Available online at www.sciencedirect.com

SCIENCE DIRECT

Genomics 87 (2006) 225 – 235

GENOMICS

www.elsevier.com/locate/ygeno

High-resolution map and imprinting analysis of the *Gtl2–Dnchc1* domain on mouse chromosome 12

Sascha Tierling [a], Simone Dalbert [a], Sandra Schoppenhorst [a,1], Chen-En Tsai [b,2], Sven Oliger [a], Anne C. Ferguson-Smith [b], Martina Paulsen [a,*], Jörn Walter [a]

[a] *Genetik/Epigenetik, FR 8.3 Biowissenschaften, Universität des Saarlandes, Postfach 151150, D-66041 Saarbrücken, Germany*
[b] *Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK*

Received 2 June 2005; accepted 27 September 2005
Available online 23 November 2005

Abstract

The imprinted *Dlk1–Dio3* region on mouse chromosome 12 contains six imprinted genes and a number of maternally expressed snoRNAs and miRNAs. Here we present a high-resolution sequence analysis of the 1.1-Mb segment telomeric to *Gtl2* in mouse and a homology comparison to the human. *Ppp2r5c* and *Dnchc1* at the telomeric end of the analyzed sequence are biallelically expressed, suggesting that the imprinted domain does not extend beyond the paternally expressed *Dio3* gene. RT-PCR experiments support the predicted presence of a maternally expressed intergenic transcript(s) encompassing *Gtl2*, *Rian*, and *Mirg*. These maternally expressed genes, and also the intergenic transcript(s), show pronounced expression in the adult mouse brain, whereas the paternally transcribed *Dio3* and the nonimprinted *Ppp2r5c* and *Dnchc1* are expressed in different tissues. Hence, tissue-specific coregulation of maternally expressed genes might be an important feature of this domain.

Keywords: Imprinting; DNA methylation; *Gtl2*; *Mirg*; *Ppp2r5c*; *Dnchc1*

To date, more than 60 genes that exhibit parental imprinting effects are known in mouse and human. An intriguing feature of imprinted genes is that they are often clustered in domains, allowing concerted allele-specific regulation of neighboring imprinted genes. Key regulatory elements are germ-line-derived differentially methylated regions (DMRs). Such an imprinted domain is located on chromosome 12 in mouse. In human and sheep, the homologous regions are also well characterized and are located on chromosome 14 and 18, respectively [1]. Uniparental disomies (UPDs) of chromosome 12 in mouse or chromosome 14 in human result in pathological phenotypes that depend on the parental origin of the disomic chromosomes [2–5].

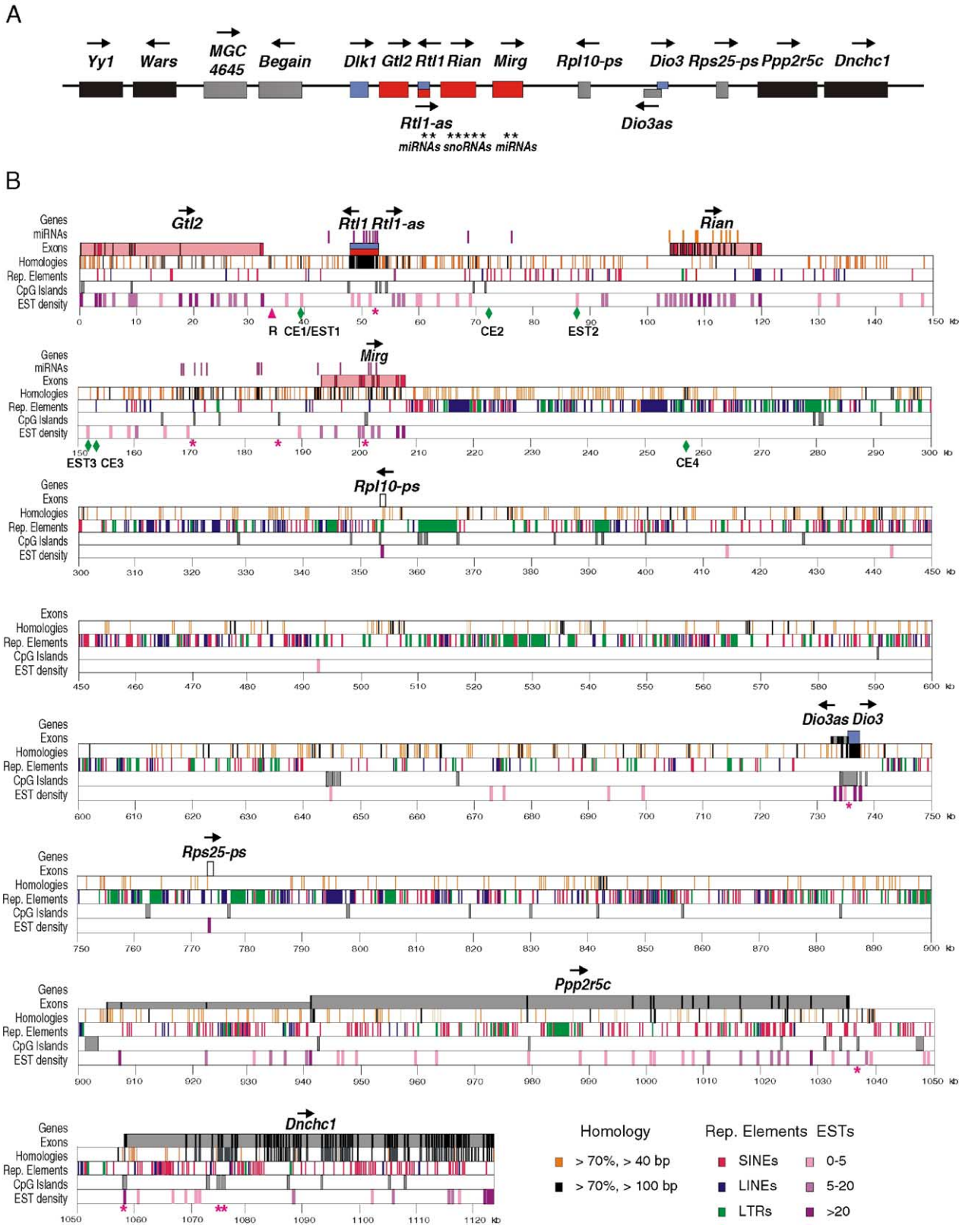
In this domain six genes have been identified to date: the paternally expressed genes Delta-like homolog 1 (*Dlk1*), retrotransposon-like 1 (*Rtl1*), and iodothyronine deiodinase 3 (*Dio3*) and the maternally expressed genes *Gtl2* (gene-trap locus 2) (*Meg3*—maternally expressed gene 3), *Rian* (RNA imprinted and accumulated in nucleus) containing multiple snoRNA genes, and *Mirg* (microRNA containing gene) (Fig. 1A). In addition, two antisense transcripts, *Rtl1-as* and *Dio3as*, overlap the *Rtl1* and *Dio3* genes [6–9]. Of these, *Rtl1-as* was shown to be maternally expressed [6,7,10].

The identified maternally expressed transcripts *Gtl2*, *Rtl1-as*, *Rian*, and *Mirg* are likely to represent nontranslated transcripts that are all transcribed in the same orientation. Three of the maternally expressed transcripts encompass the precursors for small RNAs: the *Rian* transcript contains precursors for snoRNAs, the *Rtl1-as* transcript and *Mirg* are associated with miRNAs [7,11,12]. A similar clustering of small RNAs has also been described for the imprinted Prader–Willi/Angelman syndromes region, from which snoRNAs appear to be processed from a long paternally expressed transcript that is associated with the *SNRPN* transcription unit and overlaps with the oppositely oriented maternally expressed *UBE3A* gene [13,14]. It is not known if the aforementioned maternally expressed transcripts on chromosome 12 are also part of a longer transcriptional unit that is initiated at the

* Corresponding author. Fax: +49 0 681 302 2703.
E-mail address: m.paulsen@mx.uni-saarland.de (M. Paulsen).
[1] Present address: Clinical-Research-Services, Robert-Koch-Strasse 120 a, D-65779 Kelkheim/TS, Germany.
[2] Present address: Buddhist Tzu-Chi General Hospital, 707, Section 3 Chung-Yang Road, Hualien, Taiwan.

