ELSEVIER

Gene 262 (2001) 283–290

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Gene structure, chromosomal localization and immunolocalization of chicken centromere proteins CENP-C and ZW10☆

Atsushi Okamura[a,1], Carlos Pendon[b,1], Manuel M. Valdivia[b], Toshimichi Ikemura[a], Tatsuo Fukagawa[a,*]

[a]*Department of Evolutionary Genetics, National Institute of Genetics and The Graduate University for Advanced Studies, Mishima, Shizuoka-ken 411-8540, Japan*

[b]*Department of Biochemistry and Molecular Biology, Facultad de Ciencias, Universidad de Cadiz, 11510 Puerto Real Cadiz, Spain*

Received 28 July 2000; received in revised form 10 October 2000; accepted 24 October 2000
Received by T. Gojobori

Abstract

We determined the genomic structures and complete sequences of the coding regions of the chicken CENP-C and ZW10 genes. These two genes encode proteins that are thought to be involved in maintaining the fidelity of chromosome segregation. The chicken CENP-C gene is 30 kb in length and contains 19 exons. The chicken ZW10 gene spans 10 kb and contains 15 exons. The 5′-untranslated regions of these genes contain several binding sites for transcription factors such as Sp-1, E2F, p300, and members of the GATA family. By fluorescence in situ hybridization (FISH) analysis, the CENP-C was mapped to chromosome 4 and the ZW10 gene was mapped to a microchromosome. Antibodies against the chicken ZW10 protein revealed a cell cycle-dependent staining pattern in DT40 cells. ZW10 protein was distributed throughout the cytoplasm of DT40 cells during interphase. In most metaphase cells, ZW10 proteins appeared equally divided between the centromere and the spindle apparatus. During anaphase, chicken ZW10 proteins were no longer localized near chromosomes or the mitotic apparatus but were present diffusely in the cytoplasm.

Keywords: Centromere; DT40; Genomic Structure; ZW10; CENP-C

1. Introduction

The centromere plays a fundamental role in accurate chromosome segregation during mitosis and meiosis in eukaryotes. Some of its functions include sister chromatid adhesion and separation, microtubule attachment, chromosome movement and mitotic checkpoint control (reviewed in Choo, 1997). In the vertebrate centromere, the kinetochore is a protein-DNA complex to which microtubules attach during metaphase. Centromere proteins have been identified primarily by immunological studies in vertebrate cells. Antibodies against centromeric proteins have been isolated both from patients with autoimmune disease (Moroi et al., 1980) and from immunized animals (Compton et al., 1991). This has led to the identification of several protein antigens known as centromere proteins (CENPs) (Pluta et al., 1995). In addition, analyses of mutations that cause errors in chromosome segregation have identified several centromeric proteins. This method has been applied most extensively in the *Saccharomyces* and *Drosophila* and has led to identification of several genes that encode components of the centromere (reviewed in Choo, 1997). Some of these genes have homologues in other organisms. Our recent studies have focused on CENP-C and ZW10, two centromere proteins (Fukagawa and Brown, 1997; Fukagawa et al., 1999a).

CENP-C has been localized to the inner kinetochore plate, adjacent to the centromeric DNA (Saitoh et al., 1992), and it is known to bind DNA directly (Yang et al., 1996). Disruption of CENP-C gene results in embryonic lethality (Kalitsis et al., 1998a) due to chromosome missegregation and metaphase arrest (Fukagawa and Brown, 1997; Fukagawa et al., 1999a). Metaphase arrest has been observed after microinjection of anti-CENP-C antibodies

Abbreviations: FISH, fluorescence in situ hybridization; CENP, Centromere protein, DAPI, 4′ 6-Diamidine-2-phenylindole-dihydrochloride; PBS, phosphate buffed saline

☆ The nucleotide sequence data reported in this paper have been submitted to DDBJ/EMBL/NCBI nucleotide sequence database with accession numbers: AB042323 (ZW10 gene) and AB042324 (CENP-C gene).

* Corresponding author. National Institute of Genetics, Mishima, Shizuoka-ken, 411-8540 Japan Tel.: +81-559-81-6792; fax: +81-559-81-6794.

E-mail address: tfukagaw@lab.nig.ac.jp (T. Fukagawa).

[1] These authors contributed equally to this work.

0378-1119/01/$ - see front matter
PII: S0378-1119(00)00517-5

into HeLa cells (Tomkiel et al., 1994). CENP-C protein shares a region of homology with Mif2p, an essential centromeric protein in *Saccharomyces cerevisiae* (Brown, 1995; Meluh and Koshland, 1995). CENP-C is found only at active centromeres (Earnshaw et al., 1989; Sullivan and Schwartz, 1995) and is necessary to induce formation of a functional centromere (Fukagawa et al., 1999a). These studies suggest that CENP-C is an important component for kinetochore assembly.

ZW10 was originally identified as the protein product of the *Drosophila melanogaster l(1) zw10* gene, hereafter called ZW10 (Smith et al., 1985). Mutations in the ZW10 gene disrupt chromosome segregation during mitosis and meiotic division (Williams et al., 1992). Recent evidence suggests that *Drosophila* ZW10 may act as part of the spindle checkpoint; sister chromatids separate prematurely when treated with the microtubule-depolymerizing drug in ZW10 mutants (Williams et al., 1996). ZW10 homologues are present in diverse species including *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Mus muscuclus* and *Homo sapiens* (Starr et al., 1997). This conservation suggests that ZW10 plays an important role in proper chromosome segregation.

