



HAL
open science

Genomic matrix attachment region and chromosome conformation capture quantitative real time PCR assays identify novel putative regulatory elements at the imprinted Dlk1/Gtl2 locus

Caroline Braem, Bénédicte Recolin, Rebecca C Rancourt, Christopher Angiolini, Pauline Barthès, Priscilla Branchu, Franck Court, Guy Cathala, Anne C Ferguson-Smith, Thierry Forné

► **To cite this version:**

Caroline Braem, Bénédicte Recolin, Rebecca C Rancourt, Christopher Angiolini, Pauline Barthès, et al.. Genomic matrix attachment region and chromosome conformation capture quantitative real time PCR assays identify novel putative regulatory elements at the imprinted Dlk1/Gtl2 locus. *Journal of Biological Chemistry*, 2008, 283 (27), pp.18612-18620. 10.1074/jbc.M801883200 . hal-01600983

HAL Id: hal-01600983

<https://hal.science/hal-01600983>

Submitted on 30 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Genomic Matrix Attachment Region and Chromosome Conformation Capture Quantitative Real Time PCR Assays Identify Novel Putative Regulatory Elements at the Imprinted *Dlk1/Gtl2* Locus^{*[5]}

Received for publication, March 7, 2008, and in revised form, April 17, 2008. Published, JBC Papers in Press, May 5, 2008, DOI 10.1074/jbc.M801883200

Caroline Braem^{†1}, Bénédicte Recolin[‡], Rebecca C. Rancourt[§], Christopher Angiolini[§], Pauline Barthès[‡], Priscillia Branchu[‡], Franck Court[‡], Guy Cathala[‡], Anne C. Ferguson-Smith[§], and Thierry Forne^{‡2}

From the [†]Institut de Génétique Moléculaire, UMR 5535 CNRS-Université Montpellier II, IFR 122, 1919, Route de Mende, 34293 Montpellier Cedex 5, France and the [§]Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY, United Kingdom

We previously showed that genomic imprinting regulates matrix attachment region activities at the mouse *Igf2* (insulin-like growth factor 2) locus and that these activities are functionally linked to neighboring differentially methylated regions (DMRs). Here, we investigate the similarly structured *Dlk1/Gtl2* imprinted domain and show that in the mouse liver, the G/C-rich intergenic germ line-derived DMR, a sequence involved in domain-wide imprinting, is highly retained within the nuclear matrix fraction exclusively on the methylated paternal copy, reflecting its differential function on that chromosome. Therefore, not only “classical” A/T-rich matrix attachment region (MAR) sequences but also other important regulatory DNA elements (such as DMRs) can be recovered from genomic MAR assays following a high salt treatment. Interestingly, the recovery of one A/T-rich sequence (MAR4) from the “nuclear matrix” fraction is strongly correlated with gene expression. We show that this element possesses an intrinsic activity that favors transcription, and using chromosome conformation capture quantitative real time PCR assays, we demonstrate that the MAR4 interacts with the intergenic germ line-derived DMR specifically on the paternal allele but not with the *Dlk1/Gtl2* promoters. Altogether, our findings shed a new light on gene regulation at this locus.

Genomic imprinting is a parent-of-origin gene-silencing mechanism required for normal mammalian development. It involves germ line-specific epigenetic modifications acquired

on restricted regions of the genome (imprinting control region or element) that control the imprinting of several genes often over several hundred kilobase pairs. Accumulating evidence indicates that monoallelic expression at mammalian imprinted loci largely results from allele-specific higher order chromatin structures that impair or favor gene expression. In this context, it becomes crucial to elucidate the genomic architecture associated with imprinted genes as well as to identify DNA sequences and factors involved in such higher order chromatin organization. In the present work, we examine the potential role of the so-called matrix attachment regions (MARs)³ in imprinting and gene regulation. MARs have been operationally defined in “*in vitro* MAR assays” by their ability to “attach” to a purified nuclear matrix or scaffold. In that context, they appear as A/T-rich DNA sequences that may be involved in chromatin structure and gene expression (1). They are frequently associated with enhancers and promote chromatin accessibility and histone acetylation. Using a “genomic MAR” assay based on high salt treatment of purified nucleus preparations (2, 3), we previously showed that parental genomic imprinting controls MAR activities at the mouse *Igf2* (insulin-like growth factor 2) locus. That work indicated that tissue-specific MARs and differentially methylated regions (DMRs) may act as bipartite elements controlling the long range activity of other regulatory elements such as enhancers (3), thus contributing to the regulation of gene expression. Here, we investigate MAR activities at the *Dlk1/Gtl2* imprinted domain located on mouse chromosome 12 (human chromosome 14). Both the *Dlk1* and *Gtl2* genes are expressed during embryogenesis, and they are known to function as a critical barrier against parthenogenetic development in mammals (4). The *Dlk1* (Delta-like 1 homologue) gene is expressed from the paternal allele. It encodes a protein with homology to members of the Notch-Delta family of signaling molecules, which is important for mammalian development (5–7). 80 kilobases downstream of *Dlk1*, the *Gtl2* (gene trap locus 2) promoter is active only on the maternally inherited

* This work was supported by Association pour la Recherche contre le Cancer Contract 4868, “GIS Longévité” Contract GISLO401, Agence Nationale de la Recherche Grant ANR-07-BLAN-0052-02 (to T. F.), and funds from the Centre National de la Recherche Scientifique. This work was also supported by Cancer Research-United Kingdom, the Bill and Melinda Gates Trust (Cambridge), and European Union Framework Program 6. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Tables 1 and 2.

¹ Supported by Association pour la Recherche contre le Cancer Fellowship JR/MLD/MDV-P05/2 and P06/2.

² To whom correspondence should be addressed: IGMM, UMR5535 CNRS-UMII, IFR122, 1919, Route de Mende, 34293 Montpellier Cedex 5, France. Tel.: 33-467-61-36-84; Fax: 33-467-04-02-31; E-mail: forne@igmm.cnrs.fr.

³ The abbreviations used are: MAR, matrix attachment region; 3C, chromosome conformation capture; qPCR, quantitative real time PCR; DMR, differentially methylated region; IG, intergenic germ line-derived; HRS, high salt recovered sequence; NC, negative control; *En*, embryonic day *n*; *Dn*, postnatal day *n*; GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

chromosome (8, 9). It produces large polyadenylated alternatively spliced transcripts, which may extend further downstream and be processed into numerous noncoding RNAs, including several small nucleolar RNAs and microRNAs in addition to the *Gtl2* RNA (10–13). Regulatory mechanisms involved in the imprinting of the *Gtl2* promoter may differ between the embryo and the placenta (14). In all lineages, imprinting of the *Dlk1-Gtl2* domain is dependent on the so-called intergenic germ line-derived DMR (IG-DMR) located ~13 kb upstream of the *Gtl2* gene. This region is highly methylated on the paternal chromosome and unmethylated on the maternal chromosome (15). This methylation of the IG-DMR is therefore associated with activity from *Dlk1* and repression of *Gtl2*. The pattern of differential methylation originates from the gametes and is maintained in all cell types during the entire life of the individual (primary imprinting mark). The significance of the IG-DMR was demonstrated by its deletion in the mouse (16). In the embryo, the deletion of the IG-DMR (Δ IG-DMR) on the maternally inherited chromosome leads to loss of imprinting (activation) of the normally repressed paternally expressed genes (*Dlk1*, *Rtl1*, and *Dio3*) and to the repression of the maternally expressed noncoding RNAs. However, the latter effect is not observed in the placenta to the same extent (14, 16). Furthermore, upon paternal transmission of the deletion, imprinting is unaltered. These findings demonstrate that the unmethylated IG-DMR acts as a repressor for imprinted protein coding genes on the maternal chromosome over a region of ~1 Mb. To better understand how such long range regulation may occur at this locus, we investigated MAR activities and performed chromosome conformation capture (3C-qPCR) assays (17) in both the embryonic and the postnatal mouse liver.

