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

A systematic p53 mutation library links functional impact to evolutionary conservation and cancer mutation pattern

Eran Kotler[1, 2], Guy Goldfeld[1], Maya Lotan-Pompan[2], Thomas A Hopf[3], Debora S Marks[3], Moshe Oren[1, *] and Eran Segal[1, 2, *]

[1]Department of Molecular Cell Biology, [2]Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel; [3]Department of Systems Biology, Harvard Medical School, Boston, MA, USA.

[*]Correspondence should be addressed to E.S. (eran.segal@weizmann.ac.il) or M.O. (moshe.oren@weizmann.ac.il).

The *TP53* tumor suppressor gene is very frequently mutated in human cancer. Research on p53 mutations has focused predominantly on six major "hotspot" codons which comprise only ~30% of cancer-associated p53 mutations. To systematically study the impact of a wide diversity of p53 mutations, we created a synthetically designed mutation library and measured the functional impact of ~10,000 DNA-binding domain (DBD) p53 variants, including nearly all patient-reported mutations, in human cells. Our results elucidate the selective pressure to maintain specific residues at particular positions in p53 throughout evolution. Furthermore, they highlight the differential outcome of distinct classes of p53 mutations in human patients, and suggest that loss of anti-proliferative functionality is a key selective force shaping the landscape of cancer-associated p53 mutations. Finally, we show that when combined with additional acquired p53 mutations, seemingly neutral SNPs within the DBD may modulate phenotypic outcome and presumably tumor progression.

The *TP53* tumor suppressor gene, encoding 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), leading to loss of tumor suppressive activity and possible gain of novel oncogenic functions (reviewed in (Oren and Rotter, 2010)). In its 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 capacity 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 its ability to induce apoptosis (Kakudo 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 has been largely limited to the most prevalent "hotspot" mutations, accounting for ~30% of cancer-associated mutations, leaving 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., 2015), phenotypic characterization of such p53 mutants has not been performed.

To address this important knowledge gap, we designed a synthetic library of 9,833 unique ~200 nucleotide-long single-stranded DNA oligonucleotides encoding variations in the p53 DBD (Fig.

S1). To delineate the impact of mutations derived from cancer samples and of asymptomatic natural polymorphisms, the library design included: (i) nearly all DBD mutations occurring in 28,869 tumor samples (Bouaoun et al., 2016), most of which were previously unstudied; (ii) all possible permutations at each hotspot codon (378 variants), allowing their in-depth characterization; and (iii) combinations of naturally occurring SNPs (within the DBD) with additional DBD mutations, 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, 573bp long), we generated 4 sub-libraries (labeled "A", "B", "C" and "D"), covering consecutive 141-144bp fragments of the DBD flanked by wtp53-homologues regions (Fig. 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 bicistronic 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 low multiplicity of infection (MOI=0.1), so that each transduced cell would express only a single mutp53 variant (Fig. S2a-c). 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 (Methods). Setting a minimum threshold of 200 reads per variant at the first time point (Fig. 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 (Fig. S2e and 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 its fraction 2 days PI (Fig. 1b). Importantly, the relative abundance of all 571 synonymous sequence variants in the library, encoding wtp53,

decreased within this time frame much faster 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* wtp53, as measured by flow cytometry and qRT-PCR in comparison to the entire library or to a p53-null control (Fig. S3a-c). 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 (Figs. 1c, S3d, S4). 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 $p<10^{-10}$), albeit not showing a proliferative advantage over p53-null control cells. The latter might be because H1299 cells, naturally not "addicted" to mutp53, may require excessive mutp53 expression to elicit oncogenic gain-of-function, 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 median retention (or depletion) across these three measurements ($\log_2$ fold enrichment compared to first time point). Comparing the RFS calculated for different DNA sequence variants encoding the same amino acid substitution verified the robustness of our measurements (Fig. S5a, $R=0.89$, $p<10^{-100}$). Unexpectedly, when averaging the effects of all DNA sequence variants encoding the same amino acid substitution, a bimodal distribution is observed (Figs. 1d), where the great majority of p53 alterations can be discretely categorized as either retaining wtp53 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 (Figs. S5b, S6a,b).

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 (Fig. 2a). Notably, regardless of position along the DBD, premature termination codons and frameshift mutations resulted in a similarly strong disruption of

p53 functionality (Fig. 2a bottom rows in heatmap, Fig. S6c,d). 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, 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, 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 (Figs. 2b, S7a,b) 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 variants (Fig. 1d), the majority of codons segregated into two major groups displaying opposing phenotypic responses to mutagenesis (Fig. S7b).

We then calculated an evolutionary conservation score (ECS, Methods) for each residue along the DBD from a multiple sequence alignment of 1887 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 (Fig. 2a,c, R_s=0.79, $p<2x10^{-41}$), highlighting the tight coupling between p53 protein sequence evolutionary conservation and vulnerability to functional alteration. We further utilized this conservation-functionality connection 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 (Fig. 2d, Methods). This unsupervised model, which accounts for site-specific amino acid constraints alone and does not explicitly model a specific phenotypic functionality, showed a significant correlation to our experimental measurements (R_s=-0.59, $p<2x10^{-283}$), suggesting that RFS reflects a p53 functionality which 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 H1299 cells, in line with p53's involvement in multiple non-redundant processes (Kakudo et al., 2005; Pfister and Prives, 2017).

Closer examination of the mutational effects within specific codons revealed 3 distinct response patterns of codons to mutations: (i) positions highly susceptible to mutation (i.e. substitution to nearly any amino acid abolished anti-proliferative p53 functionality), including the hotspot codons G245, R248, R249 and R273 (Fig. 3a); (ii) positions resilient to mutations, tolerating practically all substitutions (with the frequent exception of proline) without losing p53 functionality (Fig. 3b); and (iii) codons in which a continuous phenotypic spectrum is observed, with mutation outcome largely depending on the specific substitution (Fig. 3c). 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 (Fig. 3a-d; Student's T $p<10^{-38}$). 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 (Fig. 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.

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 (Fig. 4a) revealed that p53 mutation prevalence across all tumor types (Bouaoun et al., 2016) is positively correlated with RFS ($R_s=0.4$, $p<4x10^{-57}$). Intriguingly, in apparent discordance with the overall picture, we also observed mutations that are rare in tumors despite having lost p53 functionality (dashed triangle, Fig. 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 protein sequence alterations achievable by a single A-G or C-T transition, the correlation between mutation prevalence and loss of functionality increases (Fig. 4b, R_s=0.52, $p<2x10^{-38}$). 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 sigmoid relationship, and the apparent separation of the ~10 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 assay, or a mechanistic tendency towards accrual of mutations at those sites.

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 measurements obtained in our phenotypic assay. This model predicted relative mutation abundance with an R=0.72 ($p<10^{-100}$) and R=0.75 ($p<3x10^{-27}$) on CV and 10% unseen test cases, respectively (Figs. 4c, S8a). Importantly, the most contributing feature in this prediction task was our measured RFS score (Fig. S8c), 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 (Figs. 4d, S8b,d). Once again, our trained model enabled accurate prediction of mutational outcome (R=0.87, $p<10^{-100}$ and R=0.88, $p<2x10^{-47}$ 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.

TP53 germline mutations underpin the majority of cases of Li-Fraumeni syndrome (LFS), an inherited cancer predisposition resulting in early-onset tumors including sarcomas, breast cancers and adrenocortical tumors (Li et al., 1988; Malkin et al., 1990). Importantly, tumors are observed at an earlier age in LFS family members harboring p53 DBD mutations that compromise anti-

proliferative functionality (RFS>-1), when compared to those bearing *TP53* germline mutations that retain anti-proliferative capacity (RFS≤-1), highlighting the prognostic value of the RFS score (Fig. 4e, 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 (Fig. S9). Yet, individuals with the 6 most prevalent hotspot mutations exhibit an even somewhat lower ATD (Kruskal-Wallis $p<0.01$), suggesting that they elicit additional gain-of-function effects not captured by our experimental system.

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 wtp53 (Wang et al., 2014). Another rare polymorphism within the DBD is R213R (rs1800372), 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. Interestingly, while on both backgrounds nonsense and frameshift mutations yielded a similar complete loss of p53 functionality, the effects of missense mutations were found to largely depend on genetic background (Figs. 5a,b, S10): while the R213R background slightly enhanced p53 functionality, the V217M background rendered the acquisition of "mild" missense mutations more disruptive to p53 function (higher RFS on SNP background relative to wt background). These results exemplify the importance of background (asymptomatic) coding sequence polymorphisms in shaping the outcome of cancer gene mutations, underscoring the merit of personalized genetic analyses.

Our findings provide a first comprehensive catalogue for the functional consequences of thousands of p53 DBD mutations in human cells, and potentially hold important clinical implications. Our cell-based measurements are highly indicative of the *in vivo* functional outcome of p53 mutations. Thus, tumor-associated p53 mutations retaining wtp53-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. Conceivably, p53 mutations that retain wtp53-like functionality in this assay may nevertheless still endow cancer-supportive phenotypes in a context-dependent manner. Yet, the high concordance of our functional measurements 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. Future studies should further expand our understanding of context-dependent mutational effects.

Methods:

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. All cells were kept at 37°C in a humidified atmosphere containing 5% CO_2 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.

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_EMCV_ (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 supermix (Invitrogen). 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 used to assess the relative abundance of p53-positive cells: Fw- CTGTGCAGCTGTGGGTTGATTC and Rv- CCAAATACTCCACACGCAAATTTC, and for the intergenic normalization region: Fw- TTTTTCCCCATCTGTTGGCT and Rv- TCTCCAGCTCTGCAACAACCT.

