Table 1. Cont.

L. rhamnosus ATCC 53103			*L. paracasei* JCM 8130			*L. casei* ATCC 393		
Locus	**Size (kb)**	**Product description**	**Locus**	**Size (kb)**	**Product description**	**Locus**	**Size (kb)**	**Product description**
LRHM_2085–LRHM_2097	12.3	conserved hypothetical protein				LBCZ_2402–LBCZ_2414	12.5	carbohydrate utilization gene cluster
LRHM_2115–LRHM_2119	8.3	CRISPR-associated protein				LBCZ_2437–LBCZ_2492	66.7	putative cell surface protein, carbohydrate utilization gene cluster
LRHM_2193–LRHM_2198	11.8	cell surface protein, glycosyltransferase				LBCZ_2499–LBCZ_2517	21.1	transposase, conserved hypothetical protein
LRHM_2223–LRHM_2230	7.3	multidrug ABC transporter, hypothetical protein				LBCZ_2616–LBCZ_2643	31.3	carbohydrate utilization gene cluster, transposase
LRHM_2351–LRHM_2356	8.1	multidrug ABC transporter				LBCZ_2678–LBCZ_2694	15.0	transposase
LRHM_2545–LRHM_2597	57.7	carbohydrate utilization gene cluster (region-5)				LBCZ_2698–LBCZ_2704	7.6	PTS transporter
LRHM_2635–LRHM_2651	15.4	carbohydrate utilization gene cluster (region-6)						
LRHM_2779–LRHM_2793	12.5	prophage region III						

doi:10.1371/journal.pone.0075073.t001

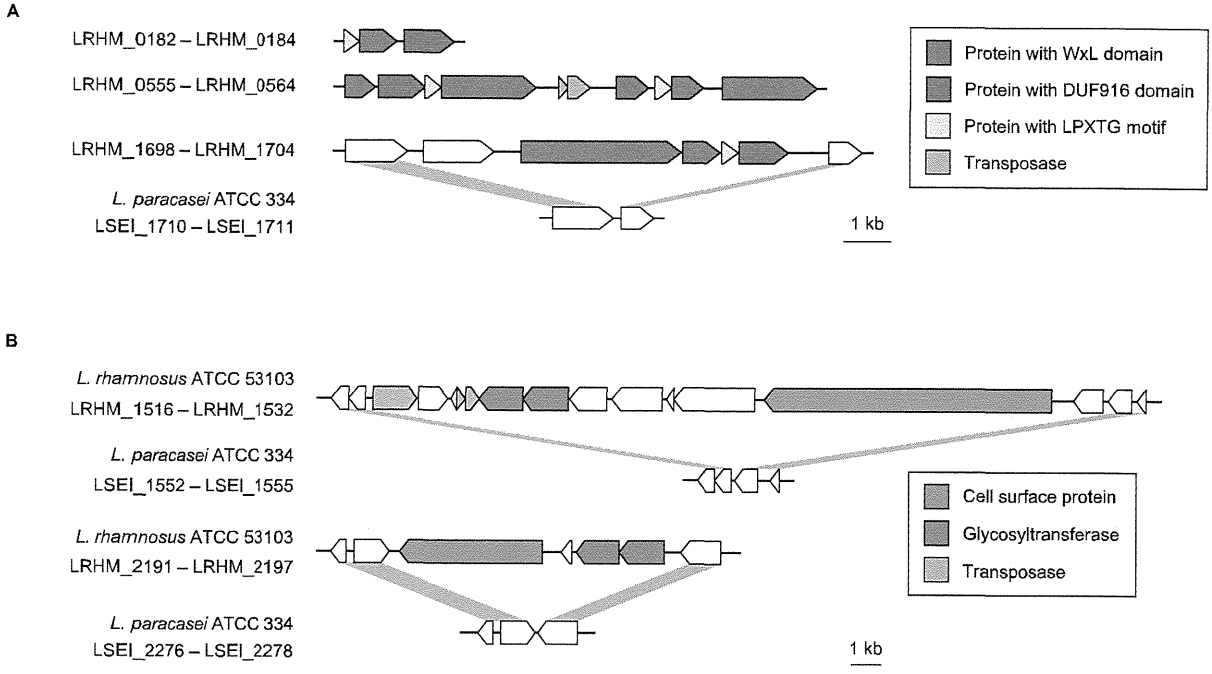


Figure 4. Gene clusters encoding cell surface proteins in *L. rhamnosus* ATCC 53103. (**A**) WxL clusters. (**B**) Putative glycosylated cell-surface protein clusters. Genes and their orientations are depicted with arrows. Gray bars indicate orthologous regions between *L. rhamnosus* ATCC 53103 and *L. paracasei* ATCC 334.
doi:10.1371/journal.pone.0075073.g004

Extracellular Components

Another group has also determined the complete genome sequence of *L. rhamnosus* GG, and revealed the presence of the SpaCBA pili on the cell surface of *L. rhamnosus* GG [9]. SpaA is a backbone-forming major pilin, SpaB is a minor pilin, and SpaC located at the pilus tip is essential for the mucus adherence of *L. rhamnosus* GG [9,30]. The *spaCBA* genes are encoded in the largest GI (LRHM_0376 to LRHM_0466) in *L. rhamnosus* ATCC 53103 (Fig. S5). The *L. paracasei* Zhang, *L. paracasei* BL23, and *L. paracasei* ATCC 334 genomes also encode the *spaCBA* genes (Fig. S5). In contrast, *L. casei* ATCC 393 completely lacks the *spaCBA* genes. The *spaCBA* genes were also encoded in *L. paracasei* COM0101, but the *spaC* gene was truncated by a nonsense mutation [25] (Fig. S5), which probably encodes a non-functional protein. Douillard *et al.*, (2013) clearly showed that the *L. paracasei* strain isolated from Yakult produced no pilus structures by an immunoelectron microscopy using immunogold staining [31]. It has been reported that the adhesion capacity of *L. rhamnosus* GG to Caco-2 cells and intestinal mucus was approximately 10 times that of strain Shirota, which was obtained from Yakult [32]. This may be because *L. rhamnosus* GG encodes the intact SpaCBA and *L. paracasei* COM0101 encodes truncated SpaC. Furthermore, *L. paracasei* JCM 8130, *L. paracasei* BD-II, and *L. paracasei* LC2W also contained truncated *spaC* gene (Fig. S5), and *L. rhamnosus* Lc 705 and ATCC 8530 completely lacked the *spaCBA* genes. The *spaCBA* genes have been found only in the *L. casei* group to date. Because different lineages in *L. casei* strains contained the *spaCBA* genes, it has been suggested that the *spaCBA* genes were not recently acquired [25]. It could thus be speculated that the ancestral strain of the *L. casei* group had encoded the intact *spaCBA* genes and then *spaCBA* may have been lost or disrupted in certain strains of the *L. casei* group.

L. rhamnosus ATCC 53103 had three gene clusters encoding proteins with a C-terminal WxL domain (Fig. 4A). The WxL domain is conserved in the surface proteins in low-GC gram-positive bacteria [33] and attaches to the peptidoglycan on the cell surface [34]. The WxL protein cluster was not found in other sequenced intestinal lactobacilli. The proteins with the WxL domain were present together with the proteins containing the DUF916 domain (PF06030) of unknown function and the small proteins with the LPXTG-like sorting motif, and their gene organizations were similar to that in *L. plantarum* WCFS1 [35]. Of the three WxL protein clusters, one (LRHM_1699 to LRHM_1702) was not conserved in the sequenced *L. paracasei* strains (Fig. 4A, Table 2). There were 14 genes encoding proteins that had both a signal sequence for secretion and an LPXTG-type motif for covalent anchoring to the peptidoglycan matrix (Table 2), and these proteins can be cleaved by sortase. The protein LRHM_1529 was composed of 3,275 amino acid residues, representing the largest protein in this genome, and it contained imperfect repeats consisting of serine, alanine, and aspartic acid. This serine-rich motif has been found in the extracellular proteins in the genomes of other gram-positive bacteria such as *L. plantarum*, *L. johnsonii*, and *Streptococcus pneumoniae* [29,36,37]. The protein LRHM_1529 was encoded in the region (LRHM_1518 to LRHM_1530), which contained two glycosyltransferase genes (Fig. 4B). It has been suggested that glycosyltransferase, encoded by the adjacent genes, caused O-linked glycosylations on the serines in the putative cell surface protein, thus producing mucin-like structures [36]. Similarly, the protein LRHM_2193 had an LPXTG-type motif, and it contained imperfect repeats consisting of serine and alanine and two adjacent glycosyltransferase genes (Fig. 4B). Thus, LRHM_1529 and LRHM_2193 could encode glycosylated cell-surface adhesives. The protein LRHM_1797 (2,357 amino acids) plays an important modulating role in

Table 2. Putative cell surface adherence proteins of *L. rhamnosus* ATCC 53103.

Locus	Size (aa)	Contained domain	SignalP	*L. paracasei* ATCC 334	*L. paracasei* BL23	*L. paracasei* Zhang	*L. paracasei* JCM 8130	*L. casei* ATCC 393	*L. rhamnosus* Lc 705	*L. rhamnosus* ATCC 8530
LRHM_0051	1,492	fibronectin-binding	+	−	−	−	−	−	+	+
LRHM_0182	106	LPXTG	+	−	−	−	−	−	+	+
LRHM_0183	268	WxL	+	+	+	+	+	+	+	+
LRHM_0184	359	DUF916	+	+	+	+	+	+	+	+
LRHM_0426	334	LPXTG (SpaA)	+	−	+	+	+	−	−	−
LRHM_0427	241	LPXTG (SpaB)	+	−	+	+	+	−	−	−
LRHM_0428	895	LPXTG (SpaC)	+	+	+	+	−	−	−	−
LRHM_0555	220	WxL1	+	+	+	+	+	+	+	+
LRHM_0556	340	DUF916	+	+	+	+	+	+	+	+
LRHM_0557	118	LPXTG	+	+	+	+	+	+	+	+
LRHM_0558	688	WxL2	+	−	+	+	−	+	+	+
LRHM_0561	238	WxL1	+	+	+	+	+	−	+	+
LRHM_0562	124	LPXTG	+	+	+	+	+	−	+	+
LRHM_0563	229	WxL1	+	+	+	+	+	−	+	+
LRHM_0564	679	WxL2	+	+	+	+	+	−	+	+
LRHM_1138	401	LPXTG	−	+	+	+	+	+	+	+
LRHM_1331	213	LysM	−	+	+	+	+	+	+	+
LRHM_1393	567	fibronectin-binding	−	+	+	+	+	+	+	+
LRHM_1528	913	Ig-like fold	+	−	−	−	−	+	+	+
LRHM_1529	3,275	LPXTG	+	−	−	−	−	+	+	+
LRHM_1699	351	DUF916	+	−	−	−	−	+	+	+
LRHM_1700	114	LPXTG	+	−	−	−	−	−	+	+
LRHM_1701	262	WxL	+	−	−	−	−	+	+	+
LRHM_1702	1,131	WxL	−	−	−	−	−	+	+	+
LRHM_1797	2,357	LPXTG	−	−	−	−	−	+	+	+
LRHM_2006	1,561	LPXTG	+	−	−	−	−	−	+	−
LRHM_2185	1,973	LPXTG	+	+	+	+	+	+	+	+
LRHM_2193	1,653	LPXTG	−	−	−	−	−	+	+	+
LRHM_2248	388	LPXTG, mucin-binding domain	−	+	+	+	+	+	+	+
LRHM_2279	517	LPXTG (SpaD)	+	+	+	+	+	−	+	+
LRHM_2281	983	LPXTG (SpaF)	+	+	+	+	+	−	+	+
LRHM_2626	1,494	LPXTG	+	−	−	−	−	−	+	+
LRHM_2815	2,603	LPXTG	+	+	+	+	+	−	+	+

*'+' indicates that the orthologous gene is present, and '−' indicates that the orthologous gene is absent.

doi:10.1371/journal.pone.0075073.t002

adhesion to intestinal epithelial cells and biofilm formation [38]. These genes (LRHM_1529, LRHM_1797, and LRHM_2193) were absent in the sequenced *L. paracasei* strains. The presence of a variety of the cell surface adherence proteins could contribute to the probiotic properties of *L. rhamnosus* ATCC 53103.

Conclusions

We determined the complete genome sequences of *L. paracasei* JCM 8130 and *L. casei* ATCC 393, and the draft genome sequence of *L. paracasei* COM0101. Furthermore, we re-annotated the genome of *L. rhamnosus* ATCC 53103. We confirmed that *L. casei* ATCC 393 is distinct from the *L. paracasei* strains previously. Comparative genome analysis revealed 1,682 core genes and genome-wide synteny in the *L. casei* group. Chromosomes of the *L. casei* group contained GIs, many of which are also found at the same loci, suggesting that the chromosomes of the *L. casei* group contain several hypervariable regions at the same loci, which may contribute to the adaptation to each ecological niche. The *spaCBA* pilus gene cluster, which was first identified in *L. rhamnosus* GG, was also found in other strains of the *L. casei* group, but several *L. paracasei* strains including COM0101 contained truncated *spaC* gene. *L. rhamnosus* ATCC 53103 encodes SpaCBA pili, proteins with WxL domain, two glycosylated cell-surface adhesives, and several large proteins with the LPXTG motif. The complete genome sequences of *L. rhamnosus*, *L. paracasei*, and *L. casei* will provide a framework that will help understand the genomic differences between strains within the *L. casei* group.

