MOLECULAR AND CELLULAR BIOLOGY, Nov. 2007, p. 7816–7827 Vol. 27, No. 22
0270-7306/07/$08.00+0 doi:10.1128/MCB.00455-07

Accurate Homologous Recombination Is a Prominent Double-Strand Break Repair Pathway in Mammalian Chromosomes and Is Modulated by Mismatch Repair Protein Msh2[▽]

Jason A. Smith, Laura A. Bannister, Vikram Bhattacharjee, Yibin Wang, Barbara Criscuolo Waldman, and Alan S. Waldman*

Department of Biological Sciences, University of South Carolina, Columbia, South Carolina 29208

Received 16 March 2007/Returned for modification 21 May 2007/Accepted 27 August 2007

We designed DNA substrates to study intrachromosomal recombination in mammalian chromosomes. Each substrate contains a thymidine kinase (*tk*) gene fused to a neomycin resistance (*neo*) gene. The fusion gene is disrupted by an oligonucleotide containing the 18-bp recognition site for endonuclease I-SceI. Substrates also contain a "donor" *tk* sequence that displays 1% or 19% sequence divergence relative to the *tk* portion of the fusion gene. Each donor serves as a potential recombination partner for the fusion gene. After stably transfecting substrates into mammalian cell lines, we investigated spontaneous recombination and double-strand break (DSB)-induced recombination following I-SceI expression. No recombination events between sequences with 19% divergence were recovered. Strikingly, even though no selection for accurate repair was imposed, accurate conservative homologous recombination was the predominant DSB repair event recovered from rodent and human cell lines transfected with the substrate containing sequences displaying 1% divergence. Our work is the first unequivocal demonstration that homologous recombination can serve as a major DSB repair pathway in mammalian chromosomes. We also found that Msh2 can modulate homologous recombination in that Msh2 deficiency promoted discontinuity and increased length of gene conversion tracts and brought about a severalfold increase in the overall frequency of DSB-induced recombination.

Homologous recombination is understood to play important roles in a number of biological processes in mammalian cells, including the repair of DNA double-strand breaks (DSBs). Recombination must be regulated to prevent unwanted rearrangements, and we had previously shown that spontaneous recombination in mammalian chromosomes normally occurs only between sequences sharing a considerable length and degree of homology (36, 60, 61, 66, 67). If recombination between imperfectly matched sequences (homeologous recombination) were to occur at an appreciable frequency, a genomic catastrophe could result when one considers the abundance of imperfectly matched repeated sequences, such as Alu family repeats, in a mammalian genome. Consistent with expectations for stringent homology requirements for recombination, we had previously reported that efficient intrachromosomal recombination between linked sequences in a mammalian genome requires that the sequences share more than 134 bp of perfect homology and that a single nucleotide mismatch can measurably reduce recombination (36, 61). For two linked sequences about 1 kb in length and displaying 19% sequence divergence, recombination was reduced over 1,000-fold compared to recombination between sequences displaying near perfect homology (60).

Although it is not yet known precisely which proteins bring about the notable sensitivity of recombination in mammalian cells to sequence divergence, a body of evidence points to the mismatch repair (MMR) machinery as a likely regulator of recombination (reviewed in references 20, 26, 30, 49, and 55). Msh2 (*M*ut*S* *h*omolog *2*) is an important player in MMR in eukaryotes, functioning as a heterodimer in association with Msh6 or Msh3. Interest in Msh2 has been sparked by the finding that inherited deficiency in Msh2 is one cause of hereditary nonpolyposis colorectal cancer, a cancer predisposition syndrome associated with instability of microsatellite sequences (27, 38, 40, 44, 63). In addition to its function in postreplicative MMR, Msh2 is involved in a number of other pathways that help to preserve genome integrity. Msh2 is involved in response to DNA damage, with a role in a signaling cascade that activates cell cycle checkpoints or apoptosis (3, 7, 8, 10, 11, 22, 26, 33, 49, 64). Msh2 is a component of the BRCA-1-associated genome surveillance complex, a multiprotein complex involved in the recognition and response to abnormal DNA structures (17, 29, 62). In yeast, Msh2 has been implicated in DSB repair by playing a role in the removal of nonhomologous DNA tails and possibly assisting in a homology search (20, 21, 31, 48, 53, 54). Msh2-deficient mouse cells display an increase in chromosomal damage and fail to form Mre11 and Rad51 foci in the G_2 phase of the cell cycle following X-irradiation (24). Further, Msh2 deficiency is associated with an elevated number of Rad51 foci in mouse and human cells in the absence of treatment with damaging agents (24, 68) and an elevated number of Rad51 foci following replication arrest of human cells with hydroxyurea (68).

The influence of Msh2 on Rad51 focus formation is consistent with a regulatory role for Msh2 in recombinational repair. Of additional relevance to a potential role for Msh2 in recombination are reports that Msh2 displays high-affinity binding to substrates resembling Holliday junctions (2). Msh2 colocalizes

* Corresponding author. Mailing address: Department of Biological Sciences, University of South Carolina, 700 Sumter St., Columbia, SC 29208. Phone: (803) 777-8405. Fax: (803) 777-4002. E-mail: awaldman@biol.sc.edu.

[▽] Published ahead of print on 10 September 2007.

with Msh6, p53, BLM, and Rad51 at sites of DSBs at stalled replication forks in human cells, where it appears to help regulate processing of recombination intermediates (68). Recent work with yeast (*Saccharomyces cerevisiae*) (20, 25, 52, 54) has suggested that the Msh2-Msh6 heterodimer, perhaps in conjunction with the Sgs1 helicase (a BLM homolog), participates in a process of heteroduplex rejection that dismantles recombination intermediates formed between imperfectly matched DNA sequences and prevents homeologous recombination. Indeed, deficiency in Msh2 homologs in yeast, moss, *Arabidopsis*, and mammalian embryonic stem (ES) cells has been shown to be associated with an increase in homeologous recombination (12, 18, 19, 25, 42, 46, 52, 54, 56, 65). It has also been reported that human cells deficient in Msh2 display a defect in the completion of accurate recombinational repair of DSBs in plasmid DNA (58).

In this work, we aimed to investigate the impact of loss of Msh2 on intrachromosomal recombination in mammalian somatic cells. To do so, we developed a novel set of recombination substrates to study spontaneous as well as DSB-induced recombination between sequences displaying 1% or 19% divergence. As alluded to above, homologous recombination has been implicated as a potential means for repairing a DSB. Nonhomologous end joining (NHEJ) represents an alternative DSB repair pathway in mammalian cells that involves no template and, in contrast to recombinational repair, is error prone because one or several nucleotides are usually deleted or inserted prior to DSB healing. NHEJ is generally considered to be a major DSB repair pathway throughout the mammalian cell cycle, with homologous recombination apparently playing a repair role in the late S and G_2 phases of the cell cycle (recently reviewed in reference 51). While developing our current studies, we came to the realization that, despite an extensive literature on recombination in mammalian cells, there was actually no published work that definitively demonstrated that accurate homologous recombination can function as a significant DSB repair pathway in mammalian chromosomes. Such a demonstration would require an experimental scheme that meets at least the following two criteria: (i) DSB repair must be monitored in an unbiased fashion involving no selection for accurate repair, and (ii) DSB repair products must be analyzed at the nucleotide level to demonstrate faithful exchange of nucleotides between interacting sequences. A survey of the literature revealed no work that met both of these criteria. The substrates developed by us in the present work allowed us to meet these criteria and provided us with the opportunity to decisively explore the use of recombination as a means for DSB repair.

In this report, we present what we believe is the first unequivocal evidence that accurate, conservative recombination can serve as a primary means for the repair of DSBs in mammalian chromosomes. Perhaps surprisingly, we found that accurate homologous recombination predominated over NHEJ as a means for DSB repair in hamster, mouse, and human somatic cells. Furthermore, by studying recombination in Msh2-deficient Chinese hamster ovary (CHO) cell line Clone B (CB) and in Msh2-proficient cell line MT+, from which CB was derived (4, 5), we found that accurate recombination is not dependent on Msh2 function. Msh2 deficiency also did not abrogate the barrier to homeologous recombination between sequences with 19% divergence, alter the rate of spontaneous recombination between sequences with 1% divergence, or change the balance of events resolving as crossovers versus noncrossovers. At the same time, our data revealed that Msh2 deficiency increased gene conversion tract length, promoted discontinuity of gene conversion tracts, and increased the overall rate of DSB-induced recombination. Thus, accurate homologous recombination can be modulated by Msh2 in several ways that may promote genome stability in mammalian somatic cells.

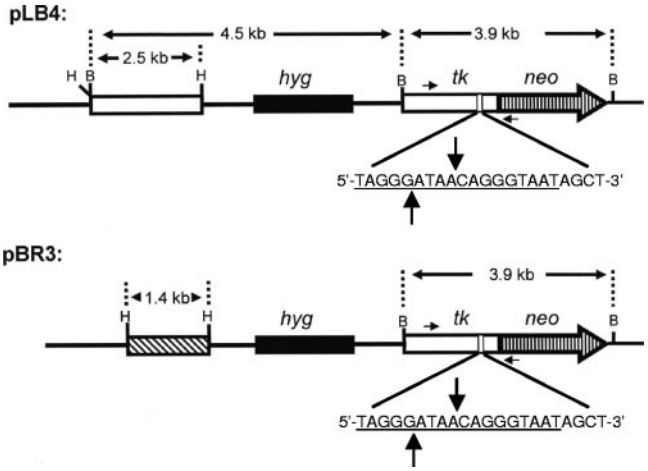

FIG. 1. Recombination substrates pLB4 and pBR3. Each substrate contains a functional hygromycin gene (*hyg*), used to select for stably transfected cells, and a *tk-neo* fusion gene that is disrupted by a 22-bp oligonucleotide containing the 18-bp recognition site for endonuclease I-SceI (underlined sequence). Sites of staggered cleavage by I-SceI are indicated. Substrate pLB4 contains a 2.5-kb HindIII fragment (open rectangle) containing a complete HSV-1 *tk* gene, while pBR3 contains a 1.4-kb HindIII fragment (striped rectangle) containing a complete HSV-2 *tk* gene. The orientation, with respect to transcription, of both the *tk* gene and *tk-neo* fusion gene in each substrate is from left to right as drawn. PCR primers AW85 and AW91 are indicated by short horizontal arrows and are located 1.4 kb apart in the *tk-neo* fusion gene. BamHI (B) and HindIII (H) sites are shown.