EXPERIMENTAL PROCEDURES

Genomic MAR Assays/HRS Assays

The genomic MAR assay, which we propose to rename the “high salt recovered sequence assay” (HRS assays: see “Discussion”), was performed as previously described (3). Briefly, the nuclei isolated from tissue samples were extracted in high salt conditions (2 M NaCl), and the resulting “nuclear halos” were digested by XbaI, HindIII, and BamHI. Matrix bound DNA (MAR fraction) was isolated from loop DNA (loop fraction) by ultrafiltration (Ultrafree-CL), and the relative enrichment of target sequences in the MAR fraction relative to the loop fraction was determined by real time quantitative PCR (see below).

Quantitative Analysis of MAR/HRS Assays

In each HRS assay, target DNA sequences were quantified by real time PCR using a SYBR Green mix (18) and a LightCycler apparatus (Roche Applied Science; software version 3.5). Amplifications were performed as previously described (3), and the enrichments were calculated as the ratio of the amount of target DNA in the MAR fraction *versus* the loop fraction. To standardize each assay, the ratio values were normalized against the ratio obtained for a negative control (NC) located 41 kb upstream of the *Gtl2* gene and lacking predicted MAR features (see Fig. 2B). For primer sequences used for quantifications, please see supplemental data.

Quantification of the Allele Proportions

To assess the allelic specificity of the IG-DMR attachment, we took advantage of its parental origin-specific methylation pattern and the presence of restriction sites for methylation-sensitive/dependent enzymes. The following real time PCR primers were designed on either side of a restriction site for the methylation-dependent McrBC enzyme, which corresponds to a HhaI restriction site that was previously reported to be fully methylated on the paternal allele and unmethylated on the maternal (15): 5'-ACTCCTGGAGTGAGGGAAGG-3' and 5'-CAGCTAACCTGAGCTCCATGC-3'. The relative allele proportions in the DNA samples were determined by real time PCR quantifications performed before and after McrBC digestion. The reliability of the method was first demonstrated by using genomic DNA samples of wild-type mice (expected IG-DMR methylation level: 50%) and of Δ IG-DMR mice (16) upon both maternal (expected IG-DMR methylation, 100%) and paternal transmission (expected IG-DMR methylation, 0%) (data not shown).

To assess the allelic specificity of the *Gtl2*-MAR attachment, we performed MAR assays on nuclei from reciprocal hybrid mice obtained by mating (C57BL/6) \times (CBA) F1 mice (*Mus musculus domesticus*) with the JF1 strain (*Mus musculus molossinus*). The following primers for real time PCR were designed on either side of a BglII polymorphic restriction site absent in *M. musculus domesticus*: 5'-GCAGTTCTTTGCTCCACTGAAG-3' and 5'-ATCTTGACACGGGGAAGAG-3'. Amplifications were equally efficient on *M. musculus molossinus* and *M. musculus domesticus* alleles (*i.e.* no PCR biases were detected) (data not shown).

The relative proportions of both alleles were deduced from quantifications of DNA digested or undigested with BglII (see Refs. 2 and 3).

To assess MAR4 allelic attachment, a primer pair, specific to *M. musculus molossinus*, was designed: 5'-GTGCCTCCTCATCTTCTCCC-3' and 5'-GTGCGCTCACTGGGTGACAA-3'. This primer pair does not significantly amplify MAR4 from the *M. musculus domesticus* alleles. The allelic attachment of MAR4 was deduced from the *M. musculus molossinus* MAR4 attachment in MAR assays performed on reciprocal hybrid mice.

Luciferase Reporter Assay

DNA fragments containing the MAR4 and F9 sequences were PCR amplified on genomic DNA with the following primer pairs: MAR4 5'-AAGAGGGTGCCTCAGTCTGG-3' and 5'-TTGGGTTCCACCCTCAGTGG-3'; F9 5'-GGAATGCCACCCACAAGGAG-3' and 5'-TCAGCCTCCGGAACAGCTAG-3'. The resulting 2248- and 2720-bp fragments were cloned in both directions in the pDrive vector (Qiagen) and subsequently cloned into the “pGL3promotor” vectors (firefly luciferase under the SV40 promoter) using XhoI and KpnI restriction sites. MAR4 constructs were then subjected to mutagenesis by using the QuikChange site-directed mutagenesis kit (Stratagene) with the following primers: 5'-GCGAAGATCTCCAGACGTTGTGCTAAGACGAGG-3' and 5'-CCTCGTCTTAGCACAACGTCTGGGAGATCTTCGC-3'.

MAR/HRS at the Mouse *Dlk1/Gtl2* Locus

The resulting Δ MAR4 constructs contain a 2169-bp insert in the sense or antisense directions with a 79-bp deletion that removes the A/T-rich MAR4 sequence from the constructs. Primary bipotential mouse embryonic liver cells, derived from E14.5 embryos, were transfected with these reporter plasmids, together with the *Renilla* luciferase (*Rluc*) control reporter vector pRL-TK vector (Promega) using the jetPEI transfection reagent (PolyPlus Transfection) according to the supplier's protocol. Luciferase activity was determined with a dual luciferase reporter assay system (Promega).

3C-qPCR Assays

3C-qPCR assays were performed as previously described (17). Briefly, the assays were performed on nucleus preparations extracted from E15.5/16.5 or D30 mouse liver samples. 1×10^7 nuclei were formaldehyde cross-linked, digested using EcoRI, and ligated as previously described (17). DNA was then decross-linked (overnight at 65 °C) and purified by classical phenol extraction procedures. An additional digestion using BamHI was then performed on each sample to allow more accurate quantifications (see Ref. 17). Real time PCR was performed on a LighCycler apparatus (Roche Applied Science) using the TaqMan technology (QuantiTect Probe PCR Master Mix; Qiagen). We used a 5'FAM-3'BHQ1 oligonucleotide probe (MWG-Bioctech). For primer sequences used for quantifications, please see the supplemental data. 3C-qPCR data processing was as follows.

Normalization 1: GAPDH Normalization ("Loading Control")

As previously described (17), Ct obtained for each chimerical ligation fragments were processed using parameters of a standard curve (slope and intercept) to obtain quantification values that were normalized to a "GAPDH loading control."

Normalization 2: Noise Band Determination and Normalization of Data to Basal Interaction Level

This normalization compensates for experimental variations and allows comparison between different 3C assays; it replaces the *Pdhb* or *Ercc3* normalizations usually used, (see Ref. 17). The procedures below (Parts A, B, and C) should be followed independently for each sample.

Part A: Removal of "Deviant" Experimental Points—Important: removal can only be done when at least three independent 3C assays have been performed for a given sample. In the present study, three independent assays were performed for both the E16.5 liver and the D30 liver samples. When only one or two independent assays have been performed, select all points and go to Part B of the procedure.

For Step A.1, for each experimental point, calculate the Log of the values (v) normalized to GAPDH (Normalization 1), $[\text{Log } v]$.

For Step A.2, for each fragment (fx), calculate the mean $[m(fx)]$ of $[\text{Log } v]$.

For Step A.3, for each fragment (fx), calculate the standard deviation $[sd(fx)]$ of $[\text{Log } v]$.

For Step A.4, for each experimental points corresponding to a given fragment (fx), calculate $x(fx) = m(fx) - (sd(fx)*k)$, where k is the tolerance factor that we systematically fix at 1.05.

For Step A.5, for each experimental point corresponding to a given fragment (fx), calculate $y(fx) = m(fx) + (sd(fx)*k)$, where k is the tolerance factor that we systematically fix at 1.05.

For Step A.6, for each fragment (fx), select all experimental points for which $x(fx) < [\text{Log } v] < y(fx)$. Remove all other points.

Part B: Determination of the Basal Interaction Level—For Step B.1, for each experimental point selected above (Part A) corresponding to a given fragment (fx), calculate the mean of values (v) normalized to GAPDH (normalization 1) $[m'(fx)]$.

