Somatic and Germline TP53 Alterations in Second Malignant Neoplasms from Pediatric Cancer Survivors

Amy L. Sherborne1, Vincent Lavergne1, Katharine Yu1, Leah Lee1, Philip R. Davidson2, Tali Mazor3, Ivan V. Smirnoff3, Andrew E. Horvai4, Mignon Loh5, Steven G. DuBois6, Robert E. Goldsby5, Joseph P. Neglia7, Sue Hammond8, Leslie L. Robison9, Rosanna Wustrack10, Joseph F. Costello3, Alice O. Nakamura2, Kevin M. Shannon3, Smita Bhatia11, and Jean L. Nakamura1

Abstract

Purpose: Second malignant neoplasms (SMNs) are severe late complications that occur in pediatric cancer survivors exposed to radiotherapy and other genotoxic treatments. To characterize the mutational landscape of treatment-induced sarcomas and to identify candidate SMN-predisposing variants, we analyzed germline and SMN samples from pediatric cancer survivors.

Experimental Design: We performed whole-exome sequencing (WES) and RNA sequencing on radiation-induced sarcomas arising from two pediatric cancer survivors. To assess the frequency of germline TP53 variants in SMNs, Sanger sequencing was performed to analyze germline TP53 in 37 pediatric cancer survivors from the Childhood Cancer Survivor Study (CCSS) without any history of a familial cancer predisposition syndrome but known to have developed SMNs.

Results: WES revealed TP53 mutations involving p53’s DNA-binding domain in both index cases, one of which was also present in the germline. The germline and somatic TP53-mutant variants were enriched in the transcriptomes for both sarcomas. Analysis of TP53-coding exons in germline specimens from the CCSS survivor cohort identified a G215C variant encoding an R72P amino acid substitution in 6 patients and a synonymous SNP A639G in 4 others, resulting in 10 of 37 evaluable patients (27%) harboring a germline TP53 variant.

Conclusions: Currently, germline TP53 is not routinely assessed in patients with pediatric cancer. These data support the concept that identifying germline TP53 variants at the time a primary cancer is diagnosed may identify patients at high risk for SMN development, who could benefit from modified therapeutic strategies and/or intensive posttreatment monitoring.

Introduction

Epidemiologic data indicate that there will be 500,000 survivors of pediatric cancers in the United States by 2020 (1). Long-term follow-up of the expanding population of pediatric cancer survivors have identified substantial late toxicities resulting from intensive treatment regimens (2–6). Second malignant neoplasms (SMNs) arise after exposure to genotoxic therapies, which include radiotherapy and some chemotherapeutic agents. Sarcomas were the first histologic cancer type recognized after clinical radiotherapy (7), and continue to arise as SMNs in pediatric cancer survivors treated with modern regimens (6, 8). Compared with their de novo counterparts, SMNs are often clinically more aggressive, more difficult to treat, and associated with worse outcomes (3, 4, 9).

This risk is elevated 35 years after treatment and appears to persist in patients treated in the moderm era (10). Whereas epidemiologic studies have identified exposure to radiation and young age at the time of therapy as major risk factors, the molecular basis for SMN development remains poorly understood.

To characterize the mutational landscape of treatment-induced sarcomas and to identify candidate SMN-predisposing germline and somatic variants, we performed whole-exome sequencing (WES) of radiation-induced sarcoma SMNs and matched germline control DNA from two pediatric cancer survivors. RNA sequencing (RNA-Seq) validated WES-identified variants at the transcriptomic level. Both SMNs exhibited a similar pattern of base substitutions, and contained mutations in the tumor suppressor gene TP53, which was somatic in one patient and present in the germline of the other. Loss of normal
TP53 Alterations in Second Malignant Neoplasms

Translational Relevance

Most pediatric cancer survivors who develop second malignant neoplasms (SMNs) have no known tumor predisposition syndrome, and accurately identifying which individuals are at greatest risk for SMN development is currently challenging. This work presents next-generation sequencing of exomes and transcriptomes of SMNs to resolve the role of TP53 in SMNs at a level of molecular detail not previously shown. We identify somatic and germline TP53 mutations in sarcoma SMNs, and in a cohort of pediatric cancer survivors known to develop SMNs, a TP53 polymorphism that is significantly more frequent than the estimated frequency in the general population. These findings strongly suggest that analyzing the TP53 gene for variants, both synonymous and nonsynonymous, may be a viable strategy for identifying individuals at the greatest risk of SMN formation. Our work suggests the utility of germline analysis at the time of diagnosis to actualize personalized medicine and offer stratified cancer survivorship care.

