MCR RapidIMPACT

A Massively Parallel Fluorescence Assay to Characterize the Effects of Synonymous Mutations on *TP53* Expression

Geetha Bhagavatula[1,2], Matthew S. Rich[1], David L. Young[1], Maximillian Marin[1], and Stanley Fields[1,2,3]

Abstract

Although synonymous mutations can affect gene expression, they have generally not been considered in genomic studies that focus on mutations that increase the risk of cancer. However, mounting evidence implicates some synonymous mutations as driver mutations in cancer. Here, a massively parallel assay, based on cell sorting of a reporter containing a segment of p53 fused to GFP, was used to measure the effects of nearly all synonymous mutations in exon 6 of *TP53*. In this reporter context, several mutations within the exon caused strong expression changes including mutations that may cause potential gain or loss of function. Further analysis indicates that these effects are largely attributed to errors in splicing, including exon skipping, intron inclusion, and exon truncation, resulting from mutations both at exon–intron junctions and within the body of the exon.

These mutations are found at extremely low frequencies in healthy populations and are enriched a few-fold in cancer genomes, suggesting that some of them may be driver mutations in *TP53*. This assay provides a general framework to identify previously unknown detrimental synonymous mutations in cancer genes.

Implications: Using a massively parallel assay, this study demonstrates that synonymous mutations in the *TP53* gene affect protein expression, largely through their impact on splicing.

Visual Overview: http://mcr.aacrjournals.org/content/molcanres/15/10/1301/F1.large.jpg. *Mol Cancer Res; 15(10); 1301–7.* ©2017 AACR.

Introduction

A central pursuit of cancer genomics is to identify somatic mutations that serve as "driver" mutations, which confer a growth advantage to cells. As high-throughput sequencing technologies now enable comprehensive identification of the mutational spectrum of tumors, many recurrently mutated driver genes and mutations have been identified (1). Although significant headway has been made in identifying such mutations, the focus of these studies typically has been missense, nonsense, and frameshift alleles. However, synonymous mutations, which do not change the protein sequence of a gene product, were enriched in some oncogenes and the tumor suppressor *TP53* (2), based on an analysis of data from The Cancer Genome Atlas. This result implies that many synonymous mutations could, in fact, be drivers of oncogenesis.

Synonymous mutations can affect gene expression and protein function in multiple ways (3). Synonymous mutations have widely studied effects on splicing (4). Those that fall in splice junctions can often abrogate splicing at that site or uncover cryptic splice sites. Similarly, synonymous mutations can fall in splicing enhancers or suppressors—exonic sequences that modulate the inclusion of an exon. The rate of protein translation can be modulated by many properties of the mRNA, including secondary structure and codon bias due to limitations in the tRNA pool. These factors alter ribosomal progression through the mRNA, which can affect the efficiency of protein folding and modification, leading to a decrease in protein expression. Synonymous mutations in miRNA-binding sites can affect protein expression; for example, in Crohn's disease (5) and melanoma (6), miRNA targeting was abrogated in the mutant copy, leading to increased protein expression. Disease-associated synonymous mutations are also enriched in sites of RNA–protein interaction (7).

Considering the range of their possible effects, methods to readily annotate synonymous mutations would be useful. One avenue to identify possible drivers is via massively parallel assays to determine the effects of mutations (8). The resulting functional information yields greater predictive power than state-of-the-art *in silico* predictions of mutation function (9), like CADD (10) or PolyPhen (11). However, massively parallel assays also have largely ignored synonymous variants, often

[1]Department of Genome Sciences, University of Washington, Seattle, Washington. [2]Howard Hughes Medical Institute, University of Washington, Seattle, Washington. [3]Department of Medicine, University of Washington, Seattle, Washington.

Corresponding Author: Stanley Fields, University of Washington, Box 355065, Seattle, WA 98195. Phone: 206-616-4523; E-mail: fields@uw.edu

using these mutations as neutral variation to evaluate the significance of the changes of functional scores caused by nonsynonymous mutations (12).

Here, we demonstrate a general method to detect the effect of synonymous mutations in a cancer-associated gene. We chose *TP53*, the gene most frequently mutated in cancer, as a test case and focused on exon 6 of this gene because synonymous mutations at the terminal 3′ splice site of this exon are enriched in tumor samples (2). Although no splicing enhancer or silencer has been functionally identified in this exon, a previous study found that synonymous mutations in *TP53* are enriched in predicted splicing enhancers (13). We created a p53 protein reporter in which fluorescence was dependent on correct splicing between exons 5 and 6, and between exons 6 and 7. Using a massively parallel fluorescence-based deep mutational scanning assay, we measured the protein level of variants containing nearly all synonymous mutations in this exon. In addition to splice site–disrupting mutations, distinct internal mutations caused defects in splicing in this reporter construct. These mutations are rare in healthy populations and enriched a few-fold in cancer patients. Our methodology provides a general workflow for measuring the effects of synonymous mutations and could be used widely to uncover and catalog possible driver mutations in other cancer genes.

