Germline mutations in telomerase and other telomere maintenance genes manifest in the premature aging short telomere syndromes. Myelodysplastic syndromes and acute myeloid leukemia (MDS/AML) account for 75% of associated malignancies, but how these cancers overcome the inherited telomere defect is unknown.

We used ultra-deep targeted sequencing to detect somatic reversion mutations in 17 candidate telomere lengthening genes among controls and patients with short telomere syndromes with and without MDS/AML, and we tested the functional significance of these mutations.

While no controls carried somatic mutations in telomere maintenance genes, 29% (16 of 56) of adults with germline telomere maintenance defects carried at least 1 ($P < 0.001$), and 13% (7 of 56) had 2 or more. In addition to $\text{TERT}$ promoter mutations, which were present in 19%, another 13% of patients carried a mutation in $\text{POT1}$ or $\text{TERF2IP}$. $\text{POT1}$ mutations impaired telomere binding in vitro and some mutations were identical to ones seen in familial melanoma associated with longer telomere length. Exclusively in patients with germline defects in telomerase RNA (TR), we identified somatic mutations in nuclear RNA exosome genes $\text{RBM7}$, $\text{SKIV2L2}$, and $\text{DIS3}$, where loss-of-function upregulates mature TR levels. […]
Somatic reversion impacts myelodysplastic syndromes and acute myeloid leukemia evolution in the short telomere disorders

Kristen E. Schratz,1,2,3 Valeriya Gaysinskaya,1 Zoe L. Cosner,1 Emily A. DeBoy,1,4 Zhimin Xiang,1 Laura Kasch-Semenza,4 Liliana Florea,4,5 Pali D. Shah,5 and Mary Armanios1,2,3,4,6

1Department of Oncology, 2Telomere Center at Johns Hopkins, 3Sidney Kimmel Comprehensive Cancer Center, 4Department of Genetic Medicine, 5Department of Medicine, and 6Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

BACKGROUND. Germline mutations in telomerase and other telomere maintenance genes manifest in the premature aging short telomere syndromes. Myelodysplastic syndromes and acute myeloid leukemia (MDS/AML) account for 75% of associated malignancies, but how these cancers overcome the inherited telomere defect is unknown.

METHODS. We used ultra-deep targeted sequencing to detect somatic reversion mutations in 17 candidate telomere lengthening genes among controls and patients with short telomere syndromes with and without MDS/AML, and we tested the functional significance of these mutations.

RESULTS. While no controls carried somatic mutations in telomere maintenance genes, 29% (16 of 56) of adults with germline telomere maintenance defects carried at least 1 (P < 0.001), and 13% (7 of 56) had 2 or more. In addition to TERT promoter mutations, which were present in 19%, another 13% of patients carried a mutation in POT1 or TERF2IP. POT1 mutations impaired telomere binding in vitro and some mutations were identical to ones seen in familial melanoma associated with longer telomere length. Exclusively in patients with germline defects in telomerase RNA (TR), we identified somatic mutations in nuclear RNA exosome genes RBM7, SKIV2L2, and DIS3, where loss-of-function upregulates mature TR levels. Somatic reversion events in 6 telomere-related genes were more prevalent in patients who were MDS/AML-free (P = 0.02, RR 4.4, 95% CI 1.2–16.7), and no patient with MDS/AML had more than 1 reversion mutation.

CONCLUSION. Our data indicate that diverse adaptive somatic mutations arise in the short telomere syndromes. Their presence may alleviate the telomere crisis that promotes transformation to MDS/AML.

FUNDING. This work was supported by the NIH, the Commonwealth Foundation, the S&R Foundation Kuno Award, the Williams Foundation, the Vera and Joseph Dresner Foundation, the MacMillan Pathway to Independence Award, the American Society of Hematology Scholar Award, the Johns Hopkins Research Program for Medical Students, and the Turock Scholars Fund.
MDS/AML. We found diverse somatic reversion mechanisms that have great specificity relative to the germline defect. In contrast to our expectations, however, these somatic mutations were rare in patients with MDS/AML and their allele burden was significantly higher in patients who were MDS/AML-free. These data have implications for fundamental understanding of how somatic adaptation in the setting of defective hematopoiesis may provide protection against MDS/AML evolution in at least some inherited bone marrow failure syndromes.

