Identification of an evolutionarily conserved insulator element at the 3' boundary of the imprinted *Igf2/H19* domain

Ko Ishihara

Doctor of Philosophy

Department of Genetics

School of Life Science

The Graduate University for Advanced Studies

2001

Contents

Abbreviations

BAC: bacterial artificial chromosome

bp: base pair

ChIP: chromatin immunoprecipitation

CTCF: CCCTC-binding factor

DMR: differentially methylated region

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

EGTA: ethylene glycol-bis (β-aminoethyl ether) N,N,N',N',-tetraacetic acid

EMSA: electrophoretic mobility shift assay

HEPES: N-2-hydroxyethylpiperazine-N'-ethanesulfonic acid

Igf2: insulin-like growth factor II

kb: kilo base pair

L23mrp: L23 (mitochondrial)-related protein

PAGE: polyacrylamide gel electrophoresis

PCR: polymerase chain reaction

PIPES: 1,4-piperazinediethanesulfonic acid

PMSF: phenylmethylsulfonyl fluoride

SDS: sodium dodecyl sulfate

Su(Hw): suppressor of hairly-wing

UV: ultraviolet

YAC: yeast artificial chromosome

Abstract

Igf2 and *H19* are closely linked imprinted genes, which lie at the centromeric end of a 1-Mb imprinted domain on mouse chromosome 7. *L23mrp* and other genes located 3' (more centromeric) to *H19* are not imprinted and do not interact with the enhancers shared by *Igf2* and *H19*. It is therefore suggested that the intergenic region between *H19* and *L23mrp* contains a boundary or an insulator element. I have identified a binding site for CTCF, a nuclear factor that mediates insulator activity in vertebrates, in the intergenic region. This site is conserved in human and mouse, associated with a major DNase I hypersensitive site, and bound by CTCF *in vivo*. Functional assays using reporter constructs demonstrated that this element functions as an insulator in transfected cells. The findings suggest that this CTCF site contributes to the 3' boundary of this imprinted domain. Thus, together with the findings on the differentially methylated CTCF sites 5' to *H19*, CTCF-dependent insulators not only regulate but also delimit the imprinted domain.

Introduction

The eukaryotic genome is partitioned into independent functional domains that are separated by DNA sequences called boundary elements or insulator elements (for a review, see Udvardy, 1999; Bell *et al.*, 2001). These elements ensure the correct expression of genes within the domain by blocking influences from regulatory elements in the neighboring domains. Thus, insulator elements are defined by their ability to block enhancer-promoter interactions when positioned between the enhancer and the promoter, or to protect transgenes from position effects. The best characterized insulators include the Su(Hw)-binding sites within the *gypsy* retrotransposon of *Drosophila* (Geyer and Corces, 1992), the *scs* and *scs'* elements flanking the 87A7 *hsp70* locus of *Drosophila* (Udvardy *et al.*, 1985), as well as the chicken β-globin insulators (Chung *et al.*, 1993; Saitoh *et al.*, 2000). Previously, Bell *et al.* (1999) showed that a protein named CCCTC-binding factor (CTCF) binds to the core sequence of the chicken β-globin insulator and that this factor plays a key role in enhancer-blocking activity. Moreover, other insulators in vertebrates also contain CTCF sites and show CTCF-dependent insulator activity (Bell *et al.*, 1999; Saitoh *et al.*, 2000; Bell and Felsenfeld, 2000; Hark *et al.*, 2000; Antes *et al.*, 2001; Filippova *et al.*, 2001).

The *Igf2* and *H19* genes are closely linked imprinted genes, which are expressed only from the paternal and the maternal allele, respectively (DeChiara *et al.*, 1991; Bartolomei *et al.*, 1991; for a review, see Sasaki *et al.*, 2000). Interestingly, the reciprocal imprinting of the two genes is dependent upon methylation-sensitive, CTCF-dependent insulators within the differentially methylated region (DMR) located 5' to *H19* (Bell and Felsenfeld, 2000; Hark *et al.*, 2000; Kanduri *et al.*, 2000b). These genes are located at the centromeric end of a 1-Mb imprinted domain, which contains at least 14

imprinted genes, in mouse chromosome band 7F4/F5. It is known that *L23mrp* (*Rpl23*) and other genes located 3' (more centromeric) to *H19* are not imprinted. In addition, a previous study by fluorescence *in situ* hybridization identified a transition from asynchronous replication at *H19* to synchronous replication at *L23mrp* (Greally *et al.*, 1998). Moreover, it was shown that the endoderm-specific enhancers located between *H19* and *L23mrp* interact with both *Igf2* and *H19* (Leighton *et al.*, 1995), but not with *L23mrp* (Zubair *et al.*, 1997). These findings suggest that an insulator element is present between *H19* and *L23mrp*.

In the present study, I have looked for an insulator element within the *H19/L23mrp* intergenic region based on sequence homology with the known CTCF-dependent insulators. Our study identified an evolutionarily conserved CTCF-binding site that showed an enhancer-blocking activity in transfected cells. The results suggest that this CTCF-dependent insulator serves to define the 3' boundaries of this imprinted domain.

Materials and methods

Isolation of nuclei

Cell nuclei from 12.5-dpc mouse embryos were isolated as described (Sasaki et al. 1992). Mouse embryos were homogenized using an all-glass Dounce homogenizer (20 strokes) in five volumes of nuclear buffer (10 mM Tris-HCl [pH 7.5], 60 mM KCl, 5 mM $MgCl_2$, 15 mM NaCl, 0.1 mM EGTA, 0.5 mM dithiothreitol (DTT), 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 2 µg/ml Aprotinin, 2 µg/ml Pepstatin A, 2µg/ml Leupeptin) containing 0.3 M sucrose (nuclear buffer/0.3 M sucrose). The homogenate was filtered through a pad of cheesecloth, layered over a 4-ml cushion of nuclear buffer/1.8 M sucrose. Nuclei were then pelleted at 18,000 rpm for 20 min in a Beckman SW41Ti rotor. Pelletted nuclei were resuspended in a small volume of nuclear buffer/0.3 M sucrose/5% glycerol and stored at -80°C.

Electrophoretic mobility shift assay (EMSA)

Nuclei isolated from embryos were resuspended in an equal volume of buffer C (10 mM HEPES-KOH [pH 7.6], 0.4 M KCl, 0.1 mM EDTA, 3 mM $MgCl_2$, 1 mM DTT, 0.2 mM PMSF, 2 µg/ml Aprotinin, 2 µg/ml Pepstatin A, 2 µg/ml Leupeptin), mixed gently for 30 min, and centrifuged for 30 min at 15,000 g. The supernatant was dialyzed overnight against buffer D (20 mM HEPES-KOH [pH 7.6], 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 20% glycerol, 0.2 mM PMSF, 2 µg/ml Aprotinin, 2 µg/ml Pepstatin A) at 4°C and centrifuged for 10 min at 15,000 g. The supernatant, designated as nuclear extract, was frozen in aliquots and stored at -80°C. EMSA was carried out using the Gel Shift Assay Core System (Promega). Double-stranded probes were labeled with [γ-^{32}P]ATP (3000 Ci/mmol) and T4 polynucleotide kinase. The DNA binding reaction was performed for 20

min at 25°C in a volume of 10-15 μl, containing 10 μg of the nuclear extract, 35 fmol of ^{32}P-labeled probe, 2-3 μl of 5 X binding buffer (1 X buffer: 10 mM Tris-HCl [pH 7.5], 50 mM NaCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 4% glycerol, 0.05 mg/ml poly(dI-dC):poly(dI-dC), 0.5 mM DTT) with or without 50-fold molar excess of unlabeled competitor. The reaction mixture was electrophoresed on a 4% polyacrylamide gel in 0.5 X TBE. For super-shift assays, 2 μl of the anti-CTCF antibody (Upstate biotechnology) was added to the reaction mixture and incubated at 4°C for 1.5 hr before the addition of radiolabeled probes. The sequences of the probes were shown in Table I. Methylation of duplexes was carried out by use of *Sss*I methylase (New England Biolabs) according to the supplier's instructions. The methylation reaction was monitored by digestion of the duplexes with the methylation-sensitive restriction enzyme *Aci*I (CCGC).

DNase I hypersensitive site assay

Nuclei equivalent to 20 μg of DNA were resuspended in 90 μl of nuclear buffer/0.3 M scrose/5% glycerol. The samples were combined with 10 μl of the same solution containing 0.25, 0.5, 1, 2, 3, or 4 U/μl of DNase I (Roche) (final concentration, 25, 50, 100, 200, 300, or 400 U/ml, respectively) and digested at 25°C for 5 min. The reaction was terminated by adding an equal volume of 20 mM EDTA [pH 8.0], 1% SDS containing 0.5 mg/ml of proteinase K. Proteinase K digestion was carried out overnight at 37°C. DNA was purified by phenol-chloroform extraction and ethanol precipitation. The purified DNA was digested with appropriate restriction enzymes and subjected to agarose gel electrophoresis, followed by blotting onto Byodine B membrane (Pall) and UV-cross-linking by Stratalinker (Stratagene). Probe DNAs were ^{32}P-radiolabeled and used for hybridization in Church's solution at 65°C. The membranes were washed twice with 2 X SSC, 0.1% SDS at 65°C and then twice in 0.2 X SSC, 0.1% SDS at 65°C. The

membranes were then exposed to X-ray films (Kodak) at -80°C.

Chromatin immunoprecipitation (ChIP) assay

The nuclei (3 X 10^6 nuclei) isolated from 12.5-dpc mouse embryos were resuspended in 360 µl of NB1 buffer (0.3 M sucrose, 15 mM Tris-HCl [pH 7.5], 60 mM KCl, 15 mM NaCl, 5 mM MgCl2, 0.1 mM EGTA) containing protease inhibitors (1 mM PMSF, 2 µg/ml aprotinin, 2 µg/ml leupeptin, 2 µg/ml pepstatin A). To cross link proteins to DNA, 10 µl of formaldehyde (37%) was added to the nuclear suspension and incubated at 25°C for 10 min. The nuclei were pelleted and then resuspended in 200 µl of nuclear wash buffer (5 mM PIPES-KOH [pH8.0], 85 mM KCl, 0.5% NP40) containing protease inhibitors for 10 min on ice. The nuclei were collected by centrifugation and resuspended in 200 µl of SDS lysis buffer (1% SDS, 50 mM Tris-HCl [pH 8.1], 10 mM EDTA) containing protease inhibitors for 10 min on ice. The lysate was sonicated to reduce the size of DNA to an average of about 600 bp. Insoluble material was removed by centrifugation. The supernatant was diluted 10-fold with dilution buffer (0.01% SDS, 16.7 mM Tris-HCl [pH 8.1], 1.1% Triton X-100, 2 mM EDTA, 167 mM NaCl) containing protease inhibitors and precleared at room temperature for 30 min with 60 µl of salmon sperm DNA/Protein A agarose slurry (Upstate Biotechnology). Two microliter of anti-CTCF antibody (Upstate Biotechnology) was added to the precleared supernatant and the mixture was incubated at room temperature for 1 hr, followed by an addition of 60 µl of salmon sperm DNA/Protein A agarose slurry (Upstate Biotechnology) and additional incubation at room temperature for 1 hr. Agarose beads were then collected and washed sequentially for 3-5 min each in L buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl [pH 8.1], 150 mM NaCl), H buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl [pH 8.1], 500 mM NaCl), and LiCl buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholate, 10

mM Tris-HCl [pH 8.1], 1 mM EDTA). Beads were then washed twice with TE buffer and extracted with 500 µl of 1% SDS/0.1 M $NaHCO_3$. The eluted samples were heated at 65°C overnight to reverse the formaldehyde cross-linking. Following proteinase K digestion, DNA was recovered by phenol-chloroform extraction and ethanol precipitation. PCR was carried out for 25 cycles with 50 µCi of [α-^{32}P]dCTP. The PCR products were separated by PAGE and visualized by autoradiography. Primers used were: ChIP4-1 (PCT4, up), 5'-AAA AGG TGC CCA TCT TGA TGG CTG-3'; ChIP4-2 (PCT4, down), 5'-TTT CTG ACT CTC CTG ATA CCA TGT-3'; ChIP12-1 (PCT12, up), 5'-GGT GGA GGA AGG CGC CAT GTG G-3'; ChIP12-2 (PCT12, down), 5'-CTG ACT TCA GGA GGG TCT GGG ACT-3'; ChIP6/14-1 (PCT6/14, up), 5'-TAG AAT CAC TCC AAC TGG CAT GTC-3'; ChIP6/14-2 (PCT6/14, down), 5'-TAA TAC CAG CTA CAT GAG ATC CTG-3'; m3-s (*H19* DMR m3, up), 5'-CTG TTA TGT GCA ACA AGG GAA-3'; m3-a (*H19* DMR m3, down), 5'-GGT CTT ACC AGC CAC TGA-3'; OLG-1 (*H19* exon 5, up), 5'-GTG AAG CTG AAA GAA CAG ATG GTG-3'; OLG-5 (*H19* exon 5, down), 5'-AAG CAC ACG GCC ACA CCC AGT-3'.

