Characterization of a CENP-C homolog in Arabidopsis thaliana

Yutaka Ogura1,2*, Fukashi Shibata2, Hiroshi Sato1,2 and Minoru Murata1,2

1Research Institute for Bioresources, Okayama University
   Kurashiki, 710-0046, Japan
2CREST, Japan Science and Technology Corporation
   Kawaguchi, 332-0012, Japan

(Received 11 February 2004, accepted 3 June 2004)

Centromere protein C (CENP-C) is a component of the kinetochore essential for correct segregation of sister chromatids in mammals. In Arabidopsis thaliana, a single-copy gene encoding a protein homologous to CENP-C has been found by homology in the whole-genome sequence. To investigate the CENP-C homolog (AtCENP-C), we cloned cDNAs by RT-PCR and determined its full-length coding sequence. Antibodies against the synthetic peptide for the C-terminal residues of AtCENP-C detected a polypeptide in Arabidopsis cell extracts on western blots. Immunofluorescence labeling with the antibodies and fluorescence in situ hybridization demonstrated clearly that AtCENP-C is present at the centromeric regions throughout the cell cycle.

Key words: Arabidopsis thaliana, CENP-C, centromere, 180-bp repeat, immunofluorescence

INTRODUCTION

The centromere is a region on eukaryotic chromosomes, required for proper segregation of chromatids during mitotic and meiotic cell division. At the centromere, a protein complex called the kinetochore is known to be assembled during mitosis and meiosis, and serves as the attachment site for spindle microtubules. Except for a few fungi including Saccharomyces cerevisiae, which has “point centromeres” ca. 125 bp long, and species with holocentric chromosomes, most eukaryotic organisms have larger, regional, centromeres (40 kb to several Mb) containing repetitive sequences, but centromeric DNA sequences are diverse among different species (Pluta et al. 1995, Clarke 1998). On the other hand, several kinetochoore proteins have sequence similarity among evolutionarily distant organisms (Dobie et al. 1999, Pidoux and Allshire 2000). In humans, CENP-A and CENP-C are components of the inner kinetochore plate, the region of the kinetochore closest to the centromeric heterochromatin, and are located at the centromeres throughout the cell cycle (Saitoh et al. 1992, Warburton et al. 1997). Although its exact function remains unknown, CENP-C is suggested to be necessary to induce formation of a functional centromere, by injection of anti-CENP-C antibodies into human cells (Tomkiel et al. 1994), gene disruption experiments in mice (Kalitsis et al. 1998) and conditional loss-of-function analysis in the chicken DT40 cell line (Fukagawa and Brown 1997). In plants, maize cDNAs for CENP-C homologs were first identified by their sequence similarity to human CENP-C (Dawe et al. 1999). Maize CENP-C proteins were present at centromeres throughout the cell cycle and colocalized with centromeric histone H3 (CENH3) in meiotic cells (Dawe et al. 1999, Zhong et al. 2002). In Arabidopsis thaliana, a hypothetical protein (Accession no. AAF71990) was identified to have global similarity to maize CENPC in amino acid sequence (Yu et al. 2000). However, the protein and its gene structure have not been characterized. In this study, we have cloned cDNAs corresponding to the A. thaliana CENP-C homolog (AtCENP-C), and investigated colocalization with the centromere-specific major satellites (180-bp family) by immunofluorescence labeling and fluorescence in situ hybridization (FISH).

MATERIALS AND METHODS

Plant materials. Arabidopsis thaliana (L.) Heynh. ecotype Columbia was used. For isolation of RNA from seedlings, the seeds were sown on a GM agar plate (Valvekens et al. 1988). Plants grown on soil in a greenhouse under natural light were used to isolate RNA and cell extracts from flower buds. A cell line of A. thaliana

* Corresponding author. E-mail: yogura@rib.okayama-u.ac.jp
ecotype Columbia was maintained as described (Mathur et al. 1998).

