Human Molecular Genetics, 2011, Vol. 20, No. 16 **3188–3197**
doi:10.1093/hmg/ddr224
Advance Access published on May 18, 2011

Methylation screening of reciprocal genome-wide UPDs identifies novel human-specific imprinted genes†

Kazuhiko Nakabayashi[1,‡,¶], Alex Martin Trujillo[3,¶], Chiharu Tayama[1], Cristina Camprubi[3], Wataru Yoshida[1], Pablo Lapunzina[4], Aurora Sanchez[5], Hidenobu Soejima[6], Hiroyuki Aburatani[7], Genta Nagae[7], Tsutomu Ogata[2], Kenichiro Hata[1] and David Monk[3,*,‡]

[1]Department of Maternal-Fetal Biology and [2]Department of Molecular Endocrinology, National Research Institute for Child Health and Development, Tokyo 157-8535, Japan, [3]Cancer Epigenetic and Biology Program (PEBC), Institut d'Investigació Biomedica de Bellvitge (IDIBELL), Hospital Duran i Reynals, Barcelona, Spain, [4]Instituto de Genética Médica y Molecular (INGEMM), CIBERER, IDIPAZ-Hospital Universitario La Paz, Universidad Autónoma de Madrid, Madrid, Spain, [5]Servei de Bioquímica i Genètica Molecular, CIBER de Enfermedades Raras, and Institut d'Investigacions Biomèdiques August Pi i Sunyer, Hospital Clínic, Barcelona, Spain, [6]Division of Molecular Genetics and Epigenetics, Department of Biomolecular Sciences, Faculty of Medicine, Saga University, Saga 849-8501, Japan and [7]Genome Science Division, Research Center for Advanced Science and Technology, the University of Tokyo, Tokyo 153-8904, Japan

Received April 11, 2011; Revised and Accepted May 13, 2011

Nuclear transfer experiments undertaken in the mid-80's revealed that both maternal and paternal genomes are necessary for normal development. This is due to genomic imprinting, an epigenetic mechanism that results in parent-of-origin monoallelic expression of genes regulated by germline-derived allelic methylation. To date, ~100 imprinted transcripts have been identified in mouse, with approximately two-thirds showing conservation in humans. It is currently unknown how many imprinted genes are present in humans, and to what extent these transcripts exhibit human-specific imprinted expression. This is mainly due to the fact that the majority of screens for imprinted genes have been undertaken in mouse, with subsequent analysis of the human orthologues. Utilizing extremely rare reciprocal genome-wide uniparental disomy samples presenting with Beckwith–Wiedemann and Silver–Russell syndrome-like phenotypes, we analyzed ~0.1% of CpG dinculeotides present in the human genome for imprinted differentially methylated regions (DMRs) using the Illumina Infinium methylation27 BeadChip microarray. This approach identified 15 imprinted DMRs associated with characterized imprinted domains, and confirmed the maternal methylation of the *RB1* DMR. In addition, we discovered two novel DMRs, first, one maternally methylated region overlapping the *FAM50B* promoter CpG island, which results in paternal expression of this retrotransposon. Secondly, we found a paternally methylated, bidirectional repressor located between maternally expressed *ZNF597* and *NAT15* genes. These three genes are biallelically expressed in mice due to lack of differential methylation, suggesting that these genes have become imprinted after the divergence of mouse and humans.

*To whom correspondence should be addressed. Tel: +34 932607500 ext. 7128; Fax: +34 2607219; Email: dmonk@idibell.cat
†Methylation array data: the data from the Illumna Infinium Human Methylation27 BeadChip microarray has been deposited with GEO database, accession number GSE28525.
‡Co-corresponding author. Tel: +81-3-3416-0181; Fax: +81-3-3417-2864; Email: knakabayashi@nch.go.jp
¶These authors contributed equally to this work.

INTRODUCTION

Genomic imprinting is an epigenetic process in which one allele is repressed, resulting in parent-of-origin specific monoallelic expression (1). To date, around 100 imprinted transcripts have been identified in mouse, including protein coding genes, long non-coding RNAs (ncRNA) and microRNAs. Approximately two-thirds show conserved imprinting status between mouse and humans, while some show imprinting restricted to humans (http://igc.otago.ac.nz/home.html).

Genomic imprinting is regulated by epigenetic modifications such as DNA methylation, along with repressive histone modifications that are transmitted through the gametes from the parental germlines (1). Many imprinted regions contain differentially methylated regions (DMRs) that exhibit parent-of-origin-dependent DNA methylation. Of the 21 known DMRs in mouse, a subset have been shown to function as *cis*-acting imprinting control regions (ICRs) orchestrating the monoallelic expression of genes over more than 100 kbp away (2). The establishment of imprinted methylation in both the maternal and paternal germlines requires the *de novo* DNA methyltransferae Dnmt3a and its related protein Dnmt3L (3,4). Maintenance of these DMRs is stable throughout somatic development and is regulated by Dnmt1 and Uhrf1 during DNA replication (5,6).

The identification of novel imprinted genes is important as it is becoming increasing evident that alterations in the fine-tuning of imprinted gene expression can influence a number of complex diseases such as obesity, diabetes, neurological diseases and cancer (7–9), in addition to the well-defined imprinting syndromes associated with severe disruption of imprinted domains.

The identification of imprinted genes has traditionally been performed in mouse owing to the ease of embryo and genetic manipulations, and has utilized gynogenetic and androgenetic embryos, or mice harboring regions of uniparental disomy (UPD), where two copies of an entire chromosome or chromosomal region is inherited from only one parent (reviewed in 10). These embryos have then been used in expression screen-based approaches such as subtractive hybridization, differential display or expression array hybridization (11,12). However, these screens are not deemed comprehensive, as imprinted gene expression can be both tissue- and developmental-stage specific. Previously, sophisticated screens have detected allelic differences in DNA methylation at imprinted DMRs present in all somatic tissues, irrespective of temporal and spatial expression. Techniques such as restriction landmark genomic screening, methylation-sensitive representation difference analysis (Me-RDA) and methylated DNA immunoprecipitation (MeDIP) have identified regions of allelic DNA methylation associated with chromosomal regions controlling several imprinted genes in mice (13–15) and humans (16,17).

In order to identify novel imprinted genes in humans, we have performed a quantitative genome-wide methylation screen comparing the methylomes of three-genome-wide paternal UPD (pUPD) samples identified with Beckwith–Wiedemann-like phenotypes and one genome-wide maternal UPD (mUPD) Silver–Russell-like syndrome case (18–21) with the methylomes of six normal somatic tissues. The genome-wide UPD samples were all mosaic, and we utilized DNA extracted from leukocytes as these presented with lowest level of the biparental cell line. The DNA methylation profiles of these samples only differ at imprinted DMRs, since they are all derived from leukocytes, making them ideal to screen for novel imprinted loci. We utilized the Illumina Infinium Human Methylation27 BeadChip microarray and were able to identify 15 imprinted DMRs associated with known imprinted transcripts, and confirm the allelic methylation within intron 2 of the *RB1* gene (22).

