Cell, Vol. 83, 655–666, November 17, 1995, Copyright © 1995 by Cell Press

The Bloom's Syndrome Gene Product Is Homologous to RecQ Helicases

Nathan A. Ellis,*† Joanna Groden,†‡ Tian-Zhang Ye,* Joel Straughen,‡ David J. Lennon,* Susan Ciocci,* Maria Proytcheva,* and James German*
*Laboratory of Human Genetics
New York Blood Center
New York, New York 10021
‡Department of Molecular Genetics
Biochemistry and Microbiology
University of Cincinnati College of Medicine
Cincinnati, Ohio 45267

Summary

The Bloom's syndrome (BS) gene, *BLM*, plays an important role in the maintenance of genomic stability in somatic cells. A candidate for *BLM* was identified by direct selection of a cDNA derived from a 250 kb segment of the genome to which *BLM* had been assigned by somatic crossover point mapping. In this novel mapping method, cells were used from persons with BS that had undergone intragenic recombination within *BLM*. cDNA analysis of the candidate gene identified a 4437 bp cDNA that encodes a 1417 amino acid peptide with homology to the RecQ helicases, a subfamily of DExH box-containing DNA and RNA helicases. The presence of chain-terminating mutations in the candidate gene in persons with BS proved that it was *BLM*.

Introduction

Bloom's syndrome (BS) (reviewed by German, 1993) is a rare genetic disorder, the major clinical manifestations of which are small size, sun-sensitive facial erythema, and immunodeficiency. Homozygosity for mutation at *BLM*, the BS locus, constitutes a mutator genotype: somatic cells from persons with BS accumulate excessive numbers of mutations at all loci examined, including both coding sequences and noncoding repetitive DNA. Excessive numbers of chromatid gaps and breaks are visible in cultured but otherwise untreated cells, and exchanges between chromatids are increased in number. Interchanges between homologous chromosomes are increased, signifying increased somatic crossing over, and an abnormally great number of sister chromatid exchanges (SCEs) are present in bromodeoxyuridine (BrdU)-treated cells. Molecular studies have confirmed the increased rate of somatic recombination in BS cells. An important clinical consequence of the excessive genomic instability of BS is an enormous predisposition to the generality of human cancers.

Cytogenetic and biochemical evidences have shown that various aspects of DNA replication are disturbed in BS cells. However, the biochemical candidate gene approach to the identification of the primary enzymatic defect had been unsuccessful, in part owing to the numerous cellular effects of the BS phenotype; therefore, positional cloning was initiated. *BLM*, already assigned to human chromosome 15 (McDaniel and Schultz, 1992), was regionally mapped by the analysis of affected persons whose parents are cousins (German et al., 1994). Tight linkage was demonstrated between *BLM* and *FES*, a gene localized to band 15q26.1 by fluorescence in situ hybridization (Mathew et al., 1993). Because the data from the homozygosity mapping indicated that *BLM* was situated within 1 cM of *FES*, a yeast artifical chromosome (YAC) and P1 contig of the region flanking *FES* was constructed, and polymorphic markers in that region were identified (Straughen et al., submitted). Recently, a discovery was made that led to a powerful approach to localize *BLM* exactly.

The hypermutability of BS cells includes hyperrecombinability. Although cells from all persons with BS exhibit the diagnostic high SCE rate, in some persons a minor population of low SCE lymphocytes exists in the blood. Lymphoblastoid cell lines (LCLs) with low SCE rates can be developed from these low SCE lymphocytes. In multiple low SCE LCLs examined from 11 persons with BS, polymorphic loci distal to *BLM* on 15q had become homozygous in LCLs from five persons, whereas polymorphic loci proximal to *BLM* remained heterozygous in all low SCE LCLs. These observations supported the hypothesis that low SCE lymphocytes arose through recombination within *BLM* in persons with BS who had inherited paternally and maternally derived *BLM* alleles mutated at different sites. Such a recombinational event in a precursor stem cell in these compound heterozygotes gave rise to a cell whose progeny had a functionally wild-type gene and phenotypically a low SCE rate (Ellis et al., 1995).

The low SCE LCLs in which reduction to homozygosity had occurred were used for localizing *BLM* by an approach we refer to as somatic crossover point (SCP) mapping. The precise map position of *BLM* was determined by comparing the genotypes of the recombinant low SCE LCLs from the five persons mentioned above with their constitutional genotypes at loci in the region around *BLM*. The strategy was to identify the most proximal polymorphic locus possible that was constitutionally heterozygous and that had been reduced to homozygosity in the low SCE LCLs, and to identify the most distal polymorphic locus possible that had remained constitutionally heterozygous in them. *BLM* would have to be in the short interval defined by the reduced (distal) and the unreduced (proximal) heterozygous markers. The power of this approach was limited only by the density of polymorphic loci available in the immediate vicinity of *BLM*. In this report, we describe the SCP mapping, isolation, and identification of *BLM*.

†The first two authors contributed equally to this work.

Results

Localization of *BLM* to a 250 kb Interval

BLM previously was localized by SCP mapping to a 1.3 cM interval bounded proximally by *D15S116* and distally by four tightly linked loci *D15S127*, *FES*, *D15S158*, and *IP15M9* (Ellis et al., 1995). The four loci are present in a 1–2 cM interval on chromosome 15 (Beckmann et al., 1993; Gyapay et al., 1994). The order of these four loci was determined by polymerase chain reaction (PCR) analysis of clones in a 2 Mb YAC and P1 contig that encompasses *BLM* (Straughen et al., submitted). The four loci were oriented with respect to the telomere by finding a recombinant chromosome in a BS family in which crossing over had occurred between *BLM* and *IP15M9*, placing *IP15M9* on the distal end of the contig (Figure 1A). Because *D15S127* was the most proximal locus that was reduced to homozygosity in low SCE LCLs, polymorphic loci in the region proximal to it were sought. There, a polymorphic locus, *D15S1108*, was identified that remained constitutionally heterozygous in the recombinant low SCE LCLs, in contrast with locus *D15S127*, which had become homozygous in them (Figure 1B). This shift from heterozygosity to homozygosity of markers indicated that *BLM* is situated in the 250 kb region between *D15S1108* and *D15S127*.

Two genes, *FES* and *FUR*, map distal to *D15S127* in this region of chromosome 15. SCP mapping thereby elim-

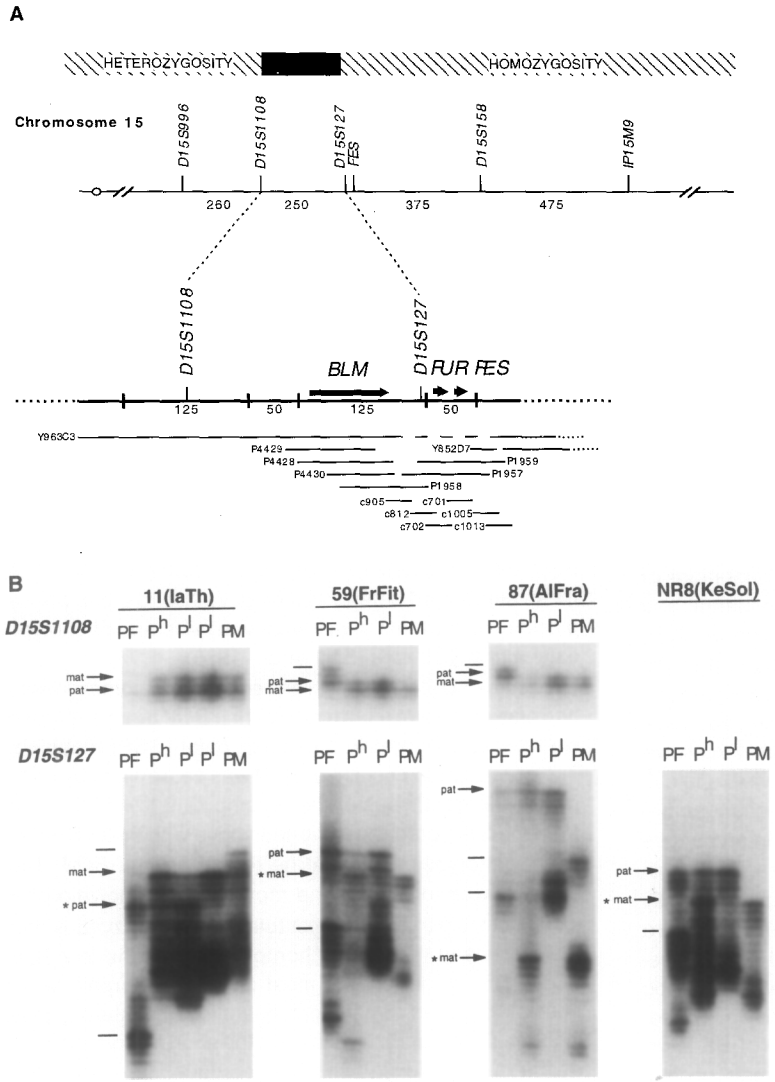

Figure 1. SCP Mapping of *BLM*

(A) Genetic map of the *BLM* region of 15q. On the upper horizontal line, the order and distances (shown in kilobases) between the polymorphic microsatellite loci were estimated by long-range restriction mapping (Straughen et al., submitted). The distance between *D15S127* and *FES* (not indicated) was determined to be 30 kb by restriction enzyme mapping of a cosmid contig (see below). Vertical lines indicate the position of the marker loci, and the circle represents the centromere. The interval between loci *D15S1108* and *D15S127* is expanded below the map. Vertical lines intersecting mark the unmethylated CpG-rich regions identified by long-range restriction mapping, and arrows indicate the direction of transcription of three genes in the region. Certain YACs (Y), P1s (P), and cosmids (c) from the contig (Straughen et al., submitted) are depicted by horizontal lines underneath the map. Dashes on the YAC lines indicate internal deletions. At the top, the horizontal cross-hatched bars indicate regions proximal to *BLM* that remained heterozygous in the low SCE LCLs and regions distal to *BLM* that had become homozygous. The minimal region to which *BLM* was thus assigned by SCP mapping is represented by a closed box.

