A Systematic p53 Mutation Library Links Differential Functional Impact to Cancer Mutation Pattern and Evolutionary Conservation

Highlights
- Multiplexed quantification of the effect of thousands of distinct p53 mutations
- Differential mutation outcomes reflect functional evolutionary constraints
- Neutral SNPs may interact with acquired mutations to modulate phenotypic outcome
- Hotspot p53 mutations confer a relative competitive advantage in vivo

Authors
Eran Kotler, Odem Shani, Guy Goldfeld, ..., Debora S. Marks, Moshe Oren, Eran Segal

Correspondence
moshe.oren@weizmann.ac.il (M.O.), eran.segal@weizmann.ac.il (E.S.)

In Brief

Data Resources
GSE115072
A Systematic p53 Mutation Library Links Differential Functional Impact to Cancer Mutation Pattern and Evolutionary Conservation

Eran Kotler,1,2 Odem Shani,1 Guy Goldfeld,1 Maya Lotan-Pompan,2 Ohad Tarcic,1 Anat Gershoni,1 Thomas A. Hopf,3 Debora S. Marks,2 Moshe Oren,1,4,* and Eran Segal1,2,*

1Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 7610001, Israel
2Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel
3Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
4Lead Contact
*Correspondence: moshe.oren@weizmann.ac.il (M.O.), eran.segal@weizmann.ac.il (E.S.)

https://doi.org/10.1016/j.molcel.2018.06.012

SUMMARY

The TP53 gene is frequently mutated in human cancer. Research has focused predominantly on six major “hotspot” codons, which account for only ~30% of cancer-associated p53 mutations. To comprehensively characterize the consequences of the p53 mutation spectrum, we created a synthetically designed library and measured the functional impact of ~10,000 DNA-binding domain (DBD) p53 variants in human cells in culture and in vivo. Our results highlight the differential outcome of distinct p53 mutations in human patients and elucidate the selective pressure driving p53 conservation throughout evolution. Furthermore, while loss of anti-proliferative functionality largely correlates with the occurrence of cancer-associated p53 mutations, we observe that selective gain-of-function may further favor particular mutants in vivo. Finally, when combined with additional acquired p53 mutations, seemingly neutral TP53 SNPs may modulate phenotypic outcome and, presumably, tumor progression.

INTRODUCTION

The TP53 tumor suppressor gene, which encodes the p53 transcription factor, is the most frequently mutated gene in human cancer (Kandoth et al., 2013; Olivier and Taniere, 2011). In response to a variety of cellular stress conditions, p53 is activated to suppress transformation by inducing cell cycle arrest, DNA damage repair, senescence, or apoptosis (Bieging et al., 2014; Levine and Oren, 2009). In concordance with its pivotal role in suppressing tumorigenesis, mutations disrupting wild-type p53 (wtp53) function are extremely common in human cancers (Brosh and Rotter, 2009), with variable prevalence between tumor types (Blons and Laurent-Puig, 2003; Iacopetta, 2003; Peller and Rotter, 2003; Schuijer and Berns, 2003). Unlike other tumor suppressors, the majority of cancer-associated mutations in p53 are missense mutations residing in its DNA-binding domain (DBD) (Bouaoun et al., 2016; Olivier et al., 2010) and leading to loss of tumor suppressive activity and possible gain of novel oncogenic functions (reviewed in Oren and Rotter, 2010). In its wild-type (WT) form, p53 binds as a homo-tetramer to DNA response elements of its target genes and orchestrates gene expression patterns to cope with cellular stress. In contrast, the effects of mutant p53 (mutp53) on target gene expression are more complex (Weisz et al., 2007) and are thought to interfere with pivotal signaling pathways (Riley et al., 2008). Previous studies in yeast have shown that different mutp53 variants exhibit altered transactivation capacities towards wtp53 target genes in a mutant- and target-dependent manner (Kato et al., 2003; Resnick and Inga, 2003). This comprises a general reduction in transactivation capacity or an altered spectrum of regulated genes, including non-wtp53 target genes (Menendez et al., 2006; Resnick and Inga, 2003). Interestingly, the residual transcriptional activity of a particular mutant is not directly indicative of the extent of its tumor suppressive functionality or of its ability to induce apoptosis (Kadox et al., 2005). Moreover, some p53 mutants (“super-trans mutants”) exhibit increased transactivation potential towards particular targets when compared to wtp53 (Resnick and Inga, 2003). Thus, in order to determine the tumor-suppressive capacity of specific p53 mutations, the phenotypic impact of each mutation should be experimentally quantified. Such knowledge is of particular importance for the personalized treatment of cancer.

However, while the availability of patients’ genomic sequences constantly increases, our understanding of mutation-specific biological effects remains a limiting factor. Specifically, detailed studies assessing the effects of p53 mutations in human cells have been largely limited to the most prevalent “hotspot” mutations, which account for ~30% of cancer-associated mutations; this leaves the remaining ~70% mostly uncharacterized. Deep mutational scanning (Fowler and Fields, 2014) offers a high-throughput approach for revealing the consequences of genetic variation both in regulatory and coding sequences (Brenan et al., 2016; Keren et al., 2016; Majithia et al., 2016; Tewhey et al., 2016; Ulirsch et al., 2016; Weingarten-Gabbay et al., 2016). Yet, although the construction of a large-scale mutp53 library has recently been reported (Kitzman et al., 2016), the ability to comprehensively characterize the consequences of the p53 mutation spectrum remains a significant challenge. In this study, we created a synthetically designed library of ~10,000 p53 variants and measured the functional impact of each mutation in human cells in culture and in vivo. Our results highlight the differential outcome of distinct p53 mutations in human patients and elucidate the selective pressure driving p53 conservation throughout evolution. Furthermore, while loss of anti-proliferative functionality largely correlates with the occurrence of cancer-associated p53 mutations, we observe that selective gain-of-function may further favor particular mutants in vivo. Finally, when combined with additional acquired p53 mutations, seemingly neutral TP53 SNPs may modulate phenotypic outcome and, presumably, tumor progression.
phenotypic characterization of such p53 mutants has not been performed. To address this important knowledge gap, we devised a massively parallel mutational scan and measured the functional impact of thousands of distinct p53 DBD variants in vitro and in vivo. This provides the first comprehensive phenotypic catalogue of p53 mutations in human cells, quantifying the relative fitness effect of each mutation in culture and its selective advantage in vivo—presumably representing context-dependent gain-of-function. Our findings demonstrate tight sequence-fitness relationships underpinning evolutionary selection and recapitulate the prevalence and clinical impact of particular mutations in human tumors.

RESULTS

Massively Parallel Measurement of the Functional Impact of p53 Mutations

To quantitatively measure the anti-proliferative functional capacity of thousands of mutp53 variants, we designed a synthetic library of 9,833 unique, ~200-nucleotide-long, single-stranded DNA oligonucleotides encoding variations in the p53 DBD (Figure S1). To delineate the impact of mutations derived from cancer samples and of asymptomatic natural polymorphisms, the library design included: (1) nearly all DBD mutations occurring in 28,869 tumor samples (Bouaoun et al., 2016), most of which were previously unstudied; (2) all possible permutations at each hotspot codon (378 variants), allowing their in-depth characterization; and (3) combinations of naturally occurring SNPs (within the DBD) with additional DBD mutations, which aimed to unravel possible genetic interactions (1,139 variants). In addition, to systematically characterize the effect of mutations across the entire DBD, we created all single-nucleotide substitutions, deletions, and insertions (3,874 variants); all single amino acid substitutions requiring up to 2 nucleotide changes (3,480 variants); and dinucleotide transitions (e.g., CC > TT and GG > AA, 97 variants), as well as premature stop codons (304 variants) and in-frame single and double (consecutive) amino acid deletions at each position (363 variants).

To achieve accurate mutagenesis of the entire DBD (residues 102–292, 573 bp long), we generated 4 sub-libraries (labeled “A,” “B,” “C,” and “D”), covering consecutive 141–144 bp fragments of the DBD flanked by wt-p53 homologues regions (Figure 1A). These sub-libraries were separately amplified and cloned into lentiviral vectors, yielding a mutp53 coding sequence followed by an internal ribosome entry site (IRES)-driven enhanced green fluorescent protein (EGFP) reporter. This bi-cistronic cassette allows estimation of relative mutp53 expression levels based on EGFP intensity. Library-encoding viruses were used to transduce p53-null H1299 (non-small-cell lung cancer) cells at a low multiplicity of infection (MOI = 0.1) so that each transduced cell would express only a single mutp53 variant (Figures S2A–S2C). Infected cultures were sampled at 2, 6, 9, and 14 days post-infection (PI), and the relative abundance of each variant at each time point was assessed using next-generation sequencing (STAR Methods). Setting a minimum threshold of 200 reads per variant at the first time point (Figure S2D), we were able to capture the dynamics of 9,516 unique DNA sequence variants (97% of the designed variants) corresponding to 5,708 protein sequence variants. To piece together information on the entire DBD, data in each sub-library was normalized according to the dynamics of synonymous (“silent”) mutation variants compared across all sub-libraries (Figure S2E and STAR Methods).

