ELSEVIER

EXPERIMENTAL HEMATOLOGY

Experimental Hematology 34 (2006) 664–671

Genetic variation in telomeric repeat binding factors 1 and 2 in aplastic anemia

Sharon A. Savage[a,*], Rodrigo T. Calado[b,*], Zhong-Tao Xin[c], Hinh Ly[c], Neal S. Young[b], and Stephen J. Chanock[a,d]

[a]*Section of Genomic Variation, Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Md., USA;* [b]*Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Md., USA;* [c]*Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Ga., USA;* [d]*Core Genotyping Facility, National Cancer Institute, National Institutes of Health, Bethesda, Md., USA*

(Received 1 December 2005; revised 26 January 2006; accepted 8 February 2006)

Objective. **Abnormal telomere shortening has been observed in a subset of individuals with aplastic anemia (AA). We hypothesized that genetic variation in two genes critical in telomere biology, *TERF1* and *TERF2*, could be a risk factor for AA.**

Methods. **The proximal promoter and all coding regions of *TERF1* and *TERF2* were sequenced in 47 individuals with acquired AA. Regions with genetic variation were sequenced in an additional 95 AA patients and 289 healthy controls. Single nucleotide polymorphism (SNP) frequencies were analyzed using co-dominant and dominant models and haplotypes determined. Functional studies evaluated telomerase activity, telomere and telomeric overhang lengths, and TRF2 protein expression in select patients.**

Results. **Two nonsynonymous amino acid changes were detected, one in exon 9 of *TERF1* and another in exon 6 of *TERF2*. These sequence variants resulted in conservative amino acid changes and were not predicted to alter TRF1 or TRF2 protein expression or function. SNP and haplotype analyses in acquired AA patients suggested that one variant allele, in intron 9 of *TERF1*, and haplotype could be associated with increased risk for aplastic anemia (OR 1.59, 95% confidence interval 1.06–2.39, $p = 0.033$). *TERF2* SNPs and haplotypes were not significantly associated with aplastic anemia.**

Conclusions. **It is possible that a common genetic variant in *TERF1* is associated with risk for AA but additional studies are required. Highly penetrant, non-synonymous, or insertion-deletion mutations in *TERF1* and *TERF2* were not identified and therefore are not likely to be major genetic risk factors for the development of AA.**

Aplastic anemia (AA) is a heterogeneous disease that is classified as either constitutional or acquired later in life [1]. In AA, the bone marrow contains mostly fat and very few hematopoietic cells. In acquired AA, an autoimmune process targets the hematopoietic tissue [1]. Approximately one third of individuals with acquired AA have abnormally short telomere lengths in leukocytes [2,3]. Mutations in genes responsible for telomere maintenance have been identified in both constitutional and acquired AA. X-linked dyskeratosis congenita (DKC) (OMIM 305000), a constitutional type of AA, is caused by mutations in *DKC1*, a gene that encodes a ribonucleoprotein, dyskerin, that associates with the telomerase complex [4,5]. Autosomal dominant DKC has been shown to result from mutations in the *TERC* gene, which encodes the RNA component of the telomerase complex that serves as a template for telomere elongation [6]. Mutations in *TERC* as well as in the telomerase reverse transcriptase (*TERT*) have been identified in patients with acquired AA without physical anomalies characteristic of DKC [7,8].

Telomeres are long nucleotide repeats with associated protiens at the ends of chromosomes that are essential for

Offprint requests to: Sharon A. Savage, M.D., Section of Genomic Variation, Pediatric Oncology Branch, National Cancer Institute, Advanced Technology Center, 8717 Grovemont Circle, Bethesda, MD 20892-4605; E-mail: savagesh@mail.nih.gov

*Sharon A. Savage and Rodrigo T. Calado contributed equally to this manuscript and should be considered co–first authors.

0301-472X/06 $–see front matter.
doi: 10.1016/j.exphem.2006.02.008

maintaining chromosomal integrity. Telomerase, a reverse transcriptase, uses an RNA template (*TERC*) to extend TTAGGG nucleotide repeats for up to 10 to 15 kb in length in human cells [9]. Telomeric attrition results in critically short telomeres, prompting cellular senescence or crisis [9,10]. Telomeric repeat binding factor 1 (TRF1 [gene name *TERF1*], 8q13, MIM# 600951) binds to telomeric DNA and can inhibit telomerase as well as induce bending, looping, and pairing of the duplex telomeric DNA [11,12]. Telomeric repeat binding factor 2 (TRF2 [gene name *TERF2*], 16q22.1, MIM# 602027) also binds directly to the telomere and protects it from degradation and chromosomal end-to-end fusion [13–15]. *TERF1* and *TERF2* are highly conserved at the genomic and protein level between species [16,17]. We recently found that there are relatively few numbers of single nucleotide polymorphisms (SNPs) in *TERF1* and *TERF2* and accordingly, nucleotide diversity in these genes is limited as compared to large catalogs of other genes [16,17]. We therefore hypothesized that rare sequence variants in *TERF1* and/or *TERF2* could be associated with acquired AA.

