

## Direct cloning and analysis of DNA sequences from a region of the Chinese hamster genome associated with aphidicolin-sensitive fragility

Andrew H. Palin<sup>1</sup>, Ricky Critcher<sup>1</sup>, Daniel J. Fitzgerald<sup>2</sup>, John N. Anderson<sup>2</sup> and Christine J. Farr<sup>1,\*</sup>

<sup>1</sup>Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, UK

<sup>2</sup>Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA

\*Author for correspondence

Accepted 3 April; published on WWW 27 May 1998

### SUMMARY

Fragile sites are reproducibly expressed and chemically induced decondensations on mitotic chromosomes observed under cytological conditions. They are classified both on the basis of the frequency with which they occur (rare and common) and in terms of the chemical agent used to induce expression in tissue culture cells. Aphidicolin-sensitive common fragile sites appear to be ubiquitous in humans and other mammals and have been considered as candidates of pathological importance. Recently DNA from FRA3B, the most highly expressed constitutive fragile site in the human genome, has been cloned although as yet the cause of the underlying fragility has not been identified. In this study we describe the isolation, using a direct cloning approach, of DNA from a region of the Chinese hamster genome associated with aphidicolin-inducible fragility. Cells of a human-hamster somatic cell hybrid were transfected with a pSV<sub>2</sub>HPRT vector while exposed to aphidicolin, an inhibitor of DNA polymerases  $\alpha$ ,  $\delta$  and  $\epsilon$ . FISH analysis of stable transfectant clones revealed that the ingoing plasmid DNA had preferentially integrated into fragile site-containing chromosomal bands. Plasmid rescue was used to recover DNA sequences flanking one such integration site in the hamster genome. We demonstrate by FISH analysis of metaphase cells induced with aphidicolin that the rescued DNA is from a region of fragility on

Chinese hamster chromosome 2, distal to the *DHFR* locus. Analysis of the DNA sequences flanking the integration site revealed the overall A+T content of the 3,725 bp region sequenced to be 63.3%, with a highly [A]•[T]-rich 156 bp region (86.5%) almost adjacent to the integration site. Computational analyses have identified strong homologies to *Saccharomyces cerevisiae* autonomous replicating sequences (ARS), polypyrimidine tracts, scaffold attachment site consensus sequences and a 24 bp consensus sequence highly conserved in eukaryotic replication origins, all of which appear to cluster around the [A]•[T]-rich sequences. This domain also possesses structural characteristics which are common to both prokaryotic and eukaryotic origins of replications, in particular an unusually straight conformation of low thermal stability flanked either side by highly bent DNA segments. Further isolation and characterisation of DNA sequences from common fragile sites will facilitate studies into the underlying nature of these enigmatic regions of the mammalian genome, leading to a greater understanding of chromatin structure.

Key words: Common fragile site, Aphidicolin, Chinese hamster, Chromosome condensation, Origin of replication

### INTRODUCTION

Fragile sites are characterised cytogenetically as distinct sites on chromosomes at which gaps, breaks and other aberrations occur when cells are cultured under specific conditions or exposed to certain chemical agents. They appear to be an inherent and universal structural feature of the mammalian genome, having been identified not only in primates, but also in a growing list of species, including mouse, rat, hamster, cow, cat and dog (Lin et al., 1984; Djalali et al., 1987; Robinson and Elder, 1987; Smeets and van de Klundert, 1990; Stone et al., 1991). In the human genome fragile sites have been identified on every chromosome with the exception of chromosome 21 and are classified by frequency of expression as either rare or common (Sutherland, 1991). The rare sites are caused by

infrequent alleles and range in incidence from 1 in 40 to 1 in several thousand individuals. The common or constitutive fragile sites are considered to be homozygous alleles present at fixation in the population; however, the expression frequency of particular sites may vary between individuals, and between different tissue types (Hunt and Burgoyne, 1987; Murano et al., 1989; Caporossi et al., 1995). Both classes of fragile site are further subdivided according to the conditions of tissue culture required to induce their expression cytogenetically. Three types of rare fragile site have been described: folate-sensitive, distamycin A-inducible and BrdU-requiring. Amongst the common fragile sites, there are also three types: the largest group is sensitive to the DNA polymerase inhibitor aphidicolin, while a small number of other sites can be induced using either 5-azacytidine, or BrdU.

As yet, the only fragile sites to have been defined at the molecular level are in the rare folate-sensitive group FRAXA (Kremer et al., 1991; Oberle et al., 1991; Verkerk et al., 1991; Yu et al., 1991), FRAXE (Sutherland and Baker, 1992), FRAXF (Hirst et al., 1993), FRA11B (Jones et al., 1994) and FRA16A (Nancarrow et al., 1994) and a single member of the distamycin A-inducible class, FRA16B (Yu et al., 1997). The molecular basis for expression of all these sites is the expansion of a polymorphic repeat sequence. The five folate-sensitive sites characterised arise as the result of expansion of p[CCG]<sub>n</sub> repeats, which normally exhibit repeat-copy number polymorphism, with subsequent methylation of the expanded repeats and the adjacent CpG islands. Only recently has the first non-folate-sensitive fragile site, FRA16B, been characterised at the DNA level. This site is associated with a polymorphic [A]•[T]-rich 33 base pair minisatellite, a finding consistent with the nature of the agents that induce its cytogenetic expression (namely distamycin A and a series of related compounds which bind in the minor groove of [A]•[T]-rich DNA (Van Dyke et al., 1982).

While expanded repeats may be the common basis for rare fragile sites, as yet the molecular nature underlying any common fragile sites has not been elucidated. The only common site to have been cloned is the aphidicolin-sensitive site at human 3p14.2 (FRA3B) (Boldog et al., 1994; Wilke et al., 1994; Paradee et al., 1996) but no specific sequence responsible for fragility has been identified (Rassool et al., 1996; Boldog et al., 1997; Inoue et al., 1997). Moreover, it appears that breakage can occur across a very broad region spanning several hundred kilobases of DNA (Paradee et al., 1995, 1996; Ohta et al., 1996; Rassool et al., 1996; Wilke et al., 1996). These data suggest that the structure of the common fragile sites may be quite different from that of rare sites.

A further characteristic of chromosomal fragile sites is their highly recombinogenic nature. They are regions frequently associated with sister chromatid exchange events (Glover and Stein, 1987, 1988), the sites of viral integration (McDougall, 1979; Durnam et al., 1988), and recurring chromosomal rearrangements in human tumours (LeBeau and Rowley, 1984; Yunis and Soreng, 1984; Miro et al., 1987). It has also been demonstrated by Rassool et al. (1991) that under tissue culture conditions of fragile site induction ingoing plasmid DNA preferentially integrates into expressed decondensations. In this report we describe the application of this direct approach for recovering DNA from the vicinity of an aphidicolin-sensitive fragile site in the Chinese hamster genome. Analysis revealed that it is [A]•[T]-rich and does not contain the repeat motifs characteristic of rare fragile sites. However, DNA sequence motifs and structurally distinct regions associated with prokaryotic and eukaryotic origins of replication have been identified immediately adjacent to the site of integration.

## MATERIALS AND METHODS

### DNA constructs

All molecular biology manipulations were carried out using standard methods (Sambrook et al., 1989). The DNA construct used in this study (pHTAG3) is based on a pSV2 plasmid backbone into which has been inserted a human *HPRT* minigene (Reid, 1990). In addition the construct contains a 4.5 kbp 'tag' of genomic DNA from human

chromosome 17 inserted into the *Bam*HI site of the SV40 DNA. This 'tag' serves to increase the overall size of the plasmid, allowing direct visualisation of each integration site by FISH. A unique *Nde*I site in the pBR322-derived sequence was used to linearise the plasmid before transfection.

### Cell culture and transfection

Routine cell culture techniques were employed throughout this work. D113JA(6TGr) is a Chinese hamster ovary-human chromosome 6 somatic cell hybrid maintained in DMEM supplemented with 10% FCS, 1× Gibco BRL non-essential amino acids, 4 mM glutamine and with 1 mg ml<sup>-1</sup> geneticin (G418 sulphate) to maintain the human chromosome which has pSV2neo integrated in the region of 6q21. D113JA(6TGr) is sensitive to HAT medium following selection for mutations in the endogenous hypoxanthine phosphoribosyl transferase gene by incubation with 6-thioguanine. Don is a pseudodiploid Chinese hamster lung fibroblast cell line maintained in DMEM supplemented with 10% FCS (Hsu and Zenzes, 1964).

Fragile site expression was generated by the incubation of the appropriate cell line with either 0.2 or 0.4 µM aphidicolin supplemented to the culture medium for 24 hours. 10 µg linearised pHTAG3 was transfected into both normal and fragile site-expressing D113JA(6TGr) cells using a Bio-Rad Gene Pulser at 300 volts and 960 µF. 24 hours after electroporation, HAT medium was added to the culture and cells maintained to allow colony formation. Individual transfectants were isolated and maintained in medium supplemented with 1× HAT.

### Fluorescence in situ hybridisation

Metaphase chromosomes were prepared from D113JA(6TGr) and Don using standard cytogenetic techniques. DNA probes were biotin-labelled using the Gibco BRL BioNick™ labelling system and all hybridisations were blocked with Gibco BRL hamster C<sub>0</sub>t-1 DNA. All protocols were performed on a Hybaid OmniSlide thermal cycler and processed in a Hybaid OmniSlide Module. In situ hybridisations of pHTAG3 to DHTAG and CTAG transfectant metaphase spreads, and cD4QC1 to D113JA(6TGr) were performed as previously described (Farr et al., 1995).