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

Synthetic library production 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-ethylenediaminetetraacetic 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 10µM forward (Fw) primer, 10µl of 10µM reverse (Rv) primer and 2µl HerculaseII 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- TGTCATCTTCTGTCCCTTCCCAGAAA, Rv-ATGGCGCGGACGCGGGT; DBD-B Fw-CTGTGCAGCTGTGGGTTGATTC, Rv-CCAAATACTCCACACGCAAATTTC; DBD-C Fw-CCCTCCTCAGCATCTTATCCGAGT, Rv-AGGATGGGCCTCCGGTT; DBD-D Fw-TGTGTAACAGTTCCTGCATGGG, Rv-GCAGCTCGTGGTGAGGCT. Products from all 24 identical reactions were pooled together and separated from non-specific fragments by electrophoresis on a 2.5% agarose gel stained with GelStar (Cambrex Bio Science Rockland), 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- GATGTCATGGATCCATGGAGGAGCCGCAGTC and Rv-GTACTGATGCGGCCGCTCAGTCTGAGTCAGGCCCTTC 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 37C to remove the dsRed sequence. Digested amplicon and plasmid were separated from nonspecific fragments by electrophoresis on a 1% agarose gel using a gel extraction kit (Nucleospin), and ligated using T4 ligase (NEB) for 2h at 24°C. Ligated plasmids were transformed into Escherichia coli (HIT-DH5α, 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 pEF1_EMCV_(Kazadi et al., 2008) with a wtp53 sequence using restriction-free (RF) cloning (Unger et al., 2010). Thus, the wtp53 coding sequence was amplified by PCR from pC53-SN3 (Baker et al., 1990) using Kapa HiFi Polymerase (KAPA Biosystems) with primers adding restriction sites (5' AscI, 3' RsrII) and flanking sequences that are homologous to the pEF1a plasmid on both sides of the mRFP sequence (Fw-CTAGCCTCGAGGTTTAAACGGTACCGGCGCGCCCACTGCCATGGAGGAGCCGCAGTCAGATC, Rv-GGGGGGGCGGAATCCTCAGGCTAGTCGGTCCGGACAATCGCCATGTCAACGCGTGAATGTCAGTCTGAGTCAGGCCCTTCTG). The product was purified using PCR purification kit (Qiagen), and used as a mega-primer for cloning into pEF1a 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 pEF1a lentiviral vector by conventional restriction-ligation cloning of the p53 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. cloni 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 (vector-only) colonies, RF no-insert control reactions were performed in parallel, replacing library amplicons with 10ul water. Following Dpn1 digestion and transformation, 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- GAGCCGCAGTCAGATCCTAG; Rv-GCAGCTCGTGGTGAGGCT). 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' AscI, 3' RsrII) for ligation into the pEF1a

vector (Fw-ACGGTACCGGCGCGCCCACTGCCATGGAGGAGCCGCAGTCAGATC; Rv-AGGCTAGTCGGTCCGGACAATCCAGATGTCAACGCGTGAATGTCAGTCTGAGTCAG GCCCTTCTG). To reduce variant 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, 0.2uM 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 30 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 (pEF1a-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.cloni (Lucigen) with 2µl of ligation mix, as described in Step 1, collecting a total of $2.75x10^5$, $1.53x10^5$, $3.57x10^5$ and $3.94x10^5$ colonies from libraries A, B, C and D, respectively. Colony PCRs were performed for validation as described above, and plasmids were extracted using NecleoBond Xtra Maxi kit (cat #740414.10).

Lentivirus 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 (Mercury), aliquoted and stored at -80C.

To determine viral titer, $6x10^5$ H1299 cells were plated in 10cm dishes 16 hours prior to infection. A single aliquot from each viral library was thawed at 37^0C, and serially diluted in RPMI1640. Diluted virus-containing medium was added to the plates in a final volume of 5ml, supplemented with 8μg/ml polybrene (Sigma, AL-118). 8hr later, cells were washed three times with PBS, and fresh RPMI1640 medium was added. 48hr post-infection, cells were harvested in trypsin, washed with PBS and the percent of EGFP-positive cells was determined for each virus dilution using a Guava EasyCyte flow cytometer (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 $6x10^5$ 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 ~$3.6x10^5$ and ~$7.2x10^5$ 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.

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 $1x10^7$, $1.3x10^7$, $2.5x10^7$, and $2.5x10^7$ cells (yielding 160, 220, 300 and 295μg of gDNA) 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 95^0C 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-NNNNNTGTCATCTTCTGTCCCTTCCCAGAAA, Rv- NNNNNATGGCGCGGACGCGGGT; sub-library B: Fw-NNNNNCTGTGCAGCTGTGGGTTGATTC, Rv-NNNNNCCAAATACTCCACACGCAAATTTC; sub-library C: Fw-NNNNNCCCTCCTCAGCATCTTATCCGAGT, Rv-NNNNNAGGATGGGCCTCCGGTT; sub-library D: Fw-NNNNNTGTGTAACAGTTCCTGCATGGG, Rv-NNNNNGCAGCTCGTGGTGAGGCT. 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 (Nucleospin). 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.

Deep sequencing, normalization and computational analysis. 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).

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:

$$RFS = Median\left(R_{t_1}, R_{t_2}, R_{t_3}\right)$$

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

$$R_{t_i} = \log_2\left(\frac{r_{t_i}}{r_{t_0}}\right)$$

r_{t_i} 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.

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.

$$ECS(i) = 1 + \frac{\sum_a f_i(a)\ln(f_i(a))}{\ln q}$$

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 jackhmmer 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 σ 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 σ_i in position *i*. These parameters are inferred from the sequence alignment using l_2-penalized maximum likelihood inference. The effect of a substitution ΔE can then quantified by the log-odds ratio of the probabilities of the mutant and wild-type sequences under the model:

$$\Delta E(\boldsymbol{\sigma}^{\mathrm{mut}}, \boldsymbol{\sigma}^{\mathrm{wt}}) = \log \frac{P(\boldsymbol{\sigma}^{\mathrm{mut}})}{P(\boldsymbol{\sigma}^{\mathrm{wt}})}$$

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/).

Mutation prevalence and RFS prediction models. All predictions were performed with Random Forest regression models, using the scikit-learn (Pedregosa et al., 2011) RandomForestRegressor class. In each of the prediction tasks, we randomly partitioned the data into a 90% training set and a 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. 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 (e.g. in Fig. 1b), or Wilcoxon signed-rank test for matched samples (e.g. Figs. 4e and f). All performed statistical tests were two-sided. To assess the difference between time-course dynamics (e.g. in Fig. 1c) we performed a within-subjects two-way analysis of variance (ANOVA). For comparison between ages of tumor

onset in patients harboring different mutation types we performed a one-way Kruskal-Wallis H test (Fig. S9). All error bars represent ±SEM (standard error of the mean), unless noted otherwise. Heatmap hierarchical clustering was performed using a Euclidean distance matrix. Statistical analyses were performed using the scipy python package.

Acknowledgements:

We thank Shira Weingarten-Gabbay, David Zeevi, Michal Levo, Leeat Keren, Noa Furth, Odem Shani, Yael Aylon and Adina Weinberger for fruitful discussions and valuable inputs. We gratefully acknowledge Zohar Yakhini for help with library synthesis, N. Rahm and A. Telenti for the lentiviral bicistronic constructs, Jonathan Weinstein for help with structural models and Yoav Peleg for cloning assistance. This work was supported by grants from the European Research Council (to E.S.) and from the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation and 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; M.L-P cloned the libraries; T.H and D.M performed analyses; E.S. and M.O. conceived the project and experiments, supervised the work, and wrote the manuscript.

Figures:

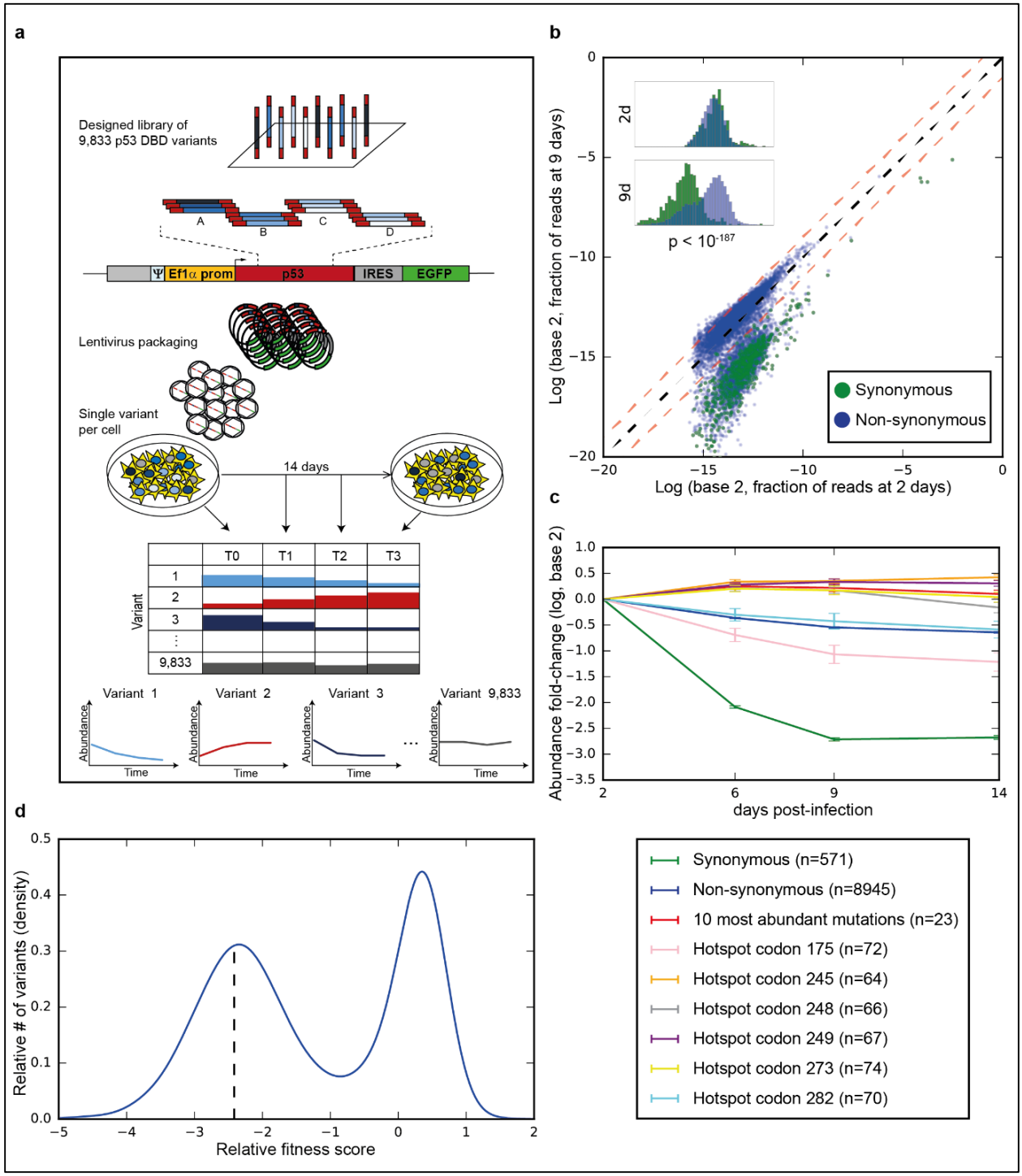


Figure 1: A high-throughput experimental system for measuring the effects of p53 variations. **(a)** Experimental design: a library of 9,833 designed p53 sequence variants was synthesized on Agilent microarrays in 4 separate pools spanning consecutive parts of the DBD (labeled A, B, C

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 Methods) for each individual variant. **(b)** Relative fraction of reads for each variant (N=9516) 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/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 6 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=6837.

