Mammalian Genome 5, 416–423 (1994).

Mammalian Genome

Generation and characterization of an ordered lambda clone array for the 460-kb region surrounding the murine *Xist* sequence

C. Rougeulle,[1] L. Colleaux,[2] B. Dujon,[2] P. Avner[1]

[1]Unité de Génétique Moléculaire Murine, Institut Pasteur, 25 rue du Dr Roux, F-75724 Paris-Cedex 15, France.
[2]Unité de Génétique Moléculaire des Levures, Institut Pasteur, 25 rue du Dr Roux, F-75724 Paris-Cedex 15, France

Received: 31 December 1992 / Accepted: 23 February 1994

Abstract. The *Xist* sequence has several characteristics that make it a potential candidate for the X-inactivation center. To investigate the role of *Xist* and adjacent sequences lying within the X-inactivation center candidate region, a 460-kb region surrounding the murine *Xist* sequence has been arrayed in lambda contigs with a combination of IRS-PCR-based hybridization and YAC fragmentation. The orientation of the *Xist* sequence in relation to the telomere and centromere of the X Chromosome (Chr) has been established with this contig and shown to be inverted compared to that in human.

Introduction

Transcription of the X Chr in mammals, including mouse and human, is subject to regulation such that activity of most X-linked genes is similar in both males and females, despite the presence of an additional copy of the X Chr in the female. This is achieved by the silencing of one of the two X Chrs present in each cell of the female adult soma, a process known as X-inactivation.

Studies of chromosomal rearrangements, and in particular reciprocal X-autosome translocations, have led to the assumption that a unique region, the X-inactivation center (*Xic*), is necessary for inactivation to be initiated. This is based on findings that in such X-autosome translocations only one of the two X Chr fragments undergoes X-inactivation (Russell 1983). With such translocations, the mouse *Xic* has been localized between the T16H translocation breakpoint, in the distal part of band XD (Rastan 1983), and the breakpoint characterized in the embryonic stem cell line HD3 (Rastan and Robertson 1985), lying close to the junction of bands XD and XE. Within this span, the region covered by the $Ta25^H$ deletion can be excluded from consideration, since $Ta25^H$ mice undergo normal X-inactivation (Cattanach et al. 1991). The candidate region for the mouse *Xic* is compatible with that in human where, in a similar manner, the existence of a larger number of rearrangements has allowed the mapping of XIC to a more restricted region, between PHKA1 and PGK1. More recently, the report of a patient with a rearranged X Chr has refined the distal limit of XIC in a region lying 200–700-kb distal to XIST (Leppig et al. 1993).

Comparative mapping has established that the murine and human X Chr are highly conserved (Amar et al. 1988), differing only by a series of intrachromosomal rearrangements involving a minimum of five breakage events, which have occurred during the period of evolutionary divergence separating primates from rodents. Within each of the segments so defined, conservation of gene order appears almost complete between the two species (Faust et al. 1992), leading to the expectation that in the mouse the inactivation center must similarly lie distal to *Phka* and proximal to *Pgk1*.

Recently, a sequence, known as XIST for X-inactive-specific-transcript, has been isolated in human (Brown et al. 1991). This sequence has the unique characteristic of being specifically transcribed from the inactive X Chr. It has been assigned to Xq13 by in situ hybridization and its localization refined by use of somatic cell hybrids, which placed it within the XIC candidate region. A murine XIST homolog, *Xist,* has been identified, and its localization and expression pattern shown to be similar to those in human (Borsani et al. 1991; Brockdorff et al. 1991). On the basis of these observations, it has been suggested that *Xist* may play a role in the X-inactivation process.

As part of our ongoing studies concerning the X-inactivation process and the *Xic* region in the mouse in particular, we report here on the analysis of a 460-kb mouse YAC containing the *Xist* sequence. Joint use of an IRS-PCR-based technique for establishing clone overlap, which we have adapted for use in mouse, and a I-*Sce*I based YAC fragmentation technique, allowing the orientation and distribution of primary contigs, has facilitated the rapid establishment of a contig array corresponding to the entire YAC, and the analysis and identification of structural features associated with the region. Such features include the apparent inversion of the murine *Xist* sequence compared to human.

Correspondence to: P. Avner

Materials and methods

YAC PA-2 subcloning

Total yeast DNA containing YAC PA-2 was partially digested with *Sau*3 A and size selected on a 0.8% agarose gel. DNA from slices containing fragments of 15–25-kb was electroeluted and purified on an elutip column (Schleicher & Schuell). 250 ng of purified DNA was ligated with 500 ng of λ DASH vector (Stratagene) and packaged by use of a Stratagene packaging extract, resulting in a total of 200,000 pfu. Half of this library was plated onto 210-mm plates at a density of 25,000 phages per plate, and colony lifts were taken with Hybone N+ membrane (Amersham). The mouse repeat probes used to identify mouse DNA containing clones were B2 and Tu96 (derived from the R repeat) (Nelson et al. 1984). The sublibrary was propagated in the *E. coli* SRB(P2) strain.

Two additional genomic libraries were used for screening: a commercially available 129/Sv female mouse cosmid library from Stratagene and a female mouse lambda library constructed and kindly provided by Bernhard Hermann.