MATERIALS AND METHODS

Cell culture. CHO cell lines MT+ and CB, an Msh2-deficient derivative of MT+, were provided by Margherita Bignami and were previously characterized (4, 5). Normal human fibroblast cell line GM00637 was obtained from the NIGMS. The CHO and human cell lines were cultured in alpha-modified minimum essential medium (Sigma) supplemented with 10% fetal bovine serum. Mouse L*tk*$^-$ fibroblasts were cultured in Dulbecco's modified Eagle medium (Sigma) supplemented with minimal essential medium nonessential amino acids and 50 μg of gentamicin sulfate per ml. All cells were maintained at 37°C in a humidified atmosphere of 5% CO_2.

Recombination substrates. Plasmids pLB4 and pBR3 (Fig. 1) were derived from pTNeo99-7 which was described previously (6) and contains a *tk-neo* fusion gene rendered nonfunctional by the insertion of a 22-bp oligonucleotide that contains the 18-bp recognition sequence for yeast endonuclease I-SceI. Substrate pLB4 was produced by inserting a 2.5-kb HindIII fragment containing a complete functional herpes simplex virus type 1 (HSV-1) thymidine kinase (*tk*) gene into the unique HindIII site on pTNeo99-7. This inserted "donor" fragment served as a potential recombination partner for the *tk-neo* fusion gene. The *tk* donor on pLB4 shares about 1.7 kb of homology with the *tk* portion of the fusion gene. The *tk* donor and the *tk-neo* fusion gene are oriented as direct repeats, with the direction of transcription going from left to right, as drawn in Fig. 1. Due to several scattered mismatches, the donor on pLB4 displays about 1% sequence divergence from the *tk* portion of the *tk-neo* fusion gene (Fig. 2).

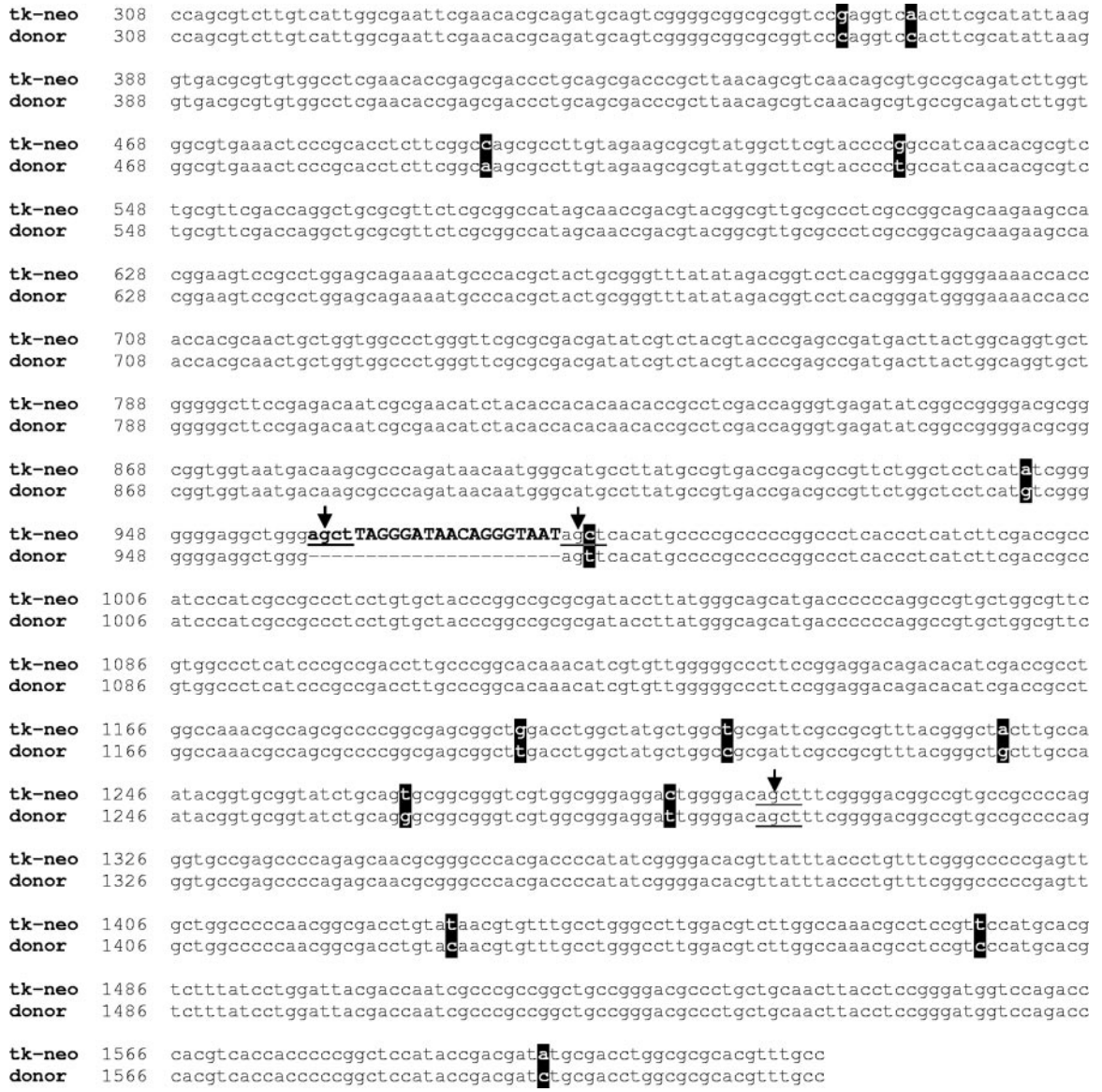

FIG. 2. Alignment of *tk-neo* fusion gene sequence with donor *tk* sequence from pLB4. Nucleotides 308 to 1622 (numbering according to reference 59) of the *tk* portion of the *tk-neo* fusion gene are aligned with the corresponding donor sequence from pLB4. The span of *tk* sequence shown comprises the *tk* portion of PCR products generated by primers AW85 and AW91 and used in subsequent analyses. Mismatches between donor and *tk-neo* fusion gene sequences are highlighted. The 22-bp oligonucleotide containing the 18-bp I-SceI recognition sequence inserted in the fusion gene (absent from donor) is depicted in boldface, with the actual I-SceI recognition sequence in uppercase. AluI recognition sites (agct) are underlined and are indicated by downward arrows.

Substrate pBR3 was produced by inserting a 1.4-kb HindIII fragment containing a complete functional HSV-2 *tk* gene into the unique HindIII site on pTNeo99-7. This inserted fragment displays 19% sequence divergence with the *tk* portion of the *tk-neo* fusion gene and served as a potential homeologous recombination partner for the *tk-neo* fusion gene. The *tk* donor on pBR3 shares 1.4 kb of homeology with the *tk* portion of the fusion gene. As in pLB4, the donor and fusion gene reside as direct repeats.

Establishing cell lines stably transfected with recombination substrates. For each cell line, 5×10^6 cells were electroporated with either 3 μg of pLB4 or 3 μg of pBR3 using a Bio-Rad Gene Pulser set at 1,000 V, 25 μF. Cells were plated into 150-cm^2 flasks and allowed to grow for 2 days under no selection. Cells were then plated at a density of 1×10^6 cells per 75-cm^2 flask into medium supplemented with hygromycin at either 500 μg/ml (CB CHO cells), 400 μg/ml (MT+ CHO cells), 100 μg/ml (GM00637 human cells), or 400 μg/ml (L*tk*$^-$ mouse cells). After 10 to 14 days, hygromycin-resistant (Hyg^R) colonies were picked. These clones were propagated and screened by Southern blot analysis to identify cell lines containing a single integrated copy of the transfected plasmid substrate.

Determination of spontaneous intrachromosomal recombination rate. Spontaneous recombination rates were determined by fluctuation tests. Starting from single cells, at least 10 independent subclones were generated from each cell line. The subclones were propagated to several million cells each and then plated at a density of 5×10^5 cells per 75-cm^2 flask into medium containing 1,000 μg/ml G418 (for CHO and human cells) or 400 μg/ml G418 (for mouse L*tk*$^-$ cells) in order to select for G418^R segregants. After 14 days of growth under selection, colonies were counted and recombination rates (in terms of events/cell/generation) were calculated by the method of the median (32).

