Development 129, 1205-1213 (2002)
Printed in Great Britain © The Company of Biologists Limited 2002
DEV5961

Novel conserved elements upstream of the *H19* gene are transcribed and act as mesodermal enhancers

Robert A. Drewell*, Katharine L. Arney*, Takahiro Arima†, Sheila C. Barton, James D. Brenton‡ and M. Azim Surani§

Wellcome/CRC Institute of Cancer and Developmental Biology and Physiological Laboratory, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK

*These authors contributed equally to this work

†Present address: Department of Reproductive Physiology and Endocrinology, Medical Institute of Bioregulation, Kyushu University, 4546, Tsurumihara, Beppu, Oita 874-0838, Japan

‡Present address: Cancer Genomics Program, Department of Oncology, University of Cambridge, Hutchison/MRC Research Centre, Hills Road, Cambridge CB2 2XZ, UK

§Author for correspondence (e-mail: as10021@mole.bio.cam.ac.uk)

Accepted 3 December 2001

SUMMARY

The reciprocally imprinted *H19* and *Igf2* genes form a co-ordinately regulated 130 kb unit in the mouse controlled by widely dispersed enhancers, epigenetically modified silencers and an imprinting control region (ICR). Comparative human and mouse genomic sequencing between *H19* and *Igf2* revealed two novel regions of strong homology upstream of the ICR termed *H19* upstream conserved regions (HUCs). Mouse HUC1 and HUC2 act as potent enhancers capable of driving expression of an *H19* reporter gene in a range of mesodermal tissues. Intriguingly, the HUC sequences are also transcribed bi-allelically in mouse and human, but their expression pattern in neural and endodermal tissues in day 13.5 embryos is distinct from their enhancer function. The location of the HUC mesodermal enhancers upstream of the ICR and *H19*, and their capacity for interaction with both *H19* and *Igf2* requires critical re-evaluation of the *cis*-regulation of imprinted gene expression of *H19* and *Igf2* in a range of mesodermal tissues. We propose that these novel sequences interact with the ICR at *H19* and the epigenetically regulated silencer at differentially methylated region 1 (DMR1) of *Igf2*.

Key words: Enhancer, Genomic imprinting, *H19*, HUC, *Igf2*, Mouse

INTRODUCTION

The imprinted *H19-Igf2* locus is subject to complex regulation involving a differentially methylated *cis*-acting silencer, an imprinting control region (ICR), comprising both boundary/insulator and silencer functions, enhancers downstream of *H19*, and other transcriptional units of unknown function (Fig. 1A) (Arney et al., 2001; Onyango et al., 2000; Surani, 1998). The *cis*-elements mediate expression of *H19* as well as regulating access to shared enhancers by the paternal *Igf2* and maternal *H19* alleles (Leighton et al., 1995). Indeed, the *H19* transcriptional unit is dispensable for correct imprinting at the locus (Ripoche et al., 1997). The differentially methylated ICR proximate to *H19* and silencers both upstream of *Igf2* and within the intergenic region are the known key regulatory regions at this locus (Ainscough et al., 2000a; Ainscough et al., 2000b; Constancia et al., 2000; Drewell et al., 2000; Srivastava et al., 2000; Thorvaldsen et al., 1998). The ICR upstream of *H19* itself is a complex multipartite regulatory region. On the unmethylated maternal chromosome, the region acts as a boundary/insulator, recruiting the CTCF protein to block access to *Igf2* by enhancers located downstream of *H19* and thus allowing *H19* expression (Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000). On the methylated paternal allele, a silencer element present in the ICR acts to silence *H19* in specific tissues (Brenton et al., 1999; Drewell et al., 2000). Methylation of the ICR prevents CTCF binding and formation of the boundary, allowing the downstream enhancers free access to the *Igf2* promoter.

Unlike the *H19* gene, which is differentially methylated at the promoter, both parental alleles of *Igf2* are unmethylated at the promoter and potentially active (Sasaki et al., 1992). To combat this, both the boundary element upstream of H19 and a silencer element located at differentially methylated region 1 (DMR1) upstream of *Igf2* acts to prevent transcription of *Igf2* from the maternal allele in specific tissues (Constancia et al., 2000). How precisely the complex regulation of the two genes is achieved by distant enhancers and a variety of *cis*-regulatory elements remains to be fully elucidated. In this study, we have identified two novel regulatory elements upstream of the ICR, which act as strong mesodermal enhancers in specific tissues. They are also bi-allelically transcribed. The location of these elements between *H19* and *Igf2* is unexpected and they, in concert with other *cis*-elements

at the locus, may drive expression of *H19* and *Igf2* in a variety of mesodermal tissues.

MATERIALS AND METHODS

RNA isolation and RT-PCR

RNA was isolated from mouse embryos and organs using Trizol™ (Life Technologies) according to the manufacturer's instructions. RT-PCR was performed using the Qiagen One-Step RT-PCR kit, or as previously described (Ainscough et al., 1997). PCR products were electrophoresed, blotted onto Hybond N+ membranes (Amersham Pharmacia Biotech) and hybridised to α^{32}P-dCTP labelled probes, according to standard procedures.

PCR primer sequences were as follows:
mouse HUC1S 5′-ATCCTGCTGGTATCCTGAGG-3′;
mouse HUC1AS 5′-ACTTATGCGTTCAGTCACTTCC-3′;
mouse HUC2S 5′-AAGAGAATGGACAGGACCCAGG-3′;
mouse HUC2AS 5′-CATTCAAAAGGGAACAAGGGC-3′;
human HUC2S 5′-AGGGGAATGGACAGGGCCCAGG-3′; and
human HUC2AS 5′-GATTCAAGAGGGAGCGAGGGC-3′.

To determine allelic expression of HUC 2, RT-PCR products were purified using GeneClean™ (Bio 101) and digested to completion with either *Tru*9I (mHUC2) or *Dde*I (hHUC2).

In situ hybridisation analysis

Sense and antisense riboprobes (relative to the direction of *H19* transcription) were prepared from a 1865 bp clone containing HUC1, HUC2 and the intervening sequence by in vitro transcription using a DIG RNA labelling kit (Boehringer Mannheim). An antisense probe spanning the *H19* transcriptional unit (Drewell et al., 2000) was used as a control. Sagittal sections (8 μm) from 13.5 days post coitum (dpc) mouse embryos were used for in situ hybridisation, essentially as described previously (Wilkinson and Nieto, 1993). Sections were counterstained with Eosin.

