



Comparative Genomics Sheds Light on Mechanisms of Genomic Imprinting

Thanh H. Vu and Andrew R. Hoffman

Genome Res. 2000 10: 1660-1663

Access the most recent version at doi:[10.1101/gr.166200](https://doi.org/10.1101/gr.166200)

References This article cites 22 articles, 7 of which can be accessed free at:
<http://genome.cshlp.org/content/10/11/1660.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

An advertisement banner with a teal background. On the left, the text reads "CRISPR and RNAi Genetic Screening. Your new superpower." In the center is a white box with the text "LEARN MORE". On the right is a woman in a red and white superhero costume with a red mask, and the Cellecta logo, which is a green molecular structure with the word "CELLECTA" below it.

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Cold Spring Harbor Laboratory Press

Comparative Genomics Sheds Light on Mechanisms of Genomic Imprinting

Thanh H. Vu¹ and Andrew R. Hoffman¹

Medical Service, VA Palo Alto Health Care System and Department of Medicine, Stanford University, Palo Alto, California 94304, USA

Genomic imprinting is an epigenetic chromosomal modification in the germ line that leads to preferential expression of one of the two parental alleles in a parent-of-origin-specific manner. A number of recent studies suggest that genomic imprinting is mediated by a set of elements in distinct chromosomal regions that have been termed imprinting centers (IC). By definition, the IC coordinate three activities: the establishment of imprint marks, the maintenance of these imprint marks throughout development, and the implementation of the preferential expression from a specific parental allele (for review, see Ben-Porath and Cedar 2000).

At least four major imprinting models that incorporate recent discoveries have been elucidated: First, the sense/antisense competition model for preferential allelic expression, as exemplified by studies examining the regulation of *Igf2r* and its associated antisense on mouse chromosome 17 (Box 1; Fig. 1). Second, the enhancer/chromosomal insulation model developed on the basis of the reciprocal imprinting of *Igf2/H19* in human and in mouse (Box 2; Fig. 2A). Third, the bipartite IC model, which describes the regulation of multiple imprinted genes in a 2-Mb Prader-Willi syndrome/Angelman syndrome (PWS-AS) region of human chromosome 15. Finally, the promoter-specific reciprocal-imprinting model as seen at the human and mouse *Gnas* locus (Box 2; Fig. 2B).

To understand the mechanisms underlying genomic imprinting, a number

of research groups have focused on the identification and characterization of the imprinting elements present in the various ICs. As the DNA sequence of the whole human genome is nearly complete and complete sequences of other species are or will soon be available, an entirely new strategy for understanding potential imprinting mechanisms has now become available. Sitting at one's computer, it will be possible to devise large-scale comparisons of genomic sequences within and/or between species that could yield new information on potentially important elements that govern preferential allelic expression.

In this issue, we see how two research groups (Wylie et al. 2000; Onyango et al. 2000) have advanced our understanding of the mechanisms of genomic imprinting by employing a comparative analysis of distinct imprinting regions in human and mouse genes.

Human Chromosome Regions Have Common Imprinting Elements

R. Jirtle and his colleagues at Duke University have identified two reciprocally imprinted genes, *DLK1* and *GTL2*, on human chromosome 14q32 (Wylie et al. 2000). The cognate murine genes, *Dlk1* and *Gtl1*, on mouse chromosome 12 are also linked and reciprocally imprinted (Schmidt et al. 2000; Takada et al. 2000). *DLK1*, which encodes a transmembrane protein with 6 EGF (epidermal growth factor) repeat motifs, is paternally expressed. In contrast, *GTL2*, which encodes an apparently nontranslated RNA, is located 102 kb downstream of *DLK1* and is expressed from the maternal allele.

A functionally similar imprinted region that harbors two reciprocally im-

printed genes, *IGF2* and *H19*, has been extensively studied in both mouse and human (Box 2). There are striking similarities in the spatial, structural, and reciprocal imprinting characteristics of these two imprinted regions (*IGF2/H19* on human chromosome 11p15 and *DLK1/GTL2* on human chromosome 14q32). *IGF2* is the paternally expressed gene, while the maternally expressed *H19* encodes for a nontranslated RNA. These similarities have prompted the Duke group to further elucidate the IC elements on 14q32 region by making careful comparisons with the 11p15-imprinted region.

