

# Imprinting of *IGF2* P0 transcript and novel alternatively spliced *INS-IGF2* isoforms show differences between mouse and human

D. Monk<sup>1,\*</sup>, R. Sanches<sup>2</sup>, P. Arnaud<sup>3</sup>, S. Apostolidou<sup>1</sup>, F.A. Hills<sup>4</sup>, S. Abu-Amero<sup>1</sup>, A. Murrell<sup>5</sup>, H. Friess<sup>6</sup>, W. Reik<sup>2</sup>, P. Stanier<sup>1</sup>, M. Constância<sup>2</sup> and G.E. Moore<sup>1</sup>

<sup>1</sup>Institute of Reproductive and Developmental Biology, Imperial College London, W12 ONN London, UK, <sup>2</sup>Laboratory for Developmental Genetics and Imprinting, The Babraham Institute, Cambridge CB2 4AT, UK, <sup>3</sup>Institute of Molecular Genetics, CNRS, UMR-5535 and University of Montpellier II, 1919 Route de Mende, 34090 Montpellier, France, <sup>4</sup>School of Health and Social Sciences, Middlesex University, Enfield EN3 4SA, UK, <sup>5</sup>Department of Oncology, Cambridge University, MRC-Hutchison Centre, Cambridge CB2 2XZ, UK and <sup>6</sup>Department of General Surgery, University of Heidelberg, Heidelberg, Germany

Received December 22, 2005; Revised and Accepted February 24, 2006

GenBank accession nos DQ104203–DQ104205

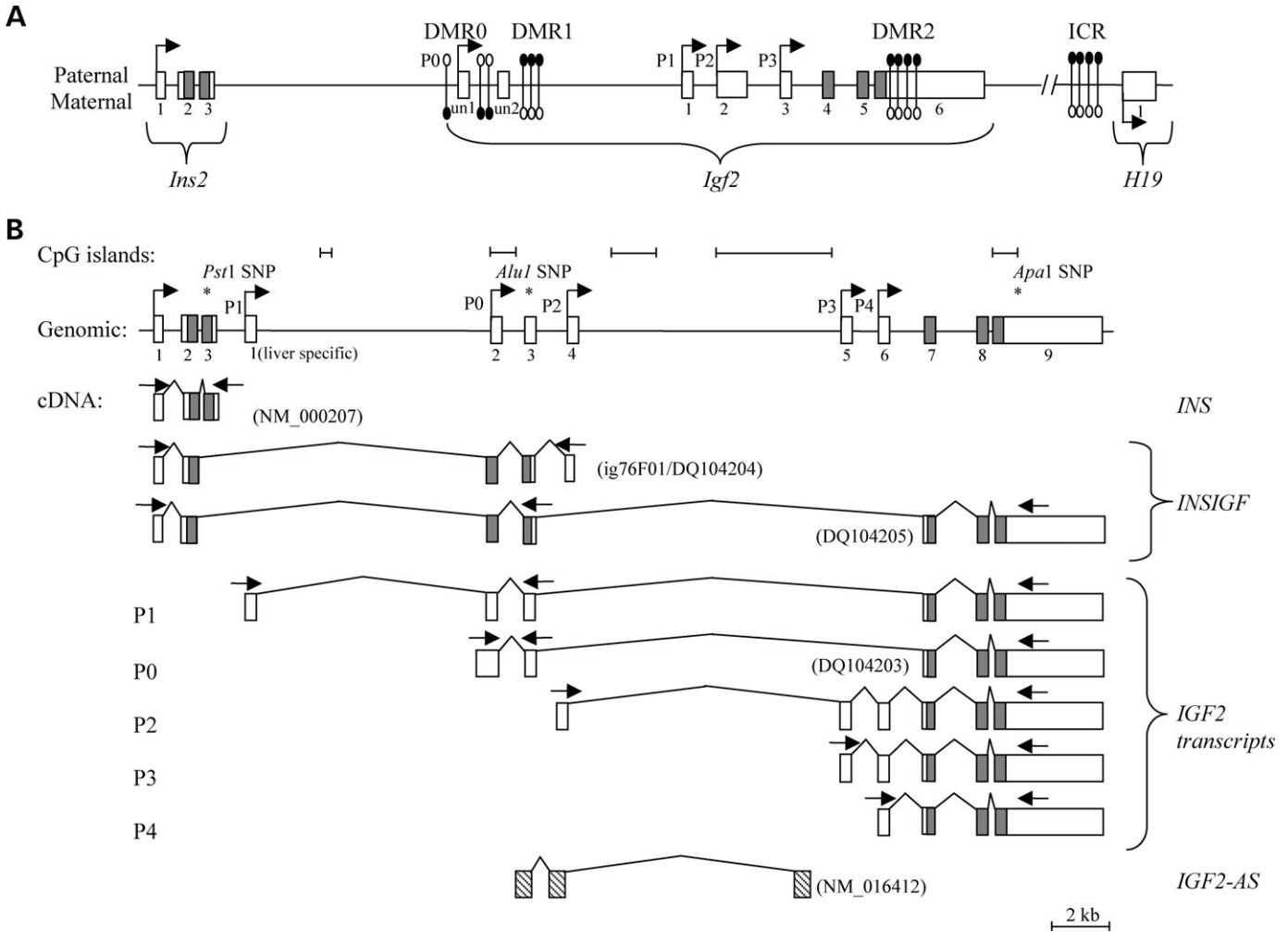
Genomic imprinting is limited to a subset of genes that play critical roles in fetal growth, development and behaviour. One of the most studied imprinted genes encodes insulin-like growth factor 2, and aberrant imprinting and DNA methylation of this gene is associated with the growth disorders Beckwith–Wiedemann and Silver–Russell syndromes and many human cancers. Specific isoforms of this gene have been shown to be essential for normal placental function, as mice carrying paternal null alleles for the *Igf2*-P0 transcript are growth restricted at birth. We report here the identification of three novel human transcripts from the *IGF2* locus. One is equivalent to the mouse *Igf2*-P0 transcript, whereas the two others (*INSIGF* long and short) originate from the upstream *INS* gene that alternatively splices to downstream *IGF2* exons. In order to elucidate the molecular mechanisms involved in the complex imprinting of these novel *IGF2* transcripts, both the allele-specific expression and methylation for all the *IGF2* promoters including P0 and the *INSIGF* transcripts were analysed in human tissues. Similar to the mouse, the human *IGF2*-P0 transcript is paternally expressed; however, its expression is not limited to placenta. This expression correlates with tissue-specific promoter methylation on the maternal allele. The two novel *INSIGF* transcripts reported here use the *INS* promoter and show highly restricted tissue expression profiles including the pancreas. As previously reported for *INS* in the yolk sac, we demonstrate complex, tissue-specific imprinting of these transcripts. The finding of additional transcripts within this locus will have important implications for *IGF2* regulation in both cancer and metabolism.

## INTRODUCTION

Genomic imprinting is an epigenetic phenomenon that results in the expression of only one allele in a parent-of-origin-dependent manner (1). To date, more than 70 imprinted genes have been described in the mouse (<http://www.mgu.har.mrc.ac.uk/imprinting/imprinting.html>). One of the first imprinted genes to be described was the paternally expressed insulin-like growth factor 2 (*Igf2*) (2). Over the

past decade, this gene has been extensively studied because it is a potent embryonic growth factor that is crucial during embryogenesis and is implicated in the human disorders Beckwith–Wiedemann syndrome (BWS) (MIM 130650), Silver–Russell syndrome (SRS) (MIM 180860) and in tumourigenesis (3–5). The *Igf2* gene, comprising many alternatively spliced transcripts from different promoters, maps to a large imprinted domain on distal mouse chromosome (Chr) 7 and is flanked 5' by insulin (*Ins2*) and 3' by the non-coding

\*To whom correspondence should be addressed. Tel: +44 2075942125; Fax: +44 2075942129; Email: d.monk@imperial.ac.uk



**Figure 1.** (A) Schematic scale diagram of the mouse *Igf2/H19* domain on distal mouse Chr 7. The mouse *Igf2* gene comprises multiple transcripts originating from unique promoters, all of which splice into the common protein-coding exons (shaded). The *H19* untranslated gene is located ~72 kb distal of *Igf2* and the *Ins2* gene ~9 kb proximal. Regions of differential methylation are shown, with the filled circles representing methylated and open circles the unmethylated regions. (B) Schematic scale diagram of the human *INS*, *IGF2* and *INSIGF* transcripts. CpG islands and the splicing patterns of the sense and anti-sense transcripts are shown. The coding exons are shaded, and the positions of the SNPs are indicated by asterisks. The arrows represent the approximate positions of the RT-PCR primers.

