Engineering mouse models to investigate the function of imprinting

Rosalind M. John

Advance Access publication date 10 May 2010

Abstract

Some insight into the developmental basis for imprinting specific genes during the evolution of mammals can be gained from conventional gene ‘knockout’ studies. However, the consequences of full loss of function are often wide-ranging and may obscure the critical, dosage-related phenotype. This review focuses on transgenic techniques employed to alter the dosage of imprinted genes, including the application of bacterial artificial chromosome transgenic mice, in imprinting research. Advantages of dosage-based techniques over conventional knockout studies will be discussed, with examples. Important applications of transgenic mice in imprinting research, including studying gene expression patterns, the identification of imprinting centres and isolating the consequences of altered gene dosage, are reviewed with a particular focus on the imprinted domain on mouse distal chromosome 7.

Keywords: Genomic imprinting; gene dosage; transgenic technologies

INTRODUCTION

Imprinted genes are those that are expressed from one parental chromosome as a consequence of epigenetic marks set down either in the father’s or the mother’s germ line [1]. While most imprinted genes are exclusively expressed from one parental allele, the imprinted expression of some genes can be seen as just a bias in expression between the alleles. The degree of imprinted expression can also vary between tissues, between developmental stages and even between splice variants. Nonetheless, the fundamental feature of genomic imprinting is that it alters the expression level of a particular gene from one parental allele without altering essential function.

The most widely accepted model proposed to rationalise the existence of genomic imprinting in mammals is known as the parental conflict hypothesis [2, 3]. This hypothesis is primarily based on data from the targeted deletion of imprinted genes in mice. This data suggests that the paternal genome has acted to silence growth-restricting genes and that the maternal genome has responded by silencing growth-enhancing genes. However, not all targeted deletions alter embryonic growth and recent reviews highlight the influence of imprinted gene products on mammalian behaviour and/or metabolism [4, 5]. While studies on loss of function are important for our understanding of gene function per se, altering the dosage of the imprinted gene may reveal more informative data on the function of the imprint. Reducing the expression of the single active allele by design is not achievable with current technologies. Increasing the level of an imprinted gene or restoring biallelic expression is more feasible and several approaches have been used to achieve this goal.

MATERNAL OR PARENTAL DUPLICATION OF THE WHOLE GENOME

The existence of imprinted genes was first suggested because both parental genomes are necessary for normal embryonic development [6–8]. Parthenogenetic embryos, which carry two maternal genomes, rarely develop beyond the 25 somite stage and are growth restricted with rudimentary extraembryonic (EE) tissues whereas androgenetic embryos, carrying two paternal genomes, are developmentally delayed.

Corresponding author. Rosalind M. John, Cardiff School of Biosciences, Cardiff University, Cardiff, CF10 3AX, UK. Tel: +44 2920870145; Fax: +44 2920876328; E-mail: johnrm@cf.ac.uk

Rosalind M. John is a Senior Lecturer in Genetics. She began her studies on Genomic imprinting in 1995 working as a post-doctoral fellow with Professor Azim Surani at Cambridge University.

© The Author 2010. Published by Oxford University Press. All rights reserved. For permissions, please email: journals.permissions@oxfordjournals.org
and die at the 6–8 somite stage but with relatively normal EE development, at least at this stage. In these monoparental embryos, the entire diploid genome is derived maternally or paternally, disturbing the dosage of all imprinted genes. Even as a conservative estimate, this probably translates as at least 80 alterations in the expression of protein-encoding imprinted genes, with both loss of expression and increased dosage.