Our experimental system is the chicken DT40 cell line (Buerstedde and Takeda, 1991). It is useful for the study of cell-autonomous functions by reverse genetic analyses because modification of endogenous genes by homologous recombination is very efficient in DT40 cells. In previous work, we used homologous recombination to engineer a conditional loss-of-function mutation into the CENP-C gene (Fukagawa and Brown, 1997; Fukagawa et al., 1999a). Further experiments involving sophisticated genetic approaches, such as temperature sensitive mutations, in the DT40 system require knowledge of the entire genomic structures of the target genes. The genomic DNA structures of the chicken CENP-C and ZW10 genes have not been described, previously. Here, we report the genomic structures and complete genomic sequences of the coding regions of the chicken CENP-C and ZW10 genes. We also show a cell cycle-dependent staining pattern for chicken ZW10 protein in DT40 cells.

2. Materials and methods

2.1. Isolation and characterization of genomic clones

To obtain genomic clones, we used cDNAs of CENP-C and ZW10 as probes. Hybridization screening and purification of genomic clones from DT40 genomic libraries were performed as described previously (Fukagawa and Brown, 1997; Fukagawa et al., 1999a). For the determination of gene structure the phage clones were subcloned into pBluescript (Stratagene) or pUC118 (Takara) and subjected to standard restriction mapping and Southern blot analyses.

2.2. DNA sequencing

Shotgun DNA sequencing of genomic regions was performed with the ABI Fluorescent BigDye Terminator Cycle Sequencing kit (Perkin-Elmer). Samples were sequenced with an ABI 377 automated sequencer (Perkin-Elmer). Assembly of sequence data for each gene was done with ATGC software (SDC). The nucleotide sequences of CENP-C (27630 nt) and of ZW10 (17419 nt) have been deposited in DDBJ, EMBL and NCBI under Accession Numbers AB042324 and AB042323, respectively.

2.3. Fluorescence in situ hybridization (FISH) and immunocytochemistry

Metaphase chromosomes were prepared by addition of hypotonic solution and fixed in methanol/acetic acid (3:1). FISH was done with a standard method (Fukagawa et al., 1999b). Biotin-labeled probe was detected with Cy3- or FITC-conjugated avidin diluted with 4× SSC, 0.05% Tween 20, 1% BSA. Chromosomes were counterstained with DAPI at 0.2 μg/ml. Chromosome numbers were determined with chromosome specific painting probes.

Immunofluorescent staining for whole cells was performed as described previously (Fukagawa et al., 1999a). Cells were collected onto slides with a cyto-centrifuge, fixed in 3% paraformaldehyde in PBS for 15 min at room temperature, permeabilized in 0.5% NP-40 in PBS for 15 min at room temperature, rinsed three times in 0.5% BSA and incubated for 1 h at 37°C with anti-ZW10 antibody. The antibody was then detected with Cy3-conjugated goat anti-rabbit IgG.

3. Results

3.1. Isolation and characterization of the ZW10 gene

The chicken ZW10 gene was isolated from a phage library made from the genomic DNA of DT40 cells. Approximately 2×10^6 phage plaques were screened with the ZW10 cDNA (Fukagawa et al., 1999a) as a probe. We identified five independent positive plaques. For determination of the gene structure, the phage clones were subcloned into plasmid vectors and subjected to restriction mapping, Southern hybridization analysis and sequencing. We sequenced a 17 kb region including coding and 5′- and 3′-flanking regions of the ZW10 gene. The contiguous 17,419 nt sequence was deposited into DDBJ, EMBL and NCBI under Accession No. AB042323. Pair-wise alignment of this genomic sequence with the cDNA sequence revealed that the complete chicken ZW10 gene spanned approximately 10 kb and contained 15 exons (Fig. 1A). The exons vary in length from 72 to 581 bp, and the intron sizes range from 106 to 997 bp. Sequence analysis across the intron/exon junctions showed that the splice acceptor and donor sites contain dinucleotide consensus sequences