Materials and Methods

Human research protection

All work was performed under a research protocol approved by the University of California San Francisco (UCSF) Committee on Human Research (IRB protocol 11-07304).

Microscopy

Pathologic review was performed on hematoxylin and eosin–stained (H&E) sections by A.E. Horvai. Photographs of histology were taken with an Olympus BX41 microscope, using Olympus UplanFl 10×/0.3 and 20×/0.5 objectives. An attached Olympus DP72 camera and Adobe Photoshop CS2 were used to capture the images.

Sample preparations

The two SMNs analyzed by whole-exome and transcriptome sequencing were isolated at UCSF (San Francisco, CA). Genomic DNA was isolated from fresh-frozen tumor samples and fresh peripheral blood samples using previously described techniques (11). RNA was isolated from fresh-frozen tumor samples using the RNeasy RNA Isolation Kit (Qiagen). H&E-stained sections of formalin-fixed or fresh-frozen tumor samples were reviewed by a pathologist (A.E. Horvai), who estimated greater than 90% tumor cell representation in the tumor samples.

WES

WES was performed using the NimbleGen Human exome v3.0 kit. Captured material was indexed and sequenced on the Illumina GAII and HiSeq2000 platform at the Institute for Human Genetics at UCSF (San Francisco, CA). Successfully sequenced reads were then mapped to the human reference genome (GRCh37) using GATK best practices, and somatic mutations were called using MuTect and Pindel as described previously (12). ANNOVAR was used for variant annotation and to filter known human mutations to produce a cleaner list of variants (Supplementary Files S1 and S2). Mutation Assessor, SIFT, and PROVEAN were used to assign functional predictions. Validation of a subset of somatic variants by Sanger sequencing confirmed 92% of SNVs (23/25 tested).

RNA-Seq

cDNA was synthesized from RNA, followed by library preparation using the Illumina TruSeq Stranded Kit with Ribodepletion-Gold (Illumina) according to the manufacturer’s instructions. Paired-end sequencing (100 bp) was performed on a HiSeq4000. Reads were aligned using Bowtie (13) and TopHat2-Fusion (14), which included analysis for fusion transcripts. Only fusions with >100 paired-end reads supporting the mapping were considered. Those with <5 individual reads spanning the fusion, or those with >10 reads contradicting the fusion were also removed. Tophat-fusion-post was used to filter out false fusions due to highly similar sequences or pseudogenes. The 50 base pairs on the left side of a fusion and the 50 base pairs on the right side are combined, and then BLASTed (15) against the human blast database. If match length (range: 0–100) + identity percent (0 to 100) was greater than 160, the fusion was filtered out, which resulted in no predicted fusions being retained for either patient. Expression levels were estimated using Cufflinks (16). TP53 expression for normal muscle control was obtained from the Genotype Tissue Expression (GTEx) project (17).

Sanger sequencing

Sanger sequencing was performed on exons 2–11 of TP53 (detailed in Supplementary File S4, with primer sequences listed in Supplementary File S5) in the germline DNA of CCSS samples.

Copy-number analysis

Using the exome sequencing data, we identified regions of copy-number variation (CNV) in tumor DNAs relative to the matched normal using Control-FREEC with default parameters (18). Analysis was restricted to the region selected by the Human exome v3.0 capture kit (NimbleGen). CNVs were smoothed into 2.5-kb segments. Significance of Control–FREEC predictions were assessed by assigning Wilcoxon test and Kolmogorov–Smirnov test P values in R.