Hybridisation of cD4QC1 to Don metaphase spreads was performed using tyramide (TSA™) in situ hybridisation (TISH) (Raap et al., 1995). Briefly, metaphase spreads were hybridised with biotinylated probes as described previously. The slides were then blocked with 100 µl 0.5% Marvel, 150 mM NaCl, 100 mM Tris-HCl, pH 7.5, for 30 minutes in a humidified chamber at 37°C before being incubated with 100 µl 1/100 streptavidin-horseradish peroxidase (HRP) conjugate for 30 minutes at 37°C and washed three times, for 5 minutes each, at room temperature in 100 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Tween-20 (TNT buffer). 100 µl of a 1/50 dilution of biotinyl tyramide (in 1× amplification buffer) was added to the slides and incubated at room temperature for 10 minutes before being washed three times in TNT buffer, for 5 minutes each, at room temperature. The slide was incubated with 100 µl of 1/500 avidin-FITC conjugate for 30 minutes at room temperature, before being washed three times in TNT buffer, for 5 minutes each, at room temperature.

Metaphase chromosomes were counterstained with 1 µg ml<sup>-1</sup> 4',6'-diamidino-2-phenyl-indole (DAPI) and 400 ng ml<sup>-1</sup> propidium iodide and mounted in 16 µl Vectorshield anti-fade solution (Vector Laboratories) before being analysed and digitally captured using a Nikon Optiphot 2 microscope and a Bio-Rad MRC 600 confocal scanning laser microscope. Images were captured using the Bio-Rad CoMOS version 7.0 system software and further manipulated under Adobe Photoshop 3.0.5.

### DNA preparation and gel electrophoresis

Genomic DNA from the appropriate cell lines was extracted as previously described (Sambrook et al., 1989). Routinely, 10 µg genomic DNA was enzymatically digested with 50 µg ml<sup>-1</sup> BSA, 1×

recommended reaction buffer and 10 units enzyme per  $\mu\text{g}$  DNA under recommended conditions. Genomic DNA digestions were electrophoresed in 0.5% agarose with 0.5 $\times$  TBE buffer at approximately 4 volts/cm. DNA molecular mass markers of appropriate size were employed on all gels.

The preparation and digestion of genomic DNA for pulsed field gel electrophoresis was performed under standard conditions. Electrophoresis was performed in 0.5 $\times$  TBE at 12°C using Bio-Rad high strength analytical grade agarose on the Bio-Rad CHEF-DRII system.

### Molecular cloning of DNA flanking plasmid integration sites

Plasmid rescue was performed on cell lines to isolate the DNA sequences flanking the integration of pHTAG3. Genomic DNA from the relevant cell line, and in parallel the parental cell line, was digested with either *Bam*HI or *Eco*RI and ligated under dilute conditions with 1 unit  $\mu\text{l}^{-1}$  T4 DNA ligase at 4°C for 12 hours. After circularisation, the DNA was ethanol precipitated and electroporated into DH5 $\alpha$  competent cells.

For each cell line three electrotransformations were performed: a control undigested, unligated genomic DNA to assay for the presence of episomal pHTAG3; a control digested, unligated genomic DNA; and a digested, ligated experimental sample of genomic DNA. Bacterial colonies were isolated and analysed by restriction digest analysis, with putative genomic DNA inserts being hybridised to genomic DNA Southern blots of the appropriate panel of cell lines.

Following plasmid rescue of genomic DNA flanking the pHTAG3 integration in DHTAG 1, an *Eco*RI-*Hind*III restriction fragment (EH9) was used to probe a cosmid library of Chinese hamster genomic DNA (a kind gift from Prof. M. Debatisse) generated by partial genomic digestion with *Mbo*I and ligation to the vector pTCF (Grosveld et al., 1982). Positive colonies were identified and analysed by both restriction analysis and FISH to confirm the correct identity of the clone.

### DNA sequencing and analyses

All DNA sequencing was performed with the ABI Prism Dye Terminator kit on an ABI 373A Automated Sequencer and analysed with the ABI software packages Factura version 1.2 and Sequence Navigator version 1.0.1 (Sequencing Facility, Dept of Genetics, University of Cambridge). The target sequence was amplified by polymerase chain reaction on a Perkin Elmer Cetus programmable thermocycling driblock (96°C, 30 seconds; 55°C, 15 seconds; 72°C, 4 minutes and 30 seconds; 25 cycles).

Other DNA amplification protocols were performed as standard. 200 mM each dNTP, 1 $\times$  reaction buffer, 1  $\mu\text{M}$  both forward and reverse primers, 1 U *Taq* polymerase and between 1 and 50 ng template DNA were mixed to a total volume of 25  $\mu\text{l}$ , 40  $\mu\text{l}$  mineral oil layered on the surface and sample cycled on a Perkin Elmer Cetus programmable thermocycling driblock at the appropriately determined conditions. Products were observed by agarose electrophoresis, under optimal conditions.

Sequence computational analyses of the data was achieved using the GCG package (Genetic Computer Group, Winsconsin, USA), MacVector 4.5 (Kodak) and with BLASTN 2.0.2 (Altschul et al., 1997) and FASTA 3.0 (Pearson and Lipman, 1988). The program for calculating DNA curvature from nucleotide sequences has been described (Eckdahl and Anderson, 1987, 1990). The program uses the parameters of the wedge model for bent DNA and assumes bending occurs only at AA/TT. The predicted path of the helical axis is converted into a set of three-dimensional coordinates and the ENDS ratio is calculated from these coordinates. The ENDS ratio is used as an index of predicted bending and is defined as the ratio of the contour length of a nucleotide segment to the shortest distance between the ends of the segment. Ends ratios were computed at a window width of 120 bp and a window step of 10 nucleotides. Ends ratios have been

shown to be reliable predictors of DNA curvature as measured by electrophoretic and cyclization studies (Eckdahl and Anderson, 1987; Wang et al., 1994; Albert et al., 1995). Duplex stability calculations were calculated using the program Thermodyn, as described by Natale et al. (1992). The method is based on nearest-neighbour dinucleotide free energy values and the computer output represents the magnitude of free energy required to unwind a 100 bp segment and a step window size of 1 bp.

## RESULTS

### Cytogenetic analysis of the parental cell line, D113JA(6TGr)

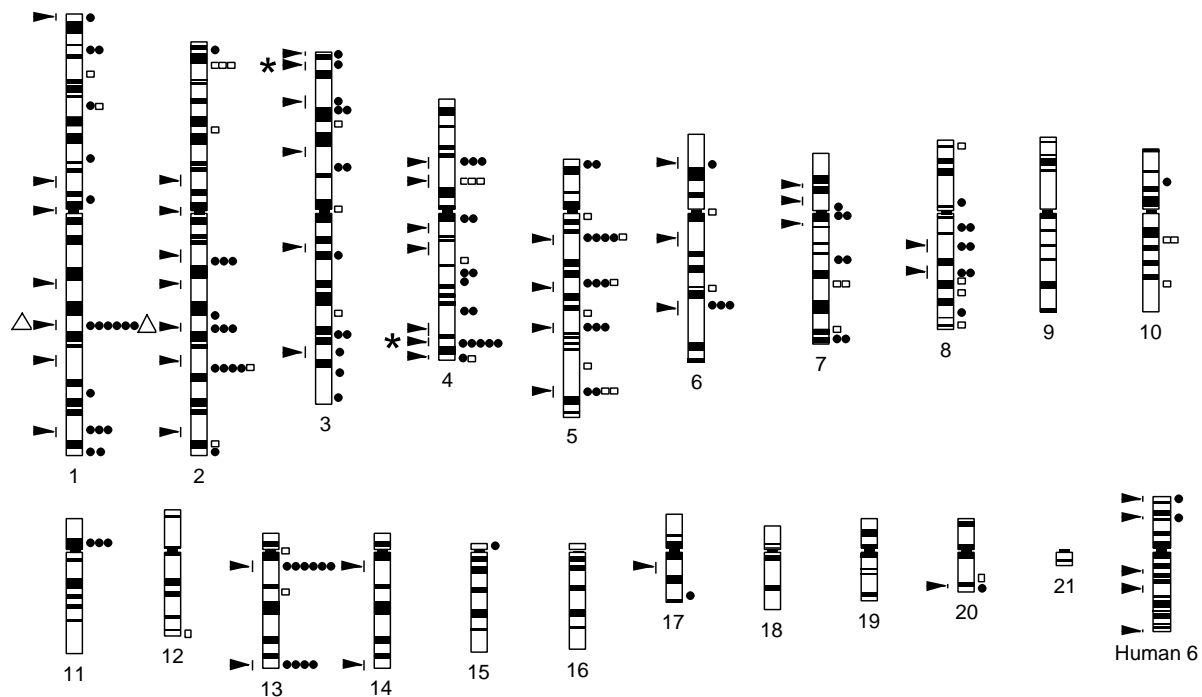
D113JA(6TGr) is an aneuploid Chinese hamster ovary (CHO)-human somatic cell hybrid which contains 21 hamster chromosomes and a single human chromosome 6. Homology between chromosomes 1 and 2 and between chromosomes 13 and 14 could be detected; however, the remainder of the genome is unidentifiable due to gross karyotypic rearrangements. Similarly, comparison of the D113JA(6TGr) karyotype with other Chinese hamster karyotypes allowed only limited identification of homologous chromosomes.

D113JA(6TGr) was G banded on both normal and fragile site-expressing metaphases. 386 Chinese hamster and 30 human chromosome 6 differentially staining bands were detected; an ideogram of the karyotype is shown in Fig. 1. Since this CHO-human hybrid has a highly aneuploid karyotype an arbitrary nomenclature has been devised: chromosomes are numbered according to decreasing size and individual light/dark bands are numbered xL/D, respectively, where x = the number of light/dark bands from the centromere. The exceptions to this rule of notation are human chromosome 6 and the Chinese hamster fragile sites previously reported at 1q31 and 2q31 (Rassool et al., 1991; Kuo et al., 1994).