Supporting Information

Figure S1 Linear representations of the plasmids of *L. casei* 393 and of *L. rhamnosus* Lc 705. Genes and their orientations are depicted with arrows. Several lines connect orthologs with the following colors: red, genes sharing over 95% amino acid identity; orange, genes sharing 70–95% amino acid identity; blue, transposase genes; and green, partially conserved genes.
(EPS)

Figure S2 Genetic relationships among *L. paracasei* strains as defined by multilocus sequence typing. (A) Concatenated sequences of five MLST loci (*ftsZ*, *metRS*, *mutL*, *pgm*, and *polA*) were analyzed as described previously [24]. **(B)** Venn diagram comparing the gene inventories of four *L. paracasei* strains. Data resulted from reciprocal BLASTP analysis. The numbers of shared and unique genes are shown.
(EPS)

Figure S3 COG classification of dispensable protein-coding genes of the *L. casei* group.
(EPS)

Figure S4 Synteny between the chromosomes in the *L. casei* group. Each plot point represents reciprocal best matches by BLASTP comparisons between orthologs.
(EPS)

Figure S5 The *spaCBA* pili cluster arrangement. Genes and their orientations are depicted with arrows.
(EPS)

Table S1 General genomic features of strains sequenced in this study.
(PDF)

Acknowledgments

We thank K. Furuya, C. Shindo, H. Inaba, K. Motomura, and Y. Hattori (The University of Tokyo), and A. Tamura and N. Itoh (Kitasato University) for technical assistance, and Dr. H. Zhang for supplying *L. paracasei* Zhang.

Author Contributions

Conceived and designed the experiments: HM. Performed the experiments: AN MT TT HN SI. Analyzed the data: HT KO MM MH HM. Contributed reagents/materials/analysis tools: AN. Wrote the paper: HT HM.

References

1. Felis GE, Dellaglio F (2007) Taxonomy of Lactobacilli and Bifidobacteria. Curr Issues Intest Microbiol 8: 44–61.
2. Dicks LM, Du Plessis EM, Dellaglio F, Lauer E (1996) Reclassification of *Lactobacillus casei* subsp. *casei* ATCC 393 and *Lactobacillus rhamnosus* ATCC 15820 as *Lactobacillus zeae* nom. rev., designation of ATCC 334 as the neotype of *L. casei* subsp. *casei*, and rejection of the name *Lactobacillus paracasei*. Int J Syst Bacteriol 46: 337–340.
3. Felis GE, Dellaglio F, Mizzi L, Torriani S (2001) Comparative sequence analysis of a *recA* gene fragment brings new evidence for a change in the taxonomy of the *Lactobacillus casei* group. Int J Syst Evol Microbiol 51: 2113–2117.
4. Dellaglio F, Felis GE, Torriani S (2002) The status of the species *Lactobacillus casei* (Orla-Jensen 1916) Hansen and Lessel 1971 and *Lactobacillus paracasei* Collins *et al.* 1989. Request for an opinion. Int J Syst Evol Microbiol 52: 285–287.
5. Acedo-Félix E, Pérez-Martínez G (2003) Significant differences between *Lactobacillus casei* subsp. *casei* ATCC 393T and a commonly used plasmid-cured derivative revealed by a polyphasic study. Int J Syst Evol Microbiol 53: 67–75.
6. Diancourt L, Passet V, Chervaux C, Garault P, Smokvina T, et al. (2007) Multilocus sequence typing of *Lactobacillus casei* reveals a clonal population structure with low levels of homologous recombination. Appl Environ Microbiol 73: 6601–6611.
7. Judicial Commission of the International Committee on Systematics of Bacteria (2008) The type strain of *Lactobacillus casei* is ATCC 393, ATCC 334 cannot serve as the type because it represents a different taxon, the name *Lactobacillus paracasei* and its subspecies names are not rejected and the revival of the name '*Lactobacillus zeae*' contravenes Rules 51b (1) and (2) of the International Code of Nomenclature of Bacteria. Opinion 82. Int J Syst Evol Microbiol 58: 1764–1765.
8. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, et al. (2006) Comparative genomics of the lactic acid bacteria. Proc Natl Acad Sci U S A 103: 15611–15616.
9. Kankainen M, Paulin L, Tynkkynen S, von Ossowski I, Reunanen J, et al. (2009) Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human– mucus binding protein. Proc Natl Acad Sci U S A 106: 17193–17198.
10. Mazé A, Boël G, Zúñiga M, Bourand A, Loux V, et al. (2010) Complete genome sequence of the probiotic *Lactobacillus casei* strain BL23. J Bacteriol 192: 2647–2648.
11. Zhang W, Yu D, Sun Z, Wu R, Chen X, et al. (2010) Complete genome sequence of *Lactobacillus casei* Zhang, a new probiotic strain isolated from traditional homemade koumiss in Inner Mongolia, China. J Bacteriol 192: 5268–5269.
12. Ai L, Chen C, Zhou F, Wang L, Zhang H, et al. (2011) Complete genome sequence of the probiotic strain *Lactobacillus casei* BD-II. J Bacteriol 193: 3160–3161.
13. Chen C, Ai L, Zhou F, Wang L, Zhang H, et al. (2011) Complete genome sequence of the probiotic bacterium *Lactobacillus casei* LC2W. J Bacteriol 193: 3419–3420.
14. Pittet V, Ewen E, Bushell BR, Ziola B (2012) Genome sequence of *Lactobacillus rhamnosus* ATCC 8530. J Bacteriol 194: 726.
15. Morita H, Toh H, Oshima K, Murakami M, Taylor TD, et al. (2009) Complete genome sequence of the probiotic *Lactobacillus rhamnosus* ATCC 53103. J Bacteriol 191: 7630–7631.
16. Doron S, Snydman DR, Gorbach SL (2005) *Lactobacillus* GG: bacteriology and clinical applications. Gastroenterol Clin North Am 34: 483–498.
17. Mahenthiralingam E, Marchbank A, Drevinek P, Garaiova I, Plummer S (2009) Use of colony–based bacterial strain typing for tracking the fate of *Lactobacillus* strains during human consumption. BMC Microbiol 9: 251.
18. Morita H, Kuwahara T, Ohshima K, Sasamoto H, Itoh K, et al. (2007) An improved DNA isolation method for metagenomic analysis of the microbial flora of the human intestine. Microbes Environ 22: 214–222.

19. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27: 4636–4641.
20. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25: 955–964.
21. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8: 785–786.
22. Grissa I, Vergnaud G, Pourcel C (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 35: W52–57.
23. Rasko DA, Myers GS, Ravel J (2005) Visualization of comparative genomic analyses by BLAST score ratio. BMC Bioinformatics 6: 2.
24. Cai H, Rodríguez BT, Zhang W, Broadbent JR, Steele JL (2007) Genotypic and phenotypic characterization of *Lactobacillus casei* strains isolated from different ecological niches suggests frequent recombination and niche specificity. Microbiology 153: 2655–2665.
25. Broadbent JR, Neeno-Eckwall EC, Stahl B, Tandee K, Cai H, et al. (2012) Analysis of the *Lactobacillus casei* supragenome and its influence in species evolution and lifestyle adaptation. BMC Genomics 13: 533.
26. Horvath P, Coûté-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, et al. (2009) Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int J Food Microbiol 131: 62–70.
27. Cai H, Thompson R, Budinich MF, Broadbent JR, Steele JL (2009) Genome sequence and comparative genome analysis of *Lactobacillus casei*: Insights into their niche-associated evolution. Genome Biol Evol 1: 239–257.
28. Makarova KS, Koonin EV (2007) Evolutionary genomics of lactic acid bacteria. J Bacteriol 189: 1199–1208.
29. Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, et al. (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. Proc. Natl. Acad. Sci U S A 99: 14422–14427.
30. Reunanen J, von Ossowski I, Hendrickx AP, Palva A, de Vos WM (2012) Characterization of the SpaCBA pilus fibers in the probiotic *Lactobacillus rhamnosus* GG. Appl Environ Microbiol 78: 2337–2344.
31. Douillard FP, Ribbera A, Järvinen HM, Kant R, Pietilä TE, et al. (2013) Comparative genomic and functional analysis of *Lactobacillus casei* and *Lactobacillus rhamnosus* strains marketed as probiotics. Appl Environ Microbiol 79: 1923–1933.
32. Lee YK, Lim CY, Teng WL, Ouwehand AC, Tuomola EM (2000) Quantitative approach in the study of adhesion of lactic acid bacteria to intestinal cells and their competition with enterobacteria. Appl Environ Microbiol 66: 3692–3697.
33. Kleerebezem M, Boekhorst J, van Kranenburg R, Molenaar D, Kuipers OP, et al. (2003) Complete genome sequence of *Lactobacillus plantarum* WCFS1. Proc Natl Acad Sci U S A 100: 1990–1995.
34. Brinster S, Furlan S, Serror P (2007) C-terminal WxL domain mediates cell wall binding in *Enterococcus faecalis* and other gram-positive bacteria. J Bacteriol 189: 1244–1253.
35. Siezen R, Boekhorst J, Muscariello L, Molenaar D, Renckens B, et al. (2006) *Lactobacillus plantarum* gene clusters encoding putative cell-surface protein complexes for carbohydrate utilization are conserved in specific gram-positive bacteria. BMC Genomics 7: 126.
36. Tettelin H, Nelson KE, Paulsen IT, Eisen JA, Read TD, et al. (2001) Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science 293: 498–506.
37. Pridmore RD, Berger B, Desiere F, Vilanova D, Barretto C, et al. (2004) The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533. Proc Natl Acad Sci U S A 101: 2512–2517.
38. Vélez MP, Petrova MI, Lebeer S, Verhoeven TL, Claes I, et al. (2010) Characterization of MabA, a modulator of *Lactobacillus rhamnosus* GG adhesion and biofilm formation. FEMS Immunol Med Microbiol 59: 386–398.

OPEN ACCESS Freely available online

PLoS one

Characterization of Novel Paternal ncRNAs at the *Plagl1* Locus, Including *Hymai*, Predicted to Interact with Regulators of Active Chromatin

Isabel Iglesias-Platas[1☯], Alex Martin-Trujillo[2☯], Davide Cirillo[3], Franck Court[2], Amy Guillaumet-Adkins[2], Cristina Camprubi[2¤a], Deborah Bourc'his[4], Kenichiro Hata[5], Robert Feil[6], Gian Tartaglia[3], Philippe Arnaud[6¤b], David Monk[2*]

1 Servicio de Neonatología, Hospital Sant Joan de Déu, Fundació Sant Joan de Déu, Barcelona, Spain, **2** Imprinting and Cancer Group, Cancer Epigenetics and Biology Program, Bellvitge Institute for Biomedical Research, L'Hospitalet de Llobregat, Catalonia, Spain, **3** Center for Genomic Regulation, Universitat Pompeu Fabra, Barcelona, Spain, **4** Institut National de la Santé et de la Recherche Médicale, Unité de Génétique et Biologie du Développement, Institut Curie, Paris, France, **5** Department of Maternal-Fetal Biology and Department of Molecular Endocrinology, National Research Institute for Child Health and Development, Tokyo, Japan, **6** Institut de Génétique Moléculaire de Montpellier, Centre National de la Recherche Scientifique and University of Montpellier, Montpellier, France

Abstract

Genomic imprinting is a complex epigenetic mechanism of transcriptional control that utilizes DNA methylation and histone modifications to bring about parent-of-origin specific monoallelic expression in mammals. Genes subject to imprinting are often organised in clusters associated with large non-coding RNAs (ncRNAs), some of which have cis-regulatory functions. Here we have undertaken a detailed allelic expression analysis of an imprinted domain on mouse proximal chromosome 10 comprising the paternally expressed *Plagl1* gene. We identified three novel *Plagl1* transcripts, only one of which contains protein-coding exons. In addition, we characterised two unspliced ncRNAs, *Hymai*, the mouse orthologue of *HYMAI*, and *Plagl1it* (*Plagl1* intronic transcript), a transcript located in intron 5 of *Plagl1*. Imprinted expression of these novel ncRNAs requires DNMT3L-mediated maternal DNA methylation, which is also indispensable for establishing the correct chromatin profile at the *Plagl1* DMR. Significantly, the two ncRNAs are retained in the nucleus, consistent with a potential regulatory function at the imprinted domain. Analysis with catRAPID, a protein-ncRNA association prediction algorithm, suggests that *Hymai* and *Plagl1it* RNAs both have potentially high affinity for Trithorax chromatin regulators. The two ncRNAs could therefore help to protect the paternal allele from DNA methylation by attracting Trithorax proteins that mediate H3 lysine-4 methylation.