Materials and Methods

Primers and other oligonucleotides

A list of oligonucleotides used in this study is found in Supplementary Table S1.

Creation of a mini-gene variant library of *TP53* exon 6

We first created a reporter construct that encodes the p53 DNA-binding domain (residues 97–292) fused to EGFP. A cassette containing *TP53* exons 5–7 (chr17:7675238-7674115) was amplified from NA12892 genomic DNA (Coriell Biorepository) with TP53_Exon5,6,7_F01 and TP53_Exon5,6,7_R01 primers. This product was cleaned and reamplified using TP53_expand_F1 and TP53_expand_R1 to add flanking sequence corresponding to the 31 N-terminal amino acids (residues 97 to 127) and 31 C-terminal amino acids (residues 262 to 292). EGFP was amplified from pAG414GAL1-EGFP (14) with primers GFP_Cloning_F01 and GFP_Cloning_R03. These products were cloned by Gibson assembly (15) into the *Afl*II and *Hind*III sites in pcDNA5/FRT and transformed into chemically competent DH5α *Escherichia coli* (*E. coli*; NEB). *Bam*HI and *Nsi*I sites flanking exon 6 were added to this plasmid by inverse PCR with TP53pcDNA5_BamHI_F02 and TP53pcDNA4_NsiI_R01, followed by phosphorylation and ligation of the PCR product. This plasmid was named TP53E-GFP-pcDNA5.

We replaced *TP53* exon 6 in our reporter with a library of synonymously mutated variants of the exon. The library was created by doped oligonucleotide synthesis (TriLink BioTechnologies, San Diego). Each position that could be mutated synonymously was doped at 3% to each possible synonymous variant. For example, position 11 in the exon (the third base of a proline codon, CCT) was doped with 1% each of A, C, and G, whereas position 17 in the exon (the third base of a glutamic acid codon, CAG) was doped with 3% A. This oligonucleotide was made double-stranded by annealing the primer TP53_libraryamp_R01 and extending with Phusion polymerase (NEB), and the double-stranded DNA was amplified with TP53dopedoligo_RE_F03 and

TP53dopedoligo_RE_R02. This product was cleaned with a Zymo Clean and Concentrator 5 kit (Zymo Research). Both *TP53*E-GFP-pcDNA5 and the doped library were digested with *Nsi*I and *Bam*HI, and ligated together using T4 DNA ligase (NEB). This ligation reaction was cleaned and electroporated into ElectroMAX DH10B *E. coli* (Invitrogen) and cultured overnight in 50 mL of LB + 100 μg/mL ampicillin. This transformation resulted in a library containing approximately 242,000 colony-forming units. DNA used for transfection was harvested using a PureLink HiPure Plasmid Midiprep Kit (Invitrogen).

Specific clonal variants, used for assay validation or splicing analysis, were created by site-directed mutagenesis using inverse PCR with a mutation-containing primer, followed by phosphorylation with T4 polynucleotide kinase and T4 DNA ligase (NEB). See Supplementary Table S1 for primer sequences.

Flp-mediated integration of variant libraries

The reporter assay that we used requires that only a single variant is integrated into each cell. Thus, we took advantage of the Flp-In-293 cell line, a derivative of the human embryonic kidney cell line (HEK-293T; Invitrogen), which has a single Flp-FRT site integrated in the genome. We integrated the library of *TP53* exon 6 variants into the genome of the cell line by cotransfection with 22 μg of pOG44 (Invitrogen) and 75 μg of the *TP53* library plasmid pool using Lipofectamine 3000 (Thermo Fisher). Media were changed after 24 hours, and selection for flipped-in cells was started after 48 hours by adding 100 μg/mL hygromycin. Cells were cultured under hygromycin selection for 14 days, at which point they were harvested. The 293 cell line library contained approximately 12,000 colony-forming units.

Fluorescence-activated cell sorting

The amount of reporter protein for each synonymous variant was measured by fluorescence-activated cell sorting. Library cells were washed, trypsinized, and resuspended in PBS + 1 μg/mL DAPI. Libraries were sorted on a FACSAria (BD Biosciences) into four bins, and cells were collected in PBS. Gates were set by comparing the distributions of fluorescence from a cell line expressing either a negative control (G113A) or the wild-type exon. The lowest gate was set to bisect the negative distribution. The highest gate was set to bisect the wild-type distribution. A final gate was set to bisect the range between the other two gates.

