YEAST VOL. 11: 659–666 (1995)

High-Resolution Cosmid Mapping of the Left Arm of *Saccharomyces cerevisiae* Chromosome XII; A First Step Towards An Ordered Sequencing Approach

PATRIK SCHOLLER, SANDRA SCHWARZ AND JÖRG D. HOHEISEL*

Molecular-Genetic Genome Analysis, German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

Received 25 November 1994; accepted 20 February 1995

For the sequencing of the left arm of chromosome XII of *Saccharomyces cerevisiae*, we fine-mapped the entire 450 kb fragment between the ribosomal DNA (rDNA) and the left telomere. Total yeast DNA in agarose blocks was digested with I-*Ppo*I, which exclusively cuts once in each repeat unit of the rDNA. The resulting fragment was isolated from pulsed-field gels, together with the equally sized chromosome IX. A cosmid library of some 30-fold chromosome coverage was generated from this material, with the cloning efficiency being around 20 000 clones per microgram genomic DNA. The chromosome XII and IX specific clones were identified by complementary hybridizations with the respective chromosomes. For the left arm of chromosome XII, a contiguous cosmid array (contig) with an average map resolution better than 9 kb was generated by clone hybridization procedures. The ordered library serves as a tool for the physical mapping of genetic markers. Also, a minimal set of 15 clones was selected that covers the entire fragment. This subset forms the basis for the generation of a template map of much higher resolution for a directed sequencing of the left arm of chromosome XII.

KEY WORDS — hybridization mapping; cosmids; genome analysis; chromosome XII

INTRODUCTION

The sequencing of the genome of *Saccharomyces cerevisiae* is progressing at a rate that will make the budding yeast, if not the first completely sequenced organism, then at least the first eukaryotic organism whose genetic information will be known entirely. The results from the analyses of four chromosomes have already been published (Oliver *et al.*, 1992; Dujon *et al.*, 1994; Johnston *et al.*, 1994; Feldmann *et al.*, 1994), and all other chromosomes are being worked at in an internationally concerted effort, with some of them near completion. The 450 kb of the left arm of chromosome XII is the last fragment to be sequenced within the framework of the European Union Yeast Sequencing Project.

The European approach so far has been based on cosmid maps whose clones have been distributed to a network of sequencing groups (Goffeau and Vassarotti, 1991). Each group has tackled the actual sequencing by the means and methods of its own choice. Besides stimulating the evolution and development of techniques, this policy has allowed the assessment and comparison of the different sequencing strategies adopted. In the chromosome XII project, simultaneously to the mere deciphering of the DNA sequence, the comparison between strategies is extended to the mapping stage, thus linking further the mapping and sequencing levels of genome analysis. The intention is to achieve significant reductions in the amount of cost- and time-extensive sequencing work by putting a comparatively small extra effort into the map generation. As a first step toward this end, we report here the generation of a contiguous array of cosmid clones using highly efficient hybridization procedures. The resulting high-resolution map can be utilized for the physical localization of genetic markers. Also, a minimal clone set was determined from the hybridization data, both to provide the

*Corresponding author.
This manuscript is dedicated to Fritz M. Pohl, a remarkable character and great teacher of molecular biology.

CCC 0749–503X/95/070659–08
© 1995 by John Wiley & Sons Ltd

standard map format of the European Union Yeast Sequencing Project and as a basic framework for the ordered sequencing approach.

