The origins of genomic imprinting in mammals

Carol A. Edwards A,B, Nozomi Takahashi A, Jennifer A. Corish A and Anne C. Ferguson-Smith A

A Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.
B Corresponding author. Email: cae28@cam.ac.uk

Abstract. Genomic imprinting is a process that causes genes to be expressed according to their parental origin. Imprinting appears to have evolved gradually in two of the three mammalian subclasses, with no imprinted genes yet identified in prototheria and only six found to be imprinted in marsupials to date. By interrogating the genomes of eutherian suborders, we determine that imprinting evolved at the majority of eutherian specific genes before the eutherian radiation. Theories considering the evolution of imprinting often relate to resource allocation and recently consider maternal–offspring interactions more generally, which, in marsupials, places a greater emphasis on lactation. In eutherians, the imprint memory is retained at least in part by zinc finger protein 57 (ZFP57), a Kruppel associated box (KRAB) zinc finger protein that binds specifically to methylated imprinting control regions. Some imprints are less dependent on ZFP57 in vivo and it may be no coincidence that these are the imprints that are found in marsupials. Because marsupials lack ZFP57, this suggests another more ancestral protein evolved to regulate imprints in non-eutherian subclasses, and contributes to imprinting control in eutherians. Hence, understanding the mechanisms acting at imprinting control regions across mammals has the potential to provide valuable insights into our understanding of the origins and evolution of genomic imprinting.

Additional keywords: epigenetics, evolution, marsupials.

Introduction

Genomic imprinting is an epigenetic process whereby genes are expressed from one of the two chromosome homologues in a parent-of-origin-specific manner. Because imprinted genes are, in effect, functionally haploid, they lose the protection that diploidy provides against deleterious mutations. This poses many questions about the evolution of this epigenetic process. Herein we discuss the current knowledge as to when imprinting arose in vertebrates, why the imprinting process may have evolved and how the mechanism of imprinting may have emerged.

Imprinting first arose in the therian lineage

To date, the only vertebrates in which genomic imprinting has been reported are mammals. In the mouse, >130 imprinted genes have been identified, most of which are found in clusters located at particular chromosomal regions in the genome. To identify when imprinting evolved in the mammalian lineage, various known eutherian imprinted genes have been tested in marsupials and monotremes (Table 1). Six imprinted genes have been identified in marsupials, namely insulin-like growth factor 2 (IGF2), the non-coding RNA gene H19, insulin (INS), IGF2 receptor (IGF2R), paternally expressed 10 (PEG10) and mesoderm specific transcript/paternally expressed 1 (MEST/PEG1; Killian et al. 2000; O’Neill et al. 2000; Suzuki et al. 2005, 2007; Ager et al. 2007; Smits et al. 2008), whereas no imprinted gene has been found in monotremes (Killian et al. 2000, 2001; Edwards et al. 2008).

The most intact imprinted gene cluster in marsupials is the IGF2/H19 domain. In eutherians, this region contains two protein coding genes, IGF2 and INS, and the long non-coding RNA (lncRNA) gene H19. In mice, Igf2 is paternally expressed in the developing embryo except in the choroid plexus and leptomeninges (DeChiara et al. 1991; Ferguson-Smith et al. 1991). IGF2 was first shown to be expressed from the paternally inherited chromosome in pouch young of the grey short-tailed opossum (Monodelphis domestica; O’Neill et al. 2000). More extensive analysis in the tammar wallaby (Macropus eugenii) found it to be paternally expressed in the fetal body and adult mammary gland, but biased paternal expression was found in the fetal brain and placenta (Suzuki et al. 2005; Stringer et al. 2012b, 2012c). The insulin gene (INS or Ins2 in mouse) is only imprinted in the human and mouse yolk sac where it too is expressed from the paternally inherited chromosome (Deltour et al. 1995; Moore et al. 2001). In tammar wallaby, INS is paternally expressed in the yolk sac membrane (YSM), the principal placenta in most marsupials (Ager et al. 2007), and in the adult mammary gland (Stringer et al. 2012c). The lncRNA gene H19 is maternally expressed in mice (Bartolomei et al. 1991) and an orthologous region has been identified in tammar...
The Igf2r gene encodes IGF2R and a mannose-6-phosphate receptor, both of which are maternally expressed in young brain (Killian et al. 2001), but imprinting is not imprinted, at least in adult prototherians. However, this scenario may not be the case for all primates because, in the Cynomolgus macaque, Igf2r was found to be imprinted in all individuals analysed (Cheong et al. 2015). In the mouse, the Igf2r gene lies within an imprinted cluster with solute carrier family 22 members 2 (Slc22a2) and 3 (Slc22a3), which are expressed from the maternally inherited alleles in the placenta. Their imprinting is regulated by the reciprocally imprinted, paternally expressed IncRNA, Airn, which is an antisense transcript to Igf2r (Zwart et al. 2000) Edwards et al. (2008). The population (Xu et al. 1993). However, this scenario may not be the case for all primates because, in the Cynomolgus macaque, Igf2r was found to be imprinted in all individuals analysed (Cheong et al. 2015). In the mouse, the Igf2r gene lies within an imprinted cluster with solute carrier family 22 members 2 (Slc22a2) and 3 (Slc22a3), which are expressed from the maternally inherited alleles in the placenta. Their imprinting is regulated by the reciprocally imprinted, paternally expressed IncRNA, Airn, which is an antisense transcript to Igf2r (Zwart et al. 2000). In marsupials, Igf2r is maternally expressed in Virginia opossum (Didelphis virginiana) but lacks the

### Table 1. Imprinting studies performed in marsupials and monotremes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Marsupial references</th>
<th>Monotreme references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Imprinted in marsupials and eutherians</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGF2</td>
<td>O’Neill et al. (2000)</td>
<td>Killian et al. (2001)</td>
</tr>
<tr>
<td>INS</td>
<td>Ager et al. (2007)</td>
<td></td>
</tr>
<tr>
<td>H19</td>
<td>Smits et al. (2008)</td>
<td></td>
</tr>
<tr>
<td>PEG1/MEST</td>
<td>Suzuki et al. (2005)</td>
<td></td>
</tr>
<tr>
<td>PEG10</td>
<td>Suzuki et al. (2007)</td>
<td></td>
</tr>
<tr>
<td><strong>Imprinted in eutherians only</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIO3</td>
<td>Edwards et al. (2008)</td>
<td></td>
</tr>
<tr>
<td>RTL1</td>
<td>Edwards et al. (2008)</td>
<td></td>
</tr>
<tr>
<td>SNRPN</td>
<td>Rapkins et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>UBE3A</td>
<td>Rapkins et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>ASB4</td>
<td>Suzuki et al. (2007)</td>
<td></td>
</tr>
<tr>
<td>SGCE</td>
<td>Suzuki et al. (2007)</td>
<td></td>
</tr>
<tr>
<td>PPP1R9A</td>
<td>Suzuki et al. (2007)</td>
<td></td>
</tr>
<tr>
<td>GRB10</td>
<td>Stringer et al. (2012a)</td>
<td></td>
</tr>
<tr>
<td>CDKN1C</td>
<td>Ager et al. (2008a)</td>
<td></td>
</tr>
<tr>
<td>KCNQ1OT1</td>
<td>Ager et al. (2008a)</td>
<td></td>
</tr>
<tr>
<td>PHLDA2</td>
<td>Suzuki et al. (2011)</td>
<td></td>
</tr>
<tr>
<td>L3MBTL</td>
<td>Aziz et al. (2013)</td>
<td></td>
</tr>
<tr>
<td><strong>Gene orthologue not in marsupials</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NNAT</td>
<td>Evans et al. (2005)</td>
<td></td>
</tr>
<tr>
<td>PEG3</td>
<td>Suzuki et al. (2011)</td>
<td></td>
</tr>
<tr>
<td>MAGEL2</td>
<td>Rapkins et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>MKRN3</td>
<td>Rapkins et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>NDN</td>
<td>Rapkins et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>AIRN</td>
<td>Weidman et al. (2006a)</td>
<td></td>
</tr>
<tr>
<td>NAPIL5</td>
<td>Wood et al. (2007)</td>
<td></td>
</tr>
<tr>
<td>INPP5F_V2</td>
<td>Wood et al. (2007)</td>
<td></td>
</tr>
<tr>
<td>Zrs1 (U2af1-rs1)</td>
<td>Wood et al. (2007)</td>
<td></td>
</tr>
<tr>
<td>MCTS2</td>
<td>Wood et al. (2007)</td>
<td></td>
</tr>
</tbody>
</table>

