922–924 Nucleic Acids Research, 1995, Vol. 23, No. 6

Functional selection for the centromere DNA from yeast chromosome VIII

Ursula Fleig, Jens D. Beinhauer and Johannes H. Hegemann*

Institute of Microbiology and Molecular Biology, Justus-Liebig-University, Frankfurter Strasse 107, 35392 Giessen, Germany

Received December 16, 1994; Revised and Accepted January 31, 1995 EMBL accession no. X83205

ABSTRACT

Centromeres are essential components of eucaryotic chromosomes. In budding yeast, up to now, 15 of the 16 centromere DNAs have been isolated. Here we report the functional isolation and characterization of *CEN8*, the last of the yeast centromeres missing. The centromere consensus sequence for the 16 chromosomes in this organism is presented.

INTRODUCTION

Transmission of genetic material in eucaryotic cells relies on the mechanisms of mitosis and meiosis. During these processes spindle fibres attach to specialized regions on the chromosomes, the centromeres, which are complex structures made up of centromere DNA and centromere proteins (for a review see 1). In the budding yeast *Saccharomyces cerevisiae* the centromere DNA (*CEN* DNA) required to assemble a functional mitotic and meiotic centromere is 125 bp long and can be subdivided into three conserved DNA elements CDEI, CDEII and CDEIII (2–4). Up to now, 15 of the 16 yeast *CEN* DNAs have been isolated and sequenced (summarized in 5). Here we describe the functional selection and sequence analysis of *CEN8*, the last of the yeast centromeres to be isolated.

RESULTS AND DISCUSSION

An assay established by Hieter and co-workers for the functional selection of *CEN* DNAs is based on ARS plasmids carrying a *SUP11* tRNA gene (6). ARS plasmids in *S.cerevisiae* are able to autonomously replicate, but show a segregation bias in that they are rarely transmitted to daughter cells during mitosis and therefore accumulate in mother cells (7). *CEN* DNA present on the plasmid will result in proper segregation of the plasmid (i.e. one copy/cell). As expression of several copies of the *SUP11* tRNA gene is lethal, only cells which have functional *CEN* DNA on the plasmid will survive. We obtained from Prof. Isono (Kobe University) non-mapped λ clones containing 13–14 kb of chromosome VIII DNA inserts (8). *Eco*RI-cleaved insert DNA was cloned into the *Eco*RI-cleaved vector YRp14-*ARS1*, which carries the *ARS1* element, the *SUP11* gene and the selectable marker *URA3*. Transformation of the resulting plasmids into yeast strain YJH11 (described in 9) yielded no transformants except for the plasmid which carried a 4.5 kb *Eco*RI fragment called A3. Further subcloning of A3 and functional testing yielded a 2.55 kb *Bgl*II–*Eco*RI fragment showing *CEN* activity. Sequence analysis of this fragment revealed the presence of the characteristic centromere DNA sequence shown in Figure 1A, which is 100% identical to the *CEN8* sequence just published as part of the sequence analysis of chromosome VIII from *S.cerevisiae* (10).

To finally prove that this sequence is able to act as *CEN* DNA *in vivo*, it was placed on an artificial chromosome fragment and tested for its segregational behavior using the cycloheximide R/S system (11). To this end, the A3 *Eco*RI DNA fragment was cloned into pBluescript (Stratagene) and then subcloned as a *Bam*HI–*Sal*I fragment into the chromosome fragmentation vector pKE5 (11). Transformation of the linearized fragmentation vector into yeast strain YJH6 (described in 12) resulted in the formation of an additional chromosome fragment of about 125 kb carrying the A3 DNA fragment. Separation of the chromosomes of the transformants via OFAGE visualized the new chromosome (Fig. 2A). Hybridization of A3 DNA to the chromosome blot identified two bands. One represents chromosome VIII, the other the newly generated chromosome fragment (Fig. 2B). Using the R/S system, the mitotic loss rate of the chromosome fragment was determined as 1.1×10^{-3}. This is similar to the loss rates obtained for other *CEN* DNAs when analyzed with this system (e.g. *CEN6* shows a loss rate of 4.8×10^{-4}; 13).

