Mechanism of imprinting on mouse distal chromosome 7

JUSTIN F-X. AINSCOUGH*, ROSALIND M. JOHN AND M. AZIM SURANI
Wellcome/CRC Institute of Cancer and Developmental Biology, and Physiological Laboratory, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK
(Received 14 May 1998 and in revised form 13 July 1998)

Summary
Genomic imprinting is an epigenetic mode of gene regulation that results in expression of the autosomal ‘imprinted’ genes from only a single allele, determined exclusively by parental origin. To date over 20 imprinted genes have been identified in mouse and man and these appear to lie in clusters in restricted regions on a subset of chromosomes. This may be a critical feature of imprinting suggesting a domain-type mode of regulation. Imprinted domains are replicated asynchronously, show sex-specific meiotic recombination frequencies and have CpG-rich regions that are differentially methylated, often associated with the imprinted genes themselves. Mouse distal chromosome 7 is one such domain, containing at least nine imprinted genes spanning over 1 Mb of DNA. For the maternally expressed \( p57^{Kip2} \) gene, passage through the female germline is essential to generate the active state, whereas passage through the male germline is needed to force the maternally expressed \( H19 \) gene into an inactive state. It is therefore possible that the mouse distal chromosome 7 imprinted domain is actually composed of two or more independently regulated subdomains.

1. Distal chromosome 7 is a maternally expressed domain

Work by Bruce Cattanach and others has shown that imprinted genes are vital for mammalian development, whereby epigenetic modifications in the germline regulate mono-allelic expression of parental alleles (McGrath & Solter, 1984; Surani et al., 1984; Cattanach & Kirk, 1985; Cattanach et al., 1992). Uniparental disomy (UPD) of some autosomes in the mouse was demonstrated to show non-complementation resulting in phenotypes ranging from early embryonic lethality to abnormalities in behaviour. This suggested the presence of genes important for both development and normal behaviour on these chromosomes which are differentially expressed from the two parental alleles.

Several of these imprinting effects are associated with chromosome (Chr) 7 (Searle & Beechey, 1990; Cattanach et al., 1992). Maternal duplication of distal Chr 7 leads to growth retardation and mid-gestational lethality, while paternal duplication results in early embryonic lethality. Distal Chr 7 maternally-deficient embryos lack placental spongiotrophoblast (McLaughlin et al., 1996), which is attributed to loss of activity of the maternally expressed \( Mash2 \) gene. However, chimeric animals that contain two paternal copies of distal Chr 7 survive longer and show striking growth enhancement, in part due to a double dose of the paternally expressed \( Igf2 \) gene (Ferguson-Smith et al., 1991). There are at least seven more genes in this region that are preferentially expressed from one parental allele and both their gene order and, in most cases, their imprinting are conserved in humans at 11p15.5 (Fig. 1, Table 1). This region is associated with the genetically complex fetal overgrowth disorder, Beckwith–Wiedemann syndrome (Ping et al., 1989). \( Igf2, \) \( Ins2, \) and some intergenic transcripts of unknown function (Moore et al., 1997), are contained within a 30 kb region in this domain and are the only genes that are expressed from the paternally inherited chromosome. The rest of the domain contains genes that are expressed almost exclusively from the maternally inherited chromosome. For those genes examined there are regions of parental-origin-specific
Distal7

Fig. 1. Gene organization on mouse distal chromosome 7/human 11p15.5. Nine imprinted genes located on distal chromosome 7 in mouse/11p15.5 in human, are bounded by non-imprinted genes. Direction of transcription, where known is indicated by a horizontal arrow. Translocation breakpoints, associated with Beckwith–Wiedemann syndrome in humans are indicated by vertical arrows. Not to scale.

