Creation of a deletion series of mouse YACs covering a 500 kb region around *Xist*

Edith Heard, Philip Avner and Rodney Rothstein^{1,*}

Unité de Génétique Moléculaire Murine, CNRS URA1445, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cédex 15, France and ¹Department of Genetics and Development, Columbia University College of Physicians and Surgeons, 701 W. 168th Street, New York, NY 10032-2704, USA

Received February 7, 1994; Revised and Accepted April 14, 1994

ABSTRACT

Two mouse YACs, PA-2 and PA-3, contain the *Xist* gene and are 460 kb and 3.3 Mb long respectively. While PA-2 is non-chimeric, PA-3 contains a substantial proportion of non-contiguous DNA. As a prerequisite to functional studies of the role of this region in X inactivation, we have created a deletion series of YACs that are spaced at approximately 50 kb intervals and were able to eliminate the unwanted chimeric sequences in YAC PA-3. For this purpose, we have constructed mouse B1 fragmentation vectors based on those described for human *Alu* fragmentation. Having created this series of YAC deletion derivatives, we were able to eliminate efficiently the 10–15% aberrant YACs that arise during the course of a fragmentation experiment by assessing their marker content. The overlap and the opposite orientation of the two YAC inserts permitted the creation of deletions on both sides of the 500 kb region around *Xist*. The use of this series of YACs in a biological assay will help us define the extent of the sequences necessary to bring about X chromosome inactivation.

INTRODUCTION

Yeast artificial chromosomes have provided a means for the cloning and characterization of genomic DNA spanning hundreds and even thousands of kilobase pairs. The assembly of large contigs of overlapping YACs spanning several megabases has greatly facilitated the mapping of complex genomes. However, the structural analysis of individual YACs of interest can pose a problem owing to their large size and potential chimerism. Although partial digestion with rare cutting enzymes, or fingerprint analysis are useful for the mapping of relatively short YACs of a few hundred kilobases in length, they are unsuitable for larger YACs. YAC fragmentation provides an alternative approach for the rapid physical mapping of yeast chromosomes or large YACs (1). In order to fragment YACs containing mouse DNA inserts, we have adapted the strategy developed by Pavan *et al.* (2) for the fragmentation of YACs containing human DNA inserts. In this report, we demonstrate that this technique also

enables the generation of shorter cloned fragments from a given YAC for functional studies on a particular region within the YAC or, in certain cases of YACs with chimeric inserts, for the elimination of the undesired sequences.

The mouse sequence we have chosen to use for the targeting of homologous recombination events throughout the YAC insert is the B1 repeat element (3). The distribution of B1 repeats in the mouse genome, if it were random, is estimated to be about once every 20 kb and the divergence of individual sequences from the B1 consensus is thought to be minimal (3).

We have applied this approach to two YACs that both contain DNA derived from the candidate region for the X-inactivation center (Xic) as defined by mapping studies of the human (4) and mouse X chromosomes (5). This region encompasses the *Xist* gene on the mouse X chromosome (6,7,8) and is currently thought to be of great interest with respect to Xic for two main reasons. Firstly, the unusual expression properties of the *Xist* gene suggest that it represents a potential candidate for the Xic (9, 10,11). Secondly, recent data from our laboratory (12) suggest that the genetically defined X chromosome controlling element (Xce), (13), which affects the random nature of X inactivation in the mouse, is tightly linked to but distinct from the *Xist* gene and thus may also lie in this 460 kb region. The nature of the X inactivation phenomenon, in other words, the transcriptional silencing of an entire chromosome, suggests perhaps that the region necessary to initiate this process, the Xic, may be quite extensive. For these reasons, we are carrying out a detailed analysis of YACs covering this region and in particular we are establishing functional assays for X-inactivation that involve the transfer of YACs into mouse cells. The availability of a series of shortened YACs covering the region around *Xist* to different extents will be useful for delimiting the sequences necessary for X-inactivation.

We chose to adapt the YAC fragmentation strategy of Pavan *et al.* (2) for work in the mouse as it fulfills certain requirements that could not necessarily be fulfilled by alternative fragmentation approaches. One of these requirements is the possibility of using PCR-based mapping of STS's (sequence tagged sites) for the positioning of the fragmentation sites along the YAC. Three out of the ten markers we used in this 500 kb region are STS's.

*To whom correspondence should be addressed

Furthermore, part of our strategy for the establishment of functional tests to assess the extent of the region necessary to produce X inactivation depends upon the introduction of different YAC fragments from the candidate region into mouse embryonic stem cells (14, for review see ref. 15). Other approaches for fragmentation, such as the random integration of I-Sce-I sites described by Colleaux *et al.* (16), which results in the *in vitro* generation of two fragments, do not readily permit these types of analyses. For example, the presence of both products of fragmentation in the I-Sce-I approach, prevents the use of mapping by PCR.

One of the two YACs we are studying is a 460 kb YAC, PA-2, which has been extensively mapped (8) and has been subcloned into a series of λ clones that have been organized into a contig (17). As a consequence of this contig, a series of probes and STS's are available across the region covered by YAC PA-2. These have enabled us to assess the integrity of the YACs produced by fragmentation at a level that is not always possible owing to the relative paucity of probes in some regions of the mouse genome.

