Genomic imprinting: employing and avoiding epigenetic processes

Marisa S. Bartolomei

Genes Dev. 2009 23: 2124-2133
Access the most recent version at doi:10.1101/gad.1841409
Genomic imprinting refers to an epigenetic mark that distinguishes parental alleles and results in a monoallelic, parental-specific expression pattern in mammals. Few phenomena in nature depend more on epigenetic mechanisms while at the same time evading them. The alleles of imprinted genes are marked epigenetically at discrete elements termed imprinting control regions (ICRs) with their parental origin in gametes through the use of DNA methylation, at the very least. Imprinted gene expression is subsequently maintained using noncoding RNAs, histone modifications, insulators, and higher-order chromatin structure. Avoidance is manifest when imprinted genes evade the genome-wide reprogramming that occurs after fertilization and remain marked with their parental origin. This review summarizes what is known about the establishment and maintenance of imprinting marks and discusses the mechanisms of imprinting in clusters. Additionally, the evolution of imprinted gene clusters is described. While considerable information regarding epigenetic control of imprinting has been obtained recently, much remains to be learned.

Mammals possess a small number of genes subject to an unusual form of gene regulation, termed genomic imprinting (Barlow and Bartolomei 2007). Genomic imprinting, which results in the parental-specific expression of these genes, is proposed to be the major block to parthenogenesis in mammals (Kono et al. 2004). Genomic imprinting also subjects mammals to a greater genomic risk because a mutation in one allele (either genetic or epigenetic) can result in the absence of one or more gene products, thereby leading to a number of well-known imprinting disorders, including Beckwith-Wiedemann syndrome, Silver-Russell syndrome, Prader-Willi syndrome, and Angelman syndrome (Horsthemke and Wagstaff 2008; Ideraabdullah et al. 2008).

At this time, there are ~100 imprinted genes identified in the mouse (see http://www.har.mrc.ac.uk/research/genomic_imprinting for a full list), with the imprinting of many of these genes conserved in humans (http://www.otago.ac.nz/IGC). Although additional genes and transcripts may be discovered using more modern sequence analysis of transcriptomes, it is notable that the first few published studies of this nature reported only a small number of novel imprinted genes, suggesting that many new widely expressed imprinted genes are not likely to be identified (Babak et al. 2008; Wang et al. 2008). Nevertheless, more detailed, allele-specific transcriptome analyses of purified cell populations should add to the growing list of tissue-specific imprinted genes.

Imprinted genes have a variety of functions; many imprinted genes are important for fetal and placental growth and development, whereas others are involved in postnatal behavior. Imprinted genes are also located throughout the genome, including on the X chromosome. Some imprinted genes currently map as singletons or in discrete pairs, but most imprinted genes reside in ~1-Mb clusters (O’Neill 2005). These larger clusters usually contain at least one noncoding RNA (ncRNA), which is often >100 kb in length, and maternally and paternally expressed genes (Fig. 1). Also present within each cluster is an imprinting control region [ICR; alternatively called imprinting control element [ICE] or differentially methylated domain [DMD]]. The region, typically a few kilobase pairs in length, exhibits parental allele-specific DNA methylation and post-translational histone modifications. Furthermore, when the ICR is deleted, loss of imprinted gene expression is observed for the linked genes (Wutz et al. 1997; Thorvaldsen et al. 1998; Yang et al. 1998; Fitzpatrick et al. 2002; Lin et al. 2003; Williamson et al. 2006). This review describes the properties of this unusual form of gene expression, the two main hypothesized mechanisms of imprinted gene expression, and studies that address the evolution of imprinted gene loci.

