The EMBO Journal Vol.17 No.1 pp.325–333, 1998

Interstitial deletions and intrachromosomal amplification initiated from a double-strand break targeted to a mammalian chromosome

E.Pipiras, A.Coquelle, A.Bieth and M.Debatisse[1]

Unité de Génétique Somatique (URA CNRS 1960), Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cédex 15, France

[1]Corresponding author
e-mail: mdebat@pasteur.fr

Interstitial deletions of tumour suppressor genes and amplification of oncogenes are two major manifestations of chromosomal instability in tumour cells. The development of model systems allowing the study of the events triggering these processes is of major clinical importance. Using the properties of the I-SceI nuclease to introduce a localized double-strand break (DSB) in a mammalian chromosome carrying its target sequence, we demonstrate here that both types of mutations can be initiated by non-conservative DSB repair pathways. In our system, I-SceI activity dissociates a transfected *gpt* gene from its promoter, allowing the isolation of gpt$^-$ clones. Our results show that intrachromatid single-strand annealing events occur frequently, giving rise to interstitial deletions not accompanied by other chromosomal rearrangements. We also observed that, when present in the cells, extrachromosomal DNA molecules are integrated preferentially at the broken locus. Taking advantage of the insertion of the I-SceI recognition sequence telomeric to and close to the dihydrofolate reductase gene, we show that a less frequent outcome of I-SceI activity is the initiation of cycles of intrachromosomal amplification of this marker, from breaks at a site merging with the enzyme target.

Keywords: deletions/double-strand breaks/gene amplification/I-SceI nuclease/tumour progression

Introduction

Several recent studies point to a major role for double-strand breaks (DSBs) in the well-documented genomic instability that contributes to tumour progression. They establish a strong correlation between breaks at genetically defined loci, designated fragile sites and chromosomal deletions of tumour suppressor genes identified in some cancers (Roth and Wilson, 1988; Jones *et al.*, 1995; Ohta *et al.*, 1996; Sozzi *et al.*, 1996; Wilke *et al.*, 1996). However, the molecular mechanisms leading to these rearrangements remain to be characterized. The study of *in vitro* model systems has shown that site-specific breaks induced by drugs that activate fragile sites are also able to trigger a complete sequence of events leading to intrachromosomal gene amplification (Coquelle *et al.*, 1997). In this work, only gene amplification could be selected for among the possible outcomes of fragile site activation. None of these studies has determined clearly whether the breaks *per se* or some particular sequences or structural features of the fragile sites contribute to initiate the different processes, a question that could also be addressed specifically with a model system. Thus, it is of special interest to design *in vitro* systems to study the various chromosomal consequences of targeted DSBs, to evaluate their relative frequencies and to analyse the pathways involved in the formation of these different outcomes.

Recently, promising experimental approaches based on the properties of I-SceI, a mitochondrial group I intron-encoded nuclease of *Saccharomyces cerevisiae* (Dujon, 1989), have been developed. Because I-SceI is not a recombinase, it constitutes a particularly interesting system to study the chromosomal consequences of a targeted DSB. The 18 bp recognition sequence of the nuclease (Colleaux *et al.*, 1988) is expected to be infrequent enough to be naturally absent from most genomes, and can be introduced in a chromosome by transfection. The I-SceI recognition sequence has indeed been introduced in the nuclear genome of various organisms, and induction of site-specific DSBs by the enzyme has been demonstrated in a broad variety of eukaryotic cells. I-SceI-targeted breaks have been shown to stimulate homologous recombination between extrachromosomal DNA substrates in yeast (Plessis *et al.*, 1992), plant (Puchta *et al.*, 1993), *Xenopus* (Segal and Carroll, 1995) and mammalian (Rouet *et al.*, 1994a) cells. Homologous recombination events between transfected substrates and a chromosomal locus are also enhanced in yeast (Fairhead and Dujon, 1993), plant (Puchta *et al.*, 1996) and mammalian (Rouet *et al.*, 1994b; Choulika *et al.*, 1995) cells. Finally, homologous recombination between two chromosomal sequences is stimulated by I-SceI-mediated DSB in yeast (Fairhead and Dujon, 1993) and in various mammalian cells (Lukacsovich *et al.*, 1994; Rouet *et al.*, 1994b; Sargent *et al.*, 1997).

In order to be able to study the outcomes of DSBs induced by I-SceI, we designed a plasmid in which the I-SceI recognition site was inserted between the bacterial guanine phosphoribosyl transferase (*gpt*) coding sequence and the cytomegalovirus (CMV) promoter. Stable transfectant clones were obtained from a Chinese hamster fibroblastic cell line, and the transgene integration sites were mapped by fluorescent *in situ* hybridization (FISH). In one clone (112), the construct had integrated close to and telomeric to the dihydrofolate reductase gene (*DHFR*), the amplification of which can be selected for in medium supplemented with metothrexate (MTX) (Schimke *et al.*, 1978). Clone 112 was chosen for further studies since it would permit analysis of the consequences of a variety of healing events—including amplification, deletions and

translocations—at the cytological and the molecular level in independent clonal cell populations. In this clone, a structure composed of two short direct repeats of plasmid sequences flanking a unique 6 kb region containing the I-SceI recognition site has been formed during the transfection experiment. We describe here two outcomes, interstitial deletions and intrachromosomal gene amplification, of a DSB occurring within a single copy locus flanked by short repeats, a situation mimicking the organization of mammalian genomes which are punctuated by repeated sequences such as LINE and SINE elements (Deininger, 1989; Hutchison *et al.*, 1989).

Results

Recovery and analysis of gpt$^+$ cell lines with a transfected I-SceI recognition sequence

Preliminary experiments had shown that GMA32 cells stably transfected with pI-SceI and expressing up to 10 ng of the nuclease per 10^6 cells, as estimated by Western blot quantification, have a normal growth rate (not shown). This strongly suggests that I-SceI recognition sequences do not exist in the genome of these Chinese hamster cells. An HGPRT$^-$ derivative of the GMA32 cell line was transfected with circular plasmid DNA (pci) which contains the I-SceI recognition sequence inserted between the selectable and counter-selectable *Escherichia coli gpt* gene and the CMV promoter, and the *neo* gene under the control of the SV40 early promoter (see Materials and methods). Two days after transfection, the cells were plated in medium containing G418 and mycophenolic acid (MPA) + xanthine (X) to select for stably transfected G418^r gpt$^+$ clones. Such clones were recovered after 10 days and were routinely grown in selective medium. The transgene was localized by FISH in several independent clones, using pci as a probe. In one clone, designated 112, the transgene was integrated on the p arm of one chromosome 2 (Figure 1A). This localization is of particular interest because chromosome 2p is easy to recognize both after trypsin/Giemsa (Figure 1B) or 4′,6′-diamidino-2-phenylindole (DAPI) banding (Coquelle *et al.*, 1997) and because it bears several markers, including the *dhfr* locus, for which cosmid probes are available. The results of a two-colour FISH experiment, in which the *DHFR* genes and the transgene were revealed with rhodamine and fluorescein respectively, are shown in Figure 1C. The transgene lies very close to and telomeric to the *dhfr* locus, an appropriate localization to test the ability of I-SceI-induced DSBs to trigger intrachromosomal amplification of this marker.