**MATERNAL OR PATERNAL DUPLICATIONS OF THE DISCRETE CHROMOSOMAL REGIONS**

On a more localised scale, Robertsonian and reciprocal translocations can be used to generate embryos carrying uniparental disomies and uniparental duplications of whole or selected chromosomal regions [9–11]. These studies have been particularly useful both in identifying specific chromosomal regions containing imprinted genes and in characterising the phenotypes associated with paternal or maternal disomy. Currently, there are 14 phenotypically distinguishable imprinted regions located on eight mouse chromosomes (2, 6, 7, 9, 11, 12, 17 and 18) each with a complex phenotype that we now know reflects the existence, in most cases, of multiple imprinted genes within the region. There are also regions of the genome where imprinted genes are known to be located, on chromosomes 1, 5, 10, 13, 14, 15 and 19, where no obvious abnormal phenotype has been reported associated with maternal or paternal duplication. This work is summarised at [http://www.mousebook.org/catalog.php?catalog=imprinting](http://www.mousebook.org/catalog.php?catalog=imprinting).

Mouse chromosome 7 contains at least five imprinted domains. Maternal or paternal disomy of the most distal domain alone is predicted to directly alter the expression of at least 20 genes contained within two mechanistically distinct imprinted domains. The IC1 domain (Figure 1) spans the maternally-expressed non-coding RNA, H19, and the
paternally-expressed embryonic growth factor, insulin-like growth factor-2 (Igf2) and insulin II (Ins2) genes [12, 13]. The second domain, IC2, encompassing the paternally-expressed non-coding RNA, Kcnq1ot1 (Lit1), the maternally expressed Phlda2, Slc22a18, Cdkn1c, Kcnq1, Ascl2 and Tsc4 genes and additional, less well-characterised genes (Figure 1). Mice with maternal duplication/paternal deficiency (MatDup.dist7) for this distal seven region are growth retarded by 50% at embryonic day (E) 17.5 with small placentae whereas those with paternal duplication/maternal deficiency are lost earlier in development [11]. It is possible to genetically normalise the expression of individual genes or groups of genes within a disomic embryo by introducing a targeted deletion into the cross. MatDup.dist7 embryos are rescued from lethality when they also carry a targeted deletion that restores Igf2 expression [14]. But given the complexity of these crosses, the low frequency of generating the required combinations and the introduction of potentially confounding genetic background effects, this is still a considerable undertaking.

MODIFYING THE ENDOGENOUS LOCUS TO ENGINEER BIALLELIC EXPRESSION

Imprinted domains are regulated by discrete DNA regions termed imprinting centres (ICs) [15]. These are functionally defined by engineering targeted deletions. Inheritance through one parental germ line releases the domain from imprinted expression (loss of imprinting, LOI) whereas inheritance through the other parent’s germ line generally, but not always, has no effect (Table 1). These models can provide excellent tools for understanding the consequences of increased gene dosage particularly where the IC controls a few, well-characterised targets.

Igf2/H19 domain: This domain encompasses three imprinted genes: two encode a protein product (Igf2 and Ins2) and the third gene produces the non-coding RNA, H19. Additionally, there is at least one imprinted microRNA within the domain [16]. In mice carrying maternally-derived deletions which eliminate the IC for this region, the normally silent maternal Igf2 and Ins2 alleles become transcriptionally active and expression of H19 transcript is reduced or eliminated. These changes result in an increased birth weight of between 8 and 30% as a consequence of enhanced embryonic growth earlier in development [12, 13] (Figure 1). In the placenta, the consequence of LOI of this domain is placental megaly with a 250% increase in the number of giant trophoblast cells and glycogen cells [17]. Igf2 encodes a potent regulator of embryonic growth [18] suggesting that excess Igf2 drives these overgrowth phenotypes. However, the H19 transcript may intrinsically possess growth-regulating properties [19] and the functional consequence of excess Ins2 is unknown.

Igf2r domain: Paternal inheritance of a targeted deletion of the IC for the Igf2r domain on mouse chromosome 17 releases Igf2r, Ooct2/Slca22a2 and Ooct3/Slc22a3 from paternal silencing [20, 21]. Biallelic expression results in growth restriction with weight difference apparent between E15.5 and E17.5 [22]. Similar to the Igf2 region, Igf2r is the only protein-coding gene within its imprinted domain that is expressed in the embryo. The growth phenotype is reciprocal overgrowth induced by loss of expression of Igf2r [23, 24]. Growth restriction is most likely due to increased lysosomal targeting and degradation of embryonic Igf2 protein induced by excess Igf2r. However, despite the normal appearance of the placenta, subtle phenotypes in response to excess Ooct2/Slc22a2 and/or Ooct3/Slc22a3 cannot be excluded.