0888-7543/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2005.09.018

transcriptional start site of *Gtl2* upstream of all the other maternally expressed genes. Lin et al. identified the IGDMR 13 kb upstream of *Gtl2* as responsible for the regulation of the imprinting status for all known genes within the domain [15]. Whereas the deletion of the IGDMR on the paternal allele does not alter imprinting of any of the genes in the cluster, maternal transmission of the deletion leads to biallelic expression of normally paternally expressed genes and repression of mater-



nally expressed transcripts such as the *Gtl2* gene, snoRNAs, and microRNAs in *Rtl1-as* and *Mirg* [15]. Whereas *Dlk1* and *Dio3* appear to be twofold upregulated in these mutants, a much higher rate of *Rtl1* transcription was observed. This suggests that in wild-type animals *Rtl1* expression is regulated by RNA interference through its maternally expressed microRNA-containing antisense transcript. Consistent with this, cleavage products of the *Rtl1* transcript that correspond to sites of microRNA-mediated RNAi have been determined [12].

For detailed analyses of coordinated regulation of genes in this region by central imprinting elements such as the IGDMR and for a better understanding of the role of noncoding RNAs in this imprinted region, it is necessary to determine transcriptional activity throughout the region and to define the boundaries between imprinted and neighboring biallelically expressed genes, since it is not known how far the influence of the IGDMR extends to neighboring regions. At the centromeric flank the biallelically expressed *Wars* and *Yy1* genes, 600 kb from *Dlk1*, are the closest genes that have been analyzed for imprinted expression [16]. At the telomeric flank *Ppp2r5c* (regulatory subunit of protein phosphatase 2) and *Dnchc1* (dynein, cytoplasmic, heavy chain 1), which neighbor the imprinted *Dio3* gene, have not yet been analyzed.

In this study we focused our analysis on the telomeric part of the imprinting domain between *Gtl2* and *Dnchc1* in mouse. We present a detailed map including all genes/transcripts within the region in relation to structural features such as sequence elements that are conserved in mouse and human, CpG islands, distribution of repetitive elements, and tandem repeats. In addition, we show data on the expression and imprinting status of yet uncharacterized genes and transcripts.

Results

Sequence conservation between human and mouse

To obtain a comprehensive overview of the genomic organization of the telomeric part of the *Gtl2/Dlk1* imprinting domain on mouse chromosome 12 (Fig. 1A) we generated a high-resolution map with a variety of informative sequence features (Fig. 1B). In mouse the analyzed region spans 1,123,497 bp starting at exon 1 of *Gtl2* and terminating at the last exon of *Dnchc1*. To determine sequence conservation between mouse and human the murine sequence was aligned to the corresponding *GTL2(MEG3)/DNCHC1* region in human, which spans 1,224,676 nt (Ensembl, version 17.33.1). Using the PIPMAKER software [17] we identified 247 highly conserved sequence elements (CEs) of at least 100 bp length and 70% identity (Fig. 1B, Supplemental Material Table S1). The highest concentration of CEs is found in exons of the protein-encoding genes *Rtl1*, *Dio3*, *Dnchc1*, and *Ppp2r5c* and to a lesser extent in the non-protein-coding genes *Gtl2*, *Rtl1-as*, *Mirg*, and *Dio3as*. The *Rian* gene is not conserved although a snoRNA-containing gene, *MEG8*, is found at a corresponding position in the human sequence [11] (Fig. 1B, Supplemental Material Fig. S1). The CEs are not evenly distributed along the domain: the first 210 kb between *Gtl2* and *Mirg* contain 72 CEs (Fig. 1B). A remarkably high proportion of these, 46 CEs, are located outside of exons. In the adjacent 533 kb between *Mirg* and *Dio3* the conservation drops significantly (65 CEs). The 390 kb telomeric of *Dio3* contain 107 CEs. The conservation in this area is almost entirely confined to the large number of exons of the *Ppp2r5c* and *Dnchc1* genes. In summary, a cluster of intra- and intergenic CEs reflects the high sequence conservation in the region of the maternally expressed *Gtl2*, *Rtl1*, *Rian*, and *Mirg* genes.

Analysis of ESTs and predicted transcripts

To identify all transcriptional units in the region we mapped all expressed sequence tags (ESTs) deposited in the mouse EST section of the GenBank database. We identified 1153 matching mouse ESTs using the BLASTN algorithm with the following search criteria: match of >100 bp and >98% identity. Of these ESTs 997 corresponded to exons of the known *Gtl2*, *Rtl1/Rtl1-as*, *Rian*, *Mirg*, *Dio3/Dio3as*, *Ppp2r5c*, and *Dnchc1* genes. Our analysis did not identify yet unannotated exons within these genes except for *Ppp2r5c*, which apparently has three additional 5′ exons. We also identified two intronless genes, *Rpl10-ps* (GenBank Accession No. NM_052835), between *Mirg* and *Dio3*, and *Rps25-ps* (GenBank Accession No. NM_024266), between *Dio3* and *Ppp2r5c*, in the mouse. Both are most likely retrotransposed pseudogenes of the ribosomal protein genes 10 and S25 and are not present in the corresponding human regions. Conversely, the human sequence contains a pseudogene, *RPL26P4* (GenBank Accession No. NG_002527), between *DIO3* and *PPP2R5C*, which has no counterpart in mouse. Imprinting and expression analyses of the pseudogenes were impossible due to the high sequence similarities to other chromosomal pseudogene copies. Among the remaining 156 ESTs we did not find

Fig. 1. Detailed graphical map of the mouse *Gtl2* and *Dnchc1* region. (A) Schematic overview of the *Dlk1/Dio3* imprinting cluster in mouse. Genes are shown as colored boxes: red, maternal expression; blue, paternal expression; black, biallelic expression; gray, unknown allele-specific expression. The locations of small RNAs are indicated by asterisks. The relative transcriptional orientation of genes is indicated by arrows. The drawing is not to scale. (B) Graphical compilation of the sequence features of the *Gtl2–Dnchc1* region in mouse. Arrows above the genes indicate the orientation of transcription. Genes are colored according to their imprinting status: red, maternally expressed; blue, paternally expressed; white, unknown. The exons of the biallelically expressed *Dio3as, Ppp2r5c*, and *Dnchc1* genes analyzed in this paper are marked in black. The 5′ extension (smaller bar) of *Ppp2r5c* contains three additional 5′ exons. Proven tandemly repeated microRNAs [6,19,35] are shown in purple, snoRNAs in orange [11]. The "Homologies" lanes show the identified conserved elements between mouse and human. Short homology matches (>70% identity, >40 bp length) are shown in orange, long matches (>70% identity, >100 bp length) in black. Repetitive elements are differentiated in lane "Rep. Elements" as LTRs (green), LINEs (blue), and SINEs (red). "EST density" is given by pink bars: light pink, 1–5 matching ESTs; bright violet, 5–20 ESTs; dark violet, >20 ESTs per position. Conserved CpG islands are marked by pink asterisks, the conserved tandem repeat (R) downstream of *Gtl2* is marked by a pink triangle. Highly conserved elements (CE 1–4) and ESTs 1–3 that have been analyzed by RT-PCRs are marked by green rhombi.

any evidence for additional yet unidentified spliced transcripts, which also argues against the existence of three predicted transcripts in the NCBI database (LOC212473, GenBank Accession No. XM_153721; LOC238398, GenBank Accession No. XM_138266; 4930511J24Rik, GenBank Accession No. XM_147703).

The remaining 156 intergenic ESTs mapped to 38 distinct intergenic positions, 26 of those are concentrated in the region between *Gtl2* and *Mirg* (Fig. 1B). The great majority of the intergenic ESTs are transcribed in the same direction as *Gtl2*, *Rtl1-as*, *Rian*, and *Mirg*. Except for one (CE1), the ESTs do not overlap with the highly conserved elements.