Table 1. Imprinted genes on mouse distal chromosome 7/human 11p15.5

<table>
<thead>
<tr>
<th>Gene (mouse/human)</th>
<th>Product</th>
<th>Expressed allele</th>
<th>Null phenotype (mouse)</th>
<th>Overexpression phenotype (mouse)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H19/H19</td>
<td>Untranslated RNA</td>
<td>Maternal</td>
<td>Growth enhancement</td>
<td>None</td>
</tr>
<tr>
<td>Igf2/Igf2</td>
<td>Insulin-like growth factor 2</td>
<td>Paternal</td>
<td>Growth retardation</td>
<td>Growth enhancement</td>
</tr>
<tr>
<td>Ins2/INS</td>
<td>Insulin</td>
<td>Paternal</td>
<td>Growth retardation</td>
<td>n.d.</td>
</tr>
<tr>
<td>Tapa/TAPA</td>
<td>Transmembrane protein</td>
<td>Maternal</td>
<td>N.d.</td>
<td>N.d.</td>
</tr>
<tr>
<td>Mash2/ASCL2</td>
<td>bHLH transcription factor</td>
<td>Maternal</td>
<td>Placental failure</td>
<td>N.d.</td>
</tr>
<tr>
<td>Kvlqt1/KVLQT1</td>
<td>Potassium channel protein</td>
<td>Maternal</td>
<td>N.d.</td>
<td>N.d.</td>
</tr>
<tr>
<td>p57kip2/p57kip2</td>
<td>Cyclin-dependent kinase inhibitor</td>
<td>Maternal</td>
<td>Perinatal lethality/growth retardation</td>
<td>Growth retardation</td>
</tr>
<tr>
<td>Impt/IMPT1</td>
<td>Multiple membrane spanning polyspecific transporter-like gene</td>
<td>Maternal</td>
<td>N.d.</td>
<td>N.d.</td>
</tr>
<tr>
<td>Ipl/IPL (TSSC3)</td>
<td>TGAD 51</td>
<td>Maternal</td>
<td>N.d.</td>
<td>N.d.</td>
</tr>
</tbody>
</table>

The genes are presented in the order in which they are located on the chromosomes. Lowercase type denotes the mouse homologue, while uppercase type denotes the human homologue. Where known the function of the gene product, the expressed allele, and phenotypes of both over-expression and gene deletion/disruption, are indicated. n.d., not determined.

1, Bartolomei et al. (1991); 2, Zhang & Tycko (1992); 3, DeChiara et al. (1991); 4, Ohlsson et al. (1993); 5, Giddings et al. (1994); 6, Guillemot et al. (1995); 7, Alders et al. (1997); 8, Gould & Pfeiffer (1998); 9, Lee et al. (1997b); 10, Hatada & Mukai (1995); 11, Chung et al. (1996); 12, Dao et al. (1998); 13, Qian et al. (1997); 14, Duvillie et al. (1997).

differential methylation at CpG islands and in general it is the active allele that is unmethylated (see Ainscough & Surani, 1996; John & Surani, 1996). However, Igf2 is an exception to this since it is the active paternal allele that shows a greater degree of methylation in two regions: one 3 kb upstream of the first promoter, and another within the gene body. The actual CpG island associated with the Igf2 gene is unmethylated on both parental alleles (Sasaki et al., 1992; Feil et al., 1994).

Asynchronous DNA replication at imprinted regions in some somatic cells, where the paternal allele
Imprinting on mouse distal Chr 7

is replicated before the maternal allele (Kitsberg et al., 1993; Bickmore & Carothers, 1995), suggests differences in chromatin structure between the parental regions. The replication asynchrony across distal Chr 7 represents a late maternal pattern rather than an early paternal pattern since in H19 null mice the two chromosomes replicate synchronously, at the same time as the Rpl23 (L23mrp) gene, whereas in wild-type mice the maternal H19 region replicates after Rpl23 (Greally et al., 1998). The replication asynchrony ends somewhere between H19 and the Rpl23 gene, which is bi-allelically expressed in both mouse and human and may mark one boundary of the imprinted domain (Tsang et al., 1995; Zubair et al., 1997). The other boundary may be marked by the NAP2 gene, which is also bi-allelically expressed (Hu et al., 1996) and within 14 kb of the maternally expressed IPL gene (Qian et al., 1997).