(B) Autoradiographic evidence showing heterozygosity proximal to *BLM* and reduction to homozygosity distal to *BLM*. The four persons of five from whom low SCE LCLs had been established that were informative at *D15S1108* or *D15S127* are shown. To determine both the constitutional and the recombinant cell line genotypes, we carried out PCRs using DNA samples prepared from high SCE cells (P^h) and low SCE LCLs (P^l) of persons with BS as well as samples from their fathers (PF) and their mothers (PM). These persons are identified by their Bloom's Syndrome Registry designations (see German and Passarge, 1989). Arrows point to DNA fragments amplified from the heterozygous alleles of the constitutional genotypes, pat (for paternal) and mat (for maternal). Asterisks mark alleles in the low SCE LCLs that are lost through somatic crossing over. Lines mark DNA fragments amplified from alleles of the parents but that were not transmitted to the offspring with BS. From one of the four persons with BS, 11 different clonal LCLs were examined; three of the 11 had undergone reduction to homozygosity at loci distal to *BLM*, as explained elsewhere (Ellis et al., 1995). Autoradiographic patterns are shown from two of the 11 low SCE LCLs from 11(laTh), one representative of cell lines in which allele losses were detected (P^l sample on right) and another of cell lines in which they were not (P^l sample on left).

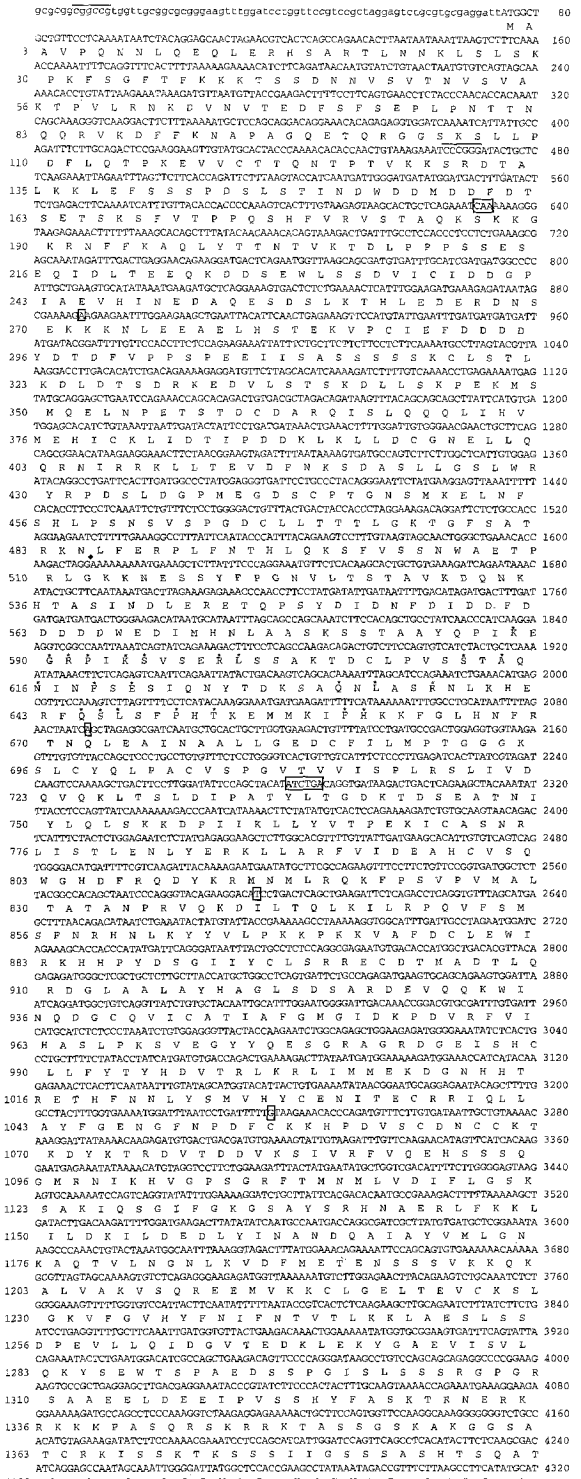

Figure 2. The 4437 bp H1-5′ Sequence, Which Represents the Merged Sequences of the H1 cDNA and the 5′ Clones, with Its Encoded 1417 Residue Amino Acid Sequence in Single Letter Code

Nucleotides in the open reading frame starting at the first in-frame ATG, 75 bp from the first nucleotide of the H1-5′ sequence, are capitalized. The in-frame nonsense codon (TAA) marked by the period is followed by 88 nt of 3′ untranslated sequence. At the initiator methionine, there is a Kozak consensus sequence (Kozak, 1989), and an acceptable polyadenylation sequence (underlined) is present 20 bp upstream of a 21 bp poly(A) tail. Sites at which substitution or deletion were detected in persons with BS (see Table 1) are boxed, and a site at which an insertion was identified is marked by a diamond. The EagI and SmaI sites used in the construction of a full-length cDNA referred to as B3 (see Experimental Procedures) are overlined. Asterisks mark amino acid identities to three motifs present in the RNA polymerase II largest subunit (see text).

inated them as candidates for *BLM*. Consistent with this conclusion, an earlier mutation search in six BS LCLs had failed to uncover mutations in *FUR* (data not shown).

Isolation of a Candidate for *BLM*

cDNAs were isolated from the 250 kb region between *D15S1108* and *D15S127* by direct cDNA selection using cDNA libraries from cultured fibroblasts and the T cell line Jurkat. Libraries from these cell lines were chosen because fibroblasts and T lymphocytes from persons with BS exhibit the high SCE phenotype, indicating that *BLM* is expressed in these cell types. In direct selection experiments using cosmid c905 (see Figure 1A), an 847 bp cDNA designated 905-28 was isolated after two rounds of direct selection. It was found in fewer than 1 in 1 × 10^6 clones screened in the fibroblast library but was present in 6 of 28 selected cDNA clones, a 250,000-fold enrichment. The six cDNAs represented by 905-28 were the only selected cDNAs that by Southern blot analysis mapped to the *BLM* region and that identified nonrepetitive sequences in the human genome (data not shown). The 905-28 cDNA identified single-copy sequences that are situated approximately 55 kb proximal to *FUR* (Figure 1A).

The 905-28 cDNA then was used to screen a HeLa cDNA library; 28 cDNAs were isolated, representing at least 15 distinct classes of overlapping clones. Each of these classes had the same sequence as the 905-28 cDNA at the 3′ ends but a different length of 5′ sequence. In the longest cDNA isolated, clone H1, a long reading frame was found that was open to the 5′ end. Additional sequences upstream of the start of the H1 cDNA were identified by a PCR cloning method (see Experimental Procedures). Clones extending 5′ of the H1 cDNA were isolated from the HeLa library, permitting the identification of 4437 bp of sequence, which we refer to as the H1-5′ sequence (Figure 2).

Starting at the first in-frame ATG 75 bp from its 5′ end, the H1-5′ sequence encodes a 1417 amino acid peptide with a predicted molecular mass of 159 kDa. No in-frame stop codons were present between this ATG and the 5′ end of the H1-5′ sequences. An extensive cDNA analysis was carried out to map the 5′ end of the candidate gene. We screened 8 × 10^6 LCL cDNA clones by hybridization with a 5′ probe. We isolated 11 clones and sequenced their 5′ ends (Figure 3). In addition, 12 fibroblast clones prepared by a 5′ rapid amplification of cDNA ends (RACE) technique were sequenced. Both analyses indicated that the H1-5′ sequence is full length.

The predicted peptide encoded in the H1-5′ sequence was used to carry out a BLASTP search of amino acid

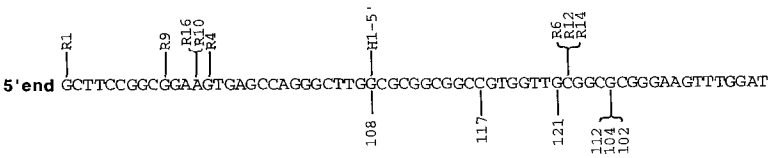

Figure 3. Nucleotide Sequence of the 5′ End of the Candidate Gene Determined by cDNA Analysis and 5′ RACE Experiments

The sequence of the longest cDNA isolated (clone R1) is shown. The sequences were obtained by analysis of 11 lymphoblastoid cDNAs (clone names prefixed by an R), identified by screening 8 × 10^6 clones with an EagI–SmaI DNA fragment from the 5′ part of the H1-5′ sequence (Figure 2) and of 12 5′ RACE clones amplified from fibroblast cDNA with nested PCR primers (see Experimental Procedures). Vertical lines mark the nucleotides at which nine lymphoblastoid cDNA (clones named above the sequences) and six cloned 5′ RACE fragments (clones named below the sequences) initiated. Three cDNA and six 5′ RACE clones not shown contained sequences that initiated less than 38 bp upstream of the first in-frame ATG. The sequences at the 5′ end are GC rich (71%), perhaps explaining the absence of in-frame nonsense codons upstream of the first in-frame ATG.

sequence databases. The searches identified significant homologies to the three known peptides in the RecQ subfamily of DExH box–containing helicases (Figure 4). The amino acid identities were concentrated in the region (residues 649–1041) containing the seven conserved helicase domains of the human *RECQL* (44%), Saccharomyces cerevisiae *SGS1* (43%), and Escherichia coli *recQ* (42%) genes. Thus, the product of the candidate gene contains motifs homologous to those found in DNA helicases, suggesting that the protein is an enzyme engaged in DNA manipulation.