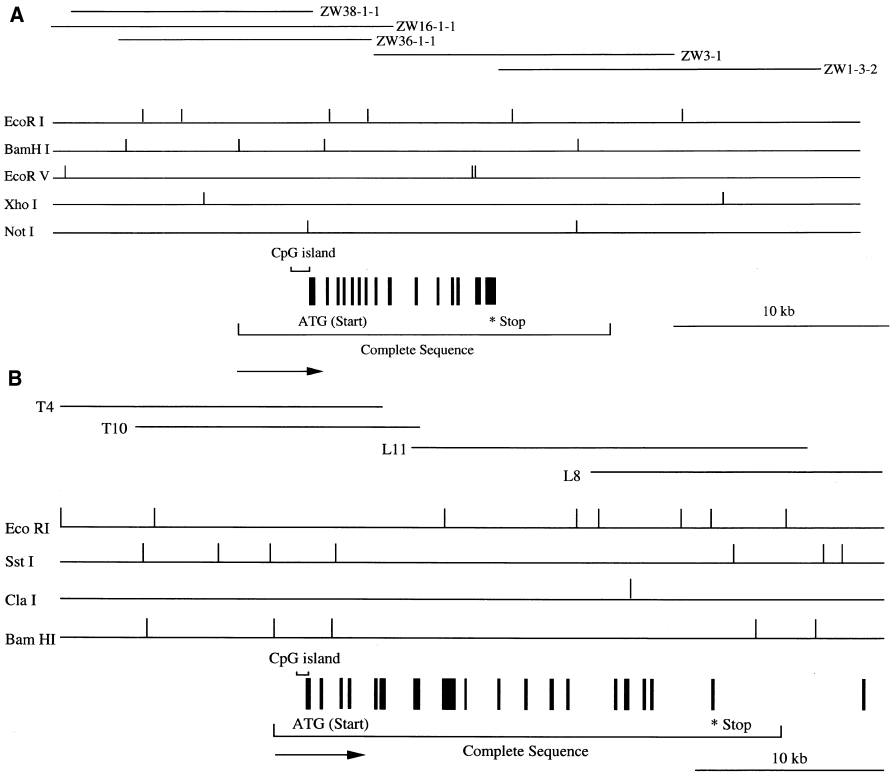

Fig. 1. (A) Organization of the chicken ZW10 gene showing positions of exons 1 to 15. Black boxes represent exons. The arrow indicates the direction of transcription of the gene. Relative positions of the genomic phage clones and the EcoRI, BamHI, EcoRV, XhoI and NotI restriction sites are shown. We sequenced the 17,419 nt genomic region presented here. (B) Organization of the chicken CENP-C gene showing positions of exons 1 to 19. Black boxes represent exons. The arrow indicates the direction of transcription of the gene. Relative positions of the genomic phage clones and the EcoRI, SstI, ClaI and BamHI restriction sites are shown. We sequenced the 27,630 nt genomic region presented here.

AG and GT, respectively (Table 1). The genomic sequence includes a 5′-untranslated region (UTR) that contains a CpG island and is quite GC-rich; the GC% of the -500 to -1 region is 71.4%, whereas the GC% of the coding region is 45.0%. Examination of a transcription-factor binding-site profile database revealed sites for several transcriptional regulatory factors, including ubiquitous factors such as Sp-1 and E2F and tissue/cell specific factors such as members of the GATA family, p300 and Elk-1 (Fig. 2A). We also sequenced the 3′-flanking region of the ZW10 gene

Table 1
Exon-intron organization of chicken ZW10 gene

Exon			Sequence at exon-intron junction and intron size (nt)		
NO.	Position	nt	5′ splice donor	nt	3′ splice acceptor
1	3390–3715	326	GAA CGA Ggt cac cgc	644	att tta gGT TCA ACG
2	4360–4461	102	GCA GGA Ggt gag gag	189	gtt gca gTT TGA TAC
3	4651–4728	78	GGA AAA Ggt gat gat	106	cca cca gGC ACG AAG
4	4835–4994	160	TCT AAA Ggt act cgc	221	tct tta gAA ACC ATT
5	5216–5365	150	ACT TTT Ggt aag gca	284	tcc tca gGC AAG TTG
6	5650–5841	192	CTG CTG Agt aag ttg	331	aat tca gAT GTG CCT
7	6173–6336	164	TAT AGA Ggt agg tct	286	att tca gGT GAT TAA
8	6623–6805	183	TGT AAA Ggt act gct	454	ccc tta gAT CAC GCC
9	7260–7495	236	ATC AGT Ggt agg aat	997	ttt gta gCT GCA TAC
10	8493–8564	72	ACC ACA Agt aag tgt	642	gag aca gAG AGA ACC
11	9207–9376	170	AGA CTT Ggt aat ttt	643	cct aaa gGG ATG GAG
12	10020–10150	131	AAG GCA Ggt agg caa	372	att cca gGT ATT GCA
13	10523–10654	132	CCT AGA Ggt aac tgc	549	att cca gGA TAT CTC
14	11204–11406	203	TAG ATC Agt agg tgt	276	ctc gta gGT GGG CGG
15	11683–12263	581			

and performed BLAST searches of this region against the NCBI non-redundant database. We found the chicken homologue of hepsin, a novel trypsin-like serine protease gene (Leytus et al., 1988) in this region. We also examined the copy number of ZW10 in the chicken genome with Southern hybridization analysis of chicken genomic DNA digested with several restriction enzymes. Five micrograms of high-molecular-weight DNA from DT40 cells was digested completely with each enzyme and size-fractionated by electrophoresis on a 1% agarose gel. Southern blots of gels containing chicken DNA were hybridized with several radiolabeled cDNA probes. Hybridization patterns were consistent with those predicted from the sequence (data not shown). Extra bands were not detected. We thus conclude that the ZW10 sequence is present only once in the haploid chicken genome.

3.2. Isolation and characterization of the CENP-C gene

To facilitate studies of the structural and functional properties of CENP-C, we determined the complete genomic sequence for the coding region of chicken CENP-C. We had previously isolated four independent genomic phage clones that covered the entire CENP-C region (Fukagawa and Brown, 1997). For determination of the gene structure, the genomic phage clones were subcloned into plasmid vectors and subjected to restriction mapping, Southern hybridization analysis and sequencing. We sequenced 27 kb that include the coding and 5′-flanking regions of the CENP-C gene. This sequence does not contain the 3′-flanking region that was assigned by hybridization using a cDNA probe specific to the 3′-UTR of the CENP-C gene (e.g. last exon in Fig. 1B). The contiguous 27,630 nt sequence was deposited in DDBJ, EMBL and NCBI under Accession No. AB042324. Pair-wise alignment of this genomic sequence the reported cDNA sequence (Fukagawa and Brown, 1997) revealed that the complete chicken CENP-C gene spans approximately 30 kb and contains 19 exons (Fig. 1B). The exons vary in length from 39 to 838 bp, and introns range from 0.1 to 8 kb. All intron/exon junctions follow the GT/AG rule (Table 2). The 5′-UTR of the CENP-C gene contains a CpG island and is quite GC-rich; the GC% of the -500 to -1 region is 62.6%, whereas the GC% of the coding region

