Imprinted expression of the canine *IGF2R*, in the absence of an anti-sense transcript or promoter methylation

Fiona M. O'Sullivan,[a] Susan K. Murphy,[b] Lauren R. Simel,[b] Amanda McCann,[c] John J. Callanan,[c] and Catherine M. Nolan[a],*

[a]School of Biology and Environmental Science, University College Dublin, Belfield, Dublin 4, Ireland
[b]Division of Gynecologic Oncology, Duke University Medical Center, Durham, NC 27708, USA
[c]Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland
*Author for correspondence (email: Catherine.nolan@ucd.ie)

SUMMARY Imprinted genes are epigenetically modified in a parent of origin-dependent manner, and as a consequence, are differentially expressed. Although the evolution of genomic imprinting is a subject of intense debate, imprinted genes have been studied primarily in mice and humans and in a small number of marsupial mammals. Comparative studies involving rodents and primates are of limited value, as they belong to the same superordinal group of eutherian mammals (Euarchontoglires). On the other hand, comparisons involving marsupials may not be informative, due to phylogenetic distance. *Canis familiaris* belongs to Laurasiatheria, a sister-group of Euarchontoglires, and should prove useful in comparative studies of imprinted genes. Using RT-PCR we demonstrate monoallelic expression of the canine *IGF2R* in several tissues, including uterus and umbilical cord. In the case of umbilical cord, we identify the expressed allele as maternally derived. The canine *IGF2R* is thus an imprinted gene. Using bisulfite sequencing, we show that the canine *IGF2R* resembles the imprinted mouse *Igf2r* in having a CpG island in intron 2 that is hemi-methylated. However, it differs from the mouse gene in that maintenance of the monoallelic expression of canine *IGF2R* does not require expression of an anti-sense transcript from the paternally derived allele, or methylation of the repressed *IGF2R* promoter. In these two important features, the imprinted canine gene resembles the imprinted opossum *IGF2R*. Our data suggest that these features were properties of the ancestral imprinted *IGF2R* and that the anti-sense transcript (*Air*) and promoter methylation observed in mouse are derived features of the mouse *Igf2r* locus.

INTRODUCTION

Genomic imprinting is a method of gene regulation whereby an epigenetic modification or "imprint" enables cells to identify a gene as maternally or paternally derived, and to either express or repress the gene in accordance with its parental origin. Imprinting has been observed in eutherian and marsupial mammals, but not in monotremes or non-mammalian vertebrates, and is thought to have evolved over 180 Ma, in a common ancestor of therian mammals (Killian et al. 2001). Imprinted genes are functionally haploid and thus have an increased vulnerability to mutations, loss of heterozygosity, and dysregulation of imprinting mechanisms. Despite this vulnerability, genomic imprinting survives in extant orders (Weidman and Jirtle 2007). Imprinting affects a small number of mammalian genes ($<1\%$), but many of these have important roles in development (Reik and Walter 2001; Smith et al. 2006), and aberrant imprinting has been implicated in developmental and neurobehavioral disorders, and in many cancers (Reik and Walter 2001; Murphy and Jirtle 2003).

The factors that drove the evolution of imprinting are much discussed, yet imprinting has been studied in relatively few mammalian orders. To date, imprinted genes have been identified in rodents, primates (human), and artiodactyls, among placental mammals, and in a small number of marsupial species (Wilkins and Haig 2003). Comparative studies of imprinted genes have mainly focused on mice and humans and may be somewhat limited by the relative phylogenetic proximity of these species, as both are members of the same superordinal group, Euarchontoglires (Madsen et al. 2001; Murphy et al. 2001). On the other hand, comparisons of rodents and primates with marsupial mammals are not ideal either, because of the evolutionary distance between marsupial and eutherian mammals. Clearly our understanding of the evolution of imprinted genes would benefit from studies in mammals that are phylogenetically distant from rodents and primates, but closer to them than marsupial species. In this respect, the domestic dog (*Canis familiaris*) has the potential to be an extremely informative species, as the order Carnivora, to which the dog belongs, is a member of

Laurasiatheria, a sister-group to Euarchontoglires (Madsen et al. 2001; Murphy et al. 2001).

In addition to its phylogenetic position, the canine offers several advantageous features for studying the biological relevance of genomic imprinting. As companion animals, dogs share man's environment and are medically investigated to a much greater extent than any other domesticated animal. They exhibit a wide range of diseases, including many of the complex diseases of humans to which imprinted genes are thought to contribute (cancer, diabetes, heart-discase and neurological disease) (Ostrander and Wayne 2005). Because of the unique population structure of modern dogs and the availability of extensive, high quality genome sequence (Kirkness et al. 2003; Lindblad-Toh et al. 2005), the canine may be an efficient system for whole-genome association studies of these diseases (Ostrander and Wayne 2005).

To initiate the study of genomic imprinting in this emerging model organism, we examined the canine ortholog of an intensely studied imprinted gene, *IGF2R*. *IGF2R* codes for the insulin-like growth factor 2 receptor (IGF2R), a multifunctional receptor involved in lysosome biogenesis and, in viviparous mammals, in regulation of fetal growth (Nolan and Lawlor 1999; Ghosh et al. 2003). The *IGF2R* is imprinted in rodents, artiodactyls, and opossum, with predominant expression of the maternally inherited allele (Barlow et al. 1991; Killian et al. 2000, 2001), but its expression is biallelic in the vast majority of humans (Killian et al. 2001; Monk et al. 2006). Although this gene is intensely studied, many questions remain as to the nature of the imprint and the mechanism by which monoallelic expression is achieved. In particular, imprinted expression in the mouse and opossum appears to involve different mechanisms (Killian et al. 2000; Weidman et al. 2006).

In mice, imprinting of *Igf2r* involves a differentially methylated CpG island in intron 2 known as DMR2. Methylation of this region takes place during oogenesis and appears to constitute the imprint associated with this gene (Stoger et al. 1993). In its unmethylated state, DMR2 functions as a promoter for a non-coding transcript termed *Air*, in anti-sense direction to *Igf2r* (Wutz et al. 2001). The *Air* transcript, or its expression, is necessary for methylation of the promoter of the paternally derived *Igf2r* and for repression of this allele (Zwart et al. 2001; Sleutels et al. 2002). The opossum *IGF2R* does not possess a DMR2 and its monoallelic expression does not require expression of an anti-sense transcript or promoter methylation (Killian et al. 2000; Weidman et al. 2006). These differences led to suggestions that monoallelic *IGF2R* expression is achieved by different means in the two species, following divergent evolution of the locus, or alternatively that both species share an imprinting mechanism whose fundamental features have not yet been defined (Killian et al. 2000). Examination of the imprint status of the canine *IGF2R* should help in elucidating evolution at the *IGF2R* locus.

In this report, we identify the canine *IGF2R* as an imprinted gene, by showing preferential expression of the maternally inherited allele and repression of the paternally derived allele. Somewhat surprisingly, we show that the imprinted canine *IGF2R* differs from the mouse ortholog and resembles the opossum *IGF2R*, in that monoallelic expression is observed in the absence of an anti-sense transcript and of promoter methylation. The most parsimonious explanation of our findings is that the ancestral imprinted *IGF2R* did not feature a paternally derived anti-sense transcript, and that the *Air* transcript associated with the mouse *Igf2r* locus is a derived feature.

MATERIALS AND METHODS

Tissues
Umbilical cord and parental hair were collected at birth and immediately submerged in RNAlater (Ambion/Applied Biosystems, Warrington, UK) in accordance with the manufacturer's guidelines. Following incubation at 4°C for 24 h, samples were subsequently stored at -20°C. Additional tissues (uterus, brain, heart, kidney, liver, lung, and skeletal muscle) were obtained during surgery, or during routine postmortem examinations of adult dogs, and processed in RNA later as described above, or snap frozen in liquid N_2 and stored at -80°C.