By comparing the methylation profiles of six somatic tissues and the genome-wide UPD cases, we identified a novel paternally methylated DMR which acts as a directional silencer resulting in the maternal expression of *ZNF597* (also known as *FLJ33071*) and *NAT15* on chromosome 16, and a maternally methylated DMRs encompassing the promoter region of the *FAM50B* retrotransposon on chromosome 6, which is paternally expressed in human tissues. Interestingly, the CpG islands of the mouse orthologues of *ZNF597*, *NAT15* and *FAM50B* are all unmethylated, resulting in biallelic expression in mid-gestation embryonic tissues.

RESULTS

Defining a hemimethylated data set

Almost all imprinted domains contain at least one region of allelic DNA methylation which is thought to regulate imprinting *in cis* (1). In order to identify new imprinted genes in humans, we performed a methyation screen of six different normal somatic tissues derived from the three germinal layers (placenta, leukocytes, brain, muscle, fat, buccal cells) and compared the data set with the methylation profiles from reciprocal genome-wide UPD samples. Genomic DNA was modified by sodium bisulfite treatment and hybridized to the Illumina Infinium Human Methylation27 platform. This array covers 27 578 CpG dinucleotides associated with 14 000 human genes. To identify novel imprinted DMRs, we took advantage of the fact that these CpG-rich sequences have a methylation profile of ~50% in all somatic tissues. We identified 78 CpG probes associated with 15 known imprinted DMRs on the array (average methylation 52%, SD 11.7) (Supplementary Material, Fig. S1). To define a range in which novel imprinted DMRs should lie, we used the mean for the known imprinted DMR ± 1.5 SD (range 34.4–69.6). After applying these defined cutoffs, we identified 3212 CpG probes for which the mean methylation value for all normal tissues was within this range. To rule out the possibility that a mean of ~52% was caused by extreme values of hyper- and hypomethylation as a result of tissue-specific methylation, we only assessed those within 1.8 times SD distance from the methylation average. This step ensures that the ~52% methylation value is representative of all tissues. Using these criteria, we reduced the data set to 1836 CpG probes, which were in addition to 72 probes mapping to known imprinted DMRs.

Determining the allelic methylation using genome-wide UPDs

To identify novel imprinted DMRs within the above hemi-methylated data set outlined earlier, we compared the tissue methylation profiles to those obtained for the samples with genome-wide UPD. Of the 1836 CpG probes, only 14 gave methylation profiles consistent with an imprinted profile (Supplementary Material, Fig. S2). We subsequently mapped the exact location of the candidate CpGs using the genomic sequence of the unconverted DNA probes in the BLAT search tool (UCSC Genome Bioinformatics http://genome.ucsc.edu/). These 14 CpG probes were located close to nine autosomal genes, *RB1* (5), *FAM50B* (2), *ZNF597* (1), *TRPC3* (1), *SYCE1* (2), *TSP50* (1), *SORD* (1) and *ZBTB16* (1). We identified five independent probes located throughout CpG 85 (the CpG island identifier on the UCSC genome browser, build GRCh37/hg19) of the recently identified *RB1* imprinted gene on chromosome 13. These probes were unmethylated with average β-values of 0.21, 0.17 and 0.18 in the three genome-wide pUPD samples but hypermethylated, having an average β-value 0.88, in the genome-wide mUPD sample (a complete unmethylated CpG has a β-value of 0, and a fully methylated dinucleotide being 1). Using bisulphite PCR incorporating the single-nucleotide polymorphism (SNP) rs2804094 and sequencing of individual DNA strands, we were able to confirm that this 1.2 kb CpG island is a maternally methylated DMR in placenta, leukocyte and kidney-derived DNA and unmethylated in sperm (Supplementary Material, Fig. S3).

We identified one probe was located close to CpG 55 of the *TRPC3* gene on human chromosome 4 that was suggestive of a maternally methylated DMR. Subsequent allelic bisulphite PCR encompassing the SNP rs13121031 revealed that this region was subject to SNP-associated methylation and not parent-of-origin methylation (data not shown). The CpG islands within the promoters of *ZBTB16*, *TSP50* and *SORD* each had one probe that was suggestive of imprinted methylation, however allele-specific bisulphite PCR analysis revealed that these regions had a mosaic methylated profile (data not shown).

Two probes mapping to CpG 124 of *SYCE1/SPRNP1* on chromosome 10 also had a methylation profile consistent with an imprinted DMR. However, these probes were unable to discriminate *SYCE1* from *SPRN*, a second region that shared 93% homology. Due to the difficulty in designing bisulphite PCR primers that could specifically target *SYCE1*, we were unable to validate our initial observations.

The *ZNF597/NAT15* CpG island is a paternally methylated DMR

To date, only seven paternally methylated DMRs have been identified, the somatic DMRs at the *NESP, IGF2-P0* and *MEG3/GTL2* promoters, the germline *H19* differentially methylated domain (DMD), *Rasgrf1* DMD, IG-DMR and *ZDBF2* DMR (15,23–26). The *RASGRF1* is not imprinted in humans due to lack of the DNA repeat elements that are involved in establishing germline methylation (27). We identify two CpG probes, one mapping to CpG 41 between the promoters of *ZNF597* and *NAT15*, the other 500 bp away, in a region flanking CpG 41. Both probes were hypermethylated in the three genome-wide pUPD samples (β-values of 0.83, 0.42, 0.75) and hypomethylated (β-value of 0.08) in the genome-wide mUPD sample. Using bisulphite PCR and subsequent sequencing of heterozygous DNA samples for the SNP rs2270499, we were able to confirm that the methylation was solely on the paternally derived allele in placenta, leukocyte and kidney (Fig. 1). This is consistent with the previous report that *ZNF597* is maternally expressed in human leukocytes (28). Bisulphite PCR and sequencing of sperm DNA revealed that this region lack methylation, indicating that CpG41 is not a germline DMR. Using allele-specific RT–PCR that incorporated coding SNPs within exon 3, we observed maternal expression in brain ($n = 1$) and placenta ($n = 3$), and confirmed imprinting in leukocytes ($n = 2$).

The gene encoding *N*-acetyltransferase 15, *NAT15*, is encoded by two different transcripts (Fig. 1A). To determine whether *NAT15* is also subject to genomic imprinting, we performed allelic RT–PCR using PCR primers that could discriminate each isoform. We find that *NAT15* isoform 1 is maternally expressed in both placenta ($n = 5$) and leukocytes ($n = 1$), whereas isoform 2 is biallelically expressed ($n = 4$) which is consistent with CpG 101 being unmethylated (Fig. 1, data not shown).

