Published online February 2, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 2 ***749–756***
DOI: 10.1093/nar/gkh244

Structure of a palindromic amplicon junction implicates microhomology-mediated end joining as a mechanism of sister chromatid fusion during gene amplification

Yukiko Okuno, Peter J. Hahn[1] and David M. Gilbert*

Department of Biochemistry and Molecular Biology and [1]Department of Radiation Oncology, SUNY Upstate Medical University, 750 East Adams Street, Syracuse, NY 13210, USA

Received October 31, 2003; Revised December 15, 2003; Accepted December 29, 2003 DDBJ/EMBL/GenBank accession no. AB124804

ABSTRACT

Amplification of the copy number of oncogenes is frequently associated with tumor progression. Often, the amplified DNA consists of large (tens to hundreds of kilobases) 'head-to-head' inverted repeat palindromes (amplicons). Several mechanisms have been proposed to explain palindrome formation but their relative contributions in nature have been difficult to assess without precise knowledge of the sequences involved at the junction of natural amplicons. Here, we have sequenced one such junction and compared this sequence to the un-rearranged structure, allowing us to pinpoint the site of sister chromatid fusion. Our results support a novel model, consistent with all described sister chromatid fusions, in which sister chromatid fusion is initiated by microhomology-mediated end joining of double strand breaks.

INTRODUCTION

Failure to properly repair DNA double-strand breaks (DSBs) is a major cause of genomic instability (1–4). Sister chromatid fusion, first described by McClintock (5) is a significant component of mis-repaired DNA damage associated with genomic instability. During anaphase, di-centric fused sister chromatids have been observed microscopically to be pulled in opposite directions, directly leading to re-breakage. In the next cell cycle, the broken chromosome is replicated, and the sister chromatids refuse. This cycle continues until telomeres are restored to the chromosomes (6). The clonal descendants display a genomic instability phenotype (7).

Sister chromatid fusion can result in gene amplification, providing a selective advantage for clonal descendants (8). For example, Chinese hamster cells growing in the presence of toxic levels of methotrexate (MTX) suffer DNA damage because the MTX inhibits the dihydrofolate reductase (DHFR) gene resulting in imbalances in nucleotide pools (9–12). Occasionally, breakage of hamster chromosome 2, containing the DHFR gene, leads to multiple copies of the DHFR gene and survival of the DHFR over-expressing clones. Trask and Hamlin (8) have demonstrated that the hallmark of this form of gene amplification is a homogeneous staining region (HSR) expanding distally from the original location of the unamplified gene.

Sister chromatid fusion is thought to result from inappropriate DNA DSB repair, but the repair mechanism has not been clear. Studies in both yeast (13,14) and mammalian cells (15) suggest that DSBs arising near inverted repeats can lead to sister chromatid fusions through intra-molecular homologous recombination. In mammalian cells, integration of plasmids containing inverted repeats and an inducible DNA DSB site beyond the inverted repeats led to sister chromatid fusion following induction of breakage near the repeats (15). The proposed mechanism is that exo-nucleolytic digestion of one strand exposes the other for homologous recombination and, rather than initiating recombination with the homologous chromosome, the exposed single strand folds back via the inverted repeats (Fig. 1). Subsequent DNA replication would produce fused sister chromatids, consisting of a palindromic di-centric chromosome with an asymmetric region flanked by inverted repeats at the junction. This is an attractive model, but there is no evidence that such inverted repeats exist at natural sister chromatid fusion sites (below) and another study has suggested that they are not necessary. Murnane and colleagues (6,16,17) studied telomere loss by inserting the Herpes simplex thymidine kinase (TK) gene near telomeres. Selection for loss of the TK marker gene suggested that telomere loss is common and generally leads to sister chromatid fusion cycles lasting until new telomeres are added (6). However, no inverted repeats or significant regions of homology were found at the recombination sites, leading these authors to conclude that sister chromatid fusion resulted from non-homologous end joining (NHEJ).

Unfortunately, very little data concerning natural sister chromatid fusions has been collected, mainly because identifying the exact breakpoints in large chromosomes is prohibitively labor intensive. Amplified DNAs ('amplicons')

*To whom correspondence should be addressed. Tel: +1 315 464 8723; Fax: +1 315 464 8750; Email: gilbertd@upstate.edu
Present address:
Yukiko Okuno, Division of Biological Science, Graduate School of Science, Nagoya University, Huro-Cho, Chikusa-ku, Nagoya, Aichi-pref. 464-8602 Japan

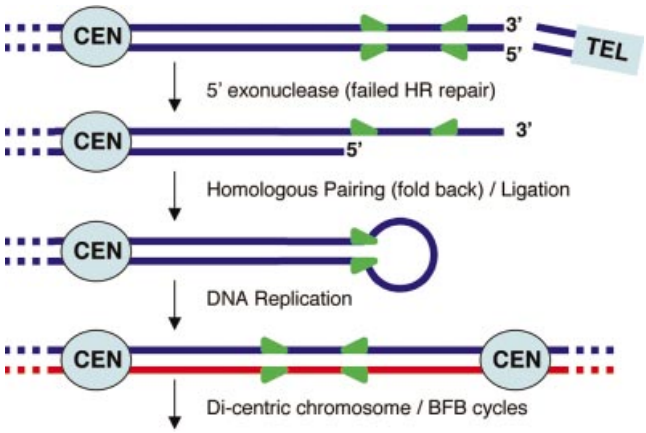


Figure 1. Intramolecular homologous recombination model for sister chromatid fusion. After a DSB, attempts to undergo homologous repair will result in exonucleolytic digestion of the 5′ strand of the centromeric fragment to facilitate single strand invasion by the 3′ strand. If this strand fails to find the homologous chromosome, intra-strand pairing can occur to generate a hairpin structure. After replication, this structure will be converted to a di-centric chromosome, which leads to gene amplification through BFB cycles (adapted from 15).

in drug resistant cells have been studied intensively (8,18–20). Amplicons can be as large as 10 Mb and are frequently organized as either inverted or tandem repeats located on expanded chromosomal arrays. The inverted repeats are thought to arise from sister chromatid fusion. Many of these amplicons have been cloned as part of an effort to understand the mechanism of their formation. However, only two such amplicon junctions have been sequenced and compared to the unamplified chromosomal sequence (21,22), but no regions of extensive homology or inverted repeat sequences were identified. Therefore, these natural sister chromatid fusions are more similar to the telomere fusion structures described by Murnane and colleagues (17) than the inverted repeat/homologous recombination model proposed by Tanaka *et al.* (15).