For Step B.2, for each experimental points selected above (Part A) corresponding to a given fragment (fx), calculate the standard deviation of the values normalized to GAPDH v (normalization 1) $[sd'(fx)]$.

For Step B.3, calculate the mean $[M]$ of all the $[m'(fx)]$ values.

For Step B.4, calculate the mean $[SD]$ of all $[sd'(fx)]$ values.

For Step B.5, for each fragment (fx), select all experimental points for which $m'(fx) < M - (SD)$.

For Step B.6, calculate the mean $[MI]$ of all the $[m'(fx)]$ values of the points selected in Step B.5.

For Step B.7, calculate the standard deviation $[SDI]$ of all the $[m'(fx)]$ values of the points selected in Step B.5.

For Step B.8, for each fragment (fx), select all experimental points for which $m'(fx) > MI + (SDI)$.

For Step B.9, calculate the mean $[M2]$ of all the $[m'(fx)]$ values of the points selected in Step B.8.

For Step B.10, calculate the standard deviation $[SD2]$ of all the $[m'(fx)]$ values of the points selected in Step B.8.

The value of $M2$ is the raw basal interaction level that we then use to normalize our data (see Part C, below). $M2 \pm SD2$ is the raw noise band.

Part C: Normalization to the Basal Interaction Level and Determination of the Noise Band—For Step C.1, for each fragment (fx), calculate the normalized mean $M(fx) = m'(fx)/M2$.

For Step C.2, for each fragment (fx), calculate the normalized standard deviation $SD(fx) = sd'(fx)/M2$.

For Step C.3, calculate the normalized basal interaction level = $M2/M2 = 1$.

For Step C.4, calculate the normalized noise band (NB) = $SD2/M2$. The basal interaction level $\pm NB$ is the normalized noise band.

For Step C.5, make a graph showing the distribution of the $M(fx)$ values ("relative cross-linking frequencies") as a function of the distance (in kb) between the fx fragments and the "constant fragment."

RESULTS

Evolutionarily Conserved MAR Features at the *Dlk1/Gtl2* Locus—Bioinformatic analysis of the sequence encompassing both the *Dlk1* and *Gtl2* genes and the entire intergenic region predicted five MARs at this locus in the mouse (Fig. 1A). Three of them are located around the *Dlk1* gene: MAR1 is juxtaposed with the *Dlk1* promoter and MAR2 and MAR3 map downstream of this gene. MAR4 is located within the intergenic region between two previously described "conserved sequences" (*CS9* and *CS10*) (19). The fifth predicted MAR, which we called the *Gtl2*-MAR, is located few kilobase pairs downstream of the IG-DMR (16). Interestingly, despite the fact

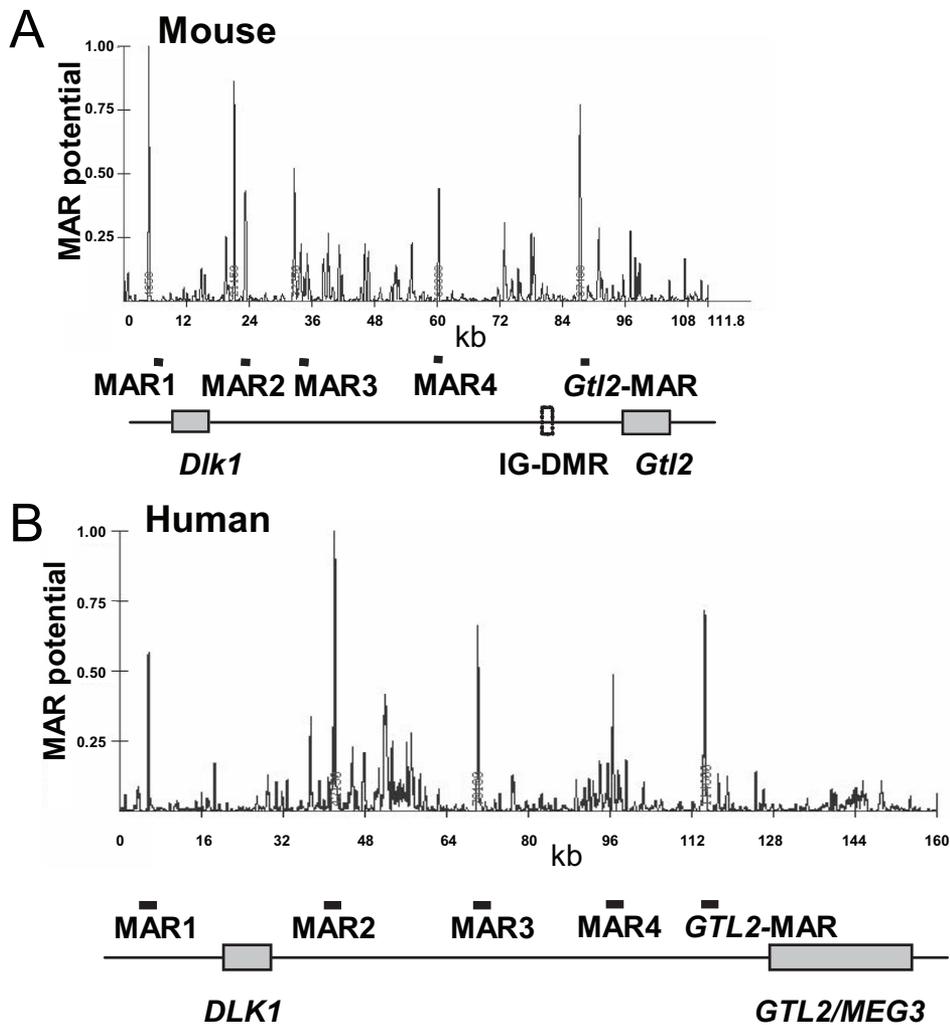


FIGURE 1. Conserved MAR features at the murine and human *Dlk1/Gtl2* loci. Murine (A) and human (B) *Dlk1/Gtl2* locus sequences were analyzed with the MAR-Wiz software (version 1.5). A window size of 300 bp stepped at 50-bp intervals was used. The line graph shows the MAR potential score versus nucleotide positions. A map of the locus is shown below the graphs (solid boxes, gene bodies; open boxes, IG-DMR).

that the human *DLK1/GTL2* locus is considerably longer (~160 kb) than the equivalent mouse locus (~110 kb), both the number and relative positions of the predicted MARs relative to the IG-DMR and the *Dlk1* gene are conserved between both species (Fig. 1). Altogether, the situation is reminiscent of MARs previously identified at the *Igf2/H19* locus (3).

Patterns of Matrix Association at the Mouse *Dlk1/Gtl2* Locus—Using a genomic MAR assay previously set up in our laboratory (3), we analyzed the nuclear matrix association of 18 restriction fragments spread throughout 95 kb of the mouse *Dlk1/Gtl2* locus. An intergenic 15.8-kb sequence that contains a high proportion of repetitive elements (19) was not investigated. The experiments were conducted in parallel on two sets of samples obtained from embryonic day 15.5 (E15.5) and 30-day-old mouse livers, when the genes are expressed or fully repressed, respectively (Fig. 2A). We first analyzed five negative controls (NC, F5, F7, F9, and F11), *i.e.* fragments lacking any particular feature. Four of them (NC, F5, F7, and F11) display attachment levels corresponding to nonspecific background (Fig. 2B). Surprisingly, the F9 fragment, which possesses a mild attachment level at E15.5, displays a higher attachment level in the 30-day-

old mouse liver, when the *Dlk1* and *Gtl2* genes are repressed ($p = 0.1251$). This fragment contains highly repetitive LINE1 elements, and its potential role in the regulation of gene expression or imprinting may need further investigation.