To investigate whether the intergenic conserved elements and ESTs are indeed transcribed we performed RT-PCR on DNase I-treated RNAs derived from embryo (12.5 dpc) and newborn brain. For amplification we selected three CEs (CE1–3) and two ESTs (EST2 and -3) between the *Gtl2* and *Mirg* genes (positions marked in Fig. 1B). We observed consistent transcription at all five positions (Fig. 2A). In contrast, the RT-PCR at a conserved position (CE4) 48 kb downstream of *Mirg* was negative. Of the tested elements we selected one, CE1 between *Gtl2* and *Rtl1-as*, for strand-specific RT-PCRs. This confirmed exclusive transcription in the same orientation as *Gtl2*, *Rtl1-as*, *Rian*, and *Mirg* (Fig. 2B). One of the transcribed elements, CE2, contains a polymorphism between *Mus musculus domesticus* (Dom) and *M. musculus molossinus* (Mol) mice, which allowed us to examine allele-specific expression in the F2 progeny of reciprocal crosses of these mouse strains. RT-PCR products were amplified from neonatal brain and embryo (12.5 dpc) RNA and analyzed using a modified single nucleotide primer extension (SNuPE) method. In brief, primers 5′ of an informative SNP were elongated by SNuPE and elongated primer products were quantitatively separated on an ion pair reverse-phase HPLC (SIRPH) (for details see Material and methods, S. Tierling et al., manuscript in preparation [18]). The analysis showed that CE2 was

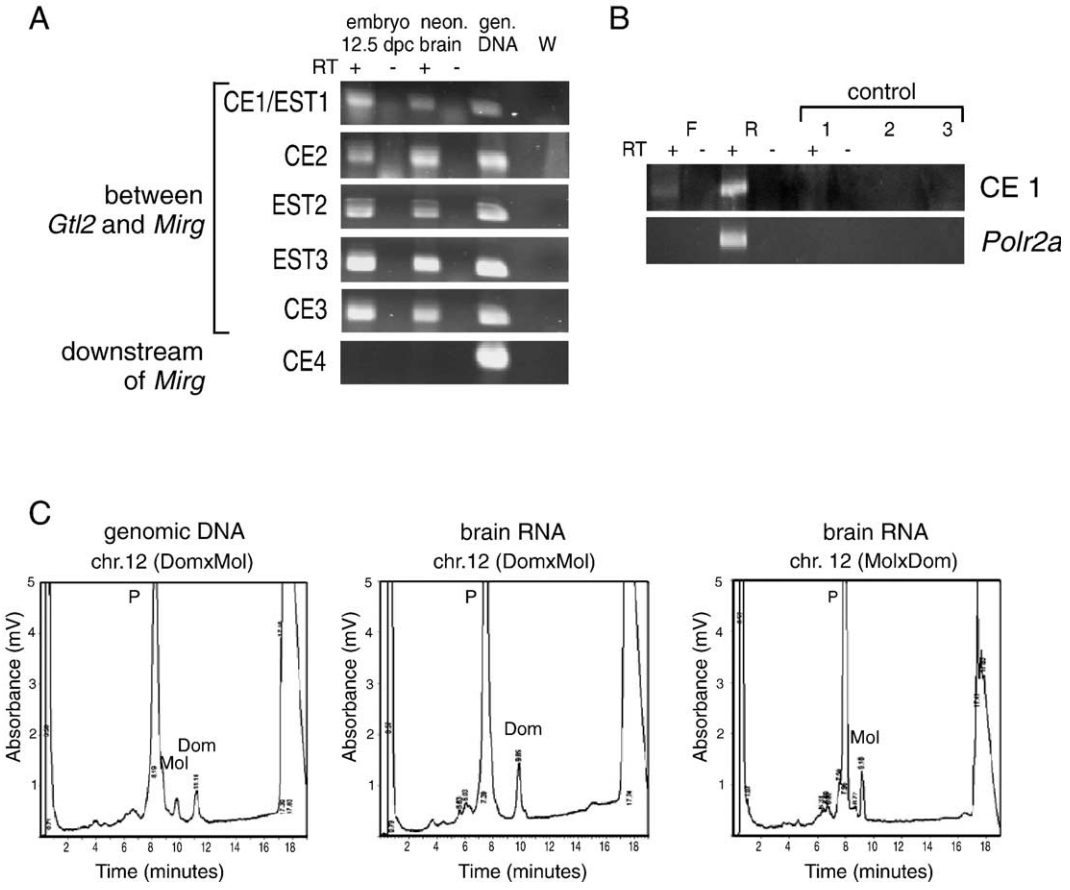


Fig. 2. Expression analyses of intergenic CEs and ESTs. (A) Expression of the CEs and ESTs. CE1 overlaps EST1. RT-PCR products were obtained from DNase I-treated total RNA isolated from embryo (12.5 dpc) and neonatal brain and separated in agarose gels. RT+ (reverse transcriptase added) and RT− (without reverse transcriptase) reactions are shown in consecutive lanes. As positive control a genomic DNA was used as template (gen. DNA). W, water/no-template control. (B) Strand-specific expression of CE1 in neonatal brain. RT-PCR products were obtained from DNase I-treated RNA. Top: RT-PCR products of CE1. Bottom: RT-PCR products of *Polr2a*. Labeling of lanes: F, reverse transcription using the RT-PCR-forward primer; R, reverse transcription using the RT-PCR-reverse primer. Controls: 1, reverse transcription without primer; 2, reverse transcriptase without template; 3, water control (no template). Products are visible only after reverse transcription with transcript-specific reverse primers. (C) UV spectra of SIRPH assays analyzing the imprinted expression of CE2 in brain. Allele-specific primer extension (ddCTP, Dom; ddTTP, Mol) was performed on PCR products obtained from genomic DNA (Dom × Mol) or from neonatal brain cDNA of genotyped F2 animals (chromosome 12, Dom × Mol, and chromosome 12, Mol × Dom). The relevant elution products are labeled with P, unextended primer; Mol, Mol-specific extension product; Dom, Dom-specific extension product. *x* axes, HPLC elution time (min); *y* axes, UV absorbance (mV). The absence of a paternal Mol peak in the middle graph and of a paternal Dom peak in the right graph reflects the exclusive maternal expression of CE2.

expressed only from the maternal alleles in the tested samples (Fig. 2C).

CpG islands in Rtl1, Mirg, and Dio3 are conserved

In human and mouse, the overall G+C contents are similar, with 46.51 and 46.94%, respectively. CpG-rich sequences are concentrated in the telomeric and centromeric gene-rich regions of the domain. Using the EMBOSS CpG plot tool (for settings see Methods) we identified 61 CpG islands in the mouse and 108 in the human sequence. Comparing the positions of CpG islands and conserved elements we found 9 CpG islands overlapping CEs (highlighted by asterisks in Fig. 1B). They are found in *Rtl1* (GenBank Accession No. AC152063.4, nt 112,700–113,040), upstream and in intron 4 of *Mirg* (GenBank Accession No. AC121784.2, nt 107,869–108,149, nt 122,709–122,969, nt 138,599–138,949), in *Dio3* (GenBank Accession No. AL591207, nt 110,185–112,364), and upstream and within *Dnchc1* (GenBank Accession No. AL773556.1, nt 36,925–37,135; AL1596265.12, nt 202,583–203,493, nt 186,593–187,023, nt 186,013–186,473). The CpG islands in *Rtl1* and *Dio3* have been shown to be biallelically methylated and unmethylated, respectively [8,19]. A DNA methylation analysis of the remaining CpG islands is in progress.

Repetitive elements

Imprinted regions have been shown to be depleted in SINEs compared to regions with biallelic gene expression [20,21]. Moreover the presence of specific tandem repeats has frequently been described as a hallmark of imprinted domains. We therefore decided to determine the amount of retrotransposed LINEs and SINEs, LTRs, and simple and tandem repeats in the entire region in mouse and human (Supplemental Material Table S2).

In mouse, retrotransposed elements amount to 28.13%, in human to 38.29%, of the entire sequence. The higher percentage in human is due mainly to a 100-kb insertion of repetitive elements in human. Whereas the overall amount of repetitive elements is comparable to average values known for mammalian genomic regions with similar G+C content, the SINE content (7.72%) is reduced in comparison to an expected average value of 13% [23,24]. This observation supports the notion that a depletion of SINEs is a hallmark of imprinted regions [20]. Moreover a detailed map of SINEs, LINEs, and LTRs reveals that their distribution is uneven (Fig. 1B): in the first 210 kb between *Gtl2* and *Mirg*, the overall repeat content is low (6.00%), followed by a striking increase between *Mirg* and *Dio3* (35.82%). In this region SINEs are still underrepresented (6.50%), while LINEs