The transgene was then mapped at the molecular level. DNA from line 112 was prepared and digested by *Eco*RI, a restriction enzyme which has no site in pci. A unique 12 kb band was detected when Southern blots were probed with pci, a result indicating, as previously suggested by FISH, that the transgene integrated the cellular genome at a single locus (Figure 4). An *Eco*RI–I-SceI double digestion disclosed a unique I-SceI site, and further analyses of restriction patterns (not shown) allowed us to draw the map presented in Figure 2. It shows that a 6 kb long region containing the I-SceI recognition site, the *gpt* and the *neo* genes and associated transcription initiation and termination sequences is present as a single copy. On the contrary, plasmid sequences (1.5 kb long at most) have been duplicated and are organized as direct repeats flanking the single copy sequence described above.

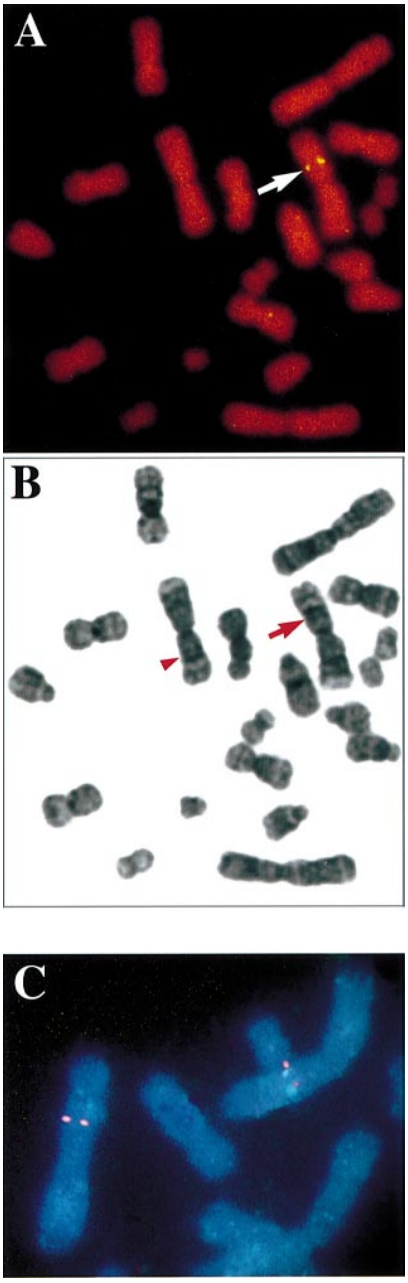

Fig. 1. Cytological analysis of clone 112. (A) and (B) represent the same metaphase plate. (**A**) FISH with a probe specific for the transgene. The arrow points to the integration site. (**B**) Giemsa banding, the arrow points to the integration site on chromosome 2p, the arrowhead points to the homologous chromosome arm. (**C**) Double colour FISH. Red spots: *DHFR* genes, green spots: transgene.

Selection of gpt$^-$ clones

The pci plasmid was constructed so as to force a DSB occurring at the I-SceI site to separate the *gpt* gene from its promoter. Imperfect repair events leading to chromosomal rearrangements such as deletions, translocations or gene amplification are thus expected to generate gpt$^-$ cells. Cells of line 112 were transfected with pI-SceI or with pcDL-SRα296 as control. After 2 days in

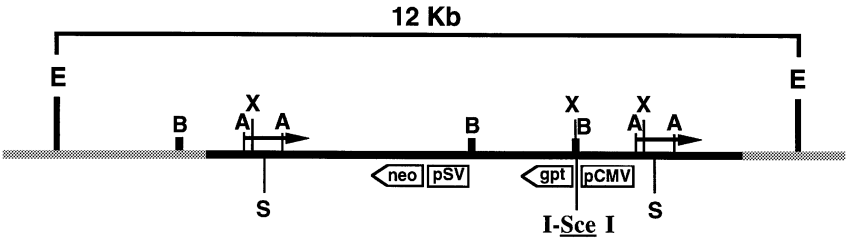

Fig. 2. Map of the transgene in clone 112. Black horizontal line, transgene sequence; grey lines, genomic sequence. Promoters and genes are represented in boxes; transcription directions are indicated. Black arrows represent the direct repeats. I-SceI: I-SceI recognition sequence. E, *Eco*RI; A, *Apa*LI; B, *Bam*HI; X, *Xmn*I; S, *Sca*I restriction sites.

normal medium, aliquots of the cells transfected with either plasmid were plated in medium containing thioxanthine (TX) (T2: 2 days after transfection) or in normal medium in which the cells were allowed to grow for 3 more days (~6 cell generations) before being plated in TX-containing medium (T5: 5 days after transfection). Ten days later, some plates were used to compare the frequencies of resistant clones after fixation and Giemsa staining, and individual colonies were picked in the remaining dishes. The results of a typical experiment are shown in Figure 3. At least 10 independent transfection experiments were performed, and similar results were obtained, leading to the conclusion that after transfection with the control plasmid, the frequency of TX-resistant colonies is reproducibly close to 10^{-3} at T2 and $2–3\times10^{-3}$ at T5. This indicates that the gpt$^-$ phenotype is acquired spontaneously at a significant frequency. After transfection with pI-SceI, the frequency of colonies was generally close to that determined in the control experiments at T2 but repeatedly was enhanced at least 20-fold at T5. Such results clearly show that I-SceI does induce the formation of TX-resistant clones but that relatively long expression time is required for the cells to become phenotypically gpt$^-$.

Characterization of the gpt$^-$ clones.

