# Science

## AAAS

# Supplementary Materials for

## A dominant-negative effect drives selection of *TP53* missense mutations in myeloid malignancies

Steffen Boettcher, Peter G. Miller, Rohan Sharma, Marie McConkey, Matthew Leventhal, Andrei V. Krivtsov, Andrew O. Giacomelli, Waihay Wong, Jesi Kim, Sherry Chao, Kari J. Kurppa, Xiaoping Yang, Kirsten Milenkowic, Federica Piccioni, David E. Root, Frank G. Rücker, Yael Flamand, Donna Neuberg, R. Coleman Lindsley, Pasi A. Jänne, William C. Hahn, Tyler Jacks, Hartmut Döhner, Scott A. Armstrong, Benjamin L. Ebert

Correspondence to: benjamin_ebert@dfci.harvard.edu

**This PDF file includes:**

Materials and Methods
Figs. S1 to S13
Table S1 and S2
References

**Other Supplementary Materials for this manuscript include the following:**

N/A

## Materials and Methods

### Cell culture

K562 and MOLM13 cell lines were obtained from Broad Institute cell line repository and were cultured in RPMI, 100 U/ml penicillin and 100 mg/ml streptomycin with 10 or 20% FBS, respectively, in a 37°C incubator at 5% carbon dioxide.

### CRISPR/Cas9-mediated homology-directed repair (HDR) or gene knock-out (KO)

The *TP53* gene in K562 and MOLM13 cells was edited by CRISPR-HDR or CRISPR-KO as previously described (*36*) with some modifications. Recombinant Cas9 protein, synthetic locus-spedific CRISPR RNAs (crRNA), and transactivating crRNAs (tracrRNA) were all purchased from Integrated DNA Technologies (IDT). Equimolar amounts (120pmol) of crRNAs and tracrRNAs were mixed in a Cas9 buffer (HEPES 20 mM, 150 mM KCl, 1 mM $MgCl_2$, 10% glycerol, 1 mM TCEP) to obtain a total volume of 5 μl. To generate crRNA:tracrRNA duplexes, the mixtures were heated to 98°C for 5 min and allowed to cool down to ambient temperature. Recombinant Cas9-3NLS (100 pmol) was diluted in Cas9 working buffer to obtain a final volume of 5 μl. Both crRNA:tracrRNA duplexes and diluted Cas9-3NLS were slowly mixed and incubated for 20 min at room temperature to allow formation of ribonucleoprotein (RNP) complexes. $2x10^5$ cells were harvested, spun down, and resuspended in 10 μl of SF nucleofection solution (Lonza). RNP and cell solutions were combined in a 20 μl nucleocuvette (Lonza). For CRISPR-HDR, single-strand oligonucleotides (ssODN) HDR templates were designed according to the rules established by Richardson et al. (*36*), purchased from IDT, and 100 pmol were added per electroporation reaction. Electroporation was carried out using a 4D-Nucleofector™ (Lonza) using cell line-specific settings according to the manufacturer's recommendations. Electroporated cells were transferred to T25 cell culture flasks and allowed to recover for 5-7 days. In some experiments to deplete non-edited, i.e. *TP53* wild-type cells, nutlin-3a at a concentration of 5-10 μM was added. Genome-editing was assessed by ultra-deep amplicon next-generation sequencing performed at the Massachusetts General Hospital Center for Computational and Integrative Biology DNA core facility (MGH CCIB DNA core, Cambridge, MA). Subcloning was performed by limiting dilution. Genome-edited clones were screened by Sanger sequencing (Eton Bioscience,

Boston, MA) and genotypes were confirmed by ultra-deep amplicon next-generation sequencing (MGH CCIB DNA core, Cambridge, MA).

**Growth curve determination**

$2x10^5$ K562 or $5x10^5$ MOLM13 cells were plated in one well of 12-well plate on day 0. Cells were counted every 2 days using a Vi-CELL™ automated cell viability analyzer and counter (Beckman Coulter) according to the manufacturer's recommendations using default parameter settings. All recovered cells were re-plated in the appropriate tissue culture flask at a density of $2x10^5$ cells/ml for K562 or $5x10^5$ cells/ml for MOLM13 cells.

**Immunoblots**

Cells were lyzed in Pierce™ IP lysis buffer (ThermoFisher, #87787) freshly supplemented with Halt™ Protease and Phosphatase Inhibitor Cocktail (ThermoFisher, #78440). Protein concentration was quantified using Pierce™ BCA Protein Assay Kit (ThermoFisher, #23225). Equal protein amounts were loaded and run on a polyacrylamide gel, transferred to polyvinylidene difluoride (PVDF) membranes and blotted for p53 (mouse monoclonal antibody DO-1, Cell Signaling Technology, #18032), p21 (rabbit monoclonal antibody 12D1, Cell Signaling Technology, #2947), and actin (mouse monoclonal antibody, abcam, #8226). Goat anti-mouse-HRP (Genesee Scientific, #20-304), goat anti-rabbit-HRP (Genesee Scientific, #20-303), and SuperSignal™ West Dura Extended Duration Substrate (ThermoFisher, # 34075) were used to visualize the blot.

**Quantitative reverse-transcription PCR**

Total RNA isolation was carried out using RNeasy kit (QIAGEN, #74106). Reverse transcription was done using SuperScript™ IV VILO™ Master Mix with ezDNase™ Enzyme (ThermoFisher, #11766500). qPCR was performed using TaqMan® Gene Expression Master Mix (ThermoFisher, #4369016) and TaqMan® Gene Expression Assays (*CDKN1A*, Hs00355782_m1; *MDM2*, Hs00540450_s1; *ACTB*, Hs99999903_m1) on a 7900HT Fast-Real-Time PCR System (Applied Biosystems).

**Apoptosis assays and cell cycle analysis**

Cells were treated with DMSO or 100 nM daunorubicin and were collected at different periods of time. Annexin V Allophycocyanin (APC) staining (Biolegend, #640920) in combination with DAPI (4',6-Diamidino-2-Phenylindole Dilactate; Biolegend, # 422801) was used to determine daunorubicin-induced apoptosis by flow cytometry using a FACSCanto II (BD Biosciences).

For cell cycle analysis, cells were treated with DMSO or 100 nM daunorubicin for 24 hours. Cells were collected and stained with cell-permeable DNA-binding dye CytoPhase™ violet (Biolegend, #425701), incubated for 90 minutes at 37°C, and analyzed by flow cytometry using a FACSCanto II (BD Biosciences).

**Drug sensitivity assays**

$2x10^4$ cells were seeded per well in a 96-well flat-bottom plate. The following compounds were used: daunorubicin (SelleckChem, #S3035), cytarabine (SelleckChem, # S1648), etoposide (SelleckChem, # S1225), mitoxantrone (SelleckChem, # S2485), nutlin-3a (SelleckChem, # S8059). Drugs were purchased in powder form, dissolved in DMSO, diluted in the appropriate cell culture medium and added in limiting dilutions to the wells using a HP® D300e Digital Dispenser (HP Inc.). After 72 hours of drug exposure, cell viability was assesses using CellTiter-Glo® luminescent assay (Promega, #G7572) using a FilterMax F5 (Molecular Devices) plate reader. Cell viabilities were calculated relative to DMSO controls.

**In vitro competition assays**

Cells were labeled with fluorescent proteins GFP or RFP657 by lentiviral overexpression and sorted to obtain pure populations. RFP657[+] cells of various *TP53* genotypes were mixed in a 1:9 ratio with GFP[+] control cells (either *TP53* wild-type or null) and cultured in the presence of DMSO or different drug concentrations in 96-well flat-bottom plates for up to 10 days. Cells were analyzed using a FACSCanto II (BD Biosciences) every 48 hours. At this timepoint cell cultures were replenished with fresh medium and the appropriate amount of drug to maintain constant drug concentrations over time.

**ChIP-seq**

$20x10^6$ cells were crosslinked using 1% methanol-free formaldehyde (ThermoFisher, #28906) in PBS for 7 min at room temperature. Remaining formaldehyde was quenched with 100 mM Tris-HCl pH8.0 and 25 mM Glycine. Cytoplasm was stripped using 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM EDTA, 1% SDS for 10 min at room temperature followed by precipitation of nuclei by centrifugation at 10,000g. The nuclei were then resuspended in 66 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM EDTA, 1.7% Triton X-100, 0.5% SDS and sheared using an E220 sonicator (Covaris). Chromatin fragmentation was quality-controlled using ScreenTape D5000 (Agilent). Immunoprecipitation with anti-p53 antibodies (DO-1, abcam, #1101) and Protein G Dynabeads™ (ThermoFisher, #10003D) was performed overnight at 4°C. Beads were washed and immunoprecipitated DNA was reverse-crosslinked in 100 mM NaHCO3, 100 mM NaCl, 1% SDS, and quantified by ScreenTape HS D1000 (Agilent) and Qubit (ThermoFisher). Illumina-compatible sequencing libraries were prepared using SMARTer® ThruPLEX® DNA-Seq Kit (Takara, #R400675) and sequenced using a NextSeq™ 550 Sequencing System (Illumina) to obtain paired-end reads.

**RNA-seq**

$0.5x10^6$ cells were collected, washed with ice-cold PBS followed by total RNA extraction using RNeasy kit (Qiagen, #74106). The quality of extracted RNA was determined using RNA ScreenTape (Agilent). $RIN^e$ values were always >8.0. RNA was quantified by Qubit (ThermoFisher). 500 ng of RNA was used to prepare Illumina-compatible 3'-end sequencing libraries using QuantSeq 3' mRNA-Seq Library Prep Kit FWD for Illumina (Lexogen) and then sequenced using a NextSeq™ 550 Sequencing System (Illumina®) to obtain single-end reads.

**ChIP-seq and RNA-seq analyses**

Raw Illumina sequencer output was converted to FASTQ format using bcl2fastq (v2.17). In order to trim read qualities, we ran trimmomatic (v0.36; minimum trimmed length 34bp). We then used FastQC (*2*) on the trimmed files to assess the quality of the sequencing data, which was then aligned to the human genome (hg19) using BWA-mem (*3*). Picard MarkDuplicates (*4*) was used to identify and remove duplicate reads in the binary alignment files. Final .bam files were indexed using SAMtools (v1.2). For ChIP-seq, we called, annotated and identified peaks in *TP53* wild-

type as well as *TP53* missense mutants with Homer v4.10 (*37*) using the *TP53*[null] sample as the control, excluding regions found in the excludable ENCODE mappability tracks from Crawford, Sidow and Batzoglou (*38*). We limited our analysis to peaks over transcriptional start sites (TSS), in intron 1 or those peaks found within 10kb upstream of the TSS. We filtered out peaks with fewer than 30 tag counts after normalizing to a 10kb region to control for non-specific binding of p53 antibodies. We noticed contamination with plasmid DNA (cDNA for human *KMT5A*) that was removed from analysis.