*H19* transcript, a region of conserved synteny with human Chr 11p15 (Fig. 1).

In both human and mouse, the paternal expression of *Igf2* and *Ins2* is closely linked to the expression of the reciprocally imprinted *H19* transcript (6). Research into the imprinting control of this domain has revealed a complicated regulatory mechanism that utilizes multiple enhancers, boundary elements, histone modifications and complex physical DNA looping, which are all allele-specific and some of which are species-specific (7,8). Many studies of the *Igf2/H19* interval have implicated specific regions of differential DNA methylation as critical for correct allelic expression of these imprinted genes. The mouse *Igf2/H19* domain is known to contain four differentially methylated regions (DMRs): DMR0 surrounding the region of the placental-specific P0 promoter (9); DMR1, situated 3 kb upstream of promoter P1, a region known to bind the transcription repressor GCF2 (10,11); DMR2, located at the 3'-untranslated region (UTR)

of *Igf2* (12,13); lastly, the *H19* imprinting control region (ICR) that has been shown to act as a methylation-sensitive boundary, binding the insulator protein CTCF (6,13,14). Sequence analysis has revealed the human *IGF2/H19* region shares many of the same features, with a few notable exceptions. These include the absence of a human DMR1 and, although both genes are associated with four different promoters, the human gene without a previously reported P0 has an upstream 'liver-specific' promoter, which has not been found in the mouse (15).

The expression of the different mouse *Igf2* transcripts is dependent on promoter usage, with P0-derived transcripts only expressed in the labyrinthine layer of the mouse placenta, whereas transcripts from the P1–P3 are found throughout the developing embryo and placenta (9,16). The reported human *IGF2* transcripts (P2–P4) are found in both the developing embryo and the extra-embryonic material, with adult expression limited to P1-derived biallelic transcription in

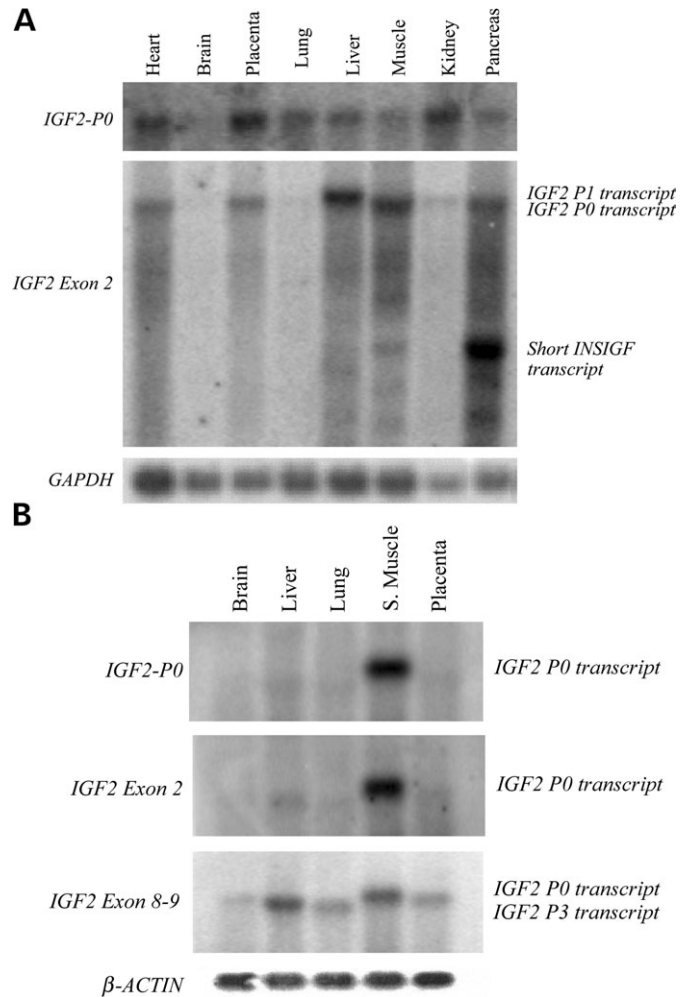
liver (17,18). The IGFII protein has been shown to be a major modulator of fetal and placental growth (19). Elegant gene targeting studies have shown that paternally inherited *Igf2* null alleles in heterozygous mice have the same growth restricted phenotype as homozygous null animals (~60% of wild-type weight) (2). Targeted disruption specifically of the mouse *Igf2*-P0 transcript, by removing the untranslated exon 2, resulted in a milder but still growth-restricted phenotype when compared with the total *Igf2* null, even though the fetal promoters (P1–P3) were unaffected and remained transcriptionally active (16). The *Igf2*-P0 null mice were reported to have a smaller than normal placenta at embryonic day (E)16 of gestation, which was shown to affect the nutrient transfer capacity of the placenta (20). However, once the placental constraints on growth were lifted at birth, postnatal catch-up growth was observed.

In the present study, we have investigated the imprinted expression of the various transcripts of the human *IGF2* gene, both in first trimester fetal and placental tissues and term placental material. In addition to the four previously known *IGF2* promoters, we describe the expression, imprinting and methylation of three novel transcripts, the human *IGF2*-P0 promoter/transcript and two *INS* linked to *IGF2* (*INSIGF*) transcripts and discuss their potential involvement in human development both *in utero* and in tumourigenesis.

## RESULTS

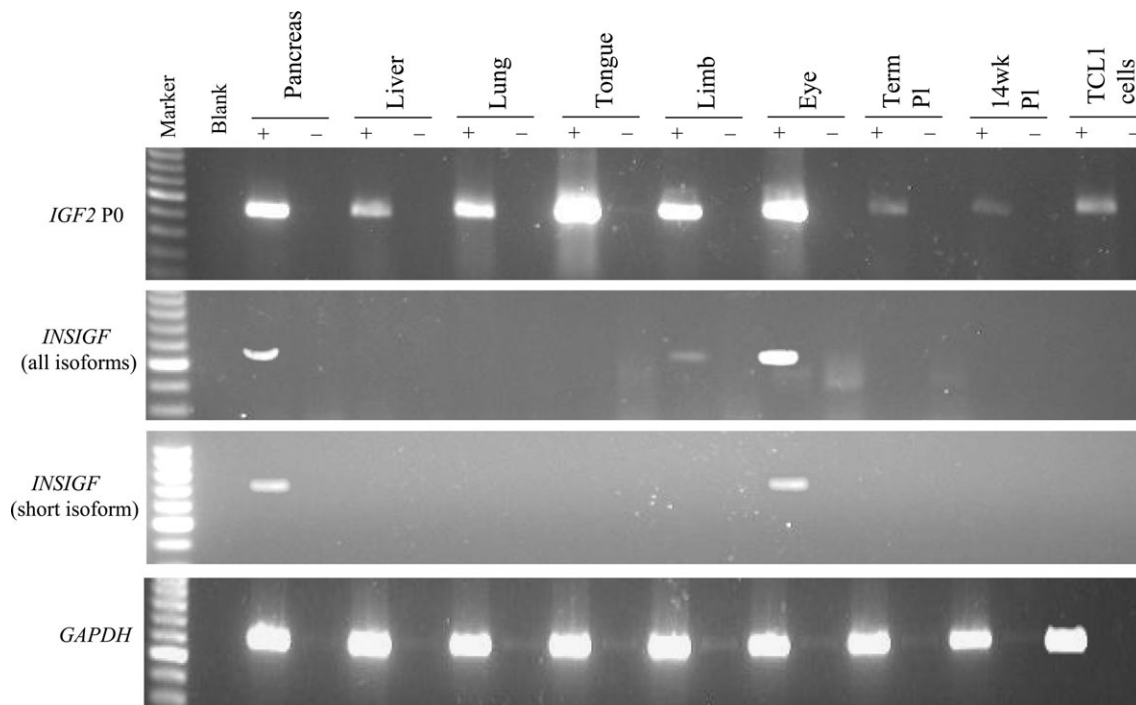
### Identification and expression of the human *IGF2*-P0 and *INSIGF* transcripts