Cell transfection assay

Regions were amplified by PCR and cloned into the pGL3-Promoter vector (Promega), upstream of an SV40 promoter driving a firefly luciferase reporter gene, and sequenced to confirm integrity and orientation. PCR primers for HUC1 and HUC2 are described above. Other PCR primers were as follows:

DMDS 5′-TGCCTACAGTTCCCGAATCACC-3′;
DMDAS 5′-CGGCATCGTCTGTCCATTTAGC-3′;
Downstream enhancers S 5′-ATCATTACATCCTGGTGCCTCC-3′; and
Downstream enhancers AS 5′-TAGGCAGTTGGATGATGG-CACC-3′.

The HUC1+2 construct was obtained by using the HUC1S and HUC2AS primers. The 1+2Δ construct was obtained by *Apa*I digestion of the 1+2 construct followed by re-ligation, removing 1.1kb of sequence between HUC1 and HUC2.

DNA constructs were transfected into HeLa cells, cultured for 20 hours, lysed and luciferase readings assayed. Firefly luciferase values were normalised against a co-transfected Renilla luciferase reporter gene driven by a Thymidine Kinase (TK-Renilla) promoter, as described in the DLR assay protocol (Promega). Each construct was tested in triplicate in each experiment, and the experiment was repeated. The *H19* differentially methylated domain was tested both unmethylated (DMD), and after in vitro methylation by *Sss*I (DMD^m). Cells transfected with TK-Renilla alone demonstrated no firefly luciferase activity (negative control).

Mouse transgenic assay

The region encompassing HUC1 and HUC2 was amplified using the following primers:

tgHUCS 5′- CAGGCAGTCAGTCATCTCAGCC-3′; and
tgHUCAS 5′- GCATTCAAAAGGGAACAAGGGC-3′.

*Spe*I sites were incorporated into the primers allowing this fragment to be cloned upstream of the H19 promoter (–816 bp to +5 bp relative to the transcriptional start site) and a *PLAP* reporter gene (see Fig. 5A). This reporter was generated by *Xba*I excision of the *H19-PLAP* transgene described by Brenton et al. (Brenton et al., 1999), followed by cloning into pBluescript. Vector sequences were removed and transgene injection, embryo recovery, subsequent fixation and staining of embryos performed as described in Brenton et al. (Brenton et al., 1999). Embryos were removed at 13.5 days following oviduct transfer and bisected to allow penetration of stain. DNA was prepared from yolk sacs using standard procedures and used for PCR genotyping of the embryos.

PCR primers for genotyping were as follows:
*PLAP*S 5′- TTGGTTGACAGAGTAGGGGC-3′; and
*PLAP*AS 5′- GAGCAAAGATCAGGTCAGCC-3′

RESULTS

Identification of conserved regions upstream of human and mouse *H19*.

To identify additional *cis*-regulatory elements involved in the control of imprinted gene expression of *H19* and *Igf2*, we used genomic DNA sequence upstream of the mouse *H19* gene from both published sequence (AF049091) and from sequencing we performed ourselves. This sequence extended from the transcriptional start site of *H19* to approximately 11.5 kb upstream (GenBank Accession Number, AF327412) (Fig. 1A). Using Nucleotide Identification X (NIX, HGMP Resource Centre, UK) sequence analysis, we identified two novel sequences upstream of *H19*, which we have termed HUC1 and HUC2 (*H19* Upstream Conserved). They exhibit very strong homology to genomic sequence from a human P1 artificial chromosome (PAC) clone (AC004556) (Fig. 1B). The PAC clone contains the equivalent *H19* upstream region and transcription unit.

The conserved sequences revealed similar genomic organisation in the same orientation upstream of the *H19* gene. They were approximately 400 bp each in size and were separated by approximately 1.4 kb of unconserved sequence (Fig. 1B,C). These conserved sequences do not contain any recognisable repetitive elements and have no homology to any other sequences in the human or mouse genome databases. HUC1 and HUC2 therefore represent two novel DNA sequence blocks, which are conserved at the *H19* upstream region in mice and humans. Cross-species Southern hybridisation analysis also revealed the presence of the HUC sequences in other mammals, including rat, dog and muntjac deer (data not shown). The HUCs are relatively poor in CpG content and are therefore unlikely to be a target for regulation by DNA methylation. We failed to detect any significant level of CpG methylation on either parental chromosome when analysed by methylation-sensitive restriction enzyme digestion (data not shown).

HUC sequences are bi-allelically transcribed

As highly conserved sequences are potentially indicative of transcribed regions, we examined whether the HUC regions are transcribed sequences. We initially examined expression from HUC1 and HUC2 using RT-PCR in 13.5 days post coitum

