doi:10.1093/jmcb/mjaa020
Published online May 5, 2020

Article

Landscape of genomic imprinting and its functions in the mouse mammary gland

Haibo Xu[1,2,†], Lina Zhao[1,2,†], Xu Feng[1], Yujie Ma[1], Wei Chen[1], Li Zou[1], Qin Yang[1], Jihong Sun[3], Hong Yu[4,*], and Baowei Jiao[1,5,6,*]

[1] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[2] Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming 650223, China
[3] Department of Radiology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, China
[4] Department of General Surgery, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, China
[5] KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[6] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China
[†] These authors contributed equally to this work.
[*] Correspondence to: Baowei Jiao, E-mail: jiaobaowei@mail.kiz.ac.cn; Hong Yu, E-mail: blueyu000@zju.edu.cn

Edited by Luonan Chen

Genomic imprinting is an epigenetic modification of DNA, whereby gene expression is restricted to either maternally or paternally inherited alleles. Imprinted genes (IGs) in the placenta and embryo are essential for growth regulation and nutrient supply. However, despite being an important nutrition delivery organ, studies on mammary gland genomic imprinting remain limited. In this study, we found that both the number of IGs and their expression levels decreased during development of the mouse mammary gland. IG expression was lineage-specific and related to mammary gland development and lactation. Meta-analysis of single-cell RNA sequencing data revealed that mammary gland IGs were co-expressed in a network that regulated cell stemness and differentiation, which was confirmed by our functional studies. Accordingly, our data indicated that IGs were essential for the self-renewal of mammary gland stem cells and IG decline was correlated with mammary gland maturity. Taken together, our findings revealed the importance of IGs in a poorly studied nutrition-related organ, i.e. the mammary gland, thus providing a reference for further studies on genomic imprinting.

Keywords: imprinted gene, imprinted gene network, mammary gland, stem cell, differentiation, metabolism

Introduction

Genomic imprinting is an epigenetic event that results in biased parental-specific allelic expression in therian mammals (Stringer et al., 2014). Imprinted genes (IGs) appear to be widespread in eutherian mammals and marsupials, though remain unreported in monotremes (Renfree et al., 2013; Stringer et al., 2014). To date, nearly 150 IGs have been identified in mice (Plasschaert and Bartolomei, 2014). There are several evolutionary explanations for the existence of genomic imprinting, including the kinship theory, the sexual antagonism theory, and maternal-offspring coadaptation theory (Brandvain et al., 2011; Patten et al., 2014). The core of parental conflict theory is related to offspring development and nutrition supply. Hence, organs that provide nutrition for the promotion of fetal growth, and thus fetal fitness, are crucial in parental conflict and IG expression (Wilkins, 2014).

As an essential mammalian organ for nutrition delivery and fetal growth, the placenta is a primary target for studying the quantity and function of IGs (Wagschal and Feil, 2006; Frost and Moore, 2010). At present, most identified IGs in mice have been found in the placenta (Tunster et al., 2013). Functional studies show that placental IGs are crucial for embryonic development and metabolism; for example, the paternally expressed *Igf2* gene is associated with nutrient transfer in the placenta and is a positive regulator of embryonic growth (Zwart et al., 2001). Conversely, *H19*, a maternally expressed gene located in the same cluster as *Igf2*, restricts placental and embryonic growth by reducing *Igf1r* expression (Keniry et al., 2012). In addition, the maternally expressed *Cdkn1c* gene is

Received January 12, 2020. Revised April 12, 2020. Accepted April 20, 2020.

associated with spongiotrophoblast restriction in the placenta and is a negative regulator of embryonic growth (Zhang et al., 1998). These placental studies provide evidence supporting the parental conflict theory.

The functions of IGs are diverse and include nutrient transport (*Slc22a2*, *Slc22a3*, and *Slc38a4*) (Jonker et al., 2003; Constancia et al., 2005), signaling (*Ins2* and *Igf2*) (Neirijnck et al., 2019), transcription (*Zac1*) (Varrault et al., 2006), and cell cycle (*Cdkn1c*) (Zhang et al., 1998). However, the previously identified molecules and associated mechanisms cannot explain the unique single-allele expression pattern of IGs. Earlier studies demonstrated that IGs are often co-expressed as an imprinted gene network (IGN) and undertake functions as a whole, including the regulation of growth and metabolic processes (Varrault et al., 2006; Gabory et al., 2009; Sandhu, 2010; Al Adhami et al., 2015; Keverne, 2015). Furthermore, several IGs have been found to modify the expression of other IGs contained within the same IGN. For example, *Zac1* regulates the expression of other co-expressed IGs in the same IGN, including *Igf2*, *H19*, *Cdkn1c*, and *Dlk1*, thus resulting in intrauterine growth restriction (Varrault et al., 2006). Furthermore, *H19* acts as a regulator within the IGN that controls growth in mice (Gabory et al., 2009). IGs can also regulate networks formed by other cell signaling components (Sandhu, 2010; Al Adhami et al., 2015). For instance, IGs are known to participate in extracellular matrix (ECM) remodeling and ECM-linked signaling, leading to the regulation of cell cycle exit and differentiation (Varrault et al., 2006; Al Adhami et al., 2015). It should be noted, however, that IGN components exhibit substantial plasticity and are highly dependent on tissue type and developmental stage. For example, the IGN targeted by *H19* is found in postnatal organs but not in the placenta (Gabory et al., 2009); furthermore, IGNs targeted by *Peg3* in the hypothalamus and placenta differ considerably (Keverne, 2015). As such, the underlying mechanism and formation of IGNs remain elusive. Deciphering IGNs in more tissues and developmental stages should shed light on these questions.

To date, genomic imprinting is well recognized in the placenta, embryo, and several adult tissues (Tunster et al., 2013; Babak et al., 2015; Perez et al., 2015; Andergassen et al., 2017). However, studies on genomic imprinting patterns during development, particularly in postnatal organs, remain limited. Similar to the placenta, the mammary gland is a unique mammalian organ and a primary source of nutrition for all newborns (Stringer et al., 2012). Although it is an obvious target for imprinting under maternal-offspring nutrition delivery, systematic study of imprinting in the mammary gland is poor. Here, we investigated the dynamic imprinting status and expression level of mouse IGs in breast tissue. Our results revealed that the core functions of the IGs were metabolism and development via a mammary gland-specific IGN. Functional assay also indicated that the IGs in the mammary gland were required for self-renewal of mammary gland stem cells (MaSCs) and inhibited the differentiation of mammary gland cells. Our study provides new evidence for the significance of genomic imprinting in postnatal development.

Results

Imprinting status of IGs in mouse mammary gland

To investigate the status of IGs in the mouse mammary gland, we performed genome-wide screening using reciprocal crosses, as is reported previously (Perez et al., 2015). Two mouse strains (i.e. PWK/PhJ and C57BL/6J) were selected due to their divergent single-nucleotide polymorphism (SNP) sites (Babak et al., 2015). We dissected breast tissue from nine F1 hybrid individuals for RNA-sequencing with relatively deep coverage (Supplementary Table S1). Lactation day 15 (LD-15) was chosen for sampling due to the maximum ratio of epithelium in breast tissue at this stage. As a mouse reference genome is only available for the C57BL/6J strain, we aligned the sequenced reads to the parental pseudogenomes as is reported previously (Crowley et al., 2015) to improve alignment quality and minimize bias caused by differences in genetic distance between the parental genomes and reference sequence. Reads that overlapped with parental SNPs with maximum consistency (alleles reported from specific parent) were assigned to the corresponding parent (Huang et al., 2014). After relative log expression normalization of sequence libraries, we calculated the parental expression bias of each gene across each individual as well as parental expression significance using the Integrative Statistics of alleLe Dependent Expression (ISoLDE) method (Reynes et al., 2020). In total, we detected 23 and 26 IGs in LD-15 mice using the 'threshold' and 'default' methods in ISoLDE, respectively (Figure 1A and B). Another four genes (*Impact*, *Kcnq1*, *Gnas*, and *Snhg14*), previously reported as IGs, did not pass low gene expression filtering (filterT step), although they showed significant allele-specific expression. Among the above 30 IGs (26 plus 4), 15 were significantly maternally expressed and 15 were significantly paternally expressed. Most of these IGs showed strong allele-specific expression (parental bias >0.75), except for *H13*/*Srgap1*/*Slc22a1* (maternally expressed) and *Impact* (paternally expressed) (Figure 1A, B, and E). Numerous IGs reported previously in other tissues (Plasschaert and Bartolomei, 2014) were identified as non-imprinted due to bi-allelic expression (20), strain bias expression (18), or undetectable expression (90) in the LD-15 mammary glands using ISoLDE (Supplementary Figures S2 and S3). At the peak lactation stage, a relatively higher percentage of genes demonstrated strain effect bias ($n = 2614$, Supplementary Figure S1F), whereas few genes showed parental expression bias ($n = 32$, Supplementary Figure S1E). Similar results were also observed at the virgin and LD-3 stages (Supplementary Figure S1A–D). This variance may be due to the existing large numbers of SNPs and INDELs in hybrid mice, as strain bias expression is mainly controlled by *cis*-regulatory elements (Crowley et al., 2015; Wang et al., 2019).