MATERIALS AND METHODS

Isolation of the left arm of chromosome XII

Intact yeast chromosomes from strain FY23/RD005 (a gift from B. Dujon) were prepared in agarose plugs as described by Tettelin *et al.* (1994). For digestion with I-*Ppo*I, the plugs were washed in 0·1 M-diethanolamine/HCl, pH 9.5, at 4°C for 1 h and equilibrated overnight under the same conditions. Sixteen agarose blocks, each of about 70 µl, were incubated in a final volume of 2·5 ml of 0·1 M-diethanolamine/HCl, pH 9·5, 2 mM-dithiothreitol, 2 mM-spermidine and 0·2 mg/ml bovine serum albumin (I-*Sce*I incubation buffer, Boehringer Mannheim). After 1 h on ice, 20 µl 1 M-$MgCl_2$ and 40 µl I-*Ppo*I (105 units/µl; Promega) were added. The sample was kept on ice for 2 h for equilibration, followed by a 3-h incubation at 37°C. The reaction was terminated by the addition of 10 ml 0·5 M-EDTA, pH 8·0.

Pulsed-field gel electrophoresis was carried out on a BioRad DRII apparatus in 1% SeaPlaque GTG low-melting-point agarose (FMC) using standard conditions for the separation of yeast chromosomes (run time, 48 h; voltage, 6 V/cm; time ramp, 50–120 s, 0·5 × TBE buffer at 14°C). The gel was stained with ethidium bromide and destained in water. The appropriate DNA band was cut out from the gel under long-wave UV light (366 nm). The DNA was carefully released from the gel by an agarase-treatment (Boehringer Mannheim) as advised by the manufacturer, with additional extractions with phenol and chloroform/isoamylalcohol (24 : 1) before the ethanol precipitation.

Generation of the cosmid library

About 2 µg of chromosomal DNA were partially digested with 1 unit of *Mbo*I (New England Biolabs) at 37°C in 100 µl 33 mM-Tris-acetate, pH 7·9, 65 mM-potassium acetate, 10 mM-magnesium acetate, 0·5 mM-dithiothreitol (TAK-buffer). A third of the sample was removed from the reaction after 1, 2 and 3 min, respectively, and immediately mixed with 1 volume of 50 mM-EDTA, followed by a heat inactivation of the enzyme. The samples were pooled again and purified by subsequent extractions with phenol and chloroform/isoamylalcohol (24 : 1). The DNA was ethanol-precipitated and taken up in 10 µl TAK-buffer containing 1 unit of calf intestine alkaline phosphatase (Boehringer Mannheim). Dephosphorylation was for 30 min at 37°C. The reaction was terminated by a 10-min incubation at 75°C, after which 1 µl of 0·5 M-EDTA, 1 µl 20% SDS and 2·5 µl of a proteinase K solution (2 mg/ml) were added. After digestion at 56°C for 30 min, protein was removed by phenol and chloroform/isoamylalcohol extractions. The DNA was precipitated with ethanol and taken up in 10 µl water; 2 µl were run on a 0·5% agarose gel to check the degree of digestion.

Vector arms of the cosmid vector Lawrist7 (de Jong *et al.*, 1989) were prepared as detailed earlier (Hoheisel *et al.*, 1991) by a linearization with *Sca*I, dephosphorylation of the ends and subsequent cleavage of the *Bam*HI cloning site. The remaining 8 µl partially cut chromosomal DNA (an estimated 1 µg) were ligated to 2 µg of Lawrist7 vector arms in 15 µl of 30 mM-Tris–HCl, pH 7·5, 6·6 mM-$MgCl_2$, 1 mM-dithiothreitol, 1 mM-spermidine, 1 mM-ATP and 0·1 unit/µl T4 DNA ligase (Amersham) at 4°C for 72 h. 5 µl ligation mix were packaged into Lambda phage particles using Gigapack II Gold extracts (Stratagene). Transfection into *Escherichia coli* DH5α was carried out as recommended by the manufacturer; the best efficiency (total clone number per packaging reaction) was achieved using a 1 : 10 dilution of the phage solution. Individual cosmid clones were picked into 384-well microtitre dishes and stored frozen at minus 70°C as described (Hoheisel *et al.*, 1991). For labelling purposes or restriction analysis, cosmid DNA was isolated following a modified alkaline lysis protocol (Pohl *et al.*, 1982).

