Genomics **60,** 20–30 (1999)
Article ID geno.1999.5906, available online at http://www.idealibrary.com on IDEAL®

Novel Human and Mouse Homologs of *Saccharomyces cerevisiae* DNA Polymerase η

John P. McDonald,* Vesna Rapić-Otrin,* Jonathan A. Epstein,† Bernard C. Broughton,‡ Xiangyuan Wang,§ Alan R. Lehmann,‡ Debra J. Wolgemuth,§ and Roger Woodgate*,[1]

*Section on DNA Replication, Repair and Mutagenesis and †Unit on Biologic Computation, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892-2725; ‡MRC Cell Mutation Unit, University of Sussex, Falmer, East Sussex, BN1 9RR United Kingdom; and §Department of Genetics and Development, Columbia University, College of Physicians and Surgeons, New York, New York 10032

Received April 21, 1999; accepted June 14, 1999

The *Saccharomyces cerevisiae RAD30* gene encodes a novel eukaryotic DNA polymerase, pol η that is able to replicate across *cis-syn* cyclobutane pyrimidine dimers both accurately and efficiently. Very recently, a human homolog of *RAD30* was identified, mutations in which result in the sunlight-sensitive, cancer-prone, Xeroderma pigmentosum variant group phenotype. We report here the cloning and localization of a second human homolog of *RAD30*. Interestingly, *RAD30B* is localized on chromosome 18q21.1 in a region that is often implicated in the etiology of many human cancers. The mouse homolog (*Rad30b*) is located on chromosome 18E2. The human *RAD30B* and mouse *Rad30b* mRNA transcripts, like many repair proteins, are highly expressed in the testis. *In situ* hybridization analysis indicates that expression of mouse *Rad30b* occurs predominantly in postmeiotic round spermatids. Database searches revealed genomic and EST sequences from other eukaryotes such as *Aspergillus nidulans, Schizosaccharomyces pombe, Brugia malayi, Caenorhabditis elegans, Trypanosoma cruzi, Arabidopsis thaliana,* and *Drosophila melanogaster* that also encode putative homologs of *RAD30,* thereby suggesting that Rad30-dependent translesion DNA synthesis is conserved within the eukaryotic kingdom.

INTRODUCTION

Cells are constantly challenged by both natural and synthetic agents that damage their genome. Although various mechanisms exist to ensure that the majority of DNA damage is recognized and the integrity of the DNA is faithfully restored, some damage often persists (for a general review see Friedberg *et al.*, 1995). Unfortunately, unrepaired DNA damage impedes basic cellular processes, such as transcription and replication, and, as a result, leads to cell death. To avoid these fatal consequences, cells invoke "damage tolerance" mechanisms. These pathways include direct bypass of the lesion by translesion DNA synthesis and avoidance mechanisms, such as postreplication daughter-strand gap repair and template-strand switching (Banerjee *et al.*, 1988; Rupp and Howard-Flanders, 1968; Rupp *et al.*, 1971; Higgins *et al.*, 1976; Koffel-Schwartz *et al.*, 1996). In all cases, the offending lesion persists in DNA, but the cell avoids the immediate consequences of blocked transcription/replication forks, providing additional opportunities for the lesion to be repaired before again being encountered by a new transcription/replication complex.

The best characterized of these tolerance mechanisms is translesion DNA synthesis, which in *Escherichia coli* requires the heterotrimeric UmuD′$_2$C complex (Woodgate and Levine, 1996; Smith and Walker, 1998). Compared to our understanding of processes like nucleotide excision repair or mismatch repair, the actual mechanism by which the Umu proteins facilitate translesion DNA synthesis is relatively poorly understood. The prevailing model, based upon genetic analyses, suggests that together with RecA protein, the Umu complex enables the cell's main replicative polymerase, pol III, to replicate across DNA lesions, with a concomitant decrease in replication fidelity (Woodgate and Levine, 1996; Smith and Walker, 1998). However, the recent reconstitution of translesion synthesis *in vitro* utilizing highly purified UmuD′$_2$C led to the striking discovery that the Umu complex possesses intrinsic DNA polymerase activity (Tang *et al.*, 1998, 1999), which results in the (mis)incorporation of nucleotides at sites of DNA damage and the extension of aberrant primer termini.

DNA sequence data from this article have been deposited with the GenBank Data Library under Accession Nos. AF140501 and AF151691.

[1] To whom correspondence should be addressed at NICHD, NIH, Building 6, Room 1A13, 9000 Rockville Pike, Bethesda, MD 20892-2725. Telephone: (301) 496-6175. Fax: (301) 594-1135. E-mail: woodgate@helix.nih.gov.

0888-7543/99 $30.00

The recent cloning of the *E. coli dinB* gene (Ohmori *et al.*, 1995), which was previously shown to play a role in untargeted mutagenesis of lambda phage (Brotcorne-Lannoye and Maenhaut-Michel, 1986), revealed that DinB shares extensive homology with UmuC, suggesting that both proteins might possess analogous biochemical functions. Indeed, further support of this notion comes from reports demonstrating that overexpression of DinB results in an increase in damage-induced mutagenesis of F-plasmids (Kim *et al.*, 1997) and that *dinB* also encodes a DNA polymerase (Wagner *et al.*, in press).

In *Saccharomyces cerevisiae,* mutagenic translesion synthesis requires DNA polymerase ζ, which is a complex of the Rev3 and Rev7 proteins and the Rev1 protein (Lawrence and Hinkle, 1996). Biochemical analysis of Pol ζ has revealed that, like UmuD′$_2$C, it too possesses the ability to replicate past thymine–thymine *cis-syn* cyclobutane dimers (Nelson *et al.*, 1996a). Furthermore, the Rev1 protein, which shares limited structural homology with UmuC and DinB (Larimer *et al.*, 1989), possesses a novel template-directed dCTP transferase activity (Nelson *et al.*, 1996b).

Finally, another homolog of the *E. coli umuC* and *dinB* genes, designated *RAD30,* has been identified and characterized in *S. cerevisiae* (McDonald *et al.*, 1997; Roush *et al.*, 1998). Similar to both *dinB* and *umuC,* the *RAD30* gene is damage inducible, and deletion of *S. cerevisiae RAD30* results in a mild UV-sensitive phenotype reminiscent of *S. cerevisiae rev* mutants (McDonald *et al.*, 1997; Roush *et al.*, 1998). Depending upon the reversion marker and the DNA damaging agent used, Rad30 was categorized as being involved in an error-free repair pathway (McDonald *et al.*, 1997) or an error-prone pathway (Roush *et al.*, 1998). This paradox can potentially be explained by the fact that *RAD30* encodes a novel eukaryotic DNA polymerase, called pol η (Johnson *et al.*, 1999a). Pol η is apparently capable of replicating past a *cis–syn* thymine dimer and preferentially inserts two dATPs opposite the thymine–thymine dimer (Johnson *et al.*, 1999a). While such findings are consistent with the idea that pol η functions primarily in an error-free bypass mechanism of UV-induced DNA lesions (McDonald *et al.*, 1997; Johnson *et al.*, 1999a), it is conceivable that at other DNA lesions, such as 6-4 photoproducts, pol η replication is, in fact, error-prone, which would also be consistent with the observations of Roush *et al.* (1998).

From the aforementioned phenotypes of mutant *S. cerevisiae RAD30* strains, and the very recently described DNA polymerase activity of the UmuC, DinB, Rad30, and human XPV/RAD30A proteins, it is clear that this family of proteins plays a pivotal role in the ability of cells to tolerate DNA damage. Database searches reveal that Rad30 is conserved in eukaryotes, and here, we report the cloning and characterization of novel *Homo sapiens* and *Mus musculus* homologs of *RAD30.*[2]

MATERIALS AND METHODS

Identification of human and mouse RAD30 homologs and screening of cDNA libraries. Utilizing the TBLASTN program (Altschul *et al.*, 1990, 1994), the expressed sequence tag database (dbEST) translated in all six frames was screened using both the *E. coli* DinB and the *S. cerevisiae* Rad30 proteins. Several human cDNA sequences that share significant homology with Rad30 were identified (unpublished observations). One such cDNA, D78887, was obtained from ATCC (American Type Culture Collection, Rockville, MD), and the insert (~2.4-kb DNA from the *Eco*RI to the *Xho*I sites) was sequenced (Lark Technologies, Houston, TX). Translation of the large continuous open reading frame (ORF) obtained from D78887 revealed that it lacks an initiation methionine (and is therefore only a partial cDNA clone), but that the remainder of the translated protein shares significant similarity to the *S. cervisiae* Rad30 protein. Similarly, a mouse cDNA, AA162008, that also encodes part of a putative mouse *Rad30* homolog was obtained from ATCC, and 1.8 kb from the *Sal*I to the *Not*I sites was sequenced (Lark Technologies).