Thirty two clones, selected at T2 or T5 from transfections with either plasmid, were analysed further by dot-blot to determine whether the *gpt* gene was present or absent in the gpt$^-$ clones. As shown in Table I, hybridization with a *gpt* probe revealed the presence of the gene in all the eight clones recovered after transfection with plasmid pcDL-SRα296, suggesting that the gpt$^-$ phenotype probably results from gene extinction or point mutations in these cases. After transfection with pI-SceI, eight out of nine T2 and one out of fifteen T5 clones were similar to those obtained with the control plasmid. In striking contrast, in the last T2 clone and in the 14 remaining T5 clones, the *gpt* gene was lost, as judged by the complete lack of hybridization signal. Thus, the specific increase in gpt$^-$ clones induced by I-SceI parallels the emergence of cells from which the gene is deleted. The DNA of various clones was purified, restricted with *Eco*RI and used to prepare Southern blots. Hybridization with the *neo* gene as probe established that this gene was absent from the cells of all the clones presenting a deletion of the *gpt* gene, while it was present in all other cases (Figure 4A). Thus, the gpt$^-$ clones induced by I-SceI exhibited a deletion of at least the *neo* and *gpt* genes. On the contrary, both genes were present in all the clones transfected with the

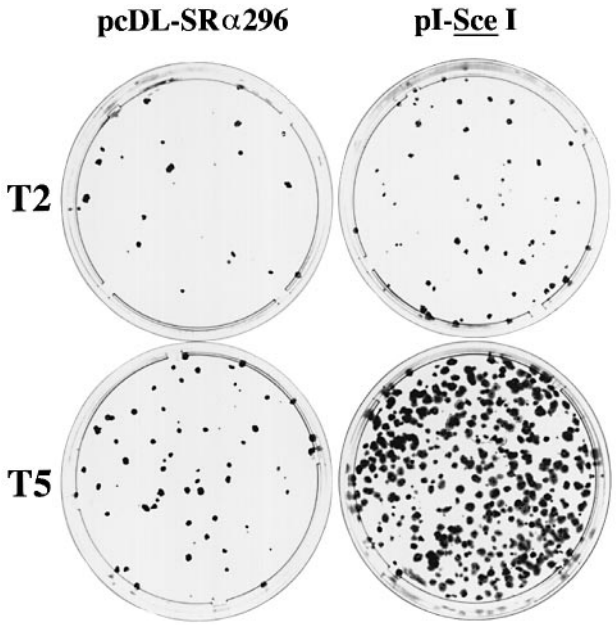

Fig. 3. Giemsa staining of gpt$^-$ colonies after transfection. Left: transfection with the control plasmid, right: transfection with the plasmid coding for I-SceI. T2 and T5 correspond to the expression time (in days) before plating in selective medium.

Table I. Summary of dot-blot analysis of TX-resistant colonies

Transfection conditions	No. of clones analysed	*gpt* gene	
		Present	Absent
pcDL-SRα T2 + T5	8	8	0
I-SceI T2	9	8	1
I-SceI T5	15	1	14

control plasmid and in some 90% of the clones obtained at T2 after transfection with pI-SceI. The DNAs were then hybridized with pT7T3, a plasmid which participated to the construction of pci (Figure 4B). With the exception of one clone I-SceI T2 (I-SceI T2-3), the 12 kb *Eco*RI band typical of line 112 was found, as expected, in the clones which have retained the *gpt* and *neo* genes. Strikingly, in seven out of seven clones analysed in which the genes are deleted, a single fragment of constant size (~6 kb) was detected, indicating that not all the plasmid sequences have been lost. The transgene sequences remaining in these clones were mapped further; the results indicated that one copy of the repeats and the whole intervening sequence have been removed (not shown).

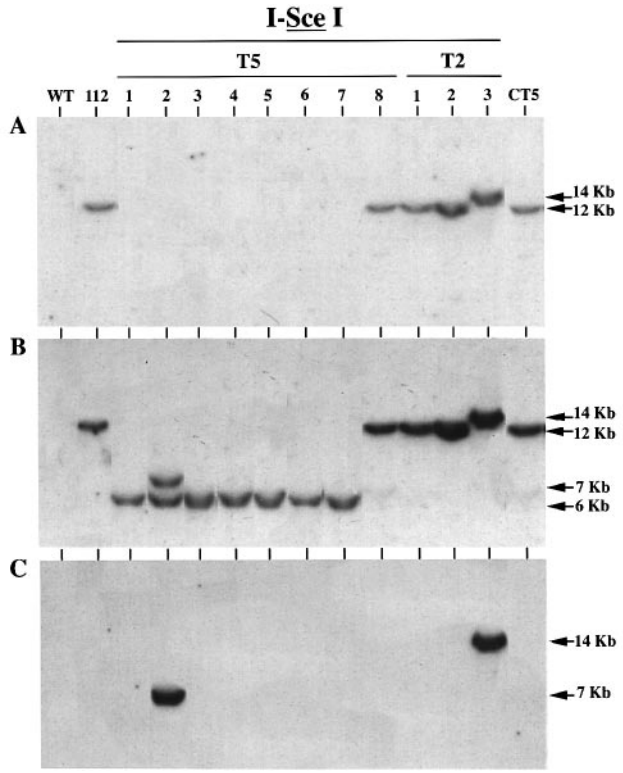

Fig. 4. Southern analysis of gpt⁻ clones. The DNA were restricted with *Eco*RI. Hybridization with (**A**) the *neo* gene (**B**) plasmid pT7T3 and (**C**) the I-SceI gene. WT: parental line; eight clones T5 and three clones T2 recovered after transfection with p I-SceI are shown. Clone T5-8 is the only undeleted T5 clone. CT5: clone T5 obtained after transfection with the control plasmid. Fragment sizes are indicated in kb on the right.

The constant size and sequence content of the deletions strongly suggest that homology-driven events are responsible for their formation. Hybridization of the same blot with the I-SceI gene as a probe (Figure 4C) showed that the 7 and the 14 kb extra bands observed respectively in clones I-SceI T5-2 and T2-3 contain the I-SceI gene sequence. This is simply explained if pI-SceI integrated in some transfected cells, even though such events were not selected for in this experiment. Moreover, the existence of a 14 kb-long *Eco*RI fragment, in place of the expected 12 kb fragment in clone I-SceI T2-3, strongly suggests that a targeted integration occurred in this case.

Finally, deleted I-SceI T5 clones were analysed by G banding. Two normal chromosomes 2 were identified in each of them at that observation level, a result indicating that the 6 kb deletion is not associated with extensive chromosomal rearrangements (not shown).

DHFR amplification in gpt⁻ clones

To determine whether gene amplification can be triggered by a site-specific break induced by I-SceI, we took advantage of the localization of the transgene in line 112. As mentioned above, it lies very close to and telomeric to the *DHFR* gene (Figure 1C). We previously demonstrated that MTX not only selects for but also induces *DHFR* amplification because it activates a folate-sensitive fragile site located relatively far away from the selected gene, near the telomere of chromosome 2p (Coquelle *et al.*, 1997). The breaks induced at fragile sites trigger the breakage–fusion–bridge cycles (BFB) mechanism of amplification, the landmark of which is a deletion of the part of the chromosome arm telomeric to the break from the amplified chromosome (Figure 6A). This mechanism operates in a cell in which broken sister chromatids have fused to form a bridge. At mitosis, the breakage of this giant inverted repeat leaves each daughter cell with a chromatid lacking one telomere. After replication, the broken sister chromatids fuse again, perpetuating the BFB cycles. Amplification occurs in one daughter cell when breakage at anaphase is asymmetric, leading to unequal distribution of the sequences lying on the dicentric chromatid in the daughter cells (Smith *et al.*, 1992a; Toledo *et al.*, 1992; Ma *et al.*, 1993).