Recovery of DSB-induced G418^R clones. Plasmid pCMV3xnls-I-SceI ("pSce") was generously provided by Maria Jasin (Sloan Kettering). This plasmid contains a gene encoding the I-SceI endonuclease under the control of the cytomegalovirus (CMV) promoter and is expressible in mammalian cells. For DSB induc-

tion, 5×10^6 cells were resuspended in 800 μl of phosphate-buffered saline containing 20 μg of pSce (or in phosphate-buffered saline alone, for controls). Cells were electroporated using a Bio-Rad Gene Pulser set to 1,000 V, 25 μF. Cells were plated into 150-cm^2 flasks and allowed to grow for 2 days under no selection. Cells were then plated at a density of 5×10^5 cells per 75-cm^2 flask into medium supplemented with G418. After 10 to 14 days, $G418^R$ colonies were counted, picked, and propagated for further analysis.

Southern blotting analysis. Genomic DNA samples (8 μg each) were digested with appropriate restriction enzymes and resolved on 0.8% agarose gels. DNA was transferred to nitrocellulose membranes and hybridized with a ^{32}P-labeled HSV-1 and/or HSV-2 *tk* probe as previously described (37).

PCR amplification and DNA sequence analysis. A segment of the *tk-neo* fusion gene spanning the I-SceI site was amplified from 500 ng of genomic DNA isolated from $G418^R$ clones using primers AW85 (5′-TAATACGACTCACTA TAGGGCCAGCGTCTTGTCATTGGCG-3′) and AW91 (5′-GATTTAGGTG ACACTATAGCCAAGCGGCCGGAGAACCTG-3′). AW85 is composed of nucleotides 308 to 327 of the coding sequence of the HSV-1 *tk* gene (numbering according to reference 59), with a T7 forward universal primer appended to the 5′ end of the primer. AW91 is composed of 20 nucleotides from the noncoding strand of the neomycin gene mapping 25 through 44 bp downstream from the neomycin start codon, with an Sp6 primer appended to the 5′ end of the primer. PCR was carried out using Ready-To-Go PCR beads (GE Healthcare) and a "touchdown" PCR protocol. The annealing temperature was initially set to 72°C and was progressively decreased in steps of 2°C down to 62°C with two cycles at each temperature. An additional 20 cycles were run at an annealing temperature of 60°C. For each experiment, the PCR products were expected to be ~1.4 kb in length unless detectable deletions or insertions had occurred. Prior to sequencing, PCR products were treated with shrimp alkaline phosphatase and exonuclease I (USB). PCR products were then sequenced from a T7 or Sp6 primer on a LICOR 4000L at the DNA Sequencing and Synthesis Core Facility in the Department of Biological Sciences at the University of South Carolina. At least 800 bp of DNA sequence was determined for each PCR product sequenced.

RESULTS

Design of substrates to study intrachromosomal recombination. We engineered substrates pLB4 and pBR3 (Fig. 1) to study homologous and homeologous recombination, respectively, in mammalian chromosomes. Each substrate contains a *tk-neo* fusion gene. Inserted within the *tk* portion of the fusion gene in each substrate is a 22-bp oligonucleotide that contains the 18-bp recognition site for endonuclease I-SceI. This insertion disrupts the function of the fusion gene and allows for the targeted introduction of a DSB by transfecting pSce (an I-SceI expression vector) into cells containing an integrated substrate. Substrates pLB4 and pBR3 are accordingly useful for studying both spontaneous and DSB-induced recombination. Furthermore, the 18-bp length of the I-SceI recognition site makes it likely that expression of I-SceI results in the introduction of only a single DSB in the genome, within the integrated substrate. Recombination events between the *tk-neo* fusion gene and a linked *tk* donor sequence in pLB4 or pBR3 can restore function to the *neo* sequence and can be recovered by selection in G418.

A seminal feature of the recombination substrates is that restoration of function to the *neo* portion of the *tk-neo* fusion gene does not require that the 22-bp insert in the *tk* portion of the fusion gene be removed accurately, since any event that restores an appropriate reading frame for *neo* expression will confer resistance to G418. This feature is of particular relevance to studies of DSB-induced recombination since selection for accurate recombinational repair is eliminated. This feature thus allows recovery of a broad spectrum of DSB events.

As shown in Fig. 2, there are several scattered nucleotide mismatches between the *tk* portion of the *tk-neo* fusion gene and the *tk* donor sequence on pLB4. These mismatches aid in the sequence analysis of products of recombination events and allow an unambiguous demonstration of transfer of an expanse of genetic information between the donor *tk* sequence and the fusion gene. The mismatches produce a low level of sequence divergence (about 1%) between the donor sequence and the *tk-neo* fusion gene, while preserving substantial stretches of perfect homology, including 410- and 231-bp segments of continuous homology. Each of these stretches of continuous homology exceeds the minimal amount of homology required for efficient homologous recombination (61). For simplicity, we refer to the donor and fusion gene sequences on pLB4 as being "homologous" despite the sparse mismatches. In contrast, the donor and fusion gene sequences on pBR3 display 19% divergence with no substantial stretches of continuous homology. We refer to these latter sequences as being "homeologous".

Recovery of spontaneous homologous recombination events from wild-type and Msh2-deficient CHO cell lines. We wanted to study the effect of Msh2 deficiency on homologous and homeologous recombination in mammalian chromosomes. To accomplish this, MT+ (Msh2 proficient) and CB (Msh2 deficient) CHO cells were stably transfected with pLB4 or pBR3 (Fig. 1). Hyg^R colonies were recovered, and genomic DNA was isolated. Cell lines containing a single integrated copy of pLB4 or pBR3 were identified by Southern blotting analysis (data not shown). Two MT+ cell lines and two CB cell lines were identified that each contained a single integrated copy of pLB4, and these cell lines were designated MT+/pLB4-3, MT+/pLB4-22, CB/pLB4-9, and CB/pLB4-20. Two MT+ cell lines and two CB cell lines were identified that each contained a single integrated copy of pBR3, and these cell lines were designated MT+/pBR3-9, MT+/pBR3-10, CB/pBR3-10, and CB/pBR3-40.

To measure homologous recombination rates, fluctuation tests were performed on cell lines containing pLB4 to determine the rate of appearance of $G418^R$ segregants (Table 1). The median number of colonies recovered per subclone of MT+/pLB4 cell lines was comparable to the median number of colonies recovered per subclone of CB/pLB4 cell lines.

To ascertain that recovered colonies indeed arose via recombination, a portion of the *tk-neo* fusion gene encompassing the original position of the disrupting 22-bp oligonucleotide was PCR amplified from genomic DNA isolated from $G418^R$ colonies. A 1.4-kb PCR product was obtained from each clone (Fig. 3A), as expected (Fig. 1). PCR products were digested with AluI (Fig. 3B). In parental cell lines, two AluI sites flank the I-SceI site, while a third AluI site is about 300 bp downstream from the I-SceI site (Fig. 2). If recombination had taken place in a given clone, the 22-bp insert containing the I-SceI site and the two flanking AluI sites would have been eliminated and the PCR product would produce a 990-bp fragment and a 381-bp fragment upon digestion with AluI (see Fig. 3B, lanes 3 to 8). On the other hand, retention of the parental AluI restriction sites in a PCR product would produce 653-, 381-, 337-, and 22-bp fragments upon AluI digestion. The latter outcome, or a substantially similar outcome, would denote that a clone arose via a mechanism other than accurate homologous recombination (see Fig. 3B, lane 9).

In total, 39 $G418^R$ clones from MT+/pLB4 cell lines and 34 $G418^R$ clones from CB/pLB4 cell lines were screened by PCR and AluI digestion. As indicated in Table 1, 32 out of 39 clones

TABLE 1. Recovery of spontaneous recombination events from CHO cell lines

Cell line	No. of cells plated (10^6)[a]	Median no. of $G418^R$ colonies recovered[b]	Total no. of colonies analyzed[c]	No. of recombinant colonies[d]	Recombination rate (10^{-7})[e]
MT+/pLB4-3	40	4	19	17	4.58
MT+/pLB4-22	40	7.5	20	15	5.90
CB/pLB4-9	36	7	18	18	7.5
CB/pLB4-20	40	4	16	16	5.12
MT+/pBR3-9	80	2.5	20	0	
MT+/pBR3-10	80	2	22	0	
CB/pBR3-10	160	5	57	0	
CB/pBR3-40	160	13.5	94	0	

[a] Total number of cells plated into G418 selection in fluctuation tests.
[b] Median number of $G418^R$ colonies recovered per subclone in fluctuation tests.
[c] Total number of $G418^R$ colonies analyzed by PCR amplification and AluI digestion as described in the text.
[d] Number of colonies analyzed that displayed the recombinant AluI restriction pattern.
[e] Recombination rates were determined for cell lines containing pLB4 as described in Materials and Methods and were then corrected to reflect the fraction of colonies analyzed that were bona fide recombinants. For cell lines containing pBR3, rates were not determined since no bona fide recombinants were found among the clones analyzed.

from the MT+/pLB4 cell lines displayed the restriction pattern expected for accurate recombinants, while the remaining 7 clones did not. The latter clones, which presumably arose from mutations that allowed expression of the fusion gene, were not analyzed further. All 34 clones recovered from the CB/pLB4 cell lines displayed the restriction pattern for recombinants. The calculated recombination rates were corrected to account for the recovery of the nonrecombinants (Table 1). The average recombination rate for MT+/pLB4 cell lines was 5.24×10^{-7}, and the average recombination rate for CB/pLB4 cell lines was 6.31×10^{-7}.