PCR amplification

Mixed oligonucleotide primers. Phage plaques were touched with a toothpick, which was then soaked in 50 µl of TE, 2 µl of which was then used for the PCR. A mix of 4 primers corresponding to the B2 and L1 sequences were used in combination with vector primers. The B2 primers used were B2A1 and B2A3, the L1 primers were R1 and R2 (Simmler et al. 1991), while the vector primers used corresponded to the T3 and T7 promoter sequences. PCR was carried out with an automated thermal cycler (TECHNE) in a total volume of 50 µl under the following conditions: 1 mM of each primer in 50 mM KCl, 10 mM Tris-HCl pH 8.0, 1.5 mM $MgCl_2$, 0.01% gelatin, 125 mM each dNTP, and 2.5 units of Thermus Aquaticus Polymerase (Perkin Elmer/Cetus). A denaturation step of 4 min at 94°C was followed by 30 amplification cycles consisting of a 94°C denaturation (1 min), a 51°C annealing (1 min), and a 72°C extension (4 min), step completed by a final 8-min extension step at 72°C.

Specific primers. For the walking step, two primers were used, one from the vector, the other corresponding to a mouse repeat. In this case, PCR was performed with 100 ng of purified phage DNA under the conditions described above. For each amplification reaction, control PCR with the B2A3 primer alone was performed to identify and exclude fragments that were due to internal IRS amplification between neighboring B2 sequences.

Labeling of PCR products, probes, and hybridization

Out of the 50 µl of PCR, 30 µl was loaded onto a 1% low-melting agarose gel. Bands corresponding to the amplification products were cut out, and the equivalent of 50 ng of DNA was used for random priming (Amersham). Prehybridization of the probe with 300 µg of total mouse DNA was performed to block repetitive sequences.

Hybridizations were carried out with 0.5 M sodium phosphate, pH 7.2, 1 mM EDTA, and 7% SDS at 65°C, modified from Church and Gilbert (1984). Filters were washed in 40 mM sodium phosphate, pH 7.2, 1% SDS at 65°C.

The HR1 and HF2 fragments correspond respectively to the *Eco*RI-*Xba*I and *Kpn*I-*Xba*I restriction fragments of the MR20 clone (Borsani et al. 1991; see Fig. 5D).

Fragmentation

E. coli and yeast strains. Integrative vectors were constructed and propagated in *E. coli* TG1 strain [(Δlac pro) thil, sup E44, hsdD5, F′ (traD36, $proA^+B^+lacI^Q lacZ\Delta M15$)]. The YAC-containing yeast strain is an AB 1380 derivative (1): a, *ura 3, ade2, trp1, lys2-1, his5, can1-100.*

*Construction of integrative vectors with I-*Sce *I sites.* Integration of I-*Sce* I sites at random positions. A 171-bp-long fragment corresponding to the B2 repeat sequence was PCR amplified with oligonucleotides B2R1: 5′GGAATTCGGGCTGGAGAGATGGCT3′ and B2Xba1: 5′GCTCTAGATTTATTATATGTAAGTACACTG3′. The amplified product was then digested with *Eco*RI and *Xba*I, and a three-way ligation of the *Hin*dIII-*Xba*I *Lys2* cassette from pAF 107 and this *Eco*RI-*Xba*I PCR fragment into the *Hin*dIII and *Eco*RI sites of the pSK^+ vector was performed (Stratagene).

Integration of I-*Sce* I sites into the *Xist* sequence. A 4832-bp *Hin*dIII-*Eco*RV fragment containing the *Lys2* gene was introduced into the *Hin*dIII-*Sma*I sites of pAF100 (18) to produce pAF107. A 1270-bp *Hin*dIII-*Kpn*I fragment internal to the *Xist* gene (Borsani et al. 1991) was subcloned into the pSK^+ vector to produce the pSKHF vector. The 4864-bp *Hin*dIII-*Xba*I fragment from pAF107 containing the *Lys2* gene and the I-*Sce* I site was then introduced into the *Hin*dIII-*Xba*I sites of pSKHF.

Construction of transgenic strains. Yeast cells containing the YAC PA-2 were transformed by electroporation (Biorad) with 2 µg of in vitro linearized integrative DNA vectors. Transformants were selected on minimal glucose medium selective for lysine. Transformants were isolated on minimal glucose medium selective for uracil and tryptophan to ensure the presence of the YAC after electroporation.

Digestion of yeast chromosomal DNA in agarose plugs. Agarose plugs and in vitro digestion by I-*Sce* I were carried out as described by Thierry and Dujon (1992).

Results

Organization into contigs of YAC subclones

YAC PA-2 was isolated by screening a total mouse genomic YAC library made from *Eco*RI partially digested C3H male DNA (Larin et al. 1991) with the *Xist* cDNA sequence (Borsani et al. 1991). Both ends of this YAC were shown to be X-derived (Heard et al. 1993). A schematic map of YAC PA-2 is shown in Fig. 1A. A total genomic library of the yeast strain carrying YAC PA-2 was constructed in the lambda DASH II vector, and clones originating from mouse genomic DNA were identified by screening the library with probes for B2 and R sequences, belonging respectively to the SINEs and LINEs families of mouse repeat elements (Hastie 1989). These two probes were chosen for screening because they are both rodent specific and of high copy number. They are expected to occur approximately once every 30 kb (Hastie 1989). A total of 450 such mouse-specific clones were identified and organized into five 96-well matrices without further plaque purification. Secondary screening of these ordered matrices with the same repeat probes eliminated false positives. Subsequently, 410 clones were used for contig construction.