The seven helicase domains identified by their homology to RecQ constitute only the middle third of the predicted peptide. Between residues 588 and 661, amino acid identities were discovered with three short motifs present in a broad phylogenetic spectrum of RNA polymerase II largest subunits (marked by asterisks in Figure 2). The function of these motifs is unknown. No other significant homologies were identified to amino acid sequences in databases.

The amino acid composition of the nonhelicase regions of the predicted peptide is unusual. The amino-terminal 648 residues of the peptide are rich in acidic (17%), basic (12%), and polar (34%) amino acids; 13% of the residues are serines. Similarly, the carboxy-terminal 376 residues also are rich in acidic (11%), basic (16%), and polar (30%) amino acids; and, again, 14% of the residues are serines. The function of these highly charged regions is unknown.

RNA Expression of the Candidate Gene in Cultured Cells

Northern blot analysis was used to determine the size of the full-length transcript from the candidate gene. The H1 cDNA was hybridized to total RNAs prepared from HeLa cells, normal diploid cultured fibroblasts, and non-BS LCLs. Two RNA bands at approximately 4.5 kb were visualized on the autoradiogram (Figure 5A). This size is consistent with the length of the longest cDNAs sequenced (see Figures 2 and 3).

In addition, Northern blot analysis was performed using total RNAs prepared from LCLs from seven unrelated persons with BS (Figure 5B). In three BS LCLs, the quantity of RNAs identified by hybridization to the H1 cDNA was decreased in comparison to that of the control LCLs. In four BS LCLs, the pattern of the two RNA bands was abnormal compared to that in normal cells: in one BS LCL, the lower band was absent; in another, the upper band is absent; and in two others, the intensity of the lower of the two bands was increased and the upper decreased. The RNA loading was equal in all the lanes as evidenced by hybridization with a probe for the glyceraldehyde-3-phosphate dehydrogenase (*G3PD*) gene. These observations suggest that RNAs identified by the H1 cDNA might be destabilized in BS LCLs as result of mutations in the candidate gene (see Surdej et al., 1994). The derivation of the two RNA bands from the gene is unknown, and the cause of these abnormalities is under investigation.

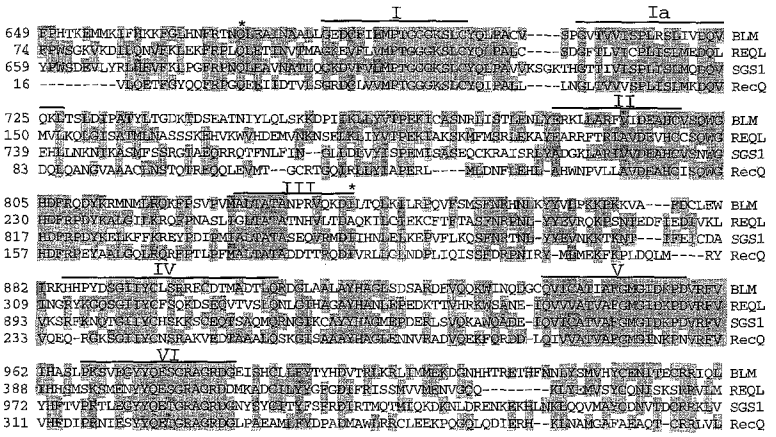

Figure 4. Amino Acid Sequence Homologies in the Seven Conserved Helicase Domains among the Putative Peptide Encoded by the H1-5′ Sequence and the Three Other Known Members of the RecQ Subfamily of Helicases

The numbers on the left indicate amino acid positions in each peptide, and gene product names are on the right. Sequence alignments were performed by the Megalign computer program (DNAStar); dashes indicate gaps inserted by the program to maintain alignment. Amino acid residues that are identical at a position between sequences are stippled. Two different stipplings are used when at a position two pairs of identical amino acids were observed. Overlined sequences mark the seven helicase domains (Gorbalenya et al., 1989). The DExH box is in helicase domain II. Asterisks denote positions at which putative missense mutations were identified. The candidate gene product is referred to here as BLM because mutations have been discovered in the gene in persons with BS (see text).

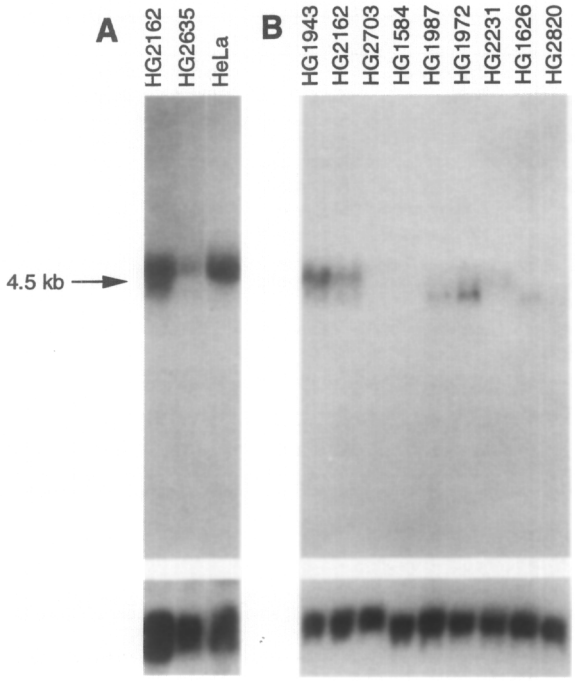

Figure 5. Northern Blot Analysis of the H1-5′ Sequences Expressed in Cultured Cells

In (A), RNA preparations were analyzed from HG2162, a normal LCL; HG2635, a normal diploid fibroblast cell line; and HeLa cells. In (B), RNA preparations were analyzed from HG1943 and HG2162, which are normal LCLs, and from HG2703, HG1584, HG1987, HG1972, HG2231, HG1626, and HG2820, which are BS LCLs. Total RNA (30 μg) from each cell line was loaded in each lane. Labeled probes (the H1 cDNA shown in the upper panels and a cDNA for *G3PD* shown in the lower panels) were hybridized to membranes of the blotted gels and, after washing, the membranes were exposed for 1–3 days (H1 DNA) or for 15 min (G3PG cDNA). On a 7 day exposure, faint bands resembling the hybridization pattern in normal cells were detected at the 4.5 kb position in HG2703, HG1584, and HG2820. The LCLs developed from persons with BS are shown in Table 1, except HG2703 (NR2[CrSpe]) and HG2820 (142[MaMatu]).

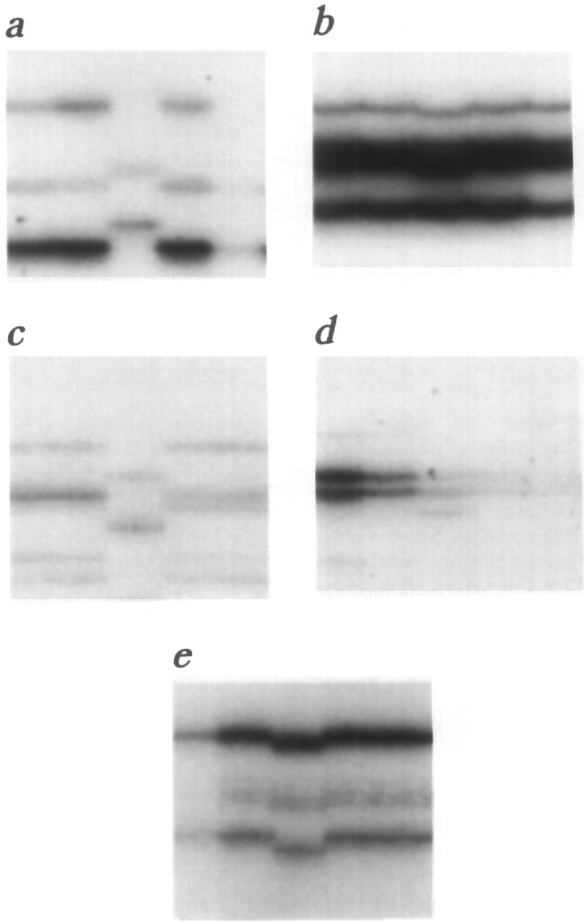

Figure 6. Novel SSCP Conformers Detected in cDNA Samples Isolated from BS LCLs after PCR Amplification of the Candidate Gene

Each panel includes five lanes of cDNAs from five unrelated persons with BS amplified with oligonucleotides designed from a unique region of the Candidate gene. The novel conformers in which mutations were detected are shown in the center lanes of each panel: BS LCL HG1514 from 15(MaRo) in (a), BS LCL HG1624 from 113(DaDem) in (b), BS LCL HG1926 from 97(AsOk) in (c), BS LCL HG2231 from 139(ViKre) in (d), and BS LCL HG1626 from 93(YoYa) in (e). Not shown are the novel conformers detected in 92(VaBi) and 112(NaSch).