To isolate a full-length human *RAD30B* cDNA clone, a human endothelium cDNA library (umbilical vein (HUVEC), one passage; Catalog No. 937223) was obtained from Stratagene (La Jolla, CA) and screened using 90-mer oligonucleotides as probes. These probes were designed to sequences at the 5′-end, the middle, and the 3′-end of the D78887 cDNA, the rationale being that any clone hybridizing to all three probes would contain, at the very least, cDNA inserts of the same size or larger than that found in D78887. Using such an approach, three new human cDNAs were identified. Based on restriction enzyme digestion analysis, two were only slightly larger than D78887, but the third clone contained significantly more upstream sequence than that found in D78887. Sequence of this clone indicated that it encodes a full-length human *RAD30B* gene and is henceforth identified by its GenBank accession number, AF140501.

The 5′-end of the mouse *Rad30b* cDNA was obtained by performing 5′ RACE PCR on "Marathon Ready" mouse testis cDNA and using the Advantage PCR amplification kit (Clontech, Palo Alto, CA) under conditions specified by the manufacturer. The primers used for this amplification were MRAD30N4 (5′TGACGGTGTGACGGGAGACTTGTCCTCTCC3′), which was designed based on sequences near the 5′-end of the mouse *Rad30b* EST, AA162008, and the AP1 primer (5′CCATCCTAATACGACTCACTATAGGGC3′) (Clontech), which is homologous to the adaptor sequence at the 5′-end of the Marathon cDNA. The approximately 1.0-kb amplification product was digested with *Xho*I (within the adaptor sequence) and *Blp*I (within the mouse *Rad30b* cDNA) and cloned into the AA162008 EST digested with *Sal*I and *Blp*I to construct pJM297, and a full-length mouse *Rad30b* cDNA was reconstructed. The full nucleotide sequence of this clone has been deposited with GenBank and is henceforth identified by its GenBank accession number, AF151691.

In situ analysis of mouse Rad30b. Adult CD-1 mice were purchased from Charles River (Wilmington, DE) and were used as the source of testis tissue. Whole testes were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) overnight at 4°C, processed, and embedded in paraffin by standard methods of histological sectioning (Chapman and Wolgemuth, 1994). The 1.8-kb AA162008 cDNA clone of mouse *Rad30b,* originally a *Sal*I to *Not*I insert in pCMV-SPORT, was subcloned into plasmid Bluescript SK II (Stratagene, La Jolla, CA). Seven-micrometer sections were cut and subjected to *in situ* hybridization analysis essentially as previously described (Rhee and Wolgemuth, 1995). Briefly, the sections were deparaffinized, rehydrated, and prehybridized at room temperature for 3 h in 50% formamide, 0.6 M NaCl, 120 mM Tris · HCl, pH 8.0, 8

[2] The designated gene symbol for the human *RAD30* homolog described herein is *RAD30B* and that of the mouse homolog is *Rad30b.*

mM EDTA, 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.1% bovine serum albumin, 0.5 mg/ml denatured salmon sperm DNA, and 0.6 mg/ml yeast total RNA. The hybridization mixture was identical to the prehybridization mixture with the addition of 10 mM dithiothretiol, 0.1% SDS, 100 mg/ml dextran sulfate, and 10^5 dpm/μl ^{35}S-labeled sense and antisense riboprobes generated from the insert in pBluescript SK II. After hybridization at 50°C for 18 h, the sections were washed in formamide wash solution, treated with RNase A solution, and subjected to a final wash at 65°C for 2 h in 0.1× SSC. The sections were dehydrated and covered with nuclear track emulsion (Kodak type NTB-2) and exposed for 10 days at 4°C. The sections were developed and counterstained with hematoxylin and eosin and viewed under either bright-field or epiluminescence optics, the latter yielding a greenish color for the light reflected by silver grains deposited by response to radioactivity.

DNA sequence analysis of the human RAD30B gene in various repair-deficient cell lines. RNA was extracted from cultured fibroblasts and reverse-transcribed into cDNA by random priming, using the First Strand cDNA synthesis kit (Amersham Pharmacia Biotech, Piscataway, NJ). The human *RAD30B* cDNA was then amplified by PCR in three sections (from nucleotides 10 to 915, 801 to 1790, and 1671 to 2377 of the sequence reported in AF140501). The PCR products were subsequently gel-purified and sequenced directly using the Thermosequenase radiolabeled terminator cycle sequencing kit (Amersham Pharmacia Biotech, Piscataway, NJ).

RESULTS

Phylogenetic Analysis of the UmuC/DinB/Rev1/Rad30 Superfamily

An iterative search of protein and nucleotide sequence databases (GenBank release 110.0), using the *E. coli* UmuC, *E. coli* DinB, *S. cerevisiae* Rev1, and *S. cerevisiae* Rad30 proteins for comparison, identified many new members of the UmuC-like, DinB-like, Rev1-like, and Rad30-like subfamilies (Fig. 1). In addition, we identified genomic and EST sequences from eukaryotic organisms such as *Aspergillus nidulans*, *Brugia malayi*, *Trypanosoma cruzi*, and *Drosophila melanogaster* (not shown) that encode putative homologs of Rad30, suggesting that it represents a *bona fide* branch of the UmuC/DinB/Rev1/Rad30 superfamily (Fig. 1). Our new phylogenetic tree contains 36 proteins and, like our previous tree (which consisted of 16 proteins) (Kulaeva *et al.*, 1996), can be broadly subdivided into the existing four subfamilies. The UmuC-like proteins are found in prokaryotes, with the majority identified in gram-negative Enterobacteriaceae and their plasmids, although there does now appear to be a subset of UmuC-like proteins from nonenteric gram-positive bacteria (Fig. 1). The most conserved and ubiquitous family consists of the DinB-like proteins, which are found in many prokaryotes, archaea, and eukaryotes, including humans (Fig. 1). In contrast, the Rev1-like and Rad30-like proteins are somewhat more diverged and are found only in eukaryotic cells (Fig. 1). Interestingly, many eukaryotic cells contain at least one member from each of the DinB/Rev1/Rad30 subfamilies, suggesting that there may be some division of biochemical functions.

Isolation of Full-Length Human and Mouse RAD30B cDNAs

Utilizing the TBLASTN program (Altschul *et al.*, 1990, 1994), we identified several cDNA sequences of human origin in dbEST that encode putative homologs of Rad30. Homology comparisons of the translated open reading frame of one clone, GenBank Accession No. D78887, confirmed that it does indeed encode a human homolog of *RAD30* (Fig. 1). However, the cDNA was obviously incomplete as the large ORF lacked an initiating methionine as well as amino acid residues at the amino-terminus that are highly conserved in other DinB/Rad30 homologs. By further screening of a human endothelium cDNA library, we were able to identify one clone (identified by its GenBank accession number, AF140501) that contained significantly more upstream sequence than that found in D78887.