### Table 1. Imprinting studies performed in marsupials and monotremes

Genes are divided into those imprinted in marsupials, those that have a marsupial orthologue that is biallelically expressed and those with no marsupial orthologue. IGF2R, insulin-like growth factor receptor 2; INS, insulin; IGFBP5, insulin-like growth factor binding protein 5; H19, imprinted non-coding RNA; PEG1, paternally expressed 1; PEG10, paternally expressed 10; CDKN1C, cyclin dependent kinase inhibitor 1C; KCNQ1OT1, KCNQ1 opposite strand/antisense transcript 1; PHLDA2, pleckstrin homology like domain family A member 2; L3MBTL1 histone methyl-lysine binding protein; GTL2/MEG3, gene-trap locus 2/maternally expressed 3; NNAT, neurexin; PEG3, paternally expressed 3; MAGEL2, melanoma antigen, family L, 2; NNAT, neuronatin; PEG3, paternally expressed 3; MAGEL2, melanoma antigen, family L, 2; MKRN3, makorin, ring finger protein, 3; NDN, nedelin; AIRN, antisense Igf2r RNA; NAPIL5, nucleosome assembly protein 1-like 5; INPP5F_V2, inositol polyphosphate-5-phosphatase F variant 2; Zrs1, (U2af1-rs1), zinc finger (CCCH type), RNA binding motif and serine/arginine rich 1; MCTS2, malignant T cell amplified sequence 2.

Virginia opossum (Didelphis virginiana) but lacks the

**Table 1. Imprinting studies performed in marsupials and monotremes**

Genes are divided into those imprinted in marsupials, those that have a marsupial orthologue that is biallelically expressed and those with no marsupial orthologue. IGF2R, insulin-like growth factor receptor 2; INS, insulin; IGFBP5, insulin-like growth factor binding protein 5; H19, imprinted non-coding RNA; PEG1, paternally expressed 1; PEG10, paternally expressed 10; CDKN1C, cyclin dependent kinase inhibitor 1C; KCNQ1OT1, KCNQ1 opposite strand/antisense transcript 1; PHLDA2, pleckstrin homology like domain family A member 2; L3MBTL1 histone methyl-lysine binding protein; GTL2/MEG3, gene-trap locus 2/maternally expressed 3; NNAT, neurexin; PEG3, paternally expressed 3; MAGEL2, melanoma antigen, family L, 2; NNAT, neuronatin; PEG3, paternally expressed 3; MAGEL2, melanoma antigen, family L, 2; MKRN3, makorin, ring finger protein, 3; NDN, nedelin; AIRN, antisense Igf2r RNA; NAPIL5, nucleosome assembly protein 1-like 5; INPP5F_V2, inositol polyphosphate-5-phosphatase F variant 2; Zrs1, (U2af1-rs1), zinc finger (CCCH type), RNA binding motif and serine/arginine rich 1; MCTS2, malignant T cell amplified sequence 2.
differentially methylated region (DMR) in intron 2 that is the promoter for *Airn* in eutherians (*Killian et al. 2000*). *M6P/IGF2R* is biallelically expressed in the echidna and platypus. Interestingly, the monotreme orthologues lack the IGF2 binding domains and therefore produce proteins that only have M6P-binding properties, perhaps negating the need to be imprinted in these species. (*Killian et al. 2000*). Evidence for a marsupial *Airn* transcript has not been found (*Weidman et al. 2006a*), hence the mechanism regulating *IGF2R* imprinting in marsupials is not known.

**Peg10** is a retrotransposon derived gene that arose in the genome after the divergence of the therians from the monotremes (*Suzuki et al. 2007*). In the mouse, **Peg10** resides in a large imprinted cluster that also contains sarcoglycan, epsilon (Sgcε), which is also paternally expressed, ankyrin repeat and SOCS box-containing 4 (*Ashb4*), which is expressed from the maternally inherited chromosome and protein phosphatase 1, regulatory subunit 9A (*Ppp1r9a*), which exhibits maternally biased expression in the placenta (*Ono et al. 2003*). In the tammar wallaby, **PEG10** is paternally expressed in the embryo and yolk sac, but the other genes are bialleically expressed (*Suzuki et al. 2007*). The final gene that has been suggested as imprinted in marsupials is **MEST/PEG1**. This gene is paternally expressed in eutherians (*Kaneko-Ishino et al. 1995*), but shows biallelic but paternally biased expression in multiple different tammar tissues (*Suzuki et al. 2005*).

Together, these data indicate that genomic imprinting first arose in the therian lineage. No imprinted genes have been identified in the monotremes, but only four genes have been experimentally assessed (*IGF2, IGF2R, DLK1* (delta like non-canonical Notch ligand 1) and *DIO3* (deiodinase, iodothyronine deiodinase, iodothyronine deiodinase, type III); *Killian et al. 2000, 2001; Edwards et al. 2008*) and, furthermore, due to scarcity of material, no genes have been tested in embryos. Hence, it is possible that there is imprinting expression in monotreme fetuses that becomes biallelic in adults, as is the case for *IGF2* in humans (*Isa et al. 1996*).

**Most imprinted clusters were established in the eutherian ancestor before radiation**

Genes that are imprinted in eutherians but not in marsupials fall into two categories: those that have a marsupial orthologue and those that do not. Thirteen genes fall into the first category and 11 fall into the second (Table 1).

The **DLK1**/*GTL2* cluster contains genes from both categories (Fig. 1). The paternally expressed protein coding genes **DLK1** and **DIO3** are both present in the marsupials and monotremes but are bialleically expressed (*Weidman et al. 2006b; Edwards et al. 2008*). In addition, remnants of the retrotransposition event that formed the retrotransposon Gag like 1/paternally expressed 11 (*RTL1/PEG11*) gene in eutherians are present in marsupials but they lack the eutherian-specific open reading frames (*Edwards et al. 2008*). In contrast, the maternally expressed lncRNA, gene trap locus 2/maternally expressed 3 (*GTL2/MEG3*) and the long arrays of imprinted microRNAs (miRNAs) and small nucleolar RNAs (snRNAs) embedded within the **DLK1/GTL2** cluster have no orthologues in marsupials or monotremes.

The **Snprn** imprinted cluster is the largest eutherian imprinted domain, being over 3 Mb in the mouse. It is not present in marsupials (*Rapkins et al. 2006*). Misregulation of imprinting in this region in humans leads to two distinct neurological disorders, Prader–Willi syndrome and Angelman syndrome (*Buiting 2010*). In eutherians, the region contains two maternally expressed genes, ubiquitin protein ligase E3A (*Ube3a*) and ATPase, class V, type 10A (*Atp10a*), four paternally expressed protein-coding genes (makorin, ring finger protein, 3 (*Mknn3*), melanoma antigen, family L, 2 (*Magel2*), necdin (*Ndn*) and small nuclear ribonucleoprotein N (*Snrnp*)) and paternally expressed non-coding (nc)RNAs including large arrays of snoRNAs. Genomic studies of this region in marsupials reveal that the region is eutherian specific. **SNRPN** is present in the tammar wallaby and opossum, but it is tandemly repeated next to its parent gene small nuclear ribonucleoprotein B (**SNRPB**) (*Rapkins et al. 2006*). In the tammar wallaby, **SNRPN** is on chromosome 1q, whereas **UBE3A** is on chromosome 5. **MKRN3**, **MAGEL2**, **Ndn** and the snoRNAs are all absent in marsupials, indicating that region only came together in the eutherians perhaps along with the evolution of its imprinting (*Rapkins et al. 2006*).