Nucleic acid preparation
Genomic DNA was extracted from frozen tissues by traditional phenol–chloroform extraction methods (Sambrook et al. 1989) or using the DNeasy Tissue Kit (Qiagen GmbH, Hilden, Germany). Total RNA was isolated using Tri-Reagent (Molecular Research Center Inc., Cincinnati, OH, USA) or RNA-STAT 60 (Tel-Test Inc., Friendswood, TX, USA) in accordance with the manufacturers' instructions.

cDNA synthesis
RNA was DNAse treated using the DNA-free kit (Ambion) according to the manufacturer's guidelines. cDNA was synthesized at 42°C with Superscript II RNAse H$^-$ reverse transcriptase (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's guide, using random hexamer primers (Invitrogen). As a control for genomic DNA contamination, an identical reaction was carried out without the reverse transcriptase (-RT).

PCR amplification
Routine PCR amplifications used Platinum Taq DNA Polymerase (Invitrogen) or the Expand Long Template PCR System with Buffer 3 (Roche Diagnostics GmbH, Mannheim, Germany). For amplification of the GC-rich region toward the 5′end of the *IGF2R* cDNA, the Advantage® cDNA PCR Kit and Polymerase Mix (Clontech Laboratories, Palo Alto, CA, USA) was used. PCR amplifications used variations of the following touchdown protocol: 94°C for 3 min, 94°C for 20 sec, 72°C for 15 sec (annealing temperature decreased by 1°C/cycle in the next six cycles), 72°C for 150 sec; followed by 30 cycles of 94°C for 20 sec, 65°C for 15 sec,

72°C for 150 sec, with a final extension of 72°C for 5 min. In some reactions, the touchdown annealing range was 70–60°C, and an annealing temperature of 60°C was used in subsequent cycles. For allelic expression studies involving exon 48 polymorphisms, the total number of cycles was 38. Allelic expression studies in intron 1 and intron 2 used a total of 50 cycles. Oligonucleotide primers were purchased from Sigma-Genosys (Haverhill, Suffolk, UK). Primer sequences used in this study are available on request.

Purification and sequencing of amplicons

PCR products were purified using the Jetquick PCR Purification Spin Kit (Genomed GmbH, Lohne, Germany) in accordance with the manufacturer's guidelines, and sequenced commercially (Macrogen Inc., Seoul, Korea; MWG-biotech, Ebersberg, Germany) using internal primers (sequences available on request).

Canine *IGF2R* cDNA sequence

A search of the NCBI EST database revealed two canine ESTs (*BQ091802* and *BQ839605*), with similarity to the 3′end of the mouse and human *IGF2R* transcripts. The sequences of these ESTs, in conjunction with cross-species *IGF2R* primers (Killian et al. 2001), and the canine genome sequence publicly available at the Broad Institute (www.broad.mit.edu/mammals/dog/) and the UCSC genome browser (http://genome.ucsc.edu), were used to amplify overlapping fragments of the canine *IGF2R* from canine liver cDNA or from genomic DNA (gene-specific primer sequences available on request). Amplified fragments were sequenced in both directions and 9032 bp of canine *IGF2R* mRNA sequence (AY965264) was confirmed in this way.

DNA methylation analysis

CpG islands were identified using the USC Norris Comprehensive Cancer Center's web-based program for CpG Island Searcher (http://www.uscnorris.com/cpgislands2/cpg.aspx) based on published criteria (Takai and Jones 2002). Genomic DNA was treated with sodium bisulfite as described previously (Huang et al. 2006) to convert all unmethylated cytosines to uracils, leaving methylated cytosines unaffected. PCR was used to generate amplicons for nucleotide sequencing, using primers designed to the bisulfite-modified sequence and that either were devoid of CpG dinucleotides or incorporated a single degenerate base where necessary. PCR primers for the *IGF2R* promoter yielded a 215 base pair (bp) amplicon and were as follows: forward, 5′-TTT TTG GTT TGT AGT TTA GTT TTY GTT TTG G -3′ and reverse, 5′- ACC AAC AAC AAC AAC TAC AAC AAA ACC -3′. PCR was performed using 30–40 ng of bisulfite modified template DNA (assuming 100% recovery) and Platinum Taq DNA polymerase (Invitrogen) in 12.5 μl reaction volumes. Cycling conditions were 3 min at 94°C followed by 40 cycles of 94°C for 30 sec, 66°C for 30 sec, and 72°C for 30 sec, followed by a 72°C extension for 5 min. PCR primers used for the CpG island within intron 2 generated a 366 bp amplicon: forward, 5′- GGA GTT GGT YGT GGG TTT GG-3′ and reverse, 5′- CCC AAA ACT CCC RAA ACT CCA C-3′. Cycling conditions for the intron 2 CpG island were 3 min at 94°C followed by 35 cycles of 94°C for 30 sec, 63°C for 30 sec, and 72°C for 30 sec and a final extension at 72°C for 5 min.

Amplicons were resolved on 2% agarose gels, purified using Sigma GenElute gel purification columns (Sigma-Aldrich, St. Louis, MO, USA) and sequenced using the ThermoSequenase Radio-labeled Terminator Cycle Sequencing Kit (USB Corporation, Cleveland, OH, USA) with the reverse primer listed above (35 cycles of 95°C for 30 sec, 66°C for 30 sec, 72°C for 60 sec) for the promoter region and with primer 5′-GGG AGG AYG GTT AGG TTG G-3′ for intron 2 (35 cycles of 95°C for 30 sec, 63°C for 30 sec, 72°C for 60 sec). Sequencing products were resolved on acrylamide sequencing gels, dried, and exposed to a phosphor screen before scanning using the Molecular Dynamics Storm Phosphorimaging System (GE Healthcare, Pittsburgh, PA, USA).

Sequence analysis

Canine genomic sequence at Broad Institute (www.broad.mit.edu/mammals/dog/) and in the UCSC genome browser (http://genome.ucsc.edu, May 2005 assembly) was used for identification of the *IGF2R* exon–intron structure. For prediction of IGF2R signal peptide, the prediction program at www.cbs.dtu.dk/services/SignalP/ was used. Sequences were compared using the ClustalW program at www.ebi.ac.uk and BLAST at www.ncbi.nlm.nih.gov/. CpG islands were identified at http://www.uscnorris.com/. For identification of tandem repeats, we used http://tandem.bu.edu/trf/trf.html as described in Benson (1999).

