Development 129, 897-904 (2002)
Printed in Great Britain © The Company of Biologists Limited 2002
DEV5989

The chicken β-globin insulator element conveys chromatin boundary activity but not imprinting at the mouse *Igf2*/*H19* domain

Piroska E. Szabó, Shih-Huey E. Tang, Michael R. Reed*, Francisco J. Silva, Walter M. K. Tsark and Jeffrey R. Mann†

Division of Biology, Beckman Research Institute of the City of Hope, 1450 East Duarte Road, Duarte, California 91010-3011, USA
*Present address: Life Sciences Division, Motorola, 757 South Raymond Avenue, Pasadena, California 91105-3250, USA
†Author for correspondence

Accepted 26 November 2001

SUMMARY

Imprinting of the mouse insulin-like growth factor 2 (*Igf2*) and *H19* genes is regulated by an imprinting control region (ICR). The hypomethylated maternal copy functions as a chromatin insulator through the binding of CTCF and prevents *Igf2* activation in *cis*, while hypermethylation of the paternal copy inactivates insulator function and leads to inactivation of *H19* in *cis*. The specificity of the ICR sequence for mediating imprinting and chromatin insulation was investigated by substituting it for two copies of the chicken β-globin insulator element, $(Ch\beta GI)_2$, in mice. This introduced sequence resembles the ICR in size, and in containing CTCF-binding sites and CpGs, but otherwise lacks homology. On maternal inheritance, the $(Ch\beta GI)_2$ was hypomethylated and displayed full chromatin insulator activity. Monoallelic expression of *Igf2* and *H19* was retained and mice were of normal size. These results suggest that the ICR sequence, aside from CTCF-binding sites, is not uniquely specialized for chromatin insulation at the *Igf2/H19* region. On paternal inheritance, the $(Ch\beta GI)_2$ was also hypomethylated and displayed strong insulator activity – fetuses possessed very low levels of *Igf2* RNA and were greatly reduced in size, being as small as *Igf2*-null mutants. Furthermore, the paternal *H19* allele was active. These results suggest that differential ICR methylation in the female and male germ lines is not acquired through differential binding of CTCF. Rather, it is likely to be acquired through a separate or downstream process.

Key words: Chromatin Insulation, Imprinting, Methylation, Epigenetics, Insulin-like growth factor 2, Gene regulation, Mouse

INTRODUCTION

In placental mammals a small subset of genes is expressed from only one of the two available alleles, according to parental origin – the imprinted genes (Bartolomei and Tilghman, 1997; Ferguson-Smith and Surani, 2001) (see http://www.mgu.har.mrc.ac.uk/imprinting/imprinting.html). As the two alleles of imprinted genes exhibit differential epigenetic states in the same cell, and these states are faithfully maintained throughout cell division, they constitute a unique resource for the understanding of epigenetic processes of gene regulation during development.

Much evidence has accumulated that the methylation of cytosine residues in CpGs of DNA is an integral part of the system of genomic imprinting. Most, if not all, imprinted genes are associated with differentially methylated regions (DMRs) (Mann et al., 2000) and disruption of this differential methylation, usually through loss of hypermethylation at one allele, results in loss of monoallelic expression (Caspary et al., 1998; Howell et al., 2001; Li et al., 1993; Tucker et al., 1996). Primary DMRs are those DMRs inherited from the gametes, and their methylation may constitute the epigenetic mark that transmits imprinting information from gamete to embryo (Nabetani et al., 1997; Shibata et al., 1997; Shibata et al., 1998; Tremblay et al., 1997; Tremblay et al., 1995). A stronger case for this can be made for sperm in comparison to egg DNA, because, being packaged by protamines, there is no opportunity for the transmission of epigenetic information by a 'histone code' (Jenuwein and Allis, 2001). Aside from DNA methylation, sperm DNA may be relatively epigenetically naïve until it enters the ooplasm.

To understand the mechanisms of genomic imprinting, much work has focused on the imprinted genes insulin-like growth factor 2 (*Igf2*), an embryonic mitogen, and *H19*, which produces an untranslated RNA with no known function. In mouse, these genes are located ~80 kb apart on distal chromosome 7 and are coordinately expressed in tissues of mesoderm and definitive endoderm origin. This coordination is due to the sharing of a suite of enhancers located downstream of *H19* (Bartolomei et al., 1993; Leighton et al., 1995b; Zemel et al., 1992). Monoallelic expression of the two genes is regulated by an intervening imprinting control element (ICR) located –2 kb to –4.4 kb relative to the transcription start site (RTSS) of *H19* (Leighton et al., 1995a; Ripoche et al., 1997; Thorvaldsen et al., 1998). On the maternal chromosome, the

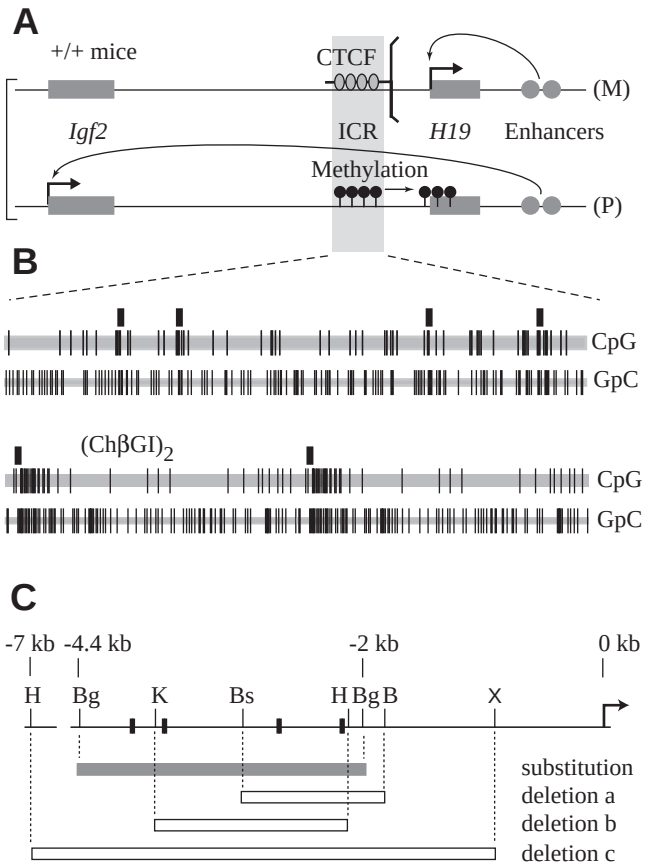

Fig. 1. (A) Gene regulation at the *Igf2* and *H19* loci. (M), Maternal allele; (P), Paternal allele. (B) Nucleotide structure of the 2.4 kb ICR and the 2.4 kb $(ChβGI)_2$. Vertical lines are CpGs or GpCs. CTCF-binding sites are indicated (gray rectangles) and are in the 5′-CCGCnnGGnGGCAG-3′ orientation. Sequence of the ICR is as published elsewhere (Ishihara et al., 1998) (GenBank Accession Number, AF049091) and is defined here as a *Bgl*II fragment at –2.0 to –4.4 kb RTSS of *H19*. To obtain the ChβGI sequence, we sequenced the outer 1.2 kb *Xba*I fragment of plasmid pJC13-1 (Chung et al., 1993). This sequence has been deposited with GenBank (Accession Number, AY040835). (C) Comparison of the present ICR substitution (gray bar) with three independent ICR deletions (open bars): deletion a (Thorvaldsen et al., 1998), deletion b (Drewell et al., 2000) and deletion c (Srivastava et al., 2000). *H19* transcription start site (arrow). X, *Xba*I; B, *Bam*HI; Bg, *Bgl*II; H, *Hin*dIII; Bs, *Bsp*EI; K, *Kpn*I.