FAM50B DMR shows graduated methylation

We identified two probes mapping to a 1.7 kbp CpG island within the *FAM50B* promoter. These probes were hypermethylated in the genome-wide mUPD (average β-values of 0.86), but hypomethylation in the three pUPD samples mUPD (β-value of 0.23, 0.39, 0.31). Allelic bisulphite sequencing showed that the methylation profile of CpG 143 differs between the 5′ and 3′ ends. The 5′ region flanking the SNP rs2239713, overlapping the *FAM50B* promoter, is a maternally methylated DMR in placenta-, leukocyte- and kidney-derived DNA, while the 3′ region near rs34635612 is fully methylated on both parental alleles. Despite this methylation gradient, the *FAM50B* gene is paternally expressed in placenta ($n = 6$) (Fig. 2).

The absence of allelic methylation at the mouse orthologues *of ZNF597, NAT15* and *FAM50B* is associated with biallelic expression

To determine whether the allelic expression of the novel imprinted transcripts was conserved in mouse, we investigated the allele-specific expression using RT–PCR amplification across transcribed SNPs. Mouse tissues were derived from interspecies crosses at both embryonic day E9.5 and post-natal day 1. The *Fam50b* gene has two isoforms with alternative first exons. We could only detect expression in testis, which was derived from both parental alleles. Exon 2 of Fam50b corresponds to an X-chromosome-derived retrogene and overlaps a methylated CpG island.

The *Nat15* and *Znf597* genes share two different promoter CpG islands, CpG 35 and CpG 87 that are orthologous to the *ZNF597* DMR and the *NAT15* isoform 2 promoters, respectively. In mouse, both of these regions are unmethylated. Both *Nat15* isoforms are predominantly expressed in

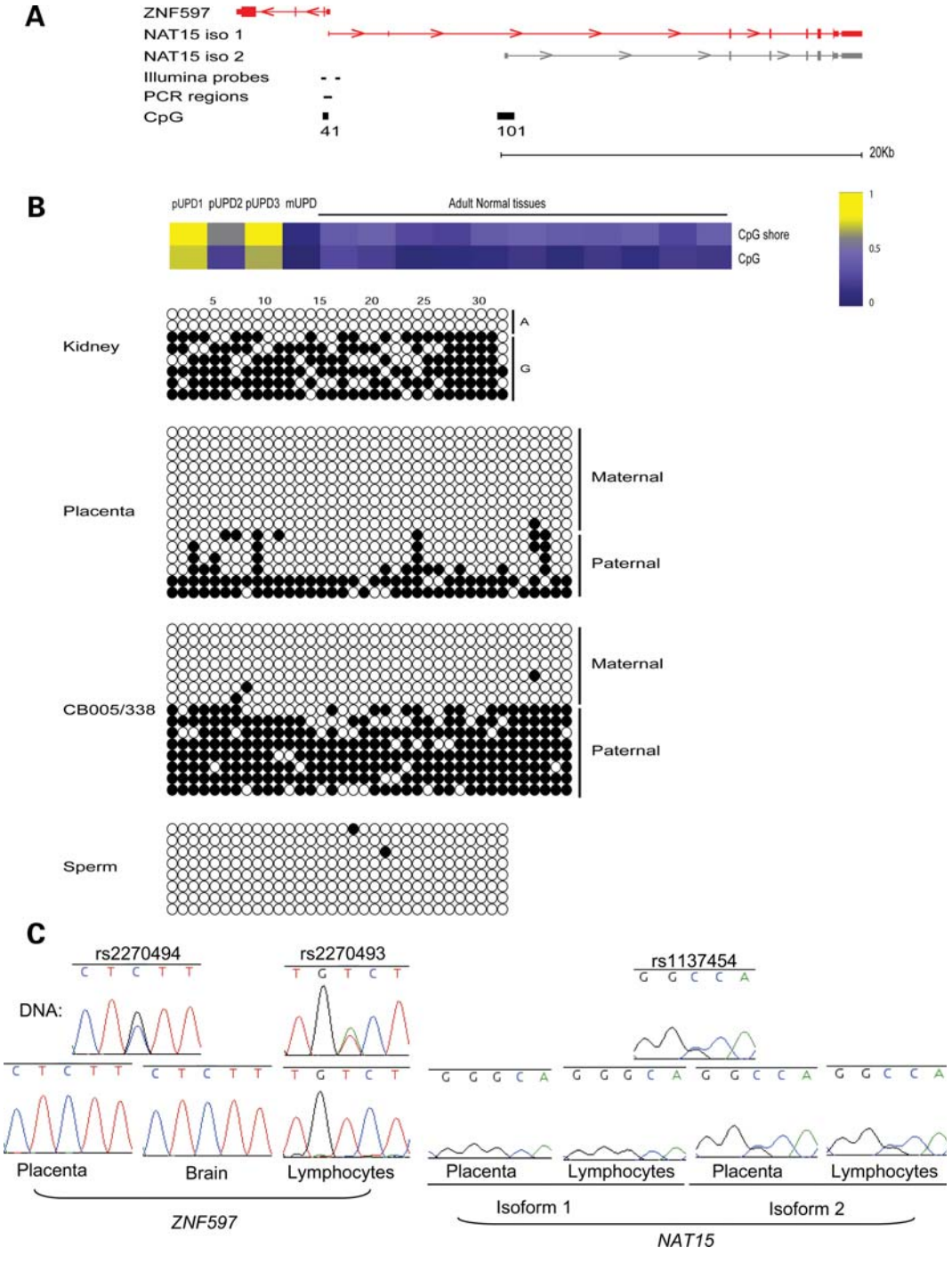

Figure 1. (**A**) Map of the *ZNF597-NAT15* locus on human chromosome 16, showing the location of the various transcripts, CpG islands, Illumina probes and bisulphite PCR regions (red transcripts are maternally expressed, blue paternally expressed and grey are expressed from both parental alleles. Arrows represent the direction of transcription) (not drawn to scale). (**B**) Heat map of the Infinium HumanMethylation27 BeadChIP for the *ZNF597* CpG probes (cg24333473 in CpG island; cg14654875 in CpG shore), with confirmation of allelic methylation in kidney, placenta and cord blood derived DNA. Each circle represents a single CpG dinucleotide and the strand, a methylated cytosine (filled circle) or an unmethylated cytosine (open circle). The same region was analyzed in sperm-derived DNA. (**C**) The sequence traces show allelic expression for the *ZNF597* and *NAT15* genes.

brain and testis, which is equally derived from both parental alleles. The variants of *Znf597* were expressed in E9.5 whole embryo, yolk sac and placenta, and in individual tissues later in development. Allelic expression analysis revealed that these transcripts were not imprinted, with equal expression from both parental chromosomes (Fig. 3).