The second YAC we have studied is a 3.3 Mb YAC, PA-3, whose insert is chimeric. Using the fragmentation approach, we have generated a series of shorter YACs, some of which have inserts derived exclusively from the X chromosome. In this way we have simultaneously defined the extent of X chromosome derived sequences in this YAC and rendered them useful for other purposes such as functional studies of the region.

MATERIALS AND METHODS

Yeast strains, crosses and yeast cell transformations

YACs PA-2 and PA-3 were isolated, as described previously (Heard *et al.*), from the library constructed by Larin *et al.* (18) in *Saccharomyces cerevisiae* strain AB1380 (*MATa ade2-1 can1-100 lys2-1 trp1 ura3 his5-u*) using pYAC4. Using standard yeast genetics (19) both YACs were transferred to a *his3* genetic background after crossing to YPH857 (20), a *his3* mutant strain (*MAT α leu2- Δ 1 trp1- Δ 63 ura3-52 ade2-101 his3- Δ 200 lys2-801 cyh^R*). The two strains used for fragmentation were: W1178-3A, a *his3* derivative containing YAC PA-2 and W1175-11C, a *his3* derivative containing YAC PA-3.

Non-fragmented YAC-containing strains were maintained in AHC selective medium lacking uracil and tryptophan (21). Yeast transformations were carried out based on the protocol of Burgers and Percival (22) using 1 μ g of linearized plasmid DNA. Transformants were selected on synthetic complete medium lacking histidine (SD COM -His). Colonies that survived this selection were replica-plated on to medium lacking tryptophan (SD COM -Trp) or uracil (SD COM -Ura) to identify those colonies that had lost the ability to grow in the absence of uracil but still grow in the absence of tryptophan. His⁺Trp⁺Ura⁻ colonies were maintained on SD COM -Trp-His medium.

Fragmentation vectors

The mouse YAC fragmentation vectors were created by first determining a B1 consensus sequence (refs. 23, 24, and Steve Best, personal communication). Four oligonucleotides were synthesized as shown in Figure 1B. The oligonucleotides were designed to overlap by 12 base pairs in the central portion of the sequence and by the central four base pairs of the *Bam*HI site that was engineered into the sequence (see legend to Fig.

1B). An *Apa*I and *Xma*I site were engineered into the 5' end of the sequence to facilitate the cloning and subsequent verification of the orientation of the B1 sequence in the vector, respectively. An *Xho*I site was engineered in to the 3' end to facilitate cloning and for subsequent use to expose the B1 sequence for use in fragmentation (see below). Five micrograms of each oligonucleotide were phosphorylated with polynucleotide kinase (25) in a 80 μ l reaction. Excess ATP was removed by salt precipitation. Next, 2.5 μ g of each oligonucleotide was mixed in a ligation reaction and incubated at 18 °C for 18 hr. Ethidium bromide staining of an agarose gel after electrophoresis confirmed successful ligation. Ligase was inactivated at 65 °C for 10 min and the mixture was digested with *Bam*HI to liberate the full length B1 fragment. Next the B1 fragment was digested separately with *Apa*I and *Xho*I to generate directionally clonable B1 fragments. Approximately 4 pmol of the B1 fragment was ligated with 0.5 pmol of pBP103 that had been digested with *Bam*HI and *Apa*I and separately with *Bam*HI and *Xho*I in 20 μ l. After 18 hr at 16 °C, the ligation mixtures were electroporated into electrocompetent *E. coli* strain TG-1 (25). Since the B1 was inserted into the polylinker of pBP103 (26), which is an in-frame fusion with β -galactosidase, successful insertions were detected as white colonies on X-gal medium.

Two clones, pWJ521 and pWJ522 were chosen as representatives of each B1 orientation. DNA sequence analysis confirmed that pWJ522 contained a complete B1 insert. On the other hand, pWJ521 had undergone extensive rearrangement during cloning. However the B1 repeat was of the correct length. A replacement, pWJ528, for pWJ521 was created by digesting pWJ522 with *Pvu*II and inverting the B1 insert (Figure 1).

Electrophoresis and hybridization

Yeast DNA was prepared in low melting point (LMP) agarose plugs using the procedure described by Carle and Olson (27). Briefly, approximately 5×10^7 cells from a 5 ml overnight culture were rinsed in 50 mM EDTA, resuspended in 150 μ l of SEM [1M sorbitol (Sigma), 20 mM EDTA, 14 mM 2-mercaptoethanol] containing 1 mg/ml Zymolyase (ICN Immunobiochemicals). The suspension was mixed with 150 μ l of molten agarose (1.5 % LMP agarose in 1M sorbitol, 20 mM EDTA) and dispensed into slots of pre-chilled 60 μ l plug moulds and allowed to sit on ice. Plugs were incubated in 3 ml S.E.M with 1 mg/ml Zymolyase and 10mM Tris-Cl at 37°C for 2 to 4 hours, and then transferred to 3 ml of 0.5 M EDTA, 1% sarkosyl containing 2 mg/ml Proteinase K (Boehringer Mannheim) and incubated at 50°C for 24 h. Plugs were stored in 0.5 M EDTA and were rinsed thoroughly in TE (10 mM Tris-Cl, 1 mM EDTA) prior to gel-loading.