Mutations in the Candidate Gene in Persons with BS

To determine whether the candidate gene is *BLM*, we prepared RNAs from LCLs from 13 unrelated persons with BS and from cell lines from four unaffected controls. These RNAs were used to generate cDNAs for mutational analysis of the expressed sequences of the candidate gene. Sequences in these 13 BS and four control non-BS cDNAs were amplified in approximately 200 bp segments using PCR primers designed from the open reading frame in the H1-5′ sequence (see Experimental Procedures). The amplified segments were analyzed by single strand conformation polymorphism (SSCP) analysis using two conditions for electrophoresis. Novel SSCP conformers (Figure 6) were identified, and the genetic changes underlying them were sequenced (Table 1).

We identified seven unique mutations in ten persons with BS (the boxed and diamond-marked nucleotides in Figure 2), as well as four polymorphic base pairs that will not be described here. Of the mutations, four introduced premature nonsense codons into the coding sequence and three introduced amino acid substitutions (see below); one of the four chain-terminating mutations arose by a 3 bp deletion, one by a nucleotide substitution, one by a 1 bp insertion that caused a frameshift, and one by a 6 bp deletion accompanied by a 7 bp insertion that also caused a frameshift. This last mutation was detected in all four persons with Ashkenazi Jewish ancestry. The potential products encoded in these four mutant alleles are 185, 271, 515, and 739 amino acids in length, respectively, and none contains a complete set of the seven helicase domains. Of these mutant alleles, three were detected in the homozygous state, indicating that the persons inheriting them in double dose probably have no active *BLM* gene product in their cells. These observations are evidence that the H1-5′ sequences are mutated in persons with BS, thereby proving that the candidate gene is *BLM*.

Finally, two putative missense mutations were identified

Table 1. Mutations Identified in the Candidate Gene in Persons with BS

Identification[a]	Ancestry	Cell Line	Position[b] (bp)	Alteration	Zygosity at Gene	Kind	Codon Change	Predicted Peptide[c]
97(AsOk)	Japanese	HG1926	631	3 bp deletion[d]	Homozygous	Nonsense	S→stop	185
112(NaSch)	American/European	HG2510	888	A→T	Heteroygous	Nonsense	K→stop	271
93(YoYa)	Japanese	HG1626	1610	1 bp insertion	Homozygous	Frameshift[e]	Q→R[f]	515
139(ViKre)	American/European	HG2231	2089	A→G	Heteroygous	Missense		1417
15(MaRo)	Ashkenazi Jewish	HG1514	2281	6 bp detetion and 7 bp insertion	Homozygous	Frameshift[g]		739
42(RaFr)	Ashkenazi Jewish	HG2522	2281	6 bp deletion and 7 bp insertion	Homozygous	Frameshift[g]		739
107(MyAsa)	Ashkenazi Jewish	HG2654	2281	6 bp deletion and 7 bp insertion	Homozygous	Frameshift[g]		739
Nr2(CrSpe)	Ashkenazi Jewish	HG2727	2281	6 bp deletion and 7 bp insertion	Homozygous	Frameshift[g]		739
92(VaBi)	Italian	HG1584	2596	T→C	Homozygous	Missense	I→T[h]	1417
113(DaDem)	Italian	HG1624	3238	G→C	Homozygous	Missense	C→S	1417

[a] Bloom's Syndrome Registry designations. Three unrelated persons with BS were examined in whom mutations have yet to be detected: 61(DoHo), in HG2122; 30(MaKa), in HG1987; 140(DrKas), in HG1972.
[b] The nucleotide positions are as identified in the H1-5′ sequence (Figure 2).
[c] Number of amino acids starting from the first in-frame ATG found in the H1-5′ sequence (Figure 2).
[d] The deletion of CAA at nucleotide positions 631–633 results in a stop codon at amino acid position 186 (Figure 2).
[e] The insertion of an A base pair causes the insertion of a novel codon for K after amino acid 514 position (taken from the H1-5′ sequence; Figure 2).
[f] At amino acid position 672.
[g] The deletion of ATCTGA and insertion of TAGATTC causes the insertion of the novel codons for LDSR after amino acid position 736, and after these codons there is a stop codon.
[h] At amino acid position 843.
[i] At amino acid position 1055.

in two persons with BS that introduced amino acid substitutions at residues conserved in RecQ helicases (residues with asterisks in Figure 4), and one was identified that introduced an amino acid substitution of cysteine to serine in the carboxy-terminal region of the peptide. Because the three genetic alterations could be polymorphisms and the actual BS-associated mutations could have gone undetected, analyses of the *BLM* gene product in vitro will be required to demonstrate whether these substitutions cause the mutant phenotype.

Discussion

In the present study, *BLM* was isolated by a positional cloning strategy. *BLM* first was localized by homozygosity mapping to a 2 cM interval flanking *FES* (German et al., 1994), a gene already mapped to chromosome band 15q26.1. A 2 Mb YAC and P1 contig encompassing *FES* was constructed, and closely spaced polymorphic DNA markers in the contig were identified (Straughen et al., submitted). *BLM* then was assigned by SCP mapping to a 250 kb interval in the contig, one bounded by the polymorphic loci *D15S1108* and *D15S127* (Figure 1). A cDNA clone (905-28) was isolated by direct cDNA selection using a cosmid clone from the interval, and cDNA analysis identified the 4437 bp H1-5′ sequence (Figure 2). This sequence encodes a putative peptide homologous to the RecQ helicases (Figure 4). RNA transcripts 4.5 kb long were identified by Northern blot analysis (Figure 5A), and electrophoretic abnormalities in RNAs were detected in cells from seven unrelated persons with BS, suggesting that these RNAs are derived from mutant *BLM* genes (Figure 5B). Finally, reverse transcription PCR and SSCP analyses disclosed seven unique mutations in ten persons with BS (Table 1; Figure 6), four that are chain terminating and three that are putative missense substitutions, two of the three affecting amino acid residues conserved in RecQ helicases and the third changing a cysteine to a serine.

SCP Mapping, a Powerful Strategy

In a recent tabulation of the 42 inherited disease-associated genes isolated by positional cloning (Collins, 1995), 19 were transmitted as autosomal dominants and 17 as X-linked recessives; however, only five were autosomal recessives. The reasons for the relative paucity of positionally cloned autosomal recessive disease-associated genes are at least twofold. First, the cloning of over half of the genes (26 of the 42 tabulated) was aided by chromosome breakpoints within or near the disease-associated gene; however, only one of these was in an autosomal recessive. Second, and of greater importance, the number of families transmitting rare autosomal recessive disease-associated genes generally is small, and the number of persons in sibships who would be informative in recombinational analysis also is small. Because a single investigator usually cannot obtain the numbers of families required for linkage analysis, the localization and subsequent positional cloning of rare autosomal recessive genes have lagged behind that of dominant and X-linked recessive genes.

Even when samples from numerous families have been collected and analyzed, usually the amount of positional information obtained is limited. In the case of BS, the Bloom's Syndrome Registry (German and Passarge, 1989), a research resource that has provided the material for all our recent genetical studies, made possible an extensive recombinational analysis of *BLM* by homozygosity mapping. This analysis permitted a minimum regional assignment of *BLM* to approximately 1.3 Mb (unpublished data). This size of minimum interval is typical of recombinational analysis. A search for and subsequent mutational analysis of genes from a 1.3 Mb region would have been laborious.

The problem of too little positional information in available families can be mitigated in exceptional situations in which linkage disequilibrium between the disease-associated gene and tightly linked polymorphisms can be detected in a genetic isolate. In these cases, localization of a gene to a short interval in the genome by haplotype analysis can be more exact than is possible using standard linkage analysis of family data (e.g., Kerem et al., 1989; Sirugo et al., 1992; Lehesjoki et al., 1993; Hastbacka et al., 1994). Linkage disequilibrium in fact was a strategy available in BS (Ellis et al., 1994a), and it permitted a minimum regional assignment of *BLM* to the same 250 kb interval reported here (Ellis et al., submitted). This approach could have allowed us to clone *BLM*. Instead, we first carried out SCP mapping.

In the SCP mapping strategy, we took advantage of recombinant cell lines established from BS somatic cells in which crossing over within *BLM* had taken place, resulting in the correction of the mutant phenotype in their progenies (Ellis et al., 1995). After a segregational event, all polymorphic loci distal to *BLM* were reduced to homozygosity in half of the cases of intragenic recombination. This mapping method was preferred to linkage disequilibrium mapping because the crossovers that permitted localization of *BLM* had occurred within the gene itself and fewer genotypes were required for the analysis. By genotyping polymorphic loci that flank *BLM* in high SCE and low SCE samples from only five persons with BS and their parents, we delimited the position of *BLM* to the short interval bounded by the marker loci *D15S1108* and *D15S127* (Figure 1). With *BLM* assigned to such a short interval, the cloning of *BLM* became straightforward. The first candidate gene isolated from the interval proved to be *BLM*.

Loss-of-Function Mutations at *BLM*

The candidate gene for *BLM* isolated from the interval identified by SCP mapping encodes a 1417 amino acid peptide, homologous to RecQ helicases. Mutational analysis of the first 13 unrelated persons with BS examined permitted the identification of seven unique mutations in ten of them (Table 1). The fact that four of the seven mutations characterized so far result in premature termination of translation indicates that the cause of most BS is the loss of enzymatic activity of the *BLM* gene product. Identification of loss-of-function mutations in *BLM* (Table 1) is consistent with the autosomal recessive transmission of BS, and the homology of BLM and RecQ suggests that BLM has enzymatic activity. Thus, we predict that most BS mutations result in loss of function of BLM.

