

# CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA

Jun-ichirou Ohzeki, Megumi Nakano, Teruaki Okada, and Hiroshi Masumoto

Division of Biological Science, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan

Centromere protein (CENP) B boxes, recognition sequences of CENP-B, appear at regular intervals in human centromeric  $\alpha$ -satellite DNA (alphoid DNA). In this study, to determine whether information carried by the primary sequence of alphoid DNA is involved in assembly of functional human centromeres, we created four kinds of synthetic repetitive sequences: modified alphoid DNA with point mutations in all CENP-B boxes, resulting in loss of all CENP-B binding activity; unmodified alphoid DNA containing functional CENP-B boxes; and nonalphoid repetitive DNA sequences with or without functional CENP-B boxes. These four synthetic repetitive DNAs were introduced into cultured human cells (HT1080), and de novo centromere

assembly was assessed using the mammalian artificial chromosome (MAC) formation assay. We found that both the CENP-B box and the alphoid DNA sequence are required for de novo MAC formation and assembly of functional centromere components such as CENP-A, CENP-C, and CENP-E. Using the chromatin immunoprecipitation assay, we found that direct assembly of CENP-A and CENP-B in cells with synthetic alphoid DNA required functional CENP-B boxes. To the best of our knowledge, this is the first reported evidence of a functional molecular link between a centromere-specific DNA sequence and centromeric chromatin assembly in humans.

## Introduction

The centromere is an essential structural domain of chromosomes and is responsible for accurate chromosome segregation in mitosis and meiosis. It has several functions, including sister chromatid adhesion, linking chromosomes to spindle microtubules, and synchronous separation of sister chromatids at anaphase onset (Choo, 1997a). These centromere functions are important for maintaining chromosomal euploidy in eukaryotic organisms.

Many protein components of the centromere have been identified, including centromere proteins (CENPs)\* A, B, C, E, H, and F (Earnshaw and Rothfield, 1985; for review see Warburton, 2001), and some of these protein components are conserved among yeast species and humans (for review see Kitagawa and Hieter, 2001). In particular, CENP-A, a centromere-specific histone H3 variant (Sullivan et al., 1994; Shelby et al., 1997), is highly conserved among

most eukaryotes (for review see Choo, 2001). This protein is required for assembly of other centromere components (Howman et al., 2000), and is considered a fundamental component of the centromere-specific chromatin structure. CENP-B, a highly conserved protein in humans and mice, is specifically localized at the centromere (Earnshaw and Rothfield, 1985). This protein binds to the 17-bp motif of the CENP-B box sequence in alphoid DNA at its amino-terminal region and forms homodimers at its carboxy-terminal region (Masumoto et al., 1989; Yoda et al., 1992). CENP-C, a constitutive centromere protein, localizes at the inner kinetochore lamina on mitotic chromosomes and is required for centromere function (Saitoh et al., 1992; Kalitsis et al., 1998). CENP-E transiently assembles on the outer surface of the kinetochore on mitotic chromosomes and is required for mitotic check point and interactions between the kinetochore and spindle microtubules (Yen et al., 1991; Abrieu et al., 2000). On the other hand, centromeric DNA organization is very divergent among species (for reviews see Murphy and Karpen, 1998; Choo, 2001; Henikoff et al., 2001). Thus, there is no single general theory of how these centromere chromatin components assemble at specific genetic loci in eukaryotes, although the roles of centromeric DNA in some species have been clarified.

In humans, the relationship between centromeric DNA organization and function is complicated. Human alphoid

Address correspondence to Hiroshi Masumoto, Graduate School of Science, Nagoya University, Chikusa-ku, Nagoya 464-8602, Japan. Tel.: 81-52-789-2985. Fax: 81-52-789-5732. E-mail: g44478a@nucc.cc.nagoya-u.ac.jp

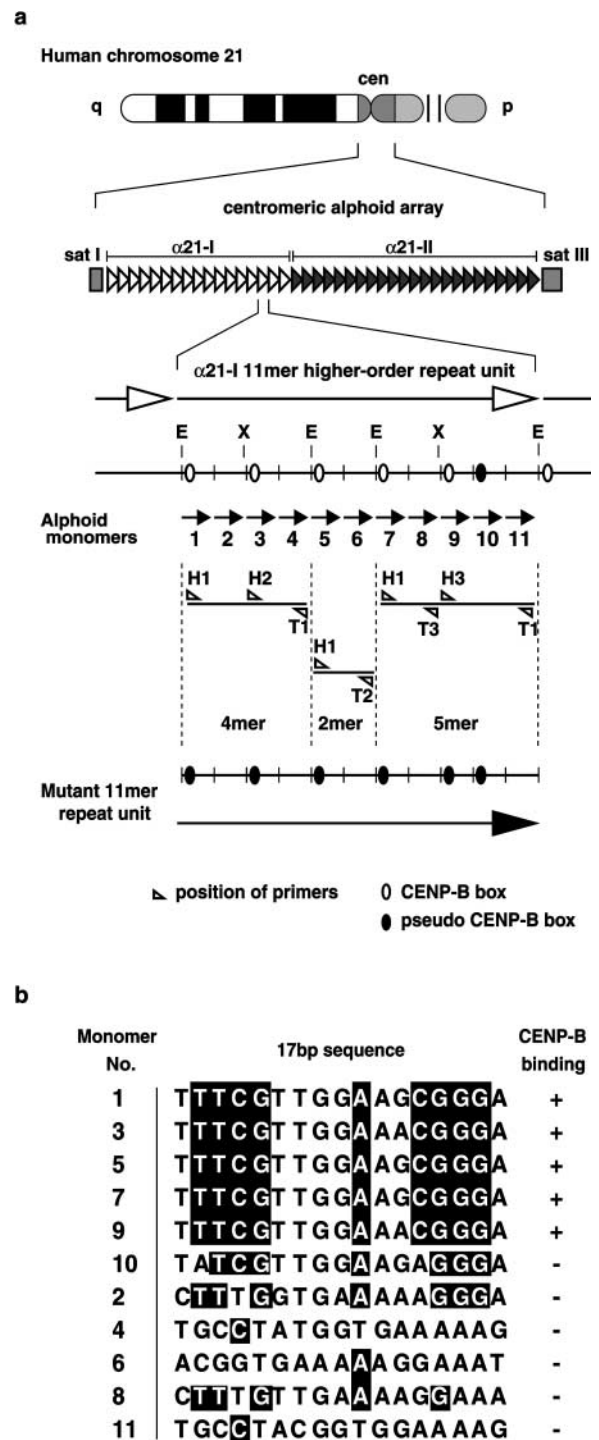
\*Abbreviations used in this paper: BAC, bacterial artificial chromosome; CENP, centromere protein; CHIP, chromatin immunoprecipitation; MAC, mammalian artificial chromosome; MTR, mutant  $\alpha$ 21-I 11mer repeat; WTR, wild-type  $\alpha$ 21-I 11mer repeat; YAC, yeast artificial chromosome.

Key words: CENP-B; CENP-B box; alphoid DNA; MAC; CENP-A

DNA contains a huge repetitive sequence, exists only at the centromeric region, and is found in all human chromosomes (Alexandrov et al., 2001). Alphoid sequences consist of tandem repeats of an AT-rich 171-bp alphoid monomer unit, and some alphoid monomers form chromosome-specific higher-order repeated units (Willard, 1985; for review see Willard and Waye, 1987). The repetitive structure of alphoid DNA can be classified into two types of repeats (Ikeno et al., 1994): units composed of several monomers (type-I alphoid repeat; Fig. 1 a,  $\alpha 21$ -I) and monomeric organization consisting of diverged alphoid monomer units (type-II alphoid repeat; Fig. 1 a,  $\alpha 21$ -II). Centromere components are mainly assembled on type-I alphoid sequences (Ikeno et al., 1994; Ando et al., 2002; Politi et al., 2002). Human artificial chromosome formation is associated only with type-I alphoid sequences (Harrington et al., 1997; Ikeno et al., 1998; Masumoto et al., 1998; Schueler et al., 2001). The CENP-B box appears only in type-I alphoid sequences (Masumoto et al., 1989; Muro et al., 1992; Haaf and Ward, 1994) of autosomes and X chromosomes. However, it has been reported that neocentromeres (a rare phenomenon in which centromeres form on fragmented chromosomes) have no significant centromeric DNA sequences, not even alphoid DNA (du Sart et al., 1997; Lo et al., 2001). Like the Y chromosome, neocentromere-containing chromosomes are stably maintained in cells that undergo mitosis (Tyler-Smith et al., 1999).

