Chromosoma (1999) 108:367–374

CHROMOSOMA
© Springer-Verlag 1999

CENP-A associated complex satellite DNA in the kinetochore of the Indian muntjac

Omid Vafa, Richard D. Shelby, Kevin F. Sullivan

Department of Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Received: 18 May 1999; in revised form: 5 July 1999 / Accepted: 20 July 1999

Abstract. The centromere/kinetochore complex is a chromosomal assembly that mediates chromosome motility and mitotic regulation by interacting with microtubules of the mitotic spindle apparatus. Centromere protein A (CENP-A) is a histone H3 homolog that is concentrated in the chromatin of the inner kinetochore plate of human chromosomes. To identify DNA sequences associated with the inner kinetochore plate, we used anticentromere autoantibodies to immunoprecipitate CENP-A associated chromatin selectively from Indian muntjac fibroblasts. DNA was cloned from immunoprecipitated CENP-A- associated chromatin and characterized by DNA sequence and hybridization analyses. A novel centromeric satellite DNA sequence was identified and shown by fluorescence in situ hybridization analysis to be present at all centromeres of the Indian muntjac. This satellite DNA constitutes a 972 bp monomer repeat and shows partial homology with satellite II DNA of the white-tailed deer. Southern blot analysis of muntjac genomic DNA suggests that this satellite DNA is present in repetitive tandem arrays and contains complex internal arrangements. In conjunction with previous work showing the association of CENP-A with human α-satellite DNA, we conclude that the mammalian inner kinetochore plate contains a unique form of chromatin that contains CENP-A in association with complex satellite DNA.

Introduction

In eukaryotes, the mitotic and meiotic segregation of chromosomes requires centromeres. Cytogenetically evident as the primary constrictions of mitotic chromosomes, centromeres interact with spindle fibers to mediate the alignment of chromosomes at the metaphase plate and the subsequent migration of chromosomes to dividing daughter cells (Pluta et al. 1995). These specialized chromosomal domains comprise a uniquely packaged form of heterochromatin that acts as a nucleus for the assembly of the kinetochore, a trilaminar proteinaceous structure on the surface of each chromatid that mediates interactions with the spindle fibers of the mitotic apparatus (Rieder and Salmon 1998). It is becoming increasingly evident that the kinetochore is a structure that not only mediates the mechanical segregation of chromosomes through its microtubule binding components, but also contains checkpoint protein complexes that serve to integrate spindle mechanics and chromosome movement with the cell cycle machinery (Allshire 1997).

It remains unclear what structural features of the chromosome specify mammalian centromere function. Budding yeast chromosomes possess a 125 bp *cis*-acting DNA element that acts as a nucleus for the protein assembly events required for centromere function (Clarke and Carbon 1980; Lechner and Ortiz 1996). The complexity of centromeric organization is extended in the fission yeast, *Schizosaccharomyces pombe*, in which centromeres span ~40–100 kb of DNA, containing irregular nucleosomal arrays flanked by various repeated DNA elements (Clarke and Baum 1990; Polizzi and Clarke 1991). In *Drosophila*, a functional centromere has been dissected to a 420 kb segment of heterochromatic DNA containing both repetitive satellite sequences and transposon-like elements (Murphy and Karpen 1995; Sun et al. 1997). However, these sequences are neither exclusive to centromeres nor present at all *Drosophila* centromeres. The relationship between centromere function and DNA sequence becomes even less clear in mammalian centromeres, where centromere function is generally associated with extremely large arrays of satellite DNA spanning hundreds to thousands of kilobase pairs (Tyler-Smith and Willard 1993; Murphy and Karpen 1998). The major centromeric satellite DNA of human chromosomes, α-satellite, has recently been

---

Present address: O. Vafa, The Salk Institute, Gene Expression Lab, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA

Edited by: J.B. Rattner

Correspondence to: K.F. Sullivan (e-mail: ksullivan@scripps.edu)

shown to promote formation of minichromosomes (Harrington et al. 1997; Ikeno et al. 1998). However, a variety of evidence indicates that α-satellite DNA is neither necessary (Voullaire et al. 1993; Barry et al. 1999) nor sufficient (Earnshaw et al. 1989) to establish a functional centromere. An alternative model has emerged in which centromere function is specified by epigenetic features of the chromosome (Karpen and Allshire 1997). Indeed, centromere activity in the fission yeast exhibits strong epigenetic effects, including stochastic activation or inactivation and position effect variegation (Allshire et al. 1994; Steiner and Clarke 1994; Ekwall et al. 1997). Complete understanding of centromere function in mammals will therefore require characterization of the specialized nucleoprotein complex that constitutes the kinetochore.

A group of chromatin associated proteins thought to play a role in mammalian kinetochore assembly have been identified as antigens recognized by human autoimmune sera from CREST patients (calcinosis, Reynaud's phenomenon, esophageal dysmotility, sclerodactly, telangiectasia) (Earnshaw and Rothfield 1985; Earnshaw et al. 1987; Saitoh et al. 1992; Sullivan et al. 1994). One of these, CENP-B, is a sequence specific α-satellite DNA binding protein that is distributed throughout the centromeric heterochromatin domain (Cooke et al. 1990; Muro et al. 1992; Yoda et al. 1992). CENP-A and -C are components of the inner kinetochore plate and specifically associate with functional centromeres (Earnshaw et al. 1989; Saitoh et al. 1992; Warburton et al. 1997). CENP-C is thought to be a DNA binding protein, although sequence specificity of this DNA binding has not been demonstrated (Sugimoto et al. 1994; Yang et al. 1996). CENP-A is an M_r 17,000 histone H3 homolog thought to assemble a unique type of chromatin fiber in the inner kinetochore plate (Palmer et al. 1987, 1991; Sullivan et al. 1994; Shelby et al. 1997). Direct isolation of nucleosomal DNA associated with CENP-A has demonstrated that α-satellite DNA is the major sequence component of the inner kinetochore plate in human chromosomes (Vafa and Sullivan 1997).

