Human Molecular Genetics, 2003, Vol. 12, No. 3 **233–245**
DOI: 10.1093/hmg/ddg028

Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, *Peg3*

Joomyeong Kim*, Angela Kollhoff, Anne Bergmann and Lisa Stubbs

Genomics Division, Biology and Biotechnology Research Program, L-441, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94551, USA

Received August 27, 2002; Revised and Accepted December 3, 2002

The 5′-ends of two paternally expressed mouse genes, *Peg3* and *Usp29*, are jointly associated with a CpG island that exhibits allele-specific methylation. Sequence comparison of the regions derived from human, mouse and cow revealed the presence of two evolutionarily conserved sequence motifs including one that is repeated multiple times within the first intron of *Peg3* in all three mammals. DNA mobility shift and chromatin immunoprecipitation (ChIP) assays clearly demonstrated that this motif is an *in vivo* binding site for the *Gli*-type transcription factor YY1. The YY1-binding site contains one CpG dinucleotide, and methylation of this CpG site abolishes the binding activity of YY1 *in vitro*. The *Peg3* YY1-binding sites are methylated only on the maternal chromosome *in vivo*, and ChIP assays confirmed that YY1 binds specifically to the paternal allele of the gene. Promoter, enhancer and insulator assays with deletion constructs of sequence surrounding the YY1-binding sites indicate that the region functions as a methylation-sensitive insulator that may influence the imprinted expression of *Peg3* and neighboring genes. The current study is the first report demonstrating the involvement of YY1 in methylation-sensitive insulator activity and suggests a potential role of this highly conserved protein in mammalian genomic imprinting.

INTRODUCTION

The two parental alleles of certain mammalian genes are functionally nonequivalent due to genomic imprinting, a process by which one allele becomes epigenetically modified and inactivated based on parental origin. More than 50 imprinted genes have been isolated from human and mouse (www.mgu.har.mrc.ac.uk/imprinting/imprin-ref.html#impregs), and many of these imprinted genes are involved in either fetal growth or animal behaviors (1,2). The known imprinted genes are clustered in discrete chromosomal regions, indicating that genomic imprinting is a long-range mechanism affecting relatively large regions of chromosomes (3,4). Several lines of evidence suggest that small DNA elements may serve to regulate the imprinting of a whole domain and coordinate the expression of individual genes in each region (5–7).

The molecular mechanisms that underlie imprinting control are still elusive but a series of studies focused on the well-known imprinted domains, including *H19*/*Igf2, SNRPN* and *Igf2r*, have further located the genomic regions that play critical roles for the imprinting control of each domain (8–10). Deletions or modifications of these genomic regions, termed imprinting control regions (ICRs), are shown to affect the transcription and imprinting of genes locally as well as those found at relatively distant genomic sites. Common features of these ICRs include tandem-repeat sequence structures and close association of, or overlap with, CpG islands that display allele-specific methylation (11,12). In the case of the H19 ICR, the sequence of the repeated region is well conserved among distantly related mammals (13,14). Analyses of the H19 ICR indicate that the core of these repeated sequences is a binding site of the transcription factor CTCF, which serves as an insulator element for the imprinting control of *H19* and *Igf2*.

Peg3 (paternally expressed gene 3) was the first imprinted gene identified in the proximal imprinted domain of mouse chromosome 7 (Mmu7) (15), which was predicted originally in breeding studies with translocation mutant mice (16). Five additional imprinted genes have been identified subsequently within the surrounding genomic regions. These include the paternally expressed genes *Usp29* (17) and *Zf264* (18), and maternally expressed genes *Zim1* (19), *Zim2* (20), and *Zim3* (18). Comparative mapping and imprinting studies of this domain indicate that human homologs of most of the mouse imprinted genes are located in the syntenically homologous region of chromosome (HSA19q13.4) (17,21,22), and that human *PEG3* and *ZIM2* are also imprinted (23). The clustering of the imprinted genes in this genomic interval suggests that the imprinting of this domain may be controlled by a small number

*To whom correspondence should be addressed. Tel: +1 9254243792; Fax: +1 9254222282; Email: kim16@llnl.gov

of ICRs. Sequence comparison of the 500 kb mouse and human imprinted regions indicate the presence of a number of evolutionarily conserved genomic fragments (J.Kim *et al.*, in preparation), including a 5 kb genomic region surrounding the first exons of *Peg3* and *Usp29*. This conserved region contains a CpG island that is methylated in an allele-specific pattern in somatic cells (23,24) as well as in germ cells (25) and includes an unusual tandem-repeat sequence structure as observed in the ICRs of other imprinted domains (11). Recent studies have also shown complete methylation of both parental alleles of this region in a large fraction of human glioma patients, suggesting that human *PEG3* might function as a tumor suppressor gene (26,27).

In the current study, we performed a series of functional studies of this 5 kb conserved, differentially methylated region (called PEG3-DMR), to investigate potential roles of this element in the transcription and imprinting control of *Peg3* and *Usp29*. Three-way sequence comparison analyses of PEG3-DMR sequences derived from human, mouse and cow indicate that this region contains at least two evolutionarily conserved sequence elements, including multiple binding sites for a *Gli*-type transcription factor YY1. Our studies show that the YY1-binding to the motif is methylation-sensitive and that all of the YY1-binding sites are differentially methylated between two parental alleles *in vivo*. Functional assays of this region suggest that the YY1-binding region acts as a methylation-sensitive insulator that may play a role in the imprinting control of this domain.

RESULTS

Identification of two conserved motifs located in the first intron of *Peg3*

As shown in other ICRs, the 5′-ends of two paternally expressed genes, *Peg3* and *Usp29*, are also part of a CpG island exhibiting allele-specific methylation (Fig. 1A). The sequence structure and evolutionary conservation of this potential ICR were analyzed by comparing genomic sequences derived from three distantly related mammals: human, mouse and cow. To isolate cow BACs containing *Peg3*, the RPCI-42 library was screened with two probes derived from different portions of the coding region of mouse *Peg3* (described in Materials and Methods). One of the identified BACs was selected and sequenced at the Joint Genome Institute (RPCI42-246L9; GenBank accession no. AC073666). Examination of the sequence of this bovine BAC indicated that the overall genomic layout of the *Peg3* region is well conserved in the three mammals (Fig. 1B). In all three species, three exons are found in the conserved 5 kb interval starting 1 kb upstream of *Peg3* and extending to the second exon of that gene. The region includes *Peg3* exons 1 and 2 and an oppositely oriented transcript that initiates within 150 bp of *Peg3*. In mouse, this sequence comprises the first non-coding exon of the paternally expressed gene, *Usp29;* the homologous human region is also included in a non-coding mRNA, called *MIM1* (Mer-repeat containing imprinted transcript 1; Kim *et al.*, unpublished data), but is not co-transcribed with *USP29* (18). The exon structure and imprinting status of the bovine gene has not yet been determined. The locations and spacing of these exons are well conserved among the three mammals.

The sequence of the 5 kb genomic region, called PEG3-DMR, was also examined for evolutionary conservation and repeated sequence structure. Three-way comparison of human, mouse and cow sequences identified several conserved subregions including the three exons and one small genomic segment, about 40 bp in length, located about 800 bp downstream of exon 1 (Fig. 2A and B). Owing to the repeated nature of the *Peg3* first intron sequence similarity, searching and two-way comparative sequence alignment programs, such as BLAST (www.ncbi.nlm.nih.gov/BLAST/) and PIPMAKER (http://nog.cse.psu.edu/pipmaker/), did not yield reliable clues regarding the locations of relevant functional sequences (Fig. 1B). To perform a more systematic analysis, we used the 'printrepeats' program for identifying and visualizing small repeat sequences within larger sequence domains. As shown in Figure 1C and D, the first intron of human, mouse and bovine *Peg3* contains several copies of tandem repeats. Although the same sequence is reiterated in all three species, the location, length and number of repeated units vary among the three mammals. In human and cow, the repeated units are clustered more compactly with high sequence identity between repeated units (Fig. 1C and D). By contrast, mouse repeats are more widely distributed and show relatively high levels of sequence variation. Sequence similarity of these repeat units is much higher within any one species than between different species, suggesting that the repeat units may have been duplicated since the divergence of rodent, ungulate and primate lineages. However, comparison of repeat units from different species clearly demonstrates that the duplicated regions share a common evolutionary history. Closer examination revealed a small core region, 11 bp in length, which is highly conserved both within and between species, whereas surrounding sequences are more species-specific (Fig. 2C). The 11 bp core sequence contains one CpG dinucleotide. As expected from the high mutation rate of CpG dinucleotides to either CpA or TpG in mammalian genomes (28), many repeat units contain one or two base changes at the CpG sites (Fig. 2D). Nevertheless, all three mammals carry 4–6 repeats units with the identical, canonical sequence. Overall, our analyses of the 5 kb genomic region identified two highly conserved sequence elements (CSEs) located in the first intron of *Peg3*, CSE1 and CSE2 (Fig. 2A).