Multiple transcripts of human *IGF2* are synthesized as a result of alternate promoter usage and splicing of the unique 5'-exons to the 3'-downstream exons encoding the IGFII precursor protein. Promoters P2–P4 are known to be transcriptionally active in many fetal tissues and show a general down-regulation after birth (21). The P1 transcript is unique to humans and incorporates the untranslated exons 1–3. It has been reported that the P1 transcript is expressed from both parental alleles in postnatal liver and fetal choroid plexus/leptomeninges (17). In the mouse, a placenta-specific *Igf2* transcript is transcribed from the P0 promoter and incorporates the first two untranslated exons (equivalent to human exons 2 and 3; Fig. 1) that subsequently splice onto the common coding exons. As a consequence of the mouse data with its strong implications as a model for intrauterine growth restriction, we set out to look for a human 'P0 transcript'. This should potentially incorporate human exon 2 (as the first exon), exon 3 and the common coding exons 7–9. A 5'-RACE revealed cDNA products originating 5' to the exon 2 donor site (ctccagCCTCA nucleotide 139620 of AC130303) and extending 258 bp into the upstream intron 1 (data not shown). Using various strategically designed forward primers within this region and reverse primers located within either exon 3 or 7, RT-PCR was used to confirm the 5'-RACE data. This demonstrated that the human *IGF2*-P0 transcript initiation site is located 5' to the splice donor site of the P1 transcript's exon 2 and that this 5'-UTR sequence is unique to the human *IGF2*-P0 transcript.



**Figure 2.** (A) A multi-tissue adult northern blot hybridized with an anti-sense riboprobe corresponding to the unique 5'-*IGF2*-P0 sequence reveals a single band of ~5 kb in all tissues. The same blot was reprobed with an anti-sense riboprobe corresponding to *IGF2* exon 2. The ubiquitously expressed *IGF2*-P0 transcript was detected, as was the liver-specific P1 transcript and the short *INSIGF* transcript in pancreas. The blot was hybridized with a human *GAPDH* cDNA fragment to ensure equal sample loading. (B) Fetal northern blots hybridized with the *IGF2*-P0 and exon 2 riboprobes reveal expression limited to skeletal muscle. The same blots were subsequently rehybridized with an *IGF2* exons 8 and 9 riboprobes to detect all *IGF2* isoforms and a  $\beta$ -Actin probe to ensure equal loading.

Northern blotting has previously shown that the mouse *Igf2*-P0 transcript is solely expressed in the placenta, whereas *in situ* hybridization shows that it is specifically restricted to the labyrinthine layer (9,16). To assess the expression of the human *IGF2*-P0 transcript, we hybridized an adult multi-tissue northern blot with an anti-sense riboprobe consisting of the *IGF2*-P0-specific 5'-UTR sequence. A single band of ~5 kb was detected in term placenta and all adult human tissues, except brain. This corresponds to a predicted *IGF2*-P0 transcript of 4998 bp (transcript accession no. DQ104203) utilizing the AATAAA polyadenylation motif (nucleotides 159837–159841 of AC130303) common in other *IGF2* transcripts (Fig. 2A). To investigate whether the expression of the *IGF2*-P0 transcript was limited to



**Figure 3.** RT-PCR on a panel of human fetal cDNA reveals ubiquitous expression of *IGF2*-P0, with weaker expression in the first trimester and term placenta samples. The expression of *INSIGF* is limited to pancreas and eye. The lower panel shows PCR, using *GAPDH* primers to ensure cDNA integrity.

placenta during early development, we hybridized a northern blot containing selected first and second trimester tissues samples with the same P0-specific riboprobe. In contrast to the mouse, this revealed that the expression of the *IGF2*-P0 transcript is not detectable in early placenta but is most abundant in fetal skeletal muscle (Fig. 2B). The same adult and fetal northern blots were then sequentially re-hybridized with riboprobes specific to *IGF2* exon 1, exon 2 and a fragment encompassing exons 8 and 9 in order to detect expression of other *IGF2* transcripts. The *IGF2* exon 2 riboprobe revealed ubiquitous expression of the *IGF2*-P0 transcript in adult tissues (except brain) as well as the slightly larger *IGF2*-P1 transcript in liver.

Interestingly, the exon 2 riboprobe revealed an ~1.5 kb pancreas-specific band that was not detected when using either the P0-specific or P1-specific riboprobes (Fig. 2A). From the investigation of expressed sequence tag (EST) databases, we suspected that this band might result from a novel transcript originating from the *INS* promoter and splicing to *IGF2* exon 2 (EST transcript accession nos ig76F01/BM312593 or ir36b02/CA866245). ESTs for this transcript were pancreas-specific and consist of *INS* exons 1 and 2, which splice on to *IGF2* exons 2–4. In addition, a precisely matching protein sequence (XP\_508220) has been predicted by the computational GNOMON method in the *Pan troglodytes* genome. We have named this transcript *INSIGF*. As the transcript lacks *INS* exon 3 and the usual stop codon, the open reading frame differs from that of *INSULIN* and extends into the *IGF2* exon 3. Translation of this ORF would result in a novel protein that not only includes the *INSULIN* leader sequence and B-chain peptide but also an

additional peptide at the C-terminal end. Analysis of the Prodom and Pfam databases does not detect any other known protein motifs within this amino acid translation. Subsequent RT-PCR amplifications revealed that a longer splice variant, which includes *INS* exons 1 and 2 and *IGF2* exons 2, 3, 7, 8 and 9 (transcript accession no. DQ104205), also exists. This transcript is bicistronic, with the open reading frames present for both the novel *INSIGF* protein and *IGFII* (Fig. 1B). Both transcripts of *INSIGF* show a restricted expression profile when analysed by RT-PCR. Primers within exon 1 of *INS* and exon 3 of *IGF2* (detecting both *INSIGF* transcripts) show expression restricted to pancreas, eye and limb. However, the specific amplification of the short *INSIGF* isoform using a reverse primer in *IGF2* exon 4 detected products only in pancreas and eye (Fig. 3).

### Mouse *Insigf* transcripts

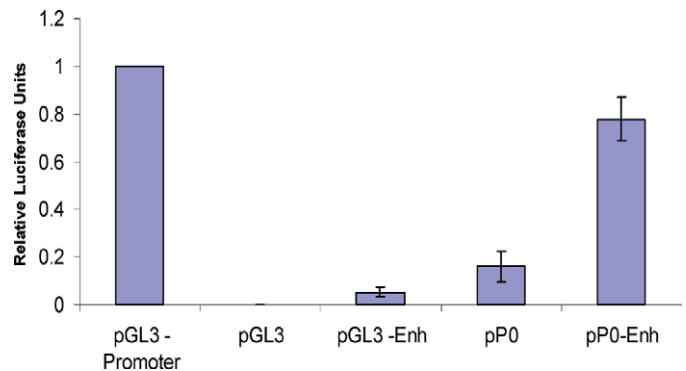
Investigation of a potential mouse *Insigf* sequence has failed to identify a transcript by RT-PCR, northern blotting or interrogation of dbEST. Nevertheless, the alignment of the mouse and human genomic sequences reveals significant homology corresponding to the exons and exon-intron boundaries, at least up to and including the mouse untranslated exon u2. Translation of this putative transcript would result in a protein similar to the human *INSIGF*. The conceptual protein translations, however, differ near the C-terminal ends where the mouse sequence incorporates a 27 bp imperfect tandem repeat sequence. With no stop codon present until after this sequence, 194 amino acids based on the PEPPEPPK repeat motif would be added (Fig. 4).