Bisulfite methylation assay

Genomic DNA (2µg) from 14.5-dpc mouse embryos was digested overnight with *Stu*I restriction enzyme in a volume of 20 µl. Bisulfite treatment was carried out as described by Paulin *et al.* (1998). The DNA was denatured by the addition of 2 µl of 3 M NaOH and incubated at 37°C for 15 min. For the bisulfite reaction, 208 µl of 6.24 M urea/2 M sodium metabisulfite (pH5.0) and 12 µl of 10 mM hydroquinone were added to the denatured DNA. The reaction was performed in a 0.5 ml PCR tube overlaid with 100 µl of mineral oil and subjected to 20 cycles of 55°C for 15 min followed by denaturation at 95°C for 30 sec on the DNA Thermal Cycler 480 (Perkin Elmer). Modified DNA was purified using the

Wizard DNA Clean-Up System (Promega) according to the manufacturer's instructions and eluted into 45 µl of water. The reaction was completed by treatment with 5 µl of 3 M NaOH at 37°C for 15 min, followed by the addition of 20 µl of 10 M ammonium acetate. The bisulfite-reacted DNA was collected by ethanol precipitation and amplified by PCR. The PCR products were subcloned into pBluescript plasmids (Stratagene) and sequenced using Big Dye Terminator Cycle Sequencing Kit and ABI PRISM 377 Sequencer (Perkin Elmer). Primers used were: bisP1 (PCT6/14, upper strand, up), 5'-TTG GTA TGT TAG TTG GTT TTG GTG ATG GG-3'; bisP2 (PCT6/14, upper strand, down), 5'-ACC TAA CTC CTA TCC TCA ATC CCA ATA AAT-3'; bisP3 (PCT6/14, lower strand, up), 5'-ATA ACT CTT CCA AAA CCC TAA CCA TCC TAA-3'; bisP4 (PCT6/14, lower strand, down), 5'-GAG ATT TTG GTT GGT AGA AGA ATA ATA GTA G-3'; bisP5 (PCT4, upper strand, up), 5'-ATT TTG ATG GTT GGT TTT GTT AGG GGT AAA-3'; bisP6 (PCT4, upper strand, down), 5'-TCC TTT CTA ACT CTC CTA ATA CCA TAT AAA-3'.

Colony assay

The reporter plasmid pHN consisted of a Neomycin-resistance gene (*Neo*) driven by the *H19* promoter (–818 bp to +6 bp) and a 1.8-kb *Aat*II-*Hin*dIII fragment containing the *H19* DMR insulator. This insulator should block the influence from the adjacent regions. A 2.5-kb *Nsi*I-*Bgl*II genomic fragment, which harbors the endoderm-specific enhancers (CS3 and CS4) for *H19* and *Igf2*, was placed downstream of the *neo* gene in pHN to generate pHNE. Test fragments were excised from the cosmid cDH2 (Ishihara *et al.*, 1998) by appropriate restriction enzymes, treated with T4 DNA polymerase to produce blunt ends and subcloned into the pBluescript plasmids at the *Eco*RV site. The fragments were liberated by double digestion with *Xho*I and *Spe*I and inserted into the

*Xho*I/*Spe*I sites of pHNE between the *neo* gene and the enhancers. Test fragment for pHNME was generated by introducing base substitutions into the CTCF consensus sequence at PCT12 (from 5'-CTG CCC CCT TTA GG-3' to 5'-AGT AAA AAG GGC TT-3') of a plasmid clone carrying a 1.5-kb *Xho*I-*Pma*CI fragment by a PCR-mediated mutagenesis method described by Imai *et al.* (1991). Primers used were: 12mut-1 (up), 5'-TTT TAC TTA GAG GAG CAA GCA TGC CCA-3'; 12mut-2 (down), 5'-AGG GCT TTA GCC CAA GGC TCA GAA CCA -3'. To make pHNEPX, the 1.5-kb *Xho*I-*Pma*CI fragment was ligated to the 3' end of the enhancer fragment and brought into pHN. The reporter constructs were linearized with *Mlu*I, and 0.2 pmol of each construct was transfected into 1.2 X 10^6 Hep3B cells using Lipofectamine Plus (Gibco BRL), together with 0.2 pmol of a plasmid carrying a hygromycin-resistance gene linearized with *Xho*I. After 48 hrs, cells were replated into two separate 90mm-dishes and subjected to drug selection by G418 (800 μg/ml) and hygromycin-B (250 μg/ml), respectively. Colonies were counted after 2 weeks of selection. The number of G418-resistant colony was corrected for transfection efficiency based on the number of hygromycin-resistant colony and normalized to that obtained with pHNE.

Results

Derivation of CTCF consensus sequence

To identify insulator elements within the *H19/L23mrp* intergenic region, I decided to look for putative binding sites for CTCF. The sequences of the CTCF sites identified prior to this study were much diverse. Among the CTCF sites, the chicken β-globin FII sequence (Bell *et al.*, 1999) and multiple (4 in mouse and 6 in human) CTCF sites within DMR at the *H19* locus (Bell and Felsenfeld, 2000; Hark *et al.*, 2000) have been studied in detail and the essential core region for CTCF-binding have been identified. I therefore compared these sequences and derived a 14-bp consensus sequence for the core region (5'-CCGCNNGGNGGCAG-3') (Figure 1).

Identification of sequences that can bind CTCF in vitro

To look for putative CTCF-binding sites, I scanned a 33-kb region between *H19* and *L23mrp* (GenBank no. AF049091, bases 8137-41680; GenBank no. AP003183, bases 8167-41712) (Figure 2A) for the CTCF consensus sequence with a criterion of 12 matches or more in the 14 nucleotides. As a result, I identified 6 candidate CTCF sites (PCT1-PCT6) on the upper strand and 9 candidates (PCT7-PCT15) on the lower strand (Figure 2A and B).

To determine which one of the candidate sequences can bind CTCF *in vitro*, electrophoretic mobility shift assays (EMSAs) were carried out with 84-bp duplex probes containing the sequences (Table I). When the probes were incubated with nuclear proteins from 12.5-dpc mouse embryo, all probes formed at least one, but usually more, complexes (Figure 3). Among the complexes, however, the major complexes formed with PCT4, PCT12 and PCT6/14 (PCT6 and PCT14 were located close each other;

Figure 5A) showed the same mobility as the CTCF complex formed with the m1 probe from the *H19* DMR (Bell and Felsenfeld, 2000). The three complexes were all super-shifted with polyclonal anti-CTCF antibodies (Figure 3). Also, formation of the complexes was greatly diminished by competition with excess unlabeled FII fragments but not with Sp1 consensus duplexes (Figure 4). These findings show that the PCT4, PCT12 and PCT6/14 sequences can specifically bind CTCF *in vitro*. A second complex formed with the PCT4 probe probably contain Sp1 because formation of this complex was inhibited by addition of unlabeled Sp1 consensus duplexes or unlabeled FII fragments (which has one Sp1 site) as competitors (Figure 4).

Cross-species conservation of the putative CTCF sites

To further assess the significance of the four potential CTCF sites (PCT4, PCT6, PCT12 and PCT14), I asked whether these sites are conserved through evolution. When I examined the human *H19* region (Genbank no. AF087017, bases 10988-40560; Genbank no. AC004556, bases 15016-20800) for these sequences, two sites corresponding to PCT12 and PCT14 were present at orthologous positions (Figure 5A). The other two potential CTCF sites as well as the 11 non-CTCF-binding sites were not conserved.

To examine whether the conserved human sequences can bind CTCF, I carried out EMSA experiments with a nuclear extract from 12.5-dpc mouse embryo. The major complexes formed with the human probes (hPCT12 and hPCT14, Table I) showed the same mobility as the CTCF complex formed with the m1 probe (data not shown). These complexes were competed with excess unlabeled FII fragments but not with Sp1 consensus duplexes (Figure 5B). Furthermore, the complexes were super-shifted with anti-CTCF antibodies (Figure 5C). Thus, the two conserved sequences (PCT12 and

PCT14) from both human and mouse are capable of binding CTCF *in vitro*.

DNase I hypersensitivity at PCT12

The known CTCF-dependent insulators, such as the chicken β-globin insulator (Bell *et al.*, 1999; Saitoh *et al.*, 2000) and mouse *H19* DMR insulator (Hark and Tilghman, 1998; Khosla *et al.*, 1999), are all associated with DNase I hypersensitive sites in chromatin. I therefore examined the DNase I sensitivity of the mouse PCT4, PCT12 and PCT6/14 regions. I first tested a 7.7-kb region containing both PCT4 and PCT12 (Figure 6A). Nuclei isolated from 12.5-dpc mouse embryo were treated with increasing concentrations of DNase I and the DNA was purified and analyzed by Southern blotting. End fragments from the 7.7-kb *Eco*RV region were used as probes (Figure 6A, top). The study revealed three distinct hypersensitive sites (Figure 6A, bottom): one located at, or very close to PCT12 and the others located in the middle and at the 3' end of the second exon of the muscle-specific transcription unit *Nctc1* (Ishihara *et al.*, 1998). No hypersensitive site was detected at or around PCT4 (Figure 6A, bottom).

I next analyzed the DNase I sensitivity of the PCT6/14 region. Two probes (PPm1 and BIP1) derived from the ends of a 4-kb *Pst*I fragment containing PCT6/14 were used (Figure 6B, top). Two hypersensitive sites were detected and mapped at both ends of CS9 (Figure 6B, bottom), which is a mesoderm-specific enhancer (Ishihara *et al.*, 2000). However, I did not observe any DNase I hypersensitive site associated with PCT6/14 (Figure 6B, bottom). These results suggest that PCT12 may be the only CTCF-dependent insulator that works *in vivo*.

In vivo binding of CTCF to PCT12

To know whether PCT12 and other potential CTCF sites are bound by CTCF *in vivo*, I

carried out chromatin immunoprecipitation (ChIP) assays with nuclei isolated from 12.5-dpc mouse embryo. I treated the nuclei with formaldehyde to cross-link protein with DNA, fragmented the chromatin by sonication, and carried out immunoprecipitation with anti-CTCF antibodies. By PCR using specific primers, I examined whether PCT12 and the other potential CTCF sites are co-immunoprecipitated with CTCF. As shown in Figure 7, the PCT12 region, as well as the positive control (m3) from the *H19* DMR, was greatly enriched in the anti-CTCF immunoprecipitates. In contrast, PCT4, PCT6/14 regions and a negative control region (*H19* exon 5) were not enriched. The results clearly demonstrate that, among the candidates, only PCT12 is bound by CTCF in the chromatin of mouse embryo.