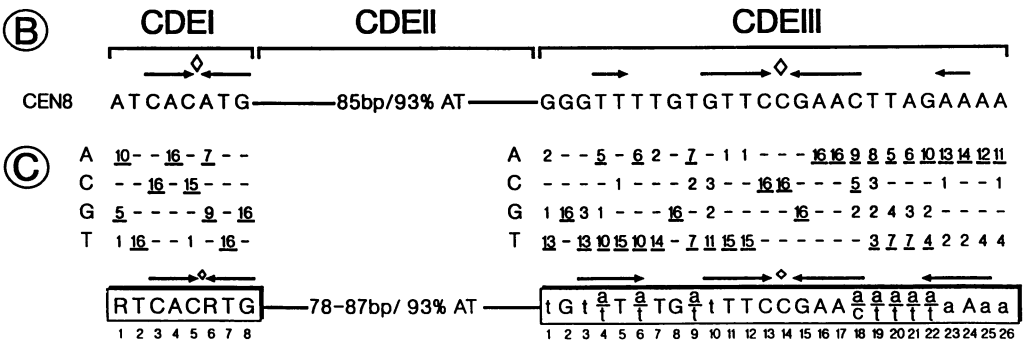
The availability of the *CEN* DNAs from all 16 yeast chromosomes allowed us to determine the final consensus sequence for yeast centromere DNA (2,3,14–19) (Fig. 1B and C). Centrally located in all *CEN* DNAs is the extremely AT-rich CDEII element (87–98% of A:T and T:A base pairs). The AT-rich DNA tends to be arranged in runs of A:T and T:A bp, but the element does not carry any detectable conserved sequence motif. It is worth mentioning that the CDEII elements from two *S.cerevisiae* strains from geographically different origins are completely conserved, indicating a possible sequence-specific role for this element in centromere assembly and/or function (18). Likewise, replacing parts of the CDEII sequence with random AT-rich sequences significantly reduces mitotic and meiotic centromere function (20). CDEII is flanked by CDEI and CDEIII, which are highly conserved and carry palindromic sequences. CDEI consists of a highly conserved 8 bp long sequence motif present in all 16 *CEN* DNAs (positions 1–8). Only positions 5 of

* To whom correspondence should be addressed

Ⓐ

```
  1  CTCTAATTAC ACACAAGTTA AAAAGTAGAT AAGAAGAACT CCAACAATTA   50
 51  CACATCCACA AAACGATTTA GTATAATAAG TAGGATGGTA ATCACATGAC  100
101  TAATAATTCT TTTAATTTTA ATTAATTTAA TAAAATTAAA ATAATATATA  150
151  TACTAAATTG TTTATTAAAA ATGATTAAAC ATTGGGTTTT GTGTTCCGAA  200
201  CTTAGAAAAT TGGTTAAGAC ATTATCTTAA AAACGCGAAA AATAAAAAAA  250
251  TGATTATTGA TGTTTATTTC ACAAGCACTA AGGGAGAGTT TTGTCAACAA  300
```

Ⓑ



Ⓒ

Figure 1. (**A** and **B**) DNA sequence of *CEN8* and surrounding area. The *CEN8* DNA is shown in bold letters. The centromere DNA elements (CDE) are indicated. The diamonds represent the axis of 2-fold symmetry and the arrowheads indicate palindromic sequences. The EMBL database accession no. is X83205. (**C**) Consensus sequence for yeast centromere DNAs. The sequences of all 16 *CEN* DNAs were aligned and the consensus sequence established. Numbers below the sequence identify individual positions. Capital letters indicate that the residue is present 14–16 times among the 16 *CEN* DNAs, lower case letters indicate that the residues appears 11–13 times. R indicates a purine. Also shown is the frequency with which each residue appears. The underlined numbers indicate the residues that were used to generate the consensus sequence.

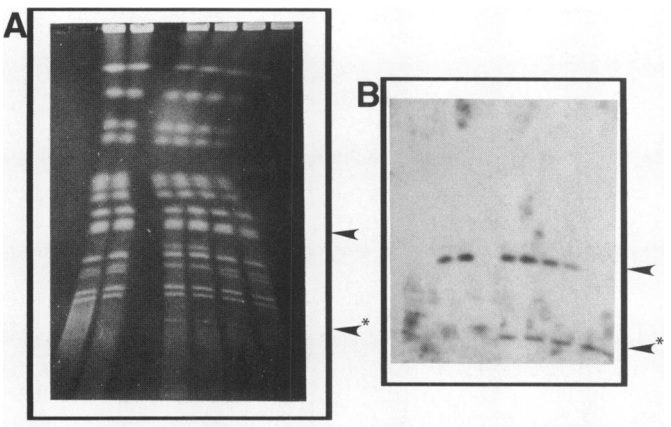


Figure 2. (**A**) OFAGE analysis of yeast transformants containing a chromosome fragment carrying a 4.5 kb *CEN8* fragment. (**B**) Southern analysis of OFAGE gel shown in (A). Fragment A3 was used as a probe. Position of chromosome VIII is indicated by an arrow, the chromosome fragment by an arrow with star.

CEN2 and 1 of *CEN9* do not fit the consensus sequence. In contrast, the 26 bp long CDEIII element shows little homology among the 16 *CEN* DNAs, except for the highly conserved central core element (positions 11–17) and two G:C bp at positions 2 and 8. The rest of this element is T:A-rich at the 5′ region and A:T-rich at the 3′ region. The overall sequence similarity among the 16 *CEN* DNAs is very low, making it unlikely that two or more of the *CEN* DNAs have an ancient predecessor in common. Alternatively, such a duplication must have occurred a long time ago, as the CDEII sequences are very different among all *CEN* DNAs. It has been proposed that the centromere region from chromosome VIII has been duplicated from the centromere region of chromosome XI, because several centromere-linked genes are found to be duplicated (10).