Multiple binding sites for methylation-sensitive YY1

The two CSEs identified in the first intron of *Peg3* represent strong candidates as regulatory sequences, conserved because they contain binding sites for unknown transcription factors. To test this possibility, the two CSEs were first searched against the Transfac database (http://transfac.gbf.de/TRANSFAC/). These searches showed that CSE1 contains binding sites for a number of transcription factors, including Sp1, GCF and HiNF-C, whereas CSE2 contains sequences corresponding to the binding site for a *Gli*-type zinc-finger protein, YY1. To test the ability of these transcription factors to bind specifically CSE1 and CSE2 sequences, we performed DNA mobility shift assays with nuclear extracts prepared from HeLa cells (Fig. 3A–D). The assays indicated the presence of two

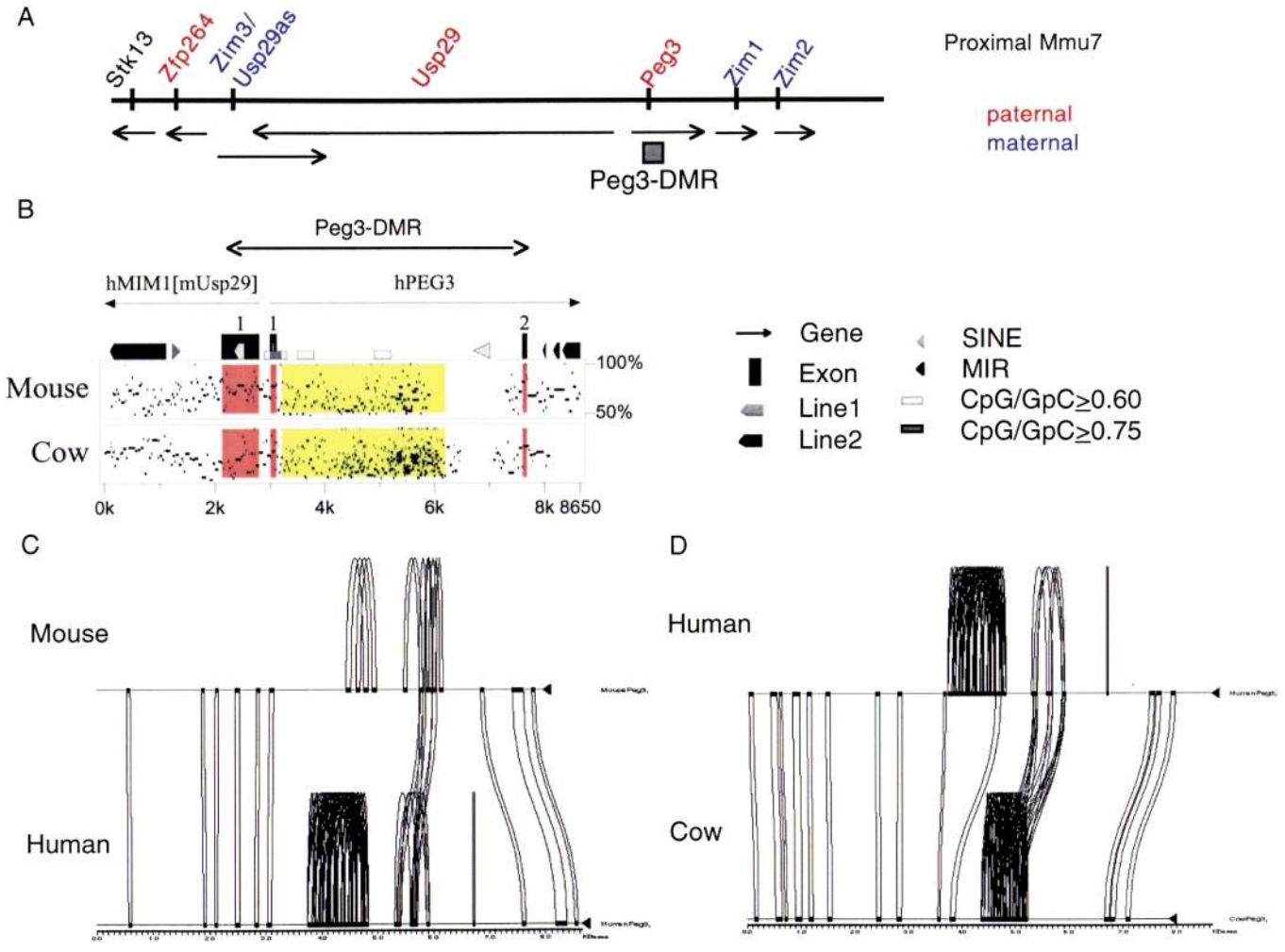
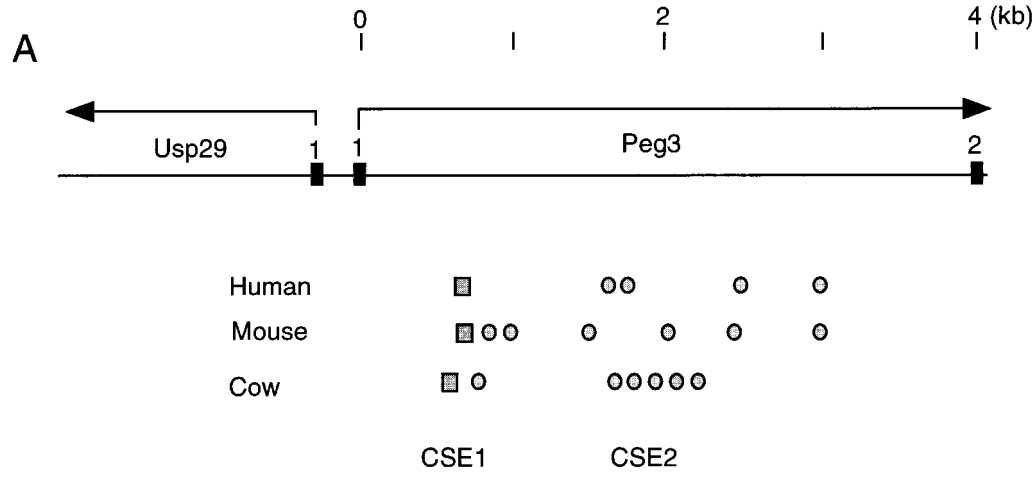


Figure 1. Sequence comparison of the surrounding regions of the first exons of *Peg3* and *Usp29/MIM1*. (**A**) The genomic organization of six imprinted genes in the proximal region of Mmu7. Maternally expressed genes are marked in blue whereas names of paternally expressed genes are shown in red. The location and orientation of each transcript are indicated by arrows. The relative position of Peg3-DMR within this imprinted domain is represented by a box below the map. (**B**) The 8 kb genomic sequences derived from human, mouse and cow *Peg3* were analyzed using Pipmaker. The pink-shaded regions correspond to exons and the yellow-shaded regions to the tandem repeat regions of the first intron of *Peg3*. The 5 kb DMR region is indicated at the top of the PIP plot. (**C**–**D**) Conservation patterns of tandem repeat sequences located within the first introns of human, mouse and cow *Peg3*. Inter-species conservation is shown as vertical lines between upper and lower sequences. Intra-species conservation (tandem repeat sequence) is shown as curved lines to regions within the same sequence.

potential complexes binding to CSE1 in a sequence-specific manner, with one set of proteins binding to the 3′-region and a second complex binding to the 5′-portion of the CSE1 sequence. The identity of the two complexes is currently being investigated.

To test binding of transcription factor(s) to CSE2, the 32 bp sequence corresponding to the first repeat unit in the human sequence was used as a probe for gel shift assays (Fig. 3B–D). As shown in Fig. 3B, an unknown transcription factor binds to this CSE2 32mer and the binding is sequence-specific, based on the selective binding to CSE2 against 100-fold molar excess of an unrelated probe (in this case CSE1, lane 4 in Fig. 3B). Since database searches suggested this sequence contained a YY1 binding site, the binding of this unknown protein from nuclear extracts was also competed against a commercially available YY1 consensus oligonucleotide (yy1) probe and a mutated version (yy2) that has the three nucleotide differences which are known to disrupt YY1 binding. As shown in Fig. 3B lanes 5 and 6, only the perfect YY1 consensus probe competed effectively against the CSE2 probe. To confirm the localization of YY1 protein within the CSE2-binding complex, we performed supershift assays with a commercially available YY1-antibody, which recognizes the zinc-finger portion of YY1. The supershift assay also included one unrelated antibody generated against the Usp29 protein, which had been prepared previously in our laboratory, as a negative control. The Usp29 antibody did not affect the binding of the complex to CSE2 but the YY1 antibody inhibited the binding of the complex to CSE2 significantly (Fig. 3B, lanes 7 and 8). Since the antigenic epitope of this YY1 antibody corresponds to the DNA-binding domain, the addition of the YY1 antibody to the supershift assay is likely to inhibit the binding of YY1 to DNA rather than shifting the binding complex. To further test this possibility, we repeated the supershift assays with another



B

```
Human GGGGGCGGGGCTTGTTT-GCCGCAGTGGTGGTGGGAGGGGCG
Mouse GGGGGCGGGTCTTGTTTTGCCGCAGTGGCGGTGGGCGGGGCT
Cow   TGGGGCGGGGCTTGCTT-GCCGCAGTGGCGGCGGGAGGAGCT
```

C

```
Human
     CCCACCCACCTGGGCGCCATCTTTAATGAAAG
     CCCGCCCGCCTGGGCGCCATCTTTGATGGGGG
     GTTGATTGCGCAGGCGCCATCTTGAGTGACAC
     GTGTACTGCGCAGGCGCCATCTTGAGTGACAG
Mouse
     CTCCACCCAGTTGGCGCCATCTTGGAGGGTGG
     CCCAGCCCGCTTGGCGCCATCTTGAAATGACC
     TCCGCCCATCCTGGCGCCATCTTTCCTGGCCA
     CCCGCCCAATTTGGCGCCATCTTTGCCAGCAA
     CCTGCCCACCTGGGCGCCATCTTTACTTGTCA
     CCTCCTCCCTCTGGCGCCATCTTTGTTGGCCA
Cow
     ACCCGCCGCCGCGGCGCCATCTTTGATGGGGG
     TCTTACTGCGCAGGCGCCATCTTGAATGACAG
     GCTTACTGCGCAGGCGCCATCTTGAATGGCAG
     GCTTACTGCGCAGGCGCCATCTTGAATGACGG
     GCTTACTGCGCAGGCGCCATCTTGAATGACCG
     GCTTACTGCGCAGGCGCCATCTTGAATGGCAG

IAP LTR SEQUENCE
     AACACAGCTGTCAGCGCCATCTTGTAACGGCG

YY1 BINDING SITE
                 GCGCCATCTTGANT
```