To determine the relative changes in clonal abundance, we compared the fraction of reads corresponding to each sequence variant at 9 days PI to the fraction at 2 days PI (Figure 1B). Importantly, the relative abundance of all 571 synonymous sequence variants encoding wt-p53 in the library decreased much faster within this time frame than that of the 8,945 non-synonymous variants (mean depletion of ~6.55-fold versus ~1.45-fold for synonymous and non-synonymous, respectively; Mann-Whitney U, p < 10^-187). This is concordant with the well-documented anti-proliferative effect of bona fide wt-p53, as measured by flow cytometry and qRT-PCR in comparison to the entire library or to a p53-null control (Figures S2E–S2H). Furthermore, in contrast to all synonymous mutations, variants encoding the ten most prevalent p53 mutations across all tumor types (Bouaoun et al., 2016) were robustly retained in the population, demonstrating the ability of our system to recapitulate in vivo behavior and simultaneously map the phenotypic outcome of thousands of unique mutations (Figures 1C, S2I, and S3A). Similarly, variants encoding missense mutations in the 6 hotspot codons (averaged across all mutations in each codon) lacked a WT-like tumor-suppressive effect and were well retained (ANOVA, p < 10^-119), although they did not show a proliferative advantage over p53-null control cells. The latter might be because H1299 cells, which are naturally not “addicted” to mutp53, may require excessive mutp53 expression to elicit oncogenic gain-of-function (GOF), while a single integrated copy of our lentiviral vector only yields relatively low p53 levels in non-stressed cells.

To allow quantitative comparison between variants and improve measurement robustness, we took advantage of the repeated measurements at 6, 9, and 14 days PI and calculated a relative fitness score (RFS) for each variant based on its retention (or depletion) across these three measurements (STAR Methods). Comparing the RFS calculated for different DNA sequence variants encoding the same amino acid substitution verified the robustness of our measurements (Figure S3B; R = 0.89, p < 10^-100). To further assess this measure’s accuracy and reproducibility, we performed independent replications of time-course experiments. Reassuringly, RFS scores of sequence variants were highly correlated between biological replicates (Figure S3C; R = 0.94, p < 10^-100).

Unexpectedly, when averaging the effects of all DNA sequence variants encoding the same amino acid substitution (“protein variants”), a bimodal distribution is observed (Figure 1D); in this distribution, the great majority of p53 alterations can be discretely categorized as either retaining wt-p53 functionality (strongly depleted in the population) or abrogating it (stably retained). A similar separation is evident when variants are displayed according to their DNA sequence (Figures S3D and S3E).

To explore the generality of our findings, we examined the distribution of effects in a different cellular context. We performed an identical time-course experiment using sub-library
Figure 1. A High-Throughput Experimental System for Measuring the Effects of p53 Variations

(A) Experimental design: a library of 9,833 designed p53 sequence variants was synthesized on Agilent microarrays in 4 separate pools spanning consecutive parts of the DBD (labeled A, B, C, and D) and cloned into a lentiviral vector expressing p53 followed by an IRES-driven EGFP reporter. p53-null H1299 cells were infected with packaged viruses at MOI = 0.1 to ensure integration of a single variant per cell. Infected cells were sampled along a 14-day time course, and the relative abundance of each variant at each time point was assessed by next-generation sequencing, allowing extraction of a growth curve and a relative fitness score (RFS, calculated as described in STAR Methods) for each individual variant.

(B) Relative fraction of reads for each variant (n = 9,516) at 9 days (y axis) versus 2 days (x axis) post-infection of H1299 cells with the p53 library. Variants along the y = x diagonal retain stable relative abundance in the population. Parallel red diagonal lines represent a 2-fold increase or decrease in abundance between day 2 and day 9. Green dots represent synonymous sequence variants (encoding wtp53 amino acid sequence). Inset: distribution of fraction of reads (log, base 2) for synonymous (green) and non-synonymous (blue) mutations at 2d and 9d.

(C) Time course growth curves comparing the dynamics of all synonymous mutations to that of all non-synonymous mutations, mean of the 10 most abundant p53 amino-acid substitutions in the IARC database, and means of all non-synonymous mutations in each of the six commonly mutated hotspot codons. Plotted lines represent means ± STE at each time point post-infection. "n" denotes the number of different DNA sequence variants averaged to calculate each mean value.

(D) Distribution of RFS values across all protein sequence variants. Dashed line shows median RFS of all synonymous variants. n = 6,837. See also Figure S3D and Table S2.
Figure 2. The Pattern of Functional Effects of p53 Mutations Correlates with Protein Structural Domains and Evolutionary Conservation

(A) RFS of p53 variants carrying mutations in the DBD. For each codon (x axis; numbers relate to amino acid positions), the RFS of all single amino acid substitutions (one letter codes on the left), deletions (“0”), premature stop codons (“*”), and frameshifts (“F.S.”) are presented. The RFS of each individual substitution is depicted according to the color bar at the bottom, with red indicating high fitness (survival in the population) and blue indicating low fitness (preferential depletion from the population). The WT residue in each position is marked in yellow; missing data points are in grey. Orange bars on top show the prevalence (% of population) of each domain.

(B) Schematic representation of the protein structure with domains labeled.

(C) Correlation between residue evolutionary conservation and mean RFS of AA substitutions. The Pearson correlation coefficient (R) is 0.79 (p < 2x10^-41).

(D) Scatter plot showing the measured RFS against predicted functionality. The Pearson correlation coefficient (R) is -0.59 (p < 2x10^-230).
“D” in HCT116 colorectal carcinoma cells in which both endogenous wtp53 alleles had been knocked out using CRISPR-Cas9 (Figure S3E; generous gift of Prof. K. Vousden). We found high correlations between RFS measurements of sequence variants (R = 0.73, p < 10^{-20}) and protein variants (R = 0.81, p < 10^{-20}) in the two contexts (Figures S3E and S3H). At the same time, the functional effects of p53 mutations span a slightly more continuous spectrum in HCT116 than in H1299 cells, suggesting that a minority of variants exhibit context-dependent effects in these settings.

**The Functional Impact of Distinct p53 Mutations Correlates with Protein Structural Domains and Evolutionary Conservation**

We next examined how amino acid sequence variations affect the RFS as an indicator of wtp53-like anti-proliferative capacity. To this end we calculated, for each assayed amino acid substitution at each position along the DBD, the median RFS measured for all DNA sequence variants encoding the same particular substitution in H1299 cells (Figure 2A). Notably, regardless of position along the DBD, premature termination codons and frameshift mutations resulted in a similarly strong disruption of p53 functionality (Figure 2A, bottom rows in heatmap; Figure S3E). In contrast, the effects of substituting or deleting a single amino acid were strongly dependent on its position within the DBD. Thus, the L1 loop (residues F113–T123) and most of the L2 loop (residues K164–C176 and C182–L194) are rather robust to alterations. Conversely, most mutations in the L3 loop (residues M237–P250) and the specific residues involved in coordination of zinc binding (R175, C176, H179, C238, and C242) compromise p53 functionality, which is in line with the documented importance of the zinc ion for the thermodynamic stability of the DBD (Bullock et al., 1997; Duan and Nilsson, 2006). Furthermore, hierarchical clustering of the relative fitness scores across the DBD (Figures 2B and S4A) grouped amino acids with similar biochemical properties close to one another. Thus, valine co-clustered with isoleucine and leucine while aspartate co-clustered with glutamate and phenylalanine co-clustered with tryptophan and tyrosine. Hence the effects of mutations on relative fitness capture the sequence-structure-function relationships in p53. Notably, as observed for specific protein variants (Figure 1D), the majority of codons also segregated into two major groups that display opposing phenotypic responses to mutagenesis (Figure S4B).

We then calculated an evolutionary conservation score (ECS; STAR Methods) for each residue along the DBD from a multiple sequence alignment of 1,887 homologous sequences, which correspond to 246 non-redundant sequences when clustering sequences that are more than 80% identical. Comparing ECS values with codon mean RFS measurements revealed a strong correlation (Figures 2A and 2C; R_{s} = 0.79, p < 2 \times 10^{-11}), highlighting the tight coupling between p53 protein sequence evolutionary conservation and vulnerability to functional alteration. Consequently, we attempted to blindly predict the effects of amino acid substitutions on protein functionality using a statistical model of sequence variation in the alignment based on evolutionary bias towards or away from specific residues at each position (Figure 2D; STAR Methods). This unsupervised model, which accounts for site-specific amino acid constraints alone and does not explicitly model a specific phenotypic functionality, showed a correlation with our experimental measurements (R_{s} = −0.59, p < 2 \times 10^{-28}), suggesting that RFS reflects a p53 functionality that is under evolutionary selection. Despite this correlation, some variants that were predicted to retain at least partial wtp53 functionality exhibited complete loss of anti-proliferative activity. This discrepancy may suggest the existence of additional context-dependent functionalities that are not evident in unstressed H1299 cells, which would be in line with p53’s involvement in multiple non-redundant processes (Kakudo et al., 2005; Pfister and Prives, 2017).