Loss of heterozygosity analysis

SNPs were called using UnifiedGenotyper (19). Only those that were heterozygous in the germline, present in dbSNP (build 132), and had 10 or more spanning reads were selected for analysis. Allelic frequency in the tumor, as represented by the proportion of reads supporting the A and B alleles, were plotted across chromosomes. Regions demonstrating loss of heterozygosity are represented by migration of the allele frequencies away from 0.5, which would represent approximately equal read depth of the maternal and paternal alleles.

www.aacrjournals.org Clin Cancer Res; 23(7) April 1, 2017 1853

© 2017 American Association for Cancer Research. Published OnlineFirst September 28, 2016; DOI: 10.1158/1078-0432.CCR-16-0610
Identification and classification of driver mutations

All somatic variants were compared against a list of potential driver genes \(\left(n = 573\right)\) composed of all genes identified in the COSMIC cancer gene census (October 2015; ref. 20). Using the criteria outlined by Muruganui and colleagues (21) we categorized variants as category 1 (high confidence driver mutation), category 2 (putative driver mutation), category 3 (low confidence driver mutation), or category 4 (representing a variant of unknown significance).

Statistical analysis

Fisher exact test was used to determine whether the number of A639G substitutions in the SMN sample (4 mutations in 37 tissue samples for which chemical reactions succeeded out of 41 samples total) was higher than the frequencies observed in the 1000 Genomes Project (5 mutations in 1,000 genomes sequenced). A one-sided \(P\) value of 0.00015 was obtained by direct calculation using standard Microsoft Excel functions.

Results

Clinical case histories

Patient 1 was an 11-year-old male with a pelvic embryonal rhabdomyosarcoma who was treated with chemotherapy and external beam radiotherapy as prescribed by Children’s Oncology Group protocol D9803 (22). Radiotherapy to the bladder delivered 41.4 Gy in 23 fractions (Supplementary Fig. S1A). Nine years after his initial therapy, an MRI demonstrated a new 6-cm × 6-cm mass in the right pubic component of the acetabulum (Supplementary Fig. S1B). A biopsy of the mass revealed pathology consistent with a high grade (3/3) chondroblastic osteosarcoma (Supplementary Fig. S1C).

Patient 2 was a 17-year-old male when first diagnosed with chondroblastic osteosarcoma of the pelvis and treated with neo-adjuvant chemotherapy followed by resection. Adjuvant intensity–modulated radiotherapy delivered 60 Gy in 25 fractions to the right pubic ramus (Supplementary Fig. S1D). Seven years after the initial diagnosis, he developed a right lateral thigh mass within the previously irradiated volume (Supplementary Fig. S1E). Resection of this tumor revealed pathology consistent with a grade 3/3 deep pleomorphic sarcoma (Supplementary Fig. S1F).

DNA and RNA isolated from freshly frozen, unfixed tumor samples of both patients were analyzed by WES and RNA-Seq, respectively. Freshly collected peripheral blood from each patient served as germline control and was analyzed by WES. While tumor material received from patient 1 was insufficient for intratumor heterogeneity to be investigated, patient 2’s tumor sample was sufficiently large enough to enable sequencing of two physically separate regions (termed “a” and “b”.

TP53 mutations in SMN specimens

TP53 is a tumor suppressor gene that defines the Li-Fraumeni tumor predisposition syndrome and also is commonly mutated in human malignancies (23, 24). WES of germline and tumor DNA from patient 1 identified a somatic TP53 mutation (rs28934574, c.844C>T, p.Arg282Trp; Fig. 1A). This mutation involves a recognized “hot spot” residue in p53’s DNA-binding domain, and is among the most common TP53 mutations found in human cancers (25). Germline analysis of patient 1 revealed no deleterious variants or mutations. In addition, there was no history of malignancy in patient 1’s parents, sibling, or grandparents, and thus no genetic testing was performed for his parents.

In contrast, patient 2 harbored a heterozygous germline mutation in the TP53 gene that has been previously associated with Li-Fraumeni syndrome (rs11540652, c.743G>A, p.Arg248Gln; Fig. 1B, refs. 26, 27). This nonsynonymous germline variant, which was also confirmed in a CLIA-approved laboratory, occurs in exon 7 and is classified as a Class I functional mutation (28) involving a residue that contacts the minor groove of the DNA helix. These types of pathogenic mutations are distinguished from Class II mutations, which destabilize the structural integrity of the DNA-binding domain. Both exomes sequenced from patient 2’s SMN (referred to as samples “a” and “b”) showed enrichment of the germline TP53 mutation, as evidenced qualitatively by Sanger sequencing chromatograms and quantitatively by read count (Fig. 1B). Importantly, patient 2 was not suspected of having a familial cancer predisposition syndrome at initial diagnosis and had not been referred for genetic evaluation before the SMN diagnosis. There was no history of malignancy in patient 2’s parents or siblings, and no genetic testing was performed for his parents.