To study the status of IGs during mammary gland development, we analyzed IG expression at the virgin, early, and peak lactation stages in the mammary gland based on our and other published datasets (Andergassen et al., 2017). In total, 67 genes were identified as imprinted in at least one

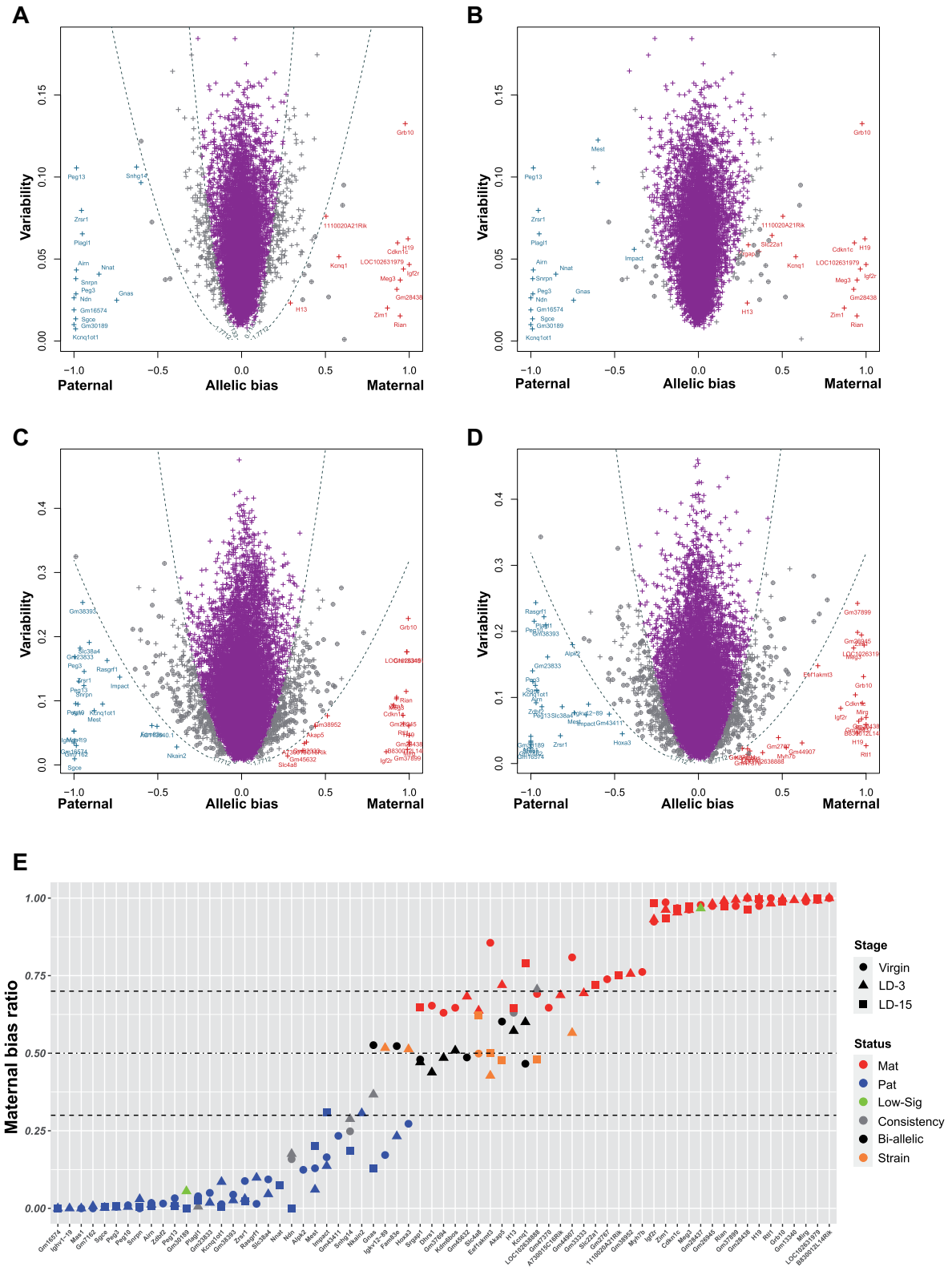


Figure 1 IGs identified by ISoLDE in mouse mammary gland. (**A** and **B**) IGs identified by ISoLDE at LD-15 stage using threshold (**A**) and default methods (**B**). (**C** and **D**) IGs identified by ISoLDE at virgin (**C**) and LD-3 (**D**) stages using threshold method. (**E**) Combination of statistical results of IGs from virgin, LD-3, and LD-15 stages. Statistical significance of parental and strain bias was calculated by ISoLDE. In **A**–**D**, red and blue indicate significant maternal and paternal bias expression, respectively. In **E**, symbols represent different sources/stages of mammary gland tissues and colors indicate status of each gene (red, blue, and orange represent significant maternal, paternal, and strain bias expression; green: significant parental bias expression, but not passing filter module; gray: bias direction agreement between replicates; black: bi-allelic expression).

developmental stage (Figure 1A–E), including 37 IGs not reported previously. In general, the number of detected IGs showed a decreasing tendency during development of the mammary gland, with 48, 42, and 30 IGs at the virgin, early lactation (LD-3), and peak milk secretion stages (LD-15). These results are consistent with trends reported in the placenta, brain, liver, and heart (Lui et al., 2008; Babak et al., 2015; Andergassen et al., 2017). The 67 IGs could be roughly divided into two sub-groups, i.e. stable IGs (parental expression bias >0.7) and unstable IGs (parental expression bias <0.7). Stable IGs tended to maintain their imprinting status or lose expression during development, whereas unstable IGs tended to alter their imprinting status to bi-allelic expression or strain bias expression (Figure 1E; Supplementary Figure S2), which may be responsible for the observed decrease in IG number in our study. Notably, the number of overall detectably expressed genes also decreased. Relative to the virgin stage, the proportions of expressed genes (FPKM $>$ 1) in LD-3 and LD-15 were 94.7% and 72.5%, respectively. In addition, the proportions of expressed IGs in LD-3 and LD-15 were 87.5% and 62.5%, respectively. Detectably expressed IGs showed a relatively higher downtrend.

The status of IGs is generally controlled by the imprinting control region (ICR) within the imprinted cluster, with IGs regulated in such clusters (Edwards and Ferguson-Smith, 2007; Bartolomei and Ferguson-Smith, 2011). Among the 37 newly identified IGs (Supplementary Figure S3), 11 were located in or in close proximity to known imprinted clusters (Supplementary Figure S3), including the *H13–Nnat* cluster (*Myh7b*), *Gnas–Nespas* cluster (*Gm30189*), *Ube3a–Ndn* cluster (*Gm38393*), *Grb10–Cobl* cluster (*LOC102631979*), *Dlk1–Dio3* cluster (*Gm26945* and *Gm37899*), and *Slc22a3–Igf2r* cluster (*Slc22a1*, *Gm16574*, *Gm23833*, *Gm7162*, and *Mas1*), as well as two new IGs (*Srgap1* and *Slc4a8*) located near the ICR.