**Results**

To understand the mechanisms of telomere maintenance in patients with short telomere MDS/AML, we compared the prevalence of somatic mutations among 84 individuals who were divided across 3 groups: healthy controls (mean age 64 years, \( n = 28 \)), patients with short telomere syndromes who did not have MDS/AML (mean age 60 years, \( n = 40 \)), and patients with short telomere syndromes who fulfilled World Health Organization criteria for MDS/AML (mean age 56 years, \( n = 16 \)). Table 1 summarizes their clinical characteristics. Both the control group and patients with short telomere syndromes who did not have MDS/AML had normal telomere length, while the other 2 groups had comparable abnormalities in telomere length (Figure 1, B and C). We used a customized ultra-deep, error-corrected sequencing platform (13, 14), and designed our experiment to detect low frequency somatic mutations down to 0.5% variant allele frequency (VAF). Using this pipeline, we targeted 16 candidate genes in which mutations could functionally promote telomere maintenance. Their functions were grouped into 4 categories: (a) promoters of the core telomerase components, telomerase reverse transcriptase and telomerase RNA, \( TERT \), and \( TR \), (b) shelterin genes, of which several are mutated in cancer-prone families hypothesized to have long telomere syndromes (15, 16), (c) nuclear RNA exosome genes involved in \( TR \) processing and where loss-of-function has been shown to increase mature TR levels, and (d) \( ATRX \) and \( DAXX \), known to be mutated in cancers that rely on alternative lengthening of telomeres (ALT) mechanisms (Figure 1A and Supplemental Table 1; supplemental material available online with this article; https://doi.org/10.1172/JCI147598DS1). In addition, we examined \( TP53 \), since its loss may allow bypass of the short telomere checkpoint (9). Among the 3 groups, we identified 36 clonal somatic mutations: 34 single nucleotide variants (SNVs) and 2 insertions/deletions. They were found in 7 of the 17 candidate genes (35%): \( TP53 \), the \( TERT \) promoter, \( POT1 \), \( TERF2IP \), \( RBM7 \), \( SKIV2L2 \), and \( DIS3 \) (Figure 2, A-D and Table 2). Only 1 was found in a healthy control, and that mutation was a low allele frequency \( TP53 \) (VAF 1%) subclinical clonal hematopoiesis mutation. Another 10 \( TP53 \) mutations were identified and fell predominantly into the DNA binding domain, and all are recurrently mutated in cancer (Figure 2A). The latter mutations were divided equally among patients with short telomere syndromes with and without MDS/AML (although both patients who carried 2 \( TP53 \) mutations had MDS/AML), and the \( TP53 \) allele burden was also higher than that found in the patients who were MDS/AML-free (Figure 2A and Figure 3, A and B).

All the remaining somatic telomere gene mutations were exclusively in patients with short telomere syndromes with 30% carrying at least 1 mutation (16 of 56 short telomere vs. 0 of 28 controls, \( P < 0.001 \), Fisher’s exact test) (Table 2). The most common mutations were in the \( TERT \) promoter. They were seen at the –146 and –124 hotspots upstream of \( TERT \)’s transcription start site (Figure 2B). Nineteen percent (11 of 56) of the individuals in the combined short telomere groups carried at least one of these mutations. This rate is higher than the 5% to 7% prevalence previously reported in patients with pulmonary fibrosis and dyskeratosis congenita, which used less sensitive sequencing methods (17, 18). These \( TERT \) promoter mutations are identical to somatically mutated nucleotides found in numerous cancers, which create de novo ETS-binding sites that upregulate transcription of the intact \( TERT \) allele in \( trans \) with the germline mutation (17, 19–21). In the 1 patient in whom both canonical sites were mutated, the SNVs were on distinct alleles, indicating they arose in 2 different clones.