Materials and methods

Patients and controls

Blood samples were obtained from 142 unrelated patients with apparently acquired AA (age range, 6–83 years; median, 39) who were treated at a single institution (Hematology Branch of the National Heart, Lung, and Blood Institute, National Institutes of Health). The diagnosis of severe AA was based on the bone marrow and blood count criteria of the International Agranulocytosis and Aplastic Anemia Study [18]. To exclude Fanconi anemia, clastogenic stress chromosome breakage tests using diepoxybuate and mitomycin C were performed on peripheral blood mononuclear cells. The first group of 47 patients was selected based on either a failure of response to at least two courses of cyclosporin and antithymocyte globulin (34 patients, 72%), a positive family history of hematologic abnormalities without physical anomalies characteristic of dyskeratosis congenita (2 patients, 4%), or short telomeres in leukocytes (20 patients, 42%), defined in previous studies as below the 90th percentile based on age- and gender-matched controls [2]. Nine patients (19%) had both failure to respond to immunosuppressive therapy and short telomeres. The second validation group consisted of 95 consecutive patients seen in the same clinic for evaluation and treatment of acquired AA. Assessment of telomere length was not available for the validation group.

Of the 142 patients, 68 were female and 74 were male. Race or ethnic background, as reported by the patients or their guardians, was as follows: Caucasian, 101 patients (71%); African-American, 17 (12%); Hispanic, 16 (11%); and Asian, 8 (6%). Patients came from the United States and from several Latin American and Asian countries. Patients or their guardians provided written informed consent for genetic testing, according to protocols approved by the institutional review board of the National Heart, Lung, and Blood Institute.

The 289 healthy controls included 109 African-Americans (23 from SNP500Cancer [19] [http://snp500cancer.nci.nih.gov], 86 from Human Variation Panel HD100AA, Coriell Cell Repositories [http://ccr.coriell.org]), 125 Caucasians (31 from SNP500Cancer and 94 from Human Variation Panel HD100CA, Coriell Cell Repositories), 23 Hispanics (SNP500Cancer), and 32 Asians (24 from SNP500Cancer and 8 from north-central China, provided by Dr. Phil Taylor, NCI).

PCR and sequencing

Genomic contigs for *TERF1* (NC_000008.9) and *TERF2* (NC_000016.8) were selected using Human Genome build 35. All exons and proximal promoter regions of *TERF1* and *TERF2* were sequenced in the first group of 47 patients, affording a greater than 95% detection rate for variants with frequency greater than 1% [20]. Regions with variation in this screen were sequenced in the second group of 95 patients and the 289 healthy controls.

Polymerase chain reaction (PCR) primers designed with Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) were appended with M13 forward or reverse tags. Genomic DNA was amplified by PCR with MJ Research model PTC-225 thermal cyclers (Bio-Rad Laboratories, Inc., Waltham, MA, USA) and the following conditions: 10 ng genomic DNA, 0.2 μM each primer, 200 μM each dNTP, 2 mM $MgCl_2$, 0.5 units AmpliTaq Gold DNA polymerase (ABI-Perkin Elmer, Foster City, CA, USA), and the manufacturer's buffer. *TERF2* exon 2 required 2.5X MasterAmp enhancer (Epicentre, Madison, WI, USA) in the AmpliTaq Gold PCR reaction. *TERF2* exon 1 was amplified using the Advantage-GC 2 PCR protocol (Clontech Laboratories, Inc., Mountain View, CA, USA) and 10 ng genomic DNA, 0.2 μM each primer, 200 μM each dNTP, 1 M GC Melt, 0.3 μL of 50X Advantage GC2 polymerase mix, and the manufacturer's buffer. Primers and annealing temperatures are shown in Table 1. Bi-directional sequencing of amplified DNA using the Big-Dye Terminator method (ABI-Perkin Elmer, Foster City, CA, USA) and M13 forward and reverse primers was analyzed on ABI-Perkin Elmer platforms (models 3100 and/or 3700) with Sequence Analysis 3.7 software (ABI-Perkin Elmer) and Sequencher 4.0.5 software (Gene Codes Corporation, Ann Arbor, MI, USA).

SNP frequencies were determined and case-control genotype data were analyzed with SAS v8.02 software (SAS Institute, Inc., Cary, NC, USA) using two genetic models: a co-dominant model, which compared homozygous wild-type, heterozygotes, and homozygous variants, and a mutation-dominant model, in which heterozygotes are considered equal to homozygous variants. Haplotypes were estimated and a case-control permutation test performed using PHASEv2.1 [21,22]. HaploStats was also used to construct haplotypes and determine the global score *p* value, haplotype frequencies, and odds ratios [23]. *In silico* assessment of the effect of amino acid changes was performed using SIFT [24] and PolyPhen [25]. GeneSplicer [26] and NNSPLICE (v0.9) [27] were used to evaluate SNPs for location in potential intron-exon splice sites.