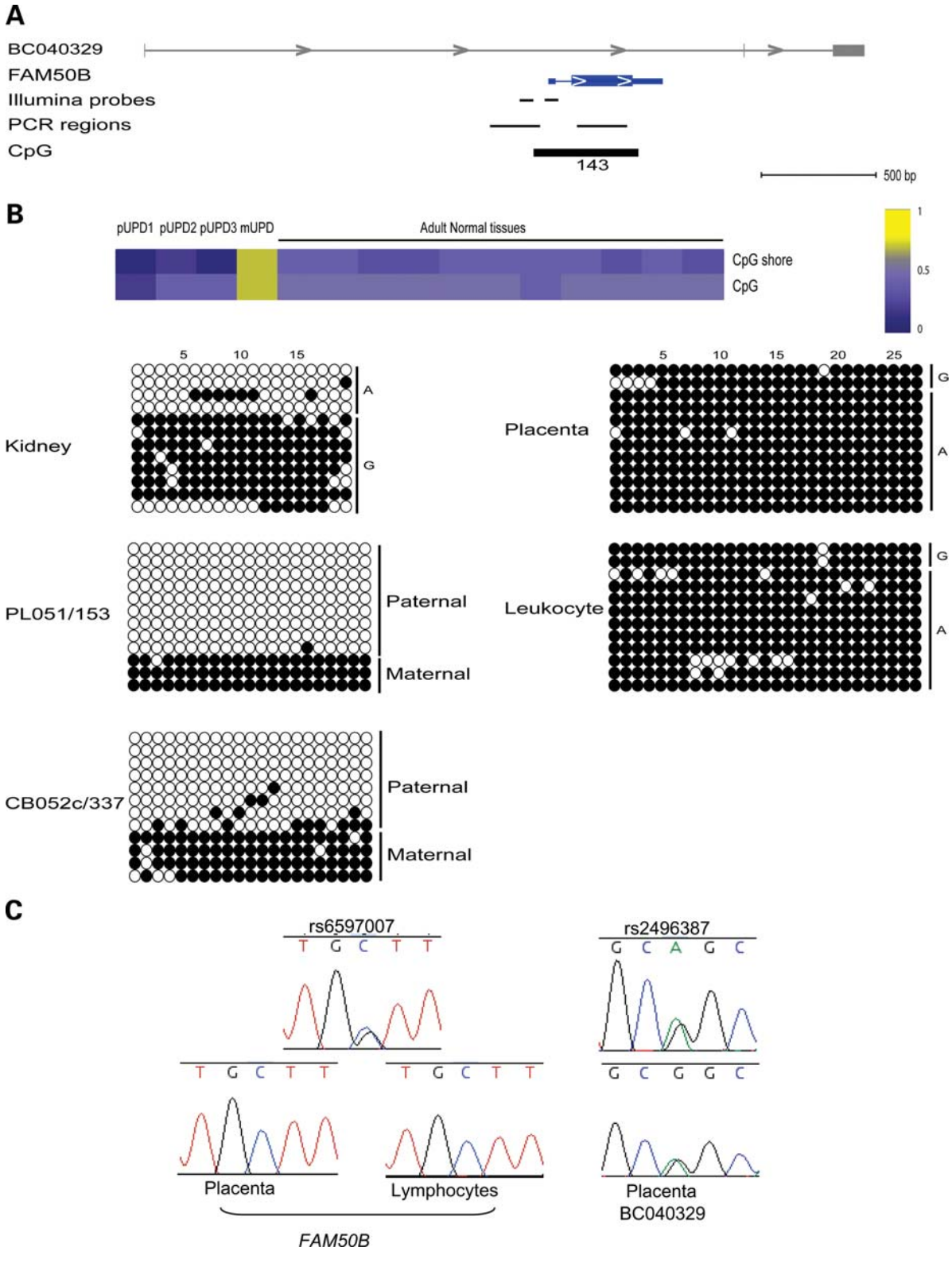

Figure 2. (**A**) A map of the *FAM50B/BC040329* locus, with the location of the of the CpG island (not to scale). (**B**) Heat map for CpG probes mapping to the *FAM50B* promoter (cg01570885; cg03202897) and the subsequent analysis of allelic methylation in various tissues. The methylation profiles on the left are from the 5′ CpG island region, while those on the right are from the 3′ region. (**C**) The allelic expression of *FAM50B* and the host gene in term placenta and leukocytes.

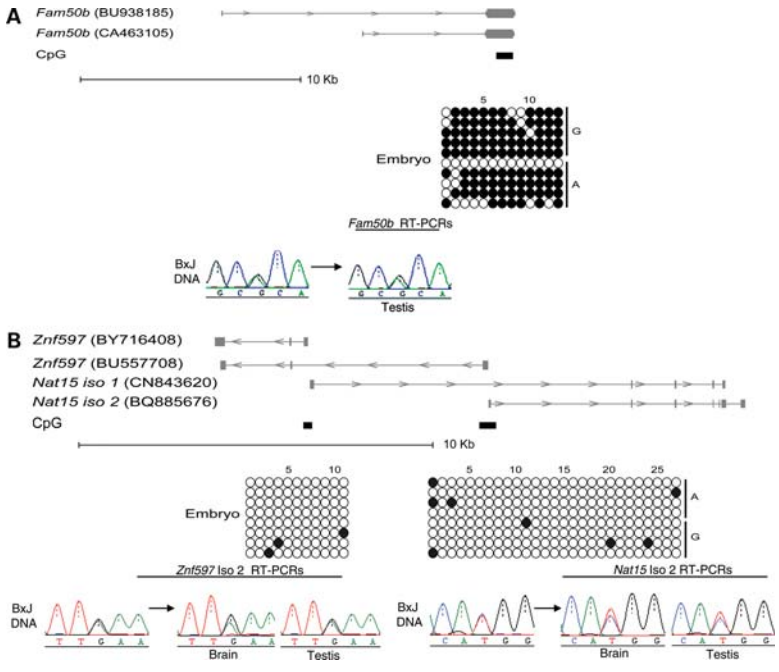

Figure 3. Schematic maps of the *Fam50b* (**A**) and *Znf597/Nat15* genes (**B**), with the location of the alternative promoter regions. The methylation status of the orthologous CpG islands associated with each domain was examined in embryo-derived DNA. The allelic expression of each gene in various mouse tissues from reciprocal mouse crosses. For clarity, only the expression in B6 × JF1 tissues is shown.

DISCUSSION

Identification of new human imprinted genes requires screening human samples

Most screens for new imprinted genes are undertaken in mouse with subsequent confirmation of the imprinting status of the human orthologues. Despite the success, this approach will not identify imprinted loci specifically imprinted in humans. To date, very few imprinted genes are human-specific, however, these rare transcripts do exist as highlighted by the paternally expressed *L3MBTL, C19MC* and *RB1* genes (22,29,30). Using DNA from Beckwith–Wiedemann and Silver–Russell-like phenotypes with reciprocal genome-wide UPDs, we have performed a comprehensive screen of ∼0.1% of the human methylome. Despite the extensive coverage of Illumina Infinium Human Methylation27 BeadChip microarray, we identified very few novel imprinted loci. However, it must be noted that paternal germline DMRs are not associated with CpG islands, and therefore maybe remote from gene promoters and promoter CpG islands present on the array.