For RNA-seq, sequencing data were aligned with STAR (*39*) and the transcript counts per gene were calculated using RSEM (*40*). We used EdgeR (*41*) to analyze which genes were differentially expressed between mutant *TP53* (null and missense) and wild-type *TP53*. If a gene had a greater than a two-fold change in transcripts per million (TPM) relative to the wild-type and a Benjamini-Hochberg q-value less than 0.05, we would consider it as differentially expressed. We only included genes that had at least 3 TPM in 3 or more mutants.

## Gene ontology analysis

GO terms (*42*, *43*) associated with the pooled top or bottom 30 most differentially expressed genes relative to *TP53* wild-type were determined using PANTHER (*44*). To account for multiple hypothesis testing, we used an FDR set at 1%. Only GO terms with more than 3 associated transcripts and a fold-enrichment over the genomic background of more than tenfold were included in our analysis.

## *TP53* MITE-seq screen

The general principle of 'Mutagenesis by Integrated TilEs' followed by deep sequencing (MITE-seq) screens has been described elsewhere (*45-47*). The generation of the *TP53* MITE library containing 7'893 cDNAs encoding for all possible single amino acid variants of full-length p53, its cloning into the lentiviral backbone pMT_BRD025 which allows for ORF expression under control of the human EF1α promoter, and lentivirus production have recently been described (*24*). To generate K562-*TP53*[wild-type]-*CDKN1A*-GFP reporter cells, the endogenous *CDKN1A* gene in K562-*TP53*[wild-type] cells was tagged at the carboxy terminus with GFP using the eFlut (endogenous Fluorescent tagging) methodology as previously described (*48*).

K562-$TP53^{\text{wild-type}}$-$CDKN1A$-GFP reporter cells were infected with the $TP53$ MITE library at a low multiplicity of infection to ensure integration of only one $TP53$ cDNA variant per cell. Infections were done in two independent replicates on separate days ($40\times10^6$ cells per replicate). Cells were selected with 2 µg/ml puromycin (ThermoFisher, #A1113802) for two days and allowed to recover and expand for another two days. Infected cells were then split into equal fractions and treated with 0.1% v/v DMSO or 5 µM nutlin-3a for 24 hours after which GFP$^{\text{hi}}$ and GFP$^{\text{lo}}$ cells from either treatment arm were collected by flow-sorting and frozen as cell pellet at -80°C ($2\times10^6$ - $10\times10^6$ cells per sample). Cells were subjected to genomic DNA (gDNA) isolation using QIAamp DNA Blood Midi kits (QIAGEN, #51183).

For ORF purification from gDNA, 24 PCR reactions per gDNA sample in a volume of 50 µl containing ~1.5 µg of gDNA were performed. Q5® High-Fidelity DNA polymerase (New England Biolab, # M0491L) was used. All 24 PCR reactions for each gDNA sample were pooled, concentrated with a QIAquick PCR Purification Kit (QIAGEN, #28104), loaded onto a 1% agarose gel, and separated by gel electrophoresis. Bands of the expected size were excised, and DNA was purified first using a QIAquick Gel Extraction Kit (QIAGEN, #28704) followed by AMPure XP beads (Beckman Coulter, #A63880).

Sequencing samples were prepared according to the Illumina® Nextera XT protocol. For each purified ORF fragment sample, we set up six Nextera reactions, each with 1 ng of purified ORF DNA. The Nextera reactions of each were indexed with unique i7/i5 index pairs. After the limited-cycle PCR step, the 6 Nextera reactions were combined and purified using AMPure XP beads (Beckman Coulter, #A63880. All samples were then quantified, pooled by equal weight, and detected using the NextSeq™ 550 Sequencing System (Illumina®) using two reads, each 150 bases in length, and 2 index reads of 8 bases long.

NextSeq™ 550 Sequencing data were processed with the ORFcall software developed by the Broad Institute (47). The ORFcall software aligns read pairs to the $TP53$ reference sequence. The codons corresponding to each variant were then tallied. At each codon position, there were counts for all 64 possible codons, including programmed variants and variants not programmed but appearing due to library synthesis errors and PCR/NGS errors. We performed quality control by looking at, first, the abundance of both programmed variant codons and un-programmed variant codons, and second, the abundance of variant codons that differ from the wild-type codons by one, or more than one nucleotide. PCR steps involved in this workflow and NGS produce errors that

can potentially add to the counts of programmed variants, particularly those variants that are 1 nucleotide 'hamming' distance from the wild-type codon. In addition, we assessed replicate reproducibility at codon level data. We tried different ways to call hits – at codon level, or at amino acid level, option to include only the programmed variants etc. In this library, we designed synonymous codons (silent mutations) at 37 codon positions. The raw read counts were normalized to the fraction of counts at each codon position.

To determine the degree of enrichment or depletion of each p53 variant, we calculated the log10 ratios of fractional read counts in GFP$^{lo}$ over GFP$^{hi}$ cell populations for all non-synonymous and synonymous p53 variants. These enrichment scores expressed as $\log_{10}(\text{GFP}^{lo}/\text{GFP}^{hi})$ were depicted as heatmaps and aligned to the *TP53* mutational data obtained from patients with myeloid malignancies.

**Mouse maintenance, transplants, and genotyping**

All mouse experiments were performed in compliance with an Institutional Animal Care and Use Committee (IACUC)-approved animal protocol (Brigham and Women's Hospital, Boston, MA). Mice with conditional *Trp53* knock-in and knock-out alleles (p53LSL.R172H, p53LSL.R270H, and p53$^{flox/flox}$) were obtained from Tyler Jacks (Massachusetts Institute of Technology, Cambridge, MA) and bred to Mx1-Cre as well as congenic C57BL/6J mice expressing pan-leukocyte markers CD45.1 or CD45.2, respectively. F1 generation mice double-positive for CD45.1 and CD45.2 were used as recipients. All mice including recipients were maintained on a mixed C57BL/6.129svj background. Genotyping was performed by Transnetyx Inc. (Cordova, TN). Donor Mx1-Cre mice expressing conditional *Trp53* knock-in and knock-out alleles were treated with 4 doses of 200 mg of high-molecular weight poly(I:C) (Invivogen, #tlrl-pic-5) at 6-10 weeks of age. Hematopoietic competition assays were performed as follows: donor mice were euthanized, whole bone marrow cells were isolated and subjected to isolation of c-Kit$^{+}$ HSPCs using anti-CD117 (c-Kit) magnetic beads (Miltenyi, # 130-091-224). c-Kit$^{+}$ cells of the different *Trp53* genotypes used in this study were enumerated and mixed at a 1:1 ratio. Prior to retro-orbital transplantation of a minimum of $2\times10^5$ c-Kit$^{+}$ cells per mouse, recipient mice were sublethally irradiated with 5 Gy twice within a 4-hour interval. After hematopoietic reconstitution was confirmed by peripheral blood (PB) chimerism using flow cytometry, mice were randomized into treatment and control groups. Mice in the treatment arm received a single dose of sublethal

irradiation (2.5 Gy). Thereafter, both treatment and control groups were assessed for PB chimerism every 4 weeks up until 16 weeks post treatment. For PB chimerism assessment, peripheral blood was collected, erythrocytes were lysed (Qiagen, # 158904) and leukocytes were stained for CD45.1 (Biolegend, #110708), CD45.2 (Biolegend, #109814), CD11b (Biolegend, # 101224), CD3ε (Biolegend, # 100306), B220 (Biolegend, # 103222), and assessed by flow cytometry using a FACSCanto II (BD Biosciences).

## Human genetics

*TP53* mutational data from a total of 1,040 patients with myeloid malignancies that underwent targeted panel sequencing at Dana-Farber Cancer Institute (with approval of the local institutional review board), or were being sequenced as part of the diagnostic work-up of the German-Austrian AML Study Group (AMLSG) trials, or that were reported in Lindsley et al. (*4*) were analyzed and visualized using MutationMapper (*49*, *50*). Similarly, *TP53* mutational data from 129 published individuals with clonal hematopoiesis of indeterminate potential (CHIP) were compiled (*27*, *28*, *30*, *51-55*) and visualized using MutationMapper (*49*, *50*).

For a total of n=164 patients with acute myeloid leukemia treated with anthracycline/cytarabine-based regimens within clinical trials performed by the German-Austrian AML-SG study group *TP53* mutational data at time of study entry and consecutive clinical outcome data were available. Kaplan-Meier survival analyses were performed and depicted as event-free survival (defined as time from study entry until treatment failure, relapse from complete remission, death from any cause or until censoring at the time of last follow-up) and overall survival (defined as the time from study entry until death from any cause or until censoring at the time of last follow-up). Differences in survival curves were assessed using log-rank (Mantel-Cox) tests.

For a subset of n=101 and n=145 of *TP53*-mutant AML patients co-mutation data and cytogenetic data were available. Co-mutation plots were generated using OncoPrinter (*49*, *50*). Fisher's exact test was used to test for statistical significance with respect to number of co-mutations or frequency of complex karyotype in patients with *TP53* missense or truncating mutations, respectively.