D

```
Human
     TGCTGCGTTGGTGCCGCCATCTTGCAGTCCGG
     CCCCACCCACTGGGTGCCATCTTTTATGAGGG
     CCCCACCCATTGGGTGCCATCTTTTATGAGGG
     CCCACCCACCTGGGCGCCATCTTTAATGAAAG
     CCCGCCCGCCTGGGCGCCATCTTTGATGGGGG
     CCCGCCCACCTGGGTGCCATCTTTGATCGGGG
     GTTGATTGCGCAGGCGCCATCTTGAGTGACAC
     GTGTACTGTGCAGGTGCCATCCTGAATGACAG
     GTGTACTGCGCAGGCGCCATCTTGAGTGACAG
     GTTCATTCATTCTGTGCCATCTTTGCCGGTTC
Mouse
     CTCCACCCAGTTGGCGCCATCTTGGAGGGTGG
     CCCAGCCCGCTTGGCGCCATCTTGAAATGACC
     TCCGCCCATCCTGGCGCCATCTTTCCTGGCCA
     CCCACCTACCCTGACGCCATCTTTGCTGGCTG
     CTGATATAGTGAGGCACAATCTTGAATGACAG
     AAGTTCTCAGCGGGTGTCATCTTGAGTGACAG
     CCCGCCCAATTTGGCGCCATCTTTGCCAGCAA
     CTTTATGGTGCAGACGCCATCTTGAATGACAG
     GTCTGAAGTAAGGGTGTCATCTTGAATGACTG
     GTTTCTGGTGCAGTAGCCATCTTGAATGACAG
     CCTGCCCACCTGGGCGCCATCTTTACTTGTCA
     ATTTATGGCGCATGAGCAATCTTGAATGACAG
     CCTCCTCCCTCTGGCGCCATCTTTGTTGGCCA
     ATTATCTGTGCAGGTGTCATCTTGAATGACAA
```

Figure 2. Two conserved sequence elements identified from the first intron of *Peg3*. (**A**) Positions of two conserved sequence elements (CSEs) within the first intron of *Peg3*. The diagram shows the positions and directions of exons within this region. The square indicates CSE1; the circle, CSE2. (**B**) Sequence alignment of CSE1s. (**C**) Sequence alignment of CSE2s. The conserved core and surrounding sequences of human, mouse and cow CSE2s are compared with the consensus sequence of YY1-binding sites. (**D**) Sequence alignment of all tandem repeats of human and mouse that show similarity to the consensus sequence of CSE2. Many tandem repeats of human and mouse show one or two base differences as compared with the core sequence of CSE2.

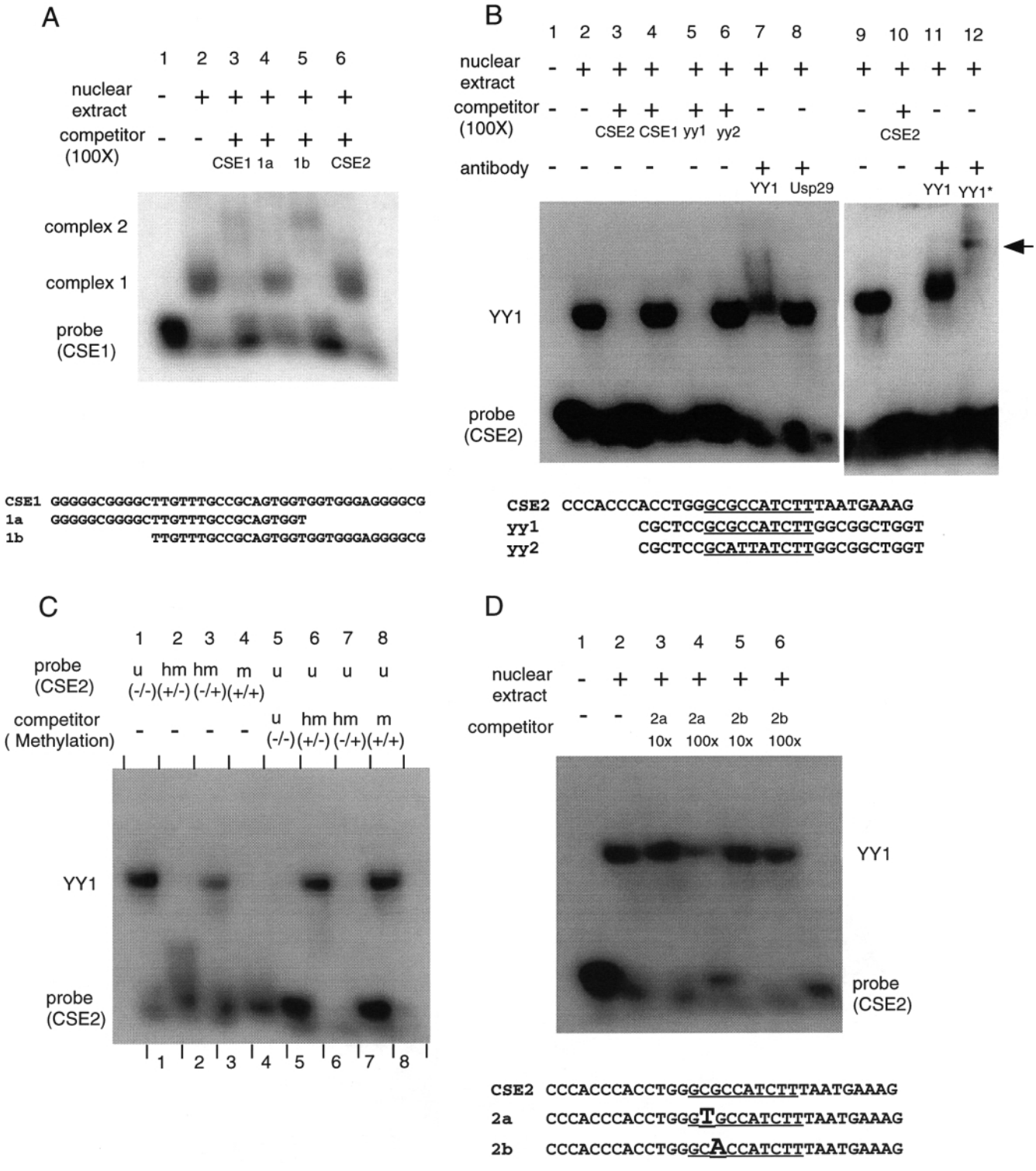
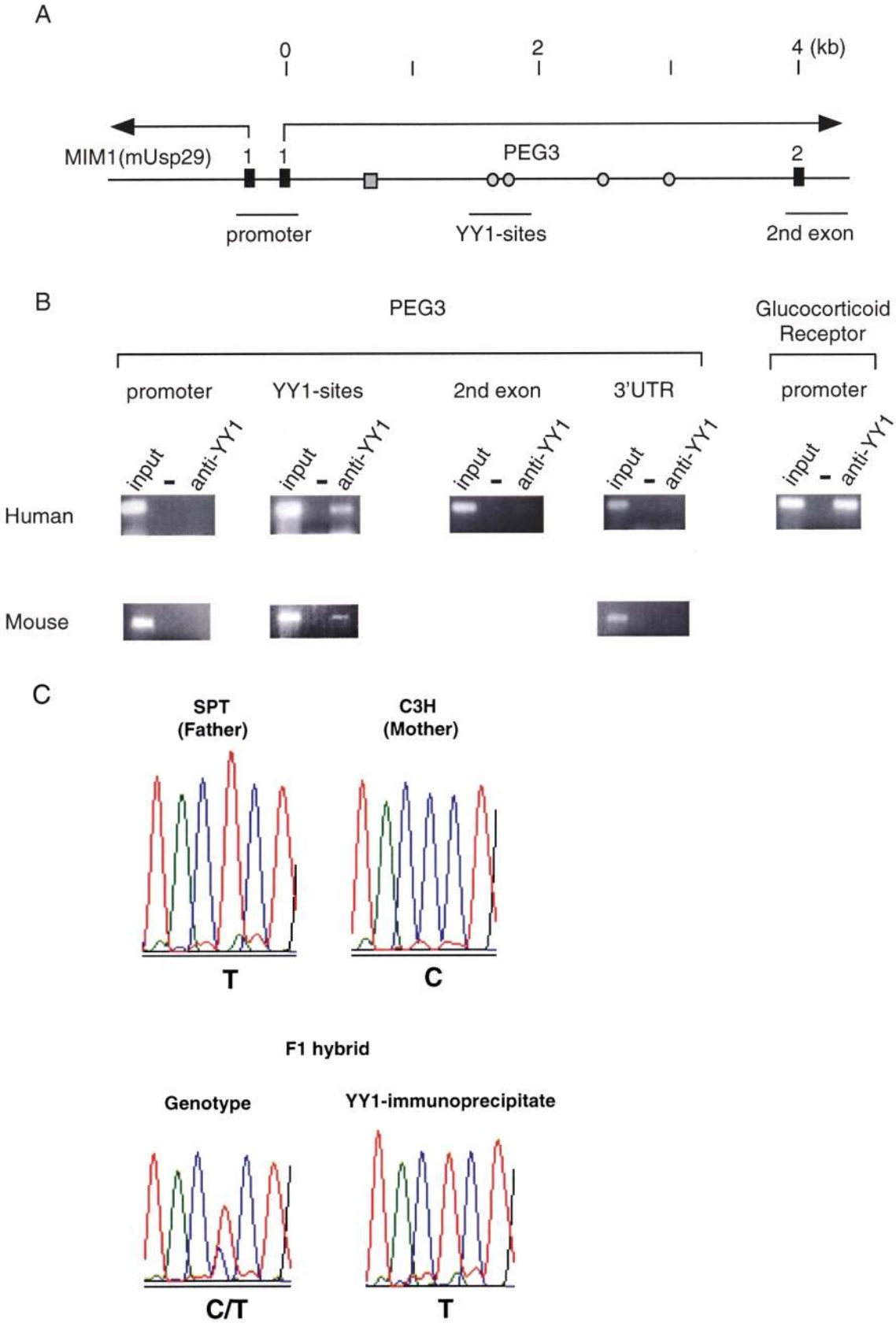