RESULTS

The canine *IGF2R* is an imprinted gene

The canine *IGF2R* is located on dog chromosome 1 and, similar to its orthologs in other mammals, consists of 48 exons (Liu et al. 1995; Killian and Jirtle 1999). It gives rise to an mRNA (AY965264) of at least 9 kb that encodes a protein of 2497 amino acids, the IGF2R. The predicted mature canine IGF2R is very similar to the receptors in opossum, mouse, human, and bovine (72%, 83%, 85%, and 80% amino acid identity, respectively). In order to determine the allelic expression status of the canine *IGF2R*, we re-sequenced exon 48 in a number of individuals and identified a G/T SNP and an AG INDEL in the *3′UTR* (nt 7768 and 8593, respectively, of AY965264; sequenced in reverse in Fig. 1). We identified individuals heterozygous for at least one of these polymorphisms and isolated RNA from these animals. Following cDNA preparation, the regions surrounding the polymorphisms were amplified by PCR and the amplicons sequenced. To ensure that amplification was specific for cDNA and did not represent contaminating genomic DNA, we used a primer pair that spanned intron 47, and included control reactions in which reverse transcriptase was omitted (-RT). Initial experiments examined cDNAs from adult uterus and neonatal umbilical cord (uterus, $n = 4$; umbilical cord, $n = 9$), and in both tissues and all individuals, expression of *IGF2R* was monoallelic (as shown by representative individuals in Fig. 1), strongly suggesting that the canine *IGF2R* is an imprinted gene.
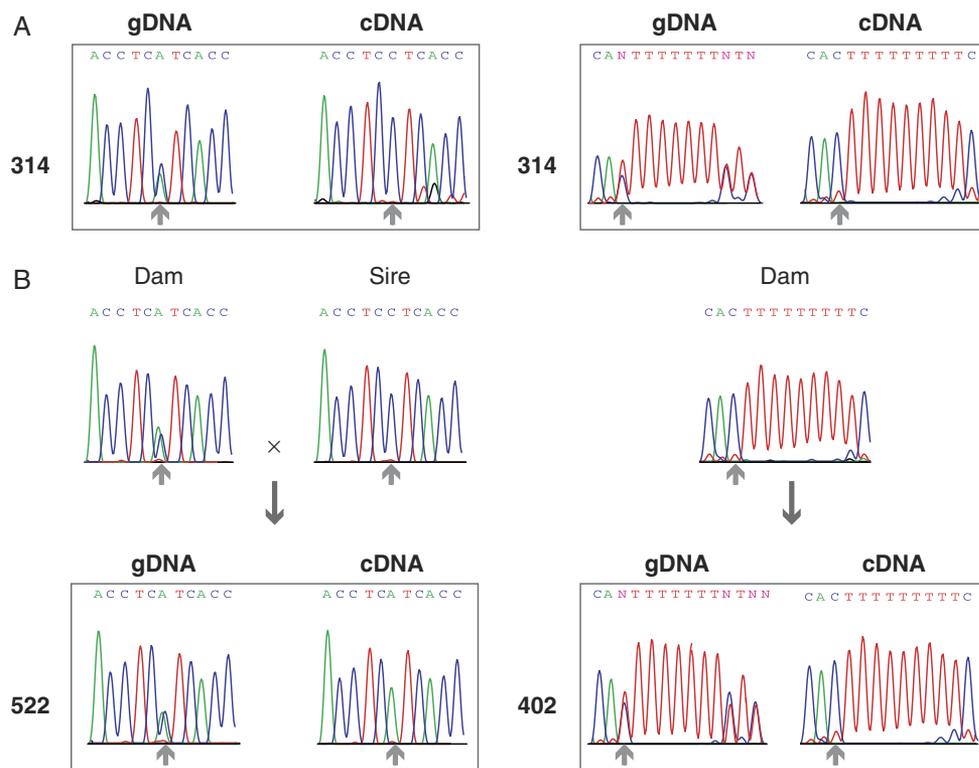
Fig. 1. The canine *IGF2R* is an imprinted gene. (A) Monoallelic expression in the uterus. Dog 314 is heterozygous for two polymorphisms in exon 48 (gDNA, arrowheads; G/T SNP and AG INDEL) (both sequenced in reverse). For both polymorphisms, only one allele is represented in the cDNA. (B) Expression of maternally derived *IGF2R* in umbilical cord. Pups 522 and 402 are heterozygous for exon 48 G/T SNP or exon 48 AG INDEL, respectively (gDNA, arrowheads; both sequenced in reverse). The sire of pup 522 is homozygous for the G allele, while the dam of pup 402 is homozygous for the AG insertion allele. Thus, parental genotypes identify the expressed allele in both offspring as maternally derived (gDNA, genomic DNA).

To determine the parental origin of the expressed allele, we genotyped the parent(s) of heterozygous pups who had contributed umbilical cord samples. Two independent parent/offspring combinations were informative, in that at least one parent was homozygous for the polymorphism for which the pup was heterozygous. The genotype of the parents clearly identifies the expressed allele in umbilical cord as maternally derived (shown for two offspring/parent combinations in Fig. 1B). Thus the canine *IGF2R* is an imprinted gene, with expression and repression, respectively, of maternally and paternally derived alleles. This allelic expression pattern is similar to that seen with the imprinted *IGF2R*s of opossum, rodents, and artiodactyls (Barlow et al. 1991; Mills et al. 1998; Killian et al. 2000, 2001), and contrasts with the situation in humans where biallelic *IGF2R* expression is apparent in the vast majority of individuals (Killian et al. 2001; Monk et al. 2006).

For some imprinted genes, the degree of allele repression varies according to tissue type, or shows temporal variation with repression restricted to certain developmental stages. For example, both alleles of murine *Igf2r* are expressed early in development, before implantation, and the paternal allele is subsequently repressed in all somatic tissues except for the brain, where the *Igf2r* expression remains biallelic (Wang et al. 1994; Lerchner and Barlow 1997; Hu et al. 1999). Biallelic *Igf2r* expression has also been demonstrated in primary cultures of mouse neurons, whereas expression in primary cultures of glial cells and fibroblasts is monoallelic (Yamasaki et al. 2005). To see whether the extent of monoallelic expression of the canine *IGF2R* varies with tissue type, we examined expression in heterozygous adult individuals for which a range of tissues was available. Although there was some variation in the expression patterns from individual to individual, expression in the brain and the kidney ($n = 2$) was biallelic, whereas expression in other tissues ($n = 2$ or 3) was predominantly or exclusively monoallelic (Fig. 2). We did not have genomic DNA from the parents of these individuals and were thus unable to confirm the parental origin of the expressed allele. However, for a given individual, the same allele was predominant in each tissue, and we assume that this represents the maternally derived allele.
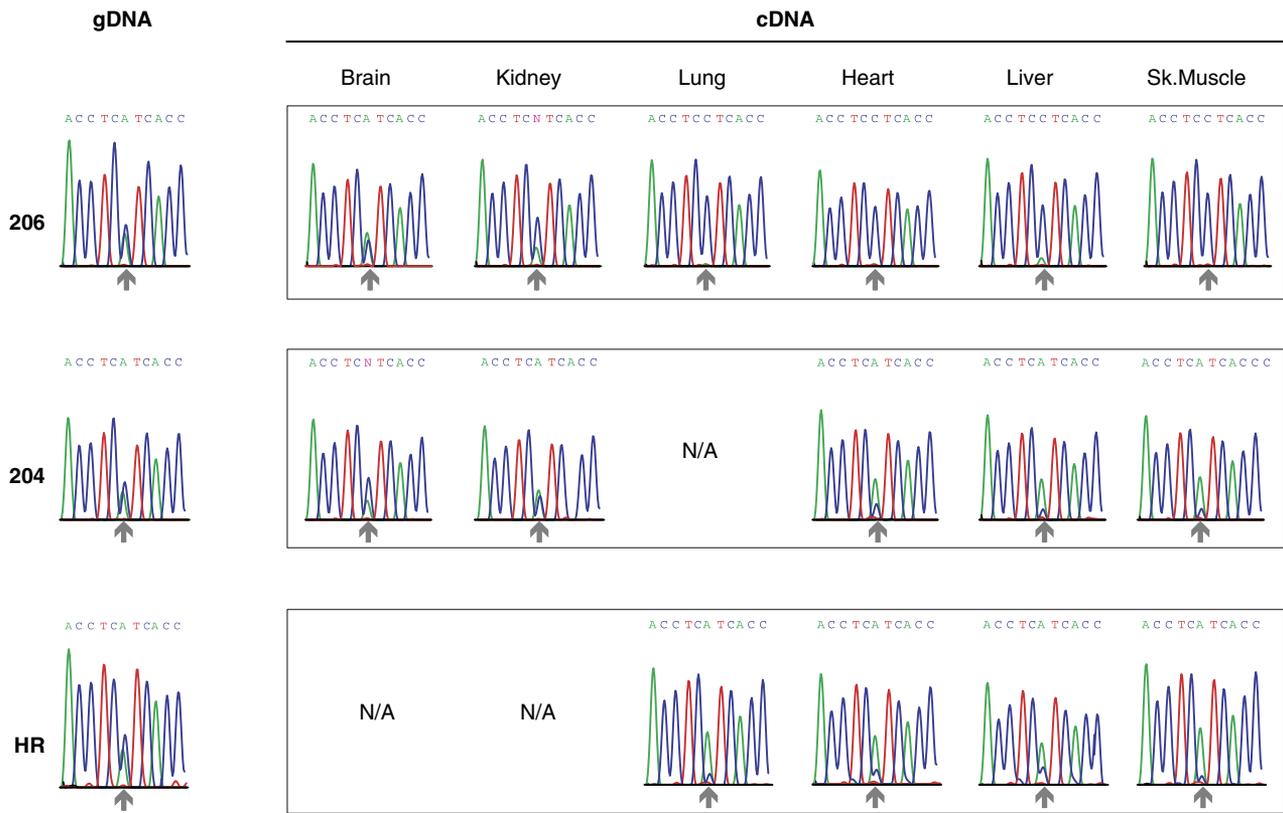
Fig. 2. Tissue-specific differences in the allelic repression of canine *IGF2R*. Allelic expression of *IGF2R* was analyzed in tissues of three dogs, each heterozygous for the exon 48 G/T SNP (arrowhead). Expression in brain and kidney is biallelic, whereas expression in lung, heart, liver and skeletal muscle is predominantly, or exclusively, monoallelic. Some individual variation in extent of repression is apparent. Dog 206 shows essentially monoallelic expression in lung, liver, skeletal muscle, and heart, while dogs 204 and HR1 show preferential expression of one allele, but the second allele is expressed to a variable extent in these tissues. All amplicons were sequenced in reverse (gDNA, genomic DNA; N/A, not available).