mInsigf	1	<u>MALWMRFLPLLALLFLWESHPTQAFVKQHL</u> <u>CGSHLVEALYLVCGERGFFYTPMSRREVED</u>	60
		MALWMR LPLLALL LW P AFV QHLCGSHLVEALYLVCGERGFFYTP +RREED	
hINSIGF	1	<u>MALWMRLPLLALLLALWGPDPAAAFVNQHL</u> <u>CGSHLVEALYLVCGERGFFYTPKTRREAED</u>	60
mInsigf	61	<i>QDSAPSSPAATSSPRLEGPGTTTRAPPAL</i> <i>EAGAANIDPASGSSSGQFRKRI</i> <i>LGTSDSPVL</i>	120
		Q SA S ++TS+ EG T RAPPAL ANI A GSSS QFR+RLGTSDSPVL	
hINSIGF	61	<u>LQASALSLSSTST-WPEGLD</u> <u>ATARAPPALVVTANIGQAGGSSSRQFRQ</u> <u>RALGTSDSPVL</u>	119
mInsigf	121	FIHRPGTSGTTRKLEGRVRIKTELTVQEEEEEEENDYDDCRPLTQGSSESPKAKLLVQPE	180
		FIH PG +GT + LEYRGR + TE +E D PQGSES	
hINSIGF	120	FIHCPGAAGTAQGLEYRGRRVTEPVWEEV-----DSSPQPGSES-----	160
mInsigf	181	LQKPKSRVPPRPEPRPEPPKPEPRPEPPKPEPRPEPPKPEPRPEPPKPEPRPEPPKPEP	240
		L P PEP EP PE	
hINSIGF	161	LPAQPPAQAPQPEPQAR---EPSPE-----	185
mInsigf	241	<i>RPVPPRPEPRVPPRHEPRVPPRPEPRPLPPRPELRPLLLRSEPRVPPRPELRPLLLR</i>	300
hINSIGF	185	-----	185
mInsigf	301	<i>PEPRVPPRPELRPLPPRPELRPLPPRPESRPLPPRPEPRPLPPRPELKPEPEPELKP</i>	360
hINSIGF	185	-----	185
mInsigf	361	EPEPEKKQARASRDSSVPKCCGLWPRRRSPRSQN	394
		CCGLWPRR RSN	
hINSIGF	153	-----VSCCGLWPRRRQRSQN	200

**Figure 4.** Clustal W alignment of human and mouse conceptual INSIGF protein translations. Amino acids corresponding to INSULIN-leader sequence and B-chain peptide are underlined, and the mouse proline-rich repeats are in italic type.

### Characterization of the human *IGF2*-P0 promoter

Analysis of the 5'-RACE products for the *IGF2*-P0 transcript revealed a consensus sequence (PyPyAN(T/A)PyPy) for an initiator element (Inr element). Inr promoter elements are generally found to encompass transcription start sites in eukaryotic promoters and in particular TATA-less promoters. To measure *IGF2*-P0 promoter activity, we utilized a transient luciferase expression assay in a human-term extravillous trophoblast cell line (Tc11) (22). A 552 bp fragment immediately upstream of the first exon and including the putative *IGF2*-P0 5'-UTR was cloned into pGL3 reporter constructs both with and without an SV40 enhancer. This enhancer is known to stimulate the transcription of many eukaryotic promoters *in vitro*. The promoter activity was measured using the firefly luciferase reporter gene and normalized for transfection efficiency against the co-transfected *Renilla* luciferase reporter plasmid (TK-*Renilla* luciferase). Promoter activity of P0 constructs in replicate experiments relative to a positive control (pGL3-promoter) is shown in Figure 5. The Tc11 cells showed weak activity for the human P0 promoter (pP0 construct), which increased significantly when associated with the SV40 enhancer (P0-Enh construct). In fact, the luciferase activity detected with the P0-Enh construct is close to the corresponding value obtained for the positive pGL3-promoter control (SV40-promoter activity) and is significantly stronger than that of the pP0 construct and the pGL3-Enh vector (the enhancer only, no promoter control). The significant difference between reporter activities between the pP0-Enh construct and the pGL3-Enh vector is due to the presence/absence of the promoter sequence and confirms that the region cloned has promoter activity *in vitro*. The promoter activity reported for the mouse *Igf2*-P0

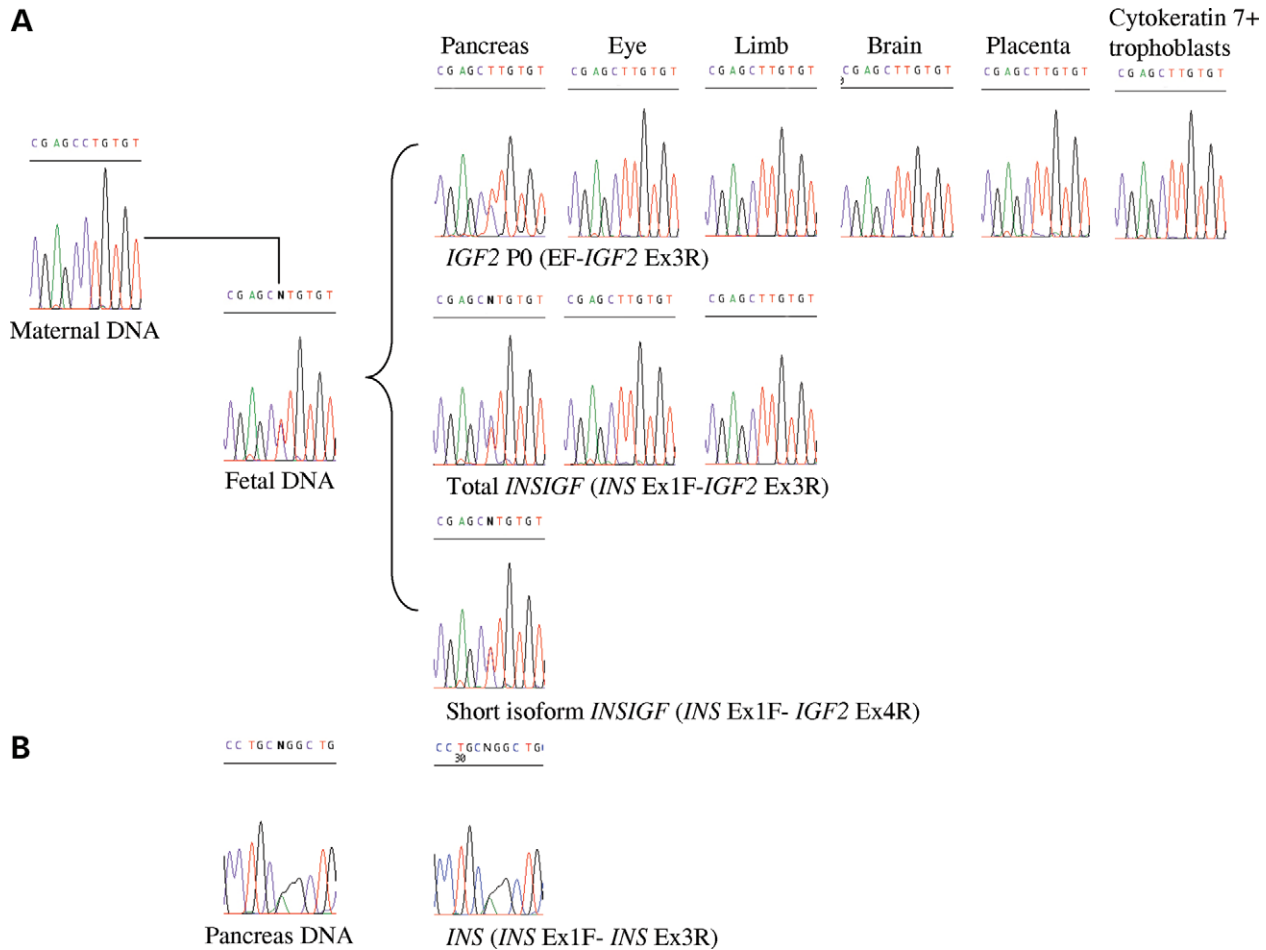


**Figure 5.** Luciferase assays showing promoter activity of P0 in the presence of enhancers. These results represent 10 separate experiments, each containing three replicates. For each assay, normalized luciferase readings were expressed as a ratio relative to the standard positive pGL3-promoter construct.

and P3 promoters in Tc11 cells is greater than the values obtained for human *IGF2*-P0 promoter, which corresponds to the higher transcription levels observed *in vivo* (Constancia and Murrell, unpublished data) (23).