CTCF is a zinc finger protein implicated in the regulation of chromosomal boundaries (Fig. 2A). In the mouse *Igf2/H19* region, CpG methylation in the CTCF binding site eliminates binding of the enhancer-blocking CTCF, thereby allowing expression of *Igf2* from the paternal allele (Bell and Felsenfeld 2000; Hark et al. 2000). Methylation analysis by bisulfite sequencing revealed a differential methylation region (DMR) within 2 kb upstream of *GTL2*, analogous to the silencer DMR1 of the IC on the *IGF2/H19* (Wylie et al. 2000). A DMR in the promoter region of *DLK1* may correspond to the DMR in the *IGF2* gene (Constancia et al. 2000). Interestingly, two unique CTCF binding sites were found in the region one kilobase upstream of *GTL2*. Bisulfite sequencing indicated that all CpG sites within and outside the binding sites were hemimethylated (equal numbers of methylated and unmethylated molecules), suggesting that CpG methylation in this DMR may serve as an epigenetic switch to turn on/off the enhancer-blocking activity of the IC by binding/unbinding the CTCF to the unmethylated/

¹Corresponding authors.

E-MAIL thanvu@stanford.edu, arhoffman@leland.stanford.edu; FAX (650) 856-8024.

Article and publication are at www.genome.org/cgi/doi/10.1101/gr.166200.

Box 1. Imprinting of *Igf2r*: Sense and Antisense Competition Model (Fig. 1)

An elegant model of sense/antisense competition has been proposed by Wutz et al. (1997). On the mouse paternal allele, the IC in region 2 is unmethylated, while promoter region 1 is methylated; antisense *Air* is transcribed. *Air* competes and turns off the *Igf2r* sense transcript. On the maternal allele, the IC is methylated and *Air* is turned off. Promoter region 1 is unmethylated, and *Igf2r* is transcribed. Two cis-regulatory elements in region 2, the de novo methylation signal (DNS) and the allele-discrimination signal (ADS), have been shown to be involved in the establishment of differential methylation of the *Igf2r* DMR (Birger et al. 1999). Imprinting of *Igf2r* is tissue specific. In the mouse central nervous system (CNS), *Igf2r* is transcribed from both parental alleles (Wang et al. 1994). On the paternal allele, the IC is unmethylated, and *Air* is transcribed. The promoter region 1 is unmethylated; *Air* transcription fails to outcompete *Igf2r* on the paternal allele (Hu et al. 1999). Imprinting of *Igf2r* is species specific. In humans, the "homologous IC" is methylated on the maternal allele. On the paternal allele, unmethylated IC fails to initiate antisense transcript (T.H. Vu and A.R. Hoffman, unpubl.). The human homologous IC is inactive. Promoter region 1 is unmethylated on both alleles; *IGF2R* is not imprinted. We propose that the active IC in the mouse coordinates the correct methylation pattern in promoter region 1 and initiates antisense transcription. In mouse CNS, IC fails to establish the correct promoter methylation, while in humans the "defunct IC" fails in both activities.

Imprinting of *Igf2r* is also observed in marsupials (mammals that give birth to altricial offspring) but this imprinting is not observed in monotremes (egg-laying mammals) (John et al. 2000; Killian et al. 2000). However, no differential methylation is detected in region 2 of the marsupial *Igf2r*. The mechanism of marsupial *Igf2r* imprinting remains to be explored.

Since the discovery of *Air*, imprinted antisense transcripts are observed in all four major imprinting models (see text). In human *KVLQT1*, targeted deletion of the *LIT1* CpG island (the putative IC) abolishes transcription of the antisense on the paternal allele, accompanied by activation of multiple normally suppressed paternal alleles in the *KVLQT1* region (Horike et al. 2000).

methylated IC. In parallel with the enhancer/chromosomal-boundary model of reciprocal imprinting, Wylie et al. (2000) also found the same consensus sequence for two enhancer elements as seen in *IGF2/H19* ~10 kb downstream from the *GTL2* transcription site.