Methylation prevents CTCF to bind to PCT4 and PCT6/14

It is known that, if CpG dinucleotides within the recognition sequence for CTCF are methylated, the factor cannot bind to this target (Bell and Felsenfeld, 2000; Hark *et al.*, 2000). PCT4, PCT6 and PCT14 each contained one CpG dinucleotide. To examine the effect of methylation on CTCF-binding, I carried out EMSA with *in vitro* methylated PCT6/14 duplexes (Figure 8). The CTCF complex formed with PCT6/14 disappeared when competed with an excess of unlabeled PCT6/14 duplexes but not with methylated duplexes (Figure 8). This indicates that the binding affinity of CTCF to PCT6/14 is greatly reduced by CpG methylation.

Subsequently, methylation status of these three sites in 14.5-dpc mouse embryo was analyzed by bisulfite genomic sequencing. As shown in Figure 9A and B, the unique CpG site in both PCT6 and PCT14 was highly methylated. Methylation was observed in 17 out of 19 sequenced clones for PCT6 (89%) and 16 out of 19 for PCT14 (84%). Similarly, the single CpG site in PCT4 was methylated in 10 out of 11 sequenced clones

(91%) (Figure 10A and B). It seems, therefore, likely that these three sites were highly methylated in 14.5-dpc embryo, suggesting that the lack of CTCF-binding to PCT4 and PCT6/14 *in vivo* is due to methylation of the target sequences.

.

PCT12 has an insulator activity

The only candidate left for an insulator is PCT12. I therfore examined its enhancer-blocking activity by a colony assay. Various constructs shown in Figure 11 were respectively transfected into Hep3B cells and their enhancer-blocking activity was assayed by counting the number of G418-resistant colony. When the *H19* DMR insulator (a control fragment), which has 4 CTCF sites, was inserted between the promoter and the enhancers (pHNIE), the colony number decreased to approximately 10%. I then tested a 1.5-kb *Xho*I-*Pma*CI fragment containing PCT12 (pHNPXE) and observed that the colony number decreased to approximately 40%. This suggests that PCT12 indeed has an enhancer-blocking activity. Similar reduction in colony number was observed in pHNXPE, which was identical to pHNPXE except that the fragment was inserted in the opposite direction. Thus the enhancer-blocking activity was independent of orientation of the fragment. To exclude that the reduced colony number was due to silencer activity, I relocated the fragment to a position 3' to the enhancer (pHNEPX). This construct affected the colony number only slightly, indicating that the fragment possesses little silencer activity. Lastly, to ask if the insulator activity was dependent on the CTCF site, base substitutions were introduced at all 14 nucleotides of the PCT12 sequence (pHNME). The result suggested that the insulator activity of the PCT12 fragment is largely (>60%) dependent upon the CTCF site, but some other sequences within the fragment could also be contributory to the insulator activity.

Discussion

I have identified a novel CTCF-dependent insulator (PCT12) in the *H19/L23mrp* intergenic region based on sequence homology with the known CTCF-binding sites. This sequence element is conserved between human and mouse, bound by CTCF *in vivo*, and exhibited an insulator activity in transfected cells. The existence of this insulator is consistent with the transition from asynchronous replication at *H19* to synchronous replication at *L23mrp* (Greally *et al.*, 1998) and with the lack of interaction of the endoderm-specific *H19* enhancers with the *L23mrp* promoter (Zubair *et al.*, 1997). Although elucidation of the exact function of this insulator awaits germline deletion in mice, it appears to be an important component of the 3' (centromeric) boundary of the imprinted *Igf2/H19* domain.

Previous studies showed that the reciprocal imprinting of *Igf2* and *H19* is regulated by a cluster of methylation-sensitive CTCF-dependent insulators in the DMR located 5' to *H19* (Bell and Felsenfeld, 2000; Hark *et al.*, 2000; Kaduri *et al.*, 2000b) and the multiple tissue-specific enhancers located 3' to *H19* (Leighton *et al.*, 1995; Ishihara *et al.*, 2000). It is proposed that the methylation status of the DMR determines the activity of the insulators and decides which gene is to be activated by the 3' enhancers on each parental chromosome. Figure 12 summarizes the locations of the DMR insulators, the tissue-specific enhancers, and the 3' insulator identified in this study (top). This domain structure is reminiscent of the chicken β-globin domain, which is flanked by a CTCF-dependent insulator on each side (Bell *et al.*, 1999; Saitoh *et al.*, 2000) (Figure 12, bottom). Thus CTCF-dependent insulators may be a common component of the domain boundaries in the vertebrate genome.

The imprinted domain defined by the 3' insulator includes a non-imprinted

transcription unit, *Nctc1* (Ishihara *et al.*, 1998) (Figure 12, top). This transcription unit is active only in the adult skeletal muscle. Some of the genes located in this 1-Mb imprinted domain escape imprinting, with a common feature of not being associated with CpG islands (for example, *Tssc6* and *Th*). The lack of a CpG island associated with *Nctc1* may be the reason why this transcription unit escapes imprinting even though it is located on the imprinted side of the insulator.

I previously identified multiple tissue-specific enhancers in the *H19/L23mrp* intergenic region (Ishihara *et al.*, 2000). Among these, CS9, which is a mesoderm-specific enhancer, is located more 3' to the 3' insulator (Figure 12, top). Other studies involving BAC and YAC transgenes showed that additional enhancers for expression in the heart, kidney and lung should be present further 3' (Ainscough *et al.*, 2000; Kaffer *et al.*, 2000). These findings pose a potential problem that the far 3' enhancers must act on *Igf2* and *H19* over the 3' insulator. I consider the following three possibilities. Firstly, despite my data that the insulator activity is independent of orientation in transfected cells, it may act in an orientation-dependent manner in the genomic context and/or in an *in vivo* situation. Then the enhancers on the 3' side may interact with the two genes. There are some examples for such an orientation-dependent insulator (Robinett *et al.*, 1997; Kanduri *et al.*, 2000a). Secondly, there is evidence that ehnancer-blocking activity of an insulator is dependent upon enhancer-promoter combination (Hagstrom *et al.*, 1996; Scott *et al.*, 1999) and thus certain enhancer-promoter pairs may not be blocked by CTCF-dependent insulators. Thirdly, a promoter targeting sequence (PTS) (Zhou and Levine, 1999), which can override the activity of an insulator to mediate interactions between genes and enhancers, may be present in a region 3' to the 3' insulator. Then the more 3' enhancers should be able to activate *Igf2* and *H19* through this element.

Mouse mutant known as *minute* (*Mnt*), which originally occurred in a population

of radiation induced mutagenesis, exhibits dwarfism only upon paternal transmission (Cattanach *et al.*, 2000). Linkage analysis mapped the region responsible for *Mnt* in distal chromosome 7. A molecular study of *Mnt* has revealed that the mutation involves an inversion of the region immediate 5' to the 3' insulator reported here to several megabases 3' to the imprinted domain (Davies *et al.*, personal communication). It is tempting to speculate that if the 3' insulator function as a 3' boundary of the imprinted domain, its removal from the authentic position would deteriorate imprinting of the genes in the domain. The fact that imprinting of *Igf2* and *H19* is disrupted in *Mnt* mice (Davies *et al.*, personal communication) supports this notion.

My study suggests that a sequence-based approach is effective in identifying CTCF-dependent insulators, despite that the factor binds to a diverse range of sequences by combinatorial use of the 11 zinc fingers (for a review, see Ohlsson *et al.*, 2001). In addition to the CTCF-dependent insulator that we identified, however, this region may contain other insulators as well. For example, one putative CTCF site that we identified in this study was conserved between human and mouse (PCT14). Although *in vivo* evidence for this site being an insulator is lacking, it could display insulator activity in a restricted tissue or at a specific developmental time. Also, there may be sites that show little homology with the consensus sequence but can in fact bind CTCF. Moreover, insulators that interact with unknown factors could be present. Further studies are needed to reveal all the sequence elements that contribute to the boundary function of this intergenic region.

Acknowledgements

I am deeply grateful to Prof. H. Sasaki for invaluable discussion and suggestions throughout this work and H. Shirohzu and T. Sado for their advice and discussion. I would like to express my sincere appreciation to all the members of the Prof. Sasaki's laboratory.

References

Ainscough,J.F-X., Dandolo,L. and Surani,M.A. (2000) Appropriate expression of the mouse *H19* gene utilises three or more distinct enhancer regions spread over more than 130 kb. *Mech. Dev.*, **91**, 365-368.

Antes,T.J., Namciu,S.J., Fournier,R.E. and Levy-Wilson,B. (2001) The 5' boundary of the human apolipoprotein B chromatin domain in intestinal cells. *Biochemistry*, **40**, 6731-6742.

Bartolomei,M.S., Zemel,S. and Tilghman,S.M. (1991) Parental imprinting of the mouse *H19* gene. *Nature*, **351**, 153-155.

Bell,A.C. and Felsenfeld,G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature*, **405**, 482-485.

Bell,A.C., West,A.G. and Felsenfeld,G. (1999) The protein CTCF is required for the enhancer-blocking activity of vertebrate insulators. *Cell*, **98**, 378-396.

Bell,A.C., West,A.G. and Felsenfeld,G. (2001) Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science*, **291**, 447-450.

Cattanach,B.M., Peters,J., Ball,S. and Rasberry,C. (2000) Two imprinted gene mutations: three phenotypes. *Hum. Mol. Genet.*, **9**, 2263-2273.

Chung,J.H., Whiteley,M. and Felsenfeld,G. (1993) A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. *Cell*, **74**, 505-514.

DeChiara,T.M., Robertson,E.J. and Efstratiadis,A. (1991) Parental imprinting of the mouse insulin-like growth factor II gene. *Cell*, **64**, 849-859.

Filippova,G.N., Thienes,C.P., Penn,B.H., Cho,D.H., Hu,Y.J., Moore,J.M., Klesert,T.R., Lobanenkov,V.V. and Tapscott,S.J. (2001) CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.*, **28**, 335-343.

Geyer,P.K. and Corces,V.G. (1992) DNA position-specific repression of transcription by a Drosophila zinc finger protein. *Genes Dev.*, **6**,1865-1873.

Greally,J.M., Starr,D.J., Hwang,S., Song,L., Jaarola,M. and Zemel,S. (1998) The mouse *H19* locus mediates a transition between imprinted and non-imprinted DNA replication patterns. *Hum. Mol. Genet.*, **7**, 91-95.

Hagstrom,K., Muller,M. and Schedl,P. (1996) Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the *Drosophila* bithorax complex. *Genes Dev.*, **10**, 3202-3215.

Hark,A.T., Schoenherr,C.J., Katz,D.J., Ingram,R.S., Levorse,J.M. and Tilghman,S.M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature*, **405**, 486-489.

Hark,A.T. and Tilghman,S.M. (1998) Chromatin conformation of the *H19* epigenetic mark. *Hum. Mol. Genet.*, **7**, 1979-1985.

Imai,Y., Matsushima,Y., Sugimura,T. and Terada,M. (1991) A simple and rapid method for generating a deletion by PCR. *Nucleic Acids Res.*, **19**, 2785.

Ishihara,K., Kato,R., Furuumi,H., Zubair,M. and Sasaki,H. (1998) Sequence of a 42-kb mouse region containing the imprinted *H19* locus: identification of a novel muscle-specific transcription unit showing biallelic expression. *Mamm. Genome*, **9**, 775-777.