Unusual properties of imprinted genes

Imprinted genes display properties that separate them from the majority of the mammalian genome. First, the alleles of these genes must be differentially modified, or epigenetically marked, with their parental origin so that appropriate patterns of expression are assumed in the
soma. It has long been hypothesized that this marking occurs in the germline, because this is the time when parental genomes are in separate compartments and can be differentially modified. Second, the germline marks must be maintained after fertilization when the vast majority of the genome is being reprogrammed. Here, DNA methylation and post-translational histone modifications are erased and reset during the process in which pluripotent cells of the preimplantation embryo give rise to cells of various fates, while imprints are maintained (Morgan et al. 2005). DNA appears actively demethylated on the paternal genome within hours of fertilization and is subsequently passively demethylated on the maternal genome (Rougier et al. 1998; Mayer et al. 2000; Oswald et al. 2000). The precise mechanism of this demethylation has been called into question with the recent description of 5-hydroxymethylcytosine in mammals and the identification of TET1, which catalyzes the conversion of 5-methylcytosine to this modified base (Kriaucionis and Heintz 2009; Tahiliani et al. 2009). However, regardless of the mechanism, imprinted genes are uniquely immune to the demethylation process during preimplantation development. Thus, an important question remains as to how imprints are protected from genome-wide demethylation. Finally, imprints are largely maintained in the soma, although tissue-specific imprinting is frequently observed, and are subsequently erased during germ cell development. In the germline, however, erasure of imprints is not unique, as most of the genome is reprogrammed at this time. More specifically, DNA demethylation, which serves as a proxy for erasure of genomic imprints, occurs concomitantly with demethylation of other parts of the genome (Hajkova et al. 2002).

Differential marking of parental alleles

As mentioned above, ICRs exhibit parental-specific epigenetic modifications that are proposed to be essential for the recognition of parental origin. Currently, only differential DNA methylation of these regions has been demonstrated in gametes and during preimplantation development. This does not mean that other epigenetic modifications are not involved in marking and maintaining parental identity; rather, the large number of cells required to assay allele-specific post-translational histone modifications and higher-order chromatin structure makes it difficult to assess the role of these modifications in germ cells and early embryos. Thus, much attention has focused on acquisition and maintenance of DNA methylation. Curiously, most ICRs are methylated on the maternal allele, and this methylation is established during the oocyte growth phase prior to ovulation (Lucifero et al. 2002). In contrast, the three paternally methylated ICRs (H19/Igf2, Dlk1/Dio3, and Rasgrf1) are methylated in gonocytes during the period between mitotic arrest and birth (Davies et al. 1999; Li et al. 2004). The use of conditional deletion alleles of the de novo family of DNA methyltransferase proteins has shown that all of the ICRs tested to date, with one exception, use the de novo DNA methyltransferase DNMT3A and its stimulatory protein, DNMT3L, to confer DNA methylation on the ICRs in the respective germ cells (Bourc’his et al. 2001; Hata et al. 2002; Bourc’his and Bestor 2004; Kaneda et al. 2004). The exception is the Rasgrf1 ICR, which is methylated in male germ cells and uses the de novo DNA methyltransferase DNMT3B. The reason for this difference is unclear, but is most likely due to the fact that the enzyme predominantly methylates repetitive DNA and the Rasgrf1 ICR and surrounding sequences are highly repetitive. In summary, we know 1) the sequences that are methylated in the germ cells (ICRs), 2) when the marks are put on the ICRs, and 3) the enzymes that confer the marks. However, it is still unclear how the epigenetic machinery recognizes ICR sequences. One clue comes from the structural analysis of the complexed C-terminal domains of DNMT3A and DNMT3L, which was obtained by X-ray crystallography (Jia et al. 2007). A tetrameric complex consisting of these two enzymes preferentially methylates a pair of CpGs that were 8–10 base pairs apart. Such spacing is found in maternally methylated, but not paternally methylated, imprinted loci. Notably, such CpG spacing is widespread in the genome (Ferguson-Smith and Greally 2007). Additional specificity was suggested by the demonstration that DNMT3L interacts with the N terminus of histone H3 and that this interaction is inhibited by H3 Lys 4 methylation (Ooi et al. 2007). More recently, Chotalia et al. [2009] described transcription [some of it oocyte-specific]