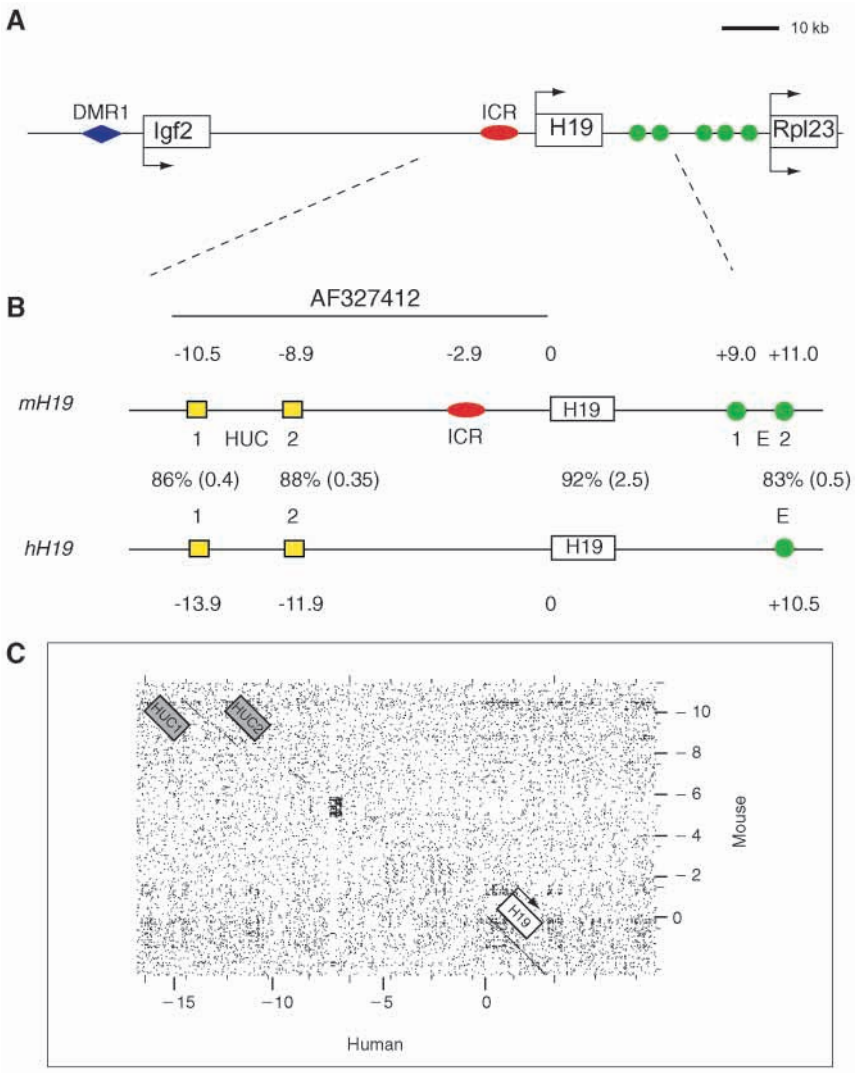

Fig. 1. Genomic organisation and conservation of HUC1 and HUC2 sequences in mouse and human. (A) Mouse *H19*/*Igf2* imprinted locus. *H19* is expressed only from the maternal allele, *Igf2* only from the paternal allele and *Rpl23* is bi-allelically expressed. Characterised enhancers downstream of *H19* (green circles) mediate expression in both endodermal and some mesodermal tissues (Leighton et al., 1995; Brenton et al., 1999; Ishihara et al., 2000). The ICR at *H19* (red ellipse) and the silencer located within DMR1 upstream of *Igf2* (blue diamond) (Constancia et al., 2000) are also shown. (B) The conserved structural organisation of sequences at the *H19* gene. The extended sequenced region upstream of *H19* (AF327412) is indicated. The top line represents mouse DNA sequence, aligned with human sequence on the bottom line (PAC clone pDJ998n23, EM:AC004556) using BlastN software. The numbers above and below the lines relate to the 5′ boundary of various regions relative to the *H19* gene transcriptional start site. Percentages shown in between the two sequences are the degree of homology at the nucleotide level and the approximate length in kb of the conserved region is shown in brackets. The *H19* transcribed region (white box) shows 92% homology between the two species. This degree of conservation is almost matched by HUC1 and HUC2 (yellow boxes), which have a greater level of sequence homology than the characterised downstream enhancer elements (green circles) (Leighton et al., 1995). Conversely, sequence analysis does not reveal a human homologue of the ICR, which comprises both silencer and boundary functions (red ellipse), located upstream of mouse *H19*. (C) Dotplot sequence comparison of human and mouse *H19* upstream region. GCG dotplot software only identified the *H19* transcribed sequence and HUC1 and HUC2 regions, as conserved between mouse and human (default parameters) in the upstream region.

(dpc) whole mouse embryos and detected transcription from these sequences using primers internal to mouse *Huc1* and *Huc2* (Fig. 2A,B). However, we did not detect a message with primers spanning the two conserved regions, indicating they are not part of the same transcript (Fig. 2B). We also attempted RT-PCR with primers from the *Huc* sequences and exon 1 of *H19* and again failed to detect a linked transcript (data not shown). This observation suggests that mouse *Huc1* and *Huc2* represented two separate novel RNAs. We also examined expression by Northern hybridisation analysis in 11.5 dpc and 13.5 dpc whole embryos, where we detect relatively weak low molecular weight bands (data not shown), indicating that the mouse *Huc* transcripts are not part of a larger abundant mRNA.

RNA in situ hybridisation analysis of 13.5 dpc embryos revealed specific expression from the HUC sequences in a number of tissues, including forebrain and midbrain, developing ear, limbs, liver, lungs and the genital eminence (Fig. 2C,D). Comparison with control embryonic sections hybridised to a *H19* probe showed that *Huc* (Fig. 2C,D) and *H19* (Fig. 2E) transcription was neither mutually exclusive, nor shared in all tissues. The direction of the probe used indicated that transcription was orientated in the same direction as the *H19* transcript. The probe in the opposite direction detected no transcripts from the HUC region (Fig. 2F). Detailed analysis of *Huc* expression revealed that the pattern was consistently restricted to the telencephalon and choroid plexus in the forebrain and the cerebellar primordium in the midbrain (Fig. 2G), the cochlea, olfactory epithelium, muscles of the tongue (Fig. 2H) and the genital tubercle (Fig. 2I).

As the HUC regions lie within an imprinted domain, we also tested whether the mouse *Huc* transcripts were generated from only one of the parental chromosomes. Using a restriction polymorphism between 129/Sv and *M. spretus* mouse strains identified in the most highly conserved region, mouse *Huc2*, we performed RT-PCR analysis on RNA isolated from 13.5 dpc whole embryos (Fig. 3A). This showed that mouse *Huc2* was transcribed from both chromosomes (Fig. 3B). Furthermore, we also found that the human HUC2 sequence was transcribed in human placental tissues. A polymorphic site was identified in the parental genomic DNAs of two families which demonstrated that *hHUC2* was also expressed from both alleles (Fig. 3C,D). Therefore, in both mice and humans, *Huc2* is bi-allelically expressed, within a chromosomal domain thought previously to contain only mono-allelically expressed genes.

Our analysis suggests that the *Huc* transcripts are not exons of a larger transcript – indeed no consensus splice acceptor or donor sites can be detected around the HUC sequences using standard analysis software. It is therefore possible that *Huc1* and *Huc2* represent small non-coding RNAs. We could also detect expression by RT-PCR (data not shown), through the