Wylie et al. (2000) have presented an elegant approach combining bioinformatics and experimental data on allelic expression and bisulfite sequencing to elucidate a model for *DLK1/GTL2* reciprocal imprinting. The *DLK1/GTL2* and the *IGF2/H19* regions are located on different chromosomes (human 14q32 and 11p15; mouse chromosomes 12 and 7), and they encode functionally diverse proteins *DLK-1* (or *FA1*) and *IGF-II*. The striking similarities in the imprinting characteristics of these two imprinting dyads suggest that many imprinting elements that are crucial for this type of reciprocal imprinting would be conserved in heterologous chromosomal locations within one species, as well as in orthologous regions between species.

The Structure of a 1-Mb Imprinted Region

Imprinted genes usually appear in pairs or clusters. As imprinting often influences large chromosomal domains and may be regulated by more than one (or by a bipartite) IC, comparison of larger imprinted domains would pinpoint the

crucial conserved, putative imprinting elements in the ICs. In this issue, A. Feinberg (2000) and his colleagues at Johns Hopkins University pioneered

such an approach (Onyango et al. 2000). They performed a comparative analysis of a mouse 1-Mb imprinted region that is orthologous to the human 11p15 do-

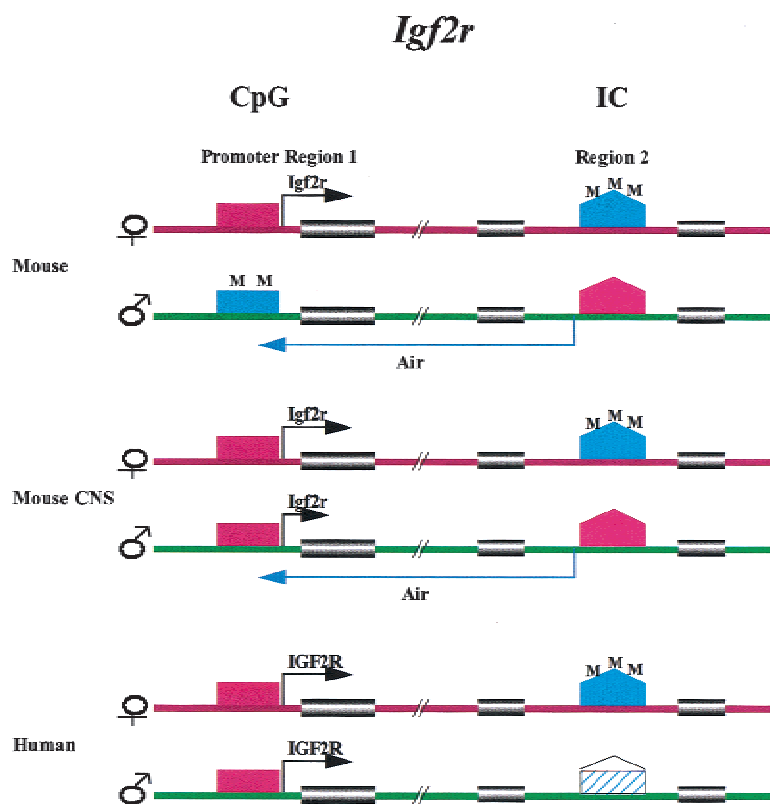


Figure 1 Coordination of IC methylation, promoter methylation, and antisense transcription in the *Igf2r* imprinting model. Red box, unmethylated and active promoter. Green box, methylated and inactive promoter. Red pentagon, unmethylated and active IC. Green pentagon, methylated IC. White pentagon, unmethylated and defunct IC. Black arrow, sense transcript. Purple arrow, antisense transcript.