Sequencing library preparation and analysis

Variant frequencies were calculated by high-throughput sequencing of the integrated library in each sorted bin. Genomic DNA and total RNA were extracted from sorted cells using Trizol. Library variants were amplified from genomic DNA using three sequential reactions. Phusion polymerase was used for all reactions, and each reaction was performed on a BioRad MiniOpticon and monitored to avoid overamplification.

The integrated library was amplified from 250 ng of genomic DNA (to avoid the endogenous *TP53* locus), using primers specific for library locus (pcDNA5/FRT_seq_F01 and pcDNA5/FRT_seq_R01) and the following PCR conditions: 98°C for 30 seconds, then 25 cycles of 98°C for 10 seconds, 62°C for 30 seconds, and 72°C for 65 seconds. This reaction was cleaned with a Zymo Clean and Concentrator 5 kit. 100 ng of this product was used as template for a second PCR to amplify a

short fragment containing only exon 6, using oligos TP53_exon6amp_F01 and TP53_exon6amp_R01 and the following PCR conditions: 98°C for 30 seconds, then 8 to 10 cycles (depending on the sample) of 98°C for 10 seconds, 65°C for 30 seconds, and 72°C for 5 seconds. This reaction was cleaned as before. Sequencing and demultiplexing adaptors were appended by a final PCR. A total of 150 ng of the previous PCR product was amplified with TP53_seqadapter_F1 and TP53_seqadapter2_R1-4 and the following PCR conditions: 98°C for 30 seconds, then 8 to 10 cycles (depending on the sample) of 98°C for 10 seconds, 67°C for 30 seconds, and 72°C for 5 seconds. Reactions were cleaned and sequenced on an Illumina MiSeq using paired-end 160 base reads, using TP53_libraryseq_F01 as a forward primer, TP53_libraryseq_R02 as a reverse primer, and TP53_indexseq_F01 as an index read primer.

The effect of each variant on fluorescence was calculated as a bin-weighted sum of inferred cell counts for each variant. Variant counts in each bin were tabulated using Enrich (16), and each variant's frequency in a bin was used to estimate the number of cells sorted in that bin for that variant. The frequency of total cell counts in each bin was multiplied by that bin's median GFP intensity. The weighted frequencies were summed to calculate a GFP score, and this score was divided by the GFP score of the wild-type sequence to create a normalized GFP score. The normalized GFP scores for each of two replicates were averaged to compute the score reported here. Nonsynonymous variants (likely due to synthesis or sequencing errors) were removed after GFP score calculations. We filtered variants with fewer than 7,500 total cell counts in each replicate, as this threshold allows for the most variants with a high correlation between replicate GFP scores (see Supplementary Fig. S1).

Downstream analyses were performed in Python. Supplementary File S1 contains all pre- and postprocessed data. An iPython notebook containing analyses is supplied in Supplementary File S2. Raw sequencing reads can be found in SRA Bioproject PRJNA384242.

Splicing analysis by PCR

RNA was extracted from clonal cell lines, reverse transcribed, and amplified. These products were separated by polyacrylamide gel electrophoresis. Fragments of interest were gel-extracted, reamplified with KAPA Robust2G (KAPA Biosystems) using primers TP53_exon6amp_F01 and TP53_exon6amp_R01, cloned into pGEM-T (Promega), and Sanger sequenced to identify isoforms.

Results

To test the effects of synonymous variants in exon 6 of *TP53*, we developed a fluorescent reporter assay based on FlowSeq (17), which uses FACS sorting to separate variants of differing expression levels followed by next-generation DNA sequencing to calculate a score for each variant based on its enrichment or depletion in each FACS bin. The reporter consisted of a cassette containing the genomic sequence spanning exons 5–7 of *TP53* (encoding residues 126 to 261), as well as flanking cDNA encoding the rest of the p53 DNA-binding domain, such that in total the translated protein product of the cassette spanned residues 97–292 of p53 (Fig. 1A). The C-terminus of this cassette was fused to GFP, and the construct was expressed from a constitutive *CMV* promoter. The pre-mRNA from this construct can undergo only two correct splicing events, fusing exons 5 to 6 and exons 6 to 7 (Fig. 1A). We created a library of exon 6 synonymous variants that fall in the third position of a codon by doped oligonucleotide synthesis (Fig. 1B), and cloned these variants in the place of the wild-type exon (i.e., chr17:7578289-7578177) in the expression construct. The library was integrated into the genome of Flp-In-293 cells to ensure that a single exon variant was expressed per cell.