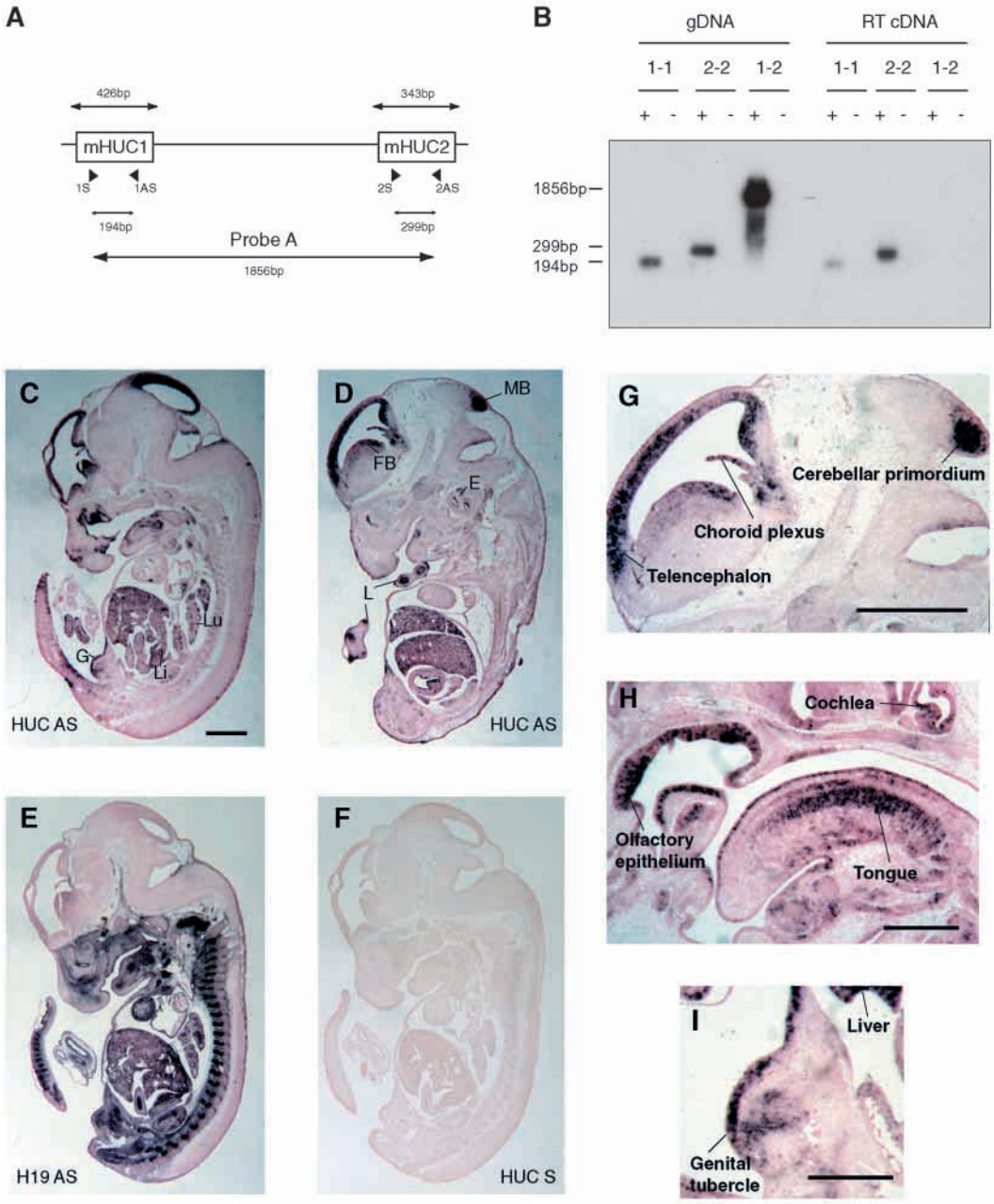

Fig. 2. Expression of HUC1 and HUC2 in the mouse. PCR primers and probes at mouse HUC1 and mouse HUC2. (A) PCR primers internal to both conserved regions were used to amplify regions specific to either mouse HUC1 (194 bp), mouse HUC2 (299 bp) or both regions and the intervening sequence (1856 bp, probe A). The PCR primers are described in the Materials and Methods. (B) RT-PCR detection of RNA transcripts at the HUC sequences. PCR amplifications were performed on either genomic DNA (gDNA) or reverse transcribed total RNA (RT cDNA) from 13.5 dpc embryos and the products hybridised to probe A. Mouse HUC1 and mouse HUC2 produced a product when amplified using primers internal to the conserved region (1-1 or 2-2). However, no product was detected from the RT cDNA sample when amplified across the two conserved regions (1-2), suggesting mouse HUC1 and mouse HUC2 are not part of the same transcription unit. Control lanes are water for gDNA (–) and RT- for RT cDNA (–). (C,D) HUC expression is detected from an antisense (HUC AS) probe spanning HUC1 and HUC2 and is restricted to specific tissues in the embryo. (E,F) *H19* expression from an antisense (H19 AS) probe (E) detects widespread expression in endodermal and mesodermal tissues. No expression was detected from a HUC sense (HUC S) probe (F), indicating that the HUC sequences are only transcribed in the same orientation as *H19*. (G-I) Detailed expression pattern detected from the HUC antisense probe in the brain (G), tongue and ear (H) and genital tubercle (I). E, ear; FB, forebrain; G, genital eminence; L, limb; Li, liver; Lu, lung; MB, midbrain. Scale bars: in C, 1 mm for C-F; 0.5 mm in G-I.