Recently, microhomology-mediated end joining was reported to be involved in amplification of the c-myc gene following non-sister chromatid fusion. Mice deficient in p53 and either Ku (23) or XRCC4 (24) develop lymphomas. RAG-mediated DSBs are mis-repaired via microhomology-mediated end joining fusion of c-myc and IgH leading to a di-centric chromosome and amplification similar to the process described above for sister chromatid fusion. Both Ku proteins and XRCC-4 are required for NHEJ. NHEJ is the primary DSB repair pathway in mammalian cells, but these results suggest that microhomology-mediated end joining uses a repair system other than NHEJ.

The head-to-head DHFR amplicon in a Chinese Hamster Ovary (CHO) cell line (CHOC 400) has been used as a model amplicon for many studies of gene amplification and its structure generally supports the sister chromatid ('break–fusion–bridge', BFB) model of gene amplification (8). CHOC 400 cells contain ~500-fold amplification of the DHFR gene (25). The amplicons are arranged in either head-to-head or head-to-tail orientation. Two types (Type I and Type II) of head-to-head variety are described by the Hamlin group, with Type II constituting 75% of the amplicons in CHOC 400 (26). Since Type II amplicons contain Type I junctions, Type I is believed to have preceded Type II (27). However, both are found in the earliest analyzed sub-clones preceding the derivation of CHOC 400, indicating that these break points arose very early during the amplification process. To determine whether this amplicon could have resulted from a chromosome break near an inverted repeat, we determined the DNA sequence near the more prevalent Type II junction of one of the primary head-to-head DHFR amplicons in CHOC 400, and compared it to the unamplified chromosomal sequence in CHO cells. We report here that the sequence at the breakpoint indicates that the head-to-head amplicons were the result of the joining of two sister chromatids via a 2 bp overlap (microhomology) at sites separated by 4 kb. We propose that microhomology-mediated end joining can account for this and all other described sister chromatid fusions.

MATERIALS AND METHODS

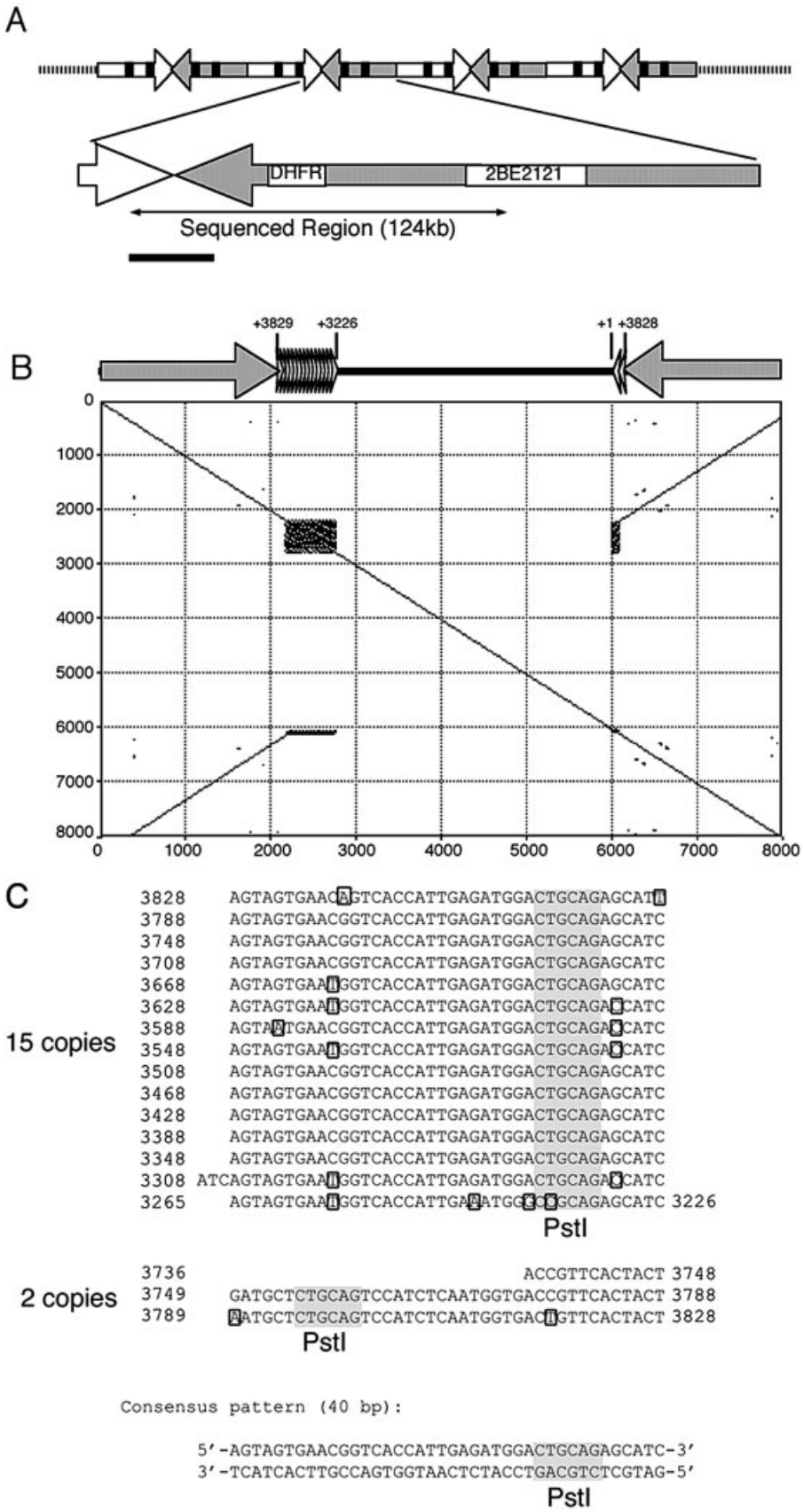
Cell culture, DNA isolation and Southern blotting

CHOC400, CHOAA8 and CHOK1 were all maintained with DMEM + 5% FBS, as previously described (28). CHOC 400 contains 1000 copies of the DHFR locus present as stable HSRs, generated by stepwise selection in MTX up to 400 μg/ml, as described in several reports, which are reviewed by Hamlin *et al.* (25). Genomic DNA was isolated as described (29). For Southern hybridization, DNA was cleaved with PstI, subjected to electrophoresis in a 1% agarose gel containing TBE (Tris-borate EDTA), transferred to Hybond N+ (Amersham) nylon membranes and hybridized with purified DNA fragments, which were labeled with ^{32}P using a random-priming kit (Ambion).