## Statistics

Statistical significance testing was performed using GraphPad Prism software (version 8.0.1).
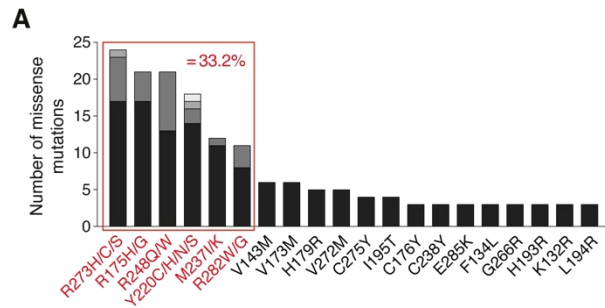
**A**



**Fig. S1. Most frequent *TP53* missense mutations in high-risk MDS patients.**

(**A**) Top 20 most frequent *TP53* missense mutations in high-risk MDS patients undergoing allogeneic hematopoietic stem cell transplantation (Lindsley et al., NEJM 217). Red square indicates the six most frequently affected amino acid residues comprising 33.2% of all mutations and that were chosen for in-depth analysis in this study.
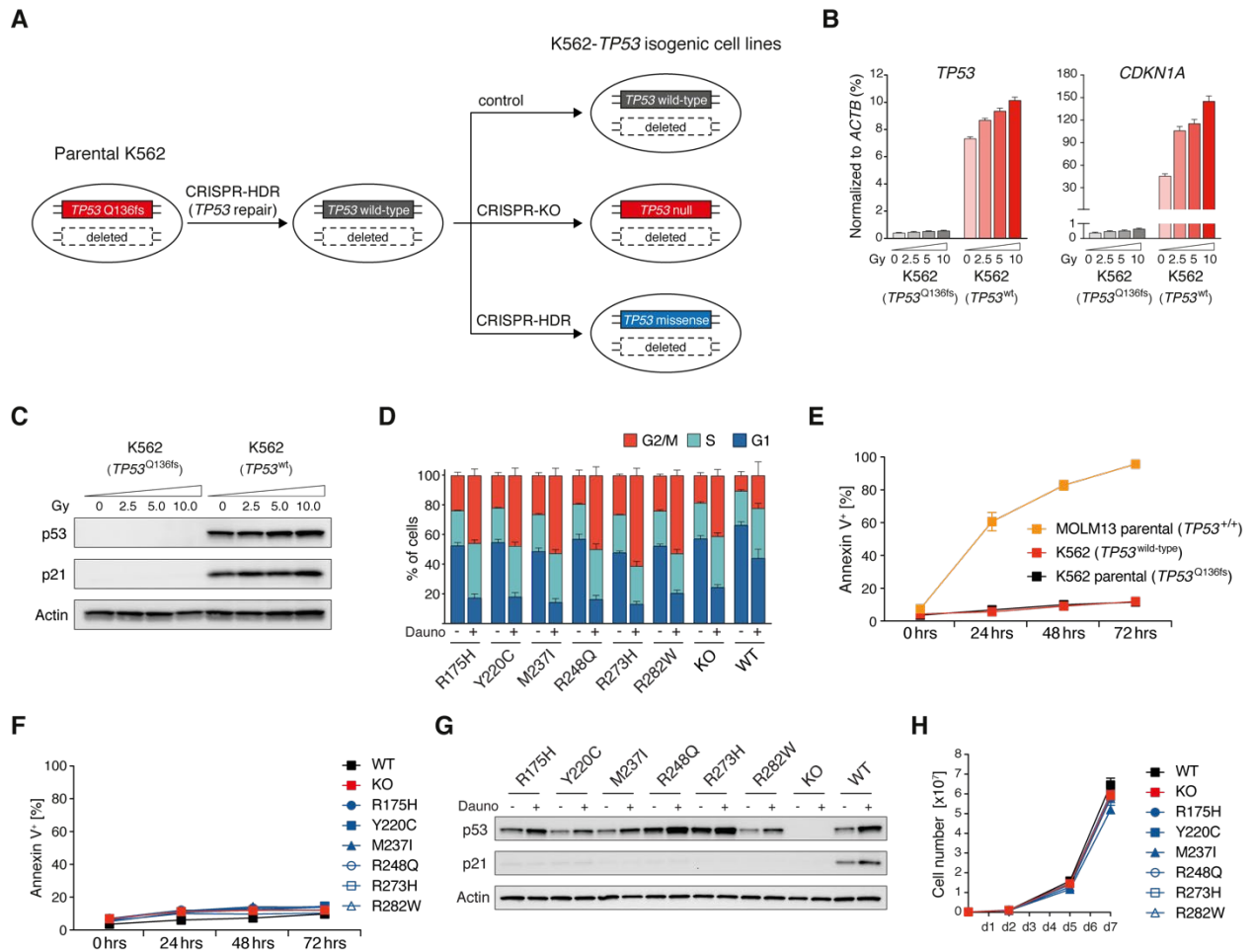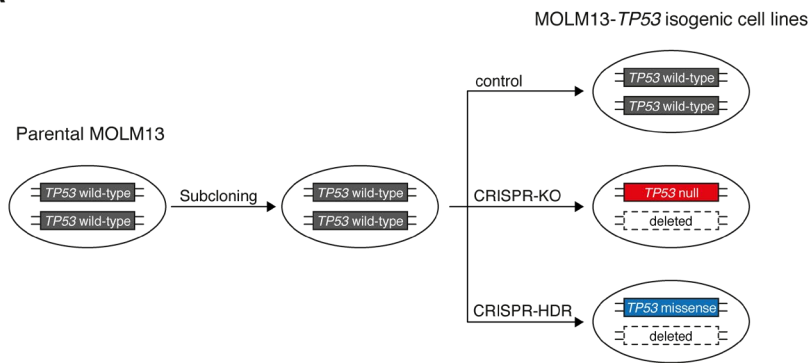
**Figure S2**



**Fig. S2. Generation and characterization of K562-*TP53* isogenic AML cell lines.**

(**A**) Schematic of the experimental workflow for generating K562-*TP53* isogenic AML cell lines. CRISPR-HDR, CRISPR-Cas9-mediated homology-directed repair. (**B**) RT-qPCR measuring expression of *TP53* and *CDKN1A* transcripts normalized to *ACTB* in parental K562 cells carrying a *TP53*[Q136fs] frame-shift mutation as well as K562 cells with repaired *TP53*. Cells were treated with increasing doses of gamma-irradiation and RNA was collected 6 hours later (replicates n=2, error bars indicate s.e.m.). (**C**) K562 cells carrying a *TP53*[Q136fs] frame-shift mutation as well as K562 cells with repaired *TP53* were treated with increasing doses of gamma-irradiation, 6 hours later after whole cell protein lysates were collected, run on a polyacrylamide gel, and immunoblotted for p53, p21 and actin (replicates n=2, representative images are shown). (**D**) K562-*TP53* isogenic AML cell lines were treated with DMSO (-) or 100nM daunorubicin (+) for 24 hours, after which cells were stained with

11

CytoPhase™ violet and analyzed by flow cytometry to assess cell cycle distribution (replicates n=3, error bars indicate s.e.m.). (**E**) Parental MOLM13 (*TP53*$^{+/+}$), parental K562 (*TP53*$^{Q136fs}$), and K562 with repaired *TP53* were treated with 100nM daunorubicin for up to 72 hours. At the indicated time points, cells were stained with Annexin V and analyzed by flow cytometry to assess total apoptotic cells (replicates n=3, symbols represent averages of experimental replicates, error bars indicate s.e.m.). (**F**) K562-*TP53* isogenic AML cells of the indicated genotypes were treated with 100nM daunorubicin for up to 72 hours. At the indicated time points, cells were stained with Annexin V and analyzed by flow cytometry to assess total apoptotic cells (replicates n=3, symbols represent averages of experimental replicates, error bars indicate s.e.m.). (**G**) K562-*TP53* isogenic AML cell lines were treated with DMSO (-) or 100nM daunorubicin (+) for 6 hours, after which whole cell protein lysates were collected, run on a polyacrylamide gel, and immunoblotted for p53, p21 and actin (replicates n=3, representative images are shown). (**H**) Growth kinetics of K562-*TP53* isogenic AML cell lines (replicates n=3, symbols represent averages of experimental replicates, error bars indicate s.e.m.).

**A**

MOLM13-*TP53* isogenic cell lines

Parental MOLM13

Subcloning

control → *TP53* wild-type / *TP53* wild-type

CRISPR-KO → *TP53* null / deleted

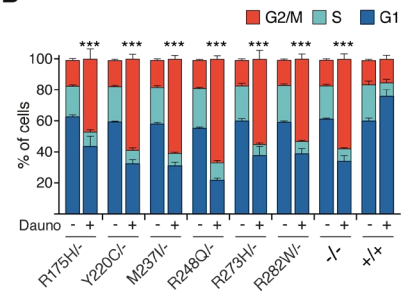CRISPR-HDR → *TP53* missense / deleted

**B**



**Fig. S3. Generation and characterization of MOLM13-*TP53* isogenic AML cell lines.**

(**A**) Schematic of the experimental workflow for generating MOLM13-*TP53* isogenic AML cell lines. CRISPR-HDR, CRISPR-Cas9-mediated homology-directed repair. (**B**) MOLM13-*TP53* isogenic AML cell lines were treated with DMSO (-) or 100nM Daunorubicin (+) for 24 hours, after which cells were stained with CytoPhase™ violet and analyzed by flow cytometry to assess cell cycle distribution (replicates n=3, error bars indicate s.e.m., two-tailed Student's *t*-test, *** $p < 0.001$).
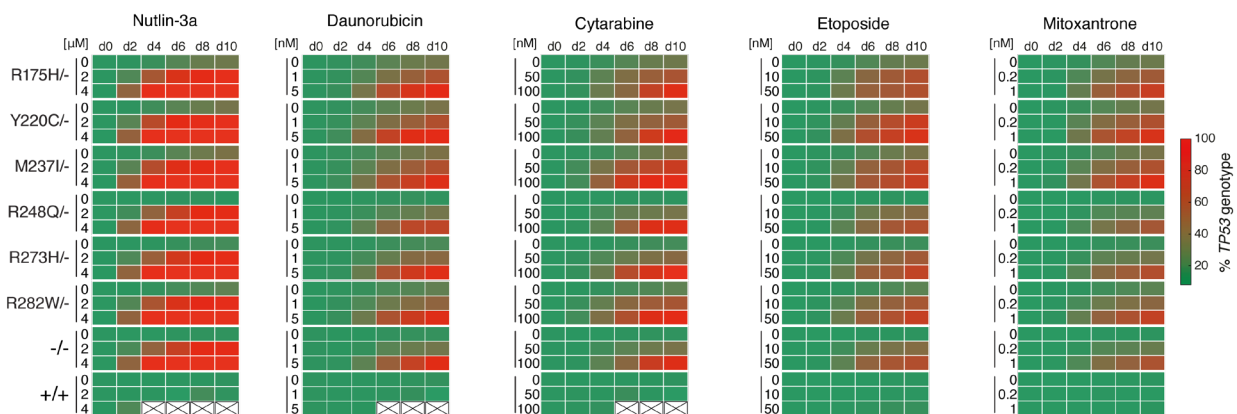
**Figure S4**