DNA sequencing of AF140501 revealed an ATG codon that likely represents the initiating codon of the human *RAD30B* open reading frame. In support of this assertion, we noted that the putative initiating methionine is positioned upstream of the first region of protein conservation, the DMDCFFAAVE region, among the DinB-like and the Rad30-like proteins (see below), and at a distance similar to that found in the Rad30 proteins from *S. cerevisiae* and *Schizosaccharomyces pombe.* Furthermore, addition of the 5′ sequences derived from AF140501 to the D78887 cDNA would yield a transcript, including the poly(A) tract, that is at least 2.5 kb in size and is in good agreement with the size of the human *RAD30B* RNA detected by Northern blot analysis (data not shown). The G+C content of the 60- to 70-bp leader sequence is 73.5% (cf., the human *RAD30B* ORF, which is 44% G+C) and lacks any upstream ATG codons. This, together with the presumptive translational start site of human *RAD30B*, GCCatgGG, is consistent with features found in other upstream sequences (Kozak, 1987, 1997). Based upon this analysis, we conclude that we have identified the ATG start codon of human *RAD30B*.

Using the translated amino acid sequence from the human *RAD30B* cDNA from D78887 to search dbEST, we identified a mouse cDNA sequence (GenBank Accession No. AA162008) that encodes a mouse *Rad30b* homolog (Fig. 1). Sequence analysis revealed that the deduced mouse Rad30b protein encoded by AA162008 was highly homologous to the human RAD30B protein with ~68% identical residues over a stretch of 500 amino acid residues (human RAD30B residues 212–715). Based upon the alignment of mouse Rad30b to human RAD30B (Fig. 2) and other homologs (see below), it was clear, however, that the cDNA encoded by A162008 was incomplete. We obtained the 5′ end of the mouse *Rad30b* cDNA by performing 5′ RACE PCR, and an ~1.0-kb PCR product was subcloned into the AA162008 vector to construct a full-length mouse *Rad30b* cDNA. Subsequent sequence analysis of this PCR product revealed that it does indeed encode the

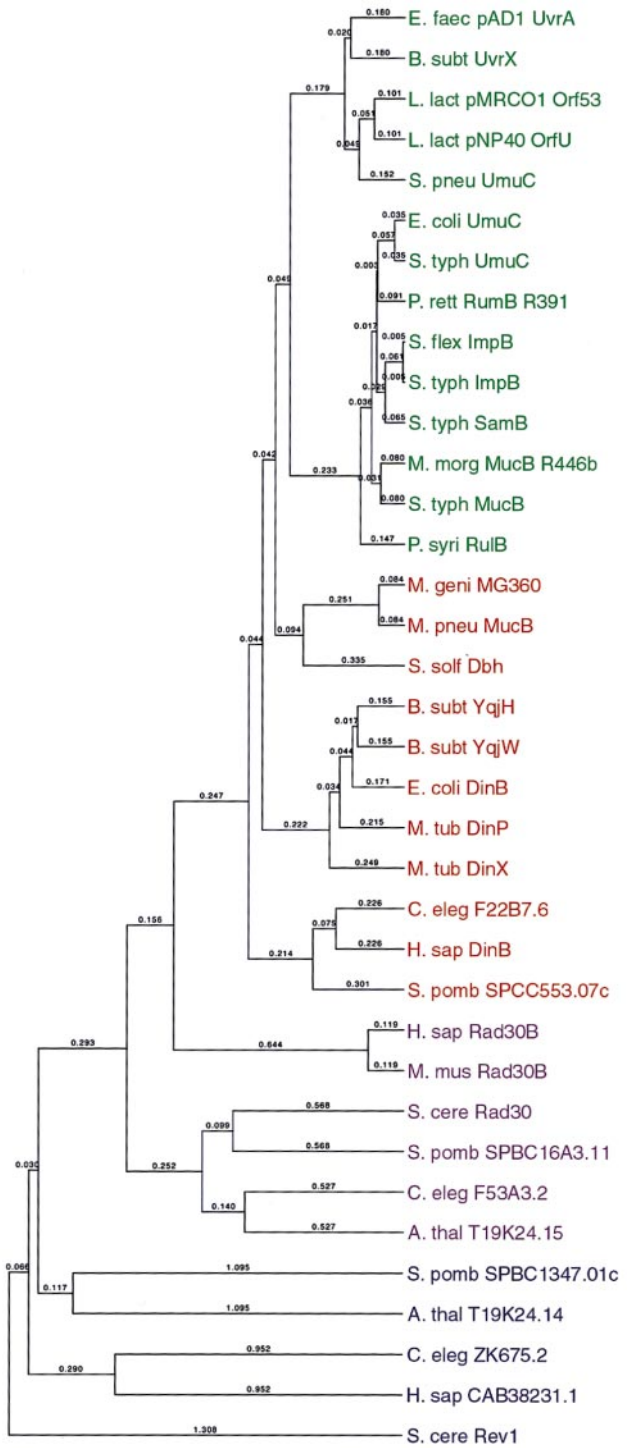

FIG. 1. Phylogenetic analysis of the UmuC/DinB/Rev1/Rad30 superfamily. GenBank (release 110.0+) was searched using the *E. coli* UmuC, *E. coli* DinB, *S. cerevisiae* Rev1, and *S. cerevisiae* Rad30 proteins as reference. Thirty-six proteins that are homologous to one or more of the aforementioned reference proteins were identified. A phylogenetic tree of these proteins was generated using the program Geneworks (version 2.51, Oxford Molecular, Campbell, CA) with the ClustalW alignment and the PAM250 table of amino acid replacements. The numbers on each branch indicate the relative distances between each branch. Based upon this alignment, we believe that the tree can be broadly divided into four subfamilies: UmuC-like (green); DinB-like (red); Rad30-like (purple); and Rev1-like (blue). Although not shown in this figure, the recently described Rad30A/XPV protein (Masutani *et al.*, 1999b) would be located in the Rad30 subfamily in a branch near the *Arabidopsis thaliana* Rad30 protein.

amino-terminus of mouse Rad30b. Thus, the mouse Rad30b protein is 717 amino acids long and shares ~74% identity with the human RAD30B protein.

Chromosomal Location of Human RAD30B and Mouse Rad30b

DNA sequence analysis of our full-length clone AF140501 revealed a 340-bp overlap with an STS (sequence taggedsite), designated WI-11064 (dbEST GenBank Accession No. R37837). The WI-11064 STS maps 415.97 cR from the top of the chromosome 18 linkage group (WICGR radiation hybrid map; Whitehead Institute for Biomedical Research, Cambridge, MA) and is within the D18S474–D18S64 interval that maps to band 18q21.1.

Using the mouse *Rad30b* cDNA, we were able to obtain a mouse genomic clone that we used in FISH analysis of mouse metaphase chromosomes. This resulted in specific labeling of a small chromosome that was believed to be mouse chromosome 18. A second experiment utilizing both the mouse *Rad30b* probe and a probe specific to the centromeric region of chromosome 18 demonstrated that it is located on chromosome 18E2. A portion of the 18E2 band of mouse chromosome 18 delineates a region of synteny between mouse chromosome 18 and the human chromosome 18q21 region and that apparently includes *RAD30B*.

Comparison of Human and Mouse Rad30B Proteins to Other Rad30-like Proteins

Based upon our earlier studies (Kulaeva *et al.*, 1996), we noted that the UmuC/DinB/Rev1 superfamily was most conserved in their amino-termini and that this region could be subdivided into five discrete regions or motifs. Alignment of the eukaryotic Rad30 proteins with the *E. coli* DinB protein suggests that the Rad30-like proteins also share the same five conserved motifs (Fig. 2). These include motif III, which contains a putative DExx-box (Mg^{2+} binding site) (Walker *et al.*, 1982), that may very well be the catalytic active site of the protein. Indeed, mutations at the same site in *E. coli* UmuC (Steinborn, 1978; Koch *et al.*, 1992; Tang *et al.*, 1999), *E. coli* DinB (Wagner *et al.*, in press), and *S. cerevisiae* Rad30 (Johnson *et al.*, 1999a, *b*) result in a complete loss of polymerase activity in each of the respective proteins. Motif V contains a helix–hairpin–helix region found in many DNA binding proteins (Aravind *et al.*, 1999). No functions have yet been ascribed to motifs I, II, and IV, but they are also presumably critical for function as they are conserved in prokaryotes, archaea, and eukaryotes. As can be seen in Fig. 3, these five conserved regions are all clustered in the amino-termini of the Rad30-like proteins. By com-