In eutherians, the **CDKN1C** imprinted domain lies directly next to the **IGF2/H19** cluster and this synteny is conserved in the tammar wallaby (*Ager et al. 2008a*). The mouse **Cdkn1c** region contains 10 imprinted maternally expressed protein coding genes including: achaete-scute family bHLH transcription factor 2 (*Ascl2*), cyclin dependent kinase inhibitor 1C (**Cdkn1c**), potassium voltage-gated channel, subfamily Q, member 1 (**Kcnq1**), and pleckstrin homology like domain family A member 2 (**Phlda2**). The paternally expressed ncRNA gene *Kcnq1ot1*, which is an antisense transcript to *Kcnq1*, is known to regulate imprinting gene expression across the cluster (*Fitzpatrick et al. 2002*). Interestingly, **KCNQ1OT1** is expressed in the tammar wallaby, but is biallelically expressed and lacks the differentially methylated promoter that acts as the imprinting control region in eutherians (*Ager et al. 2008a*). Consistent with this, **CDKN1C** and **PHLD2A** have both been shown to be bialleically expressed in the tammar wallaby (*Ager et al. 2008a; Suzuki et al. 2011*), hence imprinting is also not conserved at this region.

Data from these studies indicate that most genes acquired imprinting regulation in the eutherian lineage. However, eutherian studies have only been performed on mammals belonging to the suborders Euarchontoglires (e.g. primates and rodents) or Laurasiatheria (e.g. canines and ungulates), whereas the Xenarthra and Afrotheria have not been studied. By studying available genomes of Xenarthra and Afrotheria species we are able to ascertain when genes missing from marsupial genomes first arose in the eutherian lineage. In addition, by looking at the conservation of elements, such as known imprinting control regions (ICRs) in mouse and human through eutherian evolution, we can infer whether imprinting arose before or after the eutherian radiation.

For example, we know that the **DLK1/DIO3** region in marsupials lacks the maternally expressed ncRNAs as well as an intact copy of **RTL1** (*Edwards et al. 2008*). In contrast, by performing sequence analysis on the **DLK1/DIO3** region in the
elephant (Afrotheria) and armadillo (Xenarthra) we identified GTL2, snoRNAs and miRNA orthologues within the domain (UCSC genome browser; Karolchik et al. 2012). Both species also contain intact copies of RTL1. Furthermore, regions of homology to the ICR, the intergenic DMR (IG-DMR), were found in both elephant and armadillo (Fig. 1). Bisulfite pyrosequencing of the IG-DMR conserved region in elephant samples shows the region is approximately 50% methylated, indicating that this region is likely to be a DMR in the elephant (Fig. 1).

Using comparative sequence analysis, we can predict when the majority of imprinted domains were established (Fig. 2). For example, some sequence conservation between the ICRs of the SNRPN, GRB10 (growth factor receptor bound protein 10) and GNAS (guanine nucleotide binding protein, alpha stimulating) clusters was identified in at least one species from all eutherian clades (UCSC genome browser conservation track; Karolchik et al. 2012), suggesting that imprinting regulation was established in the common eutherian ancestor for each of these domains. However, to confirm imprinting in all these domains, expression and methylation analyses would need to be performed in relevant species. Four of the genes that have no marsupial orthologue are retrotransposed copies of X-linked genes that have arisen at various points in eutherian evolution (Wood et al. 2007). Taken together, the data indicate that most of the imprinting that has been characterised in eutherians was established after their divergence from marsupials but before the eutherian radiation (between 65 and 130 million years ago), but that imprinting at other loci has arisen subsequently, for example...
when a retrotransposition event has occurred, and perhaps governed by novel selective pressures.

**Function of imprinted genes in eutherians**

In order to understand why the imprinting process arose in mammals, it is essential to analyse the functions of known imprinted genes. Several human syndromes are caused by dysregulation of imprinted regions through either uniparental disomies (UPDs), microdeletions or defects in imprinting regulation. The phenotypes of these syndromes provide some clues to the functions of imprinted genes, such as in growth and neural function. Silver-Russell syndrome (dysregulation of multiple clusters), Temple syndrome (*DLK1/GTL2* domain) and transient neonatal diabetes mellitus (*PLAGL1* domain) all lead to intrauterine growth retardation.
(Temple et al. 2000; Abu-Amero et al. 2008; Ioannides et al. 2014), whereas patients with Beckwith–Wiedemann syndrome (IGF2 and CDKN1C clusters) have fetal and postnatal overgrowth (Weksberg et al. 2010). Prader–Willi syndrome leads to hyperphagia and obsessive compulsive behaviour with mild mental retardation and is caused by maternal UPD15 or paternal deletions at 15q11–13, which contains the SNRPN cluster. Paternal UPD15 or maternal deletions at this domain lead to Angelman syndrome, another neurological disorder, but in this case exemplified by severe mental retardation hyperactivity and constant laughter (Nicholls and Knepper 2001). Mental retardation is also evident in patients with Kagami–Ogata syndrome (paternal UPD14), and mild intellectual disability is associated with Temple syndrome (maternal UPD14; Ioannides et al. 2014; Kagami et al. 2015).

An extensive review of phenotypes of mouse models of imprinted genes found that most imprinted genes function in neonatal transitions and metabolism in addition to fetal and neonatal growth, placentation and behaviour (Cleaton et al. 2014). Brain and behaviour was the most common function of murine imprinted genes, followed by placentation and energy homeostasis (Fig. 3). Interestingly, when we look at marsupial imprinted genes we see that they do not fall into the category associated with brain and behaviour, but rather most of them have a role in placentation (Fig. 3; Suzuki et al. 2005, 2007; Ager et al. 2007, 2008b; Smits et al. 2008; Stringer et al. 2012b). This suggests that the major role of imprinting in marsupials is the control of nutritional resources, whereas in eutherians imprinting is also important in controlling behaviour and postnatal adaptations during the life course.

Imprinted genes have major roles in placenta
tion

An important role for imprinted genes in human and mouse placenta is well established (for reviews, see Coan et al. 2005; Monk 2015). In mice, early experiments using pronuclear transfer demonstrated that a full complement of both maternally and paternally derived genomes was necessary for mammalian development to term (Barton et al. 1984; McGrath and Solter 1984; Surani et al. 1984). Gynogenetic and parthenogenetic conceptuses generally develop only as far as the 25 somite stage and have poorly developed extra-embryonic tissues (Surani and Barton 1983; Surani et al. 1984). In contrast, androgenotes have well developed extra-embryonic tissues but very retarded embryos (Barton et al. 1984), indicating roles for imprinted genes in both placental and fetal development.