In situ *filters*

Using 384-pin replicators (Genetix, U.K.), bacterial material was transferred from the microtitre dishes onto Hybond N^+ membranes (Amersham) of 22 × 22 cm, which had been soaked in 2YT growth medium. For cell growth, the filters were placed onto agar plates containing 30 µg/ml kanamycin and incubated for about 16 h at 37°C. The DNA was attached to the filter *in situ* (Hoheisel *et al.*, 1991). The protocol consists of two steps of denaturation, followed by neutralization. Each filter was digested for 30 min in 500 ml of 50 mM-Tris–HCl, pH 8·5, 50 mM-ERDTA, 100 mM-NaCl, 0·9% (w/v) Na-sarkosyl (BDH/Merck) and 0·3 mg/ml pronase E, the latter replacing

proteinase K of the original protocol; there was no difference in performance compared to filters treated with proteinase K. The membranes were dried overnight blotted on 3MM Whatman paper and the DNA was crosslinked by UV-irradiation according to the manufacturer's instructions.

PCR-amplification of genetic markers

Chromosome XII sequences from the EMBL sequence database were searched for appropriate primer sequences, which had to be about 1 kb apart and of compatible melting temperatures, using the HUSAR program package (Heidelberg Unix Sequence Analysis Resources, Bioinformatics Department, DKFZ, Heidelberg, Germany). The DNA fragments were PCR-amplified from 1 ng DNA of gel-purified chromosome XII at standard conditions in an Autogene II cycler (Grant).

Hybridization procedure

For each experiment, about 50 ng DNA (chromosomes; genetic markers; cosmids) were ^{32}P-labelled by random hexamer priming (Feinberg and Vogelstein, 1983). Prior to hybridization, probe made from cosmids was competed with 125 ng/μl Lawrist7 vector-DNA (sonicated at 18 microns for 10 times 10 s) and 570 ng/μl yeast tRNA (Sigma) in 175 μl water by incubating at 95°C for 5 min, followed by the addition of 25 μl 1 M-Na-phosphate, pH 7·2, and an incubation at 65°C for 2 h. The library filters were pre-hybridized in 0·5 M-Na-phosphate, pH 7·2, 7% SDS, 1 mM-EDTA, and 0·1 mg/ml yeast tRNA at 65°C for at least 1 h. Hybridization was performed in the same buffer at 65°C overnight, the probe concentration being about 0·5 Mcpm/ml. Subsequently, the filters were briefly rinsed at room temperature in 40 mM-Na-phosphate, pH 7·2, 0·1% SDS. The same buffer at room temperature was then added and the filters were washed, rocking slowly in a waterbath of 65°C for about 20 min. The filters were briefly blotted on Whatman chromatography paper to remove excess liquid, before being exposed to film for 2 h to overnight at −70°C using intensifying screens. To strip the probe off the filters, about 1 litre of 5 mM-Na-phosphate, pH 7·2, 0·1% SDS at room temperature was poured into a box containing up to 20 filters and kept in a waterbath of about 90°C for 30 min, after which the procedure was repeated. The filters were then immediately transferred to pre-hybridization conditions and stored between the hybridizations at 65°C, before being used again. Only over long periods, filters were stored dry after stripping. The hybridization results were read manually and analysed on a SUN SPARC-II workstation using purpose-written software (Mott *et al.*, 1993).

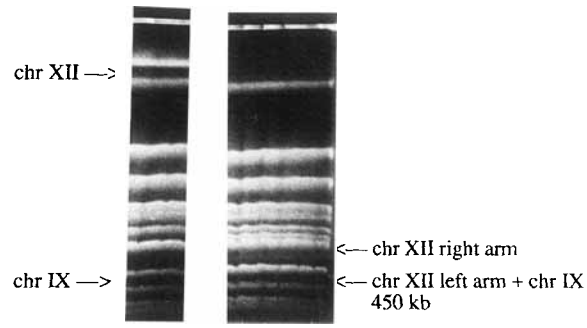

Figure 1. Pulsed-field electrophoresis of I-*Ppo*I-digested yeast chromosomes. On the left, the karyotype of yeast strain TY23/RD005 is shown. Upon digestion with I-*Ppo*I, chromosome XII is cut into its two arm fragments, marked on the right. The multiple copies of the 9·5 kb rDNA were eluted from the gel during electrophoresis.

