The EMBO Journal vol.5 no.5 pp.1011-1021, 1986

Chromosome walking shows a highly homologous repetitive sequence present in all the centromere regions of fission yeast

Yukinobu Nakaseko, Yasuhisa Adachi, Shin-ichi Funahashi, Osami Niwa and Mitsuhiro Yanagida

Department of Biophysics, Faculty of Science, Kyoto University, Sakyo-ku, Kyoto 606, Japan

Communicated by P.Philippsen

By cloning centromere-linked genes followed by partial overlapping hybridization, we constructed a 210-kb map encompassing the centromere in chromosome II and a 60-kb map near the centromere of chromosome I in the fission yeast *Schizosaccharomyces pombe* which has three chromosomes. Integration of the cloned sequences onto the chromosome and subsequent analyses of tetrads and dyads revealed an ~50 kb long domain located in the middle of the 210-kb map, tightly linked to the centromere and greatly reduced in meiotic recombination. This domain contained at least two classes of repetitive sequences. One, designated yn1, was specifically present in a particular chromosome and repeated three times in the 210-kb map of chromosome II. The other, designated dg, was located in all the centromere regions of three chromosomes. One (dgI) and two (dgIIa, dgIIb) copies of the dg were found in the maps of chromosomes I and II, respectively. The dgIIa and dgIIb were arranged with a 20-kb interval within the repetitive domain. In the centric region of chromosome III, 3-4 copies of the dg appeared to exist. By determining the nucleotide sequences of dgI and dgIIa, the dg was identified to be 3.8 kb long. The sequence homology was 99% between dgI and dgIIa. These extraordinarily homologous sequences seemed not to be transcribed into RNA nor to be encoding any protein. The larger part of the dg sequence was internally non-repetitious, a 600-bp region existed which consisted of stretches of several short repeating units. The structures in or surrounding the centromeres of *S. pombe* appear to be much more complex than those of the budding yeast *Saccharomyces cerevisiae*.

Key words: centromere/chromosome walking/domain structure/repetitive DNA/*Schizosaccharomyces pombe*

Introduction

In the eucaryotic cell cycle, chromosomal DNA duplicated in the S phase is held together at a specific site called the centromere (Blackburn and Szostak, 1984). The centromere is a strongly condensed, heterochromatic region which forms a constriction in the chromosome. By late prophase, arms of sister chromatids extending from the centromere are already separated, but the chromatids are held together by the centromere until the beginning of anaphase. The centromere is also the spindle fiber attachment site of a chromosome. To be precise, the kinetochore which is a complex proteinacous body attached laterally to the centromeric DNA (Ris and Witt, 1981; Valdivia and Brinkley, 1985) is the structure with which the ends of chromosomal microtubules are associated for chromosomal separation (Mitchison and Kirschner, 1985a, b; Murray and Szostak, 1985). Little is known about the molecular basis of the association between sister chromatids at the centromere and between centromeric DNA, kinetochore proteins and microtubules. Their structural and functional alterations during the cell cycle have to be fully characterized before centromere function is understood.

The centromere DNA segment (abbreviated CEN) was first isolated from a genomic library of the budding yeast *Saccharomyces cerevisiae* (Clarke and Carbon, 1980; Hsiao and Carbon, 1981; Fitzgerald-Hayes *et al.*, 1982b; Panzeri and Philippsen, 1982; Stinchcomb *et al.*, 1982; Maine *et al.*, 1983). The methods employed were (i) isolation of DNA sequences containing centromere-linked genes, followed by chromosome walking or overlap hybridization, and (ii) selection for sequences capable of imparting stable inheritance to autonomously replicating sequence (ARS) plasmids normally lost at high frequency during non-selective growth of yeast. Such CEN/ARS plasmids maintain a copy number of approximately one. In addition, the presence of this CEN DNA segment causes the plasmid genetic marker to segregate as a centromere-linked gene through meiosis.

S. cerevisiae has 17 chromosomes, and 10 CEN sequences have now been cloned (Hieter *et al.*, 1985a). These CEN DNAs share common properties, revealing three conserved sequence elements (Fitzgerald-Hayes *et al.*, 1982a; Carbon, 1984; Hieter *et al.*, 1985a). The centromere DNA sequence elements are localized within very small chromosome regions of < 120 bp. They contain common features in nucleotide sequences, although no cross-hybridization can be detected among the isolated CEN sequences. A prominent characteristic of all CEN sequences is the presence of a 78-86 bp high A+T sequence element (93-94% AT). This element is bound on one side by a conserved 25-bp sequence element and on the other side by the other element of a conserved 8-bp sequence. The nuclease-resistant core is centered over these three elements (Bloom and Carbon, 1982). Single microtubules appear to be attached to the centromeres of individual chromosomes (Peterson and Ris, 1976), and therefore the CEN-protein complex may act as the microtubule attachment site and may represent a primitive kinetochore structure (Bloom and Carbon, 1982).

The mitotic stability of circular CEN/ARS plasmids is strikingly increased by the order of 10^{-2}, but is still far less stable than a regular yeast chromosome ($\sim 10^{-5}$). Short linear CEN plasmids containing *Tetrahymena* telomeric ends (Szostak and Blackburn, 1982) are even more unstable, and rapidly lost in non-selectable media. This instability appears to be related to the small size of the linear mini-chromosomes, because the addition of 50-100 kb lambda phage DNA greatly improves the mitotic stability of plasmids (Murray and Szostak, 1983; Hieter *et al.*, 1985b). The mitotic stability of a linear minichromosome is possibly proportional to chromosomal length in *S. cerevisiae*.

A question to be solved is whether this elegant centromere system of the budding yeast can be applied to other organisms. The large kinetochores of higher eucaryotes might consist of repeats of a unit similar to the primitive yeast centromere. The

CEN/ARS plasmids of *S. cerevisiae* were introduced into the cells of other organisms including cultured animal cells, *Neurospora* and different yeasts but thus far have not been stably inherited (Carbon, 1984). In the present study we have attempted to isolate the centromere DNA of the fission yeast *Schizosaccharomyces pombe*. This organism is evolutionarily far distant from *S. cerevisiae* (Kaufer *et al.*, 1985; Matsumoto and Yanagida, 1985) and has only three chromosomes per haploid genome (Kohli *et al.*, 1977; Umesono *et al.*, 1983; Gygax and Thuriaux, 1984). Chromosomes of *S. pombe* are on average 6-fold longer than those of *S. cerevisiae*. They are condensed during mitosis (Toda *et al.*, 1981; Umesono *et al.*, 1983) and separated by the kinetochore microtubules (Tanaka and Kanbe, 1986) in the spindle apparatus (McCully and Robinow, 1971) which is formed immediately before mitosis (Hiraoka *et al.*, 1984).

The methods employed to isolate the *S. pombe* centromere DNAs are basically the same as those employed for *S. cerevisiae*, but the results obtained are strikingly different, as reported in this paper. We isolated DNA sequences containing centromere-linked genes of chromosomes I and II followed by partial-overlapping hybridization, and constructed a 210-kb map surrounding the centromere of chromosome II and a 60-kb partial map near the centromere of chromosome I.

Identification of the centromeric DNA by selecting sequences from the cloned DNA has failed; none of the cloned DNA fragments inserted into ARS plasmids are capable of imparting stable segregation to the plasmids in the cells of *S. pombe*. By integrating the cloned sequences onto the chromosome and subsequently analyzing tetrads and dyads, however, we found a large 50-kb domain in the middle of the 210-kb map of chromosome II that is tightly linked to the centromere and greatly reduced in meiotic recombination. Interestingly, the same domain contained different classes of repetitive DNA sequences. In contrast, the surrounding regions showed regular recombination frequencies and consisted of unique sequences.

Nucleotide sequence determinations show that there are at least two classes of the centromere-linked repetitive sequences. One is specifically present in a particular chromosome, while the other is commonly present in all the centromere regions. The latter sequence, designated dg, is ~3.8 kb long, apparently non-coding and surprisingly homologous (99%). The *S. pombe* centromere structure and function are discussed in the light of these findings.

Results

Isolation of the TPS13 gene linked to the centromere of chromosome II

A temperature-sensitive (ts) *tps13* (formerly *tsl24*) locus is tightly linked to the centromere of chromosome II (Figure 1; Kohli *et al.*, 1977; Nakaseko *et al.*, 1984; Gygax and Thuriaux, 1984). Therefore, we began chromosome walking with isolation of the *TPS13* gene. We constructed a gene library of pDB248′ (Beach and Nurse, 1981) that contained the wild-type *S. pombe* genomic DNA partially digested with *Sau*3A1. By transformation, plasmid pTS13 that complemented *tps13* was obtained. It contained a 9 kb long genomic DNA insert (Figure 2a).

By subcloning, the complementing activity (indicated by +) was found to reside in a 4.3-kb *Bam*HI–*Bgl*II fragment in pSS13. Genomic Southern blot hybridization indicated that the 9 kb long *TPS13* sequence was unique in the genome (data not shown). By integration on the chromosome and subsequent tetrad analysis, the cloned sequence was identified as being derived from the *TPS13* gene (described below).