Somatic variants in radiation-induced SMNs

Deep WES was performed for SMNs (147-fold mean exome coverage, with 99% of the exome covered by ≥ 10 reads, and 90% covered by ≥ 50 reads in all samples; Supplementary Fig. S2). A total of 202 somatic single-nucleotide variants (SNV) were identified in patient 1’s sarcoma SMN (including 41 nonsynonymous and 8 small indels; Supplementary File S1 and Supplementary Fig. S3). In patient 2, the SMN demonstrated a total of 268 somatic SNVs (102 SNVs private to sample A, 55 SNVs private to sample B, and 111 SNVs present in both tumor samples). Sixteen indels were identified, including 8 and 5 that were unique to samples a and b, respectively and two that were present in both (Supplementary File S2). SNVs were characterized by a predominance of C → T and G → A substitutions (Supplementary Fig. S3). Sixteen dinucleotide substitutions were identified across all tumor exomes, of which 10 were unique (Supplementary File S3). Nonsynonymous SNVs alter coding sequences and may have diverse consequences. ANNOVAR software (29) was used to annotate variants and estimate the functional impact of nonsynonymous somatic variants arising in the sarcoma specimens. We categorized nonsynonymous SNVs according to their likelihood of being a driver mutation as described previously (Materials and Methods; ref. 21). In addition to the high confidence TP53 driver mutations described above, patient 2 also had a putative driver mutation (Category 2 in NCOA2, a gene encoding a transcriptional coactivator for nuclear receptors [30] known to be altered in sarcomas [31]. Only one somatic variant from patient 1’s tumor was annotated to a gene in COSMIC (AFF3), and this variant was predicted to be functionally neutral.

Copy number/loss of heterozygosity

Both SMN exomes were characterized by copy-number variations (Fig. 1C and D and Supplementary Fig. S4). Each of the SMNs analyzed harbored large-scale (>25% of chromosome arm) alterations and focal chromosomal changes (Supplementary Files S1 and S2). Numerous copy-number losses occurred on chromosomes 4, 5, 6, 7, 13, 16, and 17 in both sarcomas.
Figure 1.
Somatic and germline mutations in TP53.
A, Sanger sequencing validation of a heterozygous somatic mutation identified by WES in the TP53 gene of patient 1’s SMN. WES-derived read counts supporting each SNV are shown.
B, Sanger sequencing validation of a heterozygous germline mutation identified by WES in the TP53 gene of patient 2 that has been previously associated with de novo Li-Fraumeni syndrome (rs11540652, c.743G>A, p. Arg248Gln). WES-derived read counts supporting each SNV are shown.
C and D, The top panels show genome-wide copy-number alterations for patient 1 and patient 2. Regions of normal copy-number (ploidy – 2) are plotted in green, regions showing gain in red, and loss in blue. Segments of copy-number change that were statistically significant from normal ploidy (Wilcoxon signed-rank test) are indicated by a line of the corresponding color (red, gain; blue, loss) either above or below the plot. Allelic ratio data of germline heterozygous loci are displayed in the lower panels. Variants in a normal genome would be centered at 0.5, which represents the presence of two alleles. Any deviation from 0.5 indicates allelic imbalance or LOH. Chromosomes are displayed consecutively, beginning with chromosome 1 at the far left.
E, Copy-number and allele frequency plots of chromosome 17 for patient 1, highlighting the TP53 locus.
F, Copy-number and allele frequency plots of chromosome 17 for patient 2, highlighting the TP53 locus.
We also examined SMNs for loss of heterozygosity (LOH), defined as heterozygote allele frequencies deviating from 0.5. Large-scale LOH (100 kb or greater) was present throughout the exomes of both patients, consistent with genomic instability (Fig. 1C and D and Supplementary Fig. S4). LOH frequently cooccurred with normal copy-number profiles, indicating acquired uniparental disomy (UPD), also termed copy-neutral LOH. One such region encompasses the TP53 locus, located on the p arm of chromosome 17 (Fig. 1E and F). Interrogating the read depth at this position confirmed that in both patients the wild-type allele of TP53 had been lost and the variant allele duplicated.