A mouse

```
            1       10        20        30        40        50        58
            |-------+---------+---------+---------+---------+-------|
      9928  CTCCAGCTTTTG---AAGCTCAGATGTCCAGAAA-GTGGGACCTCTTTAACGTTTCAG
      9806  CTCCAGCCTTTGATGAAACTCAGATGTCCAGAAA-GAGGGACCTTTCTGATGTTCCAG
      9735  CTCCAGCCATTGGTGAAGCTCAGATGTCCAGAAA-GAGGGACCTCTCTGATGTTCCAG
      9593  CTCCAGCCTTTGATGAAGCACAGATGTCCAGAAA-GATGGCCCTCTCTGATGTTCCAG
      9451  CTCTAGCCTTTGATGAAGCTCAGATGTCCAGAAA-GATGGCCCTCTCTGATGTTCCAG
      9522  CTCTAGCCTTTGATGAAGCTCAGATGTCCAGAAA-GATGGCCCTCTCTGATGTTCCAG
      9380  ATCTAGCCTTTGATGAAGCTCAGATGTCCTGAAA-GAGGGACCTCTCTGATGTTCCAG
      9664  CTCCAGCCATTGATGAAGCTCAGATGTCTAGAAA-GAGAGTCCTCTATGATGTTCCAG
      9308  CTCCAGCCTTTGATGAAGCACAGATGTCCAGAAAAGAGCATTCTCAATGATGTTATAG

Consensus  CTCCAGCctTTGatgAAGCtCAGATGTCCAGAAA,GaGggacCTCtcTgAtGTTccAG
```

B human

```
            1       10        20        30        40        50        60        71
            |-------+---------+---------+---------+---------+---------+---------+|
    102855  CTCAGTTCCAGCCTCCAATGAAACCCAGATGTTCAGAAAGAGGGACCTCCTTGATGTTCTAGAGCCAGGAC
    103272  TCCAGTTCCAGCCTCCAGTGAAGCTCAGATGTCCAGAAAGAGGGACTACCTTGATGTTTTAGAGCCAGGAC
    103343  TCCAGCTCCAGCCTCCGATAAAGCTCGGATGTTCAGAATGAAGGACGTCCTTGATGTTCTAGAGCCAAGAC
    103938  TCCAGCTCCAGCCTCTGATGAAGCTCAGATATTCAGAAAGGGGGACTTCCTTTATGTTCCAGAGCCAAGAC
    103201  TCCAGATCCAGTGTCTGAAGAAGCTCAGAGGTCCAGAAAGAGGGACCTCCTTGATGTTCTAGAGCCAGGAC
    103867  TCTAGCCCTAGCCTCTAATGAAGCTCAGATGTCCAGAAAGAAGGACTTTCTTGAAATTCTAGAGCCAGAAC
    102926  TCCAGATCTAGCATCAGGTGATGCTCAGAGGTCTAGAAAGAGGAACCTGCTTAATGTTCTAGAGCCAAGAC
    103628  CCCAGTTCCAACATCTGACAAAGCCCAGATTTCCAGAAAGAGGGGTCTCCTTGATGTTCCAGAGCTGAGGG
    104008  CCCAGCTCCAACCTCTGACAAAGCCCAGATGTCCAGAAAGAGGGACCCCCTTGATGTTCTAGAGCTGGAAC
    103558  TTCAGCTCTAGCTTTCGAAGAAGCGCAGATGTCCAGAAAGAGGGACCTCTTTGATGTTCTAGGGCTGGGAC
    103726  TTTAGCTTTAGTTTTTGATGAAGCTCAGATGTGCAGAAAGAGGGACTTCCTTGATGTTCTAGGGCTGGGAC
    102997  TCCAGATCCAGCATCTGAT---GCTCAGATATTCAGAAACAGGGACCTCCTTGATATTCTAGAGCCGGGAC
    103065  TCCAGATCCAGTGTCTGAT---GCTCAGATATCCAGAAAGAGGGACCTCCTTGATTTTCTAGAGCCAGGAC
    103133  TCCAGATCCAGCATCCGATGATGCTGAGATGTCCAGAAAGAGAGAC---CTTGATGTTCTAGAGCCAGGAT

Consensus  TcCAGcTCcAGccTCtGAtgAaGCtCAGATgTcCAGAAAGAGGGACctcCTTGATgTTCTAGAGCcagGAC
```

Fig. 3. Multiple alignments of tandem repeats downstream of *Gtl2/GTL2*. Alignments and consensus sequences of the (A) mouse and (B) human tandem repeats using the program MultAlin (http://prodes.toulouse.inra.fr/multalin/multalin.html). Numbers in front of individual motifs refer to the position in the genomic sequence (GenBank Accession Nos. AC152063.4 (mouse), AL117190.6 (human)). Highly conserved nucleotides (>80%) are highlighted in black, weakly conserved nucleotides (<80%) are labeled in gray.

Table 1
Tissue-specific expression of *Mirg,* CE2, *Ppp2r5c, Dnchc1,* and *Dio3as*

	Days post coitum	Placenta	Embryo	Yolk sac
Mirg	12.5	+++	+++	+++
	16.5	+++	+++	+++
CE2	12.5	+++	+++	+++
	16.5	+++	+++	+++
Ppp2r5c	12.5	+++	+++	+++
	16.5	+++	+++	+++
Dnchc1	12.5	+++	+++	+++
	16.5	+++	+++	+++
Dio3as	12.5	++	+	++
	16.5	++	+	++

	Stage	Brain	Heart	Intestine	Kidney	Liver	Lung	Limb	Muscle	Spleen	Tongue
Mirg	Neonate	+++	–	–	–	–	–	+++	nd	–	+++
	Adult	+++	–	–	–	–	–	nd	–	–	–
CE2	Neonate	+++	–	–	–	–	–	+++	nd	–	+++
	Adult	+++	–	–	–	–	–	nd	–	–	–
Ppp2r5c	Neonate	+++	++	+++	+++	–	+	+	nd	–	–
	Adult	+++	++	++	+	++	+	nd	+	–	–
Dnchc1	Neonate	+++	+++	+++	+++	+++	+	–	nd	+	–
	Adult	++	+	+	+	+	++	nd	–	++	–
Dio3as	Neonate	–	–	+	+	–	+	–	nd	–	+
	Adult	+	–	–	+	+	+	nd	–	–	–

Semiquantitative RT-PCRs were performed on random-primed cDNA derived from total RNA of embryonic-stage, newborn, and adult organs. Intensities of RT-PCR products (35 cycles) were compared to the β-*actin* RT-PCR products (30 cycles) derived from the same tissues. Genes and analyzed stages are listed on the left. Top: RT-PCR expression data for prenatal stages (12.5 and 16.5 dpc). Bottom: RT-PCR expression data for neonatal and adult organs. Relative expression levels were classified as +++, strong; ++, intermediate; +, weak; –, not detectable; nd, not determined. Muscle tissue was dissected only from adult animals; instead, hind limbs were prepared from newborn mice.

(18.19%) and LTRs (11.13%) reach levels published for nonimprinted regions of similar G+C content [24]. Downstream of *Dio3*, i.e., outside of the imprinted domain, the SINE content increases progressively and reaches 13.45% in *Ppp2r5c* and *Dnchc1*; this is similar to the level in nonimprinted regions (approx 15%) with similar G+C content [22–24].

Using the PIPMAKER and FUZZNUC software ([16], http://bioweb.pasteur.fr/seqanal/interfaces/fuzznuc.html) we identified three conserved distinct tandem repeat arrays in the region. Two were previously described, encompassing the snoRNAs in *Rian*/*MEG8* and miRNAs in and around the *Mirg* gene [11]. Our detailed analysis revealed a third conserved tandem repeat array 3.2 kb downstream of *Gtl2*/*GTL2* (GenBank Accession No. AC152063.4, nt 9308–9981) (Fig. 3). In mouse 9 repeat units cover approx 700 bp; the corresponding human array consists of 14 repeat units covering 1200 bp (GenBank Accession No. AL117190.6, nt 102,855–104,078) (Fig. 3). Human and mouse repeats share a core consensus motif of 53 bp with 88% identity. The repeated motifs do not contain conserved CpG's. Nevertheless, as structural elements the repeat arrays resemble tandem repeats found in other imprinted regions. The tandem repeat array is structurally distinct from the snoRNA and miRNA clusters in *Rian* and *Mirg* as the repeat units do not form pronounced stable hairpins in RNA structure prediction (data not shown).