**Residues within the DBD exhibit differential robustness to modification**

Closer examination of the mutational effects within specific codons revealed 3 distinct response patterns of codons to mutations; (1) positions highly susceptible to mutation (i.e., substitution to nearly any amino acid abolished anti-proliferative p53 functionality), including the hotspot codons Q245, R248, R249, and R273 (Figures 3A and S4C); (2) positions resilient to mutations, tolerating practically all substitutions (with the frequent exception of proline) without losing p53 functionality (Figures 3B and S4D); and (3) codons in which a continuous phenotypic spectrum is observed, with mutation outcome largely depending on the specific substitution (Figures 3C and S4E). Interestingly, the latter group includes the hotspot residues R175 and R282, extending earlier observations (Ory et al., 1994). Altogether, our findings demonstrate the merit of in-depth functional characterization of p53 mutations, even at hotspot positions.

Overlaying the relative evolutionary representation (i.e., the percent of species in which that particular amino acid is present at a given position) over the measured phenotypic effect for each of the substitutions reveals that the mean relative representation of variants retaining wtp53 functionality is dramatically higher than in non-functional variants (Figures 3A–3D; Student’s t test, total mutated cases) of somatic mutations in each indicated residue across all tumor types (Bouaoun et al., 2016); asterisks mark major hotspots. Grey bars depict the evolutionary conservation score of each position across 1,887 homologous sequences (see STAR Methods); black bars represent mean RFS (± STD) for each codon (mean across each heatmap column). Bottom: linear scheme of major structural motifs in the DBD. Yellow rectangles denote residues engaging the zinc ion. See also Table S3.

(B) Same data as in (A) hierarchically clustered in both dimensions, grouping together amino acid positions with similar robustness to modification. Note that most positions (columns) are either tolerant to mutations (predominantly blue) or highly sensitive (predominantly red). This is concordant with the conservation scores of the codons (grey bars). See also Figure S4.

(C) Mean RFS, calculated for each amino acid position across all missense mutations at that position and plotted against the evolutionary conservation score of that position (0, completely variable; 1, fully conserved), n = 191.

(D) Measured RFS of each amino acid substitution plotted against predicted effect of its substitution on p53 functionality, as derived from a statistical model of evolutionary sequences (log-odds ratio of mutant and WT sequence probabilities, WT = 0; see STAR Methods). Yellow diamond represents wtp53. n = 2,990.
Figure 3. Residues within the DBD Exhibit Different Patterns of Robustness to Modification

(A–C) For each indicated p53 codon, we show the median RFS measured for all assayed amino acid substitutions at that position (blue, decrease in relative abundance, presumably retaining wt-p53 activity; red, relative enrichment, presumably due to loss of wt-p53 function). Grey bars on top indicate relative representation (% of sequenced species in which the indicated amino acid is present at that particular position, according to ConSurf (Ashkenazy et al., 2016) multiple sequence alignment). The amino acid occupying that position in wt-p53 is indicated at the bottom. See also Figure S4 C–S4E.

(D) p53 variants were divided by their RFS into two groups: retaining wt-p53-like functionality (blue) or dysfunctional (red). Threshold for separation was set to $1/C_0$ (equivalent to 2-fold depletion). Bars = % relative representation (mean ± STE) for each group, calculated as in (A–C). Student’s t test $p < 10^{-38}$; $n = 1,152$ in red group, $n = 2,005$ in blue group.

(E) Mean RFS for each amino acid position superimposed over the p53 structure (a monomer bound to DNA). Blue and red colors represent tolerant (low mean RFS) and susceptible (high mean RFS) positions, respectively.
Thus, the functional impact of mutations in human cells faithfully reproduces the constraints that shape the DBD sequence during evolution. Notably, the differences in robustness of codons to modification are concordant with known p53 structure-function dependencies: superimposing the mean RFS of each codon on the protein’s 3D structure shows that residues positioned in proximity to the DNA are generally more functionally vulnerable (Figure 3E). Together, these strong associations between our functional measurements, conservation, and structure position canonical anti-proliferative p53 capacity as a pivotal property under strong evolutionary selection.

**Mutation Prevalence in Human Tumors Is Correlated with Loss of Anti-proliferative Capacity**

Assessing the contribution of specific mutations to cancer features is key to patient-specific tailoring of treatment. We therefore asked whether the relative fitness effects measured in vitro correspond with the prevalence of particular p53 mutations in human tumors. Reassuringly, this analysis (Figure 4A) revealed that p53 mutation prevalence across all tumor types protein sequence alterations achievable by a single A-G or C-T transition, the correlation between mutation prevalence and loss of functionality increases (Figure 4B, $R_\beta = 0.52$, $p < 2 \times 10^{-38}$). Fitting a logistic function to the data managed to explain a large fraction of the variability ($R^2 = 0.49$) in relative fitness of transition mutants on the basis of clinical prevalence. The sigmoidal relationship—and the apparent separation of the most prevalent mutations without a further increase in relative fitness—suggest that additional explanations underlie the high prevalence of those hotspot mutations. These may include oncogenic effects not captured by our in vitro assay or a mechanistic tendency towards accrual of mutations at those sites.

**p53 Mutations that Disrupt WT-like Functionality In Vitro Span a Broad Phenotypic Spectrum In Vivo and May Exhibit Selective GOF**

Intrigued by the fact that, in both examined cellular systems, hotspot p53 mutations did not confer a growth advantage relative to the rest of the disruptive mutations, we hypothesized that such GOF is more likely to be revealed under selective

**Figure 4. Mutation Prevalence in Human Tumors Is Highly Correlated with Loss of wtP53 Anti-proliferative Effect**

(A) The RFS of each p53 variant is plotted against its relative abundance across all human tumors (IARC p53 database). Grey circles, missense mutations; red triangles, nonsense mutations; orange triangles, frameshift mutations; green diamonds, synonymous mutations. The ten most frequent hotspot mutations are indicated. Dashed triangle indicates mutations observed at low abundance despite having lost p53 functionality. Dashed horizontal line indicates median RFS of all synonymous variants. $n = 1,465$. See also Table S3.

(B) Same as in (A), but including only amino acid substitutions achievable by a single transition mutation. Colors signify the evolutionary conservation score of each residue; see color bar on the right. Triangular dots represent nonsense mutations. Dashed line represents a sigmoidal fit of the form $y = c / (1 + e^{x - x_0}) + y_0$, created using least squares residual minimization. $R^2 = 0.49$; $n = 526$. (Bouaoun et al., 2016) is positively correlated with RFS ($R_\beta = 0.4$, $p < 4 \times 10^{-57}$). Intriguingly, in apparent discordance with the overall picture, we also observed mutations that are rare in tumors despite having lost p53 functionality (dashed triangle, Figure 4A). Thus, the importance of such variants could not be deduced from mutation prevalence, emphasizing the necessity of direct functional measurement. Closer examination of these mutations revealed a marked enrichment in variants requiring more than a single nucleotide change or a purine-pyrimidine transversion. When excluding such variants from the analysis and retaining only
Figure 5. p53 Mutations that Disrupt WT-Like Functionality In Vitro Span a Broad Phenotypic Spectrum In Vivo and May Exhibit Selective GOF

(A) Schematic representation of the experimental workflow for measuring the effects of p53 variants in vivo. 9 days following infection with sub-library “D,” positive cells were sorted by EGFP expression. 2 weeks later, cells were transplanted into nude Foxn1nu mice. After ~3 weeks, tumors were extracted, and the relative enrichment/depletion of each variant was assessed (compared to its abundance in the injected pool). Only variants that disrupted p53 functionality in vitro (high RFS scores) were included in the analysis.

(B) Enrichment scores of sequence variants measured in two different tumors. A representative example is shown.

(C) Median in vivo enrichment scores versus in vitro RFS of non-disruptive sequence variants.
pressures operating in vivo within a growing tumor. We therefore FACS-sorted mutp53-positive HCT116 cells transduced with sub-library “D” (which includes 6 of the 10 most prevalent mutations), injected them subcutaneously into nude mice, and measured the relative enrichment of each variant in the formed tumors in comparison to its abundance in the initial injected pool (Figure 5A). Of note, this pool is already largely depleted of mutants that retain anti-proliferative capacity in vitro. Reassuringly, variant enrichment scores were highly correlated between individual tumors (Figures 5B, 5D, and S5A). In contrast, no correlation was found between in vivo enrichment score and in vitro RFS (Figure 5C), highlighting the difference in selective forces between these environments. Interestingly, p53 variants that equally disrupt wild-type p53 functionality in vitro were found to span a broad phenotypic spectrum in vivo (Figures 5E and S5F). Notably, enrichment scores of sequence variants encoding amino acid substitutions that are among the 10 most prevalent cancer-associated mutations were higher than those of truncating (nonsense or frameshift) and missense mutations encoding less prevalent amino acid substitutions (Figure 5F; Kruskal-Wallis, p < 10\(^{-15}\)), suggesting that they elicit in vivo GOF in both human cancer and experimental mouse tumors.

While traditional two-dimensional (2D) cell culture is commonly used in cancer research, cultivation of cells in three-dimensional (3D) space is believed to more faithfully approximate the in vivo conditions. In fact, hotspot p53 mutations have been implicated in disruption of tissue architecture by up-regulating the mevalonate pathway (Freed-Pastor et al., 2012) as well as in enabling anchorage-independent growth of tumor cells via fibronectin production in 3D culture (Iwanicki et al., 2016). Puzzled by our observation that hotspot p53 mutants were advantageous in vivo but not in 2D culture, we asked whether such selective advantage might be revealed in 3D culture. To preserve sufficient representation of library complexity in this experimental setting, we focused on a panel of 66 p53 variants of unknown significance (VUS). Such knowledge is very valuable—for example, in assessing the possible implications of a particular TP53 germline mutation identified by pre- or postnatal genetic testing.

