Untreated	Rb+p53+ DEN-Treated	Rb+p53- DEN-Treated	Rb-p53+ DEN-Treated	Rb-p53- DEN-Treated
# of Samples	3	3	5	5	4	5

RNA Preparation for Microarray Analyses

RNA was harvested from frozen liver tissue using RNeasy Blood and Tissue kit (Qiagen, Inc., Valencia, CA) according to the manufacturer's suggested protocols. Purified RNA was quantified on a Nanodrop ND-100 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE). RNA quality was assessed by analysis on an Agilent 2100 bioanalyzer (Agilent Technologies, Palo Alto, CA). RNA amplification and labeling was performed by the WT-Ovation Pico RNA amplification system (NuGen Technologies, Inc., San Carlos, CA). Briefly, 50 ng of total RNA was reverse transcribed using a chimeric cDNA/mRNA primer, and a second complementary cDNA strand was synthesized. Purified cDNA was then amplified with ribo-SPIA enzyme and SPIA DNA/RNA primers (NuGEN Technologies, Inc.). Amplified DNA was purified using QIAquick PCR Purification Kit (Qiagen, Inc.). Sense transcript cDNA (ST-cDNA) was generated from 3 µg amplified cDNA using WT-Ovation Exon module (NuGen Technologies, Inc.). Purified ST-cDNA was assed for yield using the Nanodrop Spectrophotometer (NanoDrop Technologies, Inc.). 2.5 µg ST-cDNAs were fragmented and chemically labeled with biotin to generate biotinylated ST-cDNA using FL-Ovation cDNA biotin module V2 (NuGen Technologies, Inc.). Each Affymetrix GeneChip Mouse Gene 1.0 ST array (Affymetrix, Inc., Santa Clara, CA) was hybridized with fragmented and biotin-labeled target (2.5 µg) in 110 µl of hybridization cocktail. Target denaturation was performed a 99^{o}c for 2 min. and then 45^{o}c for 5 min., followed by hybridization for 18 hrs. Arrays were washed and stained using Genechip Fluidic Station 450, and hybridization signals were amplified using antibody amplification with goat IgG (Sigma-Aldrich) and anti-streptavidin biotinylated antibody (Vector Laboratories, Burlingame, CA, USA). Chips were scanned on an Affymetrix Gene Chip Scanner 3000, using Command Console Software.

Microarray Summarization and Normalization

The Affymetrix Mouse Gene 1.0 ST arrays were processed using the Robust Multichip Average (RMA) algorithm (Irizarry 2003), including background correction, and quantile normalization on

the core annotation subset in Affymetrix Expression Console version 1.1 (Affymetrix, Inc.). The Affymetrix probeset annotation version used was release 31, dated August 2010. Subsequent data processing and analyses were performed using MATLAB software (The Mathworks, Inc., Natick, MA).

Differential Expression Analysis

Differential expression was determined for pair-wise comparisons among the different mouse treatments, using Significance Analysis of Microarrays (SAM) (53). Differentially expressed genes were selected based on an estimated 10% FDR, and subsequently limited by an absolute fold change of 1.5 or more.

Comparative Expression Analysis

The Gene Expression Omnibus (GEO) [http://www.ncbi.nlm.nih.gov/geo/] microarray database was used to search for retrospective human HCC microarray datasets. Two Affymetrix HGU133plus2 platform datasets with accession numbers GSE6764 and GSE9843, and one custom platform dataset with accession number GSE4108 were selected. Raw data (.CEL) files were obtained for the datasets GSE6764 and GSE9843, and RMA (Irizarry 2003) preprocessing was performed on each individual dataset using *RMAExpress* software (http://rmaexpress.bmbolstad.com/) with updated library files (Entrez version 12, dated July 30, 2009) (Dai, 2005) obtained from the University of Michigan Microarray Lab website: (http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genomic_curated_CDF.asp). Raw data files for the GSE4108 dataset were not available, so this data was downloaded from GEO as an expression matrix centered on the dataset's mean expression value. Missing values in this dataset were replaced by the gene (row) median. For each human microarray dataset, the lowest 50^{th} percentile variance genes were filtered out prior to comparison with the mouse data. Genes differentially expressed in the mouse experiments were mapped to each independent human dataset based on gene symbol and used to define subset of human genes associated with Rb and p53 deficiency. Within these Rb and p53 deficiency subsets, psuedo-

expression signatures were defined by median-centering human gene expression profiles, multiplying negatively regulated genes by -1, and taking the average over all genes. These psuedo-expression signatures were used to rank human liver samples from low-to-high expression and observe phenotypic trends as a function of expression gradient. A Kolmogorov-Smirnov Test was used to test the distribution of HCC cases, ranked by the Rb/p53-deficient gene expression signature, against the null hypothesis that they are uniformly distributed among non-cancerous liver samples.

Survival data was obtained for human patient samples in the GSE4108 dataset and used to test the association of the mouse p53/Rb loss profile with time to disease recurrence or death. The average expression profile of genes up-regulated (i.e. genes altered in a synonymous fashion) in the RB/p53-deficient mouse liver samples was used to rank human samples from high to low expression, with samples above the median assigned to a high expression group, and samples below the median assigned to a low expression group. Kaplan Meier analysis was used to observe survival trends in each group, and the log rank test was used to evaluate the difference in survival trends between the high-expression (Red) and low expression (Blue) group.

Immunoblot Antibodies

Proteins were detected using the following antibodies: from Santa Cruz Biotechnology (Santa Cruz, CA)- Lamin B (M-20), MCM7 (141.2), PCNA (PC10), p130 (C-20), C/EBPα (14AA), Cyclin E (C-19), CDK1 (17) and CDK2 (M-2). Monoclonal Cyclin B1 antibody was a generous gift from Dr. Philipp Kaldis.

Comparative genomic hybridization (CGH) Analyses

R and Bioconductor including snapCGH package (Smith, Marioni, McKinney, Hardcastle and Thorne; http://www.bioconductor.org) were used for CGH data analysis. Log-ratios were median-normalized. Averaging duplicated probes and missing value imputation were done with aCGH (Fridlyand and Dimitrov, www.bioconductor.org) and snapCGH. Segmentation or

breakpoint detection was performed to translate these smoothed ratios into gain/loss/no-change states using DNAcopy method in snapCGH. Probability-based gain/loss calling was done using CGHcall package. Permutation method was used to generate minimal common regions (MCRs) among the treated samples, which computed for each probe a *P*-value that tests the significance of the alteration of that probe across the set of samples. This *P*-value is based on a permutation test that assumes that the alterations found are randomly located in the genome. Then, the consecutive probes with *P*-values lower than a cut-off were merged in a common region.

References

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Research 2003; 31(4):e15.

Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Benney WE, Myers RM, Speed TP, Akil H, Watson SJ, Meng F. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Research 2005; 10;33(20):e175.