The next most common class of mutations was found in \( POT1 \), a negative regulator of telomere elongation. Two mutations were found in canonical splice junction sequences (Figure 2C and Table 2). For \( POT1 \)’s and the missense mutations in the other 4 protein-coding genes, all but one had combined annotation-dependent depletion (CADD) scores greater than or equal to 20, signifying they represent the top 1% of damaging mutations, and 5 had a score of greater than or equal to 30, indicating they were

| Table 1. Clinical characteristics of controls and short telomere groups with and without MDS/AML |
|---------------------------------|---------------------------------|---------------------------------|------------------|
| Controls                        | Short telomere                  | Short telomere with MDS/AML     |
| Number of patients              | 28                              | 40                              | 16               |
| Median age, years (range)       | 64 (50–79)                      | 60 (40–81)                      | 56 (12–74)       |
| Male/Female                     | 13/15                           | 23/17                           | 11/5             |
| Germline mutation<sup>a</sup>   |                                 |                                 |                  |
| \( TERT \)                      | 17                              | 3                               |                  |
| \( TR \)                        | 6                               | 1                               |                  |
| \( RTEL1 \)                     | 6                               | 2                               |                  |
| \( PARN \)                      | 3                               | 2                               |                  |
| \( DHCI \)                      | 2                               | 2                               |                  |
| \( NAF1 \)                      | 0                               | 1                               |                  |
| \( TIN2 \)                      | 1                               | 0                               |                  |
| \( ZCCH8 \)                     | 1                               | 0                               |                  |
| Unknown                         | 4                               | 5                               |                  |
| Patients with pulmonary fibrosis<sup>b</sup> | 31 (78)                       | 11 (69)                         |                  |

<sup>a</sup>The autosomal genes listed were heterozygous mutant. The unknown category included 1 patient with low telomerase RNA levels for each of the short telomere groups. <sup>b</sup>Patients with pulmonary fibrosis refers to patients who had this concurrent diagnosis at the time of study recruitment and blood draw.
Notably, the p.K33E somatic mutation phenocopied a mutant that was deleted for the entire first OB fold of POT1 (Figure 2, E–G). In contrast, POT1 p.C605S, which is found in the TBP1-interacting domain (Figure 2C), showed no DNA binding defects. These collective data supported that multiple de novo somatic mutations arise relatively commonly in telomerase and telomere gene mutation carriers. They are functionally impactful and overlap with somatic telomere maintenance mutations seen in cancer.

We next examined if the telomere-related somatic mutations clustered in the patients with MDS/AML, but found a surprising and paradoxical result. Adults with short telomere syndromes who were MDS/AML-free had more mutations than those patients with MDS/AML (RR 4.4, 95% CI 1.2–16.7, \( P = 0.02 \), Fisher’s exact test). Among adults without MDS/AML, 35% had at least 1 and 18% had at least 2 somatic mutations (14 and 7 of 40, respectively, Figure 3A). But for patients with MDS/AML, only 13% (2 of 16) carried 1 and none carried 2 mutations. The somatic mutant clones were also on average 3.8-fold larger in MDS/AML-free adults (mean 14.7% vs. 3.9%). Notably, none of the MDS/AML patients had a clonal somatic mutation in a telomere gene with VAF greater than 10% (\( P = 0.048 \), Fisher’s exact test; Figure 3B), suggesting this 10% threshold may be a useful predictive threshold if validated in other studies. We assessed the incidence of MDS/AML in the patients who carried somatic reversion mutations and all of them remained MDS/AML-free with a median follow-up of 2.0 years (range 3 months–11 years).