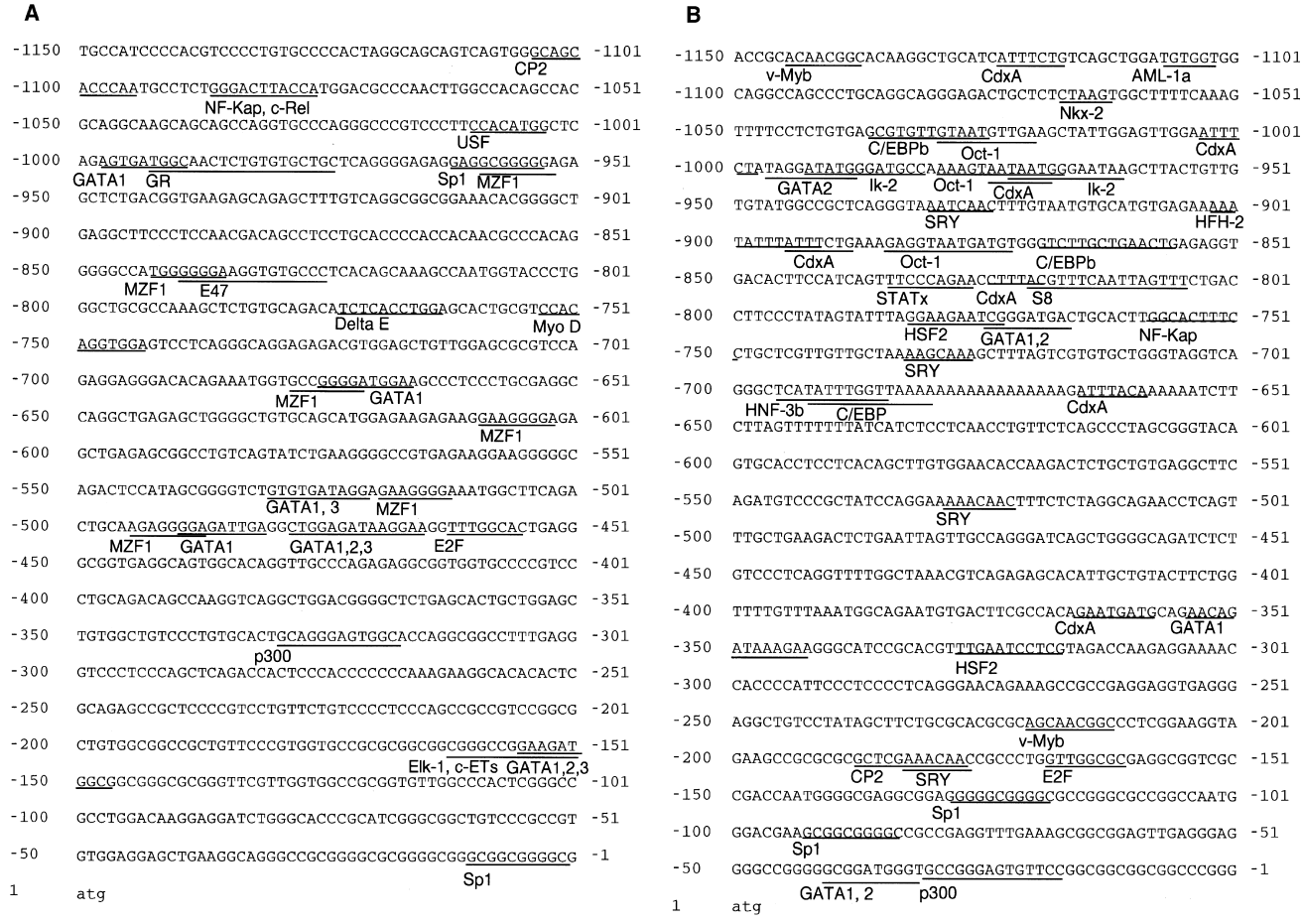

Fig. 2. Transcription-factor binding sites in the regulatory regions of the ZW10 (A) and CENP-C (B) genes. The nucleotide sequence is numbered, 1 corresponding to the 'a' of the initiating atg codon. Potential transcription-factor binding-sites are underlined.

Table 2
Exon-intron organization of chicken CENP-C gene

Exon			Sequence at exon-intron junction and intron size (nt)		
NO.	Position	nt	5′ splice donor	nt	3′ splice acceptor
1	1821–1865	45	GCG CTT Ggt gag aga	768	tta aca gGA TCA CTT
2	2634–2680	47	AAG GAG Ggt aat ttc	1063	ttt cta gAA AGA AGA
3	3744–3805	62	TTT GAA Tgt aag cct	326	ttt gta gGT GAT GAT
4	4132–4229	98	CCC TAG Ggt aag ttg	1358	atg tca gCC AGA CGC
5	5588–5681	94	AGC AAA Ggt tgg tag	94	gtg gca gGG GAG CCA
6	5776–5995	220	TAC AAA Cgt aag tag	1615	tct tta gGC GTA CGG
7	7611–7990	380	CAT GCA Ggt atg tca	1045	ttc cta gGA AAA TCT
8	9036–9873	838	GAC TGC Agt aag cta	359	ctt cta gAC ATT CCC
9	10233–10271	39	CTG AAA Ggt aat ttc	1845	ttt gaa gCG GGC TTG
10	12117–12190	74	TCT TCA Ggt act att	1257	ttt rca gAT AAT TCT
11	13448–13516	69	AAA ATA Ggt aag ctt	1359	aaa tta gTG CTG CCT
12	14876–14980	105	CCT TCA Ggt aca agt	763	ctt tca gGA AGG CTT
13	15744–15839	96	AAA ACA Agt atg aat	2516	ccc tta gGA AGT GAG
14	18356–18451	96	CTT CTA Ggt aag gta	329	cct gca gAG TGC ATT
15	18781–18935	155	TAC AAT Agt aag tac	910	tat taa gGC TTT CTA
16	19846–19936	91	CCA GCA Ggt aag tgc	354	ttt ata gGA AAT GGA
17	20291–20363	73	ATG ACA Ggt gaa gaa	3149	gtt cca gGG CTC CTG
18	23513–23630	118	TTG CTG Tgt gtg tca	>3968	
19	> 27630				