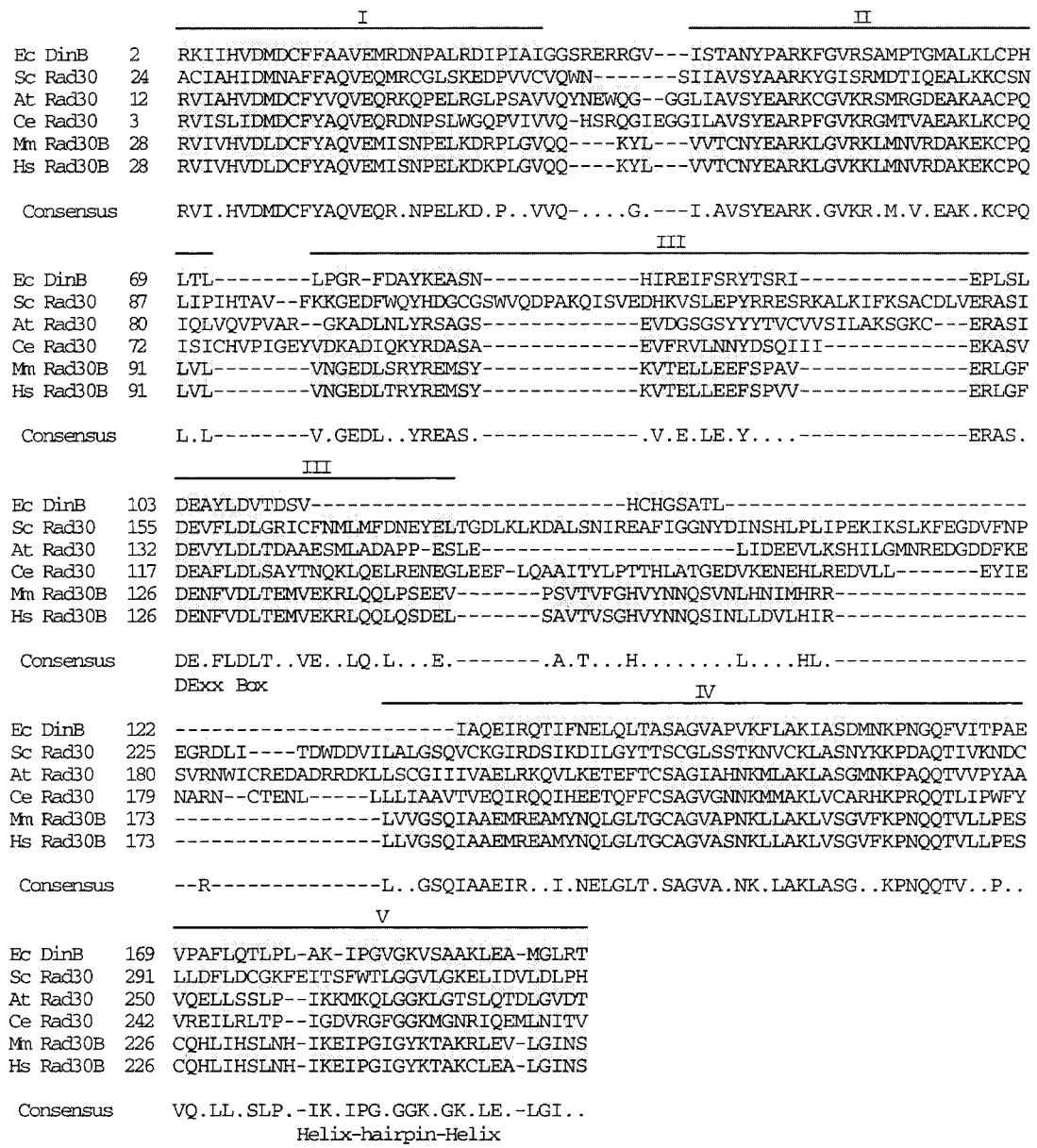

FIG. 2. Alignment of the N-terminal regions in Rad30-like proteins. *E. coli* DinB (EcDinB); *S. cerevisiae* Rad30 (ScRad30), *A. thaliana* Rad30 (AtRad30), *C. elegans* Rad30 (CeRad30), *Mus musculus* (MmRad30B), and *H. sapiens* Rad30B (HsRad30B). The alignment was performed using the MegAlign program (DNAstar, Inc., Madison, WI). The exact locations of amino acids in each protein are indicated at the left-hand side of the figure. Regions that are highly conserved are shaded in gray. The conserved motifs (I–V) identified in a previous study (Kulaeva *et al.,* 1996) are overlined. Motif III is a putative DExx- box (Mg^{2+} binding site). Motif V is a putative helix-hairpin-helix DNA binding domain.

parison, the carboxyl-termini of the Rad30-like proteins are much more diverged and vary considerably in length.

In Situ Hybridization Analysis of the Expression of Mouse Rad30b in Adult Testis

The expression of *RAD30* in various adult tissues was examined by Northern blot analysis (data not shown). In all human tissues tested, the human *RAD30B* messenger is detected as an ~3-kb transcript. Interestingly, expression of the human *RAD30B* gene is elevated in testis, and to a lesser extent in heart and pancreas, and is present at low levels in other tissues. The expression pattern of the mouse *Rad30b* gene resembles that of human *RAD30B*, with increased levels in testis and very low levels in spleen, lung, and brain tissues (data not shown).

Based upon the elevated expression of human *RAD30B* and mouse *Rad30b* in testis, we were interested in determining the stage at which mouse *Rad30b* is expressed during sperm cell development. As a consequence, adult testicular sections that had been hybridized with ^{35}S-labeled antisense (experimental) or sense (control) riboprobes for mouse *Rad30b* were ex-

amined for the cellular localization of silver grains. It was immediately obvious that not all the tubules in a given cross section of the adult testis were equally expressing mouse *Rad30b* (Fig. 4A). This suggested that only a subset of cells were expressing mouse *Rad30b* at readily detectable levels. Examination of the tubules at higher magnification revealed that round spermatids were clearly expressing relatively high levels of mouse *Rad30b* (Fig. 4B). Neither spermatogonia nor cells in the interstitial regions between the tubules appeared to express *RAD30B.*

It was of interest to determine when in the developmental pathway between the mitotic "stem cells" and the haploid spermatids detectable levels of mouse *Rad30b* were first apparent. Selected tubules were therefore examined and staged according to Oakberg (1956) as described in Russell *et al.* (1990). This analysis revealed that none of the meiotic prophase spermatocyte stages were labeled (Figs. 4C and 4D).

DNA Sequence Analysis of RAD30B in Various Repair-Deficient Human Cell Lines

Based upon the fact that the *S. cerevisiae* Rad30 protein is a DNA polymerase, it has been hypothesized that human homologs of *RAD30* might be defective in XP-V cells (Johnson *et al.,* 1999a). To test this directly, we sequenced the whole of the *RAD30B* open reading frame in two XP variants, XP30RO and XP6DU. We found that XP30RO was heterozygous and XP6DU homozygous for two mutations (Table 1). The first was an A2118G transition, resulting in a change of T706A. This alteration is, however, also present in several ESTs in the GenBank database (see Accession Nos. D78887, AA812734, AA156602, and W60419) and presumably represents a common polymorphism. The second change was a three-base insertion in the 5′-UTR. The 5′-UTR contains three repeats of CGA 23–31 bases upstream of the ATG start codon [5′(−34 bp)CGG CGA CGA CGA GGA AGA CGC CGA GGC CTG GGC CATG3′ (3 bp)]. The mutation that we have found is the insertion of a fourth CGA repeat at this point. We considered the possibility that this change could affect the translation of the gene and might conceivably be related to the disease phenotype. To investigate this further, we analyzed the 5′-UTR in two affected siblings of XP6DU, namely XP2DU and XP3DU (Table 1). Although XP3DU was, like her sibling XP6DU, homozygous for the insertion, XP2DU was heterozygous. This showed that the insertion did not segregate with the disease phenotype and suggested that it was likely to be a polymorphism. We therefore analyzed a further 14 cell lines from normal and other types of repair-deficient donors (Table 1). Of these, 4 were heterozygous for the insertion, confirming that it was a polymorphism rather than a disease-related mutation, although interestingly, the insertion was always found in association with the A706 allele. Based upon these data, we conclude that human *RAD30B* is not defective in the XP variants studied.