Cell lineage-specific expression of IGs in mammary gland

After identifying the status of IGs, we explored their expression patterns at single-cell resolution during differentiation and development of the mammary gland using published 10× Genomics data (Bach et al., 2017). Mammary gland epithelium exhibits a well-established hierarchy consisting of MaSCs at the apex and a series of progenitors and mature mammary ducts and alveolar cells (Fu et al., 2020). Based on single-cell RNA sequencing (scRNA-seq) data and the expression distribution of well-recognized lineage markers and dimension reduction clustering, we identified eight cell lineages, similar to that reported originally (Bach et al., 2017), including MaSCs, basal cells, myoepithelial cells (Myo), luminal progenitor cells (LumPro), luminal hormone sensing progenitor cells (LumHorPro), luminal hormone sensing mature cells (LumHorMat), luminal alveolar progenitor cells (LumAlvPro), and luminal alveolar mature cells (LumAlvMat) (Supplementary Figure S4A). The IGs demonstrated dynamic expression in the mammary gland and their expression profiles in the MaSCs were highly distinct from those in other lineages (Figure 2A and module 1 of Figure 2B). Each cell lineage contained specific IGs. Of note, 25 of the 71 mouse IGs, including *Peg3*, *Ndn*, and *Sgce*, were exclusively and highly expressed in MaSCs (Figure 2A), implying crucial roles of IGs for stemness maintenance in MaSCs. The IGs were mainly grouped into four co-expression modules based on Monocle-3 (Cao et al., 2019; Figure 2B), with similar modules generated using WGCNA (Langfelder and Horvath, 2008; Supplementary Figure S7A), further confirming the lineage-specific expression of IGs. Using the scRNA-seq data, our results showed that IGs were expressed in a lower proportion of mammary gland epithelial cells (Supplementary Figure S5A), although most IGs showed high expression in epithelial cells (Supplementary Figure S5B), which may account for the fewer IGs expressed or detected in adult tissues in previous research (Babak et al., 2015). The lineage-specific expression profiles suggested crucial roles of IGs in stemness maintenance, cell differentiation, and lineage determination in mammary glands.

IGNs and their functional implications in mammary glands

IGs are frequently co-regulated with each other and form co-regulated networks with positive connections. Previous studies have indicated that diverse IGs showed common functions predominately via IGNs (Varrault et al., 2006; Lui et al., 2008; Gabory et al., 2009; Al Adhami et al., 2015). Our analysis of scRNA-seq data showed that the IGs were co-expressed in two super modules by Monocle-3 and WGCNA (Figure 2B; Supplementary Figure S7), implying that separate co-expression networks may exist,which differs from that found in other tissues (Varrault et al., 2006; Lui et al., 2008; Gabory et al., 2009; Al Adhami et al., 2015). To determine whether IGs were more co-regulated/expressed than random gene pairs, we first calculated the distribution of Pearson correlation coefficients (PCCs) between gene pairs in several gene sets (i.e. IG set, same gene ontology (GO) term gene set, random gene set). Results indicated that IG pairs showed higher correlation than random gene pairs, and comparison between the gene ontology biological process (GOBP) gene sets (pairs from the same GOBP) and random gene sets was shown as a positive control (Figure 2C). Moreover, the correlations between IG pairs within a co-regulated gene module were higher than those IGs pairs within the whole IG sets, indicating that IGs in the same module had higher connectivity. Remarkably, the correlations between gene pairs within super module 1 (S1: module 1 + module 4) were greater than those within super module 2 (S2: module 2 + module 3) (Figure 2C). We then demonstrated the connectivity between IGs by Cytoscape. Clearly, the IGN was formed by two main sub-networks, which were negatively correlated with each other (Figure 2D; Supplementary Figure S7B and C). Furthermore, these sub-networks were divided by lineage-specific expression pattern, but not by imprinting status; for example, IGs highly expressed in MaSCs (e.g. *Peg3*, *Dcn*, *Ndn*, *Gnas*, and *Grb10*) were more likely to group into sub-

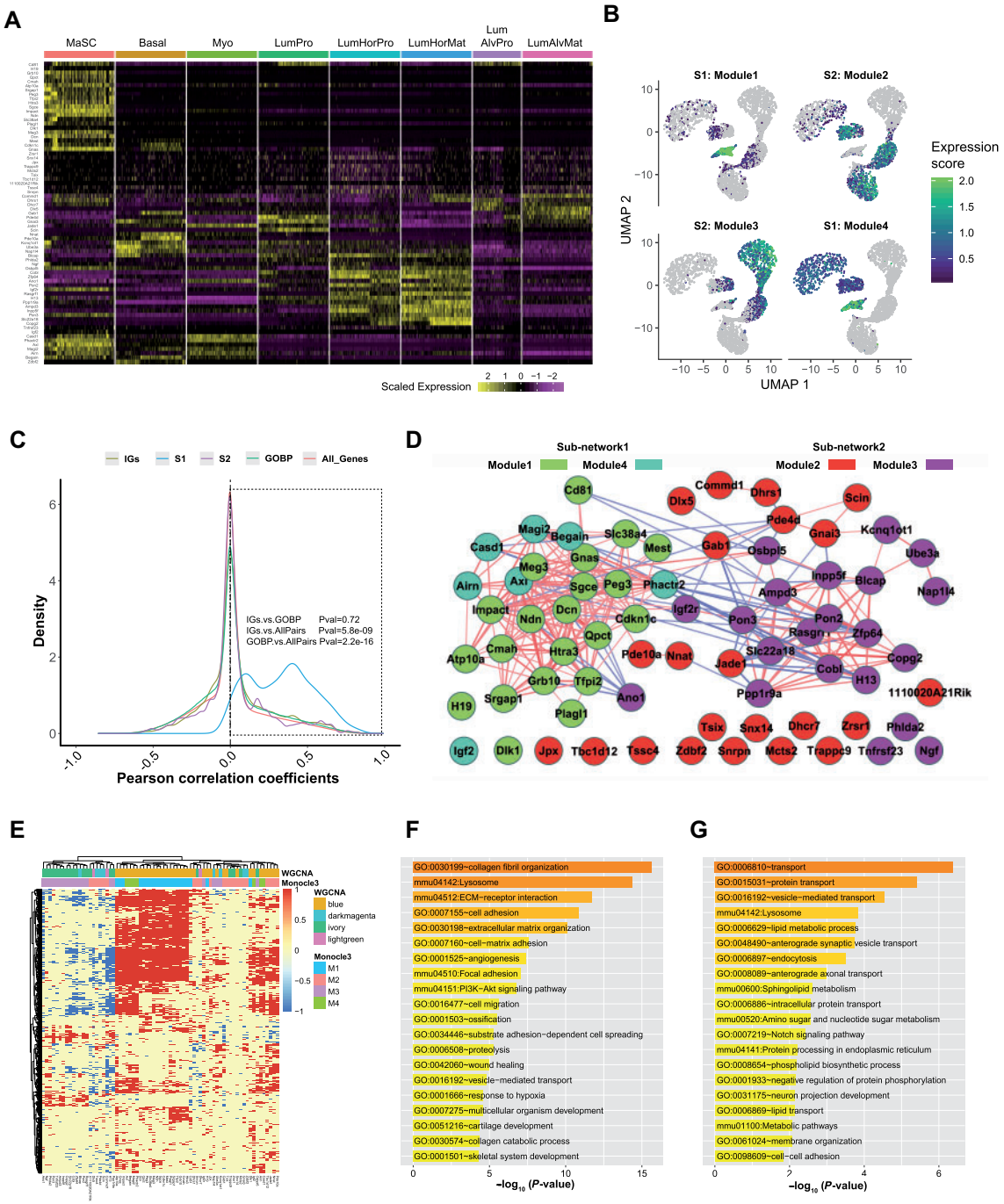


Figure 2 Expression profiles and co-expression network of IGs. Expression profiles were based on scRNA-seq data released from Walid T. Khaled's lab (Bach et al., 2017). Mammary gland IGs identified in our study and Priori mice IGs were combined for figure demonstration. (**A**) Dynamic expression patterns of IGs in each cell type in mammary gland. Genes were grouped based on expression modules calculated using Monocle-3. (**B**) Heatmap showing co-regulated gene module of IGs across different cell lineages, plotted using plot_cells function in Monocle-3. (**C**) Correlation distribution among IG, GOBP, and randomly selected gene sets, respectively. Based on expression patterns in **B**, IGs were spilt into two super modules (S1: module 1 + module 4, S2: module 2 + module 3). Differences in positive PCCs between each gene set were determined by one-tailed Wilcox test. (**D**) Co-expression network of IGs based on value of PCCs ($|PCCs| > 0.4$ and $P < 0.01$). Node color representing co-expression module and edge color being proportional to correlation (red: gene pairs with positive PCC values; blue: gene pairs with negative PCC values). (**E**) Heatmap showing functional relevance of IGs based on GSEA using IG co-expressed genes. Color in red and blue indicated that IG was positively and negatively correlated with relevant GO term, respectively. Yellow represents no correlation. (**F** and **G**) GO enrichment analysis of bi-allelically expressed genes (those genes connected with >5% of total expressed IGs) co-expressed within sub-network 1 (**F**) and sub-network 2 (**G**). Top 20 significant GOBP and KEGG terms enriched in each sub-network are shown.

network 1 and IGs in the differentiated populations were more likely to group into sub-network 2, further implying the crucial functions of IGs in MaSCs. Taken together, these results suggest that the IGN in the mouse mammary glands was composed of two sub-networks.