Canine *IGF2R* contains a potential DMR2

Having identified the canine *IGF2R* as an imprinted gene, we investigated the mechanisms involved in achieving monoallelic expression, to see whether the canine *IGF2R* resembles the mouse gene or the opossum ortholog. Intron 2 of the canine *IGF2R* contains a 2923 bp CpG island (chromosome l, 52228773–52231695, May 2005 assembly; %GC = 72.6; CpG_{obs}/CpG_{exp} = 0.91). The canine CpG island shows no sequence similarity to mouse *DMR2* or to the CpG islands in intron 2 of human, sheep, or bovine *IGF2R*s. In particular, the canine CpG island does not contain a sequence similar to a putative imprinting control element demonstrated in mouse *Igf2r* (Birger et al. 1999). However, the canine CpG island contains several tandem repeats (Table 1). Tandem repeat arrays are significantly enriched in CpG islands associated with many imprinted genes (Hutter et al. 2006) and are present in mouse DMR2 (Smrzka et al. 1995), although their significance is not yet clear. To see whether the canine intron 2 CpG island is methylated, we performed bisulfite DNA sequencing of a portion of the CpG island (beginning 1805 bp from the beginning of the CpG island). We found hemimethylation of all the CpGs in this region (Fig. 3). Methylation was observed in all tissues examined, and in all individuals (uterus, $n = 7$; umbilical cord, $n = 2$; kidney, $n = 8$; brain, $n = 3$; liver, $n = 2$; lung, $n = 2$; blood, $n = 1$; skin, $n = 1$; heart, $n = 1$). Although our analysis does not allow us to specifically associate the observed methylation with the maternally derived allele, it seems likely that the canine *IGF2R* has a DMR2 similar to that seen in mice and artiodactyls (Zwart et al. 2001; Young et al. 2003; Long and Cai 2007), and that this could constitute the canine *IGF2R* imprint.

No evidence for a putative canine *AIR* transcript

In addition to constituting the putative *Igf2r* imprint, murine DMR2 is required for monoallelic expression of *Igf2r*: in its unmethylated state, it generates a non-coding anti-sense transcript (*Air*) that is necessary for repression of the paternally derived *Igf2r* (Wutz et al. 2001; Zwart et al. 2001; Sleutels et al. 2002). The structural similarities between the canine and mouse intron 2 CpG islands suggested that the canine CpG

**Table 1. Tandem repeats in canine *IGF2R* intron 2 CpG island**

	Repeat sequence[1]	Motif length (nt)	Number of repetitions
1	GGCCTGGCGCGCAGGGTCTGCAGGACCCGGCGCGGCCTGGCCTGT	90	2.9
	GGAGCGTGGCCGGGTCTGGCGGGTACGGGCGAGCGCGCCCTGGCG		
2	GGTCTGGCGGGCCCGGGTGAGTGCAGCCT	29	2.1
3	GCGCGGTCTGGCGGACCCAGCG	22	2.8
4	TGGTCTGGCGGGCCCGGCGCGCGCAGCC	28	3.6
5	GGTCCGGCGGGCCCGGCGCGCGCAGCCT	55	2.4
	GGTCTGGAGGACCCGGCCGGGAGCCGC		
6	CTGGAGGACCCGGC	14	3.0
7	GGCCTGGCGCGCGCGGCCGGGTTGGCAG	56	2.3
	GACCCGGCCTGGAGACCGCGATCTGGAG		
8	CCCGGCGCGACCCGGAGGA	19	2.5
9	GCGCGGTCGGGCGGGCCCG	19	3.1
10	GGCGGGCCT	9	5.2
11	GGTCTGGCGGGCCCGGCGCGC	21	2.9
12	CGCTGTCCGCGGGG	14	2.6

[1]Repeats identified using http://tandem.bu.edu/trf/trf.html as in Benson (1999).

island could also function as a promoter for an anti-sense transcript. It was important therefore to test for the presence of a putative canine *AIR* transcript in tissues where canine *IGF2R* is monoallelically expressed.