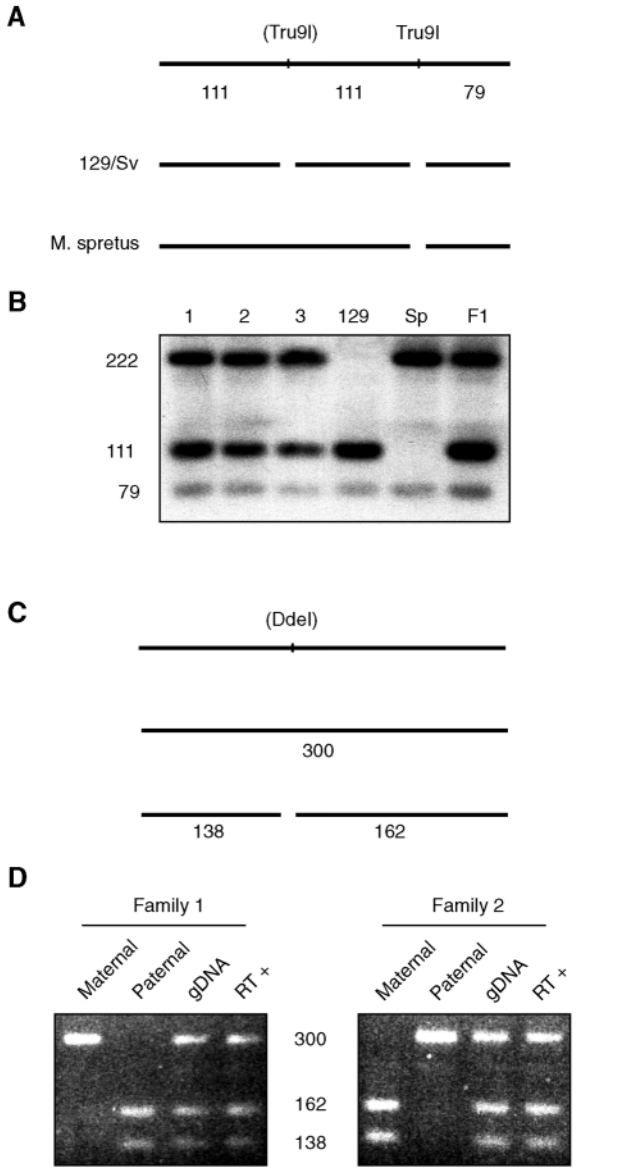

Fig. 3. Bi-allelic expression of HUC2 in mouse and human. (A) *Tru*9I polymorphism in the mouse HUC2 sequence between 129/Sv and *M. spretus* strains was used to analyse allele-specific expression by RT-PCR in 13.5dpc embryos, as described in Fig. 2. (B) In three separate embryos (1-3), transcription from mouse HUC2 was detected from both parental alleles. The pattern was the same as that from genomic DNA of the F1 hybrid (129/Sv × *M. spretus*). (C) *Dde*I polymorphism in the human HUC2 sequence was identified in the genomic DNA of parents in two independent families. (D) RT-PCR analysis was performed on total RNA isolated from post-partum placentas from these families. This demonstrated bi-allelic expression of human HUC2 in placenta (RT+). The pattern was the same as that from genomic DNA of the placental tissue (gDNA).

previously characterised silencer element at *H19* (Drewell et al., 2000), suggesting that there may be other transcripts of unknown function at the *H19/Igf2* locus. Such noncoding transcripts have been characterised at several other imprinted loci (Arima et al., 2000; Moore et al., 1997; Takada et al., 2000; Wutz et al., 1997), although in many cases their function remains enigmatic.

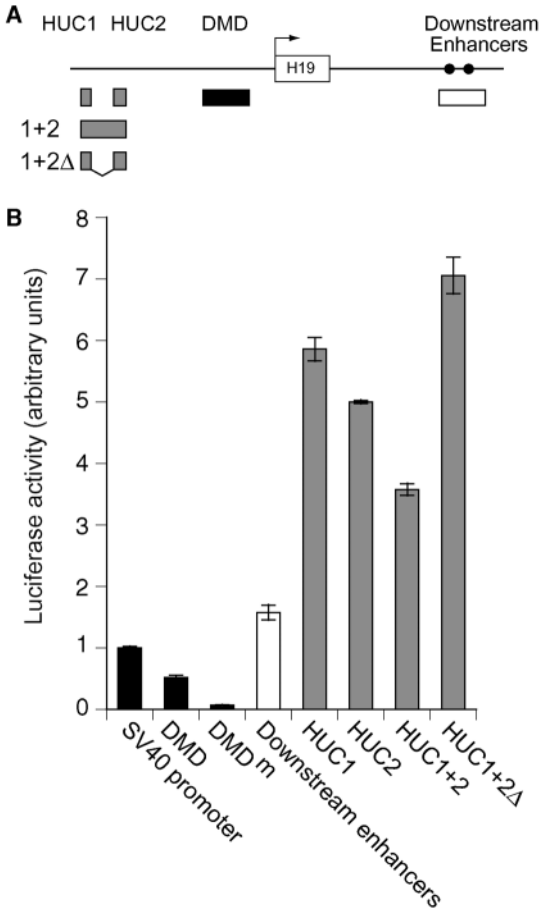

Fig. 4. Enhancer activity of HUC1 and HUC2 in vitro. (A) Map indicating the regions tested in the cell transfection assay. Regions were cloned upstream of an SV40 promoter driving a firefly luciferase reporter gene. (B) Transcriptional activity of transfected constructs. The light emission obtained from the SV40 promoter alone was normalised to a value of 1. DMD, *H19* differentially methylated domain; DMD^m, methylated DMD. HUC 1, HUC2 and both regions including the intervening sequence (HUC1+2) were tested. Maximum enhancer activity was detected from HUC1+2Δ (HUC1+2 with intervening sequence removed). Error bars show calculated standard error of the mean (s.e.m.) values.

HUC sequences demonstrate enhancer activity in vitro

As demonstrated previously, short regions of very highly conserved sequence between species may also be indicative of *cis*-regulatory elements. Long range regulatory elements (Loots et al., 2000) and transcriptional enhancers (Aparicio et al., 1995) have been identified in this way. The enhancers identified to date for the *H19* and *Igf2* genes are located downstream of the *H19* gene (Fig. 1B) (Ishihara et al., 2000; Leighton et al., 1995). It is, however, possible that the HUC sequences represent additional enhancer elements. We first investigated the ability of the HUC sequences to act as enhancers in a HeLa cell transfection system. In this assay, HUC1 and HUC2 demonstrated enhancer activity that was approximately fourfold greater than that of the previously characterised endodermal enhancers located downstream of *H19* (Fig. 4) (Leighton et al., 1995; Brenton et al., 1999).