Median fluorescence in this assay spanned an approximately 20-fold range, from a low due to a known splice site–disrupting mutation (G113A; Fig. 1C, red) to a high due to wild type (Fig. 1C, blue). The population of cells in the Flp-In library had a large peak overlapping the G113A peak, and a large shoulder of cells with intermediate fluorescence (Fig. 1C, green). We sorted these cells into four bins that were established based on the fluorescence properties of the G113A-mutant and the wild type, sequenced the exon 6 integrated in the cells of each bin, and calculated a score for each variant as a wild-type–normalized weighted sum of GFP intensity across the bins (S_{GFP}, see Materials and Methods). In total, we calculated S_{GFP} for 16,916 synonymous variants of the exon (including multiply-mutated variants), which averaged 3.2 mutations per variant (Supplementary Fig. S1A). We focused our analysis on single synonymous mutants in the exon, of which we measured 76 of the 77 expected synonymous mutations. We tested the fluorescence of 20 variants clonally and found significant correlation with our pooled data ($P = 0.0002$, Supplementary Fig. S2A).

S_{GFP} scores ranged from 0.47 to 1.31 (Fig. 2A). Sixty-eight variants were neutral (considered here as having S_{GFP} 70% or above relative to wild type), but a subset of nine mutations caused a large decrease (S_{GFP} less than 70% of wild type). As expected, two of the mutations—T2G (p.Gly187) and G113A (p.Glu224)—disrupting either of the exon's canonical splice sites caused at least a 30% decrease in S_{GFP}. In addition, seven other mutations in the interior of the exon—G32A (p.Val197), A35G (p.Glu198), A38T (p.Gly199), T47G (p.Arg202), G59A (p.Leu204), G95T (p.Val218), and G107C (p.Pro222)—had S_{GFP} scores less than 0.7. We mined our data for doubly mutated variants that contained these mutations and found that, with the exception of T47G, the negative effects of the single mutants were epistatic to other synonymous mutations in the exon (Supplementary Fig. S2B).

Using *in silico* predictions, we asked whether mutations in the library of exon 6 variants would have predicted effects on splicing (18), RNA structure (19), or miRNA targeting (20), or were predicted to have strong functional effects (10). Of these, only splicing prediction yielded a significant correlation with experimental results ($P = 1.3e{-}4$, Supplementary Fig. S3). Increased inclusion of exon 6 results in increased GFP intensity, whereas skipping of exon 6 results in GFP being out of frame and therefore decreased GFP intensity. As such, splicing defects would be apparent in our assay. To examine splicing of the reporter construct, we extracted RNA from 22 clonal samples and used RT-PCR to amplify the region between exons 5 and 7 (Fig. 2C and Supplementary Fig. S4). The wild-type sample contained two major isoforms: a correctly spliced 196 base product and an exon 6–skipped 83 base product. These two fragments were present in varying intensities across all 22 samples. In addition, mutant samples contained an unspliced, 845 base product that includes both flanking introns. As the 845 base product was not present in the sample from cells

Bhagavatula et al.
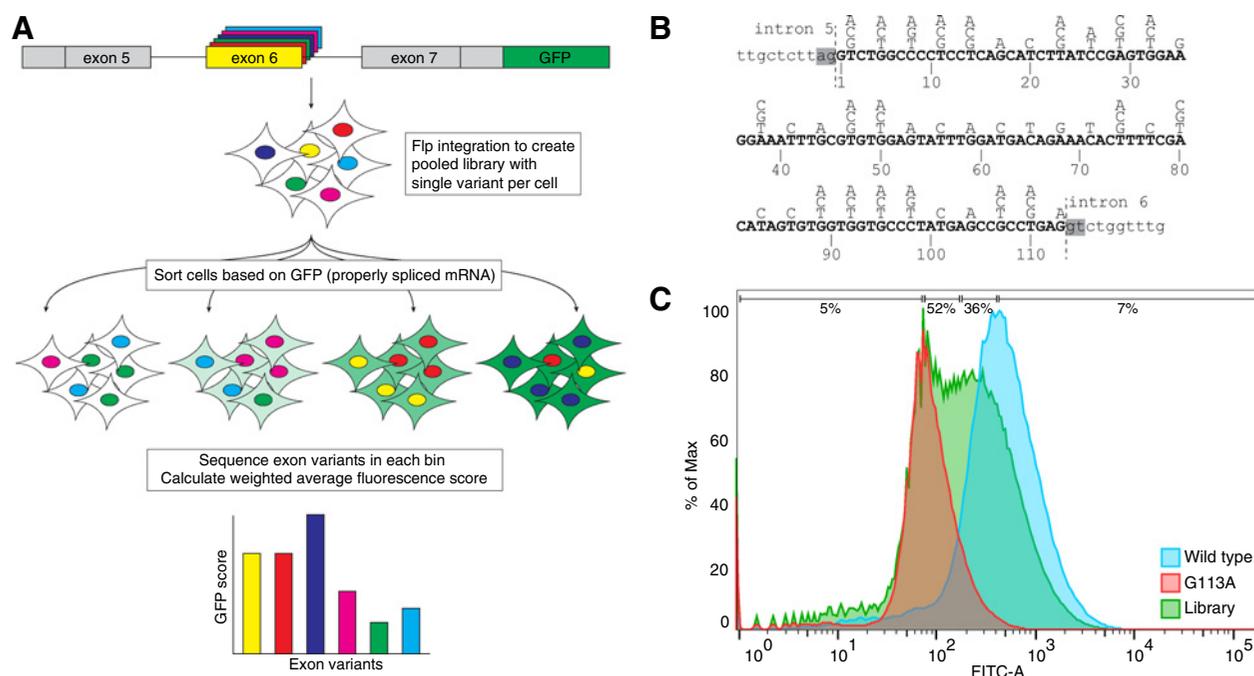
Figure 1.
A massively parallel assay to measure the splicing effects of synonymous mutations in *TP53*. **A,** Overview of the assay. A library of synonymous variants is created in a *TP53* reporter construct. This library is integrated into the genome of Flp-In-293 cells, a derivative of the human embryonic kidney cell line (HEK-293T) using Flp-mediated recombination. Cells are sorted into bins based on GFP intensity, and the exon library from each bin is sequenced and genotyped. S_{GFP} is calculated as a weighted sum based on the frequency of each variant in each sorted bin. **B,** Library design. Positions in the mutagenized exon (bold) are numbered, and degenerate positions are annotated with the possible mutant bases above the wild-type sequence. The known splice donor and acceptor sites are highlighted with gray boxes. **C,** Flow cytometry of *TP53* variant library. Cells were sorted based on GFP signal (*x* axis); blue, cells expressing a wild-type exon 6; red, negative control cells expressing the G113A variant of exon 6; green, *TP53* exon 6 synonymous variant library. Bins are represented by lines at top, annotated with the percentage of sorted cells in each bin.