Ishihara,K., Hatano,N., Furuumi,H., Kato,R., Iwaki,T., Miura,K., Jinno,Y. and Sasaki,H. (2000) Comparative genomic sequencing identifies novel tissue-specific enhancers and sequence elements for methylation-sensitive factors implicated in *Igf2/H19* imprinting. *Genome Res.*, **10**, 664-671.

Kaffer,C.R., Srivastava,M., Park,K.Y., Ives,E., Hsieh,S., Batlle,J., Grinberg,A., Huang,S.P. and Pfeifer,K. (2000) A transcriptional insulator at the imprinted *H19/Igf2* locus. *Genes Dev.*, **14**, 1908-1919.

Kanduri,C., Holmgren,C., Pilartz,M., Franklin,G., Kanduri,M., Liu,L., Ginjala,V., Ulleras,E., Mattsson,R. and Ohlsson,R. (2000a) The 5' flank of mouse *H19* in an unusual chromatin conformation unidirectionally blocks enhancer-promoter communication. *Curr. Biol.*, **10**, 449-457.

Kanduri,C., Pant,V., Loukinov,D., Pugacheva,E., Qi,C.F., Wolffe,A., Ohlsson,R. and Lobanenkov,V.V. (2000b) Functional association of CTCF with the insulator upstream of the *H19* gene is parent of origin-specific and methylation-sensitive. *Curr. Biol.*, **10**, 853-856.

Khosla,S., Aitchison,A., Gregory,R., Allen,N.D. and Feil,R. (1999) Parental allele-specific chromatin configuration in a boundary-imprinting-control element upstream of the mouse *H19* gene. *Mol. Cell. Biol.*, **19**, 2556-2566.

Leighton,P.A., Saam,J.R., Ingram,R.S., Stewart,C.L. and Tilghman,S.M. (1995) An enhancer deletion affects both *H19* and *Igf2* expression. *Genes Dev.*, **9**, 2079-2089.

Ohlsson,R., Renkawitz,R. and Lobanenkov,V. (2001) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.*, **17**, 520-527.

Paulin,R., Grigg,G.W., Davey,M.W. and Piper,A.A. (1998) Urea improves efficiency of bisulphite-mediated sequencing of 5'-methylcytosine in genomic DNA. *Nucleic Acids Res.*, **26**, 5009-5010.

Robinett,C.C., O'Connor,A. and Dunaway,M. (1997) The repeat organizer, a specialized insulator element within the intergenic spacer of the *Xenopus* rRNA genes. *Mol. Cell. Biol.*, **17**, 2866-2875.

Saitoh,N., Bell,A.C., Recillas-Targa,F., West,A.G., Simpson,M., Pikaart,M. and Felsenfeld,G. (2000) Structural and functional conservation at the boundaries of the

chicken beta-globin domain. *EMBO J.*, **19**, 2315-2322.

Sasaki,H., Ishihara,K. and Kato,R. (2000) Mechanisms of *Igf2/H19* imprinting: DNA methylation, chromatin and long-distance gene regulation. *J. Biochem.*, **127**, 711-715.

Sasaki,H., Jones,P.A., Chaillet,J.R., Ferguson-Smith,A.C., Barton,S.C., Reik,W. and Surani,M.A. (1992) Parental imprinting: potentially active chromatin of the repressed maternal allele of the mouse insulin-like growth factor II (*Igf2*) gene. *Genes Dev.*, **6**, 1843-1856.

Scott,K.C., Taubman,A.D., and Geyer,P.K. (1999) Enhancer blocking by the *Drosophila gypsy* insulator depends upon insulator anatomy and enhancer strength. *Genetics*, **153**, 787-798.

Udvardy,A. (1999) Dividing the empire: boundary chromatin elements delimit the territory of enhancers. *EMBO J*, **18**, 1-8.

Udvardy,A., Maine,E. and Schedl,P. (1985) The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J. Mol. Biol.*, **185**, 341-358.

Zhou,J. and Levine,M. (1999) A novel cis-regulatory element, the PTS, mediates an anti-insulator activity in the *Drosophila* embryo. *Cell*, **99**, 567-575.

Zubair,M., Hilton,K., Saam,J.R., Surani,M.A., Tilghman,S.M. and Sasaki,H. (1997)

Structure and expression of the mouse *L23mrp* gene downstream of the imprinted *H19* gene: biallelic expression and lack of interaction with the *H19* enhancers. *Genomics*, **45**, 290-296.

Table I. EMSA probes

PCT1	5'-TAGCCAAGGGACTGCCAACCTATACAAGATCA **GGGCAGGGGGGCAG** TGCTAGAGATATGTGGGCCAAAACTTGACAGGCATAGA-3'
2	5'-ACAAACAGCTGTCTCGACTCCCAGGCCCCGTGAAGACC **CTGGTTGGGGGCAG** GGGCTATGGGATACCTCCTGGGGATACACCAG-3'
3	5'-GACCAGGCTCTCCGTCCCCTTCTGTGCAGCCACAGGAAT **CTGCAAGGTGGCTG** CCTGAGGGAAGGCAGGCAATGGGTAGGGCAG-3'
4	5'-CCTTCTATCAGAATGTCACCCTCCCCTAAAGCTA **CCGCTGGGGTCCAG** GGGTGGCACCACACTCTTCAATCTACTTCTTTCCC-3'
5	5'-GTGGAGGAAGGCGCCATGTGGAACCAACTGGCC **TGGCTTGGGGGCAG** GAGGCCCCCCAAGCAGTGAGCCCAGCTGGGCATGCT-3'
6/14	5'-GAGGTACCCATGCCTGGCCACATT **CTGCCACACCGTGGCCGCTGGGAGGAAG** GGACAGGGACCAGGCCAGGCAGGGAGGCATTA-3'
7	5'-TCCCGAGGCATAATGTAGACTGCACAGTAAAT **CTGACTCCTTGCAG** TAAGGCTCGTAATGGGTAGGCAGCTACACATGAGGCTG-3'
8	5'-GCTTAGAGAGGGCTGCTAGACAGGCTCCAAGC **CTGCCACCCAAAGG** CACTTTCAACCTCCCCCCTCCTTGTGCGGAGAAGCTAA-3'
9	5'-GGAAGCCACCTTTGTGTACCCATTGTGAGTACC **CTGCTCCCCGGCAG** CTCTCAGCACTGCTGAGGTACAGAGGGAGCGCCTCTG-3'
10	5'-AGGAGTGGGCAACAGAAGAATGGAGAGGCTGCTTCT **CTGCCCCCAAGCCC** CCCAGCTTCCTGCCTCACAGCCCACAAAGCACTA-3'
11	5'-CCAATACTTGGGGACACCATGCTCCCACTGCCCC **CTACCACCACACGG** CCCTGGGTACACACAGCCCTCAAACCCACTAAGGCT-3'
12	5'-AAGCAGTGAGCCCAGCTGGGCATGCTTGCTCCTCTA **CTGCCCCCTTTAGG** TAGCCCAAGGCTCAGAACCAGCTCCCCCAGAGAC-3'
13	5'-TTAGGAGGCCCCTTCCAGAATTTAGA **CTGCCTCATCGCAG** GGACACAGGGAAAACTAAAAATGTTCACACACTAAGGTACCCCA-3'
15	5'-TGGCCTTCAGCCTCAGCCAGCATACCCCCCTT **CTCCCGCCCACCGG** CCTGCTCGCCCGCCCGCCCGTGCACTCCAGTGGGGGTG-3'
hPCT12	5'-AGAGGGAGGCGTCCGACCGACCGCACTGCTCTG **CTGCCCCCACCAGG** CAGGCCGAGGCCGGCATCCCCGGAGACCGGGACCTGG-3'
14	5'-TCCGGTGCCCTTGCCTGGCATGCT **CTGCCACACCGTGGCCGCGTGAGGGACAG** ACAGCGCGGGACAGAATCCCACCTGGCAGGG-3'

Only the upper strand sequence of the duplex probe is shown. Sequences simlar to the CTCF consensus (Fig. 2B) are indicated in bold.

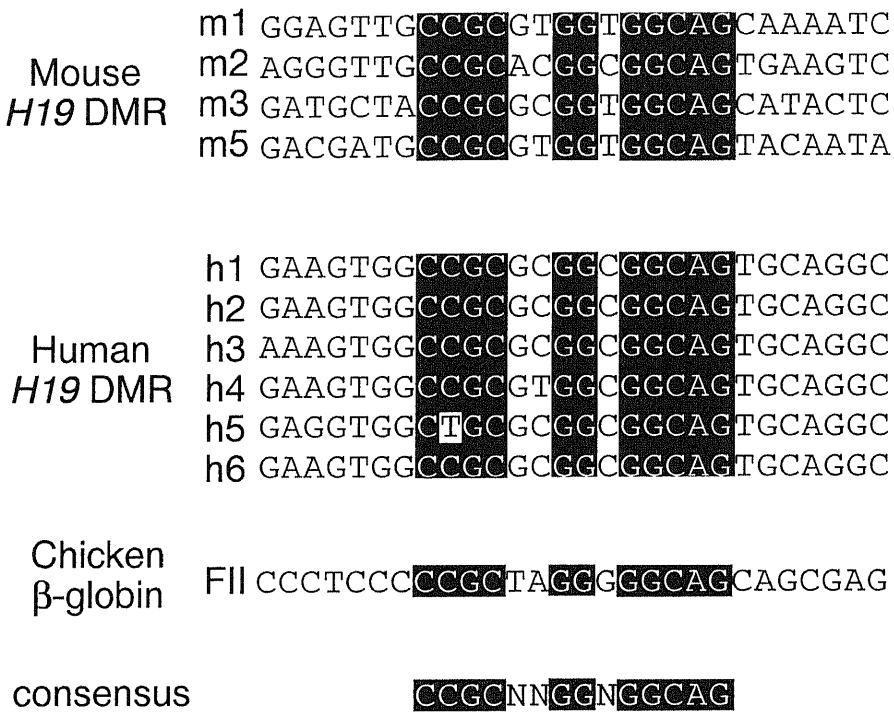


Fig. 1. Sequence alignment of the known CTCF sites. The sequence of the mouse (m1-m3, m5) and human (h1-h6) CTCF sites derived from *H19* DMR and the chicken β-globin FII site are shown. The bases conserved between the β-globin FII sequence and the *H19* DMR sequences are boxed. The 14-bp consensus sequence is shown at the bottom.

Fig. 2. Identification of potential CTCF sites within the mouse *H19/L23mrp* intergenic region. (**A**) Structure of the *H19/L23mrp* region. Segments conserved between human and mouse (CS1-10, Ishihara *et al.*, 2000) are shown by closed circles (enhancers) and open circles (function unknown). The positions of the sequences similar to the CTCF consensus are shown below (PCT1-15). (**B**) Alignment of the sequence resembling CTCF sites. The consensus sequence for CTCF is indicated at the top. Nucleotides identical to the CTCF consensus are boxed.