This loss of enzymatic activity is not lethal in cells because three of the chain-terminating mutations were detected in a homozygous state. The nonlethality could result from the existence of some residual enzymatic activity in the truncated peptides; however, this seems unlikely because one of the homozygous chain-terminating mutations results in chain termination after only 185 amino acids. Alternatively, the function of BLM may not be essential for cell survival. Other factors in the cell may be able to substitute for BLM, albeit inefficiently.

In the four persons with Jewish ancestry, a 6 bp deletion and 7 bp insertion at nucleotide 2281 were identified, and each of the four persons was homozygous for the mutation. Homozygosity was predictable because linkage disequilibrium had been detected in Ashkenazi Jews with BS between *BLM*, *D15S127*, and *FES* (Ellis et al., 1994a). Thus, a person who carried this 6 bp deletion and 7 bp insertion was a founder of the Ashkenazi Jewish population, and nearly all Ashkenazi Jews with BS inherit the mutation identical by descent from this common ancestor. Identification of the mutation now permits the screening of carriers in the Ashkenazim by a PCR test.

BS is an autosomal recessive trait with high penetrance and expressivity. The observation of loss-of-function mutations in *BLM* helps to explain these genetic characteristics. The short stature, characteristic facies, facial sun sensitivity, hyper- and hypopigmented patches on the skin, immunodeficiency, male infertility, female subfertility, premature menopause, and the predispositions to late onset diabetes and to neoplasia exist in virtually all groups of persons with the syndrome. The BS phenotype is similar in Ashkenazi Jews, the Dutch, the Flemish, Germans, Italians, Greeks, Turks, and Japanese; i.e., wherever it has been diagnosed. In addition, the elevated chromatid exchange and the hypermutability are constant cellular manifestations. No more variability in the expressivity of the mutations has been detected in persons with BS who inherit an identical mutation by descent from a common ancestor, as happens in Ashkenazi Jews with BS and in non-Ashkenazi Jewish persons with BS whose parents are cousins, than has been detected in persons who are compound heterozygotes (German et al., 1996). Nevertheless, with *BLM* cloned, it is possible to identify the mutations in any person with BS, and more subtle genotype–phenotype correlations now can be carried out.

BLM as a Putative DNA Helicase

The *BLM* gene product contains amino acid motifs that are homologous to motifs in the RecQ helicases (Figure 4), a subfamily of DExH box–containing DNA and RNA helicases. *recQ* is an E. coli gene that is a member of the RecF recombination pathway (Nakayama et al., 1984), a pathway of genes in which mutations abolish the conjugational recombination proficiency and ultraviolet (UV) resistance of a mutant strain lacking both RecBCD (part of exonuclease V) and SbcB (exonuclease I) activities (Horii and Clark, 1973). RecQ has DNA-dependent ATPase and DNA helicase activities and can translocate on single-

stranded DNA in a 3′–5′ direction (Umezu et al., 1990). Besides *BLM*, only two other *recQ*-like genes are known. First, *SGS1* is a yeast gene in which mutations suppress the slow growth of cells carrying mutations in the *TOP3* topoisomerase gene (Gangloff et al., 1994). It also was isolated in a yeast two-hybrid screen through its interactions with both the yeast Top2p and Top3p topoisomerases (Gangloff et al., 1994; Watt et al., 1995). Second, *RECQL* is a human gene isolated from HeLa cells, the product of which possesses DNA-dependent ATPase, DNA helicase, and 3′–5′ single-stranded DNA translocation activities (Puranam and Blackshear, 1994; Seki et al., 1994). The homology of BLM to RecQ and RECQL suggests that BLM also has DNA-dependent ATPase and DNA helicase activities, and studies to investigate this have been initiated.

In addition to helicase domains, BLM contains amino-terminal and carboxy-terminal regions that are composed predominantly of charged and polar amino acid residues. The presence of nonhelicase regions in BLM raises the possibility of additional enzymatic activities. The nonhelicase regions could operate to provide functional specificity to BLM, e.g., by promoting interactions with other proteins, or to provide substrates for phosphorylation that might regulate BLM activity in the cell cycle.

A Function for BLM in DNA Replication

Some genes in the DExH family have been implicated in DNA repair, and mutations in three of them, the *XPB*, *XPD*, and *ERCC6* genes, have been identified in the human disease phenotypes xeroderma pigmentosum and Cockayne's syndrome (Weber et al., 1990; Frejter et al., 1992; Troelstra et al., 1992; Sung et al., 1993; Ma et al., 1994). A universal function for the RecQ helicases, however, is not established. No abnormality in humans has been attributed to defects in *RECQL*. Even the cellular function of RecQ in bacteria is unclear, although it most likely participates in an aspect of postreplication recombinational repair (Luisi-DeLuca et al., 1989; Kusano et al., 1994; Tseng et al., 1994). The phenotype of yeast *SGS1* mutants includes slow growth, poor sporulation, chromosome nondisjunction at mitosis, missegregation in meiosis (Watt et al., 1995), and an elevated recombination frequency (Gangloff et al., 1994). Sgs1p is known to interact with topoisomerases II and Top3p and therefore may function in chromosome separation, a process in which intertwined DNA strands are resolved when replication forks converge. The predicted sizes of BLM (1417 residues) and Sgs1p (1447 residues) are similar, the two peptides have similar base compositions outside the helicase domains, and mutations in the genes encoding them result in genomic instability. In addition, an interaction between BLM and topoisomerase II in human cells was suggested by Heartlein et al. (1987), who observed that topoisomerase II activity is decreased in BrdU-treated BS cells. Although these interesting similarities are inconclusive, a possible functional homology between BLM and Sgs1p warrants further investigation.

In general, BLM has been implicated in the complex processes of DNA replication. Mutations in *BLM* have impressively pleiotropic cytogenetic and biochemical consequences. The chromosome breaks, gaps, and translocations and the high frequency of intra- and interchromosomal strand exchanges all point to a disturbance of DNA replication. In BS cells, the rate of nascent DNA chain elongation is retarded (Hand and German, 1975; Giannelli et al., 1977), and the distribution of DNA replicational intermediates is abnormal (Lonn et al., 1990). Some, though not all, cultured BS cells exhibit increased sensitivity to DNA-damaging agents, e.g., UV radiation, mitomycin C, N-nitroso-N-ethylurea, and ethyl methanesulfonate (Krepinsky et al., 1979, 1980; Ishizaki et al., 1981; Heddle et al., 1983; Kurihara et al., 1987). Disturbances in several enzymes that participate in DNA replication, DNA repair, or both have been identified in some, though, again, not all BS cell lines, including DNA ligase I (Chan et al., 1987; Willis and Lindahl, 1987), topoisomerase II in BrdU-treated BS cells (Heartlein et al., 1987), thymidylate synthetase (Shiraishi et al., 1989), uracil-DNA glycosylase (Seal et al., 1988), N-methylpurine DNA glycosylase (Dehazya and Sirover, 1986), O^6-methylguanine methyltransferase (Kim et al., 1986), and superoxide dismutase (Nicotera et al., 1989). These investigations show that certain enzymes concerned with DNA replication, repair, or both appear to be dysregulated in BS and that cultured BS cells make variously abnormal responses to DNA-damaging agents.

The evidence that BS cells have a defect in DNA repair, however, is slight (Friedberg et al., 1979; German and Schonberg, 1980). BS cells in general are not hypersensitive to UV or X-ray irradiation by standard assays, and no defect in a specific DNA repair enzyme or pathway has been reported. Although the explanation for the pleiotropic effects of BS mutations still is unknown, the predicted function of BLM as a DNA helicase suggests that the BS cell encounters greater difficulties than normal in the resolution of specific DNA structures generated during DNA replication. BLM presumably is one member of an assembly of gene products that acts in a pathway to resolve these structures. The excessive rates of chromatid exchange (homologous chromatid interchange configurations at metaphase and the SCE) might be microscopically visible manifestations of repair processes that are activated by the inability of the mutant cell to resolve the structures properly. Identification of the substrates on which BLM operates represents one of the important areas for future investigation.

Conclusions

With the cloning of the BS gene *BLM* and the inference that its gene product is a putative DNA helicase, insight has been gained into the molecular basis of the genomic instability that is the most impressive feature of BS cells. The absence of the *BLM* gene product most likely destabilizes other enzymes that participate in DNA replication and repair, perhaps through direct interactions or through more general responses to DNA damage. Elucidation of the enzymatic activities of BLM, the factors with which it interacts, and the substrates on which it operates now are required to understand the role of BLM in the maintenance of genomic stability.

Experimental Procedures

Subjects and Samples

The persons with BS in whom low SCE lymphocytes have arisen have been described previously (German et al., 1996). Epstein–Barr virus–transformed LCLs were developed from these and other persons with BS by standard culture methods using material obtained through the Bloom's Syndrome Registry (German and Passarge, 1989). The recombinant low SCE LCLs in which reduction to homozygosity had been detected and the cells used to determine the constitutional genotypes of the five persons from whom these recombinant low SCE LCLs were developed also have been described previously (Ellis et al., 1995). The polymorphic loci typed included some previously reported (Beckmann et al., 1993; Gyapay et al., 1994) and others that were identified during the physical mapping of the *BLM* region of chromosome 15 (Straughen et al., submitted). The methods of preparation of DNA samples, oligonucleotide primers, and conditions for PCR amplification of microsatellite polymorphisms on chromosome 15 have been described previously (German et al., 1994; Ellis et al., 1994a; Straughen et al., submitted).