expressing the wild-type exon, it most likely derived from unspliced mRNA, rather than trace genomic DNA contamination.

S_{GFP} quantifies the amount of p53 reporter protein present, independent of the amounts of reporter RNAs that do not give rise to detectable fluorescent protein. However, a lower proportion of correctly spliced RNA compared with other RNAs would be expected to lead to less reporter protein and thus lower S_{GFP}. The approximate intensity of the 845 base and 83 base isoforms relative to the correctly spliced 196 base isoform observed from each mutant was consistent with its S_{GFP}. For example, the known T2G splice junction mutant appeared enriched in the amount of 83 base exon-skipped product relative to the correctly spliced 196 base isoform (Fig. 2C). Eighteen of 21 mutants also contained a 277 base fragment, which includes intron 5. The intensity of this band decreased in variants with mutations farther from the 5′ splice site, consistent with this isoform being due to disruption of this splice site. The G113A splice junction mutation, which inactivates the 3′ splice site of the exon, resulted in the use of a cryptic site that shifted the splice site 5 bases (2). The A38T mutation produced a unique truncated splice isoform of 117 bases (Supplementary Fig. S4) by creating a consensus splice donor. This isoform contains the first 36 bases of exon 6, causing a frameshift and premature stop codon 8 bases later. T47G created an approximately 180 base-truncated isoform (Supplementary Fig. S4). We confirmed the sequence of these isoforms by cloning and

Sanger sequencing the cDNAs (data not shown). Many mutants contained additional alternatively spliced fragments ranging in size between approximately 350 and 600 bases, which were not easily identifiable.

Based on the abundance of the cDNA products, some mutations (most notably A38T and G95T) appeared to result in significantly less mRNA relative to other variants. We did not measure RNA stabilities of these variants, but note that RNA degradation pathways, especially the nonsense-mediated decay pathway, may affect these mRNAs differentially.

We hypothesized that if the mutations identified in our assay were truly detrimental, they should be at low frequencies in healthy populations but present in tumor samples. We searched the genomes and exomes of healthy individuals using the gnomAD database (ref. 21; Supplementary Table S2), for variants in exon 6 of *TP53*. Eleven synonymous variants in exon 6 were found in the database, with 10 of the 11 variants present in fewer than 37 individuals (out of approximately 250,000 total haploid exomes and genomes). Only one of these was a detrimental variant, G59A ($S_{GFP} = 0.57$), which was found in 9 heterozygous individuals. One neutral variant in our assay (A80G, $S_{GFP} = 1.02$) was found at a high population frequency (1.2%), and was homozygous in 33 individuals.

