Nucleic Acids Research, Vol. 20, No. 21 **5625–5631**

Nested chromosomal fragmentation in yeast using the meganuclease I-*Sce* I: a new method for physical mapping of eukaryotic genomes

Agnès Thierry and Bernard Dujon
Unité de Génétique Moléculaire des Levures, Département de Biologie Moléculaire (URA 1149 du CNRS), Institut Pasteur, 25 rue du Dr Roux, F-75724 Paris-Cedex 15, France

Received August 21, 1992; Revised and Accepted October 12, 1992

ABSTRACT

We have developed a new method for the physical mapping of genomes and the rapid sorting of genomic libraries which is based on chromosome fragmentation by the meganuclease I-*Sce* I, the first available member of a new class of endonucleases with very long recognition sequences. I-*Sce* I allows complete cleavage at a single artificially inserted site in an entire genome. Sites can be inserted by homologous recombination using specific cassettes containing selectable markers or, at random, using transposons. This method has been applied to the physical mapping of chromosome XI (620 kb) of *Saccharomyces cerevisi* and to the sorting of a cosmid library. Our strategy has potential applications to various genome mapping projects. A set of transgenic yeast strains carrying the I-*Sce* I sites at various locations along a chromosome defines physical intervals against which new genes, DNA fragments or clones can be mapped directly by simple hybridizations.

INTRODUCTION

The development of rapid and reliable methods for high resolution physical mapping of large genomes and for the assembly of contigs from genomic libraries is a critical aspect of all genome projects. Several 'bottom up' strategies have been developed. One consists in assembling contigs by means of 'fingerprinting' random clones from complete libraries (1, 2, 3, 4). Another consists in hybridizing cosmid or lambda clones organized in matrices with short oligonucleotides representing specific sequences (5) or with end-specific probes synthesized from T3 or T7 promoters of chosen or pooled clones (6, 7). 'Top down' strategies on the contrary depend upon the separation of large DNA molecules by pulsed field electrophoresis and have been largely facilitated by the possibility of cloning large inserts in YAC vectors in yeast (8). Yet the success of the top down strategies is limited by the availability of rare cutter restriction endonucleases for the construction of long range physical maps. Strategies have been developed that combine the use of synthetic oligonucleotides, the RecA protein, and methylases to allow the cleavage of large genomes at a single or at a few sites using classical bacterial restriction endonucleases (9, 10). But these very elegant strategies involve several successive and delicate steps to be successful.

A new endonuclease, I-*Sce* I, which has a recognition site extending over 18 bp and is encoded by a mobile group I intron of yeast mitochondria, has been previously reported (11, 12, 13). Although some base substitutions are tolerated within the recognition sequence, most significantly reduce the efficiency of cleavage by the enzyme under optimal conditions (13). On a random DNA sequence with 50% GC content, the probability of occurence of the I-*Sce* I site is 1.4×10^{-11} i.e. one cleavage site expected for ca. 20 times the size of the human genome. I-*Sce* I, therefore, potentially permits cleavage of a complex genome at a single site, provided its recognition site can be artificially inserted. The actual specificity of I-*Sce* I has been directly verified on the genome of *Saccharomyces cerevisiae* (1.4×10^{7} bp). Complete cleavage of total DNA at an artificially inserted I-Sce I site without digestion of the other chromosomes has been demonstrated (14). Other genomes have been tested which confirm the general absence of I-*Sce* I recognition sites.

We have now developed a new mapping strategy based on the systematic insertion of artificial I-*Sce* I sites at various positions along a chromosome. A chosen set of transgenic strains containing appropriately spaced I-*Sce* I sites define physical intervals against which new genes or clones can be mapped by simple hybridizations. The method has been applied to the chromosome XI of yeast to construct a physical map and to rapidly sort cosmid libraries into a unique contig.

MATERIALS AND METHODS

E.coli and yeast strains

Cosmids and subclones were propagated in *E.coli* TG1 strain ((Δ lac pro) thi1, supE44, hsdD5, F′(traD36, proA^{+}B^{+} lac I^{Q} lacZΔM15).

FY1679 is a diploid yeast strain issued from the cross between the two strains FY23 (MATa ura3-52, trp1Δ63, leu2Δ1 GAL2) and FY73 (MATα ura3-52, his3Δ200, GAL2) kindly provided by Fred Winston. FY23 and FY73 are derived from the strain S288C and are isogenic with it except for the markers indicated. FY23/C6 is a transgenic strain derived from FY23 and carrying

the I-*Sce* I site in chromosome III (15). Other transgenic yeast strains derived from FY1679 are described below (see construction of transgenic yeast strains).

Cosmid library

Three complete cosmid libraries of the yeast genome have been constructed by partial Sau3A digestion of total DNA from strain FY1679 ligated into *Bam*HI-digested cosmid vectors pWE15 and pOU61cos (Thierry *et al.*, in prep). A chromosome XI specific sublibrary has been sorted out by colony hybridization using PFG purified chromosome XI as a probe. This sublibrary is composed of 138 independent clones respresenting, altogether, eight times the chromosome equivalent.