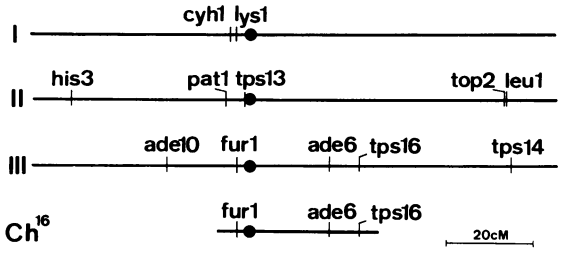

Fig. 1. Marker genes near the centromeres of the fission yeast *S. pombe*. The chromosome locations of the centromere-linked genes are shown. *S. pombe* has three chromosomes I, II and III per haploid genome. Ch^{16} is an artificially-made minichromosome containing three centromere-linked marker genes of chromosome III (Niwa *et al.*, 1986). The filled circles indicate the position of the centromeres. Centromere linkage analyses indicate that *lys1*, *tps13* and *fur1* are most linked to the centromere of I, II and III, respectively (Kohli *et al.*, 1977; Nakaseko *et al.*, 1984); the segregation patterns of these markers were mostly reductional. A detailed chromosomal map of *S. pombe* is shown in Gygax and Thuriaux (1984) and Kohli *et al.* (1977).

Cloning of 43-kb DNA extending from the TPS13 gene

A Charon 4A library containing 15–20 kb long inserts of *S. pombe* genomic DNAs (Matsumoto and Yanagida, 1985) was screened by plaque hybridization using pTS13 as the probe. Restriction maps of the positive clones were partially overlapped with each other and also with the probing *TPS13* sequence; three of the seven clones obtained, H4, H5 and H14, are shown in Figure 2b. For the second walking, we employed two probes pSS100 and pSS101 (Figure 2b). Phage clones hybridizing to these probes were obtained, and two of each of them are shown (H20 and H22 from pSS100, and H6 and H7 from pSS101).

The cloned genomic sequences were extended from the *TPS13* gene in both directions, and their restriction sites were consistent such that they could be aligned in an overlapping, linear map of 43 kb long DNA (Figure 2b). An *Eco*RI site in the *TPS13* gene was defined as the start (indicated by 0 above the site) for chromosome walking, and thereby the 43 kb long sequence spanned a region between the minus end at −19 kb and the plus end at +24 kb.

Orientation of the 43-kb map to the centromere

To determine the genetical distances from the *tps13* locus, the cloned sequences with a marker gene were integrated on the chromosome by homologous recombination. Each of the four fragments A, B, C and D (Table I; their locations indicated by the filled boxes in Figure 2b) was ligated with an integration vector YIp32 containing the *S. cerevisiae LEU2* as the marker gene (Botstein *et al.*, 1979). Resulting hybrid plasmids did not contain ARS and were employed to transform *S. pombe leu1*. Leu^+ transformants obtained were crossed with *tps13*, and the tetrads were dissected to determine the map distances between *tps13* and the integrated loci.

Tetrad data (Table I) indicated that the integrated locus for fragment A was 6.1 cM apart from *tps13*, while the loci for B, C and D were 1.3, <0.4 and 1.0 cM from *tps13*, respectively. Note that fragment C was derived from the *TPS13* gene (pTS13). The distance between A and the start was 18 kb, giving a ratio of 3 kb per cM for the direction of the minus end. On the other hand, D and the start were 21 kb apart in the opposite direction, giving a 7-fold higher ratio (21 kb/cM). The genetical distance per unit DNA length was asymetric in the two directions; ~3 cM per 10 kb for the minus end and 0.4 cM per 10 kb for the

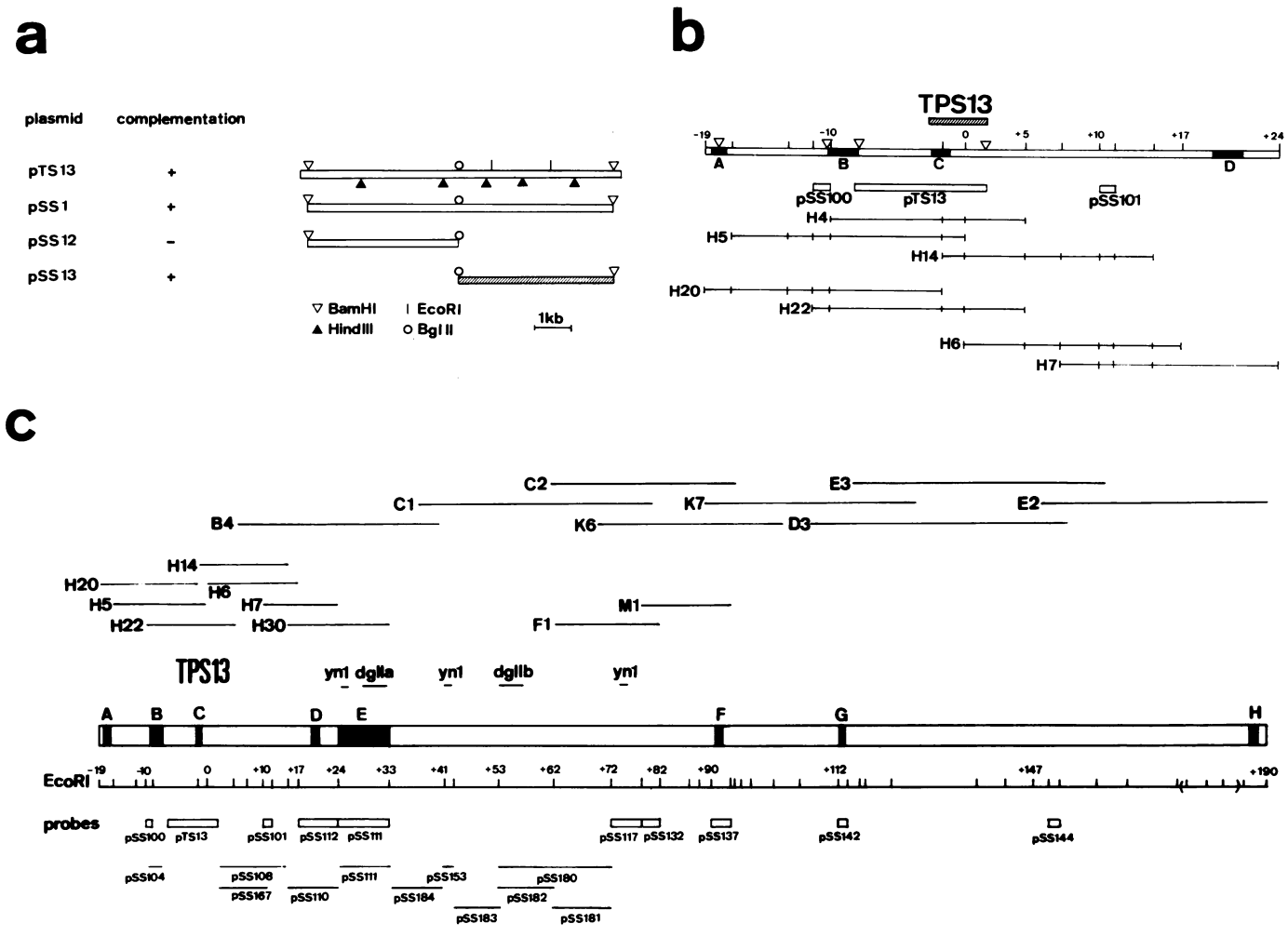

Fig. 2. Chromosome walking near the centromere of chromosome II. **(a)** Cloning of the *TPS13* gene by transformation. A gene library was made by ligating *Sau*3A1 partial digests of the wild-type *S. pombe* genomic DNA with a shuttle vector pDB248′ (Beach and Nurse, 1981) that contained the *S. cerevisiae LEU2* gene, 2 μm DNA and pBR322. A 9 kb long genomic insert of pTS13 complemented the temperature-sensitive *tps13* mutation (indicated by +). Complementing activity was restored in the 4.3-kb *Bgl*II−*Bam*HI fragment of pSS13. Restriction sites for *Eco*RI, *Bam*HI, *Bgl*II and *Hin*dIII in the cloned sequences are shown. **(b)** Cloning of a 43 kb long region surrounding the *TPS13* gene by partial-overlapping hybridization. The wild-type *S. pombe* DNA was partially digested with *Eco*RI, and the 15−20 kb long fragments were isolated by sucrose gradient centrifugation, followed by ligation to Charon 4A. The gene library thus made (Matsumoto and Yanagida, 1985) was screened by plaque hybridization using pTS13, pSS100 and pSS101 as probes. *Eco*RI restrictions maps of the clones obtained (H4, H5 and H14 from pTS13; H20 and H22 from pSS100; H6 and H7 from pSS101) were partially overlapped, and the resulting assembled map was extended in both directions from the *TPS13* gene. All of the clones obtained were consistently aligned into a linear map of 43 kb in length. The sites for *Eco*RI are indicated by vertical lines. *Eco*RI fragments indicated by A, B, C and D (the filled box) were subcloned into YIp32 and used for integration on the chromosome (see Table I). The numbers above the *Eco*RI restriction sites represent kilobase (kb) distances from the starting *Eco*RI site (indicated by 0) in the *TPS13* gene. **(c)** A 210-kb map determined by chromosome walking using cosmid and Charon 4A libraries. This map is shown to encompass the centromere of chromosome II. Ten probes used for walking are indicated by the empty boxes. Fragments A−H used for integration on the chromosome are indicated by filled boxes. Sizes of the integrated fragments are shown in Table I. Kilobase distances from the starting *Eco*RI site in the *TPS13* gene are numerically indicated. Restriction sites for *Eco*RI are represented by vertical lines below the chromosomal DNA. Individual Charon 4A and cosmid clones are indicated by the lines above the chromosome DNA. Plasmids (pSS) used for examining stability are shown below.

plus end. Meiotic recombination in the region toward the plus end seemed to be reduced.