Fragile sites scored from 200 metaphases were mapped onto the schematic karyotype. Eight frequently expressing loci were identified (Table 1). To identify the maximum number of aphidicolin-sensitive fragile sites capable of expression, D113JA(6TGr) was incubated in 0.4  $\mu\text{M}$  aphidicolin for 120 hours and all decondensations or aberrations attributable to fragile site expression were recorded. 48 expressing sites were mapped onto the D113JA(6TGr) karyotype after analysis of 250 metaphases. The locations of all fragile sites observed in D113JA(6TGr) cells, including 5 aphidicolin-sensitive fragile sites previously identified on human chromosome 6 (Glover et al., 1984; Yunis and Soreng, 1984; Sutherland, 1991) are shown in Fig. 1.

**Table 1. Expression frequencies of the most common aphidicolin-inducible fragile sites in D113JA(6TGr)**

Locus	Fragile sites scored	Expression frequency
1q3D	2	1%
1q4D	16	8%
1q7L	2	1%
5q3L	8	4%
5q5L	6	3%
12p1L	1	0.5%
15q2L	1	0.5%
13q3D	2	1%



**Fig. 1.** A schematic diagram showing the karyotype of D113JA(6TGr). Arrowheads indicate the locations of aphidicolin-induced fragile sites mapped within this genome. The asterisk indicates the chromosomal locations 3p5L and 4p9L and the open triangles indicate the homologous locations of 1q31 and 2q31. The positions of independent pHTAG3 integrations within the DHTAG and the CTAG panel of cell lines, as identified by FISH, are shown by solid circles and by open squares, respectively.

### Generation of panels of stable transfectants with and without pre-exposure to aphidicolin

A plasmid, pHTAG3, carrying an HPRT-minigene driven by the SV40 early promoter, together with several kilobases of genomic DNA originating from human chromosome 17p12.3, was linearised and electroporated into D113JA(6TGr). The recipient cell line, which is HPRT-deficient, was exposed to 0.2  $\mu$ M aphidicolin for 24 hours prior to transfection; in control transfections no aphidicolin pretreatment was administered. HAT-resistant stable transfectants were cloned and cell lines established. The sites of plasmid integration were localised using in situ hybridisation in 87 experimental cell lines (designated DHTAG) and in 37 control cell lines (designated CTAG). The distribution of the integration sites relative to the aphidicolin sensitive fragile sites previously assigned to this karyotype is shown in Fig. 1.

FISH identified one site of integration in the majority of both experimental and control cell lines (two or more sites of integration were observed in 20% of the experimental DHTAG lines and in 5% of the control set). However, Southern blot analysis revealed that in most experimental and control transfectants, multiple copies of the plasmid (ranging approximately between one and twenty copies) were present, either as a result of concatemerisation of the incoming DNA molecules and/or subsequent amplification events (data not shown). The distribution of integration sites in the experimental cell lines clearly demonstrates clustering within particular regions of the genome. For example, fourteen chromosomal bands have plasmid integrations associated with them in three or more independent transfectants. In particular the homologous loci at 1q31 and 2q31 are associated with

integration events in seven DHTAG cell lines, with two of the cell lines displaying reproducible hybridisation signals on both homologues. Intriguingly multiple integrations localised to two fragile site-containing chromosomal bands on chromosome 13q and appeared frequently to be associated with translocation and/or iso-chromatid duplications, events characteristic of an unstable region. In the control panel, integration events were infrequently found to cluster (two chromosomal bands displayed three integrations each). These rare clusters may be the result of coincidental events, actively transcribed regions of the genome serving as preferential integration targets, or spontaneously expressing fragile sites.

To determine whether the distribution of integration sites correlates with the sites of aphidicolin-induced fragility the data set was subjected to statistical analysis. Of 112 integration sites observed in the DHTAG panel, 67 were in chromosomal bands to which fragile sites had been localised (60%), whilst only 10/39 CTAG integrations were similarly linked (26%). Chi square analysis of these data indicate a highly significant bias towards integration into fragile site-expressing chromosomal bands ( $P < 0.001$  (1 df); Table 2), consistent with the idea that chemically induced decondensations of mammalian chromosomes can act as preferential sites for integration of exogenous DNA.

### Fluorescence in situ hybridisation of pHTAG3 across fragile sites

To assess the physical relationship between the sites of plasmid integration and aphidicolin-induced decondensations, fluorescein-biotin labelled pHTAG3 was hybridised to metaphases of selected, aphidicolin-induced,

**Table 2. Chi squared analysis of the pattern of pHTAG3 integration in experimental and control cell lines**

Incubation condition	Integrations into fragile regions	Integrations into non-fragile regions	Total
+ Aphidicolin (DHTAG)	67	45	<b>112</b>
– Aphidicolin (CTAG)	10	29	<b>39</b>
<b>Total</b>	<b>77</b>	<b>74</b>	<b>151</b>

$$\chi^2 = \frac{151 [(67 \times 29 - 45 \times 10) - 0.5 \times 151]^2}{112 \times 39 \times 74 \times 77}$$

$$\chi^2 = 12.1904 \text{ (1 df)}$$

This statistic is highly significant:  $P < 0.001$  (1 degree of freedom), indicating that aphidicolin-induced fragile site expression during transfection of pHTAG3 has generated a bias towards integration within fragile site-associated regions of the D113JA(6TGr) genome.

DHTAG cell lines. The cell lines chosen for study had integrations mapping to hamster chromosomes 1q31 (DHTAG 10, 13, 42, 54, 67, 81 and 98), 4q9L (DHTAG 1, 10, 54, 55 and 95), 4q10L (DHTAG 87) or to the human chromosome at 6p22 or 6p25 (DHTAG 31 and 86, respectively). Extended exposures to 0.4  $\mu$ M aphidicolin for 120 hours enhanced fragile site expression and allowed precise mapping of plasmid integration sites in relation to aphidicolin-induced chromosomal aberrations. In seven out of the thirteen cell lines examined pHTAG3 hybridisation signals were identified immediately adjacent to active fragile sites or breakpoints (DHTAG 1, 10, 31, 54, 55, 87 and 95). Furthermore, four of these seven cell lines clearly demonstrated hybridisation signals within a region of decondensation (DHTAG 10, 31, 55 and 87) (Fig. 2). The integration within DHTAG 10 (located at hamster chromosome 1q31) was shown to hybridise twice proximally, twice distally and once across a decondensation. DHTAG 31 generated hybridisation signals both distal to and within a decondensation in human 6p22. DHTAG 55 (integration site hamster 4q9L) showed pHTAG3 hybridisation proximal to (twice), distal to (fourteen times) and within (three times) a decondensation. These data therefore support the idea of preferential plasmid integration into fragile site-expressing regions of the Chinese hamster-human recipient cell line.

#### Plasmid rescue of DNA flanking an integration event associated with an aphidicolin-sensitive fragile site

Plasmid rescue was used to recover DNA flanking the 4q9L integration site in cell line DHTAG 1. Attempts to recover flanking DNA by plasmid rescue from a number of other transfectants failed, presumably due to the complexity of the integration events. Genomic DNA derived from a 4.3 kbp *Eco*RI plasmid rescued from DHTAG 1 was used to probe Southern blot filters of *Eco*RI-digested DNA from a CHO cell line, the parental somatic cell hybrid D113JA(6TGr), DHTAG 1 itself and from a series of other transfectants (including DHTAG 10, 54, 55 and 95) in which the integration sites had been assigned to the same chromosomal band. Fig. 3 clearly shows a ~3.8 kbp hamster-derived hybridisation band in all the

cell lines, while a second ~4.3 kbp *Eco*RI band is present in DHTAG 1. When this filter was rehybridised with a pBR322-derived probe, the disrupted 4.3 kbp *Eco*RI allele was found to co-localise with a pHTAG3-specific band in DHTAG 1 (data not shown). This confirms that genomic DNA flanking the plasmid integration site in DHTAG 1 has been recovered and that the plasmid integration events in the other cell lines of this 4q9L series are not associated with the same *Eco*RI restriction fragment. The ratio of the signal intensities from the normal and disrupted alleles in DHTAG 1 is close to 1, as expected (quantitated using a Packard Instant Imager).

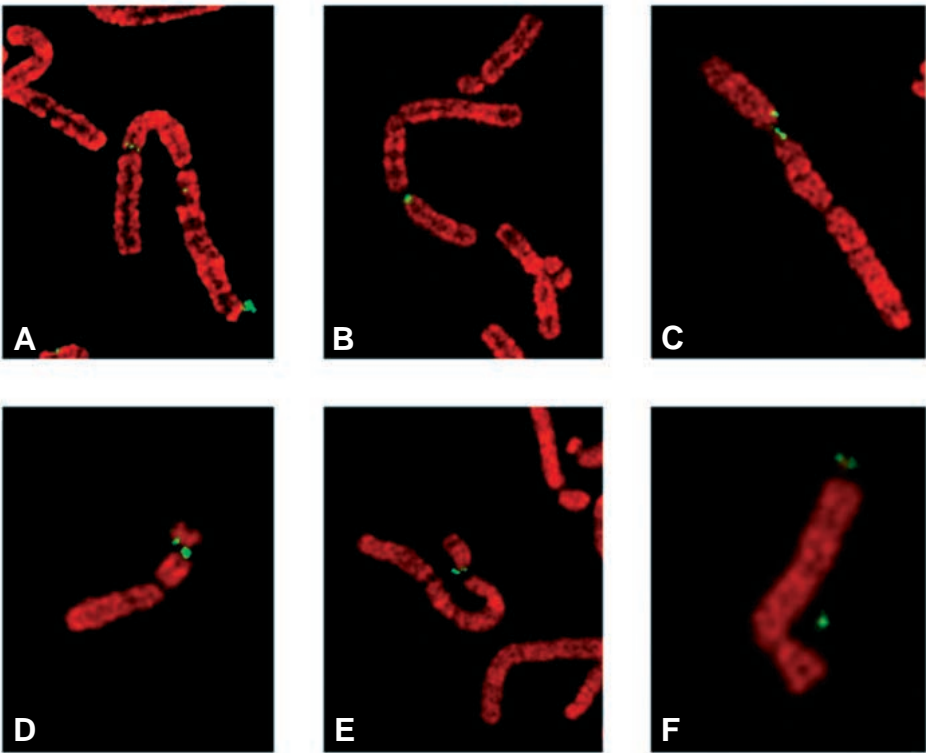
Pulsed field gel electrophoresis of the 4q9L series of DHTAG cell lines and Southern blot analysis using the rescued flanking DNA and pBR322-derived sequences as probes did not reveal any restriction fragments in common between DHTAG 10, 54, 55 or 95 and DHTAG 1 (with the exception of the pSV2neo-associated bands), indicating that the integration sites in this transfectant series are separated by at least 500 kbp of DNA (data not shown). Since fluorescence in situ hybridisation indicated a close relationship between the sites of plasmid integration and aphidicolin-induced decondensations in the 4q9L series this data suggests that several fragile sites, physically separated by many hundreds of kilobases of DNA, may be clustered within this particular chromosomal band.