Biochemical isolation of the CENP-A chromatin compartment provides a unique opportunity to identify conserved components of the mammalian kinetochore. Human CENP-A assembles at the centromeres of the Indian muntjac (Sullivan et al. 1994) as well as several other mammalian species (K.F. Sullivan, unpublished observations). Here we have used a CREST serum that specifically recognizes CENP-A in the Indian muntjac to isolate chromatin from the kinetochore of Indian muntjac chromosomes. We show that cloned CENP-A associated DNA from the Indian muntjac kinetochore is enriched in a distinct GC-rich complex satellite sequence that shares identity with cervid satellite II family (Qureshi and Blake 1995). In view of the lack of conservation among complex satellite sequences in various classes of mammals, our data strongly suggest that epigenetic features of centric chromatin containing CENP-A play a greater role than a conserved primary sequence in the assembly of the mammalian kinetochore.

Materials and methods

Cell culture and chromosome spreads

Indian muntjac fibroblasts (ATCC CCL 157) were cultured in Ham's F-12 medium with 10% FBS (GIBCO BRL, Gaithersburg, Md.) at 37°C in 5% CO_2 atmosphere. Cells were blocked in metaphase by incubation with Colcemid (10 µg/ml) for up to 6 h prior to collection and processing for chromosome spreads used in localization of CENP antigens and fluorescence in situ hybridization (FISH) analysis. Spreads for CENP localization were made after trypsinization and harvest of blocked metaphase cells, followed by washing once with medium and twice with PBS. Cells were resuspended in 2 ml of hypotonic buffer (10 mM TRIS-HCl, pH 7.4, 40 mM glycerol, 20 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$, 0.5 mM $MgCl_2$, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride for 30 min at 4°C. Then 100 µl of this supension was cytospun onto coverslips at 1000 rpm for 5 min and allowed to air dry at room temperature. For FISH analysis, metaphase-blocked cells were washed in PBS and resuspended in 75 mM KCl (37°C) for 25 min. After centrifugation, 10 ml of fixative (3:1 methanol:glacial acetic acid) was added dropwise to the cell pellet followed by incubation on ice for 10 min. Cells were collected by centrifugation at 1000 rpm for 5 min and resuspended in a series of reduced fixative volumes to 1 ml. Chromosome spreads were made by fixed cell-suspension drops from 30 cm onto glass slides.

Immunochemistry and immunofluorescence microscopy

Immunoblotting analyses were performed as previously described (Sullivan et al. 1994) using anticentromere sera (hACA) from 30 patients at a dilution of 1:300. Blots were developed using horseradish peroxidase-conjugated secondary antibodies (Amersham, Arlington Heights, Ill.) and chemiluminescence detection reagent (Pierce Chemical, Rockford, Ill.). CENP antigens were localized in chromosome spreads by immunocytochemistry using a human anticentromere serum (hACA-M) at a 1:1000 dilution as described previously (Sullivan et al. 1994), and chromosomes were stained with 4',6-diamidino-2-phenylindole. Immunoprecipitation of kinetochore chromatin from Indian muntjac fibroblast cell lines involved the isolation and washing of nuclei according to Masumoto et al. (1989), followed by a wash in buffer A (5 mM HEPES, pH 7.5, 10 µM leupeptin, 1.5 µM aprotinin, 1 mM dithiothreitol) and centrifugation at 3000 *g*. To generate chromatin fragments, the nuclear pellet was resuspended in digestion buffer (buffer A containing 1 mM $CaCl_2$, 200 U/ml micrococcal nuclease) at 0.5–1×10^8 nuclei/ml and incubated at 37°C for 3 min. Digestion was stopped by adding EDTA to a final concentration of 20 mM. After centrifugation at 8000 *g* the supernatant was collected, and the pellet was resuspended in buffer A containing 20 mM EDTA. The nuclei were subjected to two additional rounds of extraction by 10 s pulses of sonication followed by centrifugation to collect solubilized chromatin in the supernatant. Supernatants were pooled into a siliconized Eppendorf tube, supplemented with NP-40 to 0.1%, 10 µl hACA-M serum and mixed end over end on a rotor for 2 h at 4°C. A 100 µl aliquot of protein-A Sepharose (Pharmacia Fine Chemicals, Piscataway, N.J.) was blocked in NET gel buffer (150 mM NaCl, 50 mM TRIS-HCl, pH 7.5, 1 mM EDTA, 0.1% NP-40, 0.25% gelatin, 0.02% NaN_3), equilibrated in buffer A and incubated with immune complexes for an additional 2 h on a rocker at 4°C. Immunoprecipitates were washed five times (10–15 min each) with buffer A containing 10 mM EDTA and further processed for protein and DNA analysis.