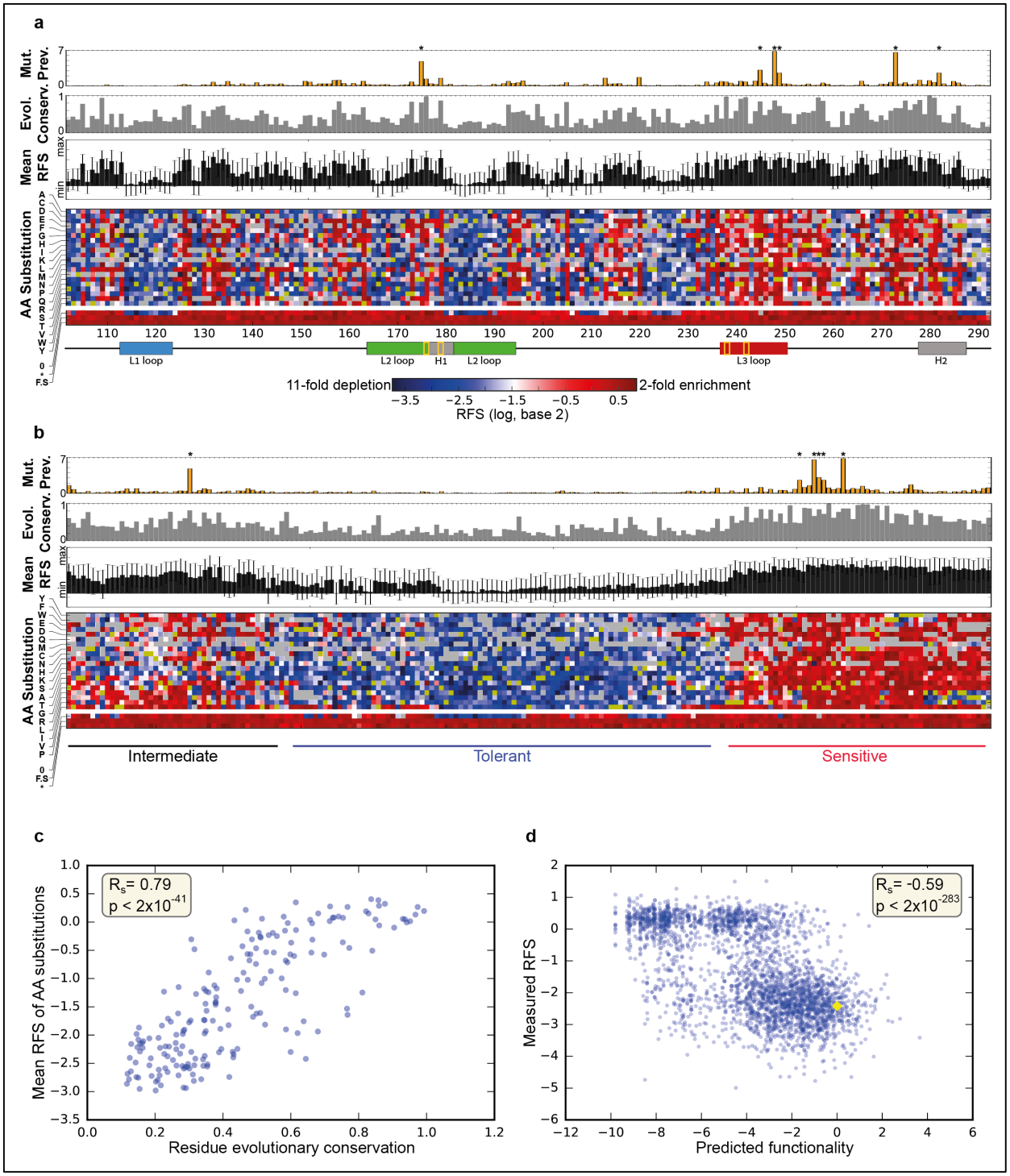


Figure 2: The pattern of functional effects of p53 mutations correlates with protein structural domains and evolutionary conservation.

(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 wild-type residue in each position is marked in yellow; missing data points are in grey. Orange bars on top show the prevalence (% of 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 1887 homologous sequences (see 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. **(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). **(c)** Mean RFS, calculated for each amino acid position across all missense mutations at that position, 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 wild-type sequence probabilities, wt=0; see Methods). Yellow diamond represents wtp53. N=2990.

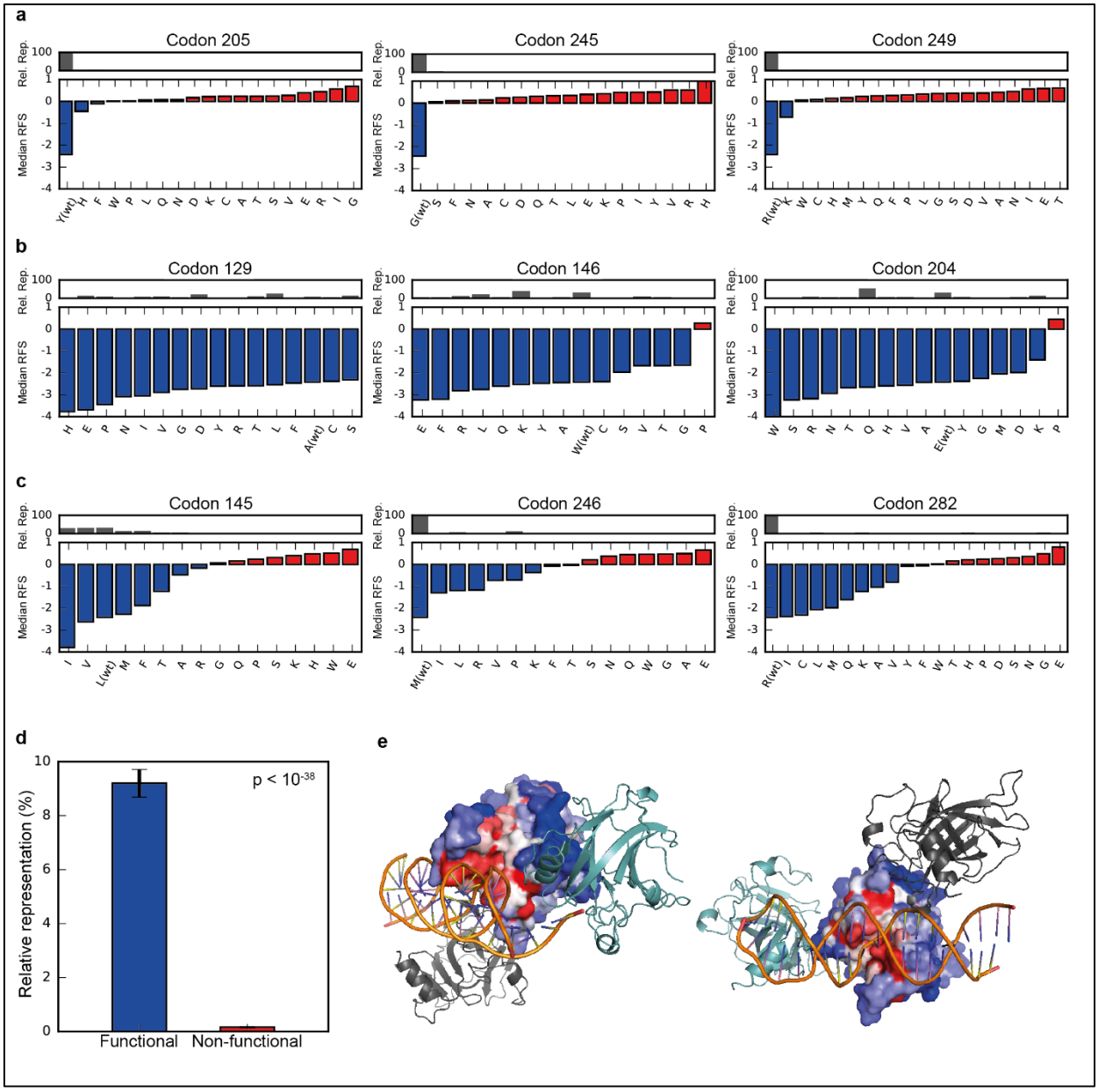


Figure 3: Residues within the DBD exhibit different patterns of robustness to modification. **(a-c)** For each indicated p53 codon, we show the median RFS measured for all assayed amino acid substitutions at that position (blue = decrease in relative abundance, presumably retaining wtp53 activity; red = relative enrichment, presumably due to loss of wtp53 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 wtp53 is indicated at the bottom. **(d)** p53 variants were divided by their RFS into two groups: retaining wtp53-like functionality (blue) or

dysfunctional (red). Threshold for separation was set to -1 (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=1152 in red group, N=2005 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.

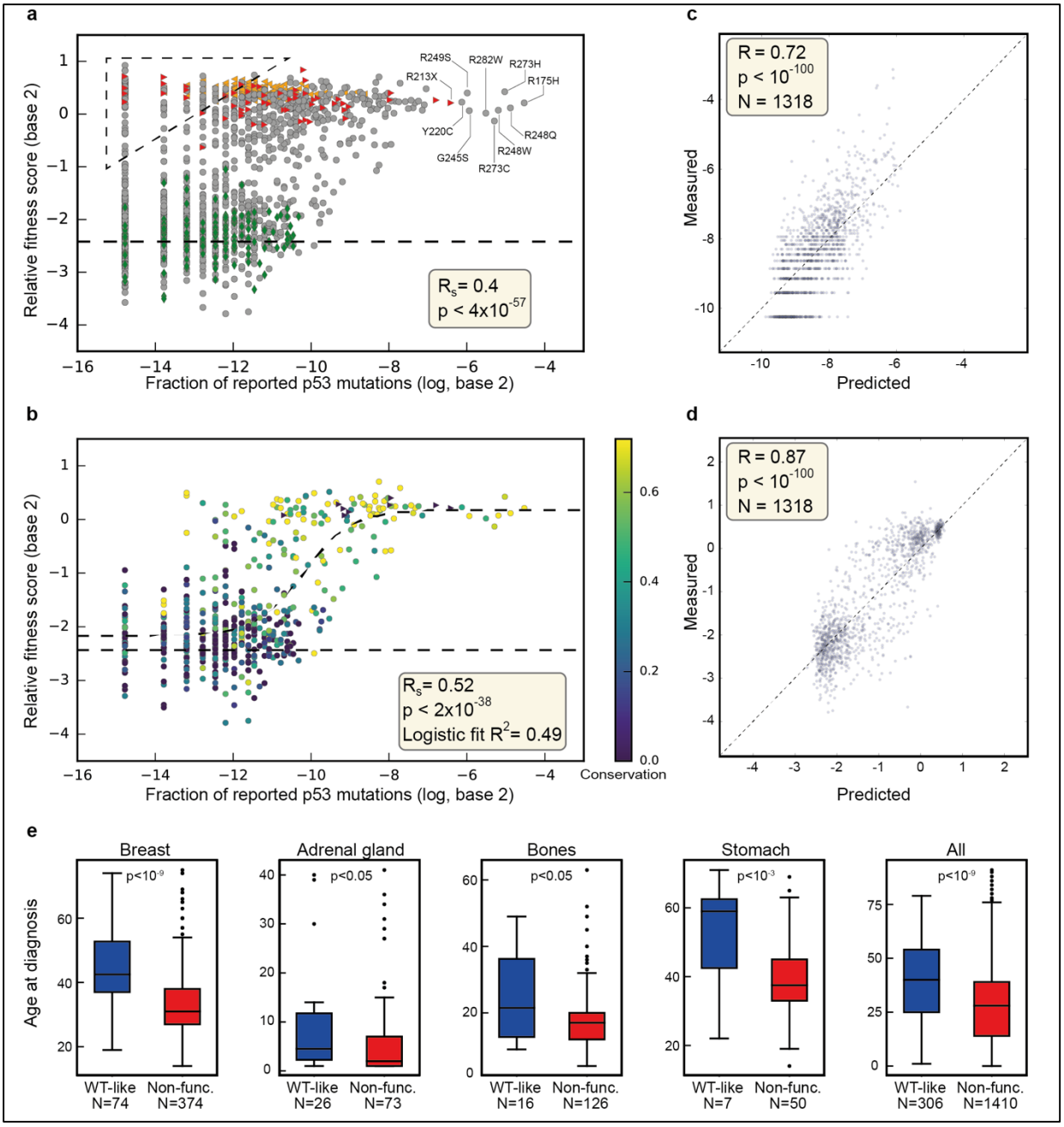


Figure 4: Mutation prevalence in human tumors is highly correlated with loss of wtp53 anti-proliferative effect.