Telomere length measurement

Telomeres of peripheral blood leukocytes were measured using the Telo *TAGGG* Telomere Length Assay kit (Roche, Mannheim, Germany), according to the manufacturer's instructions. Two μg of genomic DNA was digested with *Hinf*I and *Rsa*I restriction enzymes and then separated on agarose gels. DNA was transferred to a positively charged nylon membrane, hybridized with a digoxigenin (DIG)-labeled probe specific for telomeric repeats,

Table 1. Primer sequences and PCR conditions

Gene	Region amplified	Forward primer*	Reverse primer*	Annealing temperature (°C)
TERF1	Promoter	GGCAAATGCGTTTTCATTTT	GAACTGCTCCTCGTCGTTTC	68
	Exon 1	GTACCCAAGCGAGCCATTTA	ACCGGGATTCGAACAAAGTT	60
	Exon 2	TTTCCAACCAATGATTTTACACA	CCCCTATGCACCAGACACTT	66
	Exon 3	AATCTCATCAACCTTTTTCTATTTG	ATTAAGTTCCAGCCCCCAGT	66
	Exon 4	TGGGTCCTAAATGTGTGTGG	AAGCATTTCTTTTCTCCTCCAA	66
	Exon 5	TTAAAATGAGGAGAGGCCAGA	CCACACCTTTCTACAAAACCA	66
	Exon 6	ACCACCTTCGTAACGCTGAT	TCAGCAAGATAACATGCTACTCAAT	66
	Exon 7	TGCTTATGTTTGGAGAAACTTG	TGGGCTAACTAGGACCCTGA	68
	Exon 8	ACTAAGCAGGGAGAGCACCA	GCAGTGTACCACCATAGCAAA	68
	Exon 9	CACTCCCCCTCTGATTTCAA	AAAGATGGCAAAGGGCAAAT	68
	Exon 10	TGGGCAGCTGAGCTTACATA	CTGGTAGCACAAACCCTCATT	68
TERF2	Promoter/Exon 1	GGAACTACGGCGTCTGAGAA	TCCCGGATCTGTCTGAAGTC	56**
	Exon 2	GCTGGAAGAGGCAGTCAATC	GAAAAGCAGCCAAGACAACC	58†
	Exon 3	TGCATGCTCCATTTCCATAA	CCGAATGGTGCAGTAACCTC	58
	Exon 4	CCGACGTCAGGTGATCCA	GCTGCAGGTGTTACATGTTGA	62
	Exon 5	GGGCTTCCCTTCCTGTATGT	CTGGCCTGCTTTTGCATT	62
	Exon 6	TGCAATTCTGTGGCCTCTTC	GGGAGGCAAGGAGAGAGAAA	62
	Exon 7	TGCCTATGCTTGATTTGCAG	CCACTCCTGCGTCAAGTTCT	60
	Exon 8	CAACGGGACAGGAAGGAGTA	ACTTGAGCCCAGAAGGTTGA	60
	Exon 9	GTTTTCCCTTTGGGGACATT	GATGTCCTGCCATTTGATCC	60
	Exon 10	CAAATGGCAGGACATCAAGA	AGGAACCATGCTCCTGTGAA	66

*All PCR primers were tagged with M13 forward (TGTAAAACGACGGCCAGT) and reverse (CAGGAAACAGCTATGACC).
**2.5× MasterAmp enhancer.
†Advantage GC PCR.

incubated with a DIG-specific antibody covalently coupled to alkaline phosphatase, and detected by chemiluminescence with the addition of an alkaline phosphatase substrate. Average telomere length was determined by comparing the signals relative to the molecular weight standard using Image Quant TL software (Amersham Biosciences, Piscataway, NJ, USA).

Telomeric single-strand overhang length measurement

Telomeric overhang length was determined with the telomere-oligonucleotide ligation assay (T-OLA), as recently described [28,29]. Five μg of genomic DNA was hybridized with 0.5 pmol ^{32}P-end-labeled oligonucleotide $(CCCTAA)_4$ for 16 hours at 50°C in 10 μL, followed by ligation with 40 units of *Taq* DNA ligase (New England Biolabs, Beverly, MA, USA) for 5 hours at 50°C. Reaction products were precipitated, dried, resuspended in water, and denatured in formamide-containing loading buffer, and separated in 6% TBE-urea gel (Invitrogen, Carlsbad, CA, USA). As a DNA loading control, 10 ng of each T-OLA sample was used as template for a quantitative PCR for the *GADPH* gene (forward primer, 5′-ACC CCT TCA TTG ACC TCA AC-3′; reverse, 5′-CTT CTC CAT GGT GGT GAA GA-3′; 27 cycles, 55°C of annealing temperature for 30 seconds).

Telomerase activity in primary cells

The telomerase activity of activated T cells was evaluated by the telomeric-repeat amplification protocol with the fluorescent TRAPeze Telomerase Detection kit (Chemicon, Temecula, CA, USA) in cells from the patient and gender- and age-matched controls. Peripheral blood cells were cultured in RPMI 1640 with L-glutamine and 10% fetal calf serum in the presence of phytohemagglutinin (5 μg/mL) and interleukin-2 (40 IU/mL) for 6 days at 37°C with 5% carbon dioxide. Protein was extracted and telomerase activity was assayed according to the manufacturer's instructions, with slight modifications, as previously described [8].