Construction of integrative vectors with the I-Sce I site

a—Integration of I-Sce I sites into previously cloned genes. A 708 bp fragment internal to the gene FAS1 and a 580 bp fragment internal to the gene TIF1 were prepared by PCR amplification from FY1679 DNA and inserted at the *Eco* RI site of pAF100 (15) to produce pAF301 and pAF303, respectively. A *Sal* I cassette and a *Kpn* I cassette from pAF101 have been inserted into the unique *Sal* I site of the insert of pAF301 or the unique *Kpn* I site of the insert of pAF303 to produce pAF302 and pAF304, respectively. pAF101 is a pUC19 derivative containing a URA3 cassette with the I-*Sce* I site (15).

b—Integration of I-Sce I sites at unknown positions. The 2.5 kb *Eco* RI fragment of cosmid pEKG081 from our chromosome XI specific sublibrary has been inserted into the *Eco* RI site of a pBluescript derivative to produce pAF019. The *Nco* I-*Hpa* I fragment internal to the insert of pAF019 (16) has been replaced by a *Bam* HI cassette of pAF101 using blunt end ligation to produce pAF021.

2μg of DNA from cosmids pUKG040, pUKG041, pUKG057 and pUKG062 from our chromosome XI specific sublibrary were digested by *Bam* HI and ligated with 0.1μg of *Bgl* II digested pAF101 DNA previously dephosphorylated and gel purified. Recombinants were transformed into TG1 and selected on LB plates containing 50μg/ml of ampicillin. From each original cosmid, one plasmid clone was chosen as containing an insert of appropriate size and a unique restriction site (*Hpa* I) useful for linearization prior to yeast transformation. Plasmids obtained were called pAF305 (4.0 kb insert of pUKG040), pAF306 (4.5 kb insert of pUKG041), pAF307 (2.1 kb insert of pUKG057) and pAF308 (9.0 kb insert of pUKG062).

Construction of transgenic yeast strains

FY1679 has been transformed by electroporation (BioRad) using *in vitro* linearized DNA: respectively, pAF302, 304, 021, 307, 305, 308 and 306 for FY1679/A302, /D304, /H81, /M57, /E40, /T62 and /G41. Transgenic yeast strains FY1679/A302, D304 and H81 were constructed using the 'ends out' topology, the others (FY1679/M57, E40, T62 and G41) using the 'ends in' topology (figure 2). Transformants were selected on minimal glucose medium selective for uracile. Ca. 40 to 50 (URA3+) colonies were obtained for 0.2μg of *Eco* RI digested pAF302, pAF304 and pAF021 plasmid DNA in the 'ends out' topology; ca. 400 to 800 (URA3+) colonies were obtained for 1μg of *Hpa* I digested pAF305, pAF306, pAF307 and pAF308 plasmid DNAs in the 'ends in' topology. In all cases, integrations of transforming DNA into homologous sites on chromosome XI were verified by Southern blotting of *Ssp* I digested transgenic yeast DNAs, probed with the URA3 cassette.

Digestion of yeast chromosomal DNA in agarose plugs

Yeast cells are harvested during mid-log or late-log phase, washed and resuspended in 0.05M EDTA at 3×10^9 cells /ml. One ml of cell suspension is added with 0.33 ml of solution A (prewarmed @ 37°C) and with 1.67 ml of solution B (prewarmed @ 45°C). After gentle mixing the suspension is poured in prewarmed molds, then let @ 4°C for gelification. Agarose blocs of ca. 60 μl (6×10^7 cells) are cut out, placed in 2 to 3 volumes of solution C and incubated ca. 24h @ 37°C. Blocs are transferred in 2 to 3 volumes of solution D and incubated overnight @ 50°C, then rinsed and stored @ 4°C in 0.5M EDTA pH9.0.

Solutions are: A: Na citrate 0.1M at pH5.8, sorbitol 1M, EDTA 0.01M, β mercaptoethanol 0.6M and Zymolyase 100,000 (Kirin) 1mg/ml; B: agarose (InCert, FMC) 1% (w/v) in 0.125M EDTA; C: EDTA 0.45M,Tris-Cl 0.01M, β-mercaptoethanol 0.9M, pH 8.5; D: EDTA 0.45M ,Tris Cl 0.01M, sarkosyl 1% (v/v) and proteinase K 1 mg/ml, pH8.5.

Before digestion with I-*Sce* I, agarose blocs are washed overnight in an excess volume of diethanolamine 0.1M, pH 9.5 @ 4°C. Washed blocs are then placed in Eppendorf tubes containing 160 μl of diethanolamine 0.1M, pH 9.5 dithiotreitol 0.001M, 0.2 mg/ml of bovine serum albumin and incubated 1 h. @ 4°C. Then, 20 U of I-*Sce* I (Boehringer Mannheim cat No. 1362 399) are added to each tube together with its 'enhancer' solution. Diffusion of the enzyme in the blocs is allowed for 2 h @ 0°C. Reactions are started by addition of 1.6 μl of a 1M $MgCl_2$ solution (8 mM final) and incubation @ 37°C for 1h.

Hybridizations

Chromosome fragments were cut out from PFGE gels (1% SeaPlaque GTG agarose, FMC) under UV illumination (312 nm). Agarose blocs (cut as small as possible, e.g. 50 μl) were rinsed twice in 40 vol of H_2O for 1hr @ 0°C to eliminate the TBE buffer and ethidium bromide. Labelling was done by random priming in agarose as follows. Blocs were melted at 95°C for 5 min. and added with 6 μl of DPN6 primers (Pharmacia), 10 μl of 10× random priming buffer (0.5 M tris-Cl @ pH 7, 0.1 M Mg SO_4 and 0.001 M dithiotreitol) and 19 μl of H_2O. The tubes were placed at 95°C for 10 min. to allow mixing, quickly spun and placed at 37°C for 10 min. Tubes were added with 5 μl of a mix of dCTP, dTTP and dGTP (10 mM each), 7 μl α ^{32}P-dATP and 12 units of Klenow polymerase and incubated at 37°C for 2 hrs. Labelled DNA is denatured at 95°C and used directly as a probe. All hybrizations were carried out according to 17.