**A**



**B**
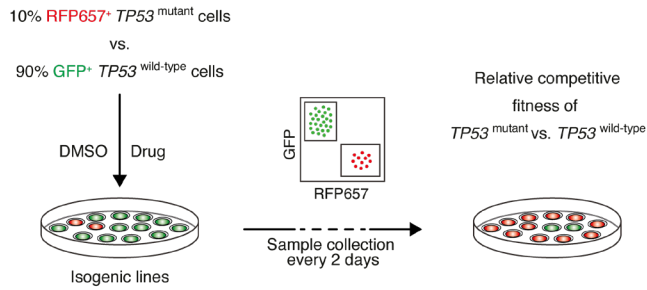
MOLM13-*TP53* isogenic cell lines  (*TP53* mutant vs. *TP53* wild-type)



**C**

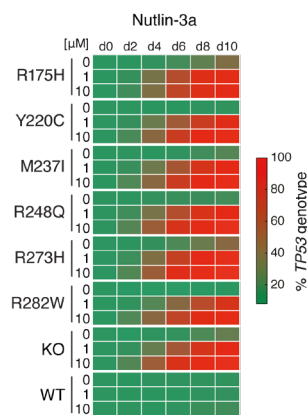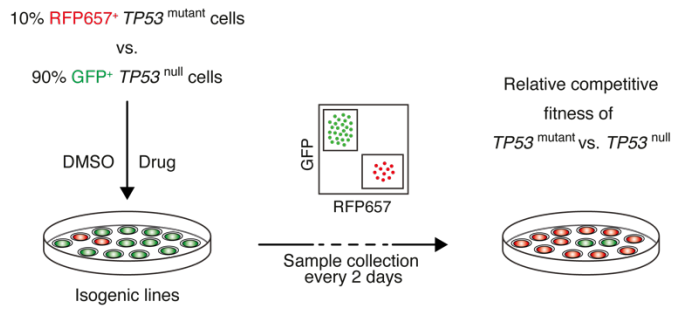K562-*TP53* isogenic cell lines  (*TP53* mutant vs. *TP53* wild-type)



**Fig. S4. In vitro competition assays (*TP53* mutant vs. *TP53* wild-type) in MOLM13-*TP53* and K562-*TP53* isogenic AML cells.**

(**A**) Schematic of the experimental workflow for *in vitro* competition assays in MOLM13-*TP53* and K562-*TP53* isogenic AML cell lines. 10% *TP53*mutant RFP657+ cells were mixed with 90%

*TP53*$^{+/+}$ GFP$^+$ cells and cultured in the presence of DMSO or indicated drugs over a period of 10 days during which repetitive flow-cytometric measurements were performed. (**B**) Heatmaps depicting results from *in vitro* competition assays in MOLM13-*TP53* isogenic AML cell lines (*TP53*$^{mutant}$ vs. *TP53*$^{wild\text{-}type}$). (**C**) Heatmaps depicting results from *in vitro* competition assays in K562-*TP53* isogenic AML cell lines (*TP53*$^{mutant}$ vs. *TP53*$^{wild\text{-}type}$). (replicates n=2-3, color shades represent increasing percentage of *TP53*$^{mutant}$ cells).

# Figure S5

## A

10% RFP657⁺ *TP53* mutant cells
vs.
90% GFP⁺ *TP53* null cells

DMSO | Drug

Isogenic lines

GFP
RFP657

Sample collection
every 2 days

Relative competitive
fitness of
*TP53* mutant vs. *TP53* null

## B

**MOLM13-*TP53* isogenic cell lines**  (*TP53* mutant vs. *TP53* null)



Nutlin-3a | Daunorubicin

% *TP53* genotype

## C

**K562-*TP53* isogenic cell lines**  (*TP53* mutant vs. *TP53* null)



Nutlin-3a

% *TP53* genotype
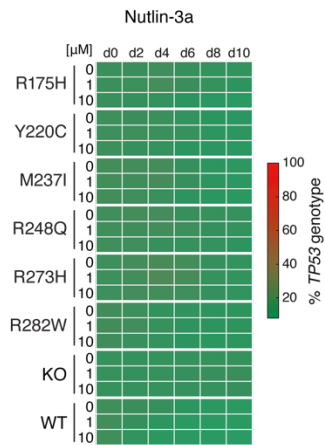
**Fig. S5. In vitro competition assays ($TP53^{mutant}$ vs. $TP53^{null}$) in MOLM13-*TP53* and K562-*TP53* isogenic AML cells.**

(**A**) Schematic of the experimental workflow for *in vitro* competition assays in MOLM13-*TP53* and K562-*TP53* isogenic AML cell lines. 10% $TP53^{mutant}$ RFP657$^+$ cells were mixed with 90% $TP53^{null}$ GFP$^+$ cells and cultured in the presence of DMSO or indicated drugs over a period of 10 days during which repetitive flow-cytometric measurements were performed. (**B**) Heatmaps depicting results from *in vitro* competition assays in MOLM13-*TP53* isogenic AML cell lines ($TP53^{mutant}$ vs. $TP53^{null}$). (**C**) Heatmaps depicting results from *in vitro* competition assays in K562-*TP53* isogenic AML cell lines ($TP53^{mutant}$ vs. $TP53^{null}$). (replicates n=2-3, color shades represent increasing percentage of $TP53^{mutant}$ cells).
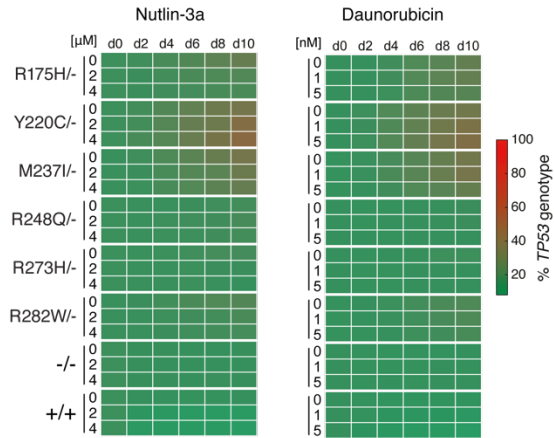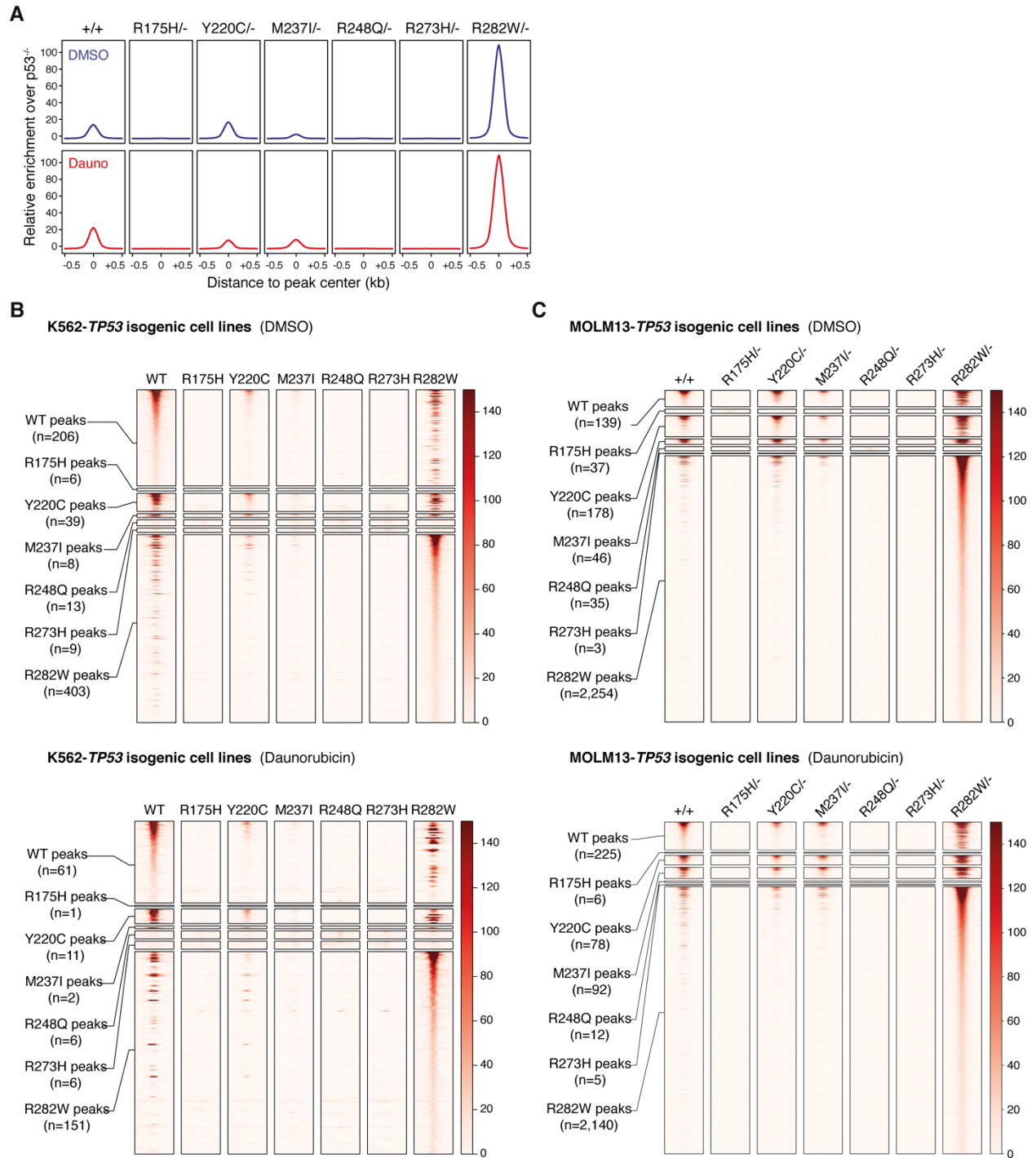
**Fig. S6. ChIP-seq in K562-*TP53* and MOLM13-*TP53* isogenic AML cells.**

(**A**) Genome-wide relative enrichment of wild-type and missense mutant p53 variants (ChIP/p53$^{-/-}$) over transcriptional start site (TSS)-proximal regions (-10kb – first intron) in