To interpret the common biological functions of IGs, we performed gene set enrichment analysis (GSEA) based on the expression correlation between IGs and non-IGs, as per previous study (Iyer et al., 2015; Green et al., 2017; Wang et al., 2017b). We estimated PCCs from expression levels between each IG and all other genes and conducted GSEA with ranked PCCs. An association matrix was built between those IGs and significant GO gene sets ($P < 0.05$). The GO gene sets could be roughly clustered into groups based on functional similarity, including transcription/translation, signaling, cell cycle, membrane, adhesion, development, cell fate, chromatin modification, immunity, metabolism, and nutrition storage (Supplementary Table S2). Of note, the IGs in the same co-regulated modules were enriched in similar biological functions (Figure 2E), indicating that IGs in the same co-regulated module may function similarly. The highly expressed IGs in MaSCs (super module 1) were predominantly enriched in functions related to development and cell signaling, whereas those highly expressed in differentiated cells (super module 2) were especially enriched in metabolism and nutrition storage.

The existence of IGNs has been proven in several studies and are considered to function as a whole by targeting a pre-existing machinery to perform regulatory functions and stabilize organ systems (Lui et al., 2008; Gabory et al., 2009; Al Adhami et al., 2015). To investigate the enriched biological functions of each sub-network, we treated the sub-network as an integral and used co-expressed non-IGs in each sub-network to perform DAVID enrichment analysis (i.e. GO terms and KEGG pathways). The first sub-network was enriched in genes associated with ECM, similar to earlier results (Al Adhami et al., 2015), and associated with development, cell proliferation, cell differentiation, and the Wnt and PI3K–AKT signaling pathways (Figure 2F; Supplementary Table S3). The second sub-network was significantly enriched in genes involved in transport, metabolic processes, and mammary epithelial cell differentiation associated pathways (e.g. Notch, Hippo, Prolactin, and Erbb signaling pathways) (Figure 2G; Supplementary Table S3). This method generated similar biological functions as that observed from GSEA for each IG.

Overall, our results showed that IGs in the mammary gland constituted an IGN composed of two less-connected sub-networks. Genes in different sub-networks targeted diverse regulatory genes and functioned differently.

Effects of IG expression changes on self-renewal and differentiation of mammary gland cells

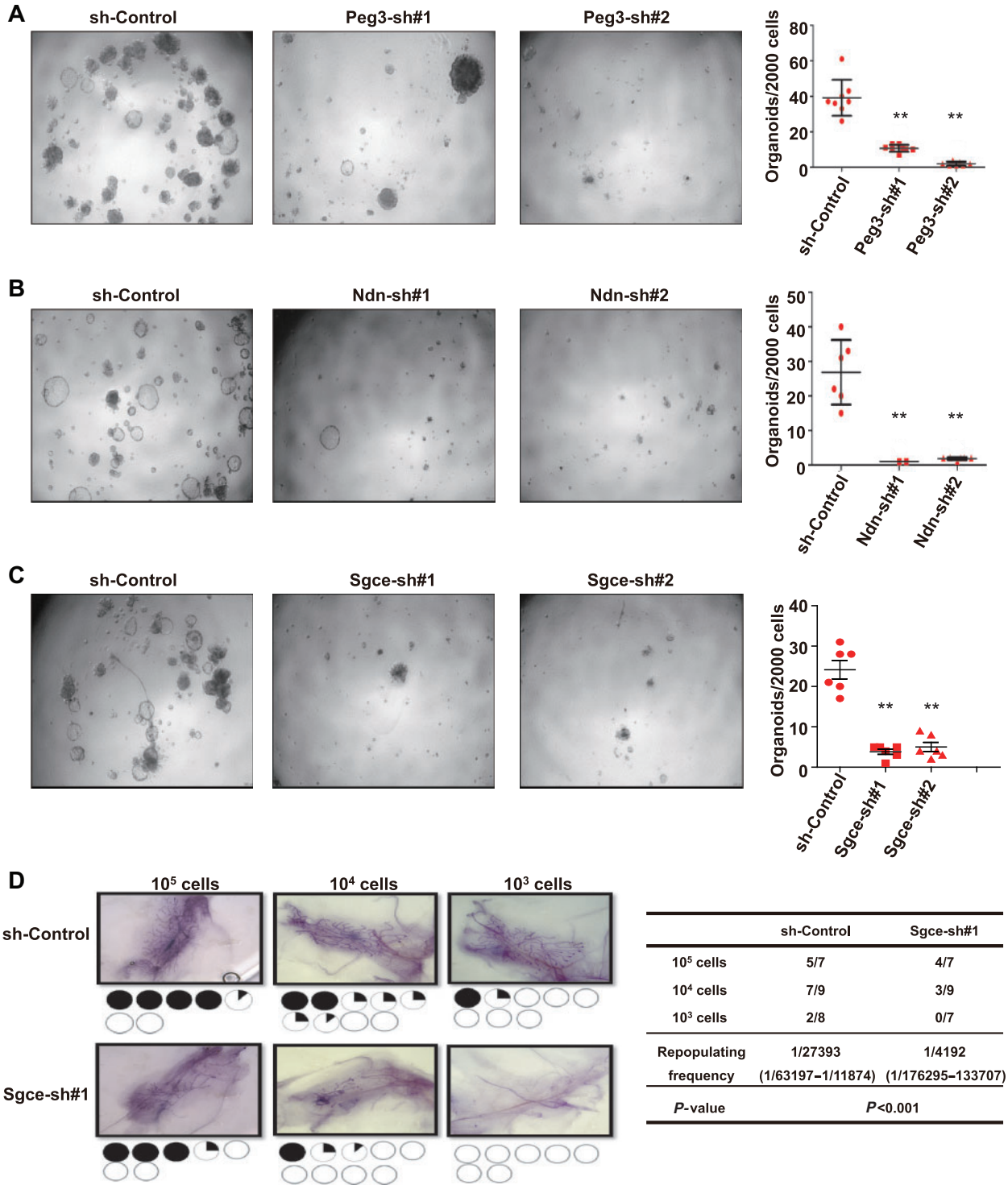
Based on the above evidence related to stem cells (Figure 2), we examined those IGs associated with the regulation of cell stemness in mouse mammary gland primary cultured epithelial cells. Using a well-recognized mammosphere forming assay for self-renewal *in vitro*, our results indicated that knockdown of *Peg3*, *Sgce*, and *Ndn* significantly reduced the self-renewal ability of MaSCs (Figure 3A–C), consistent with the higher expression levels of those IGs in the MaSC lineage (Supplementary Figure S4B–D). Moreover, we performed an *in vivo* assay to demonstrate the functional roles of IGs in MaSCs. When expression levels of *Sgce* were reduced in the mammary gland epithelial cells, MaSC activity was significantly inhibited, as reflected by decreased reconstitution and ductal outgrowth after transplantation into cleared fat pads (Figure 3D). These results were consistent with our PCC-based GSEA analyses (Figure 2E; Supplementary Table S2), in which the expression levels of *Peg3*, *Ndn*, and *Sgce* were positively associated with the Wnt signaling pathway (*Peg3*: $P = 0.002$, NES $= 1.82$; *Ndn*: $P = 0.03$, NES $= 1.57$; *Sgce*: $P = 0.05$, NES $= 1.52$), PI3K–AKT (*Peg3*: $P = 0.003$, NES $= 1.48$; *Ndn*: $P = 0.0002$, NES $= 1.70$; *Sgce*: $P = 0.003$, NES $= 1.46$), and negatively associated with epithelial cell maturation (*Peg3*: $P = 0.03$, NES $= -1.65$; *Ndn*: $P = 0.056$, NES $= -1.54$; *Sgce*: $P = 0.02$, NES $= -1.71$). Taken together, these results confirmed the essential roles of IGs in MaSCs.