Models for Predicting Cancer Prevalence and Functional Impact of Distinct p53 Mutations

To further elucidate the forces shaping the spectrum of cancer-associated p53 mutations, we trained a “random forest” learning algorithm to predict mutation prevalence in human tumors. We applied 200-fold cross-validation (CV) using 90% of our data to predict mutation abundance according to mutation type and outcome, position, the probability of occurrence of similar substitutions (genome-wide, along evolution), residue evolutionary conservation, and the phenotypic measurements we obtained in H1299 cells in vitro. This model predicted relative mutation abundance with an R = 0.72 (p < 10\(^{-100}\)) and R = 0.75 (p < 3 × 10\(^{-27}\)) on CV and 10% unseen test cases, respectively (Figures S7A and S7B). Importantly, the most contributing feature in this prediction task was our measured RFS score (Figure S7C), demonstrating the importance of direct systematic assessment of pan-mutation effects and underscoring the connection between loss of anti-proliferative capacity and prevalence in cancer.

Next, we attempted the complementary task of predicting variants’ RFS using the above mutation features, substituting all experimental measurements with mutation prevalence (Figures S7D–S7F). Once again, our trained model enabled accurate prediction of mutational outcome (R = 0.87, p < 10\(^{-100}\)) and R = 0.88, p < 2 × 10\(^{-47}\) in CV and on unseen test cases, respectively, providing a reliable estimation of the phenotypic effects of p53 variants of unknown significance (VUS). Such knowledge is very valuable—for example, in assessing the possible implications of a particular TP53 germline mutation identified by pre- or postnatal genetic testing.
the prognostic value of the RFS score (Figure 6A; Mann-Whitney U, p < 10⁻⁹ across all cancers). Of note, the age at tumor diagnosis (ATD) is similar in LFS family members with truncating TP53 mutations (frameshift or nonsense) and in members with functionally disruptive missense mutations, as predicted by our functional in vitro measurements. Yet, individuals with the six most prevalent hotspot mutations even exhibit a somewhat lower ATD (Figure 6B; Kruskal-Wallis p < 0.01), suggesting that
they elicit additional gain-of-function effects. Concordantly, tumors are observed at a younger age in LFS family members with mutations exhibiting high in vivo enrichment scores (in HCT116 xenografts) than in those bearing mutations that were depleted from mouse tumors (Figure 6C; Mann-Whitney U, p < 0.01).

The V217M SNP Affects the Phenotypic Outcome of Missense p53 Mutations

Finally, we took advantage of our assay to evaluate the significance of SNPs within the p53 DBD. Although highly conserved, the DBD nevertheless harbors several polymorphic variations. For example, V217M (rs35163653, resulting from a G > A transition), a non-synonymous validated SNP (Whibley et al., 2009), has been functionally studied in yeast, where it induced elevated expression of CDKN1A, BAX, and NOXA (Kato et al., 2003), and in human cells, where its transcriptional signature was indistinguishable from that of wt p53 (Wang et al., 2014). Another rare polymorphism within the DBD is R213R (rs1800372), which is caused by a synonymous A > G transition in exon 6. So far, these SNPs have not been associated with cancer risk (Ganci et al., 2011; Pilger et al., 2007; Sharma et al., 2014). Nevertheless, in the presence of secondary acquired mutations, these SNPs may affect cancer predisposition or aggressiveness. In search for combinatorial effects, we combined these SNPs with all single-base mutations residing in sub-library ‘C.’

DISCUSSION

Our findings provide a first comprehensive catalogue for the functional consequences of thousands of p53 DBD mutations in human cells, and they potentially hold important clinical implications. Our cell-based in vitro and in vivo measurements are highly indicative of the functional outcome of p53 mutations. Thus, tumor-associated p53 mutations retaining wt p53-like anti-proliferative functionality are rather unlikely to be driver mutations. In addition, our observations highlight the importance of direct measurement of mutation impact to determine the outcome of VUS and justify large-scale systematic scans aimed to broaden our understanding of mutation-driven phenotypic landscapes.

Of note, our assay quantifies the relative competitive advantage of cells that are fully identical to each other except for the presence of different p53 mutations. Undoubtedly, the competitive forces that operate during tumor progression are much more complex, encompassing cells that share the same TP53 mutation but differ in other aspects as well as a plethora of non-cell-autonomous effects of mutp53. Nevertheless, the fact that our relative fitness landscape is highly similar to the actual distribution of cancer-associated p53 mutants argues that we are able to capture an important determinant of the in vivo advantage conferred by TP53 mutations.

Conceivably, p53 mutations that retain wt p53-like functionality under specific conditions may nevertheless endow cancer-supportive phenotypes in a context-dependent manner. However, the high correlation of mutation phenotypic outcomes across cell types—and their concordance with human mutation prevalence, structural motifs, and evolutionary conservation—argues that biochemical features underpinning the anti-proliferative effects of p53 in this model are also seminal for its tumor suppressor activity as well as for its primordial biological functions. Furthermore, our results stress the context dependence of mutp53 GOF activity; while we did not observe a proliferative advantage for hotspot p53 mutations in vitro, these variants were significantly enriched in heterotopic mouse tumors, providing an

Figure 7. The Phenotypic Outcome of Missense Mutations Is Affected by the V217M SNP

(A) For each variant, we plotted its standardized relative fitness (Z scores) when the mutation is on a SNP R213R background as compared to a wt p53 background. n = 388. Variant type colors are shown at the bottom.

(B) Same as (A), but for SNP V217M. n = 385.
additional explanation for their high prevalence in human cancers. The fact that a similar enrichment was not observed in a 3D culture setting suggests that some factors or selective pressures that act in vivo were absent in both 2D and 3D cell culture models. These may include (but are not limited to) variable levels of nutrients such as glucose or essential amino acids, hypoxic conditions, or other metabolic factors. Plausibly, in an actual tumor microenvironment which involves interactions with the immune system and with additional factors, the variant-specific effects of p53 mutations might be further modified or augmented. Future studies should further expand our understanding of context-dependent mutational effects.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Cell culture
  - Animal models
- **METHOD DETAILS**
  - Plasmids
  - Quantitative real-time PCR (qPCR)
  - Western blot
  - Library synthesis and amplification
  - Construction of backbone plasmids
  - Library cloning into master plasmids
  - Lentivirus production and infections
  - Sample preparation for sequencing
  - In vivo assays
  - Genomic DNA extraction from mouse tumors
  - 3D culture relative proliferation assay
  - Deep sequencing
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Computational analysis
  - Mutation prevalence and RFS prediction models
  - Statistical analyses
- **DATA AND SOFTWARE AVAILABILITY**
  - Cell Resources

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and seven tables and can be found with this article online at https://doi.org/10.1016/j.molcel.2018.06.012.

ACKNOWLEDGEMENTS

We thank Shira Weingarten-Gabbay, David Zeevi, Michal Levo, Leeat Keren, Noa Furth, Yael Aylon, and Adina Weinberger for fruitful discussions and valuable inputs. We gratefully acknowledge Zohar Yakini for help with library synthesis, N. Rahm and A. Teneti for the lentiviral bicistronic constructs, Karen Vousden for providing p53-knockout HCT116 cells, Jonathan Weinstein for help with structural models, and Yoav Peleg for cloning assistance. This work was supported by grants from the European Research Council (614504, to E.S.) and from the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation (119494) and by a Center of Excellence grant (1779/11) from the Israel Science Foundation (to M.O.).

AUTHOR CONTRIBUTIONS

E.K. conceived the project, devised the experiments, designed the synthetic library, performed experiments, analyzed the data, and wrote the manuscript; G.G. performed experiments; E.K., O.S., and M.L.-P. cloned the libraries; T.A.H. and D.S.M. performed analyses; O.T. and A.G. performed in vivo experiments; E.S. and M.O. conceived the project and experiments, supervised the work, and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 16, 2017
Revised: March 23, 2018
Accepted: June 6, 2018
Published: June 21, 2018; corrected online: August 13, 2018

REFERENCES:


### REAGENT or RESOURCE | SOURCE | IDENTIFIER
--- | --- | ---
**Antibodies**
Mouse anti-p53: PAb1801 | Kind gift from David Lane | N/A
Mouse anti-p53: DO-1 | Kind gift from David Lane | N/A
Mouse anti-vinculin | Sigma-Aldrich | Cat#V9131; RRID: AB_477629
**Critical Commercial Assays**
Herculase II Fusion DNA polymerase | Agilent Technologies | Cat#600675
Phusion DNA polymerase | NEW ENGLAND BioLabs Inc. | Cat#M0530
T4 ligase | NEW ENGLAND BioLabs Inc. | Cat#M0202
Cpol (RsrII) restriction enzyme | Thermo Scientific | Cat#ER0741
SgsI (Asci) restriction enzyme | Thermo Scientific | Cat# ER1891
BamHI restriction enzyme | NEW ENGLAND BioLabs Inc. | Cat#R0136
NotI restriction enzyme | NEW ENGLAND BioLabs Inc. | Cat#R0189
DpnI restriction enzyme | NEW ENGLAND BioLabs Inc. | Cat#R0176
**FastAP Alkaline Phosphatase** | Thermo Scientific | Cat# EF0651
**E. cloni 10G electro-competent cells** | Lucigen | Cat#LC-60117-2
**Electroporation Cuvettes, 0.1 cm gap** | Bio-Rad | Cat#1652083
**KAPA HiFi Hot Start Ready Mix PCR Kit** | Novagen | Cat#KK2601
**SPRI beads Agencourt AMPure XP** | Beckman Coulter | Cat#A63881
T4 Polynucleotide Kinase | NEW ENGLAND BioLabs Inc. | Cat#M0201L
T4 DNA Polymerase | NEW ENGLAND BioLabs Inc. | Cat#M0203L
Klenow Fragment (3'-5' exo-) | NEW ENGLAND BioLabs Inc. | Cat#M0212L
**High Sensitivity D1000 ScreenTape** | Agilent Technologies | Cat#5067-5584
**Lambda Exonuclease** | Epicentre | Cat#LE032K
**Qubit dsDNA HS Assay kit, 500 Assays** | Invitrogen | Cat#Q32854
**NextSeq 500/550 Mid Output v2 kit (300 cycles)** | Illumina | FC-404-2003
**Miseq Reagent Micro Kit ,v2 (300 cycles)** | illumina | MS-103-1002
jetPEI DNA transfection reagent | Polyplus Transfection | Cat#101-10N
**Polybrene** | Sigma | Cat#AL-118
**RNaseA** | Sigma | Cat# R5500
**GelStar Nucleic Acid Gel Stain 10,000x concentrate** | LONZA | Cat#50535
**Nucleospin Gel and PCR Clean-up kit** | MACHEREY-NAGEL | Cat#740609
**QIAquick PCR Purification Kit** | QIAGEN | Cat#28104
**MinElute PCR Purification Kit** | QIAGEN | Cat#28004
**DNeasy blood and tissue kit** | QIAGEN | Cat#69504
**Gel extraction kit** | QIAGEN | Cat#28704
**NucleoBond Xtra maxi kit** | MACHEREY-NAGEL | Cat#740414
**Plasmid DNA mini kit** | RBC BioScience | Cat#YPD100
**REDEExtract-N-Amp PCR ready mix** | Sigma | Cat#R4775
**Fast SYBR Green Master Mix** | Thermo Scientific | Cat#4385614
**Spheroid microplate, 96-well, Ultra-low attachment** | Corning | Cat#4515
**CellCarrier spheroid plates, 96-well, Ultra-low attachment** | PerkinElimer | Cat#6055330

(Continued on next page)
Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Moshe Oren (moshe.oren@weizmann.ac.il) in accordance with the Weizmann Institute of Science’s material transfer agreement (MTA).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture
Human embryonic kidney cells 293T (HEK 293T) were cultured in Dulbecco’s modified Eagle’s medium ( Biological Industries, Beit-Haemek, Israel (BI) supplemented with 10% heat-inactivated fetal bovine serum (HI-FBS, BI) and 1% penicillin and streptomycin (P.S., BI). H1299 human lung carcinoma cells were cultured in RPMI1640 medium (BI), supplemented with 10% HI-FBS and 1% P.S (BI). p53 CRISPR knock-out HCT116 human colorectal carcinoma cells were a kind gift from Karen Vousden. These cells were cultured in McCoy 5A medium, supplemented with 10% HI-FBS, 1% L-Glutamine 200mM (BI) and 1% P.S. All cells were kept at 37°C in a humidified atmosphere containing 5% CO₂ and were frozen in freezing medium (90% HI-FBS + 10% dimethyl sulfoxide (DMSO, Sigma). Trypsin-EDTA solution C (BI) was used to detach cells from culture dishes.

Animal models
Procedures involving animals were approved by the Animal Ethics Committee of the Weizmann Institute (Institutional Animal Care and Use Committee number 35960617-1) and conformed to the guidelines of the Israel Council for Experiments in Animals. Athymic nude male mice (Hsd: Athymic Nude-Foxn1™) were ordered from Harlan Biotech. At time of injection, mice were 8-10 weeks old.

METHOD DETAILS

Plasmids
The pPRIG plasmid (pPRIG-Hd-HA-Red (Martin et al., 2006)) was kindly provided by the Pognonec lab (Université de Nice Sophia Antipolis, Nice, France). pEF1_EMVC_ (pSIN.EF1.cPPT.mRFP.IRESEMCV.eGFP.WPRE (Kazadi et al., 2008)) was a gift from A.
Telenti (The Institute of Microbiology of the University Hospital Center, Lausanne, Switzerland), pMDL, pVSV-G, and pRSV-Rev helper plasmids for lentivirus packaging were kindly provided by S. Lev (Weizmann Institute of Science, Israel).

Quantitative real-time PCR (qPCR)
DNA was purified using DNeasy blood and tissue kit (Qiagen) and qPCR was performed on a StepOne real-time PCR machine (Applied Biosystems) using SYBR Green PCR master mix (Thermo Scientific). Standard curve values for each amplicon were measured and the relative quantity in each sample was normalized to an intergenic region upstream of the KCNA4 gene. The following primers were used to assess the relative abundance of p53-positive cells: Fw- CTGTGCACTGTTGGGTATTG and Rv-CCAAATACTCCACAGCAAAATTC, and for the intergenic normalization region: Fw- TTTTCCCCATCTGTGGCT and Rv-TCTCCAGCTGTCAACAACCT.

Western blot
Immunoblot analysis was performed as previously described (Hoffman et al., 2014). Antibodies used were: p53 (mixture of DO1 + PAb1801); vinculin (Sigma).

Library synthesis and amplification
Initial library synthesis and amplification were based on a protocol previously used for yeast promoter libraries (Sharon et al., 2012). Pools of fully-designed ~200-residue long single-stranded DNA oligonucleotides were obtained from Agilent Technologies (Santa Clara, CA). To achieve accurate mutagenesis of the entire DBD (573bp long, total of 9,833 sequence variants), the complete library is composed of 4 sub-libraries, each covering a different 141-144bp fragment of the DBD ("DBD-A" covers residues 102-149, "DBD-B": 150-197, "DBD-C": 198-245, and "DBD-D": 246-292). Each modified segment contains common wtp53-homologous sequences (at least 20 nucleotides long) at both ends, to enable initial PCR amplification and restriction-free (RF) cloning into the vector (described below). To avoid non-specific concatamerization due to overlapping sequences, these 4 sub-libraries were obtained in two separate pools, each covering sequence variations in non-overlapping sub-fragments: parts A & C together (203 nucleotides-long), and parts B & D together (197 nucleotides-long).

Libraries were synthesized using Agilent’s on-array synthesis technology (Cleary et al., 2004; LeProust et al., 2010), and provided as DNA oligo pools in two separate tubes (10pmol). Each pool was dissolved in 200μl Tris-ethylenediaminetraacetic acid (Tris-EDTA), creating solutions of 3.35 and 3.25 ng/μl of A&C and of B&D, respectively. An aliquot of each library was diluted (1:100 and 1:50 dilutions for A&C and B&D, respectively), and used as template for PCR amplification of each of the 4 sub-libraries. To reduce PCR bias, 24 identical reactions were performed in parallel for each sub-library. Each 50μl reaction tube contained 5μl of library template, 10μl of 5×Herculase II reaction buffer, 5μl of 2.5mM deoxynucleotide triphosphate (dNTPs) each, 10μl of 10mM forward (Fw) primer, 10μl of 10μM reverse (Rv) primer and 2μl Herculase fusion DNA polymerase (Agilent Technologies, #600679). PCR parameters used were: 95°C for 1 min, 14 cycles of 95°C for 20s, and 68°C for 80s, each, and finally one cycle of 68°C for 4min. Primers used:

DBD-A Fw- TGTCATCTTGTCCCTTCCAGAAAA, Rv-ATGGGCCGAGCGGCAGGT; DBD-B Fw-CTGTGCACTGTTGGGTATTG and Rv-CCAAATACTCCACAGCAAAATTC; DBD-C Fw-CCCTCCTCAGCATCTTATCCGAGT, Rv-AGGATGGGCCTCCGGTT; DBD-D Fw-TGTGTAAAGCGTCTCCATGTTG, Rv-GACGCTGTTGGAAGCTT.

Products from all 24 identical reactions were pooled together and used as template for non-specific fragments by electrophoresis on a 2.5% agarose gel stained with GelStar (LONZA), extracted from the gel, and purified using a gel extraction kit (Nucleospin).