Sequence analysis and PCR

The sequencing of cosmid H2 (26), that includes the head-to-head junction of the CHOC 400 Type II amplicon was part of a project to sequence 120 kb of the CHO DHFR locus, and will be reported elsewhere (Y.O. and D.M.G., in preparation). To visualize the junction point, Pustell (dot) matrix analysis was performed using the software MacVector (Oxford Press) with

Figure 2. The presence of an asymmetric region flanked by unusual inverted repeats at the amplicon junction. The complete DNA sequence of 124 kb of the Type II head-to-head DHFR amplicon in CHOC 400 cells, including the amplicon junction, has been determined (Y.O. and D.M.G., in preparation). (**A**) Organization of the amplicon junction in CHOC 400 cells (27). Black bar denotes the region analyzed in (B). (**B**) Sequences containing the junction connecting the Type II amplicons, which contained a 4 kb asymmetric region, were subjected to Pustell matrix analysis, revealing that the asymmetric region is flanked on one side with 15 copies of a 40 bp repeat, each of which contains a Pst1 cleavage site, and on the other side with two copies of the identical 40 bp repeat in an inverted orientation. (**C**) Sequence of the 15mer and dimer inverted repeats. The shaded region indicates the Pst1 cleavage sites. Boxed nucleotides vary from the consensus sequence. Note that the dimer contains the same polymorphic nucleotides as the first two repeats of the 15mer.


A
DHFR
2BE2121
Sequenced Region (124kb)
B
+3829
+3226
+1
+3828
0
1000
2000
3000
4000
5000
6000
7000
8000
C
15 copies
3828 AGTAGTGAACAGTCACCATTGAGATGGACTGCAGAGCATA
3788 AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3748 AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3708 AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3668 AGTAGTGAAAGGTCACCATTGAGATGGACTGCAGAGCATC
3628 AGTAGTGAAAGGTCACCATTGAGATGGACTGCAGAGCATC
3588 AGTAATGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3548 AGTAGTGAAAGGTCACCATTGAGATGGACTGCAGAGCATC
3508 AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3468 AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3428 AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3388 AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3348 AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC
3308 ATCAGTAGTGAAAGGTCACCATTGAGATGGACTGCAGAGCATC
3265 AGTAGTGAAAGGTCACCATTGAAATGGCGCGCAGAGCATC 3226
PstI
2 copies
3736 ACCGTTCACTACT 3748
3749 GATGCTCTGCAGTCCATCTCAATGGTGACCGTTCACTACT 3788
3789 AATGCTCTGCAGTCCATCTCAATGGTGACAGTTCACTACT 3828
PstI
Consensus pattern (40 bp):
5′-AGTAGTGAACGGTCACCATTGAGATGGACTGCAGAGCATC-3′
3′-TCATCACTTGCCAGTGGTAACTCTACCTGACGTCTCGTAG-5′
PstI

the following conditions: Window size: 30, Min. % Score: 65, Hash value: 6, Jump: 1). Tandem Repeats Finder (30) was used to find a consensus sequence for the PstI Box. The positions of hamster msh gene exons were predicted by comparison of mouse cDNA (GenBank accession no. NM-010829 and M80360), mouse genomic sequence (GenBank accession no. NW_000085), and human genomic sequence (GenBank accession no. NT_006713). Hamster exon numbers were named according to that of mouse and human. For the sequence between exons 7 and 8 of MSH3, PCR products indicated in Figure 4 were sequenced and this sequence has been submitted to GenBank (accession no. AB124804). Annealing temperatures for primers are described in the figure legends. Primer sequences are as follows: primer 1, 5′-GAGTTCGTAACCACTGAGCC-3′; primer 2: 5′-TTCAAATGAGACCCAAACTC-3′; primer 3, 5′-AGGCTGTTACTTCTGTGC-3′; exon 8 primer, 5′-GGAAGCAGAGTCCTGGAA-3′.

RESULTS

The primary DNA sequence of the junction region of the DHFR type II head-to-head amplicon in CHOC 400 cells, previously localized by restriction mapping (18,26,27,31), revealed an ~4 kb asymmetric region bounded by 40 bp repeats (Fig. 2). Fifteen tandem copies of the 40 bp repeats are on one side, and two tandem copies in an inverted orientation are located on the other side. This was an intriguing finding, since this sequence organization is exactly as predicted by the intramolecular recombination model (15), if the break occurred just outside the repeats (Fig. 1). However, since the sequence shown in Figure 2 was derived from a cosmid, it was necessary to determine whether this inverted repeat organization is also present at the junctions of Type II amplicons in the CHOC 400 genome and whether it is present in the wild type CHO genome, prior to amplification. First, PCR primers (primers 1 and 2, Fig. 3A) were designed to detect an 808 bp fragment encompassing the larger of the two inverted repeats (15 copies of 40 bp). Results confirmed the presence of an amplification product of the appropriate size in both CHOC 400 (Fig. 3B) and in two independent un-amplified CHO cell lines (Fig. 3C). The presence of a PstI cleavage site within each of the 40 bp repeats allowed for the unambiguous identification of this PCR amplification product without DNA sequencing by demonstrating that it is cleaved by PstI into a prominent 40 bp band plus two fragments corresponding to the sizes of the two end PstI bands (Fig. 3D). Therefore, the larger of the two inverted repeats is present in both CHOC 400 and the unamplified CHO genome.