Figure 1. Components of an imprinted gene cluster. The maternal [top] and paternal [bottom] alleles of an imprinted gene cluster are depicted. Imprinted gene clusters contain maternally expressed [pink-filled boxes], paternally expressed [blue-filled boxes], and biallelically expressed genes [i.e., nonimprinted genes, green-filled boxes]. These nonimprinted genes can be found in the middle of a cluster surrounded by imprinted genes. The ICR [yellow] controls imprinting of multiple genes in the region; deletion of this differentially methylated element results in loss of imprinting of the linked genes. Many imprinted clusters also contain additional DMRs [orange] that acquire DNA methylation after the preimplantation stage. Not depicted here are long ncRNAs and insulators, which have an essential role in imprinted clusters.
across differentially methylated regions [DMRs] that they propose are required for the establishment of DNA methylation imprints in the oocyte. Importantly, only protein-coding transcripts traversing the germline ICRs are thought to be involved in methylation establishment, which is in contrast to the better-established role for transcription of ncRNAs in regulating the imprinting of adjacent genes [see below]. Although it is as yet unclear how this transcription may be attracting the DNA methylation machinery, Chotalia et al. [2009] suggest that transcription across ICRs is required to establish or maintain open chromatin domains that are permissive for establishment of DNA methylation. Moreover, to investigate and define the mechanism further, it is first necessary to describe the temporal relationship between transcription and de novo methylation in greater detail. Nevertheless, CpG spacing, the state of post-translational histone modifications, and transcription in oocytes could provide a starting point for the acquisition of maternal-specific DNA methylation imprints.

There is far less information regarding how paternal-specific DNA methylation imprints are established in the male germline. Shaw and coworkers [Jelinic et al. 2006] suggested that the methylation of the H19 ICR in the male germline involves the recognition of a specific chromatin signature through BORIS, also known as CTCF-like (CTCFL). CTCFL is a paralog of CTCF, which is an insulator protein that binds to the H19 ICR and prevents DNA methylation of the maternal allele [see below]. Using embryonic day 15.5 [E15.5] testes, Jelinic et al. [2006] employed chromatin immunoprecipitation experiments to demonstrate that CTCFL binds to the paternally methylated H19 ICR in male germ cells. Additionally, CTCFL expression coincides with H19 imprint establishment and the N terminus of CTCFL interacts with Protein Arginine Methyltransferase 7 [PRMT7], which methylates histones H2A and H4. However, the most definitive proof in support of a functional role for these proteins establishing methylation at the ICR derives from a cotransfection experiment of plasmids encoding CTCFL, PRMT7, DNMT3A/B/L, and the H19 ICR in Xenopus oocytes. Given the absence of H19 and imprinting in Xenopus, further support for a role of BORIS/CTCFL awaits additional biochemistry and germline-specific deletion of these genes.

In a discussion of differential DNA methylation, it is important to note that ICRs are distinguished from DMRs [Fig. 1]. Although both are differentially methylated in somatic cells, DMRs do not fulfill the criteria of an imprinting mark. That is, they are not placed on the DNA in the germline and do not survive post-fertilization reprogramming of DNA methylation marks. One example of a DMR is the promoter of the maternally expressed Cdkn1c gene. This promoter becomes methylated on the paternal allele in the post-implantation embryo, after imprinting is established at this locus [Bhogal et al. 2004]. Further, the promoter of the maternally expressed Igf2r gene is methylated on the paternal allele, and this DNA methylation also occurs after establishment of imprinting [Stöger et al. 1993].