Analyses of tissue- and allele-specific gene expression

The imprinted expression of a number of genes in the domain is controlled by the IGDMR, a CpG island imprinting control element between *Dlk1* and *Gtl2* [15]. However, so far little is known about the developmental and tissue-specific regulation of gene expression within this imprinted domain. A comparative analysis of expression patterns may help to understand if imprinted and nonimprinted genes within the domain have distinct expression profiles and may give clues about the physiological role of the genes within the cluster associated with imprinting phenotypes. In the present study we analyzed tissue-specific expression of the *Mirg, Ppp2r5c, Dnchc1*, and *Dio3as* genes and included also the intergenic element CE2 between *Rtl1* and *Rian* (Table 1 and Supplemental Material Fig. S1). All five transcripts are expressed in embryos and extraembryonic tissues at 12.5 and 16.5 dpc. *Mirg* expression decreases progressively during postnatal development: in newborn mice, expression is most pronounced in brain, limbs, and tongue and becomes limited to brain in adult mice. A similar expression pattern was seen for the CE2

Notes to Table 2:
Compiled allele-specific expression data for *Mirg*, *Ppp2r5c*, and *Dnchc1*. SIRPH assays were performed on gene-specific RT-PCR products obtained from RNA of embryonic, neonatal, and adult tissue stages, as listed. To distinguish the alleles, material of F1 animals (Dom × Mol) or of F2 animals ((Dom × Mol) × Dom) and (Dom × (Dom × Mol)) was used. Expression levels are given as ratios of allele-specific expression to total expression: Mat/(Mat + Pat) and Pat/(Mat + Pat). For *Ppp2r5c* and *Dnchc1*, results of two individuals derived from each of the different crosses are shown. Mat, maternal allele; Pat, paternal allele; nd, not determined.

Table 2
Allele-specific expression of *Mirg*, *Ppp2r5c*, and *Dnchc1*

(A) *Mirg*

Sample	Dom × (Dom × Mol)		(Dom × Mol) × Dom	
	Mat (%)	Pat (%)	Mat (%)	Pat (%)
12.5 dpc				
Placenta	100	0	100	0
Embryo	100	0	100	0
Yolk sac	100	0	100	0
16.5 dpc				
Placenta	100	0	100	0
Embryo	100	0	100	0
Yolk sac	100	0	100	0
Neonatal				
Brain	100	0	93	7
Limb	100	0	94	6
Skin	100	0	96	4
Tongue	100	0	92	8
	Dom × Mol		(Dom × Mol) × Dom	
	Mat (%)	Pat (%)	Mat (%)	Pat (%)
Adult				
Brain	100	0	100	0

(B) *Ppp2r5c*

Sample	Dom × (Dom × Mol)				(Dom × Mol) × Dom			
	Individual 1		Individual 2		Individual 1		Individual 2	
	Mat (%)	Pat (%)	Mat (%)	Pat (%)	Mat (%)	Pat (%)	Mat (%)	Pat (%)
12.5 dpc								
Placenta	76	24	60	40	9	91	41	59
Embryo	53	47	51	49	47	53	47	53
Yolk sac	56	44	56	44	25	75	43	57
16.5 dpc								
Placenta	63	37	61	39	24	76	41	59
Embryo	55	45	48	52	24	76	47	53
Yolk sac	59	41	54	46	35	65	46	54
Neonatal								
Brain	63	37	59	41	46	54	40	60
Heart	80	20	55	45	27	73	42	58
Intestine	63	37	41	59	32	68	49	51
Kidney	55	45	62	38	18	82	55	45
Lung	59	41	47	53	44	56	57	43
Limb	72	28	55	45	36	64	45	55

(C) *Dnchc1*

Sample	Dom × (Dom × Mol)				(Dom × Mol) × Dom			
	Individual 1		Individual 2		Individual 1		Individual 2	
	Mat (%)	Pat (%)	Mat (%)	Pat (%)	Mat (%)	Pat (%)	Mat (%)	Pat (%)
12.5 dpc								
Placenta	71	29	66	34	14	86	50	50
Embryo	51	49	57	43	51	49	48	52
Yolk sac	47	53	54	46	14	86	49	51
16.5 dpc								
Placenta	60	40	61	39	52	48	48	52
Embryo	49	51	56	44	50	50	40	60
Yolk sac	46	54	45	55	nd	nd	33	67
Neonatal								
Brain	50	50	52	48	53	47	51	49
Heart	nd	nd	50	50	49	51	48	52
Intestine	55	45	45	55	54	46	54	46
Kidney	53	47	52	48	51	49	49	51
Liver	56	44	43	57	53	47	54	46

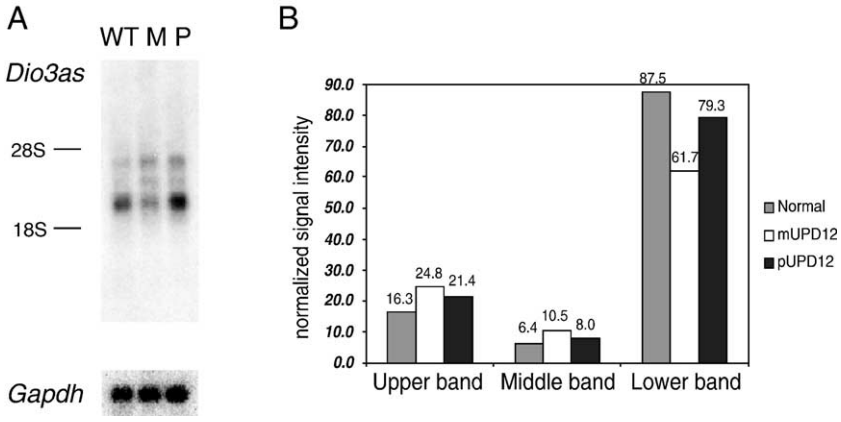


Fig. 4. Northern blot analysis probing with a *Dio3as* cDNA. (A) Northern blot of poly(A)$^+$ RNA from E15.5 wild-type (WT), mUPD12 (M), and pUPD12 (P) embryos hybridized with a *Dio3as* cDNA probe (GenBank Accession No. AY077457). The same blot was also hybridized with a *Gapdh* probe (bottom). (B) Bar diagram of the *Dio3as* signal intensities. Intensities of bands were normalized to the *Gapdh* signals. The numbers shown are the lane-specific *Dio3as*/*Gapdh* ratios multiplied by 100. The normalized expression levels of all three *Dio3as* transcripts (upper, middle, and lower bands) are very similar between mUPD12, pUPD12, and wild-type RNAs, suggesting that *Dio3as* does not exhibit imprinted expression.

element. The nonimprinted *Ppp2r5c* and *Dnchc1* genes (see below) continue to be strongly expressed in a variety of neonatal and adult tissues. Postnatal expression of *Dio3as* was barely detectable in RT-PCR analyses on postnatal tissues. In summary, *Mirg* and the intergenic transcript(s) show similar tissue-expression patterns that are distinct from those of the *Dio3as, Ppp2r5c*, and *Dnchc1* genes.

In addition, we examined the tissue- and stage-specific imprinting status of *Mirg, Dnchc1*, and *Ppp2r5c* using RNA from F1 and F2 progenies of reciprocal crosses of *M. m. domesticus* and *M. m. molossinus* mice. Allele-specific expression was determined by SIRPH on gene-specific RT-PCR products (see above and Methods). Tissues with a very low expression were excluded to avoid stochastic biases generated during RT-PCR amplification. The results are summarized in Table 2.

Mirg is predominantly expressed from the maternal allele in all tissues analyzed, confirming the maternal expression of *Mirg* as shown by the analysis of embryos with uniparental disomies of chromosome 12 [7]. Our analysis reveals that *Mirg* does not exhibit tissue-specific imprinting effects and that imprinted expression is maintained during postnatal development. *Ppp2r5c* and *Dnchc1* are apparently not imprinted in all tissues and stages analyzed. In a number of samples, predominantly in extraembryonic tissues, we observed overexpression of the *M. m. domesticus* allele for both *Ppp2r5c* and *Dnchc1* genes. These effects did not depend on the parental origin of the *M. m. domesticus* allele. In addition, we observed a variable penetrance and/or expressivity of this phenotype since not all individuals of a cross were affected and the allelic expression varied.

The RT-PCR analysis of *Dio3as* was hampered by difficulties in consistently amplifying sufficient amounts of product required for SIRPH analyses. The amplification problems were possibly caused by the high G+C content of the gene. However, the expression of *Dio3as* could readily be detected on Northern blot using polyadenylated RNA from embryos with a maternal or paternal uniparental disomy of chromosome 12 (mUPD12, pUPD12) (Fig. 4). As the histogram indicates, the expression levels of the three transcript variants did not differ significantly between mUPD12 and pUPD12 embryos. Thus, in contrast to the paternally expressed *Dio3* gene [8,9], its antisense transcript is apparently biallelically expressed and not imprinted in the mouse.