The predicted number of imprinted genes varies with estimates from 200–2000 transcripts in mouse, with one transcriptome-wide analysis, using the ultra sensitive RNA-seq technology, identifying over 1000 transcripts in brain with parent-of-origin expression bias (31). Recent studies have predicted and experimentally verified imprinted genes based on sequence and epigenetic characteristics. For example, human imprinted regions significantly lack short interspersed transposable elements in comparison with the rest of the genome and are associated with CpG islands (32,33). Using a bioinformatics approach, Luedi *et al.* (34) predicted 156 imprinted genes in humans based on similarity with known imprinted transcripts, confirming the maternal expression of *KCNK9*. In addition, the paternally expressed *MCTS2* gene was identified through a hypothesis-driven search for intronic X-chromosome-derived retrotransposons that are associated with CpG island promoters (35). Interestingly, *FAM50B* is also an imprinted X-chromosome-derived retrogene gene and was correctly identified by Luedi *et al.* (34) during their computational screening and the imprinting status recently confirmed (36).

We wished to identify additional imprinted loci based on data generated in previously published analyses. We have compared our hemimethylated data set against the 156 bio-informatically predicted imprinted genes and the 82 candidates predicted due to unequal representation of alleles in public EST libraries and expression genotype arrays (37,38). We found that fifteen out of one hundred and fifty-six and nine out of eighty-two, respectively, were present in our data set. However, none of these additional genes had a methylation profile consistent with an imprinted DMR, highlighting the high false-positive rates of bioinformatic predictions (Supplementary Material, Fig. S4). From our observations, we predict that the majority of human DMRs overlapping promoters have been identified. Following analysis of more than 14 000 genes, we identified only two new imprinted DMRs. Extrapolating this trend to the 34 702 annotated RefSeq genes, we predict that there will be around five additional unidentified DMRs in the human genome, resulting in a total of ∼35.

Parent-of-origin DNA methylation is not the only epigenetic signature associated with imprinted DMRs (reviewed in 9). Recently, a chromatin signature has been shown to mark imprinted DMRs; with trimethylation of lysine 9 of histone

H3 (H3K9me3) and trimethylation of lysine 20 of histone H4 (H4K20me3) associated with the DNA methylated allele (39), while the unmethylated allele is enriched for the transcriptionally permissive Lysine 4 methylation of histone H3 (H3K4me2/3) (40). The combination of differential DNA methylation between sperm and somatic tissues and an overlapping H3K9me3 and H3K4me3 signature has recently been used to identify 11 new candidate DMRs in mouse (41). With the availability of human ChIP-seq derived genome-wide data sets for most histone modifications (42,43), it would be interesting to determine if this histone signature recognized in mouse can be used to identify novel human imprinted DMRs. Interrogation of the NHLBI ChIP-seq data set (http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/hgtcell.aspx) revealed that the *RB1*, *ZNF597* and *FAM50B* DMRs are enriched for both H3K4me3 and H3K9me3, with the later two regions harboring functional CTCF binding sites (data not shown).

The regulation of imprinted domains on human chromosomes 13 and 16

The *RB1* DMR has previously been proposed to contain the promoter of the paternally expressed *E2B-RB1* isoform (22). We were unable to identify coding SNPs within the *RB1* gene that would allow us to determine the allelic expression in our cohort of tissues. However, we were able to show that the *LPAR6* gene, encoding lysophosphatidic acid receptor 6 located in intron 16 of *RB1* is biallelically expressed, suggesting that the *RB1* DMR does not influence the expression of this gene (Supplementary Material, Fig. S3).

The maternal expression of *ZNF597* has previously been shown in leukocytes (28). Here, we show that the *ZNF597* DMR acts as a bidirectional silencer, which orchestrates the paternal silencing of *ZNF597* and *NAT15*. This organization is reminiscent of *PEG10-SGCE* domain on human 7p22 (44). We did not observe methylation in DNA isolated from mature sperm, which suggests that this region acquires methylation during early somatic development (Fig. 1). All known somatic DMR are associated with nearby germline DMRs, which regulate the methylation in a hierarchical fashion (23,45,46), implying a yet to be identified germline DMR is situated within the vicinity of the *ZNF597* gene.

The maternally expressed *NAT15* is a highly conserved protein coding gene with two alternative first exons, with only isoform I subject to imprinting. In addition, there is evidence from EST libraries of an ncRNA (genbank: DA387972) that originates from the *NAT15* isoform 1 promoter and continues past the exon–intron splice site to produce a ~550 bp transcript. Unfortunately, we were unable to detect expression of this transcript in our tissue set, so we could not determine if this ncRNA is imprinted.

FAM50B is an imprinted retrogene

Sequence analysis revealed that the *FAM50B* transcript (previously named *X5L*) is a retrotransposon that originated from *FAM50A/XAP5* within Xq28. Unlike other classical retrogenes, this gene has an intron in the 5′ UTR in both humans and mouse, which has no counterpart in its parental gene. It is likely that the intron was inserted after retroposition, possibly during recruitment of a functional promoter region (47). Interestingly, several other imprinted genes have been shown to originate from retrotransposition from the X-chromosome genes (35,48). *FAM50B* is ubiquitously expressed, and is inserted within the intron of a host transcript *BC040329*, which is predominantly expressed in testis with low detection in brain and placenta (data not shown). This host gene is biallelically expressed in placenta ($n = 7$) (Fig. 2), of which two samples exhibited imprinted expression of *FAM50B*.

Discrepancy between imprinted DMR methylation screens

The quantitative methylation values obtained using the Illumina Infinium platform makes it suitable for comparing reference and test samples. This approach has previously been used to screen for imprinted DMRs using paternally derived androgenetic complete hydatidiform moles versus maternally derived mature cystic ovarian teratomas and in patients with maternal hypomethylation syndrome (24,49). In both cases, the genetic material analyzed is not ideally suited for comprehensive screening for novel imprinted loci. This is because it is currently unknown to what extent the DNA methylation profile is altered in ovarian teratomas, and any differences maybe due to the uniparental nature of the sample or tumorgenic changes, and candidates obtained from comparisons with complete hydatidiform moles may simply reflect tissue-specific differences. This is highlighted by the fact that of the 95 candidate probes identified by Choufani *et al.* (49), 68 overlapped with our hemimethylated data set (Supplementary Material, Fig. S4) with only *ZNF597* DMR being identified in both screens. These authors also suggest that *AXL*-promoter region is a DMR, but this was not identified using our genome-wide UPDs, and bisulphite PCR and sequencing of our samples revealed a non-allelic mosaic methylation profile (Supplementary Material, Fig. S5). In addition, the methylation profiles obtained from comparing normal and maternal hypomethylation samples will only facilitate the identification of a subset of imprinted DMRs, since *ZFP57* mutations do not effect the maintenance of all maternally methylated imprinted DMRs equally (50,51).