DISCUSSION

A Family of Novel DNA Polymerases

The molecular mechanism of damage-induced mutagenesis remains one of the major unsolved problems in the field of DNA repair. Substantial progress has, however, been recently made by the biochemical characterization of the Umu, DinB, Rev1, and Rad30 proteins, which are central to this process. Interestingly, these studies reveal that the *E. coli* Umu proteins (Tang *et al.,* 1998, 1999), *E. coli* DinB (Wagner *et al.,* in press), *S. cerevisiae* Rev1 (Nelson *et al.,* 1996b), *S. cerevisiae* Rad30 (Johnson *et al.,* 1999a), and another human homolog of *S. cerevisiae, Rad30* (Rad30A) (Masutani *et al.,* 1999b), all display limited DNA polymerase activity that is associated with the purified proteins. None of these enzymes apparently exhibits great processivity, but then again, they are presumably required only to replicate a handful of nucleotides that cannot be normally extended by the cell's main replicative enzyme. The phylogenetic study reported here (Fig. 1) of 36 related proteins suggests that they fall into four subfamilies: UmuC-like, DinB-like, Rev1-like, and Rad30-like. Given that the prototype for each family has been shown to possess DNA polymerase activity, it seems reasonable to speculate that all of the proteins shown in Fig. 1, including RAD30B, might exhibit similar DNA polymerase activities. Interestingly, many organisms carry multiple members of this superfamily. *E. coli,* for example, has both UmuC and DinB, and *S. cerevisiae* has Rev1 and Rad30; *S. pombe, Caenorhabditis elegans, M. musculus,* and *H. sapiens,* all contain DinB-like, Rev1-like, and Rad30-like homologs (Fig. 1), suggesting that they may share overlapping functions within the cell. Whether this polymerase activity is ultimately considered error-free or error-prone would presumably depend on the respective fidelity of each polymerase and the DNA lesion encountered.

Homologs of S. cerevisiae Rad30 and Their Relationship to XPV

XP variant (XP-V) patients have an increased incidence of skin cancers, similar to other XP patients, but, unlike the majority of XP patients, they show no defect in nucleotide excision repair (Cleaver and Carter, 1973) but are UV-hypermutable and have an altered mutational spectrum induced by UV light or psoralen (Raha *et al.,* 1996; McGregor *et al.,* 1999). Experiments reveal that cultured XP-V cells have a reduced ability to replicate past lesions in double-stranded DNA (Svoboda and Vos, 1995; Cordeiro-Stone *et al.,* 1997; Svoboda *et al.,* 1998), and recently it has been shown that extracts from XP-V cells have a greatly reduced capac-

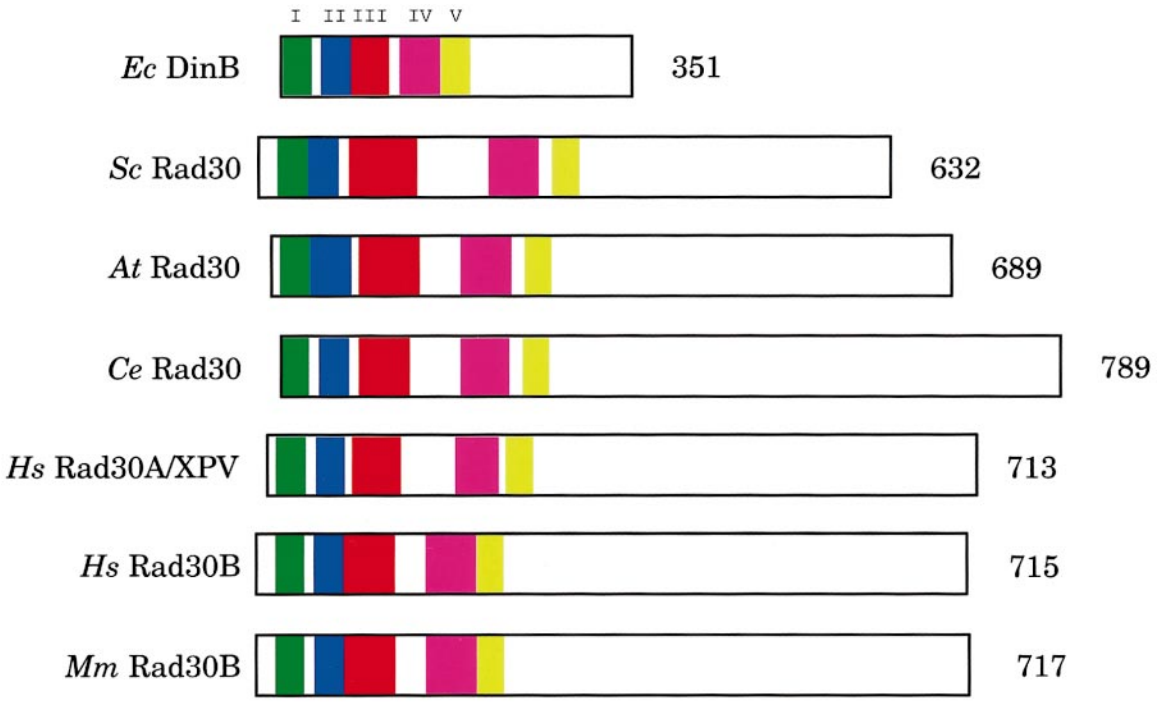

FIG. 3. Overall alignment of Rad30-like proteins. This schematic alignment of Rad30-like proteins reveals that most homology is observed in the amino-terminus and that the C-terminal tail of each protein varies in length and is much more diverged. The conserved motifs are as follows: motif 1, green; motif II, blue; motif III, red; motif IV, purple; and motif V, yellow.

ity to replicate past a site-specific acetylaminofluorene adduct in a single-stranded template. These findings clearly demonstrate that XP-V cells are deficient in translesion DNA synthesis (Cordonnier *et al.*, 1999). Such observations led Johnson *et al.* (1999a) to hypothesize that the *XPV* gene would encode a human homolog of *S. cerevisiae* Rad30.

Although DNA sequence analysis of our human *RAD30* homolog identified two common polymorphisms, neither segregated with the disease phenotype in XP variant cells (Table 1), and as a consequence, we concluded that such changes were not the cause of the XPV phenotype. However, shortly after the submission of this paper, we became aware of the work of Masutani *et al.* 1999a,b), who had purified and cloned a human protein that was able to replicate across DNA lesions. Interestingly, sequence analysis of this protein revealed that it was also related to the *S. cerevisiae* Rad30 protein, but was different from our human RAD30B homolog. Comparison of the two human homologs to *S. cerevisiae* Rad30 revealed that the human Rad30 homolog of Masutani *et al.* (in press) exhibited slightly overall higher identity (16–23% depending upon the sequence analysis program utilized) than our human Rad30 homolog (14–20% identity). As a consequence, after consultation with Dr. F. Hanaoka, we have agreed to call the human *RAD30* homolog identified by Masutani *et al.* (1999a,b) *RAD30A* and the human *RAD30* homolog reported here *RAD30B*. Interestingly, while we did not find any inactivating mutations in the human *RAD30B* gene of XP-V patients, Masutani *et al.* (1999b) did find truncating mutations in human *RAD30A* in all five XPV patients whom they analyzed. Defects in human *RAD30A*, therefore, result in the Xeroderma pigmentosum variant group phenotype (Masutani *et al.* 1999b). These findings emphasize the importance of the RAD30-related DNA polymerases in humans.

A Role for RAD30B in Spermatogenesis?