If these somatic mutations correlate negatively with MDS/AML onset, we would predict they would be limited to or enriched in the myeloid lineage. To test this directly, we used droplet digital PCR to quantify the mutant fraction in whole blood and compared it with myeloid and T cell fractions. We studied 4 mutations (POT1 and DIS3) from 3 patients and found they were all enriched in the myeloid lineage while being low or undetectable in the T cell fraction (Figure 4, A–C, POT1, and Figure 4D, DIS3). To definitively test if the somatic mutations we identified were indeed reversion events, we correlated the germline and somatic mutant pathways and found remarkable associations. While TERT promoter mutations predominated in germline TERT mutation carriers (8 of 11), POT1 and TERF2IP mutations were associated with a more heterogeneous spectrum of germline mutations including TERT, TINF2, RTELI, and PARN (Table 2).
Figure 2. Somatic mutations relative to protein functional domains and POT1 functional studies. Schema showing somatic mutations relative to mutant protein/gene with conserved domains drawn for (A) TP53, (B) the TERT promoter, (C) shelterin components, and (D) nuclear RNA exosome components, respectively. The key denotes the group in which the mutation was identified, with open circle referring to controls, black circles to patients with short telomere syndromes without MDS/AML, and red circles to patients with short telomere syndromes with MDS/AML. (E) Electrophoretic mobility shift assay (EMSA) for mutant POT1 examining binding capacity to a telomere oligonucleotide. POT1<sup>ΔOB</sup> refers to a protein deleted for the first DNA binding domain (aa127-635). (F) Immunoblot of in vitro translated products shows stable missense mutant POT1 at expressed levels, with the Myc immunoblot confirming specificity. (G) Mean intensity of binding relative to WT, with error bars representing SEM. The data shown are from 3 EMSA experiments derived from 2 independent in vitro translation reactions.
This observation is consistent with \textit{POT1} loss-of-function mutations playing a role in enhancing telomerase access or processivity (15). By contrast, mutations in the nuclear RNA exosome genes were exclusively detected in patients with germline defects in TR abundance or function. All 3 patients who carried these somatic mutations carried germline mutations in \textit{TR} itself, the ribonucleoprotein \textit{DKC1}, or had abnormally low TR levels (Table 2). These data support a model where myeloid-specific adaptation that arises in the highly replicative environment of hematopoiesis, and under the selective pressures of a germline short telomere defect, is associated with an MDS/AML-free state.

**Discussion**

In the high turnover environment of hematopoiesis, selective pressures imposed by inherited mutations allow advantageous clonal mutations to persist (28). Nearly all previously described somatic reversion mechanisms are direct in that they repair the germline mutant gene itself such as by correcting a frameshift with an insertion or through repair of the mutation by homologous recombination with the WT allele (28–30). Here, we identified several indirect somatic reversion events in the short telomere syndromes, and found they overlap with known telomere maintenance mechanisms in cancer. Their high specificity, such as is the case for the nuclear exosome components in \textit{TR} mutation carriers, underscores a remarkable versatility of the stressed hematopoietic system to acquire de novo adaptations. We did not observe gross measurable improvements in blood counts or telomere lengthening effects (Supplemental Figure 1A). Nonetheless, the recurrent mutations in the same genes highlight the competitive advantage they provide under the selective pressures of the short telomere environment. Their myeloid enrichment and persistence over 11 years in one patient we studied confirms these mutations arise in a likely primitive myeloid progenitor.

A higher CADD score denotes a more deleterious prediction, with ≥ 20 referring to the top 1% of most deleterious possible substitutions, and ≥ 30 referring to the top 0.1%. ATCL, adult T cell lymphoma.
these events were mutually exclusive. Comparing the 35% prevalence of somatic mutations we identified in adult short telomere patients without MDS/AML (Figure 3A), relative to the 10% overall incidence of MDS/AML (6), somatic reversion appears to be a more common and preferred adaptive mechanism. It is also possible that somatic reversion mutations may similarly be a favorable predictor with respect to MDS/AML risk in other inherited bone marrow failure states.

Our data raise the possibility that deep sequencing for at least some of the common mutations we identified may be a useful biomarker for assessing MDS/AML risk in some settings. Idiopathic pulmonary fibrosis is the most common indication of lung transplantation worldwide and it is estimated that 35% of familial pulmonary fibrosis patients carry a germline defect in a telomere maintenance gene (32, 33). In the setting of lung transplantation, short telomere patients are exposed to cytotoxic medications that add selective pressures on hematopoiesis, and the presence of sizable clonal reversion mutations may provide a protective advantage. The median follow-up for patients with telomere-related somatic reversion mutations in our study was 24 months (mean 36 months, range 3–130 months). Additional studies and longer follow-up will be needed to test the predictive utility of somatic reversion testing in clinical settings.