One of the first imprinted genes discovered, Igf2, has a critical role in nutrient transfer in the placenta. Mice with a global paternal deletion of the gene have small placentas with fewer glycogen cells and increased spongiotrophoblast cells in the junctional zone (Lopez et al. 1996). These placentas, in the absence of fetal IGF2, have reduced levels of the amino acid transporter solute carrier family 3 member 2 (Slc38a2) (Constância et al. 2005). In addition, deletion of a labyrinthine trophoblast-specific transcript which derives from the P0 promoter (Igf2-P0) in mice leads to reduced diffusion and permeability from embryonic day (E) 15.5. However, in the presence of fetal IGF2, these placentas upregulate the glucose transporter gene solute carrier family 2 member 3 (Slc2a3) and the imprinted amino acid transporter solute carrier family 38 member 4 (Slc38a4) (Constância et al. 2002, 2005; Sibley et al. 2004). The Igf2r gene is also expressed in the placenta. Maternal deletion of this gene leads to placentomegaly and larger fetuses that cannot survive to term (Lau et al. 1994; Wang et al. 1994). Fetuses lacking a maternal copy of Igf2r have higher levels of circulating IGF2 (Lau et al. 1994) and the overgrowth phenotype is corrected in mice also lacking a functional Igf2 gene (Filson et al. 1993). Thus, although the primary function of Igf2 is in the regulation of nutrient supply and demand in the placenta and fetus, the primary role for Igf2r is in modulating IGF2 levels (Filson et al. 1993). Other genes involved in nutrient transfer are also imprinted in the placenta, including Slc22a2 and Slc22a3 in the Igf2r cluster (Zwart et al. 2001) and solute carrier family 22 member 18 (Slc22a18) in the Kcnq1 domain (Dao et al. 1998). Genes in the Dlk1/Gtl2 cluster regulate placental development and function (Georgiades et al. 2001; Sekita et al. 2008; Ito et al. 2015). Furthermore, Grb10 has been shown to influence fetal resource acquisition because deletion of this maternally expressed gene results in placental overgrowth and increased placental efficiency (Chalambous et al. 2010).

The placenta also has a critical endocrine role inducing and maintaining physiological changes in the mother. Mothers need to undergo substantial changes in food intake and metabolism,
the cardiovascular system, immune system and mammary glands during pregnancy, all of which are controlled, in part, by endocrine secretions by the conceptus. Recently, it was shown that fetus-derived DLK1 is found at high levels in maternal blood in mouse and human, where it plays a significant role in maternal metabolic adaptations (Cleaton et al. 2016). It has also been proposed that by regulating the size of different compartments of the placenta, imprinted genes are regulating not only resource allocation, but also the amount of hormones received by the mother, which, in turn, can influence the mother’s adaptation to pregnancy (John 2013). Seven imprinted genes were identified that affect placental endocrine lineages in the mouse: Ascl2, Phlda2, Cdkn1c and Igf2r, which are expressed from the maternally inherited copy, and paternally expressed 3 (Peg3), Peg10 and Igf2, which are paternally expressed (John 2013).

Deletion of Igf2 causes growth restriction in all endocrine lineages. Placental growth restriction has been reported for mice lacking Peg3 (Curley et al. 2004), and more recent expression data suggest that lack of Peg3 does have an effect on endocrine lineages (Broad and Keverne 2011; John 2013). Peg10 has a critical role in placenta; knockout mouse placentas completely lack spongiotrophoblast and embryos die at E10.5 due to placental failure (Ono et al. 2006). Igf2r affects the placenta by acting on IGF2, as discussed above (Filson et al. 1993). The three other maternally expressed genes that influence the endocrine compartments in the placenta are all located within the coordinately regulated Cdkn1c imprinted cluster. Phlda2 encodes a pleckstrin homology-like domain protein that acts to negatively regulate the spongiotrophoblast lineage (Frank et al. 2002) and Ascl2 encodes a transcription factor that is believed to repress the formation of parietal trophoblast giant cells, as well as acting upstream of Phlda2 to repress spongiotrophoblast formation (Tunster et al. 2016). Cdkn1c also represses spongiotrophoblast and labyrinthine trophoblast proliferation (Takahashi et al. 2000). Together, these data indicate that of the major site of imprinted gene expression is the placenta and their function in eutherians is to control the supply and demand of prenatal resources.

**Placenta and imprinting evolution**

Because imprinted loci repress one of their two gene copies, they exhibit functional haploidy and lose the protection that diploidy provides against deleterious mutations. This has led to much speculation about why the process arose. In this section we discuss three of the most popular theories of imprinting evolution, the conflict/kinship theory, the supply and demand theory and the maternal–offspring coadaptation model, and how they relate to what we currently know about marsupial and eutherian imprinting.

The reciprocal functions and imprinting status of Igf2 and Igf2r led to perhaps the most prevailing theory of imprinting evolution, the conflict/kinship theory (Moore and Haig 1991). This theory argues that imprinting arose as a consequence of a conflict of interest between maternally and paternally derived genomes driven by prenatal resource control. It suggests that paternally expressed genes, such as Igf2, would be growth enhancing, favouring greater resource allocation from the mother both in utero and perinatally, perhaps to the detriment of later offspring of the mother. Conversely, imprinted genes expressed from maternally inherited chromosomes, such as Igf2r, would be growth limiting, allowing her offspring to traverse the birth canal and to conserve her resources for future offspring. Most imprinted genes with a placental phenotype in mouse knockout models follow this prediction. Deletion of six paternally expressed genes, namely Igf2, Peg1/Mest, Peg3, Peg10, Plag1 and Rtl1, causes placental growth restriction, and deletion of six imprinted genes expressed from maternally inherited chromosomes, namely Igf2r, Cdkn1c, Phlda2, Grb10, Rtl1as and H19, causes placentomegaly (Cleaton et al. 2014). Interestingly, deletion of one maternally expressed gene, namely Ascl2, leads to complete loss of spongiotrophoblast and is embryonic lethal at E10.5, similar to the Peg10-knockout mouse, which also lacks spongiotrophoblast and dies at E10.5, which would argue against the conflict hypothesis (Guillemot et al. 1994; Ono et al. 2006). However, Ascl2-null mice have an increased number of parietal trophoblast giant cells and mice expressing an Ascl2 transgene show a reduction in the spongiotrophoblast compartment, suggesting that ASCL2 does have a growth-limiting function in these cell lineages (Tunster et al. 2016). Furthermore, a larger placenta does not necessarily mean a more efficient placenta. For example, mice with paternal UPD of chromosome 12 (i.e. two paternally inherited Dlk1/Gtl2 domains including Rtl1) show placemegaly, but have defects in all three layers of the placenta, including reduced fetal capillary volume (Georgiades et al. 2000, 2001). The conflict/kinship theory seems only to apply to the subset of imprinted genes regulating growth and placentaion, including those that evolved imprinting in marsupials.

The supply and demand theory of imprinting proposes that imprinted genes in the placenta are controlling the supply of nutrients to the placenta, whereas imprinted genes in the fetus are controlling the demand for nutrients (Reik et al. 2003). For example, in the fetus paternally derived Igf2 is controlling demand by promoting growth, whereas in the placenta it is controlling supply through its effects on diffusion, permeability and transport. Conversely, maternally expressed genes Phlda2 and Cdkn1c in the adjacent imprinting cluster can counteract the effects of Igf2 by reducing nutrient supply. Maternally expressed Igf2r would act to suppress supply and demand in the fetus and placenta through its negative regulation of IGF2 levels. Reik et al. (2003) propose that the regulation of placental supply and fetal demand is a particular function of imprinted genes and suggests the coevolution of imprinting and placentaion. Again, this theory is formulated around the more ancestral placental-specific prenatal resource control functions of imprinted genes. The evolutionary pressure behind the direction of imprinting is predicted to be dependent, in part, on parental conflict, and the supply and demand theory can be seen as an extension to the conflict/kinship theory.

The maternal–offspring coadaptation model for imprinting evolution was first proposed by Wolf and Hager (2006). Coadaptation occurs when offspring genes evolve to function with a particular parentally supplied environment and when the parental genotype for this environment becomes associated with the offspring genes that are adapted to it (Wolf and Brodie 1998).
The maternal–offspring coadaptation model suggests that the expression of genes from maternally inherited chromosomes increases the adaptive integration of mother and offspring genomes where there are close maternal–offspring interactions, such as in the placenta. This theory can explain the observation that all genes that are exclusively imprinted in the placenta are maternally expressed (Wolf and Hager 2006). However, although it can include postnatal maternal–offspring relationships that do not include the placenta, it is hard to reconcile this model to imprinted genes with key roles to play in postweaning functions.