Due to the inverted repeat organization of the amplification junction, primer 1 alone can be used to amplify the entire 4 kb asymmetric region. Under these conditions, a 4 kb band can be seen following amplification with CHOC 400 genomic DNA, but not with CHO DNA (Fig. 3E). This is expected since the inverted repeat structure of sequences included in primer 1 results from the amplification process and so should not be repeated in the unamplified genome. This result also provides evidence that the PCR-amplified bands are not due to contamination by cosmid DNA but are only present in appropriate genomic DNA preparations. Further analysis of the smaller of the two inverted repeats (two copies of 40 bp; primers 1 and 3) revealed the presence of an appropriate sized PCR fragment in CHOC 400 genomic DNA (Fig. 3F). Therefore, both the 15mer and the smaller dimer of 40 bp repeats are present in CHOC 400 cells, and at least the larger of the two is present in the unamplified genome. This strongly suggests that the fusion site is directly on one side or the other of the inverted repeat dimer (between primers 1 and 3), but does not reveal whether or not an inverted repeat was present prior to chromosome breakage, as would be predicted by the model shown in Figure 1. It is possible that the fusion took place adjacent to primer 3; the smaller inverted repeat dimer could have been a fortuitous consequence of sister chromatid fusion.

To determine whether the smaller repeat dimer is present in the unamplified genome, and in an inverted orientation to the larger one, we exploited the fact that the amplicon head-to-head junction is in the highly conserved msh3 gene. The sequence is known for both mouse and human, but not for hamster. The 4 kb asymmetric region contains sequences highly similar to the mouse exon 7 and human exon 7 of the mismatch repair 3 gene. Therefore, we selected highly conserved sequences in the predicted downstream exon 8 of the msh3 gene to use as PCR primers paired with primers from the asymmetric sequence. One of the exon 8 primers, when combined with primer 3, amplified a band with CHO genomic DNA. The sequence of this PCR product was compared to the CHOC 400 sequence (Fig. 4). The CHO genomic sequence did not contain the 40 bp inverted repeat dimer. Rather, the CHOC 400 DNA sequence diverges from the CHO DNA sequence immediately (the first nucleotide) before the 40 bp repeat dimer, indicating that the dimer is part of the larger inverted repeat structure, not the asymmetric region. Therefore, the smaller 40 bp inverted repeat dimer was not present prior to sister chromatid fusion and so could not have been part of a putative fold-back structure. Note that there is a 2 bp ambiguity in the precise nucleotide position of the fusion site (shaded AC in Fig. 4), where the sequence in the contiguous unamplified DNA is the same as the repeat, consistent with microhomology-mediated end joining (see Discussion).

Since amplicons can contain DNA that is not derived from the original genomic locus (complicons) (24), it was important to confirm by Southern blotting that the asymmetric region of the head-to-head junction was present in the wild type CHO DNA, but that the smaller inverted repeat (dimer) was not. CHOC 400 and CHO DNA were both cleaved with PstI, and probed with labeled plasmid DNA containing the junction region. PstI cleaves the CHOC 400 Type II amplicon junction into four hybridizing bands, one of which (513 bp) is created by the presence of the 40 bp inverted repeat dimer. If the 40 bp dimer is not present in CHO then the 513 bp fragment will not be observed but a larger fragment will be present instead. The complete sequence of wild type CHO DNA between primer exon 8 and primer 3 (not shown) predicts that this would be a 1244 bp Pst1 fragment. Results (Fig. 5) revealed that the 513 bp band was present in CHOC 400 but not CHO DNA, while a larger band, of the size predicted by the wild type sequence (1244 bp), was present in both DNA samples. This larger band was present in molar equivalents to the other bands in CHO, but was significantly less represented in CHOC 400. Since CHOC 400 contains other amplicons (25% of total) that

are not broken in this region, the 1244 bp band in CHOC 400 represents a combination of the remaining two unamplified single copy and ~250 amplicons that are not Type II, while the 513 bp band represents the fraction (~75% or 750 copies) of amplicons organized into Type II. This experiment confirms that the asymmetric 4 kb sequence at the type II amplicon head-to-head junction is derived from the msh3 gene upstream of DHFR in CHO DNA and that it does not include an inverted