Direct cDNA Selection

Direct cDNA selection was carried out as described by Parimoo et al. (1991). In brief, DNAs (15 ng) from commercial λ cDNA libraries prepared from cultured foreskin fibroblasts (Clontech) and Jurkat cells (Stratagene) were amplified by PCR (94°C for 1 min, 55°C for 1 min, 72°C for 2 min and 10 s for 32 cycles) using primer set A (GGTGGCGACGACTCCTGGA and ACCAGACCAACTGGTAATG) for the fibroblast cDNA library and the universal forward and reverse M13 sequencing primers for the Jurkat cDNA library under standard conditions with Taq polymerase (Boehringer Mannheim). EcoRI-digested cosmid (c905) or P1 (P1958) DNAs (100 ng) bound to Hybond-N membrane in 10× SSC were denatured in 0.5 M NaOH, 1.5 M NaCl; neutralized in 0.5 M Tris–HCl (pH 7.2), 1.5 M NaCl; and fixed by UV cross-linking. Hybridization of the PCR-amplified cDNAs to repetitive sequences on the cosmid and P1 clones was blocked by prehybridizing the membranes with Cot1 DNA (25 ng/μl; GIBCO BRL), poly(dI)·poly(dC) (20 ng/μl; Pharmacia), vector DNA (pWE15 or pAD10SacBII at 25 ng/μl) in 5× SSPE, 5× Denhardt's solution, and 0.5% SDS at 65°C overnight. Hybridization of the PCR-amplified cDNAs (25 ng/μl) was at 65°C for 2 days in the same solution without poly(dI)·poly(dC). The membranes were washed, and without elution the bound cDNAs were amplified by PCR with primer set A, followed by nested PCR with primer set B (ATGGTAGCGACCGGCGCTCA and CCGTCAGTATCGGCGGAATT) for the fibroblast library and the T3 and T7 sequencing primers for the Jurkat library. A sample of the PCR product after each amplification was analyzed by agarose gel electrophoresis, and another was cloned into Bluescript. Independent clones were picked at random, plasmid DNAs were prepared, and insert sizes were determined by restriction enzyme digestion and agarose gel electrophoresis. Inserts from selected clones were purified and used as hybridization probes against all of the other clones as well as against selected genomic DNAs to determine the chromosomal origin of the sequences (see below). The enrichment procedure was repeated and the selected cDNA clones analyzed again. The fibroblast cDNA clone 905-28 was obtained after two rounds of selection (250,000-fold enriched) and was sequenced by the dideoxy chain termination technique (Sanger et al., 1977; Tabor and Richardson, 1987).

The genomic origin of clones isolated by direct selection was verified by hybridization of inserts to Southern blots of DNAs from the following: clones in the contig; human cells; and two human × hamster somatic cell hybrids, one of which contains an intact chromosome 15 as the only human chromosome present (GS89K-1; Warburton et al., 1990) and one in which the only chromosome 15 material present had, through a translocation, lost all the sequences distal to band 15q25 (GM10664; obtained from the NIGMS Human Genetic Mutant Cell Repository at the Coriell Institute for Medical Research).

cDNA Cloning, 5′ RACE, and cDNA Sequencing

The selected cDNA 905-28 was hybridized to 10^6 clones from a HeLa cDNA library (Stratagene) according to standard procedures (Sambrook et al., 1989). We isolated 28 λ clones and converted them to Bluescript plasmids by superinfection with ExAssist helper phage (Stratagene). DNA was prepared, and 15 independent size classes of clones were identified. The 5′ end of a clone from each class was sequenced with Bluescript SK sequencing primer. To extend the sequence, we synthesized two oligonucleotides from the beginning and the end of each of the 5′ sequences, and sequencing was performed on the largest cDNA clone obtained by hybridization (clone H1). This procedure provided sequences from both DNA strands for most of the H1 cDNA. Ambiguous segments were determined by sequencing with specific oligonucleotides.

Because the reading frame was open at the 5′ end of the H1 clone, additional upstream sequences were obtained by a PCR method. PCR was carried out on DNA prepared from the HeLa cDNA library using an oligonucleotide (*Y177*, TTGTGGTGTTGGGTAGAGGTT) 8 bp from the 5′ end of H1 and the T3 sequencing primer. The PCR products were cloned into pT7Blue (Novagen), 18 clones were isolated, and the eight largest inserts were sequenced. The three largest of these clones (5′-5, 5′-15, and 5′-17) extended the sequences 289 bp 5′ of the H1 cDNA. We refer to the complete cDNA sequences present in the HeLa library as H1-5′ (Figure 2). Database searches then were carried out according to the method of Altschul et al. (1990) using segments of the predicted amino acid sequence encoded in the H1-5′ sequence as queries against the collected amino acid sequence databases that are accessible through the National Library of Medicine.

A full-length clone referred to as B3 was constructed by performing PCR of HeLa library DNA using an oligonucleotide (*Y180*, GAGGCNCGGCGGCCGTGGTTGC) from the 5′ end of the H1-5′ sequence and an internal oligonucleotide (*BC13*, CCTCAGTCAAATCTATTTGCTC) that permitted amplification of a 739 bp product. EagI and SmaI sites (Figure 2) were used to clone the product into NotI–SmaI-digested H1 DNA.

The 461 bp EagI–SmaI fragment of B3 was isolated and used to probe 8 × 10^6 clones of a pREP4-cloned unidirectional cDNA library from DEB-treated lymphoblastoid cells (Strathdee et al., 1992). We identified 12 cDNA clones, and the 5′ end of 11 were sequenced. Of these, eight are apparently full-length cDNAs (Figure 3). By restriction enzyme analysis, one of the 12 clones was shown to contain a deletion 3′ of nucleotide 2897 and the insertion of about 250 bp there.

5′ RACE was performed to characterize the 5′ sequences of the candidate gene using a Clontech Marathon cDNA amplification kit according to the specifications of the manufacturer. In brief, first-strand synthesis was carried out with MMLV reverse transcriptase using poly(T)-primed RNAs prepared from cultured fibroblast, lymphoblastoid, and HeLa cells and poly(A)$^+$ RNA from placenta (provided in the kit). Then, second-strand synthesis was performed with RNase H, E. coli polymerase I, and E. coli DNA ligase. The DNA ends were made blunt with T7 DNA polymerase, and adapters with overhanging ends were ligated to the cDNA. Nested PCRs then were carried out using 5′ oligonucleotides from the adaptor (AP1 and AP2) and internal 3′ oligonucleotides from the H1-5′ sequence (*BC5*, GCCATCACCGGAACAGAAGGAA; *BC11*, TCTTCTGGAGAAGGTGGAACAA). Bands derived from the H1-5′ sequences were identified in all four of the cDNA samples. PCR products from the 5′ RACE–amplified fibroblast cDNA were cloned into Bluescript, and the 5′ ends of 12 clones were sequenced (Figure 3).

Northern Blot Analysis

RNAs were prepared from cultured cells using TRIzol reagent (GIBCO BRL) according to the instructions of the manufacturer. Total RNAs (30 μg) were size separated by electrophoresis through 6.3% formaldehyde, 1.2% agarose gels in 0.02 M MOPS, 0.05 M sodium acetate (pH 7.0), and 0.001 M EDTA. The RNAs were transferred to Hybond-N (Amersham) in 20× SSPE and fixed to the membranes by UV cross-linking. Hybridizations were performed as described previously (Ellis et al., 1994b).

SSCP Analysis

After first-strand synthesis, PCR was carried out with 200 ng of cDNA, 5.2 pmol of each oligonucleotide primer (Table 2), 3% DMSO, 0.2 mM dNTPs (Pharmacia), 1× reaction buffer from Boehringer Mannheim, 0.25 U of Taq polymerase (Boehringer Mannheim), and 1.0 μCi of [α-^{32}P]dCTP in a total volume of 10 μl. Each reaction was overlaid with mineral oil and initially denatured for 5 min at 94°C followed by 35 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min. The