Construction of backbone plasmids
As a cloning intermediate for efficient insertion of the libraries into plasmids, we used a pPRIG-wtp53-EMCV-EGFP backbone (~6.5kb) created by substituting the dsRed sequence of pPRIG-Hd-HA-Red (Martin et al., 2006) with wtp53 coding sequence. Thus, we PCR-amplified a wtp53 coding sequence from pC53-SN3 (Baker et al., 1990) using the primers Fw- GATGTCATGGATCC and Rv-ATGGCGCGGAGCCGCAGTT and Rv-GTACTGATGCGGCCGCTCAGTCTAGTGCGCAGGCCTTCTC adding 5’ BamHI and 3’ NotI restriction sites. PCR was performed using Kapa HiFi Polymerase (KAPA Biosystems). Products were purified using a PCR purification kit (Qiagen), digested with BamHI and NotI (New England Biolabs, NEB) for 75min at 37°C. 4μg of pPRIG-dsRed-EMCV-EGFP were digested with BamHI and NotI (NEB) for 1h at 37°C to remove the dsRed sequence. Digested amplicon and plasmid were separated from nonspecific fragments by electrophoresis on a 1% agarose gel stained with GelStar (LONZA), ligated using T4 ligase (NEB) for 2h at 24°C. Ligated plasmids were transformed into Escherichia coli (HIT-DHS2, RBC Bioscience) by heat shock, positive colonies were grown in Luria broth (LB) media, and the plasmids were purified using a plasmid mini-kit (RBC BioScience).

For cloning of the p53 libraries into lentiviral vectors, we first cloned a pEF1a-wtp53-EMCV-EGFP master plasmid by substituting the mRFP sequence of pEF1a-wtp53-EMCV-EGFP backbone (C14) from nonspecific fragments by electrophoresis on a 1% agarose gel stained with GelStar (LONZA), ligated using T4 ligase (NEB) for 2h at 24°C. Ligated plasmids were transformed into Escherichia coli (HIT-DHS2, RBC Bioscience) by heat shock, positive colonies were grown in Luria broth (LB) media, and the plasmids were purified using a plasmid mini-kit (RBC BioScience).

For cloning of the p53 libraries into lentiviral vectors, we first cloned a pEF1a-wtp53-EMCV-EGFP master plasmid by substituting the mRFP sequence of pEF1a-wtp53-EMCV-EGFP backbone (~6.5kb) created by substituting the dsRed sequence of pPRIG-Hd-HA-Red (Martin et al., 2006) with wtp53 coding sequence. Thus, we PCR-amplified a wtp53 coding sequence from pC53-SN3 (Baker et al., 1990) using the primers Fw- GATGTCATGGATCC and Rv-ATGGCGCGGAGCCGCAGTT and Rv-GTACTGATGCGGCCGCTCAGTCTAGTGCGCAGGCCTTCTC adding 5’ BamHI and 3’ NotI restriction sites. PCR was performed using Kapa HiFi Polymerase (KAPA Biosystems). Products were purified using a PCR purification kit (Qiagen), digested with BamHI and NotI (New England Biolabs, NEB) for 75min at 37°C. 4μg of pPRIG-dsRed-EMCV-EGFP were digested with BamHI and NotI (NEB) for 1h at 37°C to remove the dsRed sequence. Digested amplicon and plasmid were separated from nonspecific fragments by electrophoresis on a 1% agarose gel stained with a gel extraction kit (Nucleospin), ligated using T4 ligase (NEB) for 2h at 24°C. Ligated plasmids were transformed into Escherichia coli (HIT-DHS2, RBC Bioscience) by heat shock, positive colonies were grown in Luria broth (LB) media, and the plasmids were purified using a plasmid mini-kit (RBC BioScience).

For cloning of the p53 libraries into lentiviral vectors, we first cloned a pEF1a-wtp53-EMCV-EGFP master plasmid by substituting the mRFP sequence of pEF1a-wtp53-EMCV-EGFP backbone (~6.5kb) created by substituting the dsRed sequence of pPRIG-Hd-HA-Red (Martin et al., 2006) with wtp53 coding sequence. Thus, we PCR-amplified a wtp53 coding sequence from pC53-SN3 (Baker et al., 1990) using the primers Fw- GATGTCATGGATCC and Rv-ATGGCGCGGAGCCGCAGTT and Rv-GTACTGATGCGGCCGCTCAGTCTAGTGCGCAGGCCTTCTC adding 5’ BamHI and 3’ NotI restriction sites. PCR was performed using Kapa HiFi Polymerase (KAPA Biosystems). Products were purified using a PCR purification kit (Qiagen), digested with BamHI and NotI (New England Biolabs, NEB) for 75min at 37°C. 4μg of pPRIG-dsRed-EMCV-EGFP were digested with BamHI and NotI (NEB) for 1h at 37°C to remove the dsRed sequence. Digested amplicon and plasmid were separated from nonspecific fragments by electrophoresis on a 1% agarose gel stained with a gel extraction kit (Nucleospin), ligated using T4 ligase (NEB) for 2h at 24°C. Ligated plasmids were transformed into Escherichia coli (HIT-DHS2, RBC Bioscience) by heat shock, positive colonies were grown in Luria broth (LB) media, and the plasmids were purified using a plasmid mini-kit (RBC BioScience).
cloning into pEF1α plasmid. The RF reaction was carried out using 1μl Phusion DNA polymerase and 10μl Phusion buffer (NEB), 1μl dNTPs (10mM), 2.5μl DMSO, 40ng template plasmid and 200ng of the mega-primer. 30 PCR cycles were performed (95°C for 30s, 60°C for 60s, 72°C for 8 min.) followed by a final elongation step of 10 min at 72°C. To digest remaining non-amplified plasmids, 1μl DpnI (NEB) was added to the reaction and following a 2hr incubation at 37°C, products were used for bacterial transformation as described above. Single colonies were picked and sequenced (Sanger sequencing at the Weizmann institute’s Life Sciences Core Facilities) for validation of the entire wtP53 sequence.

**Library cloning into master plasmids**

Purified amplified library fragments were used as large PCR primers for amplification of the entire vector in a RF cloning reaction (Geiser et al., 2001; Unger et al., 2010) so that each sub-library replaces the equivalent part of the wtP53 DBD sequence within the target (pPRIG) plasmid (Step 1). pPRIG-cloned libraries (each sub-library as a separate pool) were then transferred to the pEF1α lentiviral vector by conventional restriction-ligation cloning of the pS3 variants (Step 2).

In Step 1, 50μl RF reactions were carried out using 1μl Phusion DNA polymerase and 10μl Phusion buffer (NEB), 1μl dNTPs (10mM), 2.5μl DMSO, 20ng template plasmid and 10μl of gel-purified library amplicons. 30 PCR cycles were performed (95°C for 30s, 60°C for 60s, 72°C for 5min.) followed by a final elongation step of 7min. at 72°C. To digest remaining non-amplified plasmids (wtP53), 12μl from each product were incubated with 1.2ul Dpn1 (NEB) for 2hr at 37°C. Then, an additional 1.2ul Dpn1 were added and samples were incubated for another 2hr at 37°C.

RF products were purified using PCR purification kit (Qiagen) and transformed into E. coli 10G electro-competent cells (Lucigen). Each bacteria tube was divided into seven aliquots (25μl each) for electroporation with 2μl of plasmids using 0.1cm gap cuvettes (Biorad) according to the manufacturer’s protocol. All transformation tubes from each sub-library were pooled together and seeded on LB agar (200 mg/ml ampicillin) 15cm plates. 16 hours after transformation, to ensure adequate preservation of library complexity, we collected a total of 73000, 76500, 113500 and 122300 colonies, representing a sampling of 44-, 44-, 38- and 35-fold over designed library sizes of sub-libraries A, B, C and D, respectively. To assess the percentage of remaining wtP53 colonies was assessed. For validation of RF products, we performed colony PCR on 24 random colonies from each sub-library. Each reaction was performed in 20μl and contained a colony picked from the plate, 8μl of double distilled water (DDW), 10μl of REDExtract-N-Amp PCR ready mix (Sigma) and 1μl of each primer (Fw- GAGCCGAGTCAG ATCTCTAG; Rv- GAGCTCGTGGTGGGCTGC). Collected colonies were scraped from plates into LB medium, and pooled plasmids were purified using a NucleoBond Xtra maxi kit (Macherey Nagel).

For Step 2, we PCR amplified each p53 sub-library from the intermediate pPRIG plasmids (Step 1 products) using primers that add restriction sites (5’ Ascl, 3’ RsrII) for ligation into the pEF1α vector (Fw-ACGTTAACGGCGCGCCACTGCGATGGAGGAGCGAG TCGAGATC; Rv-AGGCTAGTCCGCAACATCCAGATGCAACGCCTGATGTACGCTGATGCAGCGCCTTCTG). To reduce vector representation bias, for each sub-library we performed 14 identical PCR reactions using 25μl 2xKapa HiFi ready-mix (KAPA Biosystems), 50ng of pPRIG-library template, 0.2μM from each primer and DDW to a final volume of 50μl. The parameters for PCR were 95°C for 10 min, 9 cycles of 98°C for 10 sec, 68°C for 30 sec and 72°C for 1.5 min and a final elongation of 72°C for 5 min. The fourteen reactions were pooled together, purified using 3 QiAquick PCR purification columns (Qiagen), and products were pooled again after the elution step. Purified amplicons were then digested with CpoI (RsrII, Catalog No. ER0741) and SgsI (AscI, Catalog No. ER1891) (Thermo Fisher Scientific, Fermentas) restriction enzymes. Digestion reaction mixtures contained: 12 μl Fast Digest buffer (Fermentas), 3μg of the purified library, 7.5μl CpoI, 2.4μl SgsI and DDW up to a total volume of 120μl. The mixture was incubated for 2hr at 37°C, followed by 20 min inactivation at 65°C. Target vector (pEF1α-wtP53-EMCV-EGFP) digestion was performed using the same enzymes, in a reaction-mixture containing: 18μl Fast Digest buffer, 15μg of the plasmid library, 9μl of each enzymes and DDW up to a total volume of 180μl. The mixture was divided into three tubes and incubated for 2.5 hours at 37°C, followed by 20 min inactivation at 65°C. Then, alkaline phosphatase (FastAP, Thermo Fisher Scientific) was added to each tube (3μl of Fast AP buffer, 3μl of FastAP enzyme and DDW up to a total volume of 30μl). The mixture was incubated for an additional 30 min at 37°C, followed by 20 min inactivation at 65°C.