Table II. Analysis of MTX + TX-resistant clones

Transfection conditions	No. of clones analysed	DMs	BFBs		
			Total	Long[a]	Short[a]
pcDL-SRα T2 + T5	7	1	6	6	0
I-SceI T2	3	1	2	2	0
Total	10	2	8	8	0
I-SceI T5	10	1	9	1	8

[a]Refers to the distance separating the normal copy and the first extra copy of the *DHFR* gene on the amplified chromosome 2p (for illustration see Figure 5B and C).

In the case of MTX-induced *DHFR* amplification, the first extra copy of the gene is separated from the normal copy by twice the rather long distance separating the gene and the fragile site (Coquelle *et al.*, 1997). Obviously, this distance is predicted to be much shorter if the initial inverted duplication is initiated from a DSB induced at the nearby I-SceI site. Thus, we selected clones resistant to TX plus MTX after transfection with either the control or the pI-SceI plasmid. Two days after transfection, cells were plated in MTX + TX (T2) or in MTX alone, TX being added 3 days later (T5). In control experiments, the cells were challenged at T2 or T5 with TX or with MTX alone. At T2 with either plasmid and at T5 with the control plasmid, double-resistant clones appeared with a frequency corresponding roughly to the product of the frequencies of each single resistance. On the contrary, when the cells were transfected with pI-SceI and challenged in MTX and TX at T5, the frequency of double-resistant clones was higher (4–10 times depending on the experiment) than expected from the frequencies of clones resistant to TX or MTX alone (not shown). We analysed 20 double-resistant clones exhibiting an amplification of the *DHFR* gene; 10 were obtained upon transfections with the control plasmid (T2 and T5) or with pI-SceI (T2), and 10 with pI-SceI (T5). Table II summarizes the results obtained by FISH using probes for the *dhfr* locus and two cytogenetic markers; marker II-6 which lies close to the telomere of chromosome 2p (Figure 5A) and marker II-4 which is relatively close to and telomeric to both the *DHFR* gene and the transgene (compare Figure 1C with Figure 5B and C). In all experimental conditions, rare clones exhibited extrachromosomal copies of the *DHFR* gene on double-minute chromosomes (DMs), while the cytological characteristics of BFB cycles were found in the others. In the latter clones, when selected at T2 after

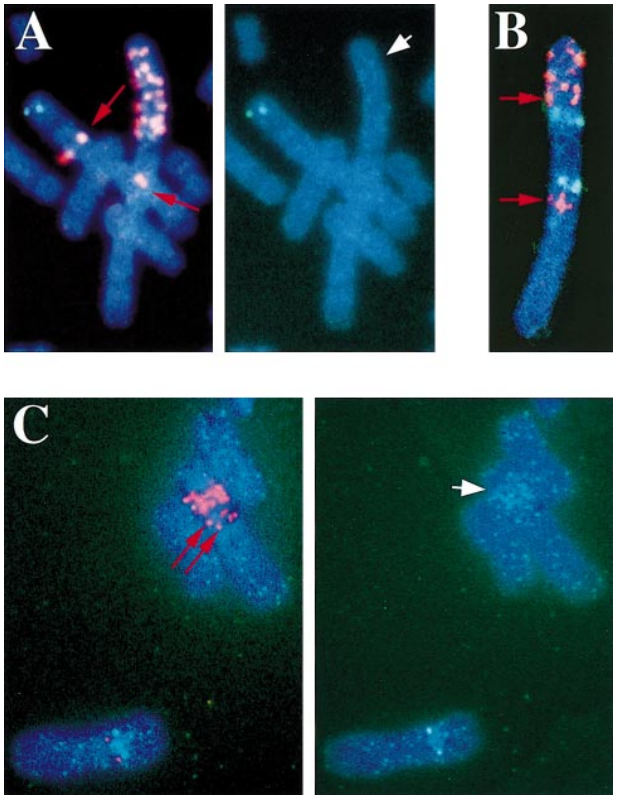

Fig. 5. Two-colour FISH of MTX + TX-resistant clones.
(A and B) Cells of clones from control transfections. (**A**) Green signal, marker II-6; red signal, *DHFR*. Left: red arrows point to the two normal *DHFR* copies. Right: only the green signal is visualized, the white arrow shows the amplified chromosome 2p deleted for marker II-6. (**B**) Green signal, marker II-4; red signal, *DHFR*. The red arrows point to the normal and to the first extra copy of the *DHFR* gene, separated by a large distance. Two-colour hybridization of co-amplified sequences discloses the palindromic organization of the initial repeat. (**C**) Cells of clones recovered after transfection with p I-SceI at T5. Green signal, marker II-4; red signal, *DHFR*. Left: the red arrows point to the normal and first extra copy of the *DHFR* gene, lying close to each other. Right: only the green signal is visualized, the white arrow shows the amplified chromosome 2p deleted for marker II-4.

transfection with either plasmid or at T5 with the control plasmid, the structure of the amplified chromosome arm was similar to the one found previously in MTX-induced *DHFR*-amplified mutants (Coquelle *et al.*, 1997). In particular, we observed a deletion of the telomeric marker II-6 from the amplified chromosome arm (Figure 5A) and, as a consequence of the localization of the fragile site, a large initial inverted duplication in which marker II-4 was co-amplified (Figure 5B). Strikingly, in eight out of nine clones recovered after transfection by pI-SceI at T5, the organization of the extra copies of the *DHFR* genes along the chromosome 2p arm was notably different from that observed in all the clones obtained in control experiments. The results are illustrated in Figure 5C: in most cells of these clones, the first extra copy of the *DHFR* gene is located close to the normal copy of the gene, while marker II-4 is deleted from the amplified chromosome arm. Thus, the initiating break occurred repeatedly between marker II-4 and the *DHFR* gene, a result indicating that the intrachromosomal amplification process was triggered in these cases by an I-SceI-induced DSB.

Discussion

The availability of rare cutting enzymes, allowing induction of DSBs in a defined chromosomal context in mammalian cells, recently has led several groups to reevaluate the relative contribution of homologous and non-homologous events in the repair process. With systems designed to reveal homologous recombination events leading to the reconstruction of a functional selectable gene from two inactivated copies or fragments of this gene (gain-of-function assays), it was demonstrated that the induction of a break within one of the repeated regions considerably increases the frequencies of homologous recombination and of some types of illegitimate events associated with small deletions correcting insertion mutations (Rouet *et al.*, 1994b). The latter events were shown to depend on Ku-mediated end joining (Liang *et al.*, 1996), a repair pathway requiring no homology and which was long considered to be the more efficient healing process in mammalian cells (Roth and Wilson, 1988). More recently, a loss-of-function assay was designed which allowed analysis of a wide spectrum of events taking place between direct repeats when a DSB is introduced in one repeated sequence. Again, it was found that both illegitimate and homologous recombination are stimulated (Sargent *et al.*, 1997).