Several approaches to contig establishment including fingerprinting, restriction mapping, and sequencing have been described (Evans 1991). All these methods require DNA extraction of individual clones and are thus time consuming, especially when lambda contigs are being constructed. We have chosen to adapt for use in the mouse a recently developed method based on the amplification of random fragments within cloned inserts with direct Alu-PCR of lysis supernatants (Pieretti et al. 1991). Amplified

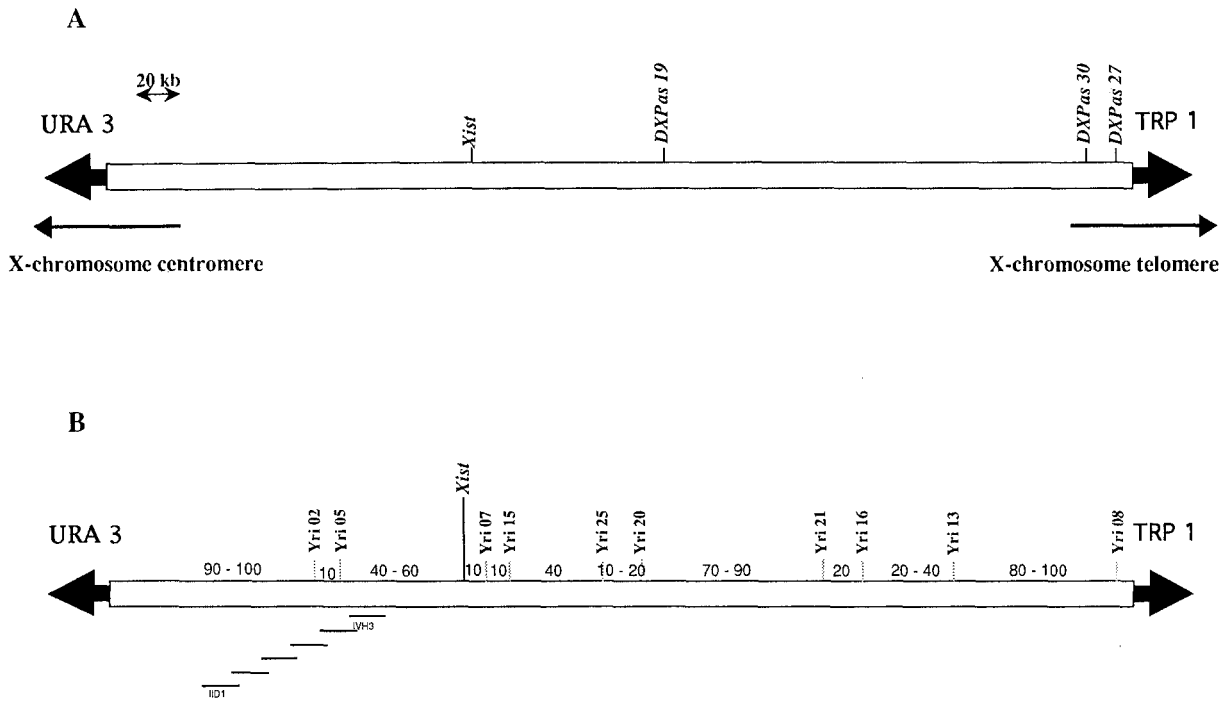

Fig. 1. Map of YAC PA-2. The **open box** represents the 460-kb insert, **arrows** the YAC's vector arms. The YAC is orientated with respect to the X Chr centromere and telomere. **(A)** Localization of known markers within the YAC. **(B)** The I-*Sce* I map deduced from PFGE measurements. Physical intervals separating the different insertions (Yri for YAC random integration) are indicated. The length variation observed for some intervals is due to the limit in the precision of PFGE measurements. As an example, the distribution of one of the initial contigs with flanking members IID1 and IVH3 is shown in relation to insertion Yri02.

fragments are then used to probe clone matrices and identify overlapping clones. An example of such a hybridization is shown in Fig. 2.

In the initial phase of contig establishment, the lambda clones used as substrates for PCR were chosen randomly from the YAC-derived lambda library, and a mix of three primer pairs was used. One set corresponded to the 3′ and 5′ ends of the B2 sequence, and another pair of primers to the ends of the R sequence. These primer pairs are known to allow efficient IRS-PCR amplification of mouse material in IFGT hybrids (Simmler et al. 1991). The other pair of primers used corresponded to the T3 and T7 promoter sequences flanking the cloning site of the vector and allow the amplification of IRS-vector products. Simultaneous use of more than a single primer pair increased significantly the number of clones yielding useful amplification products (data not shown). Using this method, we were initially able to establish a series of 12 contigs, each 30–70 kb long, covering approximately 350 kb of the 460 kb insert of the original YAC PA-2.

Random overlap strategies of contig mapping rarely result in the establishment of complete physical maps, but instead generate sets of "islands" of ordered clones separated by gaps. The closure strategies necessary to fill such gaps all require the isolation of end-specific probes. End-specific probes corresponding to the clones located at the extremities of the existing contigs were, therefore, generated by IRS-vector PCR. The efficiency of this approach depends on the representation and orientation of repeats in proximity to the vector arm, which serve as the second internal primer. Since preliminary experiments demonstrated that the B2A3 primer gave a high percentage of amplification products, it was used for such IRS-vector amplification (data not shown). Higher end-specific amplification efficiency was observed when such PCR were carried out directly on purified DNA rather than on lysis products.

Through the use of this approach, eight of the initial dispersed contigs were increased in size and then consolidated into four longer ones, covering approximately 400 out of the 460 kb of the YAC insert.

One of the three remaining gaps was just proximal to the lambda 510 (Fig. 4). We were unable to generate IRS-vector product from the proximal end of this clone, but this in itself is not surprising since this region corresponds in part to the first exon of the *Xist* sequence (Brockdorff et al. 1992). A set of primers was therefore designed from the *Xist* cDNA sequence and the probe derived used to screen the initial lambda sublibrary (without prescreening with mouse-specific repeat probes). Among the clones isolated, one was shown to overlap lambda IVH3 and thus fill the gap.