We also searched for synonymous mutations in the COSMIC database (in which *TP53* was sequenced in 127,779 samples) to estimate the frequencies of mutations in tumor samples (22).
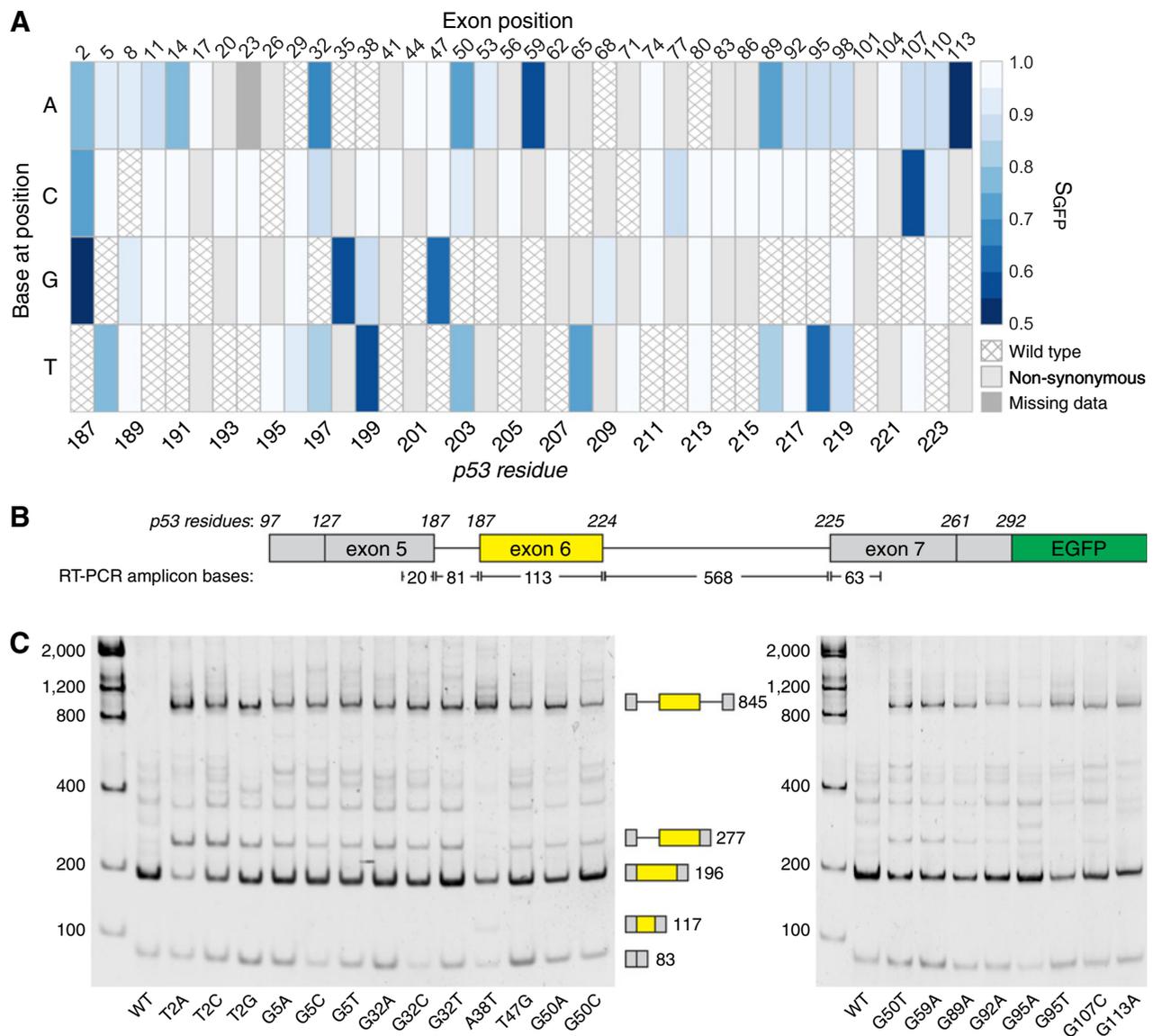
Figure 2.
Synonymous mutations cause mis-splicing. **A,** S_{GFP} scores for synonymous mutations in exon 6 of *TP53* plotted as a heatmap. At each codon's third position (*x* axis), the base present in the exon variant (*y* axis) is colored based on S_{GFP}. The wild-type base at each position is hatched, nonsynonymous mutations are light gray, and missing data are dark gray. The p53 residue of each mutation is noted below the heatmap. **B,** Model of the reporter construct. Amino acid residues in p53 are shown in italics above the model. Residue 188 spans the splice junction between exons 5 and 6. Lengths of relevant exonic and intronic fragments are below the model. **C,** Polyacrylamide gel showing cDNA for clonal samples. TP53 exon 6 mutations are shown below each well. Cartoon representations of major fragments and their lengths, matching the model in **B**, are shown between the two panels. Samples were split between two gels.