The orientation of the cloned sequences relative to the centromere was identified by the following evidence. Firstly, a Leu$^+$ transformant integrated with A was crossed with *his3* (Figure 1; Kohli *et al.*, 1977); tetrad data indicated that the integrated A did indeed exist between *his3* and *tps13*. *his3* was located in the short arm of chromosome II. Secondly, the genomic sequence that complements *pat1* (Figure 1; Iino and Yamamoto, 1985; *pat1* is identical to *ran1*; Nurse, 1985) was located between A and the *TPS13* sequence (Beach *et al.*, 1985; M.Yamamoto, personal communication). Therefore, the direction from the minus to the plus end seemed to correspond to that from the short to the long arm of chromosome II.

Chromosome walking by a cosmid library

To facilitate a faster walking beyond the plus end in the 43-kb map, we made a cosmid library containing ~40 kb long genomic DNA inserts that were partially digested with *Sau*3A1, followed by fractionation by sucrose gradient centrifugation. The cosmid vector (pSS10; Figure 3) employed had been ligated with the *S. cerevisiae LEU2* gene and thus could be used as a shuttle vector for *S. pombe*. If the genomic insert itself had ARS, it could also be used as a replication-type shuttle vector.

Results of chromosome walking by the cosmid library are shown in Figure 2c, together with those obtained by the Charon 4A library. Detailed restriction analyses showed that all of the cloned genomic DNA sequences could be consistently and linearly aligned. The entire length was 210 kb; the cloned DNA spanned

Table I. Map positions and centromere linkages of the cloned fragments integrated on chromosome II

Fragment	Size (kb)	Distance[a] (kb)	Plasmid	Tetrad analysis[b]					Dyad analysis[c]			
				PD	NPD	TT	Total	cM	Red	Equ	Total	cM
A	1.4 *Hin*dIII	−18	pSS106	94	0	3	97	6.1	80	18	98	9.2
B	2.5 *Bam*HI	−9	pSS104	113	0	3	116	1.3	121	2	123	0.8
C	1.0 *Hin*dIII	−2	pSS9	126	0	0	126	<0.4	126	0	126	<0.4
D	1.6 *Hin*dIII	+23	pSS107	98	0	2	100	1.0	128	0	128	<0.4
E	8.7 *Eco*RI	+33	pSS111	98	0	1	99	0.5	130	0	130	<0.4
F	1.2 *Hin*dIII	+93	pSS145	78	0	12	90	6.7	102	7	109	3.7
G	1.0 *Hin*dIII	+114	pSS146	78	0	23	101	11.3	68	21	89	11.8
H	1.4 *Hin*dIII	+189	pSS152	66	0	62	128	24.2		N.D.		

Each of the fragments A−H (location indicated in Figure 2c) was integrated on the chromosome by an integration vector YIp32 that contains the *S. cerevisiae LEU2* gene.

[a]Distance of the fragments from the start *Eco*RI site in *TPS13*.

[b]Integrated Leu$^+$ transformants were crossed with *tps13*, and tetrads were dissected to determine the genetical distance (cM) between *tps13* and integrated loci. The number of spores showing PD (parental ditype), NPD (non-parental ditype) and TT (tetratype) is shown.

[c]Centromere linkage of the integrated Leu$^+$ marker was analyzed by dyads containing homozygous *tws1* (see text). The number of diploid spores showing reductional (Red) and equational (Equ) are indicated. Genetical distance from the centromere was estimated by assuming that the equational-type spores represent tetratype.

the region −19 kb to +190 kb. Some of the phage and cosmid clones obtained are shown above the chromosomal DNA (represented by two parallel lines). Restriction sites for *Eco*RI are indicated below the chromosomal DNA. Some of the sites are indicated by the numbers which represent the kilobase distances from the starting *Eco*RI site in the *TPS13* gene. The filled boxes A−H are the fragments used for the integration of the chromosome, while the open boxes with the names of plasmids represent the 10 different hybridization probes used for walking.

Interestingly, the three probes (pSS112, 111 and 117) from the region +17 to +78 were found to be repetitive according to the genomic Southern blot hybridization patterns and the number of clones obtained by colony hybridization, whereas the other seven probes outside the region were unique in the genome (described below). Furthermore, this 60 kb (+17 to +78) long domain contained other repetitive DNAs and coincided with the region defined by low recombination frequency.

A domain with reduced recombination

To estimate the genetical lengths of the cloned sequences in the 210-kb map, each of the four cloned fragments, E−H (Figure 2c) was integrated on the chromosome by the same procedures used for the fragments A−D. Leu$^+$ integrants were crossed with *tps13*, and the tetrads were dissected (Table I). The genetical distances between *tps13* and the integrated positions of the marker gene were 0.5 cM for E (+33 kb), 6.7 cM for F (+93 kb), 11.3 cM for G (+114 kb) and 24.2 cM for H (+189 kb) in the direction of the long arm (the plus end). Thus the total genetical length of the 210-kb DNA map was ~30 cM (6 cM at −20 kb and 24 cM at +190 kb).

In Figure 4, the map distances between *tps13* and the eight integrated fragments (A−H) were plotted versus the physical DNA length (filled circles). In an ~60 kb long domain, meiotic recombination seemed to occur only rarely. Regular recombination took place outside the domain; the ratios of the DNA length to the map distance obtained were ~3.3 kb/cM for the direction of the minus end and 5.0 kb/cM for the plus end.

Centromere linkage analyses by the tws1 mutant

A mutant of *S. pombe tws1* which grows normally in the mitotic cycle produces in meiosis two diploid spored asci (dyad) (Nakaseko *et al.*, 1984). In this mutant, the first meiotic division takes place but the second does not, resulting in enclosure of diploid nuclei into spores. Therefore, the centromere-linked markers are

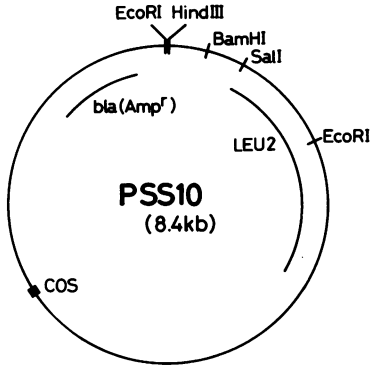

Fig. 3. A cosmid vector contructed for chromosome walking and transformation in *S. pombe*. The shuttle vector pSS10 was made as described in Materials and methods. It contains the cos site, the *S. cerevisiae LEU2* gene and the ampicillin marker. About 40 kb long genomic DNA fragments that were partially digested with *Sau*3A1 and isolated by sucrose gradient centrifugation were inserted at the *Bam*HI site of pSS10. For transformation of the *S. pombe leu1* strain, cosmid clones containing the genomic inserts were directly used.

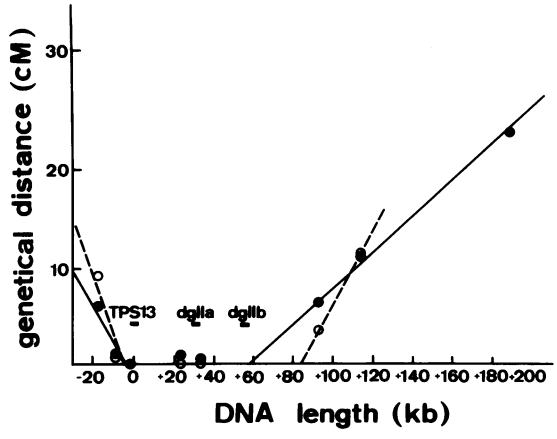

Fig. 4. Relationship between map distance and physical DNA length in the centromere-linked region of chromosome II. Each of the eight fragments (A−H) dispersedly located in the 210-kb map of chrômosome II (their locations indicated in Figure 2c) was subcloned into YIp32 and integrated in the chromosome by transformation. Integrated Leu$^+$ transformants were crossed with *tps13 tws1* (see text). The genetical distances between *tps13* and the integrated loci were estimated by tetrad analyses (the filled circles with solid line). The centromere linkages were determined by dyad analyses using *tws1* (the empty circles with broken line).