CHEF gel electrophoresis (Pulsaphor LKB) was performed using 1% agarose gels in 0.5 $\times$ Tris-borate EDTA buffer at 11°C as described by Chu *et al.* (28). For resolution of DNA fragments up to 600 kb in size, running conditions were 140 V, 60 s pulse interval, for 40 h; for fragments up to about 1 Mb, the conditions were 140 V, 150 s pulse interval for 48 h; for fragments up to 5.4 Mb, the conditions were 65 V, 30 minute pulse interval for 7 days. DNA was transferred overnight onto Hybond N+ membrane (Amersham). Hybridizations were carried out using 0.5 M sodium phosphate pH 7.2, 7% SDS, with oligo-labelled probes at 65°C for 16h. Probes generated from PCR products were pre-annealed with total mouse genomic DNA to suppress repetitive sequences.

Probes used for hybridization included the entire pBR322 plasmid for detection of YAC vector sequences; the *Xist* cDNA sequence (6); the *DXPas19* marker (5); and the *DXPas27* marker (8). In addition a series of DNA fragments were used that were generated by PCR of individual lambda DNAs derived from a lambda contig established across YAC PA-2. Using combinations of B1 and B2 repeats together, or either one with T3 or T7 primers, probes were generated from lambdas *IVA5*, *IB6*, *IG10*, *IIIA8*, *IVA9* (17).

Microsatellite marker colony PCR

The primers corresponding to the three microsatellite markers *DXPas28*, *DXPas29* and *DXPas31* were used to carry out PCR on individual yeast colonies. A colony from a freshly streaked agar plate was picked into 30 μ l of H₂O and 10 μ l of this was used in a 50 μ l PCR. The conditions used and the sequence of each primer pair were as described by Simmler *et al.* (12). 25 μ l of the PCR was run out on a 3% agarose (NuSieve) gel and the presence of product assessed by ethidium bromide staining.

RESULTS AND DISCUSSION

Construction of a mouse YAC fragmentation vector

In order to create a series of shortened derivatives from a YAC containing a mouse insert, fragmentation vectors were created that would specifically target mouse genomic DNA using a mouse repetitive element. Interspersed repetitive DNA sequences are divided into two classes, SINES and LINES. The B1 (ca 130 bp in length) and B2 (ca 190 bp in length) elements belong to the SINE class of repeats in the mouse (Krayev *et al.* 1983). From 130,000 to 180,000 copies of the B1 repeat are present in the mouse genome (30) and they share some homology with human *Alu* repeats (for review see ref. 31). We chose the B1 rather than the B2 element or a LINE as the targeting sequence for mouse YAC fragmentation because the frequency of B1 repeats is higher than that of other repeats, and the degree of sequence divergence between B1 elements is thought to be minimal (3).

The fragmentation plasmids were constructed by insertion of a B1 element into the polylinker of pBP103. This is a backbone plasmid for YAC fragmentation that carries a telomere sequence, the yeast *HIS3* gene and bacterial plasmid sequences (26). The consensus B1 element was generated using a concatemer of synthetic oligonucleotides as described in Materials and Methods.

Vectors were made carrying a B1 element in either orientation to ensure that all the B1 repeats in the YAC insert could be used as targets for fragmentation (see Figure 1C). pWJ521 and pWJ528 carry one orientation of B1 and pWJ522 carries the other orientation (Figure 1). Digestion with *SalI* and *BamHI* in the case of pWJ521 and pWJ528, and with *SalI* and *XhoI* in the case of pWJ522, linearizes the vectors leaving the 130 bp B1 sequence free at one end for homologous recombination with the B1 elements in the YAC insert, and the free yeast telomere sequence at the other end. Initially pWJ521 and pWJ522 were used for fragmentation experiments. Partial sequencing of these vectors was carried out in order to verify the integrity of their B1 repeats and some degree of rearrangement was found in pWJ521. However, this vector actually generated the same frequency of useful transformants as pWJ522, which is consistent with the fact that the rearrangement in pWJ521 does not affect the length of homologous sequences for B1 targeting (unpublished

observations). Nevertheless a second vector, pWJ528, was constructed with B1 sequences in the same orientation as pWJ521, verified by sequencing and used for later fragmentation experiments. No differences in fragmentation efficiency were detected between pWJ521 and pWJ528.

These fragmentation vectors provide the possibility of rescuing the sequences adjacent to a given YAC fragmentation site after digestion with *EcoRV* or *ClaI* and recircularization followed by transformation into bacteria, in a similar fashion to the human YAC fragmentation vectors described by (2). The appropriate end-rescue plasmids can then be selected for by transformation into BA1, a *hisB trpC* bacterial strain (32) where the yeast *HIS3* gene, which should be present on correctly recircularized YAC ends, can complement the bacterial *hisB* mutation.