### Promoter-specific imprinting of the human *IGF2*-P0 transcript

The various transcripts originating from the *IGF2* promoters were analysed for imprinted expression in human, using the RT-PCR approach. Thirty-term placentae and 18 fetal samples sets with corresponding parental blood samples were genotyped for both the *Alu1* SNP in exon 3 (rs17881514) and the *Apal* SNP in the 3'-UTR of *IGF2*



**Figure 6.** Imprinting of the *IGF2*-P0, *INSIGF* and *INS*. Parent-of-origin of expression was analysed using an SNP in exon 3 of *IGF2* or exon 3 of *INS*. Using strategically position PCR primers, each individual isoform could be analysed. (A) Genotyping of the *IGF2* exon 3 SNP reveals a heterozygous fetus (C/T) and homozygous maternal DNA sample (C/C). Expression for all transcripts arises from the paternal allele (T) in all tissues except pancreas, which is biallelic (C/T). (B) Pancreas sample heterozygous (A/G) for the *INS* exon 3 SNP and biallelic expression (A/G) of *INS* in the corresponding cDNA.

(rs680). Using both strand-specific (to avoid amplification of the paternally expressed *IGF2-AS/PEG8*) and random-primed cDNA, all heterozygous sample sets clearly showed that the expression from each individual *IGF2* promoter (P0–P4) was monoallelic, and in informative sample sets, this expression was derived from the paternal allele [three sample sets using rs17881514 (Fig. 6) and four sample sets using rs680 (data not shown)]. As the expression of the mouse *Igf2*-P0 transcript is labyrinthine-specific, we also analysed the imprinting of the human *IGF2*-P0 transcript in the first trimester ( $n = 2$ ) and term ( $n = 2$ ) cytotkeratin 7-enriched villous trophoblasts. These are the human cells equivalent to the mouse villous trophoblast cells contained within the labyrinthine zone (24). Paternal expression of the human *IGF2*-P0 transcript was also observed in these cells.

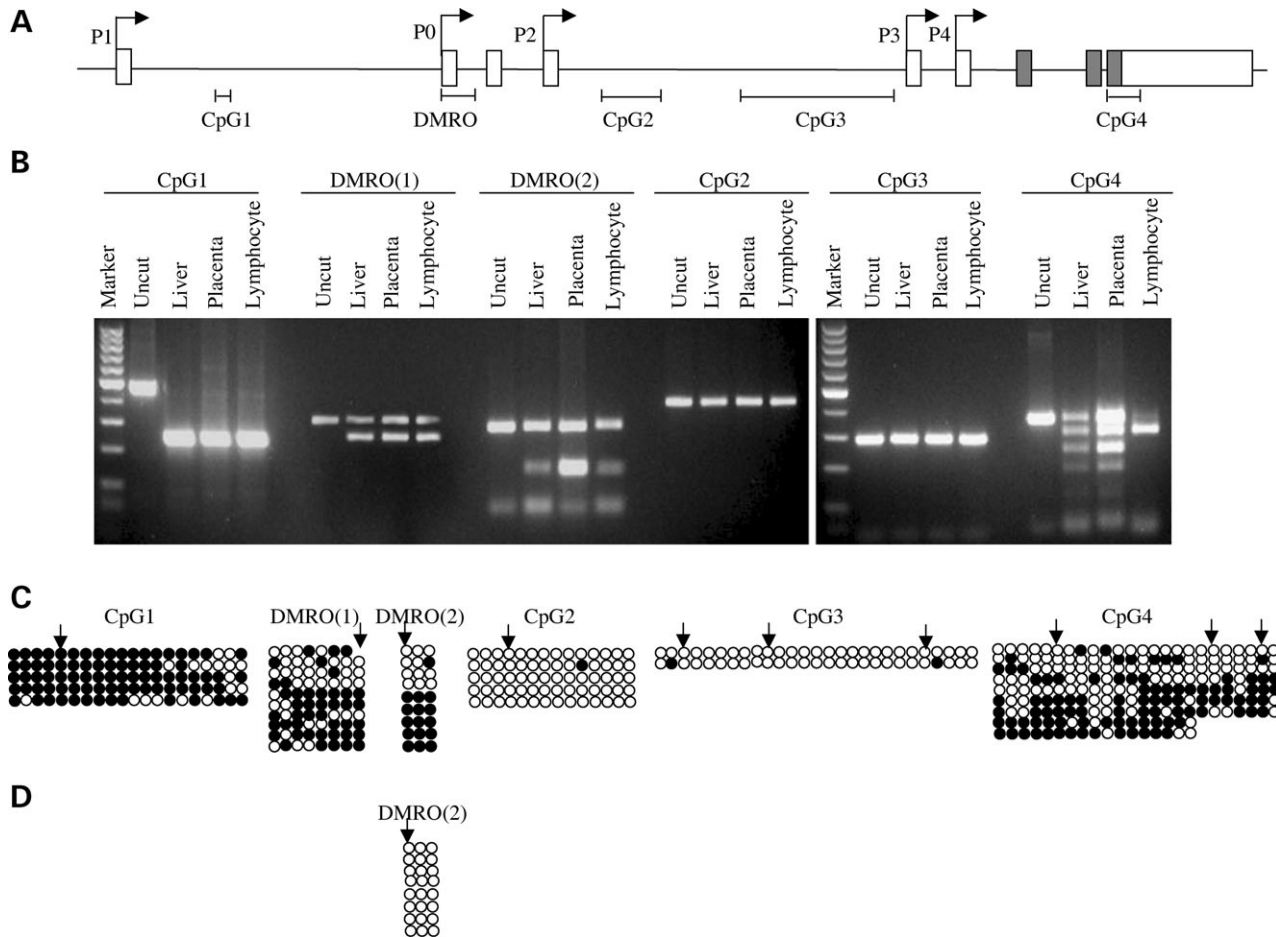
#### Imprinting of the *INSIGF* transcripts

The imprinting status of the *INSIGF* transcripts was studied in 20 normal adult pancreas samples and the same 18 fetal sample sets previously used to analyse the *IGF2* imprinting.

DNAs from these samples were also genotyped for a *Pst1* SNP in exon 3 of *INS* (rs3842757). Two pancreas samples were heterozygous for the rs17881514 and rs3842757 SNPs and were biallelic for *INS*, *INSIGF* and *IGF2*-P0 (Fig. 6). Interestingly, both *INSIGF* isoforms were monoallelic/paternally expressed in eye ( $n = 3$ ) and limb samples ( $n = 1$ ) (Fig. 6).

#### *IGF2* methylation profile

Many studies have implicated DNA methylation throughout the *IGF2/H19* region as critical for correct tissue-specific imprinting of the various *IGF2* transcripts (10–12). We, therefore, analysed all CpG islands within the *IGF2* genomic region, as well as the region surrounding the promoter of the human *IGF2*-P0 transcript (Fig. 7). Each region was analysed by treating genomic DNA from various tissues with sodium bisulphite, followed by PCR amplification that would include both methylated and unmethylated molecules. Resulting products were subjected to COBRA (combined bisulphite restriction analysis) (25) or cloned and sequenced.



**Figure 7.** (A) A schematic diagram of the human *IGF2* gene showing the relative positions of the promoters and CpG islands. (B) Restriction digestion of PCR products generated from bisulphite-treated DNA from human fetal tissues. All unmethylated C nucleotides should be converted to T nucleotides during the treatment. The restriction-enzyme recognition sequences for *Taq1* (TCGA) and *Tai1* (ACGT) will be only present in methylated DNA because methylated CpG dinucleotides are resistant to conversion. PCR products from CpG1 are completely digested by *Taq1*, indicating that this CpG island is fully methylated. The two PCR products from the P0 promoter region (DMR0) clearly have equal representation of digested and undigested molecules, indicating that this region is differentially methylated. The CpG2 and CpG3 PCR products are resistant to digestion with *Taq1*, indicating that these regions are fully unmethylated. Lastly, the CpG4 PCR products show a tissue-specific differential methylation pattern. (C) Subsequent cloning and sequencing of the various bisulphite PCR products from treated placenta DNA confirm the methylation profiles. Only the 3'-most CpG dinucleotides are shown for CpG3; however, all 50 CpG dinucleotides within the PCR product are unmethylated. The arrows represent the position of the *Taq1* or *Tai1* restriction sites within the PCR products. (D) The complete lack of methylation at DMRO (2) in a patient with pUPD11.