TRF2 Western blot

Lymphocytes (5×10^5cells) isolated from either the periphery or from cord blood were prepared in 40 μL sodium dodecyl sulfate (SDS) loading buffer and subjected to the conventional Western blotting technique using the 1:1000 dilution of the crude anti-TRF2 antibody raised in rabbit (Sambrook and Maniatis, Molecular Cloning Manual).

Results

Genetic variation in TERF1

A total of 5244 base pairs (bp) of *TERF1*, encompassing the proximal promoter and all 10 exons, were sequenced in 47 patients. Exons 1 through 8 did not have SNPs in either the coding regions or the adjacent intronic sequences and therefore were not evaluated further. Exons 9 and 10 and the proximal promoter region of *TERF1* had variants in the first 47 patients and thus were sequenced in the additional 95 patients (n = 142 patients evaluated) and 289 controls.

Comparisons of all patients and controls are shown in Table 2. SNP c.37029T>C (E3669_418) was not in Hardy-Weinberg equilibrium in the controls and was excluded from the analysis. SNP c.36912C>T (rs3863242) in intron 9 with variant allele frequency of 37.3% in AA cases and 31.8% in controls was significant in the co-dominant ($p = 0.049$) and in the mutation-dominant model, odds ratio 1.59 ($p = 0.033$, 95% confidence interval 1.06–2.39),

Table 2. Genetic variation in *TERF1* and *TERF2* in acquired aplastic anemia and healthy controls

Gene	Location	SNP descriptor[c,d]	Sequence 5′ of SNP[e]	Wild-type (%)		Heterozygous (%)		Variant (%)		Additive χ^2 p value[f]	Mutation-dominant χ^2 p value[g]
				Cases	Controls	Cases	Controls	Cases	Controls		
TERF1[a]	-246 promoter	c.-246G>A	ATGGAGTCGT	141/142 (99.3)	287/287 (100)	1/142 (0.7)				0.155	0.155
	-239 promoter	c.-239T>C	TATGGAGTCG	142/142 (100)	285/287 (99.3)		1/285 (0.3)		1/285 (0.3)	0.608	0.319
	-238 promoter	c.-238C>T	TATGGAGTCG	142/142 (100)	279/285 (97.9)		6/285 (2.1)			0.082	0.082
	-183 promoter	c.-183C>G	CGCCCAGAGG	140/142 (98.6)	276/285 (96.8)	2/142 (1.4)	9/285 (3.2)			0.282	0.282
	-107 promoter	c.-107T>A	CGCTCGCCCA	142/142 (100)	285/286 (99.8)		1/286 (0.2)			0.480	0.480
	30106 intron 8	rs2306494:A>G	GAAGGAAAAA	54/142 (38.0)	101/289 (34.9)	67/142 (47.1)	116/289 (40.1)	21/142 (14.8)	72/289 (24.9)	0.053	0.531
	30153 intron 8	rs2306493:G>T	GGAATTTCAT	54/142 (38.0)	103/289 (35.6)	66/142 (46.5)	114/289 (39.5)	22/142 (15.5)	72/289 (24.9)	0.076	0.628
	30309 exon 9 Glu-Glu	c.30309A>G	TAACTCCTGA	142/142 (100)	288/289 (99.7)		1/289 (0.3)			0.483	0.483
	30320 exon 9 Ala/Val	c.30320C>T	AAACATCGAG	141/142 (99.3)	289/289 (100)	1/142 (0.7)				0.153	0.153
	30378 intron 9	E3667_590 c.30378G>A	GGACATTAAA	142/142 (100)	284/289 (98.3)		5/289 (1.7)			0.115	0.115
	36912 intron 9	rs3863242:C>T	AAAAATGTTA	56/142 (39.4)	145/285 (50.9)	66/142 (46.5)	99/285 (34.7)	20/142 (14.1)	41/285 (14.4)	0.049	0.033[h]
	37029[i] intron 9	E3669_418 c.37029T>C	TATTTTCTTT	141/142 (99.3)	255/287 (88.9)	1/142 (0.7)	28/287 (9.8)		4/287 (1.4)	0.0007[i]	0.0001[i]
	37045 intron 9	E3669_439 c.37045G>A	AATTGTTTCT	142/142 (100)	287/288 (99.7)		1/287 (0.3)			0.482	0.482
	37119 exon 10 Arg-Arg	E3669_513 c.37119G>A	CTGGCGTGAG	142/142 (100)	278/286 (97.2)		8/286 (2.8)			0.044	0.044[j]
	37189 exon 10 Leu-Leu	E3669_583 c.37189T>C	AAGTGTCATG	142/142 (100)	278/286 (97.2)		8/284 (2.8)			0.044	0.044[j]
	37342 3′ utr	c.37342T>C	AATTTAAAAC	140/142 (98.6)	273/284 (96.1)	2/140 (1.4)	11/284 (3.9)			0.163	0.163
TERF2[b]	1507 intron 3	c.1507T>C	GCTTTGGCAT	138/142 (97.2)	272/282 (96.5)	4/142 (2.8)	10/282 (3.5)			0.692	0.692
	17455 exon 6 Val-Val	E3673_252 c.17455G>A	CAGGGCCTGT	141/142 (99.3)	277/286 (96.9)	1/142 (0.7)	9/286 (3.1)			0.115	0.115
	17476 Exon 6 Ala-Ser	c.17476G>T	CAGAGAACCC	141/142 (99.3)	287/287 (100)	1/142 (0.7)				0.155	0.155
	17508 intron 6	E3673_301 c.17508G>A	TTGTGATCTT	103/142 (72.5)	215/286 (75.2)	36/142 (25.4)	69/286 (24.1)	3/142 (2.1)	2/286 (0.7)	0.413	0.556
	17546 intron 6	c.17546C>G	GCAGTCTCCT	142/142 (100)	286/287 (99.7)		1/287 (0.3)			0.481	0.481
	17554 intron 6	c.17554C>T	CTCCTTAGGC	142/142 (100)	285/287 (99.3)		2/287 (0.7)			0.319	0.319
	17570 intron 6	c.17570A>G	AGTTCCTCTC	142/142 (100)	285/287 (99.3)		2/287 (0.7)			0.319	0.319