Analysis of mouse X chromosome YACs PA-2 and PA-3

Two of the YACs isolated from a C3H mouse YAC library using the X chromosomal *Xist* cDNA as probe (8) were used in this study. YAC PA-3 was found to be approximately 3.3 Mb long by PFGE based on comparison with *Schizosaccharomyces pombe* chromosome markers (see Figure 2). It is interesting to note that, given its length, YAC PA-3 is quite stable mitotically: 5/215 colonies lost the YAC after growth in non-selective medium. The isolation of the two extremities of its insert indicated that this YAC was chimeric, the left hand end (*TRP CEN*) being derived from the X chromosome and the right hand end (*URA*) being autosomal (8). Hybridization of the probes available at the time, namely *Xist* and *DXPas19* (Figure 3, panel E), to YAC PA-3 showed that the YAC contained the *Xist* gene but not the *DXPas19* marker. This suggested that either the left hand extremity of YAC PA-3 was situated between *Xist* and *DXPas19* or that the chimeric breakpoint lies between these two markers. Preliminary investigations with a mouse DNA probe derived from the left hand extremity of YAC PA-3 indicated the former to be the case (8). However, the generation of fragmented derivatives of YAC PA-3 described here revealed that the chimeric joint in the insert must lie between *Xist* and *DXPas19*. The orientation of YAC PA-3 is shown in Figure 3.

In order to characterize and ultimately to exploit the sequences of YAC PA-3 that were derived from the X chromosome, fragmentation of this YAC was undertaken. This would provide a series of shorter clones from YAC PA-3, some of which would be derived solely from the X chromosome and which would be suitable for functional studies by reintroduction into mouse cells.

The other YAC we have studied, PA-2, is 460 kb long and has been extensively mapped (8). Its insert is derived exclusively from the X chromosome based on analysis of its extremities and on comparative mapping with other YACs and mouse genomic DNA (8). We undertook the fragmentation of YAC PA-2 in parallel to YAC PA-3, in part as a useful control for the mouse YAC fragmentation system and also in order to generate a series of YACs for functional studies of the region surrounding the *Xist* gene. YAC PA-2 has previously been subcloned into a series of lambda clones that have been ordered into a contig (17) and has been mapped by means of B2 repeat targeted integration of *I-Sce I* sites (16). A series of mouse probes derived by PCR or from subcloned fragments of the lambda clones, as well as three microsatellite markers are available in the 460 kb region covered by YAC PA-2, as shown in Figure 3. The average distance between these probes and markers is approximately 50 kb and they have thus enabled a relatively fine analysis of the fragmentation products of YACs PA-2 and PA-3.

Fragmentation of YACs PA-2 and PA-3

Both YACs PA-2 and PA-3 were originally provided in the AB1380 strain (*his5 HIS3*). The YACs were transferred into a *his3-Δ200* background by crossing, sporulation and tetrad analysis, so that the YAC strain would now have the *HIS3*

his3-Δ200 genotype. The presence of the *his3* deletion mutation avoids the generation of spontaneous His⁺ revertants, and is necessary for use with the *HIS3* fragmentation vectors. The transfer of the YAC onto a *his3* deletion mutant background means that the frequency of His⁺ transformants due to fragmentation events can be as high as 95%. Other auxotrophic

A.

B.

C.

Figure 1. Mouse YAC fragmentation vectors. A). The pBP103 backbone plasmid of Pavan *et al.* (26) is shown with the two possible orientations of the B1 sequence inserted into the polylinker, generating pWJ522 for one orientation and pWJ528 for the other orientation (see text for details concerning pWJ521). The underlined restriction sites are used to linearize the vector: *SalI* and *XhoI* for pWJ522; *SalI* and *BamHI* for pWJ528. The *EcoRV* and *ClaI* sites indicated can be used to isolate the sequences adjacent to the fragmentation site by plasmid rescue in *E. coli*. B) The sequences of the four oligomers (A, B, C and D) used to construct the B1 repeat element are shown: the 4bp (GATC, *BamHI* site) overlap between the extremities of oligos A and C should be noted, as should the 12 bp overlap between oligo D and oligo B, which forms the central portion of the B1 sequence. The orientation of this B1 element is as shown for the pWJ528 vector above. C) Fragmentation mechanism: the linearized fragmentation vector undergoes homologous recombination at one of the B1 repeats in the parental YAC insert, introducing a telomere and deleting all the distal sequences. A vector carrying each orientation of the B1 element is shown pairing with the target YAC. In practice each vector is transformed separately. Deletion derivatives are selected based on acquisition of *HIS3* and screened for loss of *URA3*.

markers such as *lys2* and *his5* for which AB1380 contains point mutations rather than deletions tend to be much less efficient.

In order to verify the integrity of each YAC in its new genetic background prior to fragmentation and also in order to assess the stability of the two YACs after meiosis, PFGE analysis of the segregation products of different tetrads was carried out. For both YACs PA-2 and PA-3, we found evidence for unequal sister chromatid exchange leading to the simultaneous generation of larger and smaller YACs in the same tetrad (data not shown). Nevertheless, for each YAC, segregants were isolated that

contained the parental-sized YAC with the appropriate genotype (*HIS5 his3*).

Fragmentation experiments were carried out using spheroplast transformation with plasmid DNA that had been linearized between the B1 repeat and the telomere sequence (see Figure 1). The yeast strains containing the original YACs have the phenotype Ura⁺, Trp⁺, His⁻. The transformants containing the desired deletion derivatives should have the phenotype Ura⁻, Trp⁺, His⁺, although some Ura⁺, Trp⁻, His⁺ cells will also be generated. However, the fragments in the latter case should be mitotically unstable owing to the absence of a centromere (i.e. loss of the *TRP1* YAC vector arm).