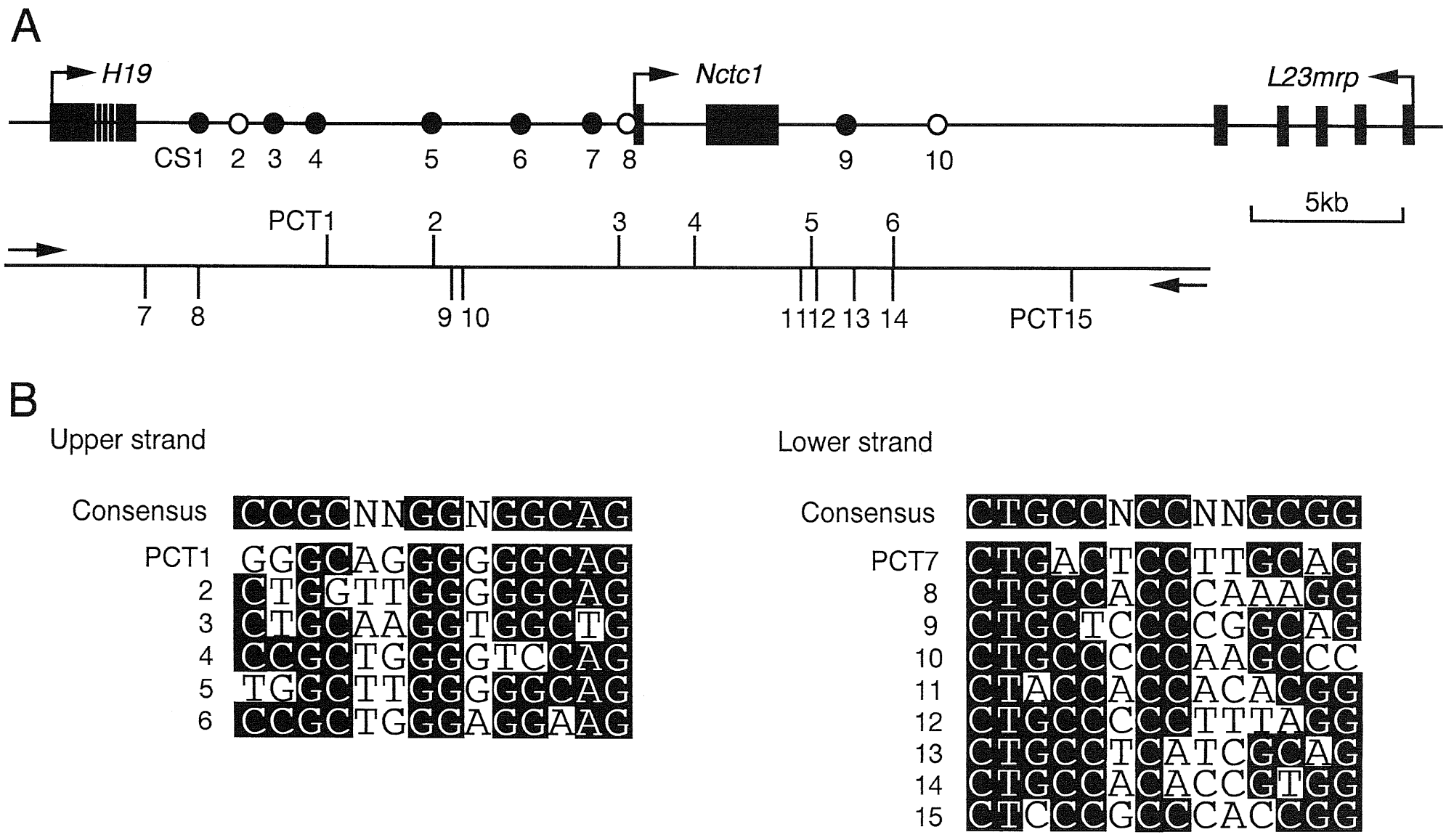


Fig. 2. Identification of potential CTCF sites within the mouse *H19/L23mrp* intergenic region.

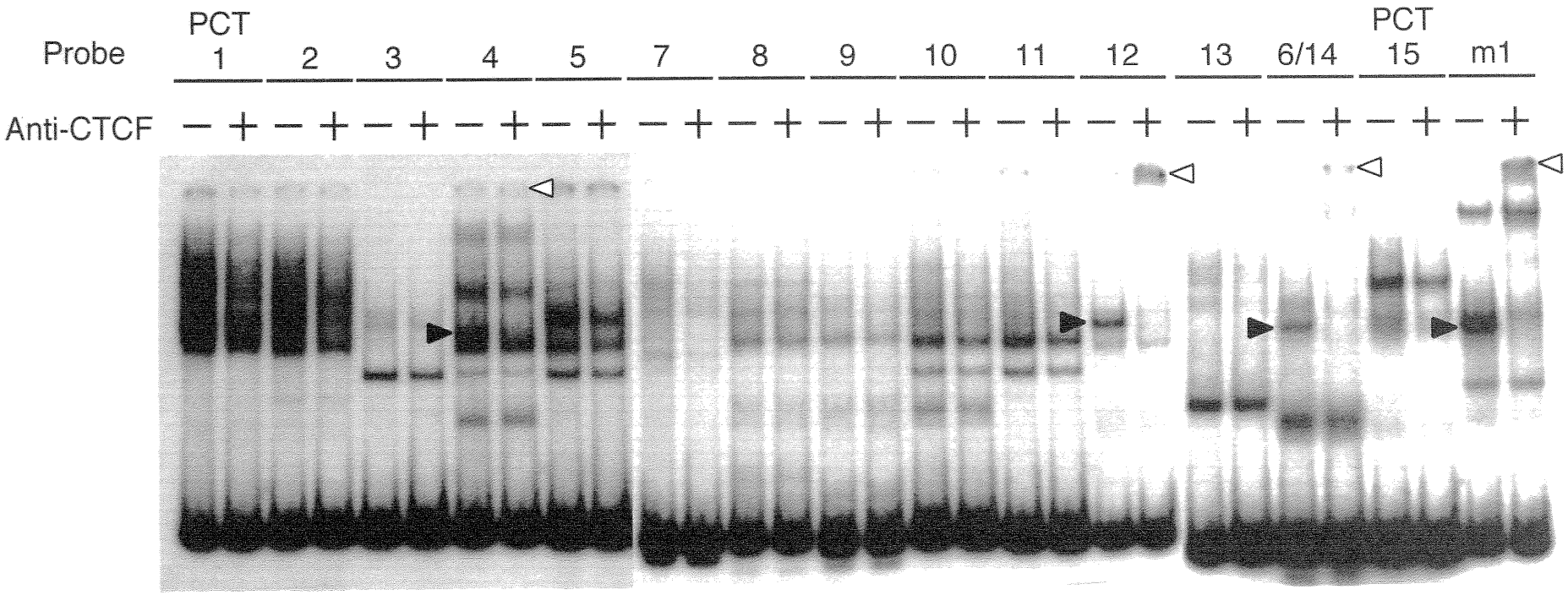


Fig. 3. EMSA for CTCF binding. Nuclear extract from 12.5-dpc mouse embryo was incubated with the indicated probe alone (-) or with anti-CTCF antibodies (+). Closed and open arrowheads indicate the CTCF complexes and super-shifted complexes, respectively.

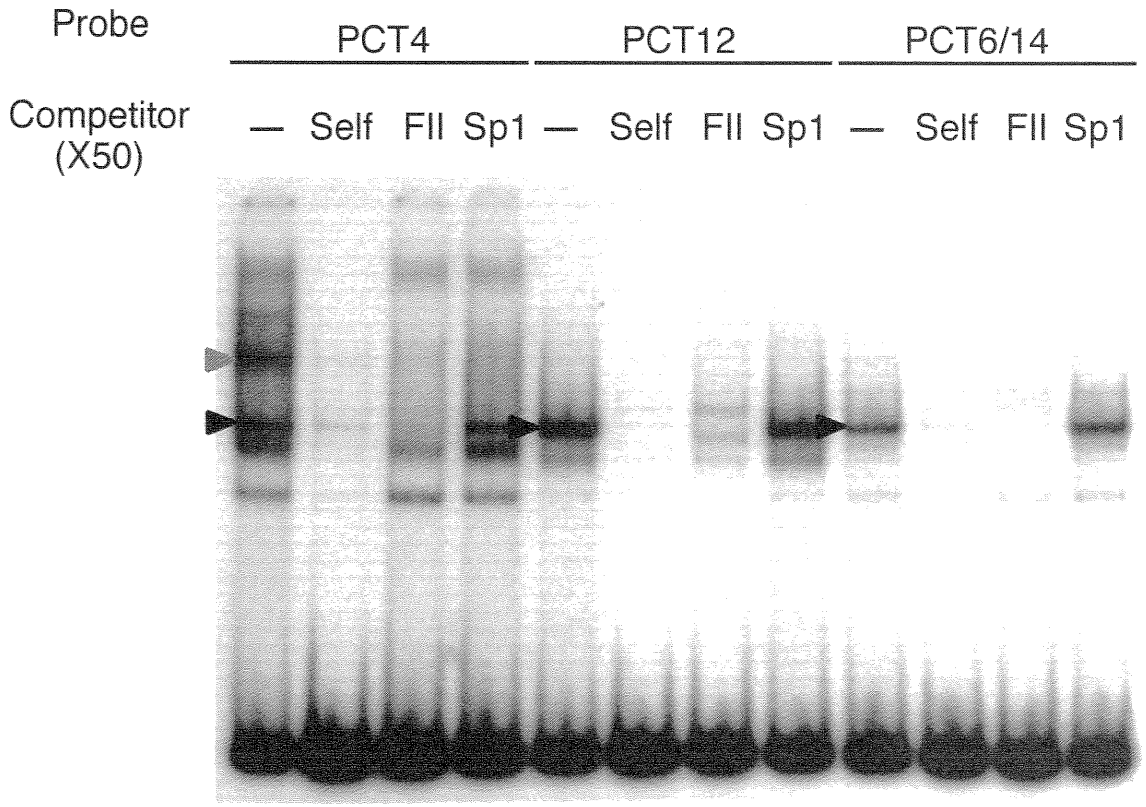


Fig. 4. The PCT4, PCT12 and PCT6/14 sequences bind CTCF. Radiolabeled probes were incubated with a nuclear extract from 12.5-dpc mouse embryo in the presence or absence of a 50-fold excess of unlabeled competitors as indicated. The black and gray arrowhead indicates the CTCF and the Sp1 complex, respectively.

Fig. 5. PCT12 and PCT14 sites are evolutionarily conserved. (**A**) Alignment of the human and mouse PCT12 and PCT6/14 sequences. Nucleotide positions corresponding to the CTCF consensus are boxed. Arrows indicate the orientation of the CTCF sites (see Figure 1A). PCT12 and PCT14, but not PCT6, are conserved. (**B**) Competition assays. The major complexes formed with the human probes (closed arrowheads) were abolished by excess unlabeled FII fragments but not by Sp1 consensus duplexes. (**C**) Super-shift assays. The complexes were super-shifted (open arrowheads) by anti-CTCF antibodies.