RESULTS

Rationale of the nested chromosomal fragmentation strategy

I-*Sce* I sites can be inserted in a genome either at random or by homologous recombination after transformation with an appropriate vector or cassette containing the site. The rationale of our method is based on the fact that each artificially introduced I-*Sce* I site provides a unique 'molecular milestone' in the genome, if total DNA is cleaved with the enzyme. Therefore, a set of transgenic strains, each carrying a single inserted I-*Sce* I site, defines physical genomic intervals between the 'milestones' and an entire genome, a chromosome or any segment of interest can in principle be mapped with any chosen degree of resolution using artificially introduced I-*Sce* I sites. Subsequentely, new genes, DNA fragments or clones of interest can be directly located relative to the I-*Sce* I map by simple hybridizations or I-*Sce* I

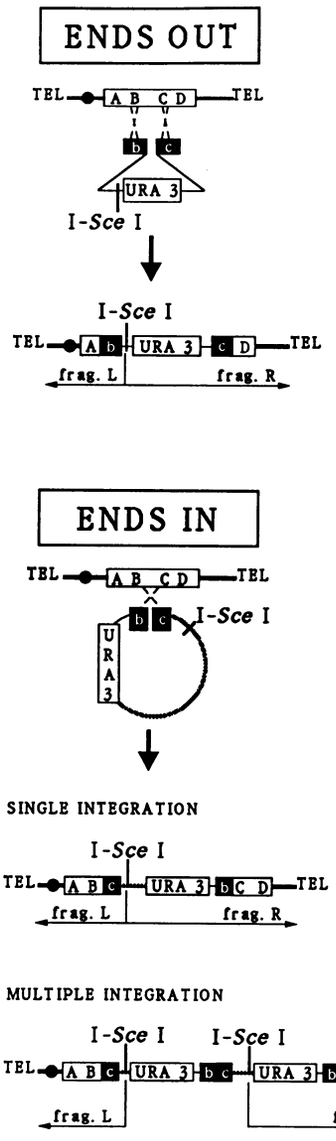

Figure 1. Insertion of I-*Sce* I sites into yeast chromosomes using a URA3 I-*Sce* I cassette. The URA3 I-*Sce* I cassette is targeted to appropriate sites of yeast chromosomes by homologous recombination after transformation of diploid yeast cells by *in vitro* linearized DNA. The first topology (ends out) results in the replacement of a given segment of yeast chromosome (the BC interval) by the linearized DNA containing the cassette. The second topology (ends in) results in the insertion of the linearized DNA into the homologous chromosome segment (between B and C). Insertion can be single or multiple resulting in tandem repeats. In all cases, digestion of yeast chromosome with I-*Sce* I generates two fragments (L and R), the length of which can be measured by PFGE and used to calculate distances between the target gene and the two telomeres. Note that the 'ends in' topology creates transgenic strains in which the vector sequence (a pUC19 derivative) is present in one of the two chromosome subfragments generated after I-*Sce* I digestion.

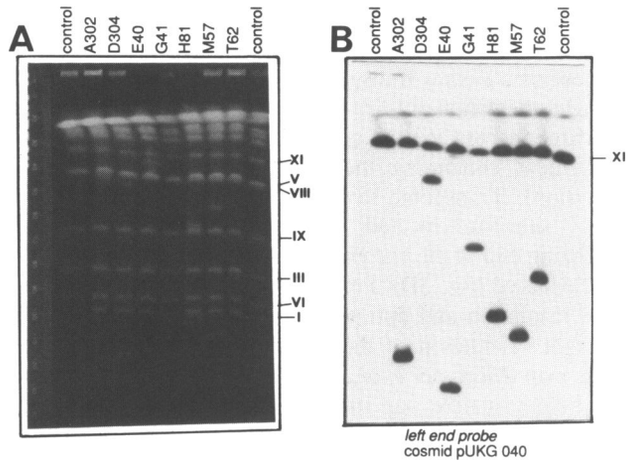

Figure 2. Set of seven transgenic yeast strains cleaved by I-*Sce* I. Chromosomes from FY1679 (control) and from seven transgenic yeast strains with I-*Sce* I sites inserted at various positions along chromosome XI were treated with I-*Sce* I under conditions described in Materials and Methods. DNA were electrophoresed on 1% agarose (SeaKem) gel in 0.25×TBE buffer at 130 V and 12°C on a Rotaphor apparatus (Biometra) for 70 hrs using 100 sec to 40 sec decreasing pulse times. (A) DNA was stained with ethidium bromide (0.2 μg/ml) and transferred to Hybond N (Amersham) membrane for hybridization. (B) ^{32}P labelled cosmid pUKG040 which hybridizes with the shortest fragment of the set was used as a probe. Positions of chromosome XI and shorter chromosomes are indicated.

sites can be used as origins from which conventional restriction mapping can be performed by partial digestion and hybridizations with each half of the I-*Sce* I cassette itself. In addition, yeast offers the decisive advantage of permitting purification of the chromosome fragments generated by I-*Sce* I digestion.

The strategy, applied here to yeast, involves the following steps: *i:* insertion of I-*Sce* I sites and selection of transgenic yeast strains; *ii:* digestion of chromosomes of transgenic yeast strains with I-*Sce* I and separation by PFGE; *iii:* selection of a probe that hybridizes with the shortest fragment of the set to orient all fragments and to derive the I-*Sce* I primary map. *iiii:* purification of chromosomal fragments from PFGE which are used, in turn, as probes to sort a genomic library.