Figure 3. DNA mobility shift assays of CSE1 and CSE2. (**A**) Identification of unknown protein complexes binding to CSE1. (**B**) Identification of YY1 as a binding factor for CSE2. Supershift assays were repeated with a second YY1-antibody that is marked with an asterisk (lane 9–12). Lanes 9 and 10 are same as lanes 2 and 3. The antibody used for lanes 7 and 11 recognizes the C-terminal 20 amino acid of YY1 whereas the antibody for lane 12 recognizes the whole protein of YY1. The shifted position of YY1 by the second antibody is indicated by an arrow. (**C**) Effect of methylation on YY1-binding. All the probes for this assay were designed to contain a sequence identical to that of the CSE2 probe, but with or without methylation of the CpG site on one or both strands; u(−/−) unmethylated, hm(+/−) methylated on top strand, hm(−/+) methylated on bottom strand, m(+/+) methylated on both strands. Lanes 1–4 show results obtained with four different combinations of methylated CSE2 sequences as probes without addition of any competitor. Lanes 5–8 shows data obtained unsing unmethylated CSE2 sequences as a probe and different combinations of methylated CSE2 sequences as competitors. Since this gel image is slanted, lane assignment is shown at both the top and bottom of the gel image for orientation. (**D**) The binding specificity of YY1. The CSE2 probe was competed against the two competitors that have single base changes in the CpG sites. The sequences of each probe used for the analyses are shown at the bottom.



anti-YY1 antibody, of which the antigenic epitope is the entire portion of human YY1. As shown in Fig. 3B lane 12, this antibody shifted the CSE2 probe sequences more significantly as compared to the first antibody.

Since the consensus-binding site of YY1 contains one CpG dinucleotide and the YY1 binding sites of *Peg3* are part of a differentially methylated region, we investigated the effect of methylation on the binding activity of YY1 (Fig. 3C). Despite several previous attempts by others (29,30), this has never been tested unambiguously, primarily due to technical difficulties involved in methylating probes for gel shift assays. To avoid this problem, the top and bottom strands for the CSE2 duplex probe were synthesized separately to contain methyl-cytosines. Four combinations of the CSE2 probe were derived: unmethylated on both top and bottom strands (−/−), methylated on both (+/+), and methylated on only top (+/−) or bottom (−/+) strands. As expected, the binding of YY1 was shown to be very sensitive to methylation. Interestingly, the methylation effect on the binding of YY1 is strand-specific; the YY1-binding to DNA is significantly affected only when the top strand is methylated, suggesting the top strand is the contact site by YY1. A similar observation has also been observed for CTCF (14).

As shown in Fig. 2D, the first intron of *Peg3* also contains many potential YY1 binding sites with one or two base differences compared to the consensus-binding site of YY1. To test the binding capability of these sites, gel shift assays were also performed with probes containing a single base difference. As shown in Fig. 3D, the CpG site is very critical for the binding of YY1; none of the probes with alternative bases at this position can serve as effective binding sites. This indicates that the motifs containing single base differences at the CpG site do not bind YY1 protein.

In vivo binding of YY1 to CSE2s

To test whether YY1 binds to the CSE2 sequences of human *PEG3 in vivo*, chromatin immunoprecipitation (ChIP) assays were performed with human cells that were grown asynchronously. Cells were treated with formaldehyde to crosslink protein with DNA, sonicated for DNA fractionation, and immunoprecipitated with anti-YY1 antibody. As shown in Fig. 4, the YY1-binding sites of *PEG3* (CSE2s) and the promoter region of the glucocorticoid receptor (GR) gene were greatly enriched in the anti-YY1 immunoprecipitates. The GR promoter region contains three YY1 binding sites and involvement of YY1 in GR expression has been documented (42); this sequence therefore serves as a positive control for YY1 binding. In contrast, sequences located immediately adjacent to the PEG3-DMR YY1-binding sites and from the 3′UTR region of *PEG3* were not enriched in the immunoprecipitated DNA. We also performed ChIP assays to test YY1-binding to the mouse *Peg3* locus using brain tissues derived from 10-day-old F_1 hybrid (*M. musculus* × *M. spretus*) mice. Consistent with the human data, sequences containing the Peg3-DMR YY1-binding sites were enriched in anti-YY1 immunoprecipitates but neighboring *Peg3* locus sequences were not enriched in this pool of DNA.

To determine whether YY1 binds preferentially to one of two parental alleles, we sequenced the PCR products generated from immunoprecipitated DNA with the YY1-region primers. As shown in Fig. 4C, the paternal allele was significantly enriched in the YY1-immunoprecipitate DNA, confirming that YY1 binds primarily to the paternal allele of mouse *Peg3 in vivo*. The above results clearly demonstrated that the CSE2s located within the first intron of human and mouse *PEG3* are bound by YY1 *in vivo* in an allele-specific manner.

Methylation status of YY1 binding sites

Since YY1 shows methylation-sensitive binding activity and the PEG3-DMR is differentially methylated in a parent-of origin specific manner (23,24), we tested the methylation status of the actual YY1-binding sites within the mouse region. Since the four nucleotides of the YY1 binding site (GCGC) are also recognition sites of the methylation-sensitive enzyme, *Hha*I, we employed a strategy similar to that used in other imprinted regions involving double digestion with one methylation-sensitive and one non-sensitive restriction enzyme (Fig. 5A). To detect allele-specific methylation, we used DNAs derived from interspecific hybrid offspring (F_1 and F_2) of two closely related mouse species, *M. spretus* and *M. musculus* (C3H). One *Bam*HI-associated RFLP (restriction fragment length polymorphism) was detected with a 1.0 kb genomic DNA probe covering part of the CpG island. Genomic DNA isolated from F_1 hybrid animals was digested with *Bam*HI, and aliquots were re-digested with *Hha*I. As shown in Fig. 5A, the 15 kb *Bam*HI fragment of F_1 DNA (corresponding to the *M. musculus* maternal allele) was resistant to *Hha*I-digestion whereas the 10 kb fragment (representing the *M. spretus* paternal allele) was effectively digested with *Hha*I. In the reciprocal test using F_2 offspring, the 15 kb fragment (the *M. musculus* paternal allele) was digested but the 10 kb fragment (the *M. spretus* maternal allele) was resistant to

Figure 4. ChIP assay of the YY1-binding sites. Asynchronously growing K562 cells were treated with 1% formaldehyde and subject to immunoprecipitation with anti-YY1 antibodies. Precipitated DNA was PCR-amplified with specific primers. (**A**) The positions of primer sets used for PCR amplification. Shown are three primer sets spanning sequences located immediately adjacent to the PEG3-DMR YY1-binding sites. The positions and sequences of the other two remaining primer sets are described in Materials and Methods. (**B**) Top panel: testing *in vivo* YY1-binding for five human target regions. Four of these sequences are derived from the human *PEG3* locus; the promoter region of the human glucocorticoid receptor (GR) gene was included as a positive control. DNA isolated from the sonicated chromatin, taken before the immunoprecipitation step, was used as 'input' DNA whereas DNA samples precipitated with or without anti-YY1 antibodies were used as 'anti-YY1' or '−' template DNA, respectively. Bottom panel: testing *in vivo* YY1-binding of the mouse *Peg3* locus. Three *Peg3* locus sequences regions were tested for enrichment in YY1-antibody immunoprecipitated DNA prepared from 10-day-old F_1 hybrid mouse brains. (**C**) Allele determination of the YY1-immunoprecipitated mouse DNA. The amplified DNA generated with primers surrounding the mouse YY1-binding sites (B), was sequenced and compared with the *M. musculus* (maternal) and *M. spretus* (paternal) parental DNA sequences. Two polymorphic base choices (C or T) that distinguish *M. musculus* from *M. spretus* alleles were detected as a mixed base sequence in PCR products generated from total F_1 genomic DNA, but YY1-immunoprecipitated DNA prepared from the same F_1 animal shows only the paternal allele at that position.