Both the supply and demand and the conflict/kinship theories predict that imprinting would exist in animals with placentas but not in egg-laying animals. Phylogenetic data are in agreement with this and suggest imprinting and placentation evolved in parallel: the number of imprinted genes has increased as mammalian placentas became more complex and able to sustain longer in utero development. To date, no imprinted genes have been identified in non-mammalian vertebrates or monotremes (Killian et al. 2000, 2001; Edwards et al. 2008). Placentation in monotremes is the most rudimental. Although these are egg-laying species, fetuses receive endometrial secretions via the shell, suggesting the level of nutrients received by the fetus is indeed dictated by the mother. Eggs are laid in the short-beaked echidna at the 18–20 somite stage (~18 days after conception) and are incubated for 10–11 days.

Imprinting is only seen in marsupials and eutherians, mammals that have a direct apposition of the endometrium and fetal membranes. Early in gestation, marsupials receive nutrients via a shell membrane, but this ruptures later in gestation at approximately the same stage that monotreme eggs are laid (18 days after removal of pouch young in the tammar wallaby, which reactivates the diapause-arrested blastocyst). After the marsupial shell membrane ruptures, a direct contact between the YSM and the endometrium is formed, possibly allowing for the fetus to have some control over the levels of nutrients it receives. Interestingly, IGF2 and growth hormone expression in the tammar wallaby placenta increase at this time, indicating that the conceptus is able to secrete hormones and signal to the mother (Menzies et al. 2011). Once the shell membrane ruptures, gestation continues in the tammar wallaby for a further 8 days.

The six identified imprinted genes in marsupials are all expressed in the marsupial placenta, and five have been shown to be imprinted (IGF2R imprinting has not been assessed in marsupial placenta yet; Suzuki et al. 2005, 2007; Ager et al. 2007, 2008b; Smits et al. 2008; Stringer et al. 2012b). Although gene manipulation technologies are evolving for non-model organisms, to date it has been difficult to perform embryonic manipulations on marsupials, so the roles of these genes in marsupial placentation have not been assessed. However, we do know from mouse models that five of these genes affect size and/or efficiency of the eutherian placenta (IGF2, H19, IGF2R, PEG1/MEST and PEG10) and fit with the supply and demand model of imprinting evolution. The imprinting of these genes also agrees with the conflict hypothesis: IGF2, PEG1/MEST and PEG10 are growth enhancing and paternally expressed, whereas IGF2R and H19 are maternally expressed and growth suppressing. The maternal–offspring coadaptation model is harder to reconcile with the marsupial data. This theory predicts genes involved in placentation would be maternally expressed; however, only two maternally expressed genes have been identified in the marsupial placenta, IGF2R and H19, and major functions of both these genes are in modulating IGF2 levels in the fetus (Filson et al. 1993; Ripoche et al. 1997; Wilkin et al. 2000; Gabory et al. 2009). Of course, it is possible that a completely different set of genes regulates marsupial placentation, and these may exhibit metatherian-specific imprinting.

Mammary gland: a site for genomic imprinting?

The theories described above propose specific roles for imprinted genes in resource acquisition from mother to child, particularly in utero. However, mammalian young continue to acquire nutrients from their mother after birth via lactation. Conflict is unlikely to influence the evolution of imprinting in the mammary gland because, unlike the placenta, the genome of the offspring’s father is not represented here and the grandpaternal genomes, which are present, are unlikely to be in conflict because both are equally likely to be present in the neonates. However, conflict could still lead to imprinting in neonates, and one would expect genes involved in suckling and appetite to be imprinted if this is the case. The coadaptation model can also be allied to imprinting in postnatal resource allocation (Renfree et al. 2013). For example, coadaptation may lead to the imprinting of genes involved in suckling, because this would enhance the genetic integration of intimate maternal–offspring interactions (Stringer et al. 2014). However, this theory also would not predict genes in the mammary gland being imprinted because it only requires genes to be expressed from maternally inherited chromosomes in the pup to exhibit increased relatedness to the mother. Therefore, none of the models proffered to date predicts the mammary gland to be a major site for imprinting, implying that if imprinting does occur here, it would be due to different evolutionary pressures or the absence of pressures in the mammary gland selecting against an imprinting status established during development.

To date, the mammary gland has not been a tissue that has been extensively studied in the imprinting field. However, Grb10 has recently been shown to be expressed from the maternally inherited allele in lactating mammary glands in mice expressing a Grb10-driven LacZ reporter (Cowley et al. 2014). It was demonstrated that Grb10 performs complimentary functions in mothers and pups: in pups, GRB10 suppresses growth, whereas in mothers Grb10 expression increases milk production. Thus, postnatal Grb10 expression fits with the supply and demand theory of imprinting because it controls the supply of nutrients from the mother and the pups’ demand for resources. The complementary and pleiotropic effects of Grb10 expression in mother and offspring also fit with coadaptation (Cowley et al. 2014). Studies specifically assessing imprinting in mammary glands have not been performed for any other genes. A recent study using RNA sequencing (RNA-seq) demonstrated that 25 known imprinted genes were imprinted in virginal and/or lactating mammary glands (Andergassen et al. 2017). The authors of that study concluded that because there was little
difference in imprinting between virginal and lactating mammary glands, there was no specific role for imprinted expression in lactation. Interestingly, in that study expression was reported for Grb10 at similar levels for virgin and lactating mammary glands (Andergassen et al. 2017), whereas Cowley et al. (2014) found no Grb10 reporter expression in virginal mammary glands. It is likely that gross RNA-seq analysis of whole mammary glands fails to take into account the cellular complexity of this organ, indicating that a more detailed systematic approach is needed to fully understand the role of imprinted genes in the mammary gland and postnatal provisioning.

In addition to Grb10, six mouse imprinted models have been shown to affect postnatal maternal provision: null models of Peg3, GnasX1, Magel2, Cdkn1c and Igf2r show impaired suckling (reviewed in Cleaton et al. 2014). Deletion of three paternally expressed genes, namely Magel2, Peg1/Mest and Peg3, has been reported to cause defects in maternal care (Lefebvre et al. 1998; Li et al. 1999; Schaller et al. 2010). Interestingly, loss of function in Peg3 causes failure in postnatal feeding when the deletion is present in the mother or her offspring, suggesting coadaptation. When mothers lack a functional copy of the gene, they fail to respond to signals from the wild-type placenta, they increase food intake in early pregnancy and they show impaired milk let down, leading to growth-retarded wild-type pups. When the pup is lacking a functional copy of the gene, it has impaired suckling efficiency compared with wild-type litter mates (Curley et al. 2004). These observations led to the proposal of a modified coadaptation model suggesting that genes that are simultaneously expressed in the maternal hypothalamus, placenta and fetal hypothalamus, such as Peg3, would tend to be switched off on the maternal allele (Keverne and Curley 2008). This would allow rapid fixation of positive traits in the population. When an advantageous mutation is inherited from the father, all offspring would benefit from efficient placental transfer in utero and good maternal care after birth, which is primed by expression in the placenta. If the advantageous mutation is inherited from the mother, it will be silenced in the offspring but they would still benefit from good maternal care and milk let down via Peg3 action in the maternal hypothalamus (Keverne and Curley 2008).