Most of the imprinted genes on distal Chr 7 are expressed predominantly from the maternal allele, consistent with the fact that most regions of differential methylation show hypomethylation of the maternal locus. Therefore mouse distal Chr 7 could be viewed as a maternally expressed, unmethylated, late replicating domain where the paternal-specific expression of Igf2, Ins2 and the intergenic transcripts is a secondary effect brought about by their relative proximity to the maternally expressed H19 gene (Leighton et al., 1995; Ripoche et al., 1997; Ainscough et al., 1997).

2. Origin of the primary imprint: maternal or paternal?

The mechanism by which the individual genes on distal Chr 7 become mono-allelically expressed is not fully understood. However, recent compelling data suggest that there are different imprints for the different genes, and that passage through both the male and the female germline is required to obtain appropriate mono-allelic expression of all the imprinted genes in this domain (Tada et al., 1998; Obata et al., 1998; Y. Kato, unpublished observations).

(i) The p57Kip2 region is maternally activated while the H19 region is paternally repressed

Embryonic germ (EG) cells are pluripotential stem cell lines derived from the primordial germ cells (PGCs) of a developing embryo, which appear to retain the appropriate epigenetic modifications of the PGCs (Labosky et al., 1994). At day (d) 11.5–12.5 of development PGCs from both male and female embryos seem to have an equivalent epigenotype, having undergone a major global demethylation event in early embryogenesis (Kafri et al., 1992). Consistent with this, the maternally expressed p57Kip2 gene is unmethylated on both alleles in EG cells derived from both male and female PGCs at this stage. However, it is intriguing that after differentiation, in the form of primary embryonic fibroblasts (PEFs) from chimeric mice, both male and female copies become methylated and are presumably not active (Tada et al., 1998). Embryos generated by the transfer of a male d 14.5–16.5 PGC into an enucleated oocyte also exhibit methylation and inactivity of p57Kip2, as judged by Southern and in situ hybridization analysis, respectively (Y. Kato, unpublished observation). This tendency towards the inactive state remains the case in the mouse germline until later in development, since the p57Kip2 gene is not activated until its passage through a growing oocyte (Obata et al., 1998). Embryos that contain one genome from a neonatal-derived non-growing oocyte (ng) and the other from a fully grown oocyte (fg) express only the fg p57Kip2 allele. The ng allele is silent, as demonstrated by allele-specific RT-PCR. This suggests that activation (or inhibition of silencing) of the maternal p57Kip2 allele results from epigenetic modifications that occur during oocyte growth, illustrated in Fig. 2.

In contrast to p57Kip2, the maternally expressed H19 gene, which is located less than 1 Mb away, appears to require transmission through the male germline to render the paternal allele inactive (Fig. 2). SSCP analysis demonstrated that H19 is expressed at equivalent levels from both alleles in ng/fg embryos. Therefore, unlike p57Kip2, the silencing mechanism for this gene is entirely male germ-line-specific (Obata et al., 1998). EG cells from d 11.5–12.5 PGCs show some degree of sex-specific differential methylation, where male cells are largely methylated at the H19 locus and female cells are substantially unmethylated (Tada et al., 1998). This suggests that the methylation imprint on H19 is either not completely erased until after this time or is in the process of being (re)established in the male cells. However, H19 is expressed in embryos derived from male d 14.5–16.5 PGCs, where the gene is also unmethylated on both alleles (Y. Kato, unpublished observation). One possibility is that the H19 gene is demethylated after d 12.5 in the male germline, and is subsequently remethylated after d 16.5 (Fig. 2). Alternatively, the hypermethylated pattern observed in male d 12.5 EG cells may be maintained throughout male germ cell development (the methylation imprint being established prior to d 11.5). In this case the demethylation and bi-allelic expression of H19 in d 16.5 male PGC embryos could be caused directly by female germ-line-specific factors following transfer of the male nuclei into the oocyte cytoplasm.