Recombination events were further classified as gene conversions (noncrossovers) or crossovers. Genomic DNA samples were digested with BamHI and displayed on a Southern blot using a *tk*-specific probe. For gene conversions, 4.5- and 3.9-kb fragments were expected while only a 3.9-kb fragment would be displayed in the case of a crossover (Fig. 4A). A

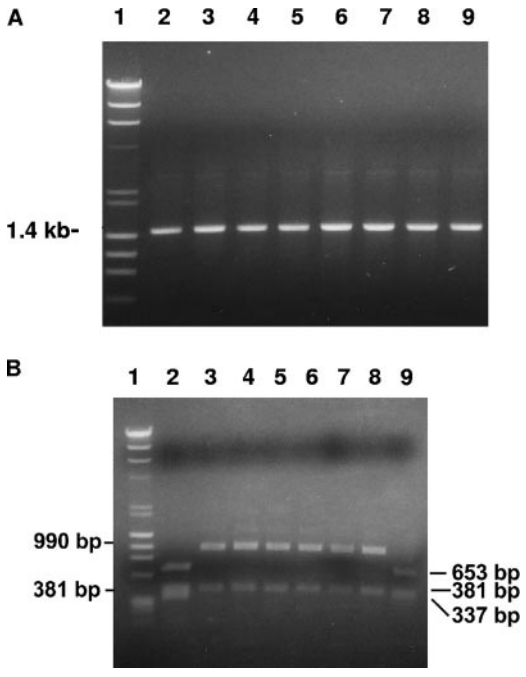

FIG. 3. Representative analysis of PCR products from spontaneous $G418^R$ segregants. (A) Lane 1 contains DNA size markers (lambda phage DNA digested with HindIII and phi X phage DNA digested with HaeIII). Lane 2 displays the PCR product generated from genomic DNA isolated from parental cell line MT+/pLB4-3 using primers AW85 and AW91; lanes 3 to 9 display PCR products generated from $G418^R$ segregants from MT+/pLB4-3. All PCR products were 1.4 kb, as expected. (B) Lane 1 displays DNA size markers. Lanes 2 to 9 display the PCR products from the corresponding lanes in panel A following digestion with AluI. The digested PCR product from parental cell line MT+/pLB4-3 in lane 2 served as a control for the nonrecombinant pattern. The clones displayed in lanes 3 to 8 produced the recombinant AluI digest pattern, while the clone in lane 9 produced a nonrecombinant pattern. See text for details.

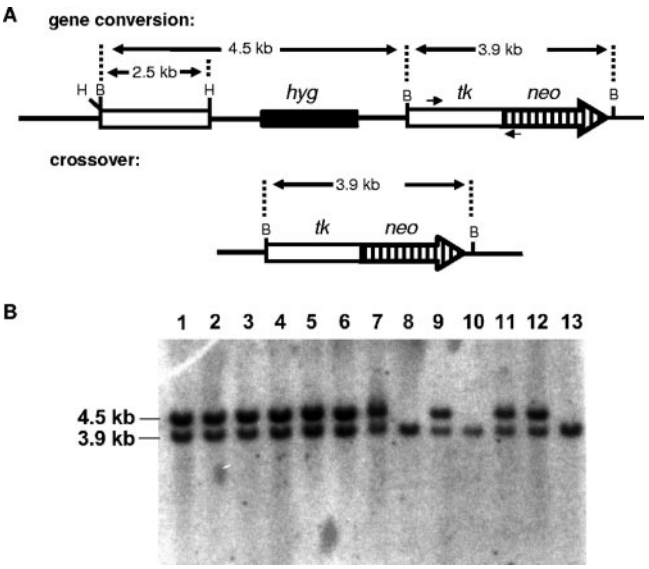

FIG. 4. Representative Southern blot analysis of recombination events. (A) Schematic illustrations of gene conversion and crossover events that produce a functional *tk-neo* fusion gene. Diagnostic expected BamHI fragments are shown. (B) Genomic DNA samples (8 μg) from $G418^R$ colonies recovered from MT+/pLB4 cell lines were digested with BamHI and hybridized with a probe specific for HSV-1 *tk* to distinguish gene conversions from crossovers. The samples in lanes 1 to 7, 9, 11, and 12 displayed the 4.5- and 3.9-kb bands diagnostic for gene conversions, while the samples in lanes 8, 10, and 13 displayed a single 3.9-kb band diagnostic for crossovers.

TABLE 2. Lengths of spontaneous gene conversion tracts in CHO cell lines

Clone	Tract length (bp)[a]
MT+ cell lines	
MT+/pLB4-22-1A	3
MT+/pLB4-22-2A	3
MT+/pLB4-3-1B	21
MT+/pLB4-3-2A	21
MT+/pLB4-3-3B	21
MT+/pLB4-3-4A	21
MT+/pLB4-3-5A	21
MT+/pLB4-3-7A	21
MT+/pLB4-3-10A	21
MT+/pLB4-22-3A	21
MT+/pLB4-22-4A	21
MT+/pLB4-22-5A	21
MT+/pLB4-22-6A	21
MT+/pLB4-22-6B	21
MT+/pLB4-22-7A	21
MT+/pLB4-22-8B	21
MT+/pLB4-3-6A	255
MT+/pLB4-22-9A	432
MT+/pLB4-22-10B	666
CB cell lines	
CB/pLB4-9-1B	21
CB/pLB4-9-2A	21
CB/pLB4-9-3A	21
CB/pLB4-9-4A	21
CB/pLB4-9-6A	21
CB/pLB4-20-2A	21
CB/pLB4-20-3A	21
CB/pLB4-20-4A	21
CB/pLB4-20-6B	21
CB/pLB4-20-7B	255
CB/pLB4-9-4B	297 (discontinuous)
CB/pLB4-9-5A	297 (discontinuous)
CB/pLB4-20-7A	325 (discontinuous)
CB/pLB4-20-8A	348 (discontinuous)
CB/pLB4-9-8B	348
CB/pLB4-9-9B	655
CB/pLB4-20-9A	744
CB/pLB4-20-1A	795

[a] Shown is the minimum gene conversion tract length, calculated as the distance between the outermost nucleotide markers converted. As indicated, four gene conversion tracts recovered from CB/pLB4 cell lines were discontinuous (see Fig. 5). All gene conversion tracts were mutation free.

representative blot is shown in Fig. 4B. Nineteen out of 25 recombinants analyzed from MT+/pLB4 cell lines displayed the expected restriction pattern for gene conversions, while 18 out of 23 recombinants from CB/pLB4 cell lines were gene conversions. All remaining events analyzed were classified as crossovers. Screening of recombinant clones revealed that all clones classified as gene conversions retained resistance to hygromycin, while all clones classified as crossovers were sensitive to hygromycin as expected (see Fig. 4A). To summarize, there was no evidence for any difference in the relative frequencies of gene conversions and crossovers in MT+ versus CB cells.

The DNA sequence was determined for the 19 clones recovered from MT+/pLB4 cell lines and for the 18 clones recovered from CB/pLB4 cell lines that, based on AluI digestions of PCR products and Southern blotting analysis, had each apparently undergone a gene conversion event. Using the nucleotide differences between the donor and *tk-neo* fusion gene sequences (Fig. 2), we determined the minimal length of gene conversion tracts (Table 2). For MT+/pLB4 cell lines, spontaneous gene conversion tracts ranged from 3 to 666 bp in length, with an average length of 87 bp. For CB/pLB4 cell lines, gene conversion tracts ranged from 21 to 795 bp in length, with an average length of 236 bp. Curiously, four recombinants recovered from CB/pLB4 cell lines displayed discontinuous gene conversion tracts (Table 2 and Fig. 5). All gene conversion tracts analyzed were mutation free. The DNA sequence data also confirmed that screening of clones by AluI digestion of PCR products was a reliable reporter for accurate recombination.

Substrate pBR3 (Fig. 1) was designed to allow recovery of intrachromosomal homeologous recombination events between sequences with 19% divergence, should such events occur at a detectable frequency. The frequency of spontaneous appearance of $G418^R$ segregants was determined for two MT+ cell lines containing pBR3 and for two CB cell lines containing pBR3 (Table 1). The colony frequencies were similar for MT+ and CB cell lines and were severalfold lower than those seen with cell lines containing pLB4 (Table 1).

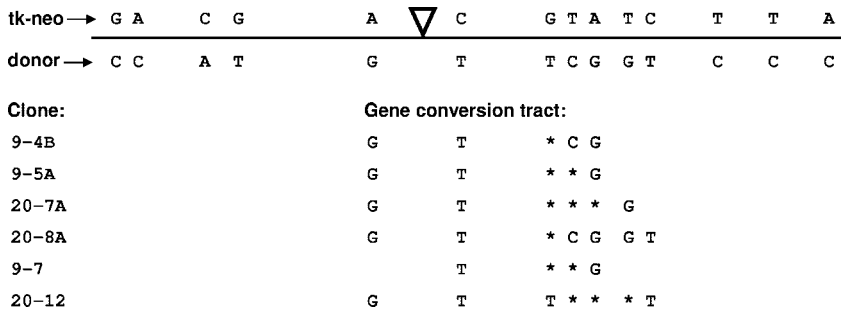

FIG. 5. Discontinuous gene conversion tracts. Shown are the structures of discontinuous gene conversion tracts displayed by six $G418^R$ clones recovered from CB/pLB4 cell lines. At the top of the figure is a schematic of the nucleotide mismatches between the parental *tk-neo* gene sequence (above the line) and the HSV-1 *tk* donor sequence (below the line). The precise sequence context of the nucleotide mismatches is illustrated in Fig. 2. The position of the I-SceI recognition site in the parental *tk-neo* fusion gene is indicated by the inverted triangle. In the lower portion of the figure is a representation of the six discontinuous conversion tracts recovered. The first four tracts were recovered from spontaneous recombinants, while the last two were recovered from DSB-induced recombinants. For each conversion tract, nucleotide mismatches originating from the donor sequence and present in the conversion tract are indicated. Asterisks indicate tract discontinuities: that is, positions at which donor nucleotides were absent from the conversion tract despite the presence of both upstream and downstream donor nucleotides in the conversion tract.