DNA fragments containing the predicted MARs were then investigated. The fragments containing the MAR1, MAR2, MAR3, and the *Gtl2*-MAR were found to be enriched in the MAR fraction in both the E15.5 and 30-day-old mouse liver (Fig. 2B). We can thus conclude that their attachment is not linked to gene expression. In contrast, MAR4 attachment appears strongly correlated with gene expression ($p = 0.0384$). Indeed, its attachment level, which is very high in the E15.5 liver when the genes are strongly expressed, falls to background level in the 30-day-old mouse liver, when the *Dlk1* and *Gtl2* genes are repressed.

We also assessed the enrichment of DNA fragments containing the previously reported conserved sequences (cs9, 10/11, 12, 13, and 14) as well as the differentially methylated regions (*Dlk1*-DMR, IG-DMR, and *Gtl2*-DMR) (15, 19). The conserved sequences are not or are only weakly enriched into the MAR fraction, with exception of the cs10/11 fragment, which displays an attachment pattern similar to the neighboring MAR4 (high attachment in the E15.5 liver; $p = 0.0384$). The attachment pattern of the *Dlk1*-DMR fragment, which essentially corresponds to exonic sequences, seems to follow the *Dlk1* gene transcription pattern. Therefore, as previously shown for other actively transcribed regions (20), attachment of the *Dlk1*-DMR fragment may thus be driven by the transcription machinery itself. The fragment containing the *Gtl2*-DMR (F12) was also found enriched in the nuclear matrix fraction. However, the *Gtl2*-DMR corresponds with the *Gtl2* promoter, and it contains several conserved sequences (cs16 to 19), and therefore it is difficult to hypothesize which feature is responsible for the nuclear matrix attachment.

The IG-DMR Is Retained within the Nuclear Matrix Fraction—Surprisingly, the fragment containing the IG-DMR was also found to be strongly and persistently attached in the mouse liver (Fig. 2B). After performing experiments at a higher resolution, we were able to isolate the attachments of the IG-DMR and of the neighboring *Gtl2*-MAR into two distinct restriction fragments separated by several other weakly attached fragments (data not shown).

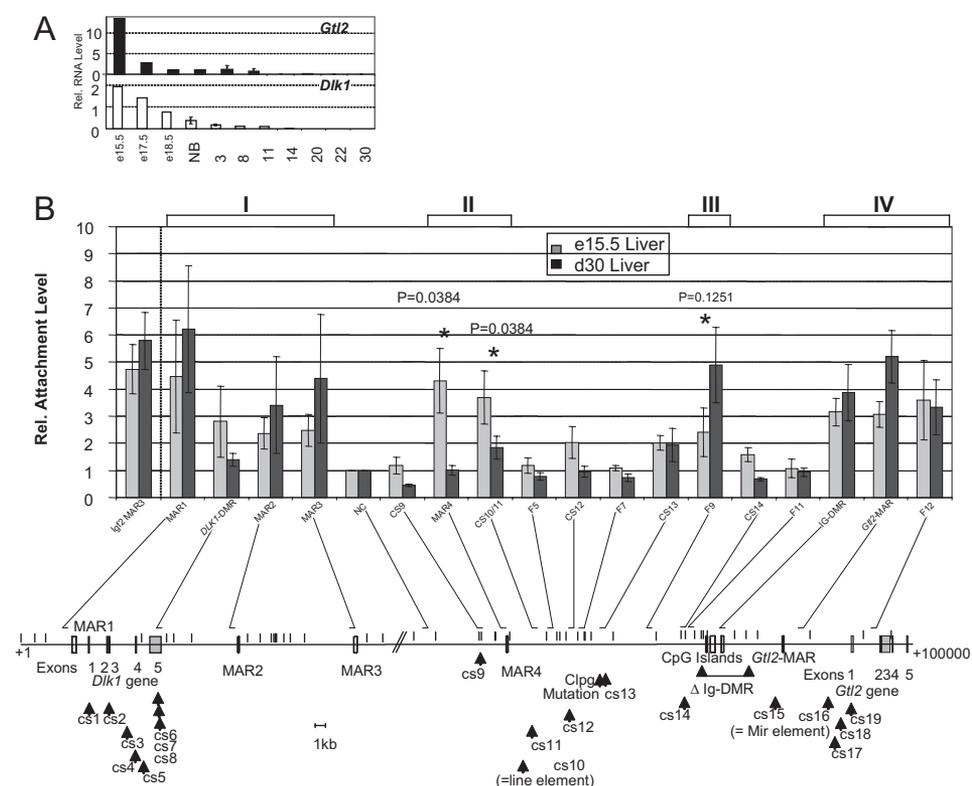


FIGURE 2. Patterns of matrix association at the *Dlk1/Gtl2* locus during the perinatal period in the mouse liver. *A*, *Dlk1* and *Gtl2* expression patterns in the mouse liver. Transcript levels were quantified by real time reverse transcription-PCR at the indicated developmental and postnatal stages (NB, newborn) and normalized to GAPDH mRNA levels as previously described (32). *B*, nuclear halos were prepared from liver nuclei at the indicated developmental stages and digested with XbaI, HindIII, and BamHI. Genomic DNA was then isolated from both the MAR and loop fractions. The graph shows the relative matrix attachment levels expressed as the ratio between the amount of target sequence quantified in the MAR and loop fractions after normalization to a NC, which is given the value 1 (see “Experimental Procedures”). Error bars, S.D. of four (E15.5) and two (D30) independent experiments. Below the bar graphs, the position of restriction sites used in the experiments (vertical bars), real time PCR amplified sequences (small horizontal bars), and potential MARs and exons (filled boxes) are indicated. Fragments that display significant differences of their attachment levels between the two sets of samples are indicated by asterisks. The *p* values were obtained by using the statistical Mann-Whitney *U* test.

The MAR4 Sequence Possesses an Intrinsic Directional Activity That Favors Transcription—Given that the MAR4 attachment pattern correlates with gene expression (Fig. 2*B*), we postulate that this sequence may favor transcription. Furthermore, because the attachment of the intergenic F9 fragment appears to increase with repression (Fig. 2*B*), we hypothesized that this fragment may possess a repressive activity. To test these hypotheses, we performed luciferase assays in primary bipotential mouse embryonic liver cells, derived from E14.5 embryos. DNA fragments containing the MAR4 or F9 sequences were cloned into a vector containing the firefly luciferase gene under the control of the strong SV40 promoter either in a sense or antisense orientation (*i.e.* a direction that is similar or opposite to the *Dlk1/Gtl2* transcriptional orientation at the endogenous locus) (Fig. 3, right part). Luciferase assays using these constructs showed that the F9 sequence displays a 2-fold enhancing activity in these cells (Fig. 3, left part). Therefore, we conclude that this sequence itself has no repressor activity. Similarly to the F9 sequence, the MAR4 fragment also displays a 2-fold enhancer activity when inserted in the sense direction. However, the antisense orientation displays a much stronger activity (2.85 ± 0.16 times the levels observed for the empty pGL3-

promotor plasmid). Interestingly, this directional effect was lost when the 79-bp A/T-rich MAR4 sequence was deleted from the construct (Δ MAR4 construct) (Fig. 3). We conclude that the 79-bp A/T-rich MAR4 fragment contains intrinsic directional activity that favors transcription.

MAR4 and *Gtl2*-MAR Attachments Occur on Both Parental Chromosomes—Because MAR4 attachment correlates with expression and because both *Dlk1* and *Gtl2* genes are monoallelically expressed, it was of interest to determine whether one or both parental copies is recovered from the nuclear matrix. To address this, we used *M. musculus domesticus* \times *M. musculus molossinus* hybrid mice and designed two real time quantitative PCR primer pairs. The first one is able to detect MAR4 independently of the genotype, *i.e.* from both *domesticus* and *molossinus* alleles (*dom* + *mol*), whereas the second primer pair is specific for the *molossinus* alleles (*mol*) (Fig. 4*A*). Only the (*dom* + *mol*) primer pair showed enrichment into the MAR DNA fraction. We conclude that genomic MAR assays retain the MAR4 sequence exclusively from the *M. musculus domesticus* chromosome. This *domesticus*-specific

finding is most likely a genetic background-specific phenomenon, because this attachment is similar in a (*dom* \times *mol*) cross compared with the reverse cross (*mol* \times *dom*) (Fig. 4*A*). We therefore conclude that attachment of the *domesticus* MAR4 allele can occur on both chromosomes and does not depend on its parental origin.