is 40.3%. Examination of the transcription-factor binding-site profile database revealed a number of potential binding sites. Search results from the profile database revealed the presence of two Sp-1 binding sites, one E2F binding site, one p300 binding site and binding sites for several GATA family members (Fig. 2B). We also tested copy number for CENP-C using the procedure outlined for the ZW10 gene. Our results indicate that CENP-C is a single-copy gene in the chicken genome (data not shown).

3.3. Chromosome in situ hybridization

To determine map positions and confirm the copy numbers of these genes, we used FISH analysis. The chicken genome is comprised of 39 chromosome pairs, 29 of which are microchromosomes. It is difficult to map a gene to a specific microchromosomes. The DT40 cell line is euploid for all chromosomes except chromosome 2, for which is present in three copies. We first used phage

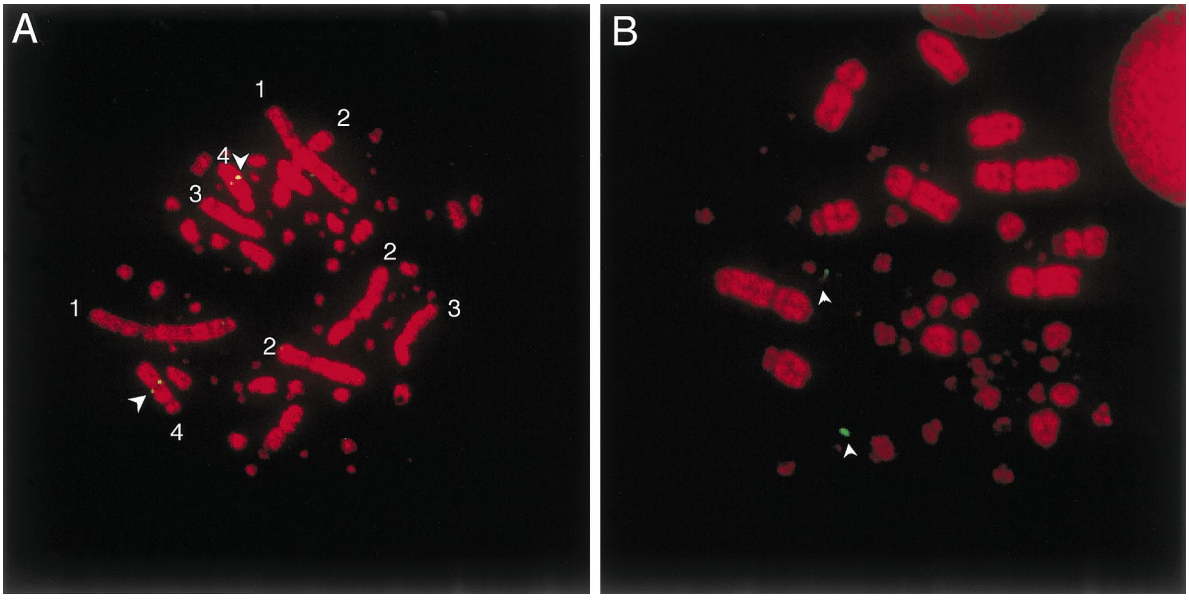

Fig. 3. Chromosomal localization of CENP-C (A) and ZW10 (B) genes by fluorescent in situ hybridization. DT40 metaphase spreads were hybridized with probes for CENP-C and ZW10. Arrows indicate probe-specific hybridization signals (green). DNA is stained with DAPI and coloured Red. CENP-C was mapped to the middle of the long arm of chromosome 4 (A), and ZW10 was mapped to an unspecified microchromosome (B).

clone L11, which contains the CENP-C sequence as a probe. We examined 100 typical (pro) metaphase spreads. Of these, approximately 90% showed complete double spots. Fluorescent signals due to L11 hybridization were localized near the center of the long arm of chromosome 4, and doublet signals were not observed elsewhere (Fig. 3A). We also used phage clone ZW38-1-1, which contains the ZW10 sequence, for FISH analysis. Fluorescent signals due to ZW38-1-1 hybridization were observed on microchromosomes (Fig. 3B). Because the microchromosomes are very small and similar, it was difficult to determine the exact number of chromosomes with ZW10 signal.