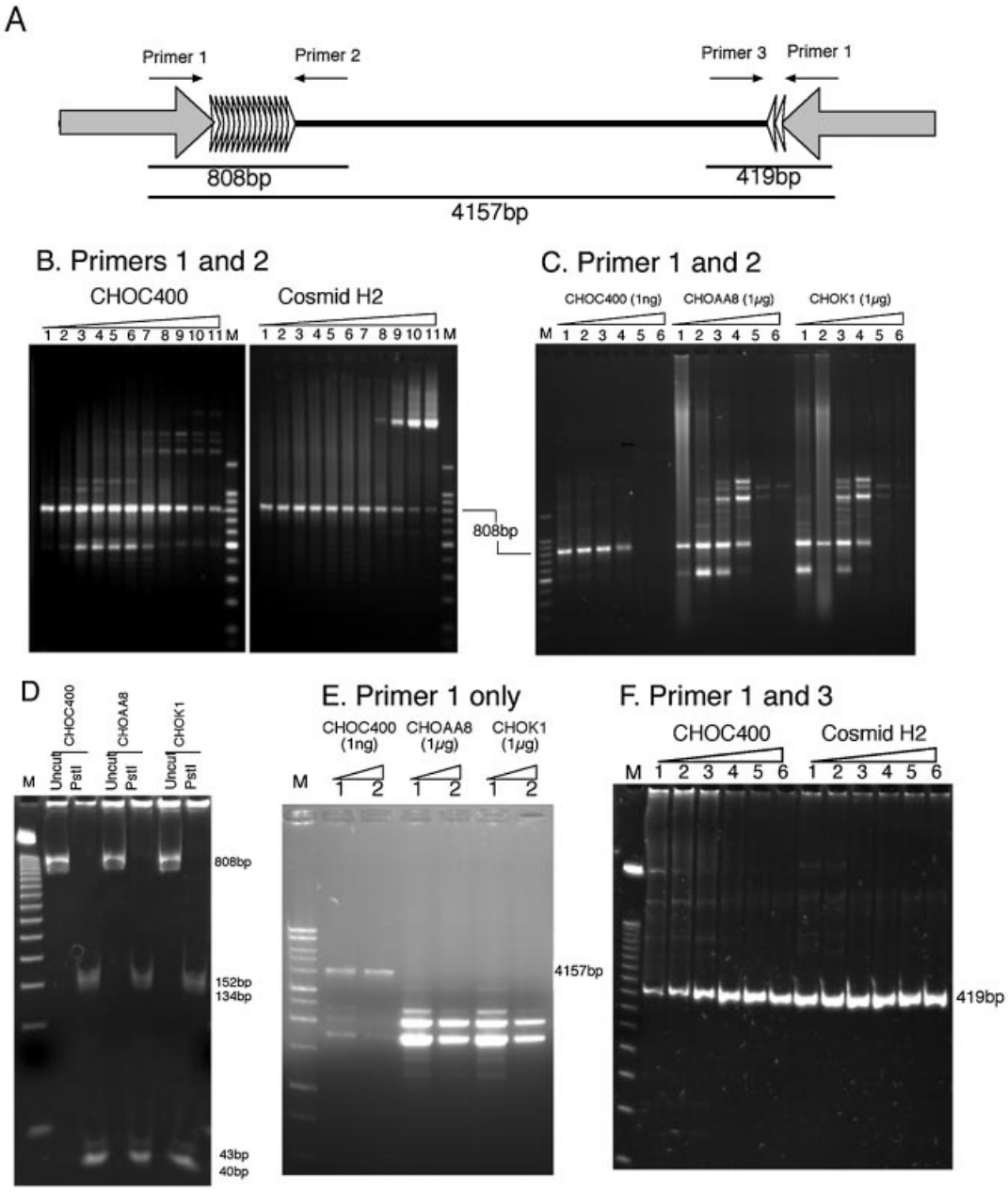


Figure 3. PCR amplification of the inverted repeats in CHOC 400 and CHO. (**A**) Primer pairs 1 and 2 were selected to flank the 15mer Pst1 repeat. These primers are expected to amplify an 808 bp fragment if the repeat is present, while primer 1 alone can amplify a 4157 bp fragment containing the entire asymmetric region, albeit much less efficiently due to its larger size. Primers 1 and 3 amplify a 419 bp fragment containing the dimer repeat. (**B**) Primers 1 and 2 amplify the 808 bp fragment in both CHOC 400 and the cosmid from which the sequenced clone was derived. Shown are results with increasing annealing temperatures: lane 1, 46.2°C; lane 2, 47.2°C; lane 3, 48.8°C; lane 4, 50.9°C; lane 5, 53.2°C; lane 6, 55.7°C; lane 7, 58.2°C; lane 8, 60.4°C; lane 9, 62.3°C; lane 10, 63.8°C; lane 11, 64.5°C. With cosmid H2 as a template, increasing annealing temperature produced a larger band consistent with the 4157 bp complete asymmetric region (lanes 9–11). (**C**) Primers 1 and 2 also amplify the 808 bp fragment in two independent CHO cell lines containing unamplified DHFR (annealing temperatures: lane 1, 46.2°C; lane 2, 48.8°C; lane 3, 53.2°C; lane 4, 58.2°C; lane 5, 62.3°C; lane 6, 64.5°C). (**D**) Pst1 cleavage of the amplified 808 bp fragment into 40 bp multimers confirms the presence of the repeat within these amplified fragments, as well as the predicted 152 and 134 bp flanking pieces. (**E**) Without competition with primer 2, primer 1 alone readily amplifies a 4 kb fragment with CHOC 400, but not CHO DNA. Annealing temperatures: lane 1, 62.3°C; lane 2, 64.6°C. (**F**) Primers 1 and 3 amplify the predicted 419 bp fragment from both CHOC 400 and the cosmid. Since sequences to the right of the repeat dimer diverge between CHOC 400 and the unamplified structure, we could not test the presence of the dimer using these primer sets. Annealing temperatures: lane 1, 46.2°C; lane 2, 48.8°C; lane 3, 53.2°C; lane 4, 58.2°C; lane 5, 62.3°C; lane 6, 64.5°C. For each panel, marker (M) is as follows: (B and C) 100 bp ladder, (D and F) 50 bp ladder, and (E) 1 kb ladder.

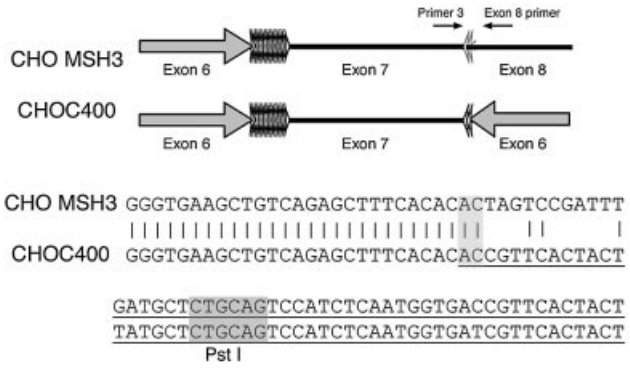


Figure 4. DNA sequence of the breakpoints. To determine whether the smaller repeat was present in CHO genomic DNA, Primer 3 and a primer based on conserved sequences from exon 8 of the msh 3 gene in human and mouse were used to PCR amplify a CHO genomic fragment. The sequence of this fragment was identical to the corresponding CHOC 400 sequence up to the point at which the CHOC 400 DNA reached the repeat dimer, and diverged thereafter. Note the 2 bp (AC) overlap (light gray box), at the point of divergence. This is the junction between the two sister chromatids, which forms a region of microhomology between this region and the repeat region 4 kb upstream. The sequence of the CHOC 400-specific dimer repeat is underlined, with the Pst1 sites more deeply shaded. Hence, the unrearranged sequence does not contain the Pst1 repeats. Only the DNA sequence in the vicinity of the breakpoint is shown.