#### Fluorescence in situ hybridisation using genomic DNA cloned directly from an aphidicolin-sensitive fragile region

An *Eco*RI-*Hind*III restriction fragment (EH9) derived from the plasmid-rescued DNA of DHTAG 1 was used to screen a Chinese hamster cosmid library. Cosmid cD4QC1 was used as a probe in FISH of normal metaphases from the parental cell line D113JA(6TGr) (Fig. 4A). The 4q9L locus associated with the integration event in DHTAG 1 was correctly identified, confirming that the rescued DNA is associated with this fragile region. In addition a second locus, designated hamster chromosome 3p5L, also hybridised reproducibly with the cosmid probe, indicating that this is the homologous chromosomal region. (Further cytogenetic analysis of the terminal regions of these 2 chromosome arms has confirmed that they are similar in their banding patterns and frequency of fragile site expression, although the extent of the chromosomal homology has not been delineated).

To determine the relationship between the genomic cosmid and the site of decondensation, cD4QC1 was biotin-fluorescein labelled and hybridised to fragile site-expressing metaphases following a 120 hour incubation with 0.4  $\mu$ M aphidicolin. The cosmid was found to hybridise to both the proximal ( $n=11$ ) and distal ( $n=7$ ) sides of induced decondensations, and occasionally ( $n=7$ ) within the gap present at either hamster 3p5L or 4q9L (Fig. 4B). These data are summarised in Table 3. Hybridisation to other loci within the D113JA(6TGr) genome was occasionally observed following fragile site induction. These hybridisation signals may be due to other sites having weak homology to the probe; alternatively, they may represent low frequency rearrangement events between the original locus and other regions of the hamster genome. An interesting observation is that 3 of the 6 unassigned hybridisation signals were associated with fragile site expression. If rearrangements have occurred then they appear

**Fig. 2.** Propidium iodide-stained metaphase chromosomes expressing aphidicolin-induced fragile sites following hybridisation with fluorescein-biotin labelled pHTAG3. (A) Chromosome 1 from DHTAG 10 demonstrating pHTAG3 hybridisation proximal to the expressing fragile site at 1q31. (B) Chromosome 1 from DHTAG 10 demonstrating pHTAG3 hybridisation distal to the same expressing fragile site. (C) Chromosome 1 from DHTAG 10 demonstrating pHTAG3 hybridisation across the aphidicolin-induced decondensation at 1q31. (D) pHTAG3 hybridisation within the decondensed/fragile region at human chromosome 6p22.2. (E) Chromosome 4 from DHTAG 55 demonstrating pHTAG3 hybridisation within the fragile site. (F) Chromosome 4 from DHTAG 95 showing an aphidicolin-induced decondensation immediately proximal to the pHTAG3 integration and terminal deletion of the chromosome.

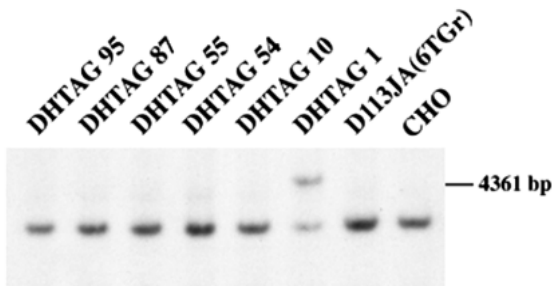


to have either created de novo fragile sites, or to have translocated fragile site-causing DNA, which would therefore have to be present within cD4QC1.

FISH of cD4QC1 was repeated on Don cells, a male-derived pseudodiploid Chinese hamster cell line. In this cell line it was possible to assign cD4QC1 to a diploid homologous region on the metacentric hamster chromosome 2 (Fig. 5A). Subsequently the distally located cD4QC1 signal was shown to be on the same chromosome arm as the *DHFR* gene (data not shown). Details of FISH to fragile site-expressing metaphases prepared after 24 hour incubation with 0.4  $\mu$ M aphidicolin are summarised in Table 4 and shown in Fig. 5B,C. Fifteen FISH signals were associated with an expressing

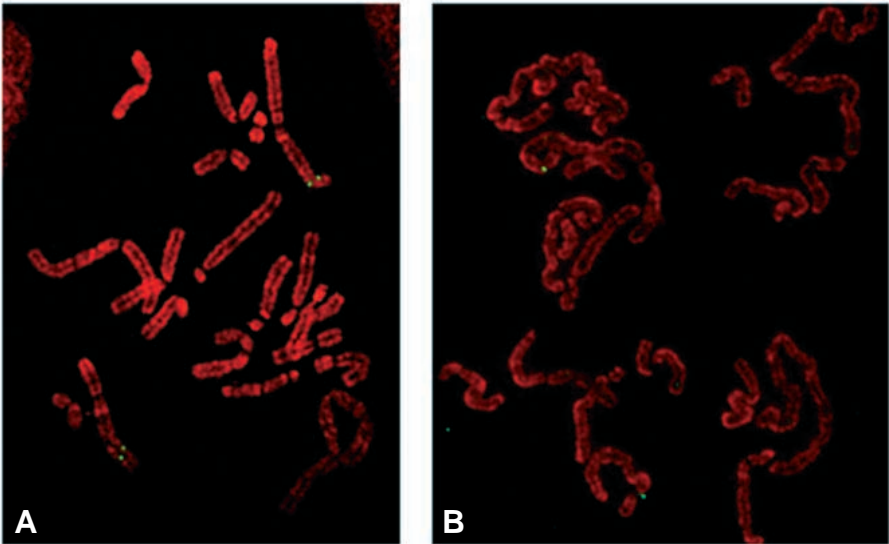
fragile site in the region, with the signal detected proximal to the decondensation nine times, distally twice and across the gap three times. Also observed was a terminal deletion of chromosome 2, with the hybridisation signal adjacent to the new chromosome end. Therefore, it can be concluded that the cosmid cD4QC1, and by implication the original plasmid integration event, is associated with an aphidicolin-sensitive fragile region present in two independent Chinese hamster cell lines.

The expression frequency of this particular fragile site appears to be significantly elevated in the Don cell line compared with the original somatic cell hybrid line, D113JA(6TGr). We estimate that 10.5% of Don metaphase chromosomes hybridising with the cD4QC1 probe showed a decondensation either adjacent to, or tightly associated with the hybridisation signal following a 24 hour aphidicolin exposure. In comparison, expression of the appropriate fragile site was detected in only 3-4% of cD4QC1-positive D113JA(6TGr) metaphase spreads following an extended 120 hour aphidicolin exposure (with even fewer being identified following a 24 hour exposure).



**Fig. 3.** An autoradiograph showing the endogenous 3.8 kb *Eco*RI allele following hybridisation to the probe EH9 in both controls (CHO and D113JA(6TGr)) and all cell lines mapping to 4q9L and 4q10L. In cell line DHTAG 1, from which the EH9 DNA was plasmid rescued, the 3.8 kb band is half the intensity of that in the control lanes. In addition, only in the DHTAG 1 cell line does the probe detect a new 4.3 kb fragment (the same size as the rescued plasmid). This therefore represents the allele into which pHTAG3 has integrated.

Table 3. Relationship between cD4QC1 and the aphidicolin-induced fragile sites at 3p5L and 4q9L in D113JA(6TGr)				
Location of cD4QC1 hybridisation	Signals scored	Fragile site-associated signals		
		Proximal	Distal	Within
Chromosome 3p5L	155	5	1	3
Chromosome 4q9L	165	6	2	3
Unassigned signals	6	2	0	1
<b>Total</b>	<b>388</b>	<b>20 (+3 unassigned loci)</b>		

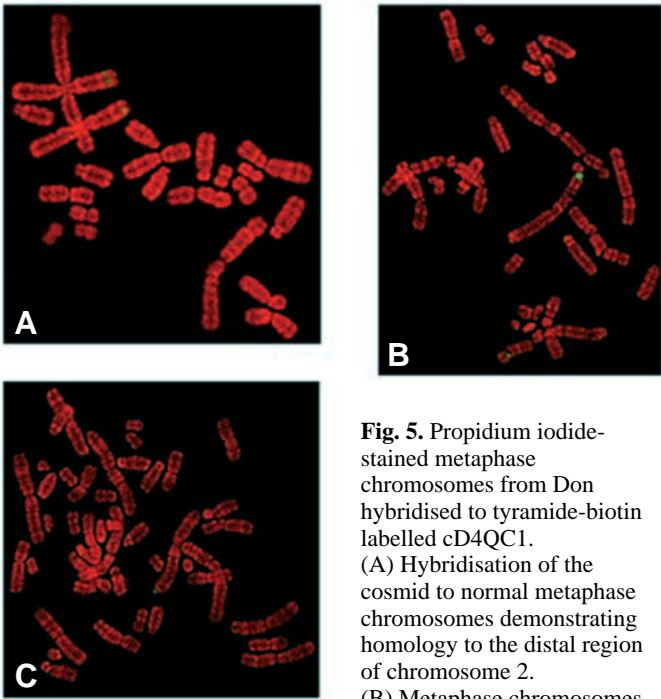


**Fig. 4.** Propidium iodide-stained metaphase chromosomes from D113JA(6TGr) hybridised to fluorescein-biotin labelled cD4QC1. (A) FISH of the cosmid to normal metaphase chromosomes demonstrating homology to the chromosomal loci 3p5L and 4q9L. (B) Metaphase chromosomes expressing aphidicolin-induced fragile sites showing the cosmid hybridising to the proximal edges of decondensations present on 3p5L and 4q9L.