Restriction-digested libraries and plasmid were separated by electrophoresis on a 1.5% agarose gel stained with ethidium bromide. Fragments of the correct size were excised from the gel and samples were purified first using Qiagen Gel extraction kit (cat #28704) and then with a Gel and PCR clean-up purification kit (NucleoSpin, cat #740609). Next, library-vector ligations were performed using 1μl of Lucigen ligase, 10xLucigen buffer, molar ratio of 1:1 (vector: library) and DDW to a final volume of 10μl. Ligation products were transformed into E.coli (Lucigen) with 2μl of ligation mix, as described in Step 1, collecting a total of 2.75x10^8, 1.53x10^8, 3.57x10^8 and 3.94x10^8 colonies from libraries A, B, C and D, respectively. Colony PCRs were performed for validation as described above, and plasmids were extracted using NucleoBond Xtra Maxi kit (cat #740414.10).

Cloning of the DBD-D mini-library was performed as described above for the large scale library, with a minimum of 4000 colonies collected after each transformation step, to ensure adequate preservation of library complexity.

**Lentiviral production and infections**

For lentiviral packaging, 2.1x10^6 HEK293T cells were seeded on 10cm plates pre-coated with poly-L-lysine 0.001% (Sigma), incubated for 20 min, and washed three times with phosphate-buffered saline (PBS). 16hr later, cells were co-transfected with three
helper plasmids (5.2μg pMDL, 3.2μg pVSVG and 2μg pRSV-Rev) and 8μg of library-encoding plasmid. Transfections were performed using jetPEI DNA transfection reagents (Polyplus Transfection) according to the manufacturer’s instructions and medium was replaced after 6-8hr. Virus containing medium was collected at 48 and 72hr post-transfection, filtered with 0.45-μm filters (Merck Millipore). We computed the virus dilution required for multiplicity of infection (MOI) of 0.1 and repeated the infection protocol accordingly in large scale. To maintain variant representation and to control for lentiviral random genomic integration, the number of cells infected with each library was planned according to the number of designed variants in that library (1665, 1743, 2974 and 3451 in libraries A, B, C and D, respectively). Thus, we plated 6 x 10cm plates with 6x10⁵ H1299 cells 16hr prior to infection with each of libraries A and B, and 12 such plates for infection with libraries C and D. A total of ~3.6x10⁵ and ~7.2x10⁵ cells were infected with each of libraries A & B, and C & D, respectively, so that on average each designed sequence was independently integrated into >200 individual cells. The percentage of infected cells was verified 48hr post-infection by flow cytometry as described above. Independent biological replications of H1299 time-course experiments were performed as described above, for libraries A and B. For HCT116 infections, titers were calibrated as described for H1299 cells, and cells were plated on 8 x 10cm plates with 1.6x10⁶ cells/plate 16hr prior to infection. Cells were infected with library D in low multiplicity (MOI = 0.05). Infection of HCT116 cells with the mini-library was performed using 2 x 10cm plates as above, at MOI = 0.08.

Sample preparation for sequencing
To maintain the complexity of the input libraries, PCR reactions were carried out on a genomic DNA (gDNA) amount calculated to contain an average of at least 200 copies of each variant included in the assayed library. Thus, at each time point along the time course experiments, gDNA was purified from a minimum of 1x10⁷, 1.3x10⁷, 2.5x10⁷, and 2.5x10⁷ H1299 cells infected with libraries A, B, C and D, respectively. gDNA was purified using DNeasy blood and tissue kit (Qiagen). For each measurement, a two-step nested PCR was performed: First, the entire DBD region was amplified (performed identically for all 4 sub-libraries). Then, in each experiment the relevant sub-library DBD fragment was amplified using primers specific to the modified region in that sub-library. In the first step (performed in multiple tubes to include the required amount of gDNA), each reaction contained a total volume of 100μl with 10μg gDNA, 50μl of Kapa Hifi ready mix X2 (KAPA Biosystems) and 5μl of each (10mM) primer. The parameters for PCR were 95°C for 5min, 18 cycles of 94°C for 30s, 65°C for 30s, and 72°C for 30s, followed by one cycle of 72°C for 5 min. The primers used for this reaction were CTGAAGACCCAGGTCCAGATGAAG (Fw) and GGAGAGGAGCTGGTGTTGTTGG (Rv). In the second PCR step each reaction contained a total volume of 50μl with 2.5μl of the first PCR product (uncleaned), 25μl of Kapa Hifi ready mix X2 (KAPA Biosystems) and 2.5μl of each (10mM) primer. The parameters for PCR were as in the first step, performing 24 cycles. Primers used for this reaction included 5 random nucleotides at their 5’-end to increase sequence complexity and facilitate cluster calling during sequencing, and were specific for each of the 4 sub-libraries: sub-library A: Fw- NNNNNTGTCACTTCTTGACTTCCTCCAGAA, Rv- NNNNATGGGGCCGACCGCGGTT; sub-library B: Fw-NNNNNTGTCACTTCCTCCAGAA, Rv- NNNNATGGGGCCGACCGCGGTT; sub-library C: Fw-NNNNNTGTCACTTCCTCCAGAA, Rv- NNNNATGGGGCCGACCGCGGTT; sub-library D: Fw-NNNNNTGTCACTTCCTCCAGAA, Rv- NNNNATGGGGCCGACCGCGGTT. Alternatively, for sequencing runs including samples from various experiments or time points, this reaction was performed using alternatively barcoded Fw primers, employed to barcode the different experimental samples. These primers included 3 random nucleotides at their 5’ to increase sequence complexity, followed by a unique 4-nucleotide barcode (see Table S1 for sequences).

Amplicons were separated from nonspecific fragments by electrophoresis on a 1% agarose gel stained with EtBr, extracted from the gel and purified using a gel purification kit (Qiagen). Products were further cleaned using a MinElute PCR Purification kit (Qiagen) and eluted in 12ul DDW. Concentration was measured using a monochromator (Tecan i-control), and sample size and purity were assessed by Tape-station using a high-sensitivity D1K screen tape (Agilent Technologies). 50ng DNA were used for library preparation for next-generation sequencing, specific Illumina adaptors were added, and DNA was enriched by 14 amplification cycles by a protocol adopted from Blecher-Gonen et al. (Blecher-Gonen et al., 2013). Samples were reanalyzed by Tape-station prior to sequencing.

In vivo assays
9 days post infection with lentiviruses expressing sub-library “D”, 3x10⁶ EGFP-positive HCT116 cells were sorted using a Becton Dickinson FACSaria III cytometer, expanded in culture for a week and frozen down. Thawed cells were then expanded for an additional week in culture and injected subcutaneously into the back of athymic nude mice. Each mouse was inoculated with a total of 2x10⁶ cells (suspended in 100ul PBS). At the time of injection, 50x10⁶ cells were pelleted, their genomic DNA was purified as described above to serve as a time-zero reference for enrichment.
Genomic DNA extraction from mouse tumors

Primary tumors were dissected manually and flash-frozen in liquid nitrogen. Although each mouse was injected in a single location, in cases where more than a single clearly separated tumor was visible during dissection, each tumor was collected separately (labeled ‘T2a’, ‘T2b’, etc.). For DNA extraction, samples were thawed and divided to 2ml tubes each containing 25mg tissue. Genomic DNA was extracted using DNeasy blood and tissue kit (Qiagen) according to the manufacturer’s protocol. For lysis, each sample was suspended in 180ul ATL buffer and homogenized for 10 minutes in a Qiagen TissueLyser LT with a single 5mm stainless steel bead (Qiagen). Following removal of the bead, 20ul proteinase K were added and samples were incubated for 1 hour at 56°C. Then, after samples cooled down to room temperature, they were supplemented with 4ul RNase and incubated for 2 minutes at RT. To avoid bias due to tumor heterogeneity, DNA from the entire tumor was purified. Thus, all lysates originating from the same tumor were pooled together and purified using the kit’s columns and the manufacturer’s protocol. Amplicons of the DBD-D region and Illumina library preparations were then prepared as described above.