A recent study questions this theory because a different deletion of Peg3 exon 9 (removing 90% of the coding sequence including all of the zinc fingers) found no maternal behaviour, lactation or suckling deficiencies (Denizot et al. 2016). This suggests that the previously reported phenotypes are due either specifically to the loss of the 5' portion of the gene or to technical differences between experiments. The original Peg3 deletion left a beta-galactosidase-Neo (β-geo) cassette in exon 5 that could affect the expression of other genes in the region, but a more recent knockout that conditionally removed exon 6 in the mammary gland leaving only flippase recognition targets and Lox sites in place also showed that on paternal transmission, mothers had problems releasing milk (Li et al. 1999; Treuy and Kim 2015). When Peg3 is truncated at exon 5 there is evidence that it maternally expressed downstream neighbour zinc finger, imprinted 1 (Zim1) is upregulated, but this is believed to be a trans process because expression of Zim1 is still predominantly from the maternal allele (Ye et al. 2014). Interestingly, no changes were seen in Zim1 expression in the Peg3 exon 9 deletion (Denizot et al. 2016). Genetic background can also influence the phenotype of mutations: the exon 9 deletion is on a C57BL/6J background, whereas the original Peg3 deletion was generated in 129Sv mice (Li et al. 1999; Denizot et al. 2016). However, the model of Li et al. (1999) was later back-crossed onto a C57BL/6J background and similar behavioural phenotypes were observed (Champagne et al. 2009). Furthermore, the mammatory-specific mutation was also on a C57BL/6J background (Frey and Kim 2015). Together, these observations indicate that the phenotypic differences in the models are most likely due the positioning of the deletion or different methods used to assess the phenotypes. Further work is necessary to confirm whether PEG3 does indeed influence maternal behaviours and lactation. This is especially relevant because the Keverne and Curley (2008) extension of the coadaptation theory has been developed around these functions.

Marsupials are altricial and rely more heavily on lactation for maternal provision of nutrients than eutherians. For example, the tammar wallaby has a 26.5-day gestation period followed by up to 350 days lactation (13.2-fold longer than gestation), whereas mice have a 20-day gestation followed by up to 24 days lactation (1.2-fold longer than gestation). By weaning, the average litter mass in marsupials is 55%, compared with 59% in eutherians, indicating that maternal investment is similar in both reproductive strategies (Haysn et al. 1985). The mammary gland not only provides nutrition to offspring, but it also provides a biochemical signalling route between the mother and her young, with milk containing many signalling molecules including insulin and IGF2 (Malven et al. 1987; Prosser 1996). This is the ancestral mechanism of signalling between mother and child, and functions that are performed by the placenta in eutherians are thought to be performed via milk in the marsupial (Power and Schulkin 2013). In agreement with this are recent data from RNA-seq that indicate that the marsupial mammary gland shares many transcripts with the eutherian placenta, including genes involved in nutrient transport and IGF-binding protein 1 (Igfbp1), which is important in IGF regulation (Guernsey et al. 2017). This suggests that placentation and lactation are performing similar functions in resource acquisition in eutherians and marsupials respectively. If this is the case, then the main site of imprinting in the marsupial would be the mammary gland (Stringer et al. 2014) and eutherians may have less dependence on imprinting in the mammary gland. Imprinting analysis in marsupial mammary glands has been performed for three genes. GRB10 was biallelically expressed in adult tammar wallaby mammary glands (Stringer et al. 2012a). This is not unexpected because GRB10 in marsupials lacks the paternal-specific promoter required for central nervous system expression in eutherians, indicating that this is an eutherian-specific imprinting gene (Garfield et al. 2011; Stringer et al. 2012a). Both IGF2 and INS are monoallelically expressed in tammar wallaby mammary gland (Stringer et al. 2012c). In eutherians, Ins2/INS imprinting has only been reported in the yolk sac (Deltour et al. 1995; Moore et al. 2001); its more sustained imprinting in tammar wallaby indicates that marsupials may have a different repertoire of imprinted genes to...
eutherians, and that by simply testing known imprinted genes in these species we may be missing key examples.

The intimate relationship between mother and fetus or neonate in mammals provides the young with all their nutritional needs. Failure to establish and maintain this relationship during pregnancy or postnatally can have implications to offspring that last for the rest of their life. Imprinted gene expression is enriched at all stages of this relationship, and the role of imprinted genes in placentation is well established. Neither the coadaptation nor conflict hypotheses predict that the mammary gland would be a site for imprinted expression; however, there is evidence of imprinted expression in this tissue, although few genes have been studied to date. This suggests that either a different evolutionary pressure is present here or that imprinting in the mammary gland simply reflects the gene’s status from earlier in development. Because monotremes also have a long lactation period, it may be that these most distantly related mammals may also exhibit imprinting in mammary glands. Clearly, a systematic analysis of this organ in monotremes, marsupials and eutherians is necessary to fully understand the roles of imprinted genes in postnatal resource provisioning in mammals and the evolution of this remarkable process.

**Evolution of the imprinting mechanisms: lessons from marsupial imprinting**

In order for the transcriptional machinery of a cell to be able to distinguish the maternally and paternally inherited copies of imprinted genes, it is necessary for the chromosomes to be marked in some way. DNA methylation is the primary ‘imprint’ in eutherians. All known imprints are germline (g) DMRs that occur over an ICR. ICRs are genomic elements that control the imprinted expression of a singleton imprinted gene or of all imprinted genes within a coordinately regulated cluster. The majority of ICRs (23 are confirmed) are maternally methylated gene promoters that gain methylation during oogenesis but remain unmethylated in spermatozoa. Only three paternally methylated ICRs have been identified; these are protein-specific guanine nucleotide-releasing factor (RAS protein-specific guanine nucleotide-releasing factor) domains and become methylated during spermatogenesis but remain unmethylated in ova.

There are only limited data indicating that the imprinting mechanism is conserved between eutherians and marsupials. The H19 differentially methylated domain (H19 DMD) in the tammar wallaby is differentially methylated in poush young samples and hypermethylated in adult testis (Smits et al. 2008). The tammar H19 DMD also contains CCCTC-binding factor (CTCF) binding sites and has insulator activity similar to those seen at the mouse ICR, indicating the region is likely to be functionally conserved between marsupials and eutherians (Smits et al. 2008). The PEG10 promoter is a maternally methylated DMR in the tammar wallaby, as in the mouse (Suzuki et al. 2007). In the tammar wallaby, PEG10 is a singleton imprinted gene, whereas in eutherians the DMR controls the imprinted expression of neighbouring genes as well (Ono et al. 2003; Suzuki et al. 2007). In the mouse, the Igf2r imprinting domain is controlled by a maternally methylated ICR that is the promoter for the antisense transcript Airn located in intron 2 of Igf2r (Wutz et al. 1997). No DMR has been identified in the corresponding intron in marsupials although a maternally methylated DMR has been reported in intron 11 (Killian et al. 2000; Weidman et al. 2006a; Das et al. 2012). Interestingly, no orthologue for Airn has been reported in marsupials, suggesting a different mechanism of imprinting control in this region. With regard to other loci, incomplete imprinting of one MEST transcript has been reported in the tammar wallaby, but no differential methylation has been observed at the promoter for this transcript, which is where the ICR is found in eutherians (Suzuki et al. 2005).

Together, these data indicate that a different imprinting control repertoire may be used by marsupials to mark the paternal and maternal copies of IGF2R and MEST. Recently, researchers have found that maternal histone 3 lysine 27 trimethylation (H3K27me3) is associated with repression of some maternal alleles of paternally expressed genes in the preimplantation embryo (Inoue et al. 2017). Imprinting in these genes was lost in the embryo by the epiblast stage, but four genes appeared to retain paternal-specific expression in the placenta. It is therefore possible that histone modifications rather than the DNA methylation may represent a more ancestral form of imprinting control. Although the establishment of imprinting by H3K27me3 had never been reported before, a role for H3K27me3 in maintaining imprinted expression in response to the germline-derived DNA methylation imprint has been reported previously (Lewis et al. 2004; Umlauf et al. 2004; Yamasaki-Ishizaki et al. 2007). Therefore, further analyses of the epigenetic profiles at marsupial IGF2R and MEST are necessary to explore possible mechanisms that may imprint these genes.