(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. 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=1465.

(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. Dashed line represents a sigmoidal fit of the form $y = c / (1+e^{-k*(x-x_0)}) + y_0$, created using least squares residual minimization. $R^2 = 0.49$. N=526. **(c)** A Random forest model was trained using 90% of the data with 200-fold cross validation (CV) for predicting relative mutation prevalence in human tumors. For each protein variant used in CV, p53 mutation prevalence (Bouaoun et al., 2016) ("Measured") was plotted against model predictions. **(d)** Same as (c), for a model trained to predict RFS (measured RFS is plotted against model-predicted RFS for each protein variant). **(e)** Age at diagnosis of LFS family members presenting with tumors of different tissue origin (Bouaoun et al., 2016). Tumors are divided according to the functional impact of the corresponding germline *TP53* mutations: blue – retaining wtp53-like anti-proliferative functionality (RFS≤-1), red – disrupting functionality (RFS>-1). p-values: Mann-Whitney U.

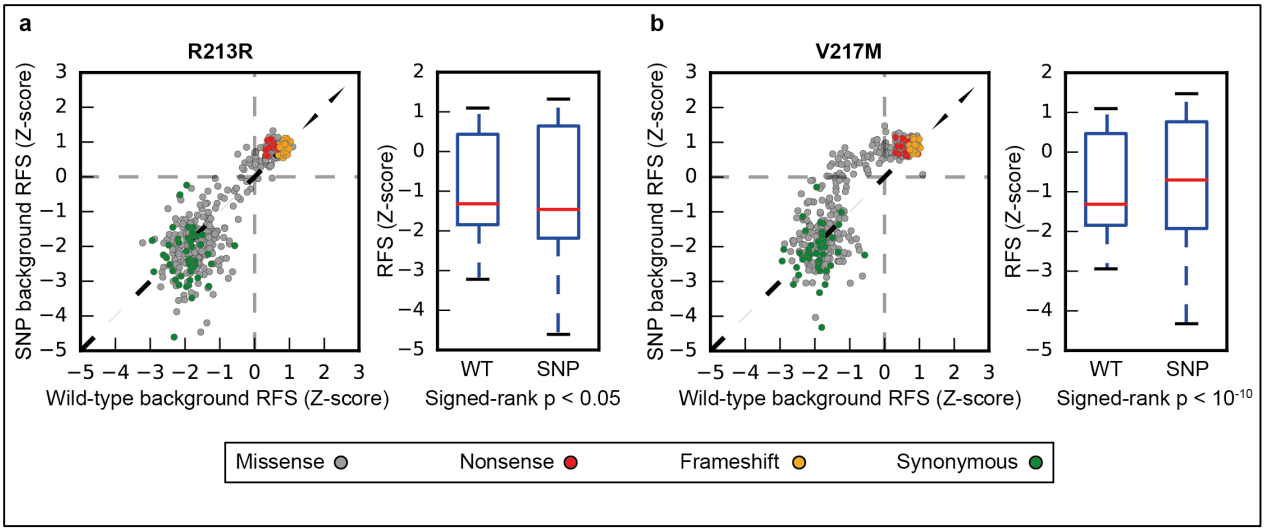
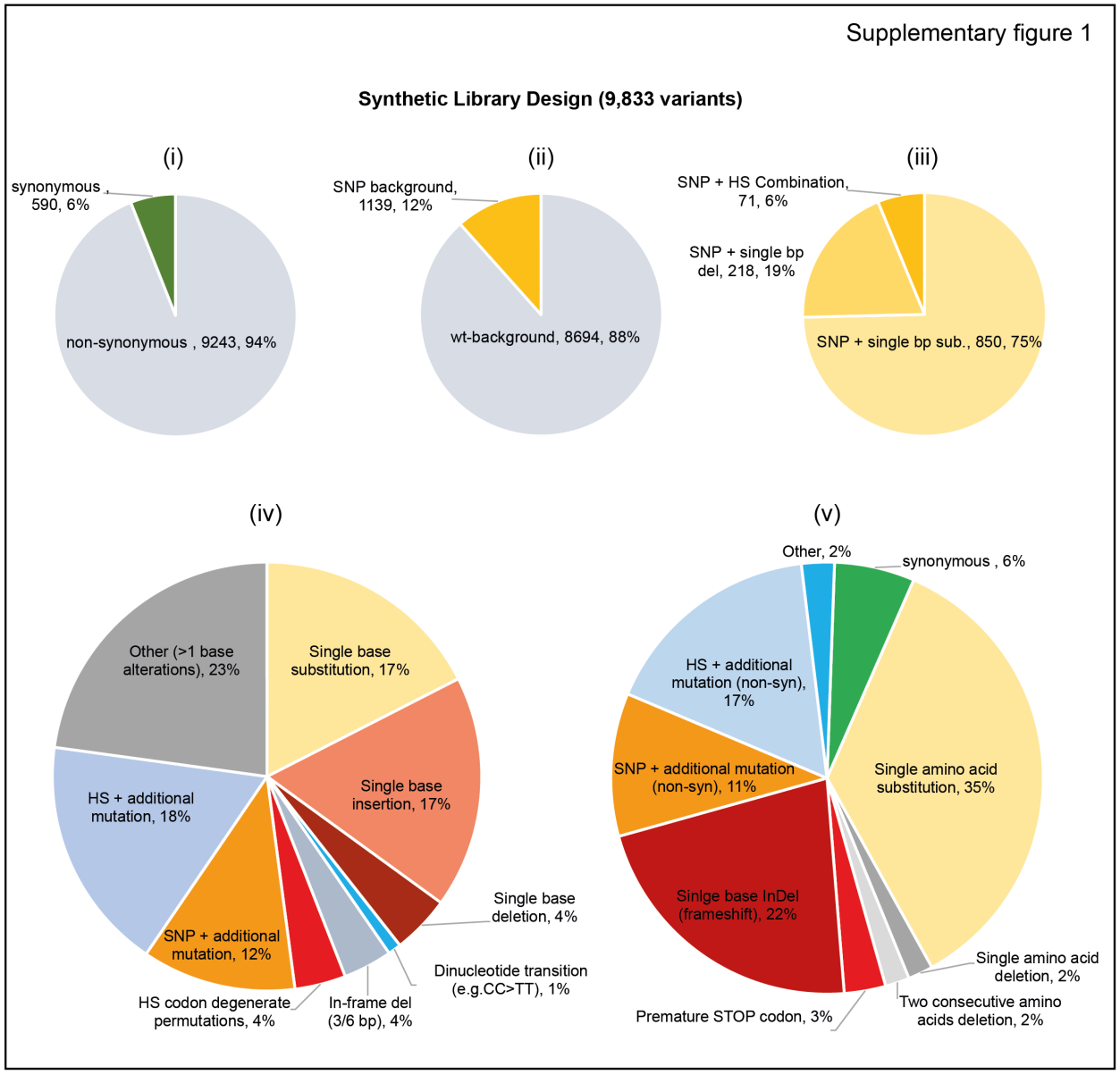


Figure 5: The phenotypic outcome of missense mutations is affected by the V217M SNP.

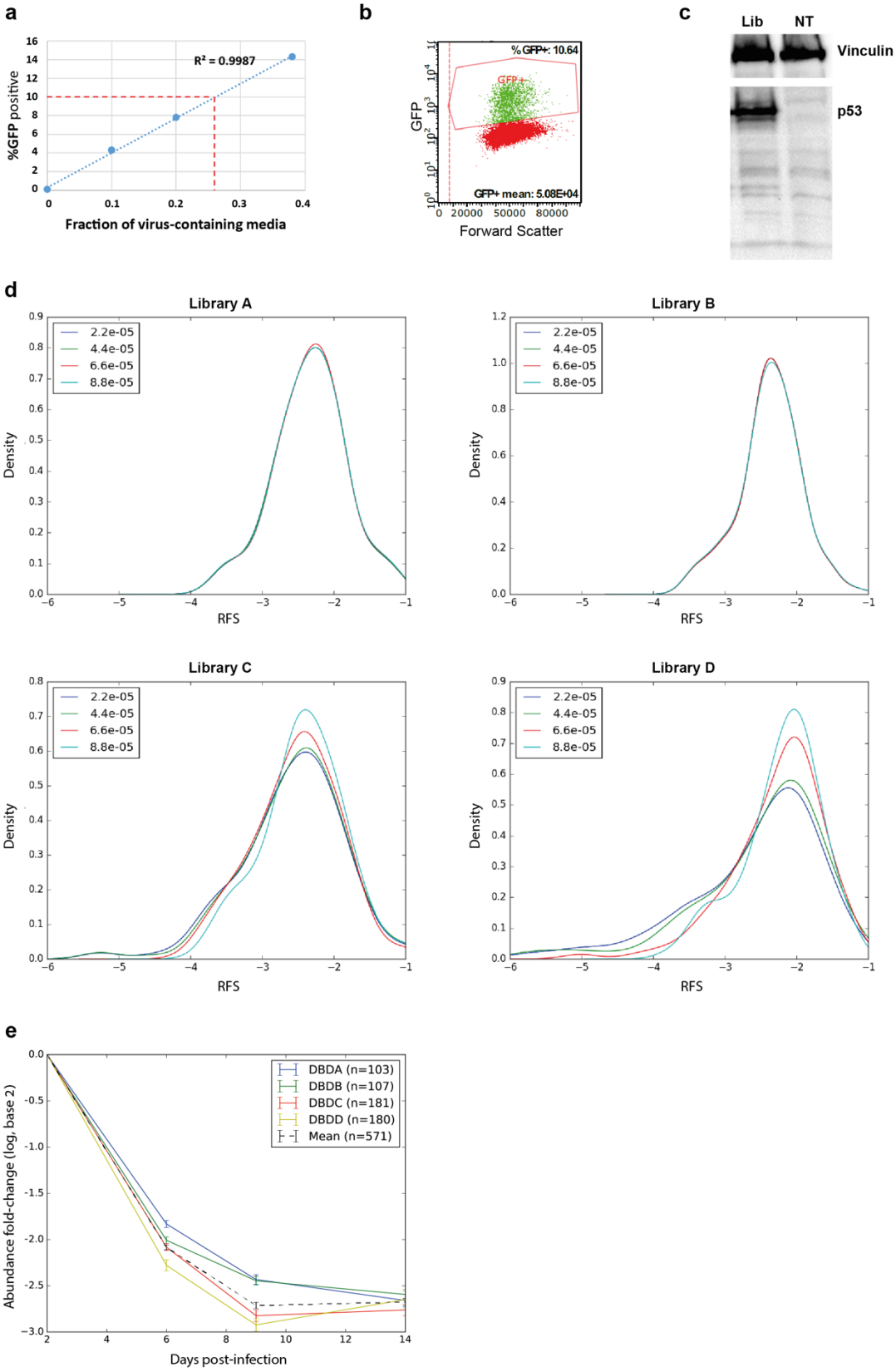
(a) For each variant, we plotted its standardized relative fitness (Z-scores) when the mutation is on a SNP R213R background as compared to wtp53 background. N=388. Variant type colors are shown at the bottom. **(b)** Same as (a), but for SNP V217M. N=385.