**Sequence analysis of DNA associated with an aphidicolin-sensitive fragile region**

A 3,725 bp *Eco*RI subclone of cosmid cD4QC1, which encompasses the plasmid integration site in cell line DHTAG 1, was sequenced and the data analysed for database similarities by BlastN 2.0.2 and FastA 3.0 searches of GenBank (release 105) and EMBL (release 53). No

significant homologies were identified, except for matches to SINE- and LINE-like elements (Fig. 6). No matches to the ~300 kb of DNA sequence from the FRA3B locus were identified (Boldog et al., 1997; Inoue et al., 1997). Interestingly, however, the area in the vicinity of the integration site displays multiple features that are characteristic of eukaryotic origins of replication (Fig. 7). The overall A+T content of the fragment is 63.6% with a highly [A]•[T]-rich 156 bp region (86.5%) immediately adjacent to the integration site itself (Fig. 7A). The [A]•[T]-rich region displays low predicted duplex stability and is flanked on both sides by intrinsically curved DNA which arises from oligo A-tracts spaced in a 10-11 bp period (Fig. 7B,C). Curvature is expressed as the ENDS ratio which is a measure of predicted bending. The magnitude of the two highest ENDS ratio peaks shown in the figure is 2-3 standard deviations greater than the mean ENDS ration of all sequences in the GenEMBL database, as well as random sequence DNA with this base composition (Fitzgerald et al., 1994). The [A]•[T]-rich region is characterised by ENDS ratios which are significantly lower than predicted from base composition. Low ENDS ratios in high [A]•[T] content DNA should give rise to an unusually straight molecule that is structurally distinct from non-curved bulk DNA (Eckdahl and Anderson, 1987). In agreement with this interpretation, autocorrelation analysis revealed that the distributions of oligo A-tracts was decidedly non-random and opposite to that which is characteristic of curved DNA (data not shown). An [A]•[T]-rich domain of low curvature and low duplex stability flanked by curved sequences is a conserved



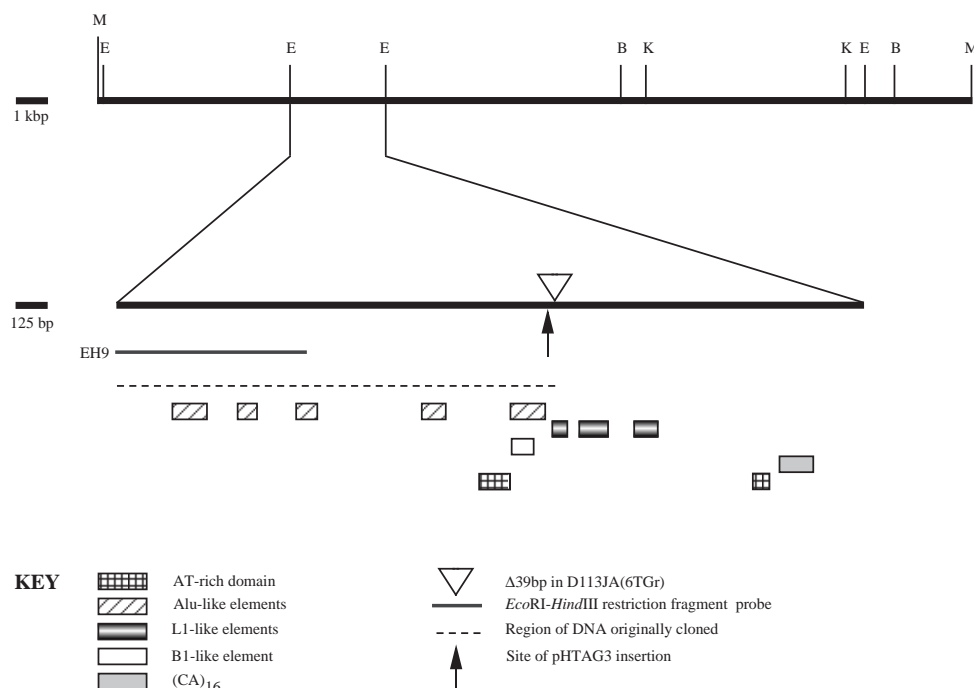
**Fig. 5.** Propidium iodide-stained metaphase chromosomes from Don hybridised to tyramide-biotin labelled cD4QC1. (A) Hybridisation of the cosmid to normal metaphase chromosomes demonstrating homology to the distal region of chromosome 2. (B) Metaphase chromosomes expressing aphidicolin-induced

fragile sites following hybridisation with cD4QC1. The cosmid has hybridised proximal to, and immediately adjacent with the induced decondensations. (C) The terminally deleted chromosome 2 (indicated by an arrow) is expressing an aphidicolin-induced fragile site with hybridisation of the cosmid to the distal edge of the decondensation.

**Table 4. Relationship between cD4QC1 and the aphidicolin-induced fragile site on Don chromosome 2**

Location of cD4QC1 hybridisation	Fragile site-associated signals
Distal to the decondensation	2
Proximal to the decondensation	9
Across/within the decondensation	3
Terminal hybridisations	1
<b>Total of scored metaphases</b>	<b>152</b>

**Fig. 6.** A schematic diagram of cD4QC1, showing in detail the 3.7 kb *EcoRI* restriction fragment flanking the pHTAG3 integration site (indicated by the open triangle). Selected enzyme restriction sites are shown (B=*Bam*HI; E=*Eco*RI; K=*Kpn*I; M=*Mlu*I). The dashed line shows the region of genomic DNA originally cloned by plasmid rescue and the solid line indicates the *EcoRI*-*Hind*III restriction fragment (EH9) used as a probe in Fig. 3 and during the screening of the cosmid library. Motifs of interest in the *EcoRI* allele, as defined by BlastN, Findpatterns and sequence analysis, are indicated in their relative positions by appropriately shaded bars (see key). The vertical arrow indicates the 39 bp deletion identified in D113JA(6TGr).



feature of replication origins (Eckdahl and Anderson, 1987, 1990). The [A]•[T]-rich region shown in Fig. 7 is enriched in short sequence motifs which have been implicated in replication activity (reviewed by Dobbs et al., 1994; Boulikas, 1996). These include multiple and frequently overlapping ARS and SAR consensus sequences, a long ( $n=39$ ) pyrimidine tract and a replication origin consensus sequence which was previously found with one mismatch in 6 of 6 origins from higher eukaryotes characterized by Dobbs et al. (1994). The observed frequency of each of the sequence motifs in the [A]•[T]-rich region was several fold higher than the expected occurrence due to chance as was determined by probability calculations that take into account the biased base composition of the region; see legend to Fig. 7 (Eckdahl and Anderson 1990).

Sequence analysis of the *EcoRI* plasmid rescued clone derived from cell line DHTAG 1 (which contains 2,207 bp of DNA flanking one side of the pHTAG3 integration site) revealed a difference between the DHTAG 1-derived DNA itself and the cosmid clone cD4QC1. PCR was used to recover DNA from this region in both the D113JA(6TGr) and Don genomes. Sequence analysis confirmed that D113JA(6TGr) has a 39 bp deletion compared with both the Chinese hamster genome used to construct the cosmid library and with the Don cell line. This deletion/insertion is immediately adjacent to the site of plasmid integration in cell line DHTAG 1 and is particularly [A]•[T]-rich (72%) (Fig. 6).

Restriction mapping of this particular aphidicolin-sensitive fragile site using rare-cutter enzymes on genomic DNA from the parental cell line D113JA(6TGr) failed to identify any CpG islands in the vicinity. The presence of [CCG]<sub>n</sub>, [CAG]<sub>n</sub> and [TTAGGG]<sub>n</sub> repeat tracts within cosmid cD4QC1 was investigated by Southern blotting using oligonucleotide probes and appropriate positive control DNAs. No hybridisation to cD4QC1 DNA was observed with any of the probes (data not