In mouse, transcription of *Air* begins in DMR2 and continues for 108 kb in an anti-sense direction to *Igf2r*, terminating in the last intron of the nearest upstream gene, the *Mas1* oncogene (Wutz et al. 1997; Lyle et al. 2000). In the dog, *MAS1* and *IGF2R* are neighboring genes on chromosome 1. However, the intergenic region in the canine contains a predicted novel gene (http://genome.ucsc.edu/Genescan 1_10.79) that could complicate expression studies in this region, and we therefore focused our search for a putative canine *AIR* in the region where it would overlap with *IGF2R*. Our strategy was to identify polymorphisms in this region that would enable us to detect expression from the paternal allele, in tissues such as uterus and umbilical cord, where *IGF2R* is monoallelically expressed. We identified three SNPs in intron 2 of the canine *IGF2R* (C/T SNPs at nt 335 and 365, and a G/T SNP at nt 883, respectively, of Accession No. EF437958).
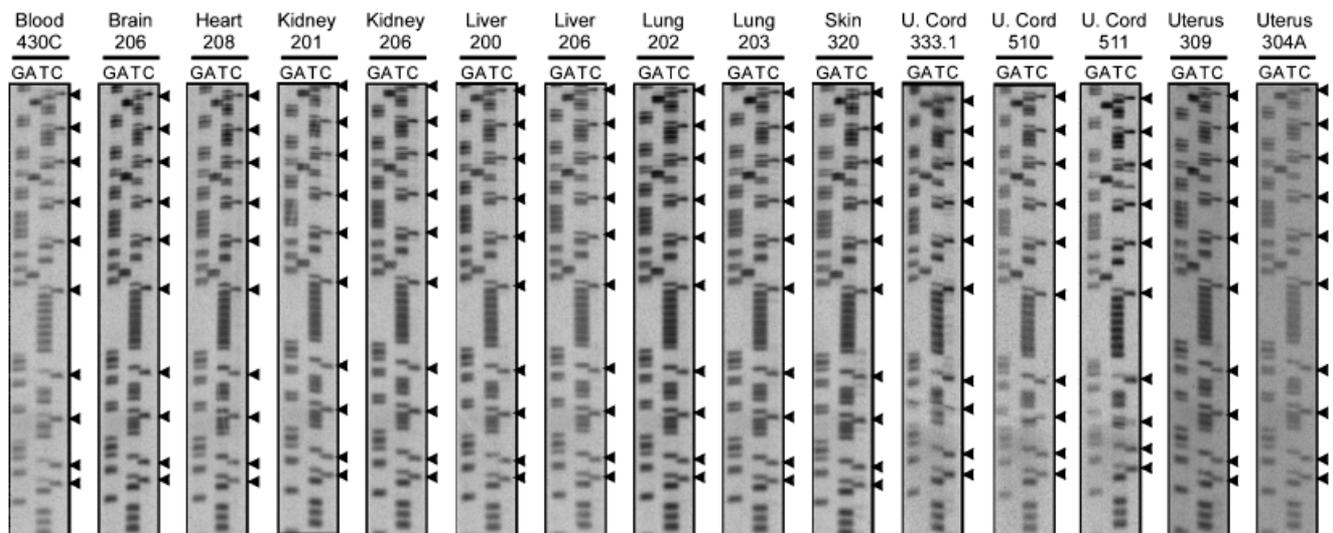
Fig. 3. Intron 2 CpG island is hemi-methylated. Genomic DNA isolated from the indicated tissues and individuals was subjected to bisulfite sequencing. PCR amplicons were generated from sequence located within the intron 2 CpG island. Lanes are loaded in the order indicated, with cytosines in the CpG dinucleotide context marked by arrowheads. Cytosines that are methylated in gDNA are unconverted by bisulfite treatment and appear in "C" lane, while unmethylated cytosines are converted and appear in "T" lane (Cs are migrating slightly faster than the corresponding Ts) (U. Cord, umbilical cord).
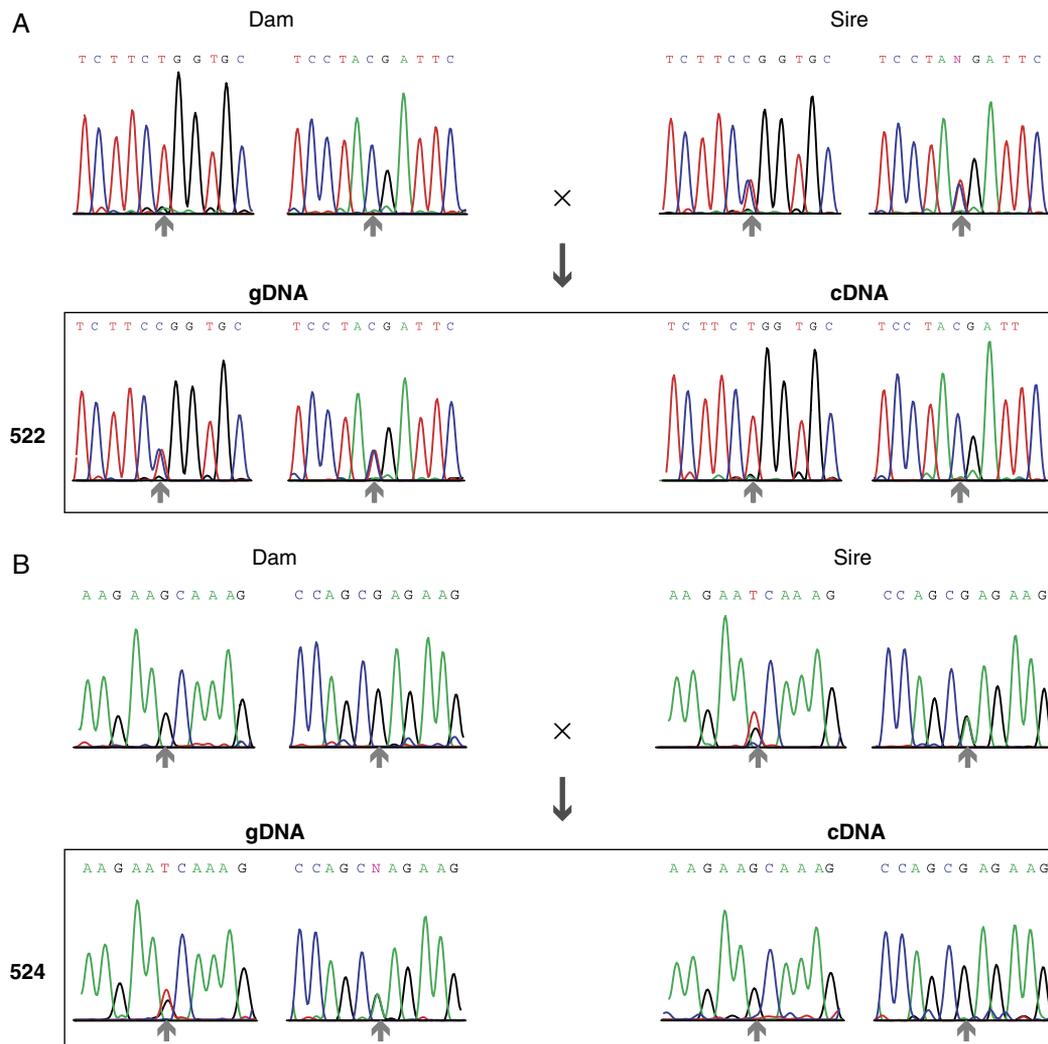
Fig. 4. Allelic expression in the putative *IGF2R/AIR* overlap region. (A) Pup 522 is heterozygous for two polymorphisms in intron 2 of the *IGF2R* (gDNA, arrowheads; C/T SNPs at 335 and 365 of Accession No: EF437958). Amplification of random hexamer-primed umbilical cord cDNA indicates expression of only one allele. The genotype of the dam, homozygous for both SNPs, identifies the expressed allele as maternal in origin. (B) Pup 524 is heterozygous for two polymorphisms in intron 1 (gDNA, arrowheads; A/C SNP and C/T SNP at 1162 and 1306, respectively, of Accession No: EF437959, both sequenced in reverse). Amplification of umbilical cord cDNA indicates expression of one allele only. The genotype of the dam identifies the expressed allele as maternal in origin. Thus, in two putative *IGF2R/AIR* overlap regions there is no evidence of a transcript derived from the paternally inherited allele (gDNA, genomic DNA).

RNA from individuals heterozygous at one or more of these sites was DNAse treated and used to generate cDNA. We used random hexamers to prime reverse transcription (so that both sense and potential anti-sense transcripts would be represented in the cDNA) and included -RT controls to confirm the absence of genomic DNA. PCR amplification of the uterus and umbilical cord-derived cDNAs (uterus, $n = 4$; umbilical cord, $n = 9$), followed by sequencing of amplicons, demonstrated the presence of a single expressed allele in each tissue and individual (shown for a representative individual, heterozygous at two positions, in Fig. 4A). Genotyping of parent(s) of two heterozygous individuals (from independent

litters) confirmed that the expressed allele in the umbilical cord was maternal in origin (shown for one individual, in Fig. 4A). This identifies the transcript as *IGF2R* pre-mRNA. No transcription was detected from the paternally inherited allele.

We also looked further upstream in the predicted overlap region for evidence of a transcript from the paternally inherited allele, and identified four polymorphisms in *IGF2R* intron 1 (C/T SNPs at nt 206, 1026, and 1306, and an A/C SNP at nt 1162 of Accession No. EF437959). Three individuals (from two independent litters) were each heterozygous for two of these SNPs. In the three individuals, PCR amplification of random hexamer-primed umbilical

cord cDNA detected transcripts from one allele only (shown for a representative individual in Fig. 4B). The genotypes of the parents of the individual in Fig. 4B showed that the transcript was derived from the maternally inherited allele. Again, there was no evidence of transcription from the paternally derived allele.