Submitted GenBank nucleotides sequences: Plagl1it: JN595789 Hymai: JN595790

Citation: Iglesias-Platas I, Martin-Trujillo A, Cirillo D, Court F, Guillaumet-Adkins A, et al. (2012) Characterization of Novel Paternal ncRNAs at the *Plagl1* Locus, Including *Hymai*, Predicted to Interact with Regulators of Active Chromatin. PLoS ONE 7(6): e38907. doi:10.1371/journal.pone.0038907

Editor: Edward E. Schmidt, Montana State University, United States of America

Received March 16, 2012; **Accepted** May 14, 2012; **Published** June 19, 2012

Funding: This work was supported by Ayuda Merck Serono- Fundación Salud 2000 de Investigación en Endocrinología 2009 (to DM and IIP); Spanish Ministerio de Educacion y Ciencia (grant number SAF2008-1578 to DM); Centre National de la Recherche Scientifique "Projects for International Scientific Cooperation" 34622 (to PA and DM); Agence National de la Recherche (to RF); 'Ligue Contre le Cancer' (to RF); Ligue contre le Cancer comité Hérault and Association pour la Recherche sur la cancer- ARC n° 4980- (for PA). DM is a Ramon y Cajal research fellow and AGA is funded by a FPU studentship. Research in the Neonatal Unit in Hospital Sant Joan de Déu is partially funded by an unrestricted grant from BebéDue Spain. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Research consumables in the Neonatal Unit at Hospital Sant Joan de Déu are partially funded by an unrestricted grant from BebéDue Spain. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials.

* E-mail: dmonk@idibell.cat

¤a Current address: Unitat de Biologia Cel·lular, Facultat de Biociències, Universitat Autònoma de Barcelona, Bellaterra Barcelona, Spain
¤b Current address: Génétique Reproduction et Developpement, Centre National de la Recherche Scientifique and Clermont Université, Clermont-Ferrand, France

☯ These authors contributed equally to this work.

Introduction

Genomic imprinting is an epigenetic form of transcriptional regulation that results in the monoallelic expression of genes from the paternal or maternal allele [1]. Currently there are around 120 confirmed imprinted genes in the mouse, with approximately 60 showing conserved imprinted expression in humans (http://igc.otago.ac.nz/home.html). Imprinted genes have been shown to play important roles in development, and code for proteins with diverse biological activities.

The allele-specific expression of imprinted genes is mediated by CpG rich sequence elements that show allelic DNA methylation [2]. These differentially methylated regions (DMRs) result from methylation deposition during oogenesis or spermatogenesis, specifically by the DNMT3A/DNMT3L *de novo* methyltransferase complex [3–5]. Following fertilization, the allelic methylation is maintained throughout development. In somatic tissues, most DMRs are also marked by allelic histone modifications, highlighting interplay between these two epigenetic systems [6]. Recently, non-coding RNAs (ncRNAs) have

been shown to be important in recruiting histone methyltransferases to imprinted gene promoters, thus revealing the diversity of epigenetic mechanisms involved in the imprinting process [7,8].

The *Plagl1* (also known as *Zac1*) imprinted gene maps to mouse chromosome 10. The human orthologue is located on human chromosome 6 [9,10]. This paternally expressed gene encodes a zinc finger transcription factor with seven C_2H_2-type zinc-fingers that regulates apoptosis and cell cycle [11]. Loss of *PLAGL1* expression is frequently observed in many human tumours, consistent with its proposed role as a tumour-suppressor gene [12]. Over-expression of the human *PLAGL1* gene is thought to be responsible for Transient Neonatal Diabetes Mellitus (TNDM), a genetic disease characterised by severe intrauterine growth restriction and insulin dependence in neonates [13]. This over-expression can result from paternal uniparental isodisomy, paternally inherited duplications of 6q24–q25 or epigenetic mutations in which the maternal allele adopts a paternal epigenotype, resulting in biallelic expression [14]. A paternally expressed ncRNA, *HYMAI*, located in the first intron of human *PLAGL1*, is also over-expressed in TNDM patients, but the function of this transcript remains unknown [13].

To explore the mechanisms regulating *PLAGL1* imprinted expression, we performed a comparative characterisation of the orthologous domain on mouse chromosome 10. We identified numerous paternally expressed ncRNAs, which we propose may be involved in maintaining the paternal allele in a transcriptionally permissive state.

Results

Novel Imprinted *Plagl1* Isoforms

To first determine the size of the *Plagl1* gene in mouse, we interrogated the working draft sequence browser (NCBI26/mm8, Feb 2006). In accordance with previous reports, we find that the *Plagl1* gene covers ~71 kb and contains 12 exons [10]. These include numerous alternatively spliced exons in the 5′UTR originating from two promoter regions embedded within two different CpG islands (Figure 1A). The majority of transcripts arise from the promoter (P1) within the DMR, whereas less abundant transcripts originate from an unmethylated CpG island ~30 kb upstream (P2) (reference EST FJ425893). The open reading frame (ORF) for these transcripts is restricted to the last two exons, resulting in a full-length protein of 705 amino acids. All full-length transcripts share a common 3′UTR, with a polyadenylation signal 24 bp from the stop codon.

As a result of expressed sequence tag (EST) alignments, we identified three additional *Plagl1* transcripts (Figure 1A). A novel *Plagl1* transcript (reference EST BM894919) originates from a unique promoter region (P3) 5′ to the exon 7 acceptor site (gtccaag//GTCTCTT or ctcacag/GTTTGAG) of P1-*Plagl1* transcript, with a 5′UTR that extends at least 300 bp into the upstream intron mapping to an interval containing a cluster of CAGE (5′Cap Analysis Gene Expression) tags. This transcript includes the last three exons and therefore incorporates the full-length *Plagl1* ORF. The remaining two transcripts (reference ESTs CJ065374 and AI607573) originate from within the *Plagl1*-DMR region but terminate after exons 4 and 5 respectively. These different RNAs contain unique 3′UTRs, extending beyond the exon boundaries into the P1-*Plagl1* introns and do not include the *Plagl1* ORF. Northern blot analysis using a *Plagl1* exon 2–3 probe revealed, in addition to the 2 major splice variants, multiple transcripts between 700 bp and 1.7 kb (Figure S1). Using various strategically designed RT-PCR primers, we were able to confirm paternal expression of all novel *Plagl1* transcripts in RNA derived from E18.5 (B x C) F1 mouse tissues (Figure 1B).

Conserved Expression of *Hymai* in Mouse

The human *PLAGL1* region contains the paternally expressed *HYMAI* transcript. This non-coding RNA has a transcription start site located within the *PLAGL1*-DMR. However, DNA sequence from this region shows only weak conservation between humans and mouse (data not shown) and no mouse *Hymai* is described on the UCSC sequence browser or in Genbank databases. We set out to determine whether this non-coding RNA is conserved in mouse. We utilised allelic RT-PCR amplifications restricted to intron 1 of P1-derived *Plagl1* transcript. We observed paternal expression of an RNA in various mouse tissues from E18.5 embryos (Figure 1B). Using 5′ and 3′ RACE, we were able to map the extent of this transcript, which we named '*Hymai*'. We identified four different transcriptional start sites (TSS) for *Hymai*, spread over a 19 bp interval embedded within the *Plagl1*-DMR (Figure S2). Using the same RACE-ready cDNA from E18.5 embryos, we were able to show that P1-*Plagl1* transcript originates from an overlapping 47 bp region, with neither P1-*Plagl1* nor *Hymai* being associated with a TATA-box. Using 3′RACE, we show that *Hymai* terminates ~5 kb from the TSS interval, with multiple 3′ RACE products (last base chr10: 12815696 and chr10: 12815706 of mouse genome NCBI37/mm9), the longest transcript terminating 46 bp after a canonical polyadenylation signal (AATAAA). We were unable to confirm a single band on northern blot analysis, since the expression of this transcript is below the detectable limits of the technique. Analysis of the open reading frame revealed that *Hymai* has no obvious ORF (Figure S 2).

Paternal Expression of a Novel *Plagl1* Internal Transcript, *Plagl1it*

Through examination of the UCSC sequence browser we identified 12 ESTs of various sizes transcribed from the same (+) strand as *Plagl1*, located within intron 5 of P1-*Plagl1*. The largest EST, AK087432, is 2964 bp, representing an intronless transcript with no ORF, that we named *Plagl1 intronic transcript* (*Plagl1it*) (Figure 1B; Figure S2). Using RACE, we found that this transcript initiates within intron 5 of P1-*Plagl1* and is at least 3.6 kb, with its 5′ end overlapping the 3′UTR of the paternally expressed EST AI607573 by ~400 bp. Northern blot analysis confirmed the presence of a faint band of between 3.5–kb (Figure S1). Using RACE and RT-PCR we were unable to link *Plagl1it* to *Plagl1*, confirming this is an independent overlapping transcript and not an alternative *Plagl1* exon or UTR (Figure S2). Using allele-specific RT-PCR, we were able to show that this transcript is expressed solely from the paternal chromosome in different mouse tissues (Figure 1B).

Expression of *Hymai* and *Plagl1it* is Uniformly Low Throughout Development

Next, we set out to analyse the tissue-specificity of expression for the novel transcripts. Using quantitative RT-PCR we determined the abundance of the transcripts in placenta, brain and decapitated embryos at E11.5, E12.5, E14.5, E18.5 and in addition to brain, liver, kidney and muscle from both newborn and adult mice (Figure S1). We observed that *Plagl1* expression was consistently higher than both *Hymai* and *Plagl1it* in all tissues and developmental stages analysed. All genes show a marked decrease in expression after birth, in both newborn and adult tissues.

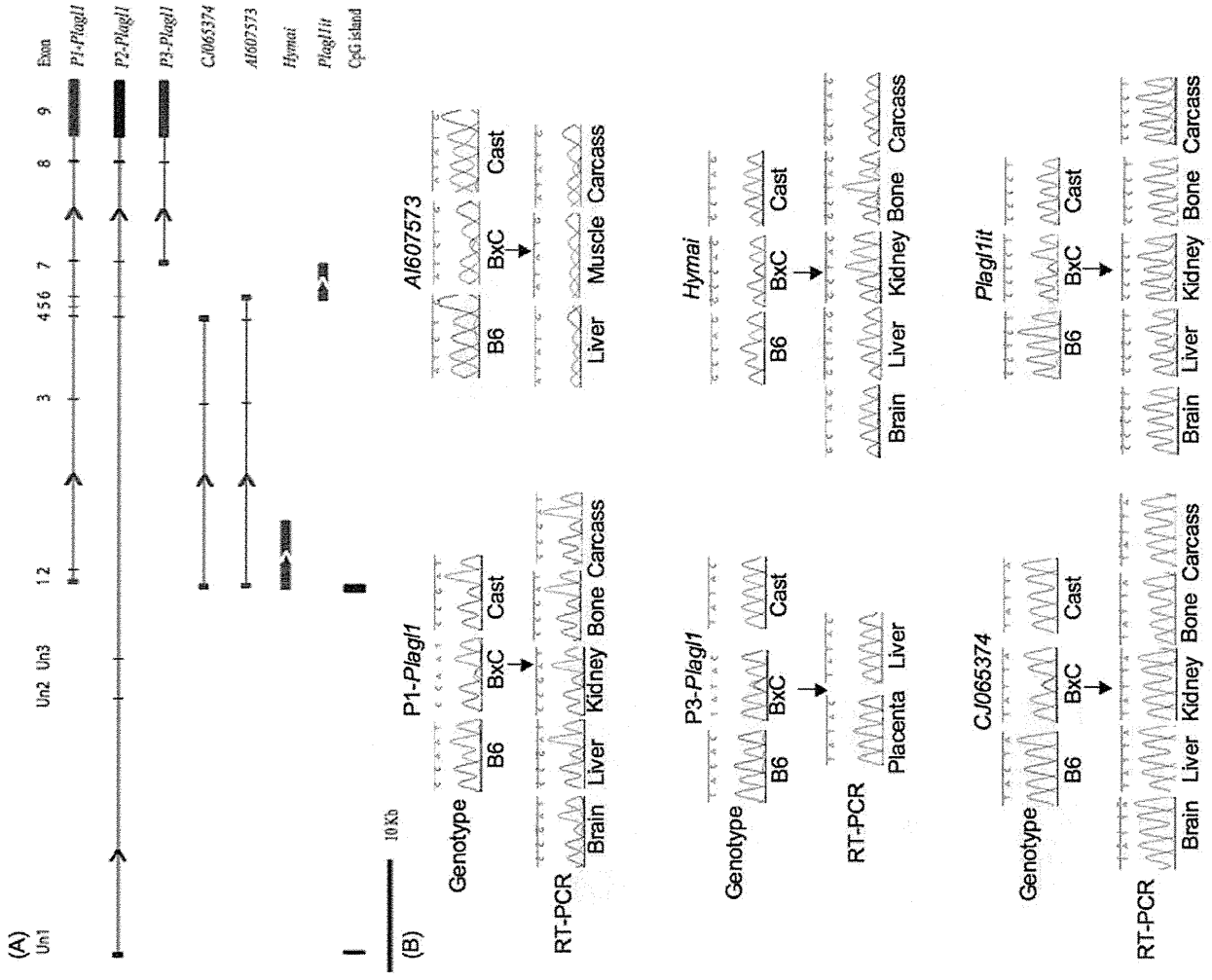


Figure 1. Schematic overview of the mouse chromosome 10 imprinted domain. (A) Map of the *Plagl1* locus, showing the location of the various imprinted transcripts and CpG islands (paternally expressed transcripts are in blue; biallelically expressed transcripts are in grey). Arrows represent direction of transcription. (B) The allelic expression of the various transcripts in embryonic tissues in reciprocal mouse crosses (for clarity only (B×C) F1 tissues are shown).
doi:10.1371/journal.pone.0038907.g001

The ncRNAs are Nuclear Retained, Unstable Transcripts

As a first step to explore whether *Hymai* and *Plagl1it* could have functional roles, we determined the cellular localisation of these ncRNAs. We performed qRT-PCR on nuclear, cytoplasmic and total RNA isolated from mouse embryonic fibroblasts (MEF) cells. The efficiency of the nuclear separation was confirmed using the *U937 snoRNA* and paternally expressed *Aim* ncRNAs that have been shown previously to not be exported to the cytoplasm. We observed residual *Aim* in the cytoplasmic fraction, suggesting slight nuclear RNA contamination only detectable when analysing highly expressed nuclear retained transcripts. The *Igf2r* mRNA was used as a control for a transcript that is exported to the cytoplasm [15]. Quantitative RT-PCR analysis revealed that the *Plagl1* transcript is efficiently exported to the cytoplasm for translation, whereas the *Hymai* ncRNA is retained in the nucleus. The *Plagl1it* transcript is present in both the nucleus and cytoplasm, but is more abundant in the nuclear fraction (Figure 2A).