Rasgrf1 domain: In this model, rather than deleting an IC, a repeat sequence located just upstream of a differentially methylated region (DMR) was replaced by an enhancer overriding the action of the IC and resulting in biallelic Rasgrf1 [25]. Expression of Rasgrf1 mRNA is apparent as early as E13.5 but an overgrowth phenotype manifests only after postnatal day 16 suggesting that the dosage of this gene is important during this critical period of postnatal maternal care. It will be interesting to determine if these individuals show an altered metabolism or whether they demand more resources from their mothers. It could also be illuminating to examine the maternal physiology and behaviour, both in animals with wild-type Rasgrf1 expression and those with biallelic expression.

Gbh10 (Megl) domain: Targeted deletion of a maternally-methylated DMR overlapping the transcriptional start site of the paternally expressed brain-specific Gbh10 transcript results in loss of expression of this transcript but biallelic expression of the major, maternally-expressed Gbh10 transcript accompanied by embryonic and placental growth.
restriction from E12.5 and E13.5, respectively [26]. The deletion also causes LOI of two adjacent genes, Ddc and Cobl, but subtle differences between two versions of the DMR deletion [26] and the inverse correlation with the phenotype due to loss of Grb10 expression [27] suggest the phenotype is due to the double dosage of Grb10.

In all these cases, although the major player in each phenotype has been suggested, subtle or masked phenotypes as a consequence of altered expression of adjacent genes within each domain cannot be discounted without genetically normalising the expression of individual genes within the LOI embryo. For the larger, more complex imprinted domains, LOI may involve distantly located genes under the control of the targeted IC that have yet to be identified.

### Table I: Summary of DMR targeted deletion models

<table>
<thead>
<tr>
<th>Region</th>
<th>Altered expression</th>
<th>Embryonic phenotype</th>
<th>Extraembryonic phenotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr. 2 Exon IA DMR</td>
<td>2.2 x Gnas</td>
<td>Reduced parathyroid hormone</td>
<td>n.d.</td>
<td>[49]</td>
</tr>
<tr>
<td>Chr. 2 Nespa DMR</td>
<td>2.5 x Nespa 0 x Nespa 0 x Gnasxl 0.3 x Exon IA 1.8 x Gnas</td>
<td>Do not feed and have a thin body</td>
<td>n.d.</td>
<td>[50]</td>
</tr>
<tr>
<td>Chr. 7 H19DMR</td>
<td>1.25–1.6 x Igf2 2 x Ins2</td>
<td>Neonates 108% WT [13] E14.5 I100% [17] Adults I17% WT [51]</td>
<td>El4.5 placenta 140% WT with 250% increase in glycogen and giant cells and 120% increase in spongiotrophoblast cells</td>
<td>[13, 17, 51]</td>
</tr>
<tr>
<td>Chr. 7 PWS-IC</td>
<td>Incr. Ube3a 0 x Snrpn 0 x Zfp217 0 x Ndn 0 x Ipw</td>
<td>Neonate 60–73% WT Neonatal lethal Low blood glucose</td>
<td>n.d.</td>
<td>[52–54]</td>
</tr>
<tr>
<td>Chr. 7 AS-IC</td>
<td>2 x Snpn 1 x Ndn</td>
<td>Neonates 85% WT</td>
<td>n.d.</td>
<td>[54]</td>
</tr>
<tr>
<td>Chr. 7 KvDMR</td>
<td>1.8–2 x Phlda2, Slc22a18, Cdkn1c, Kcnq1, Tssc4 and Ascl2</td>
<td>E14.5 embryos 80% WT E16.5 embryos 83% WT Proportional decrease in the weight of internal organs such as kidney, lung and liver</td>
<td>El6.5 placenta 80% WT</td>
<td>[36, 44]</td>
</tr>
<tr>
<td>Chr. 11 Ghrb10-DMR</td>
<td>2 x Ghrb10 type I 2 x Cabl 0 x Ghrb10 Type II 0 x Ddc</td>
<td>E12.5 embryos 80% WT E18.5 embryos 70% WT</td>
<td>El3.5 placenta 85% WT E18.5 placenta 80% WT</td>
<td>[26, 55]</td>
</tr>
<tr>
<td>Chr. 12 IG-DMR</td>
<td>2 x Dlk1 4.5 x Rtl1 2 x Dio3</td>
<td>Lethality after El6. Short body, skeletal defects. Similar growth to WT No phenotype (1.5 x Dlk1, 1.8 x Rtl1, 1.5 x Dio3)</td>
<td></td>
<td>[56, 57]</td>
</tr>
<tr>
<td>Chr. 17 Air-DMR</td>
<td>2 x Igf2r 2 x Skn22a2 2 x Skn22a3</td>
<td>E15.5 embryos 80% WT</td>
<td>E175 placental weights normal</td>
<td>[22]</td>
</tr>
</tbody>
</table>