It was initially thought that gDMRs were established specifically at ICRs, but subsequent whole-methylose sequencing studies have indicated that there are many more gDMRs than there are imprints (Kobayashi et al. 2012). Therefore, what sets ICRs apart from these other gDMRs is their resistance to reprogramming after fertilisation. There are two waves of epigenetic reprogramming in eutherian development. The first takes place in the primordial germ cells (PGCs) as they migrate towards the genital ridge early in embryogenesis. In mice, this wave of demethylation is completed by E13.5 and all imprints are erased (Hajkova et al. 2002). De novo methylation then occurs in both germ lines. In the tammar wallaby, PGCs complete migration to the genital ridge just before birth and continue to proliferate until 25 days after birth (Alcorn and Robinson 1983; Renfree et al. 1996; Ullmann et al. 1997). Analysis of the only two DMR imprints identified in marsupials found that the relative timing of reprogramming was conserved between eutherians and marsupials. The PEG10 DMR and H19 DMD became fully demethylated by Days 7 and 14 postpartum respectively, and de novo methylation at H19 DMD in the male germline started at Day 34, demonstrating that the basic mechanisms of the first wave of reprogramming are conserved within therians (Suzuki et al. 2013).

The second round of global demethylation occurs after fertilisation in the preimplantation embryo. Of considerable importance, the only gDMRs that are maintained during this
wave of reprogramming are imprints that are protected by proteins that target the methylated copy. Postfertilisation reprogramming has not been studied in non-eutherian mammals, so it is not known whether the PEG10 DMR and H19 DMD are also protected from this genome-wide wave of demethylation.

The retention of imprints during this second wave of global demethylation is the key step in imprinting control because it preserves the epigenetic memory of parental origin. Li et al. (2008) demonstrated that zinc finger protein 57 (ZFP57) was necessary for the maintenance of several imprints during post-fertilisation reprogramming. ZFP57 is a Kruppel associated box (KRAB) zinc finger protein (KZFP) that is highly expressed in the oocyte and, unlike most of the other approximately 280 KZFPs studied to date in the eutherian genome (Imbeault et al., 2017), it binds to methylated DNA. Chromatin immunoprecipitation (ChIP) analysis of C57BL/6J × Castaneous mouse hybrid embryonic stem cells demonstrated that ZFP57 binds to the methylated copy of all identified DMRs (Stroganov et al., 2015). Zfp57 is a maternal–zygotic effect gene because mutant mice lacking both the oocyte and zygotically expressed copies of the gene (MZ−/−) die prenatally (Li et al., 2008; Takahashi et al., 2015). Mice lacking a zygotic copy of the gene have a much milder phenotype, indicating that maternal oocyte-derived ZFP57 may have an important role in protecting imprints in the early preimplantation embryo. This was confirmed by a comprehensive analysis of ICR methylation in E12.5 tissues from MZ−/− mice. Of the nine ICRs studied, seven showed a significant reduction in methylation in MZ−/− mutants (Table 2). However, the degree of methylation loss varied; for example, the Snrpn DMR showed complete loss of methylation, whereas the Igf2r DMR showed only a 10–20% reduction (Takahashi et al., 2015). Two ICRs showed no loss of methylation in MZ−/− mice: (1) KvDMR, which controls the Kcnq1ot1 imprinted domain; and (2) the H19 DMR (Table 2). These analyses have since been extended, identifying two additional ICRs that remain considerably protected from demethylation in MZ−/− mouse mutants (Peg10 and Mest; N. Takahashi and A. C. Ferguson-Smith, unpubl. data). Therefore, the data indicate that murine ICRs can be divided into three classes based on their ability to retain imprints in the absence of ZFP57: those dependent on ZFP57, those partially dependent on ZFP57 and those not dependent on ZFP57 (Table 2). These data suggest that there is at least one other protein that is required to protect imprints in the preimplantation embryo. Importantly, none of the four domains that are imprinted in marsupials is completely protected by Zfp57 in eutherians, suggesting that such additional factors may be critical for the maintenance of marsupial imprinting and are likely to be more ancient than Zfp57, which is only found in eutherians (Imbeault et al., 2017).

It is possible that the more ancestral imprint protection factor belongs to the same gene family as Zfp57. The KRAB zinc finger family is one of the largest protein families in the human genome, containing over 350 members (Huntley et al., 2006; Imbeault et al., 2017). This is a rapidly evolving gene family that first arose in the Sarcopterygii lineage (tetrapods, coelacanths and lungfish; Imbeault et al., 2017). Most KZFPs have been shown to suppress transposable elements by recruiting KRAB-associated protein-1/tripartite motif containing 28 (KAP1/TRIM28) and establishing repressive chromatin marks such as H3K9me3 and DNA methylation (Wolf and Goff, 2009; Quenneville et al., 2012; Jacobs et al., 2014; Imbeault et al., 2017). It has been proposed that KZFPs play a major role in host defence and that the rapid evolution of KZFPs is the result of an ‘arms race’ between transposable elements and the host genome (Jacobs et al., 2014). In agreement with this idea is the correlation between the number of long terminal repeat (LTR) retrotransposons and the number of KRAB zinc finger genes in mammalian genomes. For example, in the platypus genome, only 43 KZFPs have been identified and its genome only consists of 0.2% LTR retrotransposons. In the therian genomes the number of LTR retrotransposons and KZFP genes is much higher; 851 KZFPs have been identified in the marsupial opossum and 10% of its genome consists of LTR retrotransposons (Imbeault et al., 2017).

It is also possible that the rapid expansion of KZFPs at the same time as the evolution of imprinting facilitated the co-option of KZFPs for imprint protection in addition to a role in host defence. Indeed, it has long been proposed that the imprint mechanism evolved from a host defence strategy. Soon after DNA methylation was first identified as the imprint, it was
suggested that the process may be an extension of its role in silencing foreign DNA (Barlow 1993). Multiple transgenes have been shown to become imprinted in a manner independent of their preintegration site (Reik et al. 1987; Sapienza et al. 1987; Swain et al. 1987; Chaillet et al. 1991; Sasaki et al. 1991). Furthermore, it was subsequently shown that several imprinted genes have themselves arisen from transposition events, including Peg10 and Rtl1/Peg11, which are both neogenes derived from Sushi-ichi retrotransposons (Ono et al. 2001; Seitz et al. 2003; Youngson et al. 2005) and Nap1L5 (nucleosome assembly protein 1-like 5), Zrsr1 (zinc finger, RNA binding motif and serine/arginine rich 1), Mcts2 (malignant T cell amplified sequence 2) and Inpp5f_v2 (inositol polyphosphate-5-phosphatase F variant 2), which are all retrocopies of X-linked genes (Wood et al. 2007). These theories of imprinting arising from a host defence mechanism are based on the fact that foreign DNA such as retrotransposons and retrocopies of genes are targeted for methylation to prevent erroneous expression and mobilisation.

Most KZFPs that have been characterised to date are involved in the establishment of repressive chromatin states and, in particular, recruiting H3K9me3. Although DNA methylation is the germline imprint, the relationship between DNA methylation and H3K9me3 is not fully defined temporally at imprints and, indeed, after fertilisation all ICRs are both DNA methylated and bound by H3K9me3. A more in-depth study is required to identify conserved KZFPs that could bind and protect gDMRs at eutherian and marsupial imprinted domains.

Concluding remarks

By comparing and contrasting when, why and how imprinting arose in marsupials and eutherians, we can get an idea of the evolutionary processes that drove its acquisition in the therian ancestor. Both imprinting and viviparity first arose after the divergence of therians from the monotremes around 160 million years ago (O’Leary et al. 2013). In marsupials, six imprinted genes have been identified that represent the most ancestral imprinted genes. All these genes are expressed in the YSM, indicating that placentation and imprinting are closely linked. Moreover, PEG10 is a neogene derived from a retrotransposon that is only found in marsupials and eutherians. Its critical role in eutherian placenta and expression in the marsupial placenta suggest the emergence of this gene may be an important event in the evolution of placenta, perhaps driving the move towards viviparity (Ono et al. 2001; Suzuki et al. 2007).