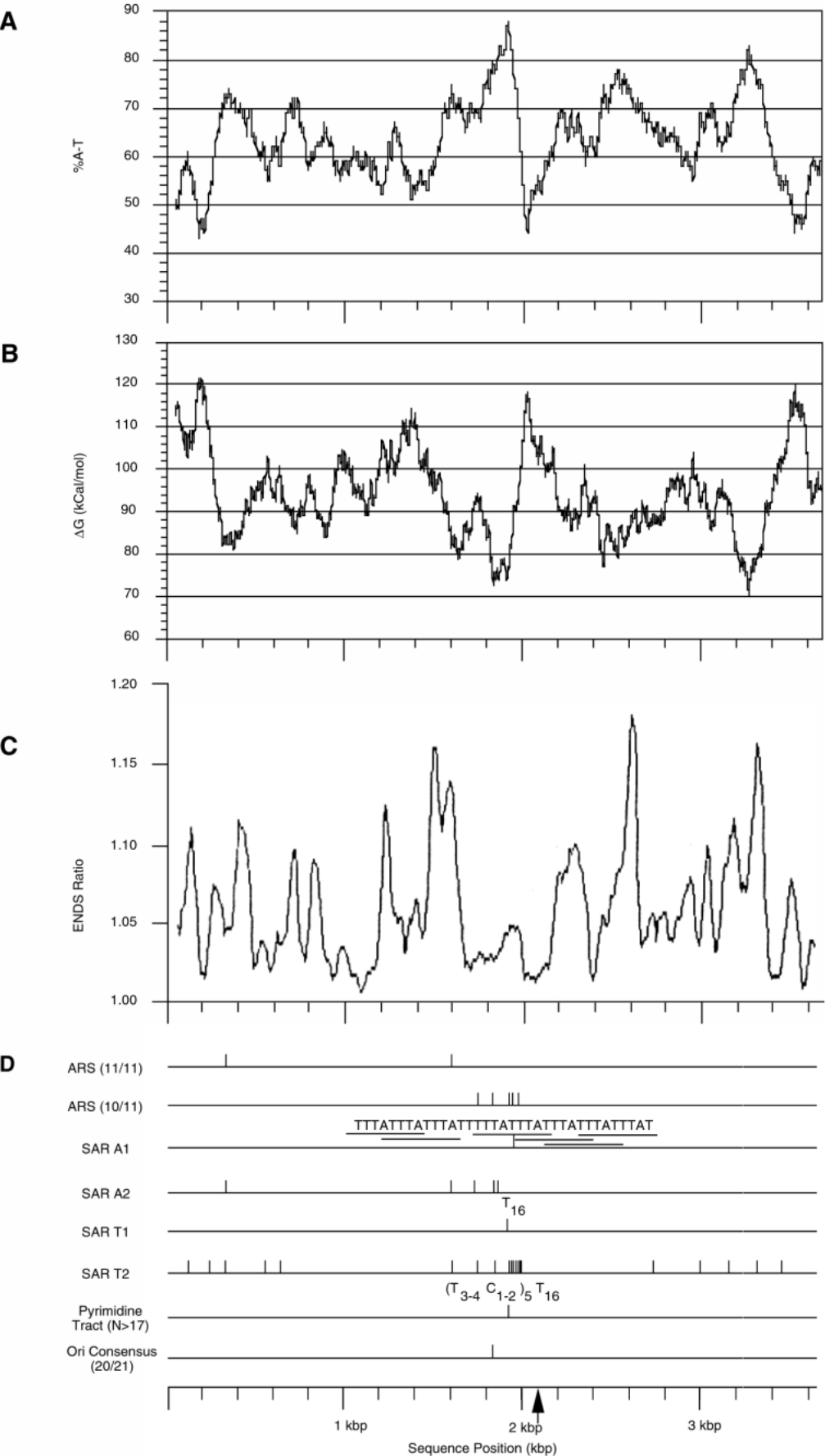
shown); it was concluded that there are no stretches of near perfect or perfect homology to these oligonucleotides within the cosmid.

## DISCUSSION

Fragile sites are known to lead to an increased rate of sister chromatid exchange and recombination in their vicinity (Rainaldi et al., 1984; Glover and Stein 1987, 1988; Gaddini et al., 1995). This suggests that fragile sites are involved with DNA breakage at some point during expression and evidence has accumulated that fragile sites can present as sites of chromosomal breakage in vivo (Jones et al., 1995; Ong et al., 1997; Zimonjic et al., 1997). In this report, a strategy for the direct cloning of common fragile sites, which exploits their highly recombinogenic nature, has been used to isolate DNA associated with a mammalian aphidicolin-sensitive fragile region. A statistically significant bias towards integration of the ingoing plasmid DNA into fragile site-associated chromosomal bands has been demonstrated with, in some instances, multiple independently derived cell lines having integrations into the same fragile site-associated band.

To assess the physical relationship between the integrations and expressing fragile sites in the immediate vicinity, the introduced plasmid DNA was itself labelled and hybridised to metaphases from selected, aphidicolin-induced, transfectants. The identification of signals which simultaneously hybridised to, or alternated between, both sides of a chromosomal breakpoint implied that the plasmid had integrated within a region susceptible to aphidicolin-induced decondensation. Plasmid rescue was used to recover 2.2 kbp of DNA flanking the introduced construct in one cell line, DHTAG 1. A cosmid, isolated using unique DNA from this rescued plasmid, was used as a probe in FISH experiments on both normal and

**Fig. 7.** Structural and sequence features of the 3.7 kbp *EcoRI* allele into which pHTAG3 integrated. (A) A+T content as a function of sequence position at a window width of 120 bp. (B) Duplex stability calculations were determined using the Thermodyn program (Natale et al., 1992). Each point represents the magnitude of free energy required to unwind a 100 bp duplex with a step interval of 1 bp. (C) ENDS ratios as a function of sequence position were computed at a window width of 120 bp. The ENDS ratio is defined as the ratio of the contour length of a DNA segment to the shortest distance between the ends of the segment. (D) Consensus sequences were identified using the GCG Findpatterns program with either 0 or 1 mismatches. Sequences are from Dobbs et al. (1994) and are: ARS WTTTATRTTTW; SAR A1 AATAAAYAAA; SAR A2 WADAWAYAWW; SAR T1 TTWTWTTWTT; SAR T2 TWWTDDTTWW; Ori consensus WAWTTDDWWWDHWGWH MAWTT. The sequence containing the 7 overlapping SAR A1 elements (underlined) is shown. The observed frequency of sequence motifs between bps 1,590-2,090 ranged from 4- to 90-fold greater than the expected occurrence due to chance as determined by probability calculations that take into account the biased base composition of the region (Eckdahl and Anderson, 1990). For each sequence motif in this region the probability of observing at least the number of matches shown in the figure are: ARS 11/11 ( $P<0.1$ ); ARS 10/11 ( $P<0.02$ ); SAR A1 ( $P<0.001$ ); SAR A2 ( $P<0.03$ ); SAR T1 ( $P<0.001$ ); SAR T2 ( $P<0.001$ ). The arrow indicates the site of pHTAG3 insertion.



fragile site-expressing metaphase chromosomes prepared from the parental cell line. This confirmed that the genomic cosmid, cD4QC1, is intimately associated with the fragile site to which the integration event in DHTAG 1 had been assigned and with the homologous chromosomal locus. On aphidicolin-exposed metaphases from the highly aneuploid parental CHO-human somatic cell hybrid D113JA(6TGr), the cosmid was demonstrated to hybridise either proximal or distal to the decondensation (7 and 3 times each, respectively) and within the gap itself on the homologous chromosomal regions of hamster 3p5L and 4q9L (7 times). These data were confirmed using an independently derived male Chinese hamster cell line, Don. These results eliminate any possibility of novel decondensations being generated through the formation of abnormal DNA structures associated with the integration of concatameric plasmid molecules.

Thus, we have demonstrated cytogenetically that the DNA sequences responsible for this aphidicolin-sensitive fragile site are closely associated with, if not at least partially within, the cosmid clone cD4QC1. The reason for the variable position of the fluorescence hybridisation signal relative to the decondensation is unknown. It may indicate that the region of fragility is very broad, with breaks and gaps occurring throughout the region. Alternatively, there may be a cluster of independent, but tightly associated, fragile sites lying within a few hundred kilobases of each other. Another possibility is that there may be differences in the extent to which the cD4QC1-associated decondensation spreads on individual chromosomes, possibly as a result of the variable use of scaffold attachment sites in the structural organisation of the chromosome (Laemmli et al., 1992; Saitoh and Laemmli 1994). Intriguingly, the expression frequency of this fragile site was considerably higher in the male lung fibroblast-derived Don cell line compared to the CHO-human somatic cell hybrid (10% compared with 3.5%). Variability in tissue-specific expression frequencies has been reported previously and may reflect differences in gene activity or DNA replication of the region and/or the involvement of undefined transacting factors (Hunt and Burgoyne, 1987; Murano et al., 1989; Caporossi et al., 1995). A strong genetic influence on site expression has also been suggested from some studies (Wulf and Niebuhr, 1985; Smeets and Arets, 1990; Austin et al., 1992). Interestingly a difference in the primary DNA sequence at the site of plasmid integration has been identified between the D113JA(6TGr) and Don genomes. The significance of this 39 bp deletion/insertion is not known, but it may reflect the general instability of this region of the hamster genome.

Whilst genomic Southern blots and DNA sequence analysis clearly showed the presence of SINEs and [A]•[T]-rich stretches of DNA within cD4QC1 and the rescued plasmid clone itself, no sequences commonly associated with non-B form DNA structures (specifically [TTAGGG]<sub>n</sub>, [CCG]<sub>n</sub> or [CAG]<sub>n</sub> repeat elements) and no internal repetition has been identified. These data are in concordance with analysis of the human FRA3B region, where no causative expanded triplet repeat, methylated CpG island or [TTAGGG]<sub>n</sub> repeats have been identified. Recent sequence analysis of almost 300 kbp of DNA encompassing FRA3B has revealed that this is also a region high in A+T content (Boldog et al., 1997; Inoue et al., 1997). However, unlike the hamster common fragile site

reported here, the FRA3B region has frequent LINE repeats and is conversely low in SINE sequences (Boldog et al., 1997; Inoue et al., 1997; Wang et al., 1997).

The most striking aspect of the primary sequence data for this hamster fragile site is the high A+T content of the region. The highly [A]•[T]-rich DNA in the immediate vicinity of the integration site in DHTAG 1 displays marked similarities to eukaryotic origins of replication (Fig. 6). The region has numerous homologies to SARs and *S. cerevisiae* ARS consensus sequences, a long palindromic tract, and a 21 bp consensus sequence that is highly conserved in eukaryotic replication origins (Dobbs et al., 1994; Boulikas, 1996). These sequence motifs reside within an [A]•[T]-rich domain of unusually straight DNA with low duplex stability which is flanked by curved DNA segments. Although matches to SAR and ARS consensus motifs occur relatively frequently in the mammalian genome the clustering of several of these motifs together with these higher-order structural domains is unusual and characteristic of origins of replication in both prokaryotic and eukaryotic organisms (Eckdahl and Anderson, 1990; Dobbs et al., 1994; Boulikas 1996). The significance of these findings in terms of chromosome fragility, if any, is not yet known; an intriguing possibility is that aphidicolin-inducible fragile sites are associated with sites of attachment of the DNA to the chromosome scaffold and with sequences at which DNA replication is initiated.