18

MOLM13-*TP53* isogenic cell lines upon treatment with DMSO or 100nM daunorubicin for 6 hours. (**B**) Heatmaps of wild-type and missense mutant p53 ChIP-seq peaks over TSS-proximal regions (-10kb – first intron) in K562-*TP53* isogenic cell lines upon treatment with DMSO (upper panel) or 100 nM daunorubicin (lower panel) for 24 hours and shown in a horizontal window 1 kb around the peak center. Heatmaps show enrichment and overlap of wild-type or missense mutant p53 ChIP-seq peaks identified from each cell line (horizontally) in every other cell line (vertically) thereby identifying shared and cell line-specific peaks, respectively. (**C**) Heatmaps of wild-type and missense mutant p53 ChIP-seq peaks over TSS-proximal regions (-10kb – first intron) in MOLM13-*TP53* isogenic cell lines upon treatment with DMSO (upper panel) or 100 nM daunorubicin (lower panel) for 6 hours and shown in a horizontal window 1 kb around the peak center. Heatmaps show enrichment of wild-type or missense mutant p53 ChIP-seq peaks as described under (B).
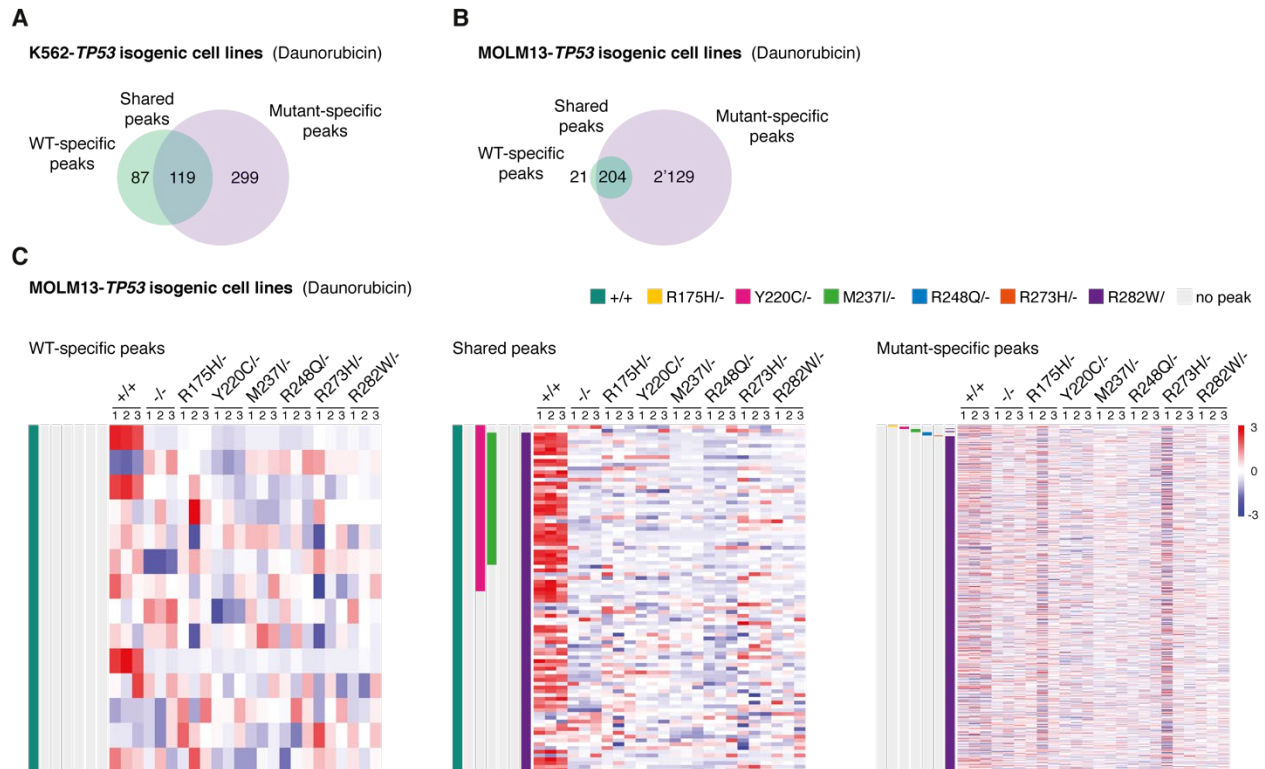
**Fig. S7. Integrated ChIP-seq and RNA-seq analysis.**

(**A**) Venn diagram showing shared, wild-type (WT)- and mutant-specific (pooled from all p53 mutant cell lines) ChIP-seq peaks in K562-*TP53* isogenic cell lines treated with 100nM daunorubicin for 24 hours. (**B**) Venn diagram showing shared, WT- and mutant-specific (pooled from all p53 mutant cell lines) ChiP-seq peaks in MOLM13-*TP53* isogenic cell lines treated with 100nM daunorubicin for 6 hours. (**C**) Heatmap depicting normalized expression of genes corresponding to WT-specific (left), shared (middle), and p53 mutant-specific (right) ChiP-seq peaks in MOLM13-*TP53* isogenic cell lines treated with 100nM daunorubicin for 6 hours (RNA-seq experimental replicates n=3).

**A**

**K562-*TP53* isogenic cell lines** (DMSO)



Higher expression than in WT — Lower expression than in WT

**B**

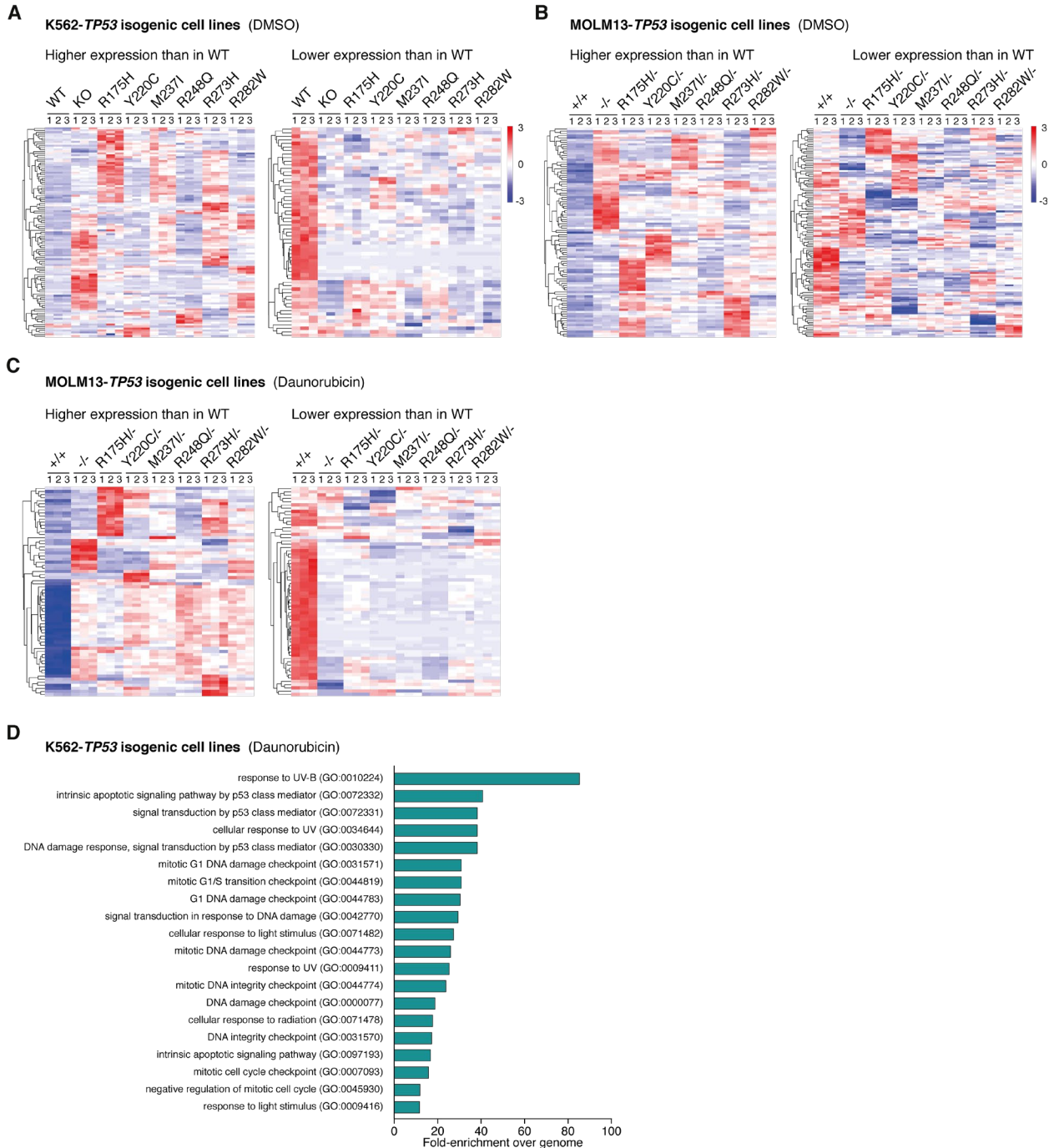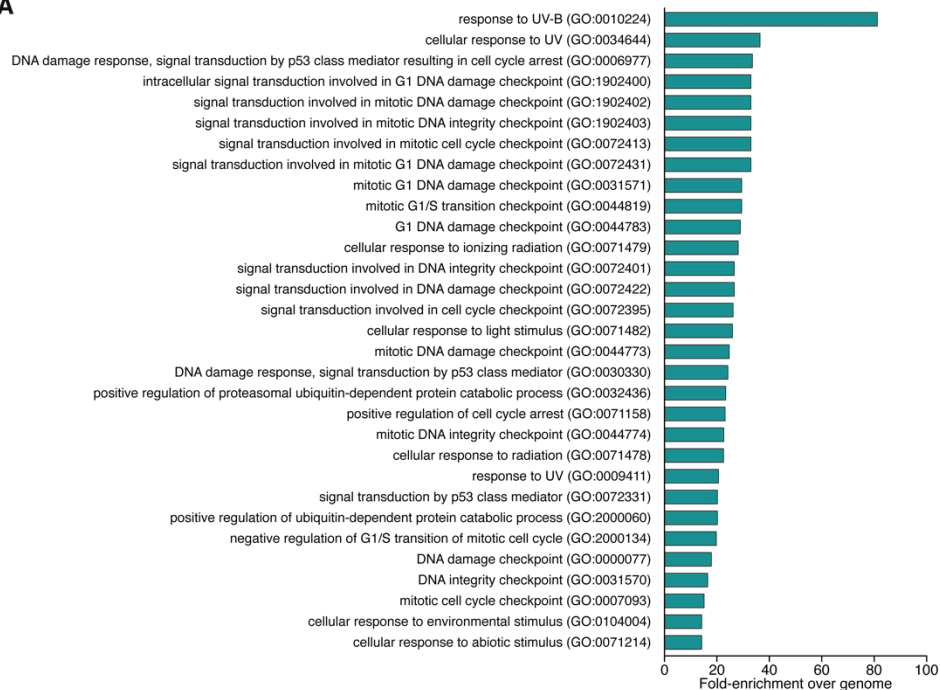**MOLM13-*TP53* isogenic cell lines** (DMSO)



Higher expression than in WT — Lower expression than in WT

**C**

**MOLM13-*TP53* isogenic cell lines** (Daunorubicin)



Higher expression than in WT — Lower expression than in WT

**D**

**K562-*TP53* isogenic cell lines** (Daunorubicin)



response to UV-B (GO:0010224)
intrinsic apoptotic signaling pathway by p53 class mediator (GO:0072332)
signal transduction by p53 class mediator (GO:0072331)
cellular response to UV (GO:0034644)
DNA damage response, signal transduction by p53 class mediator (GO:0030330)
mitotic G1 DNA damage checkpoint (GO:0031571)
mitotic G1/S transition checkpoint (GO:0044819)
G1 DNA damage checkpoint (GO:0044783)
signal transduction in response to DNA damage (GO:0042770)
cellular response to light stimulus (GO:0071482)
mitotic DNA damage checkpoint (GO:0044773)
response to UV (GO:0009411)
mitotic DNA integrity checkpoint (GO:0044774)
DNA damage checkpoint (GO:0000077)
cellular response to radiation (GO:0071478)
DNA integrity checkpoint (GO:0031570)
intrinsic apoptotic signaling pathway (GO:0097193)
mitotic cell cycle checkpoint (GO:0007093)
negative regulation of mitotic cell cycle (GO:0045930)
response to light stimulus (GO:0009416)