ICR acts as a chromatin insulator through the binding of the protein CTCF and prevents access of the *Igf2* promoter to the enhancers (Bell and Felsenfeld, 2000; Bell et al., 1999; Hark et al., 2000; Kanduri et al., 2000b; Srivastava et al., 2000; Szabó et al., 2000). On the paternal chromosome, the ICR lacks insulator activity and the paternal *Igf2* promoter now accesses the enhancers (Fig. 1A). The idea that methylation inhibits insulator activity of the ICR (Schmidt et al., 1999) has considerable support from a number of studies: (1) loss of paternal ICR methylation leads to loss of *Igf2* activity in *cis* or the gain of insulator function (Li et al., 1993); (2) no protein binding is present at the four CTCF sites of the hypermethylated paternal ICR, as shown by in vivo footprinting (Szabó et al., 2000), and no CTCF protein is associated with the paternal ICR, as shown by chromatin immunoprecipitation (Kanduri et al., 2000b); (3) pre-methylation of the ICR leads to the loss of CTCF binding in electrophoretic mobility shift assays – methylation of only the single CpG in the consensus binding sequence CCGCnnGGnGGCAG is sufficient for inhibition (Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000b); (4) gain of maternal ICR methylation leads to gain of *Igf2* activity in *cis*, or the loss of insulator function (Reed et al., 2001); and (5) methylation of the ICR in vitro leads to loss of insulator activity in an episome system (Holmgren et al., 2001).

The hypermethylated paternal ICR, while lacking insulator activity, acts to inactivate the *H19* promoter in *cis* during early development (Srivastava et al., 2000; Thorvaldsen et al., 1998). This element becomes hypermethylated and packaged into a closed chromatin structure (Bartolomei et al., 1993; Ferguson-Smith et al., 1993; Szabó et al., 2000). As this epigenetic information is inherited from sperm, it is likely that this function is dependent on the methylation in the ICR.

The germline-specific processes that determine differential ICR methylation, which could be viewed as equivalent to the imprinting mechanisms, are undefined. In addition, it is not known whether these processes are entirely separate from the later somatic ICR functions of chromatin insulation and *H19* promoter silencing, or if they are in some way connected. What is apparent is that the acquisition of differential ICR methylation is an autonomous function, at least when the ICR is moved to another location within the *Igf2/H19* region (Srivastava et al., 2000). Furthermore, the property is retained, although inconsistently, in randomly integrated multiple copy *H19* transgenes (Brenton et al., 1999; Elson and Bartolomei, 1997; Stadnick et al., 1999).

To examine these questions, a dissection of the ICR sequences through gene targeting should identify those that are important for mediating its various properties. In this study, we describe the results of a related approach in which the ICR was substituted for two copies of the chicken β-globin insulator element (ChβGI) in mice. This introduced sequence, $(ChβGI)_2$, resembles the ICR in size, and in containing CTCF-binding sites and CpGs, but otherwise lacks homology. We hoped to shed light on the answers to two questions: (1) can the $(ChβGI)_2$ effectively substitute for the ICR as a chromatin insulator; and (2) if so, could differential CTCF binding in the two germlines be a sufficient explanation for the establishment of differential methylation and imprinting in the context of the *Igf2/H19* domain? This latter possibility should not be considered unlikely on the basis that other CTCF insulator sites in the genome are not associated with paternal methylation in somatic cells. As their methylation state in sperm has not been examined, it is possible that it is acquired in the male germline, but then lost during fertilization as part of general demethylation of the paternal genome at this stage (Oswald et al., 2000).

We show that on maternal and paternal inheritance, the $(ChβGI)_2$ functioned as a strong chromatin insulator, suggesting that the ICR is not uniquely specialized for chromatin insulation at the *Igf2/H19* domain. In addition, on maternal and paternal inheritance, the $(ChβGI)_2$ was hypomethylated, i.e. the property of differential methylation acquisition or imprinting was lost. This indicates that differential ICR methylation is not mediated through differential CTCF binding.

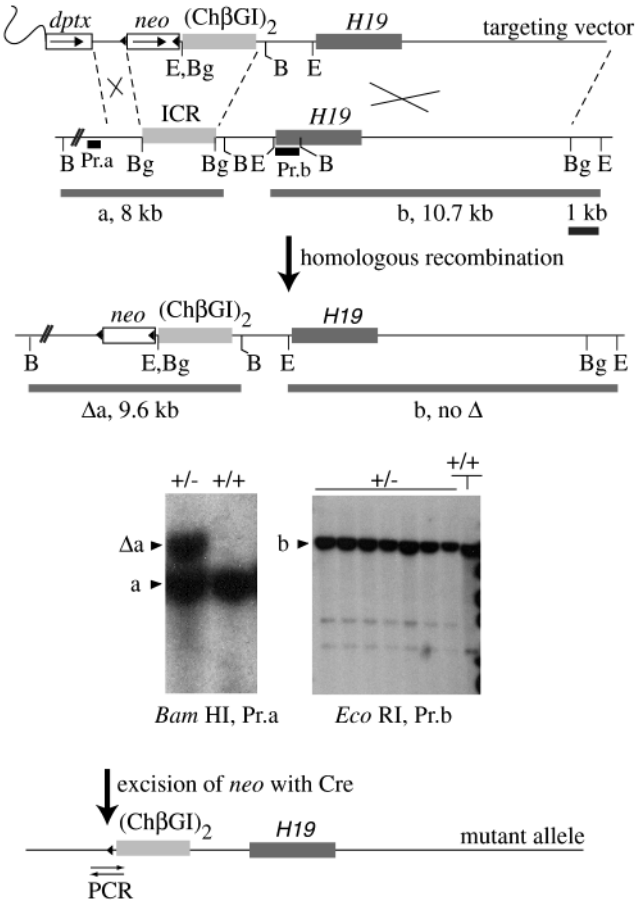

Fig. 2. Targeting at the *Igf2/H19* ICR. The ICR was replaced by the $(ChβGI)_2$, keeping the CTCF sites in the same orientation. The sequences involved are described in Fig. 1B. *Bam*HI and probe a (Pr.a) blot, one clone shown. All seven clones obtained [*Eco*RI and probe b (Pr.b) blot] underwent conservative recombination. The *neo* selection cassette was flanked by *lox*P sites allowing for excision with Cre.