Since these experiments into erasure and establishment of genomic imprints do not investigate the
germ cells directly it is still not clear what actually occurs in the germline. However, the combined data do suggest that the epigenetic modifications that lead to activation or repression of the $p57^{Kip2}$ gene differ from those of the linked $H19$ gene. In addition, all the other imprinted genes studied to date are unmethylated in both male and female EG cells from d 11:5 PGCs, and retain this state upon differentiation, suggesting that the somatic imprint for these genes has been completely erased by this stage (Tada et al., 1998). It is not clear how or why the $p57^{Kip2}$ gene shows a unique response to differentiation after d 12:5, when both male and female copies become remethylated/inactivated, this susceptibility being retained in the female germline until after the oocyte is fully developed (Obata et al., 1998). One possible explanation is that the genes on distal Chr 7 have different mechanisms of imprint erasure and re-establishment from genes in other regions. However, it is noteworthy that the other regions analysed are all methylated on the maternal allele in somatic cells, whereas distal Chr 7 has a mostly paternal-specific methylation pattern. It may be that the paternal pattern is actually a default state such that lack of a late-acting female-specific activity renders the region methylated upon differentiation, irrespective of the initial state. However, as both $H19$ and Igf2 are unmethylated in differentiated d 14:5 male PGCs this appears unlikely; the $p57^{Kip2}$ gene is therefore unique in this respect. Since the Mash2 gene is also repressed in these cells (Y. Kato, unpublished observation) it could be that the Mash2–$p57^{Kip2}$ region is regulated independently from the $H19$–Igf2 region. Thus distal Chr 7 may actually contain two separate imprinting subdomains, one of which requires maternal-specific activation late in oocyte development while the other
requires paternal-specific repression of \( H19 \) and consequent activation of \( Igf2 \).

3. Identifying imprint control elements using transgens

(i) \( Igf2/H19 \) region

To address whether the region containing the reciprocally expressed \( H19 \) and \( Igf2 \) genes is an imprinted domain in its own right, containing all the necessary elements for appropriate expression and imprinting of both genes, we constructed a reporter system using a 130 kb YAC as a transgene, and monitored activity of these genes following maternal and paternal transmission (Ainscough et al., 1997). This analysis clearly demonstrated that essential elements are located within 100 kb upstream and 30 kb downstream of the \( H19 \) gene, which allow for correct imprinting of both expression and methylation of \( H19 \) and \( Igf2 \). \( H19 \) was only ever expressed from the maternal allele irrespective of YAC copy number or position of integration. \( Igf2 \) imprinting, however, showed greater susceptibility to both position and copy number effects. This lends support to the notion that \( H19 \) is a true imprinted gene in this domain whereas \( Igf2 \) imprinting (and perhaps also that of \( Ins2 \) and the intergenic transcripts) is a secondary effect that serves to accurately control the expression levels of these genes. Such fine tuning is extremely important for the \( Igf2 \) gene since imbalance in the relative level of insulin-like growth factor II (IGF-II) protein can have gross effects on somatic growth (Wolf et al., 1994; Ward et al., 1994; Sun et al., 1997; Eggenschwiler et al., 1997). The fact that two other maternally expressed genes, \( Igf2r \) and \( Meg1 \), are known to have a role in controlling IGF-II activity lends support to this notion (Ludwig et al., 1995; Miyoshi et al., 1998).

Detailed analysis of the \( H19 \) expression pattern from the YAC, determined by \textit{in situ} hybridization on an \( H19 \) null background (Ripoche et al., 1997), showed that while enhancers for expression in many tissues of both mesodermal and endodermal origin are present, at least one further set of enhancers is located outside this region (unpublished data). However, work with smaller transgenes containing only the \( H19 \) gene, two downstream endoderm-specific enhancers, and 4 kb of upstream sequence, has indicated that a target sequence for \( H19 \) imprinting is most likely located in the \( H19 \) upstream region (Bartolomei et al., 1993; Pfeifer et al., 1996; Elson & Bartolomei, 1997), although imprinting of these transgenes was much less strictly maintained than that seen from the YAC. The target sequence identified is therefore not completely sufficient for proper imprinting of \( H19 \), since it is susceptible to both position and copy number effects.