Fold-enrichment over genome

**Fig. S8. RNA-seq and gene ontology analysis.**

(**A**) Heatmap of the pooled top 30 (left) and bottom 30 (right) genes relative to wild-type p53 in K562-*TP53* isogenic cell lines treated with DMSO for 24 hours (RNA-seq experimental replicates n=3). (**B**) Heatmap of the pooled top 30 (left) and bottom 30 (right) genes relative to wild-type

21

p53 in MOLM13-*TP53* isogenic cell lines treated with DMSO for 6 hours (RNA-seq experimental replicates n=3). (**C**) Heatmap of the pooled top 30 (left) and pooled bottom 30 (right) genes relative to wild-type p53 in MOLM13-*TP53* isogenic cell lines treated with 100nM daunorubicin for 6 hours (RNA-seq experimental replicates n=3). (**D**) Gene ontology (GO) analysis of the shared transcriptional signature among bottom 30 differentially-expressed genes in daunorubicin-treated K562-*TP53* isogenic AML cells.
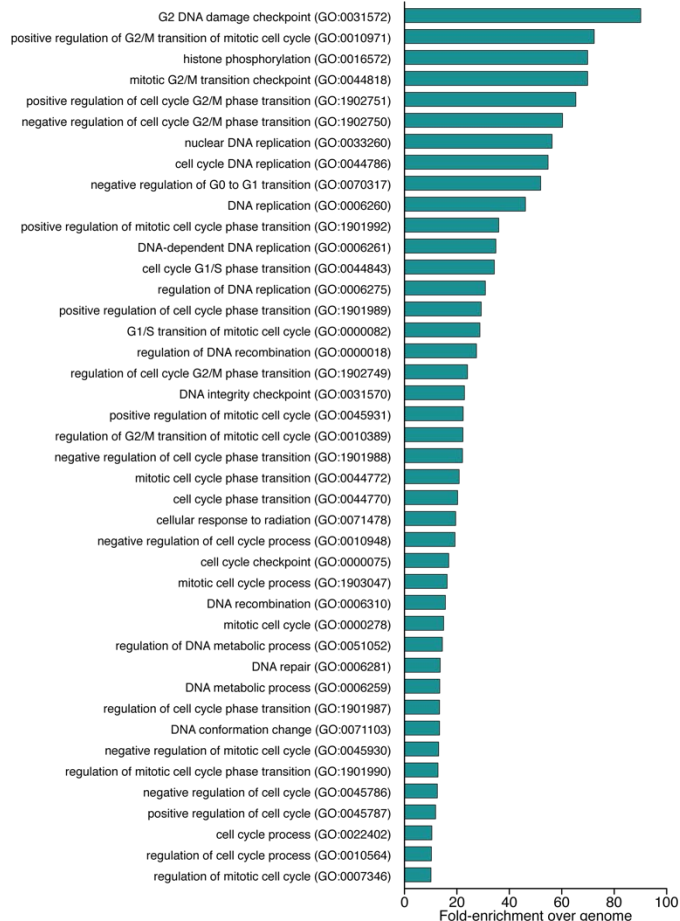
**Figure S9**

**A**



**B**

**Fig. S9. Gene ontology analysis in MOLM13-*TP53* isogenic cells.**

(**A**) Gene ontology (GO) analysis of the shared transcriptional signature among bottom 30 differentially-expressed genes in daunorubicin-treated MOLM13-*TP53* isogenic AML cells (related to Fig. 2G right panel) (**B**) Gene ontology (GO) analysis of the shared transcriptional signature among top30 differentially-expressed genes in daunorubicin-treated MOLM-*TP53* isogenic AML cells (related to Fig. 2G left panel).
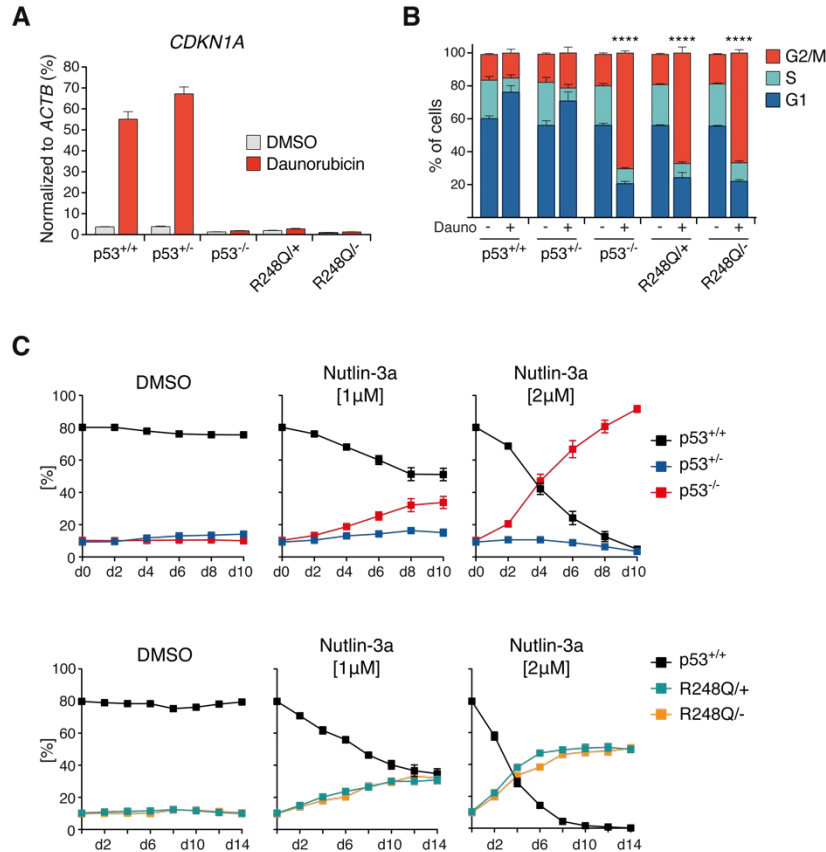
**Fig. S10. Dominant-negative phenotypes MOLM13-*TP53* isogenic cells.**

(**A**) RT-qPCR measuring expression of *CDKN1A* transcripts relative to *ACTB* in MOLM13-*TP53* isogenic cell lines of the indicated genotypes upon treatment with DMSO or 100nM daunorubicin for 6 hours (replicates n=3, error bars indicate s.e.m.). (**B**) MOLM13-*TP53* isogenic AML cell lines of the indicated genotypes were treated with DMSO (-) or 100nM Daunorubicin (+) for 24 hours, after which cells were stained with CytoPhase™ violet and analyzed by flow cytometry to assess cell cycle distribution (replicates n=3, error bars indicate s.e.m., two-tailed Student's *t*-test, *** $p<0.001$). (**C**) Results from *in vitro* competition assays in MOLM13-*TP53* isogenic AML cell lines (*TP53*^+/+ vs. *TP53*^+/- vs. *TP53*^+/+) and (*TP53*^+/+ vs. *TP53*^R248Q/+ vs. *TP53*^R248Q/-) continuously treated with DMSO or Nutlin-3a at 1 or 2 μM over a period of 10-14 days. (replicates n=3).
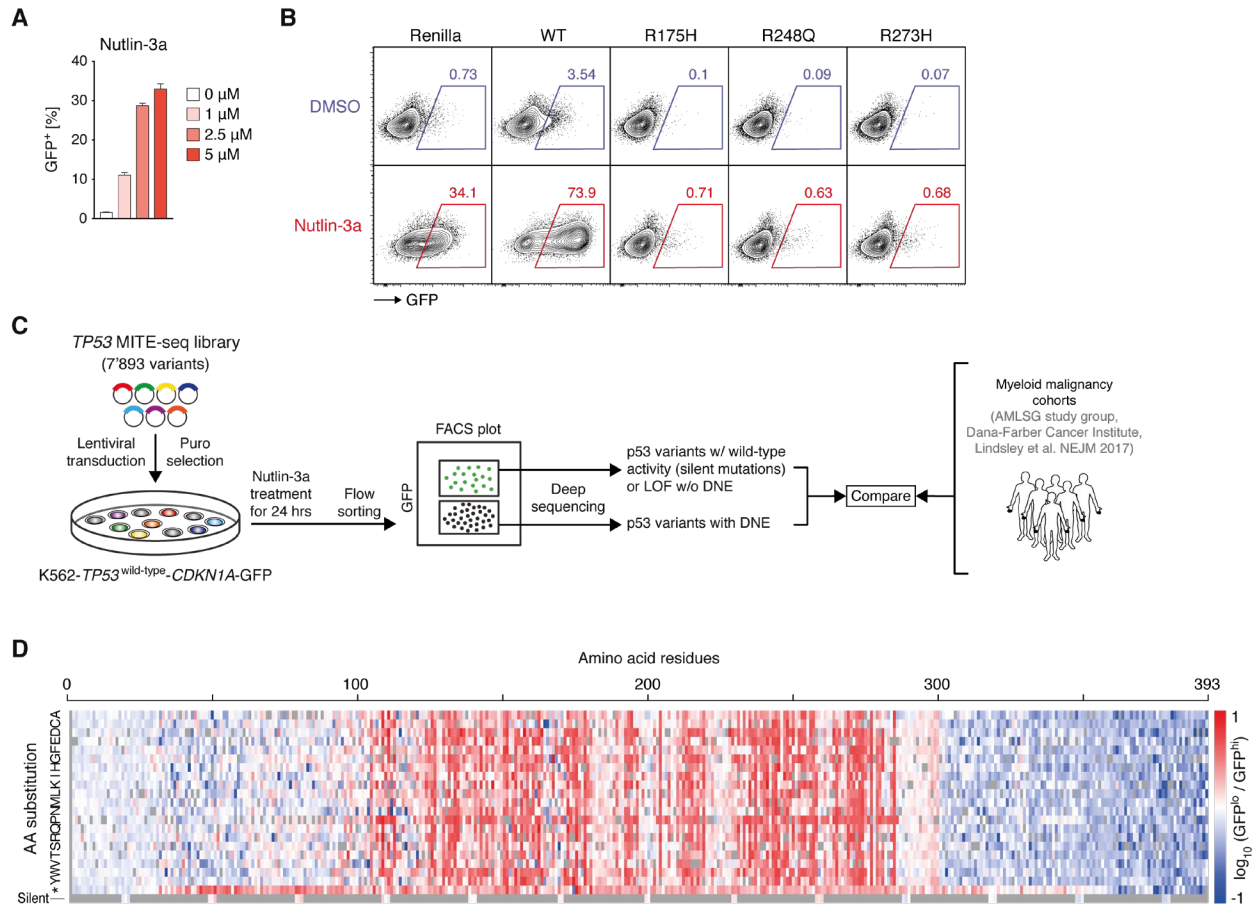
**Figure S11**