Table II. Segregation of ARS plasmids and cosmids containing the cloned sequences in the 210-kb map

Plasmids or cosmids	Fragments	ARS	Segregation of plasmid marker				
			$0^+:4^-$	$1^+:3^-$	$2^+:2^-$	$3^+:1^-$	$4^+:0^-$
pSS102	2.5 *Bam*HI	NDA2	11	1	0	0	0
pSS1	8.6 *Bam*HI	self	11	1	0	0	0
pSS2	8.6 *Bam*HI	self + NDA2	12	0	0	0	0
pSS130	16 *Bam*HI–*Sal*I	self	10	0	0	0	0
pSS108	16 *Bam*HI–*Sal*I	self + NDA2	16	0	0	0	0
pSS167	12 *Bam*HI–*Bgl*II	self + pDB	6	1	2	0	0
pSS110	9.9 *Eco*RI	TPS13	16	4	2	2	0
pSS112	7.5 *Eco*RI	NDA2	12	0	0	0	0
pSS111	10 *Eco*RI	self	12	0	0	0	0
pSS117	5.8 *Eco*RI	self	12	0	0	0	0
pSS115	5.8 *Eco*RI	self + NDA2	12	0	0	0	0
B4	40 *Sau*3A1	self	25	10	4	5	0
C1	40 *Sau*3A1	self	8	1	1	0	0
C2	40 *Sau*3A1	self	36	17	6	5	0
K7	40 *Sau*3A1	self	27	2	1	0	0
D3	40 *Sau*3A1	self	10	5	2	3	0
E2	40 *Sau*3A1	self	22	16	7	5	0

Chromosomal locations of the inserts in plasmids (pSS) and cosmids (B4, C1, C2, K7, D3, E2) are indicated in Figure 2c. The fragment DNA lengths are indicated by kb (the values for cosmids are approximate). Plasmids or cosmids were employed to transform *S. pombe leu1*, and resulting Leu^+ transformants were crossed with a *leu1* strain. The segregation of Leu^+:Leu^- was examined for each transformant.

only once segregated reductionally in the dyads, and the degree of centromere linkage can be estimated by the ratio of asci giving the reductional segregation pattern ($1^+:1^-$) to those giving the equational segregation pattern ($2^+:0^-$) in homozygous *tws1* zygotes.

Each of the Leu^+ *tws1* transformants integrated with one of the fragments A–G was crossed with *tws1* (*tws1* is recessive; Nakaseko *et al.*, 1984). Results of the dyad analyses are shown in Table I. The degree of the centromere linkage for each integrated fragment was plotted versus the physical DNA length in Figure 4 (open circles). An ~70 kb long domain covering the fragments C, D and E was very tightly linked to the centromere; no equational segregant ($2^+:0^-$ segregation) was obtained in the dyads. The centromere linkage was loosened for the other fragments B, F, A and G in this order, as the number of equational segregants increased in the crosses. Therefore, we concluded that the centromere of chromosome II must be localized within this domain which is basically the same area defined by repetitive DNAs and by low meiotic recombination.

Stability and segregation of plasmids containing the cloned sequences

We examined meiotic segregation of plasmids or cosmids containing various parts of the 210-kb cloned sequences. Sixteen different ARS plasmids or cosmids were made as shown in Table II (locations indicated in Figure 2c), and their meiotic segregation of the marker gene was investigated by tetrad analyses. The $2^+:2^-$ segregation was least frequent, while the $0^+:4^-$ segregation was most frequent. Cosmids (B4, C2, K7, D3 and E2) were generally more stable, but this appeared to be due to their longer sizes, because other cosmid clones containing the sequences unlinked to the centromere showed a similar segregation pattern.

Mitotic stability was also examined by measuring the rates of plasmid loss in non-selective rich media; none of the plasmids or the cosmids (Table II) showed marked stability. The functional assay for selecting the centromeric DNA (Clarke and Carbon, 1980; Hsiao and Carbon, 1981) has not been successful thus far in *S. pombe*.

Chromosome walk near the centromere region of chromosome I

In chromosome I, the two centromere-linked loci, namely *lys1* and *cyh1* are known (Figure 1; Kohli *et al.*, 1977; Nakaseko *et al.*, 1984). The previous dyad analysis suggested that *lys1* and *cyh1* are located in the same arm and that *lys1* is more proximal to the centromere than *cyh1*. Whether they exist in the short arm or long arm, however, is unknown.

The *LYS1* gene was cloned in pDB248′ by Beach *et al.* (1982); the 13 kb long insert in pLYS1 complements *lys1*. We cloned the *CYH1* gene (pCY1) by transformation (Materials and methods). An ~6 kb long genomic sequence complemented *cyh1* (Figure 5).

Chromosome walking was done using the Charon 4A and cosmid libraries with the probes indicated in Figure 5. Detailed restriction analyses enabled us to align the cloned DNA fragments into a 60 kb long linear map (Figure 5). The start for chromosome walking was defined as an *Eco*RI site in the *LYS1* gene sequence (indicated by 0 in Figure 5). Four probes (pBL201, pSS161, pSS162 and pSS160) were found to be repetitive DNAs by genomic Southern hybridization, whereas the other three probes were unique (data not shown). Chromosome integration has not been done to determine the map distances of the cloned fragments from *lys1* or *cyh1*. Some of the phage and cosmid clones obtained are shown in Figure 5. Stability and segregation of the cosmid clones were examined but none of them showed marked stability.

Centromere-linked repetitive sequences

Genomic Southern blot hybridization using various probes showed several different classes of repeating sequences in the centromere regions of chromosomes I and II. In chromosome II, they were localized between +17 and +78 kb, while in chromosome I, between +10 and +29 kb (the plus end). The copy number of these repetitive sequences in the centromere regions of chromosomes I and II seemed to be generally < 10.

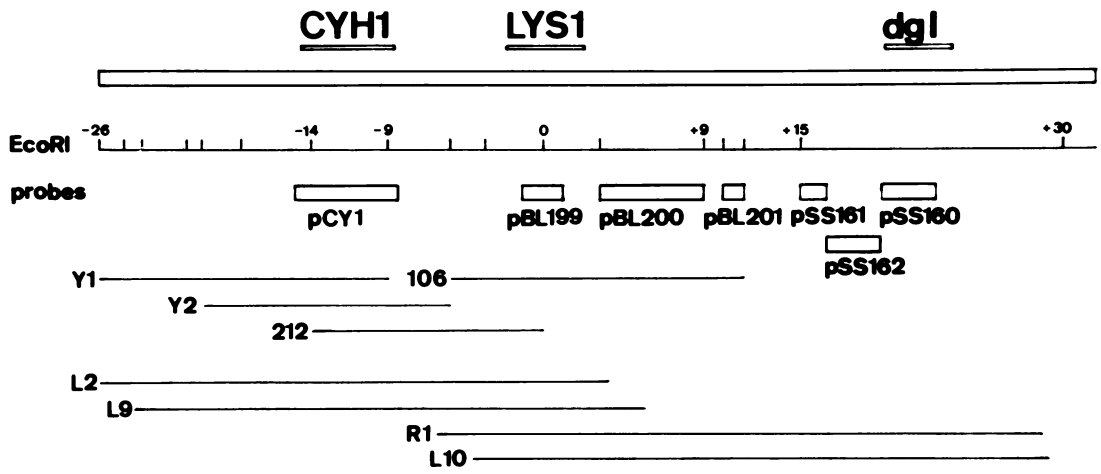

Fig. 5. Chromosome walking near the centromere region of chromosome I. An ~60-kb map near the centromere of chromosome I was constructed by partial-overlapping hybridization using Charon 4A and cosmid libraries, followed by detailed restriction analyses. The *LYS1* and *CYH1* genes which were cloned by transformation are tightly linked to the centromere of chromosome I, and *LYS1* is supposed to be more proximal to the centromere. The initial probe used for walking was the *LYS1* gene sequence; the starting *Eco*RI site is indicated by 0. At the position +20 kb from the start, the dgI sequence is present (see text). The four probes pBL201, pSS161, pSS162 and pSS160 contained repetitive DNAs whereas the other three probes contained unique sequences. Some of the Charon 4A and cosmid clones are shown.

An example of such repeating sequences, designated yn1, was present in the centromere region of chromosome II (locations indicated in Figure 2c). The nucleotide sequence of a *Sau*3A1 fragment shown in Figure 6 was a part of yn1. This 118-bp *Sau*3A1 fragment which was obtained from the 9-kb *Eco*RI fragment (+24 to +33 kb) was hybridized with the two other neighbouring 1.3- and 5.8-kb *Eco*RI fragments (positions at +41 to +42 and +72 to +78, respectively). The hybridizing *Sau*3A1 fragments were subcloned from the 1.3- and 5.8-kb *Eco*RI fragments, and their nucleotide sequences were also determined (Materials and methods). They were completely identical to that derived from the 9-kb *Eco*RI. This sequence was present neither in the cloned sequences of chromosome I nor in the minichromosome Ch[16] containing the centromere region of chromosome III (data not shown). Genomic Southern blot hybridization confirmed that only the three *Eco*RI fragments in the centromere region of chromosome II contained yn1. Thus, the yn1 existed specifically in the centromere region of chromosome II and was repeated three times at intervals of 20−30 kb.

A sequence commonly present in the centromere regions

One particular repeating sequence (designated dg) attracted our interest because it appeared to exist in all the centromere regions of the three chromosomes. Nucleotide sequence analyses showed that dg was ~3.8 kb long and had extremely high homology.