To study the function of IGs in lactation differentiation of mammary epithelial cells, we selected two IGs (*Sgce* and *Ube3a*) and detected their expression by immunofluorescence assay. Results showed that their expression levels were markedly lower during lactation (Figure 4A), as further confirmed by quantitative reverse-transcription polymerase chain reaction (qRT-PCR) results (Supplementary Figure S6A). We then used the HC11 cell line, a prolactin-responsive cell line and suitable model, to study mammary epithelial cell differentiation *in vitro* (Perotti et al., 2009). Results indicated that knockdown of *Sgce* and *Ube3a* promoted the formation of dome-like structures (Figure 4B and C), suggesting that low IG expression promoted differentiation of the HC11 mammary epithelial-like cells. Moreover, the expression levels of lactation-related genes were up-regulated in the *Sgce* knockdown HC11 cell line (Figure 4D). We also found that the HC11 cells enhanced the expression of HC11 differentiation markers after *Sgce* knockdown before DIP medium treatment for RNA-seq (Figure 4E), further confirming the above cell differentiation results.

In short, our results demonstrated that interference with IG expression impacted self-renewal and induced differentiation of mouse mammary gland epithelial cells.

Discussion

Previous studies have shown that IGs form co-regulated networks in the placenta, hypothalamus, and various cell lines (Varrault et al., 2006; Lui et al., 2008; Gabory et al., 2009; Al Adhami et al., 2015; Keverne, 2015). Our constructed IGN in the mammary gland further confirmed the existence of such networks in postnatal organs. In the mammary gland epithelial cells, this IGN showed some differences as it primarily consisted of two negatively correlated sub-networks. We also observed that sub-network 1 had similar components as the *H19-*



	sh-Control	Sgce-sh#1
10^5 cells	5/7	4/7
10^4 cells	7/9	3/9
10^3 cells	2/8	0/7
Repopulating frequency	1/27393 (1/63197–1/11874)	1/4192 (1/176295–133707)
P-value	*P*<0.001	

Figure 3 Involvement of IGs in MaSCs. (**A**–**C**)The number of colonies formed in mammosphere assay when *Peg3*, *Ndn*, and *Sgce* were knocked down in primary mammary cells ($n = 8$ replicates for **A**; $n = 6$ replicates for **B** and **C**; data represent mean ± SD). (**D**) Results of transplantation assay. In pie graph below representative images, each circle represents one mammary gland, with black area representing degree of gland filling with outgrowth. Statistical significance was calculated by Student's *t*-test compared to sh-Control group. $^{**}P < 0.01$.

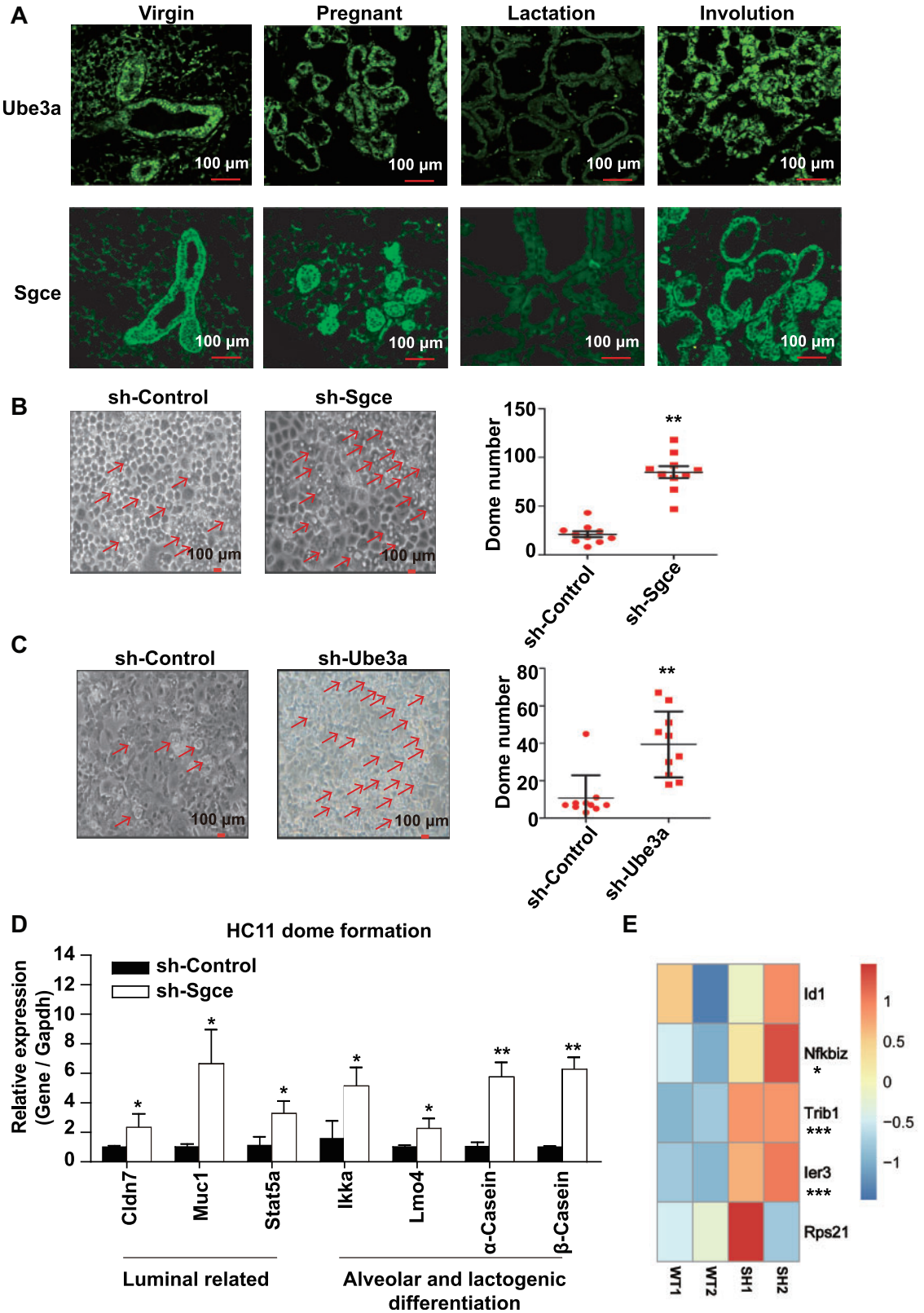


Figure 4 Involvement of IGs in differentiation of mammary cells. (**A**) Immunofluorescence of *Ube3a* and *Sgce* in different mammary gland development stages. (**B** and **C**) Dome formation assay in *Ube3a* and *Sgce* knockdown HC11 cells ($n = 10$ replicates; data represent mean ± SD). (**D**) Expression of milk-related genes in *Sgce* knockdown HC11 cells in dome formation assay. (**E**) Heatmap showing expression alteration of HC11 cell differentiation markers, adopted from previous study (Perotti et al., 2009), after *Sgce* knockdown before DIP medium treatment. In **B**–**D**, statistical significance was calculated by Student's *t*-test compared to sh-Control group. In **E**, statistical significance was calculated using edgeR with glmQLFTest. *****$P < 0.05$; ******$P < 0.01$; *******$P < 0.001$.

or *Zac1*-regulated IGNs reported previously (Varrault et al., 2006; Gabory et al., 2009). As *Zac1* (also named *Plagl1*) and *H19* were both highly co-expressed in sub-network 1 (Figure 2A and D; Supplementary Figure S7C), these two genes may be core regulators for the co-regulation of sub-network 1. Previously, Dr Zhu's lab (Liu et al., 2017) predicted that *H19* can bind to the genomic region of several IGs to form a RNA–DNA–DNA triplex structure. We confirmed the existence of this triplex structure at the genomic region of *Kcnq1ot1* and *Cdkn1c* by ChIRP assay (Supplementary Figure S7D and E), thus providing evidence of how partial IG connected with each other. The segmentation of sub-networks with negative correlation differs from that reported by Al Adhami et al. (2015), who did not observe the existence of sub-networks and found all IGN components to be positively connected. This suggests that different regulatory mechanisms or different biological functions of sub-networks may be present in mammary gland epithelial cells.