Cloning

To clone CENP associated DNA, immunoprecipitates and supernatant controls from 2–5×10^7 cells were supplemented with 0.5%

Tween 20 and Proteinase K (200 µg/ml) and incubated for 4 h at 56°C. DNA was isolated by phenol-chloroform extraction followed by precipitation with 10% sodium acetate and 3 vol. of ethanol in the presence of 1 µg of glycogen (Boehringer Mannheim, Indianapolis, Ind.). Precipitated DNA was washed in 70% ethanol, and the 3' termini of the DNA fragments were dephosphorylated with calf intestinal alkaline phosphatase (Boehringer Mannheim). DNA was reextracted, precipitated and resuspended in water. The 5' hydroxyl termini of these fragments were phosphorylated using T4 polynucleotide kinase (New England Biolabs, Beverly, Mass.). DNA fragments were treated with Taq polymerase (Fisher Scientific, Pittsburgh, Pa.) to add a single deoxyadenosine to the 3' ends of these DNA fragments. Following phenol-chloroform extraction, these DNA fragments were ligated with the linearized pCR 2.1 TA vector (Invitrogen, San Diego, Calif.), and Epicurian Coli Ultracompetent Cells (Stratagene, La Jolla, Calif.) were transformed for the generation of plasmid libraries. All transformants were plated within 1 h or kept on ice to prevent replication of individual bacterial clones prior to plating and blue-white selection using 5-bromo-4-chloro-3-indolyl-galactopyranoside on carbenicillin plates.

Nucleic acid analyses

Cloned DNA was manually sequenced (Sequenase 2.0, United States Biochemical Corp., Cleveland, Ohio) and further characterized using the Basic Local Alignment Search Tool (Altschul et al. 1990) and programs of the University of Wisconsin Genetics Computer Group suite. Probes of immunoprecipitated and cloned DNA satellite sequences radiolabeled with [^{32}P]ATP were generated by random primer labeling using exo-Klenow and the Prime-It II kit (Stratagene, La Jolla, Calif.). Probes were purified using Sephadex G25 spin columns. Hybridizations were performed as described previously (Sullivan and Glass 1991). For colony hybridization analyses, bacterial libraries were transferred to 384-well dishes (Nunc, Naperville, Ill.), replica plated, and grown on nylon filter membranes overnight. Genomic DNA isolated from Indian muntjac fibroblasts was digested with restriction enzymes according to the manufacturer's specifications (New England Biolabs) and the DNA was separated by electrophoresis on 1% agarose gels and transferred onto nylon membranes for Southern blot analysis. All blots were washed three times at 51°C for 20 min. each in 0.1×SSC, 0.1% SDS. (1×SSC is 0.15 M NaCl, 0.015 M sodium citrate.)

Fluorescence in situ hybridization analysis

The 750 bp insert of satellite clone im23 was used for in situ hybridization. DNA was biotinylated using the BioPrime DNA Labeling System (Gibco BRL Life Technologies). The FISH analysis of chromosome spreads was performed in the absence of Cot DNA by standard procedures as described by Shimizu et al. (1996).

Results

Centromere proteins of the Indian muntjac

Human autoimmune sera from scleroderma CREST patients were used to identify conserved centromere antigens in the Indian muntjac by immunoblot analysis (Fig. 1A). In nuclear lysates from Indian muntjac fibroblasts, CREST sera recognize several antigens with relative molecular masses of 19,000, 80,000, and 140,000 (Fig. 1A, lanes A–C). These cross-reacting proteins correspond in size to human CENP-A, -B and -C, respectively, and therefore are likely to be the bona fide muntjac homologs of the human CENPs. An additional band migrating at M_r 50,000 (Fig. 1A, lane D) may correspond to the less characterized CENP-D antigen previously identified in rat and Chinese hamster cells (Earnshaw and Rothfield 1985; Kingwell and Rattner 1987). Previously, we showed that kinetochore-associated DNA can be directly isolated by immunoprecipitation of CENP-A from human chromosomes (Vafa and Sullivan 1997). We reasoned that a human serum that preferentially cross-reacts with muntjac CENP-A as determined by immunoblot analysis could be used selectively to immunoprecipitate kinetochore chromatin from Indian muntjac chromosomes. This would allow us to identify DNA sequences of the kinetochore in a non-primate mammalian species. One of the examined autoimmune sera, designated hACA-M, recognizes an M_r 19,000 nuclear protein in the muntjac (Fig. 1A, lane A) and labels centromeres of all Indian muntjac metaphase chromosomes (Fig. 1B). Based on size and the known conservation of mammalian CENP-A (Sullivan et al. 1994; Kalitsis et al. 1998), we conclude that hACA-M specifically recognizes CENP-A in the Indian muntjac.