3.4. Immunolocalization of chicken ZW10 in DT40 cells

We previously showed that chicken CENP-C was localized at the centromere throughout the cell cycle, as is human CENP-C (Saitoh et al., 1992; Fukagawa et al., 1999a). Localization of human ZW10 is different from that of *Drosophila* ZW10, however. Though both *Drosophila* and human ZW10 are found primarily at kinetochores in early anaphase, human ZW10 is lost from the kinetochore and becomes cytoplasmic as anaphase progresses (Starr et al., 1997), whereas *Drosophila* ZW10 remains at the kinetochore throughout anaphase. Thus, we investigated the subcellular distribution of chicken ZW10 in the DT40 cell line. We used anti-chicken ZW10 antibody produced by us for the immunolocalization of the protein, and observed a cytoplasmic distribution of the protein during interphase (Fig. 4A). In most metaphase cells, ZW10 proteins appeared equally divided between the centromere and the spindle apparatus (Fig. 4B, C). Fig. 4C shows double-staining of cells with anti-tubulin and anti-ZW10. Some ZW10 signals were associated with kinetochore microtubules, and the others were localized to the spindles. In some metaphase cells, ZW10 proteins remained only at the centromere (Fig. 4D). We hypothesized that these cells were not ready for proper chromosome segregation and that the spindle checkpoint pathway is activated. To determine when chicken ZW10 appears at the centromere, cells were treated with colcemid, which cause depolymerization of microtubules and stained with anti-chicken ZW10 antibody. Anti-ZW10 antibody clearly showed that the protein was localized to the centromere in all chromosome spreads from DT40 cells arrested in metaphase by colcemid treatment (Fig. 4F). We also examined the distribution of ZW10 in whole cells treated with colcemid and found that ZW10 was localized to the centromere (data not shown). Throughout anaphase, ZW10 was no longer localized near the chromosomes or the mitotic apparatus but was present diffusely in the cyto-

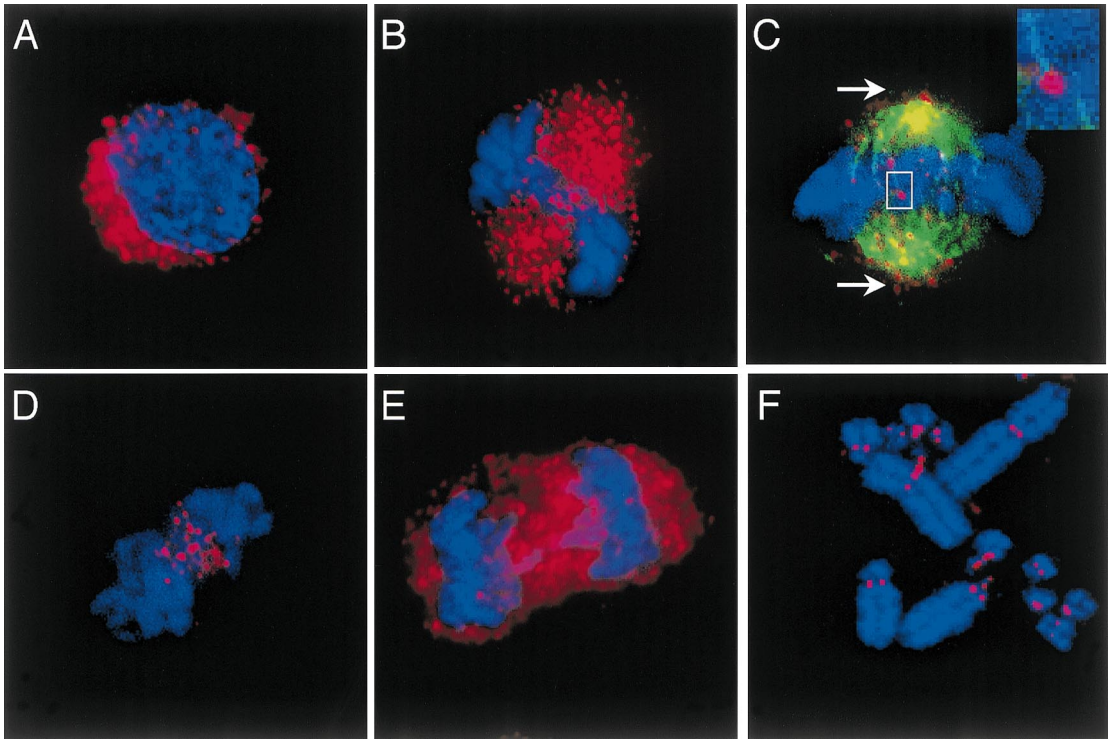

Fig. 4. Immunolocalization of chicken ZW10 in DT40 cells. DNA is blue and ZW10 signals are red. (A) In interphase cells, ZW10 is distributed in cytoplasm. (B) In metaphase cells, ZW10 appears to be divided between the centromere and the spindle apparatus. (C) Double-staining of DT40 cells with tubulin (green) and ZW10 (red). Some ZW10 signals were associated with kinetochore microtubules, and others were localized to the spindles. Inset shows that kinetochore microtubules capture ZW10. Arrows indicate the positions of the spindle poles. (D) In a metaphase cells, the proteins are localized only at the centromeres. (E) In anaphase cells, chicken ZW10 are no longer localized to chromosomes or the mitotic apparatus, but instead are present diffusely in the cytoplasm. (F) Immunolocalization of ZW10 protein in metaphase spreads. Metaphase chromosomes were prepared from cells treated with colcemid. Signals are localized to centromeres.

plasm (Fig. 4E). Although discrete sites of human ZW10 in the vicinity of the centromere in very early anaphase have been reported (Starr et al., 1997), we did not observe centromere localization of chicken ZW10, even in early anaphase.