There were two copies of dg in the centromere region of chromosome II (designated dgIIa and dgIIb). Most of the dgIIa sequence was included in the 3.4-kb *Hin*dIII subfragment of 9-kb *Eco*RI (+24 to +33). When cosmid clones B4, C1, C2, K7 and D3 covering different regions of the cloned 210-kb map of chromosome II (Figure 2c) were restricted with *Eco*RI or *Hin*dIII and probed with dgIIa (3.4-kb *Hin*dIII), only B4 and C1 showed a hybridizing band at similar positions (Figure 7a). B4 covered from +6 kb to +40 kb, whereas C1 covered from +37 kb to +80 kb. By further subcloning and subsequent hybridization with detailed restriction analyses, we could determine the precise locations of dgIIa at +29 to +33 and dgIIb at +53 to +57, as illustrated in the map of Figure 7a. The distance between dgIIa and dgIIb was ~20 kb. Because the sizes of *Eco*RI and *Hin*dIII

```
GATCAATCGTTTTTAAGTTTTTAATTTTTTACTTTTAAAAAAGAAATTATATCTGCTTC
CTAGTTAGCAAAAATTCAAAAATTAAAAAATGAAAATTTTTTCTTTAATATAGACGAAG

ATTAATTAACGATTCTTTTGGAATGAAATCATTTTTACTTTTTTAAATAAAATGAGATC
TAATTAATTGCTAAGAAAACCTTACTTTAGTAAAAATGAAAAAATTTATTTTACTCTAG
```

Fig. 6. Partial nucleotide sequence of yn1 which repeats three times in a completely identical fashion at 20−30 kb intervals in the centromere region of chromosome II. Locations of the three yn1 sequences are indicated in Figure 2c. The yn1 sequence appears to exist only in chromosome II.

fragments containing dgIIa or dgIIb were similar (~9 and 3 kb, respectively), the two hybridizing bands became overlapped in the pattern of genomic Southern hybridization (Figure 7b and c).

Genomic Southern blot hybridization probed with dgIIa showed a multiple but limited number of hybridizing bands (Figure 7b). For example, *Eco*RI digestion produced three bands at 15, 9 and 6 kb. dgIIa and IIb should give rise to the *Eco*RI bands at 9 kb. As described below, the 15-kb band originated from chromosome I, and the 6-kb band from chromosome III. The intensity of the hybridizing bands appeared to reflect the copy number of the repeating elements. The 6-kb band was most intense (roughly equivalent to the intensity of four copies), while the 9- and 15-kb bands were less (~2 and 1 copies, respectively).

We found that the cosmid clones R1 and L10 which were obtained by walking in chromosome I (Figure 5) contained the 15-kb *Eco*RI fragment that was strongly hybridized with dgIIa under a stringent hybridizing condition (Figure 7c). After restriction analyses of the subcloned 15-kb *Eco*RI fragment and subsequent hybridization with dgIIa, the homologous sequence (designated dgI) was localized in the 3.0-kb *Hin*dIII−*Bam*HI subfragment (location indicated in Figure 5). No hybridizing band was obtained in the other cosmid clones of chromosome I.

Evidence that the dg sequence is also present in the centromere region of chromosome III is shown in Figure 8a and b. Artificially-made minichromosome Ch[16] (Niwa *et al.*, 1986) deletes a large part of chromosome III except its centromere region. Ch[16] is ~500 kb long and runs as a single band on pulsed field gradient (PFG) gel electrophoresis (Schwartz and Cantor, 1984; Carle and Olson, 1984). This minichromosome band was found

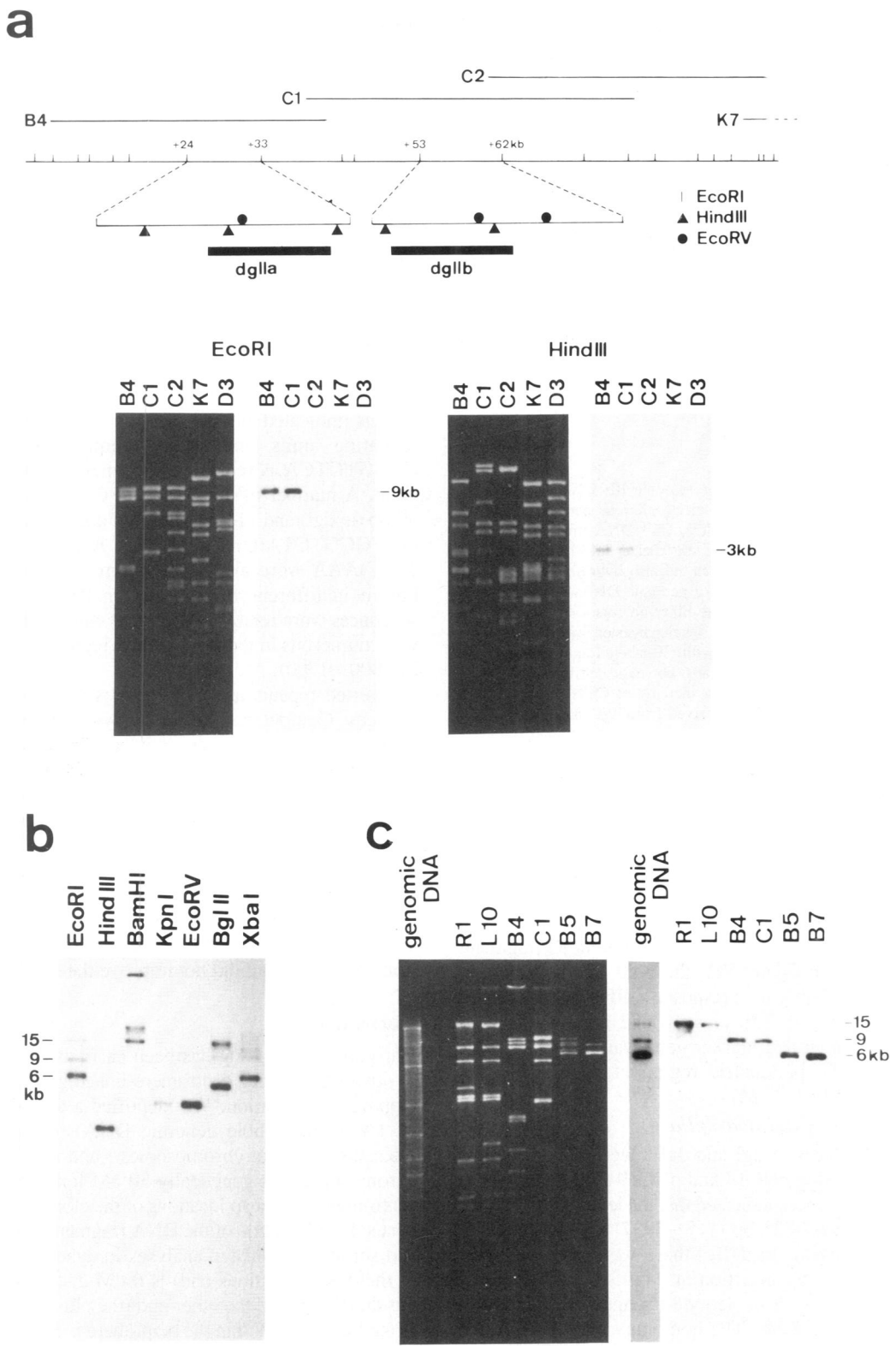

Fig. 7. The dg sequences are present in the centromere regions of different chromosomes. **(a) Top:** the restriction maps of the region containing the dgIIa and dgIIb. **Bottom:** Southern blot hybridizations of cosmid clones restricted with *Eco*RI (left) or *Hin*dIII (right) and probed with ^{32}P-labeled 3.4-kb *Hin*dIII fragment which contains dgIIa. Ethidium bromide-stained agarose gel patterns are shown together. These cosmid clones cover different regions of the 210-kb map of chromosome II as shown in the map (see also Figure 2c). Only B4 and C1 showed hybridizing bands at similar positions (at ~9 kb for *Eco*RI and ~3 kb for *Hin*dIII). The precise locations of dgIIa and dgIIb were determined by detailed restriction analyses. **(b)** Genomic Southern blot hybridizations of the wild-type *S. pombe* DNAs probed with ^{32}P-labeled dgIIa. Restriction enzymes used are indicated. A small number of hybridizing bands with different intensities are observed. **(c)** Cosmid clones derived from different chromosomes are hybridized with the dg sequence. The cosmids were digested with *Eco*RI and run in agarose gel electrophoresis. R1 and L10 which produce a hybridizing band at 15 kb are from chromosome I (Figure 5) while B4 and C1 which produce a band at 9 kb are from chromosome II. B5 and B7 producing a 6-kb band are likely to be derived from chromosome III (see Figure 8). Ethidium bromide-stained patterns are also shown.