**RT-PCR and sequencing.** RNA was isolated using Plant RNAasy Mini Kit (Qiagen). RT-PCR was performed with gene-specific primers, F1 (5'-TAACCTGAAATTCCATCGCGG-3') and R1 (5'-CTGGAGGGCAAGCGAACGTGGA-3'), using ReverTra Ace (Toyobo Co., Ltd.) and ExTaq (Takara Bio Inc.). Amplified fragments were blunt-ended, phosphorylated, separated by gel electrophoresis and cloned into pBluescript (Stratagene). For direct sequencing, fragments were amplified with gene-specific primers, F2 (5'-ATGGCGTATGGAACGGCG-3') and R2 (5'-CTTAGATGGCCGCGCTGAC-3'), using QIAGEN OneStep RT-PCR Kit (Qiagen) and purified with SUPREC-PCR (Takara Bio Inc.). DNA sequences of the clones were determined using Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems) on an automated DNA sequencer (ABI Model 310; Applied Biosystems). Database searches were performed using the BLAST service provided by NCBI (Altschul et al. 1997). Multiple sequence alignments were done using the CLUSTALW network service provided by DDBJ (Thompson et al. 1994).

**Antibody production and western blot analysis.** Anti-AtCENP-C antibodies were made and purified by Resgen (Invitrogen Corp.). A proportion of a synthetic peptide (SKVKSFVSDEYKKLVD) was conjugated with keyhole limpet hemocyanin through the N-terminus and the rest through the C-terminus, using gluteraldehyde and carbodiimide chemistries. These components were mixed, used to immunize a rabbit, and the serum was affinity purified with the peptide linked to a column utilizing two different chemistries that attach the peptide both through the N-terminus and the C-terminus. Western blot analysis was performed using an ECL plus kit (Amersham Biosciences Corp.) according to the manufacturer's instructions. The recombinant AtCENP-C expressed in Escherichia coli was prepared as follows. Restriction enzyme sites at the initiation codon and the termination codon were introduced into an RT-PCR clone (1.8Ex12SK-2) and the cDNA clone (SQ114h08F) by PCR, respectively, and they were ligated to make “full length” cDNA for AtCENP-C. It was subcloned into pET23d (Novagen) and used to transform BL21-CodonPlus(DE3)-RIL (Stratagene). The recombinant AtCENP-C was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) and the E. coli was harvested by centrifugation. For isolation of nuclei, Arabidopsis cultured cells were treated with 1% Cellulase Onozuka RS (Yakult Honsha Co., Ltd.) and 0.1% Pectolyase Y-23 (Kyowa Chemical Products Co., Ltd.) in MM buffer (0.6 M mannitol, 20 mM MES pH 5.5). Then, protoplasts were washed with MM buffer twice, suspended in MGS buffer [1 M hexyleneglycol, 10 mM PIPES pH 7.0, 10 mM MgCl₂, 10 mM KCl, 0.4 M sucrose, 0.4% Triton X-100, 0.4 mM Phenylmethanesulfonyl fluoride (PMSF)] and homogenized. After centrifugation, the supernatant containing crude nuclei was loaded on 20%–40% Percoll (Amersham Biosciences Corp.) step gradient in MGS buffer and centrifuged. The lower fraction was diluted with two-volume of MGS buffer and purified nuclei were pelleted. Flower buds were ground in liquid nitrogen. Each sample was suspended in SDS loading buffer (125 mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 5% 2-mercaptoethanol), boiled and loaded on 10% polyacrylamide gels.

**Immunolabeling and FISH.** Immunolabeling and FISH were performed as described (Shibata and Murata 2004). Anti-AtCENP-C antibodies were prelabeled with Alexa Fluor 488 using the Zenon antibody labeling system (Molecular Probes) and used for immunofluorescence labeling. After immunofluorescence detection, 180-bp repeat sequences were detected by a FISH experiment on the same glass slides. The amplified 180-bp repetitive sequences were labeled with biotin using Biotin Nick Translation Mix (Roche) and detected with Streptavidin-Cy5 (Amersham Biosciences Corp.).