TABLE 3. Recovery of DSB-induced recombination events from CHO cell lines

Cell line	No. of expts[a]	Colony frequency (10^{-3})[b]	No. of colonies		
			Analyzed[c]	Recombinant[d]	Nonrecombinant[e]
MT+/pLB4-3	2	3.01 (2.7–3.32)	45	40	5
MT+/pLB4-22	4	1.55 (0.48–2.48)	93	81	12
CB/pLB4-9	2	19.0 (10.8–27.1)	44	43	1
CB/pLB4-20	3	10.7 (9.5–13.0)	77	77	0
MT+/pBR3-9	1	0.146	27	0	27
MT+/pBR3-10	1	0.214	25	0	25
CB/pBR3-10	1	0.873	28	0	28
CB/pBR3-40	1	0.295	26	0	26

[a] Number of electroporations with pSce.
[b] Average frequency of G418^R colonies per electroporation. Colony frequency was calculated for each electroporation as number of G418^R colonies divided by number of cells plated into selection. For each cell line, the range in colony frequencies for all electroporations is shown in parentheses.
[c] Number of G418^R colonies analyzed by PCR and AluI digestion.
[d] Number of colonies analyzed that displayed the recombinant AluI restriction pattern.
[e] Nonrecombinant colonies included several clones with visible deletions or insertions based on the size of the PCR product produced.

A total of 42 G418^R colonies recovered from MT+/pBR3 cell lines and 151 G418^R colonies recovered from CB/pBR3 cell lines were analyzed by PCR and AluI digestion similar to the analysis described above. Again, nonrecombinant G418^R colonies would produce 653-, 381-, 337-, and 22-bp bands when digested with AluI. Depending on gene conversion tract length or the position of crossing over, a G418^R colony arising from homeologous recombination could produce one of several different patterns because the HSV-2 *tk* donor sequence (not shown) contains three AluI recognition sites. However, all potential recombinant AluI restriction patterns were distinct from the parental nonrecombinant pattern. Each of the 193 G418^R clones analyzed displayed the restriction pattern of the parent *tk-neo* fusion gene, indicating that none of the clones arose via homeologous recombination (data not shown). These G418^R clones presumably arose from spontaneous mutations that allowed expression of the fusion gene and were not analyzed further. Compared to the frequency of spontaneous homologous recombination observed for cell lines containing pLB4, the frequency of spontaneous homeologous recombination was reduced about 100-fold or more for MT+ and CB cells.

Recovery of DSB-induced homologous recombination events from wild-type and Msh2-deficient CHO cell lines. Substrates pLB4 and pBR3 each contain an I-SceI recognition site inserted into the *tk* portion of the *tk-neo* fusion gene, and a DSB can be induced within the fusion gene by electroporating the I-SceI expression plasmid pSce into cell lines containing an integrated copy of either pLB4 or pBR3. Cell lines MT+/pLB4-3, MT+/pLB4-22, CB/pLB4-9, and CB/pLB4-22 were each electroporated with pSce, and DSB-induced G418^R colonies were recovered (Table 3). Electroporation with pSce substantially increased colony frequencies above the spontaneous frequencies. The colony frequencies for the MT+/pLB4 cell lines were greater than 10^{-3}, while the colony frequencies for the CB/pLB4 cell lines exceeded 10^{-2}. The difference in the frequencies of DSB-induced colonies recovered from CB cell lines versus MT+ cell lines was highly statistically significant ($P = 0.003$ by t test).

Overall, 138 DSB-induced G418^R colonies from MT+/pLB4 cell lines and 121 colonies from CB/pLB4 cell lines were subjected to PCR and AluI digestion as described above. A total of 121 out of 138 colonies from MT+/pLB4 cell lines and 120 out of 121 colonies from CB/pLB4 cell lines yielded the recombinant AluI restriction pattern (Table 3).

DNA samples from G418^R clones that appeared to be recombinants based on AluI digests of PCR products were digested with BamHI and analyzed on Southern blots using a *tk*-specific probe as described above. A total of 91 out of 101 analyzed G418^R clones recovered from MT+/pLB4 cell lines displayed the pattern expected for gene conversions, while 10 clones appeared to have arisen by crossover events. Among 88 analyzed clones recovered from CB/pLB4 cell lines, 80 had apparently undergone gene conversions, while 8 arose from crossovers. As described above, screening of recombinants revealed retention of resistance to hygromycin among all clones classified as gene conversions and sensitivity to hygromycin among all clones classified as crossovers, as expected (see Fig. 4A). As observed above for spontaneous recombination events, there was no difference in the relative abundance of DSB-induced gene conversions and crossovers recovered from the MT+/pLB4 versus CB/pLB4 cell lines.

It seemed remarkable that such a high percentage of recovered DSB events (at least 87% for each cell line) appeared to have arisen via homologous recombination since pLB4 was designed to allow recovery of an unbiased spectrum of DSB events. To demonstrate unambiguously that accurate homologous recombination had occurred, DNA sequence analysis was performed on clones that were characterized as gene conversions based on the above PCR and Southern blot analyses. The DNA sequence was determined for gene conversion tracts for 22 DSB-induced G418^R colonies recovered from MT+/pLB4 cell lines and for 23 G418^R colonies recovered from CB/pLB4 cell lines (Table 4). Such analysis revealed that accurate exchange of genetic information occurred in all samples. The minimal lengths for DSB-induced conversion tracts ranged from 3 bp to 468 bp for MT+/pLB4 and CB/pLB4 cell lines. All gene conversion tracts analyzed were mutation free, and two discontinuous gene conversion tracts were recovered from CB/pLB4 cell lines (Table 4 and Fig. 5).

A greater number of nonrecombinants were recovered from MT+/pLB4 cell lines compared with CB/pLB4 cell lines, and this difference was highly statistically significant ($P = 1.4 \times 10^{-4}$ by a Fisher exact test). Sequencing of the nonrecombi-

TABLE 4. Lengths of DSB-induced gene conversion tracts in CHO cell lines

Clone	Tract length (bp)[a]
MT+ cell lines	
MT+/pLB4-3-7	3
MT+/pLB4-3-18	3
MT+/pLB4-3-4	3
MT+/pLB4-3-5	3
MT+/pLB4-3-21	3
MT+/pLB4-22-21	3
MT+/pLB4-3-3-21	
MT+/pLB4-3-11	21
MT+/pLB4-3-21	21
MT+/pLB4-3-2	21
MT+/pLB4-3-13	21
MT+/pLB4-3-17	21
MT+/pLB4-22-9	21
MT+/pLB4-22-15	21
MT+/pLB4-22-17	21
MT+/pLB4-22-20	21
MT+/pLB4-22-30	21
MT+/pLB4-22-5	21
MT+/pLB4-22-6	21
MT+/pLB4-22-8	21
MT+/pLB4-22-11	21
MT+/pLB4-22-13	468
CB cell lines	
CB/pLB4-9-5	3
CB/pLB4-9-7	3
CB/pLB4-9-15	3
CB/pLB4-9-17	3
CB/pLB4-9-22	3
CB/pLB4-20-6	3
CB/pLB4-20-7	3
CB/pLB4-20-20	3
CB/pLB4-9-8	21
CB/pLB4-9-16	21
CB/pLB4-9-4	21
CB/pLB4-9-21	21
CB/pLB4-20-10	21
CB/pLB4-20-19	21
CB/pLB4-20-20	21
CB/pLB4-20-28	21
CB/pLB4-20-8	21
CB/pLB4-20-22	21
CB/pLB4-20-14	255
CB/pLB4-9-7	279 (discontinuous)
CB/pLB4-20-12	348 (discontinuous)
CB/pLB4-20-3	305
CBpLB4-9-19	468

[a] Shown is the minimum gene conversion tract length, calculated as the distance between the outermost nucleotide markers converted. All conversion tracts were mutation free, and two conversion tracts from CB/pLB4 cell lines were discontinuous. Discontinuous tracts are illustrated further in Fig. 5.

nants revealed that they arose via NHEJ. The deletion size for the single NHEJ event recovered from a CB/pLB4 line was 22 bp, while the deletion size for the 17 NHEJ events recovered from MT+/pLB4 cell lines ranged from 1 bp to over 200 bp (data not shown). Analysis of NHEJ events recovered from MT+ and CB cells has been presented in detail previously (50). Notably, despite the fact that NHEJ events were recoverable from the CHO cell lines containing pLB4, accurate homologous recombination events were predominant among recovered DSB events.

In attempts to recover DSB-induced homeologous recombination events, cell lines MT+/pBR3-9, MT+/pBR3-10, CB/pBR3-10, and CB/pBR3-40 were each electroporated with pSce and G418^R colonies were recovered (Table 3). Colony frequencies were 10- to 100-fold lower than the frequencies recorded for recovery of DSB-induced G418^R colonies from cell lines containing the homologous recombination substrate pLB4 (Table 3).

DNA samples isolated from G418^R colonies were subjected to PCR analysis and AluI digestion as described above. In total, 52 DSB-induced G418^R clones recovered from MT+/pBR3 cell lines and 54 clones recovered from CB/pBR3 cell lines were analyzed. As recorded in Table 3, none of the DSB-induced G418^R clones arose from homeologous recombination. The majority of clones displayed the parental nonrecombinant AluI restriction pattern, and several clones had apparent insertions or deletions and correspondingly displayed PCR products larger or smaller than the expected 1.4 kb. There was no apparent difference in the recovery of deletion or insertion events from MT+ versus CB cells (data not shown).