Functional relevance of the new imprinted domains

Very little is known about the role of *FAM50B*, *ZNF597* and *NAT15*, with no previous publications describing functional studies. The three new imprinted regions we identify all map to chromosomes for which recurrent chromosomal UPDs have been reported. With the exception of pUPD and the overexpression of *PLAGL1/HYMAI* in Transient Neonatal Diabetes Mellitus, the UPDs for these chromosomes are not associated with obvious developmental phenotypes and most cases were identified because of the unmasking of mutant recessive alleles (reviewed in 52,53).

CONCLUSIONS

Our study has assisted in defining a comprehensive catalog of human imprinted genes. The use of extremely rare reciprocal

genome-wide UPD samples in unbiased methylation screens such as bisulphite genome sequencing will aid the identification of additional imprinted loci, which will facilitate study of genetic diseases associated with aberrant imprinting. The general trend until now has been that, while imprinted genes play an important role in fetal development and behavior, evolutionary forces dictated by the genetic conflict have allowed for a lack of conserved imprinting between mouse and humans (54). However, our screen has identified new human-specific imprinted transcripts, all of which have conserved gene orthologues in many taxa. These genes have selected imprinting as a mechanism of transcriptional regulation in humans despite the risk of being functional hemizygous.

MATERIALS AND METHODS

The human reciprocal genome-wide UPD samples

Genomic DNA isolated from three previously described Beckwith–Weidemann syndrome-like cases (16–18) and one Silver–Russell syndrome-like patient (19) was used in this study. Each of these cases had undergone extensive molecular characterization to confirm genome-wide UPD status and level of mosaicism. We used DNA isolated from leukocytes as these samples had minimal mosaicism of a biparental cell line. The genome-wide BWS samples had 9, 11 and 15% biparental contribution, whereas the genome-wide SRS sample had 16%.

Human tissues

Two independent tissue collections were used in this study. All tissues were collected after obtaining informed consent. The Spanish collection was from the Hospital St Joan De Deu tissue cohort (Barcelona, Spain). Normal peripheral blood was collected from adult volunteers aged between 19 and 60 years old. A selection of normal brain samples was obtained from BrainNet Europe/Barcelona Brain Bank. The Japanese tissues were collected at the National Center for Child Health and Development (Tokyo, Japan) and at the Saga University Hospital.

DNA was extracted using either the standard phenol/chloroform extraction method or the QIAamp DNA Blood Midi Kit (Qiagen). RNA was extracted using either Trizol (Invitrogen) or Sepasol®-RNA I Super G (Nacalai Tesque) and cDNA synthesis was carried out as previously described (54). Ethical approval for this study was granted by the Institutional Review Boards at the National Center for Child Health and Development and Saga University and Hospital St Joan De Deu Ethics Committee (Study number 35/07) and IDIBELL (PR006/08).

Cell lines and mouse crosses

Wild-type mouse embryos and placentas were produced by crossing C57BL/6 with *Mus musculus molosinus* (JF1) mice. C57BL/6 (B6) mice were purchased from Sankyo Labo Service Corporation, Inc. (Tokyo, Japan) and JF1/Ms (JF1) mice were obtained from the Genetics Strains Research Center at the National Institute of Genetics, Japan. All animal husbandry and breeding was approved and licensed by the National Research Institute for Child Health and Development, Japan (Approved number A2010–002).

Illumina Infinium methylation27 BeadChip microarray analysis

Approximately 1 μg DNA from the reciprocal genome-wide UPDs, placenta, leukocytes, brain, muscle, fat, buccal cells was subjected to sodium bisulphite treatment and purified using the EZ GOLD methylation kit (ZYMO, Orange, CA, USA). This DNA was then hybridized to the Illumna Infinium Human Methylation27 BeadChip microarray either at the Centro Nacional de Investigaciones Oncologicas (Madrid, Spain) or Genome Science Division, Research Center for Advanced Science and Technology (University of Tokyo, Japan) using Illumina-supplied reagents and protocols. The loci included on this array and the technologies behind the platform have been described previously (55). Before analyzing the methylation data, we excluded possible sources of technical biases that could alter the results. We discarded 109 probes because they had a false-positive rate >0.1. We also excluded 261 probes because of the lack of signal in one of the 11 DNA samples analyzed. Lastly, prior to screening for novel imprinted DMRs, we excluded all X chromosome CpG sites. Therefore, in total we analyzed 26 152 probes in all DNA samples. All hierarchical clustering and β-value evaluation was performed using the Cluster Analysis tool of the BeadStudio software (version 3).

Allelic methylation analysis

A panel of placenta-, leukocyte-, brain- and kidney-derived DNAs were genotyped to identify heterozygous samples. These DNA were converted using the EZ GOLD methylation kit. Approximately 100 ng of converted DNA was used for each bisulphite PCR. Bisulphite-specific primers (Supplementary Material, Table S1) which incorporate the SNPs were used with Hotstar Taq polymerase (Qiagen, West Sussex, UK). Amplifications were performed using either 45 cycles or a nested PCR using 35 cycles for each round. The subsequent PCR products were cloned into pGEM-T Easy vector (Promega) for subsequent sequencing.

Allelic expression analysis

Genotypes on DNA were obtained for exonic SNPs identified in the UCSC browser (NCBI36/hg18, Assembly 2006) by PCR. Sequences were interrogated using Sequencher v4.6 (Gene Codes Corporation, MI) to distinguish informative heterozygote samples. Informative samples were analyzed by RT–PCR. All primers, with the exception of those targeting *FAM50B*, are intron-crossing and incorporated the heterozygous SNP in the resulting amplicon (Supplementary Material, Table S1). RT–PCRs were performed using cycle numbers determined to be within the exponential phase of the PCR, which varied for each gene, but was between 32 and 40 cycles.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We thank Isabel Iglesias Platas for supplying DNA/cDNA from the Hospital Sant Joan de Deu placenta cohort and Professor Isidro Ferrer of the Barcelona Brain Bank for supplying human brain specimens. DNAs from normal adult tissues were a kind gift from Manel Esteller. We also thank Hiroko Meguro for technical assistance with the methylation array. We are especially grateful to Jose Martin-Subero for help and advice with the methylation array data analysis.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the Spanish Ministerio de Educación y Ciencia (SAF2008–1578 to D.M.); the Asociación Española Contra el Cancer (to D.M.); Fundació La Marató de TV3 (101130 to D.M. and P.L.); the Japan Society for the Promotion of the Science (to K.N., T.O., K.H.); the National Center for Child Health and Development of Japan (Grant 20C-1 to K.N., and Grant 22C-7 to K.H.). D.M. is a Ramon y Cajal research fellow (RYC-04548).