Application of the nested chromosomal fragmentation to the mapping of yeast chromosome XI

We have applied the nested chromosomal fragmentation strategy to the chromosome XI of yeast as a first example to examine the feasibility and performance of the method. For this purpose, we have inserted the I-*Sce* I site at 7 different locations along chromosome XI of the diploid strain FY1679, hence defining eight physical intervals in that chromosome. Sites were inserted from a URA3-I-*Sce* I cassette by homologous recombination using either of the two topologies defined in figure 1. Two sites were inserted within genetically mapped genes, *TIF1* and *FAS1*, the others were inserted at unknown positions in the chromosome from five non-overlapping cosmids of our library, taken at random (see Materials and Methods). Agarose embedded DNA of each of the seven transgenic strains was then digested with I-*Sce* I and analyzed by PFGE (fig.2A). The position of the I-*Sce* I site of each transgenic strain in chromosome XI is first deduced from the fragment sizes without consideration of the left/right orientation of the fragments. Orientation is then determined as follows. The most telomere proximal I-*Sce* I site from this set of strains is in the transgenic E40 because the 50 kb fragment is the shortest of all fragments (figure 3A). Therefore, the cosmid clone pUKG040, which was used to insert the I-*Sce* I site in the transgenic E40 (see Materials and Methods), is now used as a probe against all chromosome fragments (Fig. 2B). As expected, pUKG040 lights up the two fragments from strain E40 (50 kb and 630 kb, respectively). The large fragment is close to the entire chromosome XI and shows a weak hybridization signal due to the fact that the insert of pUKG040, which is 38 kb long, contains less than 4 kb within the large chromosome fragment (data not shown). Note that the entire chromosome XI remains visible after

I-*Sce* I digestion, due to the fact that the transgenic strains are diploids in which the I-*Sce* I site is inserted in only one of the two homologs. Now, the pUKG040 probe hybridizes to only one fragment of all other transgenic strains allowing unambiguous left /right orientation of I-*Sce* I sites (fig. 3B). No significant cross hybridization between the cosmid vector and the chromosome subfragment containing the insertion vector is visible (see Discussion). Transgenic strains can now be ordered such that I-*Sce* I sites are located at increasing distances from the hybridizing end of the chromosome (fig 3C) and the I-*Sce* I map can be deduced (fig. 3D). Precision of the mapping depends upon PFGE resolution and optimal calibration. Note that the actual left /right orientation of the chromosome with respect to the genetic map (Mortimer *et al.*, 1989) is not determined at this step.

To help visualize our strategy and to obtain more precise measurements of the sizes of intervals between I-*Sce* I sites, a new PFGE with the same transgenic strains now placed in order has been made (fig. 4). After transfer, the fragments were hybridized successively with cosmids pUKG040 and pUKG066 which lights up, respectively, all fragments from the two opposite ends of the chromosome (clone pUKG066 determines the right end of the chromosome as defined from the genetic map because it contains the *SIR1* gene, data not shown). A regular stepwise progression of chromosome fragment sizes is observed. Note some cross hybridization between the probe pUKG066 and the chromosome III, probably due to some repetitive DNA sequences (see Discussion). All chromosome fragments, taken together, now define physical intervals as indicated in fig. 3D. The I-*Sce* I map obtained here has an 80 kb average resolution.

Rapid sorting of cosmid clones by successive hybridizations with the nested chromosomal fragments

At this step, chromosomal fragments can be purified from preparative PFGE and used as probes against clones from a chromosome XI specific sublibrary. This sublibrary is composed of 138 cosmid clones (corresponding to eight times coverage) which have been previously sorted from our complete yeast genomic libraries by colony hybridization with PFGE purified chromosome XI. This collection of unordered clones has first been sequencially hybridized with chromosome fragments taken in order of increasing sizes from the left end of the chromosome (data not shown). Localization of each cosmid clone on the I-*Sce* I map could be unambiguously determined from such hybridizations. To further verify the results and to provide a more precise map, a subset of all cosmid clones, now placed in order, has been digested with *Eco*RI, electrophoresed and hybridized with the nested series of chromosome fragments in order of increasing sizes from the left end of the chromosome. Figure 5 shows the results. For a given probe, two cases can be distinguished: cosmid clones in which all *Eco*RI fragments hybridize with the probe and cosmid clones in which only some of the *Eco*RI fragments hybridize (compare for example pEKG100 to pEKG098 in fig 5B). The first category corresponds to clones in which the insert is entirely included in one of the two chromosome fragments, the second to clones in which the insert overlaps an I-*Sce* I site. Note that, for clones of the pEKG series, the *Eco*RI fragment of 8 kb visible by ethidium bromide staining is entirely composed of vector sequences (pWE15) that do not hybridize with the chromosome fragments. In cases when the chromosome fragment possesses the integration vector (ends in topology, see figure 1), a weak cross hybridization with the cosmid is observed (fig. 5E).

Examination of fig. 5 shows that the cosmid clones can unambiguously be ordered with respect to the I-*Sce* I map (fig. 3E), each clone falling either in a defined interval or across an I-*Sce* I site. In addition, clones from the second category allow us to place some *Eco*RI fragments on the I-*Sce* I map, while others remain unordered.