To determine the stability of *Hymai* and *Plagl1it* in MEFs, actinomycin (ActD) was used to inhibit transcription. We used *C-Myc* and the unspliced *Aim* transcripts as controls for RNAs with short half-life and *Gapdh* and *Igf2r* as control for RNAs with long half-lives [8,15]. Figure 2B shows that after 12 hours treatment with ActD the *C-Myc* and *Aim* mRNAs are largely depleted, whereas *Gapdh* and *Igf2r* are not affected. The *Plagl1* transcript remains abundant under these ActD conditions, suggesting that it is a highly stable transcript. However, both *Hymai* and *Plagl1it* are diminished after 12 hours to levels that are similar to *C-Myc* and *Aim*, indicating that these ncRNAs are unstable transcripts.

DNMT3L is Indispensable for *Hymai*, *Plagl1it* and *Plagl1* Imprinting

DNA methylation inherited from the maternal germline requires the DNMT3L/DNMT3A complex [3,4]. Using bisulphite DNA sequencing, we were able to confirm that the CpG island overlapping the P1-*Plagl1* and *Hymai* transcription start sites

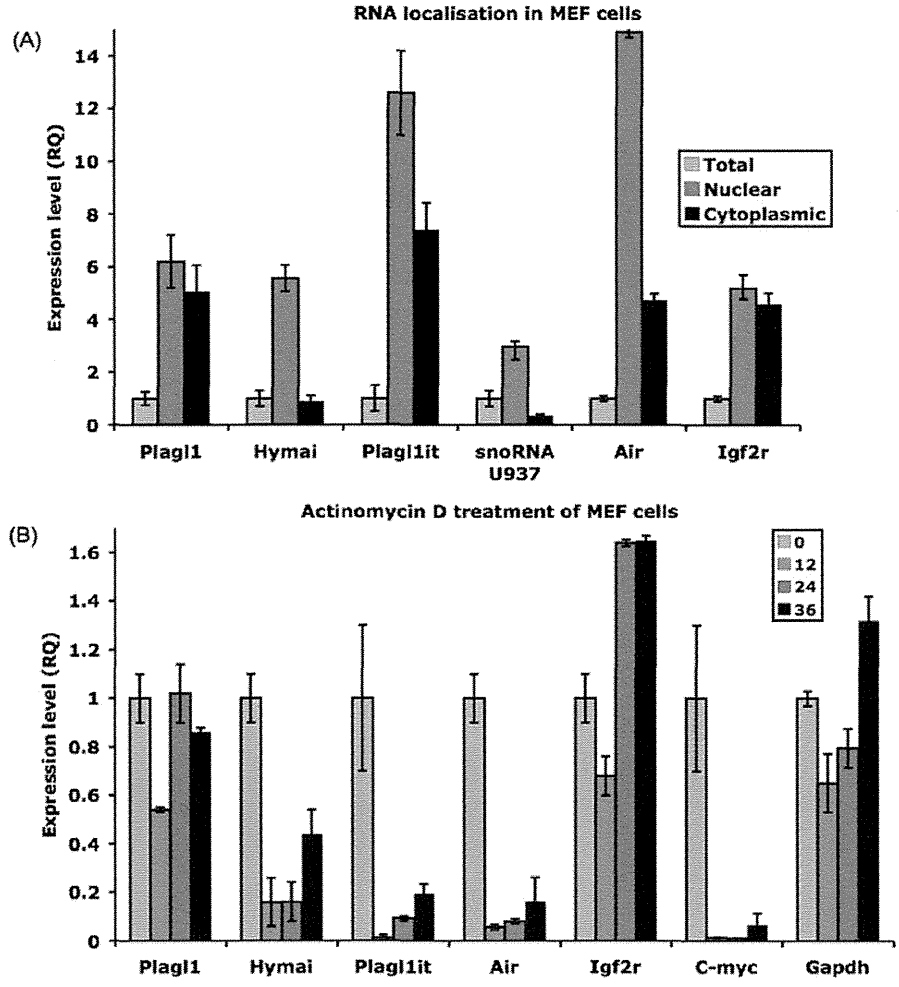


Figure 2. Cellular localization and RNA stability of the ncRNAs. (A) Distribution of the various transcripts in the nuclear (dark grey) and cytoplasmic (black) fractions, compared to total RNA (light grey). *U937 snoRNA* and *Airn* are nuclear-retained controls, whereas *Igf2* is cytoplasm-exported control. (B) Abundance of the various transcripts after exposure to Actinomycin D to determine RNA stability. The relative expression values of the control untreated samples are set to 1 (light grey bars) for each transcript. *C-Myc* and *Airn* are control transcripts for with short half-life; *Gapdh* and *Igf2r* are long half-life controls.
doi:10.1371/journal.pone.0038907.g002

is differentially methylated, whereas P2-*Plagl1* arises from an unmethylated CpG island. The promoters of *Plagl1it* and P3-*Plagl1* initiate from regions of low CpG content that display partial, but not allelic DNA methylation (Figure 3A). To assess if the maternal allelic silencing of *Hymai*, *Plagl1it* and the various *Plagl1* transcripts requires maternal germline DNA-methylation, we used qRT-PCR on mouse embryos that had inherited a deletion of the *Dnmt3l* gene from a homozygous mutant mother [3]. Lack of this essential imprinting factor led to the loss of maternal methylation at the *Plagl1*-DMR, and increased expression of all transcripts in targeted E8.5 embryos due to reactivation of the maternal allele (Figure 3B).

The *Plagl1*-DMR Chromatin Profile Requires Allelic DNA Methylation

Recent studies have suggested that there is a mechanistic link between DNA and histone methylation at imprinted DMRs [6]. To determine if there was a link between allelic DNA-methylation and any histone modifications present at the *Plagl1*-DMR, we first looked for the presence of modifications by allelic chromatin immunoprecipitation on whole embryos followed by discrimination of the parental alleles in the precipitated chromatin fractions. Our analysis focused on different modifications of histone H3 and H4; pan-acetylation of H3, acetylation of H3 lysine-9 (H3K9ac) and H3 lysine 4 dimethylation (H3K4me2) as markers of active chromatin; and the repressive marks of H3 lysine 9 trimethylation (H3K9me3) and H3 lysine 27 trimethylation (H3K27me3), along with the histone H4 lysine 20 trimethylation (H4K20me3).

We ascertained allelic enrichment using a polymorphic region between inbred mouse strains that maps within 200 bp of the CpG island associated with the *Plagl1*-DMR. Within this region H3K4me2 and H3K9ac were strongly enriched specifically on the unmethylated paternal allele (Figure 3C). The same regions showed precipitation of the repressive markers H3K9me3, H3K27me3 and H4K20me3 on the DNA-methylated maternal allele. We extended our analysis to include the promoter regions of P2-*Plagl1*, which maps within an unmethylated CpG island, and *Plagl1it*, whose promoter is not associated with a CpG island. In

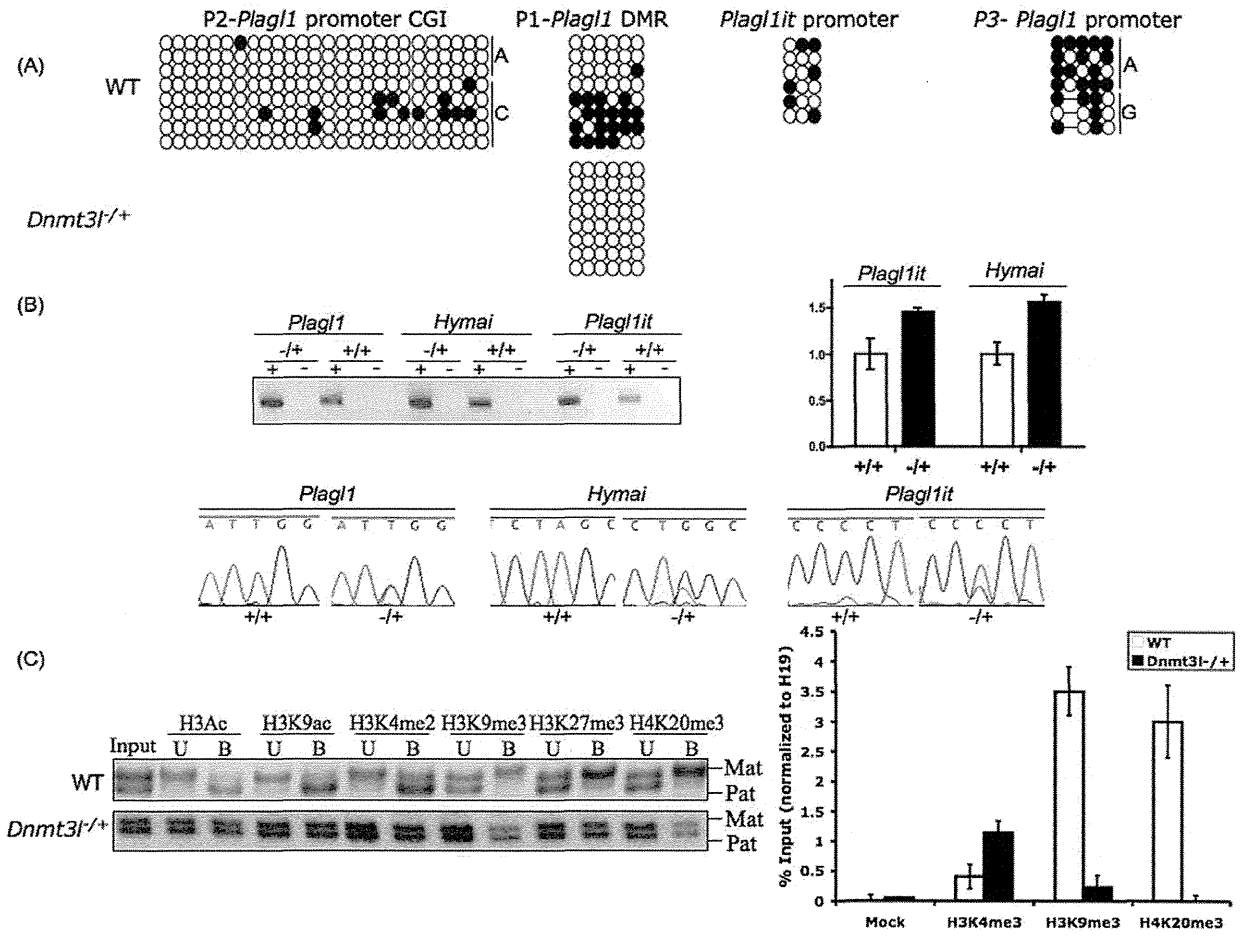


Figure 3. Analysis of *Plagl1* region in *Dnmt3l* −/+. (A) The methylation status of the *Plagl1* promoter regions in wild type +/+ and *Dnmt3l* −/+ embryos examined by bisulphite PCR. Each circle represents a single CpG dinucleotide on a DNA strand, a methylated cytosine (●) or an unmethylated cytosine (○). (B) RT-PCRs on cDNA generated with (+) and without (−) reverse transcriptase show an increase in the expression of the imprinted transcripts in *Dnmt3l*−/+ embryos as a result of reactivation of the maternal allele. (C) The histone modification signature of the *Plagl1*-DMR in wild type B×C embryos, and after targeted deletion of the *Dnmt3l* gene. DNA extracted from antibody bound (B) and unbound (U) chromatin fractions were subject to either qPCR or PCR and SSCP analysis with primers that can discriminate parental alleles.
doi:10.1371/journal.pone.0038907.g003

both cases, we failed to detect allelic precipitation, suggesting that the presence of allelic histone modifications is restricted to the DMR region (data not shown).

To assess whether the allelic histone modifications we observe at the *Plagl1*-DMR require the maternally derived DNA methylation, we performed allelic ChIP on *Dnmt3l* −/+ embryos. In agreement with observations at other imprinted DMRs [6], we detect a dramatic effect on histone modification distribution, with the lack of allelic enrichment due to "paternalization" of the maternal allele, as a result of increased H3K4me3 and a concomitant reduction of H3K9me3 and H4K20me3 (Figure 3C).

Hymai and *Plagl1it* Potentially Interact with Active Chromatin Regulatory Factors

To determine whether *Hymai* and/or *Plagl1it* could be involved in maintaining the active state of the paternal allele of the *Plagl1*-DMR, we performed a prediction of their interaction propensities against four Trithorax proteins (ASH1/KMT2H, MLL1/KTM2A, WDR5, CFP1) using the recently published catRAPID method [16]. CatRAPID allows evaluation of the interaction potential of polypeptides and RNAs using their physiochemical properties, with initial studies revealing high interactions propensities for the ncRNAs *Xist* and *HOTAIR* with Polycomb repressive complex proteins (interaction propensities 76–99% and 69–99%, respectively). In addition, CatRAPID was able to accurately predicted RNA binding of the human RNase P proteins (interaction propensities 68–99%) and discriminate RNA binding (interaction propensity >65%) and non-binding (interaction propensity <5%) proteins of the human ribonuclease mitochondrial RNA processing (MRP) complex [16].