Discussion

Our detailed analysis of a 1.1-Mb region downstream of *Gtl2* in mouse and the sequence comparison to human revealed several intriguing structural and functional features of this imprinted domain, which may help to decipher the regulation of gene expression and imprinting control in mouse and human.

Analyses of transcripts in the region downstream of Gtl2

A detailed in silico analysis of ESTs revealed that the *Gtl2–Dnchc1* region does not contain additional genes except for a few mouse- and human-specific retrotransposed ribosomal protein-like genes. Such species-specific insertions of retrotransposed genes into imprinted domains have also been observed in the BWS region of mouse [25]. It remains to be elucidated if such species-specific insertions of retrotransposed genes have effects on the regulation or imprinting of the genes in the domain.

One of the most intriguing observations from our detailed computational analysis was the presence of a significant number of strand-specific ESTs in the murine and human intergenic regions between the maternally expressed *Gtl2, Rtl1-as, Rian*, and *Mirg* genes (data not shown). In a detailed bioinformatic analysis these ESTs could not be linked to longer clearly defined spliced or unspliced transcripts. Nevertheless, our RT-PCR experiments at five intergenic positions suggest that the identified ESTs are part of longer, possibly unspliced transcript(s) that are transcribed in a *Gtl2* to *Mirg* direction. Such an intergenic transcript(s) appears to span

a region of about 210 kb including several genes, the newly identified tandem repeat, and all miRNAs and snoRNAs. This transcription apparently does not extend beyond the *Mirg* gene since a CE approx 48 kb 3′ of *Mirg* is not transcribed. Furthermore the number of intergenic ESTs and CEs 3′ of *Mirg* drops significantly. An allele-specific RT-PCR analysis of the CE2 suggests that the intergenic transcript(s) is transcribed from the maternally inherited chromosome like the *Gtl2, Rian*, and *Mirg* genes (Figs. 2B and 2C). In summary we propose that the intergenic transcript(s) might serve as an imprinted host transcript(s) for the maternally expressed intergenic noncoding RNAs, including the sno- and miRNAs in the region [6,11].

Biallelic expression of Ppp2r5c and Dnchc1

Although 750 kb away from the IGDMR and separated from the other imprinted genes by a cluster of repetitive elements, *Dio3* exhibits persistently preferential expression from the paternal allele [8,15,24], whereas the neighboring *Dio3as, Ppp2r5c*, and *Dnchc1* genes are biallelically expressed in most analyzed tissues. The biallelic expression of *Dio3as* that is centromeric to *Dio3* indicates that the imprinted domain is not separated by a well-defined boundary from nonimprinted neighboring genes. However, since examples like *Tssc6* and *Trpm5* in the BWS region show that biallelically expressed genes can be located within an imprinted domain [25,26] it cannot formally be excluded that genes telomeric to *Dnchc1* are imprinted.

For *Ppp2r5c* and *Dnchc1* we did not observe evidence for parental imprinting but repeatedly found a preferential expression of the *M. m. domesticus* allele in a number of tissues, especially during prenatal stages (Table 2). Interestingly, in some tissue samples this preference was observed for both genes. The preferential expression of the *M. m. domesticus* allele in only some F2 animals suggests strain-specific effects on allelic gene expression depending on the genetic background. Similar effects on imprinting were recently described by Croteau *et al*. [27]. The authors showed a relaxation of imprinting of the *Dlk1, Gtl2*, and *Kcnq1* genes in a high proportion of F2 animals derived from intercrossed F1 animals (*M. m. domesticus* × *M. m. molossinus*) [27]. Together these findings suggest that some imprinting effects of genes—particularly tissue-specific imprinting—might have to be reexamined in a larger set of progeny and on repeated and reciprocal backcrosses to distinguish them from strain-specific modifier effects.

Tissue-specific expression patterns

Our qualitative RT-PCR results indicate that *Mirg, Ppp2r5c*, and *Dnchc1* are expressed in different subsets of tissues. The rather widespread imprinted expression of *Mirg* in newborn mice becomes almost exclusively restricted to the brain in adult mice. This indicates that *Mirg* and its associated miRNAs might fulfill different functions during development and may be important for brain function in adult mice [6]. A similar restriction of expression to brain in adult mice has also been suggested for *Gtl2*, *Rian* and the snoRNAs of the *Rian* locus [11,28,29], and is also seen for the intergenic element CE2. These similarities suggest a coordinated expression of the maternally expressed genes in the region. Our data indicate that expression of a long polycistronic transcript that includes *Gtl2, Rtl1-as, Rian*, and *Mirg* might contribute to coexpression of these genes. Since in a microarray-based analysis of the murine transcriptome a strong expression in brain was also observed for the paternally expressed *Dlk1* gene [30], coordinated regulation might also affect to some extent the paternally expressed genes in the region. This pattern of coexpression does not extend to the telomeric part of the domain since the paternally expressed *Dio3* and the neighboring nonimprinted *Ppp2r5c* and *Dnchc1* genes have clearly distinct expression profiles [8,31].

The biological importance of the (imprinted) expression of the *Gtl2, Rian*, and *Mirg* genes in adult brain remains unclear. Most of the complex phenotypes associated with genes in this imprinting cluster have been analyzed in mice carrying uniparental disomies of chromosome 12. These mice show placental and skeletal defects and distinct muscle morphologies and die before birth [3,5]. Thus, possible brain-specific functions of these imprinted genes in postnatal tissues remain to be analyzed using alternative models.

Methods

Mice

For analyses on allele-specific gene expression, *M. m. domesticus* and *M. m. molossinus* animals were mated and tissues were dissected from F1 and F2 animals. For this, F1 (Dom × Mol) females were mated with Dom males, and reciprocally Dom females were mated with F1 males. Various tissues were prepared from newborns at day 1. Total embryo and extraembryonic tissues were taken at 12.5 and 16.5 dpc.

Genotyping of F2 mice was carried out by amplification of strain-specific microsatellites (*D12Mit8* and *D12Mit259*) on chromosome 12, both enclosing the imprinted domain. For *D12Mit8* the following primers were used: 5′-TTGCCTAACCCACTCACACC-3′ and 5′-TGGTGACTCCTTACAGAGGC-3′. The PCRs were performed in 20 μl reaction volume in the presence of 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.25 μM each primer pair, and 0.1 U *Taq* DNA polymerase. The reaction was denatured at 95°C for 3 min, subsequently 30 cycles were performed (94°C for 1 min, 51°C for 20 s, 72°C for 20 s), followed by a final extension at 72°C for 5 min. For *D12Mit259* the following primers were used 5′-TAGCAACATGTAAAAG-CATGATACC-3′ and 5-′TACCTTGAGAAAAGTATGGAGAAATG-3′. The PCRs were performed in 20 μl reaction volume in the presence of 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.25 μM each primer pair, and 0.1 U *Taq* DNA polymerase. The denaturation for 1 min at 94°C was followed by 30 cycles (94°C for 1 min, 54°C for 1 min, 72°C for 20 s) and a final extension at 72°C for 2 min.

RT-PCRs

For RNA preparation tissues were homogenized in thiocyanate solution and extracted with acidic phenol–chloroform. Total RNA was isolated from the aqueous phase according to standard protocols [32]. cDNA was synthesized by reverse transcription: 1.5 μg total RNA was reverse transcribed in 20 μl reaction volume in the presence of 50 mM Tris–HCl (pH 8.3 at 25°C), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 100 ng random primer, 0.8 U RNasin, 0.375 mM dNTPs, 7 U M-MLV reverse transcriptase (Promega). Strand-specific RT was