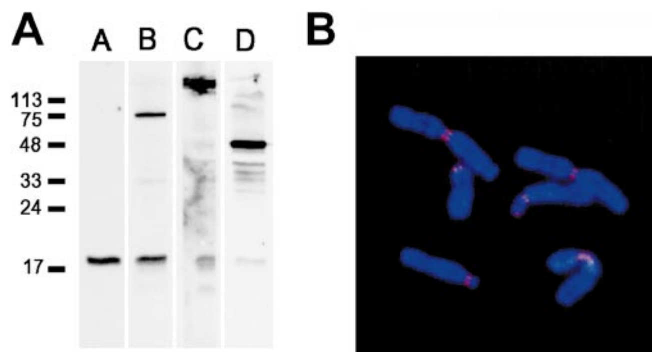
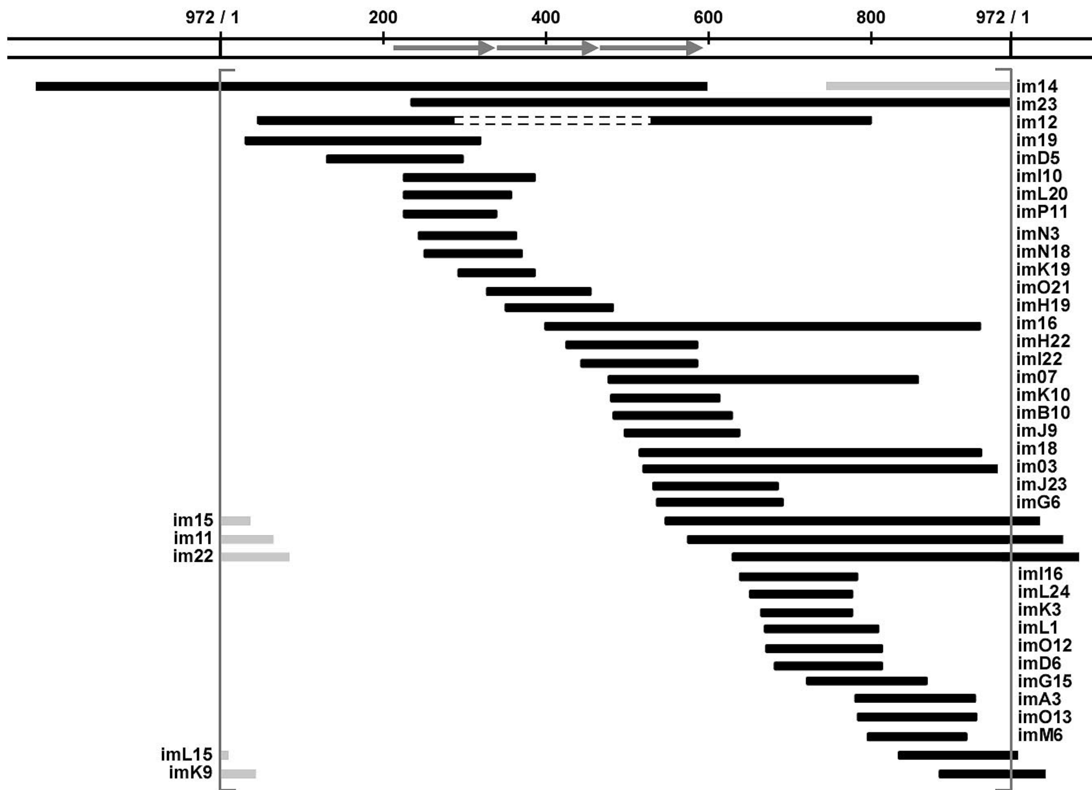


Fig. 1. Immunoblot analyses of Indian muntjac nuclear lysates (**A**) and localization by immunofluorescence microscopy of CREST antigens in metaphase chromosomes (**B**) by hACA antibodies. Representative blot analyses of Indian muntjac nuclear lysates using four different CREST autoimmune antisera from a pool of 30 patients are shown. *Lane A* shows that one of the CREST sera (hACA-M) specifically recognizes the M_r 19,000 CENP-A protein. Additional CREST antigens of M_r 80,000 (*lane B*), M_r 140,000 (*lane C*) and M_r 50,000 (*lane D*) are also observed. **B** Immunofluorescence labeling of an Indian muntjac metaphase chromosome spread using human CREST serum hACA-M detected with rhodamine labeled anti-human secondary antibody (*red*). All centromeres of Indian muntjac chromosomes are recognized. Chromosomes were counterstained with 4',6-diamidino-2-phenylindole

Complex CENP-A associated satellite DNA in the Indian muntjac

Human autoantiserum hACA-M was used to immunoprecipitate soluble chromatin prepared from Indian muntjac nuclei. DNA fragments associated with immunoprecipitated complexes were purified and cloned as described previously for human CENP-A associated DNA (Vafa and Sullivan 1997). Clones were selected at ran-



B

```
      0                                                 50
rep1  CCAAAGGCGGTTTGCAAGCATTCTCGATGGACTGCTGCCGCTCACTG.CCGAGGGAGGCTTGCTTCACGCGCAAG
rep2  CCAAGGGCCGGTTGCAAGCATTCTCGCCGG.CTGCCGCACCTCCCACCACAAGGGAGGCTTGTTTTTCCCGCAAG
rep3  CCAAGGGCCGGTTGCAAGCATTCTCGCCGG.CTGCCGAACCTCCCGCCGCAAGGGAGGCTTGTTTTTCCCGCAAG

                              100                          127
rep1  TCGCTCAAACTGCTCCCAAGCCGGGAGAAAATCGCTTCGGGGTGGTTCTTTCT
rep2  TCGCGCAATCCGCTCCCAAGCTGGGAGAAAAGCACATCGGGGTGGTTCTTTCT
rep3  TCGCGCAATCCGCTCCCAAGCCGGGAGAAAAGCACATCGGGGTGGTACTTTCT
```