We analyzed copy-number changes involving genes described in the COSMIC cancer gene census (October 2015; ref. 20), which likely contribute to tumorigenesis. Patient 1 had 216 genes affected by copy-number change that were listed as associated with cancer in the COSMIC database. A large proportion of these (n = 53) were due to the gain of the entire length of chromosome 1. For patient 2, there were 275 genes in the COSMIC database affected by copy-number change, due in a large part to gains in chromosomes 1 (n = 39), 9 (n = 27), 17 (n = 27), and 19 (n = 28) (Supplementary Files S1 and S2).

Transcriptome analysis

Transcriptome sequencing revealed differential TP53 expression in both SMNs, with mutant TP53 highly overexpressed relative to normal skeletal muscle (Fig. 2A). Expression of mutant TP53 predominated in patient 1's SMN, with 96% of reads (105/109 total reads) supporting the TP53 mutation. Similarly, in patient 2's SMN expression of the mutant germ-line, TP53 dominated, supported by 95% of the reads covering this position (407/430 total reads). We observed a significant preponderance of two transcript isoforms producing truncated p53 proteins with a shorter N-terminus region compared with the canonical tumor suppressor p53α, which is the most abundant isoform in normal tissues (ref. 32; Fig. 2B and C). Patient 2's SMN demonstrated increased expression of transcript isoform Δ40p53α (9.541 and 2.284 FPKM for patients 2 and 1, respectively), whereas the Δ133p53α isoform was more abundant in patient 1 (7.256 FPKM) compared with

**Figure 2.** Mutant TP53 expression and isoforms. A, TP53 gene expression levels quantified in FPKM (Fragments Per Kilobase of transcript per Million fragments mapped) for patient 1 (blue), patient 2 (green), and normal skeletal tissue (gray; dbGAP accession number: phs000424.v6.p1; FPKM = 4.5; n = 148 samples). Each expression value is annotated with error bars reflecting the uncertainty in assigning reads to transcripts (estimated by Cufflinks statistical model), and indicating the lower and upper bound of the 95% confidence interval on the gene FPKM value. B, Expression levels of TP53 transcript isoforms between the two patients and normal skeletal muscle (ND, not detected): Δ40p53α (variant 1), p53α (variant 2), p53β (variant 3), p53γ (variant 4), Δ133p53α (variant 5), Δ133p53β (variant 6), Δ133p53γ (variant 7), Δ40p53α (variant 8). C, Structural organization of the canonical p53 protein and its isoforms (from N-to C-termini: transactivation domains 1 and 2, proline-rich domain, DNA-binding domain, nuclear localization domain, oligomerization domain, and negative regulation domain).
patient 2 (2.943 FPKM; Fig. 2B). Variant 1 of the Δ40p53α isoform is translated from the downstream in-frame start codon and lacks the first transactivation domain, impairing p53-mediated growth suppression (ref. 33; GenBank accession: NM_001276760; Fig. 2C). Variant 5 of the Δ133p53α isoform is transcribed from an internal promoter and thus lacks all transactivation and part of DNA-binding domains. This variant stimulates angiogenesis and tumor progression in a p53-dependent and -independent manner when overexpressed (refs. 34, 35; NM_001126115; Fig. 2C).