**Maintenance of genomic imprints**

Once imprints are established in the germline, it is imperative that they survive both the reprogramming that occurs in the preimplantation embryo and the subsequent wave of de novo methylation [Morgan et al. 2005]. Not only must ICRs endure genome-wide removal of DNA methylation [and other putative epigenetic marks], but the epigenetic modifications must be placed on the newly replicated DNA. In the latter case, the maintenance DNA methyltransferase DNMT1, which has an established role in maintaining DNA methylation at imprinted loci [Li et al. 1993], is present at very low levels in preimplantation embryos. Although the details are still not entirely clear, it is likely that DNA methylation is maintained at this time through a combination of the oocyte-specific form of DNMT1 [DNMT1o] as well as the somatic form of DNMT1, which is the form of the enzyme most typically observed in mammalian cells [Howell et al. 2001; Cirio et al. 2008; Hirasawa et al. 2008]. Notably, some DMRs are differentially methylated in gametes and in somatic cells but fail to maintain this methylation right after fertilization. One example is DMR2 in the Igf2 gene [Fig. 2]; DMR2 is methylated during spermatogenesis, actively demethylated immediately after fertilization, and remethylated on the paternal allele later in development [Oswald et al. 2000]. As with other DMRs, the late differential DNA methylation is likely a downstream effect of the establishment of imprinting at the locus [Lopes et al. 2003].
Another important focus is on how imprints escape genome-wide reprogramming occurring after fertilization. The most probable scenario is that both cis- and trans-acting factors are involved in maintenance of genomic imprints. That is, DNA-binding factors specifically recognize ICRs, preventing active demethylation on the paternally methylated alleles and facilitating maintenance DNA methylation and differential chromatin modifications, as described above. A few trans-acting factors required for the maintenance of methylation have been identified recently. One is MBD3, a member of the NuRD remodeling complex. RNAi-mediated depletion of Mbd3 mRNAs in oocytes and preimplantation embryos caused biallelic expression of H19, suggesting that MBD3 plays a role in interpreting methylation to silence paternal H19 expression (Reese et al. 2007). Unexpectedly, loss of MBD3 also caused loss of DNA methylation on the paternal allele, indicating that MBD3 is involved in maintaining methylation. This was true only of the H19 gene, as MBD3 depletion had no effect on any other imprinted genes examined. In vitro studies suggest that MBD3 does not bind directly to methylated cytosines (Hendrich and Bird 1998). It is therefore hypothesized that MBD3 performs this function as part of a complex, such as the NuRD remodeling complex (Saito and Ishikawa 2002), although such complexes have yet to be characterized in mouse preimplantation embryos.

Another factor required for maintenance of methylation is ZFP57, a KrAB zinc finger protein. This class of transcription factors represses transcription by recruiting KAP-1/TIF1β corepressor complexes (Friedman et al. 1996; Abrink et al. 2001). Mutations in Zfp57 in the mouse cause both maternal and zygotic effect lethality and corresponding disruption of methylation and expression of imprinted genes (Li et al. 2008). Lack of both maternal and zygotic ZFP57 resulted in embryonic lethality and complete loss of methylation at the Snrpn, Peg1, Peg3, Peg5, and Dlk1 DMRs, while the H19 and Igf2r ICRs were unaffected. Zygotic disruption alone resulted in partial loss of methylation at these sites and partial lethality. In addition, maternal ZFP57 was required for the establishment of DNA methylation at the Snrpn ICR in oocytes, but, surprisingly, zygotic Zfp57 expression compensated for this loss and Snrpn methylation was re-established after E3.5. This result implies that despite the lack of methylation in the oocyte, Snprn harbored a residual imprint that was able to direct de novo methylation at the appropriate site in the embryo; the residual imprint was either outside the assayed region or derived from an epigenetic modification other than DNA methylation. The role of ZFP57 in the maintenance of ICR methylation was discovered independently in cases of transient neonatal diabetes (TND), which is caused by hypomethylation of the promoter of the imprinted gene PLAGL1 (Temple and Shield 2002; Mackay et al. 2008). Mutations in ZFP57 were identified in multiple affected pedigrees, and patients with these mutations often had hypomethylation of additional ICRs and DMRs, as well as clinical features not normally associated with TND. These results demonstrate that the role of ZFP57 in maintenance of DNA methylation is conserved between mice and humans.

Additional proteins have been identified that are involved in the stability of imprints, including RBBP1/ARID4A and RBBP1L1/ARID4B, which are involved in maintenance of imprinting at the Snrpn locus (Wu et al. 2006), and PGC7/STELLA (Nakamura et al. 2007). Although Stella+/− eggs have normal ICR methylation, embryos derived from these eggs are hypomethylated at multiple loci with maternally or paternally methylated ICRs. Zygotes also exhibit premature global loss of DNA methylation on the maternal pronucleus, indicating a more widespread role for STELLA. Thus, these results suggest that the maternal genome must be protected from the active demethylation occurring on the paternal genome immediately after fertilization. In summary, it is apparent that multiple proteins maintain methylation at different sets of sites, and it is likely that other proteins involved in imprint maintenance will be identified soon. In the future, it will be essential to determine how ICRs are specifically recognized over other genomic elements, and how the epigenetic machinery is recruited to these regions.