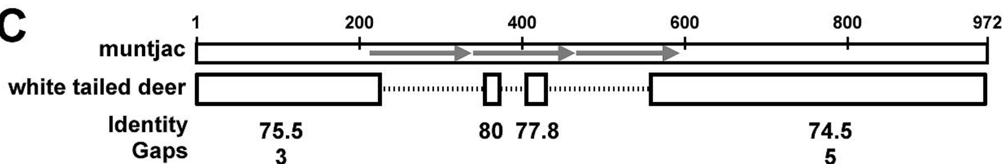


Fig. 2A–C. Analysis of cloned CENP-A associated DNA. **A** Alignment of 39 individual clones allows identification of a 972 bp repeat sequence. The features of the proposed monomer repeating unit are shown at the top with the individual sequences represented by *black bars* beneath. *Vertical lines* delimit the designated monomer repeat unit. *Gray bars* represent extensions of corresponding sequences across the monomer repeat unit, demonstrating sequence continuity across the designated boundaries. *Arrows* indicate the positions of three short internal repeats within the monomer. *Dashed line* illustrates a gap in one clone lacking the internal repeat. **B** Alignment of the three 127 bp internal tandem repeat sequences found within the 972 bp monomer repeat. **C** Schematic illustration of the homology between the muntjac 972 bp satellite II repeat and the homologous satellite II sequence reported from the white-tailed deer. The muntjac sequence is diagrammed at the *top*. Contiguous blocks of homology with the white-tailed deer satellite II sequence are shown with the percent identity and number of gaps in the alignment listed below each *block*. Note that the main difference between the two satellite II sequences is in the internal repeat region. DNA sequences have been deposited in the EMBL nucleotide sequence database under accession numbers AJ388471–AJ388510

dom from three independent experiments and subjected to DNA sequence analysis. A single DNA sequence class was recognized that was abundant in each of the libraries and showed similarity to satellite DNA characterized from the white-tailed deer (Qureshi and Blake 1995). Probes from this enriched sequence class were generated and used in colony hybridization analysis in two of these libraries. Individual satellite DNA clones were isolated for further analysis.

DNA sequences were determined for 39 independent clones, which range in length from 96 to 852 bp (Fig. 2A). The homology between individual clones ranges from 68%–98%, indicating a considerable degree of divergence within the family of CENP-A associated satellite DNAs. Comparison and pairwise alignment of the sequences revealed a repeating sequence element of 972 bp in length. The terminus of clone im23 was arbitrarily chosen as the terminus of the repeating unit for purposes of analyzing the relative positions of individual sequences within the repeat unit, as shown in Fig. 2A. The two longest clones, im14 and im23, together defined a contiguous sequence of 1224 bp, spanning a complete 972 bp repeat. We interpret the 972 bp DNA sequence repeat as corresponding to the basic repeat unit of Indian muntjac kinetochore-associated satellite DNA on the basis of this sequence alignment and Southern blot analysis of genomic DNA shown below. The 972 bp satellite monomer unit is GC rich (61% G+C) and contains an internally repetitive region consisting of three copies of a short, direct 127 bp sequence (Fig. 2B). This internally repetitive region is heterogeneous within the satellite family, as demonstrated by isolation of a clone that spans this region but contains only a single copy of the 127 bp sequence.

The Indian muntjac satellite sequence bears significant homology to the satellite II sequence characterized for the white-tailed deer (Fig. 2C; Qureshi and Blake 1995). Using a 972 bp repeat unit constructed from two clones, an overall level of 75.1% homology was seen between the two satellite DNAs, comparable to that obtained for interspecies comparisons of α-satellite DNA among primates (Haaf and Willard 1998). The major region of difference between the two satellite sequences was in the internally repetitive region of the muntjac sequence. The white-tailed deer satellite II DNA sequence possessed only a single partial copy of the 127 bp repeat unit. Thus, these repeat elements isolated in association with CENP-A correspond to the Indian muntjac satellite II DNA family. No homology was observed on comparison with human α-satellite DNA.

Organization of Indian muntjac satellite II DNA into higher order repeat families was suggested by Southern analysis of genomic DNA (Fig. 3). Digestion with PstI generates a ladder with approximately 1 kb spacing, ranging from 970 bp to 9.5 kb. A single prominent band in the EcoRI digest at 4.8 kb, also present in the PstI digest, indicates that this satellite probably exists in higher order pentameric tandem arrays of the 972 bp monomer repeat. Digestion with HindIII results in a more complex pattern, with a prominent monomer-sized band and a series of bands separated by a ca 140 bp interval between

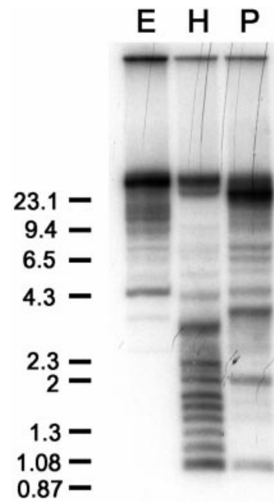


Fig. 3. Southern blot analysis of Indian muntjac genomic DNA. A CENP-A associated satellite II probe generated from clone im23 identifies a series of bands in restriction enzyme-digested genomic DNA. Genomic DNA digested by EcoRI (*E*), HindIII (*H*), and PstI (*P*) was hybridized with a radiolabeled probe generated from the 750 bp satellite clone im23 . The PstI digest reveals a monomer repeat length of ~970 bp with apparent higher order multimers