Fragmentation of YACs PA-2 and PA-3 was initially carried out using the pWJ521 and pWJ522 plasmids and subsequently using the pWJ528 plasmid. The average transformation efficiency varied between 50 and 100 colonies per μ g of linearized plasmid for all three vectors. Yeast transformants were initially selected in the absence of histidine and were then scored for the *URA3* and *TRP1* markers by growth on the appropriate omission media. The frequency of Ura⁻Trp⁺His⁺ transformants ranged from 75%–90% for each of the vectors pWJ521, pWJ522 and pWJ528 in different transformation experiments with both YACs PA-2 and PA-3. Only strains of this phenotype were analyzed further.

We believe that the high efficiency of successful transformants found with these vectors, even though the length of target sequence for homologous recombination (130 bp) is quite short, is due partly to the high conservation of B1 repeats and also to the fact that the YAC was transferred into a *his* deletion mutant strain.

Figure 2. Ethidium bromide stained gel showing the chromosomes of eight different yeast clones containing YAC PA-3, resolved by PFGE using the CHEF system with a pulse interval of 15 min. at 70 V for 5 days. The size of YAC PA-3 (3.3 Mb) was estimated by comparison with *S.pombe* chromosome markers.

Figure 3. Hybridization results of twelve representative deletion derivatives of YAC PA-3 resolved using PFGE. Lane 1, F1-k; lane 2, F1-l; lane 3, F1-m; lane 4, F1-n; lane 5, F1-o; lane 6, F1-p; lane 7, F1-r; lane 8, F1-t; lane 9, F1-u; lane 10, F1-v; lane 11, F1-w; lane 12, F1-x. Details of these YACs are shown in the top right hand panel of Figure 4. The same Southern blot was hybridized with (A) a pBR322 vector probe, which detects all YAC fragments; (B) the *IVA5* probe, (C) the *Xist* cDNA probe; (D) the *IB6* probe. (E) The relative orientations of the two YACs, PA-2 and PA-3, and the order of the mouse X chromosome markers in this region are shown. *DXPas31* represents the most proximal X chromosome marker and *DXPas27* the most distal one on this map. The solid circles represent the *TRP1* CEN ends of the YACs. The probes described above in (B), (C) and (D) are underlined. Note that as the distance between the probe used and the *TRP1* CEN arm of YAC PA-3 increases, the size of the fragments detected increases, with the exception of one aberrant fragment, PA-3 F1-x (lane 12).

Analysis of fragmentation products of YACs PA-2 and PA-3

Fragmentation products of YACs PA-2 and PA-3 were analyzed by pulsed field gel electrophoresis and Southern blot hybridization. Some fragments were also analysed by colony PCR. In this way, the sizes of YAC fragments could be estimated and the presence of various markers in the region could be assessed. An example of the hybridization results of four probes:

a pBR322 (YAC vector) probe, *Xist* (a cDNA probe), *IVA5* and *IB6* (described in Materials and Methods), on twelve pWJ521-derived fragments of YAC PA-3 is shown in Figure 3. The complete results for the analyzed fragments derived from both YACs using these fragmentation vectors are shown in Figure 4.

The size distribution of fragments from both YACs appears to be more or less random although there does seem to be a

PA-2	/kb	<i>IVA5</i>	<i>Xist</i>	<i>IB6</i>	<i>IG10</i>	<i>DXPas19</i>	<i>IVA9</i>	<i>DXPas27</i>
F1-f	50	-	-	-	-	-	-	+
F1-j	70	-	-	-	□	□	-	+
F1-p	90	-	-	-	-	-	-	+
F1-b	100	-	-	-	-	-	-	+
F1-k	100	-	-	-	-	-	-	+
F1-a	110	-	-	-	-	-	+	+
F1-n	120	-	-	-	-	-	+	+
F1-i	150	-	-	-	-	-	+	+
F1-h	170	-	-	-	-	-	+	+
F1-m	170	-	-	-	-	-	+	+
F1-c	180	-	-	-	-	-	+	+
F1-e	200	-	-	-	-	-	+	+
F1-l	220	-	-	-	-	+	+	+
F1-o	220	-	-	-	-	+	+	+
F1-d	310	-	-	+	+	+	+	+
F1-g	400	+	+	+	+	+	+	+

PA-2	/kb	<i>DXPas31</i>	<i>IVA5</i>	<i>Xist</i>	<i>DXPas29</i>	<i>IIA8</i>	<i>DXPas28</i>
F8-i	100	-	-	-	-	-	-
F8-e	150	-	-	-	-	-	+
F8-d	170	-	-	-	-	-	+
F8-f	200	-	-	-	-	-	+
F8-n	200	-	-	-	-	-	+
F8-a	250	-	-	-	-	+	+
F8-k	260	-	-	-	-	+	+
F8-m	270	-	-	-	-	+	+
F8-h	280	-	-	-	+	+	+
F8-j	280	-	-	-	+	+	+
F8-l	340	-	-	-	-	+	+
F8-g	360	-	+	+	+	+	+
F8-c	380	-	+	+	+	+	+
F8-o	390	-	+	+	+	+	+
F8-b	>460	-	-	-	-	+	+