Table 2. Pairs of Primer Sequences Used for SSCP Analysis of *BLM*

Name	Forward Sequence[a]	Reverse Sequence[a]	Product Length (bp)
C1-B	GGATCCTGGTTCCGTCCGC	GAGGTTCACTGAAGGAAAAGTC	269
C1-A	CAACTAGAACGTCACTCAGCC	GAAGTCCTTGACCCTTTGCTG	233
C1-1	GACTTTTCCTTCAGTGAACCTC	GGGATTTCTTTACAGTTGGTGTG	186
C1-2	CCAGATTTCTTGCAGACTCCG	CTCTTACAAAGTGACTTTGGGG	213
C1-3	CTTTAAGTACCATCAATGATTGGG	CCTCAGTCAAATCTATTTGCTCG	227
C1-4	GAGTAAGCACTGCTCAGAAATC	GCTTAACCATTCTGAGTCATCC	160
C1-5	CGAGCAAATAGATTTGACTGAGG	CAATACATGGAACTTTCTCAGTTG	223
C1-6	GAAGATGCTCAGGAAAGTGAC	CGTACTAAGGCATTTTGAAGAGG	215
C1-7	CAACTGAGAAAGTTCCATGTATTG	CACAGTCTGTGCTGGTTTCTG	239
C1-9	CTATTCCTGATGATAAACTGAAAC	CCTTCATAGAATTCCCTGTAGG	200
C1-10	GTGGAGATACAGGCCTGATTC	GTGTTTCAGCCCAGTTGCTAC	244
C1-11	CAGGATTCTCTGCCACCAGG	GCAGTATGTTTATTCTGATCTTTC	183
C1-12	CAGGAAATGTTCTCACAAGCAC	CCTTGATGGGTTGATAGGCAG	203
C1-13	CAGCCAGCAAATCTTCCACAG	CGCTCATGTTTCAGATTTCTGG	204
C1-14	GAATTATACTGACAAGTCAGCAC	GATCTACGATAAGTGATCTCAAG	295
C1-15	CTCCTGGGGTCACTGTTGTC	GAGTCTGTTACTTGCACAGATC	211
C1-16	CAATCATAAAACTTCTATATGTCAC	GCCATCACCGGAACAGAAGG	207
C1-17	GTGGGGACATGATTTTCGTCAAG	GATTATGTCTGTTAAAGCTCATG	175
C1-18	GACATCCTGACTCAGCTGAAG	CGTGTCAGCCATGGTGTCAC	203
C1-19	GCACCACCCATATGATTCAGG	CAGATAACCTGACAGCCATCC	179
C1-20	GATGAAGTGCAGCAGAAGTGG	CAGTCTGGTCACATCATGATAG	221
C1-21	GCAGAGCTGGAAGAGATGGG	GCTGTATTCTCCTGCATTCCG	188
C1-22	GTATAGCATGGTACATTACTGTG	CCTTGTGATGAACTATGTTCTTG	228
C1-23	GACTGACGATGTGAAAAGTATTG	CCAAAATCTTGTCAAGTATCAGC	235
C1-24	CCAGTCAGGTATATTTGGAAAAG	GGAATTTTCTGTTTCCATAAAGTC	206
C1-25	CGATCGCTTATGTGATGCTCG	CAAGCTTCTTGAGAGTGACGG	248
C1-26	GAACTTACAGAAGTCTGCAAATC	GATGTCCATTCAGAGTATTTCTG	208
C1-27	GGTGTTACTGAAGACAAACTGG	GGGTATTTCCTCGTCAAGCTC	168
C1-28	GGATAAGCCTGTCCAGCAGC	CCTAGATATCTTTCTACATGTGG	214
C1-29	GCTTCCAGTGGTTCCAAGGC	GTTATGAGAATGCATATGAAGGC	204
C1-30	CTCAAGCGACATCAGGAGCC	CAAGAATAACAGCTTTATAGTCAC	178

[a] 5′–3′.

last cycle was extended at 72°C for 5 min. PCR products were diluted in 25 μl of 0.1% SDS, 10 mM EDTA and 25 μl of 95% formamide, 20 mM EDTA, 0.5% bromophenol blue, and 0.5% xylene cyanol. Two conditions for electrophoresis were carried out for each set of reactions. In one, electrophoresis of a 90 mM Tris borate, 2 mM EDTA (pH 7.5) (GIBCO BRL), 35% MDE (AT Biochem) 10% glycerol gel was performed at room temperature, cooled by fans; in the other, electrophoresis of a 90 mM Tris borate, 2 mM EDTA (pH 7.5) (GIBCO BRL), 25% MDE (AT Biochem) gel was performed at 4°C. Electrophoresis was carried out for both conditions at 40 W constant power in 0.6× TBE running buffer. After electrophoresis, gels were transferred to 3MM paper and dried on a vacuum slab dryer. Autoradiography overnight with Kodak XAR5 film without intensifying screens was sufficient to detect bands.

DNA Sequencing of SSCP Conformers

Isolation of DNA from SSCP conformers was performed as described previously by Groden et al. (1991, 1993). Each sample was analyzed by agarose gel electrophoresis to confirm the correct size. The remainder of each sample was purified using Centricon 100 columns (Amicon) and sequenced using the dsDNA Cycle Sequencing System (GIBCO BRL) with the forward primer originally designed for SSCP analysis. Sequencing reactions were analyzed by electrophoresis through 5% denaturing polyacrylamide gels. Gels were dried and exposed to Hyperfilm-MP (Amersham) without intensifying screens.

Acknowledgments

We are grateful to Manuel Buchwald, University of Toronto, for providing the lymphoblastoid cDNA library in pREP4; Dorothy Warburton, Columbia University, for providing the GS89K-1 cell line; and Larry Gelbert, Bristol Myers Squibb, for help in searching the expressed sequence database. We thank Norma Neff and James Ray for critical reading of the manuscript and Mary Claire King for valuable discussion and encouragement. This work was supported by the Lucille P. Markey Charitable Trust (J. Groden), the Center for Environmental Genetics at the University of Cincinnati grant ES-06096 (J. Groden), and the National Foundation for Jewish Genetic Disease (J. German), the New York Blood Center, and National Institutes of Health grants HD-04134, CA-50897, and GM-47890.

Received October 12, 1995; revised October 26, 1995.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Beckmann, J.S., Tomfohrde, J., Barnes, R., Williams, M., Broux, O., Richard, I., Weissenbach, J., and Bowcock, A.M. (1993). A linkage map of human chromosome 15 with an average resolution of 2 cM and containing 55 polymorphic microsatellites. Hum. Mol. Genet. *2*, 2019–2030.

Chan, J.Y.H., Becker, F.F., and German, J. (1987). Altered DNA ligase I activity in Bloom's syndrome cells. Nature *325*, 357–359.

Collins, F.S. (1995). Positional cloning moves from perditional to traditional. Nature Genet. *9*, 347–350.

Dehazya, P., and Sirover, M.A. (1986). Regulation of hypoxanthine DNA glycosylase in normal human and Bloom's syndrome fibroblasts. Cancer Res. *46*, 3756–3761.

Ellis, N.A., Roe, A.M., Kozloski, J., Proytcheva, M., Falk, C., and German, J. (1994a). Linkage disequilibrium between the *FES*, *D15S127*, and *BLM* loci in Ashkenazi Jews with Bloom syndrome. Am. J. Hum. Genet. *55*, 453–460.

Ellis, N.A., Ye, T.-Z., Patton, S., German, J., Goodfellow, P.N., and Weller, P. (1994b). Cloning of *PBDX*, an *MIC2*-related gene that spans the pseudoautosomal boundary on chromosome Xp. Nature Genet. *6*, 394–400.

Ellis, N.A., Lennon, D.J., Proytcheva, M., Alhadeff, B., Henderson, E.E., and German, J. (1995). Somatic intragenic recombination within the mutated locus *BLM* can correct the high SCE phenotype of Bloom syndrome cells. Am. J. Hum. Genet. *57*, 1019–1027.

Frejter, W.L., McDaniel, L.D., Johns, D., Friedberg, E.C., and Schultz, R.A. (1992). Correction of xeroderma pigmentosum complementation group D mutant cell phenotypes by chromosome and gene transfer: involvement of the human *ERCC2* DNA repair gene. Proc. Natl. Acad. Sci. USA *89*, 261–265.

Friedberg, E.C., Ehmann, U.K., and Williams, J.I. (1979). Human diseases associated with DNA repair. Adv. Rad. Biol. *8*, 85–174.

Gangloff, S., McDonald, J.P., Bendixen, C., Arthur, L., and Rothstein, R. (1994). The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. Mol. Cell. Biol. *14*, 8391–8398.

German, J. (1993). Bloom syndrome: a mendelian prototype of somatic mutational disease. Medicine *72*, 393–406.

German, J., and Passarge, E. (1989). Bloom's syndrome. XII. Report from the Registry for 1987. Clin. Genet. *35*, 57–69.

German, J., and Schonberg, S. (1980). Bloom's syndrome. IX. Review of cytological and biochemical aspects. In Genetic and Environmental Factors in Experimental and Human Cancer, H.V. Gelboin, B. MacMahon, T. Matsushima, T.Sugimura, S.Takayama, and H. Takebe, eds. (Tokyo: Japan Scientific Societies Press), pp. 175–186.

German, J., Roe, A.M., Leppert, M., and Ellis, N.A. (1994). Bloom syndrome: an analysis of consanguineous families assigns the locus mutated to chromosome band 15q26.1. Proc. Acad. Natl. Sci. USA *91*, 6669–6673.

German, J., Ellis, N.A., and Proytcheva, M. (1996). Bloom's syndrome. XIX. Cytogenetic and population evidence for genetic heterogeneity. Clin. Genet., in press.

Giannelli, F., Benson, P.F., Pawsey, S.A., and Polani, P.E. (1977). Ultraviolet light sensitivity and delayed DNA-chain maturation in Bloom's syndrome fibroblasts. Nature *265*, 466–469.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P., and Blinov, V.M. (1989). Two related superfamilies of putative helicases involved in replication, recombination, repair, and expression of DNA and RNA genomes. Nucl. Acids Res. *17*, 4713–4730.

Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., Robertson, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, O., Hughes, J.P., Warrington, J., McPherson, J., Wasmuth, J., LePaslier, D., Abderrahim, H., Cohen, D., Leppert, M., and White, R. (1991). Identification and characterization of the familial adenomatous coli gene. Cell *66*, 589–600.

Groden, J., Gelbert, L., Thliveris, A., Nelson, L., Robertson, M., Joslyn, G., Samowitz, W., Spirio, L., Carlson, M., Burt, R., Leppert, M., and White, R. (1993). Mutational analysis of patients with adenomatous polyposis: identical inactivation mutations in unrelated inividuals. Am. J. Hum. Genet. *52*, 263–272.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., and Weissenbach, J. (1994). The 1993–1994 Genethon human genetic linkage map. Nature Genet. *7*, 246–339.

Hand, R., and German, J. (1975). A retarded rate of DNA chain growth in Bloom's syndrome. Proc. Natl. Acad. Sci. USA *72*, 758–762.