Mechanisms of imprinting in clusters

It was realized soon after the discovery of the first imprinted genes in the 1990s that many of these genes were located in ~1-Mbp clusters and that the imprinting of the genes was coordinately controlled by an ICR, as described above. The mechanism by which these clusters are regulated over long distances has been studied intensively in the last decade, with two prevailing models of regulation emerging. The first of these is the insulator model, in which imprinted genes share regulatory elements and the insulator controls access to these elements. In the second model, ncRNA mediates silencing of linked genes. It is important to note, however, that other mechanisms of regulation are likely and have even been reported recently. For example, Wood et al. (2008) described a new imprinting locus [H13] in which alternative polyadenylation sites are used in an allele-specific manner. The H13 gene harbors a maternally methylated internal CpG island that acquires DNA methylation in oocytes [but has not yet been tested for ICR function]. Synthesis of the full-length and functional H13 gene transcript from the maternal chromosome depends on hypermethylation of this CpG island. On the paternal allele, the unmethylated CpG island also serves as the promoter for the Mcts2 retrogene. Mcts2 expression correlates with the premature polyadenylation of H13 and expression of truncated H13 transcripts.

Insulator model of regulation

The insulator model of imprinting has been well documented for the H19/Igf2 locus [Fig. 2]. As described below, marsupials harbor an orthologous locus, with many similar attributes, including imprinted gene expression [Smits et al. 2008]. Thus, this gene cluster represents the most ancient imprinted locus identified
to date. 

**H19** encodes an ∼2.2-kb ncRNA expressed from the maternal chromosome [Bartolomei et al. 1991]. Although the precise function for this gene is still debated, a microRNA (miR-675) arises from the first exon [Mineno et al. 2006]. Moreover, Dandolo and colleagues [Yoshimizu et al. 2008] recently provided additional new data supporting an older study [Hao et al. 1993] suggesting that H19 exhibits a tumor suppressor effect. The Igf2 gene encodes a fetal growth factor expressed from the paternally inherited chromosome [DeChiara et al. 1990, 1991]. Early studies showed that these closely linked but reciprocally imprinted genes share enhancers [Leighton et al. 1995]. Additionally, an ICR was identified at the mouse locus that resides from −2 kb to −4 kb relative to the start of H19 transcription. This element was initially described as the sole region that was paternally methylated in the gametes and throughout development [Tremblay et al. 1995, 1997]. Deletion of this region at the endogenous locus results in the loss of imprinting for both H19 and Igf2 on the maternal and paternal alleles, formally demonstrating that it acts as an ICR [Thorvaldsen et al. 1998, 2002]. It was subsequently shown that the ICR binds CTCF, a protein that mediates insulator activity at the β-globin locus [Bell et al. 1999], and that the ICR itself functions as an insulator [Bell and Felsenfeld 2000; Hark et al. 2000; Kaffer et al. 2000; Kanduri et al. 2000; Szabo et al. 2000]. In this context, an insulator is defined as an element that blocks enhancer and promoter interactions when placed between them [Engel and Bartolomei 2003]. Thus, the model for imprinted gene expression at this locus is as follows [Fig. 2]: On the maternal allele, CTCF binds to the ICR and blocks the interaction of Igf2 to enhancers shared with H19, allowing H19 exclusive access to the enhancers. On the paternal allele, the ICR acquires DNA methylation in the male germline, preventing CTCF binding. Thus, on the paternal chromosome, Igf2 interacts with the enhancers and is expressed from this chromosome. The presence of DNA methylation on the paternal ICR leads to secondary methylation of the H19 promoter by an unknown mechanism, and H19 becomes silenced on the paternal chromosome.

Although it is now well established that the ICR acts as a CTCF-dependent insulator/enhancer blocker, our understanding of the mechanism of insulation remains incomplete. Chromosome conformation capture (3C) experiments in mice, which assay for physical interactions between chromosomal regions, suggest that chromosomal looping is involved in the imprinting mechanism, although the precise nature and function of the looping is debated [Lopes et al. 2003; Murrell et al. 2004; Kurukuti et al. 2006; Yoon et al. 2007; Engel et al. 2008]. While there is some consensus that the shared enhancers physically interact with the Igf2 promoters on the paternal chromosome, interactions on the maternal chromosome remain unclear. In this latter case, experiments by different groups have suggested that looping [1] physically impedes Igf2 access to the shared enhancers, [2] involves a “decoy” type of mechanism whereby the ICR forms an unproductive interaction between Igf2 promoters and the enhancers, and/or [3] allows the enhancers to interact with H19 up to but not beyond the ICR. Because these experiments were all performed in a slightly different manner by a variety of groups, further experiments are required to resolve the mechanism of insulation on the maternal allele.