The elevated level of expression of *RAD30B* in the testis suggests a possible role for Rad30b in spermatogenesis. *In situ* hybridization analysis revealed that mouse *Rad30b* expression was restricted to a subset of cells within the adult testis, namely, round spermatids. These cells have completed both meiotic replication and meiotic divisions. Although we do not have corresponding protein data to confirm when during spermatid differentiation the gene might be functioning, it clearly must be after most replication and recombination events. It is formally possible that mouse Rad30b could be functioning to repair exogenously induced lesions in the haploid cells, which could impair the function of spermatozoa in fertilization and the subsequent activation of zygotic transcription. For example, unscheduled DNA synthesis in spermatocytes and spermatids of mice exposed to chemical mutagens such as ethyl methanesulfonate has been observed, although no attempt has been made to identify the enzymes involved (Sega, 1974, 1979); Sega *et al.*, 1978). Interestingly, in addition to many other repair genes, *HR6A* and *HR6B* (homologs of *S. cerevisiae RAD6*) are expressed at the same stage of development and have

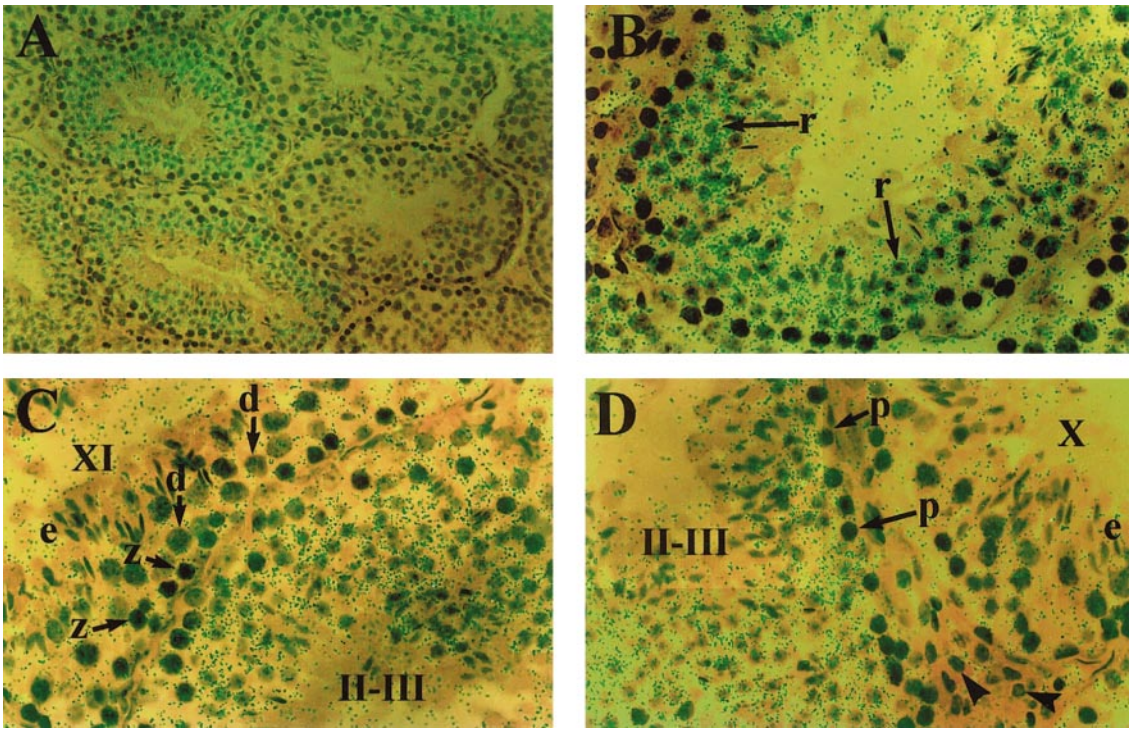

FIG. 4. Cellular localization of mouse *Rad30b* transcripts in adult testis. Histological sections of adult mouse testis were hybridized with ^{35}S-labeled mouse *Rad30b* riboprobe. Exposure time was 10 days. Photomicrographs were taken using epiluminescence optics at 20× (**A**) and at 40× magnification (**B, C,** and **D**). The stages of the seminiferous epithelium cycle are indicated by Roman numerals according to the classification of Oakberg (1956). (**A**) Mouse *Rad30b* is expressed at highest levels in only a subset of tubules in the adult mouse testis. (**B**) Round spermatids (designated "r," arrows), which are more adluminal, express high levels of mouse *Rad30b* while the spermatogonia and spermatocytes in the basal compartment do not express mouse *Rad30b* at levels above background. (**C,D**) Staged tubules show clearly that zygotene (representative cells designated "z," arrows) and diplotene spermatocytes ("d," arrows) (tubule stage XI) and pachytene and leptotene spermatocytes (tubule stage X) do not express mouse *Rad30b* at significant levels, whereas the round spermatids (tubule stage II–III) have high levels of mouse *Rad30b* transcripts. As the spermatids elongate, mouse *Rad30b* transcripts are no longer abundant (tubule stages X and XI, designated "e"). The somatic cells in the interstitial regions also do not express mouse *Rad30b* (arrowheads, lower right of panel **D**).

recently been implicated in chromatin remodeling (Koken *et al.*, 1996; Baarends *et al.*, 1999). In *S. cerevisiae,* genetic studies reveal that *RAD30* is part of the *RAD6* epistasis group (McDonald *et al.*, 1997), and we suggest that it may be no coincidence that *Rad30b* is coordinately expressed during the same stage of development as *HR6A* and *HR6B*.

Human RAD30B and a Potential Role as a Tumor Suppressor Gene

It is well established that defects in DNA repair mechanisms, such as nucleotide excision repair and mismatch repair, result in a predisposition to skin cancers (Xeroderma pigmentosum, XP) and colon cancers (*human nonpolyposis colorectal cancer*, HNPCC), respectively, demonstrating that many DNA repair genes are tumor suppressor genes. Homology of the human *RAD30B* cDNA to an STS (WI-11064) demonstrated that the human *RAD30B* gene is localized to chromosome 18q21.1. A common alteration in cancer is the loss of chromosomal regions containing tumor suppressor genes. These deletions are detected by loss of heterozygosity analyses. Interestingly, the 18q21.1 locus is often deleted in many human cancers including squamous cell carcinomas (Pearlstein *et al.*, 1998), osteosarcoma (Nellissery *et al.*, 1998), colon cancer (Lanza *et al.*, 1998), lung cancer (Takei *et al.*, 1998), and breast cancer (Yokota *et al.*, 1997), indicating the presence of tumor suppressor genes. Indeed, several cloned tumor suppressor genes have been mapped to this region, notably *DCC* (deleted in colorectal carcinomas) (Fearon *et al.*, 1990; Thiagalingam *et al.*, 1996), *SMAD2* (Riggins *et al.*, 1996), and *DPC4* (deleted in pancreatic carcinoma, locus 4), also designated *SMAD4* (Hahn *et al.*, 1996; Thiagalingam *et al.*, 1996).

However, the 18q21 region has also been implicated in the etiology of other types of human tumors for which the gene responsible has not yet been identified. For example, regions of chromosome 18q21 are frequently lost in head and neck squamous cell carcinomas (HNSCCs). Papadimitrakopoulou *et al.* (1998) identified three minimal deleted regions of chromosome 18, 18q12, 18q21.1, and 18q21.1–q21.2, in 50 primary HNSCCs. Only one of these deleted regions harbors a known tumor suppressor gene, *DCC*, and sequencing of the *SMAD2* gene revealed no mutations or polymorphisms in nine HNSCC cell lines. These findings suggest that at least two unidentified tumor

TABLE 1

Polymorphisms in the Human *RAD30B* Gene in Various Cell Lines

Cell strain	Genotype	5′-UTR insertion	A2118G
1BR	Normal	++	++
RC	Normal	++	++
XP14BR	XP-C	++	++
XP4BR	XP-C	+m	+m
XP4RO	XP-C	+m	+m
XP6BR	XP-C	++	++
XP16BR	XP-D	+m	+m
XP31BR	XP-D	++	++
XP135LO	XP-D	++	++
XP9MA	XP-D	++	++
XP1NE	XP-D	++	++
TTD10VI	XP-D	+m	+m
TTD1BEL	XP-D	++	++
CS2GO	CS-B	++	++
XP2DU	XP-V	+m	+m
XP3DU	XP-V	mm	mm
XP6DU	XP-V	mm	mm
XP7DU	XP-V	+m	+m
XP30RO	XP-V	+m	+m

Note. + indicates wildtype sequence, while m indicates the insertion of an additional CGA trinucleotide approximately 23–31 bp upstream of the human *RAD30B* gene or an A → G transition at nucleotide 2118. In this scheme, nucleotide number 1 is the A of the human *RAD30B* initiator codon.

suppressor genes are located within this region of chromosome 18 (Papadimitrakopoulou *et al.*, 1998). Similarly, Takei *et al.* (1998) have identified a novel tumor suppressor locus involved in human lung cancer on a 1-Mb region of chromosome 18q21.1 that does not encompass *DCC*, *SMAD2*, or *SMAD4*. Likewise, a putative tumor suppressor locus involved in osteosarcomas has recently been mapped to chromosome 18q21–q22, a region tightly linked to Paget disease of bone (Nellissery *et al.*, 1998). Experiments are currently in progress to explore the possibility that human *RAD30B* may be one of the novel aforementioned tumor suppressor genes at the 18q21.1 locus.