**Fig. S11. MITE-seq saturation mutagenesis screen.**

(**A**) p21-GFP expression in K562-*TP53*[wild-type]-*CDKN1A*-GFP reporter cell line upon treatment for 24 hours with increasing doses of Nutlin-3a as determined by flow-cytometry. (replicates n=2, error bars indicate s.e.m.). (**B**) Representative flow-cytometry plots showing p21-GFP expression in K562-*TP53*[wild-type]-*CDKN1A*-GFP reporter cell lines lentivirally transduced with *TP53* variants of the indicated genotypes or Renilla luciferase as control. Cells were treated with either DMSO as control (top panel) or 5μM Nutlin-3a (bottom panel) for 24 hours. (**C**) Schematic of the experimental workflow for *TP53* saturation mutagenesis screen. K562-*TP53*[wild-type]-*CDKN1A*-GFP reporter cells were lentivirally transduced at a low MOI with a library containing all 7,893 possible single amino acid *TP53* variants. After puromycin selection and treatment with either DMSO (control) or 5μM nutlin-3a for 24 hours, GFP[hi] (enriched for *TP53* variants with pure LOF or WT activity) and GFP[lo] (enriched for *TP53* variants with DNE) cells were sorted, genomic DNA

isolated, PCR-amplified and subjected to deep sequencing to quantify p53 variant representation in each population. (**D**) Heatmap depicting MITE-seq results for DMSO-treated cells shown as log10 of the ratio of normalized read counts in GFP$^{lo}$ over GFP$^{hi}$ cells per *TP53* variant.
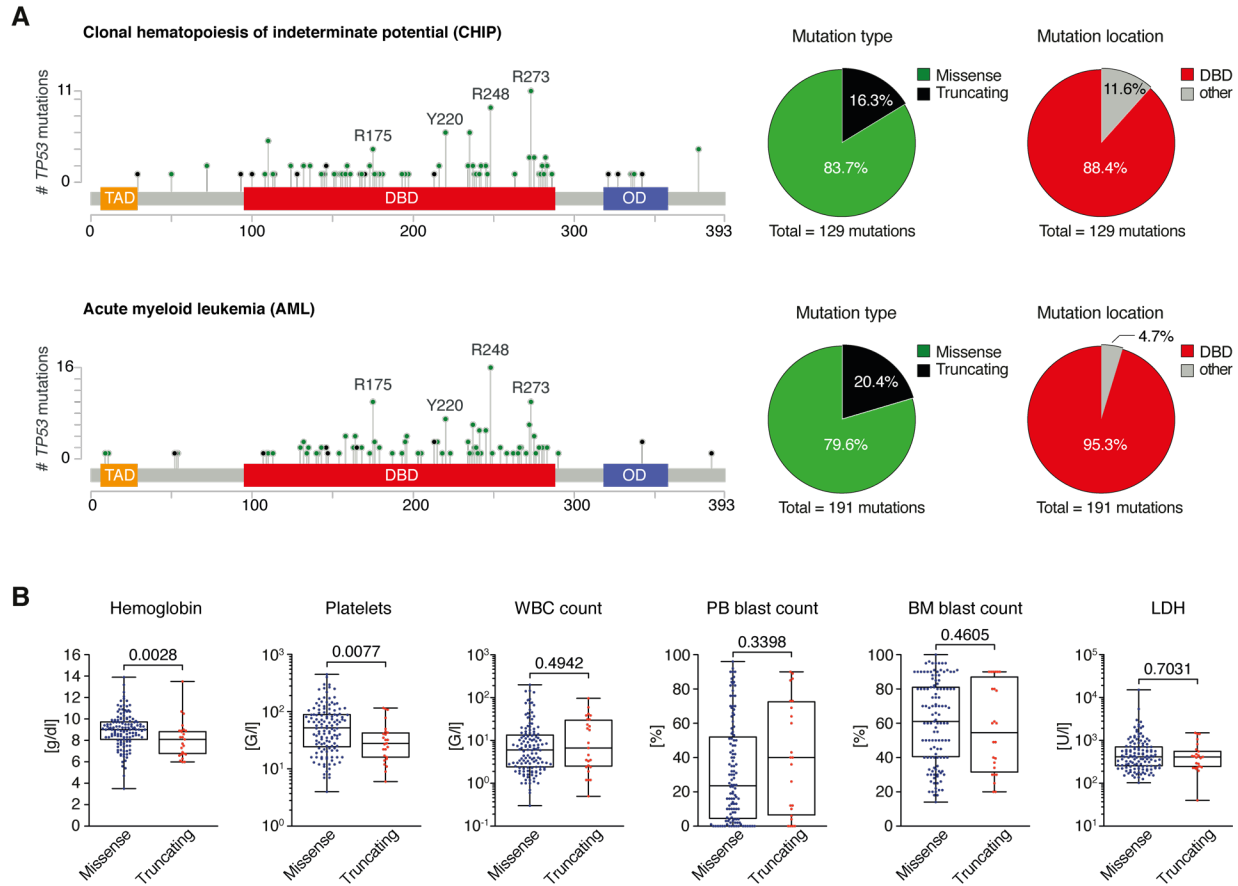
**Figure S12**

**A**



**B**



**Fig. S12.** *TP53* **mutations in clonal hematopoiesis of indeterminate potential (CHIP) vs. AML and clinicopathological parameters in AML patients.**

(**A**) Lollipop plot demonstrating *TP53* mutational data from n=129 individuals with CHIP (upper panel) and n=165 patients with AML (lower panel), missense mutations (green circles), truncating mutations (black circles) comprising frame-shift, nonsense, and splice mutations are shown aligned with the p53 domain structure. TAD, transactivation domain; DBD, DNA-binding domain; OD, oligomerization domain. Frequency of type of mutation and mutation location (right). (**B**) Hematologic parameters in AML patients with *TP53* missense or truncating mutations (the latter comprising frame-shift, nonsense, and splice mutations). (Mann-Whitney test).

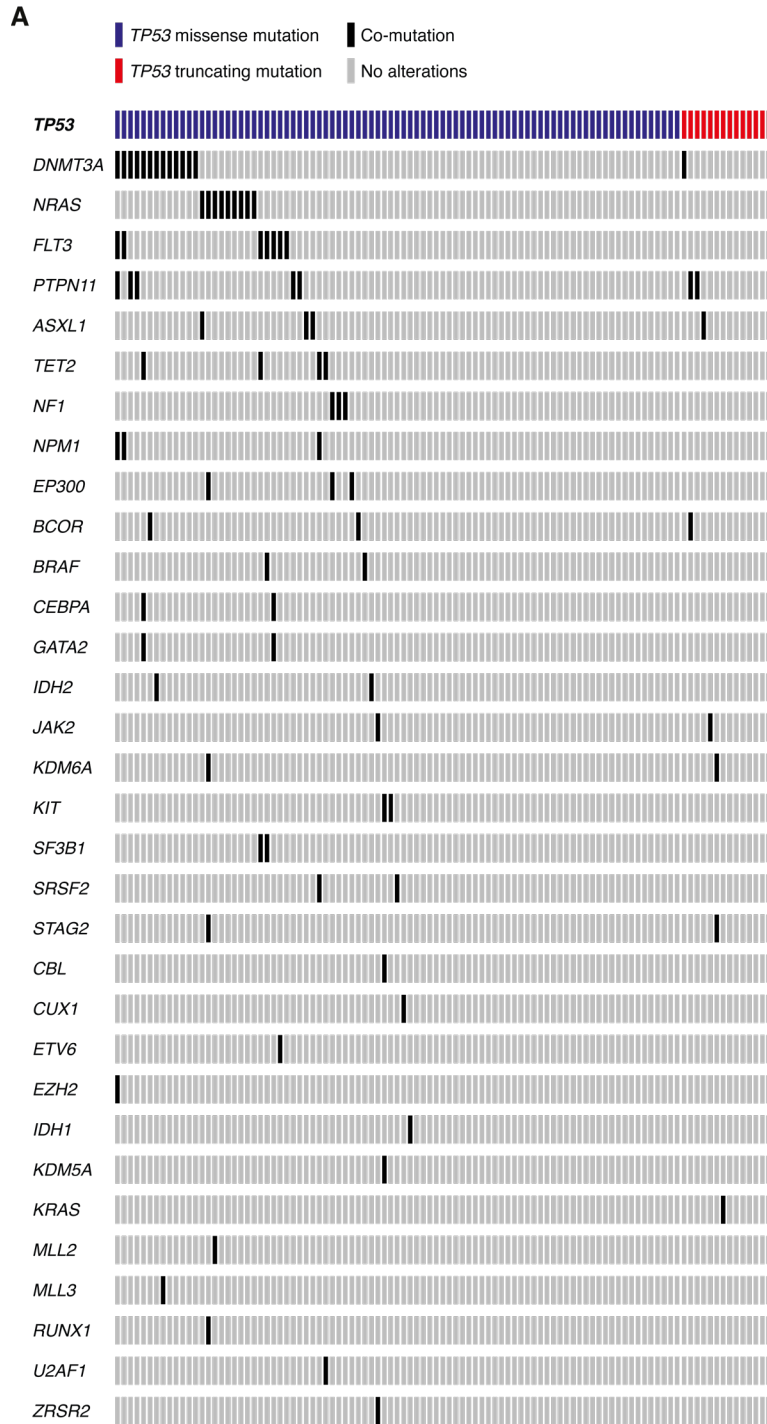**Figure S13**

**A**



**Fig. S13. Co-mutations in patients with *TP53*-mutant AML.**

(**A**) Co-mutation plot showing non-synonymous mutations in individual genes as labeled on the left. Mutations are depicted by colored bars and each column represents 1 of the 101 patients with

*TP53* mutations (n=87 missense mutations, n=14 truncating mutations) for which co-mutation data were available.