In fact only high copy number lines have been demonstrated to show imprinting, unlike the YAC which imprints \( H19 \) from both low and high copy lines. Single copy \( H19 \) mini-transgenes are most frequently silenced completely, which suggests that the imprinting seen from these transgenes when in multi-copy form could perhaps be caused by alteration of the local chromatin structure with increased transgene size, possibly reflecting the increased number of control elements (enhancers and imprint target sequences) produced.

It is unlikely that the \( H19 \) proximate element is responsible for imprinting of the whole distal Chr 7 domain, since deletion experiments on the \( H19 \) gene only affected \( Igf2 \) and \( Ins2 \), but had no notable effects on other genes located further upstream of \( Ins2 \) (Leighton et al., 1995). In addition, translocation breakpoints in the human 11p15.5 region, which are associated with Beckwith–Wiedemann syndrome, fall into two clusters, one of which disrupts the \( KVLQT1 \) gene (Fig. 1) (Hoovers et al., 1995). Since these translocations may lead toactivation of maternal \( IGF2 \), or at least up-regulation of the paternal allele, additional long-range control elements, or regions that organize differential chromatin structure, could exist in the distal Chr 7 imprinting region. Further evidence that the imprinting process on distal Chr 7 is much more complex than a single imprinting element was acquired from deletion experiments of a 13 kb region between \( Igf2 \) and \( Ins2 \) in mouse fibroblasts (Hu et al., 1997). This affected imprinting of both \( Igf2 \) and \( H19 \), indicating that another imprint control element (or target sequence) resides upstream of \( Igf2 \), which is at least involved in maintaining the reciprocal expression of these two genes.

Since the evidence suggests that imprinting of the \( p57^{kip2} \) gene is likely to be independent from that of \( Igf2 \) and \( H19 \) the ability of a series of large genomic fragments containing \( p57^{kip2} \) to imprint at ectopic loci is also being tested in this laboratory. It will be of interest to assess the role of regions analogous to those where the Beckwith–Wiedemann syndrome translocation events occur in humans, the nearest of which is within 70 kb of the \( p57^{kip2} \) gene, within the \( KVLQT1 \) gene (Lee et al., 1997).

4. Imprinting elements and gene silencers

Recent data have provided intriguing evidence that the regions identified as imprinting control elements, or gene-specific imprinting target sequences, may in fact function as silencer elements, at least in \textit{Drosophila}. This unusual series of experiments was initially set up to investigate how putative imprint control elements would act in \textit{Drosophila}, which does not show parental imprinting and has no detectable
DNA methylation, thus looking for any underlying conserved epigenetic mechanisms in the absence of DNA methylation. That the H19 upstream region target sequence may in fact have silencer activity was shown by generating transgenic flies with H19 mini-transgenes. A 1-2 kb region located 2 kb upstream of H19 had a strong silencing effect on both the LacZ reporter and the mini-white transformation marker used, irrespective of parental origin. When this region was deleted both genes became active (Lyko et al., 1997). The imprinting centre (IC) of the Prader–Willi/Angelman syndrome (PWS/AS) region (Buiting et al., 1995) was subsequently shown to have a similar silencing activity in Drosophila (Lyko et al., 1998), suggesting a possible mechanistic conservation between gene silencing in flies and genomic imprinting in mammals. We have preliminary evidence that a second such element with silencer activity in flies exists in the Igf2/H19 domain (unpublished data). Perhaps all such elements act in concert to confer the better imprinting seen from the YAC transgene. Whilst this association between gene silencing in Drosophila and genomic imprinting in mammals is intriguing it is not yet clear how these elements actually function in their proper context. Since only one of the parental alleles of an imprinted gene is silenced it may be that silencing represents the default state and maternal activation represents the true imprint. However, it must be remembered that for both the H19 gene and the PWS/AS region the element that functions to silence in flies is methylated in mice. When unmethylated, as is the case in Drosophila, the mouse genes are active.