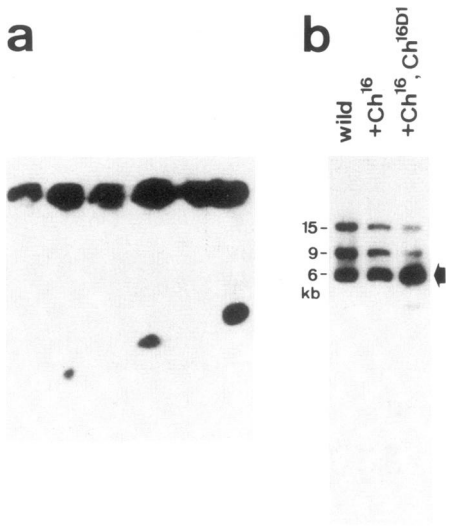

Fig. 8. Existence of the dg sequence in chromosome III. **(a)** Southern blot hybridization of a pulsed field gradient gel electrophoretic pattern of HM248 DNA which contains a minichromosome Ch^{16}. Ch^{16} DNA migrates as a single band and is strongly hybridized with ^{32}P-labeled dgI. Three regular chromosomes remain in the top wells, which are also hybridized with the dgI probe. **(b)** Southern blot hybridization of genomic DNAs derived from three different haploid strains which contain either no, one or two minichromosome Ch^{16}. The genomic DNAs were digested with *Eco*RI, run in agarose gel electrophoresis and probed with ^{32}P-labeled dgI. Note that the relative intensities of the hybridizing band at 6 kb (indicated by the arrow) are increased according to the copy number increase of Ch^{16}. The hybridizing bands at 15 kb and 9 kb are derived from dgI and dgII, respectively.

to be strongly hybridized with dgI (indicated by the arrow in Figure 8a). Furthermore, genomic Southern blot hybridization of the strains additionally containing one (Ch^{16}) or two (Ch^{16} +Ch^{16D1}) minichromosomes showed that the 6-kb *Eco*RI bands became increasingly more intense than that of the wild-type strain (indicated by the arrow in Figure 8b), suggesting that the 6-kb band originated from the centromere region of chromosome III and Ch^{16}. Cosmids containing the multiple 6-kb *Eco*RI fragments were cloned (B5, B7 in Figure 7c); the 6-kb *Eco*RI fragments might be arranged tandemly in chromosome III. The 6-kb *Eco*RI sequence was ligated with YIp plasmid and integrated on the chromosome. The integrated marker gene was tightly linked to *fur1* (Figure 1) in the pericentric region of chromosome III (PD:NPD:TT=29:0:1, 1.7 cM).

Nucleotide sequences of dgI and dgIIa

The nucleotide sequences of dgI and dgIIa were determined by the dideoxy method using pUC18 and pUC19 (Figure 9). Comparison of the two sequences indicated that the length of the highly homologous region was 3825 bp (113−3937) for dgI and 3841 bp (113−3953) for dgIIa. In dgIIa, there were 8-bp deletions (indicated by −), a 24-bp insertion at 2429, and 16 scattered replacements. A total of 48 nucleotide changes were found in sequences longer than 3.8 kb. The homology between dgI and dgIIa was nearly 99%. Most changes in deletions and additions were found in the short repetitive sequences. Many replacements were of the type purine to purine (8/16). In contrast to the extremely high homology, the sequences adjacent to dgI and dgIIa showed no resemblance at the 3′end and 30−35% homology at the 5′ end.

The dg sequences appeared to contain domains among which a 600 bp long repetitive zone between 2300 and 2900 was most prominent. This domain had different classes of short direct repeats (indicated by the arrows in Figure 9). In Table III, short repeating units in the dg sequences are summarized. CTAGTGTCA is repeated five times, forming a 45-bp subdomain. A number of the CCA and CCACT repeats constitute a 90-bp (in dgI) and 114-bp (in dgIIa) domain, while repeating units of TGCGTCTA(C)GGTATCC, AAGTGG and CATCATTA(C)AAA were also found. Some other direct repeats were present in different regions between 350 and 500. The repetitive sequences were relatively GC-rich, whereas the A and T clusters were numerous in the 1200−1800 region (AT content was 83% in 1600−1700).

Table III. Short repeating units in the dg sequence

Repeating unit	No. of repeats in dgI (dgIIa)	Total length (bp) in dgI (dgIIa)
CAAA	8 (7)	32 (28)
GATGAAAT	2 (2)	16 (16)
TGTGGT	2 (2)	12 (12)
CTAGTGTCA	5 (5)	45 (45)
GTACTAGTGTCATCTAGTGTCA	2 (2)	44 (44)
CATCA	2 (2)	10 (10)
CCAXX	13 (18)	90 (114)
TGCGTCTXGGTATCC	2 (2)	30 (30)
AAGTGG	2 (2)	12 (12)
TAGTAT	2 (2)	12 (12)
CATCATTXAAA	2 (2)	22 (22)

Inverted repeats and palindromes were scarce in the dg sequences. Computer sequence analyses indicated that the dg sequences contained no open reading frame. Furthermore, RNA blot hybridization of total *S. pombe* RNA probed with dgI showed no detectable hybridizing band. The dg sequence appeared to contain ARS; dgIIa was ligated with an integration vector YIp32 and the resulting hybrid plasmid transformed *S. pombe leu1* with a high frequency. Furthermore, the transformant cells contained a high copy number of monomeric supercoiled plasmids, but rapidly lost them. Thus, the dg sequences were non-coding, highly homologous with each other and existed only in the centromere region, but did not improve the stability of the plasmid.

Discussion

Chromosome walking has been carried out in the fission yeast *S. pombe* by cloning centromere-linked genes, followed by overlapping hybridization. We identified a total of 270 kb of DNA (>1% of the whole genomic DNA) to construct maps near the centromeres of chromosomes I and II. The 210-kb map for chromosome II is genetically 30 cM long and encompasses the centromere. The map locations of the cloned DNAs were demonstrated by integration of the DNA fragments onto the chromosome and subsequent linkage analyses in dyads and tetrads. One end of the map (the minus end) is 6 cM distal to the centromere in the short arm, and the other end (the plus end) is ~25 cM distal in the long arm. Within the map, there is an ~50 kb long domain which is closely linked to the centromere and is also reduced in meiotic recombination. Furthermore, this domain coincides with a region of approximately the same length that contains different classes of repetitive DNAs. In contrast, the adjacent regions in both sides consist of unique sequences and show regular recombination frequencies. Therefore, the centromere area of chromosome II is characterized by the presence of the 50 kb long repetitive, low-recombination domain.