REFERENCES

1. Reik, W. and Walter, J. (2001) Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.*, **2**, 21–32.
2. Tomizawa, S., Kobayashi, H., Watanabe, T., Andrews, S., Hata, K., Kelsey, G. and Sasaki, H. (2011) Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development, 138*, 811–820.
3. Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B. and Bestor, T.H. (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science*, **294**, 2536–2539.
4. Hata, K., Okana, M., Lei, H. and Li, E. (2002) Dnmt3L cooperates with Dnmt3 family of de novo methyltransferases to establish maternal imprints in mice. *Development*, **129**, 1983–1993.
5. Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R. and Sasaki, H. (2008) Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev.*, **22**, 1607–1616.
6. Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K. *et al.* (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature*, **450**, 908–912.
7. Kong, A., Steinthorsdottir, V., Masson, G., Thorleifsson, G., Sulem, P., Besenbacher, S., Jonasdottir, A., Sigurdsson, A., Kristinsson, K.T., Jonasdottir, A. *et al.* (2009) Parental origin of sequence variants associated with complex diseases. *Nature*, **462**, 868–874.
8. Davies, W., Isles, A.R. and Wilkinson, L.S. (2005) Imprinted gene expression in the brain. *Neurosci. Biobehav. Rev.*, **29**, 421–430.
9. Monk, D. (2010) Deciphering the cancer imprintome. *Brief Funct. Genomics*, **9**, 329–339.
10. Henckel, A. and Arnaud, P. (2010) Genome-wide identification of new imprinted genes. *Brief Funct. Genomics*, **9**, 304–314.
11. Kaneko-Ishino, T., Kuroiwa, Y., Miyoshi, N., Kohda, T., Suzuki, R., Yokoyama, M., Viville, S., Barton, S.C., Ishio, F. and Surani, M.A. (1995) Peg1/Mest imprinted gene on chromosome 6 identified by cDNA subtraction hybridization. *Nat Genet.*, **11**, 52–59.
12. Kuzmin, A., Han, Z., Golding, M.C., Mann, M.R., Latham, K.E. and Varmuza, S. (2008) The PcG gene Sfmbt2 is paternally expressed in extraembryonic tissues. *Gene Expr. Patterns*, **8**, 107–116.
13. Hayashizaki, Y., Shibata, H., Hirotsune, S., Sugino, H., Okazaki, Y., Hirose, K., Imoto, H., Okuizumi, H., Muramatsu, M., Komatsubara, H. *et al.* (1994) Identification of an imprinted U2af binding protein related sequence on mouse chromosome 11 using the RLGS method. *Nat. Genet.*, **6**, 33–40.
14. Kelsey, G., Bodle, D., Miller, H.J., Beechey, C.V., Coombes, C., Peters, J. and Williamson, C.M. (1999) Identification of imprinted loci by methylation-sensitive representational difference analysis: application to mouse distal chromosome 2. *Genomics*, **62**, 129–138.
15. Hiura, H., Sugawara, A., Ogawa, H., John, R.M., Miyauchi, N., Miyanari, Y., Horiike, T., Li, Y., Yaegashi, N., Sasaki, H., Kono, T. *et al.* (2010) A tripartite paternally methylated region within the Gpr1-Zdbf2 imprinted domain on mouse chromosome 1 identified by meDIP-on-chip. *Nucleic Acids Res.*, **38**, 4929–4945.
16. Hayward, B.E., Kamiya, M., Strain, L., Moran, V., Campbell, R., Hayashizaki, Y. and Bonthron, D.T. (1998) The human GNAS1 gene is imprinted and encodes distinct paternally and biallelically expressed G proteins. *Proc. Natl Acad. Sci. USA*, **95**, 10038–10043.
17. Kamiya, M., Judson, H., Okazaki, Y., Kusakabe, M., Muramatsu, M., Takada, S., Takagi, N., Arima, T., Wake, N., Kamimura, K., Satomura, K., Hermann, R. *et al.* (2000) The cell cycle control gene ZAC/PLAGL1 is imprinted–a strong candidate gene for transient neonatal diabetes. *Hum. Mol. Genet.*, **9**, 453–460.
18. Morales, C., Soler, A., Badenas, C., Rodríguez-Revenga, L., Nadal, A., Martínez, J.M., Mademont-Soler, I., Borrell, A., Milà, M. and Sánchez, A. (2009) Reproductive consequences of genome-wide paternal uniparental disomy mosaicism: description of two cases with different mechanisms of origin and pregnancy outcomes. *Fertil. Steril.*, **92**, 393.e5–e9.
19. Yamazawa, K., Nakabayashi, K., Matsuoka, K., Masubara, K., Hata, K., Horikawa, R. and Ogata, T. (2011) Androgenetic/biparental mosaicism in a girl with Beckwith-Wiedemann syndrome-like and upd(14)pat-like phenotypes. *J. Hum. Genet.*, **56**, 91–93.
20. Romanelli, V., Nevado, J., Fraga, M., Trujillo, A.M., Mori, M.Á., Fernández, L., de Nanclares, G.P., Martínez-Glez, V., Pita, G., Meneses, H. *et al.* (2010) Constitutional mosaic genome-wide uniparental disomy due to diploidisation: an unusual cancer-predisposing mechanism. *J. Med. Genet.*, **48**, 212–216.
21. Yamazawa, K., Nakabayashi, K., Kagami, M., Sato, T., Saitoh, S., Horikawa, R., Hizuka, N. and Ogata, T. (2010) Parthenogenetic chimaerism/mosaicism with a Silver-Russell syndrome-like phenotype. *J. Med. Genet.*, **47**, 782–785.
22. Kanber, D., Berulava, T., Ammerpohl, O., Mitter, D., Richter, J., Siebert, R., Horsthemke, B., Lohman, D. and Buiting, K. (2009) The human retinoblastoma gene is imprinted. *PLoS Genet.*, **12**, e1000790.
23. Takada, S., Paulsen, M., Tevendale, M., Tsai, C.E., Kelsey, G., Cattanach, B.M. and Ferguson-Smith, A.C. (2002) Epigenetic analysis of the Dlk1-Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with Igf2-H19. *Hum. Mol. Genet.*, **11**, 77–86.
24. Yoon, B.J., Herman, H., Sikora, A., Smith, L.T., Plass, C. and Soloway, P.D. (2002) Regulation of DNA methylation of Rasgrf1. *Nat. Genet.*, **30**, 92–96.
25. Reik, W., Brown, K.W., Slatter, R.E., Sartori, P., Elliott, M. and Maher, E.R. (1994) Allelic methylation of H19 and IGF2 in the Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.*, **3**, 1297–1301.
26. Murrell, A., Ito, Y., Verde, G., Huddleston, J., Woodfine, K., Silengo, M.C., Spreafico, F., Perotti, D., De Crescenzo, A., Sparago, A. *et al.* (2008) Distinct methylation changes at the IGF2-H19 locus in congenital growth disorders and cancer. *PLoS One*, **26**, e1849.
27. Pearsall, R.S., Plass, C., Romano, M.A., Garrick, M.D., Shibata, H., Hayashizaki, Y. and Held, W.A. (1999) A direct repeat sequence at the Rasgrf1 locus and imprinted expression. *Genomics*, **55**, 194–201.
28. Pant, P.V., Tao, H., Beilharz, E.J., Ballinger, D.G., Cox, D.R. and Frazer, K.A. (2006) Analysis of allelic differential expression in human white blood cells. *Genome Res.*, **16**, 331–339.
29. Li, J., Bench, A.J., Vassiliou, G.S., Fourouclas, N., Ferguson-Smith, A.C. and Green, A.R. (2005) Imprinting of the human L3MBTL gene, a polycomb family member located in a region of chromosome deleted in human myeloid malignancies. *Proc. Natl Acad. Sci. USA*, **101**, 7341–7346.