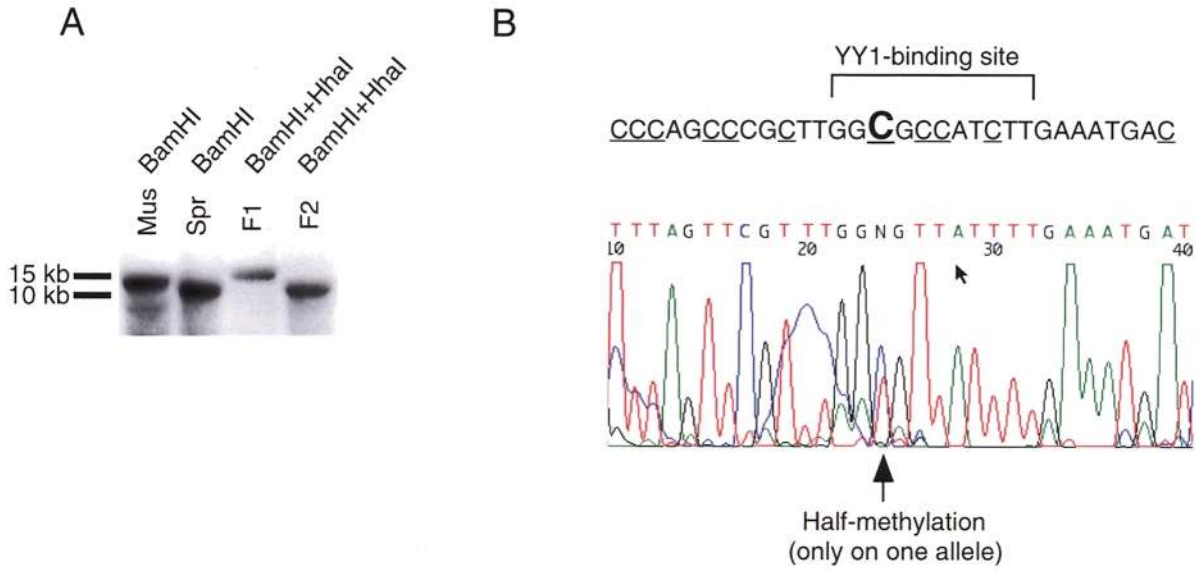


Figure 5. Methylation status of the YY1-binding sites *in vivo*. (**A**) Southern blot experiments using genomic DNAs derived from F_1 and F_2 hybrids of interspecific crossing of *M. musculus* × *M. spretus*. The paternal allele (10 kb) of F_1 animals was digested with *Hha*I (methylation-sensitive) while the maternal allele (15 kb) was resistant to the digestion, indicating methylation on the maternal allele. In the F_2 animals, the 10 kb maternal allele is resistant to *Hha*I digestion. (**B**) Bisulfite sequencing of the YY1-binding sites. The electropherogram shows the half-methylation of the second YY1-binding site.

*Hha*I-digestion. These results indicate specific methylation of maternal YY1 binding sites. To further demonstrate the methylation of the YY1 binding sites, three mouse PEG3-DMR YY1 binding sites were sequenced after bisulfite treatment, which differentiates methylated cytosine from unmethylated cytosine (31). The CpG dinucleotide of all three sites was half-methylated; results from one of these regions are shown in Fig. 5B. Together these results provide strong evidence that YY1-binding and the subsequent function of this protein is allele-specific, with effective YY1 binding occurring only to the unmethylated paternal allele.

Promoter activity of PEG3-DMR

Because of its overlap with the 5′ ends and first exons of the two mouse genes, PEG3-DMR is likely to harbor promoter elements for *Peg3* and *Usp29*. Similarly, the human region is expected to contain elements for regulation of *PEG3* and the noncoding transcript, *MIM1*. To characterize the promoter activity of this region, transient reporter assays were performed using a series of constructs that contained different portions of the human and mouse PEG3-DMR. Since the two genes are transcribed in opposite directions, all constructs were tested in both directions. Because of their close juxtaposition, the region upstream of *Peg3* contains the promoter and first exon of *Usp29*, and vice versa. All the constructs were tested with 4 different cell lines, two of human origin (293, PA-1) and two from rodents (Neuro2A, P19). The promoter assays were repeated at least three different times in each of the four different cell lines. The results are summarized in Fig. 6 and Table 1.

In this series of promoter assays, one DNA segment consistently displayed significant levels of combined promoter and enhancer activity. Similar constructs were made with homologous mouse and human sequences and tested in parallel in all cases summarized below. Constructs containing mouse and human DNA extending ∼1 kb upstream from the start site of *Peg3* and oriented so that the *Peg3* start site was positioned nearest the luciferase reporter in the PGL3-basic vector (construct h2 and m2) showed high levels of promoter activity in all cell lines. By contrast, the same inserts cloned in the opposite orientation in the same vector (construct h1 and m1) did not drive significant luciferase expression in any of the four cell lines. These results indicate that the 1 kb region contains the functional *Peg3* promoter.

In contrast, significant reporter activity was not detected in cells transfected with 1 and 4 kb constructs extending from within the first exon of *Usp29/MIM1* into the first intron of *PEG3* (construct set 3–6 in Fig. 6). The 5′-ends of *Usp29* and *MIM1* have been cloned and sequenced, and lie less than 150 bp from the start of *PEG3* in both species (20); it is therefore highly likely that these DNA segments contain the promoters for *Usp29* and *MIM1*, respectively. It is unclear presently why these constructs did not show promoter activity. Both sets of 1 and 4 kb segments contain sequences within the *PEG3* DMR and it is possible that elements within these regions may be affecting activity.

A 2 kb segment within the first intron of human *PEG3*, surrounding the series of clustered YY1 sites, was also tested for potential enhancer activity by subcloning into the pGL3-promoter vector that has the SV40 promoter (data not shown). The constructs containing the clustered YY1-binding sites did not show any significant level of enhancer activity in most cell lines tested.

Insulator activity of the YY1-binding region

The YY1-binding sites of *Peg3* and the CTCF-binding sites of *H19* show many similarities, such as localization of multiple

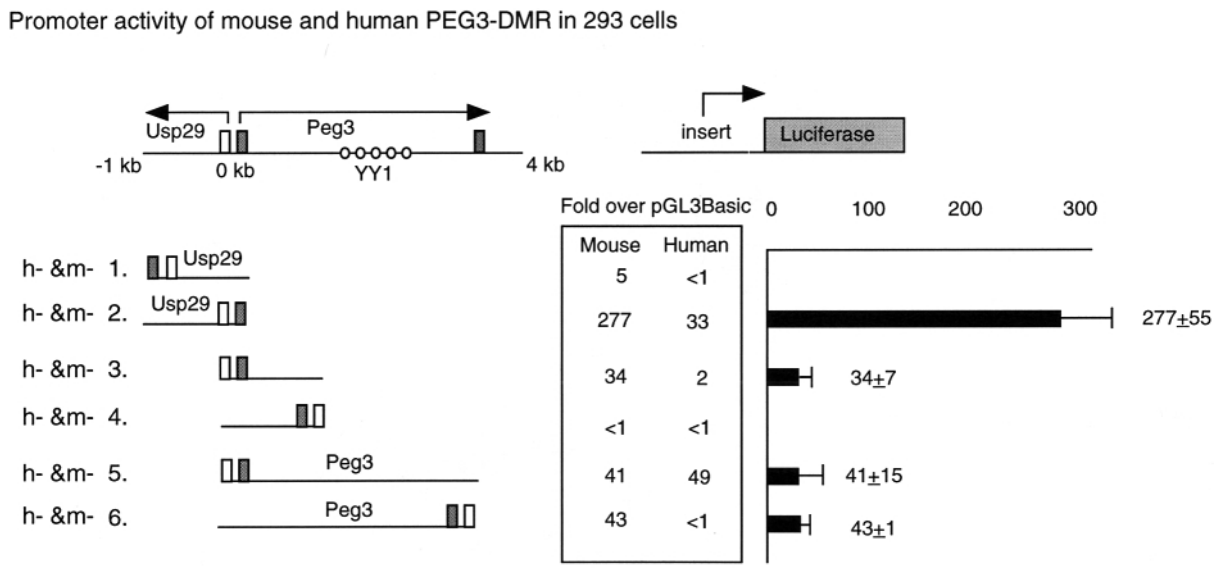


Figure 6. Promoter assays of Peg3-DMR. Several different genomic fragments covering human and mouse Peg3-DMR were subcloned into the pGL3 basic vector that lacks its own promoter and enhancer. The promoter activities of these constructs were measured and compared with a negative control, the pGL3 basic without insert. The relative values, indicating the promoter activity of each fragment, are shown in the middle column, and the mouse values are also shown as graphs. Constructs containing human and mouse genomic DNA fragments are named h1-6 and m1-6, respectively.

Table 1. Promoter assays (increase relative to pGL3Basic vector)

Construct		293	Neuro2A	PA-1	P19
Human	h1	—	—	—	—
	h2	++	++	++	++
	h3	+	+	+	++
	h4	—	—	+	—
	h5	++	++	++	++
	h6	—	—	—	—
Mouse	m1	+	—	NA	NA
	m2	+++	+++	NA	NA
	m3	+	+	+	+
	m4	—	—	+	+
	m5	++	++	++	++
	m6	++	+	+	+

+++, greater than 100-fold; ++, 100–10-fold; +, 10–1-fold; —, <1-fold; NA, not available.

binding sites within differentially methylated CpG islands (DMRs) and methylation-sensitive binding activity of the two transcription factors. We therefore tested whether the YY1-binding region of *Peg3* can function as an insulator (enhancer-blocking element) as has been shown for the CTCF-binding sites of H19 DMR. To test this possibility, we used the assay system that was previously developed to characterize insulator activity within the chicken globin locus (32,33). This assay system utilizes a testing vector, pNI-CD, which carries a neomycin resistance gene (NeoR) reporter under the control of an erythroid-specific enhancer (32).

To test the YY1-binding region of the PEG3-DMR for potential insulator activity, genomic fragments spanning different portions of the human region were subcloned into the testing vector pNI-CD in both orientations. As a positive control, the vector construct containing the 250 bp insulator derived from the chicken globin locus, cHS4, was also included in our assay (construct 2 in Fig. 7). The insulator activity of each construct was measured at least three different trials; the mean values of these measurements with standard errors are summarized for each construct in Fig. 7. The 2 kb genomic fragment including the human YY1-binding sites, hPEG3-YY1, displayed significant levels of insulator activity in both orientations (constructs 7 and 8), with activity levels similar to that of the positive control (construct 2, cHS4). The activity of the construct cloned in one orientation (construct 8) was slightly higher than that of the same fragment inserted in the opposite direction (construct 7). To narrow down the position of active insulator elements within this genomic interval, two subclones containing different portions of hPEG3-YY1 were tested individually (Fragments hPEG3-YY1U and hPEG3-YY1L; constructs 3–6). Both fragments maintained insulator activity in the pNI-CD vector, but both sub-fragments showed insulator activity only in one orientation (constructs 4 and 6) and not in the other (constructs 3 and 5). The activity of hPEG3-YY1U (construct 4) was the highest of all fragments tested, including the chicken globin locus insulator (construct 2, cHS4). The orientation dependence of the smaller fragments remains to be studied further but a similar pattern has been observed from other insulators, including the insulator located within the DMR of the Kcnq1 locus (34). These results indicate that the YY1-binding region may function primarily as an insulator.