performed using the EndoFree RT Kit (Ambion) according to the manufacturer's protocol. For subsequent PCRs the following primer pairs were used: β-*actin*, 5′-GCTGTGCTATGTTGCTCTAGACTTC-3′ and 5′-CTCAGTAACAGTCCGCCTAGAAGC-3′; *Mirg*, 5′-CCTGATGGAGGCTCGTCCAT-3′ and 5′-TAAATCCTGAGGGCAAACACTC-3;′; *Ppp2r5c*, 5′-ACTCCTCGATGACTGCACTCAGC-3′ and 5′-AGGGTGCTTTCCTACAGCTCTG-3′; *Dnchc1*, 5′-ACAAACAAGCGCCGAGAAGAAG-3′ and 5′-ACTAAACCCAGCCATTCGGTCA-3′; *Dio3as*, 5′-AGCACTCACAGGGGC-CTTCTCT-3′ and 5′-TCCTTCAGGTGGGAAGTGCTGA-3′; CE1/EST1, 5′-TGTTCTGTTCTGAGAGGGGCGG-3′ and 5′-CCAGAGTGAGCCCAGAAGCGAG-3′; CE2, 5′-AAACAAAAGGCTCCTGGCAGGC-3′ and 5′-CTGCTACTGGGCTGGGAGGGAT-3′; EST2, 5′-TGGAACCCACAGGCTTAAATCCTTT-3′ and 5′-ATTAGGGGACCTGAGATCGCTGTTG-3′; CE3, 5′-TGGCCTGGTAGCTGCTCTTTGG-3′ and 5′-TCCCATTTTGGCAGAGGCTAGGA-3′; EST3, 5′-GGGGTCCTGGAAAGCCACAGTCAGTC-3′ and 5′-CAGCCAAACACAACAGGGGGAG-3′; CE4, 5′-TGGCACCAACCAAAAACCACAA-3′ and 5′-GCTCTGGTCTCAAGGGTC-CTTGG-3′; *Polr2a*, 5′-ACCAAAGAGAAGGGCCATGGCG-3′ and 5′-TTCTGCATGCGA-CCGGGTAAGC-3′. The PCRs were performed in 20 μl reaction volume in the presence of 1 μl of the synthesized cDNA, 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.25 μM each primer pair, and 0.1 U *Taq* DNA polymerase. In strand-specific RT-PCRs 1.25 M BETAIN and 1.5% DMSO were added. For amplification of *Dio3as* LUNATAQ *Taq* DNA polymerase (Bioline) was used. The denaturation (94°C, 3 min) was followed by 35 PCR cycles (94°C for 30 s, 30 s annealing (β-*actin* 62°C, *Mirg* 55°C, *Ppp2r5c* 60°C, *Dnchc1* 61°C, *Dio3as* 61°C, all CEs and ESTs 60°C), 72°C for 30 s) and a final extension at 72°C for 5 min. The β-*actin* RT-PCR and the strand-specific RT-PCR on CE1 encompassed only 30 cycles.

SIRPH analysis

Sequencing of gene-specific RT-PCR products derived from homozygous *M. m. domesticus* and *M. m. molossinus* revealed the following SNPs: for *Mirg* an A:T SNP (nt 827 in GenBank Accession No. AJ517767), for *Ppp2r5c* a G:A SNP (nt 37367 in GenBank Accession No. AL773556.1), for *Dnchc1* a G:A SNP (nt 14186 in GenBank Accession No. NM_030238), and for CE2 a T:C SNP (nt 10226 in GenBank Accession No. AC121784.2).

For the SNuPE reaction 100–130 ng/μl gel-purified RT-PCR products were used as templates. SNuPE primers were placed immediately adjacent to the polymorphic sites and had the following sequences: *Mirg*, 5′-TCAAGGAACCCTGCCTATGC-3′; *Ppp2r5c*, 5′-CCCAGCACAGCCCTGA-3′; *Dnchc1*, 5′-TTAACGGTGTGGAAGGGTTG-3′; CE2, 5′-GGGTCGCCTGCACTCC-3′. Before the SNuPE reaction *Dnchc1* PCR products were digested with 0.2 U *FokI*. The primers were extended in 20-μl reactions under the following conditions: 3.6 μM SNuPE primer, 0.05 mM ddNTPs, 0.15 U Thermo-Sequenase (Amersham) in reaction buffer supplied by the manufacturer. After denaturation for 2 min at 96°C, 50 cycles (96°C for 15 s, 37°C for 30 s, 60°C for 2 min) were performed. Extension products were separated on a dHPLC system (WAVE DNA Fragment Analysis System, Transgenomics). Extension products were separated at 75 (*Mirg, Ppp2r5c, Dnchc1*) or 50°C (CE2) by the following acetonitrile gradients, which were generated by continuously mixing buffer A (0.1 M triethylammonium acetate (TEAA)) and buffer B (0.1 M TEAA, 25% acetonitrile): *Mirg*, 18–28% for 15 min; *Ppp2r5c*, 14–22% for 15 min; *Dnchc1*, 18–30% for 15 min; CE2, 18–30% for 15 min. After estimation of peak areas or heights allele-specific expression levels were calculated as the ratio of allele-specific expression to total expression of both alleles. Since we observed some variability in our analyzed samples we randomly selected samples for a second reverse transcription and subsequent PCR. In addition we examined RNAs from a second individual.

Northern blot analysis

Poly(A)$^+$ RNA was extracted from total RNA using the Dynabeads Oligo(dT)$_{25}$ kit (Dynal, Merseyside, UK) as per the manufacturer's protocol. Poly(A)$^+$ RNA (0.5–1 μg) was separated on formaldehyde–agarose gels, blotted, and hybridized with radiolabeled probes as described previously [28]. The *Dio3as* probe was derived from a cDNA clone (GenBank Accession No. AY077457) [31]. Hybridization signal intensities were quantified relative to *Gapdh* on an Amersham Storm 860 PhosphorImager using ImageQuant software.

Computational DNA sequence analyses

Mouse genomic sequences were taken from NCBI (www.ncbi.nlm.nih.gov; GenBank Accession No. NT_039553.3, nt 5,697,858–6,820,806; October 10, 2003). The human sequence was obtained from the Ensembl Genome Browser (www.ensembl.org; October 14, 2003; version 17.33.1). The analyzed region spans 1,123,497 bp in mouse and 1,224,676 bp in human. Pairwise alignments were generated using the PIPMAKER software at Pennsylvania State University ([17], http://bio.cse.psu.edu/cgi-bin/pipmaker). Repetitive elements were detected using the RepeatMasker software (Smit and Green, unpublished data, http://repeatmasker.org). CpG islands were identified by the CpG plot software at the European Bioinformatics Institute (www.ebi.ac.uk/emboss/cpgplot/) choosing the following settings: window = 200, step = 10, obs/exp = 0.6, MinPC = 50, length = 200. EST density was determined by BLASTN searches against mouse ESTs in the NCBI database (http://www.ncbi.nlm.nih.gov). Only EST matches that were longer than 100 bp and at least 98% identical were selected for analysis. The analysis for *Gtl2–Ppp2r5c* was done on June 22, 2004, and for *Ppp2r5c–Dnchc1* on December 14, 2004. Tissue-specific EST data were obtained from the ExQuest database ([33], http://lena.jax.org/~dcb/xquest.html). Tandem repeats were identified using the PIPMAKER software and the FUZZNUC program (http://bioweb.pasteur.fr/seqanal/interfaces/fuzznuc.html) applying a 20-nucleotide window and allowing six mismatches. Identified tandem repeats were aligned by MultAlin ([34], http://prodes.toulouse.inra.fr/multalin/multalin.html).

Acknowledgments

We thank Francois Bonhomme, Montpellier, for providing *M. mus molossinus* mice; Christina Daniel for DNA sequencing; Osman El-Maarri and Manfred Küpper for helpful discussions concerning dHPLC analysis; Arturo Hernandez for providing a *Dio3as* probe; and Gilles Gasparoni for critically reading the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft DFG (Grant no. WA1029/3-2), the EU STREP "Forsak", and the EU FP6 Network of Excellence "The Epigenome." We highly appreciate the work of numerous sequencing centers that made the mouse and human genomic DNA sequences for the analyzed region publicly available.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ygeno.2005.09.018.

References

[1] M. Georges, C. Charlier, N. Cockett, The callipyge locus: evidence for the trans interaction of reciprocally imprinted genes, Trends Genet. 19 (2003) 248–252.
[2] P. Georgiades, M. Watkins, M.A. Surani, A.C. Ferguson-Smith, Parental origin-specific developmental defects in mice with uniparental disomy for chromosome 12, Development 127 (2000) 4719–4728.
[3] P. Georgiades, M. Watkins, G.J. Burton, A.C. Ferguson-Smith, Roles for genomic imprinting and the zygotic genome in placental development, Proc. Natl. Acad. Sci. USA 98 (2001) 4522–4527.
[4] V.R. Sutton, L.G. Shaffer, Search for imprinted regions on chromosome 14: comparison of maternal and paternal UPD cases with cases of chromosome 14 deletion, Am. J. Med. Genet. 93 (2000) 381–738.