A potential explanation for the absence of a putative canine *AIR* transcript is a lack of sensitivity of our assay. We consider this unlikely, however. *Air* has been detected in mouse (embryos and adult tissues) by RT-PCR approaches that use 35–40 cycles (Hu et al. 1999; Sotomaru et al. 2002; Pauler et al. 2005; Yamasaki et al. 2005; Seidl et al. 2006). In order to maximize the detection of a putative low-abundance canine *AIR* transcript, we used 50 cycles in our assays (routine detection of mature canine *IGF2R* transcripts involved 35–40 cycles). In mouse, the levels of *Air* relative to mature *Igf2r* vary between tissues, from 17% in adult kidney to 67% in adult lung (Seidl et al. 2006). Our RT-PCR assays routinely detected canine *IGF2R* pre-mRNA, as well as the more-abundant mature *IGF2R*, and would have detected canine *AIR*, if its relative levels were comparable with those of mouse *Air*. Another possibility is that a putative canine *AIR* is spliced in such a way that our PCR primers do not facilitate its amplification. Again we consider this unlikely: the majority of murine *Air* transcripts are unspliced (Seidl et al. 2006) and, if a putative canine *AIR* had a similar splicing pattern to that of the mouse, we would have detected it. We conclude therefore that canine *AIR* transcripts are absent in the uterus and the umbilical cord, where *IGF2R* is monoallelically expressed, and are therefore not required for maintenance of monoalleleic expression. In this property, the imprinted canine *IGF2R* differs from the mouse *Igf2r*, and resembles the opossum ortholog.
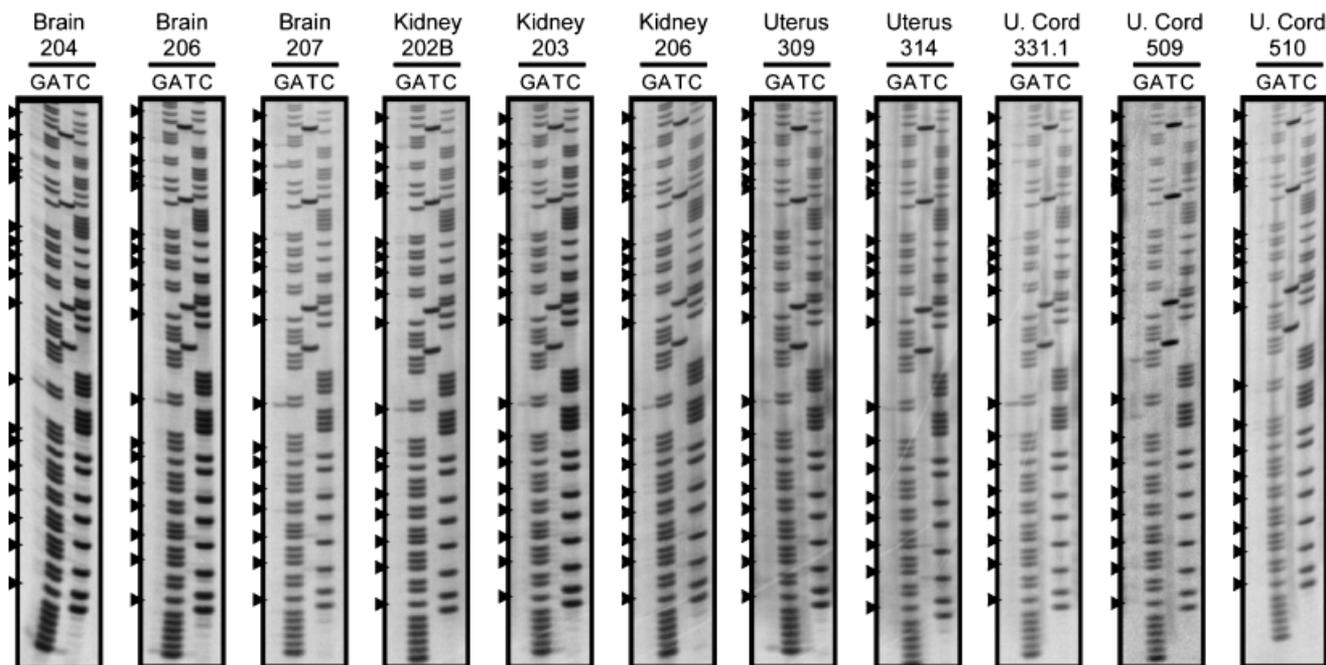
Hypomethylation of promoter-associated CpG island of canine *IGF2R*

All mammalian *IGF2R*s described to date have a CpG island associated with the promoter and first exon (Stoger et al. 1993; Riesewijk et al. 1996; Hu et al. 2006; Weidman et al. 2006). In mouse, this region is differentially methylated (methylated on the repressed allele, unmethylated on the transcribed allele) and for this reason is referred to as DMR1. Methylation at DMR1 occurs post-fertilization, following silencing of the paternally derived allele (Stoger et al. 1993), and is dependent on *Air* expression (Wutz et al. 2001; Sleutels et al. 2002). In opossum, however, the promoter-associated CpG island remains unmethylated on both alleles even though the paternally derived allele is repressed (Weidman et al. 2006). In view of the absence of a canine *AIR* transcript in tissues where *IGF2R* is monoallelically expressed, we examined the methylation status of a portion of the 2363 bp CpG island at the canine *IGF2R* promoter/exon 1 region

(chr 1, 52202808 to 52205170, May 2005 assembly; %GC = 61.6%, ObsCpG/ExpCpG = 1.012). The sequence examined began 1131 bp within the CpG island, and encompassed sequence immediately downstream of the transcription start site and spanning the translational start codon. We found a marked hypomethylation at all CpGs examined: for 17 sites quantified, the average methylation was <6.4% (Fig. 5). Hypomethylation was apparent in all tissues and individuals examined, including the uterus ($n = 6$) and the umbilical cord ($n = 6$), where expression of *IGF2R* is monoallelic, as well as in the brain ($n = 3$) and the kidney ($n = 7$), where expression is biallelic. Thus, maintenance of monoallelic *IGF2R* expression in the dog does not require methylation of the repressed promoter, and in this respect, the canine *IGF2R* resembles the imprinted opossum *IGF2R* rather than the murine ortholog.

DISCUSSION

In this study we demonstrate the existence of genomic imprinting in a representative of the order Carnivora, a mammalian order in which this epigenetic phenomenon has not previously been described. We show that the canine *IGF2R* is an imprinted gene, with the predominant expression of a single allele in most tissues. The expressed allele is maternally derived, as had been previously observed with *IGF2R* in opossum, mouse, and artiodactyls (Barlow et al. 1991; Killian et al. 2000). This contrasts with the situation in humans, where the majority of individuals express both *IGF2R* alleles (Killian et al. 2001; Monk et al. 2006). Our data support the hypothesis that *IGF2R* imprinting arose in a common ancestor of eutherian and marsupial mammals and was lost in the lineage leading to primates (Killian et al. 2001; Wilkins and Haig 2003). In addition to extending the known phylogenetic range of genomic imprinting, our demonstration of imprinting in *Canis familiaris* is timely, in that the dog is fast becoming a model system for the study of complex human diseases (Ostrander and Wayne 2005) and may therefore be useful in understanding how imprinted genes contribute to these diseases. In addition, the unique population structure of modern dogs offers significant potential for understanding how imprinted genes contribute to evolutionary processes: dogs have been subjected to strong selective pressures during the generation of the numerous "breeds" that currently exist, and exhibit more morphological diversity than any other mammal, as well as a wide range of behaviors (Lindblad-Toh et al. 2005; Ostrander and Wayne 2005). This diversity must have its origin in growth and differentiation patterns facilitated by the underlying genome and epigenome and, given the roles of imprinted genes in growth and behavior (Weidman and Jirtle 2007), they may have contributed to the generation of canine diversity. Finally, as we show here, the dog can offer insights into the evolution of imprinted genes.

Fig. 5. The CpG island associated with the *IGF2R* promoter is hypomethylated. Genomic DNA was isolated from indicated tissues and individuals and subjected to bisulfite sequencing. PCR amplicons were generated from within the CpG island, immediately downstream from the putative transcription start site and centrally located within the CpG island. Amplicons were sequenced using a reverse primer. Lanes are loaded in the order indicated, with cytosines in CpG dinucleotide context indicated by arrowheads. When sequenced in reverse, cytosines that are methylated in genomic DNA should be apparent as "G," and non-methylated cytosines as "A." The absence of a band in the G lane, together with the presence of a band in the A lane, indicates a non-methylated cytosine in native DNA (U. Cord, umbilical cord).