In our analysis we used ncRNAs *Evx1as* and *HOTTIP* as controls because they are known from experimental work to directly recruit MLL1 and WDR5 proteins to *HOX* gene loci [17,18]. We observed moderate to high interaction propensities between *Evx1as* and various functional domains of the MLL1 protein, and between *HOTTIP* and WDR5 (Figure 4A). Interestingly both are predicted to interact strongly with the CFP1 PHD and Ash1 SET-postSET regions. Subsequent analysis using our imprinted ncRNAs revealed that *Hymai* and *Plagl1it* are highly prone to interaction; in particular they have strong binding

propensity with Trithorax proteins. We observe that *Hymai* and *Plagl1it* have negligible propensity for interaction with the Polycomb repressive complex protein EZH2, which trimethylates H3K27 to repress transcription (Figure 4B). Finally, we compared the interaction propensities for *Hymai* and the human orthologue *HYMAI*. We observe that despite having different sequences, and *HYMAI* being subject to splicing, the two transcripts have similar potential interactions (Figure 4C), with 3′ regions having the highest interaction propensities (data not shown). Overall the murine *Hymai* could interact with MLL1 slightly less than human *HYMAI*, but both display high interaction propensities for ASH1 SET-postSET domains and for CFP1 (Figure 4C). Taken together, our results suggest that both *Hymai* and *Plagl1it* may interact with chromatin machinery that confers a permissive chromatin state.

Discussion

Here we show a detailed investigation of the genomic organisation of the mouse *Plagl1* domain. As in humans, *Plagl1* transcripts can originate from multiple promoters, one of which is a DMR previously shown to be methylated in the female germline and therefore likely to be the ICR for this region [10,19]. A second alternative promoter located ~30 kb upstream is within an unmethylated CpG island. This promoter is orthologous with the human P2-*PLAGL1* which gives rise to biallelically expressed transcripts in lymphocytes and pancreas [20]. In mouse, transcription from this promoter is low in somatic tissues, however the primary function of this promoter may be to allow transcription across the P1-*Plagl1* promoter CpG island in growing oocytes. This has been proposed to be important for the establishment of the allelic DNA-methylation at this DMR [21]. In addition to the alternative transcripts of *Plagl1*, we show the presence of two additional ncRNAs, *Hymai* and *Plagl1it*. In keeping with other reported ncRNAs, these are expressed at a lower level than nearby mRNAs, consistent with the hypothesis that ncRNAs may fulfil a regulatory function [22]. We were able to successfully map the TSS and polyadenylation sites for both *Hymai* and *Plagl1it* using RACE-ready cDNAs, indicating that these transcripts comprise rare ncRNAs that are polyadenylated and have 5′-Caps. The reason for the nuclear enrichment of these ncRNA is unknown, as the majority of polyadenylated RNAs are exported to the cytoplasm [23,24]. However, the lack of RNA splicing may be a significant factor in the nuclear retention, as has been described for the various full-length and spliced isoforms of *Airn* [16] and other mRNAs [24].

The precise roles of *Hymai/HYMAI* and *Plagl1it* are unclear, but it is likely that they have a different function to the other known imprinted long ncRNAs such as *Airn* and *Kcnq1ot1* due to their different affinities for chromatin remodelling enzymes. *Airn* and *Kcnq1ot1* have been shown to attract histone methyltransferases G9a/KTM1C and EZH2/KMT6, and are involved in *cis*-silencing of nearby genes [8,24,25]. However, recent studies demonstrated that large ncRNAs can also guide the permissive H3K4 histone methyltransferase machinery to target genes in mouse ES cells and MEFs [17,18] and can act as local enhancers [26]. Thus, unlike other imprinted "repressive" ncRNAs, our data suggests that *Hymai* and *Plagl1it* could act to keep the paternal allele unmethylated and in a transcriptionally permissive state. In fitting with this hypothesis, we observe that *Hymai* and *Plagl1it* are unstable transcripts, which presumably ensures they stay near the site of transcription, preventing their action in *trans* on the maternal allele within the same nucleus. Our *in silico* analysis using catRAPID suggests that *Plagl1it* and the mouse and human *Hymai/HYMAI* may interact with various components of the Trithorax group proteins, with potentially the highest specificity for SET-proSET and zinc finger CXXC domains, in agreement with previous *in vitro* experiments showing that these domains can bind RNA [27,28]. *In vitro* demonstration of these interactions is technically challenging since *Hymai* and *Plagl1it* are not expressed at the levels required for RNA-ChIP in MEF cells. However, we observe that WDR5 does precipitate preferentially on the paternal unmethylated allele of the *Plagl1*-DMR (Figure S3) substantiating our hypothesis.

Conclusions

Germline loss of methylation at the maternal allele of the *PLAGL1*-DMR is known to result in TNDM [13,29]. In addition, *PLAGL1* has been suggested to play a role in numerous cancers, including ovarian, breast and pituitary adenomas, with somatic deletions or gains in methylation resulting in loss of expression of this tumour suppressor gene [30]. We hypothesise that the newly identified ncRNA could potentially guide the H3K4 methylation machinery to the paternal allele of the *PLAGL1*-DMR, and thus protect this region from pathological hypermethylation.

Materials and Methods

Mouse Crosses and Cell Lines

For the analysis of expression, wild type mouse embryos and placentas were produced by crossing C57BL/6 (B) with *Mus musculus castaneus* (C) mice. RNA and DNA from *DNMT3L*$^{-/+}$ mice (B×C) was isolated and extracted as previously described [3]. Animal husbandry and breeding were licensed by Direction Departementale des Services Veterinaires (authorization number 34–104). Homozygous C57BL/6 mice of various gestational ages were used for expression analysis. Mouse embryonic fibroblast cell lines were established from both wild-type (B x C) F1 (Bourc'his laboratory) and C57BL/6 (B) with *Mus musculus molossinus* (JF1) F1 (Feil laboratory) mice. The Institutional Review Board of Bellvitge Institute for Biomedical Research granted scientific and ethical approval for this study (PR232/09).

RNA Preparations

Total RNA from (B×C) F1 wild type embryos, *Dnmt3l*−/+ embryos and MEF cells was isolated using Trizol reagent (Invitrogen) and subjected to double DNase 1 treatment to ensure preparations were free of contaminating DNA. 1 ug of RNA was used for first strand cDNA synthesis using Promega reagents according to the manufacturer's instructions. Nuclear and cytoplasmic RNA was isolated from MEF cells using the Norgen kit (Biotek corporation, Ontario, Canada) following manufacturers instructions. cDNA was generated using 0.5 ug of cytoplasmic, nuclear and total RNA.

Actinomycin Treatment

5×10^5 MEF cells seeded per 10 cm dish were cultured for 36 hrs. At time point 0, the medium was removed; cells were washed with PBS and then incubated with medium supplemented with 10 mg/ml Actinomycin D (dissolved in ethanol). At each time point (0, 12, 24 and 36 hrs) cells from a treated dish were harvested for RNA using Trizol (Invitrogen).

5′ and 3′ RACE

Mouse E18.5 embryo Marathon-Ready cDNA (Clontech) was used for RACE using the Advantage 2 polymerase kit (Clontech). The PCR step was performed with the gene-specific primers located in ESTs for *Plagl1* and *Plagl1it* in combination with nested adaptor oligonucleotides following manufacturers recommenda-

(A)

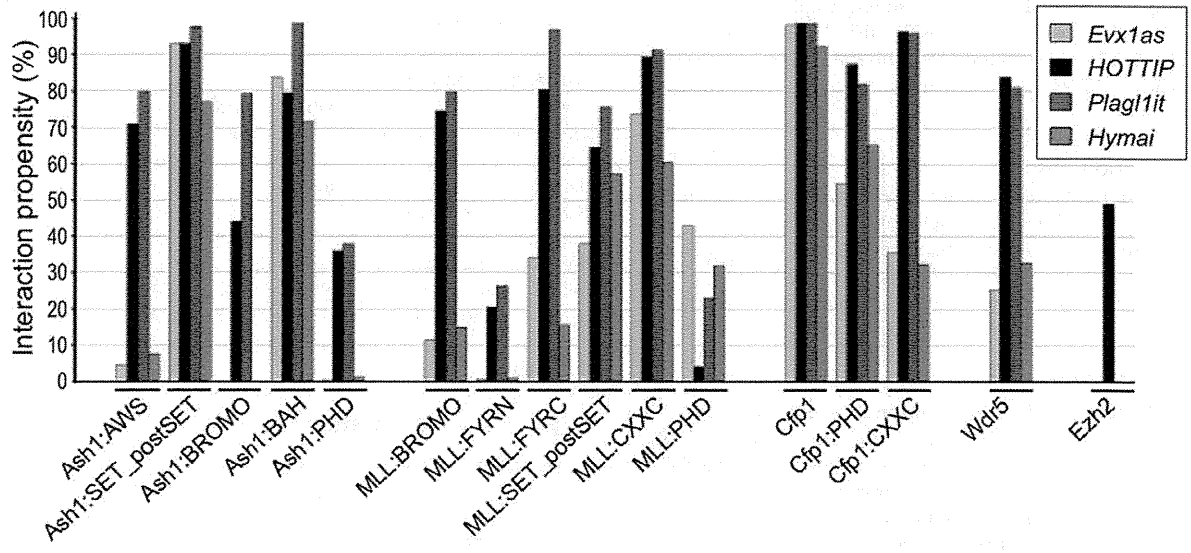
Evx1as interaction propensities

	Binding
MLL1	
SET domain	(moderate) 38
CXXC domain	(high) 74
PHD domain	(moderate) 43
Ash1	
SET domain	(high) 93
PHD domain	(absent) 0
Wdr5	(low) 25
Cfp1	(high) 98

HOTTIP interaction propensities

	Binding
MLL1	
SET domain	(moderate) 64
CXXC domain	(high) 90
PHD domain	(low) 43
Ash1	
SET domain	(high) 93
PHD domain	(modertate) 36
Wdr5	(high) 84
Cfp1	(high) 99

(B)



(C)

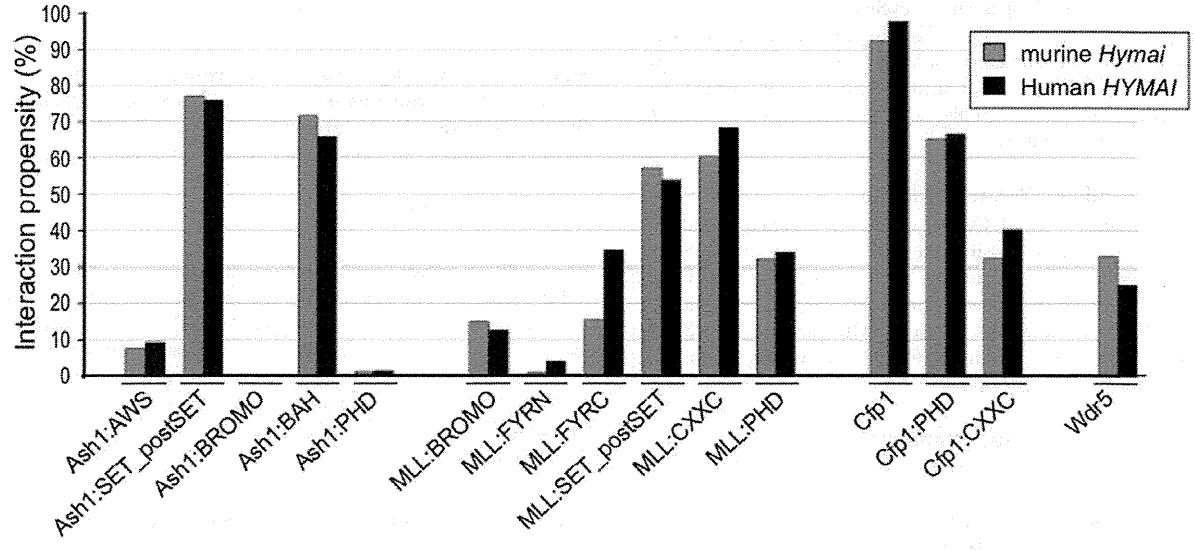


Figure 4. CatRAPID analysis of ncRNA-protein interactions. (A) CatRAPID analysis reveals the interaction propensities of the control ncRNAs; *Evx1as* with MLL1/KTM2A and *HOTTIP* with WDR5. (B) The interaction propensities for the control ncRNAs and for *Hymai* and *Plagl1it* with various components (and sub-domains) of the H3K4 and H3K27 methylation machinery. (C) Similar ncRNA-protein interactions revealed by CatRAPID analysis for *Hymai* (in black) and the human orthologue *HYMAI* (in red).
doi:10.1371/journal.pone.0038907.g004

tions. The PCR products were subcloned into pGEM T-easy vector (Promega) and 20 colonies were sequenced using an ABI prism 3100 DNA sequencer (Applied Biosystems). The full-length sequences of Plagl1it and Hymai have been deposited in GenBank and have been assigned the accession numbers JN595789 and JN595790 respectively.