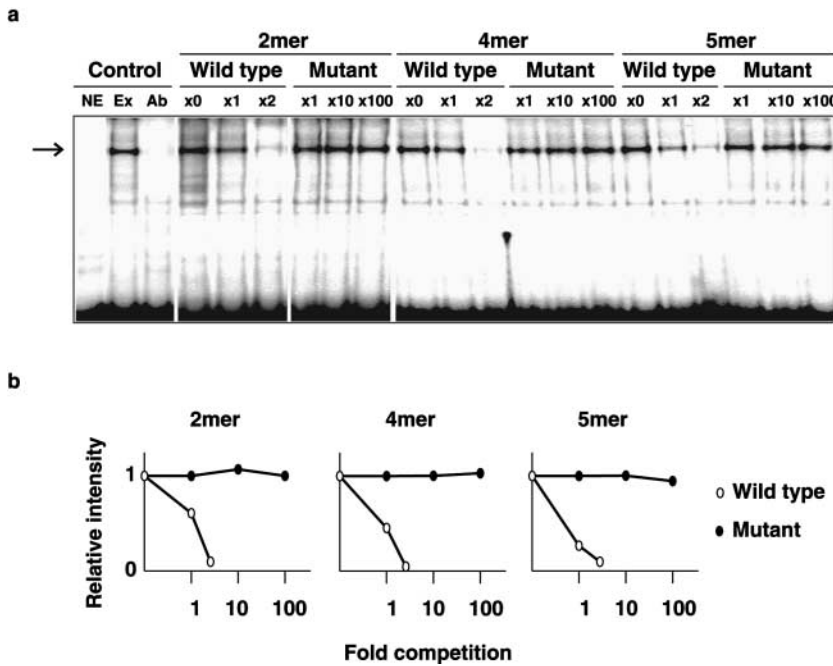
In addition, in CENP-B knockout mouse cells, functional structures of kinetochores are maintained without CENP-B (Hudson et al., 1998). Although in vivo function of CENP-B and the CENP-B box is still unclear, it appears that centromere-specific DNA sequences and CENP-B may not be necessary (at least in chromosome segregation for preexisting functional centromeres). However, in de novo assembly of centromere proteins, assembly capacity significantly depends on a specific DNA sequence: type-I alphoid DNA including the CENP-B box sequence. Studies have shown that the following types of DNA have little or no capacity for de novo centromere formation: Y chromosome alphoid DNA (Harrington et al., 1997; Grimes et al., 2002; Mejia et al., 2002); type-II alphoid DNA (Ikeno et al., 1998); and a neocentromere-specific sequence (Saffery et al., 2001). The fact that CENP-B boxes exist in centromeres of all normal human and mouse autosomes and X chromosomes (i.e., not on the Y chromosome or neocentromeres) suggests that they have functional significance.

In this paper, to determine whether information carried by the type-I alphoid DNA sequence is involved in assembly of human centromeres, we created four kinds of homogeneous synthetic repeated sequences: a type-I alphoid DNA sequence and a nonalphoid sequence, with and without CENP-B boxes. Using the mammalian artificial chromosome (MAC) formation assay (Harrington et al., 1997; Ikeno et al., 1998; Masumoto et al., 1998; Ebersole et al., 2000) and a newly developed competitive chromatin immunoprecipitation (CHIP) assay, we obtained the first known molecular evidence of a functional relationship between de novo centromeric chromatin assembly and centromeric satellite DNA.



**Figure 1. Configuration of alphoid arrays of the chromosome 21 centromere, and construction of the synthetic alphoid repeat unit with modified CENP-B boxes (mutant  $\alpha 21$ -I 11mer).**

(a) Capital letters E and X indicate EcoRI and XbaI cutting sites, respectively. The numeric characters below the  $\alpha 21$ -I 11mer unit indicate monomer numbers. (b) The 17-bp motif of the CENP-B box and its position in  $\alpha 21$ -I 11mer monomer units. The nucleotide sequence shaded black is necessary for CENP-B binding. The 17-bp sequence motif of monomer 10 has two nucleotide variations, and lacks CENP-B binding activity. All five CENP-B box motifs in the mutant  $\alpha 21$ -I 11mer were identical to this monomer 10 motif.



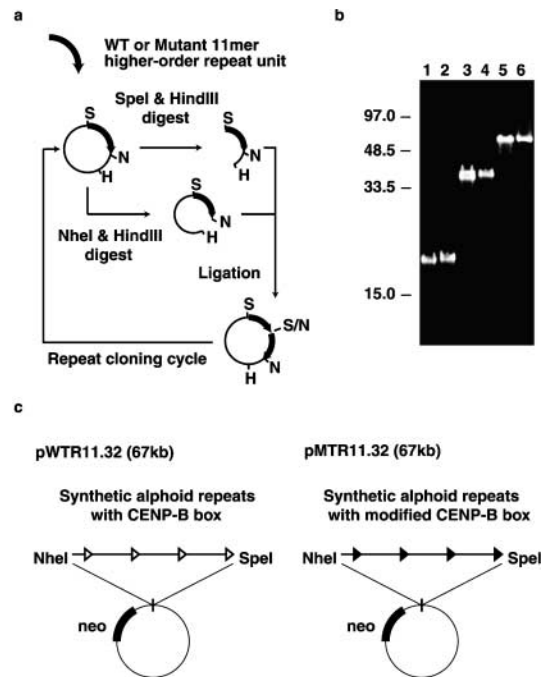
**Figure 2. Gel mobility-shift competition analysis of wild-type and mutant  $\alpha$ 21-I 11mer subunits.** (a)  $\alpha^{32}$ P-CTP-labeled oligonucleotide probe (CB59) containing a CENP-B box sequence was mixed with HeLa nuclear extracts. The mobility shift of the probe with the CENP-B protein dimer complex, indicated by the arrow point (lane Ex), was blocked by preincubation with anti-CENP-B antibodies (lane Ab). Unlabeled PCR products of wild-type and mutant 2-, 4-, and 5mer subunits were used as competitors for the labeled CB59 probe with  $\times 0$ ,  $\times 1$ , and  $\times 2$  and  $\times 1$ ,  $\times 10$ , and  $\times 100$  excess of molecules, respectively. Capital letters indicate no extract (NE), probe with extract (Ex), and probe with extract preincubated with anti-CENP-B antibody (Ab). (b) Relative intensities of shifted bands are plotted in these panels.

**Results**

**Construction of synthetic alphoid repeats with wild-type and mutant CENP-B boxes**

In a previous study, we demonstrated that a yeast artificial chromosome (YAC) containing a 70-kb alphoid DNA fragment from the  $\alpha$ 21-I site of the chromosome 21 centromere, equipped with human telomere sequences, has the capacity for de novo MAC and centromere chromatin assembly in cultured human cells. However, a YAC containing a 100-kb alphoid DNA fragment from the  $\alpha$ 21-II site (adjacent to the  $\alpha$ 21-I region) does not have that capacity (Ikeno et al., 1998). DNA organization of the 70-kb alphoid insert from  $\alpha$ 21-I was a simple homogeneous repeat of an  $\alpha$ 21-I-specific 11 alphoid monomer higher-order repeating unit ( $\alpha$ 21-I 11mer) with five CENP-B boxes. In contrast, DNA organization of the  $\alpha$ 21-II fragment was a heterogeneous repeat of diverged alphoid monomers with no CENP-B boxes (Ikeno et al., 1994). Therefore, we hypothesized that either the homogeneous repeat organization or CENP-B box sequences are necessary for de novo centromere formation. To confirm this hypothesis using a MAC-forming assay, we constructed an  $\alpha$ 21-I 11mer repeat unit with modified CENP-B boxes. This construct had no CENP-B binding activity.