Recovery of DSB-induced homologous recombination events from mouse and human cell lines. The predominance of accurate homologous recombination events among DSB repair events recovered from the CHO cell lines was striking, and so we were curious if a similar outcome would be observed for other types of mammalian cell lines transfected with pLB4. We derived a cell line from mouse L*tk*$^-$ fibroblasts and two cell lines from normal human GM00637 fibroblasts that were stably transfected with a single copy of pLB4. The transfected cell lines were designated L*tk*$^-$/pLB4, GM637/pLB4#1, and GM637/pLB4#2. Cells were transfected with pSce, and G418^R colonies were recovered (Table 5). The frequency of G418^R colonies recovered from L*tk*$^-$/pLB4 was 6×10^{-5}, while colonies were recovered from GM637/pLB4#1 and GM637/pLB4#2 at frequencies of 4.34×10^{-3} and 1.68×10^{-3}, respectively. PCR and Southern blot analysis revealed that, similar to the findings for the CHO cell lines, DSB events arose

TABLE 5. Recovery of DSB-induced recombination events from mouse and human cell lines

Cell line	No. of expts[a]	Colony frequency (10^{-3})[b]	No. of colonies		
			Analyzed[c]	Recombinant[d]	Nonrecombinant
L*tk*$^-$/pLB4	2	0.060 (0.058–0.062)	25	20	5
GM637/pLB4#1	6	4.34 (2.01–7.31)	147	109	38
GM637/pLB4#2	2	1.68 (1.46–1.89)	76	61	15

[a] Number of electroporations with pSce.
[b] Average frequency of G418^R colonies per electroporation. Colony frequency was calculated for each electroporation as the number of G418^R colonies divided by the number of cells plated into selection. For each cell line, the range in colony frequencies for all electroporations is shown in parentheses.
[c] Number of G418^R colonies analyzed by PCR and AluI digestion.
[d] Number of colonies analyzed that displayed the recombinant AluI restriction pattern.

TABLE 6. Lengths of DSB-induced gene conversion tracts in mouse and human cell lines

Cell line	Tract length (bp)[a]	No. of clones[b]
L*tk*$^-$/pLB4	3	1
	21	14
	348	1
	535	1
	598	2
	924	1
GM637/pLB4#1	21	7
	795	1
GM637/pLB4#2	3	3
	21	12
	255	2
	348	1
	468	1
	924	3

[a] Shown is the minimum gene conversion tract length, calculated as the distance between the outermost nucleotide markers converted. All conversion tracts were mutation free.

[b] Number of independent clones displaying the indicated gene conversion tract length.

from the mouse and human cell lines predominantly via homologous recombination. For the mouse cell line, recombinants comprised 80% of the recovered DSB repair events, while recombinants represented 74% and 80% of the events recovered from the two human cell lines (Table 5).

PCR and Southern blot analysis (not shown) revealed that all of the recombinants isolated from the mouse cell line and about 80% of the recombinants from the human cell lines were gene conversions, with the balance being crossovers. DNA sequence analysis of representative clones characterized as gene conversions confirmed unambiguously that accurate genetic exchange had occurred in such clones. The lengths of gene conversion tracts recovered from mouse and human cell lines ranged from 3 bp to 924 bp (Table 6). Sequence analysis also showed that the G418^R clones classified as nonrecombinants arose from NHEJ. We were thus able to use substrate pLB4 to show that accurate homologous recombination can serve as a major DSB repair pathway in several different types of mammalian somatic cells.

DISCUSSION

We examined homologous and homeologous recombination in CHO cell lines that were wild type for Msh2 (MT+ cells) or Msh2 deficient (CB cells). Using cell lines stably transfected with pLB4 (Fig. 1), we found that deficiency in Msh2 did not have a discernible impact on the rate or accuracy of spontaneous homologous recombination or on the balance between spontaneous gene conversions versus crossovers. Based on studies indicating a role for MMR (and MutS homologs in particular) in the suppression of homeologous recombination in lower and higher eukaryotes as well as prokaryotes (12, 18, 19, 25, 42, 46, 52, 54, 56, 65), we anticipated that we might see an increased level of homeologous recombination in Msh2-deficient CB cells. Experiments using cell lines transfected with pBR3 revealed, however, that the barrier to spontaneous recombination between 19% diverged sequences could not be overcome by Msh2 deficiency alone (Table 1). We also failed to recover any DSB-induced homeologous recombination events from cell lines containing pBR3, indicating that homeologous recombination was reduced well over 100-fold regardless of Msh2 status (Table 3). In yeast, suppression of recombination between sequences displaying only a few percent sequence divergence is MMR dependent while suppression of recombination between sequences displaying a level of divergence approaching 20% is only partly MMR dependent (14, 16). Notably, however, it has been reported that a deficiency in Msh2 in yeast can increase recombination between sequences displaying 18% divergence more than 40-fold and, in some experiments, Msh2 deficiency restored the rate of such homeologous recombination to a level that was only 10- to 30-fold lower than that of recombination between sequences sharing 100% identity (14, 16). This contrasts with our results, which suggest that in a mammalian genome, recombination between sequences with 19% divergence can effectively be precluded without the involvement of Msh2.

Inducing a genomic DSB by transfecting pSce into cell lines containing pLB4 brought about a large increase in the frequency of G418^R colonies relative to the spontaneous colony frequency (Table 3). For MT+/pLB4 cell lines, there was about a 10,000-fold increase in G418^R colonies following pSce transfection, while for CB/pLB4 cell lines the increase in colony frequency was even greater, about 100,000-fold. About 85% of the DSB-induced colonies recovered from MT+/pLB4 cell lines and virtually all DSB-induced colonies recovered from CB/pLB4 cell lines were found to be recombinants (Table 3). In our previous work (50) using a construct similar to pLB4 but containing no *tk* donor sequence, we reported that the frequency of NHEJ events recovered from MT+-derived cell lines and from CB-derived cell lines was about 3×10^{-4}. In contrast, in our present work the frequency of DSB-induced colonies recovered from MT+/pLB4 cell lines was about 3×10^{-3} and for CB/pLB4 cell lines the colony frequency was about 2×10^{-2}. Our collective results indicate that a significant portion of DSBs are processed in ways that do not produce viable G418^R clones when a recombination donor sequence is absent. We infer that when a donor sequence is available, as in pLB4, recovery of G418^R colonies is substantially enhanced by the channeling of DSBs into a preferred pathway of repair via recombination.

The latter observation touches on what is perhaps the most striking observation to come from our work, namely, that accurate homologous recombination was the predominant DSB pathway among recovered events in mammalian chromosomes. Historically, NHEJ has been viewed as the major pathway for healing DSBs in mammalian genomes (reviewed in references 28, 35, and 57). In our experimental system, a DSB was induced in the *tk* portion of the *tk-neo* fusion gene and any NHEJ event that restored an appropriate reading frame to allow *neo* function would have been recovered. Despite the lack of selection for accurate DSB repair, our results indicated a substantially greater recovery of accurate recombination events compared with NHEJ. In addition to our initial observation of a predominance of accurate homologous recombination among DSB events recovered from CHO cell lines, we

reproduced such results in mouse and human cell lines transfected with pLB4.

Previous reports of DSB repair via recombination in mammalian cells either involved genetic selection for accurate repair or did not include analysis of repair products at the nucleotide level. For example, in a previous report in which DSB repair events were recovered from mammalian cells without any genetic selection (34), 3 out of 22 recovered DSB events appeared consistent with gene conversions possibly arising from conservative homologous recombination. However, no DNA sequence analysis was provided to ascertain that an accurate exchange of genetic information had occurred. In our present work, we recovered 431 recombinants out of a total of 507 DSB repair events from cell lines containing pLB4 and analyzed 95 of the recombinants at the nucleotide level. All 95 sequenced clones displayed accurate recombinational repair of the DSB within gene conversion tracts extending up to several hundred base pairs in length. We believe, therefore, that our current report is the first unequivocal demonstration that accurate, conservative homologous recombination can be a major DSB repair pathway in mitotically dividing mammalian cells. Our work indicates that pLB4 is a useful tool for probing homologous recombination in mammalian chromosomes. Specifically what conditions, DNA sequences, and/or genetic background is critical in prompting the use of homologous recombination as the DSB repair pathway of choice is an area of intense interest awaiting further investigation.

Our data show that accurate homologous recombination is not dependent on Msh2 function and that gene conversion tracts can encompass hundreds of base pairs in the absence of Msh2. It was previously reported that Msh2 colocalizes with BLM in Rad51 foci which are believed to represent sites of recombinational repair of DSBs, perhaps at sites of stalled replication forks (68). In light of BLM's antirecombination activity (9, 25, 41, 52, 54), the reported stimulation of BLM's helicase activity by Msh2 (68), and the increase in Rad51 foci observed in Msh2-deficient human cells (68), one may posit that Msh2 plays a role in suppressing recombination at lesions such as DSBs. Our finding of a significantly greater recovery of DSB-induced recombinants from CB/pLB4 cell lines compared with MT+/pLB4 cell lines (Table 3) directly reinforces this notion. Msh2 may therefore protect genome integrity by reducing the likelihood of inappropriate recombination at DSBs generated by stalled replication forks or generated by other means in somatic cells. We cannot rule out the formal possibility that pLB4 integrated into different chromosomal contexts in the CB and MT+ cell lines such that DSB-induction was more efficient in the CB lines. However, the similar DSB-induced colony frequencies for the two MT+ cell lines, the similar DSB-induced colony frequencies for the two CB cell lines, and the significantly higher proportion of recombinants among DSB-induced colonies recovered from CB cell lines compared with MT+ lines suggest quantitative and qualitative differences between MT+ and CB cell lines that are not so readily explainable by simple differences in cutting efficiency by I-SceI.