ACKNOWLEDGMENTS

We thank Mary McLenigan for assistance with the Northern blots; Weng Kong Sung for help in the staging of mouse testis tubules; Agnès Tisser and Arthur S. Levine for helpful comments during the course of this work; Fumio Hanaoka, Errol Friedberg, Valerie Gerlach, Chris Lawrence, and Robert Fuchs for generously sharing unpublished data prior to publication; and Elspeth Bruford for considerable patience in sorting out the designated nomenclature for the *RAD30A* and *RAD30B* genes. This work was supported, in part, by the NIH intramural research program and by an NIH R01 grant (HD34915) to D.J.W.

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.

Altschul, S. F., Boguski, M. S., Gish, W., and Wootton, J. C. (1994). Issues in searching molecular sequence databases. *Nat. Genet.* **6:** 119–129.

Aravind, L., Walker, D. R., and Koonin, E. V. (1999). Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res.* **27:** 1223–1242.

Baarends, W. M., Hoogerbrugge, J. W., Roest, H. P., Ooms, M., Vreeburg, J., Hoeijmakers, J. H., and Grootegoed, J. A. (1999). Histone ubiquitination and chromatin remodeling in mouse spermatogenesis. *Dev. Biol.* **207:** 322–333.

Banerjee, S. K., Christensen, R. B., Lawrence, C. W., and LeClerc, J. E. (1988). Frequency and spectrum of mutations produced by a single cis-syn thymine–thymine dimer in a single-stranded vector. *Proc. Natl. Acad. Sci. USA* **85:** 8141–8145.

Brotcorne-Lannoye, A., and Maenhaut-Michel, G. (1986). Role of RecA protein in untargeted UV mutagenesis of bacteriophage λ: Evidence for the requirement for the *dinB* gene. *Proc. Natl. Acad. Sci. USA* **83:** 3904–3908.

Chapman, D. L., and Wolgemuth, D. J. (1994). Expression of proliferating cell nuclear antigen in the mouse germ line and surrounding somatic cells suggests both proliferation-dependent and -independent modes of function. *Int. J. Dev. Biol.* **38:** 491–497.

Cleaver, J. E., and Carter, D. M. (1973). Xeroderma pigmentosum variants: Influence of temperature on DNA repair. *J. Invest. Dermatol.* **60:** 29–32.

Cordeiro-Stone, M., Zaritskaya, L. S., Price, L. K., and Kaufmann, W. K. (1997). Replication fork bypass of a pyrimidine dimer blocking leading strand DNA synthesis. *J. Biol. Chem.* **272:** 13945–13954.

Cordonnier, A. M., Lehmann, A. R., and Fuchs, R. P. P. (1999). Impaired translesion synthesis in xeroderma pigmentosum variant extracts. *Mol. Cell. Biol.* **19:** 2206–2211.

Fearon, E. R., Cho, K. R., Nigro, J. M., Kern, S. E., Simons, J. W., Ruppert, J. M., Hamilton, S. R., Preisinger, A. C., Thomas, G., and Kinzler, K. W. (1990). Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science* **247:** 49–56.

Friedberg, E. C., Walker, G. C., and Siede, W. (1995). "DNA Repair and Mutagenesis," Am. Soc. Microbiol., Washington, DC.

Hahn, S. A., Schutte, M., Hoque, A. T., Moskaluk, C. A., da Costa, L. T., Rozenblum, E., Weinstein, C. L., Fischer, A., Yeo, C. J., Hruban, R. H., and Kern, S. E. (1996). *DPC4*, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* **271:** 350–353.

Higgins, N. P., Kato, K., and Strauss, B. (1976). A model for replication repair in mammalian cells. *J. Mol. Biol.* **101:** 417–425.

Johnson, R. E., Prakash, S., and Prakash, L. (1999a). Efficient bypass of a thymine–thymine dimer by yeast DNA polymerase, polη. *Science* **283:** 1001–1004.

Johnson, R. E., Prakash, S., and Prakash, L. (1999b). Requirement of DNA polymerase activity of yeast Rad30 protein for its biological function. *J. Biol. Chem.* **274:** 15975–15977.

Kim, S. R., Maenhaut-Michel, G., Yamada, M., Yamamoto, Y., Matsui, K., Sofuni, T., Nohmi, T., and Ohmori, H. (1997). Multiple pathways for SOS-induced mutagenesis in *Escherichia coli:* An overexpression of *dinB*/*dinP* results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc. Natl. Acad. Sci. USA* **94:** 13792–13797.

Koch, W. H., Ennis, D. G., Levine, A. S., and Woodgate, R. (1992). *Escherichia coli umuDC* mutants: DNA sequence alterations and UmuD cleavage. *Mol. Gen. Genet.* **233:** 443–448.

Koffel-Schwartz, N., Coin, F., Veaute, X., and Fuchs, R. P. P. (1996). Cellular strategies for accommodating replication-hindering adducts in DNA: Control by the SOS response in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **93:** 7805–7810.

Koken, M. H., Hoogerbrugge, J. W., Jasper-Dekker, I., de Wit, J., Willemsen, R., Roest, H. P., Grootegoed, J. A., and Hoeijmakers, J. H. (1996). Expression of the ubiquitin-conjugating DNA repair

enzymes HHR6A and B suggests a role in spermatogenesis and chromatin modification. *Dev. Biol.* **173:** 119–132.

Kozak, M. (1987). An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15:** 8125–8148.

Kozak, M. (1997). Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. *EMBO J.* **16:** 2482–2492.

Kulaeva, O. I., Koonin, E. V., McDonald, J. P., Randall, S. K., Rabinovich, N., Connaughton, J. F., Levine, A. S., and Woodgate, R. (1996). Identification of a DinB/UmuC homolog in the archeon *Sulfolobus solfataricus. Mutat. Res.* **357:** 245–253.

Lanza, G., Matteuzzi, M., Gafa, R., Orvieto, E., Maestri, I., Santini, A., and del Senno, L. (1998). Chromosome 18q allelic loss and prognosis in stage II and III colon cancer. *Int. J. Cancer* **79:** 390–395.

Larimer, F. W., Perry, J. R., and Hardigree, A. A. (1989). The *REV1* gene of *Saccharomyces cerevisiae:* Isolation, sequence and functional analysis. *J. Bacteriol.* **171:** 230–237.

Lawrence, C. W., and Hinkle, D. C. (1996). DNA polymerase ζ and the control of DNA damage induced mutagenesis in eukaryotes. *In* "Cancer Surveys: Genetic Instability in Cancer" (T. Lindahl, Ed.), pp. 21–31, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Masutani, C., Araki, M., Yamada, A., Kusumoto, R., Nogimori, T., Maekawa, T., Iwai, S., and Hanaoka, F. (1999a). Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *EMBO. J.* **18:** 3491–3501.

Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanoaka, F. (1999b). The *XPV* (xeroderma pigmentosum variant) gene encodes human DNA polymerase η. *Nature.* **399:** 700–704.