Compared with the genomes of higher eucaryotes, the yeast genome is very AT-rich. The complete sequence analysis of three yeast chromosomes has shown an average AT content of 61.7% (10,21,22). Interestingly, along their length the chromosomes consist of so called (G+C)-rich and (G+C)-poor segments of ~50 kb each (21,22). The centromeres are always located within a (G+C)-poor segment, in which the AT content can be up to 65%. Analysis of the first 40 bp left and right of each of the 16 *CEN* DNAs revealed that the three conserved centromere elements are embedded in an area with an average AT content of 75%. The next 50 bp left of CDEI and right of CDEIII have AT contents of 71 and 68% respectively (Heck and Hegemann, unpublished data). Thus the sequences surrounding yeast *CEN* DNA have a significantly higher AT content than other areas of the chromosomes. Exceptions are the sequences left of *CEN2* and those surrounding *CEN14*. We had shown previously that in the case of the *CEN6* centromere, function does not depend on a particular AT content within its surrounding DNA. Lowering the AT content left and right of *CEN6* from 75–80% to 38–44% did not reduce its full mitotic and meiotic centromere activity (4). Thus the natural variability in AT content of the sequences surrounding *CEN* DNA indicates the self-sufficiency of the 120 bp containing CDEI, CDEII and CDEIII in terms of centromere assembly and function. It is therefore unclear why yeast *CEN* DNA in general is embedded in a region with a particularly high AT content.

ACKNOWLEDGEMENTS

We are grateful to Prof. Isono (Kobe University, Japan) for providing us with the λ phage clones 5-S1H3 and 5-1D6. We thank H. Interthal for cosmid DNA, M. Sen-Gupta for DNA sequencing, S. Heck for excellent technical assistance and synthesis of oligonucleotides, C. Reitz for photographic work and A. Wilmen for help with the Southern blot analyses. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SFB272, A1) to J. H. Hegemann.

REFERENCES

1 Vig,B.K. (ed), (1993) *Chromosome Segregation and Aneuploidy.* Springer-Verlag, Berlin, NATO ASI Series H: Cell Biology, Vol. 72.
2 Fitzgerald-Hayes,M., Clarke,L. and Carbon,J. (1982) *Cell*, **29**, 235–244.
3 Hieter,P., Pridmore,D., Hegemann,J.H., Thomas,M., Davis,R.W. and Philippsen,P. (1985) *Cell*, **42**, 913–921.
4 Cottarel,G., Shero,J., Hieter,P. and Hegemann,J.H. (1989) *Mol. Cell. Biol.*, **9**, 3342–3349.
5 Hegemann,J.H. and Fleig,U.N. (1993) *BioEssays*, **15**, 451–460.
6 Hieter,P., Mann,C., Snyder,M. and Davis,R.W. (1985) *Cell*, **40**, 381–392.
7 Murray,A.W. and Szostak,J.W. (1983) *Cell*, **34**, 961–970.
8 Tanaka,S. and Isono,K. (1992) *Nucleic Acids Res.*, **20**, 3011–3020.
9 Niedenthal,R. (1993) PhD Thesis, University of Giessen, Giessen, Germany.
10 Johnston,M., Andrews,S., Brinkman,R., Cooper,J., Ding,H., Dover,J., Du,Z., Favello,A., Fulton,L., Gattung,S. *et al.* (1994) *Science*, **265**, 2077–2082.
11 Jehn,B., Niedenthal,R. and Hegemann,J.H. (1991) *Mol. Cell. Biol.*, **11**, 5212–5221.
12 Hegemann,J.H., Shero,J.H., Cottarel,G., Philippsen,P. and Hieter,P. (1988) *Mol. Cell. Biol.*, **8**, 2523–2535.
13 Niedenthal,R., Stoll,R. and Hegemann,J.H. (1991) *Mol. Cell. Biol.*, **11**, 3545–3553.
14 Panzeri,L. and Philippsen,P. (1982) *EMBO J.*, **1**, 1605–1611.
15 Maine,G.T., Sukorsky,R.T.,and Tye,B.-K. (1984) *Mol. Cell. Biol.*, **4**, 86–91.
16 Nietz,M. and Carbon,J. (1985) *Mol. Cell. Biol.*, **5**, 2887–2893.
17 Mann,C. and Davis,R.W. (1986) *Mol. Cell. Biol.*, **6**, 241–245.
18 Huberman,J. Pridmore,D.R., Jäger,D., Zonnefeld,B. and Philippsen,P. (1986) *Chromosoma*, **94**, 162–168.
19 Wustinger,K. and Spevak,W. (1993) *Nucleic Acids Res.*, **21**, 3321.
20 Murphy,M.R., Fowlkes,D.M. and Fitzgerald-Hayes,M. (1991) *Chromosoma*, **101**, 189–197.
21 Sharp,P.M. and Lloyd,A.T. (1993) *Nucleic Acids Res.*, **21**, 179–183.
22 Dujon,B., Alexandraki,D., Andre,B., Ansorge,W., Baladron,V., Ballesta,J.P.G., Banrevi,A., Bolle,P.A., Bolotin-Fukuhara,M., Bossier,P. *et al.* (1994) *Nature*, **369**, 371–378.