A

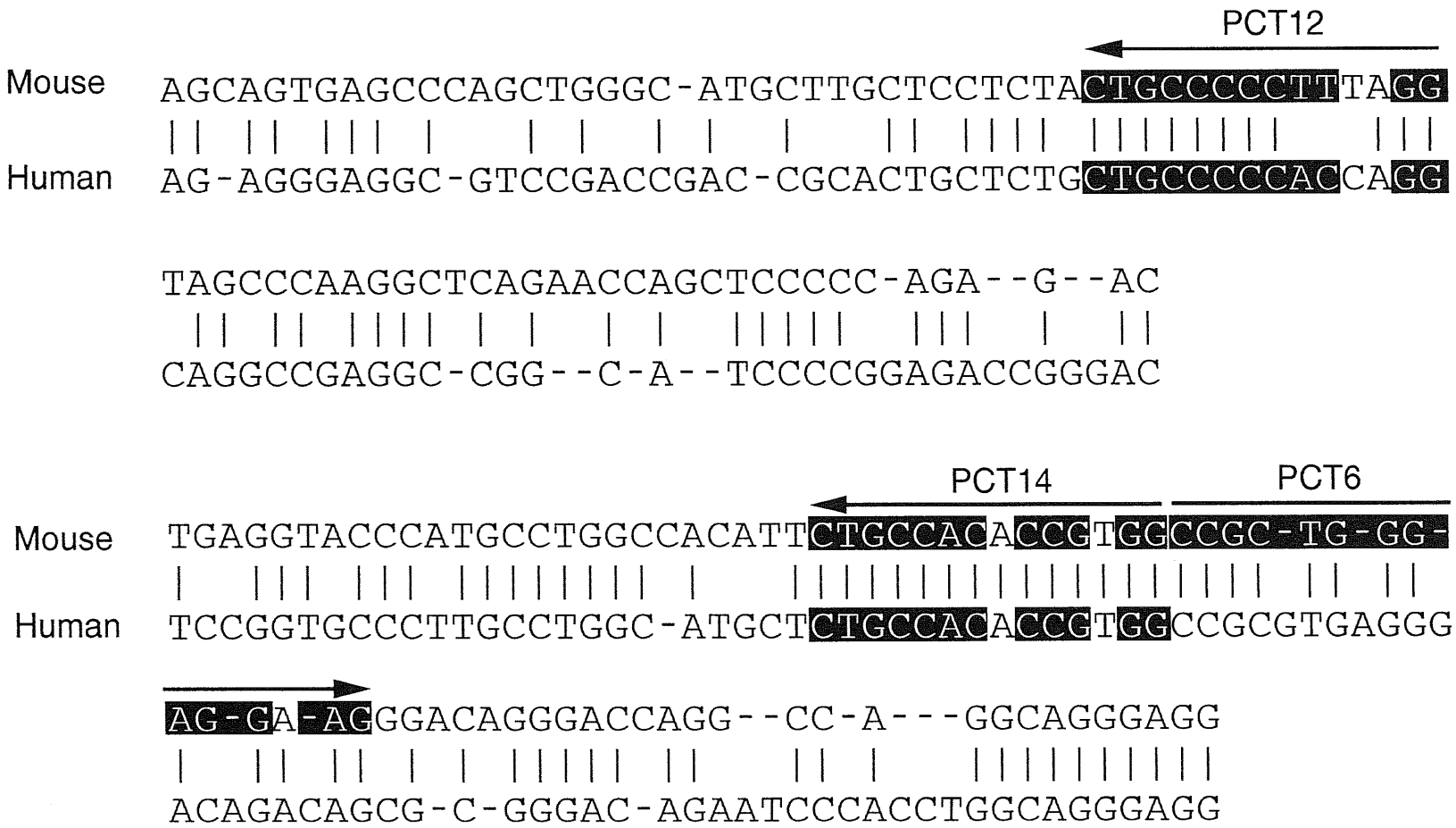


Fig. 5. PCT12 and PCT14 sites are evolutionarily conserved.

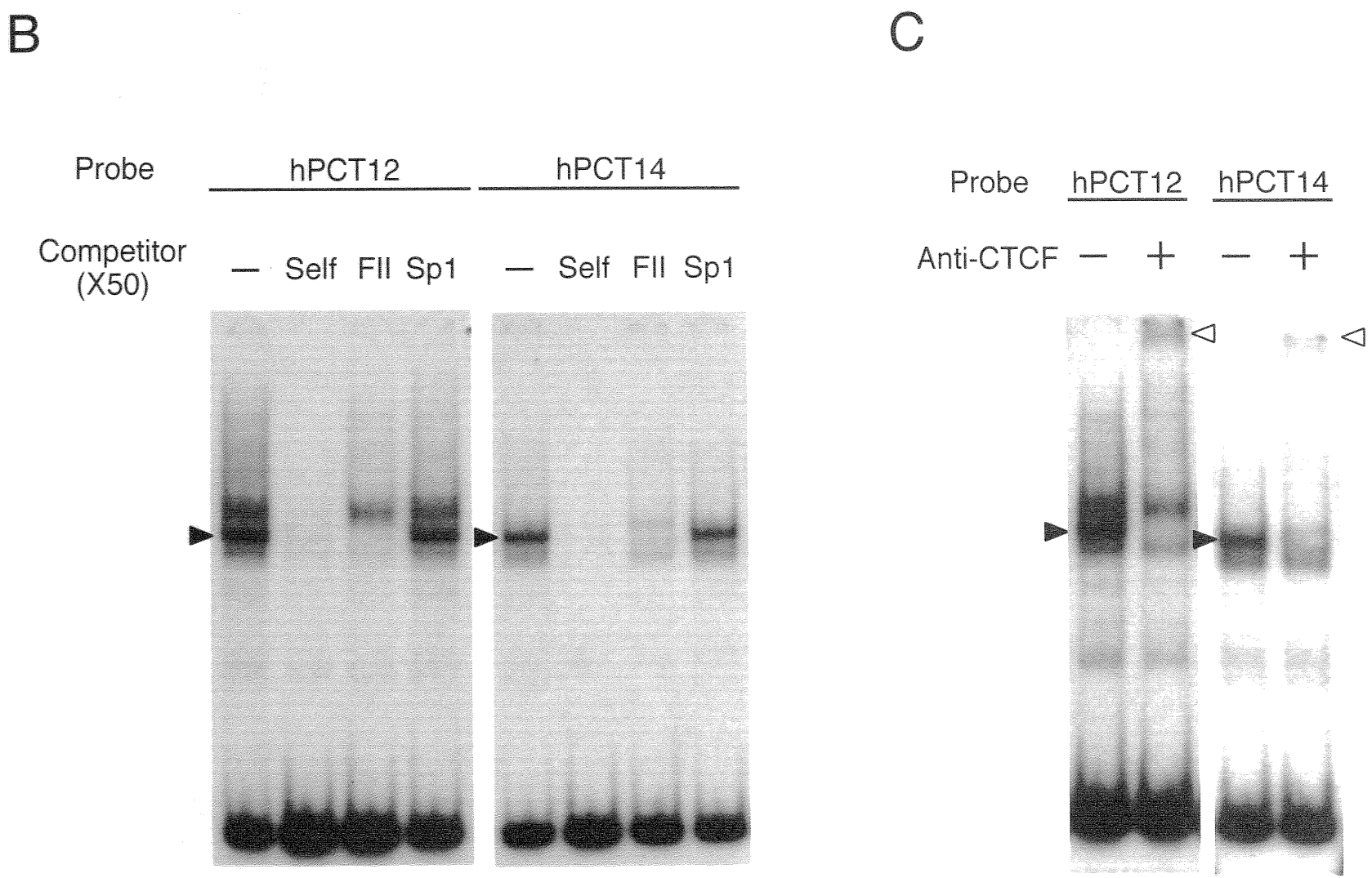


Fig. 5. PCT12 and PCT14 sites are evolutionarily conserved.

Fig. 6. DNase I hypersensitive site assay of the potential CTCF sites. Nuclei isolated from 12.5-dpc mouse embryo were digested with increasing amounts of DNase I. DNA samples were analyzed by Southern blotting. (**A**) DNase I hypersensitivity of the PCT12 region. Map of the mouse region containing PCT4 and PCT12 is shown at the top. The solid box indicates the second exon of *Nctc1* and the oval indicates a mesoderm enhancer (CS9, Ishihara *et al.*, 2000). Closed arrowheads show the positions of hypersensitive sites. The position of PCT4 is shown by an open arrowhead. Gray bars show the probes used: VBg0.6, a 0.6-kb *Eco*RV-*Bgl*II fragment; SV1, a 1-kb *Sac*I-*Eco*RV fragment. (**B**) DNase I hypersensitive sites of the PCT6/14 region. Two hypersensitive sites were detected (closed arrowheads) at both ends of CS9 using PPm1 (a 1-kb *Pst*I-*Pma*CI fragment) and BIP1 (a 1-kb *Bln*I-*Pst*I fragment) probes. An open arrowhead indicates the position of PCT6/14. Sizes are in kb.

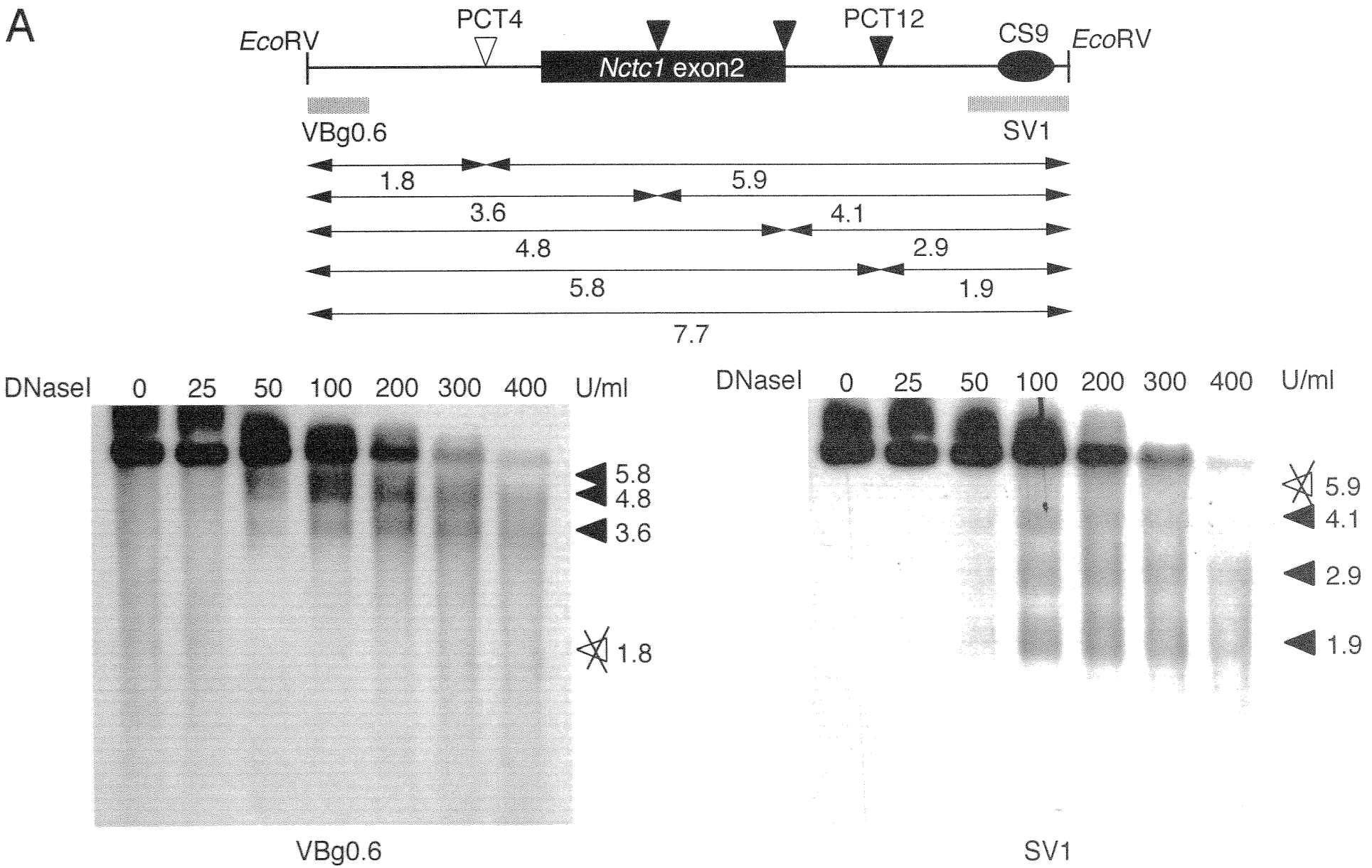


Fig. 6. DNase I hypersensitive site assay of the potential CTCF sites.