Regardless of the detailed mechanism of insulator activity, it is clear that distal elements interact. Consistent with this model, recent studies demonstrate that cohesins, which mediate sister chromatid cohesion, colocalize with CTCF at thousands of sites in the genome [Parelho et al. 2008]. Specifically, two components of the cohesin complex, RAD21 and SMCM, associate with CTCF sites at the H19/Igf2 ICR in an allele-specific manner comparable with CTCF binding [Stedman et al. 2008]. Here again, it is unclear what function cohesins serve at the H19/Igf2 locus, but it is tempting to speculate that the sister chromatid adhering properties of the cohesin complex work at other genomic loci, including imprinted loci, to stabilize looping and higher-order chromatin structure [Peric-Hupkes and van Steensel 2008].

Despite the large amount of information obtained addressing the mechanism of insulation at the H19/Igf2 locus, the question remains whether other imprinted loci use a similar mechanism. CTCF-binding sites have been identified at other imprinted genes such as Rasgrf1 and Kcnq1ot1 [Yoon et al. 1995, 1997, 2005; Fitzpatrick et al. 2007], but given the large number of CTCF-binding sites in the genome and the diverse proposed functions for CTCF [Barski et al. 2007; Kim et al. 2007; Xie et al. 2007, Filippova 2008], it is unclear whether the presence of CTCF infers insulator function. There is evidence for insulator activity at the Rasgrf1 locus [Yoon et al. 2005], but further experiments are required at this and other loci to determine whether they use an H19/Igf2-type model of imprinting regulation. Most importantly, additional cis-regulatory sequences, such as enhancers, must be identified to assess the suitability of an insulator model at these loci.

**Long ncRNA model of regulation**

In contrast to the seemingly limited use of CTCF-dependent insulation at imprinted loci, the majority of imprinted loci appear to use a second, ncRNA mechanism of regulation in clusters [Koerner et al. 2009]. The first, and perhaps the best described, cluster in this class is the Igf2r cluster, which resides on mouse chromosome 17A [Braidotti et al. 2004]. Two additional well-defined clusters include the Kcnq1 cluster, which is adjacent to H19/Igf2 in mouse and human, and the Gnas cluster. It is also likely that a long ncRNA model is employed by the Prader-Willi/Angelman syndrome and Dlk1/Dio3 loci, but more work remains to be done to define the role and properties of the long ncRNAs in these cases. For the sake of simplicity, the Igf2r locus will be described in detail here [Fig. 3].

Igf2r and two neighboring genes, Slc22a2 and Slc22a3 (solute carrier 22a2 and 22a3), are expressed maternally. This region harbors one paternally expressed transcript, Airm [antisense Igf2r RNA] that overlaps Igf2r. Several

**Long ncRNA model of regulation**

In contrast to the seemingly limited use of CTCF-dependent insulation at imprinted loci, the majority of imprinted loci appear to use a second, ncRNA mechanism of regulation in clusters [Koerner et al. 2009]. The first, and perhaps the best described, cluster in this class is the Igf2r cluster, which resides on mouse chromosome 17A [Braidotti et al. 2004]. Two additional well-defined clusters include the Kcnq1 cluster, which is adjacent to H19/Igf2 in mouse and human, and the Gnas cluster. It is also likely that a long ncRNA model is employed by the Prader-Willi/Angelman syndrome and Dlk1/Dio3 loci, but more work remains to be done to define the role and properties of the long ncRNAs in these cases. For the sake of simplicity, the Igf2r locus will be described in detail here [Fig. 3].