The first CpG island maps between the first two untranslated exons that comprise the human P1 transcript and is ~1.8 kb downstream of two *ALU* repeat elements. Using bisulphite mutagenesis and PCR on treated DNA from placenta, liver, kidney and lymphocytes, this CpG island was found to be fully methylated on both alleles in all tissues. The second and third CpG islands map to the promoter regions of P2–P4 and the *IGF2-AS* (*PEG8*) transcripts, essentially representing a single large CpG island. Two independent, non-overlapping PCR products revealed that this region was fully unmethylated in all tissues. The final CpG island maps to exon 9. Using Southern blot analysis, this region has previously been defined as a DMR in both human and mouse, with the paternal allele being more methylated than the maternal allele (10). However, the differential methylation of this region is disputed because detailed studies utilizing bisulphite analysis have revealed a mosaic methylation

profile (26). Our results also show that the methylation profile for this region is not uniform and that tissue-specific differences in methylation exist (Fig. 7).

The region surrounding the mouse P0 promoter is a maternally methylated DMR in placenta, even though it is not statistically defined as a CpG island (9). The equivalent genomic region surrounding human exon 2 also fails to meet the criteria for a CpG island. Nevertheless, differential methylation for this region has previously been reported, and hypomethylation is associated with the loss of imprinting (LOI) of *IGF2* in cancers (27–29). We analysed two regions within this interval, one overlapping with the human P0 promoter, which contains eight CpG dinucleotides, and the other between exons 2 and 3, incorporating the same CpG dinucleotides described by the Feinberg group (29). Both regions contained both fully methylated and unmethylated PCR products in all tissues analysed, indicating that the differential methylation of this region

is not limited to extra-embryonic material as seen in the mouse. Further analysis of PCR products obtained from bisulphite-treated pUPD11 DNA confirmed that the human P0 promoter is embedded within a maternally methylated DMR (Fig. 7D).

## DISCUSSION

The paternal expression of *IGF2* gene is regulated in a developmental-dependent and tissue-specific manner. Previously, the expression of *IGF2* in the fetus was thought to arise from four different promoters, utilizing alternative splicing and polyadenylation signals (15). Although P2, P3 and P4 promoters have counterparts in the rodent, the P1 promoter is unique to humans (18). Recently, a placental-specific transcript originating from the P0 promoter in mouse was discovered and investigated by the generation of a knockout (9,16). Paternal transmission of the P0 null allele resulted in a severe growth restriction of the placenta at E12, leading to fetal growth restriction at E18 due to disrupted nutrient transfer across the mouse labyrinthine layer (16,20). As the mouse *Igf2*-P0 transcript has a major effect on controlling growth and function of the placenta, we undertook an investigation to see whether such a mechanism is conserved in humans. We successfully identified a human equivalent to the mouse P0 transcript that, like the other *IGF2* transcripts from P2–P4, is paternally expressed in all fetal and placental material except pancreas. Using 5'-RACE and RT-PCR analysis, we have found that the human *IGF2*-P0 transcript start site is ~258 bp upstream of exon 2, approximately the same position as in the mouse (9), and that this region has promoter activity as demonstrated in an *in vitro* transient promoter assay in a trophoblast cell line.

Even though the human and mouse P0 transcripts originate from similar regions and have identical splicing patterns, there are subtle differences between species. The mouse *Igf2*-P0 transcript is expressed exclusively in the placenta and is specifically detected in the labyrinthine trophoblast layer (16). In contrast, the human fetal *IGF2*-P0 transcript is expressed at high levels in fetal skeletal muscle and later, at a lower transcribed level, it is found ubiquitously in adult tissues and term placenta. This marked change in spatial expression may result from the epigenetic differences we found between species in the P0 promoter regions. The paternal expression of the *Igf2*-P0 transcript in the mouse originates from the only site of maternal allele methylation in the *Igf2* locus. This maternal methylation is limited to extra-embryonic material, suggesting that the transcriptional silencing of the maternal allele may be controlled by promoter methylation (9). However, maternal methylation was observed in all human tissues, which correlates with maternal repression of the human *IGF2*-P0 transcript in the adult tissues.

It is not currently known whether the promoter methylation alone at the human DMR0 is sufficient to bring about maternal silencing of the P0 promoter, independent of the other *IGF2* promoters. Although it would have been of great interest to investigate the imprinting status of the *IGF2*-P0 transcript in relation to the other *IGF2* transcripts in BWS fibroblasts with *H19* DMR hypomethylation, the *IGF2*-P0 transcript is

not expressed in the available adult cell lines tested (data not shown). Nevertheless, evidence that the DMR0 plays a critical role in the regulation of *IGF2*-imprinted expression has come from the observation that this region is hypomethylated in Wilms' tumour and that colorectal cancers are associated with the reactivation of the maternal allele of *IGF2* (27–29). Indeed, other groups have proposed that the human DMR0 is a methylation-dependent silencer element that may be regulating *IGF2* expression independent of the *H19* DMR (29). It is, therefore, possible that the human P0 transcript has a role to play in cancer. If LOI of *IGF2*-P0 was found in cancerous tissues, specifically colorectal cancer with loss of methylation DMR0, it could act independently of the other *IGF2* isoforms that are mainly expressed and imprinted in the fetus. It remains to be studied whether the human *IGF2*-P0 transcript becomes biallelic in tumours and whether this over-expression might aid the overgrowth in these tumours (27,28).

We also report here our evidence for two novel transcripts originating from the *INS* promoter. Both contain a novel ORF altering the normal INSULIN protein. It is not yet clear whether this protein is translated *in vivo* and requires further studies to examine its potential function. The transcripts are biallelically expressed in pancreas, which would be expected because *INS*, sharing the same promoter, is also biallelic in this tissue. However, similar to *INS* that shows evidence of tissue-specific imprinting in the yolk sac and possibly also the thymus (30,31), *INSIGF* transcripts are monoallelically/paternally expressed in the limb and eye.

There may be interesting implications for a mouse *Insigf* transcript because the previously reported *Igf2*-P0 knockout mouse removes sequences potentially common to both transcripts. However, we can find no evidence for the expression of this sequence in the mouse. Also, the conceptually translated mouse *Insigf* protein is different than that of the human/monkey because of the presence of an additional simple sequence repeat found in the mouse un2 exon. This encodes a proline-rich insertion likely to render it non-functional.

Here, we have identified and characterized the human *IGF2*-P0 transcript as well as two novel imprinted transcripts running from *INSULIN* into *IGF2*. The human *IGF2*-P0 transcript is unlikely to have the same role in fetal growth as it does in the mouse because it is found in different tissues and is not specific to the placenta. However, it has a clear imprinted expression pattern in adult tissue, and we speculate that it may play a role in growth regulation due to LOI in adult cancers. Similarly, the two novel *INSIGF*-imprinted transcripts generated from this important locus also merit further study for their possible role in normal development and disease.

## MATERIALS AND METHODS

### Collection of human material

A total of 18 fetal tissue sets (8–18 weeks) with first trimester placental samples and corresponding maternal blood samples were obtained from the termination of pregnancies at Queen Charlotte's and Chelsea Hospital. Samples were washed in sterile PBS and processed. Local ethical approval for obtaining



fetal tissues was granted by the Research Ethics Committee of Hammersmith, Queen Charlottes's and Chelsea and Acton Hospitals Research Ethics Committee (2001/6028). Thirty placental trio samples consisting of multiple site placental tissues with corresponding maternal and paternal blood samples were collected from consecutive consenting pregnancies at Queen Charlotte's and Chelsea Hospital (local ethics approval 2001/6029). Twenty histologically normal adult pancreatic samples (20–46 years) were obtained from the MRC tissue bank at Hammersmith Hospital and University Hospital Heidelberg. The collection of pancreatic samples was approved by the Ethical Committee of the University of Heidelberg. Written informed consent was obtained from all patients.