Based on the co-expression connectivity between IGs and non-IGs, as well as the GSEA results, our data revealed that the IGs were connected to non-IGs in both sub-networks, which were related to different biological functions. For example, in sub-network 1, many genes were enriched in cell signaling pathways (e.g. WNT, PI3K–AKT, VEGF), cell differentiation, cell proliferation, development processes, and ECM organization. These processes, including ECM remodeling, are known to control stem cell behavior (Al Adhami et al., 2015). These enrichment results are consistent with previous reports that *Zac1*- and *H19*-regulated IGNs are critically involved in embryonic and postnatal growth and stemness (Varrault et al., 2006; Lui et al., 2008; Gabory et al., 2009; Berg et al., 2011). Interestingly, genes in sub-network 2 were enriched in transport and metabolic processes, which has not been described in previous IGNs. Earlier studies focusing on single IGs have demonstrated that various IGs are associated with metabolism, including *Nnat* (Millership et al., 2018) and *Rasgrf1* (Font de Mora et al., 2003). Notably, these IGs were also found in sub-network 2. Considering that the mammary gland is a nutrition delivery organ and highly correlated with lipid and protein metabolism (Bequette et al., 1998; Rudolph et al., 2007; Osorio et al., 2016), these results suggest that sub-network 2 may be associated with lactation function. Taken together, in the mammary gland, IGs were co-expressed or co-regulated within two sub-networks by targeting different pre-existing machinery to execute corresponding biological processes in gland development and lactation.

Previous studies have indicated that IGs are dynamic in terms of temporal and spatial regulation during embryogenesis and brain development (Babak et al., 2015; Perez et al., 2015; Andergassen et al., 2017). In this study, based on bulk-seq data from the mammary gland, we demonstrated a decreasing tendency in imprinting number and expression levels in mouse IGs during mammary gland development. Earlier research reported that IGs exhibit lower expression and loss of imprinting status in human breast cancer, and that the loss of imprinting is often regulated by the down-regulation of the previously activated allele, leading to higher relative expression of the (imperfectly) silenced allele (Goovaerts et al., 2018). In this study, however, although IGs in the mouse mammary gland showed low expression and loss of imprinting status during development, this loss was mainly due to inconsistent parental allele bias across replicates rather than a reduction in allele bias. These results imply different mechanisms in regulating the imprinting status of IGs between normal tissue development and disease. It should be noted, however, that mammary gland tissue is heterogeneous and IGs show cell lineage-specific expression. Thus, further investigation on IG status at the single-cell level may offer insight into the links between imprinting status regulation and the biological function of IGs.

IGs play regulatory roles in fetal growth and early embryogenesis (Plasschaert and Bartolomei, 2014). Various studies have demonstrated the importance of normal IG expression in disease and adult stem cells. For example, *H19* is involved in the maintenance of adult hematopoietic stem cell populations in mice (Venkatraman et al., 2013; Plasschaert and Bartolomei, 2014) and loss of *Dlk1*, a paternally expressed gene, in neural stem cells is important for postnatal neurogenesis (Ferron et al., 2011). The mammary gland is a special organ, which experiences rapid development after birth. In this study, we detected the expression profiles of IGs in general and found that they were lowly expressed during lactation and in differentiated mammary epithelial cells (Figure 2A; Supplementary Figure S6). These results imply that lower expression of partial IGs is required for differentiation of mammary epithelial cells and IGs are important for stemness maintenance. The overexpression of *Peg1*, a paternal gene showing low expression during lactation, has been shown to inhibit β-casein induction during HC11 lactogenic differentiation, and thus decrease differentiation capacity (Yonekura et al., 2019). Our data further confirmed that lower expression of *Sgce* and *Ube3a* increased differentiation ability, as observed in the dome formation assay and by marker gene levels (Figure 3). Thus, these results demonstrate that IGs play important roles in MaSC maintenance and mammary gland cell differentiation.

Materials and methods

Mice

The F1 hybrids were generated by reciprocally crossing C57Bl/6J and PWK/PhJ mouse strains. Mouse mammary gland tissues from F1 mice were collected at LD-15 and dissected. In total, five C57BL/6 × PWK/PhJ F1 hybrids and four PWK/PhJ × C57BL/6 F1 hybrids were sampled. All experimental and animal care and handling procedures were performed per the protocols approved by the Ethics Committee of the Kunming Institute of Zoology, CAS.

RNA extraction and sequencing

Total RNA was extracted using TRIzol reagent (Life Technologies). The RNA concentrations and A260 nm/A280 nm

ratios were quantified with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific). The RNA samples were sequenced on an Illumina X-Ten platform following the specified protocols for mRNA sequencing sample preparation.

Read alignment and assignment

Sequenced reads were quality controlled using Trimmomatic (v0.36) (Bolger et al., 2014). To improve alignment quality and minimize bias caused by differences in genetic distances between parental genomes and the reference sequence, we adapted a previous analysis strategy (Crowley et al., 2015) by aligning the sequenced reads to the parental 'pseudogenomes'. Later, we used pylapels and Suspenders to transfer the coordinates each mapping result back to mm10 reference genome and identify the reads parental origination (Huang et al., 2014). The IGs were identified using the ISoLDE script following the recommended workflow with the threshold and/or default methods (if more than 3×2 replicates were available). The assignment of parental reads and identification of IGs is described in the Supplementary material.

The RNA-seq data from the HC11 cell line were directly mapped to the mouse genome (mm10) using Hisat2 (Kim et al., 2019) with the Ensembl gene annotation file (Release 96) after read quality control. The featureCounts pipeline was used to produce a table of read counts at the gene level (Liao et al., 2014). Differential expression analysis was performed using edgeR (Robinson et al., 2010).

Transcriptome annotation file

Gene annotation file for IGs identification was generated by merging three annotation sources, i.e. the Ensembl (Wang et al., 2011), RefSeq (Andergassen et al., 2017), and transcripts newly assembled using RNA-seq data from the hybrid individuals sequenced in this study and RNA-seq data fromAndergassen et al. (2017). Details are described in the Supplementary material.

The priori IGs were retrieved from www.geneimprint.com and www.har.mrc.ac.uk/research/genomic_imprinting. Information on protein coding gene-associated GOBP terms was retrieved from the Ensembl BioMarts database for GSEA.

scRNA-seq data analysis

The scRNA-seq data were downloaded from the NCBI GEO database released from Walid T. Khaled's lab (Bach et al., 2017). The input files were preprocessed using the Seurat package (v3.1) (Stuart et al., 2019). Cells were quality controlled by the total number of genes, unique molecular identifiers, and percentage of molecules mapped to mitochondrial genes. Cells underwent preliminary dimensionality reduction (linear and non-linear dimensional reductions were performed by RunPCA and RunUMAP, respectively) to remove non-epithelial cells using corresponding cell markers (e.g. epithelial cells: Epcam, Cd24a, Cd29, Krt14, and Krt18; immune cells: Cd74, Cd72, and Cd54; endothelial cells: Eng, S1pr1, and Emcn; and pericyte cells: Des and Cspg4). Clusters that showed high expression of different epithelial lineage markers (e.g. clusters that expressed basal and luminal markers concurrently) were classified as doublet clusters and were removed to reduce potential impact on bioinformatics calculations (Bach et al., 2017). Finally, the MaSC, basal, myoepithelium, and luminal cells were included for downstream analysis. The scRNA-seq data only sequenced a small fraction of transcripts present in each cell, which can result in unreliable quantification of genes with low or moderate expression and can thus hinder downstream analysis (Huang et al., 2018), including PCC analysis between gene pairs. Hence, we used SAVER to recover gene expression. To reduce computational burden and time consumption, we re-sampled 500 cells from each lineage if feasible. In total, 3843 cells were used for gene expression recovery and downstream PCC calculation.

Functional annotation of IGs using GSEA

PCCs between each gene were calculated in R using the Hmisc package. Genes significantly correlated ($P \leq 0.01$) with specific IGs were ranked by PCCs and were chosen for GO enrichment assessment using the fgsea package (Sergushichev, 2016), which meant using the ranked PCCs instead of ranked logFC for downstream analysis. We established an association matrix between the IGs and significant GOBP gene sets ($P \leq 0.01$), with 1, -1, and 0 corresponding to positive, negative, and no significant correlation, respectively. The heatmap of GO enrichment was plotted using pheatmap in R.