Although the evolution of genomic imprinting is the subject of much speculation, imprinting has been studied in relatively few mammalian orders, with demonstrations of imprinted genes limited to rodents, primates (human), and artiodactyls, among placental mammals, and to a small number of marsupial species. This restricted analysis means that little is known of the evolution of imprinted genes. We chose to study the canine *IGF2R* because of well-characterized differences in imprinting between orthologs of this gene in mouse and opossum (Killian et al. 2000; Weidman et al. 2006). These differences suggested that the monoallelic expression of *IGF2R* was achieved by different means in the two species, following divergent evolution of the locus, or alternatively that both species share an as-yet uncharacterized imprinting mechanism (Killian et al. 2000). The phylogenetic position of the dog (a eutherian mammal) but in a different superordinal group to primates and rodents (Madsen et al. 2001; Murphy et al. 2001) suggested that characterization of the canine *IGF2R* could help resolve ancestral features of this imprinted gene from derived features. Although we found some similarities between the imprinted canine and mouse *IGF2R*s, our major finding is that the canine *IGF2R* differs from the mouse *Igf2r* in two important features, and resembles the opossum ortholog.

In mice, imprinting of *Igf2r* involves a differentially methylated CpG island in intron 2 (DMR2). Methylation of this region appears to constitute the imprint associated with this gene, as it takes place during oogenesis (Stoger et al. 1993) and is required for monoallelic expression of the gene (Wutz and Barlow 1998). The canine *IGF2R* also contains a CpG island in intron 2, although it bears no sequence similarity to intron 2 CpG islands in mouse, human, sheep, or bovine *IGF2R*s. We showed that the canine intron 2 CpG island is hemi-methylated, although we were unable to determine the parental origin of the methylation, due to the absence of appropriate polymorphisms. Nevertheless, from studies of mouse, human, and bovine *IGF2R*s, it seems likely that it is oocyte derived and constitutes a canine DMR2 (Young et al. 2003; Long and Cai 2007). However, the association of DMR2 with imprinting at the *IGF2R* cluster is not straightforward, as shown by cross-species studies. The imprinted *IGF2R* in sheep and cow contains a DMR2 (Young et al. 2001, 2003; Long and Cai 2007), but there is no DMR2 in the imprinted opossum *IGF2R* (Killian et al. 2000); within Euarchonta, the non-imprinted *IGF2R*s of tree shrew and ringtail lemur lack a CpG island in intron 2 (Killian et al. 2001) whereas human *IGF2R* contains a DMR2, although the human gene is not imprinted in the majority of

individuals (Killian et al. 2001; Monk et al. 2006). Thus, DMR2 does not appear to be either necessary or sufficient for imprinting *IGF2R* (Vu et al. 2006), and the functional significance of a DMR2 in the canine *IGF2R* is unclear.

For imprinted genes, inheritance of an imprint typically results in the expression of one allele, in a parent of origin-dependent manner, and repression of the other. Having identified the canine *IGF2R* as an imprinted gene with a probable DMR2, it was reasonable to expect that its monoallelic expression would correlate with the expression of an anti-sense transcript (a putative canine *AIR* generated from the intron 2 CpG island) and with methylation of the repressed promoter. Our demonstration of the monoallelic expression of canine *IGF2R* in the absence of an anti-sense transcript originating in intron 2, or of methylation of the promoter/exon1 CpG island, was therefore somewhat unexpected. We cannot formally exclude the possibility that an *AIR* transcript is required in canine embryos for the initiation of monoallelic *IGF2R* expression (although we suggest below that this is unlikely). However, it is clear that the maintenance of the imprinted expression of canine *IGF2R* does not require the expression of an anti-sense transcript originating in intron 2. Thus, in two important features (promoter hypomethylation and absence of anti-sense transcription), imprinting of the canine *IGF2R* differs from the mouse gene and resembles the phylogenetically distant opossum *IGF2R* (Weidman et al. 2006), suggesting that these features represent properties of the ancestral imprinted *IGF2R*.

Our study raises questions about the function and evolutionary origin of anti-sense transcription at the *IGF2R* locus. The involvement of murine *Air* in the monoallelic expression of *Igf2r*, and in methylation of the repressed *Igf2r* promoter, has been well established by genetic experiments involving the targeted deletion of DMR2 or truncation of the *Air* transcript (Zwart et al. 2001; Sleutels et al. 2002). However, the molecular mechanisms by which *Air* contributes to monoallelic expression are not understood. Specifically it is not known whether *Air* is needed to induce silencing, or acts at a later stage to maintain the silent state (Seidl et al. 2006). Expression studies describe the onset of monoallelic *Igf2r* expression in mouse embryos at E6.5 (Lerchner and Barlow 1997), while the earliest report of *Air* expression is at E9.5 (Sotomaru et al. 2002; Seidl et al. 2006). These data are more consistent with a role for *Air* in maintaining, rather than inducing, *Igf2r* silencing. The imprinted opossum *IGF2R* further dissociates *AIR* expression from initiation of *IGF2R* repression, as the gene does not contain an intron 2 CpG island that could function as a promoter for an *AIR* transcript (Killian et al. 2000). Furthermore, in the very small minority of human individuals where monoallelic expression of *IGF2R* was observed, no anti-sense transcripts corresponding to murine *Air* were observed (Oudejans et al. 2001; Monk et al. 2006), in spite of the existence of a DMR2 in the human *IGF2R* (Smrzka et al. 1995).

We suggest therefore that the most parsimonious interpretation of all these expression data (mouse, human, opossum, and now dog) is that the ancestral imprinted *IGF2R* did not require an anti-sense transcript originating in intron 2 for either initiation or maintenance of monoallelic *IGF2R* expression. We further propose that the promoter of the ancestral imprinted *IGF2R* did not require methylation for maintenance of repression. In this scenario, murine *Air* is a derived feature of the locus required for maintaining the repressed state in this species. These hypotheses can be tested experimentally by extending the study of *IGF2R* to additional mammalian orders. If supported, they are consistent with the idea that all species that possess an imprinted *IGF2R* may share a common mechanism for initiation of monoallelic expression (Killian et al. 2000). The nature of this mechanism remains to be identified. Post-translational modification of histone tails can contribute to the establishment of monoallelic expression patterns, including those observed with imprinted genes (Rougelle and Heard 2002). Allele-specific histone modifications have been detected in the mouse *Igf2r* promoter and in mouse *DMR2* but were not detected in the human gene (Yang et al. 2003; Vu et al. 2004; Yamasaki et al. 2005). It is possible therefore that histone modifications are involved in the initiation of imprinted *IGF2R* expression. Characterization of the histone modifications associated with the imprinted canine and opossum *IGF2Rs* is thus warranted.

CONCLUSION

In summary, our demonstration of genomic imprinting in the canine and our observation that molecular aspects of the imprinted canine *IGF2R* resemble the phylogenetically distant opossum *IGF2R* rather than the murine *Igf2r* illustrate the value of the dog as an informative species for comparative studies of imprinted genes.

REFERENCES

Barlow, D. P., Stoger, R., Hermann, B. G., Saito, K., and Schweifer, N. 1991. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the *Tme* locus. *Nature* 349: 84–87.

Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27: 573–580.

Birger, Y., Shemer, R., Perk, J., and Razin, A. 1999. The imprinting box of the mouse *Igf2r* gene. *Nature* 397: 84–88.

Ghosh, P., Dahms, N. M., and Kornfeld, S. 2003. Mannose 6-phosphate receptors: new twists in the tale. *Nat. Rev. Mol. Cell Biol.* 4: 202–212.

Hu, J. F., Balaguru, K. A., and Ivaturi, R. D., et al. 1999. Lack of reciprocal genomic imprinting of sense and antisense RNA of mouse insulin-like growth factor II receptor in the central nervous system. *Biochem. Biophys. Res. Commun.* 257: 604–608.

Hu, J. F., Oruganti, H., Vu, T. H., and Hoffman, A. R. 2006. Tissue-specific imprinting of the mouse insulin-like growth factor II receptor gene correlates with differential allele-specific DNA methylation. *Mol. Endocrinol.* 12: 220–232.