These data, together with the DNA sequence analysis of the human FRA3B region, indicate that the mechanism underlying fragility associated with aphidicolin exposure is quite different from that associated with the rare fragile sites characterised to date. Unlike the rare fragile sites, common fragile sites may not correspond to unique DNA sequences, or to specific sites, and may extend for several kilobases. DNA sequence analysis of other constitutive fragile sites, isolated either through direct or positional cloning approaches, may help clarify the underlying nature of aphidicolin-sensitive fragile sites and their relationship to biological processes such as DNA replication or gene activity. However, a fuller understanding of these enigmatic phenomena may depend upon the generation of de novo aphidicolin-sensitive fragile sites following the integration of cloned candidate sequences (such as the one described in this report) ectopically into the mammalian genome. A similar approach has been used to demonstrate the DNA sequences underlying the Adenovirus type 12-induced fragile site at human 17q21-22 (Bailey et al., 1995; Gargano et al., 1995).

Finally, the contribution, if any, made by common fragile sites to genome rearrangements arising *in vivo* has for a long time been controversial (LeBeau and Rowley, 1984; Miro et al., 1987). Direct evidence that fragile sites can be activated *in vivo* and trigger specific chromosomal breaks, albeit infrequently, has been provided by the demonstration that the autosomal rare folate-sensitive site FRA11B is implicated in the genesis of Jacobsen syndrome, a chromosomal deletion syndrome (Jones et al., 1995). Moreover, the FRA3B region appears to be a hot-spot for chromosomal rearrangement and may be involved in variable deletions detected in the *FHIT* gene in tumours (Ohta et al., 1996; Boldog et al., 1997; Ong et al., 1997; Wang et al., 1998). A role for breakage at fragile sites has also been postulated in gene amplification in multidrug-resistant CHO cells (Kuo et al., 1994) and more

recently evidence that expression of fragile sites trigger intrachromosomal mammalian gene amplification and sets boundaries to early amplicons has been provided (Coquelle et al., 1997; Pipiras et al., 1998). Such a mechanism could have important implications for the development of drug resistance and gene amplification in tumour cells, since many of the drugs frequently used in human cancer therapy can be used to select for gene amplification in vitro. The particular fragile site isolated in this work has been shown to localise distal to the *DHFR* locus on hamster chromosome 2. Further studies of this region on chromosomes on which the *DHFR* locus is amplified may reveal whether breakage at this telomerically located fragile site is an initial event in triggering DNA amplification in mammalian cells.

A.H.P. was supported by a MRC-HGMP studentship and MRC project grant, R.C. is a MRC-funded research associate and C.J.F. is a MRC Senior Research Fellow. We thank our colleagues Tom Ebersole, Jonathan Konrad and Walter Mills for their interest and advice. Michelle Debatisse kindly provided the Chinese hamster total genomic cosmid library. Joyce Hamlin is thanked for cosmids spanning the *DHFR* locus. Jeffrey Trent is thanked for supplying us with the Chinese hamster-human somatic cell hybrid, D113JA. David Kowalski generously provided the Thermodyn sequence analysis software package. DNA sequence data has been deposited in the EMBL Database library (accession number AJ000678).

## REFERENCES

- Albert, F. G., Bronson, E. C., Fitzgerald, D. J. and Anderson, J. N. (1995). Circular structures in retroviral and cellular genomes. *J. Biol. Chem.* **270**, 23570-23581.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhnag, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389-3402.
- Austin, M. J., Collins, J. M., Corey, L. A., Nance, W. E., Neale, M. C., Schieken, R. M. and Brown, J. A. (1992). Aphidicolin-inducible common fragile site expression: Results from a population survey of twins. *Am. J. Hum. Genet.* **50**, 76-83.
- Bailey, A. D., Li, Z., Pavelitz, T. and Weiner, A. M. (1995). Adenovirus type 12-induced fragility of the human RNU2 locus requires U2 small nuclear RNA transcriptional regulatory elements. *Mol. Cell. Biol.* **15**, 6246-6255.
- Boldog, F. L., Waggoner, B., Glover, T. W., Chumakov, I., Le Palsier, D., Cohen, D., Gemmill, R. M. and Drabkin, H. A. (1994). Integrated YAC contig containing the 3p14 hereditary renal cell carcinoma 3;8 translocation breakpoint and the fragile site FRA3B. *Genes, Chromosomes and Cancer* **11**, 216-221.
- Boldog, F., Gemmill, R. M., West, J., Robinson, M., Robinson, L., Li, E., Roche, J., Todd, S., Waggoner, B., Lundstrom, R., Jacobson, J., Mullokandov, M. R., Klinger, H. and Drabkin, H. A. (1997). Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. *Hum. Mol. Genet.* **6**, 193-203.
- Boulikas, T. (1996). Common structural features of replication origins in all life forms. *J. Cell. Biochem.* **60**, 297-316.
- Caporossi, D., Vernole, P., Nicoletti, B. and Tedeschi, B. (1995). Characteristic chromosomal fragility of human embryonic cells exposed *in vitro* to aphidicolin. *Hum. Genet.* **96**, 269-274.
- 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.
- Djalali, M., Adolph, S., Steinbach, P., Winking, H. and Hameister, H. (1987). A comparative mapping study of fragile sites in the human and murine genomes. *Hum. Genet.* **77**, 157-162.
- Dobbs, D. L., Shaiu, W.-L. and Benbow, R. M. (1994). Modular sequence elements associated with origin regions in eukaryotic chromosomal DNA. *Nucl. Acids Res.* **22**, 2479-2489.
- Durnam, D. M., Menninger, J. C., Chandler, S. H., Smith, P. P. and McDougall, J. K. (1988). A fragile site in the human U2 small nuclear RNA gene cluster is revealed by Adenovirus type 12 infection. *Mol. Cell. Biol.* **8**, 1863-1867.
- Eckdahl, T. T. and Anderson, J. N. (1987). Computer modelling of DNA structures involved in chromosome maintenance. *Nucl. Acids Res.* **15**, 8531-8545.
- Eckdahl, T. T. and Anderson, J. N. (1990). Conserved DNA structures in origins of replication. *Nucl. Acids Res.* **18**, 1609-1612.
- Farr, C. J., Bayne, R. A. L., Kipling, D., Mills, W., Critcher, R. and Cooke, H. J. (1995). Generation of a human X-derived minichromosome using telomere-associated chromosome fragmentation. *EMBO J.* **14**, 5444-5454.
- Fitzgerald, D. J., Dryden, G. L., Bronson, E. C., Williams, J. S. and Anderson, J. N. (1994). Conserved patterns of bending in satellite and nucleosome positioning DNA. *J. Biol. Chem.* **269**, 21303-21314.
- Gaddini, L., Belliccia, F., Limongi, M. Z. and Rocchi, A. (1995). Study of the relationship between common fragile sites, chromosome breakages and sister chromatid exchanges. *Mutagenesis* **10**, 257-260.
- Gargano, S., Wang, P., Rusanganwa, E. and Bacchetti, S. (1995). The transcriptionally competent U2 gene is necessary and sufficient for Adenovirus type 12 induction of the fragile site at 17q21-22. *Mol. Cell. Biol.* **15**, 6256-6261.
- Glover, T. M., Berger, C., Coyle, J. and Echo, B. (1984). DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum. Genet.* **67**, 136-142.
- Glover, T. W. and Stein, C. K. (1987). Induction of sister chromatid exchanges at common fragile sites. *Am. J. Hum. Genet.* **41**, 882-890.
- Glover, T. W. and Stein, C. K. (1988). Chromosome breakage and recombination at fragile sites. *Am. J. Hum. Genet.* **43**, 265-273.
- Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M. and Flavell, R. A. (1982). The construction of cosmid libraries which can be used to transform eukaryotic cells. *Nucl. Acids Res.* **10**, 6715-6732.
- Hirst, M. C., Barnicoat, A., Flynn, G., Wang, Q., Daker, M., Buckle, V. J., Davies, K. E. and Bobrow, M. (1993). The identification of a third fragile site, FRAXF in Xq27-q28 distal to both FRAXA and FRAXE. *Hum. Mol. Genet.* **2**, 197-200.
- Hsu, T. C. and Zenzen, M. T. (1964). Mammalian chromosomes *in vitro* XVII: Ideogram of the Chinese hamster. *J. Nat. Cancer Inst. (Bethesda)* **32**, 857-869.
- Hunt, P. A. and Burgoyne, P. S. (1987). A tissue specific fragile site associated with the sex reversed (SXR) mutation in the mouse. *Chromosoma* **96**, 67-71.
- Inoue, H., Ishii, H., Alder, H., Snyder, E., Druck, T., Huebner, K. and Croce, C. M. (1997). Sequence of the FRA3B common fragile region: Implications for the mechanism of *FHIT* deletion. *Proc. Nat. Acad. Sci. USA* **94**, 14584-14589.
- Jones, C., Slijepcevic, P., Marsh, S., Baker, E., Langdon, W. Y., Richards, R. I. and Tunnacliffe, A. (1994). Physical linkage of the fragile site FRA11B and a Jacobsen syndrome chromosome deletion breakpoint in 11q23.3. *Hum. Mol. Genet.* **3**, 2123-2130.
- Jones, C., Penny, L., Mattina, T., Yu, S., Baker, E., Voullaire, L., Langdon, W. Y., Sutherland, G. R., Richards, R. I. and Tunnacliffe, A. (1995). Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene *CBL2*. *Nature* **376**, 145-149.
- Kremer, E. J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S. T., Schlessinger, D., Sutherland, G. R. and Richards, R. I. (1991). Mapping of DNA instability at the Fragile X to a trinucleotide repeat sequence p(CCG)<sub>n</sub>. *Science* **252**, 1711-1714.
- Kuo, M. T., Vyas, R. C., Jiang, L.-X. and Hittelman, W. N. (1994). Chromosome breakage at a major fragile site associated with P-glycoprotein gene amplification in multidrug resistant CHO cells. *Mol. Cell. Biol.* **14**, 5202-5211.
- Laemmli, U. K., Kas, E., Poljak, L. and Adachi, Y. (1992). Scaffold associated regions: Cis-acting determinants of chromatin structural loops and functional domains. *Curr. Opin. Genet. Dev.* **2**, 275-285.
- LeBeau, M. M. and Rowley, J. D. (1984). Heritable fragile sites in cancer. *Nature* **308**, 607-608.
- Lin, M. S., Takabayashi, T., Wilson, M. G. and Marchese, C. A. (1984). An *in vitro* and *in vivo* study of a BrdU sensitive fragile site in the Chinese hamster. *Cytogenet. Cell Genet.* **38**, 211-215.
- McDougall, J. K. (1979). The interactions of Adenovirus with host cell gene loci. *Cytogenet. Cell Genet.* **25**, 183.
- Miro, R., Clemente, I. C., Fuster, C. and Egozcue, J. (1987). Fragile sites, chromosome evolution and human neoplasia. *Hum. Genet.* **75**, 345-349.