IGN construction

The co-regulation or co-expression modules of IGs were analyzed using Monocle-3 and WGCNA, respectively. The gene expression matrix recovered by SAVER was used for this analysis. For Monocle-3, find_gene_modules were used for module construction with default parameters. For WGCNA, the co-expression profile was calculated with appropriate parameters (e.g. softPower = 12, type='signed', MEDissThres = 0.7) using step-by-step network construction and module detection strategies. The IGs distributed in each module were then retrieved. The connectivity between each IG was calculated three ways: (i) Directly using an appropriate cut-off value (e.g. $|\text{PCCs}| > 0.4$ and $P < 0.01$); (ii) Constructing connectivity using WGCNA and 'threshold = 0.1' to export the network to Cytoscape; (iii) Using mutual rank (MR) cut-off ($\text{MR} < 575$). The calculation of MR and construction of IGN is described in the Supplementary material. The connections within the network were demonstrated by Cytoscape. The biological functions of the sub-networks were analyzed using DAVID with the corresponding linked non-IGs in each sub-network, respectively. For each IG, the indicated biological functions were inferred by co-expressed non-IGs using the enrichGO and enrichKEGG functions in clusterProfiler (Yu et al., 2012).

Dome formation assay

The HC11 cells were grown in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS), 5 μg/ml insulin, 5 μg/ml gentamicin sulfate, and 10 ng/ml epidermal growth factor (EGF). For lactogenic differentiation assay, confluence HC11 cells were starved for 24 h without EGF, followed by the addition of DIP medium (1 μM dexamethasone, 5 μg/ml insulin, and 5 μg/ml prolactin).

Immunofluorescence

Mammary glands at different developmental stages were fixed in 4% paraformaldehyde (PFA) and embedded in paraffin. For immunofluorescence, sections were rehydrated in graded alcohol, with antigen retrieval then carried out in 10 mM sodium citrate buffer for 20 min. The slides were blocked for 2 h in 10% goat serum, then incubated with primary antibodies at 4°C overnight. Slides were washed with phosphate-buffered saline (PBS) three times (5 min each time) and incubated with secondary antibodies for 1 h at room temperature. The primary antibodies used were Sgce (1:100, Abcam) and Ube3a (1:100, Proteintech) and the secondary antibody used was fluorescein-labeled anti-rabbit (1:200, KPL).

qRT-PCR

Total RNA was extracted using TRIzol reagent and then converted to cDNA using a PrimeScript™ RT reagent kit in accordance with the provided instructions. After this, qRT-PCR was performed on a QuantStudio 3 instrument using SYBR Green PCR Master Mix. Primer sequences used for qRT-PCR analysis are listed in Supplementary Table S4.

Mammosphere formation assay

MaSCs can be maintained and passaged *in vitro* as spheres cultured in suspension and mammosphere counts can represent self-renewal ability (Guo et al., 2012). Mammary gland epithelial cells were cultured as is reported previously (Chakrabarti et al., 2012). The primary mammary gland epithelial cells were transfected with vectors and then plated into 96-well, ultra-low attachment plates at a density of 20000 cells/ml and then incubated at 37°C for 2 weeks. After this, the number of mammospheres was counted.

Transplantation assay

The single-cell suspension of mammary epithelial cells with *Sgce* knockdown was prepared for injection into cleared fat pads. The transplantation assay was performed with different numbers of *Sgce* knockdown mammary primary cells re-suspended in 50% PBS and 50% Matrigel. The MaSC frequency was calculated using extreme limiting dilution assay (ELDA) (Lim et al., 2010).

ChIRP assay

Triplex forming oligonucleotide (TFO) information was retrieved from the LongTarget website (http://lncrna.smu.edu.cn/show/download) (Liu et al., 2017). The RNA sequences at position 4–50 nt (TFO1) and 1870–1926 nt (TFO2) of H19 labeled with biotin were used. This assay was performed following previously published protocols (Postepska-Igielska et al., 2015) and our earlier research (Wang et al., 2018). In this assay, cell nuclei of HC11 cell line were sonicated (20 cycles, 30 sec ON and 45 sec OFF) and spun at 10000 rpm for 5 min at 4°C.

Data access

The raw sequence data reported in this paper were deposited in the Genome Sequence Archive (Wang et al., 2017a) in the BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession numbers CRA001791 and CRA002258, which are publicly accessible at http://bigd.big.ac.cn/gsa. The accession numbers of other data adopted from the NCBI GEO and SRA data repository are GSE106273 and SRP067322, respectively.

Supplementary material

Supplementary material is available at *Journal of Molecular Cell Biology* online.

Acknowledgements

We thank Dr Jiang Liu for constructive suggestions and Dr Christine Watts for English editing.

Funding

This work was supported by the National Key Research and Development Program of China (2016YFA0100900), the National Natural Science Foundation of China (U1802285, 31970612, 81871403, and 81571738), Yunnan Applied Basic Research Key Projects (2018FA002 and 2015HA026), Key Research and Development Program of Zhejiang Province (2019C03014), Open Project from the State Key Laboratory of Genetic Resources and Evolution (GREKF20-04), and Fundamental Research Funds for the Central Universities.

Conflict of interest: none declared.

Author contributions: H.X., L. Zhao., and B.J. interpreted the results and wrote the manuscript. H.X. conducted all bioinformatics analyses. X.F. performed all mouse work. H.X. and L. Zhao performed the experiments. Y.M., W.C., J.S., Q.Y., and L. Zou provided experimental assistance. H.Y. discussed and interpreted the results.

References

Al Adhami, H., Evano, B., Le Digarcher, A., et al. (2015). A systems-level approach to parental genomic imprinting: the imprinted gene network includes extracellular matrix genes and regulates cell cycle exit and differentiation. Genome Res. *25*, 353–367.

Andergassen, D., Dotter, C.P., Wenzel, D., et al. (2017). Mapping the mouse Allelome reveals tissue-specific regulation of allelic expression. eLife *6*, e25125.

Babak, T., DeVeale, B., Tsang, E.K., et al. (2015). Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. Nat. Genet. *47*, 544–549.

Bach, K., Pensa, S., Grzelak, M., et al. (2017). Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. Nat. Commun. *8*, 2128.

Bartolomei, M.S., and Ferguson-Smith, A.C. (2011). Mammalian genomic imprinting. Cold Spring Harb. Perspect. Biol. *3*, a002592.

Bequette, B.J., Backwell, F.R.C., and Crompton, L.A. (1998). Current concepts of amino acid and protein metabolism in the mammary gland of the lactating ruminant. J. Dairy Sci. *81*, 2540–2559.

Berg, J.S., Lin, K.K., Sonnet, C., et al. (2011). Imprinted genes that regulate early mammalian growth are coexpressed in somatic stem cells. PLoS One *6*, e26410.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Brandvain, Y., Van Cleve, J., Ubeda, F., et al. (2011). Demography, kinship, and the evolving theory of genomic imprinting. Trends Genet. *27*, 251–257.

Cao, J., Spielmann, M., Qiu, X., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502.

Chakrabarti, R., Hwang, J., Andres Blanco, M., et al. (2012). Elf5 inhibits the epithelial–mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. Nat. Cell Biol. *14*, 1212–1222.

Constancia, M., Angiolini, E., Sandovici, I., et al. (2005). Adaptation of nutrient supply to fetal demand in the mouse involves interaction between the Igf2 gene and placental transporter systems. Proc. Natl Acad. Sci. USA *102*, 19219–19224.

Crowley, J.J., Zhabotynsky, V., Sun, W., et al. (2015). Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. Nat. Genet. *47*, 353–360.

Edwards, C.A., and Ferguson-Smith, A.C. (2007). Mechanisms regulating imprinted genes in clusters. Curr. Opin. Cell Biol. *19*, 281–289.

Ferron, S.R., Charalambous, M., Radford, E., et al. (2011). Postnatal loss of Dlk1 imprinting in stem cells and niche astrocytes regulates neurogenesis. Nature *475*, 381–385.