MATERIALS AND METHODS

ICR targeting vector

A standard replacement vector for targeting in embryonic stem (ES) cells was designed to enable substitution of the ICR. The ICR as defined here is delineated by two *Bgl*II sites at –2.0 kb and –4.4 kb RTSS of *H19*. The 3′ arm was a 11.6 kb *Bgl*II genomic fragment from a lambda library of strain 129/SvJ (Stratagene, La Jolla, CA) extending from –2.0 kb to +9.6 kb RTSS of *H19*. The 5′ arm was a 1.1 kb fragment extending from –4.4 kb to –5.5 kb RTSS of *H19* and was made by PCR using primers 5′-AAAAATCGATCCTTCC-TGGGTAATACCTCTGAGCCTG-3′ (upper) and 5′-AAAAATCGA-TTCTTCTTCCAGAAACAAGTTAGGCATG-3′ (lower) (GenBank Accession Number, AF049091). The *Cla*I site near the end of both primers allowed for subcloning. At the end of the 5′ arm the RNAPolII promoter-diphtheria toxin, a chain-SV40 polyA (*dptx*) negative selection cassette (Lexicon Genetics, The Woodlands, TX), was cloned; between the two arms, the *Pgk1* promoter-neomycin-bovine polyA (*neo*)-positive selection cassette flanked by *lox*P sites (floxed) was cloned for later excision using Cre recombinase. This base-targeting vector contained a unique *Bgl*II site just 3′ of the *neo* cassette and between the two arms of *H19* sequence – this represented the site of deletion of the ICR. At this *Bgl*II site the 2.4 kb $(ChβGI)_2$, as excised from plasmid pJC13-1 (Chung et al., 1993), was inserted using *Eco*RI and *Bam*HI. Upon homologous recombination, followed by excision of the floxed *neo* cassette with Cre, the ICR was substituted for the $(ChβGI)_2$, together with a single *lox*P site flanked by a small amount of polylinker sequence directly adjacent to the 5′ end of the $(ChβGI)_2$ (Fig. 2).

Production of transgenic mouse lines with the ICR substitution

The vector was linearized at the end of the long arm using *Not*I, then introduced into W9.5 ES cells of strain 129S1/SvImJ (Szabó and Mann, 1994) by electroporation. Probe a, 0.5 kb, was made by PCR using primers 5′-GGTGCCATCAAGCTACTACAC-3′ (upper) and 5′-CTGGATAGGACATGGGCACAG-3′ (lower) (sequences from GenBank Accession Number, AC013548). Probe b was an *Eco*RI-*Bam*HI fragment as indicated (Fig. 2). Seven clones were obtained from 144 screened. Excision of the *neo* cassette was achieved by mating male chimeras with Cre-deleter females. The latter transgenic line, a mixture of 129S1 and Swiss in this study, was made by the insertion of a *cre* expression cassette into the X-linked *Hprt* gene by gene targeting. When males carrying a floxed sequence in a reporter transgenic line (Soriano, 1999) are mated to hemizygous *cre*/0 females, excision of the sequence occurs regardless of *cre* inheritance and without mosaicism (J. M., unpublished). Positive mice were identified by PCR: a 0.24 kb product spanning the one remaining *lox*P site after *neo* cassette excision was amplified using primers 5′-GCCCACCAGCTGCTAGCCATC-3′ (upper) and 5′-CCTAGAGAA-TTCGAGGGACCTAATAAC-3′ (lower). Two clones were used to make chimeras, both giving germline transmission. Mice derived from both clones had the same phenotype, and mice derived from one clone were used in this study. Full designation of the targeted mutation described here is *H19.ICR*$^{tm1(ChβGI)MdCoh}$.

Breeding of fetuses carrying the $(ChβGI)_2$ for analysis

To produce all fetuses analyzed, one set of parents were females and males carrying the $(ChβGI)_2$ and were born from chimeras mated to *cre*/0 females. These mice therefore lacked the *neo* cassette and were heterozygous for the mutation. Females were not typed for the presence of X-linked *cre*, while males were typed as negative for *cre*. The other set of parents, unless stated otherwise, were homozygous for the *Mus musculus castaneus* form of distal chromosome 7, as derived from strain CAST/Ei (CS). More specifically, these mice were of strain FVB/NJ.CS(N3)-*distalChr.7*$^{CS/CS}$. The use of this cross allowed for allele-specific analysis of expression and methylation. Hereafter, heterozygous fetuses maternally and paternally inheriting the $(ChβGI)_2$ are designated –(M)/+ and +/–(P), respectively.

Gene expression

Northern blots were performed as described (McLaughlin et al., 1996; Szabó and Mann, 1994). Allele-specific expression was determined using RT-PCR single nucleotide primer extension (SNuPE) assays as described (Szabó and Mann, 1995), except that a different *H19* SNuPE primer was used (5′-TGAATGTATACAGCGAGTGTG-3′), incorporating the radionucleotide C for the B6-type allele and T for the CS-type allele. Each assay relies on a single known sequence difference between allelic RNAs, as provided by CS mice. Assays were conducted on the same samples used in northern blots.

Methylation in germ cells

Southern blots for methylation analysis were performed as described (Szabó and Mann, 1994). Female mice of the transgenic line TgOG2 were mated to +/–(P) males, and from the resulting 18.5 days post coitum (dpc) female and male fetuses of small size – all presumptive +/–(P) fetuses – the gonads were removed, trypsinized and triturated, and cells positive for enhanced green fluorescent protein ($EGFP^+$) were sorted using a MoFlo flow cytometer (Cytomation, Fort Collins, CO). TgOG2 mice, of mixed C57BL/6J and CBA/CaJ strain

background are homozygous for a transgene in which the *egfp* reporter (Clontech, Palo Alto, CA) is driven by the *Pou5f1* promoter. This line expresses EGFP specifically in the germline (J. M., unpublished). Two other transgenic lines made with the same construct have been reported (Anderson et al., 1999; Yoshimizu et al., 1999). Bisulfite sequencing was performed as described (Reed et al., 2001), except that bisulfite conversion was carried out at 60°C for 6 hours. In each PCR reaction, the DNA from 2000 cells was amplified using primers 5′-CAACCCCCCCCCAAAACCCCCAAAAATA-3′ (upper) and 5′-GGGTTGTTGGTTGATGATTTTGTATATAGT-3′ (lower). Two PCRs were performed on DNA from each of two conversion reactions, the products subcloned and individual clones sequenced.