RESULTS

Library generation

For the fragmentation of chromosome XII, total yeast DNA in agarose plugs was digested with I-*Ppo*I. The enzyme cleaves specifically once in each of the approximately 100 repeat units of the rDNA, which are clustered in a 1 Mbp region in the centre of chromosome XII. A digest therefore releases the two unique chromosomal regions flanking the rDNA and many copies of the 9·5 kb repeat itself. The left chromosomal arm is identical in size to chromosome IX and can thus not be separated by electrophoresis (Figure 1). With the aim of obtaining a fragment-specific cosmid library, we had initially tried to separate, by a brief gel electrophoresis, the two by far largest yeast chromosomes—IV and XII—from the rest, before performing the I-*Ppo*I digestion on the agarose slice containing these two chromosomes and a subsequent second size selection. The DNA-yield from such a procedure, however, was extremely low, so that the cloning would have required many (enzyme-)costly and time-consuming preparations. Hence, I-*Ppo*I digests done on all yeast chromosome were run on pulsed-field gels, and the 450 kb band containing both the left arm of chromosome XII and the whole of chromosome IX was cut from the gel.

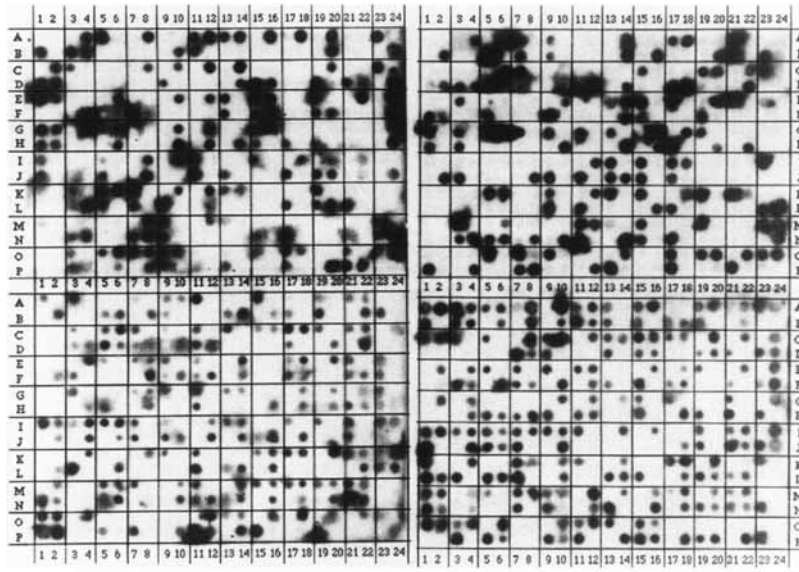

Figure 2. Selection of chromosome XII specific cosmids. Replicas of the filter pair which contains the 768 randomly picked clones of the cosmid library (384 per filter) were hybridized with probe made from gel-purified chromosomes IX (top row) or XII (bottom row). The high degree of complementarity between the respective signals is evident.

After isolation, the DNA was partially cut with *Mbo*I and ligated into the cosmid vector Lawrist7. Transfected into *E. coli* strain DH5α, the overall number of clones would have been about 20 000, if all of the ligated material worth an estimated microgram of chromosomal DNA had been used. This number is in agreement with the result of an analogous experiment done on chromosome IV (unpublished data), in which an even higher efficiency of about 30 000 cosmids per microgram of chromosomal DNA was achieved. Only 768 individual cosmids were actually picked into two 384-well microtitre dishes. With an average insert size of 37 kb, they represent a 30-fold coverage of both chromosome IX and the left arm of chromosome XII. For the mapping analysis, replica nylon filters were generated from the library, with the cosmid DNA bound *in situ*.