Supplementary figures:

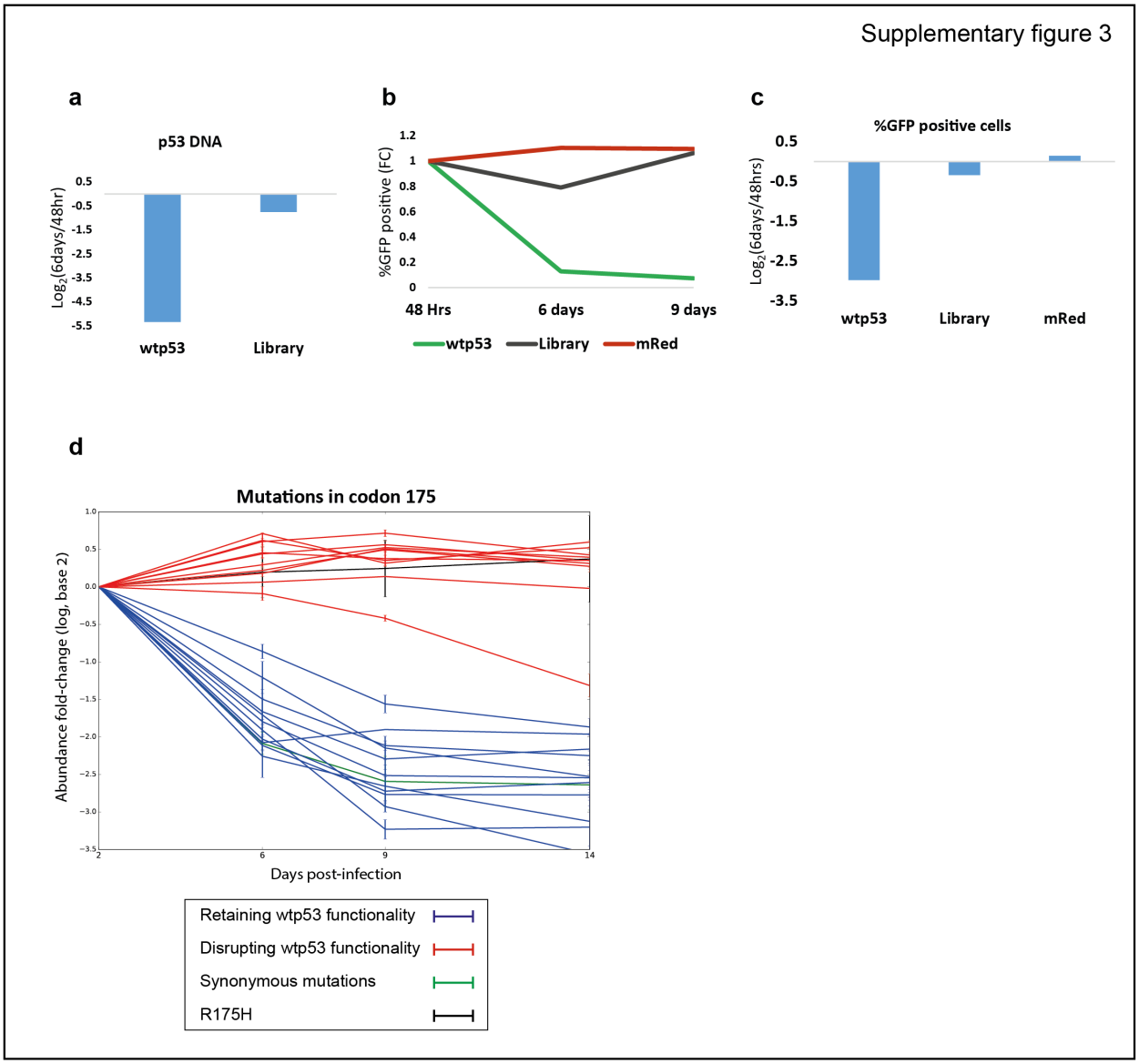


Supplementary figure 1: Synthetic design of the p53 DBD mutation library (total of 9,833 unique sequence variants). Variants are stratified by: (i) Synonymous vs. non-synonymous mutation variants; (ii) variant genetic background (i.e. mutations on a wtp53 or SNP background; (iii) mutation types on a SNP-background; (iv) nature of sequence alteration; (v) mutation outcome at the protein level. Abbreviations: HS=hotspot codon; sub=substitution; del=deletion; InDel=insertion or deletion.

Supplementary figure 2



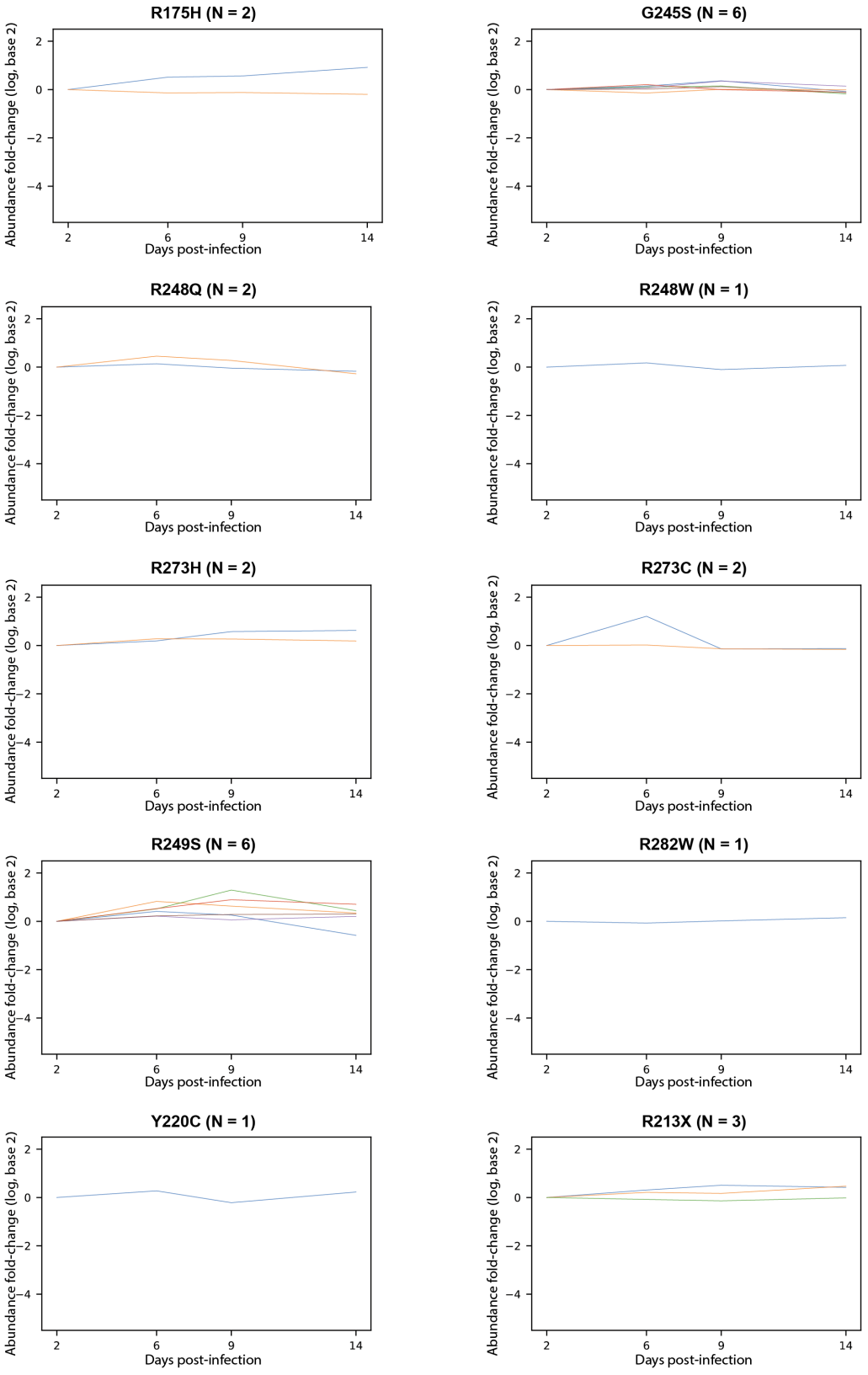
Supplementary figure 2: (a) H1299 cells were infected with serial dilutions of the p53-library lentivirus stock. 48 hours post-infection, the percentage of cells expressing EGFP was determined by flow cytometry. The dashed red line represents the interpolated fraction of viral stock required to achieve a MOI of 0.1. **(b)** Percentage of EGFP-positive H1299 cells as measured by flow cytometry 48 hours post-infection (PI) with lentiviruses encoding the p53 DBD library (data from sub-library D is shown as a representative example). **(c)** 72 hours post-infection of H1299 cells with the p53 DBD library (or no-virus control medium, NT), cells were subjected to Western blot analysis with anti-p53 antibodies (mixture of D01+PAb1801). Vinculin was used for loading control (data from sub-library D is shown as a representative example). **(d)** Threshold for the minimal number of reads at first time-point was determined by examining the RFS calculated across all synonymous mutations in each sub-library. As the minimal number of reads increases, better accuracy is achieved and variants that pass the threshold converge towards the actual value. Graphs depict the distribution of RFS measured across all silent mutations in the 4 sub-libraries for increasing thresholds of minimal read fractions used. Note that as sub-libraries C and D are of higher complexity (contain a larger number of designed variants), a larger minimal threshold is required to converge to an accurate estimation. **(e)** Time course dynamics (means +/- STE) of all synonymous mutation variants in each of the 4 sub-libraries, and the overall mean (dashed black line) across the entire DBD library. Measured read fractions from each of the 4 libraries at each time point were normalized to merge with the overall mean (Methods).