**Methods**

**Patients.** Patients were recruited from 2003 to 2020 as part of the Johns Hopkins Telomere Syndrome Study (6, 34). They were enrolled if they carried a pathogenic mutation in a telomere-related gene and had a personal/family history of a short telomere phenotype, had classic familial short telomere syndromes as defined (7), or had sporadic disease with low telomerase RNA levels (35). The MDS/AML diagnosis was assigned using WHO classification (36). Healthy controls were recruited as previously (37), and were selected for having normal telomere lengths. Cytogenetic abnormalities for the patients with MDS/AML were inferred by karyotype (6). Race data were not collected for this study. Follow-up was last assessed on April 25, 2021.

**Germline sequencing, functional studies, and telomere length measurement.** Germline DNA sequencing was performed on peripheral blood DNA using whole genome, exome, or targeted analyses (25, 30, 37). Germline mutations not reported in Alder et al. (37) or Schratz et al. (6) are listed in Supplemental Table 2. Telomerase RNA quantification and DKCI functional analyses have been detailed previously (25, 38). Peripheral blood telomere length was measured by flow cytometry and fluorescence in situ hybridization (flowFISH) as described (37, 39). All previously unreported germline mutations were deposited in the Telomerase Database (www.telomerase.asu.edu) (40).
Capture design, library preparation, and sequencing. DNA for somatic analyses was derived from peripheral blood for all 84 individuals in this study, except for 4 individuals from whom it was isolated from a bone marrow aspirate (indicated in Figure 3A). We designed 150 bp read length probes (SureDesign) targeting 37 kb across 17 genes including 2 promoters (Supplemental Table 2). Libraries were prepared following the manufacturer’s instructions (Agilent HaloPlex™ for 1–500 kb target region). Briefly, 400 ng genomic DNA diluted to 14.4 ng/μL were digested in 8 restriction enzyme reactions and hybridized to the custom biotinylated HaloPlex HS probe library incorporating sample indexes and a degenerate 10-nucleotide unique molecular barcode (i.e., UMI, to identify each individual captured DNA molecule) to facilitate base calling accuracy and quantification. Targets were ligated to form circularized fragments, captured using streptavidin beads, and polymerase chain reaction amplified (24 cycles). Target libraries were validated using the Fragment Analyzer (Advanced Analytical Technologies) followed by quantification of fragment smear analysis of 190–640 bp range. All samples were subsequently pooled at a 4 nM concentration. A final AMPure XP magnetic bead cleanup was performed on the pooled library to eliminate any adapter-dimer traces.

HaloPlex library sequencing. Sequencing was performed on an Illumina HiSeq 2500 using a 150 bp paired-end protocol, and data were demultiplexed following standard protocols. Reads were trimmed and aligned to the reference hg19 genome using the Burrows-Wheeler aligner BWA-MEM in the Agilent Alissa Align and Call v.1.1.2.2 platform. Of 86 samples designed for HaloPlex platform sequencing, 81 eventually fulfilled quality control criteria (2 failed at library preparation, 3 had low coverage). The mean depth of coverage was 12,522× (± 4591 SD) with a median of 10,952× (± 5959 SD). The coverage depth for the \( TERT \) promoter was relatively lower and was calculated separately using Samtools (version 1.10, using htslib 1.10). Within the targeted interval (chr5:1295085-1295385, hg19), its coverage was limited to chr5:1295085-1295115 and chr5:1295224-1295385, and these intervals had a mean coverage of 2457× (± 1204 SD) with a median 2421×. Supplemental Table 3 lists the mutations with the number of WT and mutant reads shown.

Three additional patients with MDS/AML were consecutively recruited after the HaloPlex sequencing protocol started, and their DNA was sequenced using a clinically validated dual-indexed targeted method that applied a background error rate analysis to allow robust detection of somatic mutations with allele frequency of at least 1% (6, 41). This pipeline included the 3 most commonly mutated genes in the HaloPlex analysis — \( TP53 \), \( TERT \) promoter, and \( POT1 \) — as well as \( TINF2 \). Supplemental Figure 1B lists the mutation types at the base pair level and shows 58% of SNVs were cytosine to thymidine transitions.