Northern Blot Analysis

To determine the size of *Plagl1it, Hymai* and the truncated *Plagl1* transcripts, we used custom-made northern blots containing 20 μg of total RNA extracted from CD1 embryos (Zyagen, San Diego, USA). The blots were hybridised with an *β-Actin* probe prior to use to confirm equal loading. Unique sequences for each transcript were amplified by PCR, and the resulting amplicon probes were radiolabelled with (^{32}P)CTP using the Ready-To-Go DNA labelling Beads (Amersham). Hybridizations were carried out overnight at 65°C and washed according to manufacturer's instructions.

RT-PCR Conditions

Allelic RT-PCRs, reactions were performed using primers that flanked polymorphisms. The amplification cycle numbers for each transcript were determined to be within the exponential phase of the PCR, which varied for each gene, but was between 32–42 cycles. The subsequent amplicons were sequenced using both the forward and reverse primers (Table S1 for primer sequences).

Real-time RT-PCR

All PCR amplifications were run in triplicate on a 7900 Fast real-time PCR machine (Applied Biosystems) following the manufacturers' protocol. All primers were optimized using SYBR Green (see additional data file 5 for primer sequences) and melt curve analysis to ensure that amplicons were specific and free of primer-dimer products. Thermal cycle parameters included Taq polymerase activation at 95°C for 10 min for 1 cycle, repetitive denaturation at 95°C for 15 sec, and annealing at 60°C for 1 min for 40 cycles. All resulting triplicate cycle threshold (Ct) values had to be with 1 Ct of each other. The quantitative values for each triplicate were determined as a ratio with the level of *Gapdh* expression (B-actin for actinomycin experiments), which was measured in the same sample, and then averaged to provide relative expression values.

Analysis of Allelic DNA-methylation

Approximately 1 μg DNA was subjected to sodium bisulphite treatment and purified using the EZ GOLD methylation kit (ZYMO, Orange, CA). Bisulphite PCR primers for each region were used with Hotstar Taq polymerase (Qiagen, West Sussex, UK) at 40 cycles and the resulting PCR product cloned into pGEM-T easy vector (Promega) for subsequent sequencing (see Table S1 for primer sequences).

Chromatin Immunoprecipitation (ChIP)

ChIP was carried out on wild type embryos, MEF cells and *Dnmt3l* −/+ embryos. ChIP was performed as previously described [6] using the following Upstate Biotechnology antisera directed against H3ac (06-599), H3K9ac (07-352), H3K4me2 (07-030), H3K9me3 (060904589), H3K27me3 (07-449) and H4K20me3 (07-463) (Upstate Biotechnology). DNA extracted from precipitated chromatin fractions was PCR amplified, and parental alleles were discriminated by either SSCP (*PLAGL1*-DMR) or by direct sequencing. Polymorphisms within 1 kb of the CpG islands were identified by interrogating SNP databases or through genomic sequencing (see Table S1 for primer sequences and location). Only ChIP sample sets that showed enrichment for additional imprinting control regions were used in the analysis. Precipitation levels in the ChIP samples were determined by real-time PCR amplification, using SYBR Green PCR kit (Applied Biosystems). Each PCR was run in triplicate and results are presented as percentage precipitation and normalised to the level of the *H19*-DMD, since methylation at this paternally methylated DMR is unaffected after maternal transmission of the Dnmt3l deleted allele.

catRAPID Analysis

We employed the catRAPID algorithm to predict potential interactions between ncRNAs and proteins [16]. This algorithm was trained using RNA-protein pairs described in the NPInter database. We calculated the average interaction propensity of each RNA species (fragmented into ~1 kb segments because of sequence length restrictions) against complete protein and unique functional domains. Multiple domains adjacent in sequence were joined together (e.g. three PHD domains in MLL1 and SET/proSET regions). In the case of domain association with a size <50 amino acids additional flanking amino acids were added upstream and downstream.

Supporting Information

Figure S1 (**A**) Expression of *Plagl1*, *Hymai* and *Plagl1it* in various tissues from embryos at different gestational stages (e = embryonic day; NB = new born). (**B**) Northern blot analysis using probes specific for *Plagl1* exon 2–3, *Plagl1it* and *Hymai*. A single transcript of less than 4 kb is detected for *Plagl1it* consistent with RACE and RT-PCRs results. Truncated *Plagl1* transcripts, between 700–1.7 kb, correspond to CJ065374 and AI607573.
(TIF)

Figure S2 Mapping of the RACE products to determine the extents of the novel transcripts and open reading frame analysis. (A) The overlapping start sites for P1-*Plagl1* and *Hymai*. (B) analysis for open reading frame using DNA Strider for *Hymai*. (C and D) The 5′ and 3′ ends of *Plagl1it* in relation to *Plagl1* transcripts, and ORF analysis.
(TIF)

Figure S3 Chromatin immunoprecipitation of WDR5 in MEF cells. (A) The upper panel shows PCR amplification of the β-actin promoter control region and *Plagl1*-DMR in the WDR5-ChIP. The lower panel is the genotypes of the input and IP (B x C), showing preferential precipitation of the paternal allele compared to input as calculated from relative area under the nucleotide curve at the SNP position. (B) Confirmation of preferential paternal enrichment by *Hinf*1 RFLP analysis.
(TIF)

Table S1 Table of PCR primer sequences.
(DOC)

Acknowledgments

We are grateful to Humberto Ferreira for help and advice with northern blots, and thank Valeria Romanelli for critical reading of the manuscript.

URLs.

http://igc.otago.ac.nz/home.html.

Author Contributions

Conceived and designed the experiments: IIP GT PA DM. Performed the experiments: IIP AMT DC FC AGA CC PA DM. Analyzed the data: IIP FC DC PA DM. Contributed reagents/materials/analysis tools: DB KH RF GT. Wrote the paper: IIP AMT DC FC AGA CC DB KH RF GT PA DM.

References

1. Reik W, Walter J (2001) Genomic imprinting: parental influence on the genome. Nat Rev Genet 2: 21–32.
2. Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, et al. (2011) Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. Development 5: 811–820.
3. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH (2003) Dnmt3L and the establishment of maternal genomic imprints. Science 21: 2536–2539.
4. Hata K, Okano M, Lei H, Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 129: 1983–1993.
5. Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, et al. (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature 24: 900–903.
6. Henckel A, Nakabayashi K, Sanz LA, Feil R, Hata K, et al. (2009) Histone methylation is mechanistically linked to DNA methylation at imprinting control regions in mammals. Hum Mol Genet 18: 3375–3383.
7. Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, et al. (2008) The Airn noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science 322: 1717–1720.
8. Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, et al. (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell 32: 232–246.
9. Kamiya M, Judson H, Okazaki Y, Kusakabe M, Muramatsu M, et al. (2000) The cell cycle control gene ZAC/PLAGL1 is imprinted–a strong candidate gene for transient neonatal diabetes. Hum Mol Genet. 9: 453–460.
10. Smith RJ, Arnaud P, Konfortova G, Dean WL, Beechey CV, et al. (2002) The mouse Zac1 locus: basis for imprinting and comparison with human ZAC. Gene 292: 101–112.
11. Abdollahi A (2007) LOT1 (ZAC1/PLAGL1) and its family members: mechanisms and functions. J Cell Physiol 210: 16–25.
12. Abdollahi A, Pisarcik D, Roberts D, Weinstein J, Cairnns P, et al. (2003) LOT1 (PLAGL1/ZAC1), the candidate tumor suppressor gene at chromosome 6q24–25, is epigenetically regulated in cancer. J Biol Chem. 2 278: 6041–6049.
13. Mackay DJ, Coupe AM, Shield JP, Storr JN, Temple IK, et al. (2002) Relaxation of imprinted expression of ZAC and HYMAI in a patient with transient neonatal diabetes mellitus. Hum Genet. 110: 139–144.
14. Mackay DJ, Temple IK (2010) Transient neonatal diabetes mellitus type 1. Am J Med Genet C Semin Med Genet.154C: 335–342.
15. Seidl CI, Stricker SH, Barlow DP (2006) The imprinted Airn ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. EMBO J. 25: 3565–3575.
16. Bellucci M, Agostini F, Masin M, Tartaglia GG (2011) Predicting protein associations with long noncoding RNAs. Nat Methods. 8: 444–445.
17. Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, et al. (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res. 18: 1433–1445.
18. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, et al. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472: 120–124.
19. Arima T, Wake N (2006) Establishment of the primary imprint of the HYMAI/PLAGL1 imprint control region during oogenesis. Cytogenet Genome Res. 113: 247–252.
20. Valleley EM, Cordery SF, Bonthron DT (2007) Tissue-specific imprinting of the ZAC/PLAGL1 tumour suppressor gene results from variable utilization of monoallelic and biallelic promoters. Hum Mol Genet. 16: 972–981.
21. Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, et al. (2009) Transcription is required for establishment of germline methylation marks at imprinted genes. Genes Dev. 23: 105–117.
22. Mattick JS, Makunin IV (2006) Non-coding RNA. Hum Mol Genet. 15 Spec No1: R17–29.
23. Fuke H, Ohno M (2008) Role of poly (A) tail as an identity element for mRNA nuclear export. Nucleic Acids Res. 36: 1037–1049.
24. Carmody SR, Wente SR (2009) mRNA nuclear export at a glance. J Cell Sci. 122: 1933–1937.
25. Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, et al. (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell. 40: 939–953.
26. Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, et al. (2010) Long noncoding RNAs with enhancer-like function in human cells. Cell 143: 46–58.
27. Krajewski WA, Nakamura T, Mazo A, Canaani E (2005) A motif within SET-domain proteins binds single-stranded nucleic acids and transcribed and supercoiled DNAs and can interfere with assembly of nucleosomes. Mol Cell Biol. 25: 1891–1899.
28. Hall TM (2005) Multiple modes of RNA recognition by zinc finger proteins. Curr Opin Struct Biol. 15: 367–373.
29. Arima T, Drewell RA, Arney KL, Inoue J, Makita Y, et al. (2001) A conserved imprinting control region at the HYMAI/ZAC domain is implicated in transient neonatal diabetes mellitus. Hum Mol Genet. 10: 1475–1483.
30. Monk D (2010) Deciphering the cancer imprintome. Brief Funct Genomics. 9: 329–339.

CLINICAL REPORT

AMERICAN JOURNAL OF medical genetics PART A

Aberrant Methylation of H19-DMR Acquired After Implantation Was Dissimilar in Soma Versus Placenta of Patients With Beckwith–Wiedemann Syndrome

Ken Higashimoto,[1] Kazuhiko Nakabayashi,[2] Hitomi Yatsuki,[1] Hokuto Yoshinaga,[1] Kosuke Jozaki,[1] Junichiro Okada,[3] Yoriko Watanabe,[3] Aiko Aoki,[4] Arihiro Shiozaki,[4] Shigeru Saito,[4] Kayoko Koide,[1] Tsunehiro Mukai,[5] Kenichiro Hata,[2] and Hidenobu Soejima[1]*

[1]Division of Molecular Genetics & Epigenetics, Department of Biomolecular Sciences, Faculty of Medicine, Saga University, Saga, Japan

[2]Department of Maternal–Fetal Biology, National Research Institute for Child Health and Development, Setagaya, Tokyo, Japan

[3]Department of Pediatrics, Kurume University, Kurume, Japan

[4]Department of Obstetrics and Gynecology, University of Toyama, Toyama, Japan

[5]Nishikyushu University, Kanzaki, Saga, Japan

Manuscript Received: 7 October 2011; Manuscript Accepted: 19 January 2012

Gain of methylation (GOM) at the H19-differentially methylated region (H19-DMR) is one of several causative alterations in Beckwith–Wiedemann syndrome (BWS), an imprinting-related disorder. In most patients with epigenetic changes at H19-DMR, the timing of and mechanism mediating GOM is unknown. To clarify this, we analyzed methylation at the imprinting control regions of somatic tissues and the placenta from two unrelated, naturally conceived patients with sporadic BWS. Maternal H19-DMR was abnormally and variably hypermethylated in both patients, indicating epigenetic mosaicism. Aberrant methylation levels were consistently lower in placenta than in blood and skin. Mosaic and discordant methylation strongly suggested that aberrant hypermethylation occurred after implantation, when genome-wide de novo methylation normally occurs. We expect aberrant de novo hypermethylation of H19-DMR happens to a greater extent in embryos than in placentas, as this is normally the case for de novo methylation. In addition, of 16 primary imprinted DMRs analyzed, only H19-DMR was aberrantly methylated, except for NNAT DMR in the placental chorangioma of Patient 2. To our knowledge, these are the first data suggesting when GOM of H19-DMR occurs.

Key words: Beckwith–Wiedemann syndrome; H19-DMR; aberrant DNA methylation; after implantation

How to Cite this Article:
Higashimoto K, Nakabayashi K, Yatsuki H, Yoshinaga H, Jozaki K, Okada J, Watanabe Y, Aoki A, Shiozaki A, Saito S, Koide K, Mukai T, Hata K, Soejima H. 2012. Aberrant methylation of H19-DMR acquired after implantation was dissimilar in soma versus placenta of patients with Beckwith–Wiedemann syndrome.
Am J Med Genet Part A 158A:1670–1675.