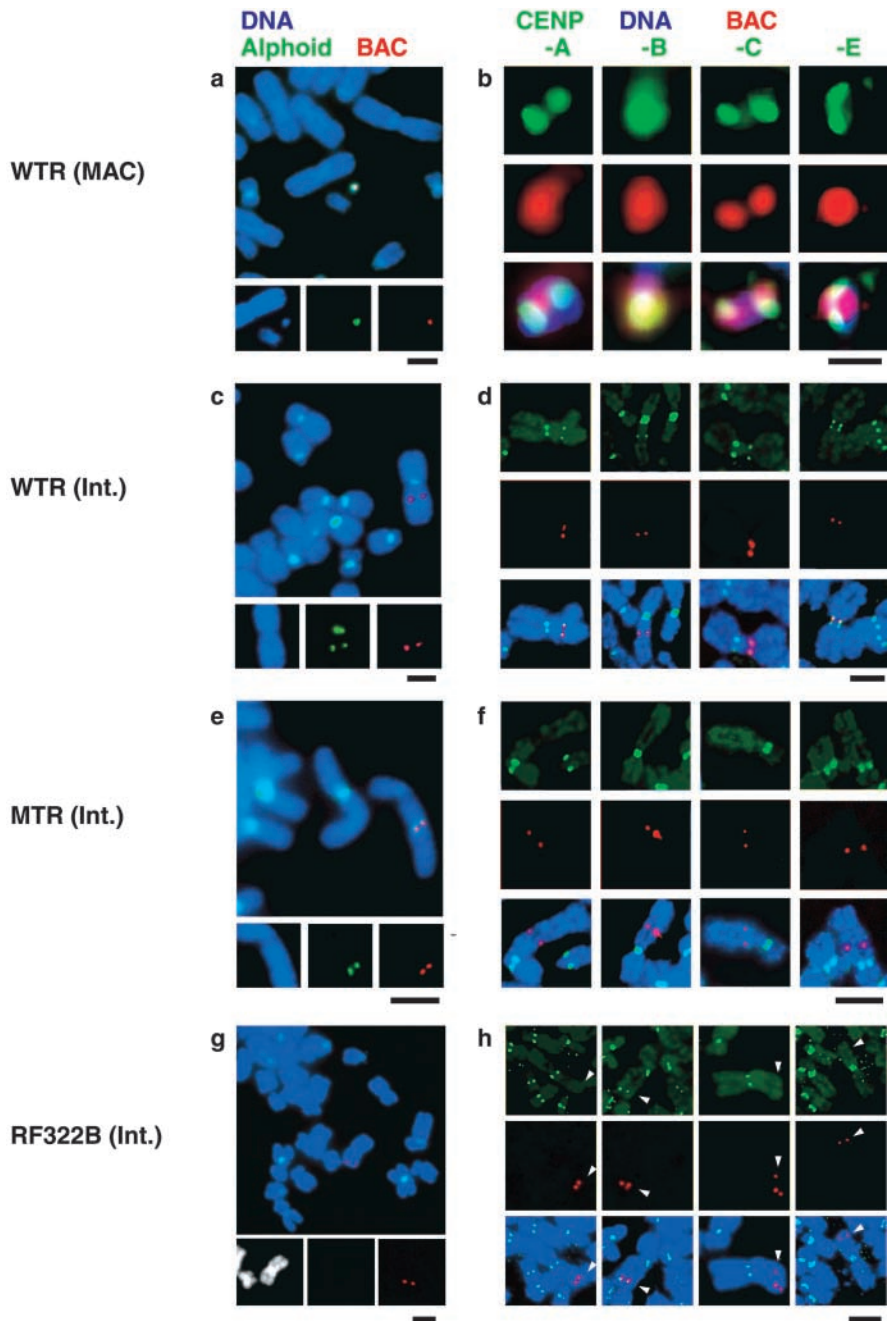
We divided the  $\alpha$ 21-I 11mer unit into 4mer, 2mer, and 5mer fragments by treating it with the restriction endonuclease EcoRI. Using PCR, we introduced a two-nucleotide substitution into all five active CENP-B box sequences, making them all identical to an inactive sequence at the CENP-B box position of the number 10 alphoid monomer unit (Fig. 1, a and b; see Materials and methods). Loss of interaction between CENP-B protein and the fragments with modified CENP-B boxes was confirmed using a gel mobility shift competition assay. The 2mer, 4mer, and 5mer fragments with modified CENP-B boxes did not compete with labeled DNA containing the CENP-B box for complex formation with CENP-B, even with a 100-fold



**Figure 3. Construction of synthetic alphoid repeats with wild-type and mutant  $\alpha$ 21-I 11mer repeat units.** (a) Schematic diagram of tandem concatenation of wild-type and mutant  $\alpha$ 21-I 11mer repeat units. (b) Pulse field gel electrophoresis of tandemly concatenated  $\alpha$ 21-I 11mer repeat arrays. Lanes 1, 3, and 5, endonuclease HindIII-cut BAC plasmids containing 8, 16, and 32 copies of extended wild type  $\alpha$ 21-I 11mer repeat units, respectively. Lanes 2, 4, and 6, BAC plasmids containing 8, 16, and 32 copies of extended mutant  $\alpha$ 21-I 11mer repeat units, respectively. (c) Circular BAC constructs of pWTR11.32 and pMTR11.32 have a common vector sequence containing a neomycin resistance gene. The only differences between the two plasmids are two nucleotide substitutions at each CENP-B box.

excess of these fragments (Fig. 2). Then, the fragments were reintegrated to form a modified  $\alpha$ 21-I 11mer repeating unit (mutant  $\alpha$ 21-I 11mer).

**Figure 4. Examples of extrachromosomal events and centromere protein assembly on introduced DNAs.** In all panels, DNA was counterstained with DAPI (blue), and introduced BAC DNA was detected by FISH with a fluorescence-labeled BAC probe (red). (a, c, e, and g) Chromosomal events detected by FISH. Green signals indicate  $\alpha 21$ -I 11mer DNA probe. (b, d, f, and h) Centromere protein assembly on the BAC DNA was analyzed with indirect immunofluorescent staining combined with FISH. Green signals indicate the indirect immunofluorescence of centromere proteins (CENP-A, -B, -C, and -E). In panel h, white arrowheads indicate the integration site of pRF322B.192. The following data from cell lines were used as examples: W0210R-8 (a and b), W0210R-1 (c and d), M1319 (e and f), and B0106 (g and h). Bars: (a and c–h) 5  $\mu\text{m}$ ; (b) 1  $\mu\text{m}$ .



In the MAC formation assay, input alphoid DNA of  $\sim 30$  kb clearly exhibited less MAC formation activity than 50-kb or 70-kb alphoid DNA (unpublished data). We therefore constructed a tandem array of 32 wild-type  $\alpha 21$ -I 11mer repeats (WTR, 60 kb) or mutant  $\alpha 21$ -I 11mer repeats (MTR, 60 kb) on a bacterial artificial chromosome (BAC) (pWTR11.32 and pMTR11.32, respectively) (Fig. 3), using a modified version of a method described by Harrington et al. (1997).

#### MAC-forming assay with synthetic $\alpha 21$ -I 11mer repeats containing active CENP-B boxes

MAC formation efficiency of circular constructs containing  $\alpha 21$ -I alphoid DNA without a telomere sequence is reported similarly to that of YAC-based linear constructs containing

$\alpha 21$ -I alphoid DNA with telomere sequences (Ikeno et al., 1998; Ebersole et al., 2000). We introduced BAC plasmid DNA containing synthetic alphoid arrays (pWTR11.32 or pMTR11.32) into HT1080 cells by lipofection, and then analyzed MAC formation efficiency of the resulting G418-resistant cell lines by FISH using  $\alpha 21$ -I 11mer alphoid and BAC DNA probes. High MAC formation efficiency as a predominant effect was observed in 12 of the 27 cell lines created by transfection with pWTR11.32 (Fig. 4 a; Table I). No host chromosomal sequences were detected on the pWTR11.32-derived MACs by FISH analysis with inter- and intra-Alu probes, and pan-alphoid probes competed with excess nonlabeled  $\alpha 21$ -I 11mer alphoid DNA, although these probes had a high affinity for arms and centromeres of host chromosomes (unpublished data). These results indicate that the ob-

Table I. Chromosomal events in the stable cell lines with synthetic 11mer repeats

Synthetic repeats inserted in BAC (Input DNA)	Experiments	Analyzed cell lines	Chromosomal events of introduced DNA <sup>a</sup>	
			De novo artificial chromosome	Integration into host chromosome (centromeric heterochromatin/arm)
Alphoid repeats with functional CENP-B boxes (60 kb) (pWTR11.32)	1st	15	7	8 (7/1)
	2nd	12	5	7 (6/1)
Alphoid repeats with modified CENP-B boxes (60 kb) (pMTR11.32)	1st	20	0	20 (9/11)
	2nd	18	0	18 (7/11)
RF322 repeats with functional CENP-B boxes (69 kb) (pRF322B.192)	1st	14	0	14 (5/9)
	2nd	21	0	21 (3/18)
RF322 repeats without CENP-B boxes (69 kb) (pRF322L.192)	1st	11	0	11 (3/8)
	2nd	11	0	11 (2/9)

<sup>a</sup>Chromosomal events were determined according to the predominant pattern (>50%) of introduced DNA containing synthetic  $\alpha$ 21-I 11mer alphoid repeat units or synthetic RF322 repeats, using FISH analysis of chromosome spreads. These patterns were counted in cases in which all three signals, DNA (DAPI), BAC, and  $\alpha$ 21-I alphoid signals, were colocalized. A  $\chi^2$  test of the predominant chromosomal events between pWTR11.32 and other cell lines with introduced plasmids showed highly significant P values (<0.0002).

served MAC formation occurred via a de novo mechanism without acquiring host centromeres.

Copy number of the BAC DNA in the cell lines containing MACs was analyzed. Quantitative competitive PCR analyses (Wang et al., 1989) showed that the neomycin resistance gene and synthetic alphoid sequence of the initial BAC construct were amplified 10–29 times and 14–36 times, respectively, in a single MAC. This BAC copy number corresponded well to the DNA content of the MAC, estimated by quantification of the fluorescence intensity of propidium iodide staining (Table II). In the 15 remaining pWTR11.32-derived cell lines, integration into host chromosomes occurred as a predominant event (Fig. 4 c; Table I).

#### No MAC formation for synthetic $\alpha$ 21-I 11mer repeats with modified CENP-B boxes

In contrast, MAC formation was not observed in any of the cell lines obtained by transfection of the BAC plasmid containing the  $\alpha$ 21-I 11mer repeats with mutated CENP-B boxes (pMTR11.32). All the input DNAs were integrated into host chromosomes with a level of BAC and alphoid amplification that was similar to, or higher than, that of the wild-type alphoid (38/38 cell lines; Fig. 4 e; Table I). The mutant sequence completely lacked the capacity for MAC assembly (0%).