RESULTS

Nucleotide structure of the ICR and (ChβGI)$_2$

The nucleotide structure of the 2.4 kb ICR in respect to CTCF binding sites, CpGs and GpCs is compared with the 2.4 kb (ChβGI)$_2$ (Fig. 1B). For the latter, a high density of CpGs is present in the vicinity of the CTCF-binding sites (Chung et al., 1997); therefore, there is potential for methylation of these regions. Comparison of the two 2.4 kb sequences gave 38% identity using 'Gap' global alignment (Seqweb Version 1.2, GCG Wisconsin Package version 10.1; Accelrys, San Diego, CA). They therefore lack homology – comparison of the (ChβGI)$_2$ with three other 2.4 kb regions selected at random from a 30 kb region downstream of *H19* (sequences from GenBank Accession Number, AF049091) gave identities of 37-40%.

Maternal inheritance of the (ChβGI)$_2$

Substituting the (ChβGI)$_2$ for the ICR had little or no effect on the normal allele specific expression pattern of *Igf2* and *H19*. Both genes were expressed monoallelically in the liver and kidney of 18.5 dpc –(M)/+ fetuses. Indeed, the low amount of paternal *H19* RNA in the liver of +/+ fetuses was not present in –(M)/+ fetuses (Fig. 3A). Thus, the maternal (ChβGI)$_2$ successfully substituted for the function of the maternal ICR as a chromatin insulator. Insulator activity is dependent on the lack of methylation. Therefore, as expected, the maternal (ChβGI)$_2$ was hypomethylated in somatic tissues of perinatal –(M)/+ mice (see Fig. 4A). In addition, CpGs within and adjacent the CTCF binding sites of the (ChβGI)$_2$ in germ cells of 18.5 dpc female +/–(P) fetuses were hypomethylated (Fig. 4B).

Northern blots showed that the total amount of *Igf2* RNA in the liver and kidney of –(M)/+ fetuses was ~1.3 times that in +/+ fetuses (Fig. 3C). This is suggestive of some degree of hyperactivation of the paternal *Igf2* allele in *trans* to the (ChβGI)$_2$, as RT-PCR SNuPE assays conducted on the same samples showed that all *Igf2* RNA was derived from this allele (Fig. 3A). In any event, any increase in total *Igf2* RNA did not result in an increase in weight of –(M)/+ fetuses (Table 1).

Paternal inheritance of the (ChβGI)$_2$

18.5 dpc +/–(P) fetuses were 50% or 61% of the weight of +/+

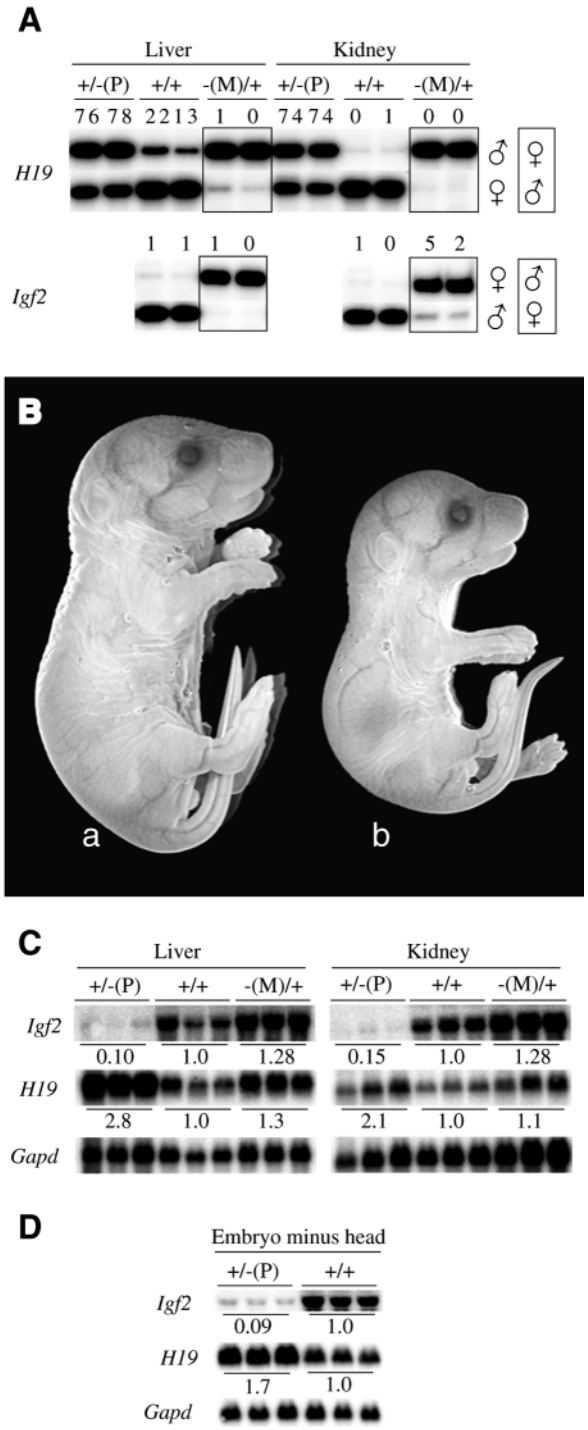

Fig. 3. Phenotype of 18.5 dpc fetuses inheriting the (ChβGI)$_2$. (A) RT-PCR SNuPE assays. +/–(P) and +/+ lanes; top row is presumptive inactive allele. –(M)/+ lanes (in boxes) (these fetuses were obtained from the reciprocal mating, hence bottom row is presumptive inactive allele). Parental origin of alleles is on the right. Value above each band is the amount of RNA contributed by the presumptive inactive allele as a percent of the total. (B) (a) +/+ fetus; (b) +/–(P) fetus. (C) Northern blots. Values under bands are the mean relative amounts of RNA standardized according to *Gapd* mRNA. All four *Igf2* transcripts were reduced to the same extent in +/–(P) mice (only the largest transcript is shown). Number of embryos used: liver, one in all lanes; kidney, two in +/+ and –(M)/+ lanes, and four in +/–(P) lanes. (D) Northern blots, 11.5 dpc midgestation embryos. One embryo was used for each lane. Other details as in C.

Table 1. Weight of 18.5 dpc fetuses on maternal and paternal inheritance of the $(ChbGI)_2$

	Mean wet weight (g)±s.d. (*n*) (range) (% of +/+ weight)	
	Fetus	Placenta
–(M)/+*	1.42±0.01 (14) (1.20-1.60) (103%)	ND
+/+*	1.38±0.09 (12) (1.29-1.55)	ND
+/–(P)†	0.75±0.07 (21)¶ (0.62-0.86) (50%)	50±7 (10)¶ (45-64) (60%)
+/+†	1.49±0.09 (11) (1.35-1.66)	83±11 (10) (73-107)
+/–(P)‡	0.70±0.02 (14)¶ (0.66-0.73) (61%)	ND
+/+‡	1.14±0.11 (7) (0.92-1.24)	ND
+/*Igf2*^{null}(P)§	0.94 (62%)	ND
+/+§	1.50	ND