In chromosome I, classes of repetitive DNA sequences are

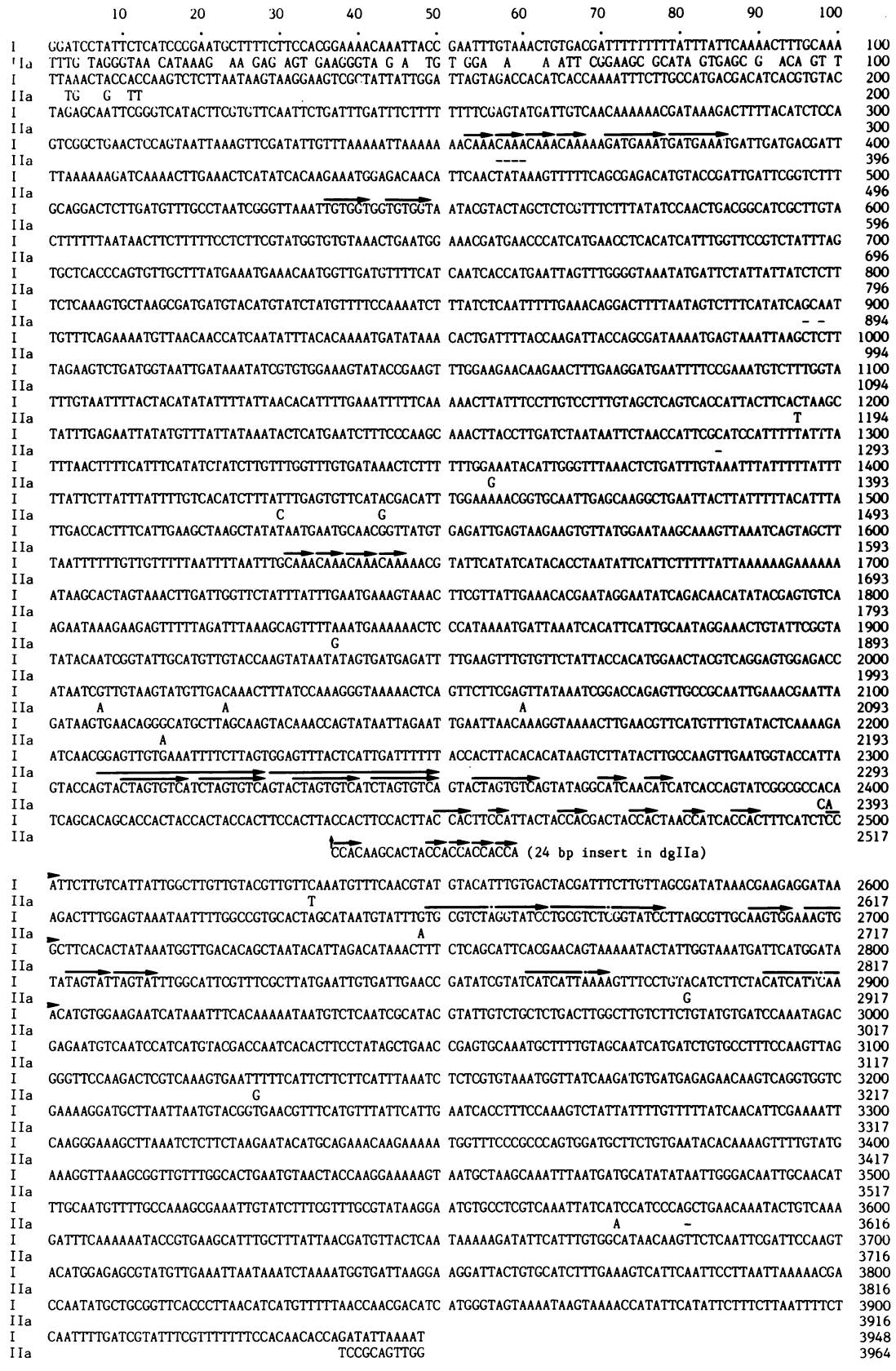

Fig. 9. Nucleotide sequences of dgI and dgIIa. 3948 nucleotides of dgI and flanking sequences are compared with 3964 nucleotides of dgIIa and flanking sequences. For dgIIa, only altered nucleotides are indicated (− indicates deletion). The dgI and dgIIa are present in different chromosomes (I and II, respectively), and their locations are indicated in Figures 2c and 7a. The highly homologous parts are 113−3937 for dgI and 113−3953 for dgIIa; the homologous nucleotide lengths are 3825 bp in dgI and 3841 bp in dgIIa. Forty-eight nucleotide changes exist between dgI and dgIIa, making up 99% homology. The arrows on the nucleotide sequences indicate different repeating units found in the dg sequences (see also Table III).

present in the 60-kb map. Chromosome walking was also undertaken in the centromere region of chromosome III by using fragments of the minichromosome Ch^{16} DNA as the probes (T.Matsumoto, unpublished). Ch^{16} contains the pericentric region of chromosome III and migrates as a single band in PFG gel electrophoresis (Niwa *et al.*, 1986). Fragments of Ch^{16} DNA were extracted from the band in the agarose gel and subsequently integrated onto the chromosome. Those fragments tightly linked to the centromere of chromosome III (the *fur1* locus as the marker) were found to be bound with the repetitive DNA sequences. Thus all three chromosomes of *S. pombe* seem to contain repetitive DNAs in the centromere regions. The meiotic recombination frequencies in these regions of chromosomes I and III, however, have not been determined.

We showed that there are at least two classes of the centromere-linked repetitive DNAs. One is located in the centromere region of a particular chromosome. The yn1 sequence which repeats three times within the 50-kb repetitive domain of chromosome II belongs to this class. The exact size of yn1 is not known. Partial nucleotide sequences of the three yn1 are identical and highly A+T rich (89%). The other class of repetitive DNA is localized in all the centromere regions of the three chromosomes; the dg sequences belong to this class. There may be other different classes of repetitive DNAs, but only future study can clarify this point.

The two dg sequences (dgIIa and dgIIb) are localized at an interval of 20 kb in the middle of the repetitive DNA domain of chromosome II. On the other hand, the one dg (dgI) is found in chromosome I. Chromosome III appears to contain 3–4 copies of the dg sequences. Evidence for the centromeric location of dgIII was obtained by Southern hybridization of the electrophoretic band of the minichromosome Ch^{16} probed with dgI. In the whole genomic DNAs containing no, one or two copies of Ch^{16}, the relative intensities of the hybridizing 6-kb *Eco*RI fragments were increased depending upon the number of minichromosomes, while the intensities of the other two bands corresponding to the dg sequences in chromosomes I and II remained the same. In short, the dg sequence repeats rarely (6–7 times) and specifically (only at the centromere region) in the genome of *S. pombe*. The copy number of the dg sequence per chromosome seems to be inversely related to the size of the chromosome; chromosome I is the largest while chromosome III is the smallest.

We determined the nucleotide sequences of dgIIa and dgI in different chromosomes. Their sequence homology is very high. The two dg sequences are ~3830 bp long and differ in 48 nucleotides, making up 98.7% homology. Preliminary sequence data of dgIIb give similar sequence homologies to dgI and dgIIa. Considering that the dg sequences seem neither to encode any protein nor to synthesize RNA, the conservation of such high homology along the 3.8-kb DNA length is surprising. To our knowledge, this homology is exceptionally high. The dg sequence consists of mostly unique-type sequences, but a 600-bp subdomain contains stretches of several different small repeating units (Table III). None of the repeating elements are similar to the *Saccharomyces* centromere sequence elements. The role of the dg sequences in the centromeric functions remains to be clarified.

The centromere regions of higher eucaryotic chromosomes are known to contain highly repetitive sequences (Singer, 1982). Human α-satellite sequences which are classified into subgroups I–IV are localized in heterochromatic centromere regions. Their unit length is short (~170 bp) and homology is low (~60%). Mouse and bovine satellite DNAs also consist of short units (234 bp and 5–8 bp, respectively), and the extent of homology is similar to human α-satellite. Therefore, the dg sequences are not like the satellite DNA because they have longer unit length, lower copy number per genome and higher nucleotide sequence homology. A sequence similar to dg might exist in the centromere region of higher eucaryotes although Southern hybridization of mouse and chicken genomic DNAs probed with the dg sequences did not show any distinct hybridizing bands.

The failure to select DNA segments that impart stability to ARS plasmids or cosmids might be explained by the assumption that an exceedingly large DNA segment which cannot be incorporated even into a cosmid may be required for the centromere function in the fission yeast. The ~50-kb repetitive domain is a candidate for such a functional centromeric segment. Alternatively, the functional centromere DNA sequence is possibly very short as in the budding yeast *S. cerevisiae*, but may only be expressed in linear plasmids having telomeric sequences at the ends or in plasmids without repetitive sequences which might have inhibitory effects on proper chromosome segregation. Another hypothesis that *S. pombe* cannot carry any additional chromosomes is untenable, because the haploid cells normally grow with two additional minichromosomes (Niwa *et al.*, 1986). These minichromosomes constructed by deletion behave as normal chromosomes and are being dissected to identify the domains essential for the centromere functions. Our knowledge of the *S. pombe* centromeres is still very limited, but the present study showed that the structures in or surrounding the centromeres of *S. pombe* are much more complex than those of *S. cerevisiae*.

Materials and methods

Strains and media

S. pombe haploid strains used were HM123 (h^- *leu1*), MS32 (h^- *tws1 leu1*), MS48 (h^- *tps13 leu1*) and MS31 (h^+ *his2 tws1 leu1*). h^- and h^+ are mating type alleles. *tws1* is a meiotic mutation that produces two diploid asci due to sporulation without second meiosis (Nakaseko *et al.*, 1984). *tps13* is a temperature-sensitive locus tightly linked to the centromere of chromosome II (Kohli *et al.*, 1977). The *leu1* mutation can be complemented by the *S. cerevisiae LEU2* gene (Beach and Nurse, 1981). HM248 and ON288-2A are partial aneuploid strains which contain one and two artificially-made minichromosomes Ch^{16}, respectively (Niwa *et al.*, 1986).

Escherichia coli strains used are as follows (Maniatis *et al.*, 1982; Beach *et al.*, 1982): HB101 for transformation and preparation of plasmid, DH1 for preparation of cosmid, LE392 for preparation of Charon 4A and BJ5183 for recovery of plasmid from *S. pombe*.

S. pombe cells were grown in the complex YPD medium or the minimal SD medium. YPD contains 2% polypeptone, 1% yeast extract and 2% glucose. SD contains 0.67% yeast nitrogen base and 2% glucose. *E. coli* was grown in LB medium and, for preparation of lambda phage DNA, NZCYM medium was used. Standard genetical procedures described for *S. pombe* (Gutz *et al.*, 1974; Kohli *et al.*, 1977) were followed.

Plasmids and cosmids

pDB248′ (Beach and Nurse, 1981) is a multi-copy shuttle vector for *S. pombe* and was used for cloning *TPS13* and *CYH1* genes by transformation and also for examining stability of cloned sequences in *S. pombe*. A cosmid pHC79 (Hohn and Collins, 1980) was ligated to the *S. cerevisiae LEU2* gene, and the resulting cosmid pSS10 (Figure 3) was used for construction of a genomic library for *S. pombe*. YIp32 and 33 are integration vectors, containing the *S. cerevisiae LEU2* gene and pBR322 (Botstein *et al.*, 1979), and were used for integration of the cloned DNA sequences on the chromosome of *S. pombe*. pUC18 and pUC19 (Yanisch-Perron *et al.*, 1985) were employed for nucleotide sequence determinations.

Transformation of fission yeast

The protoplasting method using zymolyase and Novo mutanase (Beach *et al.*, 1982) and the lithium method without protoplasting (Ito *et al.*, 1983) were employed.

Construction of S. pombe genomic libraries

Procedures for construction of the Charon 4A library for *S. pombe* genomic DNA were previously described (Matsumoto and Yanagida, 1985). Briefly, *S. pombe* genomic DNA partially digested with *Eco*RI was run in a sucrose gradient and

15–20 kb fragments were collected, followed by ligation to Charon 4A and packaging.