We next examined the allelic attachment of the *Gtl2*-MAR. Here, no allele-specific primer pair could be successfully designed; however, a BglII restriction site specific for the *M. musculus molossinus* allele was identified. We thus used a previously described technique that allows determination of allele ratios using such polymorphic restriction sites and the real time quantitative technique (2). After assessing for the absence of any PCR bias (data not shown), we showed that the *M. musculus domesticus* allele was preferentially retained in our assays (Fig. 4*B*). However, again, this preferential attachment was found to be independent of its parental origin because both the (*dom* \times *mol*) and (*mol* \times *dom*) crosses display similar allele ratios (Fig. 4*B*). We conclude that the *Gtl2*-MAR is equally attached on both parental chromosomes.

Parental Imprinting Influences IG-DMR Attachment—Finally, it was particularly important to determine the relative

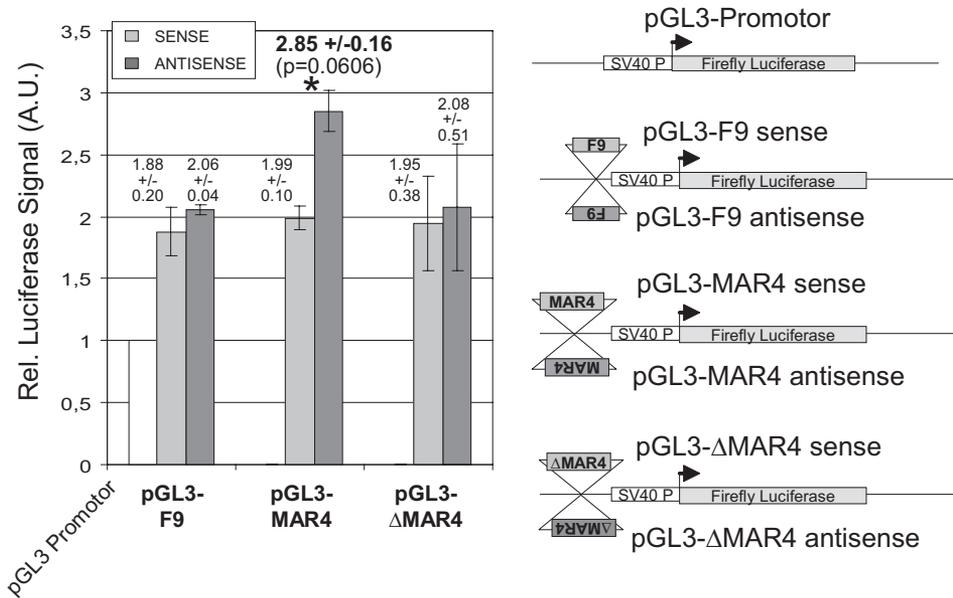


FIGURE 3. Luciferase assays of the MAR4 sequence. Fragments containing the F9, MAR4, or Δ MAR4 sequences were inserted into the pGL3 vector containing the firefly luciferase under the control of the SV40 promoter (right part of the figure). Transfections were performed on bipotential mouse embryonic liver cells derived from E14.5 embryonic liver as described under "Experimental Procedures." The histogram shows the luciferase signals obtained for each construct inserted in sense (light gray bars) or antisense (dark gray bars) orientation relative to the activity of the empty pGL3-Promotor vector (value 1) (white bar). Error bars, S.D. of two independent experiments. The asterisk indicates the significant enhancement observed for the MAR4 antisense construct. The *p* value was obtained by using the statistical Mann-Whitney *U* test.

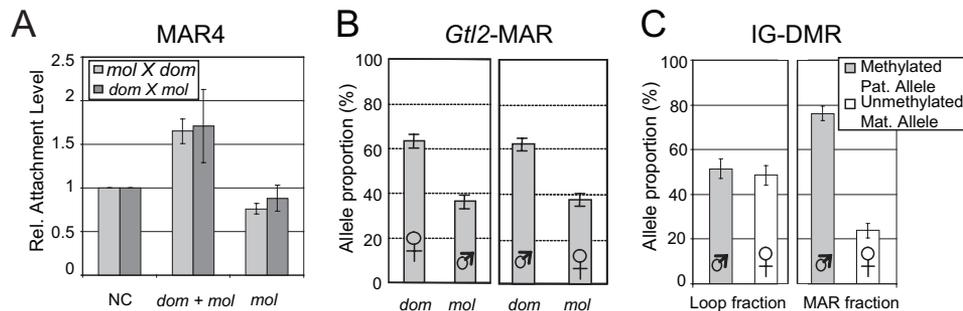


FIGURE 4. Allelic attachment of the MAR4, *Gtl2*-MAR, and IG-DMR regions. A, MAR4 allelic attachment was determined on hybrid mice (*M. musculus molossinus* females \times *M. musculus domesticus* males: light gray bars; reverse cross: dark gray bars) using primer pairs that detected either both alleles (*dom*+*mol*) or only the *molossinus* allele (*mol*). Error bars, S.D. of three independent experiments. B, *Gtl2*-MAR allele ratio was determined on hybrid mice (*M. musculus domesticus* females \times *M. musculus molossinus* males: left panel; reverse cross: right panel) as previously described (2) using a *molossinus*-specific BglIII restriction site. Error bars, S.D. of two independent experiments. C, IG-DMR allele ratio was determined on inbred *M. musculus domesticus* mice (C57BL/6J \times CBA F1) using the methylation-dependent McrBC restriction enzyme as described under "Experimental Procedures." Error bars, S.D. of two independent experiments.

allelic MAR attachment level of the IG-DMR sequence because this sequence displays allele-specific features characteristic of imprinting control regions. This sequence is hypermethylated on the paternal chromosome and completely unmethylated on the maternal copy (15). We used this imprinted allele-specific methylation pattern as a means to determine the parental origin of the alleles. We applied the DNA methylation-dependent McrBC restriction enzyme and determined the ratio between the methylated versus unmethylated alleles that correspond to the allele ratio (2). Therefore, for this experiment, classical inbred mouse strains (*M. musculus domesticus*) could be used instead of hybrid mice, thus precluding any potential PCR or genetic background related bias. We first checked that, in the loop fraction of genomic MAR assays, this region possesses a

mean methylation level of 50%, as a result of its monoallelic methylation pattern (Fig. 4C, left panel). We next examined the MAR fraction and found that the methylated paternal allele is strongly enriched in this fraction (76.2% \pm 3.1) (Fig. 4C, right panel). Because the signal observed in Fig. 2B for the IG-DMR sequence is four times higher than the control sequence (NC), the remaining signal observed on the maternal allele (23.8% \pm 3.1) can be entirely attributed to nonspecific background. Therefore, the IG-DMR sequence is attached exclusively on the methylated paternal allele.