DISCUSSION

We have identified two evolutionarily conserved sequence elements in the first intron of *Peg3*, and have demonstrated that these sequences serve as methylation-sensitive binding sites for the *Gli*-type zinc-finger gene, YY1. Our results show that methylation of the CpG sequence inhibits the binding of YY1

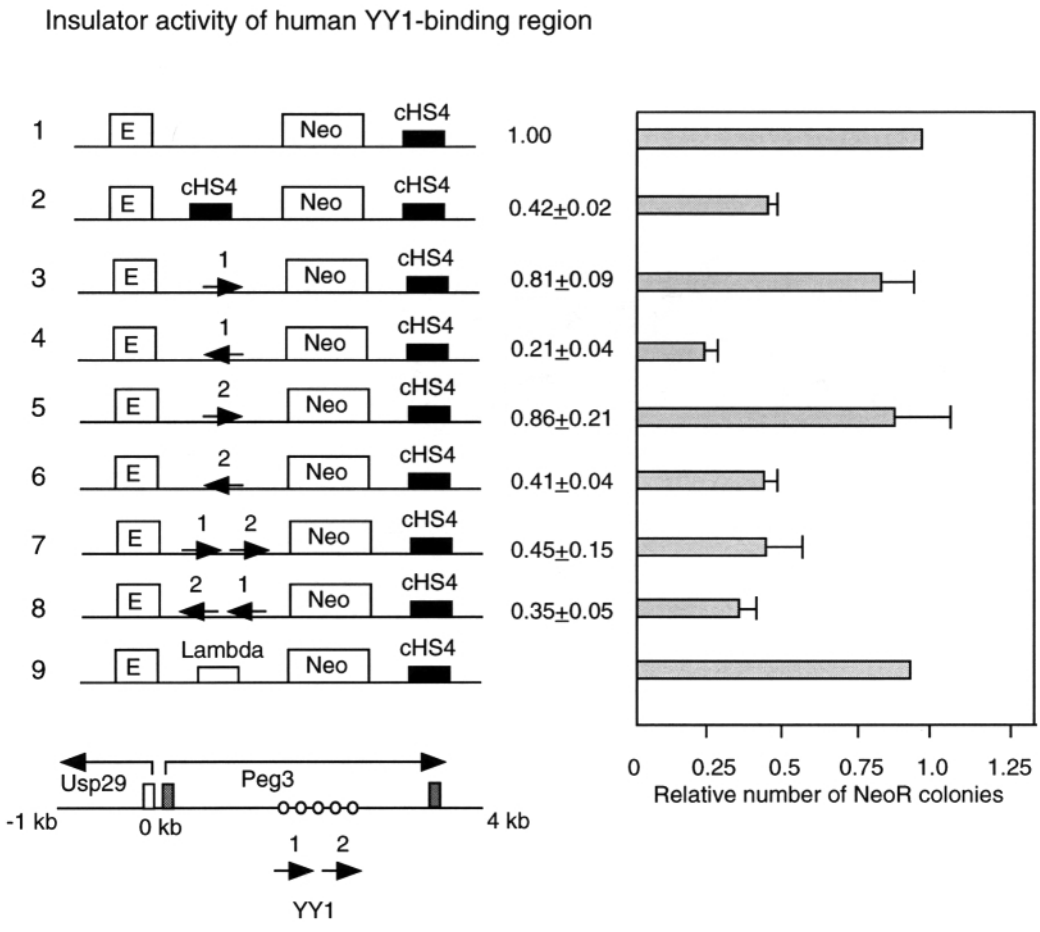


Figure 7. Insulator assay of the YY1-binding region. Several genomic fragments covering different parts of the YY1-binding region within human PEG3-DMR were subcloned into the pNI-CD vector to test insulator activity. The orientation and coverage of the fragments are shown as arrows. The testing fragments were located between an erythroid-specific enhancer (E) and a chicken insulator (cHS4) identified from chicken β-globin locus. The average values, with standard errors, are shown in the middle column and presented as graphs.

protein to the PEG3-DMR sequence and to consensus binding sequences *in vitro*. We have demonstrated that the YY1-binding sites of PEG3-DMR are methylated *in vivo*, specifically on the repressed maternal allele. We have also demonstrated that effective YY1-binding is restricted to the active, paternal allele of the PEG3-DMR. The YY1-binding element functions as a powerful insulator sequence in cell line-based assays. Together these data suggest that the YY1-binding sequence serves as a methylation-sensitive insulator for expression control of *Peg3* and neighboring genes.

The tandemly repeated YY1 binding motifs within the PEG3-DMR are highly conserved within species but comparison of human, mouse and cow repeats reveals a high degree of interspecies divergence outside the highly conserved 11 bp YY1-binding core sequence. These data suggest the independent duplication and continuous homogenization of these repeat units since the divergence of mammalian lineages. Human, mouse and cow each carry more than 10 copies of the repeat unit within the 5 kb PEG3-DMR, including four to six copies with functional binding sites for YY1. Similar types of tandem repeat structures have been observed within DMRs in several other imprinted domains (11,12). Given the unstable nature of such sequence structures in mammalian genomes, their maintenance suggests that the repeats serve an important function. It has been postulated that the repeated sequences may play a role in establishment or maintenance of genomic imprinting, for example by attracting site-specific methylation (11,12). In the case of the PEG3-DMR, it appears that continuous duplication of the short sequence repeats may have served as a mode for maintaining multiple, active copies of the YY1-binding sequence within the PEG3-DMR despite instability and sequence drift during mammalian evolution.

What function does YY1 serve in the *Peg3* region and why are multiple copies of YY1-binding sites required? Results presented here suggest that the YY1-binding region can function as an insulator element. Although these data do not rule out the possibility that the YY1-binding element might also function as a cell type-specific enhancer, none of the enhancers described to date comprise tandem arrays of binding sites for a single transcription factor. In contrast, all of the known invertebrate and vertebrate insulators contain tandem arrays of sites similar to those observed within the PEG3-DMR. Examples include the gypsy and scs/scs' insulators of *Drosophila* that contain multiple binding sites

for SU(HW) and BEAF-32, respectively, and Repeat Organizer of *Xenopus* rRNA genes and the ICR of mammalian *H19* which contain multiple binding sites for CTCF (35). The actual mechanisms are still under study but it is thought that binding of multiple transcription factors might be necessary to induce changes in local chromatin structures that are required for insulator activity (36). Based on these observations, we hypothesize that the YY1-binding region of the PEG3-DMR functions as an insulator influencing expression control of *Peg3* and neighboring imprinted genes.

Since YY1 was initially identified as a binding factor to the LTR region of IAP elements, it has been shown to be involved in the transcriptional control of a large number of viral and cellular genes (37). YY1 functions as a repressor, activator, or transcriptional initiator, depending upon the sequence context of binding sites with respect to other regulatory elements. The *Drosophila* YY1 homolog has been identified as a DNA-binding factor for the Polycomb genes, which are involved in heritable silencing mechanisms (38). The vertebrate YY1 protein has also been shown to interact with a member of the Polycomb group (Pc-G) of genes, *EED* (embryonic ectoderm development) (39,40), suggesting that YY1 also plays a role in vertebrate gene silencing mechanisms. Results presented here add another potential dimension to the functional versatility of this highly conserved *Gli*-type zinc-finger protein factor, by implicating YY1 in methylation-sensitive insulator activity and control of imprinted gene expression.

Based on the shared features of this region with other ICRs, we predict that the PEG3-DMR plays a significant role in control of imprinted expression. It is interesting to note that in mouse a maternally expressed gene, *Zim1*, lies directly downstream of *Peg3* and the YY1-binding domain is positioned between paternally expressed and maternally expressed promoters (Fig. 1A). Although there are some similarities between PEG3-DMR and *H19* ICR, understanding the potential function of the PEG3-DMR in regulating imprinted expression of neighboring genes will require further studies. The PEG3-DMR may also serve functions that are not directly related to regulating expression of imprinted genes. The borders of this imprinted domain are filled with clustered gene families, including an MKR5-like zinc-finger gene cluster in HSA19 and the *OLFR5* cluster of olfactory receptor genes in mouse (41) (www.ensembl.org/). Consistent with the function of insulator sequences within the β-globin gene cluster (33,35), the PEG3-DMR might alternatively act as a 'boundary' element that shields these gene clusters from the influence of elements involved in establishing the chromatin structures and allele-specific methylation associated with imprinting control. The present study sets the stage for informed design of targeted mutations in mice, to delineate the precise functional role of allele-specific YY1 binding in regulation of this conserved imprinted domain.

MATERIALS AND METHODS

Cow BAC isolation and sequence analyses

Two different DNA fragments corresponding to the coding region of mouse *Peg3* (GenBank accession no. AB003040, 181–401 and 4705–5080, respectively) were used for screening the cow BAC libraries (RPCI42; CHORI, Oakland, CA, USA). Four identified BACs (RPCI42-91A11, 246L9, 184C9, 194K5) were further analyzed with pulse-field-gel electrophoresis and subsequent Southern blot experiments to check the integrity of identified BACs. One of these BACs, 246L9, was submitted for sequencing to the Joint Genome Institute of DOE (Walnut Creek, CA, USA). The 8 kb genomic sequences surrounding the first exon of *Peg3* from human, mouse and cow (GenBank accession no. AC006115, AC020961, AC073666) were initially analyzed using 2Blast (www.ncbi.nlm.nih.gov/BLAST/), VISTA (www.gsd.lbl.gov/vista/), and PIP plot (http://nog.cse.psu.edu/pipmaker/). Later, the three sequences were further analyzed using the printrepeat program of the UNIX computer to visualize tandem-repeat sequence structures of the first intron of *Peg3*.