### Enrichment for uncultured human villous trophoblasts

For the isolation of villous trophoblasts, which are equivalent in function to the mouse labyrinthine trophoblasts, we used a protocol of placental tissue digestion and negative immunoselection (32). Briefly, dissected tissues from both termination (8–14 weeks gestation) and term placentas were thoroughly washed in PBS and digested with a cocktail of trypsin and DNase to release free cells. Unwanted erythrocytes were removed from the resulting cell suspension by centrifugation through a 40% percoll solution. The resulting trophoblasts are subjected to negative immunoselection using monoclonal anti-HLA class 1 (clone w6/32). All trophoblast cell preparations were subjected to cytokeratin 7 immunocytochemistry (>98% cytokeratin 7 positive cells) and HLA class 1 and Vimentin to assess cell contamination (<1.3%).

### PCR amplification of genomic DNA

Genomic DNA was extracted from placental/fetal tissue samples and peripheral blood, using standard phenol–chloroform separation. All amplifications were carried out on 200 ng DNA for 32 cycles. The following PCR primers were used for genotyping: *IGF2 Alu1* F-CCTGACAAGGTGACCCTGAAG and *IGF2 Alu1* R-GCCAAAGCTGCTTGCAGAAG; *IGF2 Apa1* F-CTTGGACTTTGAGTCAAATTGG and *IGF2 Apa1* R-CTCCTTTGGTCTTACTGGG; *INS Pst1* F-TGGTGCAGG CAGCCTGCAG and *INS Pst1* R-GTTCAAGGGCTTTATT CCATCTCTC.

### RNA isolation and RT–PCR

Total RNA was isolated from homogenized tissues using Trizol (Invitrogen). RNA was treated with amplification grade DNase 1 (Invitrogen) to degrade any contaminating genomic DNA present in the sample. First-strand cDNAs were synthesized from 2 µg DNase-treated RNA with MMLV-RT and random hexamers (Promega). Duplicate sets of samples were processed with MMLV-RT omitted to detect genomic contamination of the RNA. The presence of cDNA was confirmed using *GAPDH* primer sets. For strand-specific RT–PCR, the *IGF2 Apa1* R primer was used to synthesize the first strand of cDNA. The presence of *IGF2* sense transcript was confirmed using an *IGF2* Ex 8 and Ex 9 RT–PCR.

### Imprinting assays

For the imprinting analysis of *IGF2*, *INSIGF* and *INS*, an RT–PCR approach was used as described previously (33,34). For *IGF2*, each individual isoform of *IGF2* was amplified using a promoter-specific forward primer in conjunction with a common *IGF2* 3'-UTR reverse primer incorporating the *Apa1* SNP in the PCR product. The following promoter-specific forward primers used were: P1-CTCGCACGGGTA GAGACAGG; P0 (EF)-GTAAACCCATCATTTCAGTTGG CATT; P2-CTTCAGGCTGATTTGCCCATC; P3-GGACAA TCAGACGAATTCTCC; P4-CTTCTCCTGTGAAAGAGAC TTCC and *IGF2* 3'-UTR R-CTTTTTAGGATGGGAAT TGAG. The products were subsequently sequenced using the *IGF2 Apa1* F as a nested sequence primer. The imprinting of the *IGF2*-P1 and P0 transcripts was also confirmed using the promoter-specific forward primer and an *IGF2* exon 3 R-GCTGAGCTGGCAGCGATTTCAGAG, incorporating the *Alu1* SNP in the PCR product. The imprinting of the *INS* gene utilized an *INS* Ex1 F-CTGCATCAGAAGAGGCCATC and *INS* Ex3 R-GTTCAAGGGCTTTATTCCATC. The imprinting of the shorter *INSIGF* used *INS* Ex1 F and *IGF2* Ex4 R-GACACCTTTGAAAGAACACTCCC. All RT–PCRs used HotStar Taq DNA polymerase (Qiagen), following manufacture's instructions, and products were subjected to both restriction enzyme digestion and directly sequenced in both orientations using an ABI prism 3100 DNA sequencer (Applied Biosystems).

### Northern blotting

To analyse relative expression levels of *IGF2*-P0 in adult tissues, a multiple human blot (FirstChoice 1) containing polyA + RNA (2 µg per lane) was purchased from Ambion. Custom-made fetal northern blots were obtained from Biochain. The *IGF2* RNA probes, which encompassed the P1-specific promoter region (130501–130800 of AC0130303), P0-specific promoter region (139201–139620 of AC0130303) or the coding sequence of exons 8 and 9 (155336–156732 of AC0130303) were generated as PCR products cloned into the pGEMT-easy vector (Promega). Single-stranded sense and antisense probes were radiolabelled with  $\alpha$ [<sup>32</sup>P]UTP using the T7/SP6 MAXIscript *in vitro* transcription kit (Ambion). Hybridizations were carried out overnight at 65°C and washed according to manufacturer's instructions. The *GAPDH* or  $\beta$ -actin PCR probe was used to confirm equal loading of RNA in each lane.

### Luciferase promoter assay

Luciferase assays for *IGF2* promoter activity were carried out as described previously (23), using the Promega Dual Luciferase kit and the luciferase pGL3 reporter vectors and Qiagen Effectine reagents. The P0 promoter construct was made by amplifying bases 139039–139638 (AC130303) using the following primers: forward 5'-GCTAGGTACCTTGCTCATGG AAGATGG-3' and reverse 5'-TCAGGATCCAGCTGAGGC TGGAGAAACG-3'. These primers contained restriction sites at their 5'-ends (Fwd, *Kpn1*; Rev, *BamH1*) so that the PCR product could be directly cloned into multiple cloning site of the pGL3 vectors.

### Methylation analysis by bisulphite sequencing

Approximately 1 µg DNA was treated with sodium bisulphite as described previously (35). For the bisulphite treatment of trophoblast DNA, ~1000 cells were resuspended in 32.5 µl of solution containing 10 µg glycogen, 1 mM SDS and 280 µg/ml proteinase K, incubated for 90 min at 37°C, followed by 15 min at 96°C in a thermocycler. This lysate was used directly for bisulphite treatment. HotStar DNA polymerase (Qiagen) was used to amplify CpG island products (annealing temperature of 58°C) using the following specific primers: CpG1 F-GTTTGGAGTTGGGGTTTGG and CpG1 R-CTACCTCAACAATCCCTAAAATC; DMRO (1)F-TTG GTGTTGGAAAGTGTTTG and DMRO (1)R-CTATAACRT CCAAACCCTCTA; DMRO (2)F-GTTAAGGTAGTTTTT TTGG and DMRO (2)R-AATAACCCRCCTTAAAAAATC (27); CpG2 F-GATTTTTGGGGGATTTTAAATTTAT and CpG2 R-CCACATCCTAAAAACCAAAC; CpG3 F-GGA GGYGGTTTTTTTTYGTAGG and CpG3 R-TATAATTAATC CACTTTAATTC; CpG4 F-GTAGGGGTTTGTGTTGTTTT TTG and CpG4 R-CTACTATACTTCTCAACCC. PCR amplification, cloning and sequencing were performed as described previously (20). In addition, all PCR products were subjected to COBRA that involves digesting with *TaqI* or *TaqI* restriction enzyme to discriminate between methylated and unmethylated CpG dinucleotides to control for any potential cloning bias.

### ACKNOWLEDGEMENTS

We would like to thank Sally Newman for technical advice and sequencing, Professor Eammon Maher for the pUPD11 cell line and Dr Gavin Kelsey for critical reading of the manuscript. This work was supported by The Wellcome Trust, SPARKS, IOG Trust, WellBeing and the MRC. P.A. holds a Marie Curie European Reintegration Grant (MERGT-CT-2004-510972). This paper is dedicated to the memory of William Monk.

*Conflict of Interest statement:* None declared.