Thirty-seven synonymous variants were found in COSMIC in 79 samples, ranging from 1 to 10 samples per mutation. Four of these variants were detrimental in our assay—G32A (found in 1 carcinoma and 1 lymphoid neoplasm), G59A (1 lymphoid neoplasm), G107C (2 carcinomas and 1 benign melanocytic nevus), and G113A (9 carcinomas and 1 glioma). Thirty-three others were neutral ($S_{GFP} > 0.7$) and present in 63 samples, with T74C and A80G both found in 8 samples. The 33 neutral mutations came from 39 carcinomas, 11 hematopoietic cancers, 5 sarcomas, 2 gliomas, 2 malignant melanomas, and 4 others (Supplementary Table S2). The 3.4x enrichment of detrimental mutations in COSMIC compared with gnomAD correlates with the analysis of samples from The Cancer Genome Atlas showing that synonymous mutations were enriched in *TP53* approximately 4-fold compared with a matched, nononcogenic gene set (2).

Discussion

Using a massively parallel approach, we measured the effects of nearly all synonymous mutations in exon 6 of *TP53*, some of

Bhagavatula et al.

which are potentially loss of function. These mutations appeared to alter the splicing of the exon and surrounding introns, consistent with an accumulation of unspliced pre-mRNA or mis-spliced products that included introns. As a result, less full-length p53 protein would be expressed compared with the construct containing the wild-type sequence of exon 6. Our assay measures mutations not in their native genomic context in *TP53*, and therefore mutation effects should be verified by assaying for splicing effects in the native gene context. The synonymous mutations are found at low frequencies in healthy human populations and are a few-fold more prevalent in tumor samples. Thus, synonymous mutations—including mutations that fall at canonical splice sites as well as at least 6 bases away from a splice site—may be drivers of oncogenesis.

A qualitative effect on splicing of the reporter gene was evident for almost all of the severe mutations. As a general trend, variants with low fluorescence scores showed increased intensity for the unspliced 845 base or exon-skipped 83 base isoform relative to the intensity of the correctly spliced 196 base isoform, although some variants did not follow this trend. Our analysis here does not take RNA stability into account, although some variants (e.g., A38T and G95T) appeared to accumulate less RNA than most.

It is possible that our reporter construct, which has cDNA rather than genomic sequence flanking exons 5 and 7, differs in its splicing properties compared with the genomic copy of *TP53*. Although the wild-type version of this construct spliced correctly, we cannot rule out the possibility that the effects of mutations in our construct may be more or less severe if they were present in the genomic copy. The artificial nature of the construct may partly explain the accumulation of the 845 base isoform in the mutants, as this isoform may not enter the splicing pathway, rather than being mis-spliced. However, it would be possible in future experiments to use endogenous gene structures or perform mutagenesis *in situ*, which could mitigate these effects.

A p53-negative tumor arises from two inactivating mutations of *TP53*, such as by mutation in one copy and complete deletion in the other (23). Could synonymous mutations be responsible for one of these "hits"? In most of the variants we assayed, the amount of exon-included splice isoform was qualitatively less abundant than in wild-type cells but still present. As the tumor-suppressor function of p53 is known to be haploinsufficient (24), a reduction in p53 protein from one allele, even if not complete, could render the cell susceptible to oncogenesis in the event of a mutation to the other allele. Based on our splicing data, it might be expected that several of the synonymous mutations would cause a decrease in total p53 protein levels. Decreased p53 activity is found in many sarcomas (25), yet only five sarcoma samples had exon 6 synonymous mutations, and none of these mutations were detrimental in our reporter assay.

Alternatively, variants of *TP53* in which the protein is truncated in exon 6 or before exon 7 have been shown to promote cell proliferation and metastasis through a gain-of-function activity (26, 27). Synonymous mutations in exon 6 that cause p53 truncations would produce similarly structured proteins and may also have gain-of-function activities. These include mutations that lead to inclusion of intron 6, which creates a premature stop codon 39 bases into the intron. The A38T mutation mis-splices to delete the last 77 bases of exon 6 and leads to termination 12 bases after the aberrant splice junction with exon 7. The majority of exon 6 synonymous mutations found in the COSMIC database were from carcinoma samples, a cancer type in which dominant negative p53 mutations are common (28). Whether these isoforms are gain-of-function or dominant-negative variants remains to be tested.

Our assay measured the effects of mutations in a single *TP53* exon on expression, but the method should be general to studying other exons in *TP53* or other tumor suppressor or oncogenes. The major effects we found in our assay were due to splicing errors. Although other studies have directly measured splice isoforms of libraries of mutated exons using massively parallel RNA sequencing (18, 29), which may be a more quantitative assay than our sorting methodology, these methods do not account for other mechanisms by which synonymous mutations can affect protein expression. Our approach, on the other hand, is agnostic to the molecular mechanism of protein expression variation and should allow for a more complete understanding and characterization of synonymous mutations.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: G. Bhagavatula, D.L. Young, S. Fields
Development of methodology: G. Bhagavatula, D.L. Young
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): G. Bhagavatula, M.S. Rich, D.L. Young
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): G. Bhagavatula, M.S. Rich, D.L. Young, M. Marin
Writing, review, and/or revision of the manuscript: M.S. Rich, S. Fields

References

1. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. Science 2013;339:1546–58.
2. Supek F, Miñana B, Valcárcel J, Gabaldón T, Lehner B. Synonymous mutations frequently act as driver mutations in human cancers. Cell 2014; 156:1324–35.
3. Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet 2011;12:683–91.
4. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 2002; 3:285–98.
5. Brest P, Lapaquette P, Souidi M, Lebrigand K, Cesaro A, Vouret-Craviari V, et al. A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. Nat Genet 2011;43:242–5.