The complete set of chromosome XI-specific cosmid clones, covering altogether eight times the equivalent of the chromosome, has been sorted with respect to the I-*Sce* I map as shown in fig. 6. Ninety one cosmids fall into intervals and 47 across given sites. The number of cosmids falling in the different intervals is in agreement with the estimated sizes of the intervals, indicating

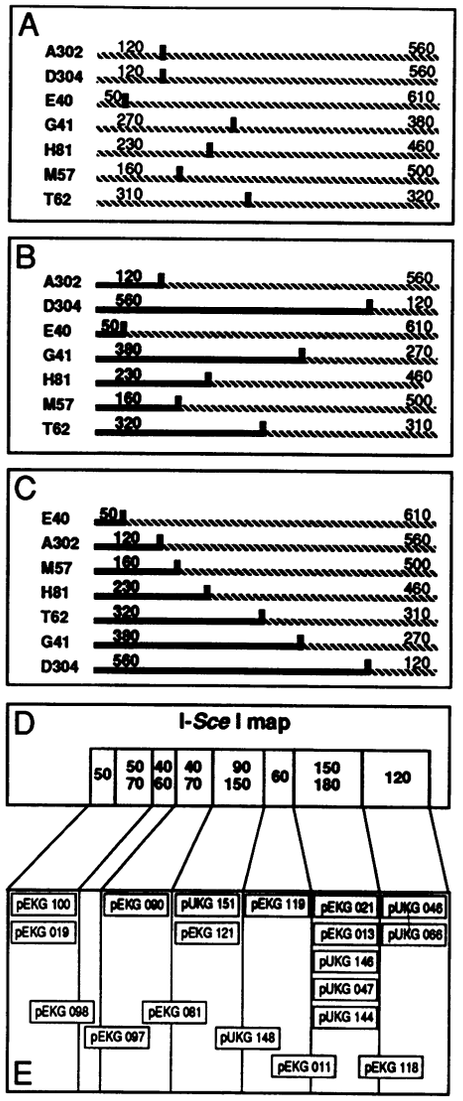

Figure 3. Rationale of the nested chromosomal fragmentation. (A) Positions of I-*Sce* I sites are placed on the map, irrespective of the left /right orientation (shorter fragments are arbitrarily placed on the left). Fragment sizes as measured from PFGE (fig. 2A) are indicated in kb (note that the sum of the two fragment sizes varies slightly due to the limit of precision of each measurement). (B) Hybridization with the probe that hybridizes the shortest fragment of the set determines the orientation of each fragment (see fig. 2B). Fragments that hybridize with the probe (full lines) have been placed arbitrarily to the left. (C) Transgenic yeast strains have been ordered with increasing sizes of hybridizing chromosome fragments. (D) Deduced I-*Sce* I map with minimal and maximal size of intervals indicated in kb (variations in some intervals are due to the limit of precision of PFGE measurements). (E) Chromosome subfragments are used as probes to assign each cosmid clone to a given map interval or across a given I-*Sce* I site (see fig. 6 for results).

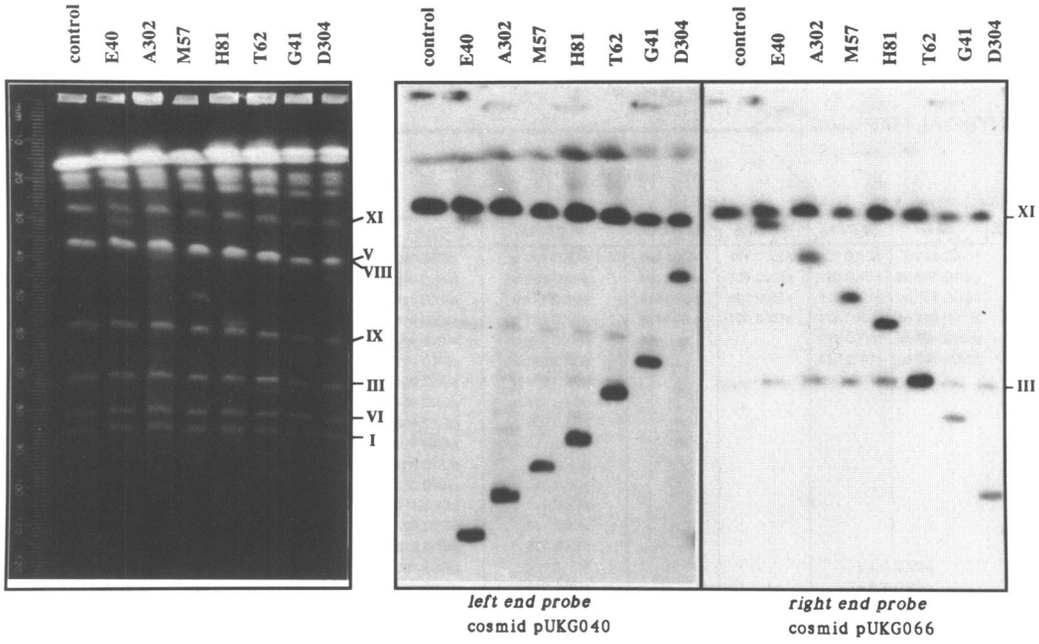

Figure 4. Mapping the I-*Sce* I sites of transgenic yeast strains by hybridization with left end and right end probes of chromosome XI. Chromosomes from FY1679 (control) and the seven transgenic yeast strains were digested with I-*Sce* I under conditions described in Materials and Methods. Transgenic strains were placed in order as explained in fig. 3. Electrophoresis conditions were as in figure 2. ^{32}P labelled cosmids pUKG040 and pUKG066 were used as left end and right end probes, respectively.