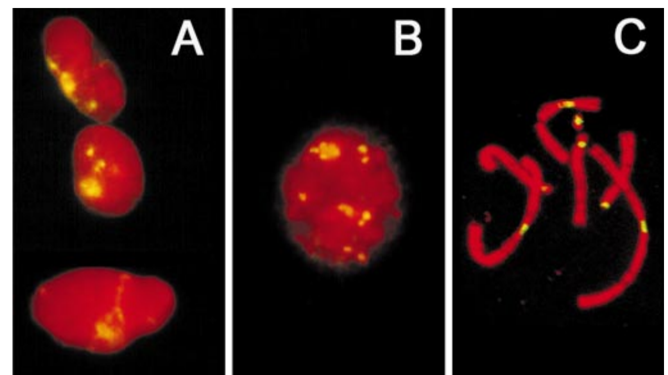


Fig. 4A–C. Fluorescence in situ hybridization with CENP-A associated satellite II DNA. Biotinylated probes generated from clone im23 were hybridized with cells and metaphase chromosomes and subsequently detected using fluorescein isothiocyanate-conjugated avidin. Representative cell images show the differential distribution of centromeric satellite DNA in interphase (**A**) and prophase (**B**), as well as its centromeric localization in metaphase chromosomes (**C**). DNA was stained with propidium iodide

1–2 kb. Together, these restriction enzyme patterns are consistent with a tandem organization of satellite II DNA based on a 972 bp monomer within a complex higher order sequence array. In addition, high molecular weight DNA >23 kb in size is observed in each digest. This suggests that additional satellite II-related DNA sequence families may exist that possess different higher-order sequence organization.

Satellite II DNA is present at all centromeres of Indian muntjac chromosomes

To determine the chromosomal distribution of the CENP-A associated satellite II DNA, we performed FISH analysis using a satellite II DNA probe prepared from the 750 bp insert of clone im23 (Fig. 4). In interphase cells, this probe hybridizes to discrete foci or linear tracks similar to those reported by Brinkley and co-

workers for CREST staining in Indian muntjac nuclei (Brinkley et al. 1984) (Fig. 4A). Occasionally these loci are also seen with preferential distribution at the nuclear periphery. A distinct double-dot hybridization pattern was evident in cells judged to be in late G2 or prophase on the basis of nuclear size or the initiation of chromosome condensation (Fig. 4B). In metaphase chromosome spreads, the centromeres of each of the seven Indian muntjac chromosomes hybridized with probe im23 (Fig. 4C). Unlike the previously characterized cervid satellite I DNA from the muntjac that hybridize with multiple loci along arms of chromosomes or preferentially with the X chromosome (Bogenberger et al. 1987), no interstitial or chromosome-specific hybridization was detected by the satellite II probe. Clearly, the satellite II sequence enriched in CENP-A chromatin is a conserved kinetochore component of all seven chromosomes of the Indian muntjac.

Discussion

This work was directed at identifying conserved components of the mammalian kinetochore. Consistent with the structural and functional conservation of mammalian kinetochores, several centromere-specific proteins are conserved among mammals (Sullivan and Glass 1991; Sullivan et al. 1994; Lanini and McKeon 1995; Kalitsis et al. 1998). In particular, the chromatin-associated CREST antigens are widely conserved and our immunoblot analysis of Indian muntjac has revealed proteins of similar size to each of the major CREST antigens, CENP-A, -B and -C. CENP-A was previously shown to be a component of the inner kinetochore plate in human chromosomes, where it is found associated with α-satellite DNA (Vafa and Sullivan 1997; Warburton et al. 1997). Using a human CREST serum with specific reactivity with the Indian muntjac CENP-A polypeptide, we have isolated a class of CENP-A associated DNA from the Indian muntjac. DNA sequence analysis reveals this to be a complex satellite DNA with a repeating unit of 972 bp. In situ hybridization demonstrates that this satellite sequence is present at all centromeres of the Indian muntjac and is located exclusively at centromeres. We propose that the 972 bp muntjac sequence represents the functional homolog of primate α-satellite DNA.

One of the most perplexing issues for understanding mammalian centromeric function remains the question of how centromeric identity is related to the sequence of the underlying DNA (Karpen and Allshire 1997; Willard 1998). Previously, we showed that human CENP-A assembles onto kinetochores in the Indian muntjac (Sullivan et al. 1994) as well as several other mammalian species (K.F. Sullivan, unpublished). CENP-A therefore recognizes a conserved feature of the mammalian centromere. We reasoned that direct isolation of CENP-A associated DNA sequences in two mammals could yield insight into conserved DNA motifs of the kinetochore itself, independently of the bulk of centromeric DNA. The 972 bp CENP-A associated satellite sequence is distinct from the satellite I family of centromere-associated satellite DNA previously described for the Indian muntjac and other ungulate species (Bogenberger et al. 1985; Lin et al. 1991; Lee et al. 1997). It shares homology with a different centromeric satellite family, satellite II of the white-tailed deer (Qureshi and Blake 1995). Cervid centromeres may therefore resemble those of primates, which possess complex collections of repetitive DNA sequence families (Tyler-Smith and Willard 1993). Direct comparison of muntjac satellite II with human α-satellite DNA failed to reveal sequence homology when several different algorithms were used for sequence comparison. The main similarity between the two sequence families is that they are both complex satellite DNA sequences. The long repeating units of α-satellite and satellite II, 171 and 972 bp, respectively, distinguish them from the simple satellite DNA sequences based on repeating units of 5–10 bp that are found in heterochromatic regions in several species (Beridze 1982). Thus, the only obviously conserved features of these kinetochore-associated DNAs are the repetitive nature of the sequences and their large repeating structure.