We designed an experimental system allowing us to analyse the different chromosomal consequences of non-conservative DSB repair events, and to compare their relative frequencies. A clone of Chinese hamster fibroblasts in which a plasmid containing the I-SceI site between the *gpt* gene and its promoter was integrated on the p arm of a chromosome 2, close to and telomeric to the *DHFR* gene, was retained for study. In this clone, designated 112, integration was accompanied by the formation of a short duplication of vector sequences, organized as direct repeats flanking a 6 kb long single copy region containing the I-SceI target. We selected gpt^- colonies in order to study the mechanisms operating when a break occurring in a single copy region flanked by such repeats is repaired imprecisely. We have shown that a sequence of constant size, including the *gpt* gene, was deleted in 14 out of 15 independent gpt^- clones recovered after transfection by a plasmid coding for I-SceI. The last clone presented the same characteristics as the clones recovered in control experiments, in which deletions were never observed. Further studies of the deleted region showed that in all 14 clones, one copy of the direct repeats was retained while the second one and all the intervening sequence were lost. At the chromosomal level, the events were not accompanied by cytologically detectable rearrangements. The fact that the very same structure was formed repeatedly indicates that the repair pathway responsible for healing is homology driven. The whole set of observations suggests that the non-conservative repair mechanism known as the single-strand annealing pathway (SSA) (Lin *et al.*, 1984) (Figure 6B) was triggered at high frequency in our experimental system. It is supposed to result from exonuclease degradation of one strand on each side of the break, or from unwinding, until complementary sequences of the repeats are exposed and anneal. This mechanism has been detected frequently in yeast between chromosomal repeats and in vertebrate cells between extrachromosomal transfected or injected sequences (Haber, 1992). Our

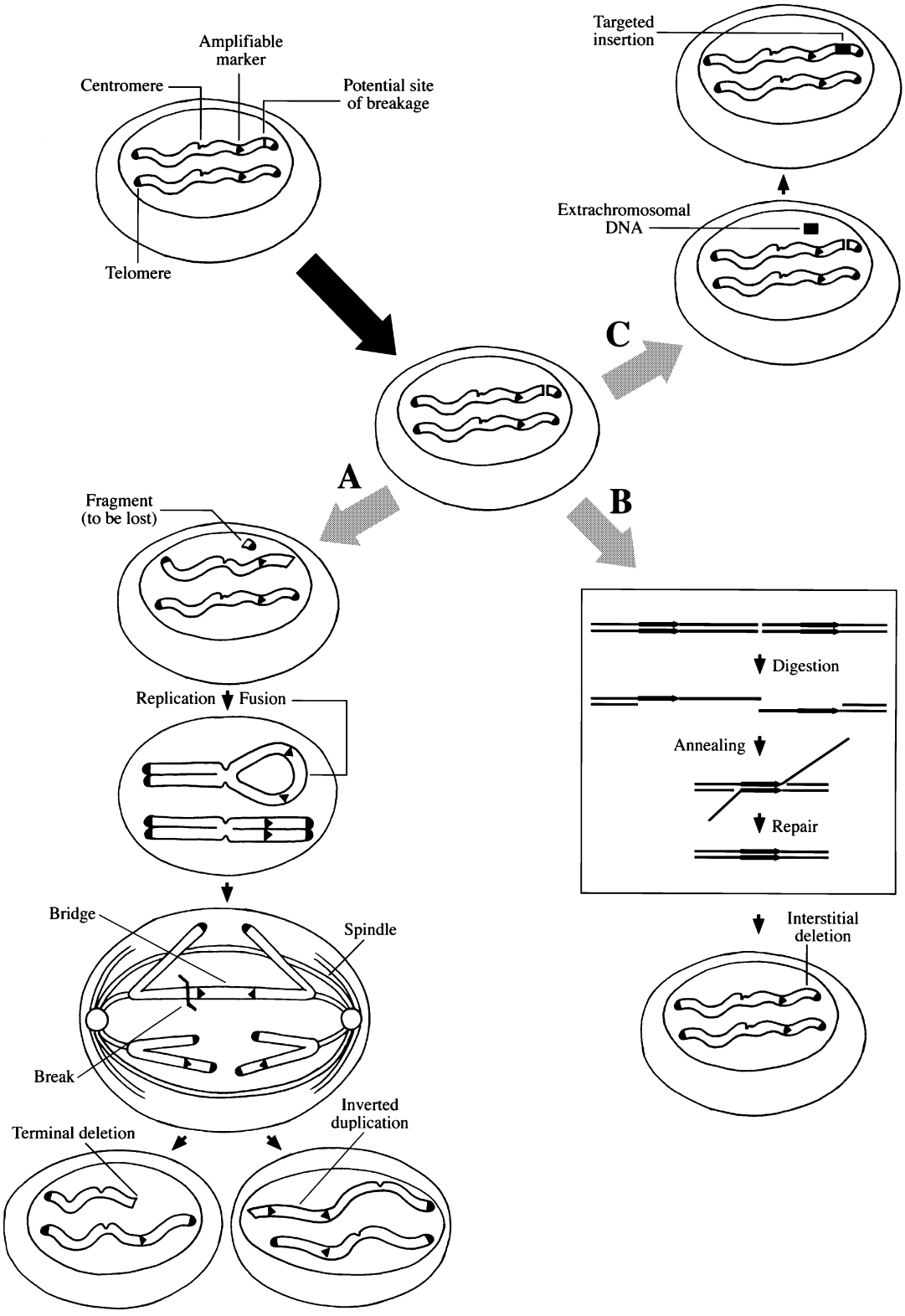

Fig. 6. Different outcomes of DSB. See text for a detailed description of the mechanisms. (**A**) BFB cycles. The chromosome fragment telomeric to the break is usually lost. The highly rearranged chromosomes resulting from a single cycle are represented: the telomeric part of the broken chromosome is deleted in one daughter cell (left), and an intrachromosomal inverted duplication is formed in the other daughter cell (right). (**B**) SSA mechanism. Short direct repeats are represented by the horizontal arrows and the intervening sequence by the black line. One repeat and the whole intervening sequence are deleted without cytologically detectable chromosomal rearrangement. (**C**) When present in the cell, extrachromosomal DNA molecules are preferentially integrated within the break.

results strongly suggest that SSA also constitutes a major homology-mediated healing mechanism of chromosomal DSBs in the Chinese hamster line studied here: ~1% of the cells became gpt^- after transfection by the plasmid encoding I-SceI while, as juged from control transfections run in parallel, only 10–20% of the cells were transfected in the experimental conditions we used. Thus, SSA accounts for some 10% of the repair events if a break occurs in every transfected cell, or more if a break occurs in <100% of the transfected cells.