Gene conversion tracts recovered from Msh2 mutants were longer than those recovered from wild-type cells. Among all gene conversions recovered from CB/pLB4 cell lines, 14 out of 41 events analyzed displayed gene conversion tracts that encompassed 3 or more mismatched nucleotide markers. This was significantly different ($P = 0.0076$ by a chi-square test) from conversion tracts recovered from MT+/pLB4 cell lines, for which only 4 out of 41 conversion tracts encompassed 3 or more nucleotide markers. Our finding of a lengthening of gene conversion tracts in Msh2-deficient CHO cells is similar to findings reported for the impact of Msh2 deficiency in mouse ES cells (18) as well as MMR deficiency in yeast (1, 13, 14). Our work adds to a body of data that suggests that MMR machinery regulates the extension of heteroduplex DNA (hDNA) during recombination.

We made the curious observation that a total of six gene conversion tracts recovered from Msh2-deficient cells were discontinuous, while none of the conversion tracts recovered from Msh2-proficient cells were discontinuous (Tables 2 and 4 and Fig. 5). The greater number of discontinuous conversion tracts recovered from Msh2-deficient versus Msh2-proficient cells was statistically significant ($P = 0.011$ by a chi-square test). A previous report revealed an apparent elevation in the frequency of discontinuous gene conversion tracts in Msh2-deficient mouse ES cells (18), although the significance of discontinuous tracts was not discussed in the previous report. Discontinuous gene conversion tracts were also recovered in Msh2- or Pms1-deficient yeast and in Msh6-deficient *Drosophila* (15, 45). Gene conversion tracts are generally viewed as products of MMR acting on hDNA, and discontinuity of tracts therefore seemingly implies some form of patchy MMR. It is thus challenging to explain discontinuous conversion tracts in the absence of Msh2. A possibility is that discontinuous gene conversion tracts produced in the absence of Msh2 are produced by an alternate MMR system. Indeed, good evidence has been presented that discontinuous gene conversion tracts produced in yeast in the absence of functional Msh2-dependent "long-patch" MMR are products of a "short-patch" MMR system operating on hDNA (15). In *Saccharomyces cerevisiae*, short-patch repair is independent of nucleotide excision repair (NER) functions (15), while short-patch repair in *Schizosaccharomyces pombe* is dependent on NER (23). Evidence for NER-independent short-patch MMR in mammalian cells has been reported (39, 43). Short-patch MMR pathways may be unmasked in long-patch repair-deficient genetic backgrounds.

In a previous report, it was shown that homologous recombination and NHEJ are often coupled in mouse ES cells (47). Specifically, it was reported that a 3′ terminus from one end of DSB can initiate homologous recombination by invading a homologous sequence and become extended by DNA synthesis using the invaded sequence as a template. The nascent strand of DNA can then be released from the template and joined to a DNA terminus from the other side of the DSB by NHEJ. In our experimental system, events of this type (initiation by gene conversion and completion by NHEJ) would be recoverable had they occurred. However, our analyses revealed no evidence for such events. Our results therefore demonstrate that, in the somatic cell lines used in our work, events that initiated as homologous recombination were completed as homologous recombination events. Our present results also appear somewhat at odds with a report that Msh2 is needed for the accurate completion of homologous recombination events in human cells (58). Aside from the use of different cell types, a possible

explanation for the discrepancy between our work and the earlier report is that the earlier work involved the use of plasmid substrates, while we studied intrachromosomal recombination events.

In summary, we used the novel DNA substrate pLB4 to provide the first definitive demonstration that accurate recombination can serve as a major DSB repair pathway in mammalian somatic cells. We also found that Msh2 plays roles in suppressing DSB-induced recombination, limiting the length of gene conversion tracts, and minimizing the number of recombinant sequence junctions in recombination events. The application of homologous recombination as a DSB repair pathway and the modulation of this pathway by Msh2 undoubtedly coalesce to enhance genome integrity. Efforts aimed at better understanding the molecular mechanisms in which Msh2 and other recombination regulatory proteins participate will increase our understanding of how genome stability is maintained and how it may be compromised.

ACKNOWLEDGMENT

This work was supported by Public Health Service grant GM47110 from the National Institute for General Medical Sciences to A.S.W.

REFERENCES

1. **Alani, E., R. A. Reenean, and R. D. Kolodner.** 1994. Interaction between mismatch repair and genetic recombination in Saccharomyces cerevisiae. Genetics **137:**19–39.
2. **Alani, E., S. Lee, M. F. Kane, J. Griffith, and R. D. Kolodner.** 1997. Saccharomyces cerevisiae MSH2, a mispaired base recognition protein, also recognizes Holliday junctions in DNA. J. Mol. Biol. **265:**289–301.
3. **Aquilina, G., and M. Bignami.** 2001. Mismatch repair in correction of replication errors and processing of DNA damage. J. Cell. Physiol. **187:**145–154.
4. **Aquilina, G., A. Zijno, N. Moscufo, E. Dogliotti, and M. Bignami.** 1989. Tolerance to methylnitrosourea-induced DNA damage is associated with 6-thioguanine resistance in CHO cells. Carcinogenesis **10:**1219–1223.
5. **Aquilina, G., G. Frosina, A. Zijno, A. Di Muccio, E. Dogliotti, A. Abbondandolo, and M. Bignami.** 1988. Isolation of clones displaying enhanced resistance to methylating agents in O6-methylguanine-DNA methyltransferase-proficient CHO cells. Carcinogenesis **9:**1217–1222.
6. **Bannister, L. A., B. C. Waldman, and A. S. Waldman.** 2004. Modulation of error-prone double-strand break repair in mammalian chromosomes by DNA mismatch repair protein Mlh1. DNA Repair **3:**465–474.
7. **Bellacosa, A.** 2001. Functional interactions and signaling properties of mammalian DNA mismatch repair proteins. Cell Death Differ. **8:**1076–1092.
8. **Bernstein, C., H. Bernstein, C. M. Payne, and H. Garewal.** 2002. DNA repair/pro-apoptotic dual-role proteins in five major DNA repair pathways: fail-safe protection against carcinogenesis. Mutat. Res. **511:**145–178.
9. **Bertrand, P., Y. Saintigny, and B. S. Lopez.** 2004. p53's double life: transactivation-independent repression of homologous recombination. Trends Genet. **20:**235–243.
10. **Brown, K. D., A. Rathi, R. Kamath, D. I. Beardsley, Q. Zhan, J. L. Mannino, and R. Baskaran.** 2003. The mismatch repair system is required for S-phase checkpoint activation. Nat. Genet. **33:**80–84.
11. **Buermeyer, A. B., S. M. Deschenes, S. M. Baker, and R. M. Liskay.** 1999. Mammalian DNA mismatch repair. Annu. Rev. Genet. **33:**533–564.
12. **Chambers, S. R., N. Hunter, E. J. Louis, and R. H. Borts.** 1996. The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. Mol. Cell. Biol. **16:**6110–6120.
13. **Chen, W., and S. Jinks-Robertson.** 1998. Mismatch repair proteins regulate heteroduplex formation during mitotic recombination in yeast. Mol. Cell. Biol. **18:**6525–6537.
14. **Chen, W., and S. Jinks-Robertson.** 1999. The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. Genetics **151:**1299–1313.
15. **Coic, E., L. Gluck, and F. Fabre.** 2000. Evidence for short-patch mismatch repair in Saccharomyces cerevisiae. EMBO J. **19:**3408–3417.
16. **Datta, A., M. Hendrix, M. Lipsitch, and S. Jinks-Robertson.** 1997. Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. Proc. Natl. Acad. Sci. USA **94:**9757–9762.
17. **De La Torre, C., J. Pincheira, and J. F. Lopez-Saez.** 2003. Human syndromes with genomic instability and multiprotein machines that repair DNA double-strand breaks. Histol. Histopathol. **18:**225–243.
18. **Elliot, B., and M. Jasin.** 2001. Repair of double-strand breaks by homologous recombination in mismatch repair-defective mammalian cells. Mol. Cell. Biol. **21:**2671–2682.
19. **Emmanuel, E., E. Yehuda, C. Melamed-Bessudo, N. Avivi-Ragolsky, and A. A. Levy.** 2006. The role of AtMSH2 in homologous recombination in Arabidopsis thaliana. EMBO Rep. **7:**100–105.
20. **Evans, E., and E. Alani.** 2000. Roles for mismatch repair factors in regulating genetic recombination. Mol. Cell. Biol. **20:**7839–7844.
21. **Evans, E., N. Sugawara, J. Haber, and E. Alani.** 2000. The Saccharomyces cerevisiae MSH2 mismatch repair protein localizes to recombination intermediates in vivo. Mol. Cell **5:**789–799.
22. **Fedier, A., and D. Fink.** 2004. Mutations in DNA mismatch repair genes: implications for DNA damage signaling and drug sensitivity. Int. J. Oncol. **24:**1039–1047.
23. **Fleck, O., E. Lehmann, P. Schar, and J. Kohli.** 1999. Involvement of nucleotide-excision repair in msh2 pms1-independent mismatch repair. Nat. Genet. **21:**314–317.
24. **Franchitto, A., P. Pichierri, R. Piergentili, M. Crescenzi, M. Bignami, and F. Palitti.** 2003. The mammalian mismatch repair protein MSH2 is required for correct MRE11 and RAD51 relocalization and for efficient cell cycle arrest induced by ionizing radiation in G2 phase. Oncogene **22:**2110–2120.
25. **Goldfarb, T., and E. Alani.** 2005. Distinct roles for the Saccharomyces cerevisiae mismatch repair proteins in heteroduplex rejection, mismatch repair and nonhomologous tail removal. Genetics **169:**563–574.
26. **Harfe, B. D., and S. Jinks-Robertson.** 2000. DNA mismatch repair and genetic instability. Annu. Rev. Genet. **34:**359–399.
27. **Heinen, C. D., C. Schmutte, and R. Fishel.** 2002. DNA repair and tumorigenesis: lessons from hereditary cancer syndromes. Cancer Biol. Ther. **1:**477–485.
28. **Jackson, S. P., and P. A. Jeggo.** 1995. DNA double-strand break repair and V(D)J recombination: involvement of DNA-PK. Trends Biochem. Sci. **20:**412–415.
29. **Jhanwar-Uniyal, M.** 2003. BRCA1 in cancer, cell cycle and genomic stability. Front. Biosci. **8:**s1107–s1117.
30. **Jiricny, J.** 2006. The multifaceted mismatch-repair system. Nat. Rev. Mol. Cell Biol. **7:**335–446.
31. **Kijas, A. W., B. Studamire, and E. Alani.** 2003. Msh2 separation of function mutations confer defects in the initiation steps of mismatch repair. J. Mol. Biol. **331:**123–138.
32. **Lea, D. E., and C. A. Coulson.** 1949. The distribution of the number of mutants in bacterial populations. J. Genet. **49:**264–285.
33. **Li, G. M.** 2003. DNA mismatch repair and cancer. Front. Biosci. **8:**d997–d1017.
34. **Liang, F., M. Han, P. J. Romanienko, and M. Jasin.** 1998. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc. Natl. Acad. Sci. USA **95:**5172–5177.
35. **Lieber, M. R., Y. Ma, U. Pannicke, and K. Schwarz.** 2003. Mechanism and regulation of human non-homologous DNA end-joining. Nat. Rev. Mol. Cell Biol. **4:**712–720.
36. **Lukacsovich, T., and A. S. Waldman.** 1999. Suppression of intrachromosomal gene conversion in mammalian cells by small degrees of sequence divergence. Genetics **151:**1559–1568.
37. **Lukacsovich, T., D. Yang, and A. S. Waldman.** 1994. Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-*Sce*I. Nucleic Acids Res. **22:**5649–5657.
38. **Mitchell, R. J., S. M. Farrington, M. G. Dunlop, and H. Campbell.** 2002. Mismatch repair genes hMLH1 and hMSH2 and colorectal cancer: a HuGE review. Am. J. Epidemiol. **156:**885–902.
39. **Muheim-Lenz, R., T. Buterin, G. Marra, and H. Naegeli.** 2004. Short-patch correction of C/C mismatches in human cells. Nucleic Acids Res. **32:**6696–6705.
40. **Muller, A., M. Korabiowska, and U. Brinck.** 2003. DNA-mismatch repair and hereditary nonpolyposis colorectal cancer syndrome. In Vivo **17:**55–59.
41. **Myung, K., A. Datta, C. Chen, and R. D. Kolodner.** 2001. SGS1, the Saccharomyces cerevisiae homologue of BLM and WRN, suppresses genome instability and homeologous recombination. Nat. Genet. **27:**113–116.
42. **Nicholson, A., M. Hendrix, S. Jinks-Robertson, and G. F. Crouse.** 2000. Regulation of mitotic homeologous recombination in yeast. Functions of mismatch repair and nucleotide excision repair genes. Genetics **154:**133–146.
43. **Oda, S., O. Humbert, S. Fiumicino, M. Bignami, and P. Karran.** 2000. Efficient repair of A/C mismatches in mouse cells deficient in long-patch mismatch repair. EMBO J. **19:**1711–1718.
44. **Peltomaki, P.** 2001. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum. Mol. Genet. **10:**735–740.
45. **Radford, S. J., M. M. Sabourin, S. McMahan, and J. Sekelsky.** 2007. Meiotic recombination in Drosophila Msh6 mutants yields discontinuous gene conversion tracts. Genetics **176:**53–62.
46. **Rayssiguier, C., D. S. Thaler, and M. Radman.** 1989. The barrier to recombination between Escherichia coli and Salmonella typhimurium is disrupted in mismatch repair mutants. Nature **342:**396–401.
47. **Richardson, C., and M. Jasin.** 2000. Coupled homologous and nonhomolo-