Marsupial paternally expressed genes are growth enhancing, whereas maternally expressed genes are growth limiting, suggesting conflict/kinship and supply and demand were the major drivers of imprinting evolution at this stage in mammalian evolution. However, marsupials rely more heavily on lactation than eutherians to support the development of their young. It is of note that the INS gene is imprinted in the marsupial mammary gland, whereas in eutherians it is only imprint ed in the yolk sac, supporting the idea that the mammary gland may be a particularly important site for imprinting in marsupials.

In eutherians, there is a switch of the major site of maternal resource allocation from the mammary gland to the placenta. That the majority of imprinted domains are evident in all four eutherian superorders along with long-lived, invasive chorionallantoic placentas once again points to the parallel evolution of imprinting with placentation. In this major wave of imprinting acquisition, more genes involved in resource allocation in the placenta became imprinted, including Cdkn1c, Phlda2, Ascl2, Grb10, as did transport genes such as Slexa4 and Slc22a18. However, in addition to placentation, imprinted genes in eutherians have vital roles in maternal adaptation to pregnancy (Dlk1, Peg3), maternal care (Peg3, Magel2), metabolism (Gnas cluster) and behaviour (Srrm cluster). This wide range of functions cannot easily be explained by a single evolutionary theory and suggests that more than one selective pressure may be acting at imprinted loci.

The mechanisms by which imprinted genes are marked and maintained appear to be different between marsupials and eutherians. In eutherians, all verified imprinted genes are marked by differential DNA methylation, which is established in the germline, but only two DMRs have been identified in marsupials, suggesting that a different more ancestral imprint mark may exist. The mechanisms by which methylation is maintained at ICRs appear to differ too; the majority of eutherian-specific ICRs are protected in the preimplantation embryo by Zfp57, whereas the more ancestral gDMRs are not. This also suggests that another more ancestral imprinting protection factor exists in marsupials.

We therefore hypothesise that mammalian clade-specific imprinted genes may exist that have evolved alongside the evolution of clade-specific mechanisms that target and maintain parental origin-specific epigenetic states at such loci and that the evolution of epigenetic pathways designed to control repressive states has contributed to their emergence. Further genome-wide systematic analysis of parental origin-specific gene expression in multiple species alongside detailed characterisation of epigenetic targeting mechanisms has the potential to test this hypothesis.

Materials and methods

Elephant samples

Elephant samples were collected at the Elephant Research Unit in the Lower Save Conservancy in Zimbabwe. Ele5 and Ele9 are placental tissues from two individuals, whereas Ele13 and Ele14 are amnion and umbilical cord from the same individual. DNA was extracted using standard phenol–chloroform protocols.

Bisulfite pyrosequencing

A 1-µg sample of genomic DNA was bisulfite converted using the Imprint DNA Modification Kit (Sigma Aldrich) according to the manufacturer’s instructions. Purified samples were amplified by polymerase chain reaction (PCR) performed in a final reaction volume of 10 mL containing 250 nM forward and reverse primers (forward, 5'-GGAATGAGGAGATGTTG GATGAA-3'; reverse, 5'-[Bnn]CCCCAACTACTCCCAT CCTAAACC-3'), 0.25 U Taq (HotStarTaq DNA Polymerase; Qiagen) and 0.2 mM dNTPs. The PCR conditions were as follows: 95°C for 15 min, followed by 40 cycles of 94°C for 30 s, 48°C for 30 s and 72°C for 30 s, with a final step at 72°C for 10 min. The PCR products were subsequently purified (Qiagen) and cloned into pCR2.1-TOPO vectors (Invitrogen). At least 10 independent clones were sequenced for each primer pair to determine the methylation status of all possible CG dinucleotides.
5 min. Single-strand PCR products were purified (PyroMark Q96 Vacuum Prep Workstation; Qiagen) and pyrosequencing was performed on a PyroMark Q96MD (Qiagen) using PyroMark Gold Q96 Reagents (Qiagen) and the pyrosequencing primer 5′-GGGATGTTGGATGAGAAT-3′ in accordance with the manufacturer’s instructions.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors thank Twink Allen and Fiona Stansfield for supplying the elephant tissues. The work presented in the paper and the writing of the paper was supported by the Wellcome Trust (Grant no. WT095606).


GRB10 imprinting is eutherian mammal specific. Mol. Biol. Evol. 29,
3711–3719. doi:10.1093/MOLEHR/MS173
Promoter-specific expression and imprint status of marsupial IGF2.
PLoS One 7, e41690. doi:10.1371/JOURNAL.PONE.0041690
Selected imprinting of INS in the marsupial. Epigenetics Chromatin 5,
imprinting: evidence from marsupials. Heredity 113, 145–155. doi:
10.1018/HDY.2014.10
Stroganetz, R., Krueger, F., Yamazawa, K., Shi, H., Gould, P., Goldman-
Takahashi, N., Gray, D., Stroganetz, R., Noon, A., Delahaye, C., Skarnes, W.
gene modulates a major endocrine compartment of the placenta to
Takashashi, N., Gray, D., Stroganetz, R., Noon, A., Delahaye, C., Skarnes, W.
C., Tate, P. H., and Ferguson-Smith, A. C. (2015). ZFP57 and the targeted
Migration of primordial germ cells to the developing gonadal ridges in
the tammar wallaby Macropus eugeni. J. Reprod. Fertil. 110, 135–143.
doi:10.1530/JRF.0.1100135
Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y., and
Feil, R. (2004). Imprinting along the Kcnq1 domain on mouse chromo-
some 7 involves repressive histone methylation and recruitment of
NG1467
of embryonic growth and lysosomal targeting by the
(2006a). Imprinting of opossum Igf2r in the absence of differential
methylation and air. Epigenetics 1, 50–55. doi:10.4161/EPI.1.1.2592
phylogenetic analysis reveals multiple non-imprinted isoforms of opos-
0116-X
Weksberg, R., Shuman, C., and Beckwith, J. B. (2010). Beckwith–Wiede-
106
Wilkin, F., Paquette, J., Ledru, E., Mamelin, C., Pollak, M., and Deal, C. L.
Wolf, J. B., and Brodie, E. D. (1998). The coadaptation of parental and
1998.TB01632.X
Wolf, D., and Goff, S. P. (2009). Embryonic stem cells use ZFP809 to silence
retrorviral DNAs. Nature 458, 1201–1204. doi:10.1038/NATURE07844
1371/JOURNAL.PBIO.0040380
Wood, A. J., Roberts, R. G., Monk, D., Moore, G. E., Schulz, R., and Oakey,
association between X chromosome homology and maternal germ-line
0030020
Wutz, A., Smrzka, O. W., Schweifer, N., Schellander, K., Wagner, E. F., and
Barlow, D. P. (1997). Imprinted expression of the Igf2r gene depends on
Functional polymorphism in the parental imprinting of the human
10.1016/BBRC.1993.2542
Yanamaki-Ishizaki, Y., Kayashima, T., Mapendak, C. K., Soejima, H.,
Ohta, T., Masuzuki, H., Kinoshita, A., Urano, T., Yoshuira, K.-i.,
Matsumoto, N., Ishimaru, T., Mukai, T., Nukawa, N., and Kishino, T.
(2007). Role of DNA methylation and histone H3 lysine 27 methylation
in tissue-specific imprinting of mouse Grb10. Mol. Cell. Biol. 27,
732–742. doi:10.1128/MCB.01329-06
doi:10.1371/JOURNAL.PONE.0108596
Youngson, N. A., Kocialkowski, S., Peel, N., and Ferguson-Smith, A. C.
(2005). A small family of sushi-class retrotransposon-derived genes in
mammals and their relation to genomic imprinting. J. Mol. Evol. 61,
481–490. doi:10.1007/S00239-004-0332-0
Bidirectional action of the Igf2r imprint control element on upstream and
downstream imprinted genes. Genes Dev. 15, 2361–2366. doi:10.1101/
GAD.206201

www.publish.csiro.au/journals/rfd