n.d. = no data. Deletions that result only in loss of expression are not included.

*a* Indicates predicted expression levels, no published data.

**USING TRANSGENES TO INCREASE GENE DOSAGE**

A more directed approach to engineering increased gene dosage is to provide mice with extra copies of a specific gene locus. A major advantage of a transgenic approach is that information can be obtained on both imprinting mechanism and the functional consequences of increased gene expression in a single model (Tables 2 and 3). Transgenes can be used to delimit ICs and then modified to test the function of specific regions [20]. Transgenes that contain both the natural promoter and enhancers of a particular gene are ideal as they accurately recapitulate tissue- and temporal-specific expression to permit the assessment of the phenotypic consequence of over expression in the appropriate setting.
The clear advantage of the smaller transgene (<50 kb) is ease of handling, manipulation and the ease of generating of multiple transgenic lines. The first example of a transgene that replicated the imprinting process was a 15-kb transgene spanning the H19 gene. Maternal expression and paternal methylation of the transgene was reported for 3/7 lines [28]. A disadvantage of small transgenes is that the site of integration can influence both their expression and their imprint and in this same study 4/7 lines did not imprint the H19 transgene [28]. Additionally, 10–20% of sequences not associated with imprinted loci can imprint at random [29]. Small transgenes are also less likely to carry all the
required elements to recapitulate expression of the endogenous locus.

These disadvantages can be overcome by increasing the size of the genomic region included in the transgene. The first examples of such an approach employed yeast artificial chromosomes (YAC). An 130-kb YAC spanning the H19/Igf2 locus was used to generate multiple lines of transgenic mice and 4/4 lines that were tested accurately imprinted H19 [30] (Figure 2). Similarly, 4/4 lines carrying a 300-kb YAC spanning the Igf2r locus showed appropriate imprinted expression [20]. YACs contain an average insert of 500–600 kb of DNA but can be much larger. The merits of a YAC based approach have been discussed previously [31] but essentially rely on the size of the insert and the amenability of YACs to genetic modification by homologous combination in yeast. For example, the H19/Igf2 YAC was modified to include a β-galactosidase marker under the control of the Igf2 promoter facilitating the analysis of both spatial and imprinted expression [30] and the Igf2r YAC was modified to remove a DMR within the Igf2r gene, now known as the Air-DMR, ablating imprinted expression [20].