To distinguish chromosome XII from chromosome IX clones, probes made from the respective, gel-purified chromosome were hybridized to the filters (Figure 2). Most clones (88·9%) were positive in only one experiment, producing no signal in the other (407 with chromosome XII; 276 with chromosome IX), while only a minority was positive in both (11; 1·4%) or neither (74; 9·7%) hybridization. Of those ambiguous cosmids, eight of the former group were positively assigned to the left arm of chromosome XII during the analysis; six of them overlap and map near the telomeric end. Only six clones of the second group were found in the eventual contig.

Hybridization mapping

Clones from the library were chosen at random and probe-hybridized back to the library, thus identifying cosmids with overlapping inserts (e.g. Figure 3). The signal intensity is a rough indicator of the scale of an overlap, although other factors like the variation in the copy number of the cosmids have an effect, also. For the actual analysis, however, the strength of the hybridization signals was not taken into account but scored on a yes-or-no basis. To minimize the number of hybridizations, new probes were selected from clones that had not been positive in one of the earlier experiments (Hoheisel *et al.*, 1993). The data were analysed using the program package developed by Mott *et al.* (1993), which orders the probes rather than the clones. Since there are many fewer probes than library clones, ordering the former is a more efficient procedure. Moreover, a comparison of two probes is based on the hybridization frequencies on a relatively large number

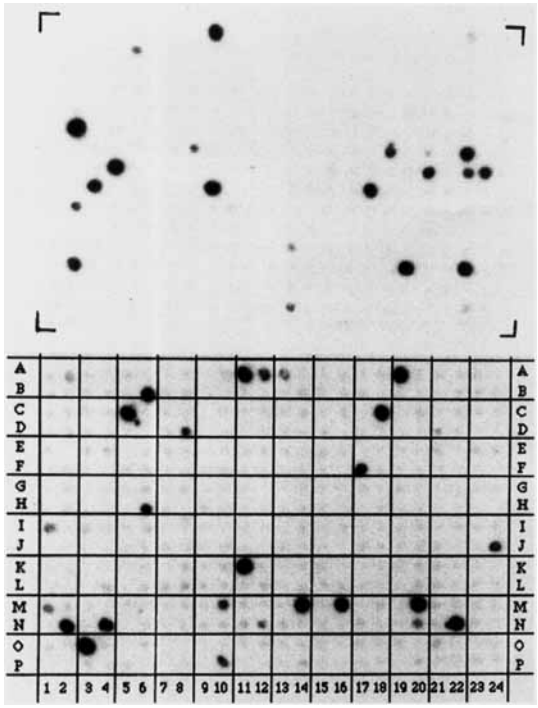

Figure 3. Mapping by clone hybridization. Cosmid 2C5 was radioactively labelled and hybridized to the library filters. Positive clones were detected by autoradiography. Along with the original clone, overlapping cosmids were identified by this procedure.

of clones. It has therefore a higher information content and, hence, higher reliability than a clone–clone comparison, which relies on information from a small number of probes. As a consequence, the map is actually defined as a sequence of probes to which the clones are fitted.

Following this strategy, only 19 cosmid hybridizations were sufficient to construct a single, continuous, and largely correct, contig of the area. Eventually, 52 probes were used (Figure 4), mainly in order to define a good minimal set of clones covering the fragment. This group of 15 cosmids is not a real minimal set in terms of the merely smallest possible number of cosmids, though. Since they form the basis for all subsequent analyses, their localization was affirmed by the extra requirement that each cosmid had to identify both its neighbouring clones in the subset, when used as a hybridization probe. For the purpose of illustration rather than verification, the 15 cosmids and some interleaving clones were digested with *Eco*RI (e.g., Figure 5).