4. Discussion

4.1. The chicken ZW10 gene

The genomic structures and complete sequences of the coding regions of the chicken CENP-C and ZW10 genes were successfully identified. These two proteins are thought to be involved in maintaining the fidelity of chromosome segregation. The chicken ZW10 gene spans 10 kb and contains 15 exons. The chicken gene has high homology to the human and mouse gene with overall identities of 69 and 81%, respectively. These high similarities suggest that ZW10 plays an important role in centromere function that has been conserved across species. Although there is no report of the genomic organization of the ZW10 gene in other organisms, a 5′-UTR sequence of the human cDNA has been reported (Starr et al., 1997). The human gene and chicken gene are organized differently in the first exon containing the 5′-UTR; the position of the initiating ATG codon is different. The first 150-nt region from initiating ATG codon of the human gene corresponds to 5′-UTR of the chicken gene, though we can find significant homology at the DNA level in this region. There are two possible explanations for this difference. One is alternative splicing; the human and chicken ZW10 genes may be translated from different regions. Very recently, we found a chicken genomic sequence in the DDBJ/EMBL/NCBI databases (Accession No. AJ250458) that contained the 5′-region of ZW10. The authors predicted some exons of the chicken ZW10 gene based on the human cDNA sequence. We could not find their predicted exons in our cDNA clones, however. There may be other transcripts expressed by alternative splicing in other type cells. We confirmed that our cDNA was functional in the DT40 cell line by a transfection experiment (data not shown). The second possibility is a sequencing error; this region contains many GC-repeat sequences, and it is easy to misread the sequencing gels.

We also examined subcellular distribution of chicken ZW10 in the DT40 cell line. During interphase, chicken ZW10 proteins were distributed in the cytoplasm. A similar phenomenon occurs with human and *Drosophila* ZW10 (Starr et al., 1997). During metaphase, ZW10 proteins appeared to be divided equally between the centromere and the spindle apparatus. However in some metaphase cells, ZW10 proteins remained only at the centromere. We believe that this distribution is related to the function of the ZW10. Recent evidence suggests that ZW10 may act as part of the spindle checkpoint in *Drosophila* cells (Williams et al., 1996). If ZW10 participates in the spindle checkpoint pathway, it is likely involved in monitoring the state of chromosome alignment or spindle assembly. Our hypothesis is that ZW10 protein is localized to the centromere during early metaphase and that it moves to the spindle after monitoring chromosomes and the spindle assembly. To test our theory, DT40 cells were treated with colcemid and stained with anti-chicken ZW10 antibody. In colcemid-treated cells, ZW10 protein localized to the centromere and not to the spindle. These findings are consistent with our hypothesis and suggest that ZW10 may monitor spindle assembly. Recently, Basu et al. reported a relation between ZW10 and checkpoint proteins Bub1 and Bub3 (Basu et al., 1998). Localization of the Bub3 to the centromere requires Bub1 but not ZW10 in *Drosophila*. ZW10 may function independent of the Bub pathway. A conditional knockout approach using the DT40 system will clarify the function of ZW10 in the checkpoint pathway.

Although localization of human ZW10 in the vicinity of the centromere in very early anaphase has been reported (Starr et al., 1997), we did not observe centromere localization of chicken ZW10, even in early anaphase. We can not explain why distribution of ZW10 is slightly different in chicken cells compared with human cells.

4.2. The chicken CENP-C gene

Compared to ZW10, the chicken CENP-C gene is larger, spanning a distance of approximately 30 kb and containing 19 exons. The chicken CENP-C protein was only 23% identical to human CENP-C over its entire length. In the region homologous to *S. cerevisiae* Mif2p (Brown, 1995), however, chicken CENP-C shows significant homologies to human CENP-C and *S. cerevisiae* Mif2p with identities of 66.7 and 34.6%, respectively. There is no report of the genomic organization of human CENP-C, but Kalitsis et al. reported the organization of the mouse CENP-C gene, which spans 60 kb and contains 19 exons (Kalitsis et al., 1998b). Although the intron-exon organization is not conserved between chicken and mouse, organization is conserved in the region homologous to *Mif2* (exons 12–18 of the chicken gene). This evolutionarily conserved organization suggests that the *Mif2* homology region plays a critical role in CENP-C function.

4.3. Long-range GC% mosaic structure and chicken chromosomes

Genomes of higher vertebrates are composed of long segments of DNA containing a mosaic structure of varying GC content (Bernardi et al., 1985; Ikemura, 1985; Fukagawa et al., 1995). Gene density, transcription patterns, DNA replication timing and other chromosomal behaviors are related to this long-range GC% mosaic structure. GC% of the ZW10 sequence (17419 nt) in this paper is 51.3%, which is relatively GC-rich for the chicken genome. We found a chicken homologue of hepsin, a novel trypsin-like serine protease gene (Leytus et al., 1988), very close to the 3′-UTR of ZW10. McQueen et al. reported that GC-rich regions, gene rich regions and CpG islands are concentrated

on microchromosomes in the chicken genome (McQueen et al., 1996). We mapped the chicken ZW10 gene to a microchromosome using FISH analysis. In contrast, the GC% of the CENP-C genomic region is 40.9%, and CENP-C was mapped to a macrochromosome (chromosome 4). This finding was consistent with that of McQueen et al. (1996). Long-range GC% mosaic structure is related to several biological features including expression timing of genes. The structural differences between ZW10 and CENP-C may be related to their functions.

4.4. Further experiments in DT40 cell lines

The genomic characterization of the chicken CENP-C and ZW10 genes and their promoter regions presented here should assist in further studies of their roles in centromere function. Chicken cell line DT40 is useful for the study of cell-autonomous functions by gene-knockout analysis because modification of endogenous genes by homologous recombination is very efficient in this cell line. We previously produced a conditional knockout cell line for CENP-C (Fukagawa and Brown, 1997; Fukagawa et al., 1999a). Further experiments with sophisticated genetic approaches, such as temperature sensitive mutations, in the DT40 system will require knowledge of the entire genomic structures of the target genes. These genomic sequences will be useful for designing point mutations or conditional gene disruptions in future experiments using DT40 cells.