Box 2. Imprinting of *Igf2/H19* and *Gnas*: Reciprocal Imprinting and Chromosomal Insulation Model (Fig. 2)

Reciprocal imprinting of *Igf2/H19* has been studied extensively. The imprinting of human *DLK1/GTL2* (mouse *Dlk1/Gtl2*) is similar to *Igf2/H19* imprinting. On the maternal allele, the IC is unmethylated. CTCF binds to the IC, which insulates enhancer activity from *Igf2*, only *H19* is transcribed. On paternal allele, the IC is methylated, which prevents CTCF binding. The IC also spreads CpG methylation along the *H19* promoter. Only *Igf2* is transcribed (sense and antisense). The *Ihit1* transcript is located 20 kb upstream of *H19* (Onyango et al. 2000). The imprinting status of *Ihit1* has not been determined. Reciprocal imprinting in *Gnas* is more complex. *Nesp* and *Xlas* transcripts are reciprocally imprinted, while *Gsα* is transcribed from both parental alleles (Liu et al. 2000). A putative IC, 2 kb upstream of *Gsα*, is methylated on the maternal allele, which may act on the methylation of the maternal *Xlas* promoter. *Nesp* is transcribed from the unmethylated maternal promoter. On the paternal allele, the IC is unmethylated and so is the *Xlas* promoter. Both *Xlas* and the IC (exon 1A) are transcribed (Liu et al. 2000). Antisense (*Gnas-AS* or *Nesp-AS*) is also transcribed from the unmethylated *Xlas* promoter region (Li et al. 2000; Wroe et al. 2000), while the methylated *Nesp* promoter turns off *Nesp* transcription. Enhancer activity and a chromosomal boundary have not yet been found.

main. They constructed a BAC contig containing ~1 Mb of the imprinted region on mouse chromosome 7. By comparing this region with its orthologous region on human chromosome 11p15, they were able to identify 33 conserved CpG islands (with sequence identity $\geq 50\%$) and 49 conserved nonexonic, non-CpG island sequences (at least 70% sequence identity in 100 bp). This group and others (Smilnich et al. 1999; Paulsen et al. 2000) have previously shown that the human imprinted region harbors two imprinted subdomains: the *IGF2/H19* domain and the domain centered around *KvLQT1*. These domains are separated by a nonimprinted domain containing *TSSC6-TAPA1-TSSC4-TRPC51*. Similar findings pertain to the mouse.

All of the imprinted genes in the mouse contain at least two CpG islands, except for the antisense *Lit1* transcript, and they are clustered by intronic sequence elements that are conserved compared with the human gene cluster. By doing this comparison, Onyango et al. (2000) identified two transcripts unique to mouse and five transcripts unique to human. In the *Kvlqt1* domain, a conserved novel transcript, *Tssc8*, was identified; it lies within both the *Kvlqt1* and antisense *Lit1* transcripts. *Tssc8* has no obvious ORF and is transcribed in the opposite orientation from *Lit1*, suggesting that it represents a novel antisense of an antisense (*Lit1*) transcript. The imprinting status and potential function of *Tssc8* will require further investigation. In addition, a mouse-specific paternally imprinted transcript (*Msuit*) was discovered within the *Kvlqt1* domain upstream of *p57kip2*.

Within the *Igf2/H19* imprinted domain, they identified a novel conserved transcript 20 kb downstream of *H19* and a novel *Igf2-H19* interval transcript (called *Ihit1*) 20 kb upstream of *H19* (Fig. 2A). Determination of the imprinting status of the novel downstream transcript will help elucidate the boundary of the imprinted/nonimprinted subdomains. The imprinting status of the interval transcript *Ihit1* will provide further evidence to test the enhancer/chromosome boundary model in the *Igf2/H19* imprinted domain (Box 2; Fig. 2A).

Future Developments

Sequence comparisons of large DNA re-

gions across species have been used recently to identify long-range regulatory elements that are difficult to discover by conventional methods. A coordinate regulator of interleukins 4, 13, and 5 was identified by large-scale sequence comparison of ~1-Mb human 5q31 region and the murine chromosome 11 orthologous region (Loots et al. 2000). By comparing imprinted domains in the mouse to the orthologous human sequences, Onyango et al. (2000) have identified multiple putative imprinting elements in the *Igf2/H19* and *Kvlqt1* imprinted domains. The full characterization of these putative imprinting elements will undoubtedly be forthcoming. Further comparisons of orthologous