PA-2	/kb	<i>DXPas31</i>	<i>IVA5</i>	<i>Xist</i>	<i>DXPas29</i>	<i>IB6</i>	<i>DXPas19</i>	<i>DXPas28</i>	<i>IVA9</i>	<i>DXPas27</i>
F2-g	110	-	-	-	-	-	-	-	+	+
F2-n	120	-	-	-	-	-	-	-	+	+
F2-f	120	-	-	-	-	-	-	□	+	+
F2-d	160	-	-	-	-	-	-	+	+	+
F2-l	190	-	-	-	-	-	-	+	+	+
F2-i	200	-	-	-	-	-	-	+	+	+
F2-e	250	-	+	+	□	+	-	-	-	-
F2-h	250+ 480	-	-	+	+	-	□	-	-	-
F2-j	280	-	+	+	+	+	-	-	-	-
F2-k	310	-	-	-	+	+	+	+	+	+
F2-c	320	-	-	+	+	+	+	+	+	+
F2-m	330	-	+	-	-	+	+	+	+	+
F2-a	360	-	+	+	+	+	□	+	+	+
F2-b	420	-	+	+	+	+	□	+	+	+

PA-3	/kb	<i>DXPas31</i>	<i>IVA5</i>	<i>Xist</i>	<i>DXPas29</i>
F1-c	40	-	-	-	-
F1-b	70	+	-	-	-
F1-f	110	+	-	-	-
F1-m	120	+	-	-	□
F1-v	120	+	-	-	□
F1-t	130	+	-	-	□
F1-p	150	+	+	-	□
F1-w	150	+	+	-	□
F1-x	200	+	-	-	□
F1-o	205	+	+	-	□
F1-l	210	+	+	+	□
F1-n	320	+	+	+	+
F1-r	320	+	+	+	□
F1-u	320	+	+	+	□
F1-k	350	+	+	+	□
F1-a	360	+	+	+	+
F1-h	400	+	+	+	+
F1-e	450	+	+	+	+
F1-d	>1000	+	+	+	□
F1-g	>1000	-	-	+	+
F1-i	>1000	+	+	+	+
F1-j	>1000	+	+	+	+
F1-s	>1000	+	+	+	□
F1-q	>2000	+	+	+	□

PA-3	/kb	<i>DXPas31</i>	<i>IVA5</i>	<i>Xist</i>	<i>DXPas29</i>
F8-f	50	□	-	-	□
F8-h	60	□	-	-	□
F8-c	190	□	+	+	□
F8-a	200	□	+	+	□
F8-b	200	□	+	+	□
F8-d	260	+	+	+	□
F8-e	280	□	+	+	□
F8-g	>1000	□	+	+	□

PA-3	/kb	<i>DXPas31</i>	<i>IVA5</i>	<i>Xist</i>	<i>IB6</i>	<i>IIA8</i>
F2-e	45	+	-	-	-	-
F2-r	70	□	-	-	-	-
F2-f	120	+	-	-	-	-
F2-g	190	+	+	+	-	-
F2-m	200	+	+	+	-	-
F2-o	210	+	+	+	-	-
F2-q	220	+	+	+	-	-
F2-i	280	+	+	+	+	+
F2-a	>800	+	+	+	□	+
F2-d	>800	+	+	+	□	+
F2-h	900	+	+	+	+	+
F2-j	1000	+	+	+	+	+
F2-b	>1000	+	+	+	□	+
F2-c	>1000	+	+	+	□	+
F2-k	>1000	+	+	+	+	+
F2-l	>1000	+	+	+	+	+
F2-n	>1000	+	+	+	+	+
F2-p	>1000	+	+	+	+	+

Figure 4. Marker content of the deletion derivatives produced from YACs PA-2 (left hand side) and PA-3 (right hand side) using fragmentation vectors pWJ521 (products F1-a, F1-b, etc); pWJ522 (products F2-a, F2-b, etc); and pWJ528 (products F8-a, F8-b, etc). Presence of a marker is shown as +; absence of a marker is shown as -; marker not tested is shown as □.

Figure 5. The twelve YAC fragments derived from YACs PA-3 (above) and PA-2 (below) chosen to cover the approximately 500 kb region around *Xist* at intervals of about 50 kb. These YACs were selected on the basis of their size and correct marker content. X chromosome derived sequences are shown as solid bars; the chimeric portion of YAC PA-3 is represented as a stippled bar. The solid circles represent the *TRP1* *CEN* ends of the YACs.

greater number of fragments derived from YAC PA-2 in the 100 kb–200 kb range (20 out of a total of 39 non-aberrant fragments). There do not, however, appear to be any outstanding B1 hotspots in the region covered by YACs PA-2 and PA-3. There is some suggestion that the size of the fragmented YACs is inversely proportional to the amount of vector DNA used in the transformation. For example, we found in one experiment with YAC PA-2 that increasing the vector concentration 2.5-fold, resulted in an overall shorter average length of fragments (170 kb vs. 250 kb).