McDonald, J. P., Levine, A. S., and Woodgate, R. (1997). The *Saccharomyces cerevisiae RAD30* gene, a homologue of *Escherichia coli dinB* and *umuC*, is DNA damage inducible and functions in a novel error-free postreplication repair mechanism. *Genetics* **147:** 1557–1568.

McGregor, W. G., Wei, D., Maher, V. M., and McCormick, J. J. (1999). Abnormal, error-prone bypass of photoproducts by xeroderma pigmentosum variant cell extracts results in extreme strand bias for the kinds of mutations induced by UV light. *Mol. Cell. Biol.* **19:** 147–154.

Nellissery, M. J., Padalecki, S. S., Brkanac, Z., Singer, F. R., Roodman, G. D., Unni, K. K., Leach, R. J., and Hansen, M. F. (1998). Evidence for a novel osteosarcoma tumor-suppressor gene in the chromosome 18 region genetically linked with Paget disease of bone. *Am. J. Hum. Genet.* **63:** 817–824.

Nelson, J. R., Lawrence, C. W., and Hinkle, D. C. (1996a). Thymine–thymine dimer bypass by yeast DNA polymerase ζ. *Science* **272:** 1646–1649.

Nelson, J. R., Lawrence, C. W., and Hinkle, D. C. (1996b). Deoxycytidyl transferase activity of yeast REV1 protein. *Nature* **382:** 729–731.

Oakberg, E. F. (1956). A description of sperminogenesis in the mouse and its use in analysis of the life cycle of the seminiferous epithelium and germ cell renewal. *Am. J. Anat.* **99:** 391–413.

Ohmori, H., Hatada, E., Qiao, Y., Tsuji, M., and Fukuda, R. (1995). *dinP*, a new gene in *Escherichia coli*, whose product shows similarities to UmuC and its homologues. *Mutat. Res.* **347:** 1–7.

Papadimitrakopoulou, V. A., Oh, Y., El-Naggar, A., Izzo, J., Clayman, G., and Mao, L. (1998). Presence of multiple incontiguous deleted regions at the long arm of chromosome 18 in head and neck cancer. *Clin. Cancer Res.* **4:** 539–544.

Pearlstein, R. P., Benninger, M. S., Carey, T. E., Zarbo, R. J., Torres, F. X., Rybicki, B. A., and Dyke, D. L. (1998). Loss of 18q predicts poor survival of patients with squamous cell carcinoma of the head and neck. *Genes Chromosomes Cancer* **21:** 333–339.

Raha, M., Wang, G., Seidman, M. M., and Glazer, P. M. (1996). Mutagenesis by third-strand-directed psoralen adducts in repair-deficient human cells: High frequency and altered spectrum in a xeroderma pigmentosum variant. *Proc. Natl. Acad. Sci. USA* **93:** 2941–2946.

Rhee, K., and Wolgemuth, D. J. (1995). Cdk family genes are expressed not only in dividing but also in terminally differentiated mouse germ cells, suggesting their possible function during both cell division and differentiation. *Dev. Dyn.* **204:** 406–420.

Riggins, G. J., Thiagalingam, S., Rozenblum, E., Weinstein, C. L., Kern, S. E., Hamilton, S. R., Willson, J. K., Markowitz, S. D., Kinzler, K. W., and Vogelstein, B. (1996). Mad-related genes in the human. *Nat. Genet.* **13:** 347–349.

Roush, A. A., Suarez, M., Friedberg, E. C., Radman, M., and Siede, W. (1998). Deletion of the *Saccharomyces cerevisiae* gene *RAD30* encoding an *Escherichia coli* DinB homolog confers UV radiation sensitivity and altered mutability. *Mol. Gen. Genet.* **257:** 686–692.

Rupp, W. D., and Howard-Flanders, P. (1968). Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet radiation. *J. Mol. Biol.* **31:** 291–304.

Rupp, W. D., Wilde, C. E., Reno, D. L., and Howard-Flanders, P. (1971). Exchanges between DNA strands in ultraviolet-irradiated *Escherichia coli. J. Mol. Biol.* **61:** 25–44.

Russell, L. D., Ettlin, R. A., Sinha Hikim, A. P., and Clegg, E. D. (1990). "Histological and Histopathological Evaluation of the Testis," Cache River Press, Clearwater, FL.

Sega, G. A. (1974). Unscheduled DNA synthesis in the germ cells of male mice exposed *in vivo* to the chemical mutagen ethyl methanesulfonate. *Proc. Natl. Acad. Sci. USA* **71:** 4955–4959.

Sega, G. A. (1979). Unscheduled DNA synthesis (DNA repair) in the germ cells of male mice: Its role in the study of mammalian mutagenesis. *Genetics* **92:** s49–s58.

Sega, G. A., Sotomayor, R. E., and Owens, J. G. (1978). A study of unscheduled DNA synthesis induced by X-rays in the germ cells of male mice. *Mutat. Res.* **49:** 239–257.

Smith, B. T., and Walker, G. C. (1998). Mutagenesis and more: *umuDC* and the *Escherichia coli* SOS response. *Genetics* **148:** 1599–1610.

Steinborn, G. (1978). Uvm mutants of *Escherichia coli* K12 deficient in UV mutagenesis. I. Isolation of *uvm* mutants and their phenotypical characterization in DNA repair and mutagenesis. *Mol. Gen. Genet.* **165:** 87–93.

Svoboda, D. L., and Vos, J. M. (1995). Differential replication of a single, UV-induced lesion in the leading or lagging strand by a human cell extract: Fork uncoupling or gap formation. *Proc. Natl. Acad. Sci. USA* **92:** 11975–11979.

Svoboda, D. L., Briley, L. P., and Vos, J. M. (1998). Defective bypass replication of a leading strand cyclobutane thymine dimer in xeroderma pigmentosum variant cell extracts. *Cancer Res.* **58:** 2445–2448.

Takei, K., Kohno, T., Hamada, K., Takita, J., Noguchi, M., Matsuno, Y., Hirohashi, S., Uezato, H., and Yokota, J. (1998). A novel tumor suppressor locus on chromosome 18q involved in the development of human lung cancer. *Cancer Res.* **58:** 3700–3705.

Tang, M., Bruck, I., Eritja, R., Turner, J., Frank, E. G., Woodgate, R., O'Donnell, M., and Goodman, M. F. (1998). Biochemical basis of SOS-induced mutagenesis in *Escherichia coli:* Reconstitution of *in vitro* lesion bypass dependent on the $UmuD'_2C$ mutagenic complex and RecA. *Proc. Natl. Acad. Sci. USA* **95:** 9755–9760.

Tang, M., Shen, X., Frank, E. G., O'Donnell, M., Woodgate, R., and Goodman, M. F. (1999). $UmuD'_2C$ is an error-prone DNA polymerase, *Escherichia coli* DNA pol V. *Proc. Natl. Acad. Sci. USA* **96:** 8919–8924.

Thiagalingam, S., Lengauer, C., Leach, F. S., Schutte, M., Hahn, S. A., Overhauser, J., Willson, J. K., Markowitz, S., Hamilton, S. R., Kern, S. E., Kinzler, K. W., and Vogelstein, B. (1996). Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. *Nat. Genet.* **13:** 343–346.

Wagner, J., Gruz, P., Kim, S. R., Yamada, M., Matsui, K., Fuchs, R. P. P., and Nohmi, T. The *dinB* gene encodes an novel *Escherichia coli* DNA polymerase (DNA pol IV) involved in mutagenesis. *Mol. Cell.*, in press.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982). Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases, and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1:** 945–951.

Woodgate, R., and Levine, A. S. (1996). Damage inducible mutagenesis: Recent insights into the activities of the Umu family of mutagenesis proteins. *In* "Cancer Surveys: Genetic Instability in Cancer" (T. Lindahl, Ed.), pp. 117–140, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Yokota, T., Matsumoto, S., Yoshimoto, M., Kasumi, F., Akiyama, F., Sakamoto, G., Nakamura, Y., and Emi, M. (1997). Mapping of a breast cancer tumor suppressor gene locus to a 4-cM interval on chromosome 18q21. *Jpn. J. Cancer Res.* **88:** 959–964.