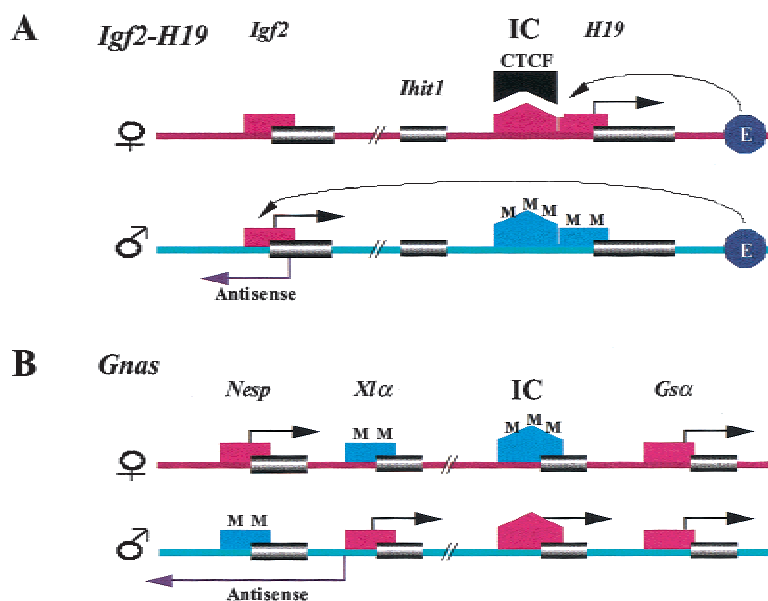


Figure 2 (A) Coordination of IC methylation, promoter methylation, and enhancer blocking activity in the *Igf2/H19* reciprocal imprinting model. CTCF, enhancer-blocking protein. *Ihit1*, novel *Igf2-H19* interval transcript. (B) Coordination of IC methylation, promoter methylation, and antisense transcription in *Nesp/Xlas/exon 1A* reciprocal imprinting model.

sequences in other mammals may help to pinpoint crucial imprinting elements, as demonstrated by the recent innovative studies that have elucidated the evolution of *Igf2r* imprinting (John et al. 2000; Killian et al. 2000).

An alternative yet complementary approach is to compare sequences of similar imprinted domains located in different loci but within the same genome, as reported by Wylie et al. (2000). In the next few years, large-scale sequence comparisons of various imprinted regions within one species and sequence comparisons across species are certain to provide important new insights on various imprinting elements of the IC.

The IC implements preferential expression of a specific parental allele through imprint marks that are established in the germ line during gametogenesis and are maintained throughout development. Although the gametic imprints can be modulated by differential chromatin structure and/or by differential DNA replication timing, the primary imprints often involve DNA methylation.

Methylation of cytosine in CpG dinucleotides creates the fifth base, a methylated cytosine, along with the four conventional DNA-building blocks, G, A, T, and C. Bisulfite sequencing, as used by Wylie et al. (2000) and others (e.g., Vu et al. 2000), helps to elucidate the imprint marks of the methylated CpG in the IC.

The methylation of a CpG island in the promoter region of a gene is a hallmark of gene regulation. The notion by Onyango et al. (2000) that almost all imprinted genes are associated with two or more CpG islands suggests the complexity of the imprinting mechanisms. Gametic methylation of IC elements located far away from the promoter region

can induce methylation/demethylation of the promoter CpG islands, thus allowing “working from a distance.” IC methylation acting in *cis* and *trans* can coordinate sense/antisense transcripts through direct promoter methylation, chromatin remodeling (e.g., by recruiting histone deacetylase), or acting on a chromosomal boundary with enhancer-blocking activity.

DNA methylation plays a pivotal role in genomic imprinting. Further innovative developments in quantifying CpG methylation genome wide (such as the development of CpG DNA chips) or quantifying methylation at specific sites (such as methylation restriction [MR], MR-PCR; T.H. Vu and A.R. Hoffman, unpubl.) would be of great interest.

ACKNOWLEDGMENTS

We thank all members of our laboratory for technical support, especially Tao Li for designing the figures. This paper was supported by NIH grant DK36054 and the Research Service of the Department of Veterans Affairs.