In recent years, transgenes based on genomic clones that can be propagated in Escherichia coli have largely superseded YACs due to the development of technologies to insert or remove specific sequences. Bacterial artificial chromosomes (BACs) are single copy plasmids based on the E. coli fertility plasmid (F Factor) and P-1 derived artificial chromosomes (PACs) are based on the bacteriophage P1. Methods for modifying BACs have been reviewed recently [32, 33] and, similar to YACs, insertion of reporters such as β-galactosidase has been useful in the analysis of both spatial and imprinted expression (Tables 2 and 3). These vectors are capable of holding large genomic inserts, up to 300 kb. Importantly, genes on BACs are expressed from their endogenous promoters at the appropriate level and time, predominantly without being affected by the site of integration [34]. The advantage of these E. coli-based vectors over YACs is their relative stability in culture, ease of DNA preparation and the ability to obtain sequence information easily [35].

ICs are classically defined by targeted deletion (Table 1). YACs/BACs that span known ICs reproduce imprinted gene expression more reliably than smaller transgenes and can be used to exclude or refine ICs (Table 2, Figure 2). The IC2 region on mouse distal chromosome 7 contains several maternally-expressed genes and one paternally

![Figure 2: Map of the mouse distal chromosome 7 imprinted domains and extent of the transgenes used to study autonomous imprinting. Black rectangles represent genes. Arrows above the rectangles represent expression from the maternal allele and those below the rectangles represent expression from the paternal allele. Filled lollipops represent regions of DNA methylation. Filled lollipop above the rectangles represent methylation the maternal allele and those below represent methylation the paternal allele as reported in a somatic cell. Location of the transgenes (black lines) is depicted below the map of the locus. Only transgenes driving expression with spatial and temporal accuracy are included and a percentage figure is given which represent the ability of the transgene to recapitulate imprinted expression when integrated elsewhere in the genome with 100% indicating that all transgenic lines showed imprinted expression and 0% indicating that no transgenic lines showed imprinted expression. (A) 15-kb H19 plasmid-based transgene [28], (B) 130-kb YAC-based transgene [30], (C) 105-kb P1-based transgene [48], (D) and (E) 85- and 270-kb BAC-based transgenes [38], and (F) 800-kb YAC-based transgene [37]. Grey triangle indicates insert site for a GFP transgene that is expressed only from the maternal allele [42].]
expressed non-coding RNA, Kcnqlot1/Lit1 (Figure 2). An IC, KvdMR1, has been functionally defined for this region located at a DMR within the Kcnql gene, which marks the origin of Kcnqlot1/Lit1 transcription [36]. Paternal inheritance of a deletion of this region results in biallelic expression of all the normally paternally repressed genes, including Cdknlc and Ascl2. An 800-kb YAC containing this IC2 domain recapitulates imprinted expression for Phlda2, Slc22a18, Cdknlc, Kcnql and Kcnqlot1, although not Ascl2 and Tsc4 [37] (Figure 2). A shorter region of 270 kb contained within a BAC transgene also spans KvdMR1 and several target genes, including Cdknlc, Phlda and Slc22a18, but does not drive imprinted expression [38] (Figure 2). This suggests that the KvdMR1 region requires additional sequence to function as an IC. The Kcnqlot1/Lit1 transcript is 108 kb in length but only the first 10 kb is present on the 270-kb BAC transgene. A truncation of this transcript, leaving KvdMR1 intact, also leads to loss of maternal silencing [39, 40] suggesting that the length of the Kcnqlot1 transcript is important for imprinting. There is a second DMR present within the IC2 domain encompassing the Cdknlc gene. Differential methylation at Cdknlc-DMR is established after fertilisation [41]. A BAC transgene encompassing this region does not imprint, excluding an autonomous role for this secondary DMR in establishing imprinted expression of Cdknlc [38] (Figure 2).

Useful information can be gleaned from artificial transgenes that integrate into imprinted domains and acquire allele-specific expression providing potentially important information on the DNA sequence that characterise the targets of the IC [29]. A recent example is that of an engineered GFP insertion between Th and Ius2 on mouse distal chromosome 7 (Figure 2 for relative position) that acquires maternal-allele expression [42]. In this example, the transgene has integrated near the IC1 domain but behaves like an IC2 domain gene, being paternally silenced and, like Cdknlc, acquiring paternal allele DNA methylation after fertilisation.