In vivo, this mechanism would be dependent on breaks occurring at random or within hot spots, and on the presence of naturally occurring direct repeats flanking the breaks. In mammalian cells, repeated sequences represent as much as 20–30% of the chromosomal DNA. In rodent genomes, for example, the B1 and B2 elements belonging to the SINEs family of repeats are each present in $>10^5$ copies per genome. The L1 members of the LINE family are also repeated ~10^5 times. Each of these repeated elements is expected to arise every 30 kb on average (Deininger, 1989; Hutchison *et al.*, 1989) and could offer a molecular substrate for SSA-driven intrachromatid repair. A limitation for such homologous events could be due to the fact that individual sequences diverge from the consensus. However, it has been shown that in yeast cells, B2 elements efficiently drive site-specific integration of plasmids into a yeast artificial chromosome containing mouse DNA (Colleaux *et al.*, 1993). The operation of the SSA pathway in human cells is suggested by the results of several studies which have shown that interstitial deletions involving repeated elements such as Alu1 sequences are frequent in human tumours (Smidt *et al.*, 1990; Tadokoro *et al.*, 1992; Nehls *et al.*, 1995; Nystrom-Lahti *et al.*, 1995; Mauillon *et al.*, 1996) and that recurrent deletions of tumour suppressor genes can be correlated with conveniently localized fragile sites (Roth and Wilson, 1988; Jones *et al.*, 1995; Ohta *et al.*, 1996; Sozzi *et al.*, 1996; Wilke *et al.*, 1996).

The unselected integration of pI-SceI has been observed in two out of 11 clones analysed. Strikingly, one of these events occurred within the 12 kb *Eco*RI fragment containing the transgene. The structure created by this integration was not studied further. However, a classical homologous integration event occurring between extrachromosomal and intrachromosomal vector sequences would require an extensive degradation of the intervening sequence around the chromosomal break, including the *neo* gene. The fact that the *neo* gene was retained in the genome of this clone rules out this hypothesis. As observed by others (Ellis and Bernstein, 1989; Berinstein *et al.*, 1992; Rouet *et al.*, 1994b), the transfected DNA could have been integrated by a one-sided homologous recombination event. It is also possible that a homology-independent mechanism preferentially drives the extrachromosomal DNA molecules to DSBs (Figure 6C). This could account for the recurrent integration of transfected DNA in cultured cells (Rassool *et al.*, 1991; Smith *et al.*, 1992b) and of viral genomes in tumours (Popescu, 1994; Choo *et al.*, 1996; Wilke *et al.*, 1996) at loci previously characterized as fragiles sites.

We were able to analyse also the contribution of DSBs to the initiation of gene amplification. In some cases, amplification proceeds through an entirely intrachromosomal pathway: the chromatid BFB cycles (McClintock, 1951) (Figure 6A), operating in cells in which two sister chromatids have fused. It was suggested initially that BFB cycles can be initiated by random breakage of chromatids, illegitimate recombination events or telomere fusion (Smith *et al.*, 1992a; Ma *et al.*, 1993; Toledo *et al.*, 1993; Poupon *et al.*, 1996). More recent work demonstrated that most BFB cycles are actually triggered by site-specific events induced at fragile sites by some clastogenic drugs (Coquelle *et al.*, 1997). Here we show that a complete sequence of events leading to *DHFR* gene amplification through BFB cycles can be initiated by an I-SceI-induced DSB telomeric to this gene. This indicates that amplification from fragile sites does not reflect any peculiarity other than their high probability of double-strand breakage. Moreover, in this model system, as in the case of fragile site activation by drugs, the efficiency of induction of amplification is low, and amplified mutants have to be selected for among the TX-resistant mutants.

Deletion of tumour suppressor genes and oncogene amplification have long been recognized to play a key role in tumour progression (Brison, 1993; Stark, 1993). The whole set of results presented here shows that, when imperfectly repaired, a DSB can trigger both an interstitial deletion and a large terminal deletion coupled to a gene amplification cycle. We previously have demonstrated that dicentric and ring chromosomes are secondary products of BFB cycles (Toledo *et al.*, 1993). This further points to DSBs having a key role in the triggering of various types of rearrangements observed frequently in tumour cells. Moreover, hot spots of breakage could contribute to targeted viral genome integration *in vivo*. Thus, DSBs appear to contribute in a major way to oncogenesis, through multiple pathways.

Materials and methods

Cell lines and growth conditions

The wild-type GMA32 cell line, its $HGPRT^-$ derivative and normal culture medium have been described previously (Debatisse *et al.*, 1977). Transfection experiments were performed using lipofectamine™ reagent (Life Technologies), using the protocol recommended by the supplier with a 5 h treatment. Routinely, 10–20% of the cells were transfected in these conditions. The selection and counterselection of gpt^+ cells were performed respectively in medium supplemented with 0.05 mM MPA (Sigma) plus 1 mM X (Sigma) (Mulligan and Berg, 1981), and with 0.6 mM TX (Sigma) (Besnard *et al.*, 1987). G418 (Gibco BRL) was used at 2 mg/ml to select for *neo*-resistant clones. Two or 5 days after transfection with pI-SceI, the cells were plated to select for mutants with amplified *DHFR* genes; cells were plated in 9 or 18 ng/ml MTX (Sigma). The independent colonies obtained in MTX were expanded to 10^5 cells in selective medium. Cells from the different clones were treated with colcemide or nocodazole, then spread on slides as previously described (Toledo *et al.*, 1992) and used for FISH studies.

Plasmid constructions

Plasmid pci was constructed in two steps: (i) the *Hin*dIII–*Sca*I fragment of plasmid pMSG (Pharmacia) containing the *E.coli gpt* gene and the SV40 polyadenylation sequence was inserted in the bacterial vector pAF100 (a generous gift of B.Dujon) restricted with *Hin*dIII and *Sca*I. In this construction, the *gpt* gene is brought to 45 bp from the I-SceI recognition site. (ii) The *Bam*HI fragment containing the I-SceI recognition site, the *gpt* gene and the SV40 polyadenylation sequence were recovered and cloned at the unique *Bam*HI site of pCB6 (a generous gift of D.Russel); pci is one of the clones in which the *Bam*HI fragment was inserted in the orientation allowing *gpt* transcription from the CMV promoter in mammalian cells.

Plasmid pI-SceI was constructed by inserting a 734 bp *Bam*HI

fragment containing the I-SceI synthetic coding sequence (Colleaux *et al.*, 1986) in *Bam*HI-restricted pcDL-SRα296 plasmid (Takebe *et al.*, 1988). A clone in which the I-SceI gene was inserted in the correct orientation with respect to the very efficient SRa promoter was used for transient transfection experiments.