3D culture relative proliferation assay

Twenty days post-infection with lentiviruses expressing the mini-library (of sub-library “D”), 1.3x10^6 EGFP-positive HCT116 cells were sorted using a Becton Dickinson FACSAria III cytometer, expanded in culture for a week and frozen down. Thawed cells were then expanded for an additional week in culture, harvested in trypsin, filtered to create a single-cell suspension and seeded in ultra-low adherence spheroid plates (1000 cells/well in 200ul McCoy culture media). To preserve library complexity, each experimental replicate was composed of a 96-well plate, providing a factor of ~1000-fold over the designed library size. A total of 4 replicates were performed (two using Corning spheroid microplates and two using PerkinElmer Cell-Carrier spheroid plates). Spheroid formation and expansion was monitored using a Nikon eclipse Ti-S microscope. In parallel to spheroid seeding, 1.5x10^6 cells were plated in each of two (regular) 10cm plates for 2D growth. At time of seeding, 10^7 cells were pelleted and their genomic DNA was purified as described above to serve as a time-0 reference for enrichment. Spheroid plates were harvested at 7 and 9 days after seeding (two plates at each time-point), pooling together all spheroids in each plate. At each of these time-points, 5x10^6 cells were sampled from each 2D-culture plate for comparison. DNA purification and sample preparation were performed as described above.

Deep sequencing

Amplified DBD fragments were sequenced on an Illumina NextSeq-500 sequencer using NextSeq 500/550 Mid Output kits (300 cycles), producing paired-end reads in the length of 150nt. For time course measurements, we obtained a minimum of ~3.5x10^6 reads for each time point for libraries A-C and ~8.5x10^6 reads for library D (mean of ~6.5x10^6 reads per sample in libraries A and B, and ~9.5x10^6 in C and D). Sequencing of 3D culture experiments (mini-library) was performed on an Illumina MiSeq desktop sequencer using a MiSeq reagent kit (300 cycles) with a minimum of 200,000 reads per sample.

QUANTIFICATION AND STATISTICAL ANALYSIS

Computational analysis

Paired-end reads were merged using Usearch (http://www.drive5.com/usearch/) setting a minimal overlap of 80 and merge length of 180-220. Merged reads in each sample were mapped to the library design, requiring a perfect sequence match, and the number of reads corresponding to each sequence variant were counted. We required a minimal coverage of >200 reads for each sequence variant at the first experimental time point. This stringent threshold, together with the large number of cells infected with each library variant as described above, enabled to average out the effects of random lentivirus integration and reduce measurement noise. ~97% of the designed sequence variants (9,516 of 9,833) passed this threshold and were further used in the analysis.

As each of the 4 sub-libraries was measured separately, data in each sub-library was normalized according to the dynamics of synonymous mutation variants, compared across all sub-libraries. Thus, for each sequence variant, we calculated the log (base 2) fold-change (FC) at each time point and normalized the log-FC of all variants in each sub-library so that the mean log-FC across synonymous variants (encoding a wtp53 amino acid sequence) in that sub-library will equal the mean log-FC of all synonymous variants (across all libraries).

The relative fitness (RFS) score for each variant was calculated as:

\[ \text{RFS} = \text{Median}(R_{t_1}, R_{t_2}, R_{t_3}) \]

Where \( R_{t_i} \) is the enrichment/depletion of the variant at time point \( t_i \) :

\[ r_6 \] represents the fraction of reads corresponding to a variant at the given time point \( t_1, t_2 \) and \( t_3 \), represent sampling at 6, 9 and 14 days post-infection, respectively.
Similarly, an in vivo enrichment score for each variant in each tumor was calculated by:

\[ R_{iT} = \log_2 \left( \frac{r_{iT}}{r_{i0}} \right) \]

where \( r_{iT} \) represents the fraction of reads corresponding to a variant (i) in tumor T, and \( r_{i0} \) represents the fraction of reads corresponding to that variant at time 0 (i.e. at time of injection). We then calculated the median enrichment for each variant across all mouse tumors. Only variants with high RFS scores in cultured HCT116 cells (\( \geq 0.5 \)) were used in the in vivo analyses.

For the 3D culture experiment, enrichment scores for each variant in the mini-library in each sample (either in 2D or 3D culture) was calculated by:

\[ R_{ij} = \log_2 \left( \frac{r_{ij}}{r_{i0}} \right) \]

where \( r_{ij} \) represents the fraction of reads corresponding to a variant (i) in sample j, and \( r_{i0} \) represents the fraction of reads corresponding to that variant at time 0 (i.e. at time of seeding). Since variant enrichment scores within each growth condition were similarly correlated regardless of their harvesting time at 7 or 9 days, we regarded each 4 samples from the same growth condition as replicates and calculated the median enrichment for each variant in each growth condition (2D or 3D) across these replicates. Only missense mutants with high RFS scores in cultured HCT116 cells (\( \geq 0.5 \)) and for which an in vivo enrichment score was available were used in the analysis.

The evolutionary conservation score (ECS) measures how conserved each position is in a multiple sequence alignment of the protein family (Sander and Schneider, 1991). It is defined as the normalized entropy of the distribution of amino acid frequencies \( f_i \) in position i, i.e.

\[ ECS(i) = 1 + \sum_a f_i(a) \ln f_i(a) / \ln q \]

and sums over all amino acid characters a in the alignment including the gap, and q=21. ECS(i) ranges from 0 (completely variable) to 1 (fully conserved). Amino acid frequencies were obtained from a sequence alignment of 1887 homologous sequences obtained by a jackhammer search against the UniRef100 sequence database (5 iterations, E-value threshold: 1E-01). To reduce sequence redundancy when calculating the frequencies, the counts for each sequence were weighted by 1/m, where m is the number of sequences in the alignment that are at least 80% identical (redundancy-reduced number of sequences: 246) (Hopf et al., 2017). Relative evolutionary representation of amino acid residues was calculated using the ConSurf (Ashkenazy et al., 2016) tool using default parameters (homologous sequences taken from UniRef90 database, and filtered for sequence homology ranging between 35-95%). Relative representation is defined as the percent out of the 150 examined species in which that particular amino acid is present at a given position.

To predict the effects of individual amino acid substitutions from evolutionary sequences, a statistical model of the family sequence alignment was inferred (Hopf et al., 2017). Since there is only limited evolutionary sequence diversity in the family alignment (redundancy-reduced number of sequences: 246), we chose to infer a simple site-independent model rather than an epistatic model that considers amino acid dependencies between pairs of positions, as described previously (Hopf et al., 2017). This choice is supported by the observation that only 29 significant long-range evolutionary couplings between pairs of positions could be detected at a 90% probability cutoff when inferring an epistatic model (Toth-Petroczy et al., 2016).

Briefly, the independent model describes the probability of any amino acid sequence \( \sigma \) in the family by

\[ P(\sigma) = \frac{1}{Z} \exp \left\{ \sum_i h_i(\sigma_i) \right\} \]

with single-site amino acid constraints \( h_i(\sigma_i) \) capturing the preference for amino acid \( \sigma_i \) in position i. These parameters are inferred from the sequence alignment using l2-penalized maximum likelihood inference. The effect of a substitution \( \Delta E \) can then quantified by the log-odds ratio of the probabilities of the mutant and wild-type sequences under the model:

\[ \Delta E(\sigma^{mut}, \sigma^{wt}) = \log \frac{P(\sigma^{mut})}{P(\sigma^{wt})} \]

A score of 0 putatively corresponds to neutral substitutions, scores < 0 to deleterious substitutions, and scores > 0 to beneficial substitutions.

The structural model of p53 was created with PyMol using the 1TSR p53 structure downloaded from the Protein Data Bank (http://www.rcsb.org/pdb).
10% test set which was left aside and not used for model fitting and optimization. To tune model parameters and assess its performance on training data we used a 200-fold cross validation scheme. Model parameters were adjusted to maximize the Pearson correlation between predicted and measured values. These parameters were then used for fitting the model using the entire training data set and final prediction on unseen test data.

Model features include: mutation prevalence (Bouaoun et al., 2016), measured RFS; variant enrichment at 6, 9 and 14 days (log-FC of read fractions over the 2d time point) and relative abundance at 2 days; mean enrichment at 6, 9 and 14 days; position within the DBD; residue evolutionary conservation and residue percent variability; “independent model” prediction of protein functionality; ‘epistatic model’ prediction of protein functionality calculated using EVMutation (Hopf et al., 2017); minimal number of transitions and transversions required to achieve the given amino acid alteration; type of mutations (substitution, deletion, insertion or tandem base transitions) and their outcome (missense, nonsense, silent); and PAM250 and BLOSUM62 substitution matrix values. Feature importance was assessed by mean decrease in impurity, as implemented in scikit-learn (Pedregosa et al., 2011).

**Statistical analyses**

Statistical tests performed are reported in the Main text, Figures, and Figure legends. All performed statistical tests were two-sided. All error bars represent ±SEM (standard error of the mean), unless specifically noted otherwise. Heatmap hierarchical clustering was performed using a Euclidean distance matrix. To assess the difference between two groups of values that are distributed approximately normal, we used Student’s t test. When this was not the case we performed nonparametric tests: Mann–Whitney U test for independent samples, or Wilcoxon signed-rank test for matched samples. To assess the difference between the dynamics of different mutation groups (e.g. in Figure 1C) we performed a within-subjects two-way analysis of variance (ANOVA). For comparison between multiple groups of values that were not distributed normally (e.g. Figures 5F and 6B) we performed a one-way Kruskal-Wallis H test. Statistical analyses were performed using the scipy python package.

**DATA AND SOFTWARE AVAILABILITY**

**Data Resources**

Raw data has been deposited to NCBI GEO: GSE115072. Processed data are available in Tables S2, S3, S4, S5, S6, and S7.