30. Noguer-Dance, M., Abu-Amero, S., Al-Khtib, M., Lefevre, A., Coullin, P., Moore, G.E. and Cavaille, L. (2010) The primate-specific mircoRNA gene cluster (C19MC) is imprinted in the placenta. *Hum. Mol. Genet.*, **19**, 3566–3582.
31. Gregg, C., Zhang, J., Weissbourd, B., Luo, S., Schroth, G.P., Haig, D. and Dulac, C. (2010) High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science*, **329**, 643–648.
32. Greally, J.M. (2002) Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome. *Proc. Natl Acad. Sci. USA*, **99**, 327–332.
33. Ke, X., Thomas, N.S., Robinson, D.O. and Collins, A. (2002) A novel approach for identifying candidate imprinted genes through sequence analysis of imprinted and control genes. *Hum. Genet.*, **111**, 511–520.
34. Luedi, P.P., Dietrich, F.S., Weidman, J.R., Bosko, J.M., Jirtle, R.L. and Hartemink, A.J. (2007) Computational and experimental identification of novel human imprinted genes. *Gen. Res.*, **17**, 1723–1730.
35. Wood, A.J., Roberts, R.G., Monk, D., Moore, G.E., Schulz, R. and Oakey, R.J. (2007) A screen for retrotransposed imprinted genes reveals an association between X chromosome homology and maternal germ-line methylation. *PLoS Genet.*, **3**, e20.
36. Zhang, A., Skaar, D.A., Li, Y., Huang, D., Price, T.M., Murphy, S.K. and Jirtle, R.L. (2011) Novel retrotransposed imprinted locus identified at human 6p25. *Nucleic Acids Res*. **[Epub ahead of print]**, doi:10.1093/nar/gkr108.
37. Seoighe, C., Nembaware, V. and Scheffler, K. (2006) Maximum likelihood inference of imprinting and allele-specific expression from EST data. *Bioinformatics*, **22**, 3032–3039.
38. Pollard, K.S., Serre, D., Wang, X., Tao, H., Grundberg, E., Hudson, T.J., Clark, A.G. and Frazer, K. (2008) A genome-wide approach to identifying novel imprinted-genes. *Hum. Genet.*, **122**, 625–634.
39. Henckel, A., Nakabayashi, K., Sanz, L.A., Feil, R., Hata, K. and Arnaud, P. (2009) Histone methylation is mechanistically linked to DNA methylation at imprinting control regions in mammals. *Hum. Mol. Genet.*, **18**, 3375–3383.
40. Fournier, C., Goto, Y., Ballestar, E., Delaval, K., Hever, A.M., Esteller, M. and Feil, R. (2002) Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J.*, **21**, 6560–6570.
41. Dindot, S.V., Person, R., Striven, M., Garcia, R. and Beaudet, A.L. (2009) Epigenetic profiling at mouse imprinted gene clusters reveals novel epigenetic and genetic features at differentially methylated regions. *Genome Res.*, **19**, 1374–1383.
42. Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823–837.
43. Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q. *et al.* (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.*, **40**, 897–903.
44. Monk, D., Wagschal, A., Arnaud, P., Muller, P.S., Parker-Katiraee, L., Bourc'his, D., Scherer, S.W., Feil, R., Stanier, P. and Moore, G.E. (2008) Comparative analysis of the human chromosome 7q21 and mouse proximal chromosome 6 reveals a placental-specific imprinted gene, TFPI2/Tfpi2, which requires EHMT2 and EED for allelic-silencing. *Genome Res.*, **18**, 1270–1281.
45. Lopes, S., Lewis, A., Hajkova, P., Dean, W., Oswald, J., Forné, T., Murrell, A., Constância, M., Bartolomei, M., Walter, J. *et al.* (2003) Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. *Hum. Mol. Genet.*, **12**, 295–305.
46. Coombes, C., Arnaud, P., Gordon, E., Dean, W., Coar, E.A., Williamson, C.M., Feil, R., Peters, J. and Kelsey, G. (2003) Epigenetic properties and identification of an imprint mark in the Nesp-Gnasxl domain of the mouse Gnas imprinted locus. *Mol. Cell. Biol.*, **23**, 5475–5488.
47. Sedlacek, Z., Münstermann, E., Dhorne-Pollet, S., Otto, C., Bock, D., Schütz, G. and Poustka, A. (1999) Human and mouse XAP-5 and XAP-5-like (X5L) genes: identification of an ancient functional retroposon differentially expressed in testis. *Genomics*, **61**, 125–132.
48. McCole, R.B. and Oakey, R.J. (2008) Unwitting hosts fall victim to imprinting. *Epigenetics*, **3**, 258–260.
49. Choufani, S., Shapiro, J.S., Susiarjo, M., Butcher, D.T., Grafodatskaya, D., Lou, Y., Ferreira, J.C., Pinto, D., Scherer, S.W., Shaffer, L.G. *et al.* (2011) A novel approach identifies new differentially methylated regions (DMRs) associated with imprinted genes. *Genome Res.*, **21**, 465–476.
50. Mackay, D.J., Callaway, J.L.A., Marks, S.M., White, H.E., Acerini, C.L., Boonen, S.E., Dayanikli, P., Firth, H.V., Goodship, J.A., Haemers, A.P. *et al.* (2008) Hyomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. *Nat. Genet.*, **40**, 949–951.
51. Li, X., Ito, M., Zhou, F., Yougson, N., Zuo, X., Leder, P. and Ferguson-Smith, A.C. (2008) A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev. Cell*, **15**, 547–557.
52. Lapunzina, P. and Monk, D. (2011) The consequences of uniparental disomy and copy number neutral loss-of-heterozygosity during human development and cancer. *Biol. Cell*, in press.
53. Kotzot, D. and Utermann, G. (2005) Uniparental disomy (UPD) other than 15: phenotypes and bibliography updated. *Am. J. Med. Genet. A.*, **136**, 287–305.
54. Monk, D., Arnaud, P., Apostolidou, S., Hills, F.A., Kelsey, G., Stanier, P., Feil, R. and Moore, G.E. (2006) Limited evolutionary conservation of imprinting in the human placenta. *Proc. Natl Acad. Sci. USA*, **103**, 6623–6628.
55. Bibikova, M., Le, J., Barnes, B., Saedinia-Melnyk, S., Zhou, L., Shen, R. and Gunderson, K. (2009) Genome-wide DNA methylation profiling using Infinium assay. *Epigenomics*, **1**, 177–200.