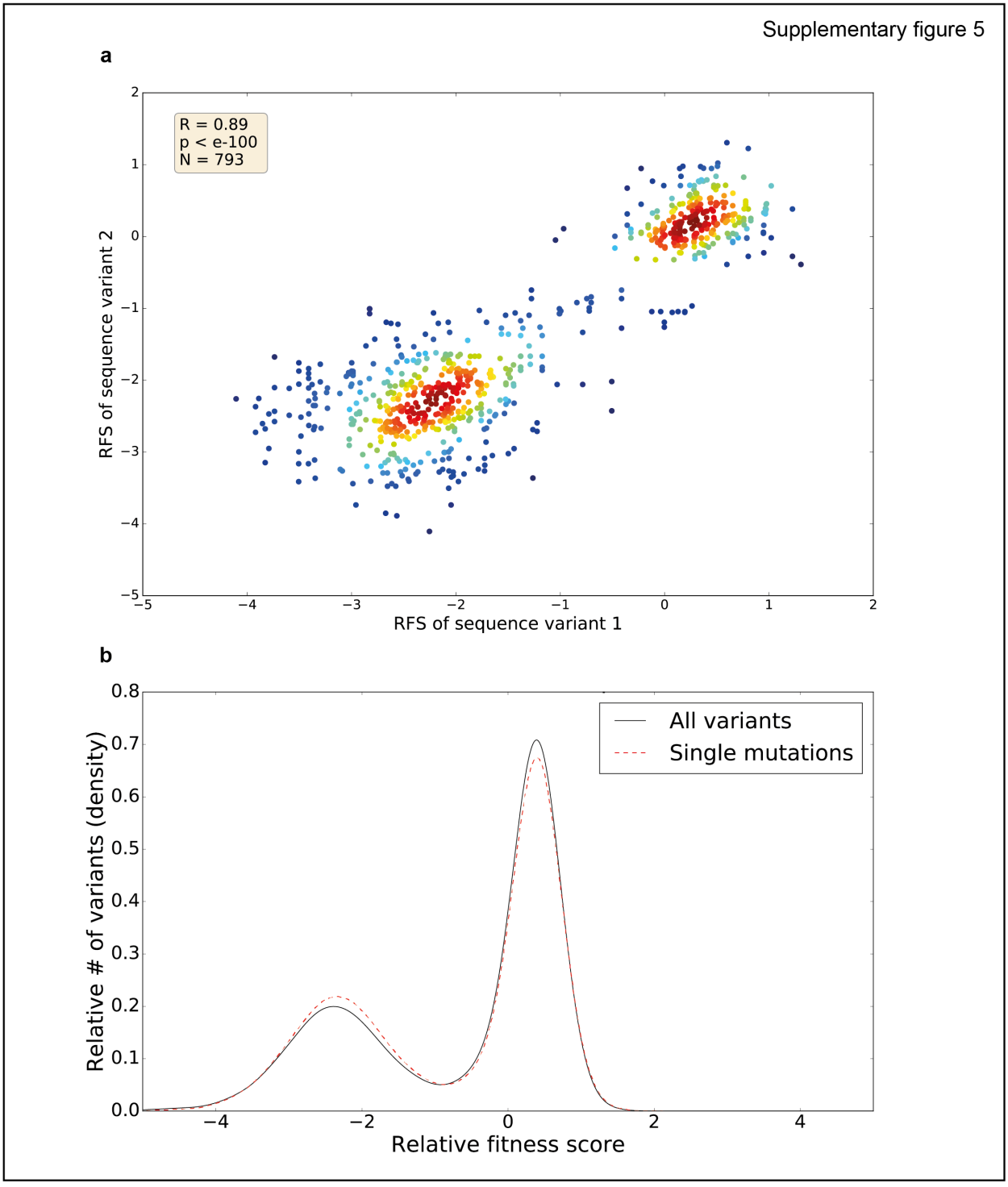
Supplementary figure 3: (a) H1299 cells were infected with the lentiviral p53 library or with wtp53-IRES-EGFP (wtp53 in the same lentiviral construct as the library), and analyzed for p53 DNA levels (measured by qPCR with p53-specific primers and normalized to an intergenic region upstream of the *KCNA4* gene), 48 hours and 6 days PI. Note that while wtp53 DNA is rapidly diluted out, the DNA of the mutant p53 library is only slightly reduced. **(b, c)** Cells were infected as in (a) or with a mRed-IRES-EGFP lentivirus (p53-null control vector), and sampled at the indicated time points PI. At each time point, the percentage of EGFP positive cells was measured by flow cytometry and compared to the 48 hour time point; fold-change values are presented. **(d)** Time course growth curves comparing the dynamics of all different protein variants at codon 175.

Each plotted line represents the mean (+/- STE) of all DNA sequence variants encoding the same particular amino acid. The R175H hotspot mutation is indicated in black.

Supplementary figure 4

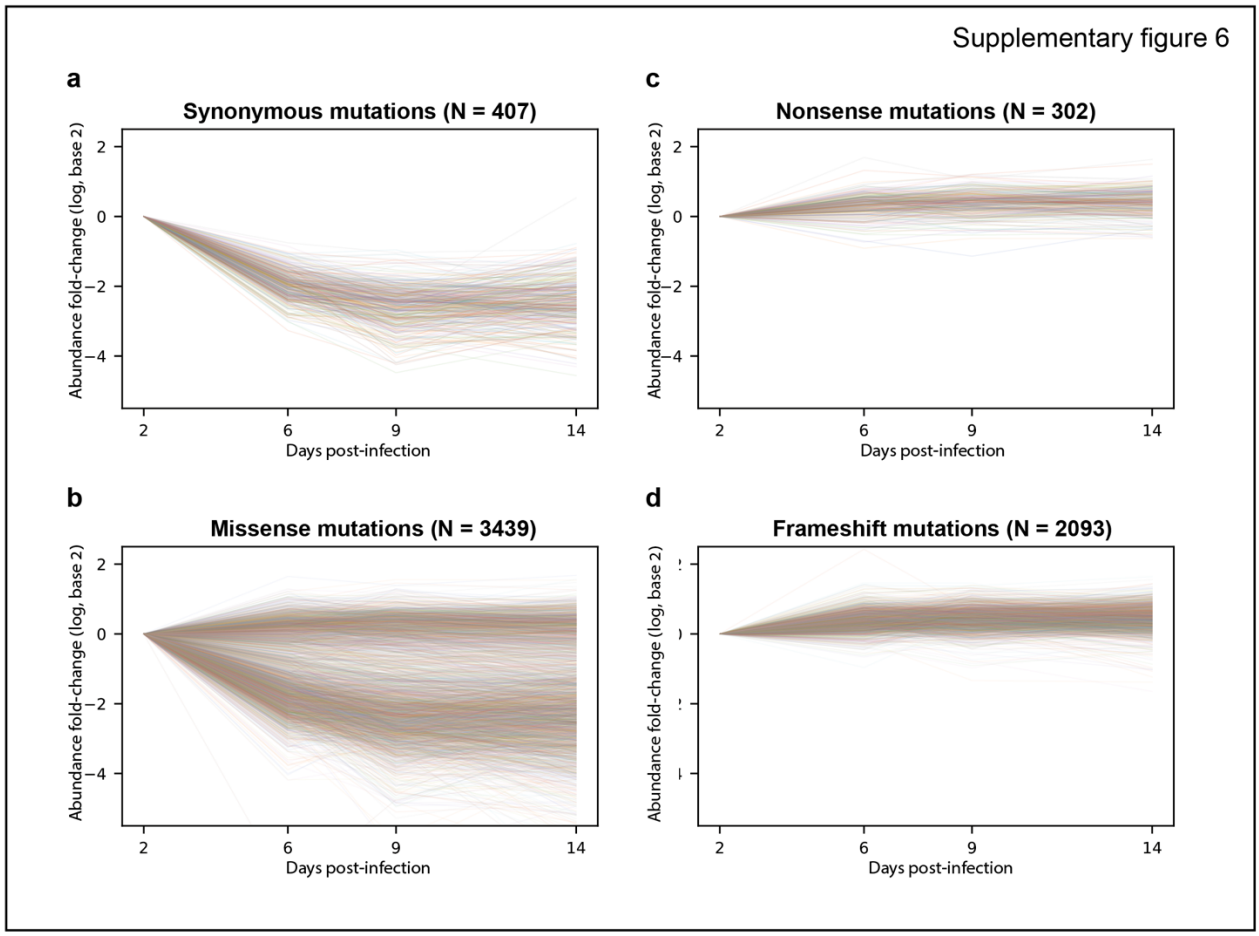


Supplementary figure 4: Time course growth curves showing the dynamics of DNA sequence variants encoding the ten most prevalent protein variations in p53. Each panel shows a specific protein variant. Lines represent unique sequence variants encoding the particular amino acid alteration.

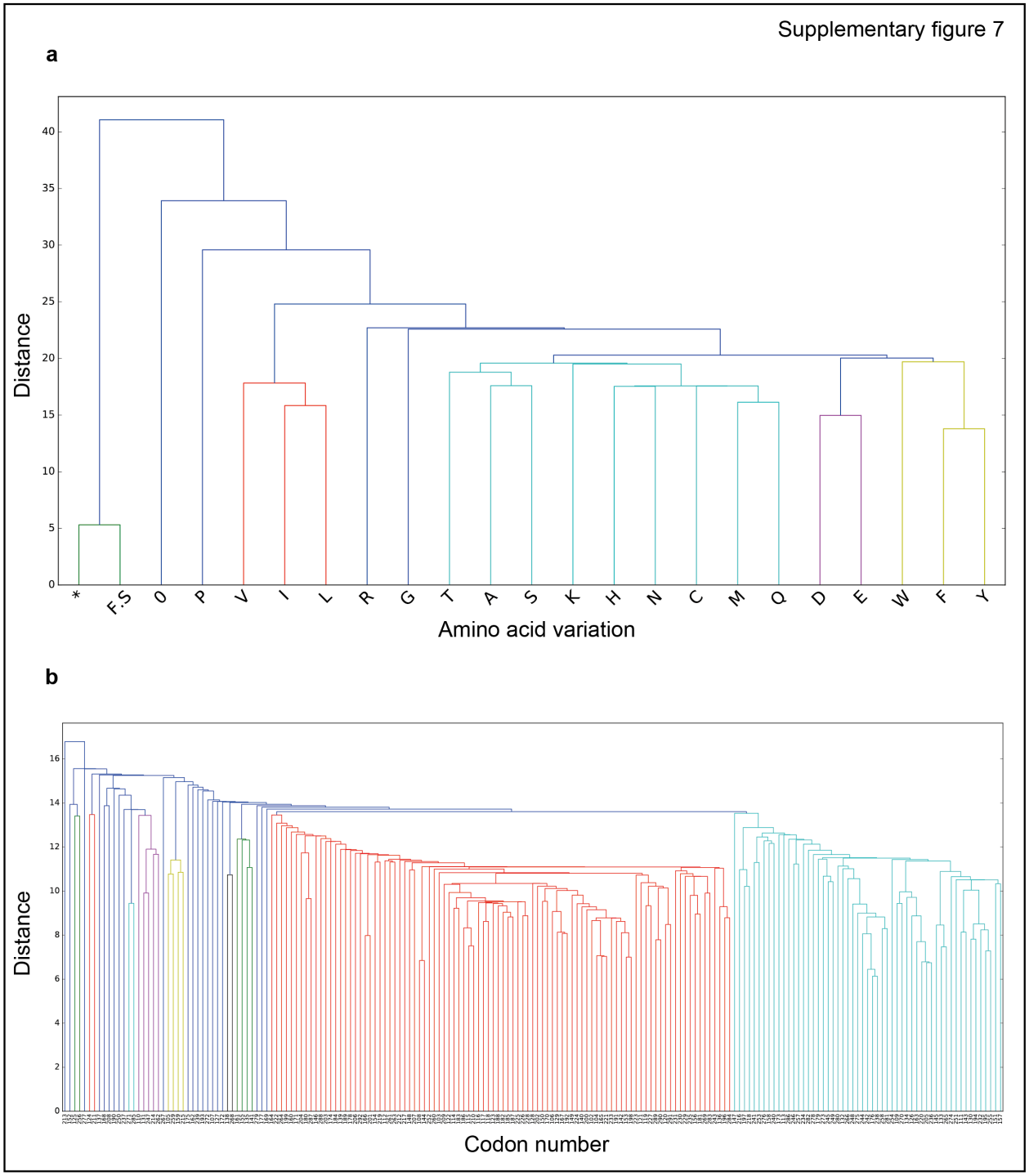


Supplementary figure 5: (a) Scatter plot showing the correlation between RFS values of different DNA sequence variants encoding the same amino acid substitution (including synonymous variants). For each codon, all pairwise combinations of sequence variants encoding the same protein sequence are plotted (RFS of one variant against the other). Colors indicate point density