Fluorescence in situ hybridization analyses using CENP-A associated satellite II DNA indicate that this complex satellite DNA is found localized at all Indian muntjac centromeres. Moreover, the nuclear distribution of this sequence is highly similar to that of CREST staining patterns previously described for the various phases of the Indian muntjac cell cycle (Brinkley et al. 1984). In interphase these seqences appear as bead-like arrays of smaller unit kinetochores, become more condensed aggregates during prophase, and assume distinct kinetochore localization patterns in metaphase chromosomes. CENP-A associated satellite II contrasts with other minisatellite and heterochromatic centric satellite sequences, which were previously shown also to occupy interstitial domains within muntjac chromosomes by in situ hybridization studies (Yu et al. 1986; Lin et al. 1991; Lee et al. 1993). The localization patterns of these sequences make them unlikely candidates for specific association with centromeric antigens. Thus, our studies reveal a novel and distinct kinetochore DNA sequence that biochemically co-purifies with CENP-A and localizes exclusively to the centromeres of all Indian muntjac chromosomes.

The fact that human CENP-A is able to assemble at centromeres in the Indian muntjac presents a conundrum. What is the conserved feature of the kinetochore that is being recognized in this heterotypic assembly event? There are at least three possibilities. One possibility is that CENP-A localizes to centromeres by DNA recognition, but that recognition involves DNA structures that have a degenerate relationship with DNA sequence and that have so far eluded identification (Shelby et al. 1997). Secondly, a conserved nucleating sequence, yet to be identified, may enable CENP-A assembly at kinetochores, which subsequently spreads into the abundant complex satellite DNA. Finally, CENP-A may not recognize a specific DNA sequence or structure, but, rather, assemble at centromeres through a chromatin assembly mechanism that derives specificity through unique replication timing or through recognition of pre-

existing chromatin proteins at the centromere (Karpen and Allshire 1997; Vafa and Sullivan 1997; Csink and Henikoff 1998; Murphy and Karpen 1998). Clearly, characterization of the evolutionarily conserved mechanisms of assembly of centromeric chromatin will be a key step in unraveling the enigma of mammalian centromere specification.

Acknowledgements. O.V. thanks Teru Kanda and Geoff Wahl at the Salk Institute and Karine Monier at Scripps for technical guidance and reagents used in FISH analysis. This work was supported by a grant from the National Institute for General Medical Sciences to K.F.S.

References

Allshire RC (1997) Centromeres, checkpoints and chromatid cohesion. Curr Opin Genet Dev 7:264–273

Allshire RC, Javerzat JP, Redhead NJ, Cranston G (1994) Position effect variegation at fission yeast centromeres. Cell 76:157–169

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Barry AE, Howman EV, Cancilla MR, Saffery R, Choo KH (1999) Sequence analysis of an 80 kb human neocentromere. Hum Mol Genet 8:217–227

Beridze T (1982) Satellite DNA. Springer, Berlin Heidelberg New York

Bogenberger JM, Neumaier PS, Fittler F (1985) The muntjak satellite IA sequence is composed of 31-base-pair internal repeats that are highly homologous to the 31-base-pair subrepeats of the bovine satellite 1.715. Eur J Biochem 148:55–59

Bogenberger JM, Neitzel H, Fittler F (1987) A highly repetitive DNA component common to all Cervidae: its organization and chromosomal distribution during evolution. Chromosoma 95:154–161

Brinkley BR, Valdivia MM, Tousson A, Brenner SL (1984) Compound kinetochores of the Indian muntjac. Evolution by linear fusion of unit kinetochores. Chromosoma 91:1–11

Clarke L, Baum MP (1990) Functional analysis of a centromere from fission yeast: a role for centromere-specific repeated DNA sequences. Mol Cell Biol 10:1863–1872

Clarke L, Carbon J (1980) Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature 287:504–509

Cooke CA, Bernat RL, Earnshaw WC (1990) CENP-B: a major human centromere protein located beneath the kinetochore. J Cell Biol 110:1475–1488

Csink AK, Henikoff S (1998) Something from nothing: the evolution and utility of satellite repeats. Trends Genet 14:200–204

Earnshaw WC, Rothfield N (1985) Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. Chromosoma 91:313–321

Earnshaw WC, Sullivan KF, Machlin PS, Cooke CA, Kaiser DA, Pollard TD, Rothfield NF, Cleveland DW (1987) Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen. J Cell Biol 104:817–829

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

Ekwall K, Olsson T, Turner BM, Cranston G, Allshire RC (1997) Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. Cell 91:1021–1032

Haaf T, Willard HF (1998) Orangutan alpha-satellite monomers are closely related to the human consensus sequence. Mamm Genome 9:440–447

Harrington JA, Van Bokkelen G, Mays RW, Gustashaw K, Willard HF (1997) Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat Genet 15:345–355