*Siblings from +/–(P) ♀ × +/+ ♂ matings.
†Siblings from +/+ ♀ × +/–(P) ♂ matings.
‡Siblings from +/+ ♀ × +/–(P) ♂ matings. +/+ females were from transgenic line TgOG2.
§Data from Baker et al. (1993). Matings were of outbred mice.
¶P<0.01, versus +/+ using Student's t-test.
ND, not done.
(M), Maternal allele; (P), Paternal allele

siblings, depending on the mouse strain used, and were of normal proportions and viable (Fig. 3B; Table 1). In addition, small size persisted into adulthood (data not shown). This phenotype was similar to that seen for paternal inheritance of a targeted null mutation of *Igf2* – perinatal mutants was 62% of the weight of wild-type siblings, of normal proportions and viable (Baker et al., 1993). This effect indicated that the paternal $(ChβGI)_2$ was inhibiting activity of the paternal *Igf2* promoter. As expected, the amount of *Igf2* mRNA in perinatal and midgestation +/–(P) embryos was found to be very low (Fig. 3C,D). Thus, the paternally inherited $(ChβGI)_2$, rather than reiterating the behavior of the paternal ICR, functioned as a chromatin insulator, as does the maternal ICR. Consistent with insulator activity of the paternal $(ChβGI)_2$ was its hypomethylated state in somatic tissues of +/–(P) mice (Fig. 4A). In addition, the $(ChβGI)_2$ was hypomethylated in male germ cells of perinatal +/–(P) mice – only 11% of CpGs analyzed were methylated (Fig. 4B). In comparison, the paternal ICR is hypermethylated in male germ cells of perinatal mice (Davis et al., 1999; Ueda et al., 2000).

In the liver and kidney of +/–(P) fetuses, *H19* was biallelically expressed (Fig. 3A), demonstrating that the paternal $(ChβGI)_2$ was unable to induce post-fertilization inactivation of the *H19* promoter in *cis* – as does the paternal ICR. Interestingly, in both tissues, the normally silent paternal *H19* allele had been activated to a level greater than that of the maternal allele, accounting for ~75% of the total H19 RNA (Fig. 3A). This may represent a level of hyperactivation of the paternal *H19* promoter at the perinatal stage: total *H19* RNA measured in liver and kidney was 2.8 and 2.1 times the normal level, respectively (Fig. 3C); thus, paternal *H19* RNA was 2.1 and 1.6 times the normal maternal level, respectively. However it was noted that in 11.5 dpc +/–(P) embryos, total *H19* RNA was only 1.7 times that of normal levels (Fig. 3D).

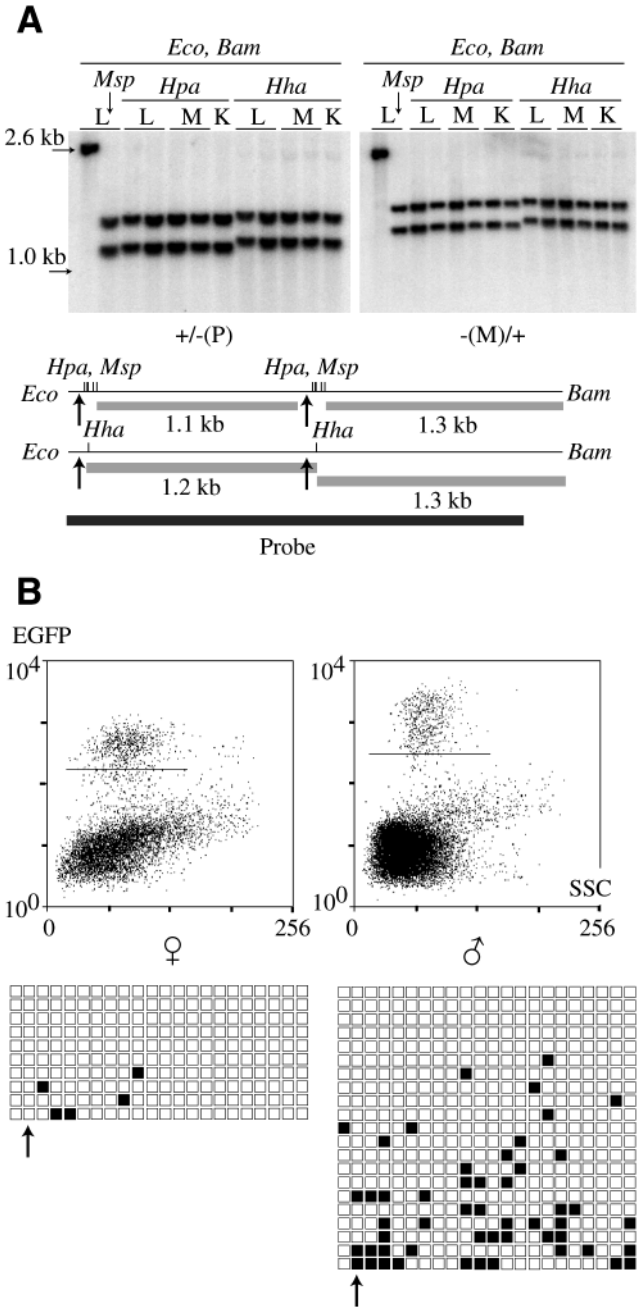

Fig. 4. $(ChβGI)_2$ methylation. (A) Southern blots of tissues of 18.5 dpc fetuses. L, liver; M, carcass (mostly muscle); K, kidney. This assay tests for methylation at the central methylation-sensitive *Hpa* (*Hpa*II) (CCGG) and *Hha* (*Hha*I) (GCGC) sites indicated below the blots (if methylated, the 2.6 kb *Eco* (*Eco*RI) and *Bam* (*Bam*HI) band should remain intact). *Msp* (*Msp*I) cuts at CCGG, regardless of CpG methylation. The CTCF sites are indicated (vertical arrows). The 2.6 kb *Eco*, *Bam* fragment consists of the 2.4 kb $(ChβGI)_2$ plus an additional 0.2 kb of downstream sequence (see Fig. 2). Major bands expected on complete cutting are shown below blots. The complete ChβGI sequence was the probe used. Number of embryos used: +/–(P), one embryo per lane for L and M, and three embryos per lane for K; –(M)/+, one embryo per lane for L and M, and two embryos per lane for K. (B) Purification and methylation state of germ cells of 18.5 dpc +/–(P) fetuses. In the flow cytometry graph, EGFP$^+$ cells above the line were sorted for analysis. Each row of squares represents the methylation state of each of 22 sequential CpGs, 5′→3′ orientation as in Fig. 1B (white square, unmethylated; black square, methylated). Each row represents a separate sequencing reaction. The CpG of the CTCF site is indicated (arrow).