MAR4 Interacts with the IG-DMR but Not the *Dlk1/Gtl2* Promoters—Because MAR4 possesses an intrinsic activity that favors transcription and because genomic MAR assays show that this sequence is retained into the nuclear matrix fraction on both chromosomes in a transcription-dependent manner, we postulated that its potential *in vivo* activity may depend on direct physical interactions with the promoters of the *Dlk1* and *Gtl2* genes. To test this hypothesis, we conducted 3C-qPCR assays (17) centered on the MAR4 sequence in livers of both E16.5 embryos and 30-day-old mice (Fig. 5). In both assays, strong "side effects," extending \sim 15–20 kb on each side of the MAR4, prevented detection of any specific interactions that may occur close to this sequence. However, from 20 to 80 kb away from MAR4, long range interaction frequencies appear to stabilize to a basal level that corre-

sponds to random interactions between the fragment containing the MAR4 and other fragments throughout the locus. Therefore, for each sample (E16.5 or D30 liver), a mean basal interaction level covering the whole locus, was calculated and fixed to a value of 1. The mean of the experimental variations of the points used to calculate this value was then determined and used to define a "noise band," within which cross-linking frequencies can be considered as reflecting random interactions (see "Experimental Procedures"). Local peaks observed outside the side effect zone and above the noise band represent favored (nonrandom) interactions. Weak interactions peaks were found between the restriction fragment containing the MAR4 and fragments -6 , -3 , and $+7$ in the E16.5 liver and fragments -10 , -6 , -3 , and $+9$ in the D30 liver. Interestingly, MAR4

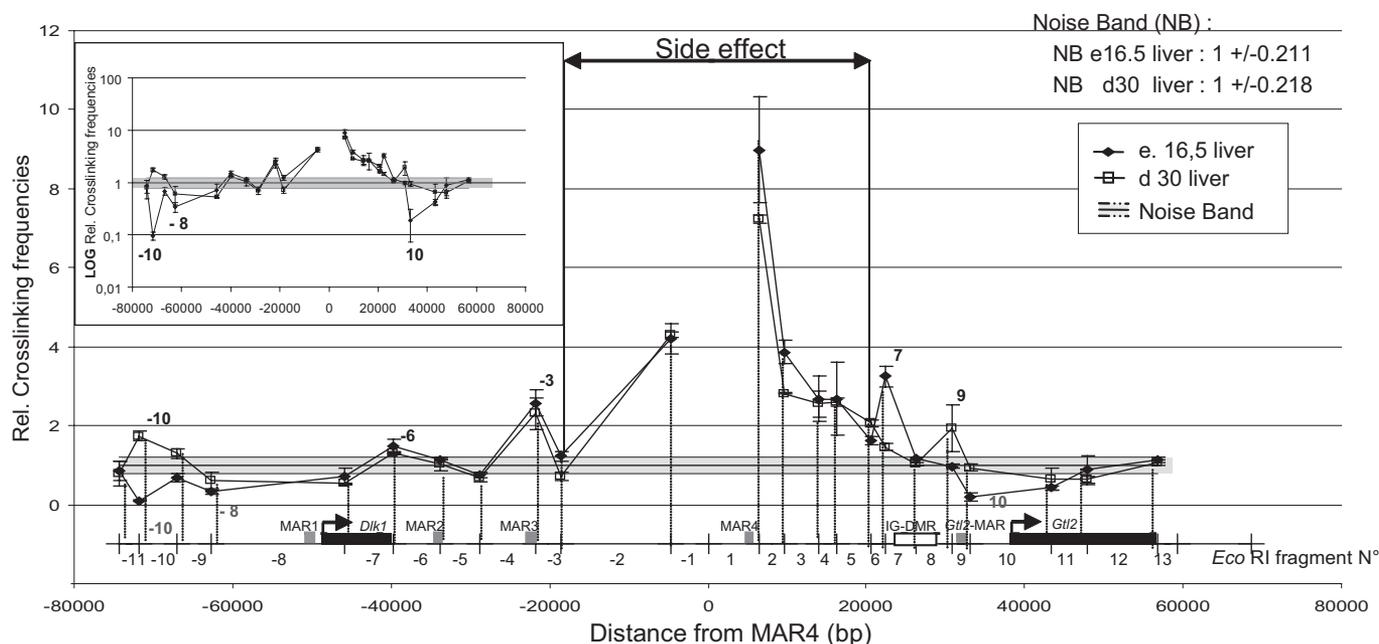


FIGURE 5. Analysis of MAR4 interactions at the mouse *Dlk1/Gtl2* locus by 3C-qPCR assays. The relative cross-linking frequencies between the EcoRI fragment containing the MAR4 sequences and the other fragments of the locus were determined by using the previously described 3C-qPCR procedure (17). 3C-qPCR assays were performed on liver samples issued from E16.5 embryos (filled diamonds) or 30-day-old mice (open squares). Three independent assays were performed for each sample and is depicted in the graph as a gray zone. The region of the graphs that corresponds to side effects is also indicated. The inset shows the same curves where a logarithmic scale has been used to represent the relative cross-linking frequencies; this representation help to stress the local effects observed on the interaction levels of fragments -10 , -8 , and $+10$ containing the *Dlk1/Gtl2* promoters. Error bars, S.D. of three independent experiments.

interacts with the 5' part of the IG-DMR (fragment $+7$) in E16.5 liver, but not in D30 liver. Conversely, interactions with fragments -10 and $+9$ (containing the *Gtl2*-MAR) are only observed in D30 liver. However, in the absence of any functional data about these unexpected interactions, their potential roles remain unclear. Finally, no interaction was found with the fragments containing the *Dlk1/Gtl2* promoters (fragments -8 and $+10$, respectively). Furthermore, in the E16.5 liver, interaction frequencies with these fragments (as well as with fragment -10) even fall below the noise band (see inset in Fig. 5). This means that the frequencies of random interactions between the MAR4 sequence and the *Dlk1/Gtl2* promoters are lower than those observed for the other fragments of the locus. Noticeably, this effect is not observed in D30 liver when the genes are repressed, and it may thus simply reflect local steric impairment around the promoters because of the high transcriptional activity in E16.5 liver.

MAR4/IG-DMR Interaction Occurs on the Paternal Chromosome—To assess whether the MAR4/IG-DMR interaction occurs preferentially on one of the two parental chromosomes, we performed 3C-qPCR assays on liver samples issued from E16.5 mouse embryos heterozygous for the Δ IG-DMR deletion (16). When the deletion was paternally inherited (Fig. 6A), no interaction peak was detected between the fragments containing the MAR4 and the remaining maternally inherited IG-DMR (fragment 7); the interaction level falls to the basal level of random interaction. In contrast, when the deletion was maternally inherited, the interaction is kept to the same level as observed for wild-type mice. The measured value of the relative cross-linking frequency for fragment 7 is 4 for both wild-type and maternal heterozygous mice (Fig. 6B). We conclude that

the MAR4/IG-DMR interaction occurs exclusively on the methylated paternal allele.

Upon maternal inheritance of the deletion, novel interaction peaks are also observed between the MAR4 and fragments 6 and 9 (containing the *Gtl2*-MAR). Because such interactions are not detected upon paternal inheritance or even in the wild-type mouse, we conclude that, when maternally inherited, the deletion chromosome adopts a peculiar genomic organization where the MAR4 is not only interacting with the IG-DMR but also with other regions including the *Gtl2*-MAR.

DISCUSSION

We provide the first detailed analyses of genomic MAR and 3C-qPCR assays at the *Dlk1-Gtl2* locus. We experimentally investigate matrix attachment of five evolutionarily conserved *in silico* predicted MARs at the *Dlk1/Gtl2* locus in mouse liver, at two stages where the genes are either strongly expressed (E15.5/E16.5) or fully repressed (D30). We show that four MARs (MAR1, MAR2, MAR3, and *Gtl2*-MAR) display a constitutive attachment that is independent of gene expression. We also identify MAR4, an intergenic A/T-rich sequence that displays an expression-related attachment. Previously reported conserved sequences (cs) (19) were only mildly attached. Most interestingly, the G/C-rich IG-DMR, which controls imprinting at this locus, displays high matrix attachment levels exclusively on the methylated paternal copy.