Variant calling and analysis. Variant calling was performed using Agilent Alissa Align and Call v1.1.2.2 software and annotated using ANNOVAR v4162018. To utilize the HaloPlex error-corrected method of analysis, we first collapsed reads originating from the same sample molecule into read families and duplicates were removed to create consensus reads. Variants were filtered to include only those present in at least 3 unique read families. To exclude sequencing artifacts, we used a quality score of greater than 90 and filtered for a variant score threshold of greater than 0.3 in Alissa. Variants that were present in more than 5% of controls were additionally excluded as alignment artifacts. Variants with allele frequencies of 40% to 60% or greater than 95% were deemed germline and excluded. We finally filtered for variants with minor allele frequencies greater than 0.0005 in gnomAD v2.1.1 (any VAF), and these were excluded as they were thought to be possibly germline (performed November 1, 2020). The \( TERT \) promoter was
manually analyzed and curated separately to include variants present in all reads of one or more consensus families at the canonical c.-124 and c.-146 positions, irrespective of variant allele frequency. All of the variants reported were manually verified in Integrative Genomics Viewer. Somatic variants that passed filtering and manual review were submitted to the COSMIC database (July 6, 2021, COSP49927).

**Variant interpretation.** Protein schema were drawn based on the longest isoform curated in the Gene Database (National Center for Biotechnology Information, accessed October 21, 2020). Linear protein structures were visualized in SnapGene (v5.1.6), and conserved domains constructed were based on Uniprot (release 2020_05) and the literature. Variants were analyzed for their presence in the COSMIC database (v.92) and CADD score (v1.6) (42).

**POT1 missense mutations, we cloned Myc-FLAG tagged POT1 into T7 expression vector (pcDNA 5/FRT, Addgene) and performed site-directed mutagenesis. A control with POT1ΔOB was generated by deleting the first OB fold (amino acids 127-635). T7 expression vectors were used in an in vitro translation reaction using the TNT Quick Coupled Transcription/Translation System (Promega), according to the manufacturer’s instructions. Briefly, a 50 μL reaction containing 1 μg of expression vector, 20 μM methionine, and 40 μL TNT Quick Master mix containing rabbit reticulocyte lysate was incubated at 30°C for 90 minutes and used immediately for the gel shift assay.

The POT1 binding assays were performed using the Odyssey Infrared EMSA Kit (LI-COR). Briefly, a 20 μL reaction mixture was prepared containing 2 μL of the IVT product, 10 nM IRDye 800-labeled, single-stranded telomeric oligonucleotide, 2.5 μM nonspecific ssDNA, and 1 μg poly(dI-dC) in binding buffer containing 100 mM Tris, 500 mM KCl, 12.5 mM DTT, 10 mM EDTA, and 0.25% Tween 20. We used the telomere oligonucleotide 5’GGTTAGGGTTAGGGTTAGGG and the nonspecific ssDNA oligonucleotide 5’TAAATACCGGGGATCCGCCTGTAGCATAAGGATGATCC as per Baumann et al. (43). Reactions were incubated for 30 minutes at room temperature and protein-DNA complexes were resolved by gel electrophoresis on a 10% polyacrylamide Tris-Borate EDTA at 100V for 2 hours at 4°C. Gels were visualized using the Li-Cor Odyssey Imaging System. Quantification was done using ImageJ (2.1.0).