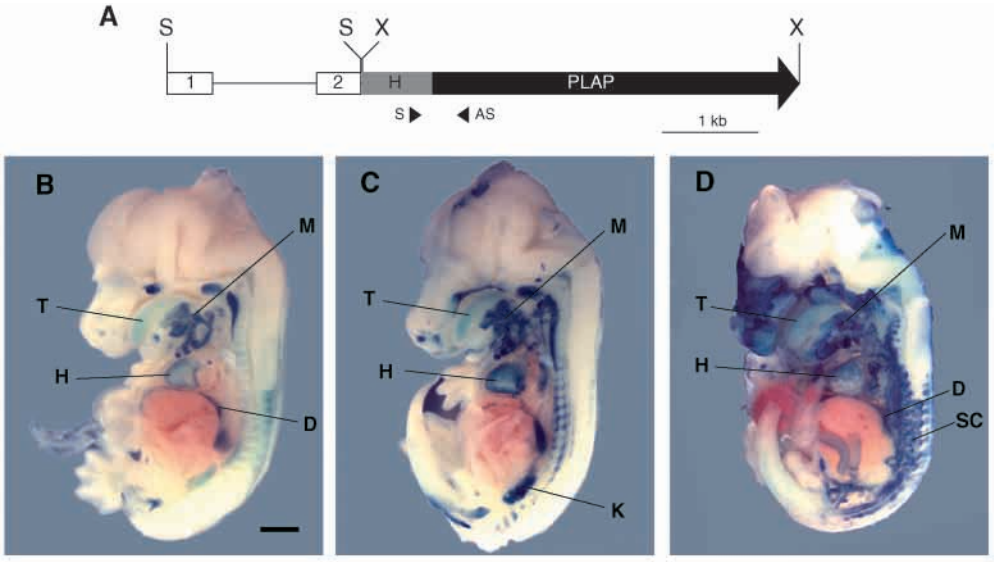

Fig. 5. Enhancer activity of the HUC sequences in vivo. (A) The endogenous locus and HUC-*PLAP* transgene. HUC 1+2 are indicated by the white boxes, the *H19* promoter region by the grey box (H) and the PLAP reporter by the black arrow. S (*Spe*I) and X (*Xba*I) are restriction sites used for construction. PCR primers used for genotyping the embryos are indicated. Embryos were confirmed transgenic by PCR genotyping (data not shown). (B-D) Three embryos showing typical *PLAP* staining pattern obtained. This expression pattern was highly consistent between all transgenic embryos. Non-transgenic embryos, as identified by PCR, failed to show any *PLAP* expression. D, diaphragm; H, heart; K, kidney; M; Meckel's cartilage; SC, spinal cartilage; T, tongue. Scale bar in B: 1 mm.

Interestingly, a drop in enhancer activity was observed when both HUC1 and HUC2 and the intervening sequence was tested. However, when this intervening sequence was removed, the enhancer activity detected was significantly stronger than that of HUC1 or HUC2 alone. In comparison, the *H19* differentially methylated domain (DMD) or ICR, which is responsible for silencing the paternal *H19* allele when methylated (Drewell et al., 2000), mediates transcriptional repression in this assay (Fig. 4). Therefore, this in vitro assay appears to be a reliable indicator of both enhancer and silencer activity of *cis*-regulatory elements from imprinted loci, in agreement with our previous studies (Arima et al., 2001).

HUC sequences can act as enhancers in vivo

To test whether the enhancer activity we detected in vitro represents a genuine in vivo function, we used a transgenic approach. A region containing both HUC1 and HUC2, plus the intervening sequence, was placed upstream of a reporter gene comprising a 0.8 kb *H19* promoter and the placental alkaline phosphatase (*PLAP*) gene (Henthorn et al., 1988). It has been previously demonstrated that transgenes containing 3.7 kb of 5′ flank upstream of the *H19* gene have no transcriptional activity in the absence of enhancer elements (Elson and Bartolomei, 1997).

Six pre-germline transgenic embryos, carrying varying numbers of copies of the transgene, were recovered at day 13.5 of gestation. The *PLAP* staining patterns were remarkably consistent between the transgenic embryos, with reporter expression principally in the developing cartilage in the ribs and spinal column, heart and developing kidney (Fig. 5B-D). Expression could also be detected in the diaphragm and tongue, and in the lung and skeletal muscle of the head. This highly tissue specific expression pattern strongly suggests that the HUCs are exclusively mesodermal enhancers, as opposed to harbouring an intrinsic general transcriptional activation activity associated with a promoter sequence. Significantly, *PLAP* expression was never detected in the liver where both endogenous *H19* and *Igf2* are highly expressed. Expression of these genes in the liver and other endodermal tissues has previously been shown to be controlled by enhancers located downstream of *H19* (Leighton et al., 1995).

It is important to note that the expression pattern observed is appropriate for a subset of tissues when compared with the full expression pattern of the endogenous *Igf2* and *H19* genes (Fig. 2E) (Leighton et al., 1995). This rules out the possibility that the transgene displays ectopic expression, and supports the notion that the HUCs are major enhancers for mesodermal tissues and can interact with the *H19*, and probably, *Igf2* promoters. Furthermore, we also note that the previously identified enhancers at the *H19/Igf2* locus do not account for the full expression patterns of these two genes (Ishihara et al., 2000; Leighton et al., 1995). What is unusual, however, is the location of these HUC enhancers. For the first time, such enhancers have been detected upstream of *H19* and the ICR. This is significant because the ICR contains a proposed insulator element that is suggested to play a crucial role in regulating promoter-enhancer communication (Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000). The location of major mesodermal enhancers upstream of the ICR necessitates re-evaluation of the mechanism of imprinted expression of the *H19* and *Igf2* genes in the mesodermal tissues described here.

DISCUSSION

The HUC sequences represent novel DNA elements within the mouse and human *H19* and *Igf2* imprinted domain. We show

that HUC1 and HUC2 are enhancers both in vitro and in vivo. Indeed, in HeLa cells, the HUC sequences have enhancer activity which is significantly greater than that of previously characterised enhancers downstream of *H19*. This may be a reflection of differences in their activity or, more likely, a demonstration of the underlying differences in their tissue specificity.

HUC enhancer tethering to the *H19* promoter

The HUC enhancers are clearly capable of driving expression of the *H19-PLAP* reporter gene from a minimal *H19* promoter in a wide range of mesodermally derived tissues in day 13.5 embryos. This expression is in a subset of tissues in which the endogenous maternal *H19* gene is normally expressed (see Fig. 2E), strongly suggesting that they interact with the endogenous *H19* promoter. It is interesting to note that a 140 kb *H19* BAC transgene extending only –6 kb upstream of the *H19* transcriptional start site (and therefore not containing the HUCs) shows significantly reduced expression in the heart and kidney (Kaffer et al., 2000), tissues in which we detect strong transcriptional activation by the HUCs. However, the HUCs and the endogenous *H19* gene are located on opposite sides of the ICR, which harbours a proposed insulator on the maternal chromosome. (Fig. 6A) (Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000). We propose that a key function of the ICR is to tether enhancers that are active in these distinct tissues to the *H19* promoter. In this case, the interaction of the HUCs with the maternal *H19* gene may be mediated by the insulator itself. The insulator may potentially act as a tethering element by recruiting a protein complex capable of directing these enhancers (and possibly the other downstream mesendodermal enhancers) to the *H19* promoter (Fig. 6A). Such a positive regulatory function for the unmethylated ICR was revealed by an extensive deletion of this region at the endogenous locus, which resulted in a decrease in the level of *H19* expression (Thorvaldsen et al., 1998).