MAR sequences have been operationally defined in *in vitro* assays as being able to attach to a purified nuclear matrix or scaffold. Therefore, most analysis of genomic MAR assays have focused on classical A/T-rich MAR sequences. However, after having adapted the sensitive real time PCR technology to the

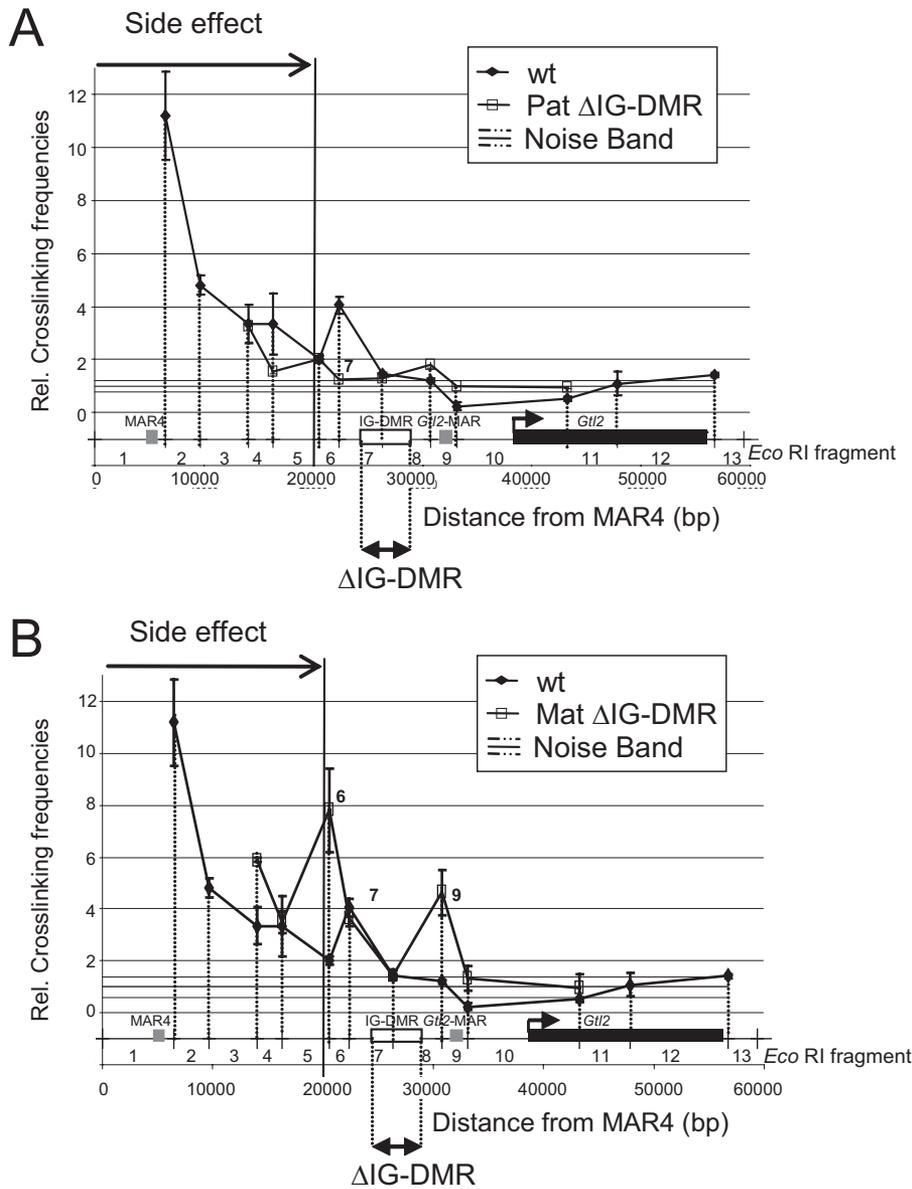


FIGURE 6. Analysis of the MAR4/IG-DMR interaction at the mouse *Dlk1/Gtl2* locus in Δ IG-DMR mutant mice. 3C-qPCR assays were performed on E16.5 liver samples issued from mice that inherited the Δ IG-DMR deletion (16) either paternally (A) or maternally (B) (open squares). The noise band is depicted in the graphs as a gray zone, and the region of the graphs that corresponds to side effects is also indicated. For easier comparisons, data from Fig. 5 corresponding to the E16.5 wild-type liver samples have also been included (filled diamonds). Error bars, S.D. of three independent experiments.

analysis of genomic MAR assays, it now clearly appears that not only such classical A/T-rich MAR sequences can be retained but also other very important regulatory sequences like enhancers (3) or G/C-rich differentially methylated regions/imprinting control regions (this work). In genomic MAR assays, such genomic sequences have the potential to be trapped (in a tissue- or developmental stage-specific manner) into the so-called nuclear matrix upon high salt treatment of nuclei and can thus be recovered. Therefore, we propose a strict and wider operational definition and to rename such sequences the HRSs.

Here, we show that, at the mouse *Dlk1/Gtl2* locus, the IG-DMR is an intergenic HRS specific to the paternal chromosome and that its attachment reflects epigenetic features related to allele-specific expression and imprinting. Such epigenetic fea-

tures may simply result from parent-specific histone modifications or from the paternal-specific allelic methylation. However, we previously showed at the *Igf2/H19* locus that the methylation status does not systematically reflect retention into the nuclear matrix. For example, the placenta-specific *Igf2* MAR0, which is located next to the maternally methylated DMR0, is preferentially attached on the unmethylated paternal allele, whereas the endodermic-specific MAR2, which is located next to the paternally methylated DMR2, is preferentially attached on the hypermethylated paternal allele (3). In the liver, the paternal-specific retention of the IG-DMR into the nuclear matrix compartment likely reflects (and perhaps contributes to) its inactivation on this chromosome.

Altogether, these findings lead us to conclude that, just like the DNase sensitivity assays, the HRS assays reveal epigenetic features of the chromatin. Because the only feature in common with all the HRS identified so far is that they appear to act at a long distance (several tens to a few hundred kb), we propose that the HRS assays most likely reflect epigenetic features involved in higher order chromatin organization of the mammalian genome. Because of technical limitations, the organization of the mammalian genome at that scale remains largely unknown. However, a recent technological breakthrough allowed the identification of chromatin loops of several hundred kb in the mammalian genome (21). Such chromatin loops appear to be linked to gene

expression, and therefore it was suggested that specific higher order chromatin architectures should be associated with gene activity or repression. Furthermore, it becomes increasingly clear that A/T-rich sequences and factors binding such regions play important roles in tissue-specific higher order organization (22). Therefore, MARs, as other HRS, appear as versatile regulatory elements that could combine with differentially methylated regions and other imprinting control regions to confer cell-specific higher order chromatin organization and to control monoallelic expression (3, 23). As proposed previously, it may well be that such higher order chromatin architecture corresponds to a real “genome format” that specifies the transcriptional status of the genes in a tissue- and developmental stage-specific manner (24).

Interestingly, we also identified a 79-bp A/T-rich sequence that we called MAR4 as a novel putative regulatory element that possesses an intrinsic activity that favors transcription. This element is located 9136 bp upstream from a conserved dodecamer motif that contains a single nucleotide polymorphism causing the Callipyge phenotype in sheep (29, 33). Interestingly, the Callipyge mutation was recently shown to enhance bidirectional long range *Dlk1-Gtl2* intergenic transcription in *cis* (25). Therefore, because the novel MAR4 activity appears to be stronger when the MAR4 is inserted in the antisense direction (with respect to the endogenous locus) in a reporter construct, its action may potentially favor transcription of antisense transcripts. Antisense transcripts have been found at numerous loci, and they play important roles in several imprinting mechanisms (26–28); however, such antisense RNAs are very weakly expressed and, although they have been described in the sheep (CLPG1 transcript) (25, 29), they remain to be identified at the *Dlk1/Gtl2* locus in mouse. We also show that the MAR4 sequence does not physically interact with the *Dlk1/Gtl2* promoters. This element could therefore be similar to the DMR2/MAR2 sequence that maintains high *Igf2* transcription levels on mouse chromosome 7 (30). The DMR2/MAR2 was not found to interact with the *Igf2* promoters but was rather proposed to act through a paternal-specific interaction with the imprinting control region (31) and endodermic enhancers (3). Similarly, we could propose that the MAR4 element, which interacts with the IG-DMR on the paternal allele (Figs. 5 and 6), may act at the endogenous *Dlk1/Gtl2* locus by recruiting distant enhancers. However, none of the 3C assays performed at the MAR4 sequence revealed interaction peaks that may correspond to such interactions. Therefore, we speculate that sequestering of the IG-DMR by MAR4 may help to ensure the activity of the *Dlk1* gene on the paternal chromosome. Alternatively, paternally specific MAR4/IG-DMR interaction might contribute to the repression of MAR4-mediated transcriptional activation of the *Gtl2* gene on that chromosome, thus contributing to the process of genomic imprinting at this locus. Inactivation of MAR4 by homologous recombination in the mouse is now required to prove the functional role of this element *in vivo*.