DNA mobility shift assays

Gel Shift Assay systems (Promega, Madison, WI, USA) were used for all our DNA mobility shift assays. In brief, about 4 μg of nuclear extracts were first mixed with binding buffer and unlabeled competitor probes with varying amounts (10–100 pmol) for 20 min at room temperature and later mixed with the P^{32}-labeled duplex probes (1 pmol) for an additional 10 min at room temperature. The reaction mixtures were separated on 5% native polyacrylamide gel (acryl : bis = 37.5 : 1) in 0.5× Tris buffer at pH 8.0 for 2 h, and the separated gels were exposed to X-ray films for 1 h. For supershift assays, two polyclonal antibodies raised against the C-terminal and entire regions of human YY1, respectively, were obtained from a commercial firm (SantaCruz Biochem, Santa Cruz, CA, USA) and the antibody against human Usp29 was prepared by another commercial firm (SynPep, Dublin, CA, USA). All the gel shift assays used nuclear extracts prepared from human HeLa cell lines (Promega, Madison, WI, USA).

The following oligonucleotides are the probes for our gel shift assays: CSE1 (5′-GGGGGCGGGGCTTGTTTGCCGCAGTGGTGGTGGGAGGGGCG-3′, 5′-CGCCCCTCCCACCACCACTGCGGCAAACAAGCCCCGCCCCC-3′); CSE1a (5′-GGGGGCGGGGCTTGTTTGCCGCAGTGGT-3′, 5′-ACCACTGCGGCAAACAAGCCCCGCCCCC-3′); CSE1b (5′-TTGTTTGCCGCAGTGGTGGTGGGAGGGGCG-3′, 5′-CGCCCCTCCCACCACCACTGCGGCAAACAA-3′); CSE2 (5′-CCCACCCACCTGGGCGCCATCTTTAATGAAAG-3′, 5′-CTTTCATTAAAGATGGCGCCCAGGTGGGTGGG-3′); CSE2a (5′-CCCACCCACCTGGGTGCCATCTTTAATGAAAG-3′, 5′-CTTTCATTAAAGATGGCACCCAGGTGGGTGGG-3′); CSE2b (5′-CCCACCCACCTGGGCACCATCTTTAATGAAAG-3′, 5′-CTTTCATTAAAGATGGTGCCCAGGTGGGTGGG-3′); yy1 (5′-CGCTCCGCGCCATCTTGGCGGCTGGT-3′, 5′-ACCAGCCGCCAAGATGGCGCGGAGCG-3′); yy2 (5′-CGCTCCGCATTATCTTGGCGGCTGGT-3′, 5′-ACCAGCCGCCAAGATAATGCGGAGCG-3′). The methylated oligonucleotides for CSE2 were prepared using methyl-cytosines (NEB, Beverly, MA, USA).

Chromatin immunoprecipitation assay

Asynchronously grown K562 cells were used for our ChIP assays to test the YY1-binding on the human *PEG3* locus.

ChIP assays were performed according to the protocol provided by Upstate Biotechnology (Upstate Biotech, Lake Placid, NY, USA) with minimal modifications. Briefly, formaldehyde was added into the K562 cells (1.0×10^7) at the final concentration of 1% and incubated at 37°C for 10 min. The treated cells were sonicated for DNA fractionation and further precipitated with anti-YY1 antibodies (SantaCruz Biochem, Santa Cruz, CA, USA). The precipitated DNAs were used as template DNAs for PCR amplification with the following primer sets: PEG3-promoter (5′-CAGCCCAGGGTGGACATCTC-3′, 5′-CAGCCTTCGCGCCAAACTGTT-3′), PEG3-YY1 sites (5′-CAAGGAGCGCGGCACTCCAC-3′, 5′-GGCAGATGCGGCGGGGCAAG-3′), PEG3-2nd exon (5′-TGACACTTTCCTACTGCATATG-3′, 5′-CCACAGCTTTAATTCCCAAATAG-3′), PEG3-3′UTR (5′-ACCAACTCAATATGTGTTTG-3′, 5′-GATGAAAGGTTATTAGCCTG-3′), and GR-promoter (5′-CCCCCTGCTCTGACATCTT-3′, 5′-CTTTTCCGAGGTGGCGAGTATC-3′). PCR amplifications were carried out for 33 cycles using standard PCR conditions.

To test *in vivo* YY1 binding at the mouse *Peg3* locus, brain tissue from 10-day-old F_1 hybrid animals were used for ChIP assays. One gram of tissue was homogenized in 10 ml of PBS buffer and treated with 1% formaldehyde for 15 min at 37°C. The following steps for ChIP assays were same as the ChIP experiments of the human *PEG3* locus. The primer sets used for the mouse *Peg3* locus are as follows: Peg3-promoter (5′-CAGCCCAGGGTGGACATCTC-3′, 5′-GCATCCCTTCACACCCACATC-3′), Peg3-YY1 (5′-GGAACTGGTTTCCAATCAATCATGAC-3′, 5′-GCAGGAGGTGGGCTTTAGGGA-3′), and Peg3-3′UTR (5′-CTTCTGGAAGCCGACATTAT-3′, 5′-CCTGATCAATGGGTTCCTTG-3′). Both human and mouse ChIP assays were repeated at least three different times.

Methylation studies of the YY1-binding sites

The methylation status of the YY1-binding sites was analyzed using genomic DNAs derived from the F_1 and F_2 hybrid offspring of the cross of C3H (*M. musculus*) and *M. spretus*. About 10 µg of DNAs were first digested with *Bam*HI for 12 h and later digested with methylation-sensitive *Hha*I for an additional 4 h. The digest DNAs were separated on 0.8% agarose gel electrophoresis, blotted onto nylon membranes, hybridized with one P^{32}-labeled probe derived from the 1 kb genomic region surrounding the first exon of mouse *Peg3* (GenBank accession no. AC020961, 106127-107242). For bisulfite sequencing of the YY1-binding site, we followed an established protocol (31) with minimum modifications. In brief, mouse genomic DNAs (4 µg) was first digested with *EcoRI* for 16 h, denatured with 0.3 M NaOH for 30 min. The prepared DNAs were treated with sodium bisulfite for 16 h at 55°C. After all the chemical treatments, the DNAs were washed with the Wizard DNA clean-up system (Promega, Madison, WI, USA) and dissolved with 50 µl of TE for storage. One microliter of the treated DNA was used as a template for the PCR amplification. The three YY1-binding sites of mouse *Peg3* were amplified and analyzed for sequencing. For the amplification of the YY1-binding sites, nested PCRs for each YY1-binding site were performed at the annealing temperature ranging from 45 to 50°C. The oligonucleotide sequences for the YY1-binding site presented in Fig. 4B are as follows; the first round (5′-GGTGTTATTATGAGTGATAGTTTAG-3′, 5′-TCAATCAAAACCTAAACCCCACCCTC-3′) and the second round (5′-TTTAGATGGTTGTAGGAGGTGGGTT-3′, 5′-CCTCAAACCCACCCCCTAAAACCAA-3′). To monitor the efficiency of the bisulfite treatment, the PCR products were subcloned into the TA cloning vector and 10 different clones were sequenced individually. If more than 90% of cytosine was converted into thymidine, those DNA samples were selected for our bisulfite sequencing analyses. Our analyses on the three YY1-binding sites were repeated at least three different times. The sequence information regarding oligonucleotides for the other two YY1-binding sites is available upon request from the authors.

Promoter and enhancer assays

DNA fragments of interest were generated by PCR for subsequent cloning into the pGL3basic and pGL3promoter vectors (Promega, Madison, WI, USA). Each fragment was cloned in both orientations. We list here the locations of the oligonucleotides used for subcloning experiments. Human DNA fragments were derived from BAC CIT-B-470F8 (GenBank accession no. AC006115): clone h1 and h2 (163502/164513), h3 and h4 (163794/162832), h5 and h6 (163794/161700). Mouse DNA fragments were from BAC RG-MBAC247N19 (GenBank accession no. AC020961): clone m1 and m2 (106891/107836), m3 and m4 (107242/106127), m5 and m6 (107242/103099). All cell lines were co-transfected with the construct of interest and pTK-Renilla as an internal control, using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA, USA) as per the manufacturer's instruction. Twenty-four hours post transfection, cells were analyzed for luciferase activity using Dual-Luciferase Assay system (Promega, Madison, WI, USA) as per the manufacturer's instruction. Plates were read in triplicate by a LumiCount plate reader (Packard, Meriden, CT, USA). Experiments were performed in 96-well plates in duplicate and each experiment was performed more than three times.

Insulator assays

DNA fragments of interest were cloned into the *AscI* site of pNI-CD (generous gift from Dr Gary Felsenfeld and Dr Adam West). Each fragment was cloned in both orientations. The locations of the tested DNA fragments within human BAC CIT-B-470F8 (GenBank accession no. AC006115) are hPEG3-YY1U (161973/161189) and hPEG3-YY1L (161209/160414). Constructs were transfected into K562 cells by electroporation at 200 V, 1000 mF (double pulse) using a Bio-Rad Gene Pulser II. After a 10 min recovery on ice, cells were plated into RPMI supplemented with 10% FBS, 200 mM L-glut and Pen/Strep. Twenty-four hours post transfection, cells were washed and resuspended in Improved MEM zinc option (GibcoBRL, Rockville, MD, USA). Cells were plated into 0.3% soft agar with 1050 µg/ml Geneticin (GibcoBRL, Rockville, MD, USA) and incubated at 37°C for 18–21 days.