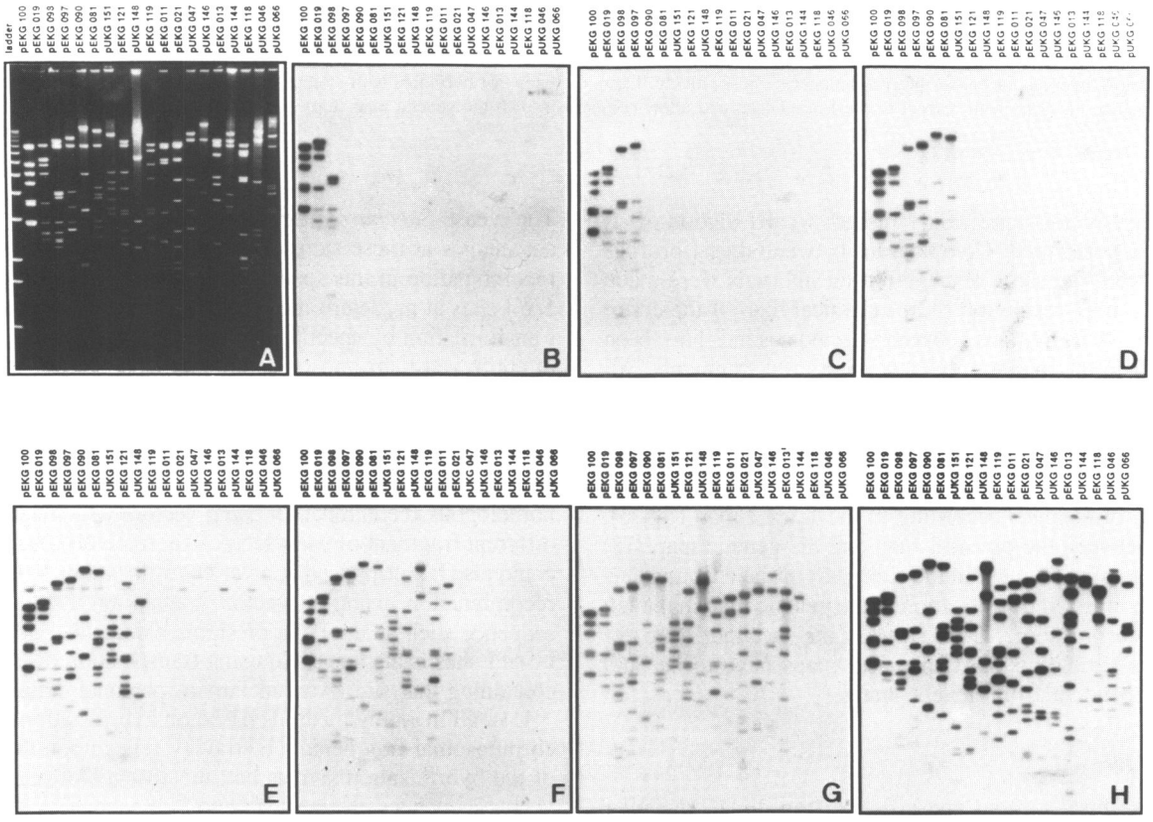

Figure 5. Mapping of a cosmid collection using the nested chromosomal fragments as probes. Cosmid DNAs were digested with EcoRI and electrophoresed on 0.9% agarose (SeaKem) gel at 1.5 V/cm for 14 hrs, stained with ethidium bromide and transferred to Hybond N membrane. Cosmids were placed in order from previous hybridization (data not shown) to help visualize the strategy. Hybridizations were carried out serially on three identical membranes using left end nested chromosome fragments purified on PFG (see fig. 4) as probes. **A**: ethidium bromide staining (ladder is the BRL '1kb ladder') , **B**: membrane #1, probe: Left tel to A302 site, **C**: membrane #1, probe: Left tel to M57 site, **D**: membrane #2, probe: Left tel to H81 site, **E**: membrane #2, probe: Left tel to T62 site, **F**: membrane #3, probe: Left tel to G41 site, **G**: membrane #3, probe: Left tel to D304 site, **H**: membrane #3, probe: entire chromosome XI.