INTRODUCTION

Beckwith–Wiedemann syndrome (BWS) is an imprinting-related condition characterized by macrosomia, macroglossia, and abdominal wall defects (OMIM #130650). The relevant imprinted chromosomal region in BWS, 11p15.5, consists of two independent imprinted domains, *IGF2/H19* and *CDKN1C/KCNQ1OT1*. Imprinted genes within each domain are regulated by two imprinting control regions (ICR), the H19-differentially methylated region (H19-DMR) or KvDMR1 [Weksberg et al., 2010]. Several causative alterations have been identified in patients with BWS: loss of methylation (LOM) at KvDMR1, gain of methylation (GOM) at H19-DMR, paternal uniparental disomy (UPD), *CDKN1C* mutations, and chromosomal abnormality involving 11p15 [Sasaki et al., 2007; Weksberg et al., 2010].

Additional supporting information may be found in the online version of this article.

Grant sponsor: Japan Society for the Promotion of Science; Grant number: 20590330; Grant sponsor: Ministry of Health, Labor, and Welfare; Grant sponsor: National Center for Child Health and Development.

*Correspondence to:
Hidenobu Soejima, M.D., Ph.D., Professor, Division of Molecular Genetics & Epigenetics, Department of Biomolecular Sciences, Faculty of Medicine, Saga University, 5-1-1 Nabeshima, Saga 849-8501, Japan.
E-mail: soejimah@med.saga-u.ac.jp

Article first published online in Wiley Online Library (wileyonlinelibrary.com): 10 May 2012
DOI 10.1002/ajmg.a.35335

Methylation of H19-DMR is erased in primordial germ cells (PGCs) but becomes reestablished during spermatogenesis [Li, 2002; Sasaki and Matsui, 2008]: this methylation regulates the expression of *IGF2* and *H19* by functioning as a chromatin insulator, restricting access to shared enhancers [Bell and Felsenfeld, 2000; Hark et al., 2000]. GOM on the maternal H19-DMR leads to expression of both *IGF2* alleles and silencing of both *H19* alleles. Dominant maternal transmissions of microdeletions and/or base substitutions within H19-DMR have recently been reported in a few patients of BWS with H19-DMR GOM [Demars et al., 2010]. However, when and how the GOM on the maternal H19-DMR occurs is not clear.

Here, we found epigenetic mosaicism in two BWS patients. We also found that GOM at H19-DMR was discordant in blood and skin versus placenta; specifically, methylation levels were lower in placental samples. These findings strongly suggest that aberrant methylation of H19-DMR occurred after implantation. As a result, we expect aberrant de novo methylation happens to a greater extent in embryos than in placentas.

MATERIALS AND METHODS

Patients

Two unrelated patients with sporadic BWS, Patient 1 (BWS047) and Patient 2 (bwsh21-015), were delivered by cesarean in the third trimester of pregnancy. The mothers of both patients conceived naturally. Patient 1 and Patient 2 met clinical criteria for BWS as described by Elliott et al. [1994] and Weksberg et al. [2001], respectively (Table I). The placenta of Patient 1 was large and weighed 1,065 g, but was without any pathological abnormality. The placenta of Patient 2 was also large, weighing 1,620 g, and had an encapsulated placental chorangioma, as reported previously [Aoki et al., 2011]. The standard G-banding chromosome analysis using peripheral blood samples showed no abnormalities in either patient. This study was approved by the Ethics Committee for Human Genome and Gene Analyses of the Faculty of Medicine, Saga University.

Southern Blot Analysis

Genomic DNA was extracted from embryo-derived somatic tissues and the placentas of the patients (Fig. 1). Methylation-sensitive Southern blots with *Bam*HI and *Not*I were employed for KvDMR1, and blots with *Pst*I and *Mlu*I were employed for H19-DMR, as described previously [Soejima et al., 2004]. Band intensity was measured using the FLA-7000 fluoro-image analyzer (Fujifilm, Tokyo, Japan). The methylation index (MI, %) was then calculated (Fig. 1). Southern blots with *Apa*I were used to identify the microdeletion of H19-DMR as described previously [Sparago et al., 2004].

Bisulfite Sequencing and Combined Bisulfite Restriction Analysis (COBRA)

Bisulfite sequencing covering the sixth CTCF binding site (CTS6) was performed. For COBRA, PCR products of each primary imprinted DMR were digested with the appropriate restriction endonucleases and were then separated using the MultiNA Microchip Electrophoresis System (Shimazdu, Japan). The methylation index was also calculated. All PCR primer sets used in this study have been listed in Supplementary Table SI (See Supporting Information online).

DNA Polymorphism Analyses

For quantitative polymorphism analyses, tetranucleotide repeat markers (*D11S1997* and *HUMTH01*) and a triplet repeat marker (*D11S2362*) from 11p15.4–p15.5 were amplified and separated by electrophoresis on an Applied Biosystems 3130 genetic analyzer (Applied Biosystems, NY); data were quantitatively analyzed with the GeneMapper software. The peak height ratios of paternal allele to maternal allele were calculated. A single nucleotide polymorphism (SNP) for the *Rsa*I recognition site in *H19* exon 5 (rs2839703) was also quantitatively analyzed using hot-stop PCR [Uejima et al., 2000]. Band intensity was measured using the FLA-7000 fluoro-image analyzer (Fujifilm).

Mutation Search of H19-DMR

To search for mutations in the binding sites of CTCF, OCT4, and SOX2, we sequenced a genomic region in and around H19-DMR, which included seven CTCF-binding sites, three OCT4 sites, and one SOX2 site.

TABLE I. Clinical Information of BWS Patients

Patient ID	Conception	Birth weight (gestational age)	Clinical features	Karyotype	Placental weight and pathology	Placental–fetal weight ratio
Patient 1 (BWS047)	Natural	4,506 g (36w2d)	macrosomia macroglossia abdominal wall defect hypoglycemia	46,XY	1,065 g no pathological findings	0.236
Patient 2 (bwsh21-015)	Natural	2,540 g (33w5d)	macrosomia macroglossia hypoglycemia renal malformation hepatosplenomegaly	46,XX	1,620 g placental chorangioma	0.638

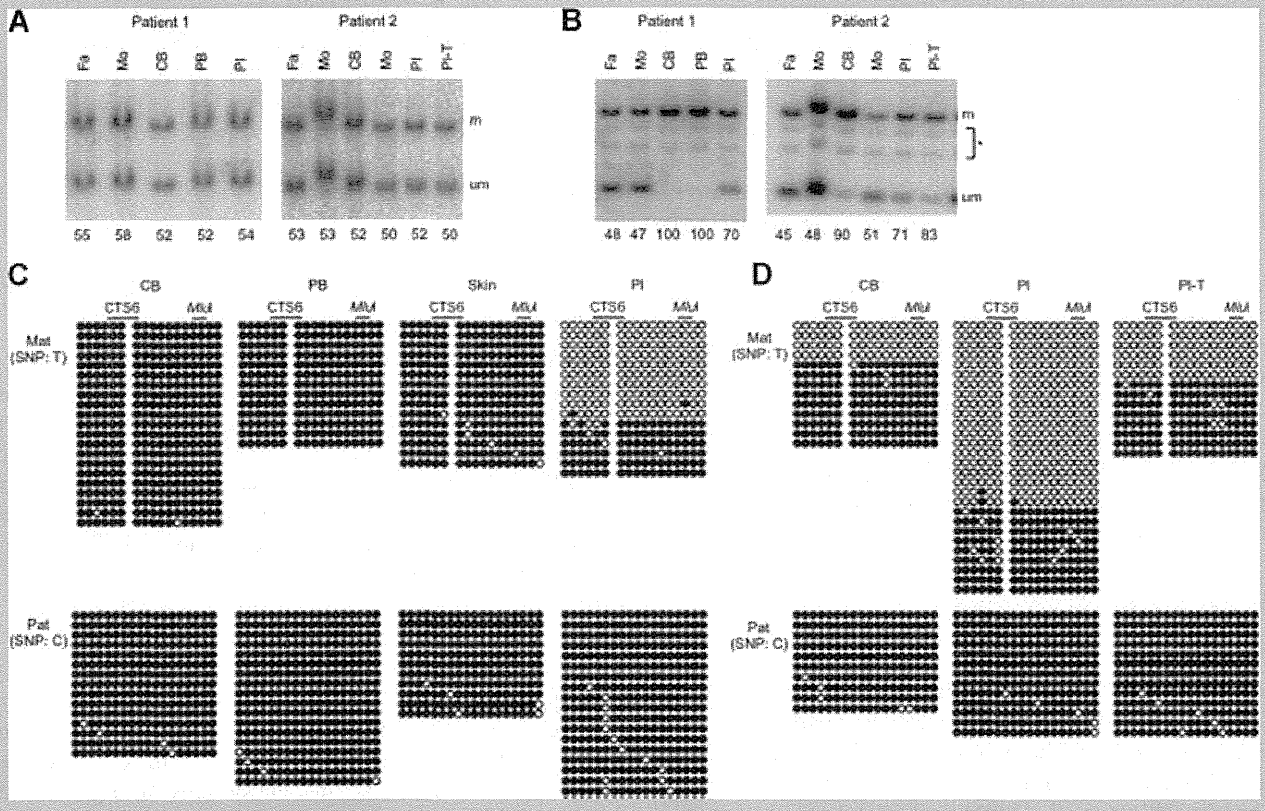


FIG. 1. Methylation analyses of KvDMR1 and H19-DMR. A: Methylation-sensitive Southern blots for KvDMR1. Genomic DNA was extracted from the cord blood, peripheral blood, skin, and placenta of Patient 1 and from the cord blood, placenta, and placental chorangioma of Patient 2. Methylation at KvDMR1 was normal in all samples analyzed. Methylation indices (MI, %) are shown under the figure. B: Methylation-sensitive Southern blots for H19-DMR. The MIs of blood samples were higher than the MIs of placental samples. MI was calculated using the equation [M/(M + U)] × 100, where M is the intensity of the methylated band and U is the intensity of the unmethylated band. C: Bisulfite sequencing of H19-DMR in Patient 1. The two parental alleles were distinguishable by differences in SNPs. Both parental alleles were completely methylated in the cord blood, peripheral blood, and skin samples, and the maternal allele, which is normally unmethylated, was partially methylated in the placenta. D: Bisulfite sequencing of H19-DMR in Patient 2. Methylation of the maternal allele was higher in the cord blood than in the placenta or placental chorangioma. These results were consistent with the results of the Southern blot analysis. We confirmed complete methylation of paternal H19-DMR alleles and complete demethylation of maternal H19-DMR alleles in four normal control placentas that were heterozygous for identifiable SNPs (data not shown). Fa, father; Mo, mother; CB, cord blood; PB, peripheral blood; PI, placenta; PI-T, placental chorangioma; m, methylated band; um, unmethylated band; *, nonspecific bands; Mat, maternal allele; Pat, paternal allele; CTS6, sixth CTCF binding site; *MluI*, a restriction site approximately 80 bp downstream of CTS6 assayed by methylation-sensitive Southern blot and COBRA.

RESULTS

We first examined the methylation status of the two ICRs, KvDMR1, and H19-DMR, at 11p15.5 using methylation-sensitive Southern blot analysis. Methylation at KvDMR1 was normal in all samples collected (Fig. 1A); however, methylation at H19-DMR was aberrant (Fig. 1B). In Patient 1, hypermethylation at H19-DMR was complete in cord blood and peripheral blood samples (MI = 100%), and hypermethylation in the placenta was partial (MI = 70%). In Patient 2, H19-DMR was partially hypermethylated in cord blood (MI = 90%) but less so in the placenta and placental chorangioma (MI = 71% and MI = 83%, respectively). For further investigation of differences in methylation between the patients' somatic tissues and placentas, the CTS6 site was subjected to bisulfite sequencing (Fig. 1C and D). We could distinguish the two parental alleles in each patient sample using informative SNPs (rs10732516 and rs2071094). The maternal allele, which is normally unmethylated, was completely methylated in the cord blood, peripheral blood, and skin from Patient 1. However, in placental samples from Patient 1, the maternal allele was only partially methylated: 36% of all CpGs analyzed were methylated. Similar results were observed in Patient 2: the maternal allele in the cord blood was 68% methylated; however, the maternal allele was only 31% and 55% methylated in the placenta and chorangioma samples, respectively. The paternal alleles, which are normally fully methylated, were fully methylated in all samples. These findings supported the results of the Southern blots. Furthermore, we could not find any microdeletions or mutations in or around H19-DMR,

including seven CTCF-binding sites, three OCT4 sites, and one SOX2 site, indicating that there was no genetic cause of the hypermethylation (Fig. 2A and data not shown).