Comparing the results obtained for pWTR11.32 and pMTR11.32, it is clear that binding of CENP-B to the CENP-B boxes in alphoid DNA is important for MAC formation. Although it has been suggested that input DNA multimerization, coupled with MAC formation, is caused by binding of CENP-B to CENP-B boxes (Ebersole et al., 2000), the pMTR11.32 plasmid was multimerized at a level similar to that of the pWTR11.32 plasmid and a control BAC plasmid containing only the neomycin resistance gene without alphoid DNA insert (unpublished data). We did not detect any significant relationships between CENP-B binding activity and input DNA amplification.

#### Centromere protein assembly on the MAC and MAC stability in cell divisions

We assessed whether CENP-B and the essential centromere components CENP-A, -C, and -E were assembled on intro-

duced DNA, by simultaneous detection of indirect immunofluorescent staining and FISH on metaphase chromosomes using specific antibodies and a digoxigenin-labeled BAC DNA probe. In all analyzed cell lines, assembly of all four of these centromere-specific proteins was detected on all the MACs in metaphase spreads (Fig. 4 b; Table II).

Next, using FISH analysis, we analyzed mitotic stability of these MACs after 60 d of culture without the selective drug G418. The MACs were very stable for an extended period, with a loss rate per cell division of <0.18% of a copy in most cases (Table II). Thus, we conclude that these MACs obtained a functional centromere. MAC formation and centromere protein assembly were dependent on the existence of CENP-B box sequences in alphoid DNA.

#### Variegated assembly of centromere proteins on ectopically integrated synthetic $\alpha$ 21-I 11mer repeats

To confirm that MAC formation and centromere protein assembly are dependent on CENP-B binding activity, we analyzed the distribution of centromere proteins on the ectopic pWTR11.32 and pMTR11.32 integration sites. Two cell lines in which pWTR11.32 integrated into a host chromosomal arm (2/27 pWTR11.32 cell lines) were used for the wild-type sequence. We recloned these cell lines in order to obtain cell lines with homogeneous ectopic integration of pWTR11.32. In these subcloned cell lines, 8–38% of cells exhibited partial assembly of CENP-A, -B, -C, and -E at the ectopic integration sites of pWTR11.32, and the CENP-A, -C, and -E signals at the ectopic sites always colocalized with CENP-B in all analyzed chromosome spreads (Fig. 4 d; Table II).

Our subcloned cell lines containing ectopically integrated pWTR11.32 exhibited a variegated assembly pattern similar to that observed for pseudo-dicentric and dicentric chromosomes (Earnshaw et al., 1989; Wandall, 1994). Although the molecular mechanisms are still unclear, a phenomena called centromere inactivation, which is known to be one of the epigenetic mechanisms of centromere regulation, may involve this variegated assembly (Fisher et al., 1997; unpublished data).

In contrast, in cell lines with ectopic integration of alphoid repeats with mutant CENP-B boxes (pMTR11.32) into host

Table II. Centromere protein assembly on introduced DNA

Input DNA (Chromosomal event)	Cell line (MAC loss rate <sup>a</sup> )	Spread	Centromere protein assembly <sup>b</sup> (%)				Multiplicity of BAC input DNA			
			CENP-A	CENP-B	CENP-C	CENP-E	Synthetic aliphoid	Neo	PI intensity (DNA amount) <sup>c</sup>	
pWTR11.32 (artificial chromosome)	W0203 (0.00175) <sup>d</sup>	40	100	100	100	100	14	16	18 (1.2 Mb)	
	W0206 (ND) <sup>d</sup>	50	100	100	+	+	26	10	30 (2.0 Mb)	
	W0210R-8 (0.00105) <sup>d</sup>	50	100	100	+	+	23	16	24 (1.6 Mb)	
	W0212 (<0.001) <sup>d</sup>	50	100	100	+	+	36	29	39 (2.6 Mb)	
	(ectopic integration)	W0210R-1	50	24	38	12	22	17	14	
		Colocalization with CENP-B <sup>e</sup>	50	100	–	100	100			
		W1203R-8	50	10	22	24	8	20	7	
Colocalization with CENP-B <sup>e</sup>	25	100	–	100	100					
pMTR11.32 <sup>f</sup> (ectopic integration)	M1319	50	0	0	0	0	19	12		
	M1318	50	0	0	0	0	43	29		
pRF322B.192 <sup>g</sup> (ectopic integration)	B0104	50	0	12	0	0	20	16		
	B0106	30	0	8	0	0	48	36		
pRF322L.192 (ectopic integration)	L0102	50	0	0	0	0	42	38		
	L0203	50	0	0	0	0	38	26		

<sup>a</sup>MAC loss rate (R) was calculated using the following formula:  $N_{60} = N_0 \times (1 - R)^{60}$ .  $N_0$  and  $N_{60}$  are the rates for MAC-containing cells in these cell lines at the time points of 0 and 60 d, respectively. During the 60 d, cells were cultured in the absence of G418, a selective drug.

<sup>b</sup>Percentage of centromere protein assembly was assessed based on colocalized signals from indirect immunofluorescence of centromere proteins and FISH analysis of BAC DNA probe.

<sup>c</sup>See Materials and methods.

<sup>d</sup>FISH analysis using inter-/intra-Alu, other aliphoid, and Sat III probes indicated that all MACs from the four cell lines in this table were negative.

<sup>e</sup>Colocalization with CENP-B was determined from indirect immunofluorescence of CENP-B and other centromere proteins.

<sup>f</sup>In addition to the results for these two cell lines, results for another four cell lines obtained by transformation with pMTR11.32 also showed no centromere protein assembly on introduced DNA sites.

<sup>g</sup>In another four cell lines obtained by transformation with pRF322B, only CENP-B assembly was observed.

chromosomal arms, no CENP-B signal was observed at the integration sites. Furthermore, there was no assembly of functional centromere components CENP-A, -C, or -E at the ectopic sites (six cell lines were analyzed) (Fig. 4 f; Table II).

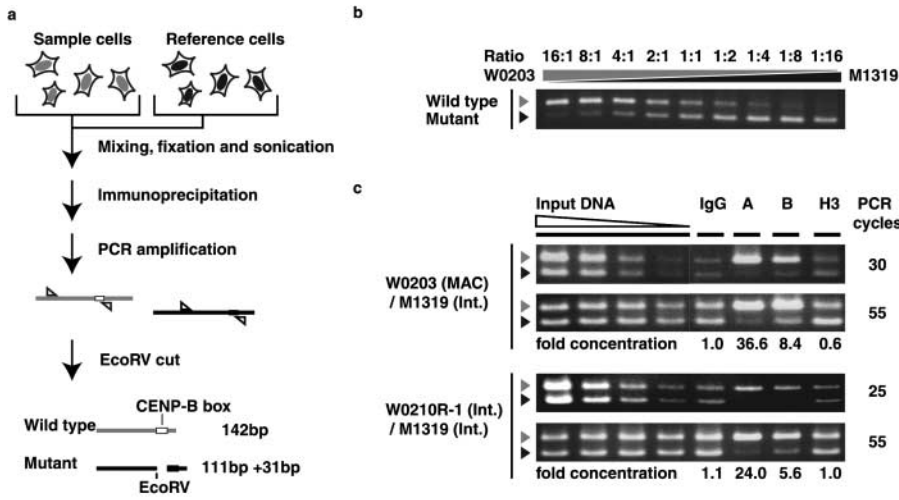
### Direct assembly of CENP-A and CENP-B on synthetic aliphoid sequences

In this study and previous studies, using cytological observation, we demonstrated that centromere protein assembly and MAC formation occur on the introduced aliphoid DNA via a de novo mechanism without acquiring host centromeres. However, it was not easy to detect direct assembly of centromere components on the introduced aliphoid DNA at the molecular level because of the difficulty in distinguishing the introduced aliphoid DNA from endogenous aliphoid DNA sequences.

In the present study, we used synthetic aliphoid DNA to generate artificial chromosomes. We designed a specific primer set that discriminates our synthetic DNA (pWTR11.32 and pMTR11.32) from endogenous aliphoid sequences. We then tested whether the centromeric chromatin components CENP-A and CENP-B assembled on WTR DNA directly, using the CHIP assay and competitive PCR detection (Becker-Andre and Hahlbrock, 1989). In our CHIP and competitive PCR assays, in order to standardize individual preparations and PCR for each sample, we used as an internal reference pMTR11.32-transformed cells

(M1319) that did not exhibit centromere protein assembly at the synthetic aliphoid integration site. We mixed the internal reference cells with sample cells, pWTR11.32-transformed cells (W0203) containing MACs.