**Table S1**

| | | Cell line name | R175H | Y220C | M237I | R248Q | R273H | R282W | KO |
|---|---|---|---|---|---|---|---|---|---|
| *TP53* allele | #1 | Genotype | R175H | Y220C | M237I | R248Q | R273H | R282W | P191fs |
| | | # of reads | 31,564 | 53,971 | 24,507 | 14,867 | 51,224 | 37,376 | 60,173 |
| | | VAF [%] | 95.96 | 98.13 | 50.03 | 48.14 | 100 | 100 | 94.97 |
| | #2 | Genotype | n/a | n/a | M237I & S227F | WT | n/a | n/a | n/a |
| | | # of reads | n/a | n/a | 23,451 | 12,559 | n/a | n/a | n/a |
| | | VAF [%] | n/a | n/a | 47.87 | 40.67 | n/a | n/a | n/a |
| cDNA | | Genotype | n/a | n/a | M237I | R248Q | n/a | n/a | n/a |
| | | # of reads | n/a | n/a | 66,579 | 57,994 | n/a | n/a | n/a |
| | | VAF [%] | n/a | n/a | 94.45 | 97.97 | n/a | n/a | n/a |

**Table S1. Next-generation sequencing results of K562-*TP53* isogenic cell lines.**

Results from genetic validation of K562-*TP53* isogenic cell lines using ultra-deep amplicon sequencing. VAF, variant allele frequency.

**Table S2**

| | | Cell line name | **R175H/-** | **Y220C/-** | **M237I/-** | **R248Q/-** | **R273H/-** | **R282W/-** | **-/-** |
|---|---|---|---|---|---|---|---|---|---|
| *TP53* allele | #1 | Genotype | R175H | Y220C | M237I | R248Q | R273H | R282W | P191fs |
| | | # of reads | 41,441 | 17,434 | 58,460 | 28,177 | 46,712 | 40,169 | 27,304 |
| | | VAF [%] | 48.5 | 34.32 | 47.76 | 46.53 | 51.51 | 53.7 | 44.38 |
| | #2 | Genotype | T170fs | Y220fs | M243fs | R248fs | L265fs | C277fs | A189fs |
| | | # of reads | 43,466 | 18.216 | 55,445 | 31,236 | 43,959 | 34,633 | 23,394 |
| | | VAF [%] | 50.87 | 35.86 | 45.29 | 51.58 | 48.48 | 46.29 | 38.03 |

**Table S2. Next-generation sequencing results of MOLM13-*TP53* isogenic cell lines.**

Results from genetic validation of MOLM13-*TP53* isogenic cell lines using ultra-deep amplicon sequencing. VAF, variant allele frequency.

**References and Notes**

1. A. J. Levine, M. Oren, The first 30 years of p53: growing ever more complex. *Nat. Rev. Cancer*. **9**, 749–758 (2009).

2. R. Bejar *et al.*, Clinical effect of point mutations in myelodysplastic syndromes. *N. Engl. J. Med*. **364**, 2496–2506 (2011).

3. F. G. Rücker *et al.*, TP53 alterations in acute myeloid leukemia with complex karyotype correlate with specific copy number alterations, monosomal karyotype, and dismal outcome. *Blood*. **119**, 2114–2121 (2012).

4. R. C. Lindsley *et al.*, Prognostic Mutations in Myelodysplastic Syndrome after Stem-Cell Transplantation. *N. Engl. J. Med*. **376**, 536–547 (2017).

5. E. R. Kastenhuber, S. W. Lowe, Putting p53 in Context. *Cell*. **170**, 1062–1078 (2017).

6. Y. Liu *et al.*, Deletions linked to TP53 loss drive cancer through p53-independent mechanisms. *Nature*. **531**, 471–475 (2016).

7. R. Brosh, V. Rotter, When mutants gain new powers: news from the mutant p53 field. *Nat. Rev. Cancer*. **9**, 701–713 (2009).

8. E. H. Baugh, H. Ke, A. J. Levine, R. A. Bonneau, C. S. Chan, Why are there hotspot mutations in the TP53 gene in human cancers? *Cell Death Differ*. **25**, 154–160 (2018).

9. D. Dittmer *et al.*, Gain of function mutations in p53. *Nat. Genet*. **4**, 42–46 (1993).

10. K. P. Olive *et al.*, Mutant p53 Gain of Function in Two Mouse Models of Li-Fraumeni Syndrome. *Cell*. **119**, 847–860 (2004).

11. S. Di Agostino *et al.*, Gain of function of mutant p53: The mutant p53/NF-Y protein complex reveals an aberrant transcriptional mechanism of cell cycle regulation. *Cancer Cell*. **10**, 191–202 (2006).

12. P. M. Do *et al.*, Mutant p53 cooperates with ETS2 to promote etoposide resistance. *Genes Dev*. **26**, 830–845 (2012).

13. J. Xu *et al.*, Gain of function of mutant p53 by coaggregation with multiple tumor suppressors. *Nat. Chem. Biol*. **7**, 285–295 (2011).

14. J. Zhu *et al.*, Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. *Nature*. **525**, 206–211 (2015).

15. D. Walerych *et al.*, Proteasome machinery is instrumental in a common gain-of-function program of the p53 missense mutants in cancer. *Nat. Cell Biol*. **18**, 897–909 (2016).

16. S. Srivastava, S. Wang, Y. A. Tong, Z. M. Hao, E. H. Chang, Dominant negative effect of a germ-line mutant p53: a step fostering tumorigenesis. *Cancer Res*. **53**, 4452–4455 (1993).

17. M. E. Hegi *et al.*, p53 transdominance but no gain of function in mouse brain tumor model. *Cancer Res*. **60**, 3019–3024 (2000).

18. A. de Vries *et al.*, Targeted point mutations of p53 lead to dominant-negative inhibition of wild-type p53 function. *Proc. Natl. Acad. Sci. USA*. **99**, 2948–2953 (2002).

19. M. K. Lee *et al.*, Cell-type, dose, and mutation-type specificity dictate mutant p53 functions in vivo. *Cancer Cell*. **22**, 751–764 (2012).

20. A. J. McGahon *et al.*, Downregulation of Bcr-Abl in K562 cells restores susceptibility to apoptosis: characterization of the apoptotic death. *Cell Death Differ*. **4**, 95–104 (1997).

21. K. Chylicki *et al.*, p53-mediated differentiation of the erythroleukemia cell line K562. *Cell Growth Differ.* **11**, 315–324 (2000).

22. A. M. A. Di Bacco, T. G. Cotter, p53 expression in K562 cells is associated with caspase-mediated cleavage of c-ABL and BCR–ABL protein kinases. *Br. J. Haematol*. **117**, 588–597 (2002).

23. S. J. Baker, S. Markowitz, E. R. Fearon, J. K. Willson, B. Vogelstein, Suppression of human colorectal carcinoma cell growth by wild-type p53. *Science*. **249**, 912–915 (1990).

24. A. O. Giacomelli *et al.*, Mutational processes shape the landscape of TP53 mutations in human cancer. *Nat. Genet*. **50**, 1–12 (2018).

25. E. Kotler *et al.*, A Systematic p53 Mutation Library Links Differential Functional Impact to Cancer Mutation Pattern and Evolutionary Conservation. *Mol. Cell*. **71**, 178–190.e8 (2018).

26. T. N. Wong *et al.*, Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. *Nature*. **518**, 552–555 (2015).

27. S. Jaiswal *et al.*, Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N. Engl. J. Med.* **26**, 2488-2498 (2014).

28. G. Genovese *et al.*, Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N. Engl. J. Med.* **26**, 2477-2487 (2014).

29. S. Abelson *et al.*, Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature*. **559**, 1–23 (2018).

30. P. Desai *et al.*, Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nat. Med*. **24**, 1–12 (2018).

31.    I. Martincorena *et al.*, High burden and pervasive positive selection of somatic mutations in normal human skin. *Science*. **348**, 880–886 (2015).

32.    I. Martincorena *et al.*, Somatic mutant clones colonize the human esophagus with age. *Science*. **362**, 911–917 (2018).

33.    M. Adorno *et al.*, A Mutant-p53/Smad complex opposes p63 to empower TGFbeta-induced metastasis. *Cell*. **137**, 87–98 (2009).

34.    C. J. Di Como *et al.*, p63 expression profiles in human normal and tumor tissues. *Clinical Cancer Res*. **8**, 494–501 (2002).

35.    M. P. Kim, G. Lozano, Mutant p53 partners in crime. *Cell Death Differ*. **25**, 161–168 (2018).

36.    C. D. Richardson, G. J. Ray, M. A. DeWitt, G. L. Curie, J. E. Corn, Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).

37.    S. Heinz *et al.*, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*. **38**, 576–589 (2010).

38.    T. Derrien *et al.*, Fast computation and applications of genome mappability. *PLoS One*. **7**, e30377 (2012).

39.    A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. **29**, 15–21 (2013).

40.    B. Li, C. N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. **12**, 323 (2011).

41.    M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. **26**, 139–140 (2010).

42.    M. Ashburner *et al.*, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet*. **25**, 25–29 (2000).

43.    The Gene Ontology Consortium, The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res*. **47**, D330–D338 (2019).

44.    H. Mi *et al.*, PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*. **45**, D183–D189 (2017).

45.     A. Melnikov, P. Rogov, L. Wang, A. Gnirke, T. S. Mikkelsen, Comprehensive mutational scanning of a kinase in vivoreveals substrate-dependent fitness landscapes. *Nucleic Acids Res*. **42**, e112–e112 (2014).

46.     L. Brenan *et al.*, Phenotypic Characterization of a Comprehensive Set of MAPK1/ERK2 Missense Mutants. *Cell Rep*. **17**, 1171–1183 (2016).

47.     UK Monogenic Diabetes Consortium *et al.*, Prospective functional classification of all possible missense variants in PPARG. *Nat. Genet*. **48**, 1570–1575 (2016).

48.     J. Stewart-Ornstein, G. Lahav, Dynamics of CDKN1A in Single Cells Defined by an Endogenous Fluorescent Tagging Toolkit. *Cell Rep*. **14**, 1800–1811 (2016).

49.     J. Gao *et al.*, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. **6**, pl1–pl1 (2013).

50.     E. Cerami *et al.*, The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. **2**, 401–404 (2012).

51.     M. Xie *et al.*, Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **12**, 1472-1478 (2014).

52.     F. Zink *et al.*, Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood*. **130**, 742–752 (2017).

53.     R. Acuna-Hidalgo *et al.*, Ultra-sensitive Sequencing Identifies High Prevalence of Clonal Hematopoiesis-Associated Mutations throughout Adult Life. *Am. J. Hum. Genet*. **101**, 50–64 (2017).

54.     S. Jaiswal *et al.*, Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N. Engl. J. Med*. **377**, 111–121 (2017).

55.     C. M. Arends *et al.*, Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. *Leukemia*. **32**, 1–12 (2018).