### REFERENCES

- Brannan, C.I. and Bartolomei, M.S. (1999) Mechanisms of genomic imprinting. *Curr. Opin. Genet. Dev.*, **9**, 164–170.
- DeChiara, T.M., Robertson, E.J. and Efstratiades, A. (1991) Parental imprinting of the mouse insulin-like growth factor II gene. *Cell*, **64**, 849–859.
- Weksberg, R., Shen, D.R., Fei, Y.L., Song, Q.L. and Squire, J. (1993) Disruption of insulin-like growth factor 2 imprinting in Beckwith–Wiedemann syndrome. *Nat. Genet.*, **5**, 143–150.
- Gicquel, C., Rossignol, S., Cabrol, S., Houang, M., Steunou, V., Danton, F., Thibaud, N., Merrer, M.L., Burglen, L., Bertrand, A. M., Netchine, I. and Bouc Y.L., (2005) Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver–Russell syndrome. *Nat. Genet.*, **37**, 1003–1007.
- Ogawa, O., Eccles, M.R., Szeto, J., McNoe, L.A., Yun, K., Maw, M.A., Smith, P.J. and Reeve, A.E. (1993) Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. *Nature*, **362**, 749–751.
- 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.
- Grandjean, V., O'Neill, L., Sado, T., Turner, B. and Ferguson-Smith, A. (2001) Relationship between DNA methylation, histone H4 acetylation and gene expression in the mouse imprinted Igf2-H19 domain. *FEBS Lett.*, **488**, 165–169.
- Murrell, A., Heeson, S. and Reik, W. (2004) Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat. Genet.*, **36**, 889–893.
- 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 putative imprinting control region upstream of mouse *Igf2*. *Proc. Natl Acad. Sci. USA*, **94**, 12509–12514.
- Feil, R., Walter, J., Allen, N.D. and Reik, W. (1994) Developmental control of allelic methylation in the imprinted mouse *Igf2* and *H19* genes. *Development*, **120**, 2933–2943.
- Eden, S., Constancia, M., Hashimshony, T., Dean, W., Goldstein, B., Johnson, A.C., Keshet, I., Reik, W. and Cedar, H. (2001) An upstream repressor element plays a role in *Igf2* imprinting. *EMBO J.*, **20**, 3518–3525.
- Forne, T., Oswald, J., Dean, W., Saam, J.R., Bailleul, B., Dandolo, L., Tilghman, S.M., Walter, J. and Reik, W. (1997) Loss of the maternal *H19* gene induces changes in *Igf2* methylation in both *cis* and *trans*. *Proc. Natl Acad. Sci. USA*, **94**, 10243–10248.
- Bell, A.C. and Felsenfeld, G. (2000) Methylation of a CTFT-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature*, **405**, 482–485.
- 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.
- van Dijk, M.A., van Schaik, F.M., Bootsma, H.J., Holthuisen, P. and Sussenbach, J.S. (1991) Initial characterization of the four promoters of the human insulin-like growth factor II gene. *Mol. Cell. Endocrinol.*, **81**, 81–94.
- Constancia, M., Hemberger, M., Hughes, J., Dean, W., Ferguson-Smith, A., Fundele, R., Stewart, F., Kelsey, G., Fowden, A., Sibley, C. and Reik, W. (2002) Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature*, **417**, 945–948.
- Ohlsson, R., Hedborg, F., Holmgren, L., Walsh, C. and Ekstrom, T.J. (1994) Overlapping patterns of IGF2 and H19 expression during human development: biallelic IGF2 expression correlated with a lack of H19 expression. *Development*, **120**, 361–368.
- Ekstrom, T.J., Cui, H., Li, X. and Ohlsson, R. (1995) Promoter-specific IGF2 imprinting status and its plasticity during human liver development. *Development*, **121**, 309–316.
- Fowden, A.L. (2003) The insulin-like growth factors and fetoplacental growth. *Placenta*, **24**, 803–812.
- Sibley, C.P., Caon, P.M., Ferguson-Smith, A.C., Dean, W., Hughes, J., Smith, P., Reik, W., Burton, G.J., Fowden, A.L. and Constancia, M. (2004) Placental-specific insulin-like growth factor 2 (*Igf2*) regulated the diffusional exchange characteristics of the mouse placenta. *Proc. Natl Acad. Sci. USA*, **101**, 8204–8208.
- Wu, H.K., Squire, J.A., Qianli, S. and Weksberg, R. (1997) Promoter-dependent tissue-specific expressive nature of imprinting gene, insulin-like growth factor II in human tissues. *Biochem. Biophys. Res. Commun.*, **233**, 221–226.
- Lewis, M.P., Clements, M., Takeda, S., Kirby, P.L., Seki, H., Lonsdale, L.B., Sullivan, M.H., Elder, M.G. and White, J.O. (1996) Partial characterization of an immortalized human trophoblast cell-line, TCL-1, which possesses a CSF-1 autocrine loop. *Placenta*, **17**, 137–146.
- Murrell, A., Heeson, S., Bowden, L., Constancia, M., Dean, W., Kelsey, G. and Reik, W. (2001) An intragenic methylated region in the imprinted *Igf2* gene augments transcription. *EMBO Rep.*, **2**, 1101–1106.
- Haigh, T., Chen, C., Jones, C.J. and Aplin, J.D. (1999) Studies of mesenchymal cells from 1st trimester human placenta: expression of cytokeratin outside the trophoblast lineage. *Placenta*, **20**, 615–625.
- Xiong, Z. and Laird, P.W. (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.*, **25**, 2532–2534.
- Vu, T.H., Li, T., Nguyen, D., Nguyen, B.T., Yao, X.M., Hu, J.F. and Hoffman, A.R. (2000) Symmetric and asymmetric DNA methylation in the human IGF2-H19 imprinted region. *Genomics*, **64**, 132–143.
- Sullivan, M.J., Taniguchi, T., Jhee, A., Kerr, N. and Reeve, A.E. (1999) Relaxation of IGF2 imprinting in Wilms' tumours associated with specific changes in IGF2 methylation. *Oncogene*, **18**, 7527–7534.

28. Cui, H., Onyango, P., Brandenburg, S., Wu, Y., Hsieh, C.L. and Feinberg, A.P. (2002) Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. *Cancer Res.*, **62**, 6442–6446.
29. Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R. and Feinberg, A.P. (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science*, **299**, 1753–1755.
30. Moore, G.E., Abu-Amero, S.N., Wakeling, E.L., Bell, G., Wilson, A., Stanier, P., Jauniaux, E. and Bennett, S.T. (2001) Evidence that insulin is imprinted in the human yolk sac. *Diabetes*, **50**, 193–203.
31. Pugliese, A., Zeller, M., Fernandez, A., Jr, Zalcberg, L.J., Bartlett, R.J., Ricordi, C., Pietropaolo, M., Eisenbarth, G.S., Bennett, S.T. and Patel, D.D. (1997) The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDD2 susceptibility locus for type 1 diabetes. *Nat. Genet.*, **15**, 293–297.
32. Hills, F.A., Elder, M.G., Chard, T. and Sullivan, M.H. (2004) Regulation of human villous trophoblast by insulin-like growth factors and insulin-like growth factor-binding protein-1. *J. Endocrinol.*, **183**, 487–496.
33. Lighten, A.D., Hardy, K., Winston, R.M. and Moore, G.E. (1997) IGF2 is parentally imprinted in human preimplantation embryos. *Nat. Genet.*, **15**, 122–123.
34. Joyce, J.A., Lam, W.K., Catchpole, D.J., Jenks, P., Reik, W., Maher, E.R. and Schofield, P.N. (1997) Imprinting of the IGF2 and H19: lack of reciprocity in sporadic Beckwith–Wiedemann syndrome. *Hum. Mol. Genet.*, **6**, 1543–1548.
35. Arnaud, P., Monk, D., Hitchins, M., Gordon, E., Dean, W., Beechey, C.V., Peters, J., Craigen, W., Preece, M., Stanier, P., Moore, G.E. and Kelsey, G. (2003) Conserved methylation imprint in the human and mouse *GRB10* genes with divergent allelic expression suggests differential reading of the same mark. *Hum. Mol. Genet.*, **12**, 1005–1019.