Hastbacka, J., de la Chapelle, A., Mahtani, M.M., Clines, G., Reeve-Daly, M.P., Daly, M., Hamilton, B.A., Kusumi, K., Trivedi, B., Weaver, A., Coloma, A., Lovett, M., Buckler, A., Kaitila, I., and Lander, E.S. (1994). The diastrophic dysplasia gene encodes a novel sulfate transporter: positional cloning by fine-structure linkage disequilibrium mapping. Cell *78*, 1073–1087.

Heartlein, M.W., Tsuji, H., and Latt, S.A. (1987). 5-Bromodeoxyuridine-dependent increase in sister chromatid exchange formation in Bloom's syndrome is associated with reduction in topoisomerase II activity. Exp. Cell Res. *169*, 245–254.

Heddle, J.A., Krepinsky, A.B., and Marshall, R.R. (1983). Cellular sensitivity to mutagens and carcinogens in the chromosome-breakage and other cancer-prone syndromes. In Chromosome Mutation and Neoplasia, J. German, ed. (New York: Alan R. Liss), pp. 203–234.

Horii, Z., and Clark, A.J. (1973). Genetic analysis of the RecF pathway of genetic recombination in *Escherichia coli* K-12. J. Mol. Biol. *80*, 327–344.

Ishizaki, K., Yagi, T., Inoue, M., Nikaido, O., and Takebe, H. (1981). DNA repair in Bloom's syndrome fibroblasts after UV irradiation or treatment with mitomycin C. Mutat. Res. *80*, 213–219.

Kerem, B.-S., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Charkravarti, A., Buchwald, M., and Tsui, L.-C. (1989). Identification of the cystic fibrosis gene: genetic analysis. Science *245*, 1073–1080.

Kim, S., Vollberg, T.M., Ro, J.Y., Kim, M., and Sirover, M.A. (1986). O^6-methylguanine methyltransferase increases before S phase in normal human cells but does not increase in hypermutable Bloom's syndrome cells. Mutat. Res. *173*, 141–145.

Kozak, M. (1989). The scanning model for translation: an update. J. Cell Biol. *108*, 229–241.

Krepinsky, A.B., Heddle, J.A., and German, J. (1979). Sensitivity of Bloom's syndrome lymphocytes to ethyl methanesulfonate. Hum. Genet. *50*, 151–156.

Krepinsky, A.B., Rainbow, A.J., and Heddle, J.A. (1980). Studies on the ultraviolet light sensitivity of Bloom's syndrome fibroblasts. Mutat. Res. *69*, 357–368.

Kurihara, T., Inoue, M., and Tatsumi, K. (1987). Hypersensitivity of Bloom's syndrome fibroblasts to N-ethyl-N-nitrosourea. Mutat. Res. *184*, 147–151.

Kusano, K., Sunohara, Y., Takahashi, N., Yoshikura, H., and Kobayashi, I. (1994). DNA double-strand break repair: genetic determinants of flanking crossing-over. Proc. Natl. Acad. Sci. USA *91*, 1173–1177.

Lehesjoki, A.E., Koskiniemi, M., Novro, R., Tirrito, S., Sistonen, P., Lander, E., and de la Chapelle, A. (1993). Localization of the *EPM1* gene for progressive myoclonus epilepsy on chromosome 21: linkage disequilibrium allows high resolution mapping. Hum. Mol. Genet. *2*, 1229–1234.

Lonn, U., Lonn, S., Nylen, U., Winblad, G., and German, J. (1990). An abnormal profile of DNA replication intermediates in Bloom's syndrome. Cancer Res. *50*, 3141–3145.

Luisi-DeLuca, C., Lovett, S.T., and Kolodner, R.D. (1989). Genetic and physical analysis of plasmid recombination in *recB recC sbcB* and *recB recC sbcA Escherichia coli* K12 mutants. Genetics *122*, 269–278.

Ma, L., Westbroek, A., Jochemsen, A.G., Weeda, G., Bosch, A., Bootsma, D., Hoeijmakers, J.H.J., and van der Ed, A.J. (1994). Mutational analysis of *ERCC3*, which is involved in DNA repair and transcription initiation: identification of domains essential for the DNA repair function. Mol. Cell. Biol. *14*, 4126–4134.

Mathew, S., Murty, V.V.V.S., German, J., and Chaganti, R.S.K. (1993). Confirmation of 15q26.1 as the site of the *FES* protooncogene by fluorescence *in situ* hybridization. Cytogenet. Cell Genet. *63*, 33–34.

McDaniel, L.D., and Schultz, R.A. (1992). Elevated sister chromatid exchange phenotype of Bloom syndrome cells is complemented by human chromosome 15. Proc. Natl. Acad. Sci. USA *89*, 7968–7972.

Nakayama, H., Nakayama, K., Nakayama, R., Irino, N., Nakayama, Y., and Hanawalt, P.C. (1984). Isolation and genetic characterization of a thymineless death-resistant mutant of *Escherichia coli* K12: identification of a new mutation (*recQ1*) that blocks the RecF pathway. Mol. Gen. Genet. *195*, 474–480.

Nicotera, T.M., Notaro, J., Notaro, S., Schumer, J., and Sandberg, A.A. (1989). Elevated superoxide dismutase in Bloom's syndrome: a genetic condition of oxidative stress. Cancer Res. *49*, 5239–5243.

Parimoo, S., Patanjali, S.R., Shukla, H., Chaplin, D.D., and Weissman, S.M. (1991). cDNA selection: efficient PCR approach for the selection of cDNAs encoded in large chromosomal fragments. Proc. Natl. Acad. Sci. USA *88*, 9623–9627.

Puranam, K.L., and Blackshear, P.J. (1994). Cloning and characterization of RECQL, a potential human homologue of the *Escherichia coli* DNA helicase RecQ. J. Biol. Chem. *47*, 29838–29845.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing

with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Seal, G., Brech, K., Karp, S.J., Cool, B.L., and Sirover, M.A. (1988). Immunological lesions in human uracil DNA glycosylase: association with Bloom's syndrome. Proc. Natl. Acad. Sci. USA *85*, 2339–2343.

Seki, M., Miyazawa, H., Tada, S., Yanagisawa, J., Yamaoka, T., Hoshino, S., Ozawa, K., Eki, T., Nogami, M., Okumura, K., Taguchi, H., Hanaoka, F., and Enomoto, T. (1994). Molecular cloning of cDNA encoding human DNA helicase Q1 which has homology to *Escherichia coli* RecQ helicase and localization of the gene at chomosome 12p12. Nucl. Acids Res. *22*, 4566–4573.

Shiraishi, Y., Taguchi, T., Ozawa, M., and Bamezai, R. (1989). Different mutations responsible for the elevated sister-chromatid exchange frequencies in Bloom syndrome and X-irradiated B-lymphoblastoid cell lines originating from acute leukemia. Mutat. Res. *211*, 273–278.

Sirugo, G., Keats, B., Fujita, R., Duclos, F., Purohit, K., Koenig, M., and Mandel, J.L. (1992). Friedreich ataxia in Louisiana Acadians: demonstration of a founder effect by analysis of microsatellite-generated extended haplotypes. Am. J. Hum. Genet. *50*, 559–566.

Strathdee, C.A., Gavish, H., Shannon, W.R., and Buchwald, M. (1992). Cloning of cDNAs for Fanconi's anemia by functional complementation. Nature *356*, 763–767.

Sung, P., Bailly, V., Weber, C., Thompson, L.H., Prakash, L., and Prakash, S. (1993). Human xeroderma pigmentosum group D gene encodes a DNA helicase. Nature *365*, 852–855.

Surdej, P., Riedl, A., and Jacobs-Lorena, M. (1994). Regulation of mRNA stability in development. Annu. Rev. Genet. *28*, 263–282.

Tabor, S., and Richardson, C.C. (1987). DNA sequence analysis with a modified T7 DNA polymerase. Proc. Natl. Acad. Sci USA *84*, 4767–4771.

Troelstra, C., van Gool, A., de Wit, J., Vermeulen, W., Bootsma, D., and Hoeijmakers, J.H.J. (1992). *ERCC6*, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. Cell *71*, 939–953.

Tseng, Y.-C., Hung, J.-L., and Wang, T.-C. (1994). Involvement of RecF pathway recombination genes in postreplication repair in UV-irradiated *Escherichia coli* cells. Mutat. Res. *315*, 1–9.

Umezu, K., Nakayama, K., and Nakayama, H. (1990). *Escherichia coli* RecQ protein is a DNA helicase. Proc. Natl. Acad. Sci. USA *87*, 5363–5367.

Warburton, D., Gersen, S., Yu, M.-T., Jackson, C., Handelin, B., and Housman, D. (1990). Monochromosomal rodent-human hybrids from microcell fusion of human lymphoblastoid cells containing an inserted dominant selectable marker. Genomics *6*, 358–366.

Watt, P.M., Louis, E.J., Borts, R.H., and Hickson, I.D. (1995). Sgs1: a eukaryotic homolog of E. coli RecQ that interacts with topoisomerase II in vivo and is required for faithful chromosome segregation. Cell *81*, 253–260.

Weber, C.A., Salazar, E.P., Stewart, S.A., and Thompson, L.H. (1990). *ERCC2*: cDNA cloning and molecular characterization of a human nucleotide excision repair gene with high homology to yeast *RAD3*. EMBO J. *9*, 1437–1447.

Willis, A.E., and Lindahl, T. (1987). DNA ligase I deficiency in Bloom's syndrome. Nature *325*, 355–357.

GenBank Accession Number

The accession number for the sequence reported in this paper is U39817.