It is not possible to directly attribute a tethering/boundary function to the ICR in mesodermal tissues, as only a very limited number of tissues (predominantly neonatal liver) were studied in mice carrying a comprehensive deletion of the ICR (Thorvaldsen et al., 1998). The hypersensitive sites that map to the ICR when maternally inherited have been found in all tissues examined to date; including, liver, brain and ES cells, irrespective of the transcriptional status of *H19* (Hark and Tilghman, 1998; Khosla et al., 1999). This consistent chromatin organisation suggests that the maternal ICR has the

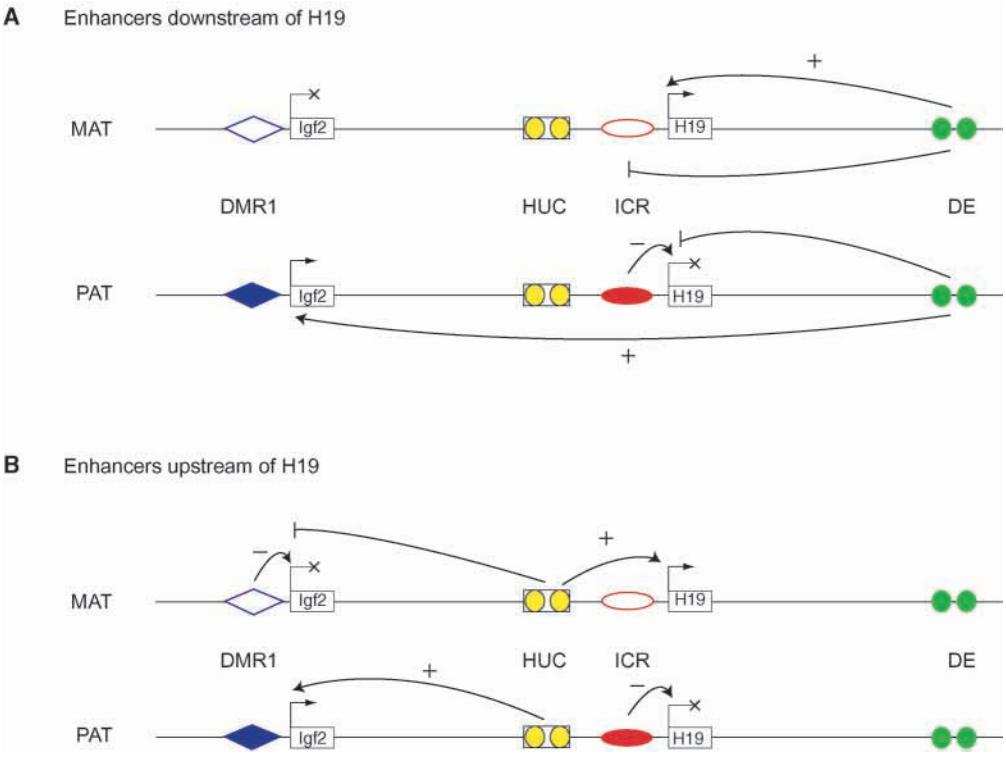

Fig. 6. (A,B) Model of proposed functional interaction of the HUC sequences with *Igf2* and *H19*. Activation of gene expression is represented by (+), repression is represented by (–) and inhibition of enhancer function is indicated by a vertical bar. *H19* and *Igf2* expression in some tissues is controlled by enhancers located downstream of *H19* (green circles) (Leighton et al., 1995; Ishihara et al., 2000). On the maternal allele (MAT), the enhancers have access to the *H19* promoter but are prevented from activating *Igf2* by the boundary function present in the unmethylated ICR (open red ellipse), mediated by CTCF. On the paternal allele (PAT), the ICR is methylated (closed red ellipse), acting to silence *H19* (Drewell et al., 2000). The absence of the boundary means the downstream enhancers are free to activate *Igf2*. In a subset of mesodermal tissues, expression is controlled by the HUC enhancers located upstream of *H19* (yellow circles). On the maternal allele, interaction of the HUCs with the H19 promoter may be mediated by a tethering activity within the unmethylated ICR (open red ellipse). Expression of *Igf2* by the HUC enhancers in mesodermal tissues is prevented by the unmethylated tissue-specific silencer element located at the *Igf2* DMR1 (open blue diamond) (Constancia et al., 2000). On the paternal allele, the HUCs are free to activate expression of *Igf2* while *H19* is silenced as before.

potential to function similarly in all tissues. Taken together with the absence of sufficient data about the role of the ICR in a variety of tissues this by no means rules out a role for the ICR in mesodermal tissues.

There is a precedent for similar promoter-enhancer interactions at complex genetic loci in *Drosophila*, where facilitator proteins bound to insulator elements are thought to specifically direct enhancers to promoters over long distances (Dorsett, 1999; Sipos et al., 1998), although the mechanisms have yet to be fully elucidated. Indeed, in the case of GAGA-mediated tethering activity at the *eve* promoter the stimulatory effect is disrupted when GAGA is separated from the promoter. The situation at the *H19* ICR appears to be different, in that the tethering activity acts over 2 kb. However, similar to the complex situation at the *H19* ICR, *Drosophila* insulators are also believed to be multipartite with the capacity to interact with enhancers (Geyer, 1997).

The role of the HUCs in regulating *Igf2* expression

On the paternal chromosome, the *H19* gene is silenced by DNA methylation through the recruitment of a repressive protein complex (R. A. D., unpublished). This would release the enhancers to interact with the closely linked paternal *Igf2* gene (Fig. 6B). As noted before, unlike the *H19* gene, both parental alleles of *Igf2* are potentially active (Sasaki et al., 1992). Therefore, the mechanism by which the maternal *Igf2* gene remains repressed is different from that governing *H19* silencing by DNA methylation. The regulation of the *Igf2* gene is governed by accessibility to the shared enhancers. The enhancers downstream of *H19* are apparently prevented from interacting with the maternal *Igf2* allele through the presence of an active insulator function within the ICR (Fig. 6A). This mechanism would not apply to the HUCs, which are located upstream of the ICR. However, there is, in addition, an epigenetically regulated silencer element upstream of *Igf2* at DMR1 that ensures silencing of the maternal *Igf2* gene in specific tissues (Fig. 6A). This was demonstrated by deletion of DMR1, resulting in the activation of the maternal *Igf2* gene. However this activation of *Igf2* was observed only in certain mesodermal tissues such as the heart and kidneys (Constancia et al., 2000). Expression of *H19*, and by inference the function of the *H19* ICR, was unaffected in these animals. This reactivated expression pattern shows striking resemblance to that conferred by the HUC enhancers described here. This suggests that there is an interaction between the HUC enhancers and the *Igf2* DMR1 in these tissues.