6. Gartner JJ, Parker SC, Prickett TD, Dutton-Regester K, Stitzel ML, Lin JC, et al. Whole-genome sequencing identifies a recurrent functional synonymous mutation in melanoma. Proc Natl Acad Sci U S A 2013;110:13481–6.

7. Silverman IM, Li F, Alexander A, Goff L, Trapnell C, Rinn JL, et al. RNase-mediated protein footprint sequencing reveals protein-binding sites throughout the human transcriptome. Genome Biol 2014;15:R3.

8. Fowler DM, Fields S. Deep mutational scanning: a new style of protein science. Nat Methods 2014;11:801–7.

9. Starita LM, Young DL, Islam M, Kitzman JO, Gullingsrud J, Hause RJ, et al. Massively parallel functional analysis of BRCA1 RING domain variants. Genetics 2015;200:413–22.

10. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310–5.

11. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–9.

12. Melamed D, Young DL, Gamble CE, Miller CR, Fields S. Deep mutational scanning of an RRM domain of the Saccharomyces cerevisiae poly(A)-binding protein. RNA 2013;19:1537–51.

13. Lamolle G, Marin M, Alvarez-Valin F. Silent mutations in the gene encoding the p53 protein are preferentially located in conserved amino acid positions and splicing enhancers. Mutat Res 2006;600:102–12.

14. Alberti S, Gitler AD, Lindquist S. A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. Yeast 2007;24:913–9.

15. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 2009;6:343–5.

16. Fowler DM, Araya CL, Gerard W, Fields S. Enrich: software for analysis of protein function by enrichment and depletion of variants. Bioinformatics 2011;27:3430–1.

17. Kosuri S, Goodman DB, Cambray G, Mutalik VK, Gao Y, Arkin AP, et al. Composability of regulatory sequences controlling transcription and translation in Escherichia coli. Proc Natl Acad Sci U S A 2013;110:14024–9.

18. Rosenberg AB, Patwardhan RP, Shendure J, Seelig G. Learning the sequence determinants of alternative splicing from millions of random sequences. Cell 2015;163:698–711.

19. Lorenz R, Bernhart SH, Höner Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol 2011; 6:26.

20. Miranda KC, Huynh T, Tay Y, Ang Y-S, Tam W-L, Thomson AM, et al. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 2006;126:1203–17.

21. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–91.

22. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res 2015;43:D805–11.

23. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759–67.

24. Berger AH, Pandolfi PP. Haplo-insufficiency: a driving force in cancer. In: Ladanyi M, Hogendoorn PC, editors. J Pathol 2011;223:137–46.

25. Vogelstein B, Kinzler KW. p53 function and dysfunction. Cell 1992;70: 523–6.

26. Senturk S, Yao Z, Camiolo M, Stiles B, Rathod T, Walsh AM, et al. p53ψ is a transcriptionally inactive p53 isoform able to reprogram cells toward a metastatic-like state. Proc Natl Acad Sci U S A 2014;111: E3287–96.

27. Shirole NH, Pal D, Kastenhuber ER, Senturk S, Boroda J, Pisterzi P, et al. TP53 exon-6 truncating mutations produce separation of function isoforms with pro-tumorigenic functions. Elife 2016;5:e17929.

28. Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P, et al. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. Hum Mutat 2007;28:622–9.

29. Mueller WF, Larsen LS, Garibaldi A, Hatfield GW, Hertel KJ. The silent sway of splicing by synonymous substitutions. J Biol Chem 2015;290: 27700–11.

Molecular Cancer Research

A Massively Parallel Fluorescence Assay to Characterize the Effects of Synonymous Mutations on *TP53* Expression

Geetha Bhagavatula, Matthew S. Rich, David L. Young, et al.

Updated version	Access the most recent version of this article at: doi:10.1158/1541-7786.MCR-17-0245

Visual Overview	**A diagrammatic summary of the major findings and biological implications:** http://mcr.aacrjournals.org/content/15/10/1301/F1.large.jpg

Cited articles	This article cites 29 articles, 7 of which you can access for free at: http://mcr.aacrjournals.org/content/15/10/1301.full#ref-list-1

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://mcr.aacrjournals.org/content/15/10/1301. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.