During the analysis of the fragmented YACs, two types of aberrant YAC were detected. In the case of YAC PA-2, one out of forty-five fragmented YACs (PA-2 F8-b) was actually larger than the parental YAC. This must represent an unusual rearrangement as only two of the six mouse DNA markers tested were present on this YAC (see Figure 4). The second aberrant class of YAC fragments is comprised of deletions that are sometimes accompanied by complex rearrangements within the retained segments. In addition, in three of the strains carrying aberrant YAC fragments (PA-2 F2-e, F2-h and F2-j), yeast chromosome I increased in size by about 30 kb. No obvious changes were seen in the yeast karyotype for other strains carrying aberrant YAC fragments. Intriguingly, all 6 of the rearranged PA-2 derived fragments involved a breakpoint somewhere between the *DXPas29* and *DXPas19* markers (an approximately 90 kb region).

Only 2 of the 50 YAC PA-3 fragments analysed (4%) appear to be aberrant (PA-3 F1-g and F1-x, the latter appears in Fig. 3). However, this low frequency compared to YAC PA-2 fragments may reflect the fact that the order of markers could only be assessed across 270 kb of the 3.3 Mb YAC insert.

Furthermore the resolution of the PFGE analysis did not allow the detection of any fragmentation products that might be larger than the parental YAC.

Previous analysis has shown that YAC PA-3 contains X chromosomal material proximal to *Xist*, as it does not contain the *DXPas19* marker. The relationship between the size of PA-3 fragments and the order of the markers in this region showed that the *TRP1* *CEN* end of the YAC PA-3 insert must be located proximally along the X chromosome with respect to *Xist* (see Fig. 5) and the chimeric breakpoint must lie somewhere between the *IIIA8* and *DXPas19* markers. The extent of X chromosomal material in YAC PA-3 must therefore be the distance from the *TRP1* *CEN* end to *DXPas19*, a maximum of 300 kb. It is worth noting that the chimeric breakpoint of YAC PA-3 lies in the same 120 kb region as the rearrangements in all six of the aberrant YAC fragments detected which were derived from YAC PA-2 (see above). This is suggestive of a region that is highly recombinogenic.

Although YAC fragmentation represents a rapid and easy method for obtaining a series of deletion derivatives across a genomic region of interest, our study has demonstrated the need for a certain amount of caution in the choice of fragments to be used for any subsequent analysis. 13% (6 out of 45) of the YAC fragments generated from YAC PA-2 were aberrant but only one of these aberrant YACs could have been detected by its size alone. This demonstrates the importance of having a high density of markers within the YAC being fragmented. In the absence of numerous markers, a simple fingerprint analysis using mouse repeat sequences as probes can be conducted. Alternatively, a more thorough approach involving restriction mapping of the YAC fragments for comparison with the original YAC (2) can

be undertaken. Finally, isolation of end probes from a selection of YAC fragments is a useful way of generating new markers in the region that can be used to assess the integrity of the panel of YAC fragments generated.

CONCLUSIONS

In this report we describe the construction and use of mouse YAC fragmentation vectors to generate a series of derivatives from two YACs that lie within the candidate region for the mouse X inactivation center and that contain the *Xist* sequence. We eliminated aberrant fragmentation products after assessment of their marker content and selected the nested series of deletions shown in Figure 5 as a representative sample of the region.

One of the two YACs studied, PA-3, is 3.3 Mb long and contains a chimeric insert. By means of fragmentation, we generated derivatives of this YAC that contain some or all of its X-derived sequences. By mapping these fragments with the available markers in the region, we were able to conclude that YAC PA-3 contains a maximum of 300 kb of X-chromosomal sequence: 170 kb ($\pm$ 20 kb) proximal and 130 kb ($\pm$ 20 kb) distal to *Xist*. The other YAC analyzed, PA-2, which is 460 kb long and has been mapped (8) and subcloned (17), is oriented in the opposite direction to YAC PA-3 along the X-chromosome. Thus the fragmentation of these YACs has provided us with a series of deletions spanning the approximately 500 kb region of interest from both directions (see Figure 5).

Increasingly, the transfer of YACs into mammalian tissue culture cells or the mouse germ line is being used for functional studies of regions containing very large genes or domains, gene clusters, or genes with distant regulatory elements necessary for their correct expression (for reviews see refs. 33 and 34). Furthermore, the ease with which marker genes, such as neomycin, can be retrofitted onto a given YAC by homologous recombination opens up numerous possibilities for studies of gene activity and regulation. Some methods for YAC transfer into mammalian cells such as microinjection (35) have an upper size limit. Thus, the ability to eliminate unnecessary regions of a YAC by fragmentation, whether chimeric or not, is indispensable. Fragmentation thus enables large regions to be broken down sufficiently to enable analysis and manipulation, while maintaining the advantage of a larger cloning capacity than cosmids, for example. In this context, given that the nature and the extent of the X-inactivation center remain unknown, a deletion series of YACs, such as the one we have described here, can be used to define the region necessary to bring about X inactivation in a biological assay.

ACKNOWLEDGEMENTS

We thank Zoia Larin, Tony Monaco and Hans Lehrach for providing YACs PA-2 and PA-3, Steve Best for help in designing the B1 consensus sequence, Roger Reeves for providing pBP103, Phil Hieter for providing YPH857, Ivana Sunjevaric for help in construction of pWJ528, Claire Rougeulle for providing probes and Marie-Christine Simmler for help to R.R. and the use of her microsatellite markers. This work was supported by grants from the Association Française contre les Myopathies (A.F.M) and the Ministère de la Recherche et de l'Espace (M.R.E) (Contrats 89C0734 et 92C0040) to P.A. E.H. was supported by a fellowship from the Wellcome Trust and R.R. was supported by the Ligue Contre le Cancer and NIH grant HG00452.