[a]*TERF1* genomic contig NC_000008.9.

[b]*TERF2* genomic contig NC_000016.8.

[c]SNPs denoted with E numbers (i.e., E3555_374) are available at http://www.ncbi.nlm.nih.gov/SNP/ and will be assigned dbSNP numbers with the next download to the dbSNP database.

[d]SNP descriptors show the location of the SNP based on nomenclature guidelines from the Human Genome Variation Society (http://www.hgvs.org/mutnomen/). If a dbSNP number is not assigned, the Genbank sequence is indicated and SNP numbering is based on coding DNA. The A of the ATG start codon based on the NCBI Evidence Viewer is base number 1.

[e]10 base pairs of sequence 5′ to the SNP but not including the SNP is shown.

[f]SASv8.02 was used to generate 2 × 3 tables consisting of cases and controls, compared to homozygous wild-type, heterozygous, and homozygous variants, resulting in 2 degrees of freedom χ^2 values.

[g]In the mutation-dominant model, one copy of the variant was considered to be a risk factor. Therefore, 2 × 2 tables were generated comparing homozygous wild-type to heterozygous plus homozygous variant.

[h]Odds ratio = 1.59, 95% confidence interval 1.06–2.39.

[i]SNP c.37029T>C (E32669_418) was not in Hardy-Weinberg equilibrium in the controls and thus excluded from analyses.

[j]Low frequency of the SNP precluded calculations of odds ratios.

suggesting that this SNP could be a risk factor for aplastic anemia. SNPs 37119G>A (E3669_513) and 37189T>C (E3699_583), both synonymous in exon 10, did not occur in patients but each had an allele frequency of 1.4% in controls ($p = 0.044$ for each). The mutation-dominant model for these two SNPs also suggested an association ($p = 0.044$) but the low frequencies precluded further conclusions. Comparison of Caucasian patients (n = 101) with Caucasian controls (n = 125) did not show statistically significant associations (data not shown). Separate analysis of patients from the different ethnic groups, African-American (n = 17), Hispanic (n = 16), and Asian (n = 8), was not performed because of the relatively few patients representing these groups. The intronic SNPs identified in *TERF1* did not appear to affect intron-exon splice sites by analysis with GeneSplicer [26] and NNSPLICE (v0.9) [27].

Haplotype analysis of patients and controls from all ethnic groups (Table 3) showed statistically significant differences in a global test using HaploStats and PHASEv2.1 ($p = 0.03$ and 0.01, respectively). The haplotype, GTCCTAGACGTGGTT, showed borderline significance in ordinal trait analysis and comparison to the most common haplotype ($p = 0.057$ and 0.079, respectively). Analysis of more refined haplotypes proximal to this significant SNP as well as analysis of haplotypes containing only common SNPs (>10% variant allele frequency) did not show statistical significance (data not shown).

One patient, a 15-year-old Hispanic female, had a sequence variation that resulted in an alanine-to-valine change in codon 377 of exon 9 (c.30320C>T) of *TERF1*. This is a moderately conservative amino acid change (Grantham D value = 64) [30]. *In silico* analysis using SIFT and PolyPhen predicted that this amino acid change would be benign [24,25]. This patient, from the group of 95 consecutive patients, had normal telomere lengths in both the double-stranded and single-stranded telomeric sequences compared to age-matched controls (Fig. 1A and B). Therefore, we conclude that this amino acid substitution could be a rare but inconsequential sequence variant.