REFERENCES

- Bell, A.C. and G. Felsenfeld. 2000. *Nature* **405**: 482–485.
- Ben-Porath, I. and H. Cedar. 2000. *Curr. Opin. Genet. Dev.* **10**: 550–554.
- Birger, Y., R. Shemer, J. Perk, and A. Razin. 1999. *Nature* **397**: 84–88.
- Constancia, M., W. Dean, S. Lopes, T. Moore, G. Kelsey, and W. Reik. 2000. *Nat. Genet.* **26**: 203–206.
- Hark, A.T., C.J. Schoenherr, D.J. Katz, R.S. Ingram, J.M. Levarso, and S.M. Tilghman. 2000. *Nature* **405**: 486–489.
- Horiike, S.I., K. Mitsuya, M. Meguro, N. Kotobuki, A. Kashiwagi, T. Notsu, T.C. Schulz, Y. Shirayoshi, and M. Oshimura. 2000. *Human Mol. Genet.* **9**: 2075–2083.
- Hu, J.F., K.A. Balaguru, R.D. Ivaturi, H. Oruganti, T. Li, B.T. Nguyen, T.H. Vu, and A.R. Hoffman. 1999. *Biochem. Biophys. Res Comm.* **257**: 604–608.
- John, R.M., and M.A. Surani. 2000. *Cell* **101**: 585–588.
- Killian, J.K., J.C. Byrd, J.V. Jirtle, B.L.

- Munday, M.K. Stoskopf, R.G. MacDonald, and R.L. Jirtle. 2000. *Mol. Cell.* **5**: 707–716.
- Li, T., T.H. Vu, Z.L. Zeng, B.T. Nguyen, B.E. Hayward, D.T. Bonthron, J.F. Hu, and A.R. Hoffman. 2000. *Genomics* **69**: 295–304.
- Liu, J., S. Yu, D. Litman, W. Chen, and L.S. Weinstein. 2000. *Mol. Cell. Biol.* **20**: 5808–5817.
- Loots, G.G., R.M. Locksley, C.M. Blankespoor, Z.E. Wang, W. Miller, E.M. Rubin, and K.A. Frazer. 2000. *Science* **288**: 136–140.
- Onyango, P., W. Miller, J. Lehoczy, C. Leung, B. Birren, S. Wheelan, K. Dewar, and A.P. Feinberg. 2000. *Genome Res.* **10**: 1697–1710.
- Paulsen, M., O. El-Maarri, S. Engemann, M. Stroedicke, O. Franck, K. Davies, R. Reinhardt, W. Reik, and J. Walter. 2000. *Hum. Mol. Genet.* **9**: 1829–1841.
- Schmidt J.V., P.G. Matteson, B.K. Jones, X.J. Guan, and S.M. Tilghman. 2000. *Genes & Dev.* **14**: 1997–2002.
- Smilnich, N.J., C.D. Day, G.V. Fitzpatrick, G.M. Caldwell, A.C. Lossie, P.R. Cooper, A.C. Smallwood, J.A. Joyce, P.N. Schofield, W. Reik, et al. 1999. *Proc. Natl. Acad. Sci.* **96**: 8064–8069.
- Takada, S., M. Tevendale, J. Baker, P. Georgiades, E. Campbell, T. Freeman, M.H. Johnson, M. Paulsen, and A.C. Ferguson-Smith. 2000. *Curr. Biol.* **10**: 1135–1138.
- Vu, T.H., T. Li, D. Nguyen, B.T. Nguyen, X.M. Yao, J.F. Hu, and A.R. Hoffman. 2000. *Genomics* **64**: 132–143.
- Wang, Z.Q., M.R. Fung, D.P. Barlow, and E.F. Wagner. 1994. *Nature* **372**: 464–467.
- Wroe, S.F., G. Kelsey, J.A. Skinner, D. Bodle, S.T. Ball, C.V. Beechey, J. Peters, and C.M. Williamson. 2000. *Proc. Natl. Acad. Sci.* **97**: 3342–3346.
- Wutz, A., O.W. Smrzka, N. Schweifer, K. Schellander, E.F. Wagner, and D.P. Barlow. 1997. *Nature* **389**: 745–749.
- Wylie, A.A., S.K. Murphy, T.C. Orton, and R.L. Jirtle. 2000. *Genome Res.* **10**: 1711–1718.

[For more information on comparative genome analysis and imprinting, see Okamura et al. in the December issue of *Genome Research*.]