In situ hybridization and Giemsa–trypsin banding

FISH was performed essentially as described (Pinkel *et al.*, 1988; Tkachuk *et al.*, 1990). Cosmids KP 454 and KZ 381 (a gift of J.Hamlin) were used to probe the *dhfr* locus. Cosmids II-4 and II-6 [isolated from a Chinese hamster chromosome 2-specific library (M.Debatisse, B.Labidi and P.Metezeau, unpublished)] were used as markers of chromosome 2p. The probes were biotinylated by nick translation (BioNick Labeling System kit; Gibco-BRL) or labelled with digoxigenin by using the DIG DNA labelling kit (Boehringer). The labelled probes were purified by filtration through Quick Spin Columns (Sephadex G-50; Boehringer). Slides, treated with RNase and proteinase K, were hybridized with different combinations of the four cosmids as previously described (Toledo *et al.*, 1992, 1993). For hybridization with the pci probe, the hybridization mixture [10 µl per slide containing 50 ng of labelled plasmid and 100 ng of Chinese hamster competitor DNA, 50% formamide, 1% Tween-20 (Pierce), 10% dextran sulfate (Pharmacia) in 2× SSC] was heated for 10 min at 70°C and immediately applied to slides. For two-colour FISH experiments with both cosmid and pci probes, incubation was performed for 8 h with the pci probe alone as just described, then the cosmid probe treated as previously described (Toledo *et al.*, 1992, 1993) was added and incubation was continued for a further 16 h. Biotinylated probes were developed with alternating layers of fluoresceinated avidin (Vector Labs) and biotin-conjugated goat anti-avidin antibody (Vector Labs); digoxigenin-labelled probes were revealed by successive treatments with mouse anti-digoxigenin (Sigma), rhodamine-conjugated rabbit anti-mouse (Sigma) and rhodamine-conjugated goat anti-rabbit antibodies (Sigma) as previously described (Toledo *et al.*, 1992). Chromosomes were stained with 0.1 mg/ml of DAPI in phosphate-buffered saline (PBS) and mounted in Vectashield (Biosys S.A.). Slides were observed with an Axiophot (Zeiss) equipped with a precision pass-band filter, the Zeiss filter combinations 487709 and 487715 and a double band-pass filter (Omega opticals) or the Zeiss filter combinations 487701, 487709 and 487715. Image processing was achieved using Photoshop.

Giemsa-trypsin banding was performed by incubating the slides for 12 s in trypsin 0.8 mg/ml in 1× PBS at 37°C. The slides were then rinsed quickly in 1× PBS, stained for 7 min in 3% Giemsa, pH 6.8 and washed in water.

Acknowledgements

We thank Drs B.Dujon for making available to us the I-SceI gene and the protein recognition sequence, and J.Hamlin for the cosmids allowing probing of the *dhfr* locus. We also thank Drs M.C.Weiss and C.Fairhead for critical reading of the manuscript. This work was supported in part by the Université Pierre et Marie Curie, the Ligue Nationale Française contre le Cancer (Comité de Paris), the Association pour la Recherche sur le Cancer, the E.E.C. (Biomed II, CT-96-0894 to M.D.) and the G.R.E.G. (décision no. 86/94). E.P. was supported by a CNRS Foreign Researcher contract (no. 6151462) and A.C. is a Fellow of the MENRS.

References

Berinstein,N., Pennell,N., Ottaway,C.A. and Shulman,M.J. (1992) Gene replacement with one-sided homologous recombination. *Mol. Cell. Biol.*, **12**, 360–367.

Besnard,C., Monthioux,E. and Jami,J. (1987) Selection against expression of the *Escherichia coli* gene gpt in hprt+ mouse teratocarcinoma and hybrid cells. *Mol. Cell. Biol.*, **7**, 4139–4141.

Brison,O. (1993) Gene amplification and tumor progression. *Biochim. Biophys. Acta*, **1155**, 25–41.

Choo,K.B., Chen,C.M., Han,C.P., Cheng,W.T. and Au,L.C. (1996) Molecular analysis of cellular loci disrupted by papillomavirus 16 integration in cervical cancer: frequent viral integration in topologically destabilized and transcriptionally active chromosomal regions. *J. Med. Virol.*, **49**, 15–22.

Choulika,A., Perrin,A., Dujon,B. and Nicolas,J.F. (1995) Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **15**, 1968–1973.

Colleaux,L., D'Auriol,M., Betermier,M., Cottarel,G., Jacquier,A., Galibert,F. and Dujon,B. (1986) A universal code equivalent of a yeast mitochondrial intron reading frame is expressed into *Escherichia coli* as a specific double strand endonuclease. *Cell*, **44**, 521–533.

Colleaux,L., D'Auriol,L., Galibert,F. and Dujon,B. (1988) Recognition and cleavage site of the intron-encoded omega transposase. *Proc. Natl Acad. Sci. USA*, **85**, 6022–6026.

Colleaux,L., Rougeulle,C., Avner,P. and Dujon,B. (1993) Rapid physical mapping of YAC inserts by random integration of I-SceI sites. *Hum. Mol. Genet.*, **2**, 265–271.

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.

Debatisse,M. and Buttin,G. (1977) The control of cell proliferation by preformed purines: a genetic study. I. Isolation and preliminary characterization of Chinese hamster lines with single or multiple defects in purine 'salvage' pathways. *Somatic Cell Genet.*, **3**, 497–511.

Deininger,P.L. (1989) SINEs: short interspersed repeated DNA elements in higher eucaryotes. In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA.* American Society for Microbiology, Washington, DC, pp. 619–636.

Dujon,B. (1989) Group I introns as mobile genetic elements: facts and mechanistic speculations–a review. *Gene*, **82**, 91–114.

Ellis,J. and Bernstein,A. (1989) Gene targeting with retroviral vectors: recombination by gene conversion into regions of nonhomology. *Mol. Cell. Biol.*, **9**, 1621–1627.

Fairhead,C. and Dujon,B. (1993) Consequences of unique double-stranded breaks in yeast chromosomes: death or homozygosis. *Mol. Gen. Genet.*, **240**, 170–178.

Haber,J.E. (1992) Exploring the pathways of homologous recombination. *Curr. Opin. Cell Biol.*, **4**, 401–412.

Hutchison,C.A.,III, Hardies,S.C., Loeb,D.D., Shehee,W.R. and Edgell,M.H. (1989) LINEs and related retroposons: long interspersed repeated sequences in the eucaryotic genome. In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 593–617.

Jones,C. *et al.* (1995) Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene CBL2. *Nature*, **376**, 145–149.