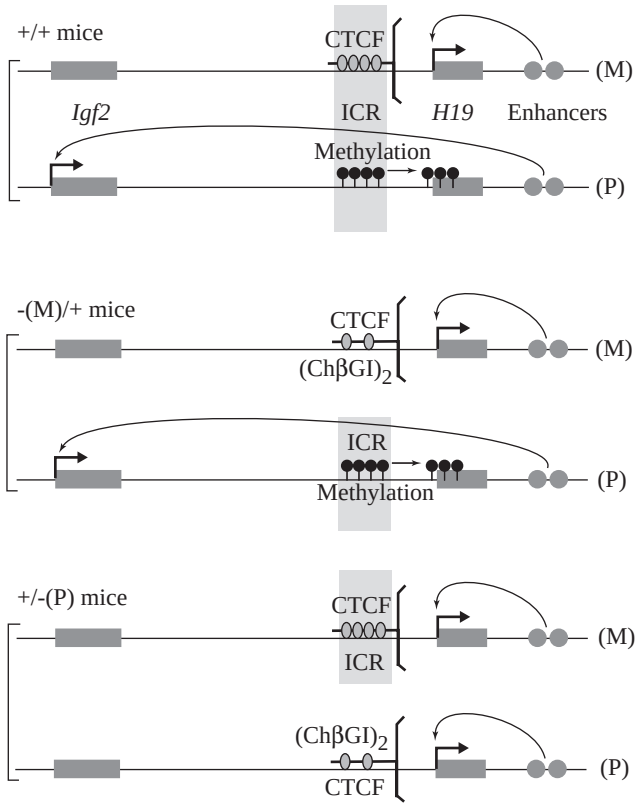

Fig. 5. Summary of results. In –(M)/+ fetuses (middle), the maternal $(ChβGI)_2$ acts as a chromatin insulator, substituting for this property of the maternal ICR in +/+ fetuses. Allele-specific expression of *Igf2* and *H19* was unaffected. In +/–(P) fetuses (bottom) the paternal $(ChβGI)_2$ acted as a chromatin insulator as did the maternal $(ChβGI)_2$. Consequently both *Igf2* alleles were strongly repressed and total *Igf2* RNA was very low. Both *H19* promoters had access to the enhancers in *cis*, and *H19* was expressed biallelically. Other details are as described in the legend to Fig. 1.

DISCUSSION

In this study, we substituted the *Igf2/H19* ICR for the $(ChβGI)_2$ by gene targeting and examined the effects in mice. The former is an element with properties of chromatin insulation, promoter silencing and imprinting (differential methylation), while the latter is a duplication of a chromatin insulator element, which, in its native state, lies at the 5′ end the chicken β-globin locus. The ICR and $(ChβGI)_2$ are of the same size, possess CTCF insulator binding sites and CpGs, but otherwise lack sequence homology. The aim of this study was to determine if the $(ChβGI)_2$ could duplicate one or more of the three properties of the ICR. The results should shed light on the mechanisms behind these properties – in particular, the potential role of CTCF binding (see Fig. 5).

The $(ChβGI)_2$ and chromatin insulation

Much evidence has accumulated that the maternal ICR mediates silence of the *Igf2* allele in *cis* through chromatin insulator activity: deletion of part or all of the maternal ICR leads to expression of the maternal *Igf2* allele, and therefore increased concentration of *Igf2* RNA and large mice (Leighton et al., 1995a; Srivastava et al., 2000; Thorvaldsen et al., 1998). In addition, the maternal ICR binds the vertebrate insulator protein CTCF in vivo (Kanduri et al., 2000b; Szabó et al., 2000), and the ICR displays insulator activity in enhancer blocking assays in cell culture and transgenic mice (Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000a; Srivastava et al., 2000). In the present study, when the ICR was replaced by the $(ChβGI)_2$, the normal state of maternal *Igf2* inactivity was retained. As the ChβGI has previously been characterized as an insulator, this finding lends even further support for the idea that maternal *Igf2* inactivity is brought about through insulator activity in *cis*.

The insulator activity of the maternal $(ChβGI)_2$ was at least as strong as that of the maternal ICR. This was seen in the complete inactivity of maternal *Igf2* in *cis* in the tissues examined, and in the normal size of –(M)/+ fetuses. This suggests that the ICR is not uniquely specialized for chromatin insulation at the *Igf2/H19* domain, and that other insulators, even from a different class of vertebrate (*Aves*), can perform equally well in this role. This result is consistent with a high degree of conservation of function of CTCF-mediated chromatin insulation, in keeping with the high level of conservation in vertebrates of the insulator binding site core and CTCF amino acid sequence (Ohlsson et al., 2001). Furthermore, the present results, in which full insulation was obtained with two ChβGI CTCF-binding sites, are consistent with previous studies that suggest only two CTCF sites are required for effective insulation at the *Igf2/H19* domain: deletion of two of the four CTCF-binding sites (deletion a, Fig. 1C) had no affect on insulator activity (Drewell et al., 2000), while deletion of three (deletion b, Fig. 1C) or all (deletion c, Fig. 1C) of the four CTCF sites eliminated much, if not all, insulator activity (Srivastava et al., 2000; Thorvaldsen et al., 1998). Thus, although the CTCF insulator binding site core of the ChβGI is slightly different from those in the ICR, and differs, at least in vitro, in the binding pattern of zinc fingers relative to the ICR sites (Ohlsson et al., 2001), it is apparently equally effective in mediating insulation.

The paternal ICR is methylated, does not bind CTCF and lacks insulator activity. By contrast, the paternal $(ChβGI)_2$ was hypomethylated and also functioned as a chromatin insulator. This was seen in the substantial reduction of total *Igf2* activity in embryos paternally inheriting the $(ChβGI)_2$. Indeed, *Igf2* expression was so low in +/–(P) 18.5 dpc fetuses that they were as small as *Igf2* null mutants. This result is consistent with the idea that paternally inherited methylation inactivates potential insulator function of the paternal ICR.

The $(ChβGI)_2$ and *H19* promoter silencing

The paternal ICR, while possessing no insulator activity, establishes silence of the *H19* promoter in *cis* during early embryogenesis. The paternal $(ChβGI)_2$ clearly lacked this property. Indeed, the paternal *H19* promoter was even more active than the maternal in the liver and kidney of –(M)/+ 18.5 dpc fetuses, as shown by the analysis of allele-specific expression, and this may represent some degree of hyperactivation of the paternal *H19* promoter. The ChβGI, in its native state, is a center for hyperacetylation, although it is not clear if this property is associated with insulator function or CTCF binding (Litt et al., 2001). If the ChβGI retained such a property in transgenic mice, then it might further potentiate

the activity of nearby promoters in *cis*. Nevertheless, it is noted that the maternal $(Ch\beta GI)_2$ had no detectable effect on *H19* expression: in tissues of –(M)/+ and +/+ fetuses, the amounts of *H19* RNA were the same, and in both cases all *H19* RNA was derived from the maternal allele.

The $(Ch\beta GI)_2$ and acquisition of differential methylation or imprinting

The ICR becomes differentially methylated during female and male germ cell development (acquisition or establishment of the imprint); these epigenetic states are inherited and retained by the maternal and paternal copies during embryonic development (maintenance of the imprint) and further epigenetic change in *cis* is induced (development of the imprint). The $(Ch\beta GI)_2$ was unable to duplicate this function of the ICR. It was hypomethylated in female and male germ cells – at the stage by which differential ICR methylation has taken place – and remained as such on the maternal and paternal chromosomes during embryonic development. Furthermore, the expression patterns of *Igf2* and *H19* on the maternal and paternal mutant chromosomes were similar. Thus, imprinting at the *Igf2/H19* domain was lost.