Screening not only of the lambda sublibrary but also of a cosmid library with flanking probes to the two remaining gaps was unsuccessful. We have, therefore, constructed a "mini-YAC" to cover these gaps in the contig. A series of shortened derivatives of YAC PA-2 have been created by means of B1-sequence targeted mouse YAC fragmentation based on the technique described by Pavan and his colleagues (Pavan et al. 1990) for human YACs (Heard et al. in press). One of these YAC PA-2 derivatives, 1PAS9, which is about 150 kb long and covers the distal (TRP1) portion of YAC PA-2, covers the uncloned region. The finalized map is shown in Fig. 4.

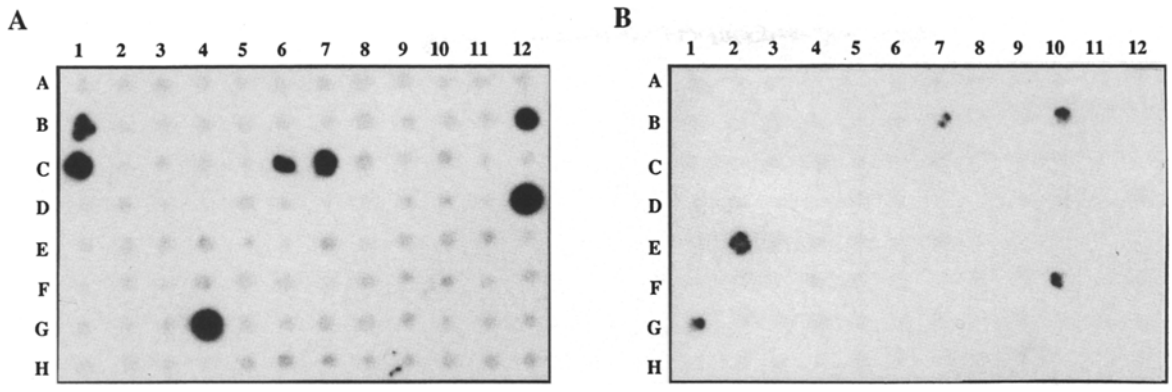

Fig. 2. Replica filters of a lambda clone matrix hybridized with IRS-PCR probes. **(A)** A IIIC3-derived probe and **(B)** A IB6-derived probe. A different subset of clones is detected by each of the probes.

Localization and orientation of the initial contigs

Fragmentation of endogenous yeast chromosomes by the introduction of artificial restriction sites for the endonuclease I-*Sce* I has been described previously (Thierry and Dujon 1992) and the feasibility of using a similar fragmentation approach for mammalian inserts cloned in YACs established (Colleaux et al. 1993). In this method, yeast cells carrying such YACs are transformed by a plasmid in which a I-*Sce* I site has been artificially inserted close to a cloned B2 repeat sequence. Homeologous recombination between the cloned B2 sequence and B2 repeats present in the YAC insert leads to the integration of the plasmid within the YAC. It has been shown that YACs having undergone such recombination normally carry a unique I-*Sce* I site, suggesting that multiple recombination events rarely occur. Digestion of such recombinant YACs with I-*Sce* I generates two fragments whose size depends on the localization of the plasmid insertion and can be determined by successive hybridizations with YAC arm-specific probes (URA3 and TRP1).

A series of ten recombinant YACs was obtained from YAC PA-2 and an I-*Sce* I map constructed (Fig. 1B). The position and left/right orientation of the contigs with respect to the YAC arms has been deduced from hybridization experiments with PCR probes corresponding to the most terminal clones in each primary contig. In all cases, only one arm of the digested YAC gave a positive signal, allowing easy positioning of each probe and the corresponding contig. Figure 3 shows the results obtained with probes from clones IID1 and IVH3, corresponding to opposite ends of a 70-kb-long contig (see Fig. 1B). The hybridization pattern of these two probes on recombinant YACs appears similar except in the case of modified YAC Yri 02. Probe $IID1_r$ hybridized with the URA 3 arm of the Yri02 (Fig. 3B), while $IVH3_l$ hybridized to its TRP1 arm (Fig. 3C). It can be deduced from this that the contig must span Yri 02. The rescue of a short genomic fragment adjacent to the I-*Sce* I site in YAC YRI 02 has confirmed this result. The clone identified with this genomic fragment as probe was found to lie in the middle of the above-mentioned contig (data not shown).

Phage DNA from four of the initial contigs was shown during such screening experiments to hybridize to yeast chromosomes, indicating that they contained yeast DNA.

Orientation of Xist using an integrative vector approach

With experiments similar to those performed with B2, an I-*Sce* I site was integrated within the *Xist* sequence of the YAC. The digestion of the recombinant YAC with I-*Sce* I generates two fragments whose size can be determined by hybridization with URA3 and TRP1 probes. Hybridization has demonstrated that *Xist* is positioned 150 kb from the URA3 arm, a position consistent with data previously obtained from pulse field gel electrophoresis (PFGE; Heard et al. 1993). The configuration of the *Xist*/I-*Sce* I integrative plasmid used was such that the I-*Sce* I site was integrated between two sub-sequences of the murine *Xist* locus, HR1 and HF2 (Fig. 5D). These sequences had previously been shown to lie 5′-3′ respectively according to the direction of transcription of *Xist*. Hybridization of the fragmented YAC with probes for HR1 and HF2 has allowed the orientation of *Xist* with respect to the YAC, since HR1 hybridized with the short arm of the digested YAC (Fig. 5B) while HF2 hybridized to its long arm (Fig. 5C). The orientation of *Xist* with respect to the YAC must therefore be URA3 arm–5′*Xist*3′–TRP1 arm. Since the YAC itself had previously been oriented with respect to the centromere and telomere of the X Chr (see Fig. 1A), the overall orientation of the murine *Xist* sequence can be shown to be centromere–5′*Xist*3′–telomere.