Next, we analyzed polymorphic markers at 11p15.4–p15.5 to determine whether copy number abnormalities or paternal UPD might be involved in these BWS patients. Although smaller PCR products were more easily amplified, paternal–maternal allele ratios in blood samples were between 0.92 and 1.33, indicating that both parental alleles were equally represented in both patients (Fig. 2B). Therefore, we could rule out copy number abnormality and paternal UPD within the patients' blood. We also investigated maternal contamination in the placenta. *D11S1997* and *HUMTH01* for Patient 1 and the *RsaI* polymorphism in *H19* (rs2839703) for Patient 2 were used for this investigation because the mothers were expected to be homozygous for such polymorphisms. Thus, we investigated contamination of our samples by assessing the homozygosity of the polymorphisms in the mothers. The paternal–maternal ratios in Patient 1 were 0.94 and 1.03, indicating an equal contribution of both parental alleles and suggesting no contamination (Fig. 2B). In Patient 2, the ratios were 0.77 and 0.78 in the placenta and chorangioma, respectively, suggesting a small amount of contamination (Fig. 2C). However, such contamination was too

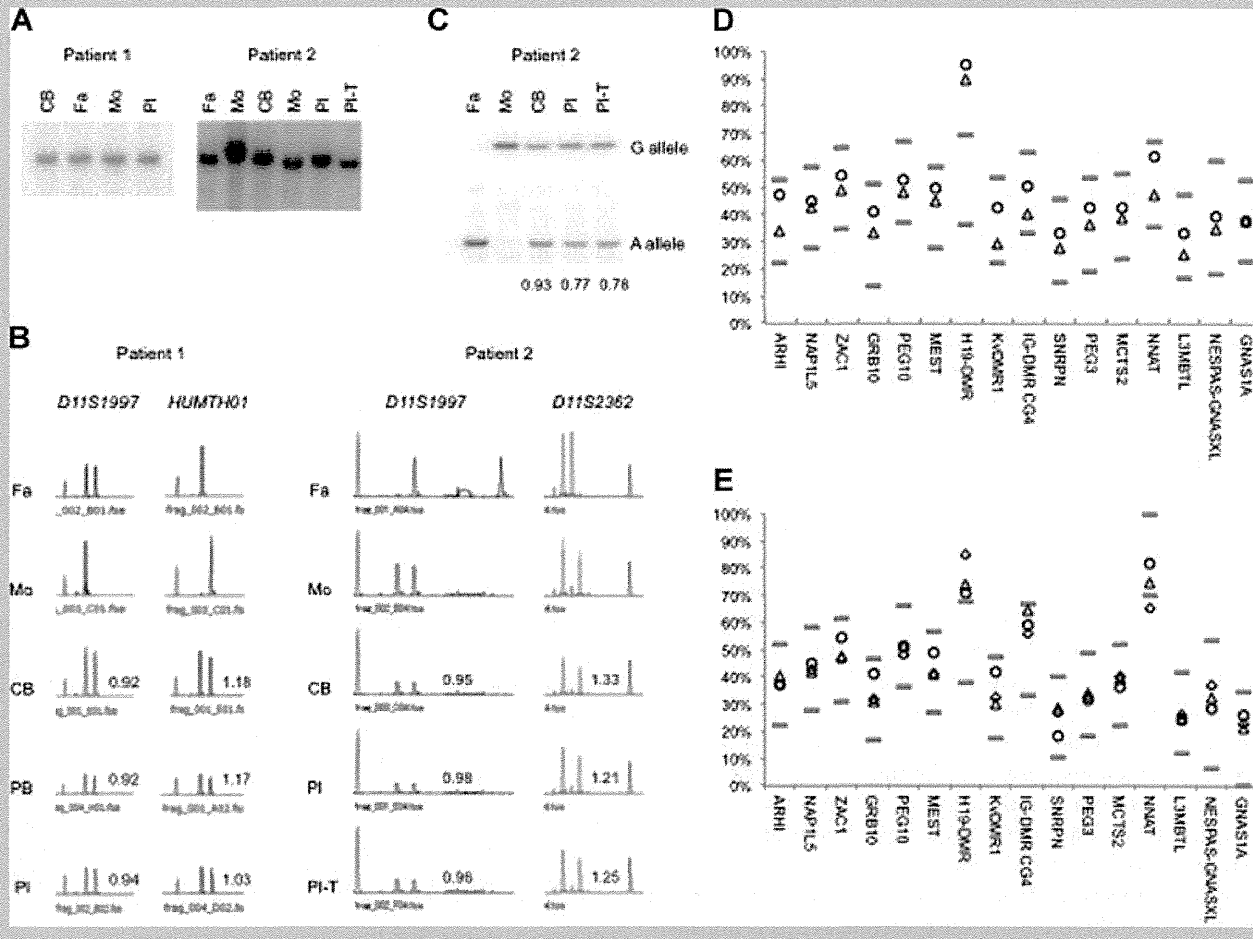


FIG. 2. Microdeletion analysis of H19-DMR, polymorphism analyses, and COBRA of primary imprinted DMRs in embryo-derived and placental samples. A: Southern blots identifying a microdeletion of H19-DMR. A genomic fragment (7.7 kb) generated by *ApaI* digestion, which included the entire H19-DMR, was evident in all samples, indicating that there was no microdeletion in this DMR. B: Microsatellite markers at 11p15.4–p15.5. The peak heights associated with each parental allele in all samples were quantitatively analyzed. The results indicated that both parental alleles were present and equally represented. C: Hot-stop PCR of an *RsaI* polymorphic site in Patient 2. The ratios of paternal allele to maternal allele are shown under the figure. Although the ratios in the placenta and placental chorangioma are lower than in the cord blood, suggesting a small amount of maternal contamination, this was not enough to affect the results of the methylation analyses. COBRA of cord blood (D) and placentas (E), demonstrating that H19-DMR was hypermethylated. CTS6 is contained within H19-DMR. Methylation at other DMRs was normal in all samples, except for methylation at NNAT, which was aberrant in the placental chorangioma. Cord blood and placentas from 24 normal individuals were used as controls. The upper limit of normal methylation was defined as the higher of these two values: (1) the average of controls + 3 SD, or (2) the average + 15%. Similarly, the lower limit of normal methylation was definite as the lower of these two values: (1) the average of controls − 3 SD, or (2) the average − 15%. The upper and lower limits are indicated by gray bars. ○: Patient 1; △: Patient 2; ◇: placental chorangioma of Patient 2.

small to affect the results of the methylation analyses. In addition, sequence analysis did not show any mutations in *CDKN1C* (data not shown). These findings indicated that H19-DMR was aberrantly hypermethylated in both BWS patients and their associated placentas, but the aberrant methylation was consistently lower in the placenta, and that the H19-DMR GOM was strictly an isolated epimutation.

Finally, we analyzed the methylation status of 16 primary imprinted DMRs scattered throughout the genome-using COBRA (Fig. 2D and E). Only H19-DMR showed aberrant methylation among all primary DMRs in all samples, except for NNAT DMR, which was abnormal only in the placental chorangioma, indicating that the *IGF2/H19* imprinted domain was targeted for aberrant methylation in both somatic tissues and the placenta.

DISCUSSION

Methylation associated with parental imprints are erased in PGC and reestablished during gametogenesis in a sex-specific manner. The paternal pronucleus in the zygote undergoes active demethylation; extensive passive demethylation then ensues on maternal and paternal chromosomes during the pre-implantation period. After implantation, de novo methylation results in a rapid increase in DNA methylation in the inner cell mass (ICM), which gives rise to the entire embryo; in contrast, de novo methylation is either inhibited or not maintained in the trophoblast, which gives rise to the placenta [Li, 2002; Sasaki and Matsui, 2008]. The imprinted DMRs, however, escape these demethylation and de novo methylation events that occur in early embryogenesis. H19-DMR GOM in BWS patients is considered an error in imprint erasure in female PGCs [Horsthemke, 2010]. However, H19-DMR GOM, whether with or without microdeletions within H19-DMR, was partial, indicating a mosaic of normal cells and aberrantly methylated cells [Sparago et al., 2007; Cerrato et al., 2008]. These findings demonstrated that aberrant hypermethylation at H19-DMR was acquired after fertilization, although the precise timing was unknown.

Both participants in this study had isolated GOM at H19-DMR. The partial and variable hypermethylation among samples suggested epigenetic mosaicism. Furthermore, methylation levels in the placentas were lower than those in the blood and skin, suggesting that the aberrant methylation was acquired after implantation—when genome-wide de novo methylation normally occurs. Aberrant de novo methylation at H19-DMR is expected to be more widespread in the embryo than in the placenta, as this is normally the case for de novo methylation [Li, 2002; Sasaki and Matsui, 2008]; this disparity in efficiency could lead to the discordance between hypermethylation in trophoblast-derived placenta and that in embryo-derived blood and skin. This hypothesis is supported by a mouse experiment in which a mutant maternal allele harboring a deletion of four CTCF binding sites was hypomethylated in oocytes and blastocysts, yet was highly methylated after implantation [Engel et al., 2006]. To our knowledge, this is the first evidence demonstrating that aberrant hypermethylation of maternal H19-DMR is acquired after implantation in humans.

We found that of 16 primary imprinted DMRs analyzed, only H19-DMR showed aberrant methylation; even methylation at IG-DMR CG4, another paternally methylated, primary imprinted DMR, was normal in our patients. Although we only studied two patients, this finding indicated that the *IGF2/H19* imprinted domain in both the embryo and placenta was more susceptible than other imprinted domains to aberrant methylation acquired after implantation.

In conclusion, we found that methylation of H19-DMR was discordant in embryo-derived somatic tissue and placenta, strongly suggesting that the aberrant de novo methylation occurred after implantation. However, the precise mechanism of isolated H19-DMR GOM is still unknown. Since no mutations in *CTCF*, an important trans-acting imprinting factor, were found in these patients with isolated GOM at H19-DMR, the potential for mutations in the OCT and SOX transcription factors should be investigated because mutations of OCT-binding sites have previously been found in a few patients with H19-DMR GOM [Cerrato et al., 2008; Demars et al., 2010].

ACKNOWLEDGMENTS

This study was supported, in part, by a Grant-in-Aid for Scientific Research (C) (No. 20590330) from the Japan Society for the Promotion of Science, a Grant for Research on Intractable Diseases from the Ministry of Health, Labor, and Welfare, and a Grant for Child Health and Development from the National Center for Child Health and Development.

REFERENCES

Aoki A, Shiozaki A, Sameshima A, Higashimoto K, Soejima H, Saito S. 2011. Beckwith–Wiedemann syndrome with placental chorangioma due to H19-differentially methylated region hypermethylation: A case report. J Obstet Gynaecol Res 37:1872–1876.

Bell AC, Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 405:482–485.

Cerrato F, Sparago A, Verde G, De Crescenzo A, Citro V, Cubellis MV, Rinaldi MM, Boccuto L, Neri G, Magnani C, D'Angelo P, Collini P, Perotti D, Sebastio G, Maher ER, Riccio A. 2008. Different mechanisms cause imprinting defects at the IGF2/H19 locus in Beckwith–Wiedemann syndrome and Wilms' tumour. Hum Mol Genet 17:1427–1435.

Demars J, Shmela ME, Rossignol S, Okabe J, Netchine I, Azzi S, Cabrol S, Le Caignec C, David A, Le Bouc Y, El-Osta A, Gicquel C. 2010. Analysis of the IGF2/H19 imprinting control region uncovers new genetic defects, including mutations of OCT-binding sequences, in patients with 11p15 fetal growth disorders. Hum Mol Genet 19:803–814.

Elliott M, Bayly R, Cole T, Temple IK, Maher ER. 1994. Clinical features and natural history of Beckwith–Wiedemann syndrome: Presentation of 74 new cases. Clin Genet 46:168–174.

Engel N, Thorvaldsen JL, Bartolomei MS. 2006. CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus. Hum Mol Genet 15:2945–2954.

Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature 405:486–489.

Horsthemke B. 2010. Mechanisms of imprint dysregulation. Am J Med Genet C Semin Med Genet 154C:321–328.

Li E. 2002. Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 3:662–673.

Sasaki H, Matsui Y. 2008. Epigenetic events in mammalian germ-cell development: Reprogramming and beyond. Nat Rev Genet 9:129–140.

Sasaki K, Soejima H, Higashimoto K, Yatsuki H, Ohashi H, Yakabe S, Joh K, Niikawa N, Mukai T. 2007. Japanese and North American/European patients with Beckwith–Wiedemann syndrome have different frequencies of some epigenetic and genetic alterations. Eur J Hum Genet 15: 1205–1210.

Soejima H, Nakagawachi T, Zhao W, Higashimoto K, Urano T, Matsukura S, Kitajima Y, Takeuchi M, Nakayama M, Oshimura M, Miyazaki K, Joh K, Mukai T. 2004. Silencing of imprinted CDKN1C gene expression is associated with loss of CpG and histone H3 lysine 9 methylation at DMR-LIT1 in esophageal cancer. Oncogene 23:4380–4388.

Sparago A, Cerrato F, Vernucci M, Ferrero GB, Silengo MC, Riccio A. 2004. Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith–Wiedemann syndrome. Nat Genet 36:958–960.

Sparago A, Russo S, Cerrato F, Ferraiuolo S, Castorina P, Selicorni A, Schwienbacher C, Negrini M, Ferrero GB, Silengo MC, Anichini C, Larizza L, Riccio A. 2007. Mechanisms causing imprinting defects in familial Beckwith–Wiedemann syndrome with Wilms' tumour. Hum Mol Genet 16:254–264.

Uejima H, Lee MP, Cui H, Feinberg AP. 2000. Hot-stop PCR: A simple and general assay for linear quantitation of allele ratios. Nat Genet 25: 375–376.

Weksberg R, Nishikawa J, Caluseriu O, Fei YL, Shuman C, Wei C, Steele L, Cameron J, Smith A, Ambus I, Li M, Ray PN, Sadowski P, Squire J. 2001. Tumor development in the Beckwith–Wiedemann syndrome is associated with a variety of constitutional molecular 11p15 alterations including imprinting defects of KCNQ1OT1. Hum Mol Genet 10:2989–3000.

Weksberg R, Shuman C, Beckwith JB. 2010. Beckwith–Wiedemann syndrome. Eur J Hum Genet 18:8–14.