Liang,F., Romanienko,P.J., Weaver,D.T., Jeggo,P.A. and Jasin,M. (1996) Chromosomal double-strand break repair in Ku80-deficient cells. *Proc. Natl Acad. Sci. USA*, **93**, 8929–8933.

Lin,F.L., Sperle,K. and Sternberg,N. (1984) Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Mol. Cell. Biol.*, **4**, 1020–1034.

Lukacsovich,T., Yang,D. and Waldman,A.S. (1994) Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. *Nucleic Acids Res.*, **22**, 5649–5657.

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.

Mauillon,J.L. *et al*. (1996) Identification of novel germline hMLH1 mutations including a 22 kb Alu-mediated deletion in patients with familial colorectal cancer. *Cancer Res.*, **56**, 5728–5733.

McClintock,B. (1951) Chromosome organization and genic expression. *Cold Spring Harbor Symp. Quant. Biol.*, **16**, 13–47.

Mulligan,R.C. and Berg,P. (1981) Selection for animal cells that express the *Escherichia coli* gene coding for xanthine–guanine phosphoribosyltransferase. *Proc. Natl Acad. Sci. USA*, **78**, 2072–2076.

Nehls,M., Schorpp,M. and Boehm,T. (1995) An intragenic deletion in the human PTPN6 gene affects transcriptional activity. *Hum. Genet.*, **95**, 713–715.

Nystrom-Lahti,M. *et al.* (1995) Founding mutations and Alu-mediated recombination in hereditary colon cancer. *Nature Med.*, **1**, 1203–1206.

Ohta,M. *et al.* (1996) The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3; 8) breakpoint, is abnormal in digestive tract cancers. *Cell*, **84**, 587–597.

Pinkel,D., Landegent,J., Collins,C., Fuscoe,J., Segraves,R., Lucas,J. and Gray ,J. (1988) Fluorescence *in situ* hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. *Proc. Natl Acad. Sci. USA*, **85**, 9138–9142.

Plessis,A., Perrin,A., Haber,J.E. and Dujon,B. (1992) Site specific recombination determined by I-Sce-1, a mitochondrial intron-encoded endonuclease expressed in the yeast nucleus. *Genetics*, **130**, 451–460.

Popescu,N.C. (1994) Chromosome fragility and instability in human cancer *Crit. Rev. Oncogen.*, **5**, 121–140.

Poupon,M.F., Smith,K.A., Chernova,O.B., Gilbert,C. and Stark,G.R. (1996) Inefficient growth arrest in response to dNTP starvation stimulates gene amplification through bridge–breakage–fusion cycles. *Mol. Biol. Cell*, **7**, 345–354.

Puchta,H., Dujon,B. and Hohn,B. (1993) Homologous recombination in plant cells is enhanced by *in vivo* induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res.*, **21**, 5034–5040.

Puchta,H., Dujon,B. and Hohn,B. (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc. Natl Acad. Sci. USA*, **93**, 5055–5060.

Rassool,F.V., McKeithan,T.W., Neilly,M.E., van Melle,E., Espinosa,R.d. and Le Beau,M.M. (1991) Preferential integration of marker DNA into the chromosomal fragile site at 3p14: an approach to cloning fragile sites. *Proc. Natl Acad. Sci. USA*, **88**, 6657–6661.

Roth,D. and Wilson,J. (1988) Illegitimate recombination in mammalian cells. In Kucherlapati,R. and Smith,G.R.(eds), *Genetic Recombination*. American Society for Microbiology, Washington, pp. 621–653.

Rouet,P., Smih,F. and Jasin,M. (1994a) Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl Acad. Sci. USA*, **91**, 6064–6068.

Rouet,P., Smih,F. and Jasin,M. (1994b) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.*, **14**, 8096–8106.

Sargent,R.G., Brenneman,M.A. and Wilson,J.H. (1997) Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. *Mol. Cell. Biol.*, **17**, 267–277.

Schimke,R.T., Kaufman,R.J., Alt,F.W. and Kellems,R.F. (1978) Gene amplification and drug resistance in cultured murine cells. *Science*, **202**, 1051–1055.

Segal,D.J. and Carroll,D. (1995) Endonuclease-induced, targeted homologous extrachromosomal recombination in *Xenopus* oocytes *Proc. Natl Acad. Sci. USA*, **92**, 806–810. [Published erratum appears in *Proc. Natl Acad. Sci. USA* 1995 Apr 11; **92**(8):3632.]

Smidt,M., Kirsch,I. and Ratner,L. (1990) Deletion of Alu sequences in the fifth c-*sis* intron in individuals with meningiomas. *J. Clin. Invest.*, **86**, 1151–1157.

Smith,K.A., Stark,M.B., Gorman,P.A. and Stark,G.R. (1992a) Fusion near telomeres occurs very early in the amplification of CAD genes in Syrian hamster cells. *Proc. Natl Acad. Sci. USA*, **89**, 5427–5431.

Smith,P.P., Friedman,C.L., Bryant,E.M. and McDougall,J.K. (1992b) Viral integration and fragile sites in human papillomavirus-immortalized human keratinocyte cell lines. *Genes Chromosomes Cancer*, **5**, 150–157.

Sozzi,G. *et al*. (1996) The FHIT gene 3p14.2 is abnormal in lung cancer. *Cell*, **85**, 17–26.

Stark,G.R. (1993) Regulation and mechanisms of mammalian gene amplification. *Adv. Cancer Res.*, **61**, 87–113.

Tadokoro,K. *et al*. (1992) Intragenic homozygous deletion of the WT1 gene in Wilms' tumor. *Oncogene*, **7**, 1215–1221.

Takebe,Y., Seiki,M., Fujisawa,J., Hoy,P., Yokota,K., Arai,K., Yoshida,M. and Arai,N. (1988) SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol. Cell. Biol.*, **8**, 466–472.

Tkachuk,D.C. *et al*. (1990) Detection of bcr–abl fusion in chronic myelogeneous leukemia by *in situ* hybridization. *Science*, **250**, 559–562.

Toledo,F., LeRoscouet,D., Buttin,G. and Debatisse,M. (1992) Co-amplified markers alternate in megabase long chromosomal inverted repeats and cluster independently in interphase nuclei at early steps of mammalian gene amplification. *EMBO J.*, **11**, 2665–2673.

Toledo,F., Buttin,G. and Debatisse,M. (1993) The origin of chromosome rearrangements at early stages of *AMPD2* gene amplification in Chinese hamster cells. *Curr. Biol.*, **3**, 255–264.

Wilke,C.M., Hall,B.K., Hoge,A., Paradee,W., Smith,D.I. and Glover,T.W. (1996) FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: direct evidence for the coincidence of viral integration sites and fragile sites. *Hum. Mol. Genet.*, **5**, 187–195.

Received September 23, 1997; revised and accepted October 14, 1997