These results clearly demonstrate that it is possible to separate the two functions of CTCF-mediated chromatin insulation and the establishment of differential methylation in the germline, i.e., these are not inseparable functions at the *Igf2/H19* domain. Therefore, any potential for differential CTCF binding is not the primary difference between the two germlines, which determines their differential acquisition of ICR methylation. This conclusion assumes the ChβGI is not in some unknown way resistant to methylation. Furthermore, this conclusion does not preclude the possibility that differential CTCF binding may be the maintenance mechanism for differential methylation of the ICR in somatic cells, i.e. binding to the maternal ICR may be required for excluding methyltransferases from the region. It has been noted that only 42 bp of the ChβGI, which includes the CTCF site, retains insulator activity in enhancer blocking assays (Bell et al., 1999). Thus, as CTCF binding is inhibited by methylation, this 42 bp region must remain hypomethylated over the course of the assay. This result, together with the observation that the ChβGI possesses a very high CpG density and C+G content around the CTCF site, as seen in CpG islands, raises the possibility that CTCF binding itself may inhibit methylation of the surrounding region. CpG islands are coincident with promoters and their hypomethylated state might result from the exclusion of methyltransferases by bound transcription factors (Brandeis et al., 1994; Macleod et al., 1994).

What other conclusions can be drawn from these results regarding the mechanism of acquisition of differential ICR methylation? As previously pointed out by Jones and Takai (Jones and Takai, 2001), 'two not-necessarily-conflicting models could account for the varying patterns (of de novo methylation), (i) an exclusion of access to methylation sites by proteins bound to specific DNA or (ii) a methylation-targeting mechanism steered by sequence-specific binding proteins'. Accordingly, differential methylation of the ICR could be set up by a female germ cell-specific protein that binds to the ICR and excludes methyltransferases, or alternatively by a male germ cell-specific protein that binds to the ICR and attracts methyltransferases. In addition, either event could be downstream of CTCF binding – in which case, mutagenesis of the binding sites would be expected to affect the process. Unfortunately, the results of the present ICR substitution do not shed light on either of these two possibilities: the lack of methylation of the $(Ch\beta GI)_2$ in both germlines and in somatic cells could be accounted for by the possibility that its sequence is not attractive to methyltransferases or cannot bind proteins that can attract methyltransferases.

A third possible mechanism for the establishment of differential ICR methylation is that there a methyltransferase specific to the male germline exists that is attracted to the ICR sequence. A role for the recently discovered DNA methyltransferases IIIA and IIIB in imprinting has been postulated (Okano et al., 1999), and an investigation of their expression in female and male germ cell development would be of interest. Again, such a mechanism could be downstream of CTCF binding. It is of interest that in ICR deletion experiments, when two CTCF sites were left intact, differential methylation was still acquired (Drewell et al., 2000), yet when one was left, differential methylation in somatic cells was not observed (Thorvaldsen et al., 1998). Thus, it is conceivable that grouped CTCF-binding events could invoke a chromatin structure that is attractive to a male germ cell-specific methyltransferase. This possibility is consistent with the arrangement of the CTCF binding sites into two pairs in mouse and rat, and into two triplets in humans, where spacing between any two sites in a pair or triplet is no greater than 0.45 kb (Stadnick et al., 1999). In the present study, the required chromatin structure may not have been initiated in the $(Ch\beta GI)_2$ because the two CTCF sites were too far apart, or because of sequence.

Further mutagenesis of the ICR in mice should lead to an identification of the sequences that mediate the acquisition of differential germline methylation and shed light on how this is maintained through somatic cell division.

We thank Hans Schöler for the *Pou5f1-egfp* construct, Adam Bell and Gary Felsenfeld for plasmid pJC13-1, and Lucy Brown, Claudio Spalla and Jim Bolen for flow cytometry. This work was supported by NIH grant 2R01GM48103-04 and NSF grant BIR-9220534.

REFERENCES

Anderson, R., Fassler, R., Georges-Labouesse, E., Hynes, R. O., Bader, B. L., Kreidberg, J. A., Schaible, K., Heasman, J. and Wylie, C. (1999). Mouse primordial germ cells lacking beta1 integrins enter the germline but fail to migrate normally to the gonads. *Development* **126**, 1655-1664.

Baker, J., Liu, J. P., Robertson, E. J. and Efstratiadis, A. (1993). Role of insulin-like growth factors in embryonic and postnatal growth. *Cell* **75**, 73-82.

Bartolomei, M. S. and Tilghman, S. M. (1997). Genomic imprinting in mammals. *Annu. Rev. Genet.* **31**, 493-525.

Bartolomei, M. S., Webber, A. L., Brunkow, M. E. and Tilghman, S. M. (1993). Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. *Genes Dev.* **7**, 1663-1673.

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**, 387-396.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A. and Cedar, H. (1994). Sp1 elements protect a CpG island from *de novo* methylation. *Nature* **371**, 435-438.

Brenton, J. D., Drewell, R. A., Viville, S., Hilton, K. J., Barton, S. C., Ainscough, J. F. and Surani, M. A. (1999). A silencer element identified in *Drosophila* is required for imprinting of *H19* reporter transgenes in mice. *Proc. Natl. Acad. Sci. USA* **96**, 9242-9247.

Caspary, T., Cleary, M. A., Baker, C. C., Guan, X. J. and Tilghman, S. M. (1998). Multiple mechanisms regulate imprinting of the mouse distal chromosome 7 gene cluster. *Mol. Cell. Biol.* **18**, 3466-3474.

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

Chung, J. H., Bell, A. C. and Felsenfeld, G. (1997). Characterization of the chicken β-globin insulator. *Proc. Natl. Acad. Sci. USA* **94**, 575-580.

Davis, T. L., Trasler, J. M., Moss, S. B., Yang, G. J. and Bartolomei, M. S. (1999). Acquisition of the *H19* methylation imprint occurs differentially on the parental alleles during spermatogenesis. *Genomics* **58**, 18-28.

Drewell, R. A., Brenton, J. D., Ainscough, J. F., Barton, S. C., Hilton, K. J., Arney, K. L., Dandolo, L. and Surani, M. A. (2000). Deletion of a silencer element disrupts *H19* imprinting independently of a DNA methylation epigenetic switch. *Development* **127**, 3419-3428.

Elson, D. A. and Bartolomei, M. S. (1997). A 5′ differentially methylated sequence and the 3′-flanking region are necessary for *H19* transgene imprinting. *Mol. Cell. Biol.* **17**, 309-317.

Ferguson-Smith, A. C. and Surani, M. A. (2001). Imprinting and the epigenetic asymmetry between parental genomes. *Science* **293**, 1086-1089.