5. Lessons from Beckwith–Wiedemann syndrome

Beckwith–Wiedemann syndrome (BWS) is often associated with paternal duplication and/or maternal deficiency of 11p15.5, strongly implicating the activity of imprinted genes in this disorder. Many sporadic non-UPD patients show bi-allelic IGF2 expression (Weksberg et al., 1993; Joyce et al., 1997) and some also show altered methylation patterns around H19 and IGF2 (Reik et al., 1994, 1995), implicating the involvement of these genes. Over-expression of IGF2 in mice causes fetal overgrowth with some similarities to BWS (Sun et al., 1997; Eggenschwiler et al., 1997). Intriguingly these experiments showed that the Igf2 locus is susceptible to trans activating signals, resulting in over-expression of the endogenous gene to levels greater than 2-fold. A similar trans activation effect on Igf2 was also demonstrated between two unlinked copies of the Igf2/H19 YAC transgene (Ainscough et al., 1997). One possible explanation for these epigenetic effects is that maternal repression of the Igf2 gene can be readily destabilized, perhaps because of the chromatin environment where the genomic region is primarily set up for maternal expression. However, not all cases of BWS are associated with increased levels of IGF2 (Nystrom et al., 1994), indicating that other nearby genes may also play a role. Importantly, mutations in the p57Kip2 coding sequence have recently been identified in 10% (5/54) of non-UPD patients (Hatada et al., 1996; O’Keefe et al., 1997; Lee et al., 1997a), and deletion of p57Kip2 in mice leads to perinatal lethality associated with abdominal wall defects and cleft palate, both of which are found in BWS patients (Yan et al., 1997; Zhang et al., 1997). Overgrowth is not seen in p57Kip2 null mice, however, although in humans, unlike in mice, some 10% of p57Kip2 expression comes from the paternal allele (Matsuoka et al., 1996; Chung et al., 1996). It is therefore possible that this low-level residual expression in humans who have mutations in their maternally inherited p57Kip2 allele leads to overgrowth, since the p21 family of cyclin-dependent kinase inhibitors, to which p57Kip2 belongs, may function at low levels to stimulate cell cycle progression, whereas at high levels they inhibit cell cycle progression (LaBaer et al., 1997). Since altered expression has been observed at both the IGF2 and the p57Kip2 loci it is probable that BWS is an imprinted contiguous gene syndrome covering over 500 kb, whereby more than one gene is responsible for the array of phenotypes seen.

It is noteworthy that five germline balanced maternally inherited chromosomal translocations fall within the 11p15.5 domain, but they do not physically disrupt either the p57Kip2 gene or the IGF2 gene (Hoovers et al., 1995). Instead they fall within a third gene in the region, KVLQT1 (Lee et al., 1997b). Although KVLQT1 cannot be excluded as a candidate gene for BWS, it seems unlikely since mutations are known to cause long QT syndrome and are not associated with growth defects. KVLQT1 spans 350 kb and four transcripts are produced, two of which (isoforms 3 and 4) are apparently non-coding. This is reminiscent of the paternally expressed SNRPN gene that lies in the region associated with PWS/AS (Dittrich et al., 1996).

If both IGF2 and p57Kip2 play a role in BWS, one possibility is that the KVLQT1 region contains a single imprinting centre similar to that proposed for PWS/AS (Buiting et al., 1995) that is directly disrupted by the translocation events. However, due to the distance between the different translocations (Lee et al., 1997b) such an element would have to be very large. In addition, the evidence discussed in this review suggests that a single imprinting centre is not likely to control all the genes in this region, and that the Igf2/H19 region may in fact be controlled independently from the p57Kip2 region. An alternative possibility to account for this is that disruption of the region, whereby the telomeric half that contains MASH2, IGF2 and H19 is translocated to another
chromosome, may separate the genes from specific chromatin structures required for correct establishment of maternal and/or paternal imprinting in the germline, or maintenance of imprinting in somatic cells. Exposure of this internal region of distal Chr 7/11p15.5 may be sufficient to upset the finely balanced control of mono-allelic gene activity throughout the entire domain.

References


gene, is biallelically expressed in fetal and adult tissues. Human Molecular Genetics 5, 1743–1748.