Genetic variation in TERF2

A total of 4515 bp of *TERF2* were sequenced in the screening sample of 47 acquired AA patients. The proximal promoter, exons 1, 2, 4, 5, and 7 through 10 did not have SNPs or other sequence changes. Sequence variation was observed in exons 3 and 6 (Table 2) and was thus evaluated in the additional 95 patients and 289 controls. Comparisons of the frequency of these SNPs did not show statistical significance in all comparisons performed. Haplotype analysis using global tests in PHASE and HaploStats were also not significant (Table 3). Comparisons of Caucasian patients and controls also did not reveal a statistically significant association. Analysis with GeneSplicer [26] and NNSPLICE (v0.9) [27] of the *TERF2* SNPs did not suggest any splice site variants.

We identified a Caucasian female with short telomeres and an alanine-to-serine change in codon 273 of exon 6 (c.17476G>T). This is a moderately conservative amino acid change (Grantham D value = 99) [30]. PolyPhen and SIFT analyses also predicted this amino acid change to be benign [24,25]. Despite having very short overall telomere lengths (Fig. 1A), this patient had normal telomerase activity and a typical single-stranded telomeric overhang

Table 3. Haplotype analysis of SNPs in *TERF1* and *TERF2* in acquired aplastic anemia and healthy controls

Gene	Haplotype	Case frequency (%)	Control frequency (%)	*p* value[a]	*p* value[b]	Global test *p* value, Haplostats	Global test *P* value, PHASEv2.1
TERF1	GTCCTGTACGCGGTT	38.02	43.34	0.154	Reference	0.03	0.01
	GTCCTAGACGTGGTT	35.56	28.86	0.057	0.079		
	GTCCTAGACGCGGTT	23.24	20.46	0.376	0.226		
	GTCGTAGACGCGGTT	0.35	1.22		0.280		
	GTCCTAGACGTGGTC	0.70	0.90		0.920		
TERF2	TGGGCCA	83.16	83.10	0.982	Reference	0.415	0.18
	TGGACCA	12.62	14.79	0.367	0.509		
	TAGGCCA	1.17	1.57	0.115	0.157		
	CGGGCCA	1.78	1.41	0.688	0.696		
	TGTGCCA	0	0.35				
	TGGGCCG	0.35	0				
	TGGGCTA	0.21	0				
	TGGGGCA	0.17	0				
	TGGACTA	0.14	0				

Haplotypes and global tests comparing cases and controls were determined using PHASEv2.1 and Haplostats. Haplostats was used for the additional statistical calculations. *TERF1* SNPc.36912C>T (rs3863242 in intron 9) is shown in bold. *TERF1* SNP c.37029T>C (E32669_418) was not used in this analysis since it was not in Hardy-Weinberg equilibrium. Data from all ethnic groups was combined to construct these haplotypes. Haplotypes less than 1% in frequency were excluded from this analysis.

[a] *p* values for haplotypes were determined for ordinal trait analysis.

[b] *p* values obtained using the most common haplotype as the reference to estimate odds ratios. *p* values were not less than 0.05; therefore odds ratios are not shown.