[5] P. Georgiades, C. Chierakul, A.C. Ferguson-Smith, Parental origin effects in human trisomy for chromosome 14q: implications for genomic imprinting, J. Med. Genet. 35 (1998) 821–824.
[6] H. Seitz, H. Royo, M. Bortolin, S.P. Lin, A.C. Ferguson-Smith, J. Cavaille, A large imprinted microRNA gene cluster at the mouse Dlk1(Gtl2 domain, Genome Res. 14 (2004) 1741–1748.
[7] H. Seitz, N. Youngson, S.P. Lin, S. Dalbert, M. Paulsen, J.P. Bachellerie, A.C. Ferguson-Smith, J. Cavaille, Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene, Nat. Genet. 34 (2003) 261–262.
[8] C.E. Tsai, S.P. Lin, M. Ito, N. Takagi, S. Takada, A.C. Ferguson-Smith, Genomic imprinting contributes to thyroid hormone metabolism in the mouse embryo, Curr. Biol. 12 (2002) 1221–1226.
[9] A. Hernandez, S. Fiering, E. Martinez, V.A. Galton, D. St. Germain, The gene locus encoding iodothyronine deiodinase type 3 (Dio3) is imprinted in the fetus and expresses antisense transcripts, Endocrinology 143 (2002) 4483–4486.
[10] C. Charlier, K. Segers, L. Karim, T. Shay, G. Gyapay, N. Cockett, M. Georges, The callipyge mutation enhances the expression of coregulated imprinted genes in cis without affecting their imprinting status, Nat. Genet. 27 (2001) 367–369.
[11] J. Cavaille, H. Seitz, M. Paulsen, A.C. Ferguson-Smith, J.P. Bachellerie, Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region, Hum. Mol. Genet. 11 (2002) 1527–1538.
[12] E. Davis, F. Caiment, X. Tordoir, J. Cavaille, A.C. Ferguson-Smith, N. Cockett, M. Georges, C. Charlier, RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus, Curr. Biol. 15 (2005) 743–749.
[13] J. Cavaille, K. Buiting, M. Kiefmann, M. Lalande, C.I. Brannan, B. Horsthemke, J.P. Bachellerie, J. Brosius, A. Huttenhofer, Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization, Proc. Natl. Acad. Sci. USA 97 (2000) 14311–14316.
[14] M. Runte, A. Huttenhofer, S. Gross, M. Kiefmann, B. Horsthemke, K. Buiting, The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A, Hum. Mol. Genet. 10 (2001) 2687–2700.
[15] S.P. Lin, N. Youngson, S. Takada, H. Seitz, W. Reik, M. Paulsen, J. Cavaille, A.C. Ferguson-Smith, Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1(Gtl2 imprinted cluster on mouse chromosome 12, Nat. Genet. 35 (2003) 97–102.
[16] A. Yevtodiyenko, M.S. Carr, N. Patel, J.V. Schmidt, Analysis of candidate imprinted genes linked to Dlk1(Gtl2 using a congenic mouse line, Mamm. Genome 13 (2002) 633–638.
[17] S. Schwartz, Z. Zhang, K.A. Frazer, A. Smit, C. Riemer, J. Bouck, R. Gibbs, R. Hardison, W. Miller, PipMaker—a web server for aligning two genomic DNA sequences, Genome Res. 10 (2000) 577–586.
[18] O. El-Maarri, SIRPH analysis: SNuPE with IP-RP-HPLC for quantitative measurements of DNA methylation at specific CpG sites, Methods Mol. Biol. 278 (2004) 195–205.
[19] N.A. Youngson, S. Kocialkowski, N. Peel, A.C. Ferguson-Smith, A small family of Sushi-class retrotransposon-derived genes in mammals and their relation to genomic imprinting, J. Mol. Evol. 61 (2005) 481–490.
[20] J.M. Greally, Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome, Proc. Natl. Acad. Sci. USA 99 (2002) 327–332.
[21] M. Paulsen, S. Takada, N.A. Youngson, M. Benchaib, C. Charlier, K. Segers, M. Georges, A.C. Ferguson-Smith, Comparative sequence analysis of the imprinted Dlk1–Gtl2 locus in three mammalian species reveals highly conserved genomic elements and refines comparison with the Igf2–H19 region, Genome Res. 11 (2001) 2085–2094.
[22] Z. Zhang, N. Carriero, M. Gerstein, Comparative analysis of processed pseudogenes in the mouse and human genomes, Trends Genet. 20 (2004) 62–67.
[23] R.H. Waterston, K. Lindblad-Toh, E. Birney, J. Rogers, J.F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, et al., Initial sequencing and comparative analysis of the mouse genome, Nature 420 (2002) 520–562.
[24] A.F. Smit, Interspersed repeats and other mementos of transposable elements in mammalian genomes, Curr. Opin. Genet. Dev. 9 (1999) 657–663.
[25] M. Paulsen, O. El-Maarri, S. Engemann, M. Strodicke, O. Franck, K. Davies, R. Reinhardt, W. Reik, J. Walter, Sequence conservation and variability of imprinting in the Beckwith–Wiedemann syndrome gene cluster in human and mouse, Hum. Mol. Genet. 9 (2000) 1829–1841.
[26] M.P. Lee, S. Brandenburg, G.M. Landes, M. Adams, G. Miller, A.P. Feinberg, Two novel genes in the center of the 11p15 imprinted domain escape genomic imprinting, Hum. Mol. Genet. 8 (1999) 683–690.
[27] S. Croteau, D. Roquis, M.C. Charron, D. Frappier, D. Yavin, J.C. Loredo-Osti, T.J. Hudson, A.K. Naumova, Increased plasticity of genomic imprinting of Dlk1 in brain is due to genetic and epigenetic factors, Mamm. Genome 16 (2005) 127–135.
[28] S. Takada, M. Tevendale, J. Baker, P. Georgiades, E. Campbell, T. Freeman, M.H. Johnson, M. Paulsen, A.C. Ferguson-Smith, Delta-like and gtl2 are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12, Curr. Biol. 10 (2000) 1135–1138.
[29] I. Hatada, S. Morita, Y. Obata, Y. Sotomaru, M. Shimoda, T. Kono, Identification of a new imprinted gene, Rian, on mouse chromosome 12 by fluorescent differential display screening, J. Biochem. (Tokyo) 130 (2001) 187–190.
[30] A.I. Su, T. Wiltshire, S. Batalov, H. Lapp, K.A. Ching, D. Block, J. Zhang, R. Soden, M. Hayakawa, G. Kreiman, et al., A gene atlas of the mouse and human protein-encoding transcriptomes, Proc. Natl. Acad. Sci. USA 101 (2004) 6062–6067.
[31] A. Hernandez, E. Martinez, W. Croteau, D. St. Germain, Complex organisation and structure of sense and antisense transcripts expressed from the DIO3 gene imprinted locus, Genomics 83 (2004) 413–424.
[32] P. Chomczynski, N. Sacchi, Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction, Anal. Biochem. 162 (1987) 156–159.
[33] A.C. Brown, K. Kai, M.E. May, D.C. Brown, D.C. Roopenian, ExQuest, a novel method for displaying quantitative gene expression from ESTs, Genomics 83 (2004) 528–539.
[34] F. Corpet, Multiple sequence alignment with hierarchical clustering, Nucleic Acids Res. 16 (1988) 10881–10890.
[35] M. Lagos-Quintana, R. Rauhut, A. Yalcin, J. Meyer, W. Lendeckel, T. Tuschl, Identification of tissue-specific microRNAs from mouse, Curr. Biol. 12 (2002) 735–739

Online references

National Center for Biotechnology Information National Center for Biotechnology Information: http://www.ncbi.nlm.nih.gov.
Sanger Institute Ensembl Genome Browser Sanger Institute Ensembl Genome Browser: http://www.ensembl.org.
Pennsylvania State University Pennsylvania State University: http://bio.cse.psu.edu/cgi-bin/pipmaker.
RepeatMasker RepeatMasker: http://repeatmasker.org.
European Bioinformatics Institute European Bioinformatics Institute: http://www.ebi.ac.uk/emboss/cpgplot.
Institute Pasteur Institute Pasteur: http://bioweb.pasteur.fr/seqanal/interfaces/fuzznuc.html.
Laboratoire de Genetique Cellulaire Laboratoire de Genetique Cellulaire, INRA: http://prodes.toulouse.inra.fr/multalin/multalin.html.
The Jackson Laboratory The Jackson Laboratory: http://lena.jax.org/~dcb/xquest.html.