Igf2r and two neighboring genes, Slc22a2 and Slc22a3 (solute carrier 22a2 and 22a3), are expressed maternally. This region harbors one paternally expressed transcript, Airm [antisense Igf2r RNA] that overlaps Igf2r. Several
nonimprinted genes also reside within this cluster, including Slc22a1, Mas1, and Plg [Plasminogen]. This locus contains at least two DMRs. DMR1, which encompasses the Igf2r promoter, is methylated on the repressed paternal chromosome (Stöger et al. 1993), but this methylation is acquired progressively during embryogenesis and is a consequence of silencing. DMR2, located in the second intron of Igf2r, is methylated on the expressed maternal chromosome. Because this methylation is present in female germ cells and persists throughout embryonic development (Brandis et al. 1993; Stöger et al. 1993; Lucifero et al. 2002), the region fulfills the initial criteria for serving as a locus-wide ICR. Consistently, deletion of DMR2 leads to reactivation of paternal Igf2r expression (Wutz et al. 1997).

DMR2, or the ICR for the Igf2r locus, contains the promoter for the Airn ncRNA, an unspliced and polyadenylated transcript that extends 108 kb in an antisense orientation to Igf2r (Wutz et al. 1997; Lyle et al. 2000). Airn is reciprocally imprinted to Igf2r and expressed exclusively from the unmethylated paternal allele. Because deletion of the ICR eliminates Airn expression and causes loss of Igf2r imprinting, the Airn ncRNA has been proposed to play a role in silencing Igf2r expression on the paternal chromosome. In fact, Igf2r is imprinted in almost all tissues where Airn is expressed (Hu et al. 1999), and is biallelically expressed in the absence of Airn mRNA (Latos et al. 2009). Surprisingly, deletion of the paternal ICR results in loss of imprinting of the neighboring, but not overlapping, Slc22a2 and Slc22a3 genes, in addition to Igf2r [Zwart et al. 2001]. Maternal expression of Slc22a2 and Slc22a3 is restricted to placenta at specific times during development, and no differential methylation has been detected around either Slc22a2 or Slc22a3 genes, indicating that imprinting of these genes may be regulated by elements located at a distance (Zwart et al. 2001).

Data from this deletion reinforce the proposal that it contains an ICR for the entire cluster, or, alternatively, that the Airn RNA itself or its transcription is required for silencing of Igf2r, Slc22a2, and Slc22a3. To distinguish between these two possible mechanisms, a mutant Airn allele was generated by insertion of a polyadenylation cassette that truncates the transcript to 4% of its length [Air-T] (Sleutels et al. 2002). This mutation does not perturb Airn imprinting because expression of Air-T and hypomethylation of the Airn promoter remain exclusively paternal. In contrast, Igf2r is biallelically expressed, with loss of paternal methylation in the Igf2r promoter region, directly implicating the Air RNA or its transcription in the repression of Igf2r. Unexpectedly, imprinting of Slc22a2 and Slc22a3 is also lost in midgestation placenta as a result of the Air-T mutation. This result was surprising, given that Airn does not overlap Slc22a2 and Slc22a3 but is transcribed in the opposite direction. Overall, these results demonstrate that the full-length Airn transcript, or transcription through the region, is required for the silencing of imprinted genes on the paternal allele.

Consistent with the results reported for Airn, truncation of Kcnq1ot1, which encodes a paternally expressed long ncRNA that mediates imprinting at the Kcnq1 locus, also results in a loss of imprinting for the adjacent genes (Mancini-Dinardo et al. 2006; Shin et al. 2008). However, a notable exception was reported for one of the truncation alleles [Shin et al. 2008]. Cdkn1c was still imprinted in a subset of embryonic tissues, demonstrating that Cdkn1c imprinting can be regulated by a mechanism independent of the Kcnq1ot1 ncRNA. Given that deletion of the locus ICR KvDMR1 results in loss of imprinting of Cdkn1c in all tissues [Fitzpatrick et al. 2002], the ncRNA-independent mechanism may rely on element(s) within KvDMR1. Accordingly, two CTCF-binding sites within KvDMR1 have been identified that are occupied in vivo only on the unmethylated paternal allele [Fitzpatrick et al. 2007]. Thus, it is possible that, for some genes, redundant mechanisms are in place to ensure appropriate imprinting.