Before starting the CHIP analyses, to examine the linearity of our competitive PCR assay as a control experiment, W0203 and M1319 cells were mixed at different ratios of cell number (from 16:1 to 1:16), and genomic DNAs were prepared from these mixtures. Using the competitive PCR assay with a specific primer set, we confirmed that synthetic aliphoid DNA fragments from WTR11.32 and MTR11.32 were amplified parallel to each other and competitively even after the DNA preparation, maintaining the initial ratio of WTR and MTR fragments (Fig. 5, a and b). Then, a WTR11.32 and MTR11.32 cell mixture (1:1) was fixed, sonicated, and immunoprecipitated in a single tube and used as competitive PCR start material. The WTR11.32 DNA fragment was immunoprecipitated ~37 times and eight times more efficiently than the reference MTR11.32 DNA fragment with anti-CENP-A and anti-CENP-B antibodies, respectively; whereas these two synthetic aliphoid fragments were nonspecifically immunoprecipitated at the same ratio to the input DNA with normal IgG (Fig. 5 c). However, when anti-histone H3 antibody was used, the ratio of immunoprecipitated WTR11.32 fragment to MTR11.32 decreased to ~60%. These results indicate that considerable numbers of nucleosomes on



**Figure 5. CHIP analysis of centromere protein assembly on introduced alphoid DNA.** (a) pWTR11.32-transformed cells (W0203 or W0210R-1) were mixed with reference pMTR11.32-transformed cells (M1319). (b) EcoRV-digested PCR products from genomic DNA templates. W0203 and M1319 cells were mixed at different ratios of cell number (from 16:1 to 1:16), and genomic DNAs for competitive PCR were prepared from these mixtures. Gray arrowheads indicate the 142-bp PCR fragment from WTR. Black arrowheads indicate the 111-bp PCR fragment from MTR. (c) EcoRV-digested PCR products from immunoprecipitated DNA template. Left four lanes show PCR products from diluted input DNA at  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ , and  $10^{-4}$  dilution, from left to right. PCR products

from immunoprecipitated DNAs are shown in the right four lanes. Capital letters indicate antibodies used in immunoprecipitation: normal mouse IgG (IgG), anti-CENP-A (A), anti-CENP-B (B), and anti-histone H3 (H3). Two different rates of amplification, logarithmic amplification state (25–30 cycles) and plateau state (55 cycles), are shown in the top and bottom panels, respectively. DNA samples comprised 75% and 15% of total PCR products in the top and bottom panels, respectively. Relative concentrations of the WTR DNA fragment versus the MTR DNA fragment, as indicated by CHIP assay using each antibody, are shown below the panels.

WTR11.32 DNA on the MAC were replaced with nucleosomes containing CENP-A.

In the mixture of cell lines in which either WTR11.32 or MTR11.32 was integrated ectopically into the host chromosome arms, the WTR11.32 DNA fragment was still immunoprecipitated efficiently but at a reduced rate with anti-CENP-A and anti-CENP-B antibody. This is consistent with the cytological observation of partial assembly of CENP-A, -B, -C, and -E at the sites of ectopic alphoid integration. Thus, we conclude that the centromere chromatin components CENP-A and CENP-B were directly assembled on the WTR synthetic alphoid DNA sequence but not on the MTR alphoid sequence; i.e., assembly of these centromere chromatin components was dependent on the CENP-B binding capacity of alphoid DNA.

### No MAC formation and no centromere chromatin assembly on nonalphoid repeated sequences containing CENP boxes

Using two kinds of synthetic  $\alpha$ 21-I 11mer alphoid repeats (~60% AT rich), we demonstrated that the CENP-B box sequence is required for high-efficiency de novo centromeric chromatin assembly. However, this raises the question of whether the CENP-B box is sufficient for de novo assembly without depending on the AT-rich alphoid sequence. The AT-rich trend is common in many eukaryotic centromere DNAs (Lo et al., 2001; for reviews see Choo, 1997b; Koch, 2000; Henikoff, 2002), therefore, suggested to be one of possible properties for efficient centromeric chromatin assembly. However, our results clearly showed that the synthetic AT-rich  $\alpha$ 21-I 11mer alphoid repeats with inactive CENP-B boxes alone could not induce any assembly of centromere chromatin (Fig. 4 f; Fig. 5 c; Table II). If the CENP-B box is sufficient, de novo centromere formation may occur even on GC-rich spacer sequences. We thus tested whether the CENP-B binding site induces de novo centromeric chromatin assembly on an exogenous nonal-

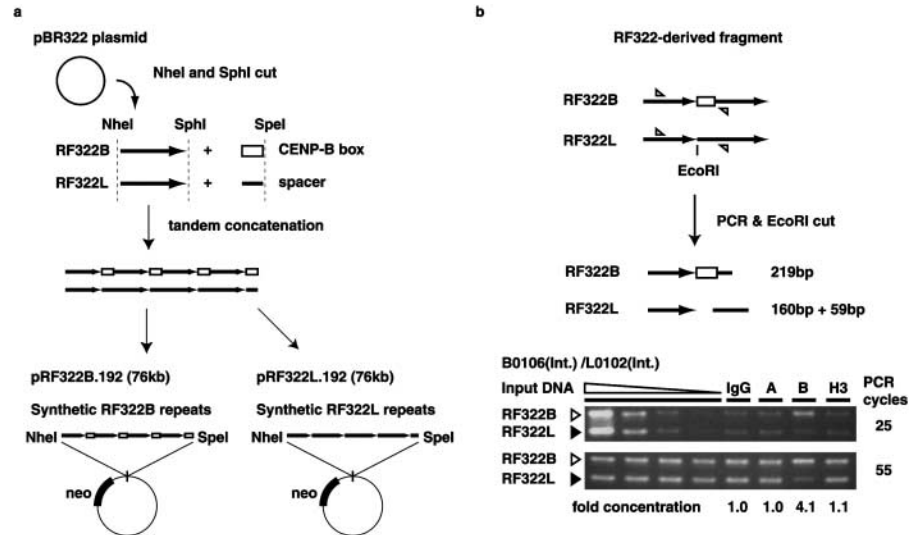
phoid GC-rich (~60%) sequence, the 340-bp NheI-SphI restriction fragment (corresponding to the length of alphoid dimer) from plasmid pBR322 (RF322). CENP-B boxes appear on every other monomer unit (~170 bp  $\times$  2) in the  $\alpha$ 21-I 11mer alphoid unit, and we joined the RF322 fragment to a CENP-B box or a spacer sequence at the SphI site. These DNA constructs, with CENP-B boxes (RF322B) or without CENP-B boxes (RF322L), were cloned into the pBACNX vector between the NheI and SpeI sites, and were concatemerized tandemly to a length of 69 kb (pRF322B.192 and pRF322L.192, respectively), as described above (Fig. 3 a; Fig. 6 a).

In the 35 cell lines obtained by transfection of pRF322B.192 BAC, FISH analysis showed that no MAC formation occurred and that all input DNAs were integrated into host chromosomes (Table I). Competitive PCR assays showed that input DNA multiplicity of pRF322B.192 BAC was similar to that of the synthetic  $\alpha$ 21-I 11mer repeat DNAs (Table II). In all 22 cell lines derived from pRF322L.192 BAC transfection, there was no MAC formation, and input DNA was integrated into host chromosomes.

Using FISH analysis and simultaneous detection of indirect immunofluorescent staining on metaphase chromosomes, we detected no assembly of the essential centromere component proteins CENP-A, -C, or -E at the integration sites of synthetic RF322B. Only CENP-B protein was detected at the RF322B sites (Fig. 4, g and h; Table II). At the RF322L integration sites, no signals were detected for any of these four centromere proteins (Table II). In competitive CHIP analysis, significant relative concentration of the RF322B fragment versus RF322L was only detected with anti-CENP-B antibody; no significant relative concentration was observed with anti-CENP-A or anti-histone H3 antibodies (Fig. 6 b). Thus, we concluded that interaction between CENP-B and the CENP-B box is not sufficient for de novo centromere chromatin assembly on the nonalphoid

### Figure 6. Construction of nonaliphoid sequences with or without CENP-B boxes and CHIP analysis.

(a) Schematic diagram of the construction of nonaliphoid repetitive sequences containing CENP-B boxes (RF322B) and not containing CENP-B boxes (RF322L). pBR322-derived fragment (RF322) was ligated to the CENP-B box (RF322B) or spacer oligonucleotide sequence (RF322L). The resulting fragments were tandemly concatenated using the same method used to construct the  $\alpha$ 21-I 11mer aliphoid repeat unit. (b) Competitive CHIP analysis of a pRF322B.192-transformed cell (B0106) using a pRF322L.192-transformed cell (L0102) as reference. Using the specific primer set, both synthetic RF322 fragments were amplified with competitive PCR, as in Fig. 5. The PCR products were separated by agarose gel electrophoresis after endonuclease EcoRI treatment. The ratio of concentration of the RF322B fragment to that of the RF322L fragment was calculated and is indicated below the panels.



GC-rich sequence, although it is required for centromere chromatin assembly on aliphoid DNA.