Germline TP53 polymorphisms in pediatric cancer survivors developing SMNs

While the development of an SMN raises concern for a genetic susceptibility to tumorigenesis, most of these cancers arise in individuals without a known familial predisposition syndrome. To determine whether germline TP53 variants are present at higher frequencies in pediatric cancer survivors developing SMNs compared with the general population, we investigated a validation cohort of 37 pediatric cancer survivors registered in the CCSS known to have developed solid SMNs. These survivors had primary malignancies that included Hodgkin lymphoma, sarcomas, and CNS tumors and developed infiltrating ductal carcinoma of the breast and sarcomas as SMNs (Supplementary Table S1). None of these patients were reported to have a germline tumor predisposition syndrome (Li-Fraumeni, neurofibromatosis I, tuberous sclerosis, neurofibromatosis 2, or ataxia telangiectasia). Importantly, individuals in the validation cohort had no family history of cancer predispositions. These patients had both germline DNA and matched solid SMN tissue available. The SMN histologies were reviewed to confirm malignancy. Sanger sequencing was performed on exons 2–11 of TP53 in the germline DNA. Four of 37 individuals (11%) harbored a germline variant A639G located in exon 6 (R213R in the protein; Fig. 3A; Table 1), a synonymous SNP rs1800372 (36) with a minor allele frequency in the general population of 0.5% (P = 0.00015 by Fisher exact test). This variant resides within the DNA-binding domain of the p53 protein (Fig. 3B and C). A second germline variant identified in this cohort was G215C in exon 4 (R72P in the protein, rs1042522; Supplementary Fig. S5; Table 1), which was present in 6 patients and mutually exclusive with the germline A639G variant, resulting in 10 of 37 evaluable patients (27%) harboring a germline TP53 variant. Exon 4 encodes the proline-rich domain, which is important for the apoptotic activity of p53 by nuclear exportation via MAPK (37). The R72P variant is not known to be deleterious, and in fact is present in 20% to 60% of the general population depending on ethnicity (38). Six of 37 germline DNAs (16%) harbored this polymorphism, which is slightly below the low range for various ethnicities described in the literature and thus could not be considered enriched for any ethnicities. Table 1 shows the proportion of the various primary cancer groups who had a TP53 variant. Frequencies for all cancer types except HD were too small to analyze. The frequencies of each variant in HD were compared with the total of all other types combined. Proportions of HD with p53 variants did not differ from other types (P value approximately 0.9). Most patients received both chemotherapy and radiotherapy (treatment exposures for each individual case shown in Supplementary Table S1).

Figure 3.
Germline TP53 variants. A, Sanger sequencing of exons 2–11 of TP53 was performed on the germline DNA of 37 pediatric cancer survivors who subsequently developed radiation-induced SMNs. Shown are chromatograms from the four individuals demonstrating the germline variant c.A639G (p.R213R), which is a synonymous SNP rs1800372. The trinucleotide codon is shown in color below the single letter amino acid symbol. B, Schematic of the p53 protein, indicating the locations of germline and somatic variants associated with SMNs. C, Three-dimensional structure of the p53 protein indicating the positions of the SMN-associated mutated amino acids.
Table 1. Malignancies in the validation cohort by TP53 variant

<table>
<thead>
<tr>
<th>Primary cancer</th>
<th>No G215C Variant</th>
<th>G215C Present</th>
<th>G215C Sequencing Failed</th>
<th>No A639G Variant</th>
<th>A639G Present</th>
<th>A639G Sequencing Failed</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Osteosarcoma</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Wilms' tumor</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Soft tissue sarcoma</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Neuroblastoma</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Astrocytoma</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ewing sarcoma</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>AML</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ALL</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NHL</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
</tr>
</tbody>
</table>

NOTE: Table displaying the breakdown of genotyping for the two p53 variants (G215C, A639G). A total of 39 variants were available for analysis. The numbers of samples for each primary cancer diagnosis harboring G215C, A639G, or neither are shown. “Sequencing failed” indicates samples for which we were unable to establish a genotype due to failure of sequencing reaction. Of note, no sample harbored both variants. Frequencies for all cancer types except HD were too small to analyze. The frequencies of each variant in HD were compared with the total of all other types combined. Proportions of HD with p53 variants did not differ from other types (P value ~ 0.9).

Abbreviations: AML, acute myeloid leukemia; ALL, acute lymphoid leukemia; HD, Hodgkin disease; NHL, non-Hodgkin lymphoma.

Given the low frequencies of patients receiving radiation only or chemotherapy only, these data do not permit testing for significant interactions between TP53 variants and specific therapies.

Discussion

Detecting somatic mutations in cancer specimens can inform prognosis and management, and defining germline mutations also informs individual cancer susceptibility. Pediatric cancers typically harbor fewer somatic alterations than adult malignancies (39). Nontherapeutic mutagen exposure (e.g., tobacco or ultraviolet light) and aging also play little or no role in the pathogenesis of pediatric malignancies. In contrast, known inherited predispositions account for a much higher proportion of cancers arising during the first 2 decades of life. Thus, evaluating potential germline susceptibilities is particularly important in pediatric patients with cancer, in whom a germline tumor suppressor gene mutation may underlie the primary cancer formation and potentially contribute to the risk of SMNs (5, 40).