The combined data relating to the *cis*-regulation of the endogenous *Igf2* and *H19* genes provides a compelling explanation for how access to the HUCs and other enhancers may be regulated by the ICR upstream of *H19* and DMR1 upstream of *Igf2*. However, expression of *H19/Igf2* from a 130 kb YAC transgene is relatively low or absent in the heart, kidney and other tissues (Ainscough et al., 2000a; Ainscough et al., 1997). These are tissues in which the HUC enhancers play a significant role in driving expression. The YAC clone does not contain the *Igf2* DMR1 silencer element or further sequence upstream of *Igf2*. This suggests that the complete imprinted expression of *Igf2* and *H19* is dependent on the interaction of a complex and extensive network of *cis*-regulatory elements at the locus, some of which may remain to be identified. In addition, the role of the transcripts at the HUC sequences (mouse *Huc1* and mouse *Huc2*) is also unclear. Although the presence of small non-coding RNAs has long been established at imprinted regions (Arima et al., 2000; Moore et al., 1997; Takada et al., 2000) and other complex loci (Ashe et al., 1997; Zhou et al., 1999), their role remains largely enigmatic. We are currently generating a targeted deletion at the endogenous mouse locus to elucidate the role(s) of the HUCs as transcribed sequences, as enhancers or other as yet unknown functions in their normal in vivo context.

The authors thank S. Khosla for assistance with the sequence alignment and J. F.-X. Ainscough for advice. R. A. D. was funded by a Wellcome Trust Prize Fellowship. K. L. A. was supported by an Elmore Research Studentship from Gonville and Caius College, Cambridge. T. A. was funded by a Newton Trust grant. J. D. B. received a Research Fellowship for a Clinician (CRC). This work was supported by a grant from the Wellcome Trust to M. A. S.

REFERENCES

Ainscough, J. F.-X., Koide, T., Tada, M., Barton, S. and Surani, M. A. (1997). Imprinting of *Igf2* and *H19* from a 130 kb YAC transgene. *Development* **124**, 3621-3632.

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

Ainscough, J. F.-X., John, R. M., Barton, S. C. and Surani, M. A. (2000b). A skeletal muscle-specific mouse *Igf2* repressor lies 40 kb downstream of the gene. *Development* **127**, 3923-3930.

Aparicio, S., Morrison, A., Gould, A., Gilthope, J., Chaudhuri, C., Rigby, P., Krumlauf, R. and Brenner, S. (1995). Detecting conserved regulatory elements with the model genome of the Japanese puffer fish, *Fugu rubripes*. *Proc. Natl. Acad. Sci. USA* **92**, 1684-1688.

Arima, T., Drewell, R. A., Oshimura, N., Wake, N. and Surani, M. A. (2000). A novel imprinted gene, *HYMAI*, is located within an imprinted domain on human chromsome 6 containing *ZAC*. *Genomics* **67**, 248-255.

Arima, T., Drewell, R. A., Arney, K. L., Inoue, J., Makita, Y., Hata, A., Oshimura, M., Wake, N. and Surani, M. A. (2001). A conserved imprinting control region at the *HYMAI/ZAC* domain is implicated in transient neonatal diabetes mellitus. *Hum. Mol. Genet.* **10**, 1475-1483.

Arney, K. L., Erhardt, S., Drewell, R. A. and Surani, M. A. (2001). Epigenetic reprogramming of the genome- from the germline to the embryo and back again. *Int. J. Dev. Biol.* **45**, 509-516.

Ashe, H. L., Monks, J., Wijgerde, M., Fraser, P. and Proudfoot, N. J. (1997). Intergenic transcription and transinduction of the human beta-globin locus. *Genes Dev.* **11**, 2494-509.

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

Brenton, J. D., Drewell, R. A., Viville, S., Hilton, K. J., Barton, S. C., Ainscough, J. F.-X. 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

Constancia, M., Dean, W., Lopes, S., Moore, T., Kelsey, G. and Reik, W. (2000). Deletion of a silencer element in *Igf2* results in loss of imprinting independent of *H19*. *Nat. Genet.* **26**, 203-206.

Dorsett, D. (1999). Distant liaisons: long-range enhancer-promoter interactions in *Drosophila*. *Curr. Opin. Genet. Dev.* **9**, 505-514.

Drewell, R. A., Brenton, J. D., Ainscough, J. F.-X., 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.

Geyer, P. K. (1997). The role of insulator elements in defining domains of gene expression. *Curr. Opin. Genet. Dev.* **7**, 242-248.

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

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.

Henthorn, P., Zervos, P., Raducha, M., Harris, H. and Kadesch, T. (1988). Expression of a human placental alkaline phosphatase gene in transfected cells: Use as a reporter for studies of gene expression. *Proc. Natl. Acad. Sci. USA* **85**, 6342-6346.

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

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

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

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

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

Loots, G. G., Locksley, R. M., Blankespoor, C. M., Wang, Z. E., Miller, W., Rubin, E. M. and Frazer, K. A. (2000). Identification of a coordinate regulator of interleukins 4, 13 and 5 by cross-species sequence comparison. *Science* **288**, 136-140.

Moore, T., Constancia, M., Zubair, M., Bailleul, B., Feil, R., Sasaki, H. and Reik, W. (1997). Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse *Igf2*. *Proc. Natl. Acad. Sci. USA* **94**, 12509-12514.

Onyango, P., Miller, W., Lehockzy, J., Leung, C., Birren, B., Wheelan, S., Dewar, K. and Feinberg, A. P. (2000). Sequence and comparative analysis of the mouse 1-megabase region orthologous to the human 11p15 imprinted domain. *Genome Res.* **10**, 1697-1710.

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.

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

Sipos, L., Mihaly, J., Karch, F., Schedl, P., Gausz, J. and Gyurkovics, H. (1998). Transvection in the *Drosophila* Abd-B domain: extensive upstream sequences are involved in anchoring distant cis-regulatory regions to the promoter. *Genetics* **149**, 1031-1050.

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.

Surani, M. A. (1998). Imprinting and the initiation of gene silencing in the germ line. *Cell* **93**, 309-312.

Takada, S., Tevendale, M., Baker, J., Georgiades, P., Campbell, E., Freeman, T., Johnson, M. H., Paulsen, M. and Ferguson-Smith, A. C. (2000). *Delta-like* and *Gtl2* are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12. *Curr. Biol.* **10**, 1135-1138.

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.

Wilkinson, D. G. and Nieto, M. A. (1993). Detection of messenger RNA by *in situ* hybridisation to tissue sections and whole mounts. *Methods Enzymol.* **225**, 361-373.

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.

Zhou, J., Ashe, H., Burks, C. and Levine, M. (1999). Characterisation of the transvection mediating region of the abdominal-B locus in *Drosophila*. *Development* **126**, 3057-3065.