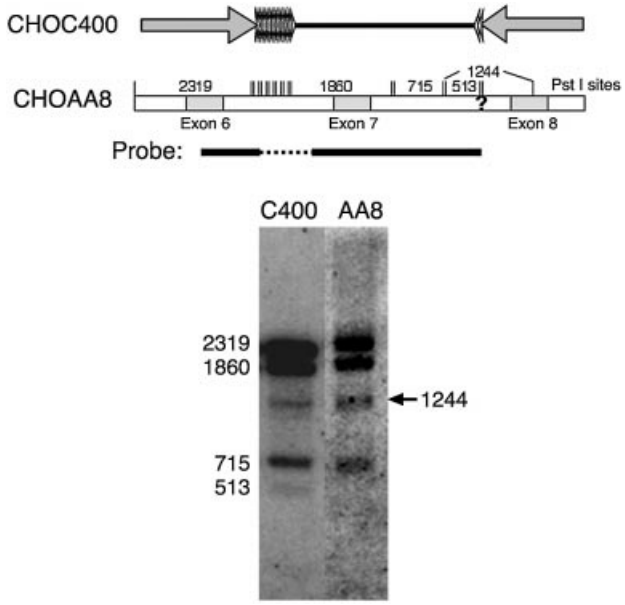


Figure 5. Southern hybridization of CHOC 400 and CHO genomic DNA in the region of the sister chromatid fusion. CHOC 400 and CHO genomic DNA were digested with Pst1 and subjected to Southern hybridization, using the probe indicated by the dark lines. The probe is derived from CHOC 400 DNA and spans the junction region; hence, the dashed interruption indicates where the probe will not hybridize in the unrearranged state. The Pst1 cleavage sites in the region are shown in the map above the hybridized membrane. The question mark indicates the Pst1 cleavage sites that would be present if the second 40 bp repeat is present in the CHO genomic DNA. The complete sequence of both CHO AA8 and CHO K1 DNA between primer Exon 8 and primer 3 (not shown) predicts a 1244bp Pst1 fragment for wild-type DNA, which is consistent with what is found by Southern hybridization. Pst1 fragments smaller than 100 bp are not visible in this experiment.

repeat structure. Note that the probe in this experiment contains the Pst1 repeat, and no other bands were noted, indicating that this unusual repeat is unique in the hamster genome.

DISCUSSION

Two very different models for sister chromatid fusion have been presented based on the results of independent studies with ectopically integrated plasmids containing engineered sites at which a DSB could be induced. In one case (15), inverted repeat structures contained within these plasmids promoted the formation of large palindromes, leading to a model in which such structures present near a break point would lead to intramolecular homologous recombination through a fold-back hairpin structure which after replication would produce di-centric fused sister chromatids (Fig. 1). In the second study (17), sister chromatid fusions were proposed to be the result of extensive degradation (kilobases) and NHEJ. The purpose of the study described here was to determine whether a naturally occurring amplicon consisting of large palindromes would support either of these models. Comparison of the cloned sequence from one such amplicon junction with the unamplified chromosomal DNA sequence revealed that the sister chromatids were joined at a site of 2 bp overlap, with no significant region of inverted repeat homology present prior to fusion. For the reasons elaborated below, we propose that sister chromatid fusion is initiated by microhomology-mediated end joining of DSBs.

The intramolecular homologous recombination model of palindrome formation (Fig. 1) predicts that amplified genes would be organized in head-to-head amplicons, with an asymmetric junction containing the inverted repeat which initiates sister chromatid fusion. Our data indicate that the head-to-head junction in the DHFR amplicon in CHOC 400 cells contains a relatively long asymmetric region, but no inverted repeat longer than a 2 bp AC overlap. A literature search for other studies that have sequenced naturally occurring inverted repeat amplicon junctions revealed, surprisingly, only two (21,22). Both amplicon junctions share similar features to the one described here. Both have asymmetric sequences (157 and 862 bp) and no inverted repeat. Importantly, both have trinucleotide microhomologies joining the two asymmetric sister chromatids. Therefore, although the asymmetric region predicted by the intramolecular recombination model (Fig. 1) is present in all sister chromatid fusions, none contains a significant region of inverted repeat homology. However, all these junctions are consistent with microhomology-mediated end joining.

Genetic studies strongly suggest that microhomology-mediated end joining is accomplished by the homologous repair system rather than the NHEJ repair system (albeit these systems may share some common subunits). Mice deficient in key components of the NHEJ system (Ku proteins, or XRCC4) that are also p53-deficient develop amplified c-myc genes with inverted repeats joined by microhomology-mediated end joining (23,24). In contrast, murine embryonic fibroblasts deficient in genes required for homologous recombinational repair (e.g. Brca1) cannot carry out microhomology-mediated end joining (32). Therefore, there is genetic precedent for the idea that homologous recombination can lead to sister

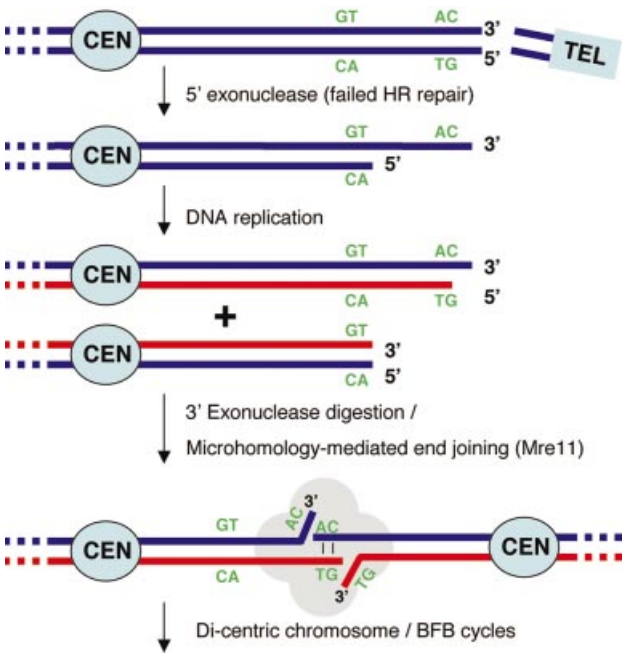


Figure 6. Microhomology-mediated recombination model for palindrome formation. The first two steps are identical to the intramolecular homologous recombination model (Fig. 1). However, instead of intramolecular recombination, DNA replication produces two sister chromatids of unequal length. The Mre11 complex then binds to the ends of the sister chromatids, and its 3′ exonuclease activity exposes the 5′ strands until regions of microhomology are reached, which stabilize the complex (33,34). At this point, enzyme systems that repair single strand gaps could complete the repair, as described (38). This produces a di-centric chromosome, which leads to gene amplification through BFB cycles as proposed by others (8,15,19).

chromatid fusion, but through microhomology rather than intra-molecular homologous recombination through inverted repeats.