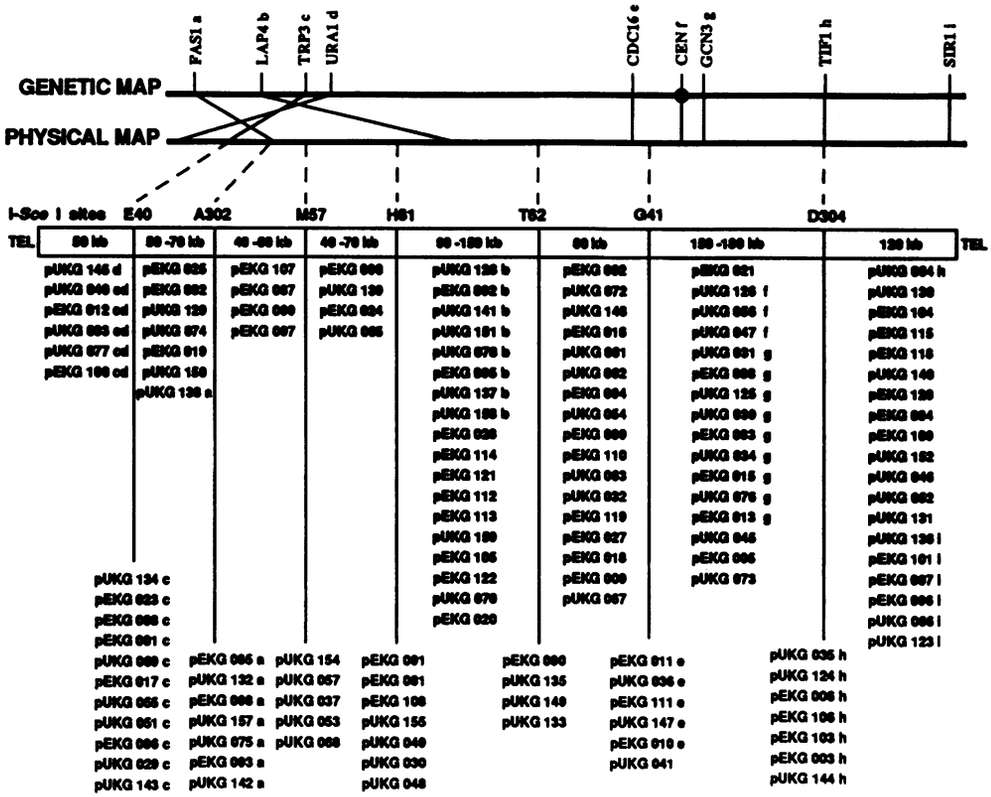

Figure 6. Map of the yeast chromosome XI as determined from the nested chromosomal fragmentation strategy. The chromosome is devided into eight intervals (with sizes indicated in kb, see fig. 3D) separated by seven I-*Sce* I sites (E40, A302 ...). Cosmid clones falling either within intervals or across a given I-*Sce* I site are listed below intervals or below interval boundaries, respectively. Cosmid clones that hybridize with selected genes used as probes are indicated by letters (a−i). They localize the gene with respect to the I-*Sce* I map and allow comparison with the genetic map (top).

that the library is not significantly biased. *Eco*RI digests of all cosmids were performed. Comparisons between digest profiles of cosmids from the same or neighboring intervals were made to construct a high resolution (3 kb) physical map of the entire chromosome. Hybridizations between selected cosmids have been used to confirm the map and resolve unordered fragments. All cosmids fall into a unique contig that covers the entire chromosome. The map will be published separately (Thierry *et al.*, in preparation). Hybridizations with a selected set of probes representing genetically mapped genes further confirm the classification of cosmids according to the I-*Sce* I map (fig. 6). Alignment between the physical map and the genetic map (18) reveals a general agreement in the right part of the chromosome but shows an inverted translocation in the genetic map in the left arm of the chromosome. This stresses the fact that even for organisms as simple as yeast, physical maps have to be built entirely independent from genetic maps.

DISCUSSION

We have developed a rapid and efficient 'top down' mapping strategy based on the application of a new endonuclease, I-*Sce* I, to fragment eucaryotic chromosomes at single artifical sites. The methods allows the direct sorting of genomic libraries. The success of the method depends upon three critical steps: the insertion of the artificial I-*Sce* I sites, the purification of the chromosome fragments and the reliability of hybridization signals. The yeast *Saccharomyces cerevisi* offers several significant advantages at these steps. First, the efficiency of homologous recombination in this species facilitates the integration of the I-*Sce* I sites at predetermined locations in the chromosomes after transformation by specific DNA cassettes. Second, purification of entire yeast chromosomes or fragments thereof can be easily achieved by PFGE, due to their size. Third, the very low abundance of repetitive sequences in the yeast genome facilitates the systematic use of hybridization as a generalized mapping tool.

In the present work, I-*Sce* I sites have been introduced by homologous recombination using vectors which each contain a different fragment of yeast DNA. Alternatively, the I-*Sce* I sites could also be introduced at a variety of locations by homologous recombination using a vector containing a repeated DNA sequence such as the delta or sigma elements. But insertion of I-*Sce* I sites is also possible using transposition of a Ty element containing that site (Arnaud Perrin, personal communication).

Like all mapping methods based on hybridizations, the nested chromosomal fragmentation strategy relies upon the sensitivity of the hybridization signals and the absence of interference by vector sequences or repeated genomic sequences. In the present work, vector sequences (pUC19) were present in one of the two yeast chromosome subfragments used as probe when the I-*Sce* I site was introduced by the 'ends in' topology but cross-hybridization with the cosmid vector used (pOU61cos and pWE15) was sufficiently low not to interfere significantly with hybridization signals generated by yeast sequences. In other

instances, it may be necessary to compete out vector sequences. Similarly, false positive due to repeated DNA sequences are not a significant problem in yeast due to the rare occurrence of such sequences (yet an example of this can be observed in figure 4 but the signal to noise ratio is adequate for correct interpretation of the results). Application of our method to YAC mapping or to higher organisms with frequent occurrence of repeated sequences will necessitate competition with unlabelled DNA to eliminate false positive.

The efficiency of our method to sort genomic clone libraries depends upon the optimal distribution of I-*Sce* I sites. The average resolution of the I-*Sce* I map must be adapted to the average insert size of the library used. In the present work, we constructed an 80 kb average resolution I-*Sce* I map to sort a cosmid library with an average insert size of ca. 35 kb. The proportion of cosmid clones expected to overlap an I-*Sce* I site is, therefore, 43% (35/80), in agreement with the actual figure observed (47 out of 138 clones tested in total) considering that the intervals were not of equivalent sizes. The efficiency of sorting the genomic clones should be maximum when half of them overlap the I-*Sce* I sites and half fall within intervals. It follows that one should introduce, on the average, one I-*Sce* I site every 70 kb to sort a cosmid library of 35 kb inserts, every 40 kb on average for a λ library and one I-*Sce* I site every 200 kb for P1 library. To facilitate the european yeast sequencing program (19, 20), we are now in the process of inserting enough I-*Sce* I sites at random to cover the entire yeast genome with an average resolution of ca. 80 kb. At present, we have several sites in each of the sixteen yeast chromosomes and are, therefore, already in a position to fragment them as needed. Physical mapping of yeast chromosomes chosen for the next phases of the sequencing program will be done using the present strategy since, as we have experienced by the present work, a physical map must be constructed without consideration of the genetic map.