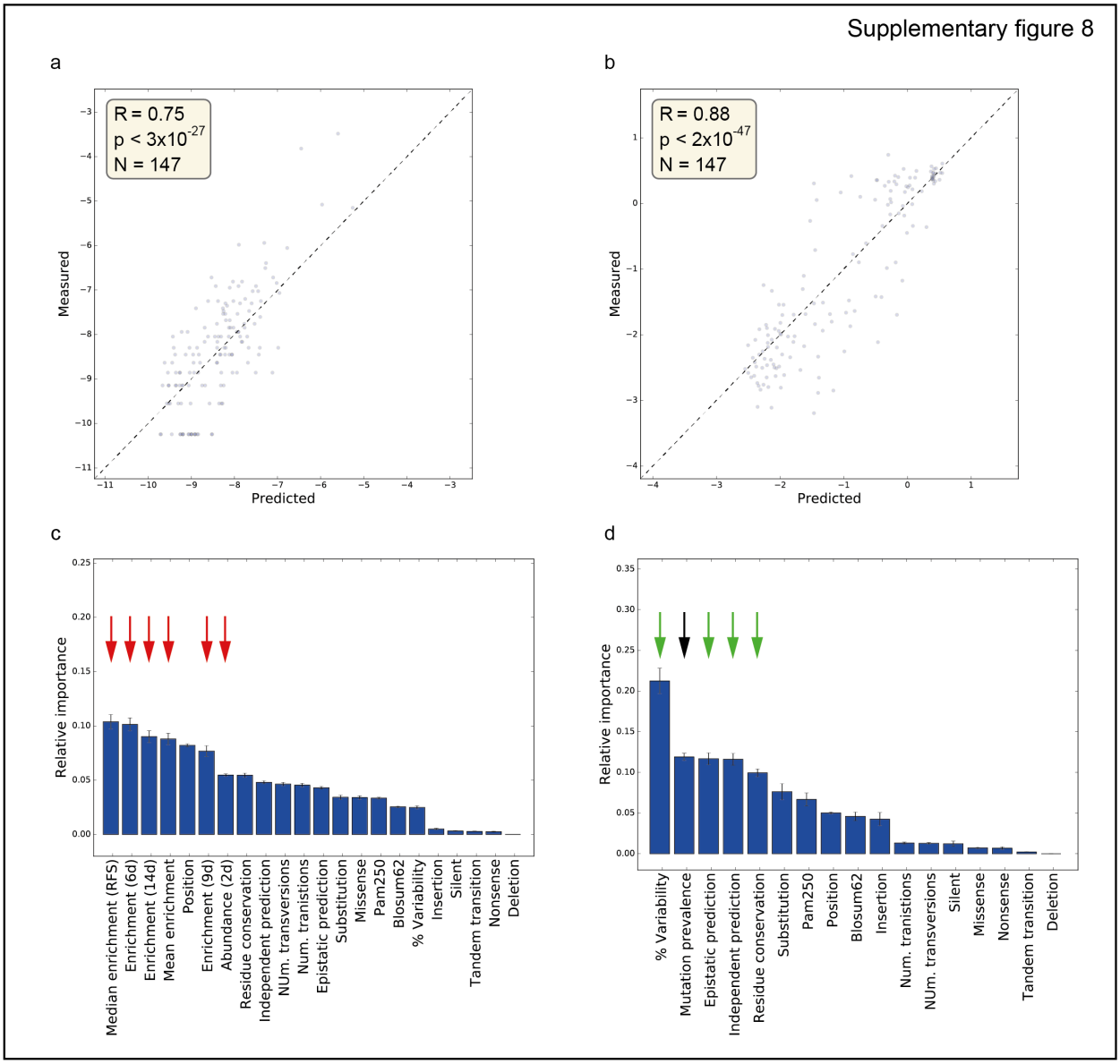
from high (red) to low (blue). **(b)** Distribution of measured relative fitness scores across all DNA sequence variants in the library (black line) or across sequence variants encoding only single mutations (dashed red line).



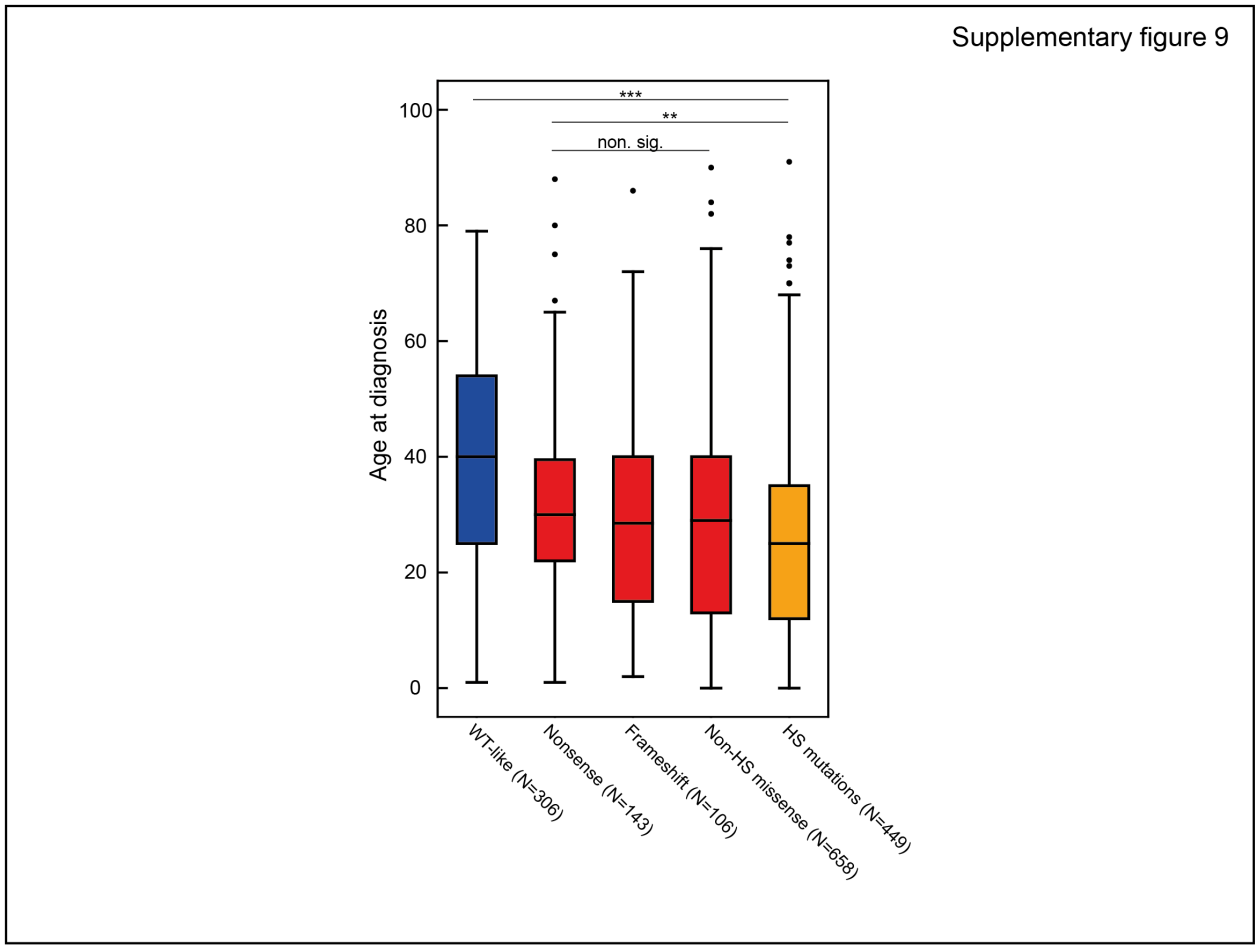
Supplementary figure 6: Time course growth curves showing the dynamics of individual sequence variants resulting in synonymous mutations **(a)**, missense mutations **(b)**, nonsense mutations **(c)** or frameshifts **(d)**.



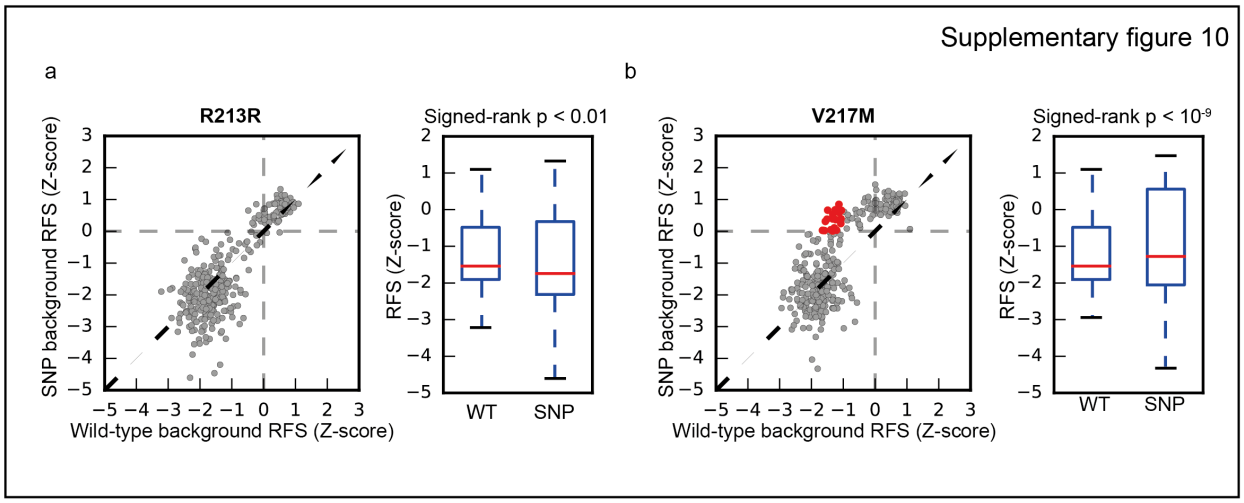
Supplementary figure 7: Hierarchical clustering of RFS values measured for p53 DBD variants. Dendrograms depict clustering by amino acid variation **(a)**, and by codon number **(b)**, corresponding to rows and columns in Fig. 2b, respectively.



Supplementary figure 8: (a, b) Prediction model performances on held-out test data. Random forest predictors were trained using 90% of the data (training set). Then, learned parameters were used for prediction on held-out test data (10%). Scatter plots show measured values of p53 mutation prevalence (a) or RFS (b) for each p53 protein variant in the test set plotted against model predictions for that variant. (**c, d**) Bars represent the relative importance of features (+/- STE) used for predicting p53 mutation prevalence (c) and RFS (d). Arrows indicate features based on our competition assay measurements (red), evolutionary conservation (green) and p53 mutation prevalence in human cancers (black).



Supplementary figure 9: Box plots show the age at diagnosis of tumors in LFS family members across all tumor types (11). Tumors are stratified by mutation type and effect: blue – missense mutations retaining wtp53-like anti-proliferative functionality (RFS≤-1); red – nonsense, frameshift and missense mutations that disrupt wtp53-functioanlity (RFS>-1), excluding the 6 most prevalent hotspot mutations; orange – hotspot mutations (R175H, R273H, R248Q, R248W, R273C and R282W). ***$p<10^{-16}$; **$p<0.01$ (Kruskal-Wallis H).