## Discussion

### Reverse genetic approach of aliphoid DNA using MAC assay

Recently, using human centromeric satellite DNAs, Harrington et al. (1997) and our groups succeeded in inducing de novo centromere assembly on naked input DNA molecules in cultured human cells and succeeded in constructing a first-generation human artificial chromosome (Ikeno et al., 1998). However, this de novo-created MAC has some unsolved problems, compared with a YAC created from an identical molecule of introduced DNA into yeast cells (Murray and Szostak, 1983). Aliphoid YAC DNA introduced into cultured human cells multimerized to generate a single MAC. Although MAC formation efficiency was high (>50% of metaphase cells showed MAC signals in 11 of 30 transformed cell lines; Ikeno et al., 1998; Masumoto et al., 1998), MAC formation did not occur in 100% of transformed cells. However, these first-generation MACs have a crucial advantage in that they consist entirely of introduced DNA molecules (Ebersole et al., 2000; Mejia et al., 2001; Grimes et al., 2002). This feature makes reverse genetic analysis (mutational analyses) of human centromeric satellite DNA possible.

In the present study, using the MAC formation assay, we analyzed the relationship between de novo centromere chromatin assembly and the sequences of centromeric DNA, aliphoid DNA, and CENP-B boxes. Using synthetic repetitive DNAs, we demonstrated that only aliphoid DNA and CENP-B boxes, but no other sequences from the human chromosomes, are required for de novo centromere assembly and MAC formation. These results not only provide important information for constructing human artificial chromosomes but also provide evidence supporting the theory that CENP-B–CENP-B box interaction is involved in the cen-

tromere assembly mechanism. Also, these results indicate that there is an element or a character in  $\alpha$ 21-I 11mer aliphoid DNA, other than the CENP-B box, that is important for de novo centromere formation. Further analysis using reverse genetic methods (mutational analyses) such as the MAC formation assay may be needed to identify this important element. MAC technology has gradually been advancing through clarification of relationships between DNA sequence elements and chromosome functions. In the near future, an ideal human artificial chromosome with full chromosomal functions (like YACs) may be created.

### Selective assembly of CENP-A chromatin on introduced synthetic aliphoid DNA with CENP-B binding capacity

In this study, using cytological analyses, we demonstrated that the functional centromere components CENP-A, -C, and -E assemble on synthetic aliphoid DNA fragments, and that CENP-B binding capacity of aliphoid DNA sequences has a major effect on this assembly. Moreover, using a CHIP-based competitive PCR assay, we demonstrated direct assembly of CENP-A on synthetic aliphoid DNA that contained functional CENP-B boxes. In contrast, at integration sites of an artificial repeated sequence based on RF322 (derived from pBR322), no CENP-A chromatin assembly was detected, although CENP-B assembly was detected on sequences that contained CENP-B boxes. Therefore, we hypothesize that there are important elements or characters in the  $\alpha$ 21-I 11mer aliphoid DNA sequence that aid in stabilizing or assembling centromere chromatin, including CENP-A chromatin.

As described in the Results, AT-rich regions of aliphoid DNA may contain such important elements or characters. Although CENP-A folded DNA sequences are not conserved among centromeres of eukaryotes, these sequences tend to be AT rich (AT content is usually at least 60%) in many eukaryotes, including yeast and humans (for reviews see Choo, 1997b; Koch, 2000; Henikoff, 2002). No sequence conservation is observed between human aliphoid DNA and mouse



centromeric minor satellite DNA (120-bp repeats), except CENP-B boxes and AT-rich nucleotide composition. We speculate that this AT-rich nucleotide composition may facilitate CENP-A nucleosome folding. Another possibility, that the relationship between regular higher-order structure and nucleosomal phasing pattern is affected by type-I alphoid DNA sequence and CENP-B/CENP-B box function, has previously been suggested (Muro et al., 1992; Romanova et al., 1996; Yoda et al., 1998). It is possible that compartmentalization of alphoid DNA by CENP-B–CENP-B box interaction restricts contact between, for example, the CENP-A–histone H4 heterotetramer (or histone H3/H4) and alphoid DNA (Sullivan et al., 1994; Shelby et al., 1997; Ando et al., 2002), and that this restriction of alphoid DNA is the basis of selectivity for CENP-A–histone H4 tetramer incorporation. A third possibility is that alphoid DNA contains binding sites for other centromere component proteins. We have detected protein–protein interactions between CENP-B and CENP-C using yeast two-hybrid analysis and an immunoprecipitation assay with HeLa chromatin (unpublished data). It has been reported that CENP-C exhibits nonspecific DNA binding activity (Yang et al., 1996), and that CENP-C associates with type-I alphoid DNA in vivo (Politi et al., 2002). This suggests that CENP-B and CENP-C play a cooperative functional role in de novo centromere formation, and that there is specific interaction between the kinetochore and type-I alphoid DNA containing CENP-B boxes.

### Absence of CENP-B boxes from the Y centromere and neocentromere

CENP-B or CENP-B box has not been detected in the alphoid DNA on the Y chromosome or the neocentromere-specific DNA sequence. In CENP-B knockout mouse cells, functional structures of kinetochores are maintained without CENP-B (Hudson et al., 1998). How do these centromeres maintain their activity without de novo assembly? It might be explicable with a functional redundancy. On the other hand, the epigenetic character of centromeres has been well described. The other possibility is that once established, the centromere activity is maintained by the chromatin assembly mechanism (Williams et al., 1998; for reviews see Murphy and Karpen, 1998; Wiens and Sorger, 1998). However, in all normal human chromosomes, functional centromere structures are formed and maintained on the alphoid DNA arrays. This study, indeed, provides evidence of a molecular link between a centromere-specific DNA and centromeric chromatin assembly in humans. The mechanisms that specify the centromere location in each species are not straightforward; thus, further investigations on centromere DNA structures and functions and its molecular evolution correlating both with the epigenetic chromatin assembly and the de novo assembly are needed (Laurent et al., 1999; Alexandrov et al., 2001; Schueler et al., 2001; for review see Henikoff, 2002).

## Materials and methods

### Construction of $\alpha$ 21-I 11mer alphoid units with modified CENP-B boxes

Fragments with modified CENP-B boxes were generated by PCR using EcoRI-digested  $\alpha$ 21-I 11mer (p11-4) subunit DNA (2-, 4-, and 5mer) and

the following primers: T1, 5'-ATACTCGGGAATTCAAATAAAAGGTAGACAGC-3'; T2, 5'-ATACTCGGGAATTCGTATAAAAAGGTAGACAGC-3'; H1, 5'-TATCCCGAGAATTCCCTCTTCCAACGATATCCTC-3'; H2, 5'-CTGTCTAGATTTTATGATGATATTCCTTTTCCAACGATATCGTTAGAGC-3'; T3, 5'-AAAATCTAGACAGAAGCACTATTAG-3'; and H3, 5'-CATGTATCCGCTAGCTTTTATGATGATATTCCTTTTCCAACGATATCGTTAGAGC-3'. The mutant 4mer sequence was constructed using two-step mutagenesis (Fig. 1 a). First, PCR product obtained using H1 and T1 primers was subcloned into the cloning vector pUC119. Next, the second half of this insert, including the wild-type CENP-B box, was replaced with H2 and T1 PCR product by using EcoRI and XbaI sites. The mutant 2mer was produced by single-step PCR using H1 and T2 primers. Using XbaI and NheI cohesive ends, the mutant 5mer was constructed from two PCR fragments: one amplified with H1 and T3 primers, and the other amplified with H3 and T1 primers. After confirmation of sequence and loss of CENP-B binding activity, these fragments were re-integrated at EcoRI sites to form  $\alpha$ 21-I 11mer repeating units with mutated CENP-B boxes.

### Tandem concatenation of synthetic alphoid repeat units

We subcloned wild-type and modified  $\alpha$ 21-I 11mer repeat units into a BAC vector (pBACNX), which includes a neomycin resistance gene, between NheI and SpeI endonuclease sites. Plasmid DNA with the inserted  $\alpha$ 21-I 11mer unit was digested using two different endonuclease sets: NheI-HindIII and SpeI-HindIII. When the two resulting fragment types were ligated, the inserted fragments were concatenated tandemly, with no change in the physical relationships between the inserts and the endonuclease sites. Repeating this cycle of cutting and ligation resulted in tandem multiple repeat arrays of the inserted fragments (Fig. 3 a).

### Construction of nonalphoid repeats

A 340-bp NheI-SphI restriction fragment from the pBR322 plasmid (RF322) was ligated at the SphI site to the CENP-B box sequence, GCATGCGATATCTTCGTTGGAAACGGGACTAGT (RF322B), or a spacer oligonucleotide sequence, GCATCGGAATTCCTGTATATAAAACCAGACTAGT (RF322L). The two resulting DNA constructs were cloned into the pBACNX vector between the NheI and SpeI sites, and were concatenated to form 69-kb tandem repeats (pRF322B.192 and pRF322L.192, respectively), as shown in Fig. 3 a and Fig. 6 a.