Discussion

On the basis of data obtained both in human and mouse, the candidate region for the murine *Xic* has been located between the *Phka* and *Pgk1* loci. Recently, the identification of the human and murine *Xist* sequence and the definition of some of its attributes, such as transcription from the inactive X Chr and localization within the candidate region for *Xic,* have led to the suggestion that *Xist* and *Xic* may be one and the same. In order to investigate systematically the involvement in X-inactivation of *Xist* and other sequences lying within the *Xic* candidate region, we have established lambda contigs covering almost the entire 460-kb region surrounding the murine *Xist* sequence. The lambda subclones were derived from a YAC isolated after screening of a mouse YAC library with the *Xist* sequence (Heard et al. 1993; Larin et al. 1991).

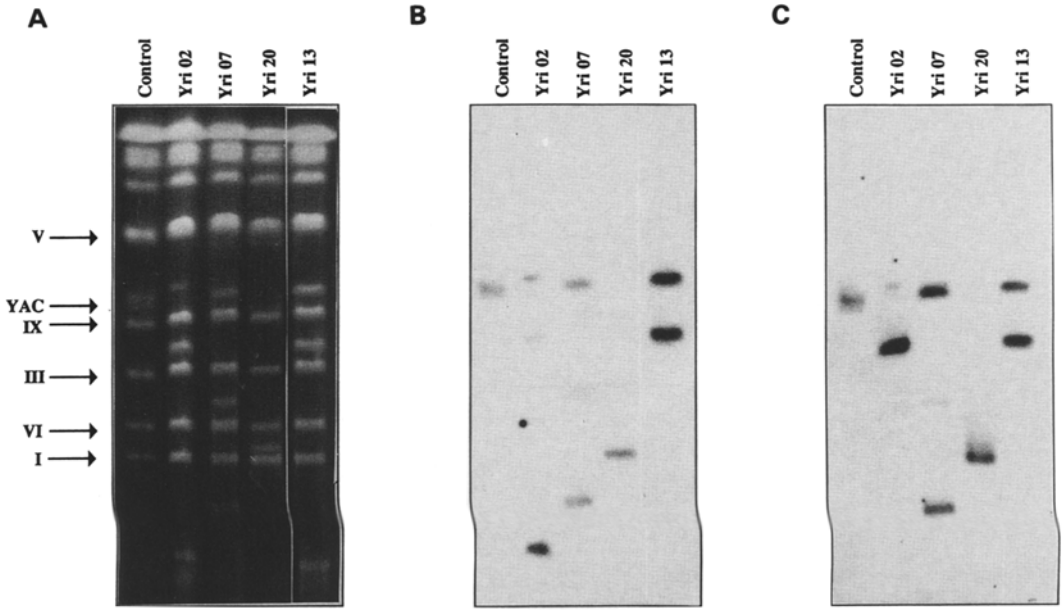

Fig. 3. Hybridization of lambda clones to the fragmented YACs. Chromosomes from strains Yri02, 07, 20, and 13 were digested with I-*Sce* I under conditions described in methods. DNAs were electrophoresed on a 1% agarose (Seakem) gel in 0.25 × TBE buffer at 130 V and 12°C on a Rotaphor apparatus (Biometra) for 70 h by use of 30- to 100-s increasing pulse times. **(A)** Ethidium bromide-stained gel. Chromosomes from AB1380 are indicated as control. The numbers on the **left** of the figure indicate the yeast chromosomes used as calibration standards; other yeast chromosomes were unresolved. **(B)** The IID1-derived probe hybridized with the URA3 arms of Yri02, 07, 20, and 13. **(C)** The IVH3-derived probe hybridized with the URA3 arms of Yri07, 20, and 13 and the TRP1 arm of Yri02.

The strategy developed here to array the lambda clones into contigs is based on the adaptation and combination of two previously described methods. In an initial phase, a series of eight contigs was established with a technique based on the serial hybridization of IRS-PCR probes obtained from lambda phage plaques against lambda clone arrays. This method was initially developed and applied to human DNA-containing clones with probes being generated by Alu-PCR (Pieretti et al. 1991). We have successfully adapted it for use in the mouse with a panel of primers derived from mouse-specific repeats. We have, moreover, shown that the simultaneous use of several repeat derived primers increases the overall efficiency of amplification. This increased efficiency may in part be a reflection of the low PCR annealing temperature used, which was chosen in order to allow coamplification from the T3 and T7 primers. This may have had the effect of increasing priming efficiency with respect to degenerate B2 and L1 consensus sequences present in the genomic DNA. The use of IRS-PCR confers a number of advantages, since it can be carried out directly on phage supernatants and eliminate time-consuming procedures such as plaque purification of lambda phage clones, isolation of phage DNA, and the subcloning of insert termini. Moreover, IRS-PCR amplification generates efficient probes since 70% of the amplification products gave usable hybridization signals. IRS-vector PCR was used for contig closure rather than RNA probes, which are known not only to be unstable but also to require a large amount of matrix DNA for their synthesis. As the bacterial host strain used [SRB(P2)] is deficient for a number of loci involved in recombination, obtaining sufficient lambda DNA for RNA probe synthesis proved to be problematic. Despite this drawback associated with the SRB(P2) strain, we believe its use is helpful in ensuring the representativeness of libraries constructed from repeat rich regions such as the one under study here. The IRS-vector PCR closure strategy proved to be particularly efficient in that several clones identified by this method showed less than 3 kb of overlap (unpublished results).