**RESULTS AND DISCUSSION**

In the A. thaliana genome, there is a gene (At1g15660) predicted to have eleven exons encoding a 710-aa protein, which has similarity with the whole maize CENPCA and with the plant of human CENP-C (Yu et al. 2000). Since BLAST searches revealed no other CENP-C-like sequence in the database, it was suggested to be a single-copy gene. In maize, genomic DNA blot analysis suggested that there are at least five different CENPC genes (Dawe et al. 1999), while CENP-C gene in human and mouse are thought to be present as a single-copy gene. We also found one EST (Accession no. AV559306) homologous to the gene by EST database searches. The cDNA clone (SQ114h08F) for the EST contained about a 1.2 kb long insert, and its coding sequence was identical to that predicted, but the N-terminal region corresponding to exon 1 to 5 and a part of exon 6 was missing (Fig. 1A). In order to clone the region, RT-PCR was performed with a set of primers F1 and R1, which are located at the upstream region of the initiation codon and within the coding region of the cDNA clone corresponding to exon 9, respectively (Fig. 1A). Using total RNA from seedlings or flower buds, we detected amplified fragments of expected size from both materials, and cloned that from the flower buds. The sequence of an RT-PCR clone was identical to the predicted one except for a deletion of the first 15 bp of exon 4. The other four RT-PCR clones also lacked the 15-bp sequence and the same result was obtained by direct sequencing of RT-PCR fragments amplified with another primer set using total RNA from flower buds and...
Characterization of *Arabidopsis* CENP-C homolog

As the 3’-terminal dinucleotide of the deleted sequence was AG, which conformed to the conserved dinucleotide for 3’ splice sites, the boundary of intron 3 and exon 4 was suggested to be 15 bp downstream of the predicted one (Fig. 1B). In the deduced amino acid sequence from the RT-PCR clone, a proline residue at position 137 of AAF71990 was changed to a leucine residue and the following five amino acids (GFETV: positions 138 to 142 of AAF71990) were deleted but no frame shift occurred compared with that from the predicted gene (Fig. 1B). When the deduced 705 amino acid sequence for AtCENP-C was used as a query for database searches, it had high similarities to maize CENP-Cs, the deduced amino acid sequence from rice cDNA and those from ESTs of several plant species, in the C-terminal region (Fig. 2).

Polyclonal antibodies were raised against a synthetic peptide (SKVKSFVSDYEYKLD) derived from the amino acid sequence near the C-terminus of AtCENP-C (Fig. 2), and were affinity purified with the peptide. Immunoblotting with the antibodies detected an approximately 105-kD band in extract from *E. coli* expressing recombinant AtCENP-C and a 70-kD band in total extract from flower buds or nuclear extract from cultured cells (Fig. 3). For human CENP-C and its homologs, the apparent molecular masses on western blots are greater than the predicted molecular masses (Saitoh et al. 1992, Dawe et al. 1999, Moore and Roth 2001). However, the 70-kD band in total and nuclear extracts was smaller than the predicted molecular mass for AtCENP-C (79 kD). This may result from its post translational modifications, processing or degradation during preparation.

To determine the subcellular localization of AtCENP-C with respect to centromeric 180-bp repeat, we performed immunolabeling and FISH on *Arabidopsis* cultured cells. Immonosignals were detected from anti-AtCENP-C on the centromeric regions of all chromosomes (Fig. 4). On prometaphase chromosomes the majority of AtCENP-C signals overlapped with signals of 180-bp sequences but a large portion of signals of 180-bp sequ-

---

**Fig. 1.** Structure of AtCENP-C gene and its cDNA. (A) Schematic representation of the predicted gene (At1g15660), the cDNA clone for the EST (SQ114h08F) and an RT-PCR clone (1.8Ex12SK-2), for AtCENP-C. Protein-coding sequences are indicated by filled boxes. The 3’ untranslated regions are indicated by open boxes. The exon numbers of the predicted gene are shown above the boxes. Arrowheads show the positions of primers used for RT-PCR. The asterisk indicates the region different from that of the predicted gene. (B) The upper sequence is the nucleotide sequence around the region of the gene, indicated by the asterisk in (A) and the lower sequence is the deduced amino acid sequence. Letters in upper case and lower case in the nucleotide sequence indicate predicted exons and introns, respectively. The nucleotide sequence of intron 3 removed from this figure is represented by dots. The 15-bp sequence not found in the RT-PCR clone (1.8Ex12SK-2) is shown in white on a black background. Vertical lines indicate splice junctions of intron 3 supported by the sequence of RT-PCR clones. A proline residue changed to a leucine residue is circled and five amino acid residues deleted by the 15-bp deletion are boxed. Number at either side of the nucleotide sequence is its position at AC013453 and that of the deduced amino acid sequence is its position at AAF71990.
ences did not overlap with AtCENP-C signals. Similar signal patterns were observed on metaphase chromosomes and interphase nuclei. This sort of differential colocalization with 180-bp repeats was also found in Arabidopsis CENP-A homolog (HTR12) (Shibata and Murata 2004). These findings indicate that AtCENP-C is present together with HTR12 at centromeres throughout the mitotic cell cycle.