FUNCTIONAL STUDIES

Studies testing the imprinting capacity of a particular sequence can also generate mice with increased dosage of an imprinted gene(s). In particular, studies on transgenes that drive expression with spatial and temporal accuracy at levels similar to the endogenous gene are particularly significant (Table 3).

Due to the nature of imprinted loci, large transgenes often span multiple genes and, similar constraints apply as with LOI models. The critical advantage that transgenes provide over LOI models for exploring function is that the exact nature and number of genes is defined by the transgenic sequence. Phenotypes can be assigned unequivocally to the gene sequences within the transgene, providing more than one line is characterised to exclude position-of-integration influences.

A second compelling reason for adopting a transgene-based approach is that these transgenes can be rapidly modified. Once the targeting construct is built, homologous recombination with the endogenous sequence in ES cells and the identification of correctly targeted clones can take several months whereas engineering a similar modification in a BAC can take just a few weeks. This technique is particularly helpful in situations where the altered dosage of one gene masks a subtle phenotype associated with altered expression of other genes within the same domain. Cdknlc and Phlda2 provide a textbook example of this phenomenon. Expression of Cdknlc from a single copy BAC transgene intrinsically restricts embryonic growth from E13.5 and this growth restriction persists into adulthood [43] (Figure 1). Two additional intact genes, Phlda2 and Slc22a18, are present on the same BAC transgene and over expressed in this model. The intrinsic growth restriction phenotype was assigned to excess Cdknlc and not Phlda2 or Slc22a18 by engineering a modification to abolish Cdknlc expression from the BAC (Figure 1). Mice carrying this modified version of the BAC over express only Phlda2 and Slc22a18 and embryonic growth at E13.5 is unaffected. In addition to clarifying the role of Cdknlc in regulating embryonic growth, this modification exposed a second embryonic growth restriction phenotype. The growth of embryos over expressing only Phlda2 and Slc22a18 was normal at E13.5 but progressively slows down as development proceeds with a 13% loss of embryonic weight by birth [44, 45]. Thus, a careful transgenic dissection of this domain revealed two distinct growth restriction phenotypes.

Many imprinted genes are expressed in the placenta but the expression status of transgenes in EE tissues and the functional consequence has largely been neglected (Table 3). In the case of Igf2,
Cdkn1c, Nnat and Dlk1, the transgenes are known not to include the placental enhancers which must lie at a significant distance from the gene body [30, 38, 46, 47]. Accurately driving over expression of these genes in EE tissues will be challenging. An exception is the Phlda2/Slc22a18 locus. The 85-kb transgene spanning this locus contains placental and embryonic enhancers for both Phlda2 and Slc22a18, accurately driving over expression of both genes [44, 45]. By combining mice carrying a single copy of the transgene with double transgenic mice also carrying a maternally inherited targeted deletion of Phlda2, the key role for Phlda2 in regulating placental weight and glycogen storage was identified [45] (Figure 1).

More sophisticated approaches combining conditional targeting either of ICs or of individual genes with tissue-specific cre recombinases will be useful in examining the relationship between embryonic growth potential and placental modifications driven by imprinting. It should be possible to use these techniques to address the relevance of gene dosage after birth, for example during the preweaning period where maternal care is still important, and also after this time point, when the rational for imprinting genes solely as a consequence of parental conflict is less convincing.

Key Points

- Mice can be engineered that over express imprinted genes under their endogenous promoters with temporal and spatial accuracy.
- Transgenic approaches that alter the dosage of an imprinted gene are invaluable tools for dissecting the functional consequences of genomic imprinting.
- Where LOI models or transgenes involve more than one imprinted gene, modification approaches or a judicious combination with gene deletion models can be useful in assigning phenotypes to individual genes and revealing phenotypes masked by adjacent loci.

FUNDING

Biotechnology and Biological Sciences Research Council grant to R.M.J.

Acknowledgements

The author would like to thank Prof. Louis Lefebvre for helpful discussions.

References