The efficiency of the contig assembly approach we have used depends on (a) the degree of cloning bias across the region; (b) the distribution of mouse-specific sequences within the region, which will influence the identification efficiency of lambda clones containing mouse insert; and (c) the efficiency of IRS-PCR or IRS-vector PCR product generation from individual lambdas, which is also influenced by the distribution of the repeat sequences.

Among the 410 secondarily screened clones derived from the YAC PA-2 sublibrary, 14 (3.4%) were negative with all of the probes used. These clones may be derived from yeast genomic DNA. The 396 remaining clones represent approximately 6000-kb of insert, or close to 13-fold coverage of the region of interest. Cloning bias was almost certainly occurring, because not all the regions of the YAC were represented equally. Probes derived from lambda IB6, for example (see Fig. 4), recognized 11% of the total matrix of clones, while probes from IVE3 recognized only 1% of the matrix.

Although we are not in a position to ascertain precisely the reasons for our failure to close the two remaining gaps, the λ3 clone, which closed a previous contig gap between IVH3 and 510, appears to be deficient in repeat sequence and would have been picked only up by the type of direct selection with a unique probe that was finally used for gap closure between the 510 and IVH3 clones.

Our efforts to apply a similar approach, involving isolation of a unique sequence probe, to the lambdas surrounding the two remaining gaps has, however, been ham-

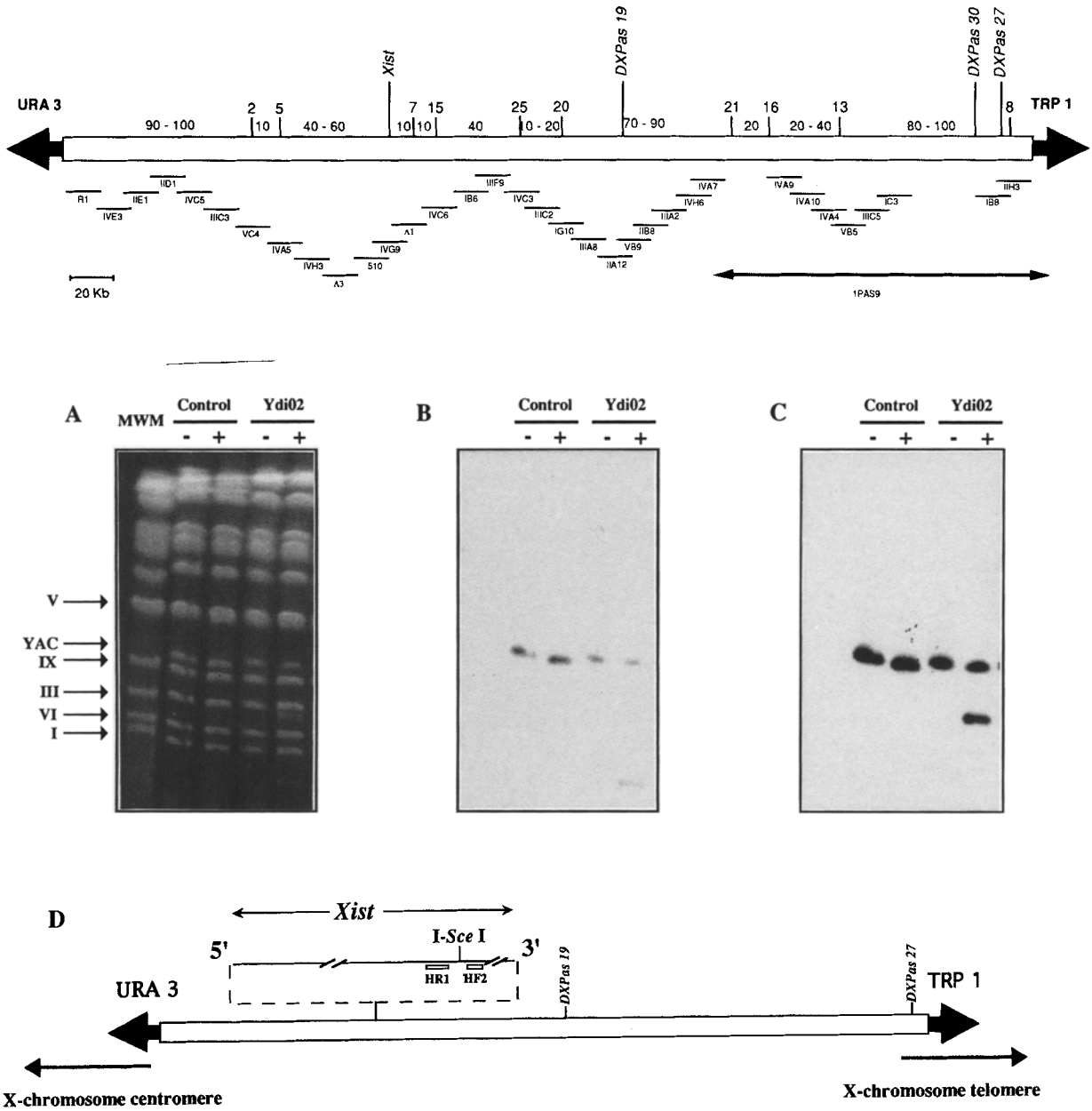

Fig. 5. *Xist* orientation. Chromosomes from AB1380/PA-2 (control) and from a transgenic strain with the I-*Sce* I site inserted within the *Xist* sequence (Ydi 02 for YAC directed integration). −, undigested DNA; +, DNA digested with I-*Sce* I. **(A)** Ethidium bromide-stained gel. MWM: molecular weight markers. The numbers on the **left** of the figure indicate the short yeast chromosomes used as calibration standards. **(B)** HR1 probe hybridizes to the URA3 arm of Ydi02. **(C)** HF2 hybridizes with the TRP1 arm. **(D)** Representation of the inserted I-*Sce* I site between the HR1 and HF2 fragments of the *Xist* sequence.

pered by the extremely repeat-rich nature of these lambdas. FISH (fluorescence in situ hybridization) analysis of lambdas IB8 and IIH3 has confirmed the unusually repeat-rich nature of lambda clones in this region (C. Vourch, personal communication).