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

Kalitsis P, MacDonald AC, Newson AJ, Hudson DF, Choo KH (1998) Gene structure and sequence analysis of mouse centromere proteins A and C. Genomics 47:108–114

Karpen GH, Allshire RC (1997) The case for epigenetic effects on centromere identity and function. Trends Genet 13:489–496

Kingwell B, Rattner JB (1987) Mammalian kinetochore/centromere composition: a 50 kDa antigen is present in the mammalian kinetochore/centromere. Chromosoma 95:403–407

Lanini L, McKeon F (1995) Domains required for CENP-C assembly at the kinetochore. Mol Biol Cell 6:1049–1059

Lechner J, Ortiz J (1996) The *Saccharomyces cerevisiae* kinetochore. FEBS Lett 389:70–74

Lee C, Sasi R, Lin CC (1993) Interstitial localization of telomeric DNA sequences in the Indian muntjac chromosomes: further evidence for tandem chromosome fusions in the karyotypic evolution of the Asian muntjacs. Cytogenet Cell Genet 63:156–159

Lee C, Court DR, Cho C, Haslett JL, Lin CC (1997) Higher-order organization of subrepeats and the evolution of cervid satellite I DNA. J Mol Evol 44:327–335

Lin CC, Sasi R, Fan YS, Chen ZQ (1991) New evidence for tandem chromosome fusions in the karyotypic evolution of Asian muntjacs. Chromosoma 101:19–24

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

Muro Y, Masumoto H, Yoda K, Nozaki N, Ohashi M, Okazaki T (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 TD, Karpen GH (1995) Localization of centromere function in a *Drosophila* minichromosome. Cell 82:599–609

Murphy TD, Karpen GH (1998) Centromeres take flight: alpha satellite and the quest for the human centromere. Cell 93:317–320

Palmer DK, O'Day K, Wener MH, Andrews BS, Margolis RL (1987) A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J Cell Biol 104:805–815

Palmer DK, O'Day K, Trong HL, Charbonneau H, Margolis RL (1991) Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proc Natl Acad Sci U S A 88:3734–3738

Pluta AF, Mackay AM, Ainsztein AM, Goldberg IG, Earnshaw WC (1995) The centromere: hub of chromosomal activities. Science 270:1591–1594

Polizzi C, Clarke L (1991) The chromatin structure of centromeres from fission yeast: differentiation of the central core that correlates with function. J Cell Biol 112:191–201

Qureshi SA, Blake RD (1995) Sequence characteristics of a cervid DNA repeat family. J Mol Evol 40:400–404

Rieder CL, Salmon ED (1998) The vertebrate cell kinetochore and its roles during mitosis. Trends Cell Biol 8:310–318

Saitoh H, Tomkiel J, Cooke CA, Ratrie H, Maurer M, Rothfield NF, Earnshaw WC (1992) CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. Cell 70:115–125

Shelby RD, Vafa O, Sullivan KF (1997) Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. J Cell Biol 136:501–513

Shimizu N, Kanda T, Wahl GM (1996) Selective capture of acentric fragments by micronuclei provides a rapid method for purifying extrachromosomally amplified DNA. Nat Genet 12:65–71

Steiner NC, Clarke L (1994) A novel epigenetic effect can alter centromere function in fission yeast. Cell 79:865–874

Sugimoto K, Yata H, Muro Y, Himeno M (1994) Human centromere protein C (CENP-C) is a DNA-binding protein which possesses a novel DNA-binding motif. J Biochem (Tokyo) 116:877–881

Sullivan KF, Glass CA (1991) CENP-B is a highly conserved mammalian centromere protein with homology to the helix-loop-helix family of proteins. Chromosoma 100:360–370

Sullivan KF, Hechenberger M, Masri K (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

Sun X, Wahlstrom J, Karpen GH (1997) Molecular structure of a functional *Drosophila* centromere. Cell 91:1007–1019

Tyler-Smith C, Willard HF (1993) Mammalian chromosome structure. Curr Opin Genet Dev 3:390–397

Vafa O, Sullivan KF (1997) Chromatin containing CENP-A and alpha-satellite DNA is a major component of the inner kinetochore plate. Curr Biol 7:897–900

Voullaire LE, Slater HR, Petrovic V, Choo KH (1993) A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? Am J Hum Genet 52:1153–1163

Warburton PE, Cooke CA, Bourassa S, Vafa O, Sullivan BA, Stetten G, Gimelli G, Warburton D, Tyler-Smith C, Sullivan KF, Poirier GG, Earnshaw WC (1997) Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Curr Biol 7:901–904

Willard HF (1998) Centromeres: the missing link in the development of human artificial chromosomes. Curr Opin Genet Dev 8:219–225

Yang CH, Tomkiel J, Saitoh H, Johnson DH, Earnshaw WC (1996) Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. Mol Cell Biol 16:3576–3586

Yoda K, Kitagawa K, Masumoto H, Muro Y, Okazaki T (1992) A human centromere protein, CENP-B, has a DNA binding domain containing four potential alpha helices at the NH2 terminus, which is separable from dimerizing activity. J Cell Biol 119:1413–1427

Yu LC, Lowensteiner D, Wong EF, Sawada I, Mazrimas J, Schmid C (1986) Localization and characterization of recombinant DNA clones derived from the highly repetitive DNA sequences in the Indian muntjac cells: their presence in the Chinese muntjac. Chromosoma 93:521–528