We suggest the following model for sister chromatid fusion (Fig. 6). First, after a chromosome break, a failed attempt at homologous recombinational repair takes place, during which single strand exonuclease digestion normally occurs to facilitate strand invasion. If a telomere is not placed on this free end prior to DNA replication, two uncapped sister chromatids are formed, and the exonuclease-digested 5′ strand becomes slightly shorter than the other (4 kb in the case described here). The two sister chromatids, which are in close sub-nuclear proximity after replication, now form a substrate for Mre11-mediated homologous recombinational repair through microhomologies, the proposed mechanism for microhomology-mediated end joining (33,34). This process would result in an asymmetric junction joined by microhomologies, which is consistent with the three known naturally occurring amplicon junctions. As discussed for non-sister chromatid fusion, this model requires that cells first be rendered p53 deficient to avoid checkpoint responses during the initial failed attempt at HR repair (23,24), which is consistent with the genotype of CHOC 400 cells (35) and the fact that active p53 inhibits gene amplification (36,37). Interestingly, this model is also consistent with the results of artificially induced telomere fusions of Murnane and colleagues, who identified small homologies at the sites of sister chromatid fusion (17). In their study, most of the fused sister chromatids differed in length by several kilobases, and in all cases the joining site was many nucleotides from the break site, also as predicted by this model. Finally, this model is also consistent with sister fusions induced by breakage near a region of inverted repeat homology (15), except that the region of homology is larger than what would normally be considered microhomology.

In conclusion, the DNA sequence for the junction between the inverted repeats in CHOC 400 cells is consistent with a mechanism relying on microhomology-mediated end joining. We do not know why the fusion point in this cell line took place within a conspicuous and apparently unique direct repeat (Pst1 box). This repeat does not hybridize to any other genomic locations in the Chinese hamster genome (Fig. 5), or in the human and mouse genomes (not shown), indicating that it cannot be a common signal for chromosome break sites. Its presence may be fortuitous. Alternatively, the nature of the sequences contained within the repeat may have slowed the exonuclease processing of the 5′ end after the initial break (step 2 in Fig. 6). However, the general resulting structure predicted by this model is that of two sister chromatids joined by an asymmetric region with the junction consisting of a short region of microhomology. It will be interesting to test this model by examining the structure of additional natural amplicon junctions compared in cells with and without the functional proteins involved in this mechanism, such as Mre11.

ACKNOWLEDGEMENTS

We would like to thank J. L. Hamlin, for the cH2 cosmid that facilitated sequencing of this amplicon junction. This work was supported by NIH grant GM-57233-01 to D.M.G.

REFERENCES

1. Jackson,S.P. (2002) Sensing and repairing DNA double-strand breaks. *Carcinogenesis*, **23**, 687–696.
2. Gao,Y., Ferguson,D.O., Xie,W., Manis,J.P., Sekiguchi,J., Frank,K.M., Chaudhuri,J., Horner,J., DePinho,R.A. and Alt,F.W. (2000) Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. *Nature*, **404**, 897–900.
3. Nyberg,K.A., Michelson,R.J., Putnam,C.W. and Weinert,T.A. (2002) Toward maintaining the genome: DNA damage and replication checkpoints. *Annu. Rev. Genet.*, **36**, 617–656.
4. Tauchi,H., Matsuura,S., Kobayashi,J., Sakamoto,S. and Komatsu,K. (2002) Nijmegen breakage syndrome gene, NBS1 and molecular links to factors for genome stability. *Oncogene*, **21**, 8967–8980.
5. McClintock,B. (1941) The stability of broken ends of chromosomes in *Zea mays*. *Genetics*, **41**, 234–282.
6. Fouladi,B., Sabatier,L., Miller,D., Pottier,G. and Murnane,J.P. (2000) The relationship between spontaneous telomere loss and chromosome instability in a human tumor cell line. *Neoplasia*, **2**, 540–554.
7. Marder,B.A. and Morgan,W.F. (1993) Delayed chromosomal instability induced by DNA damage. *Mol. Cell. Biol.*, **13**, 6667–6677.
8. Trask,B.J. and Hamlin,J.L. (1989) Early dihydrofolate reductase gene amplification events in CHO cells usually occur on the same chromosome arm as the original locus. *Genes Dev.*, **3**, 1913–1925.
9. Ciullo,M., Debily,M.A., Rozier,L., Autiero,M., Billault,A., Mayau,V., El Marhomy,S., Guardiola,J., Bernheim,A., Coullin,P. *et al.* (2002) Initiation of the breakage-fusion-bridge mechanism through common fragile site activation in human breast cancer cells: the model of PIP gene duplication from a break at FRA7I. *Hum. Mol. Genet.*, **11**, 2887–2894.