Mapping of markers

The complete map as given in Figure 4 has an average resolution of 8·6 kb. Its extent was determined by probes known to be located at the extremities of the examined genomic region. On the centromeric side, the 9·5 kb rDNA repeat overlaps with some end clones containing part of its sequence. The presence of the telomeric end was proved with the probe pEL30 (Louis and Haber, 1992). It contains the ends of the single Y′ element found at the left terminus of chromosome XII (E. Louis, personal communication).

In addition to the corroboration of the map's extent, using the above marker hybridizations, the usefulness of this technique for the physical localization of internal genetic markers was demonstrated with five such sequences. In comparison to the genetic map (Mortimer *et al.*, 1992), their order was unchanged except for the position of *trx1*. Rather than falling between *spa2* and *rad5*, it was mapped between *rad5* and *put1*, next to *rad5* (Figure 4). We repeatedly tried to map a 1·1 kb fragment which was amplified from gel-enriched chromosome XII DNA with primers specific for *gal2*, which is supposed to be located between *rad5* and *put1*, but could not observe any binding of the probe to the cosmid filters. Hybridizations to Southern blots of yeast chromosomes indicated that the probe is located elsewhere in the genome (results not shown). These data are confirmed by a significant variation in the amount of *gal2* PCR-product depending on the starting material: PCR on purified chromosome XII yielded much less product than a reaction with an equimolar amount of total yeast DNA, while the PCR-amplification of the other markers showed no such effect. This suggests that the amplification with gel-purified chromosome XII occurred due to contaminating material from other chromosomes.

DISCUSSION

Ordered cosmid or bacteriophage P1 libraries form the clonal backbone of virtually every current sequencing project. Sorting such clones by means of hybridization has been shown, in this and other projects, to be an efficient way of providing this basic resource. While hybridization mapping with unique probes on the cosmid level does not seem practicable for very large genomes on account of the vast number of experiments necessary to accumulate the required amount of information, the