The question remains, however, as to how these long ncRNAs function. While it is clear that the intact ncRNA is required, it is not known if the RNA itself or transcription through the complex is necessary. Moreover, the ncRNAs silence overlapping and nonoverlapping genes alike, suggesting that RNAi-based mechanisms may not be viable [Paufer et al. 2007]. A recently published set of studies proposes a function for the RNA in recruitment of histone post-translational modification machinery in placental tissues. RNA fluorescence in situ hybridization experiments showed that Airn and Kcnq1ot1 form RNA clouds at their site of transcription [Nagano et al. 2008; Pandey et al. 2008; Terranova et al. 2008; Redrup et al. 2009]. Terranova et al. (2008) show that these long ncRNAs are associated with a repressive histone compartment and Polycomb group proteins. This nuclear compartment is also devoid of RNA polymerase II and
exists in a three-dimensionally contracted state. Other studies on the Airn ncRNA go further in suggesting that the ncRNAs actively recruit repressive histone modifications [Nagano et al. 2008]. In this latter case, data are presented that are consistent with the notion that Airn actively recruits the histone H3 Lys 9 methyltransferase G9a. Use of the Air-T truncation allele results in reduced recruitment of G9a and biallelic expression. Although these experiments are consistent with an active recruitment function for the long ncRNAs, the studies point to an association between the RNA and the repressive chromatin machinery, and more studies, possibly temporal in nature, are required to prove a causal relationship between expression of the long ncRNA and subsequent recruitment of the epigenetic machinery followed by silencing. An additional important question concerns how some genes adjacent to the ncRNA are silenced, while others are immune to this process.

Nevertheless, the above experiments suggest a mechanism only in extraembryonic tissues. Interestingly, for both the Igf2r and Kcnq1 domains, a larger number of genes are imprinted in the placenta than in the embryo. Thus, it is plausible that the mechanism differs between extraembryonic and embryonic tissues. Barlow and colleagues [Pauler et al. 2007] propose that Airn ncRNA may silence Igf2r because of transcriptional interference, and that Airn might mediate imprinting through the process of being transcribed. Here again, more experiments are required to address and resolve these mechanistic considerations.

Evolution of imprinted gene clusters

Efforts to understand the origin and mechanism of imprinting have led investigators to study genomic organization and allelic expression patterns in divergent species. Because parthenogenesis can occur in species outside of mammals, it has been suggested that mammals alone have imprinted genes. These ideas were reinforced by the genomic analysis of imprinted loci [Smits et al. 2000]. Although imprinting in this region is well conserved in mice and humans, all orthologous genes in marsupials appear to be biallelically expressed. The DLK1/DIO3 imprint domain shows stronger conservation at the genome level but also harbors no imprinted genes in marsupials [Edwards et al. 2008]. Curiously, as with the Prader-Willi locus, the acquisition of nonprotein coding RNAs at the DLK1/DIO3 locus coincides with the appearance of imprinting. Thus, imprinting has been acquired at different times in different domains, after the evolution of viviparity. Studies of other imprinted loci in comparison to the already published evolutionary studies will most certainly provide insight into the critical elements that are required for the acquisition of imprinting.

In conclusion, a lot has been learned regarding genomic imprinting mechanisms in the last 20 years, but much remains to be elucidated. First, the full catalog of imprinted genes must be defined. It is likely that tissue-specific imprinted genes remain to be identified. The comparison of such genes in tissues with and without imprinting could be extremely helpful in defining and exploring mechanisms of transcriptome diversity as well as building theories regarding the evolution and purpose of imprinted gene expression. Second, it is clear that our knowledge of the epigenetic machinery that establishes, maintains, and protects imprints is incomplete. Here, advances in technology that enable the analysis of small numbers of germ cells and cells of the preimplantation embryo are essential. Third, a more complete understanding of insulator function and mechanism of large ncRNA action is required. Finally, the cloning and characterization of additional imprinted loci in marsupials and monotremes is necessary. Given the wealth of information the early studies have provided, it is certain that more evolutionary studies will contribute greatly to our knowledge of genomic imprinting.

Acknowledgments

I acknowledge the stimulating discussions from past and present members of my laboratory and I appreciate comments on the manuscript from Drs. Trish Labosky and Joanne Thorvaldsen. Work in my laboratory was supported by U.S. Public Health Service grants GM51279 and HD42026.

References