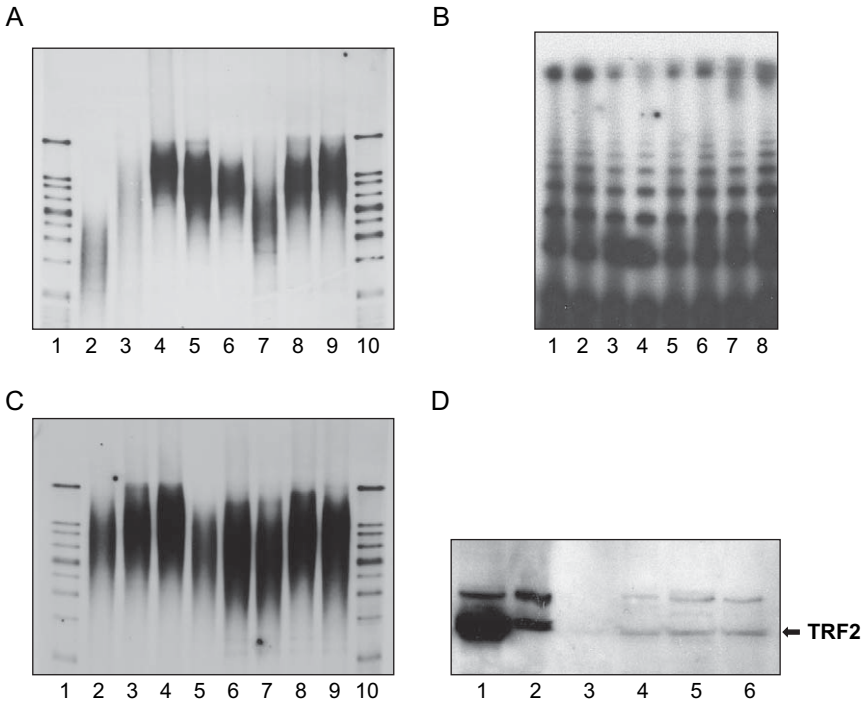

Figure 1. **(A)** Overall telomere length in peripheral blood leukocytes from *TERF1* and *TERF2* mutant patients and age- and gender-matched controls, measured by Southern blot. Lanes 1 and 10, molecular weight marker; lane 2, control DNA for long telomeres; lane 3, control DNA for short telomeres; lane 4, Ala377Val *TERF1* mutant patient; lanes 5 and 6, age- and gender-matched controls; lane 7, Ala273Ser *TERF2* mutant patient; lanes 8 and 9, age- and gender-matched controls. Ala377Val *TERF1* patient has normal telomeres in length (8.5 kb) in comparison to controls, and Ala273Ser *TERF2* patient has short telomeres compared to controls (5.3 kb vs 8.0 kb, respectively). **(B)** Length of telomeric single-strand overhangs by the T-OLA assay in *TERF1* and *TERF2* patients and controls. Lane 1, Ala377Val *TERF1* patient; lanes 2, 3, and 4, age- and gender-matched controls; lane 5, Ala273Ser *TERF2* patient; lanes 6, 7, and 8, age- and gender-matched controls. Telomeric overhangs are similar in length in patients and controls. **(C)** Telomere length of relatives of Ala273Ser *TERF2* patient. Lanes 1 and 10, molecular weight marker; lane 2, sister (noncarrier); lanes 3 and 4, controls; lane 5, mother (carrier); lane 6, father (noncarrier); lanes 7, 8, and 9, controls. **(D)** Western blot analysis for TRF2 protein expression in peripheral blood lymphocytes of Ala273Ser *TERF2* patient and controls. Lane 1, 293T cells transfected with pCMV-TRF2 vector; lane 2, 293T cells transfected with pCMV empty vector; lane 3, empty lane; lane 4, umbilical cord blood from a healthy donor; lane 5, healthy control; lane 6, Ala273Ser *TERF2* patient.

length (Fig. 1B). In addition, protein expression levels of TRF2 did not appear to be affected by the amino acid substitution (Fig. 1D). This patient had moderate aplastic anemia that did not require treatment until pregnancy, when she became transfusion dependent. Transfusions were not required after delivery. Interestingly, this patient developed chromosomal abnormalities, with 8% of monosomy 7 in cells isolated from the bone marrow. Sequence analysis of her immediate family members showed that her mother carried the same mutation while her father and sister did not. The Ala273Ser sequence variant did not appear to correlate with short telomeres, since the patient's sister also had shorter telomeres than healthy age- and gender-matched controls, while their parents had normal telomere lengths (Fig. 1C). However, the patient's telomeres are significantly shorter than her sister's, suggesting a multifactorial genetic component.

Discussion

The leukocytes of patients with AA often have shortened telomeres. Since *TERF1* and *TERF2* play critical roles in the maintenance of telomeres, we screened these two genes for mutations and determined common variation in patients with AA. The initial screening samples (n = 47) were chosen on the basis of diagnosis of acquired AA and either short leukocyte telomeres, lack of response to immunosuppressive therapy, or positive family history of hematologic abnormalities without physical anomalies characteristic of dyskeratosis congenita. If a sequence variant was present in the first 47 patients, that region was sequenced in an additional 95 patients and 289 controls. An analysis of the controls confirmed that variation in *TERF1* and *TERF2* is limited [16]. Some of the patients in this study overlapped with prior studies of mutations in *TERC* and *TERT* [7,8]. One of our patients had a *TERC* mutation and another had a *TERT* mutation (Val1090Met). However, neither of these patients had the *TERF1* intron 9 c.36912C>T (rs3863242) or the nonsynonymous mutations described.

This study suggests that the SNP in intron 9, c.36912C>T (rs3863242), of *TERF1* may occur more commonly in aplastic anemia and was associated with an increased risk for AA. This SNP is part of a less common haplotype that was seen more commonly in AA patients.

Haplotypes constructed around the significant SNP or using *TERF1* SNPs with greater than 10% variant allele frequency did not show statistical significance. However, overall, the statistical significance of the *TERF1* SNP and haplotype association is modest, given the limited population size and statistical tests performed. It is still possible that the genetic variation described in *TERF1* is an important risk factor for AA and our study should be followed up with a larger study.

Functional SNPs, common mutations, or mutations that caused radical amino acid changes, insertions, or deletions were not identified in *TERF1* and *TERF2* in this study. The 15-year-old female patient with the alanine-to-valine change in codon 377 of *TERF1* had normal telomere and single-stranded telomeric overhang lengths. This conservative amino acid change, therefore, is not likely to be causative. However, we cannot exclude the possibility that in combination with variation in other genes of the telomere complex, this *TERF1* sequence variant could prove to be significant.

An alanine-to-serine change in codon 273 of *TERF2* was found in another female patient with very short telomeres and the patient's mother. Interestingly, the patient's sister also had short telomeres when compared to the age- and gender-matched controls but her telomeres were not as short as the patient's, although there was only a one-year age difference. It is also notable that the parents had relatively short telomeres compared to the age- and gender-matched controls, although they were still within the expected range. Therefore, it is possible that this Ala273Ser sequence variant in *TERF2* may still play a role in telomere maintenance in this patient, and in combination with other, currently unidentified genetic factors could contribute to short telomeres and AA.