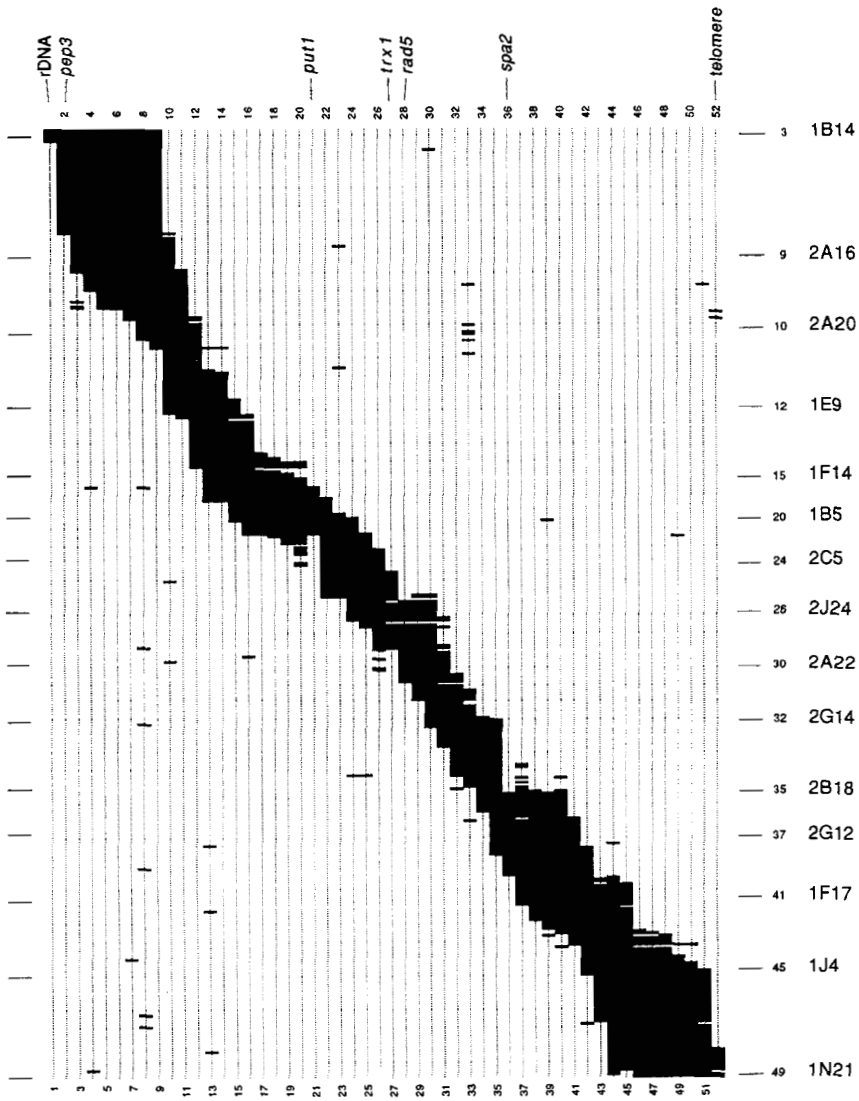

Figure 4. Cosmid map of the left arm of chromosome XII. The results of the hybridization experiments are shown in a two-dimensional matrix. The probes correspond to columns (vertical lines 1 to 52) and the clones to rows. Every hybridization event is recorded by a black bar. All probes that are not explicitly named at the top were cosmids taken from the library. The probe name *telomere* stands for plasmid pEL30 (see text). The names of the 421 cosmid clones are not given on the margin, except for the minimal subset, whose position in the map is indicated. Since all clones of the minimal set were also used as hybridization probes, their probe numbers are also shown (e.g., no. 3 for clone 1B14). Upon request, the authors can provide a much more detailed version of the map.

relative simplicity in application, data recording and analysis makes unique probes a favourable instrument for mapping projects of up to several million base pairs. This is true especially for regions that do not contain many repetitive elements, as is the case for the whole of the yeast genome, since probe isolation is fairly effortless. The only apparent repeat sequence encountered during the mapping of the left arm of chromosome XII was a sub-telomeric sequence, causing slight cross-hybridization between chromosomes XII and IX (and other chromosomes; not shown). The

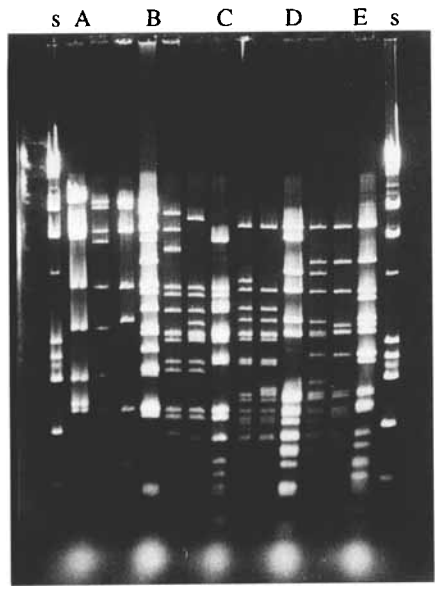

Figure 5. *Eco*RI digests of overlapping cosmids. Clones from the minimal cosmid set (A: 1F14; B: 1E9; C: 2A20; D: 2A16; E: 1B14) and a couple each of interleaving clones from the library were digested. As was expected from the hybridization data, common fragments can be seen between neighbouring cosmids. In the marginal lanes, a mixture of Lambda-DNA cut with *Hin*dIII and pUC18 digested with *Fsp*I was run as a size standard (s).

ambiguities could be resolved readily, however, when the results of the chromosome hybridizations were combined with the single-clone hybridization data. The congruence between the physical and genetic map was easily checked, without a change of methodology.