**Western blot.** Using the NuPAGE SDS-PAGE gel system, 1 μL of in vitro translation reaction was run on 8% Bis-Tris gels with MOPS-SDS running buffer at 150V and transferred to a PVDF membrane using the iBlot 2 Dry Blotting System on setting P0 (Thermo Fisher Scientific). Membranes were blocked for 1 hour before antibody staining. Membranes were stained first with anti-POT1 antibody (rabbit, NSB00-176, 1:500, Novus) and subsequently with anti-rabbit HRP antibody (goat, 1:20,000, Cell Signaling Technology) before visualization via chemiluminescence (SuperSignal West Pico Plus Chemiluminescent substrate kit) using ImageQuant. Membranes were then stripped for 15 minutes, using New Blot IR Stripping buffer, and blocked again for 1 hour before staining with anti-Myc antibody (mouse, clone 4A6, 1:1000; Millipore). Membranes were subsequently stained with anti-mouse IRDye secondary antibody as appropriate (IRD80, donkey, 1:10,000, LI-COR) before visualization on an Odyssey scanner (LI-COR). Quantification was done using ImageJ (2.1.0).

**Cell fractionation.** Aliquots of fresh blood or frozen Ficoll-separated peripheral blood cells were fractionated using the EasySep Human Myeloid Positive Selection and EasySep Human T Cell Enrichment kits (StemCell Technologies). DNA was extracted from the fractionated subsets using the MagMAX DNA Multi-Sample Ultra 2.0 Kit (Thermo Fisher Scientific) and quantified using the Qubit dsDNA HS Assay kit (Thermo Fisher Scientific).

**Droplet digital PCR**. Droplet digital PCR (ddPCR) was performed according to the manufacturer’s specifications on a QX200 system (Bio-Rad). Sequence-specific primers and probes were designed using the Bio-Rad digital assay design engine: POT1 p.K33E (dHsaMDS726207399), POT1 p.M159V (dHsaMDS784505022), POT1 p.I266Y (dHsaMDS401027863), and DIS3 p.D488N (HsaMDS2512362). An annealing temperature of 58°C was used except for the POT1 p.K33E (54°C) and DIS3 p.D488N (54°C). The WT and mutant probes were conjugated with HEX and FAM reporters, respectively. PCR amplicons were sequence-verified by cloning the products (TOPO TA Cloning, Invitrogen). For each target, the mean number of droplets analyzed was 11,890 ± 2081 SD and all reactions had more than 9000 events analyzed. Data were analyzed in QX Manager (version 1.2-STD, Bio-Rad).

**Statistics.** Graphs were generated and statistical calculations were performed using GraphPad Prism 8.0. The respective statistical tests are included in the text and figure legends and a P value less than 0.05 was considered significant. All P values shown are 2-sided.

**Study approval.** The research was approved by Johns Hopkins Medicine’s Institutional Review Board, and patients gave written informed consent.

**Author contributions**

KES designed the study and analyzed the clinical and genetic data. KES and VG selected the candidate genes. VG prepared the libraries with LKS. ZLC analyzed the POT1 functional studies and analyzed ddPCR data. EAD performed computational analyses. ZX performed the fractionation and ddPCR experiments. LF supported the sequencing analysis pipeline design. PDS analyzed clinical data. MA oversaw the project and wrote the paper with KES. All the authors reviewed and approved the manuscript.

**Acknowledgments**

We would like to acknowledge helpful conversations with Dustin Gable, and technical support from the Johns Hopkins Genomics team, the Johns Hopkins Genetics Resources Core Facility, and the Bloomberg School of Public Health Flow Cytometry Core. We are grateful for technical support from the Agilent team of Josh Zhiyong Wang, Gregory Miles, and Weiwei Liu. This work was supported by R01CA225027 and R01HL119476, the Commonwealth Foundation, the S&R Foundation Kuno Award, and the Williams Foundation (to MA); T32HL007525, a grant from the Vera and Joseph Dresner Foundation, the MacMillan Pathway to Independence Award, and the American Society of Hematology (ASH) Scholar Award (to KES); F32HL144207 (to VG); the Johns Hopkins Research Program for Medical Students (to ZLC); T32 GM136577 (to EAD); and the Turock Scholars Fund to the Telomere Center at Johns Hopkins (to KES and EAD). The authors would also like to acknowledge a gift in the name of P. Godrej.

Address correspondence to: Mary Armanios, 1650 Orleans Street, CRB 1 Room 186, Baltimore, Maryland 21287, USA. Phone: 410.502.3817; Email: marman11@jhmi.edu.