Supplementary figure 10: (a) For each p53 variant, we plotted its standardized relative fitness (Z-scores) when the mutation is on a SNP R213R background as compared to wtp53 background. Only variants encoding a non-truncated p53 protein are shown (excluding nonsense and frameshift mutations). **(b)** Same as (a), but for SNP V217M. Mutations that retain p53 functionality on a wtp53 background (at least 1 STD below the mean effect) yet are non-functional on the SNP background (positive Z-score), are colored red. Box-plots compare the overall effect of mutations on the indicated backgrounds.

References:

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res *44*, W344-350.

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

Bieging, K.T., Mello, S.S., and Attardi, L.D. (2014). Unravelling mechanisms of p53-mediated tumour suppression. Nat Rev Cancer *14*, 359-370.

Blecher-Gonen, R., Barnett-Itzhaki, Z., Jaitin, D., Amann-Zalcenstein, D., Lara-Astiaso, D., and Amit, I. (2013). High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. Nat Protoc *8*, 539-554.

Blons, H., and Laurent-Puig, P. (2003). TP53 and head and neck neoplasms. Hum Mutat *21*, 252-257.

Bouaoun, L., Sonkin, D., Ardin, M., Hollstein, M., Byrnes, G., Zavadil, J., and Olivier, M. (2016). TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data. Hum Mutat *37*, 865-876.

Brenan, L., Andreev, A., Cohen, O., Pantel, S., Kamburov, A., Cacchiarelli, D., Persky, N.S., Zhu, C., Bagul, M., Goetz, E.M.*, et al.* (2016). Phenotypic Characterization of a Comprehensive Set of MAPK1/ERK2 Missense Mutants. Cell Rep *17*, 1171-1183.

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

Bullock, A.N., Henckel, J., DeDecker, B.S., Johnson, C.M., Nikolova, P.V., Proctor, M.R., Lane, D.P., and Fersht, A.R. (1997). Thermodynamic stability of wild-type and mutant p53 core domain. Proc Natl Acad Sci U S A *94*, 14338-14342.

Cleary, M.A., Kilian, K., Wang, Y., Bradshaw, J., Cavet, G., Ge, W., Kulkarni, A., Paddison, P.J., Chang, K., Sheth, N.*, et al.* (2004). Production of complex nucleic acid libraries using highly parallel in situ oligonucleotide synthesis. Nat Methods *1*, 241-248.

Duan, J., and Nilsson, L. (2006). Effect of Zn2+ on DNA recognition and stability of the p53 DNA-binding domain. Biochemistry *45*, 7483-7492.

Fowler, D.M., and Fields, S. (2014). Deep mutational scanning: a new style of protein science. Nat Methods *11*, 801-807.

Ganci, F., Conti, S., Fontemaggi, G., Manciocco, V., Donzelli, S., Covello, R., Muti, P., Strano, S., Blandino, G., and Spriano, G. (2011). Allelic expression imbalance of TP53 mutated and polymorphic alleles in head and neck tumors. OMICS *15*, 375-381.

Geiser, M., Cebe, R., Drewello, D., and Schmitz, R. (2001). Integration of PCR fragments at any specific site within cloning vectors without the use of restriction enzymes and DNA ligase. Biotechniques *31*, 88-90, 92.

Hoffman, Y., Bublik, D.R., Pilpel, Y., and Oren, M. (2014). miR-661 downregulates both Mdm2 and Mdm4 to activate p53. Cell Death Differ *21*, 302-309.

Hopf, T.A., Ingraham, J.B., Poelwijk, F.J., Scharfe, C.P., Springer, M., Sander, C., and Marks, D.S. (2017). Mutation effects predicted from sequence co-variation. Nat Biotechnol *35*, 128-135.

Iacopetta, B. (2003). TP53 mutation in colorectal cancer. Hum Mutat *21*, 271-276.

Kakudo, Y., Shibata, H., Otsuka, K., Kato, S., and Ishioka, C. (2005). Lack of correlation between p53-dependent transcriptional activity and the ability to induce apoptosis among 179 mutant p53s. Cancer Res *65*, 2108-2114.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A.*, et al.* (2013). Mutational landscape and significance across 12 major cancer types. Nature *502*, 333-339.

Kato, S., Han, S.Y., Liu, W., Otsuka, K., Shibata, H., Kanamaru, R., and Ishioka, C. (2003). Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. Proc Natl Acad Sci U S A *100*, 8424-8429.

Kazadi, K., Loeuillet, C., Deutsch, S., Ciuffi, A., Munoz, M., Beckmann, J.S., Moradpour, D., Antonarakis, S.E., and Telenti, A. (2008). Genomic determinants of the efficiency of internal ribosomal entry sites of viral and cellular origin. Nucleic Acids Res *36*, 6918-6925.

Keren, L., Hausser, J., Lotan-Pompan, M., Vainberg Slutskin, I., Alisar, H., Kaminski, S., Weinberger, A., Alon, U., Milo, R., and Segal, E. (2016). Massively Parallel Interrogation of the Effects of Gene Expression Levels on Fitness. Cell *166*, 1282-1294 e1218.

Kitzman, J.O., Starita, L.M., Lo, R.S., Fields, S., and Shendure, J. (2015). Massively parallel single-amino-acid mutagenesis. Nat Methods *12*, 203-206, 204 p following 206.

LeProust, E.M., Peck, B.J., Spirin, K., McCuen, H.B., Moore, B., Namsaraev, E., and Caruthers, M.H. (2010). Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process. Nucleic Acids Res *38*, 2522-2540.

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

Li, F.P., Fraumeni, J.F., Jr., Mulvihill, J.J., Blattner, W.A., Dreyfus, M.G., Tucker, M.A., and Miller, R.W. (1988). A cancer family syndrome in twenty-four kindreds. Cancer Res *48*, 5358-5362.

Majithia, A.R., Tsuda, B., Agostini, M., Gnanapradeepan, K., Rice, R., Peloso, G., Patel, K.A., Zhang, X., Broekema, M.F., Patterson, N.*, et al.* (2016). Prospective functional classification of all possible missense variants in PPARG. Nat Genet *48*, 1570-1575.

Malkin, D., Li, F.P., Strong, L.C., Fraumeni, J.F., Jr., Nelson, C.E., Kim, D.H., Kassel, J., Gryka, M.A., Bischoff, F.Z., Tainsky, M.A*., et al.* (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science *250*, 1233-1238.

Martin, P., Albagli, O., Poggi, M.C., Boulukos, K.E., and Pognonec, P. (2006). Development of a new bicistronic retroviral vector with strong IRES activity. BMC Biotechnol *6*, 4.

Menendez, D., Inga, A., and Resnick, M.A. (2006). The biological impact of the human master regulator p53 can be altered by mutations that change the spectrum and expression of its target genes. Mol Cell Biol *26*, 2297-2308.

Olivier, M., Hollstein, M., and Hainaut, P. (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol *2*, a001008.

Olivier, M., and Taniere, P. (2011). Somatic mutations in cancer prognosis and prediction: lessons from TP53 and EGFR genes. Curr Opin Oncol *23*, 88-92.

Oren, M., and Rotter, V. (2010). Mutant p53 gain-of-function in cancer. Cold Spring Harb Perspect Biol *2*, a001107.

Ory, K., Legros, Y., Auguin, C., and Soussi, T. (1994). Analysis of the most representative tumour-derived p53 mutants reveals that changes in protein conformation are not correlated with loss of transactivation or inhibition of cell proliferation. The EMBO journal *13*, 3496-3504.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V*., et al.* (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research *12*, 2825-2830.

Peller, S., and Rotter, V. (2003). TP53 in hematological cancer: low incidence of mutations with significant clinical relevance. Hum Mutat *21*, 277-284.

Pfister, N.T., and Prives, C. (2017). Transcriptional Regulation by Wild-Type and Cancer-Related Mutant Forms of p53. Cold Spring Harb Perspect Med *7*.

Pilger, D.A., Lopez, P.L., Segal, F., and Leistner-Segal, S. (2007). Analysis of R213R and 13494 g-->a polymorphisms of the p53 gene in individuals with esophagitis, intestinal metaplasia of the cardia and Barrett's Esophagus compared with a control group. Genomic Med *1*, 57-63.

Resnick, M.A., and Inga, A. (2003). Functional mutants of the sequence-specific transcription factor p53 and implications for master genes of diversity. Proc Natl Acad Sci U S A *100*, 9934-9939.

Riley, T., Sontag, E., Chen, P., and Levine, A. (2008). Transcriptional control of human p53-regulated genes. Nature reviews Molecular cell biology *9*, 402-412.

Sander, C., and Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins *9*, 56-68.

Schuijer, M., and Berns, E.M. (2003). TP53 and ovarian cancer. Hum Mutat *21*, 285-291.

Sharma, S., Sambyal, V., Guleria, K., Manjari, M., Sudan, M., Uppal, M.S., Singh, N.R., Bansal, D., and Gupta, A. (2014). TP53 polymorphisms in sporadic North Indian breast cancer patients. Asian Pac J Cancer Prev *15*, 6871-6879.

Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., Keren, L., Yakhini, Z., Weinberger, A., and Segal, E. (2012). Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. Nat Biotechnol *30*, 521-530.

Tewhey, R., Kotliar, D., Park, D.S., Liu, B., Winnicki, S., Reilly, S.K., Andersen, K.G., Mikkelsen, T.S., Lander, E.S., Schaffner, S.F.*, et al.* (2016). Direct Identification of Hundreds of Expression-Modulating Variants using a Multiplexed Reporter Assay. Cell *165*, 1519-1529.

Toth-Petroczy, A., Palmedo, P., Ingraham, J., Hopf, T.A., Berger, B., Sander, C., and Marks, D.S. (2016). Structured States of Disordered Proteins from Genomic Sequences. Cell *167*, 158-170 e112.

Ulirsch, J.C., Nandakumar, S.K., Wang, L., Giani, F.C., Zhang, X., Rogov, P., Melnikov, A., McDonel, P., Do, R., Mikkelsen, T.S.*, et al.* (2016). Systematic Functional Dissection of Common Genetic Variation Affecting Red Blood Cell Traits. Cell *165*, 1530-1545.

Unger, T., Jacobovitch, Y., Dantes, A., Bernheim, R., and Peleg, Y. (2010). Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. J Struct Biol *172*, 34-44.

Wang, B., Niu, D., Lam, T.H., Xiao, Z., and Ren, E.C. (2014). Mapping the p53 transcriptome universe using p53 natural polymorphs. Cell Death Differ *21*, 521-532.

Weingarten-Gabbay, S., Elias-Kirma, S., Nir, R., Gritsenko, A.A., Stern-Ginossar, N., Yakhini, Z., Weinberger, A., and Segal, E. (2016). Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. Science *351*.

Weisz, L., Oren, M., and Rotter, V. (2007). Transcription regulation by mutant p53. Oncogene *26*, 2202-2211.

Whibley, C., Pharoah, P.D., and Hollstein, M. (2009). p53 polymorphisms: cancer implications. Nat Rev Cancer *9*, 95-107.