### Gel mobility shift competition assay

Gel mobility shift competition assays were performed using the radiolabeled CB59 probe, which includes a single CENP-B box sequence in a 59-bp oligonucleotide (Muro et al., 1992). Fragments of wild-type and modified  $\alpha$ 21-I 11mer subunits were generated for the competition assay by PCR, using plasmids containing cloned wild-type or mutant 2-, 4-, and 5mer fragments with M13 forward and reverse primers (TAKARA). More details are provided in a previous report (Masumoto et al., 1989; Yoda et al., 1992).

### Cell culture and transformation

HT1080 human fibroblast cells were cultured in DME supplemented with 10% (vol/vol) FCS, penicillin, streptomycin, and L-glutamine at 37°C in a 5% CO<sub>2</sub> incubator. Introduced BAC DNAs were purified using a QIAGEN large construction kit. Using 4.5  $\mu$ l of Lipofectamine reagent (GIBCO BRL), 0.4  $\mu$ g of purified DNA was transfected into 80% confluent HT1080 cells in 3.5-cm dishes.

### Cytological detection

Templates of in situ hybridization probes were prepared by PCR. Primers for p11-4, pan-alphoid, and intra-/inter-Alu sequences are described in our previous reports (Ikeno et al., 1998; Masumoto et al., 1998). The BAC vector template was generated by PCR using a pBAC108L template and the following primers: BACX, 5'-CCCTCGAGTGAGCGAGGAAGCACAGGG-3', and BACS, 5'-GCTCGTCGACAGCGACACACTTGCATCGG-3'. PCR products were labeled using a nick translation kit with digoxigenin-11dUTP or biotin-16dUTP (Roche Diagnostics). For simultaneous indirect immunofluorescence and FISH staining, we used anti-CENP-A (monoclonal antibody A1; Ando et al., 2002), anti-CENP-B (monoclonal antibody 2D8D8), anti-CENP-C (CGp2; Ikeno et al., 1998), and anti-CENP-E (monoclonal antibody 177; Yen et al., 1991).

### Quantification of multiplicity of introduced DNA

Multiplicity of introduced DNA was determined using a quantitative competitive PCR assay (Wang et al., 1989). Competitor DNAs for the host genomic sequence, the neomycin resistance gene sequence, junctions of syn-

thetic  $\alpha$ 21-I 11mer sequences, and synthetic RF322 sequences were prepared using a competitive DNA construction kit (TAKARA). Genomic DNA samples were isolated by agarase treatment and sonication from agarose plugs, which were prepared as described elsewhere (Ikeno et al., 1994). The primers were as follows: G1 (host genome), 5'-CTCCATTG-GAGTGAGCCCGG-3'; G2 (host genome), 5'-ACGGTTCAAATTCG-CACCC-3'; N1 (neomycin resistance gene), 5'-TGATTAGGGGTGATGGT-TCACGTAG-3'; N2 (neomycin resistance gene), 5'-CTGCATTCTAG-TTGTGGTTGTCC-3'; SA1 (junctions of synthetic aliphoid sequences), 5'-TCTGAGAATCCTTCTGTCTC-3'; SA2 (junctions of synthetic aliphoid sequences), 5'-GGGAATTCGCTAGTGAATTC-3'; RF1 (RF322 synthetic repeat), 5'-TGGGTATGGTGGCAGGCC-3'; and RF2 (RF322 synthetic repeat), 5'-CGCATGATCGCGTAGTCCG-3'. The amount of synthetic aliphoid DNA was designated as equivalent to multiplicity of the inserted aliphoid fragment (60 kb) on the BAC. DNA content of the MAC was estimated by quantification of the fluorescence intensity of the propidium iodide (5  $\mu$ g/ml)-stained metaphase spread image (Ikeno et al., 1998). Multiplicity was designated as equivalent to the copy number of 67-kb BAC DNA.

### Competitive CHIP assay

A sample cell line produced by pWTR11.32 transformation and a reference cell line produced by pMTR11.32 transformation were harvested. These cell lines contained similar levels of BAC DNA. They were mixed together and then fixed with 0.25% formalin in the same tube at 22°C for 5 min. After fixation, the cells were centrifuged and recovered in buffer S (100 mM Hepes, pH 8.0, aprotinin, leupeptin, 1% BSA). Recovered cells were sonicated for 7 min with Biorupter (Cosmobio). After sonication, samples were supplemented with an equal volume of 2 $\times$  buffer W (aprotinin, leupeptin, 1% BSA, 0.4% NP-40, 0.4% SDS, 20 mM MgCl<sub>2</sub>, 10 mM ATP, 0.02 mg/ml RNase A) and sonicated again for 2 min. This sample was centrifuged, and the supernatant was used as input material.

Input material prepared from 5  $\times$  10<sup>5</sup> cells was immunoprecipitated by using 5  $\mu$ g of antibody with 25  $\mu$ g protein G-Sepharose (Amersham Biosciences). Immunoprecipitated DNA was washed three times with buffer W (10 mM Hepes, aprotinin, leupeptin, 1% BSA, 0.2% NP-40, 0.2% SDS, 10 mM MgCl<sub>2</sub>, 5 mM ATP, 0.01 mg/ml RNase A) and then eluted by buffer E (2.5% SDS, 10 mM Tris, 0.1 mM EDTA). The antibodies used were anti-H3 (Upstate Biotechnology), anti-CENP-A, and anti-CENP-B.

The recovered DNA sample was used as a template in PCR with the amplification primers for the synthetic wild-type sequence and the  $\alpha$ 21-I 11mer with modified CENP-B boxes (SA3, TTTTGTATGTGTGTAC-CCAGCC, and SA4, ACCTTTTATTGAATTC) and the amplification primers for the synthetic RF322 repeats (RF1 and RF2). After PCR, amplified DNA was digested by the endonuclease EcoRV (for the synthetic  $\alpha$ 21-I 11mer repeats) or EcoRI (for the synthetic RF322 repeats). Digested samples were analyzed by agarose gel electrophoresis. Gel images were quantified using Advanced Quantifier (Bio Image).

We thank K. Yoda (Nagoya University) and N. Nozaki (Kanagawa Dental College, Yokosuka, Japan) for producing the anti-CENP-A antibody, T. Yen (Fox Chase Cancer Center, Philadelphia, PA) for the gift of the anti-CENP-E antibody, N. Shimizu (Keio University, Tokyo, Japan) for the pBAC108L plasmid, and M. Hirano for secretarial work.

This research was supported by a grant-in-aid for Scientific Research on Priority Areas (B), Core Research for Evolutional Science and Technology, and Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Science, Sports, and Culture of Japan, as well as a grant-in-aid for Basic Research 21 for Breakthroughs in Info-Communications and a grant-in-aid from the Cell Science Research Foundation. J. Ohzeki is a research fellow of the Japan Society for the Promotion of Science.

Submitted: 22 July 2002

Revised: 25 October 2002

Accepted: 25 October 2002

## References

Abrieu, A., J.A. Kahana, K.W. Wood, and D.W. Cleveland. 2000. CENP-E as an essential component of the mitotic checkpoint in vitro. *Cell*. 102:817–826.

Alexandrov, I., A. Kazakov, I. Tumenova, V. Shepelev, and Y. Yurov. 2001.  $\alpha$ -Satellite DNA of primates: old and new families. *Chromosoma*. 110:253–266.

Ando, S., H. Yang, N. Nozaki, T. Okazaki, and K. Yoda. 2002. CENP-A, -B, and

-C chromatin complex that contains the I-type  $\alpha$ -satellite array constitutes the prekinetochore in HeLa cells. *Mol. Cell. Biol.* 22:2229–2241.

Becker-Andre, M., and K. Hahlbrock. 1989. Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by a PCR aided transcript titration assay (PATTY). *Nucleic Acids Res.* 17:9437–9446.

Choo, K.H.A. 1997a. The Centromere. Oxford University Press, Oxford, UK. 320 pp.

Choo, K.H. 1997b. Centromere DNA dynamics: latent centromeres and neocentromere formation. *Am. J. Hum. Genet.* 61:1225–1233.

Choo, K.H. 2001. Domain organization at the centromere and neocentromere. *Dev. Cell*. 1:165–177.

du Sart, D., M.R. Cancilla, E. Earle, J.I. Mao, R. Saffery, K.M. Tainton, P. Kalitsis, J. Martyn, A.E. Barry, and K.H. Choo. 1997. A functional neo-centromere formed through activation of a latent human centromere and consisting of non- $\alpha$ -satellite DNA. *Nat. Genet.* 16:144–153.

Earnshaw, W.C., and N. Rothfield. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma*. 91:313–321.

Earnshaw, W.C., H. Ratrie III, and G. Stetten. 1989. Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. *Chromosoma*. 98:1–12.

Ebersole, T.A., A. Ross, E. Clark, N. McGill, D. Schindelbauer, H. Cooke, and B. Grimes. 2000. Mammalian artificial chromosome formation from circular aliphoid input DNA does not require telomere repeats. *Hum. Mol. Genet.* 9:1623–1631.