SMNs in pediatric cancer survivors are predominantly solid tumors, including soft tissue and bone sarcomas (8, 41), which represent approximately 15% of solid tumor SMNs (6). Our analysis of two radiation-induced sarcoma SMNs underscore the importance of TP53, a gene involved in DNA damage response and maintenance of genomic integrity (42). Using next-generation sequencing approaches, this study identifies TP53 mutations in the exomes and transcriptomes of SMNs, which can arise as somatic alterations (as in the case of patient 1) or germline mutations (as exemplified by patient 2).

Coordinated WES and RNA-Seq also indicate how mutant TP53 can be represented in SMN transcriptomes. We found that in the presence of either somatic or germline TP53 mutations, wild-type TP53 is lost and mutant TP53 duplicated and over-expressed in the SMN. Copy-number alterations were common in both SMNs; however, in both cases, copy-number–neutral LOH occurred with TP53, suggesting a shared genetic mechanism distinguishable from the remainder of the exome.

The relationship of TP53 status to SMNs in pediatric cancer survivors without a known history of a cancer predisposition has been undefined. Prior analysis of an adult patient with cancer with multiple primary tumors and a therapy-related acute myeloid leukemia identified a germline TP53 mutation, supporting the idea that cryptic variants in cancer susceptibility genes can define susceptibility to primary and secondary cancers (43). In our study, patient 2 had Li-Fraumeni syndrome but no first-degree relatives with malignancy, raising the possibility that TP53 variants are more common in pediatric survivors who develop SMNs as compared with the general population.

Individuals with Li-Fraumeni syndrome, who have a known germline TP53 mutation, are at high risk for developing a diverse array of malignancies, including sarcomas, which can present in early childhood (44). To manage this life-long risk, biobehavioral and imaging surveillance of individuals with Li-Fraumeni syndrome have been devised and shown to improve survival (45). Our findings in radiotherapy-induced sarcomas strongly argue that these individuals are also susceptible to SMNs. Using next-generation sequencing of both SMN exomes and transcriptomes, this work presents molecular evidence establishing that the germline TP53 variant is exclusively represented in the transcriptome in specific mRNA isoforms.

This analysis identified the recurrent SNP rs1800372 in a validation cohort of pediatric cancer survivors who subsequently developed SMNs. Although rs1800372 is a synonymous variant, synonymous variants are not necessarily silent or neutral in cancer evolution, as TP53 synonymous mutations can inactivate function (46). Indeed, rs1800372 is associated with poor outcome in primary breast cancer (47), and synonymous variants in the mouse Trp53 gene have been shown to influence binding of the Trp53 transcript to MDM2 and subsequent p53 function (48). Similarly, the R72P p53 variant (rs1042522), which is not classified as a deleterious mutation, has also been implicated in cancer susceptibility in humans (49). To our knowledge, this analysis is the first multiomics analysis of SMNs and determination of germline TP53 variants in survivors of pediatric malignancies.

Similar to patient 2, the CCSS validation cohort that we analyzed also did not have histories of familial tumor predisposition syndromes. However, clinical histories can be inadequate for identifying individuals with germline mutations in cancer predisposition genes, and germline mutational analysis may play an important role when estimating SMN risk. Our data suggest that analyzing germline TP53 exons at the time of diagnosis will identify patients at high risk for SMN development. These individuals may then benefit from dedicated surveillance protocols to screen for SMN formation (4).
The precise value of germline TP53 sequencing of all new sarcomas must be assessed by the analysis of a larger number of tumors than presented in this work. Patient 1 did not harbor a germline mutation in a tumor suppressor gene, but epidemiologic data suggest that children diagnosed with rhabdomyosarcoma (embryonal and pleomorphic subtypes) have a heightened risk of developing SMNs (50). Clearly, the molecular basis for SMN risk remains to be defined for many pediatric cancer survivors.