REFERENCES

- Vollrath, D., Davis, R. W., Connelly, C. and Hieter, P. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6027–6031.
- Pavan, W. J., Hieter, P. and Reeves, R. H. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1300–1304.
- Kramerov, D. A., Grigoryan, A. A. and Georgiev, G. P. (1979) *Nucleic Acids Res.*, **6**, 697–713.
- Brown, C. J., Lafrenière, R. G., Powers, V. E., Sebastio, G., Ballabio, A., Pettigrew, A. L., Ledbetter, D. H., Levy, E., Graig, I. W., Willard, H. F. (1991) *Nature*, **349**, 82–84.
- Cattanach, B. M., Rasberry, C., Evans, E. P., Dandolo, L., Simmler, M., C. and Avner, P. (1991) *Cytogenet. Cell Genet.*, **56**, 137–143.
- Borsani, G., Tonlorenzi, R., Simmler, M. C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., Willard, H. F., Avner, P., and Ballabio, A. (1991) *Nature*, **351**, 325–329.
- Brockdorff, N., Ashworth, A., Kay, G. F., Cooper, P., Smith, S., McCabe, V. M., Norris, D. P., Penny, G. D., Patel, D., and Rastan, S. (1991) *Nature*, **351**, 329–331.
- Heard, E., Simmler, M.-C., Larin, Z., Rougeulle, C., Courtier, B., Lehrach, H., and Avner, P. (1993) *Genomics*, **15**, 559–569.
- Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafrenière, R. G., Xing, Y., Lawrence, J. and Willard, H. F. (1992) *Cell*, **71**, 527–542.
- Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S. and Rastan, S. (1992) *Cell*, **71**, 515–526.
- Kay, G. F., Penny, G. D., Patel, D., Ashworth, A., Brockdorff, N. and Rastan, S. (1993) *Cell*, **72**, 171–182.
- Simmler, M.-C., Cattanach, B. M., Rasberry, C., Rougeulle, C., and Avner, P. (1993) *Mammalian Genome*, **4**, 523–530.
- Cattanach, B. M., Pollard, C. E. and Perez, J. N. (1969) *Genet. Res.*, **14**, 223–235.
- Pachnis, V., Pevny, L., Rothstein, R. and Constantini, F. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5109–5113.
- Huxley, C. and Gnirke, A. (1991) *Bioessays*, **13**, 545–550.
- Colleaux, L., Rougeulle, C., Avner, P. and Dujon, B. (1993) *Hum. Mol. Genet.*, **2**, 265–271.
- Rougeulle, C., Colleaux, L., Dujon, B. and Avner, P. *Mammalian Genome*, submitted.
- Larin, Z., Monaco, A. P. and Lehrach, H. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4123–4127.
- Rose, M. D., Winston, F., and Hieter, P. (1990) 'Methods in Yeast Genetics: A laboratory course manual'. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Reeves, R. H., Pavan, W. J. and Hieter, P. (1992) *Methods Enzymol.*, **216**, 584–603.
- Brownstein, B. H., Silverman, G. A., Little, R. D., Burke, D. T., Korsmeyer, S. J., Schlessinger, D. and Olson, M. V. (1989) *Science*, **244**, 1348–1351.
- Burgers, P. M. J. and Percival, K. J. (1987) *Anal. Biochem.*, **163**, 391.
- Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov, A. P., Bayev, A. A., and Georgiev, G. P. (1980) *Nucleic Acids Res.*, **6**, 1201–1215.
- Quentin, Y. (1989) *J. Mol. Evol.*, **28**, 299–305.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) 'Molecular Cloning: A laboratory manual'. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Pavan, W. J., Hieter, P., Sears, D., Burkhoff, A. and Reeves, R. H. (1991) *Gene*, **106**, 125–127.
- Carle, C. F., and Olson, M. V. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3756–3760.
- Chu, G., Vollrath, D., and Davis, R. W. (1986) *Science*, **234**, 1582–1585.
- Krayev, A. J., Markusheva, T. V., Kramerov, D. A., Ryskov, A. P., Skryabin, K. G., Bayev, A. A. and Georgiev, G. P. (1983) *Nucleic Acids Res.*, **10**, 7461–7475.
- Bennett, K. L., Hill, R. E., Pietras, D. F., Woodworth-Gutai, M., Kane-Haas, C., Houston, J. M., Heath, J. K. and Hastie, N. (1984) *Mol. Cell. Biol.*, **4**, 1561–1571.
- Hastie, N. D. (1989) in *Genetic Variants and Strains of the Laboratory Mouse*, eds. Lyon, M. F. and Searle, A. G. (Oxford Univ. Press, Oxford), 2nd Ed., pp. 559–573.
- Murray, A. W., Schultes, N. P. and Szostak, J. W. (1986) *Cell*, **45**, 529–536.
- Forget, B. G. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 7909–7911.
- Capecchi, M. R. (1993) *Nature*, **362**, 205–206.
- Gnirke, A., Huxley, C., Peterson, K. and Olson, M. V. (1993) *Genomics*, **15**, 659–667.