On the basis of our experience, some 10% of a given region may be expected to prove refractory to contig assembly by the sole use of the IRS-PCR technique. The necessity of using several independent approaches if contig closure is to be achieved completely has been amply demonstrated in other experimental systems.

The initial contigs were subsequently mapped and orientated with respect to one another and to the URA3 and TRP1 arms of the YAC vector by use of a panel of YAC derivatives obtained by I-*Sce* I-based fragmentation of the initial YAC. The construction of this I-*Sce* I map, with an average resolution of 45-kb, allowed positioning of the eight initial contigs on this map and showed five of them to be regularly spaced across the YAC. These observations are consistent with the distribution of the B2 and L1md repeats (from which the primers used in IRS-PCR were derived) deduced from hybridization experiments (see below). Four other initial contigs were identified by this mapping step as being yeast derived, presumably owing to contamination of mouse lambda clones by other clones containing yeast DNA.

The combined use of two independent strategies to order YAC-specific lambda subclones significantly increased

the rapidity and feasibility of contig establishment. The mouse repeat B2 has proved to be efficient not only as a primer in the IRS-PCR approach, but also as target during fragmentation of the YAC, since it enabled the construction of the 45-kb average resolution I-*Sce* I map. However, the combined use of different repeats, such as B1 and L1, as targets would undoubtedly increase the resolution of the map.

Short and long interspaced nuclear elements have been shown by in situ hybridization (Boyle et al. 1990) to be non-uniformly distributed in most metaphase chromosomes, with SINEs being concentrated within R bands and LINEs within G bands. The X Chr, however, failed to show significant R banding by this technique. On this basis, it has been concluded by Boyle and her colleagues that the X Chr is so rich in L1 sequences that R bands are virtually obscured. At the level of resolution afforded by this technique, LINEs and SINEs on the X Chr do not appear necessarily to be clustered into mutually exclusive chromosomal subregions. The establishment of the contig allowed investigation of the distribution of B2 and L1 repeats by hybridizing B2 and R probes to phage clones constituting the entire contig. Out of the 34 lambda clones required to cover the 460-kb insert, only three (9%) proved to be negative upon hybridization with the B2 probe. Both these clones localize proximal to *Xist*. Nineteen clones (56%) were positive with the R probe. The R-positive clones did not, however, fall within a single region, but are instead organized into several clusters dispersed along the YACs. Similar results have been reported for human YACs by Arveiler and Porteous (Arveiler et al. 1992). As expected, no clones were found to be hybridization negative with both probes, with the exception of the clone λ3, isolated by direct screening of the initial sublibrary. The murine X-inactivation center has been mapped to the distal part of the D band, a Giemsa-negative, or R band. Such a region would be expected to be rich in B2 sequences, as observed in our hybridization experiments. If the X Chr is particularly rich in L1 and such LINEs are not organized into well-separated subregions of the X Chr, our finding that over 50% of the lambda clones from this region contain L1 sequences is not surprising.

The establishment of this contig surrounding the *Xist* sequences has enabled the identification of several features specific to this part of the X-inactivation center candidate region and in particular the immediate vicinity of *Xist*. It has, for example, enabled a *Sal*I site, known from PFGE to lie approximative 10–40-kb distal to *Xist* and which possesses an intriguing pattern of methylation with specific methylation associated with the active rather than the inactive X Chr, to be located in lambda λ1 (see Fig. 4). The identification of this clone has allowed experiments to be undertaken to prove whether this *Sal*I site is indicative of a transcriptionally active region restricted to the inactive X Chr and whether such a region is contiguous with *Xist* itself (Courtier et al. submitted).

The contig has also facilitated the identification of $(CA)_n$ microsatellites in this 460-kb region, which were subsequently used to map the *Xce* locus (Simmler et al. 1993).

Studies concerning the comparative organization of the human and murine X Chr have generally shown a high degree of conservation, and it has been shown that rearrangements between the human and the mouse can be explained by only a limited number of intrachromosomal breakage events. Within the blocks so defined, the gene order appears to be conserved, although an exception involving the mouse homolog of the human factor VIII has been recently described by Faust and her colleagues (Faust et al. 1992). The *Xist* sequence provides a second example of such apparent local non-conservation. Insertion by homologous recombination of a I-*Sce* I site within the 3′ end of the *Xist* sequence enabled the orientation of *Xist* to be established as centromere-5′ *Xist* 3′-telomere, an orientation that appears to be the opposite to that found in human (Brown et al. 1992). This finding raises questions as to the extent of this inversion. Is it specific to *Xist,* or does it involve a larger region? Although the absence of cross-reacting conserved markers in the immediate region of *Xist* does not currently allow this question to be answered, the finding of several HTF islands within the YAC PA-2 (Heard et al. 1993) and their localization to restricted parts of the contig array described here suggest that coding sequences which would be expected to show cross-species hybridization will rapidly become available.