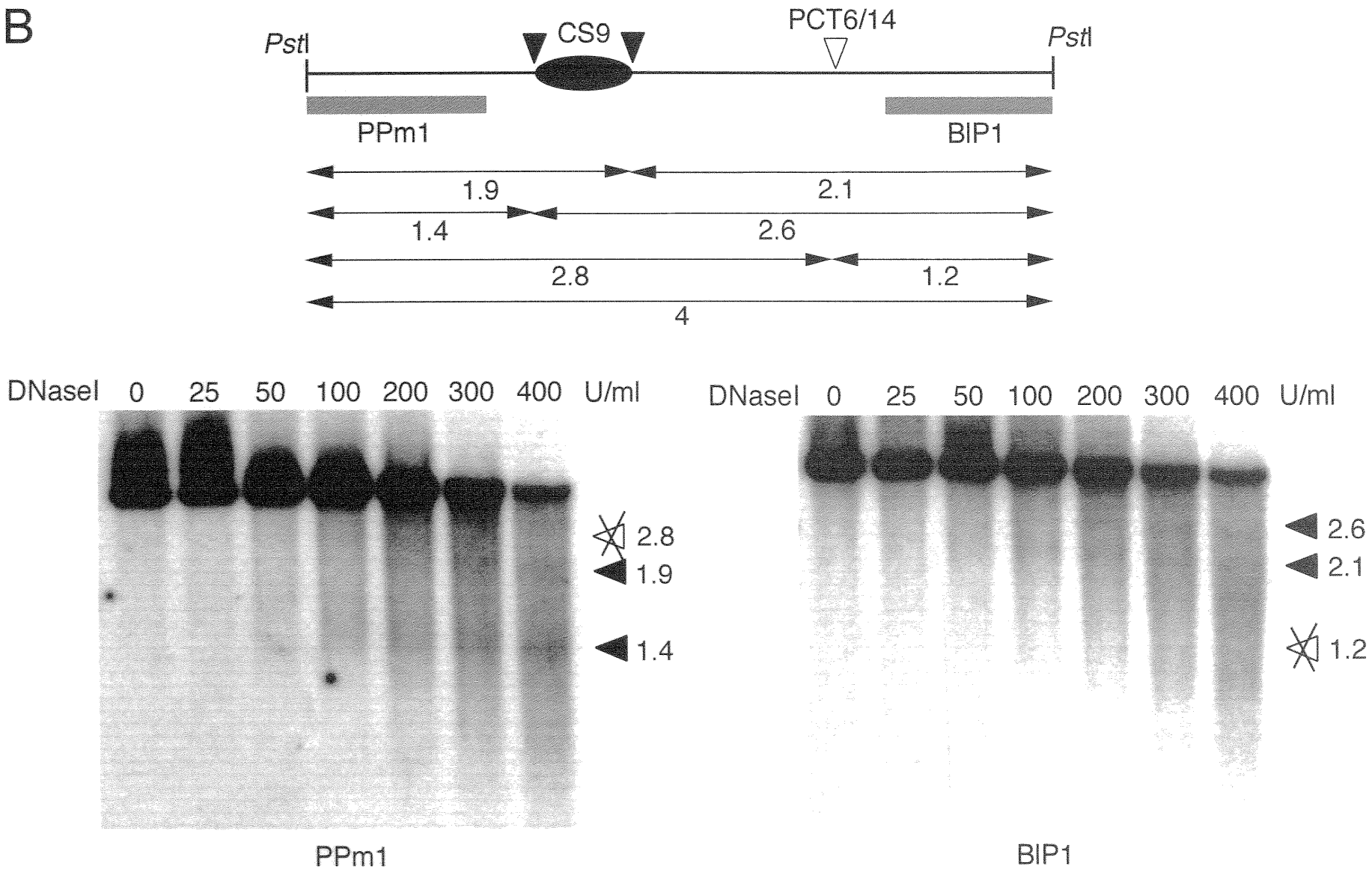


Fig. 6. DNase I hypersensitive site assay of the potential CTCF sites.

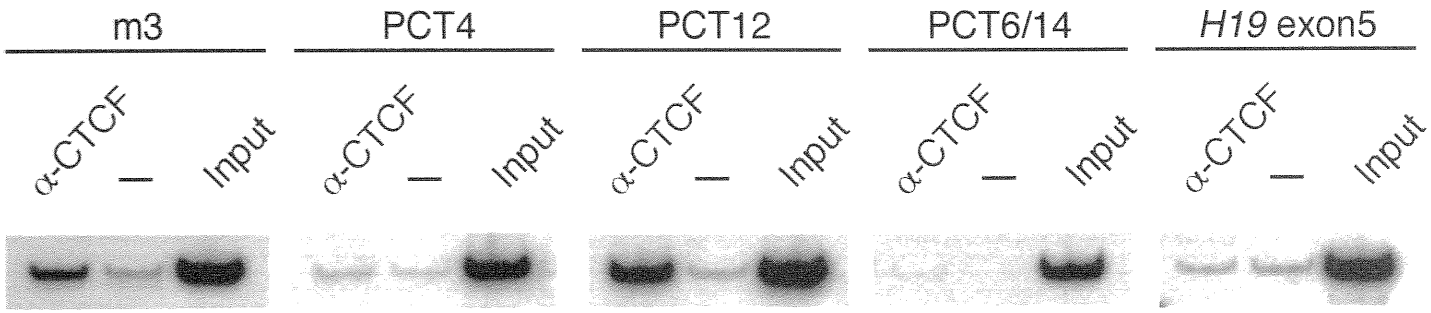


Fig. 7. PCT12 is bound by CTCF *in vivo*. Nuclei isolated from 12.5-dpc mouse embryo were treated with formaldehyde. Then sonicated chromatin was subjected to immunoprecipitation with or without anti-CTCF antibodies. Immunoprecipitated DNA was PCR-amplified with specific primers. DNA isolated from the supernatant without antibodies was used as "input" DNA. The m3 and *H19* exon 5 regions were used as a positive and a negative control, respectively.

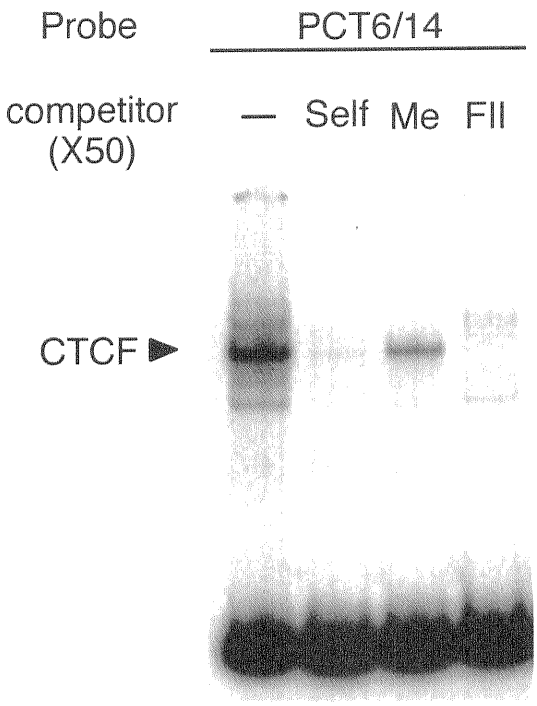


Fig. 8. Methylation of PCT6/14 sequence inhibits CTCF binding. CTCF complex formed with a PCT6/14 probe was greatly diminished by competition with excess unlabeled PCT6/14 (Self) and FII fragments (FII) but not with methylated PCT6/14 duplexes (Me).

Fig. 9. Bisulfite methylation analysis of PCT6 and PCT14. (A) Sequence of the region analyzed. The positions of CpG sites are numbered and boxed. Primers used for amplification are indicated by arrows. Bars indicate the positions of PCT6 and PCT14. (B) Methylation status of each CpG site (labeled 1-7 in A) determined by sequencing bisulfite-treated genomic DNA from 14.5-dpc mouse embryo. Methylation status is shown by open (unmethylated) or filled (methylated) circles.

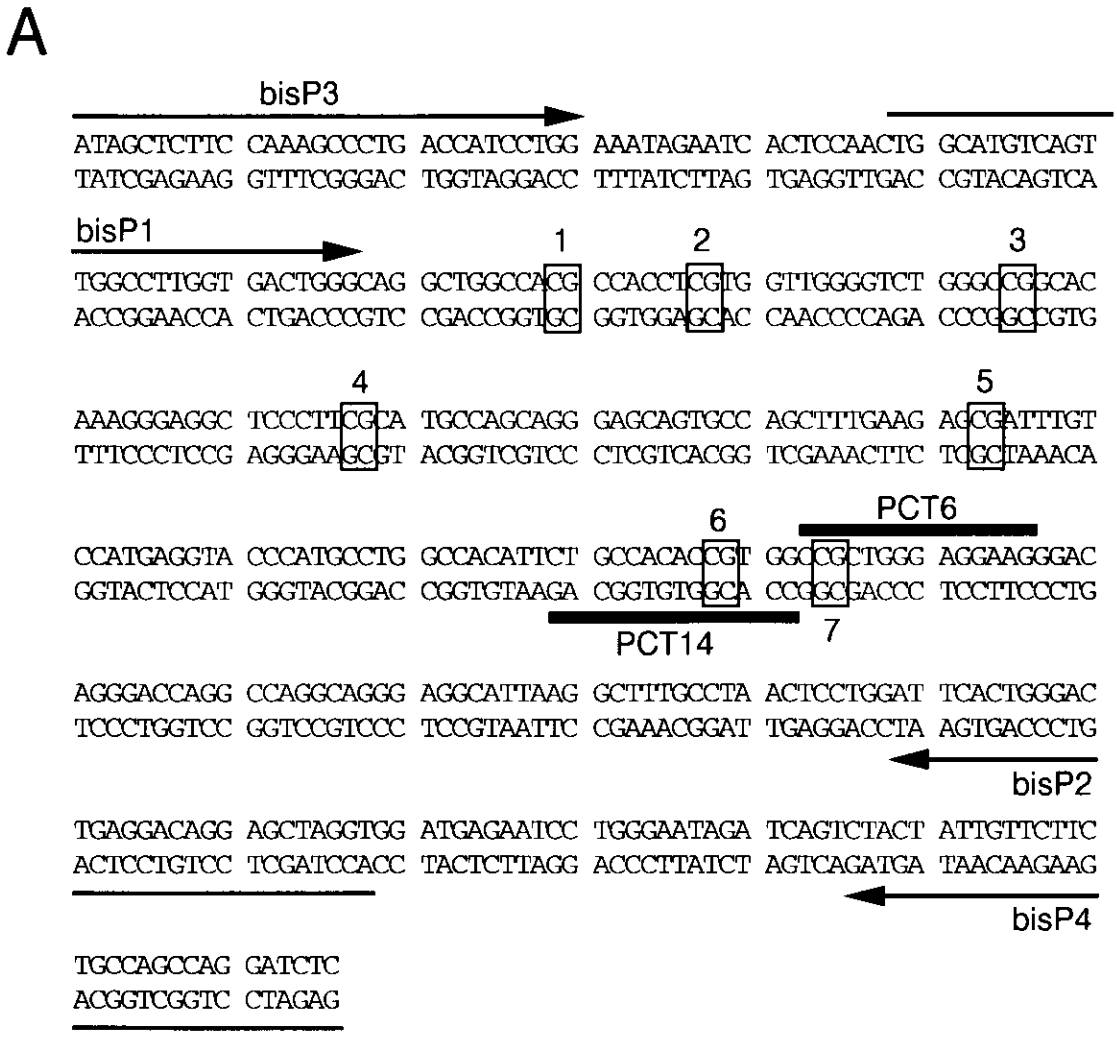
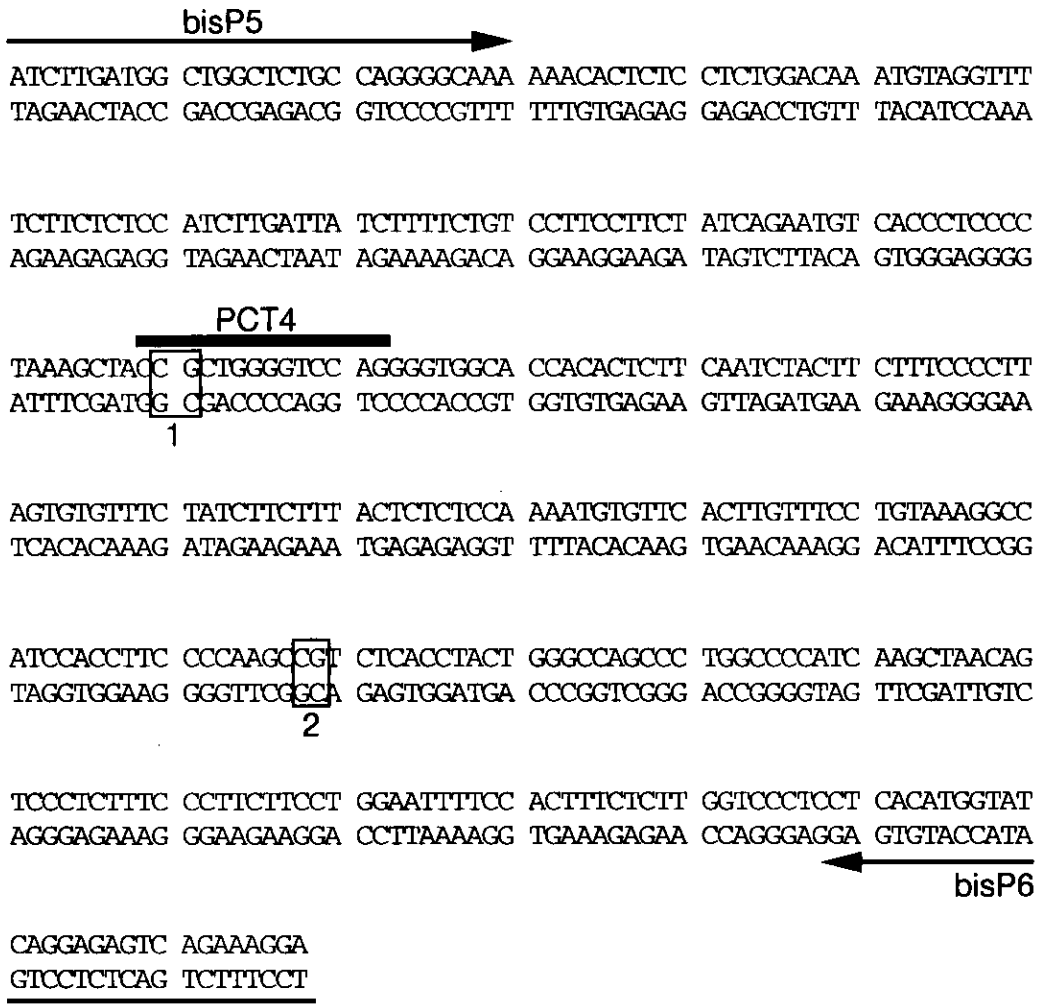