Huang, Z., Wen, Y., Shandilya, R., Marks, J. R., Berchuck, A., and Murphy, S. K. 2006. High throughput detection of *M6P/IGF2R* intronic hypermethylation and LOH in ovarian cancer. *Nucleic Acids Res.* 34: 555–563.

Hutter, B., Volkhard, H., and Paulsen, M. 2006. Tandem repeats in the CpG islands of imprinted genes. *Genomics* 88: 323–332.

Killian, J. K., et al. 2000. *M6P/IGF2R* imprinting evolution in mammals. *Mol. Cell* 5: 707–716.

Killian, J. K., and Jirtle, R. L. 1999. Genomic structure of the mammalian *M6P/IGF2* receptor. *Mamm. Genome* 10: 74–77.

Killian, J. K., et al. 2001. Divergent evolution in *M6P/IGF2R* imprinting from the Jurassic to the quaternary. *Hum. Mol. Genet.* 10: 1721–1728.

Kirkness, E. L., et al. 2003. The dog genome: survey sequencing and comparative analysis. *Science* 301: 1898–1903.

Lerchner, W., and Barlow, D. P. 1997. Paternal repression of the imprinted mouse *Igf2r* locus occurs during implantation and is stable in all tissues of the post-implantation mouse embryo. *Mech. Dev.* 61: 141–149.

Lindblad-Toh, K., et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438: 803–819.

Liu, Z., Mittanck, D. W., Kim, S., and Rotwein, P. 1995. Control of insulin-like growth factor-II/mannose 6-phosphate receptor gene transcription by proximal promoter elements. *Mol. Endocrinol.* 9: 1477–1487.

Long, J. E., and Cai, X. 2007. *Igf2r* expression regulated by epigenetic modification and the locus of gene imprinting disrupted in cloned cattle. *Gene* 388: 125–134.

Lyle, R., et al. 2000. The imprinted antisense RNA at the *Igf2r* locus overlaps but does not imprint *Mas1. Nat. Genet.* 25: 19–21.

Madsen, O., et al. 2001. Parallel adaptive radiations in two major classes of placental mammals. *Nature* 409: 610–614.

Mills, J. J., Fall, J. G., DeSouza, A. T., and Jirtle, R. L. 1998. Imprinted *M6P/Igf2* receptor is mutated in rat liver tumors. *Oncogene* 16: 2797–2802.

Monk, D., et al. 2006. Limited evolutionary conservation of imprinting in the human placenta. *Proc. Natl. Acad. Sci. USA* 103: 6623–6628.

Murphy, S. K., and Jirtle, R. L. 2003. Imprinting evolution and the price of silence. *Bioessays* 25: 577–588.

Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryderk, O. A., and O'Brien, S. J. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* 409: 614–618.

Nolan, C. M., and Lawlor, M. 1999. Variable accumulation of insulin-like growth factor II in mouse tissues deficient in insulin-like growth factor II receptor. *Int. J. Biochem.* 31: 1421–1433.

Ostrander, E. A., and Wayne, R. K. 2005. The canine genome. *Genome Res.* 15: 1706–1716.

Oudejans, C. B., et al. 2001. Allelic *IGF2R* repression does not correlate with expression of antisense RNA in human extraembryonic tissues. *Genomics* 73: 331–337.

Pauler, F. M., Stricker, S. H., Warczok, K. E., and Barlow, D. P. 2005. Long-range DNase 1 hypersensitivity mapping reveals the imprinted *Igf2r* and *Air* promoters share cis-regulatory elements. *Genome Res.* 15: 1379–1387.

Reik, W., and Walter, J. 2001. Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.* 2: 21–32.

Riesewijk, A. M., et al. 1996. Maternal-specific methylation of the human *IGF2R* gene is not accompanied by allele-specific transcription. *Genomics* 31: 158–166.

Rougelle, C., and Heard, E. 2002. Antisense RNA in imprinting: spreading silence through *Air. Trends Genet.* 9: 434–437.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual.* 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Seidl, C. I., Stricker, S. H., and Barlow, D. P. 2006. The imprinted *Air* ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J.* 25: 3565–3575.

Sleutels, F., Zwart, R., and Barlow, D. P. 2002. The noncoding *Air* RNA is required for silencing autosomal imprinted genes. *Nature* 415: 810–813.

Smith, F. M., Garfield, A. S., and Ward, A. 2006. Regulation of growth and metabolism by imprinted genes. *Cytogenet. Genome Res.* 113: 279–291.

Smrzka, O. W., et al. 1995. Conservation of a maternal-specific methylation signal at the human *IGF2R* locus. *Hum. Mol. Genet.* 4: 1945–1952.

Sotomaru, Y., Katsuzawa, Y., Hatada, I., Obata, Y., Sasaki, H., and Kono, T. 2002. Unregulated expression of the imprinted genes *H19* and *Igf2r* in mouse uniparental fetuses. *J. Biol. Chem.* 277: 12474–12478.

Stoger, R., et al. 1993. Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. *Cell* 73: 61–71.

Takai, D., and Jones, P. A. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci. USA* 99: 3740–3745.

Vu, T. H., Jirtle, R. L., and Hoffman, A. R. 2006. Cross-species clues of an epigenetic imprinting regulatory code for the *IGF2R* gene. *Cytogenet. Genome Res.* 113: 202–208.

Vu, T. H., Li, T., and Hoffman, A. R. 2004. Promoter-restricted histone code, not the differentially methylated DNA regions or antisense transcripts, marks the imprinting status of *IGF2R* in human and mouse. *Hum. Mol. Genet.* 13: 2233–2245.

Wang, Z. Q., Fung, M. R., Barlow, D. P., and Wagner, E. F. 1994. Regulation of embryonic growth and lysosomal targeting by the imprinted *Igf2/Mpr* gene. *Nature* 372: 464–467.

Weidman, J. R., Dolinoy, D. C., Maloney, K. A., Cheng, J. F., and Jirtle, R. L. 2006. Imprinting of opossum *Igf2r* in the absence of differential methylation and *Air. Epigenetics* 1: 49–54.

Weidman, J. R., and Jirtle, R. L. 2007. Imprinted and more equal. *Am. Sci.* 95: 143–149.

Wilkins, J. F., and Haig, D. 2003. What good is genomic imprinting: the function of parent-specific gene expression. *Nat. Rev. Genet.* 4: 359–368.

Wutz, A., and Barlow, D. P. 1998. Imprinting of the mouse *Igf2r* gene depends on an intronic CpG island. *Mol. Cell Endocrinol.* 140: 9–14.

Wutz, A., Smrzka, O., Schweifer, N., Schelllander, K., Wagner, E. T., and Barlow, D. P. 1997. Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. *Nature* 389: 745–754.

Wutz, A., Theussl, H. C., Dausman, J., Jaenisch, R., Barlow, D. P., and Wagner, E. F. 2001. Non-imprinted *Igf2r* expression decreases growth and rescues the *Tme* mutation in mice. *Development* 128: 1881–1887.

Yamasaki, Y., et al. 2005. Neuron-specific relaxation of *Igf2r* imprinting is associated with neuron-specific histone modifications and lack of its antisense transcript *Air. Hum. Mol. Genet.* 14: 2511–2520.

Yang, Y., Li, T., Vu, T. H., Ulaner, G. A., Hu, J. F., and Hoffman, A. R. 2003. The histone code regulating expression of the imprinted mouse *Igf2r* gene. *Endocrinology* 144: 5658–5670.

Young, L. E., et al. 2001. Epigenetic change in *IGF2R* is associated with fetal overgrowth after sheep embryo culture. *Nat. Genet.* 27: 153–154.

Young, L. E., et al. 2003. Conservation of *IGF2-H19* and *IGF2R* imprinting in sheep: effects of somatic cell nuclear transfer. *Mech. Dev.* 120: 1433–1442.

Zwart, R., Sleutels, F., Wutz, A., Schinkel, A. F., and Barlow, D. P. 2001. Bidirectional action of the *Igf2r* imprint control element on upstream and downstream imprinted genes. *Genes Dev.* 15: 2361–2366.