Acknowledgments. We thank our colleagues from the Unité de Génétique Moléculaire Murine for stimulating discussion and Iona Millwood for help with isolation of some lambda clones. This work was supported by grants from the Association Française contre la Myopathie and the Ministère de le Recherche et de l'Espace to P.A (92C0040). L. Colleaux and B. Dujon were supported by the Human Genome Analysis Programmes of the Commission of the European Communities and by grants from Université Pierre et Marie Curie.

References

Amar, L.C., Dandolo, L., Hanauer, A., Ryder Cook A., Arnaud, D., Mandel, J.-L., Avner, P. (1988). Conservation and reorganization of loci on the mammalian X chromosome: a molecular framework for the identification of homologous subchromosomal regions in man and mouse. Genomics 2, 220–230.

Arveiler, B., Porteous, D.J. (1992). Distribution of Alu and L1 repeats in human YAC recombinants. Mamm. Genome 3, 661–668.

Borsani, B., Tonlorenzi, R., Simmler, M.-C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzni, D., Lawrence, C., Willard, H.F., Avner, P., Ballabio, A. (1991). Characterization of a murine gene expressed from the inactive X chromosome. Nature 351, 325–329.

Boyle, A.L., Ballard, S.G., Ward, D.C. (1990). Differential distribution of long and short interspaced element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization. Proc. Natl. Acad. Sci. USA 87, 7757–7761.

Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., Rastan, S. (1991). Conservation of position and exclusive expression of mouse *Xist* from the inactive X chromosome. Nature 351, 329–331.

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., Rashan, S. (1992). The product of the mouse *Xist* gene is a 15 kb inactive x-specific transcript containing no conserved ORF and located in the nucleus. Cell 71, 515–526.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniè̀e, R.G., Grompe, M., Tonlorenzi, R., Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 349, 38–44.

Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J., Willard, H.F. (1992). The human *XIST* gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 71, 527–542.

Cattanach, B.M., Rasberry, C., Evans, E.P., Dandolo, L., Simmler, M.C., Avner, P. (1991). Genetic and molecular evidence of an X-chromosome

deletion spanning the tabby (*Ta*) and testicular feminization (*Tfm*) loci in the mouse. Cytogenet. Cell Genet. 56, 137–143.

Church, G.M., Gilbert, W. (1984). Genomic sequencing. Proc. Natl. Acad. Sci. USA 81, 1991–1995.

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

Evans, G. (1991). Combinatoric strategies for genome mapping. BioEssays 13, 39–44.

Faust, C.J., Levinson, B., Gitschier, J., Herman, G. (1992). Extension of the physical map in the region of the mouse X chromosome homologous to human Xq28 and identification of an exception to conserved linkage. Genomics 13, 1289–1295.

Hastie, N.D. (1989). Highly repeated DNA families in the genome of *Mus musculus*. In *Genetics Variants and Strains of the Laboratory Mouse*. M.F. Lyon, A.G. Searle, eds. (Oxford: Oxford University), pp. 559–573.

Heard, E., Simmler, M.C., Larin, Z., Rougeulle, C., Courtier, B., Lehrach, H., Avner, P. (1993). Physical mapping and YAC contig analysis of the region surrounding *Xist* on the mouse X chromosome. Genomics 15, 559–569.

Larin, Z., Monaco, A.P., Lehrach, H. (1991). Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. Proc. Natl. Acad. Sci. USA 88, 4123–4127.

Leppig, K.A., Brown, C.J., Bressler, S.L., Gustashaw, K., Pogon, R.A., Willard, H.F., Disheche, C.M. (1993). Mapping of the distal boundary of the X-inactivation center in a reaaranged X chromosome from a female expressing *Xist*. Hum. Mol. Genet. 2, 883–887.

Nelson, D.L., Weis, J.H., Przyborski, M.-J., Mulligan, R.C., Seidman, J.G., Housman, D.E. (1984). Metaphase chromosome transfer of introduced selectable markers. J. Mol. Appl. Genet. 2, 563–577.

Pavan, W.J., Hieter, P., Reeves, R.H. (1990). Generation of deletion derivatives by targeted transformation of human-derived yeast artificial chromosomes. Proc. Natl. Acad. Sci. USA 87, 1300–1304.

Pieretti, M., Tonlorenzi, R., Ballabio, A. (1991). Rapid assembly of λ phage contigs within YAC clones. Nucleic Acids Res. 19, 2795–2796.

Rastan, S. (1983). Non-random X-chromosome inactivation in mouse X-autosome translocation embryos—location of the inactivation center. J. Embryol. Exp. Morphol. 78, 1–22.

Rastan, S., Robertson, E.J. (1985). X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J. Embryol. Exp. Morphol. 90, 379–388.

Russell, L.B. (1983). X-autosome translocations in the mouse. Their characterization and use as tools to investigate gene inactivation and gene action. In *Cytogenetics of the Mammalian X Chromosome, Part A: Basic Mechanisms of X Chromosome Behaviour*. Sandberg A.A., ed. (New York: A.R. Liss) pp. 205–250.

Simmler, M.-C., Cox, R.D., Avner, P. (1991). Adaptation of interspersed repetitive sequence polymerase chain reaction to the isolation of mouse DNA probes from somatic cell hybrids on a hamster background. Genomics 10, 770–778.

Simmler, M.-C., Cattanach, B.M., Rasberry, C., Rougeulle, C., Avner, P. (1993). Mapping the murine *Xce* locus with (CA)n repeats. Mamm. Genome 4, 523–530.

Thierry, A., Dujon, B. (1992). Nested chromosomal fragmentation in yeast using the meganuclease I-*Sce* I: a new method for physical mapping of eukaryotic genome. Nucleic Acid Res. 20, 5625–5631.