Ferguson-Smith, A. C., Sasaki, H., Cattanach, B. M. and Surani, M. A. (1993). Parental-origin-specific epigenetic modification of the mouse *H19* gene. *Nature* **362**, 751-755.

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.

Holmgren, C., Kanduri, C., Dell, G., Ward, A., Mukhopadhya, R., Kanduri, M., Lobanenkov, V. and Ohlsson, R. (2001). CpG methylation regulates the *Igf2/H19* insulator. *Curr. Biol.* **11**, 1128-1130.

Howell, C. Y., Bestor, T. H., Ding, F., Latham, K. E., Mertineit, C., Trasler, J. M. and Chaillet, J. R. (2001). Genomic imprinting disrupted by a maternal effect mutation in the *Dnmt1* gene. *Cell* **104**, 829-838.

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.

Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science* **293**, 1074-1080.

Jones, P. A. and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068-1070.

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.

Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. and Tilghman, S. M. (1995a). Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature* **375**, 34-39.

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

Li, E., Beard, C. and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* **366**, 362-365.

Litt, M. D., Simpson, M., Recillas-Targa, F., Prioleau, M. N. and Felsenfeld, G. (2001). Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J.* **20**, 2224-2235.

Macleod, D., Charlton, J., Mullins, J. and Bird, A. P. (1994). Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev.* **8**, 2282-2292.

Mann, J. R., Szabo, P. E., Reed, M. R. and Singer-Sam, J. (2000). Methylated DNA sequences in genomic imprinting. *Crit. Rev. Eukaryot. Gene Expr.* **10**, 241-257.

McLaughlin, K. J., Szabo, P., Haegel, H. and Mann, J. R. (1996). Mouse embryos with paternal duplication of an imprinted chromosome 7 region die at midgestation and lack placental spongiotrophoblast. *Development* **122**, 265-270.

Nabetani, A., Hatada, I., Morisaki, H., Oshimura, M. and Mukai, T. (1997). Mouse *U2af1-rs1* is a neomorphic imprinted gene. *Mol. Cell. Biol.* **17**, 789-798.

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.

Okano, M., Bell, D. W., Haber, D. A. and Li, E. (1999). DNA methyltransferases *Dnmt3a* and *Dnmt3b* are essential for *de novo* methylation and mammalian development. *Cell* **99**, 247-257.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W. and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475-478.

Reed, M. R., Huang, C. F., Riggs, A. D. and Mann, J. R. (2001). A complex duplication created by gene targeting at the imprinted *H19* locus results in two classes of methylation and correlated *Igf2* expression phenotypes. *Genomics* **74**, 186-196.

Ripoche, M. A., Kress, C., Poirier, F. and Dandolo, L. (1997). Deletion of the *H19* transcription unit reveals the existence of a putative imprinting control element. *Genes Dev.* **11**, 1596-1604.

Schmidt, J. V., Levorse, J. M. and Tilghman, S. M. (1999). Enhancer competition between *H19* and *Igf2* does not mediate their imprinting. *Proc. Natl Acad. Sci. USA* **96**, 9733-9738.

Shibata, H., Ueda, T., Kamiya, M., Yoshiki, A., Kusakabe, M., Plass, C., Held, W. A., Sunahara, S., Katsuki, M., Muramatsu, M. et al. (1997). An oocyte-specific methylation imprint center in the mouse *U2afbp-rs/U2af1-rs1* gene marks the establishment of allele-specific methylation during preimplantation development. *Genomics* **44**, 171-178.

Shibata, H., Yoda, Y., Kato, R., Ueda, T., Kamiya, M., Hiraiwa, N., Yoshiki, A., Plass, C., Pearsall, R. S., Held, W. A. et al. (1998). A methylation imprint mark in the mouse imprinted gene *Grf1/Cdc25Mm* locus shares a common feature with the *U2afbp-rs* gene: an association with a short tandem repeat and a hypermethylated region. *Genomics* **49**, 30-37.

Soriano, P. (1999). Generalized *lacZ* expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70-71.

Srivastava, M., Hsieh, S., Grinberg, A., Williams-Simons, L., Huang, S. P. and Pfeifer, K. (2000). *H19* and *Igf2* monoallelic expression is regulated in two distinct ways by a shared *cis* acting regulatory region upstream of *H19*. *Genes Dev.* **14**, 1186-1195.

Stadnick, M. P., Pieracci, F. M., Cranston, M. J., Taksel, E., Thorvaldsen, J. L. and Bartolomei, M. S. (1999). Role of a 461-bp G-rich repetitive element in *H19* transgene imprinting. *Dev. Genes Evol.* **209**, 239-248.

Szabó, P. and Mann, J. R. (1994). Expression and methylation of imprinted genes during in vitro differentiation of mouse parthenogenetic and androgenetic embryonic stem cell lines. *Development* **120**, 1651-1660.

Szabó, P. E. and Mann, J. R. (1995). Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment, and mechanisms of genomic imprinting. *Genes Dev.* **9**, 1857-1868.

Szabó, P. E., Tang, S.-H. E., Rentsendorj, A., Pfeifer, G. and Mann, J. R. (2000). Maternal-specific footprints at putative CTCF sites in the *H19* imprinting control region give evidence for insulator function. *Curr. Biol.* **10**, 607-610.

Thorvaldsen, J. L., Duran, K. L. and Bartolomei, M. S. (1998). Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. *Genes Dev.* **12**, 3693-3702.

Tremblay, K. D., Saam, J. R., Ingram, R. S., Tilghman, S. M. and Bartolomei, M. S. (1995). A paternal-specific methylation imprint marks the alleles of the mouse *H19* gene. *Nat. Genet.* **9**, 407-413.

Tremblay, K. D., Duran, K. L. and Bartolomei, M. S. (1997). A 5′ 2-kilobase-pair region of the imprinted mouse *H19* gene exhibits exclusive paternal methylation throughout development. *Mol. Cell. Biol.* **17**, 4322-4329.

Tucker, K. L., Beard, C., Dausmann, J., Jackson-Grusby, L., Laird, P. W., Lei, H., Li, E. and Jaenisch, R. (1996). Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev.* **10**, 1008-1020.

Ueda, T., Abe, K., Miura, A., Yuzuriha, M., Zubair, M., Noguchi, M., Niwa, K., Kawase, Y., Kono, T., Matsuda, Y. et al. (2000). The paternal methylation imprint of the mouse *H19* locus is acquired in the gonocyte stage during foetal testis development. *Genes Cells* **5**, 649-659.

Yoshimizu, T., Sugiyama, N., De Felice, M., Yeom, Y. I., Ohbo, K., Masuko, K., Obinata, M., Abe, K., Scholer, H. R. and Matsui, Y. (1999). Germline-specific expression of the *Oct-4*/green fluorescent protein (GFP) transgene in mice. *Dev. Growth Differ.* **41**, 675-684.

Zemel, S., Bartolomei, M. S. and Tilghman, S. M. (1992). Physical linkage of two mammalian imprinted genes, *H19* and insulin-like growth factor 2. *Nat. Genet.* **2**, 61-65.