In addition to highly conserved region I (Yu et al. 2000), the region used for production of anti-AtCENP-C antibodies was well conserved among these species (Fig. 2). The cross-reactivities of antibodies against maize CENPCA with the centromeric regions of metaphase chromosomes of barley and field bean have been reported, although the immunosignals of the latter were dispersed around the centromeres and along the arms (ten Hoopen et al. 2000). The anti-AtCENP-C antibodies produced in this study were thought to be applicable to detect CENP-C homologs in other plant species, but no signals appeared on chromosomes of tomato cells (Shibata et al. unpublished results).

In human cells, the relationship between CENP-C and CENP-B has been extensively studied (Sugimoto et al. 1999). It was shown that CENP-C remained as discrete dots, whereas CENP-B displayed larger surrounding materials in interphase nuclei, and that centromeric alphoid DNA, targeted by CENP-B, was highly dispersed, whereas CENP-C remained as small dots on the DNA fibers. As shown in Fig. 4, the localization of AtCENP-C to the centromeric 180-bp repetitive sequences was similar to that of human CENP-C to the centromeric alphoid DNA. Human CENP-C has an internal domain with DNA-binding activity in vitro but the target sequence has not been characterized (Sugimoto et al. 1994, Sugimoto et al. 1997). Human CENP-C also has a C-terminal dimerization domain, but AtCENP-C and other plant CENP-C homologs do not have regions corresponding to the domain and they did not show significant similarity to human CENP-C, except for region I. Human CENP-B binds to a specific 17-bp sequence, CENP-B box, which is
Characterization of Arabidopsis CENP-C homolog

found in some alphoid DNA (Masumoto et al. 1989), and chromatin immunoprecipitation suggests that CENP-A nucleosomes form a complex with CENP-B and CENP-C through interaction with DNA (Ando et al. 2002). Recently, Suzuki et al. (2004) also reported that human CENP-C interact with CENP-B and one of the two interaction domains of CENP-C included region I. Neither the DNA-binding activity of AtCENP-C nor expression of CENP-B homolog has been reported in *Arabidopsis thaliana*, but further analyses using the anti-AtCENP-C and other antibodies would provide information on the interaction(s) of AtCENP-C with specific DNA sequences and/or other proteins.

Nucleotide sequence data reported are available in the DDBJ/EMBL/GenBank databases under the accession numbers AB128986 (1.8Ex12SK-2) and AB128987 (SQ114h08F).

We thank J. S. (Pat) Heslop-Harrison and T. Schwarzacher for critical reading of the manuscript and S. Miyasako for excellent technical assistance. We also thank Kazusa DNA Research Institute for the cDNA clone (SQ114h08F), C. Koncz and M. Umeda for the cell culture of *Arabidopsis thaliana*.

**REFERENCE**


---

**Fig. 4.** Colocalization of AtCENP-C with the centromeric 180-bp repeats. Chromosomes and nucleus were stained with anti-AtCENP-C antibodies prelabeled with Alexa Fluor 488. Then they were hybridized with the biotin-labeled centromeric 180-bp sequences and detected with Streptavidin-Cy5. Immunosignals of AtCENP-C (green), FISH signals of the 180-bp sequences (red) on prometaphase chromosomes (left), metaphase chromosomes (middle) and interphase nucleus (right). Bottom panels are merged images of upper panels. DAPI-counterstained chromosomes and nucleus are shown in gray. Bars = 5 µm.