Fig. 9. Bisulfite methylation analysis of PCT6 and PCT14.

Fig. 10. Bisulfite methylation analysis of PCT4. Genomic DNA from 14.5-dpc mouse embryo was bisulfite-treated, PCR-amplified with specific primers (arrows), and sequenced. (A) Sequence of the amplified PCT4 region. The position of PCT4 is shown by bar. (B) Methylation status of the PCT4 region. Open and filled circles represent unmethylated and methylated CpGs, respectively.

A



B

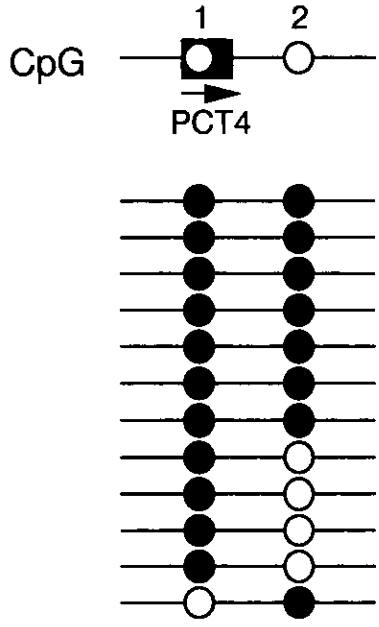


Fig. 10. Bisulfite methylation analysis of PCT4.

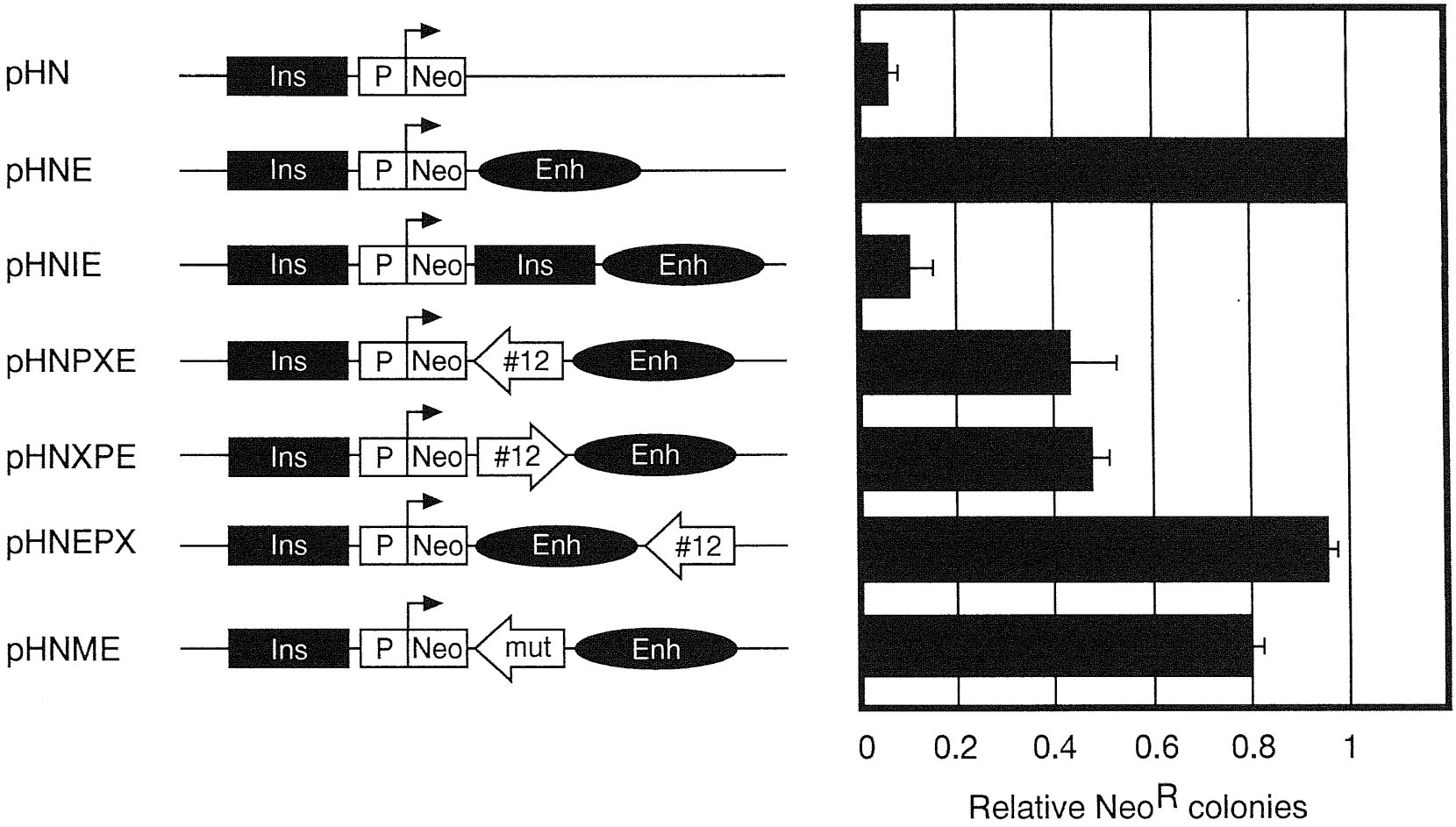


Fig. 11. Enhancer-blocking activity of PCT12. The human hepatoma cell line Hep3B was stably transfected with the indicated construct and grown in medium containing G418. The number of neomycin-resistant colony obtained with pHNE was set as 1. A 1.5-kb *Xho*I-*Pma*CI fragment containing PCT12 is shown by open arrows, indicating the orientation of transcription. A fragment with mutations at PCT12 is indicated as "mut". Neo, neomycin-resistance gene; P, mouse *H19* promoter; Enh, *H19* endoderm enhancers; Ins, *H19* DMR insulator.

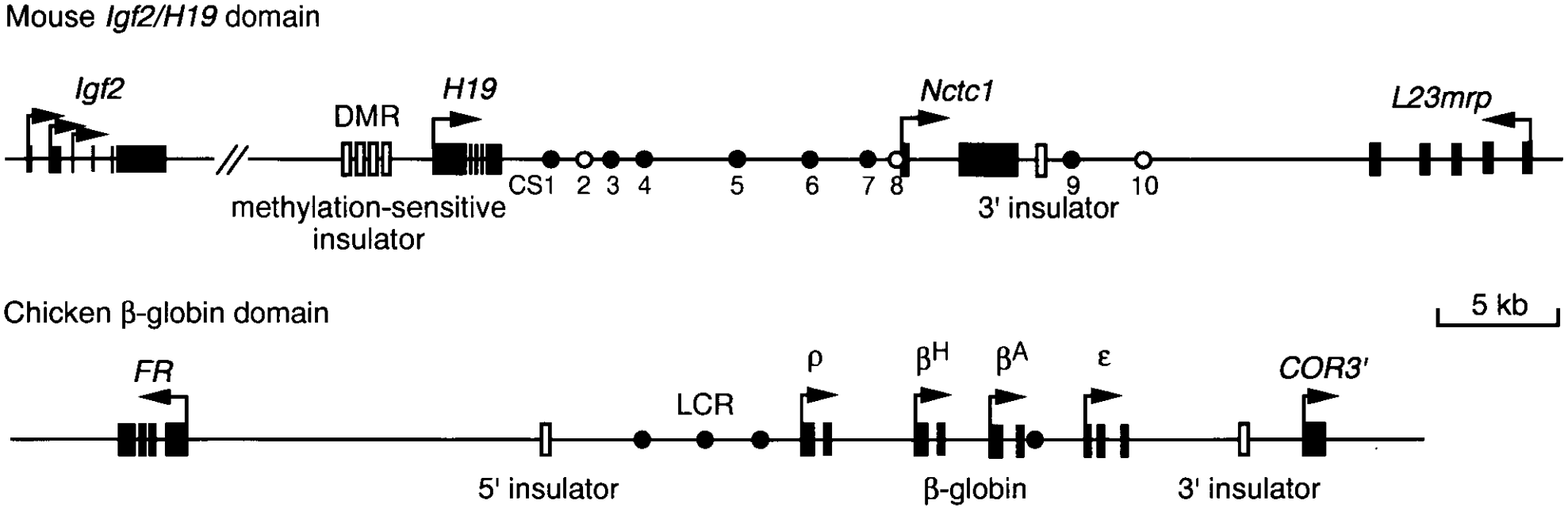


Fig. 12. Comparison of the mouse *Igf2/H19* domain with the chicken β-globin domain. Each domain is flanked by CTCF-dependent insulators. The four CTCF-dependent insulators in the 5' flanking region of *H19* are located within the DMR, which serves as a methylation-sensitive imprinting control region. *Nctc1* is a muscle-specific non-imprinted transcription unit, which is not present in the human sequence. Filled boxes and circles represent the exons and the enhancers, respectively. Open boxes indicate the CTCF sites that function as insulators. LCR, locus control region.