Compared to more conventional mapping methods such as fingerprinting or hybridization of clones using oligonucleotide sequences or end specific probes, our method is at least as efficient or even more efficient, contrary to intuitive expectation. For exemple, a set of transgenic yeast strain can be constructed in less than four weeks with only part time occupation. Sorting clone libraries by successive hybridizations with the nested fragments is then a very quick step and is amenable to automation. But in addition, our method present the very interesting advantage to permit direct measurement of the physical distance of a given site to the ends of the chromosome or to another site irrespective of the availability of a complete contig of clones covering that segment. In addition, our method automatically provides a means to examine internal consistency of the physical map irrespective of its degree of resolution. The only limitation of our method is probably that it requires more skills from the experimentators as construction and analysis of transgenic yeasts is less prone to automation than purely molecular methods.

The chromosome fragmentation strategy should be directly applicable to other eucaryotic organisms with genome sizes comparable to that of yeast. Bacterial genomes are also appropriate as integration of a few I-*Sce* I sites should be sufficient to allow purification of the chromosome subfragments by PFGE. In other organisms with larger genomes, purification of chromosome fragments is not possible but I-*Sce* I sites could also be introduced at random by transformation using a variety of vectors in which that site is inserted. Insertion of several I-*Sce* I sites at appropriate distances could be envisaged to generate chromosome subfragments of sizes compatible with PFG resolution. In all cases, it is possible to map the segments flanking a given I-*Sce* I site by cleavage by I-*Sce* I followed by partial secondary digestions with bacterial endonucleases. We have applied this procedure for partial *Bam*HI mapping on yeast chromosome XI.

But the most direct application of the nested chromosomal fragmentation is certainly the physical mapping of YACs which contain long inserts of DNA from higher eucaryotes. Preliminary experiments indicate that I-*Sce* I sites can be specifically introduced along the YAC insert by transformation of yeast with an integrative plasmid containing a repetitive mammalian sequence. Advantages of the present method compared with the generation of deletion derivatives by targeted transformation using fragmentation plasmids (21) are that the average resolution of the I-*Sce* I map depends solely on the number of transgenic strains available in the set and that the two fragments of each chromosome are simultaneously available from each individual transgenic strain.

ACKNOWLEDGMENTS

We thank Fred Winston and Guy-Franck Richard for the generous gift of strains FY23 and FY73 and of the transgenic yeast FY1679/H81, respectively. We also thank our colleagues from the Unité de Génétique Moléculaire des Levures for stimulating discussions and critical reading of the manuscript. B.D. is Professor of Molecular Genetics at University Pierre and Marie Curie. This work was supported in part by the BRIDGE and the Human Genome Analysis programs of the Commision of the European Communities and by grants from University Pierre and Marie Curie.

REFERENCES

1. Coulson, A., Sulston, J., Brenner, S. and Karn, J. (1986) Proc. Natl. Acad. Sci. 83, 7821–7825
2. Kohara Y., Akiyama, K., Isono, K. (1987) Cell 50, 495–508
3. Olson, M.V., Dutchik, J.E., Graham, M.Y., Brodeur, G. M., Helms, C., Frank M., MacCollin, M., Scheinman, R. and Frank T. (1986) Proc. Natl. Acd. Sci. 83, 7826–7830
4. Lander, E.S. and Waterman, M.S. (1988) Genomics 2, 231–239
5. Poutska, A., Pohl, T., Barlow, D.P. Zehetner, G., Craig, A., Michiels, F., Ehrich, E., Frischauf, A.M. and Lehrach H. (1986) Cold Spring Harbor Symposium on Quantitative Biology 51, 131–139
6. Evans, G.A., Lewis, K., Rothenberg, B.E. (1989) Gene 79, 9–20
7. Evans, G.A. and Lewis, K.A. (1989) Proc. Natl. Acad. Sci. 86, 5030–5034
8. Burke, D. T., Carle, F. G. and Olson, M.M. (1987) Science 236, 806–812
9. Koob, M. and Szybalski, W. (1990) Science 250, 271–273
10. Ferrin, L.J. and Camerini-Otero, R.D. (1991) Science 254, 1494–1497
11. Colleaux, L., d'Auriol, L., Betermier, M., Cottarel, G., Jacquier, A., Galibert, F. and Dujon, B. (1986) Cell 44, 521–533
12. Colleaux, L., d'Auriol, L., Galibert, F. and Dujon, B. (1988) Proc. Natl. Acad. Sci. 85, 6022–6026
13. Monteilhet, C., Perrin, A., Thierry, A., Colleaux, L. and Dujon, B. (1990) Nuc. Ac. Res. 18, 1407–1413
14. Thierry, A., Perrin, A., Boyer, J., Fairhead, C. and Dujon, B. (1991) Nuc. Ac. Res. 19 189–190
15. Thierry, A., Fairhead, C. and Dujon, B. (1990) Yeast 6, 521–534
16. Colleaux, L., Richard, G. F., Thierry, A. and Dujon, B. (1992) Yeast 8, 325–336
17. Church, G. M. and Gilbert, W. (1984) Proc. Natl. Acd. Sc. 81, 1991–1995
18. Mortimer, R.K., Shild, D, Contopoulou, C.R. and Kans, J.A. (1989) Yeast 5, 321–404
19. Vassarotti, A. and Goffeau, A. (1992) Trends in Biotechnology 10, 15–18
20. Dujon, B. (1992) Current Biology 2, 279–281
21. Pavan, W.J., Hieter, P. and Reeves, R.H. (1990) Proc. Natl. Acad. Sc. 87, 1300–1304