For construction of the cosmid gene library, *S. pombe* genomic DNA was purified and partially restricted with *Sau*3A1, and run in a linear sucrose gradient (10–40%) for centrifugation (26 000 r.p.m. for 20 h). The fractions containing ~40-kb DNA were collected and the DNA fragments were ligated to a *Bam*HI site of a cosmid vector pSS10 (Figure 3).

Plaque hybridization, colony hybridization and Southern blot hybridization

Standard procedures described by Maniatis *et al.* (1982) were followed. Hybridization was carried out under the conditions of 5 × SSPE and 0.3% SDS containing 100 μg/ml salmon sperm DNA at 65°C.

Cloning of TPS13 and CYH1 genes

A pDB248′ DNA library containing *Sau*3A1 partial digests of *S. pombe* DNA was used for transformation. For cloning of the *TPS13* gene, the recipient strain MS48 was transformed at 26°C on a minimal SD medium, and Leu^+ transformants obtained were replica plated and incubated at 36°C. Plasmid DNA (pTS13) was recovered from a Ts^+Leu^+ transformant. By chromosomal integration of the cloned sequence (pSS9; see Table III), it was determined to be derived from the *TPS13* locus.

For cloning *CYH1*, the host strain that had cycloheximide-resistant *cyh1* and *leu1* markers was used as a recipient for transformation. Leu^+ transformants grown on SD at 30°C were replica plated on SD containing cycloheximide (100 μg/ml) at 30°C. A Leu^+Cyh^s (cycloheximide-sensitive) transformant was obtained, and its plasmid was recovered (pCY1). Because *cyh1* is a recessive mutation, the wild-type *CYH1* gene makes the resistant mutant become sensitive to the drug. The cloned *CYH1* sequence was integrated on the chromosome, and its Leu^+ marker was found to be tightly linked to Cyh^r (resistant); tetrad data was PD:NPD:TT=20:0:0 (<2.4 cM).

Nucleotide sequence determination

Nucleotide sequence determination was done by the dideoxy method (Sanger *et al.*, 1977) using M13 (Messing *et al.*, 1981) or pUC plasmids (Yanisch-Perron *et al.*, 1985). Three different DNA fragments containing the yn1 sequence were prepared as follows. (i) The 9-kb *Eco*RI fragment containing dgIIa was digested with *Sau*3A1, and the resulting fragments were cloned into the M13 mp8 *Bam*HI site. Plaque hybridization was done using ^{32}P-labeled pSS117 (Figure 2c) containing a 5.8-kb *Eco*RI insert. The nucleotide sequences of positive clones were determined. (ii) Inversely, the 5.8-kb *Eco*RI fragment was digested with *Sau*3A1, and the resulting M13 clones were screened by ^{32}P-labeled pSS111 (Figure 2c) containing the 9-kb *Eco*RI fragment. The nucleotide sequences of the positive clones were determined. (iii) The nucleotide sequence of pSS153 (Figure 2c) containing the 1.3-kb *Eco*RI fragment was determined using M13. For determination of dgI, the 3.0-kb *Bam*HI–*Hin*dIII and the 2.7-kb *Hin*dIII in the 15-kb *Eco*RI fragment of chromosome I (Figure 5) were cloned into pUC plasmids, and their nucleotide sequences were determined by the dideoxy method. For determination of dgIIa, the 3.4-kb *Hin*dIII and the 2.6-kb *Hin*dIII in the 9-kb *Eco*RI fragment of chromosome II (Figure 2c) were cloned in pUC plasmids, and their nucleotide sequences were determined.

Acknowledgements

We are greatly indebted to Paul Nurse for plasmid, to Mitsuru Takanami for cosmid, to Masayuki Yamamoto for information and to Katsunori Noguchi for construction of a pDB248′ library. This work was supported by grants from the Mitsubishi Science Foundation, the Ministry of Education, Science and Culture and the Science and Technology Agency of Japan.

References

Beach,D. and Nurse,P. (1981) *Nature*, **290**, 140–142.
Beach,D., Piper,M. and Nurse,P. (1982) *Mol. Gen. Genet.*, **187**, 326–329.
Beach,D., Rodgers,L. and Gould,J. (1985) *Curr. Genet.*, **10**, 297–311.
Blackburn,E.H. and Szostak,J.W. (1984) *Annu. Rev. Biochem.*, **53**, 163–194.
Bloom,K.S. and Carbon,J. (1982) *Cell*, **29**, 305–317.
Botstein,D., Falco,S.C., Stewart,S.E., Brennan,M., Schera,S., Stinchcomb,D.T., Struhl,K. and Davis,R.W. (1979) *Gene*, **8**, 17–24.
Carbon,J. (1984) *Cell*, **37**, 351–353.
Carle,G.F. and Olson,M.V. (1984) *Nucleic Acids Res.*, **12**, 5647–5664.
Clarke,L. and Carbon,J. (1980) *Nature*, **287**, 504–509.
Clarke,L. and Carbon,J. (1983) *Nature*, **305**, 23–28.
Fitzgerald-Hayes,M., Clarke,L. and Carbon,J. (1982a) *Cell*, **29**, 235–244.
Fitzgerald-Hayes,M., Buhler,J.-M., Cooper,T.G. and Carbon,J. (1982b) *Mol. Cell Biol.*, **2**, 82–87.
Gutz,H., Heslot,H., Leupold,U. and Loprieno,N. (1974) In King,R.C. (ed), *Handbook of Genetics 1*. Plenum Press, NY, pp. 395–446.
Gygax,A. and Thuriaux,P. (1984) *Curr. Genet.*, **8**, 85–92.
Hieter,P., Pridmore,D., Hegemann,J.H., Thomas,M., Davis,R.W. and Philippsen,P. (1985a) *Cell*, **42**, 913–921.
Hieter,P., Mann,C., Snyder,M. and Davis,R.W. (1985b) *Cell*, **40**, 381–392.
Hiraoka,Y., Toda,T. and Yanagida,M. (1984) *Cell*, **39**, 349–358.
Hohn,B. and Collins,J. (1980) *Gene*, **11**, 291–298.
Hsiao,C.-L. and Carbon,J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3760–3764.
Iino,Y. and Yamamoto,M. (1985) *Mol. Gen. Genet.*, **198**, 416–421.
Ito,H., Fukuda,Y., Murata,K. and Kimura,A. (1983) *J. Bacteriol.*, **153**, 163–168.
Kaufer,N.F., Simanis,V. and Nurse,P. (1985) *Nature*, **318**, 78–80.
Kohli,J., Hottinger,H., Muntz,P., Strauss,A. and Thuriaux,P. (1977) *Genetics*, **87**, 471–489.
Maine,G., Surosky,R.T. and Tye,B.-K. (1983) *Mol. Cell Biol.*, **4**, 86–91.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
Matsumoto,S. and Yanagida,M. (1985) *EMBO J.*, **4**, 3531–3538.
McCully,E.K. and Robinow,C.F. (1971) *J. Cell Sci.*, **9**, 475–507.
Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309–321.
Mitchison,T.J. and Kirschner,M.W. (1985a) *J. Cell Biol.*, **101**, 755–765.
Mitchison,T.J. and Kirschner,M.W. (1985b) *J. Cell Biol.*, **101**, 766–777.
Murray,A.W. and Szostak,J.W. (1983) *Nature*, **305**, 189–193.
Murray,A.W. and Szostak,J.W. (1985) *Annu. Rev. Cell Biol.*, **1**, 341–367.
Nakaseko,Y., Niwa,O. and Yanagida,M. (1984) *J. Bacteriol.*, **157**, 334–336.
Niwa,O., Matsumoto,T. and Yanagida,M. (1986) *Mol. Gen. Genet.*, in press.
Nurse,P. (1985) *Mol. Gen. Genet.*, **198**, 497–502
Nurse,P., Thuriaux,P. and Nasmyth,K. (1976) *Mol. Gen. Genet.*, **146**, 167–178.
Panzeri,L. and Philippsen,P. (1982) *EMBO J.*, **1**, 1605–1611.
Peterson,J.B. and Ris,H. (1976) *J. Cell Sci.*, **22**, 219–242.
Ris,H. and Witt,P.L. (1981) *Chromosoma*, **82**, 153–170.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schwartz,D.C. and Cantor,C.R. (1984) *Cell*, **37**, 67–75.
Singer,M.F. (1982) *Int. Rev. Cytol.*, **76**, 57–111.
Stinchcomb,D.T., Mann,C. and Davis,R.W. (1982) *J. Mol. Biol.*, **158**, 157–179.
Szostak,J.W. and Blackburn,E.H. (1982) *Cell*, **29**, 245–255.
Tanaka,K. and Kanbe,T. (1986) *J. Cell Sci.*, in press.
Toda,T., Yamamoto,M. and Yanagida,M. (1981) *J. Cell Sci.*, **52**, 271–287.
Umesono,K., Hiraoka,Y., Toda,T. and Yanagida,M. (1983) *Curr. Genet.*, **7**, 123–128.
Valdivia,M.M. and Brinkley,B.R. (1985) *J. Cell Biol.*, **101**, 1124–1134.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.

Received on 28 February 1986