ACKNOWLEDGEMENTS

We would like to extend special thanks to Susan Lucas and the sequencing staff of the D.O.E. Joint Genome Institute (Walnut

Creek, CA, USA) for sequencing the bovine BACs. We thank Drs Gary Felsenfeld and Adam West at NIH for providing us two insulating vectors, pJC-13 and pNI-CD. This work was performed under the auspices of the US Department of Energy, Office of Biology and Environmental Research, by LLNL under contract no. W-7405-ENG-48.

REFERENCES

1. Bartolomei, M.S. and Tilghman, S.M. (1997) Genomic imprinting in mammals. *A. Rev. Genet.*, **31**, 493–525.
2. Tilghman, S.M. (1999) The sins of the fathers and mothers: genomic imprinting in mammalian development. *Cell*, **96**, 185–193.
3. Nicholls, R.D., Saitoh, S. and Horsthemke, B. (1998) Imprinting in Prader–Willi and Angelman syndromes. *Trends. Genet.*, **14**, 194–200.
4. Reik, W. and Maher, E.R. (1997) Imprinting in clusters: lessons from Beckwith-Wiedemann syndrome. *Trends. Genet.*, **13**, 330–334.
5. Nicholls, R.D. (1994) New insights reveal complex mechanisms involved in genomic imprinting. *Am. J. Hum. Genet.*, **54**, 733–740.
6. Dittrich, B., Buiting, K., Korn, B., Rickard, S., Buxton, J., Saitoh, S., Nicholls, R.D., Poustka, A., Winterpacht, A., Zabel, B. and Horsthemke, B. (1996) Imprint switching on human chromosome 15 may involve alternative transcripts of the *SNRPN* gene. *Nat. Genet.*, **14**, 163–170.
7. Reik, W. and Walter, J. (2001) Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.*, **2**, 21–32.
8. Leighton, P.A., Ingram, R.S., Eggenschwiler, J., Efstratiadis, A. and Tilghman, S.M. (1995) Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature*, **375**, 34–39.
9. Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F. and Barlow, D.P. (1997) Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. *Nature*, **389**, 745–749.
10. Yang, T., Adamson, T.E., Resnick, J.L., Leff, S., Wevrick, R., Francke, U., Jenkins, N.A., Copeland, N.G. and Brannan, C.I. (1998) A mouse model for Prader-Willi syndrome imprinting-centre mutations. *Nat. Genet.*, **19**, 25–31.
11. Constancia, M., Pickard, B., Kelsey, G. and Reik, W. (1998) Imprinting mechanisms. *Genome Res.*, **8**, 881–900.
12. Reik, W. and Walter, J. (1998) Imprinting mechanisms in mammals. *Curr. Opin. Genet. Devl.*, **8**, 154–164.
13. Bell, A.C. and Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature*, **405**, 482–485.
14. 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.
15. Kuroiwa, Y., Kaneko-Ishino, T., Kagitani, F., Kohda, T., Li, L.-L., Tada, M., Suzuki, R., Yokoyama, M., Shiroishi, T., Wakana, S., Barton, S.C., Ishino, F. and Surani, M.A. (1996) *Peg3* imprinted gene on proximal chromosome 7 encodes for a zinc finger protein. *Nat. Genet.*, **12**, 186–190.
16. Cattanach, B.M. and Kirk, M. (1985) Differential activity of maternaly and paternally derived chromosome regions in mice. *Nature*, **315**, 496–498.
17. Kim, J., Noskov, V., Lu, X., Bergmann, A., Ren, X., Warth, T., Richardson, P., Kouprina, N. and Stubbs, L. (2000) Discovery of a novel, paternally expressed ubiquitin-specific processing protease gene through comparative analysis of an imprinted region of mouse chromosome 7 and human chromosome 19q13.4. *Genome Res.*, **10**, 1138–1147.
18. Kim, J., Bergmann, A., Wehri, E., Lu, X. and Stubbs, L. (2001) Imprinting and evolution of two Kruppel-type zinc-finger genes, *Zim3* and *ZNF264*, located in the *PEG3*/*USP29*-imprinted domain. *Genomics*, **77**, 91–98.
19. Kim, J., Lu, X. and Stubbs, L. (1999) *Zim1*, a maternally expressed mouse Kruppel-type zinc-finger gene located in proximal chromosome 7. *Hum. Mol. Genet.*, **8**, 847–854.
20. Kim, J., Bergmann, A. and Stubbs, L. (2000) Exon sharing of a novel human zinc-finger gene, *ZIM2*, and paternally expressed gene 3 (*PEG3*). *Genomics*, **64**, 114–118.
21. Kim, J., Ashworth, L., Branscomb, E. and Stubbs, L. (1997) The human homologue of a mouse imprinted gene, *Peg3*, maps to a zinc finger gene-rich region of human chromosome 19q13.4. *Genome Res.*, **7**, 532–540.
22. Kim, J., Gordon, L., Dehal, P., Badri, H., Christensen, M., Groza, M., Ha, C., Hammond, S., Vargas, M., Wehri, E., Wagner, M., Olsen, A. and Stubbs, L. (2001) Homology-driven assembly of a sequence-ready mouse BAC contig map spanning regions related to the 46-MB gene-rich euchromatic segments of human chromosome 19. *Genomics*, **74**, 129–141.
23. Murphy, S.K., Wylie, A.A. and Jirtle, R.L. (2001) Imprinting of *Peg3*, the human homologue of a mouse gene involved in nurturing behavior. *Genomics*, **71**, 110–117.
24. Li, L.L., Szeto, I.Y., Cattanach, B.M., Ishino, F. and Surani, M.A. (2000) Organization and parent-of-origin-specific methylation of imprinted *Peg3* gene on mouse proximal chromosome 7. *Genomics*, **63**, 333–340.
25. Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H. and Trasler, J.M. (2002) Methylation dynamics of imprinted genes in mouse germ cells. *Genomics*, **79**, 530–538.
26. Kohda, T., Asai, A., Kuroiwa, Y., Kobayashi, S., Aisaka, K., Nagashima, G., Yoshida, M.C., Kondo, Y., Kagiyama, N., Kirino, T., Kaneko-Ishino, T. and Ishino, F (2001) Tumor suppressor activity of human imprinted gene *PEG3* in a Glioma cell line. *Genes Cells*, **6**, 237–247.
27. Maegawa, S., Yoshioka, H., Itaba, N., Kubota, N., Nishihara, S., Shirayoshi, Y., Nanba, E. and Oshimura, M. (2001) Epigenetic silencing of *PEG3* gene expression in human Glioma cell lines. *Mol. Carcinogen*, **31**, 1–9.
28. Bird. A.P. (1986) CpG-rich islands and the function of DNA methylation. *Nature*, **321**, 209–213.
29. Satyamoorthy, K., Park, K., Atchison, M.L. and Howe, C.C. (1993) The intracisternal A-particle upstream element interacts with transcription factor YY1 to activate transcription: pleiotropic effects of YY1 on distinct DNA promoter elements. *Mol. Cell. Biol.*, **13**, 6621–6628.
30. Gaston, K. and Fried, M. (1995) CpG methylation has differential effects on the binding of YY1 and ETS proteins to the bi-directional promoter of the Surf-1 and Surf-2 genes. *Nucl. Acids Res.*, **23**, 901–909.
31. Feil, R., Charlton, J., Bird, A.P., Walter, J. and Reik, W. (1994) Methylation analysis on individual chromosomes: improved protocol for bisulfite genomic sequencing. *Nucl. Acids Res.*, **22**, 695–696.
32. Chung, J.H., Bell, A.C. and Felsenfeld, G. (1997) Characterization of the chicken beta-globin insulator. *Proc. Natl Acad. Sci. USA*, **94**, 575–580.
33. 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**, 387–396.
34. Kanduri, C., Fitzpatrick, G., Mukhopadhyay, R., Kanduri, M., Lobanenkov, V., Higgins, M. and Ohlsson, R. (2002) A differentially methylated imprinting control region within the Kcnq1 locus harbors a methylation-sensitive chromatin insulator. *J. Biol. Chem.*, **277**, 18106–18110.
35. Bell, A.C., West, A.G. and Felsenfeld, G. (2001) Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science*, **291**, 447–450.
36. Gerasimova, T.I. and Corces, V.G. (2001) Chromatin insulator and boundaries: Effects on transcription and nuclear organization. *A. Rev. Genet.*, **35**, 193–208.
37. Thomas, M.J. and Seto, E. (1999) Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene*, **236**, 197–208.
38. Brown, J.S., Mucci, D., Whiteley, M., Dirkensen, M.L. and Kassis, J.A. (1998) The drosophila polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol. Cell*, **1**, 1057–1064.
39. Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A. and Magnuson, T. (2001) Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat. Genet.*, **28**, 371–375.
40. Satijn, D.P.E., Hamer, K.M., Blaawen, J.D. and Otte, A.P. (2001) The polycomb group protein EED interact with YY1, and both proteins induce neural tissue in *Xenopus* embryos. *Mol. Cell. Biol.*, **21**, 1360–1369.
41. Dehal, P., Predki, P., Olsen, A.S., Kobayashi, A., Folta, P., Lucas, S., Land, M., Terry, A., Ecale, Zhou, C.L. *et al.* (2001) Comparison of human chromosome 19 and related regions in mouse: conservative and lineage specific evolution. *Science*, **293**, 104–111.
42. Breslin, M.B. and Vedeckis, W.V. (1998) The human glucocorticoid receptor promoter upstream sequences contain binding sites for the ubiquitous transcription factor, Yin Yang 1. *J. Steroid. Biochem. Mol. Biol.*, **67**, 369–381.