Font de Mora, J., Esteban, L.M., Burks, D.J., et al. (2003). Ras–GRF1 signaling is required for normal beta-cell development and glucose homeostasis. EMBO J. *22*, 3039–3049.

Frost, J.M., and Moore, G.E. (2010). The importance of imprinting in the human placenta. PLoS Genet. *6*, e1001015.

Fu, N.Y., Nolan, E., Lindeman, G.J., et al. (2020). Stem cells and the differentiation hierarchy in mammary gland development. Physiol. Rev. *100*, 489–523.

Gabory, A., Ripoche, M.A., Le Digarcher, A., et al. (2009). H19 acts as a *trans* regulator of the imprinted gene network controlling growth in mice. Development *136*, 3413–3421.

Goovaerts, T., Steyaert, S., Vandenbussche, C.A., et al. (2018). A comprehensive overview of genomic imprinting in breast and its deregulation in cancer. Nat. Commun. *9*, 4120.

Green, C.D., Huang, Y., Dou, X., et al. (2017). Impact of dietary interventions on noncoding RNA networks and mRNAs encoding chromatin-related factors. Cell Rep. *18*, 2957–2968.

Guo, W., Keckesova, Z., Donaher, J.L., et al. (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. Cell *148*, 1015–1028.

Huang, M., Wang, J., Torre, E., et al. (2018). SAVER: gene expression recovery for single-cell RNA sequencing. Nat. Methods *15*, 539–542.

Huang, S., Holt, J., Kao, C.Y., et al. (2014). A novel multi-alignment pipeline for high-throughput sequencing data. Database *2014*, bau057.

Iyer, M.K., Niknafs, Y.S., Malik, R., et al. (2015). The landscape of long noncoding RNAs in the human transcriptome. Nat. Genet. *47*, 199–208.

Jonker, J.W., Wagenaar, E., Van Eijl, S., et al. (2003). Deficiency in the organic cation transporters 1 and 2 (Oct1/Oct2 [Slc22a1/Slc22a2]) in mice abolishes renal secretion of organic cations. Mol. Cell. Biol. *23*, 7902–7908.

Keniry, A., Oxley, D., Monnier, P., et al. (2012). The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. Nat. Cell Biol. *14*, 659–665.

Keverne, E.B. (2015). Genomic imprinting, action, and interaction of maternal and fetal genomes. Proc. Natl Acad. Sci. USA *112*, 6834–6840.

Kim, D., Paggi, J.M., Park, C., et al. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. *37*, 907–915.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics *9*, 559.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Lim, E., Wu, D., Pal, B., et al. (2010). Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast Cancer Res. *12*, R21.

Liu, H., Shang, X., and Zhu, H. (2017). LncRNA/DNA binding analysis reveals losses and gains and lineage specificity of genomic imprinting in mammals. Bioinformatics *33*, 1431–1436.

Lui, J.C., Finkielstain, G.P., Barnes, K.M., et al. (2008). An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs. Am. J. Physiol. Regul. Integr. Comp. Physiol. *295*, R189–R196.

Millership, S.J., Da Silva Xavier, G., Choudhury, A.I., et al. (2018). Neuronatin regulates pancreatic beta cell insulin content and secretion. J. Clin. Invest. *128*, 3369–3381.

Neirijnck, Y., Papaioannou, M.D., and Nef, S. (2019). The insulin/IGF system in mammalian sexual development and reproduction. Int. J. Mol. Sci. *20*, 4440.

Osorio, J.S., Lohakare, J., and Bionaz, M. (2016). Biosynthesis of milk fat, protein, and lactose: roles of transcriptional and posttranscriptional regulation. Physiol. Genomics *48*, 231–256.

Patten, M.M., Ross, L., Curley, J.P., et al. (2014). The evolution of genomic imprinting: theories, predictions and empirical tests. Heredity *113*, 119–128.

Perez, J.D., Rubinstein, N.D., Fernandez, D.E., et al. (2015). Quantitative and functional interrogation of parent-of-origin allelic expression biases in the brain. eLife *4*, e07860.

Perotti, C., Wiedl, T., Florin, L., et al. (2009). Characterization of mammary epithelial cell line HC11 using the NIA 15k gene array reveals potential regulators of the undifferentiated and differentiated phenotypes. Differentiation *78*, 269–282.

Plasschaert, R.N., and Bartolomei, M.S. (2014). Genomic imprinting in development, growth, behavior and stem cells. Development *141*, 1805–1813.

Postepska-Igielska, A., Giwojna, A., Gasri-Plotnitsky, L., et al. (2015). LncRNA Khps1 regulates expression of the proto-oncogene SPHK1 via triplex-mediated changes in chromatin structure. Mol. Cell *60*, 626–636.

Renfree, M.B., Suzuki, S., and Kaneko-Ishino, T. (2013). The origin and evolution of genomic imprinting and viviparity in mammals. Philos. Trans. R. Soc. Lond. B Biol. Sci. *368*, 20120151.

Reynes, C., Kister, G., Rohmer, M., et al. (2020). ISoLDE: a data-driven statistical method for the inference of allelic imbalance in datasets with reciprocal crosses. Bioinformatics *36*, 504–513.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Rudolph, M.C., McManaman, J.L., Phang, T., et al. (2007). Metabolic regulation in the lactating mammary gland: a lipid synthesizing machine. Physiol. Genomics *28*, 323–336.

Sandhu, K.S. (2010). Systems properties of proteins encoded by imprinted genes. Epigenetics *5*, 627–636.

Sergushichev, A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. bioRxiv, https://doi.org/10.1101/060012

Stringer, J.M., Pask, A.J., Shaw, G., et al. (2014). Post-natal imprinting: evidence from marsupials. Heredity *113*, 145–155.

Stringer, J.M., Suzuki, S., Pask, A.J., et al. (2012). Selected imprinting of INS in the marsupial. Epigenet. Chromatin *5*, 14.

Stuart, T., Butler, A., Hoffman, P., et al. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21.

Tunster, S.J., Jensen, A.B., and John, R.M. (2013). Imprinted genes in mouse placental development and the regulation of fetal energy stores. Reproduction *145*, R117–R137.

Varrault, A., Gueydan, C., Delalbre, A., et al. (2006). Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. Dev. Cell *11*, 711–722.

Venkatraman, A., He, X.C., Thorvaldsen, J.L., et al. (2013). Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. Nature *500*, 345–349.

Wagschal, A., and Feil, R. (2006). Genomic imprinting in the placenta. Cytogenet. Genome Res. *113*, 90–98.

Wang, S., Ke, H., Zhang, H., et al. (2018). LncRNA MIR100HG promotes cell proliferation in triple-negative breast cancer through triplex formation with p27 loci. Cell Death Dis. *9*, 805.

Wang, X., Soloway, P.D., and Clark, A.G. (2011). A survey for novel imprinted genes in the mouse placenta by mRNA-seq. Genetics *189*, 109–122.

Wang, Y., Gao, S., Zhao, Y., et al. (2019). Allele-specific expression and alternative splicing in horse × donkey and cattle × yak hybrids. Zool. Res. *40*, 293–304.

Wang, Y.M., Xu, H.B., Wang, M.S., et al. (2017b). Annotating long intergenic non-coding RNAs under artificial selection during chicken domestication. BMC Evol. Biol. *17*, 192.

Wang, Y., Song, F., Zhu, J., et al. (2017a). GSA: Genome Sequence Archive*. Genomics Proteomics Bioinform. *15*, 14–18.

Wilkins, J.F. (2014). Genomic imprinting of Grb10: coadaptation or conflict? PLoS Biol. *12*, e1001800.

Yonekura, S., Ohata, M., Tsuchiya, M., et al. (2019). Peg1/Mest, an imprinted gene, is involved in mammary gland maturation. J. Cell. Physiol. *234*, 1080–1087.

Yu, G., Wang, L.G., Han, Y., et al. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287.

Zhang, P.M., Wong, C., DePinho, R.A., et al. (1998). Cooperation between the Cdk inhibitors $p27^{KIP1}$ and $p57^{KIP2}$ in the control of tissue growth and development. Gene Dev. *12*, 3162–3167.

Zwart, R., Verhaagh, S., Buitelaar, M., et al. (2001). Impaired activity of the extraneuronal monoamine transporter system known as uptake-2 in Orct3/Slc22a3-deficient mice. Mol. Cell. Biol. *21*, 4188–4196.