Fisher, A.M., L. Al-Gazali, T. Pramathan, R. Quaife, A.E. Cockwell, J.C. Barber, W.C. Earnshaw, J. Axelman, B.R. Migeon, and C. Tyler-Smith. 1997. Centromeric inactivation in a dicentric human Y;21 translocation chromosome. *Chromosoma*. 106:199–206.

Grimes, B.R., A.A. Rhoades, and H.F. Willard. 2002.  $\alpha$ -Satellite DNA and vector composition influence rates of human artificial chromosome formation. *Mol. Ther.* 5:798–805.

Haaf, T., and D.C. Ward. 1994. Structural analysis of  $\alpha$ -satellite DNA and centromere proteins using extended chromatin and chromosomes. *Hum. Mol. Genet.* 3:697–709.

Harrington, J.J., G. Van Bokkelen, R.W. Mays, K. Gustashaw, and H.F. Willard. 1997. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat. Genet.* 15:345–355.

Henikoff, S. 2002. Near the edge of a chromosome's "black hole." *Trends Genet.* 18:165–167.

Henikoff, S., K. Ahmad, and H.S. Malik. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science*. 293:1098–1102.

Howman, E.V., K.J. Fowler, A.J. Newson, S. Redward, A.C. MacDonald, P. Kalitsis, and K.H. Choo. 2000. Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proc. Natl. Acad. Sci. USA*. 97:1148–1153.

Hudson, D.F., K.J. Fowler, E. Earle, R. Saffery, P. Kalitsis, H. Trowell, J. Hill, N.G. Wreford, D.M. de Kretser, M.R. Cancilla, et al. 1998. Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights. *J. Cell Biol.* 141:309–319.

Ikeno, M., H. Masumoto, and T. Okazaki. 1994. Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range  $\alpha$ -satellite DNA arrays of human chromosome 21. *Hum. Mol. Genet.* 3:1245–1257.

Ikeno, M., B. Grimes, T. Okazaki, M. Nakano, K. Saitoh, H. Hoshino, N.I. McGill, H. Cooke, and H. Masumoto. 1998. Construction of YAC-based mammalian artificial chromosomes. *Nat. Biotechnol.* 16:431–439.

Kalitsis, P., K.J. Fowler, E. Earle, J. Hill, and K.H. Choo. 1998. Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. *Proc. Natl. Acad. Sci. USA*. 95:1136–1141.

Kitagawa, K., and P. Hieter. 2001. Evolutionary conservation between budding yeast and human kinetochores. *Nat. Rev. Mol. Cell Biol.* 2:678–687.

Koch, J. 2000. Neocentromeres and  $\alpha$ -satellite: a proposed structural code for functional human centromere DNA. *Hum. Mol. Genet.* 9:149–154.

Laurent, A.M., J. Puechberty, and G. Roizes. 1999. Hypothesis: for the worst and for the best, L1Hs retrotransposons actively participate in the evolution of the human centromeric aliphoid sequences. *Chromosome Res.* 7:305–317.

Lo, A.W., D.J. Magliano, M.C. Sibson, P. Kalitsis, J.M. Craig, and K.H. Choo. 2001. A novel chromatin immunoprecipitation and array (CIA) analysis identifies a 460-kb CENP-A-binding neocentromere DNA. *Genome Res.* 11:448–457.

Masumoto, H., H. Masukata, Y. Muro, N. Nozaki, and T. Okazaki. 1989. A human centromere antigen (CENP-B) interacts with a short specific sequence in aliphoid DNA, a human centromeric satellite. *J. Cell Biol.* 109:1963–1973.

- Masumoto, H., M. Ikeno, M. Nakano, T. Okazaki, B. Grimes, H. Cooke, and N. Suzuki. 1998. Assay of centromere function using a human artificial chromosome. *Chromosoma*. 107:406–416.
- Mejia, J.E., A. Willmott, E. Levy, W.C. Earnshaw, and Z. Larin. 2001. Functional complementation of a genetic deficiency with human artificial chromosomes. *Am. J. Hum. Genet.* 69:315–326.
- Mejia, J.E., A. Alazami, A. Willmott, P. Marschall, E. Levy, W.C. Earnshaw, and Z. Larin. 2002. Efficiency of de novo centromere formation in human artificial chromosomes. *Genomics*. 79:297–304.
- Muro, Y., H. Masumoto, K. Yoda, N. Nozaki, M. Ohashi, and T. Okazaki. 1992. Centromere protein B assembles human centromeric  $\alpha$ -satellite DNA at the 17-bp sequence, CENP-B box. *J. Cell Biol.* 116:585–596.
- Murphy, T.D., and G.H. Karpen. 1998. Centromeres take flight:  $\alpha$  satellite and the quest for the human centromere. *Cell*. 93:317–320.
- Murray, A.W., and J.W. Szostak. 1983. Construction of artificial chromosomes in yeast. *Nature*. 305:189–193.
- Politi, V., G. Perini, S. Trazzi, A. Pliss, I. Raska, W.C. Earnshaw, and G.D. Valle. 2002. CENP-C binds the  $\alpha$ -satellite DNA in vivo at specific centromere domains. *J. Cell Sci.* 115:2317–2327.
- Romanova, L.Y., G.V. Deriagin, T.D. Mashkova, I.G. Tumeneva, A.R. Mushegian, L.L. Kisselev, and I.A. Alexandrov. 1996. Evidence for selection in evolution of  $\alpha$  satellite DNA: the central role of CENP-B/pJ  $\alpha$  binding region. *J. Mol. Biol.* 261:334–340.
- Saffery, R., L.H. Wong, D.V. Irvine, M.A. Bateman, B. Griffiths, S.M. Cutts, M.R. Cancilla, A.C. Cendron, A.J. Stafford, and K.H. Choo. 2001. Construction of neocentromere-based human minichromosomes by telomere-associated chromosomal truncation. *Proc. Natl. Acad. Sci. USA*. 98:5705–5710.
- Saitoh, H., J. Tomkiel, C.A. Cooke, H. Ratrie III, M. Maurer, N.F. Rothfield, and W.C. Earnshaw. 1992. CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. *Cell*. 70:115–125.
- Schueler, M.G., A.W. Higgins, M.K. Rudd, K. Gustashaw, and H.F. Willard. 2001. Genomic and genetic definition of a functional human centromere. *Science*. 294:109–115.
- Shelby, R.D., O. Vafa, and K.F. Sullivan. 1997. Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. *J. Cell Biol.* 136:501–513.
- Sullivan, K.F., M. Hechenberger, and K. Masri. 1994. Human CENP-A contains a histone H3–related histone fold domain that is required for targeting to the centromere. *J. Cell Biol.* 127:581–592.
- Tyler-Smith, C., G. Gimelli, S. Giglio, G. Florida, A. Pandya, G. Terzoli, P.E. Warburton, W.C. Earnshaw, and O. Zuffardi. 1999. Transmission of a fully functional human neocentromere through three generations. *Am. J. Hum. Genet.* 64:1440–1444.
- Wandall, A. 1994. A stable dicentric chromosome: both centromeres develop kinetochores and attach to the spindle in monocentric and dicentric configuration. *Chromosoma*. 103:56–62.
- Wang, A.M., M.V. Doyle, and D.F. Mark. 1989. Quantitation of mRNA by the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA*. 86:9717–9721.
- Warburton, P.E. 2001. Epigenetic analysis of kinetochore assembly on variant human centromeres. *Trends Genet.* 17:243–247.
- Wiens, G.R., and P.K. Sorger. 1998. Centromeric chromatin and epigenetic effects in kinetochore assembly. *Cell*. 93:313–316.
- Willard, H.F. 1985. Chromosome-specific organization of human  $\alpha$  satellite DNA. *Am. J. Hum. Genet.* 37:524–532.
- Willard, H.F., and J.S. Waye. 1987. Hierarchical order in chromosome-specific human alpha satellite DNA. *Trends Genet.* 3:192–198.
- Williams, B.C., T.D. Murphy, M.L. Goldberg, and G.H. Karpen. 1998. Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nat. Genet.* 18:30–37.
- Yang, C.H., J. Tomkiel, H. Saitoh, D.H. Johnson, and W.C. Earnshaw. 1996. Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. *Mol. Cell Biol.* 16:3576–3586.
- Yen, T.J., D.A. Compton, D. Wise, R.P. Zinkowski, B.R. Brinkley, W.C. Earnshaw, and D.W. Cleveland. 1991. CENP-E, a novel human centromere-associated protein required for progression from metaphase to anaphase. *EMBO J.* 10:1245–1254.
- Yoda, K., K. Kitagawa, H. Masumoto, Y. Muro, and T. Okazaki. 1992. A human centromere protein, CENP-B, has a DNA binding domain containing four potential  $\alpha$  helices at the NH<sub>2</sub> terminus, which is separable from dimerizing activity. *J. Cell Biol.* 119:1413–1427.
- Yoda, K., S. Ando, A. Okuda, A. Kikuchi, and T. Okazaki. 1998. In vitro assembly of the CENP-B/ $\alpha$ -satellite DNA/core histone complex: CENP-B causes nucleosome positioning. *Genes Cells*. 3:533–548.