The National Comprehensive Cancer Network (NCCN) guidelines for Adolescent and Young Adult (AYA) Oncology (51) define second cancer screening recommendations based on dose delivered to at-risk anatomy. These guidelines address breast cancer, thyroid cancer, and colorectal cancer. Germline screening–based risk stratification is not integrated. Recommendations are radiation dose-dependent, with cancer screening recommended for patients receiving higher radiation doses (51). The Children’s Oncology Group (COG) has also outlined comprehensive recommendations for the long-term follow-up of pediatric cancer survivors (52). Similar to NCCN guidelines, COG’s guidelines also highlight higher radiation dose as a risk factor for second malignancies. In addition, for specific types of second malignancies, germline mutations in TP53, RB1, and NF1 are listed as risk factors. However, germline mutational analysis is not routinely performed in pediatric patients with cancer. Furthermore, it is likely that germline-mediated SMN risk varies between different germline variants. Additional sequencing analyses of SMNs from pediatric cancer survivors may lead to additional tiers of recommendations based on the presence of specific germline variants. Insights into the biologic basis for SMN formation, such as germline-based risk factors, may similarly permit rational strategies to manage or mitigate this risk in cancer survivors (Fig. 4).

This study is limited by the small size of the analyzed cohort, which reflects the challenges of obtaining high quality SMN tissue for analyses, given the long latency of SMN development. Analysis of a larger cohort of pediatric cancer survivors developing SMNs may identify additional genetic variants that significantly correlate with second cancer development. Another limitation is that the corresponding SMNs for the validation cohort were not analyzed by whole-exome or transcriptome sequencing.

New approaches for mitigating SMN induction are needed. Appropriately applying such approaches requires accurately identifying which individuals are at greatest risk for SMN development, and our findings strongly suggest that analyzing the TP53 gene for variants, both synonymous and nonsynonymous, may be a viable strategy for identifying individuals at greatest risk of SMN formation. Identifying at-risk individuals is important because advances in radiotherapy delivery alone will not eliminate SMNs in future survivors. Notably, both patients received conformal radiotherapy, which is representative of modern radiotherapy and is employed to limit radiation to normal tissues and reduce late toxicities. Proton-based radiotherapy, which produces conformal dose deposition due to the unique physical characteristics of heavy particles, has been proposed as an approach that will reduce SMN risk in pediatric patients with cancer. However, SMNs caused by neutrons

---

**Figure 4.** Potential clinical implications. Assessing SMN susceptibility may be integrated into the workup for an initial cancer diagnosis. Germline sequencing in patients may differentiate individuals with elevated risk. These individuals might be offered adjusted therapies and more focused surveillance.
produced during proton radiotherapy has been raised as a concern [31]. The clinical experience to date are incomplete and indicate no significant difference in SMN rate between protons and photon radiotherapy, although with longer follow-up, differences may be detected [33].

As oncologic care continues to improve survival, considerations of late toxicity become increasingly important, and our work suggests that the identification of germline-mediated risk should be an integral component of strategies to reduce SMN formation.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: J.P. Neglia, K.M. Shannon, J.L. Nakamura
Development of methodology: A. Sherborne, J.P. Neglia, A. Nakamura, J.L. Nakamura
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M. Loh, S.G. Dubois, R. Goldsby, J.P. Neglia, S. Hammond, L.L. Robison, J.F. Costello, J.L. Nakamura
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A. Sherborne, V. Lavergne, P. Davidson, T. Mazor, I.V. Smirnov, S.G. Dubois, S. Hammond, J. Costello, A. Nakamura, K.M. Shannon, S. Bhatia, J.L. Nakamura
Writing, review, and/or revision of the manuscript: A. Sherborne, V. Laverigne, P. Davidson, T. Mazor, M. Loh, S.G. Dubois, R. Goldsby, J.P. Neglia, S. Hammond, L.L. Robison, R. Wustrack, A. Nakamura, K.M. Shannon, S. Bhatia, J.L. Nakamura
Study supervision: A. Sherborne, J.F. Costello, J.L. Nakamura
Other: L. Lee (assisting K. Yu, organizing data)

Grant Support
J.L. Nakamura was supported by grants from Alex’s Lemonade Stand Foundation, St. Baldrick’s Scholar Award, and the Hagar Family Foundation. The CCSS is supported by U24 CA55727 (C.T. Armstrong, principal investigator).


The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 10, 2016; revised July 16, 2016; accepted August 25, 2016; published OnlineFirst September 28, 2016.

References
30 Hong H, Kohli K, Gazabedian MJ, Stallesup MB. GRIP1, a transcriptional coactivator for the AF-2 transactivation domain of steroid, thyroid, retinoid, and vitamin D receptors. Mol Cell Biol 1997;17:2735–44.
Somatic and Germline TP53 Alterations in Second Malignant Neoplasms from Pediatric Cancer Survivors