For sequencing, cosmids are usually sub-cloned as small, random fragments for a shotgun strategy, or primer walking techniques are employed in an ordered approach. We intend to combine the advantages of both these methods (Hoheisel, 1994). Based on the minimal cosmid set that was selected as a first step in the work reported here, we are now in the process of ordering random sub-libraries made from each cosmid. In principle, the applied technique is identical to the cosmid mapping but with the difference that short oligonucleotides (octamers to dodecamers) are being used, in addition to clone pool hybridizations. Most oligonucleotides should occur uniquely in each sub-library, despite their shortness, but have the advantage of an overall higher output of information, since they should hybridize to more than just one such library. For sequencing, both the redundant sampling of anonymous templates and the necessity of synthesizing many sequencing primers will be replaced by an ordered sequencing with a single primer system.

ACKNOWLEDGEMENTS

The generous help of Agnes Thierry and Bernard Dujon in obtaining the chromosomal fragment is gratefully acknowledged. We also thank Edward Louis for providing telomeric probes and related, unpublished information, Jens Hanke for his help with the PCR and Richard Mott for his support on the software. This work was funded by a contract with the European Commission for the European Yeast Genome Sequencing Project.

REFERENCES

de Jong, P. J., Chen, C. and Garnes, J. (1989). Application of PCR for the construction of vectors and the isolation of probes. In Erlich, H. A., Gibbs, R. A. and Kazazian, H. H. (Eds), *Polymerase Chain Reaction*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 205–210.

Dujon, B, *et al.* (1994). Complete DNA sequence of yeast chromosome XI. *Nature* **369,** 371–378.

Feldmann, H, *et al.* (1994). Complete DNA sequence of yeast chromosome II. *EMBO J.* **13,** 5795–5809.

Goffeau, A. and Vasarotti, A. (1991). The European project for sequencing the yeast chromosome. *Res. Microbiol.* **142,** 901–903.

Hoheisel, J. D. (1994). Application of hybridisation techniques to genome mapping and sequencing. *Trends Genet.* **10,** 79–83.

Hoheisel, J. D., Lennon, G. G., Zehetner, G. and Lehrach, H. (1991). Use of high coverage reference libraries of *Drosophila melanogaster* for relational data analysis; a step towards mapping and sequencing of the genome. *J. Mol. Biol.* **220,** 903–914.

Hoheisel, J. D., Maier, E., Mott, R., *et al.* (1993). High-resolution cosmid and P1 maps spanning the 14-Mbp genome of the fission yeast *Schizosaccharomyces pombe. Cell* **73,** 109–120.

Johnston, M. *et al.* (1994). Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome VIII. *Science* **265,** 2077–2082.

Louis, E. J. and Haber, J. E. (1992). The structure and evolution of subtelomeric Y′ repeats in *Saccharomyces cerevisiae. Genetics* **131,** 559–574.

Mortimer, R. K., Contopoulou, R. and King, J. S. (1992). Genetic and physical maps of *Saccharomyces cerevisiae.*, edition 11. *Yeast* **8,** 817–902.

Mott, R., Grigoriev, A., Maier, E., Hoheisel, J. D. and Lehrach, H. (1993). Algorithms and software tools from ordering clone libraries; application to the mapping of the genome of *Schizosaccharomyces pombe. Nucl. Acids Res.* **21,** 1965–1974.

Oliver, S. G., *et al.* (1992). The complete DNA sequence of yeast chromosome III. *Nature* **357,** 38–46.

Pohl, F. M., Thomae, R. and Karst, A. (1982). Temperature dependence of the activity of DNA-modifying enzymes: endonucleases and DNA ligase. *Eur. J. Biochem.* **123,** 141–152.

Tettelin, H., Thiery, A., Fairhead, C., Perring, A. and Dujon, B. (1994). *In vitro* fragmentation of yeast chromosomes and YACs at artificially inserted sites and applications to genome mapping. *Meth. Mol. Genet.*, in press.