10. Coquelle,A., Pipiras,E., Toledo,F., Buttin,G. and Debatisse,M. (1997) Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell*, **89**, 215–225.
11. Coquelle,A., Toledo,F., Stern,S., Bieth,A. and Debatisse,M. (1998) A new role for hypoxia in tumor progression: induction of fragile site triggering genomic rearrangements and formation of complex DMs and HSRs. *Mol. Cell*, **2**, 259–265.
12. Debatisse,M., Coquelle,A., Toledo,F. and Buttin,G. (1998) Gene amplification mechanisms: the role of fragile sites. *Recent Results Cancer Res.*, **154**, 216–226.
13. Butler,D.K., Gillespie,D. and Steele,B. (2002) Formation of large palindromic DNA by homologous recombination of short inverted repeat sequences in *Saccharomyces cerevisiae*. *Genetics*, **161**, 1065–1075.
14. Lobachev,K.S., Gordenin,D.A. and Resnick,M.A. (2002) The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell*, **108**, 183–193.
15. Tanaka,H., Tapscott,S.J., Trask,B.J. and Yao,M.C. (2002) Short inverted repeats initiate gene amplification through the formation of a large DNA palindrome in mammalian cells. *Proc. Natl Acad. Sci. USA*, **99**, 8772–8777.
16. Lo,A.W., Sabatier,L., Fouladi,B., Pottier,G., Ricoul,M. and Murnane,J.P. (2002) DNA amplification by breakage/fusion/bridge cycles initiated by spontaneous telomere loss in a human cancer cell line. *Neoplasia*, **4**, 531–538.
17. Lo,A.W., Sprung,C.N., Fouladi,B., Pedram,M., Sabatier,L., Ricoul,M., Reynolds,G.E. and Murnane,J.P. (2002) Chromosome instability as a result of double-strand breaks near telomeres in mouse embryonic stem cells. *Mol. Cell. Biol.*, **22**, 4836–4850.
18. Ma,C., Looney,J.E., Leu,T.H. and Hamlin,J.L. (1988) Organization and genesis of dihydrofolate reductase amplicons in the genome of a methotrexate-resistant Chinese hamster ovary cell line. *Mol. Cell. Biol.*, **8**, 2316–2327.
19. Ma,C., Martin,S., Trask,B. and Hamlin,J.L. (1993) Sister chromatid fusion initiates amplification of the dihydrofolate reductase gene in Chinese hamster cells. *Genes Dev.*, **7**, 605–620.
20. Singer,M.J., Mesner,L.D., Friedman,C.L., Trask,B.J. and Hamlin,J.L. (2000) Amplification of the human dihydrofolate reductase gene via double minutes is initiated by chromosome breaks. *Proc. Natl Acad. Sci. USA*, **97**, 7921–7926.
21. Legouy,E., Fossar,N., Lhomond,G. and Brison,O. (1989) Structure of four amplified DNA novel joints. *Somat. Cell Mol. Genet.*, **15**, 309–320.
22. Hyrien,O., Debatisse,M., Buttin,G. and de Saint Vincent,B.R. (1988) The multicopy appearance of a large inverted duplication and the sequence at the inversion joint suggest a new model for gene amplification. *EMBO J.*, **7**, 407–417.
23. Difilippantonio,M.J., Petersen,S., Chen,H.T., Johnson,R., Jasin,M., Kanaar,R., Ried,T. and Nussenzweig,A. (2002) Evidence for replicative repair of DNA double-strand breaks leading to oncogenic translocation and gene amplification. *J. Exp. Med.*, **196**, 469–480.
24. Zhu,C., Mills,K.D., Ferguson,D.O., Lee,C., Manis,J., Fleming,J., Gao,Y., Morton,C.C. and Alt,F.W. (2002) Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell*, **109**, 811–821.
25. Hamlin,J.L., Mosca,P.J. and Levenson,V.V. (1994) Defining origins of replication in mammalian cells. *Biochim. Biophys. Acta*, **1198**, 85–111.
26. Looney,J.E. and Hamlin,J.L. (1987) Isolation of the amplified dihydrofolate reductase domain from methotrexate-resistant Chinese hamster ovary cells. *Mol. Cell. Biol.*, **7**, 569–577.
27. Ma,C., Looney,J.E., Leu,T.H. and Hamlin,J.L. (1988) Organization and genesis of dihydrofolate reductase amplicons in the genome of a methotrexate-resistant Chinese hamster ovary cell line. *Mol. Cell. Biol.*, **8**, 2316–2327.
28. Li,F., Chen,J., Izumi,M., Butler,M.C., Keezer,S.M. and Gilbert,D.M. (2001) The replication timing program of the Chinese hamster beta-globin locus is established coincident with its repositioning near peripheral heterochromatin in early G1 phase. *J. Cell Biol.*, **154**, 283–292.
29. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
30. Benson,G. (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.*, **27**, 573–580.
31. Looney,J.E., Ma,C., Leu,T.H., Flintoff,W.F., Troutman,W.B. and Hamlin,J.L. (1988) The dihydrofolate reductase amplicons in different methotrexate-resistant Chinese hamster cell lines share at least a 273-kilobase core sequence, but the amplicons in some cell lines are much larger and are remarkably uniform in structure. *Mol. Cell. Biol.*, **8**, 5268–5279.
32. Zhong,Q., Chen,C.F., Chen,P.L. and Lee,W.H. (2002) BRCA1 facilitates microhomology-mediated end joining of DNA double strand breaks. *J. Biol. Chem.*, **277**, 28641–28647.
33. Paull,T.T. and Gellert,M. (1998) The 3′ to 5′ exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. *Mol. Cell*, **1**, 969–979.
34. Paull,T.T. and Gellert,M. (2000) A mechanistic basis for Mre11-directed DNA joining at microhomologies. *Proc. Natl Acad. Sci. USA*, **97**, 6409–6414.
35. Lee,H., Larner,J.M. and Hamlin,J.L. (1997) A p53-independent damage-sensing mechanism that functions as a checkpoint at the G1/S transition in Chinese hamster ovary cells. *Proc. Natl Acad. Sci. USA*, **94**, 526–531.
36. Paulson,T.G., Almasan,A., Brody,L.L. and Wahl,G.M. (1998) Gene amplification in a p53-deficient cell line requires cell cycle progression under conditions that generate DNA breakage. *Mol. Cell. Biol.*, **18**, 3089–3100.
37. Yin,Y., Tainsky,M.A., Bischoff,F.Z., Strong,L.C. and Wahl,G.M. (1992) Wild-type p53 restores cell cycle control and inhibits gene amplification in cells with mutant p53 alleles. *Cell*, **70**, 937–948.
38. Gottlich,B., Reichenberger,S., Feldmann,E. and Pfeiffer,P. (1998) Rejoining of DNA double-strand breaks *in vitro* by single-strand annealing. *Eur. J. Biochem.*, **258**, 387–395.