In conclusion, it is possible that a common genetic variant in *TERF1* is associated with risk for AA but additional studies are required. Highly penetrant, nonsynonymous mutations and insertion-deletion mutations in *TERF1* and *TERF2* were not identified and therefore are not likely to be major genetic risk factors for the development of AA.

Acknowledgments

We are much appreciative to the patients and their families for participation in the study. The Core Genotype Facility at the National Cancer Institute assisted with sequencing samples. This research was supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

References

1. Young NS. Acquired aplastic anemia. Ann Intern Med. 2002;136:534–546.
2. Brummendorf TH, Maciejewski JP, Mak J, Young NS, Lansdorp PM. Telomere length in leukocyte subpopulations of patients with aplastic anemia. Blood. 2001;97:895–900.
3. Ball SE, Gibson FM, Rizzo S, et al. Progressive telomere shortening in aplastic anemia. Blood. 1998;91:3582–3592.
4. Knight SW, Heiss NS, Vulliamy TJ, et al. X-linked dyskeratosis congenita is predominantly caused by missense mutations in the DKC1 gene. Am J Hum Genet. 1999;65:50–58.
5. Knight SW, Vulliamy TJ, Morgan B, et al. Identification of novel DKC1 mutations in patients with dyskeratosis congenita: implications for pathophysiology and diagnosis. Hum Genet. 2001;108:299–303.
6. Vulliamy T, Marrone A, Szydlo R, et al. Disease anticipation is associated with progressive telomere shortening in families with dyskeratosis congenita due to mutations in TERC. Nat Genet. 2004;36: 447–449.
7. Yamaguchi H, Baerlocher GM, Lansdorp PM, et al. Mutations of the human telomerase RNA gene (TERC) in aplastic anemia and myelodysplastic syndrome. Blood. 2003;102:916–918.
8. Yamaguchi H, Calado RT, Ly H, et al. Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. N Engl J Med. 2005;352:1413–1424.
9. Shay JW, Zou Y, Hiyama E, Wright WE. Telomerase and cancer. Hum Mol Genet. 2001;10:677–685.
10. Maser RS, DePinho RA. Connecting chromosomes, crisis, and cancer. Science. 2002;297:565–569.
11. van Steensel B, de Lange T. Control of telomere length by the human telomeric protein TRF1. Nature. 1997;385:740–743.
12. Smith S, de Lange T. TRF1, a mammalian telomeric protein. Trends Genet. 1997;13:21–26.
13. van Steensel B, Smogorzewska A, de Lange T. TRF2 protects human telomeres from end-to-end fusions. Cell. 1998;92:401–413.
14. Karlseder J, Smogorzewska A, de Lange T. Senescence induced by altered telomere state, not telomere loss. Science. 2002;295:2446–2449.
15. Smogorzewska A, van Steensel B, Bianchi A, et al. Control of human telomere length by TRF1 and TRF2. Mol Cell Biol. 2000;20:1659–1668.
16. Savage SA, Stewart BJ, Eckert A, et al. Genetic variation, nucleotide diversity, and linkage disequilibrium in seven telomere stability genes suggest that these genes may be under constraint. Hum Mutat. 2005; 26(4):343–350.
17. Broccoli D, Smogorzewska A, Chong L, de Lange T. Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. Nat Genet. 1997;17:231–235.
18. The International Agranulocytosis and Aplastic Anemia Study. Incidence of aplastic anemia: the relevance of diagnostic criteria. Blood. 1987;70:1718–1721.
19. Packer BR, Yeager M, Staats B, et al. SNP500Cancer: a public resource for sequence validation and assay development for genetic variation in candidate genes. Nucleic Acids Res. 2004;32(Database issue):D528–D532.
20. Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet. 2001;27:234–236.
21. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 2001; 68:978–989.
22. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet. 2003;73:1162–1169.
23. Lake SL, Lyon H, Tantisira K, et al. Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. Hum Hered. 2003;55:56–65.
24. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31:3812–3814.
25. Sunyaev S, Ramensky V, Koch I, et al. Prediction of deleterious human alleles. Hum Mol Genet. 2001;10:591–597.
26. Pertea M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. Nucleic Acids Res. 2001;29:1185–1190.

27. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. J Comput Biol. 1997;4:311–323.
28. Stewart SA, Ben-Porath I, Carey VJ, et al. Erosion of the telomeric single-strand overhang at replicative senescence. Nat Genet. 2003; 33:492–496.
29. Cimino-Reale G, Pascale E, Battiloro E, et al. The length of telomeric G-rich strand 3′-overhang measured by oligonucleotide ligation assay. Nucleic Acids Res. 2001;29(7):E35.
30. Grantham R. Amino acid difference formula to help explain protein evolution. Science. 1974;185:862–864.