- Murano, I., Kuwano, A. and Kajii, T. (1989). Fibroblast specific common fragile sites induced by aphidicolin. *Hum. Genet.* **83**, 45-48.
- Nancarrow, J. K., Kremer, E., Holman, K., Eyre, H., Doggett, N. A., Le Paslier, D., Callen, D. F., Sutherland, G. R. and Richards, R. I. (1994). Implications of FRA16A structure for the mechanism of chromosomal fragile site genesis. *Science* **264**, 1938-1941.
- Natale, D. A., Schubert, A. E. and Kowalski, D. (1992). DNA helical stability accounts for mutational defects in a yeast replication origin. *Proc. Nat. Acad. Sci. USA* **89**, 2654-2658.
- Oberle, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M. F. and Mandel, J. L. (1991). Instability of a 550-base pair DNA segment and abnormal methylation in Fragile X syndrome. *Science* **252**, 1097-1102.
- Ohta, M., Inoue, H., Cotticelli, M. G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., Croce, C. M. and Huebner, K. (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.
- Ong, S. T., Fong, K. M., Bader, S. A., Minna, J. D., Le Beau, M. M., McKeithan, T. W. and Rassool, F. V. (1997). Precise localization of the *FHIT* gene to the common fragile site at 3p14.2 (FRA3B) and characterization of homozygous deletions within FRA3B that affect *FHIT* transcription in tumor cell lines. *Genes Chromosomes Cancer* **20**, 16-23.
- Paradee, W., Mullins, C., He, Z., Glover, T., Wilke, C., Opalka, B., Schutte, J. and Smith, D. I. (1995). Precise localization of aphidicolin induced breakpoints on the short arm of human chromosome 3. *Genomics* **27**, 358-361.
- Paradee, W., Wilke, C. M., Wang, L., Shridhar, R., Mullins, C. M., Hoge, A., Glover, T. W. and Smith, D. I. (1996). A 350 kb cosmid contig in 3p14.2 that crosses the t(3;8) hereditary renal cell carcinoma translocation breakpoint and 17 aphidicolin induced FRA3B breakpoints. *Genomics* **35**, 87-93.
- Pearson, W. R. and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Nat. Acad. Sci. USA* **85**, 2444-2448.
- Pipiras, E., Coquelle, A., Bieth, A. and Debatisse, M. (1998). Interstitial deletions and intrachromosomal amplification initiated from a double-strand break targeted to a mammalian chromosome. *EMBO J.* **17**, 325-333.
- Raap, A. K., van de Corput, M. P. C., Vervenne, R. A. W., van Gijlsuijk, R. P. M., Tanke, H. J. and Wiegant, J. (1995). Ultra-sensitive FISH using peroxidase-mediated deposition of biotin- or fluorochrome tyramides. *Hum. Mol. Genet.* **4**, 529-534.
- Rainaldi, G., Sessa, M. R. and Mariani, T. (1984). Inhibitors of DNA synthesis induce sister chromatid exchanges at the early S phase of the cell cycle. *Chromosoma* **90**, 46-49.
- Rassool, F. V., McKeithan, T. W., Neilly, M. E., van Melle, E., Espinosa, R. and LeBeau, M. M. (1991). Preferential integration of marker DNA into the chromosomal fragile site at 3p14: An approach to cloning fragile sites. *Proc. Nat. Acad. Sci. USA* **88**, 6657-6661.
- Rassool, F. V., LeBeau, M. M., Shen, M.-L., Neilly, M. E., Espinosa III, R., Ong, S. T., Boldog, F., Drabkin, H., MacCarroll, R. and McKeithan, T. W. (1996). Direct cloning of DNA sequences from the common fragile site region at chromosome band 3p14.2. *Genomics* **35**, 109-117.
- Reid, L. H. (1990). Regulatory elements in the introns of the human *HPRT* gene are necessary for its expression in embryonic stem cells. *Proc. Nat. Acad. Sci. USA* **87**, 4299-4303.
- Robinson, T. J. and Elder, F. F. B. (1987). Multiple common fragile sites are expressed in the genome of the laboratory rat. *Chromosoma* **96**, 45-49.
- Saitoh, Y. and Laemmli, U. K. (1994). Metaphase chromosome structure: Bands arise from a differential folding path of the highly AT-rich scaffold. *Cell* **76**, 609-622.
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbour Laboratory Press, NY.
- Smeets, D. and Arets, A. (1990). Genetic determination of fragile site expression. *Am. J. Hum. Genet.* **47**, 196-201.
- Smeets, D. F. C. M. and van de Klundert, F. A. J. M. (1990). Common fragile sites in man and three closely related primate species. *Cytogenet. Cell Genet.* **53**, 8-14.
- Stone, D. M., Jacky, P. B., Hancock, D. D. and Prieur, D. J. (1991). Chromosomal fragile site expression in dogs: 1. Breed specific differences. *Am. J. Med. Genet.* **40**, 214-222.
- Sutherland, G. R. (1991). Chromosomal fragile sites. *GATA* **8**, 161-166.
- Sutherland, G. R. and Baker, E. (1992). Characterisation of a new rare fragile site easily confused with the Fragile X. *Hum. Mol. Genet.* **1**, 111-113.
- Van Dyke, M. W., Hertzberg, R. P. and Dervan, P. B. (1982). Map of distamycin, netropsin and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA•Fe(II). *Proc. Nat. Acad. Sci. USA* **79**, 5470-5474.
- Verkerk, A. J. M. H., Pieretti, M., Sutcliffe, J. S., Fu, Y.-H., Kuhl, D. P. A., Pizzuti, A., Reiner, O., Richards, S., Victoria, M. F., Zhang, F. et al (1991). Identification of a gene (*FMRI*) containing a CCG repeat coincident with a breakpoint cluster region exhibiting length variation in Fragile X syndrome. *Cell* **65**, 905-914.
- Wang, L., Paradee, W., Mullins, C., Shridhar, R., Rosati, R., Wilke, C. M., Glover, T. W. and Smith, D. I. (1997). Aphidicolin-induced FRA3B breakpoints cluster in two distinct regions. *Genomics* **41**, 485-458.
- Wang, L., Darling, J., Zhang, J.-S., Qian, C.-P., Hartmann, L., Conover, C., Jenkins, R. and Smith, D. I. (1998). Frequent homozygous deletions in the FRA3B region in tumor cell lines still leave the *FHIT* exons intact. *Oncogene* **16**, 635-642.
- Wang, Q., Albert, F. G., Fitzgerald, D. J., Calvo, J. M. and Anderson, J. N. (1994). Sequence determinants of DNA bending in the *ilvH* promoter and regulator region of *E. coli*. *Nucl. Acids Res.* **22**, 5753-5760.
- Wilke, C. M., Guo, S.-W., Hall, B. K., Boldog, F., Gemmill, R. M., Chandrasekharappa, S. C., Barcroft, C. L., Drabkin, H. A. and Glover, T. W. (1994). Multicolour FISH mapping of YAC clones in 3p14 and identification of a YAC spanning both FRA3B and the t(3;8) associated with hereditary renal cell carcinoma. *Genomics* **22**, 319-326.
- 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.
- Wulf, H. C. and Niebuhr, E. (1985). Different sister chromatid exchange rates in XX and XY cells of a pair of human chimeric twins. *Cytogenet. Cell Genet.* **39**, 105-108.
- Yu, S., Pritchard, M., Kremer, E., Lynch, M., Nancarrow, J., Baker, E., Holman, K., Muley, J. C., Warren, S. T., Schlessinger, D., Sutherland, G. R. and Richards, R. I. (1991). Fragile X genotype characterised by an unstable region of DNA. *Science* **252**, 1179-1181.
- Yu, S., Mangelsdorf, M., Hewett, D., Hobson, L., Baker, E., Eyre, H. J., Lapsys, N., le Paslier, D., Doggett, N. A., Sutherland, G. R. and Richards, R. I. (1997). Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. *Cell* **88**, 367-374.
- Yunis, J. J. and Soreng, A. L. (1984). Constitutive fragile sites and cancer. *Science* **226**, 1199-1204.
- Zimonjic, D. B., Druck, T., Kastury, K., Croce, C. M., Popescu, N. C. and Huebner, K. (1997). Positions of chromosome 3p14.2 fragile sites (FRA3B) within the *FHIT* gene. *Cancer Res.* **57**, 1166-1170.