gous repair of a double-strand break preserves genomic integrity in mammalian cells. Mol. Cell. Biol. **20:**9068–9075.
48. **Saparbaev, M., L. Prakash, and S. Prakash.** 1996. Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination in Saccharomyces cerevisiae. Genetics **142:**727–736.
49. **Schofield, M. J., and P. Hsieh.** 2003. DNA mismatch repair: molecular mechanisms and biological function. Annu. Rev. Microbiol. **57:**579–608.
50. **Smith, J. A., B. C. Waldman, and A. S. Waldman.** 2005. A role for mismatch repair protein Msh2 in error-prone double-strand break repair in mammalian chromosomes. Genetics **170:**355–363.
51. **Sonoda, E., H. Hochegger, A. Saberi, Y. Taniguchi, and S. Takeda.** 2006. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. DNA Repair **5:**1021–1029.
52. **Spell, R. M., and S. Jinks-Robertson.** 2004. Examination of the roles of Sgs1 and Srs2 helicases in the enforcement of recombination fidelity in Saccharomyces cerevisiae. Genetics **168:**1855–1865.
53. **Sugawara, N., F. Paques, M. P. Colaiacovo, and J. E. Haber.** 1997. Role of S. cerevisiae MSH2 and MSH3 repair proteins in double-strand break repair-induced recombination. Proc. Natl. Acad. Sci. USA **94:**9214–9219.
54. **Sugawara, N., T. Goldfarb, B. Studamire, E. Alani, and J. E. Haber.** 2004. Heteroduplex rejection during single-strand annealing requires Sgs1 helicase and mismatch repair proteins Msh2 and Msh6 but not Pms1. Proc. Natl. Acad. Sci. USA **101:**9315–9320.
55. **Surtees, J. A., J. L. Argueso, and E. Alani.** 2004. Mismatch repair proteins: key regulators of genetic recombination. Cytogenet. Genome Res. **107:**146–159.
56. **Trouiller, B., D. G. Schaefer, F. Charlot, and F. Nogue.** 2006. MSH2 is essential for the preservation of genome integrity and prevents homeologous recombination in the moss Physcomitrella patens. Nucleic Acids Res. **34:**232–242.
57. **Valerie, K., and L.F. Povirk.** 2003. Regulation and mechanisms of mammalian double-strand break repair. Oncogene **22:**5792–5812.
58. **Villemure, J.-F., C. Abaji, I. Cousineau, and A. Belmaaza.** 2003. MSH2-deficient human cells exhibit a defect in the accurate termination of homology-directed repair of DNA double-strand breaks. Cancer Res. **63:**3334–3339.
59. **Wagner, M. J., J. A. Sharp, and W. C. Summers.** 1981. Nucleotide sequence of the thymidine kinase of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **78:**1441–1445.
60. **Waldman, A. S., and R. M. Liskay.** 1987. Differential effects of base-pair mismatch on intrachromosomal versus extrachromosomal recombination in mouse cells. Proc. Natl. Acad. Sci. USA **84:**5340–5344.
61. **Waldman, A. S., and R. M. Liskay.** 1988. Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology. Mol. Cell. Biol. **8:**5350–5357.
62. **Wang, Y., D. Cortez, P. Yazdi, N. Neff, S. J. Elledge, and J. Qin.** 2000. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. Genes Dev. **14:**927–939.
63. **Watson, P., and H. T. Lynch.** 2001. Cancer risk in mismatch repair gene mutation carriers. Fam. Cancer **1:**57–60.
64. **Wei, K., R. Kucherlapati, and W. Edelmann.** 2002. Mouse models for human DNA mismatch-repair gene defects. Trends Mol. Med. **8:**346–353.
65. **Worth, L., S. Clark, M. Radman, and P. Modrich.** 1994. Mismatch repair proteins MutS and MutL inhibit RecA catalyzed strand transfer between diverged DNAs. Proc. Natl. Acad. Sci. USA **91:**3238–3241.
66. **Yang, D., and A. S. Waldman.** 1997. Fine-resolution analysis of products of intrachromosomal homeologous recombination in mammalian cells. Mol. Cell. Biol. **17:**3614–3628.
67. **Yang, D., E. B. Goldsmith, Y. Lin, B. C. Waldman, V. Kaza, and A. S. Waldman.** 2006. Genetic exchange between homeologous sequences in mammalian chromosomes is averted by local homology requirements for initiation and resolution of recombination. Genetics **174:**135–144.
68. **Yang, Q., R. Zhang, X. W. Wang, S. P. Linke, S. Sengupta, I. D. Hickson, G. Pedrazzi, C. Perrera, I. Stagljar, S. J. Littman, P. Modrich, and C. C. Harris.** 2004. The mismatch DNA repair heterodimer, hMSH2/6, regulates BLM helicase. Oncogene **23:**3749–3756.