Acknowledgements

The authors are very grateful to Ms Y. Miyauchi, Ms K. Suzuki and Ms K. Kita for technical assistance. This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas (A) 'Genome Science' and (C) 'Genome Biology' and 'Cancer Biology' from the Ministry of Education, Science, Sports and Culture of Japan.

References

Basu, J., Logarinho, E., Herrman, S., Bousbaa, H., Li, Z., Chan, G.K., Yen, T.J., Sunkel, C.E., Goldberg, M.L., 1998. Localization of the *Drosophila* checkpoint control protein Bub3 to the kinetochore requires Bub1 but not Zw10 or Rod. Chromosoma 107, 376–385.

Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotivial, M., Rodier, F., 1985. The mosaic genome of warm-blooded vertebrates. Science 228, 953–958.

Brown, M.T., 1995. Sequence similarities between the yeast chromosome segregation protein Mif2 and the mammalian centromere protein CENP-C. Gene 160, 111–116.

Buerstedde, J.M., Takeda, S., 1991. Increased ratio of targeted to random integration after transfection of chicken B cell lines. Cell 67, 179–188.

Choo, K.H.A., 1997. 'The Centromere'. Oxford University Press, Oxford, New York, Tokyo, pp. 1–11.

Compton, D.A., Yen, T.J., Cleveland, D.W., 1991. Identification of novel centromere/kinetochore-associated proteins using monoclonal antibodies generated against human mitotic chromosome scaffolds. J. Cell Biol. 112, 1083–1097.

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

Fukagawa, T., Sugaya, K., Matsumoto, K., Okumura, K., Ando, A., Inoko, H., Ikemura, T., 1995. A boundary of long-range G + C% mosaic domains in the human MHC locus: pseudoautosomal boundary-like sequence exists near the boundary. Genomics 25, 184–191.

Fukagawa, T., Brown, W.R.A., 1997. Efficient conditional mutation of the vertebrate CENP-C gene. Hum. Mol. Genet. 6, 2301–2308.

Fukagawa, T., Pendon, C., Morris, J., Brown, W., 1999a. CENP-C is necessary but not sufficient to induce formation of functional centromere. EMBO J. 18, 4196–4209.

Fukagawa, T., Hayward, N., Yang, J., Azzalin, C., Griffin, D., Stewart, A.F., Brown, W., 1999b. The chicken HPRT gene: a counter selectable marker for the DT40 cell line. Nucleic Acids Res. 27, 1966–1969.

Ikemura, T., 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. 2, 13–34.

Kalitsis, P., Fowler, K.J., Earle, E., Hill, J., Choo, K.H.A., 1998a. Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. Proc. Natl. Acad. Sci. USA 95, 576–582.

Kalitsis, P., MacDonald, A.C., Newson, A.J., Hudson, D.F., Choo, K.H.A., 1998b. Gene structure and sequence analysis of mouse centromere proteins A and C. Genomics 47, 108–114.

Leytus, S.P., Loeb, K.R., Hagen, F.S., Kurachi, K., Davie, E.W., 1988. A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells. Biochemistry 27, 1067–1074.

McQueen, H.A.M., Fantes, J., Cross, S.H., Clark, V.H., Archibald, A.L., Bird, A.P., 1996. CpG island of chicken are concentrated on microchromosomes. Nat. Genet. 12, 321–324.

Meluh, P.B., Koshland, D., 1995. Evidence that the Mif2 gene of Saccaromyces cervisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. Mol. Biol. Cell 6, 793–807.

Moroi, Y., Peebles, C., Fritzler, M.J., Steigerwald, J., Tan, E.M., 1980. Autoantibody to centromere (kinetochore) in scleroderma sera. Proc. Acad. Sci. USA 77, 1627–1631.

Pluta, A.F., Mackay, A.M., Ainsztein, A.M., Goldberg, I.G., Earnshaw, W.C., 1995. The centromere: hub of chromosomal activities. Science 270, 1591–1594.

Saitoh, H., Tomkiel, J., Cooke, C.A., Ratrie, H., Maurer, M., Rothfield, N.F., Earnshaw, W.C., 1992. CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. Cell 70, 115–125.

Smith, D.A., Baker, B.S., Gatti, M., 1985. Mutations in genes controlling essential mitotic functions in *Drosophila melanogaster*. Genetics 110, 647–670.

Starr, D.A., Williams, B.C., Li, Z., Etemad-Moghadam, B., Dawe, R.K., Goldberg, M.L., 1997. Conservation of centromere/kinetochore protein ZW10. J. Cell Biol. 138, 1289–1301.

Sullivan, B.A., Schwartz, S., 1995. Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of functional centromeres. Hum. Mol. Genet. 4, 2189–2197.

Tomkiel, J., Cooke, C.A., Saitoh, H., Bernat, R.L., Earnshaw, W.C., 1994. CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. J. Cell Biol. 125, 531–545.

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

Williams, B.C., Karr, T.L., Montgomery, J.M., Goldberg, M.L., 1992. The *Drosophila l(1)zw10* gene product, required for accurate chromosome segregation, is redistributed at anaphase onset. J. Cell Biol. 118, 759–773.

Williams, B.C., Gatti, M., Goldberg, M.L., 1996. Bipolar spindle attachments after redistributions of ZW10, a *Drosophila* centromere/kinetochore component required for accurate chromosome segregation. J. Cell Biol. 134, 1127–1140.