Acknowledgments—We thank Delphine Haouzi from the Urszula Hibner's laboratory at Institut de Génétique Moléculaire who provided the bipotential mouse embryonic liver cells; Jacques Piette, Michaël Weber, and Julie Miro for critical reading of the manuscript; and the staff from the animal unit at Institut de Génétique Moléculaire for technical assistance.

REFERENCES

- Purbowasito, W., Suda, C., Yokomine, T., Zubair, M., Sado, T., Tsutsui, K., and Sasaki, H. (2004) *DNA Res.* **11**, 391–407
- Weber, M., Hagège, H., Lutfalla, G., Dandolo, L., Brunel, C., Cathala, G., and Forné, T. (2003) *Anal. Biochem.* **320**, 252–258
- Weber, M., Hagège, H., Murrell, A., Brunel, C., Reik, W., Cathala, G., and Forné, T. (2003) *Mol. Cell. Biol.* **23**, 8953–8959
- Kono, T., Obata, Y., Wu, Q., Niwa, K., Ono, Y., Yamamoto, Y., Park, E. S., Seo, J. S., and Ogawa, H. (2004) *Nature* **428**, 860–864
- DaRocha, S. T., Tevendale, M., Knowles, E., Takada, S., Watkins, M., and Ferguson-Smith, A. C. (2007) *Dev. Biol.* **306**, 810–823
- Moon, Y. S., Smas, C. M., Lee, K., Villena, J. A., Kim, K. H., Yun, E. J., and Sul, H. S. (2002) *Mol. Cell. Biol.* **22**, 5585–5592
- Raghunandan, R., Ruiz-Hidalgo, M., Jia, Y., Ettinger, R., Rudikoff, E., Riggin, P., Farnsworth, R., Tesfaye, A., Laborda, J., and Bauer, S. (2008) *Stem Cells* **17**, in press
- Schmidt, J. V., Matteson, P. G., Jones, B. K., Guan, X. J., and Tilghman, S. M. (2000) *Genes Dev.* **14**, 1997–2002
- Takada, S., Tevendale, M., Baker, J., Georgiades, P., Campbell, E., Freeman, T., Johnson, M. H., Paulsen, M., and Ferguson-Smith, A. C. (2000) *Curr. Biol.* **10**, 1135–1138
- Davis, E., Caiment, F., Tordoir, X., Cavaille, J., Ferguson-Smith, A. C., Cockett, N., Georges, M., and Charlier, C. (2005) *Curr. Biol.* **15**, 743–749
- Schuster-Gossler, K., Bilinski, P., Sado, T., Ferguson-Smith, A. C., and Gossler, A. (1998) *Dev. Dyn.* **212**, 214–228
- Seitz, H., Youngson, N., Lin, S. P., Dalbert, S., Paulsen, M., Bachellerie, J. P., Ferguson-Smith, A. C., and Cavaille, J. (2003) *Nat. Genet.* **34**, 261–262
- Tierling, S., Dalbert, S., Schoppenhorst, S., Tsai, C. E., Oliger, S., Ferguson-Smith, A. C., Paulsen, M., and Walter, J. (2006) *Genomics* **87**, 225–235
- Lin, S. P., Coan, P., da Rocha, S. T., Seitz, H., Cavaille, J., Teng, P. W., Takada, S., and Ferguson-Smith, A. C. (2007) *Development* **134**, 417–426
- Takada, S., Paulsen, M., Tevendale, M., Tsai, C. E., Kelsey, G., Cattanaach, B. M., and Ferguson-Smith, A. C. (2002) *Hum. Mol. Genet.* **11**, 77–86
- Lin, S. P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., Cavaille, J., and Ferguson-Smith, A. C. (2003) *Nat. Genet.* **35**, 97–102
- Hagège, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forné, T. (2007) *Nat. Prot.* **2**, 1722–1733
- Lutfalla, G., and Uzé, G. (2006) *Methods Enzymol.* **410**, 386–400
- Paulsen, M., Takada, S., Youngson, N. A., Benchaib, M., Charlier, C., Segers, K., Georges, M., and Ferguson-Smith, A. C. (2001) *Genome Res.* **11**, 2085–2094
- Ciejek, E. M., Tsai, M. J., and O'Malley, B. W. (1983) *Nature* **306**, 607–609
- Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., and de Laat, W. (2002) *Mol. Cell.* **10**, 1453–1465
- Cai, S., Lee, C. C., and Kohwi-Shigematsu, T. (2006) *Nat. Genet.* **38**, 1278–1288
- Kurukuti, S., Tiwari, V. K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenkova, V., Reik, W., and Ohlsson, R. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10684–10689
- Weber, M., Hagège, H., Aptel, N., Brunel, C., Cathala, G., and Forné, T. (2005) *Prog. Mol. Subcell. Biol.* **38**, 207–236
- Takeda, H., Caiment, F., Smit, M., Hiard, S., Tordoir, X., Cockett, N., Georges, M., and Charlier, C. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 8119–8124
- Kanduri, C., Thakur, N., and Pandey, R. R. (2006) *EMBO J.* **25**, 2096–2106
- Mancini-Dinardo, D., Steele, S. J., Levors, J. M., Ingram, R. S., and Tilghman, S. M. (2006) *Genes Dev.* **20**, 1268–1282
- Sleutels, F., Zwart, R., and Barlow, D. P. (2002) *Nature* **415**, 810–813
- Freking, B. A., Murphy, S. K., Wylie, A. A., Rhodes, S. J., Keele, J. W., Leymaster, K. A., Jirtle, R. L., and Smith, T. P. (2002) *Genome Res.* **12**, 1496–1506
- Murrell, A., Heeson, S., Bowden, L., Constancia, M., Dean, W., Kelsey, G., and Reik, W. (2001) *EMBO Rep.* **2**, 1101–1106
- Murrell, A., Heeson, S., and Reik, W. (2004) *Nat. Genet.* **36**, 889–893
- Millington, L., Forné, T., Antoine, E., Weber, M., Hemonnot, B., Dandolo, L., Brunel, C., and Cathala, G. (2002) *EMBO Rep.* **3**, 774–779
- Smit, M., Segers, K., Carrascosa, L. G., Shay, T., Baraldi, F., Gyapay, F., Snowder, G., Georges, M., Cockett, M., and Charlier, C. (2003) *Genetics* **163**, 453–456

**Genomic Matrix Attachment Region and Chromosome Conformation Capture
Quantitative Real Time PCR Assays Identify Novel Putative Regulatory Elements
at the Imprinted *Dlk1/Gtl2* Locus**

Caroline Braem, Bénédicte Recolin, Rebecca C. Rancourt, Christopher Angiolini,
Pauline Barthès, Priscillia Branchu, Franck Court, Guy Cathala, Anne C.
Ferguson-Smith and Thierry Forné

J. Biol. Chem. 2008, 283:18612-18620.

doi: 10.1074/jbc.M801883200 originally published online May 5, 2008

Access the most updated version of this article at doi: [10.1074/jbc.M801883200](https://doi.org/10.1074/jbc.M801883200)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

Supplemental material:

<http://www.jbc.org/content/suppl/2008/05/07/M801883200.DC1>

This article cites 32 references, 14 of which can be accessed free at

<http://www.jbc.org/content/283/27/18612.full.html#ref-list-1>